

Ora Matushansky, CNRS/Université Paris-8/UiL OTS/Utrecht University
 email: O.M.Matushansky@uu.nl
 homepage: <http://www.let.uu.nl/~Ora.Matushansky/personal/>

ENGENDERING PHI-FEATURES
Typology of Morphosyntactic Parameters, МГГУ & Ин. языкознания РАН,
October 16-18, 2013

φ-features are **inherent** properties of an NP that affect its **agreement** patterns (and anaphora)

The **declension class** of a noun (a.k.a. its *word class*), while being inherent, only affects the morphological realization of the noun itself

Usually phi-features are assumed to be linked to dedicated loci in the extended NP:

- Gender & animacy: introduced by the head noun (or by nP/GenP)
- Number: determined by the presence of number morphology (NumP)
- Person: exclusively property of pronouns (i.e., of D°)
- Definiteness: correlated with D°

The actual picture is more complicated:

- gender: mixed agreement requires gender higher in the NP
- number: pluralia tantum are lexically determined

Do we need two mechanisms of engendering phi-features: in the syntax and in the lexicon?

1. INTRODUCTION: THE RUSSIAN φ

Animacy: governs the accusative case syncretism: accusative surfaces as genitive for animate nouns and as nominative for inanimate nouns in the plural of all declension classes and in the singular for nouns ending in a consonant on the surface (type *ř*, a.k.a. the second declension)

- singular *-o-* and *-ř-* nouns are inanimate irrespective of their denotation
- nouns denoting supernatural or human beings, dolls (including marionettes) and animals are animate (even if their denotatum isn't, e.g., *pokojnik*, *mertvec* 'a dead person')
- all other nouns are inanimate

Gender: the Russian gender system is a mixed one (Corbett 1991) with the following default gender assignment rules:

- nouns denoting males are masculine
- nouns denoting females are feminine
- declinable nouns of the declension class *ř* are masculine
- declinable nouns of the declension classes *a* and *ř* are feminine
- declinable nouns of the declension class *o* are neuter
- animate indeclinable nouns are masculine (with some exceptions)
- inanimate indeclinable nouns are neuter (again, with some exceptions)
- the gender of indeclinable acronyms is determined by the gender of the head

Individual loanwords can be lexically specified with a given gender, overriding assignment from the declension class (e.g., *koala* 'koala' is masculine whereas *panda* 'panda' is feminine)

Number: NPs denoting plural individuals are plural, but plurality can also be specified in the lexical entry (*pluralia tantum*; for both count and mass nouns)

Summary: two potential sources of a phi-feature: lexical specification and semantics. Lexical specification can in fact come in two forms:

- declension class (for gender and arguably for number)
- featural specification

When is the connection between the declension class and gender established?

1.1. Syntactic gender and mixed agreement

Gender can be semantic and formal (syntactic, morphological...):

- Semantic gender of a noun is determined by the properties of entities belonging to its extension (e.g., natural gender/sex, animacy, non-flesh food in Dyirbal (Dixon 1982:178) – see Corbett 1991)
- Formal gender is random or determined by some morphological properties of the noun (e.g., its declension class)
- Gender systems seem to always have a semantic core

Mixed systems may give rise to mixed agreement:

- (1) Naša vrač – umnica. Russian
 our-FSG doctor.M clever.person
Our doctor is very clever.

Although the noun *vrač* ‘doctor’ is inherently masculine, feminine marking is possible on the determiner, attributive modifiers and the predicate (constrained by Agreement Hierarchy, see Corbett 1979 and later work) if our doctor is female.

Importantly, it is impossible to claim that the gender feature of the noun itself is different:

- (2) Naša zubnoj vrač – umnica. Russian
 our-FSG dental-MSG doctor.M clever.person
Our dentist is very clever.

So not only is gender a syntactically active feature (unlike declension class), but there exists a mechanism for establishing it in syntax

1.2. Code-switching

The (forced) use of a foreign noun (not a borrowing) in contexts where it requires agreement necessitates the existence of a syntactically active process of gender assignment (since it is impossible to assume that such online use relies on gender assignment in the lexicon).

Gender assignment in code-switching relies on a number of criteria (Leisiö 2001, Chirsheva 2009 for Russian, Poplack et al. 1982, Poplack et al. 1988, Fuller and Lehnert 2000, Violin-Wigent 2006 for other languages)

In Russian:

- **phonology:** consonant-final words are usually assigned to the masculine gender
- **semantics:** nouns denoting female and male humans are assigned to the feminine and masculine gender, respectively
- **semantic analogy:** nouns are assigned the gender of the corresponding noun (or of the noun denoting the containing taxonomic category) in the matrix language
- orthography, statistic generalizations over the vocabulary of the matrix language, cognate in the matrix language, etc.

Like in the native vocabulary, **semantic gender overrides phonology**.

Many of these factors also determine gender assignment of loanwords (cf., e.g., Rabeno and Repetti 1997)

Important: in function of the language and the experimental technique different factors play more or less important roles.

1.3. Indeclinable proper names

Russian proper names follow exactly the same rules as common nouns, with the exception of indeclinable proper names, which are assigned gender by semantic analogy in function of the associated sortal instead of the expected neuter (Rozental et al. 1998:204-205):

- (3) a. solnečnyj [gorod] Tbilisi Rozental et al. 1998:204
 sunny-MSG city.M Tbilisi
the sunny city of Tbilisi
- b. polnovodnoe [ozero] Èri
 full-flowing-NSG lake.N Erie
the full-flowing lake of Erie
- c. trudno.dostupnaja [gora] Jungfrau
 hard.accessible-FSG mountain.F Jungfrau
the nearly inaccessible Jungfrau

The same proper name may be assigned different genders when considered under different guises:

- (4) a. V period krizisa Somali stradalo ot nexvatki prodovol'stvija.
 in period crisis-GEN Somali suffered-NSG from lack victuals-GEN
During the crisis (the state of) Somali suffered from the lack of food.
- b. Somali prinjala s blagodarnost'ju gumanitarnuju pomošč'.
 Somali accepted-FSG with gratitude humanitarian aid
(The country of) Somali accepted the humanitarian aid with gratitude.

Gender assignment by semantic analogy is considerably more frequent in languages that have no declension classes (e.g., in French; see also Fraurud 1999 on the gender of toponyms in Swedish, Lang 1976, Hickey 1999 on German; De Clercq 2008), though crucially for proper names it is not always the gender of the associated sortal (e.g., for German city names)

1.4. Conclusion

Semantic gender assignment has to be available outside the lexicon.

Is there any reason to believe that formal gender assignment occurs *in* the lexicon?

- Is it really necessary to have two systems of gender assignment?
- And if there were two systems, wouldn't we expect two gender paradigms (e.g., noun class vs. natural gender) instead of one, in at least some languages?

However, if gender assignment is syntactic, how can the gender of indeclinable loanwords be exceptional? (Or is this too peripheral to matter?)

The \$1,000,000 question: **why is there only one gender system?**

2. THE MECHANISM OF GENDER ASSIGNMENT IN RUSSIAN

Core hypothesis: the so-called "formal" or "grammatical" gender is interpretable at LF

It virtually has to be:

- if it weren't, gender would be a narrow-syntactic property; distinguishing it from declension classes would be purely a matter of being syntactically active
- if it weren't, we wouldn't expect formal gender (of inanimate nouns) to depend on their semantics
- co-variation in pronominal gender is not determined in the narrow syntax (but see Elbourne 2002 for a hypothesis that pronouns involve NP-ellipsis)

Proposal: while "semantic gender" classifies individuals (most frequently, by natural gender), **"formal gender" classifies kinds**

The basis for this classification is the declension class or (primarily in its absence in Russian) semantic analogy (on the common semantic core of human first declension nouns see Nesset 2001).

The connection between natural gender and semantic gender is not a bijection or a function:

- masculine diminutives of feminine first names (e.g., *Lizok*)
- feminine-agreeing "common gender" nouns with male referents (e.g., *svoloč'*) and the use of feminine agreement to indicate male homosexuality or insignificance

Cross-linguistically feminine gender frequently contains diminutives (e.g., in the Ethiosemitic language Amharic (Kramer 2012), in the Omotic language Dizi, a.k.a. Maji (Allan 1976 via Corbett 1991) and in the Salish language Halkomelem (Steriopolo and Wiltschko 2008))

Formalization: the features [marked] (masculine) and [minor] (feminine)

Assuming that neuter is the default gender, formal gender assignment looks as follows:

- (5) a. declinable nouns of the declension class \check{r} are masculine
 declinable nouns of the declension classes *a* and \check{i} are feminine
 other declinable nouns are neuter
- b. [minor] nouns are feminine
 [marked] nouns (including animate nouns) are masculine
 other nouns are neuter

Some questions:

- how are the two sets of rules ordered?
- how to handle exceptions?
- **what about number?**

2.1. Number as a formal feature

Neutralization of gender in the plural is very common cross-linguistically

Hypothesis (for Russian): number can be a function of the declension class (cf. the "paired" gender of Zaliznjak 1967):

- *pluralia tantum* nouns belong to the plural declension classes (e.g., the adjectival *ozimye* 'winter crops' or the nominal *sani* 'sleigh')
- in the absence of a singular declension marker gender is not established

NB: The cardinal *dva/dve* 'two' and the quantifier *oba/obe* 'both' agree for gender, but combine with a noun in what is morphologically the genitive singular form. We set these aside (but see Ionin and Matushansky 2006 for the hypothesis that cardinals combine with a semantically singular lexical NP)

Russian *pluralia tantum* nouns are compatible with cardinals (with the exception of the lower cardinals combining with the paucal (= morphologically genitive singular) form of the noun) precisely because semantically they are not plural

If nominal plural specification can be purely formal, mass *pluralia tantum* (such as *černila* 'ink', *osadki* 'precipitation', *den'gi* 'money', etc.) are no longer a mystery without the need to assume that the number feature can be semantically vacuous, as it will be interpreted as kind-level classification:

- (6) a. nouns of the declension class PL are [plural]
 b. NPs denoting plural entities are [plural]

On the other hand, for most nouns the plural noun classes result from semantics, showing that the process can go in both directions

2.2. Animacy as a formal feature

Animacy is reflected in the realization of the accusative case:

- (7) accusative is realized:
 a. as a special form (-u-) in the declension class -a-
 b. as **genitive** for animate NPs that are masculine or plural
 c. as **nominative** elsewhere

In other words, for nouns in the declension classes -o- and -ĭ-, including the 10 nouns in -mja- and the noun *ditja* 'child', the animacy distinction is lost in the singular, but retained in the plural

NB: Corbett 1980 proposes that the semantic feature [animate] is carried over to syntax only with [masculine] or [plural] nouns, which is merely a restatement of the facts with no explanatory value

If plural is a declension class, the syntax of animacy may depend directly on the declension class rather than gender/number:

- (8) semantic animacy is realized as [animate] in the declension classes PL and -ĭ-

This is unlikely, as cardinal-containing NPs with a lexically animate head noun can behave as inanimates when denoting a measure (Mel'čuk 1980a, b, Matushansky and Ruys 2012):

- (9) a. [siloj rovno] v **tri medvedja** Mel'čuk 1980b
 strength-INS exactly in three-ACC=NOM bears-GEN
as strong as exactly three bears
 b. [bol'se] na **dva mal'čika**
 more on two-ACC=NOM boys-GEN
two boys more

Conclusion: the formal feature of **animacy** is a DP-level property established **very late** (as it interacts with the semantic number, but also with the denotation of the entire NP)

2.3. The phenomenon of mixed agreement

Crockett 1976: agreement can be determined by the semantic properties of the referent of a DP rather than by the formal features of the DP itself:

- (10) presupposition: our doctor is a woman
- a. Naš vrač prišël vovremja.
 our-MSG doctor.M arrived-MSG on.time
Our doctor arrived on time.
- b. %Naš vrač prišla vovremja.
 our-MSG doctor.M arrived-FSG on.time
- c. *Naša vrač prišël vovremja.
 our-FSG doctor.M arrived-MSG on.time

- d. Naša vrač prišla vovremja.
 our-FSG doctor.M arrived-FSG on.time

Cross-linguistically a pattern known as *Agreement Hierarchy* (Corbett 1979) emerges:

(11) **Agreement Hierarchy**

DP-internal < predicate < relative pronoun < personal/possessive pronoun

The likelihood of semantic agreement increases rightwards in (11), both within one language and cross-linguistically.

NB: Mixed agreement patterns are not limited to gender: number (cf. Sauerland and Elbourne 2002 on British English collective nouns) and person (Harmer and Norton 1957:270 via Corbett 2006:132) may also show such effects; see also Neeleman 2008.

DP-internal mixed agreement is likewise possible:

- (12) U nas byla ocen' xorošaja zubnoj vrač. Crockett 1976
 with us was-FSG very good-FSG dental-MSG doctor.M
We had a very good dentist.

Rothstein 1980, Nikunlassi 2000, Asarina 2008: reference-modifying and referent-modifying adjectives do not behave the same with respect to mixed agreement:

- (13) a. zubnoj/*zubnaja vrač Asarina 2008
 dental-MSG doctor.M
dentist
- b. Umelyj/umelaja vrač bystro postavila plombu.
 skillful-MSG/FSG doctor.M quickly stood-up-FSG filling
The skillful doctor quickly put in the filling. (i.e. skillful as a doctor)
- c. Umelaja/*umelyj vrač bystro postroila lodku.
 skillful-FSG/MSG doctor.M quickly built-FSG boat
The skillful doctor quickly built a boat. (i.e. skillful as a person, e.g., on a reality TV show)

Kind-level modification (cf. McNally and Boleda 2004) does not allow mixed agreement

NB: Rothstein 1980 distinguishes two types of referent-modification, but the grammaticality of both agreement patterns in (13b) can also be attributed to the fact that both reference-modification and referent-modification are compatible with the *skillful as a doctor* reading.

Mixed agreement is compatible with all genders and declension classes:

- (14) a. Ego vysočestvo/svetlost' nedovolen. styles
 his Highness.N/Highness.F not.content-MSG
His Highness (a prince/other dignitary) is not satisfied.
- b. Ja znaju ètogo nesčastnogo/ ètu nesčastnuju sirotu.
 I know this-MSG-ACC poor-MSG-ACC/ this-FSG-ACC poor-FSG-ACC orphan-ACC
I know this poor orphan.

Observation: semantic agreement with style nouns cannot be DP-internal, whereas semantic agreement with the so-called "common gender" *-a*-declension nouns is available in oblique cases as well

Additional wrinkle: mixed agreement in Russian is only possible in the nominative (Švedova 1980:§1819); in Polish epicene nouns denoting females may fail to decline (Rothstein 1980), as do Russian feminine first and last names ending in a consonant (e.g., *Èdit* 'Edith')

2.4. The phenomenon of declension class assignment

Poplack et al. 1982, Rabeno and Repetti 1997, Chirsheva 2009: even in code-switching (and more so in borrowing) a noun may acquire a word-class/declension class marker absent in the original phonological representation:

- (15) On guljaet so svo-ej dog-oj. Chirsheva 2009
 he walks with REFL-FSG.INS dog-FSG.INS
He walks with his dog.

The only reason for this can be that the declension class can arise from the underlying gender specification (in addition to being itself an underlying nominal property influencing gender assignment)

Support: loanwords are not assigned into the minor *-ŷ-* declension class

Furthermore, the nominative case endings of the three major declension classes are identical to the corresponding gender markers (*-a-* for feminine, *-ŷ-* for masculine and *-o-* for neuter) as identified by gender marking on adjectives and verbs

So a noun can be specified for any combination of the following:

- declension class
- grammatical gender
- lexical-semantic gender (natural gender for humans, semantic analogy for others)

Ideally all three match and fit in with the natural gender of the referent of the noun phrase

3. MATCHES AND MISMATCHES

Is Russian like German or French?

3.1. Simple case: indeclinable nouns

As indeclinable nouns have no declension class, their gender features can only be specified in the lexicon

Inanimate indeclinable nouns are mostly **neuter** (which we treat as the default).

Exceptionally inanimate indeclinable nouns can also be masculine (e.g., *sirokko* 'Sirocco') or feminine (e.g., *kol'rabi* 'kohlrabi')

Following Heim and Kratzer 1998, gender features are introduced as identity functions:

- (16) a. $[[F]] = \lambda f . \lambda x : x \text{ is classified as minor} . f(x)$
 b. $[[M]] = \lambda f . \lambda x : x \text{ is classified as marked} . f(x)$

Animate indeclinable nouns are mostly masculine. Given that the formal feature of animacy is always an entailment of the lexical semantics of the noun, we hypothesize that [+animate] gives rise to [marked] (in the absence of a declension class)

Exceptionally animate indeclinable nouns can also be feminine (e.g., *cece* 'a tsetse fly') and thus bear the feature [minor] (which naturally does not conflict with the feature [marked] that arises from animacy)

Some animate indeclinable nouns agree according to the natural gender of the referent (e.g., *šimpanze* 'chimpanzee')

To the best of my knowledge there are **no animate neuter indeclinable nouns**

So for indeclinable nouns Russian behaves like German or French.

3.2. Inanimate declinable nouns

For inanimate nouns the declension class usually determines the gender (but crucially, not the other way around):

(17) declension → gender

- a. nouns of the declension class *-ř-* are [marked] (masculine)
- b. nouns of the declension classes *-a-* and *-ĩ-* are [minor] (feminine)
- c. nouns of the declension classes *-o-* and *-0-* are [-animate] (neuter)

The minor declension classes *-ř-* (the third declension) and *-0-* (the 10 nouns in *-mja-* and the noun *ditja* 'child') show that even for inanimate nouns the declension class cannot be derived from gender

Systematic exceptions to (17): several diminutive and augmentative suffixes that preserve the gender of the noun they combine with, but not its declension class (Sitchinava 2011):

- (18) a. *dom* 'house' (M, declension class *-ř-*) → *dom-iřk-a* (M, dim, decl. class *-a-* or *-o-*)
 b. *pal'to* 'coat' (N, indeclinable) → *pal't-iřk-o* (N, dim, decl. class *-o-*)
 c. *sluřba* 'service' (F, declension class *-a-*) → *sluřb-iřk-a* (F, dim, decl. class *-a-*)

NB: The declension class of the output noun is predictable from the gender of the source noun and its animacy.

However, for the major declension classes it is also possible to assume the underlying gender specifications [marked], [minor] and [-animate]

3.3. Animate declinable nouns

The existence of neuter animate nouns requires the existence of the rule in (17c): otherwise *nasekomoe* 'insect', being animate, would be incorrectly predicted to be classified as [marked] (i.e., formally masculine)

It also becomes **impossible to treat neuter as the absence of a feature** (because some value must be assigned to block the assignment of [marked])

For animate nouns that are not [human] the declension class generally determines the gender, as in (17)

Systematic exceptions: nouns formed with the diminutive and augmentative suffixes as above

Lexical exceptions: masculine first declension nouns *koala* 'koala' (by semantic analogy with the masculine noun *medved'* 'bear'), *zajka* 'hare-DIM', and perhaps others

NB: For some speakers *koala* agrees according to the natural gender of the referent

3.4. Human declinable nouns

There is generally a match between the lexical semantics of a [human] noun and its formal gender: nouns that denote females are feminine, while nouns that denote males are masculine.

Systematic exceptions: declension class *-ř-* caritatives formed from proper names (e.g., *Lizok* from Liza, *Mařkin* from Masha, etc.), which may trigger masculine predicate agreement (cf. Sitchinava 2011)

Agreement with the natural gender of the referent is systematically only possible for [human] nouns (and not with all of them):

Zaliznjak 1967:70-71: the feature [human] could also be relevant for collective cardinals *dvoe* 'two', *troe* 'three', etc. (which prescriptive grammars restrict to nouns denoting male humans)

- common gender nouns: declension class *-a-*, agreement by the natural gender of the referent NP-internally and NP-externally in both direct and oblique cases.

- hybrid nouns: declension class *-ŕ-*, agreement by the natural gender of the referent only possible in the nominative case
- style nouns: declension classes *-o-* and *-ŕ-*, agreement by the natural gender of the referent only possible for the predicate and the relative pronoun. It is not obvious, however, that such nouns are [human]

3.4.1. Common gender

Iomdin 1980: "common gender" nouns actually divide into two subclasses:

- (19) a. Petja *izvestnaja/izvestnyj* zanuda. Iomdin 1980
 Peter-DIM well-known-FSG/MSG bore
Peter is a well-known bore.
- b. Žena Peti *izvestnaja/*izvestnyj* zanuda.
 wife Peter-DIM-GEN well-known-FSG/MSG bore
Peter's wife is a well-known bore.
- c. Ne obraščaj *vnimanija* na ètu staruju zanudu!
 NEG direct-IMP attention on this-FSG old-FSG bore.F
Pay no attention to this old bore (male or female).
- (20) a. Našego novogo/našu novuju starostu zovut Maša.
 our-MSG new-MSG/our-FSG new-FSG monitor call-3PL Masha
Our new class monitor is called Masha.
- b. Našego byvšego/*našu byvšuju starostu zvali Vanja.
 our-MSG previous-MSG/our-FSG previous-FSG monitor call-PL Vanya
Our previous class monitor was called Vanya.
- c. Stoilo li vybirat' novogo starostu?
 be.worth-PAST-NSG Q elect-INF new-MSG monitor
Was electing a new class monitor (male or female) worth it?

Iomdin proposes that common gender nouns are listed in the lexicon twice, with the marked agreement option identified as attributing to the referent the corresponding natural gender:

- (21) a. $[[zanuda_F]] = \lambda x . x$ is a bore expressive epithets
 b. $[[zanuda_M]] = \lambda x : x$ is male . x is a bore
- (22) a. $[[starosta_M]] = \lambda x . x$ is a monitor profession nouns
 b. $[[starosta_F]] = \lambda x : x$ is female . x is a monitor

Problem: there is no connection between these two lexical entries

More than one factor contributes to gender assignment, and each noun can fit into more than one lexical-semantic class, whose assignments may compete:

- [+animate] → [marked] (i.e., masculine)
- *-a-* declension class → [minor] (i.e., feminine)

Intuition: [human] nouns in the *-a-* declension class switch gender very easily because their formal gender assignment is handled by conflicting rules

Their gender in the absence of the natural gender is determined by their lexical semantics (cf. Nessel 2001):

- profession nouns are [marked]
- expressive epithets are [minor]

Animate declinable nouns show that this declension class (*-a-*) is compatible with non-human masculine nouns that are grammatically specified for the masculine gender

Puzzle: there are some masculine nouns in the *-a-* declension class that do not have feminine counterparts (e.g., (*gorodskoj*) *golova* 'the mayor')

The mechanism of assigning gender on the basis of the natural gender of the referent remains to be determined

3.4.2. Hybrid nouns

Unlike in the declension class *-a-*, [human] profession-denoting nouns in the declension class *-ŷ-* are assigned the same gender by their lexical semantics and their phonology:

- [+animate] → [marked] (i.e., masculine)
- *-ŷ-* declension class → [marked]

They are therefore correctly predicted to be more resistant to natural gender assignment (the mechanism still to be determined)

Intuition to develop: the inability of hybrid nouns to appear in mixed agreement patterns in non-nominative case positions is akin to the similar restriction on the use of case-deficient elements (Testelefs 2013).

3.4.3. Style nouns

Like hybrid nouns, style nouns can give rise to mixed agreement, but only NP-externally:

- (23) *Ego korolevskoe/*korolevskij vysočestvo nedovolen/?nedovol'no.*
 his royal-NSG/MSG highness dissatisfied-MSG/NSG
His Royal Highness is dissatisfied.

Their animacy (only potentially observable in the plural) is unclear

Their behavior in non-nominative case positions is impossible to determine

3.5. The mechanism of mixed agreement

Where is gender introduced structurally? Different answers, depending on who you ask:

- Sauerland 2004: all phi-features are interpretable only on the dedicated functional head ϕ (which is the highest functional head in the extended NP projection)
- Picallo 2005: gender is introduced in the dedicated functional projection GenP
- Kihm 2005, Kramer 2012: gender is introduced on n°
- Neeleman 2008, Matushansky 2013: while normally phi-features are interpreted only on nouns, under some circumstances they can be on other elements
- Steriopolo and Wiltschko 2010: gender can be introduced on the root (for nouns that lexically encode gender), on n° (for grammatical gender) or on D° (for mixed agreement)
- Percus 2011: gender is introduced on n° , but sometimes not interpreted
- Pesetsky [to appear]: the interpreted [feminine] feature in mixed agreement cases is introduced on the functional head \mathcal{K} in the extended NP projection

Assuming the last option (not my previous proposal): mixed agreement involves a dedicated functional head \mathcal{K} that introduces a presupposition on the external argument, which gives rise to the [minor] feature:

- (24) $[[\mathcal{K}]] = \lambda f . \lambda x : x \text{ is a female} . f(x)$

Novel assumption: the gender feature on \mathcal{K} syntactically interacts with the gender feature of the NP it combines with: for full syntactic integration they must match

Independent motivation: case-agreement with proper names

4. CASE-MARKING IN CLOSE APPPOSITION

Depending on the lexical-semantic category, the proper name can appear in the same case as the sortal (i.e., the case assigned to the entire NP) or in the default nominative case:

- (25) a. o russk-om poèt-e Blok-e/*Blok [+animate]
 about Russian-MSG.LOC poet.M-LOC Blok.M-LOC/*NOM
about the Russian poet Blok
- b. o roman-e "Gorod/*Gorod-e" man-made object
 about novel.M-LOC City.M-NOM/*LOC
about the novel The City
- c. na ulic-e Jakimank-a/Jakimank-e toponym
 in street.F-LOC Yakimanka.F-NOM/LOC
on the Yakimanka street

Even when the lexical-semantic category is fixed, the lexical category and formal features (ϕ -features) of the proper name can affect case-marking:

- (26) a. na ulic-e Jakimank-a/Jakimank-e ✓phi-congruent
 in street.MSG-LOC Yakimanka.FSG-NOM/LOC
on the Yakimanka street
- b. na ulic-e Balčug/*Balčug-e *non-phi-congruent
 in street.MSG-LOC Balčug.MSG-NOM/LOC
on the Balčug street
- (27) a. ot stanci-i Moskva/*Moskvy nominal proper name
 from station.FSG-GEN Moscow.FSG-NOM/GEN
from the station Moscow
- b. ot stanci-i Tixoreckaja/Tixoreckoj adjectival proper name
 from station.FSG-GEN Tixoreckaja.FSG-NOM/GEN
from the station Tixoreckaja

Empirical generalization: case-agreement depends on ϕ -congruence between the sortal and the proper name.

Number congruence is required, except for indeclinable last names (see Graudina et al. 1976)

For [+animate] proper names case-agreement is obligatory, regardless of gender congruence:

- (28) a. My govorili o russkom poète Cvetaevj/*Cvetaeva.
 we spoke about Russian-MSG-LOC poet-MSG-LOC Tsvetaeva.FSG-LOC/NOM
We spoke about the Russian poet Tsvetaeva.
- b. pro sobaku Trezora/*Trezor
 about dog-ACC Trezor-ACC=GEN/NOM
about the dog Trezor
- c. o kosmonavtax Tereškovej/*Tereškova i Gagarine/*Gagarin
 about astronauts-LOC Tereshkova.FSG-LOC/NOM and Gagarin.MSG-LOC/NOM
about the astronauts Tereshkova and Gagarin

Otherwise the degree of gender congruence is determined by the sortal:

- gender congruence not required (city, country, river names)
- gender congruence required (street names, syntactically complex city names with internal agreement)
- only with phi-congruent adjectival proper names (railway station, cape, peninsula, etc., names)

Note: there is notable cross-speaker variation in assigning different lexical-semantic categories of toponyms to these classes. I don't know whether there is any variation for individual proper names

Lack of familiarity makes case-agreement less likely.

Names of man-made objects do not allow case-agreement.

4.1. The relevance of phi-congruence

All declinable proper names show appropriate case-marking in argument positions.

Table 1: Case-agreement in close apposition

sortal	case-agreement	example
man-made objects	impossible	(25)
paths	impossible	
[+animate]	obligatory for number-congruent proper names	(28)
<i>stancija</i> 'station', etc.	with adjectival phi-congruent proper names only	(27)
<i>gorod</i> 'city', etc.	simple proper name: number congruence required	(29)
	complex proper name: phi-congruence required	(32)
<i>ulica</i> 'street', etc.	phi-congruence required	(26)

Reified quotations behave like plain quotations: no case-marking is possible

4.1.1. Case-agreement on the condition of number congruence

For **syntactically simplex** city and town names, as well as for names of countries and rivers, number congruence is required for case agreement but gender congruence is not:

- (29) a. v gorode Gagry/*Gagrax
 in city.MSG-LOC Gagry.PL-NOM/LOC
in the city of Gagry
- b. v gorode ?Tallinn/Tallinne
 in city.MSG-LOC Tallinn.MSG-NOM/LOC
in the city of Tallinn
- c. v gorode Moskva/Moskve
 in city.MSG-LOC Moscow.FSG-NOM/LOC
in the city of Moscow
- d. o strane Francija/Francii
 about country.FSG-LOC France.FSG-NOM/LOC
about the great country France
- e. o strane Kitaj/Kitaje
 about country.FSG-LOC China.MSG-NOM/LOC
about the great country China

Phi-congruent toponyms are more likely to agree.

Lack of case-agreement is more likely with recognizably foreign toponyms, which is usually associated with the lack of familiarity:

- (30) a. My govorili o malen'koj strane Gabon/??Gabone.
 we spoke about small-FSG-LOC country.FSG-LOC Gabon.MSG-NOM/LOC
We spoke about the small country Gabon.
- b. My govorili o malen'koj strane Birma/Birme.
 we spoke about great-FSG-LOC country.FSG-LOC Burma.FSG-NOM/LOC
We spoke about the small country Burma.

Assuming that the agreeing case results from concord, the proper name should agree with the sortal.

Question: if matching phi-features are enough for case-agreement, why can inanimate proper names fail to agree, unlike animate proper names?

Intuition: animate NPs must have the feature [α animate]. Inanimate NPs may (fail to) have it ([-animate] being the lexical default).

4.1.2. Case-agreement on the condition of phi-congruence

Street names and syntactically complex toponyms do not agree in case unless phi-congruent (Graudina et al. 1976:142):

- (31) a. na ulic-e Jakimank-a/Jakimank-e ✓ phi-congruent
 in street.MSG-LOC Yakimanka.FSG-NOM/LOC
on the Yakimanka street
- b. na ulic-e Balčug/*Balčug-e *phi-congruent
 in street.MSG-LOC Balčug.MSG-NOM/LOC
on the Balčug street
- (32) a. v poseleni-i Dolgij Most/*Dolgom Moste *phi-congruent
 in settlement. MSG-LOC Long Bridge.MSG-NOM/LOC
in the settlement of Dolgij Most (lit. Long Bridge)
- b. v gorod-e Belaya Cerkov/*Beloj Cerkvi *phi-congruent
 in city.MSG-LOC White Church.FSG-NOM/LOC
in the city of Belaya Cerkov (lit. White Church)
- c. v gorod-e Petropavlovsk-Kamčat-sk-ij/Petropavlovsk-e-Kamčat-sk-om
 in city.MSG-LOC Petropavlovsk-Kamčatka-ADJ-MSG-NOM/LOC
in the city of Petropavlovsk-Kamčatskij (lit. Petropavlovsk of Kamchatka)

Intuition: syntactically complex proper names containing agreeing modifiers necessarily bear formal gender features (to enable agreement internally to the proper name)

Syntactically or morphologically complex foreign toponyms always appear in the nominative case, with the exception of the pattern "city on the river":

- (33) a. v gorode Santa Barbara/*Santa Barbare
 in town.MSG-LOC Santa Barbara.FSG-NOM/LOC
in the town of Santa Barbara
- b. v gorode Frankfurte na Majne/Frankfurt na Majne
 in town.MSG-LOC Frankfurt am Main.MSG-NOM/LOC
in the city of Frankfurt am Main

Intuition: examples like (33b), while syntactically complex, can be clearly recognized to not involve agreement precisely due to their syntax (PPs do not agree).

Obligatory phi-congruence also constrains case-agreement with some other lexical-semantic classes of proper names, which we hypothesize to bear lexically specified gender features

4.1.3. Case agreement with phi-congruent adjectival proper names only

For some categories of proper names case agreement is possible only with **morphologically adjectival** toponyms on the condition of both gender and number congruence:

- (34) a. do stancii Bologoe/*Bologogo
 until station.FSG-GEN Bologoe.NSG-NOM/GEN
until the station Bologoe
- b. na stancii Moskva/*Moskvy
 on station.FSG-GEN Moscow.FSG-NOM/GEN
on the station Moscow
- c. na stancii Tixoreckaja/Tixoreckoj
 on station.FSG-GEN Tixoreckaja.FSG-NOM/GEN
on the station Tixoreckaja

Can this be a case of obligatory extraposition?

Most likely not, as in toponyms involving adjectives extraposition seems ungrammatical:

- (35) a. na Krasnoj ploščadi
 on Red-FSG-LOC Square.F-LOC
on the Red Square
- b. *na ploščadi Krasnoj
 on Square.F-LOC Red-FSG-LOC
- (36) a. na Nevskom (prospekte)
 on Nevsky-MSG-LOC avenue.M
on the Nevsky (Prospekt)
- b. *na prospekte Nevskij/Nevskom
 on avenue.M-LOC Nevsky-MSG-NOM/LOC

An incomplete list of such proper names includes boroughs (*mestečko*), villages (*selo*), ports, lakes, bays, volcanoes (*vulkan*, *sopka*), mountains, planets and railway stations. Prescriptive grammars may insist that case-agreement is impossible with such proper names or include in it islands, republics, etc. Thus toponyms preceded by the common nouns *aúl* ‘a village in the Caucasus and Central Asia’ and *kišlák* ‘a village in Central Asia’ are claimed to never agree for case, but this most likely is due to the fact that the names of such villages are extremely unlikely to be adjectival: when they are, case-agreement becomes possible on the condition of phi-congruence:

- (37) a. v kišlake/aule Čimgan/*Čimgane nominal phi-congruent
 in kishlak/aul.MSG-LOC Northern.MSG-NOM/-LOC
in the kishlak/aul Čimgan
- b. v kišlake/aule Severnom/Severnyj adjectival phi-congruent
 in kishlak/aul.MSG-LOC Northern.MSG-LOC/NOM
in the kishlak/aul Severnyj

Intuition: the difference between adjectives and nouns is that the former must have unvalued phi-features (and their declension class is determined by gender)

4.2. The syntax of phi-congruence

Hypothesis: phi-feature agreement always triggers case-agreement

- phi-feature agreement ⇒ obligatory case-agreement
- phi-feature congruence ⇒ potential case-agreement

The core intuition that we want to capture is that phi-feature congruence **can be a necessary condition** for case-agreement, but it doesn't have to be.

Formal tool: valuation of inherent phi-features for proper names in function of their lexical-semantic category

Intuition: in close apposition the proper name appears in the minimal syntactic structure that is allowed

4.2.1. Adjectival proper names

Core insight: adjectival proper names cannot not have unvalued gender and animacy features (except when they are nominalized, cf. null-derived deadjectival nouns: *zapjataja* 'comma.F', *portnoj* 'tailor.M'):

- (38) a. o russk-om poët-e Matve-ev-oj adjectival proper name
 about Russian-MSG.LOC poet.M-LOC Matvej-POSS-F.LOC
 about the Russian poet Matveeva
- b. v gorode Grozn-om nominal (deadjectival) proper name
 in city.M-LOC Fearsome-MSG.LOC
 in the city of Grozny

We run here into a very interesting issue of the formal interaction between proper names and their inflectional and derivational morphology

We generally assume that nouns are stored in the lexicon without inflectional morphology. With proper names, however, this inflectional morphology can clearly determine not only the pronunciation, but also the reference: *Puškino* (the village) is clearly distinct from *Puškin* (the city); adjectival last names, on the other hand, do not seem to have this property

DM: lexicon vs. Encyclopedia

Hypothesis: adjectival proper names are introduced with their inflectional morphology (and therefore with a valued gender feature). Their animacy feature, however, can be still unvalued and trigger agreement with the sortal. As a result, full syntactic agreement can be established between the sortal and the proper name

4.2.2. Nominal proper names

Normally Russian proper names are assigned gender on the basis of their declension class:

- (39) a. I pered mladšuju stolicej Pomerkla staraja Moskva
 and before younger capital dimmed-FSG old-FSG Moscow
 The old Moscow waned before the younger capital.
- b. Gordyj Sankt-Peterburg prevratilsja v obydennoj Leningrad.
 haughty-MSG St. Petersburg changed-MSG in pedestrian-MSG Leningrad
 The haughty St. Petersburg turned into a pedestrian Leningrad.

However, indeclinable proper names get their gender from their lexical semantics (in function of the appropriate sortal)

The two gender assignment strategies should not lead to a contradiction

Since two phi-features, animacy and gender, are involved, it seems natural to assume that one or both of them can remain unvalued

See also Bobaljik and Zocca 2011, Merchant [to appear] for optional gender specification on animate nouns

We have the following three patterns of case-agreement for (singular) nominal proper names to account for:

- no agreement: villages, mountains, volcanoes, etc.: lexically specified for gender; animacy either lexically specified ([-animate]) or absent altogether
- agreement possible with matching gender: street names, complex city names: the gender feature lexically specified, the animacy feature unvalued
- agreement possible regardless of gender: names of cities, towns, countries, rivers, etc.: the gender feature is absent, the animacy feature is unvalued
- agreement required regardless of gender: names of animate individuals: only their gender feature is introduced unvalued

Question: why is lack of case-agreement impossible for [+animate] proper names?

- incongruence in [α animate] is impossible: this feature is always interpreted
- incongruence in [α gender] is possible but overridden by the natural gender at the DP level
- total lack of phi-features is not an option

Is there any independent evidence for partial phi-feature specification of proper names?

4.3. German proper names

Moltmann 2013 distinguishes several categories for proper names in German:

- names of people: no overt article in standard German, plural anaphora possible, d-series in the relative pronoun choice
- names of churches and palaces: no overt article, plural anaphora possible, d-series in the relative pronoun choice (i.e., just like names of people, but inanimate)
- most toponyms (cities, villages, countries, continents, churches, palaces): no overt article, plural anaphora impossible, w-series in the relative pronoun choice
- names of mountains, lakes, temples: obligatory definite article (the gender of the corresponding sortal), d-series in the relative pronoun choice

Hypothesis: there is no need to postulate a hidden sortal: presence or absence of pre-specified formal gender is enough

German doesn't have declensional classes, so the gender of common nouns is not predictable from the surface form (though see Lang 1976, Corbett 1991:84-86 for references on gender assignment rules in German)

The gender of proper names is more complicated, but at least the following seems true:

- proper names denoting females are feminine, while proper names denoting males are masculine
- some proper names have a lexically fixed gender (e.g., *der Nil* 'the Nile' (M), *die Wolga* 'the Volga' (F))
- most [-animate] proper names have their gender fixed by semantic analogy (e.g., rivers, mountains, lakes, temples are masculine unless specified otherwise).

Hypothesis: just like in Russian, in German proper names of animate entities or cities are not specified for gender; the same is true for names of palaces and churches

The feature [α animate] may be valued, but does not have to be

Two more hypotheses:

- the overt definite article in standard German tracks inherent gender specification on proper names
- the w-series of relative pronouns is used in the absence of any phi-features

The difference between names of humans and names of cities comes from the fact that at the DP level the former acquire both gender and animacy from real-world reference

5. PUTTING IT ALL TOGETHER

Hypothesis: mixed agreement arises in a close apposition structure, where \mathcal{K} plays the role of the sortal

It introduces the feature [minor], but no n° . In order for the resulting structure to be nominal, \mathcal{K} must enter full syntactic agreement with the NP it combines with (and as a result inherit its nominal features). This is only possible for *-a-* nouns (which do have a [minor] feature due to their declension class)

The so-called hybrid nouns end up non-nominal in the presence of \mathcal{K} and therefore can only appear in the nominative case position (where they are not assigned case), similarly to other case-deficient elements (Testelefs 2013)

Why no full theory?

Because the notion of a declension class ends up as something of an oversimplification:

	N	M, inanimate		M, animate	F
nominative	<i>pal't-išk-o</i>	<i>dom-išk-o</i>	<i>?dom-išk-a</i>	<i>mal'č-išk-a</i>	<i>služb-išk-a</i>
accusative	o	o	*u	u	
genitive	a	?a	i	i	
dative	u	?u	e	e	
instrumental	om	om	oj	oj	
locative			e		

Similar effects arise for the diminutive suffix *-išč-* and the augmentative suffix *-in-* that also retain the gender of the source noun

In other words, it doesn't seem like the declension class is in fact a category pre-established in the lexical entry

To be continued...

6. BIBLIOGRAPHY

- Allan, Edward. 1976. Dizi. In *The Non-Semitic Languages of Ethiopia*, ed. by Lionel Bender, 377-392. East Lansing, Michigan: African Studies Center, Michigan State University.
- Asarina, Alya. 2008. Gender and adjectives in numeral constructions in Russian. Ms., MIT.
- Bobaljik, Jonathan, and Cynthia Levart Zocca. 2011. Gender markedness: the anatomy of a counter-example. *Morphology* 21:141-166.

- Chirsheva, Galina. 2009. Gender in Russian-English code-switching. *International Journal of Bilingualism* 13:63-90.
- Corbett, Greville. 1980. Animacy in Russian and other Slavonic languages: where syntax and semantics fail to match. In *Morphosyntax in Slavic*, ed. by Catherine V. Chvany and Richard D. Brecht, 43-61. Columbus, Ohio: Slavica Publishers.
- Corbett, Greville G. 1979. The agreement hierarchy. *Journal of Linguistics* 15:203-224.
- Corbett, Greville G. 1991. *Gender*. Cambridge Textbooks in Linguistics. Cambridge: Cambridge University Press.
- Corbett, Greville G. 2006. *Agreement*. Cambridge: Cambridge University Press.
- Crockett, Dina B. 1976. *Agreement in Contemporary Standard Russian*. Bloomington, Indiana: Slavica Publishers, Inc.
- De Clercq, Karen 2008. Proper names used as common nouns in Belgian Dutch and German. In *Linguistics in the Netherlands*, vol. 25, ed. by Bert Botma and Marjo van Koppen, 63-74. Amsterdam: John Benjamins.
- Dixon, R. M. W. 1982. *The Dyirbal Language of North Queensland*. Cambridge: Cambridge University Press.
- Elbourne, Paul. 2002. Situations and Individuals, Doctoral dissertation, MIT.
- Fraurud, Kari. 1999. Proper names and gender in Swedish. In *Gender in Grammar and Cognition*, ed. by Barbara Unterbeck and Matti Rissanen, 167-219. The Hague: Mouton de Gruyter.
- Fuller, Janet M., and Heike Lehnert. 2000. Noun phrase structure in German-English codeswitching: Variation in gender assignment and article use. *International Journal of Bilingualism* 4:399-420.
- Graudina, Ljudmila Karlovna, Viktor Aleksandrovič Ickovič, and Lia Pavlovna Katlinskaja. 1976. *Grammatičeskaja pravil'nost' russkoj reči. Stilističeskij slovar' variantov*. Moscow: Nauka.
- Harmer, L. C., and F. J. Norton. 1957. *A Manual of Modern Spanish*. London: University Tutorial Press.
- Heim, Irene, and Angelika Kratzer. 1998. *Semantics in Generative Grammar*. Oxford: Blackwell.
- Hickey, Raymond. 1999. The phonology of gender in Modern German. In *Gender in Grammar and Condition*, ed. by Matti Rissanen and Barbara Unterbeck, 621-663. Berlin: Mouton de Gruyter.
- Iomdin, L.L. 1980. O russkix suščestvitel'nyx tak nazyvaemogo obščego roda. *Izvestija Akademii Nauk SSSR. Serija literatury i jazyka* 39:456-461.
- Ionin, Tania, and Ora Matushansky. 2006. The composition of complex cardinals. *Journal of Semantics* 23:315-360.
- Kihm, Alain. 2005. Noun class, gender, and the lexicon-syntax-morphology interfaces: a comparative study of Niger-Congo and Romance languages. In *The Oxford Handbook of Comparative Syntax*, ed. by Guglielmo Cinque and Richard S. Kayne, 459-512. Oxford: Oxford University Press.
- Kramer, Ruth. 2012. Gender in Amharic: a morphosyntactic approach to natural and grammatical gender. Ms., Georgetown University.
- Lang, Adrienne. 1976. The semantic base of gender in German. *Lingua* 40:55-68.

- Leisiö, Larisa. 2001. *Morphosyntactic convergence and integration in Finland Russian*. Tampere: University of Tampere.
- Matushansky, Ora. 2013. Gender confusion. In *Diagnosing Syntax*, ed. by Lisa L.-S. Cheng and Norbert Corver, 271-294. Oxford: Oxford University Press.
- Matushansky, Ora, and E.G. Ruys. 2012. Numeral NPs, to a degree. Paper presented at *RALFe 2012*, Université Paris VIII, Saint-Denis, November 29-30, 2012.
- McNally, Louise, and Gemma Boleda. 2004. Relational adjectives as properties of kinds. In *Empirical Issues in Formal Syntax and Semantics 5*, ed. by Patricia Cabredo Hofherr and Olivier Bonami.
- Mel'čuk, Igor. 1980a. Animacy in Russian cardinal numerals and adjectives as an inflectional category. *Language* 56:797-811.
- Mel'čuk, Igor. 1980b. O padeže čislovogo vyraženiya v russkix slovosočetańx tipa (*bol'she*) na dva mal'čika ili po troe bol'nyx. *Russian Linguistics* 5:55-74.
- Merchant, Jason. [to appear]. Gender mismatches under nominal ellipsis. *Lingua*.
- Moltmann, Friederike. 2013. Names, sortals, and the mass-count distinction. Ms., Université Paris I.
- Neeleman, Ad. 2008. Subset controllers. Paper presented at *Words don't come easy*, University of Verona, November 17-18, 2008.
- Nesset, Tore. 2001. How pervasive are sexist ideologies in grammar? In *Language & Ideology: Cognitive Theoretical Approaches*, ed. by René Dirven, Bruce Hawkins and Esra Sandikcioglu, 197-227. Amsterdam/Philadelphia: John Benjamins.
- Nikunlassi, Ahti. 2000. On gender assignment in Russian. In *Gender in Grammar and Cognition*, ed. by Barbara Unterbeck and Matti Rissanen, 771-791. Berlin: Mouton de Gruyter.
- Percus, Orin. 2011. Gender features and interpretation: a case study. *Morphology* 21:167-196.
- Pesetsky, David. [to appear]. *Russian Case Morphology and the Syntactic Categories*. Cambridge, Massachusetts: MIT Press.
- Picallo, M. Carme. 2005. Some notes on the grammatical gender and *l*-pronouns. *Universität Konstanz Fachbereich Sprachwissenschaft, Arbeitspapier* 19:107-121.
- Poplack, Shana, Alicia Pousada, and David Sankoff. 1982. Competing influences on gender assignment: variable process, stable outcome. *Lingua* 57:1-28.
- Poplack, Shana, David Sankoff, and Christopher Miller. 1988. The social correlates and linguistic processes of lexical borrowing and assimilation. *Linguistics* 26:47-104.
- Rabeno, Angela, and Lori Repetti. 1997. Gender assignment of English loan words in American varieties of Italian. *American Speech* 72:373-380.
- Rothstein, Robert A. 1980. Gender and reference in Polish and Russian. In *Morphosyntax in Slavic*, ed. by Catherine V. Chvany and Richard D. Brecht, 79-97. Columbus, Ohio: Slavica.
- Rozental, D. E., E. V. Džandžakova, and N. P. Kabanova. 1998. *Spravočnik po pravopisaniju, proiznošeniju, literaturnomu redaktirovaniju*. Moscow: CeRo.
- Sauerland, Uli. 2004. A comprehensive semantics for agreement. Ms., ZAS, Berlin.
- Sauerland, Uli, and Paul Elbourne. 2002. Total reconstruction, PF movement, and derivational order. *Linguistic Inquiry* 33:283-319.

- Sitchinava, D. V. 2011. Rod. Materialy dlja proekta korpusnogo opisanija ruskoj grammatiki (Rusgram, <http://rusgram.ru>). Ms., Moscow.
- Steriopolo, Olga, and Martina Wiltschko. 2008. Distributed GENDER hypothesis. Paper presented at *FDSL 7.5*, Independent University of Moscow, December 6-8, 2008.
- Steriopolo, Olga, and Martina Wiltschko. 2010. Distributed GENDER hypothesis. In *Formal Studies in Slavic Linguistics. Proceedings of Formal Description of Slavic Languages 7.5*, ed. by Gerhild Zybatow, Philip Dudchuk, Serge Minor and Ekaterina Pshehotskaya. *Linguistik International* 25, 155-172. Frankfurt: Peter Lang.
- Švedova, N. Ju. ed. 1980. *Russkaja grammatika*. Moscow: Nauka.
- Testeleets, Yakov. 2013. Case deficient elements and the direct case condition in Russian. Ms., Moscow.
- Violin-Wigent, Anne. 2006. Gender assignment to nouns codeswitched into French: Observations and explanations. *International Journal of Bilingualism* 10:253-276.
- Zaliznjak, A. A. 1967. *Russkoe imennoe slovoizmenenie*. Moscow: Nauka.