

ON THE SEMANTICS OF INANIMATE GENDER

1. RUSSIAN GENDER: AN OVERVIEW

The Russian gender system is a mixed one (Crockett 1976, Corbett 1991, Doleschal 1996, etc.) combining semantic (male/female) and formal (declension class) criteria for gender assignment:

- (1) Russian gender assignment rules
 - a. nouns denoting males are masculine
 - b. nouns denoting females are feminine
 - c. declinable nouns of the declension class *-ŕ-* are masculine
 - d. declinable nouns of the declension classes *-a-* and *-ŕ-* are feminine
 - e. declinable nouns of the declension class *-o-* are neuter
 - f. animate indeclinable nouns are masculine (with some exceptions)
 - g. inanimate indeclinable nouns are neuter (again, with some exceptions)
 - h. the gender of indeclinable initialisms (not acronyms) is determined by the gender of the head

There is nothing special about such a system: for gender-differentiated individuals semantic gender trumps formal criteria and otherwise gender is determined by the declension class. There are exceptions here and there: for instance animate class *-a-* nouns can be underlyingly formally feminine or masculine without being specified as female or male in their semantics (Iomdin 1980), and human-denoting NPs may trigger agreement based not on their formal gender, but on that of their denotatum (so-called semantic agreement, Corbett 1979). Another class of exceptions that I would like to focus here is the gender of inanimate indeclinables. I will argue that their gender assignment can be based on semantic analogy offering evidence for semantic interpretability of formal gender and for computation of gender on the fly, with gender regarded as a syntactic property on the probe rather than the goal.

2. THE GENDER OF RUSSIAN INDECLINABLE NOUNS

Indeclinable nouns are borrowings and loanwords that do not fit into the Russian declension classes. Thus (singular) nouns ending in [i] and [u] are always indeclinable, nouns ending in [o] (e.g., *pal'tó* ‘coat’) and [e] (*kašné* ‘headscarf’) often are, and [a]-final and C-final animate nouns are if they denote males and females, respectively (although [a]-final surnames (e.g., *D'urá* ‘Duras’) are often indeclinable even if they denote females). Setting aside indeclinable nouns that denote sexed individuals, whose assigned gender is based on their semantic gender (feminine if female, masculine otherwise), how is gender assigned to inanimate indeclinable nouns?

It turns out (see Doleschal 1996, Murphy 2000, Wang 2014 for recent discussions in English) that indeclinable nouns are assigned gender by semantic criteria, and exceptions to these criteria are themselves semantically motivated. Starting with the general rules detailed in (1), animate indeclinables are assigned feminine or masculine gender, the latter being the default. Thus the noun *šimpanze* ‘chimpanzee’ can be in principle masculine or feminine, but the NP containing it in (2a) denotes a female chimpanzee so the feminine agreement can be viewed

Acknowledgments: Hagit’s personality, her drive, her curiosity, her interest and daring to be different and to be herself have inspired generations of students no less than her theories and her feminism, and this work on gender is but a small tribute to her fascinating presence in the field.

Examples marked as RNC come from The Russian National Corpus (ruscorpora.ru), other sources are marked as GIK for Graudina, Ickovič and Katlinskaja 1976, E for Edwardsson 1968, RDK for Rozental, Džandžakova and Kabanova 1998, and www marks examples found online.

as semantic.¹ Inanimate indeclinables would appear to be assigned inherent non-interpretible gender. While most indeclinable nouns are neuter, *ivasi* ‘South American pilchard (Sardinops sagax)’ in (2b) is feminine (cf., e.g., Graudina, Ickovič and Katlinskaja 1976:78 (henceforth, GIK) and Doleschal 1996:122), although some variation is possible. Such variation is even more prominent for masculine indeclinables, which are far more likely than feminine ones to also agree in the neuter (cf. GIK:79-82):

- (2) a. [...] umerla estestvennoj smert'ju legendarnaja šimpanze Uošo. RNC
 died.FSG natural death legendary.FSG chimpanzee Washoe
 The legendary chimpanzee Washoe died a natural death...
- b. Čelovek — [...] ne besslovesnaja ivasi v korotkix štanax. RNC
 human NEG dumb pilchard in short pants
 Homo sapiens is not some dumb pilchard in short pants.

These inherent gender specifications are usually claimed to be motivated by the hypernyms of the indeclinable nouns in question. Thus *ivasi* in (2b) is a kind of *sel'd'* ‘pilchard’, which is feminine (declension class *-ŭ*). Likewise, *aven'u* ‘avenue’ is a kind of *ulica* ‘street’, which is feminine (declension class *-a-*), *kol'rabi* ‘kohlrabi’ is a kind of *kapusta* ‘cabbage’ (declension class *-a-*), etc., and most (not all) indeclinable dance names are or can be masculine because the word *tanec* ‘dance’ is masculine.² This type of gender assignment is subject to variation as noted in Unbegaun 1947, Doleschal 1996, Murphy 2000, Wang 2014 and Baranova 2016, among many others.

The question arises of the status of such semantically based generalizations. Because they are non-systematic, they can be claimed to result from social rather than linguistic factors. After all, formal gender can be a stored lexical property of a noun and the fact that there are some semantic generalizations about exceptions may reveal nothing about language itself. Far more telling are indeclinable toponyms, which, as we will now see, are productively assigned gender based on their semantics.

3. INDECLINABLE PROPER NAMES

While most Russian toponyms are declinable and, like common nouns, are assigned gender in function of their declension class, this route is closed for indeclinable toponyms, which are systematically assigned gender by semantic analogy. The algorithm is well-known and can be found in GIK, Doleschal 1996, Rozental, Džandžakova and Kabanova 1998 (henceforth, RDK), Murphy 2000, etc.): indeclinable toponyms are assigned the gender of their associated sortal. Thus cities are masculine (because the noun *gorod* ‘city, town’ is masculine by virtue of its declension class), states (*gosudarstvo*) or lakes (*ozero*) are neuter and mountains (*gora*) and rivers (*reka*) are feminine:

- (3) a. solnečnyj [gorod] Tbilisi RDK204
 sunny-MSG city.M Tbilisi
 the sunny city of Tbilisi

¹ Semantic agreement is subject to the Agreement Hierarchy (Corbett 1979 and later work), making it possible for a semantically agreeing noun to also trigger formal agreement on a closer target, as in (i). For animal-denoting nouns such mixed agreement seems impossible.

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

² Indeclinable ethnicity nouns (e.g., *komi* ‘Komi’) are probably underlyingly indeclinable adjectives, which are masculine as linguonyms (cf. the masculine *russkij* ‘Russian’) and are assigned semantic gender as ethnonyms.

- 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

Because of this semantically driven gender assignment, one and the same proper name may be assigned different genders. Thus Somali may be assigned feminine if regarded as a country (*strana*, a feminine noun) or neuter if treated as a state (*gosudarstvo*, a neuter noun), as in (4), from Rozental, Džandžakova and Kabanova 1998, see also Crockett 1976). This apparent optionality may, however, be an illusion caused by the fact that states and countries are usually felt to be co-extensional. Perhaps we are dealing with two different entities here, as in (5), where the state of Congo and the Congo River happen to share the same name:³

- (4) a. Somali prinjala s blagodarnost'ju gumanitarnuju pomošč'. RDK
Somali accepted-FSG with gratitude humanitarian aid
(The country of) Somali accepted the humanitarian aid with gratitude.
- b. V period krizisa Somali stradalo ot nexvatki prodovol'stviya.
in period crisis-GEN Somali suffered-NSG from lack victuals-GEN
During the crisis (the state of) Somali suffered from the lack of food.
- (5) a. gde katit svoi vody širokaja i polnovodnaja Kongo. www
where rolls self's waters wide.FSG and full-flowing.FSG Congo
where the wide and full-flowing Congo rolls its waters.
- b. Kongo stalo bel'gijskoj kolonij v načale 20-go veka.
Congo became.NSG Belgian colony in beginning 20th century.GEN
Congo became a Belgian colony in the beginning of the 20th century.

In other words, the gender of an indeclinable Russian toponym depends on what it denotes. The double-barreled question arises of the precise status of this generalization from the point of view of its localization both in cognition and in syntax.

3.1. Is the gender of indeclinable nouns linguistic?

The question of the gender of a given toponym, even such quite well-known ones as *Tbilisi* or *Somali*, is a frequent one and the general issue is discussed in many Russian grammars. If the semantic generalization above were inherent, would it also not be internalized? Can it be merely a prescriptivist rule with no standing in the language itself, like that of the collective nouns used for different animals in English?

I have run some searches online and in the Russian National Corpus (ruscorpora.ru) and asked informal questions of some five Russian speakers about gender agreement with indeclinable toponyms. The results that I got mostly coincide with the semantically motivated generalization above, but variation is also attested. Thus the city of *Great Zimbabwe* can agree as both the expected masculine (6a) (*gorod* 'city' is masculine) and the unexpected neuter (6b)), the opposite is true for *the sunny Peru*, which appears with masculine agreement

³ Gender assignment by semantic analogy is frequent in languages that have no declension classes or where the declension class does not determine the gender (see Herschensohn 1978 on French deadjectival nouns, Fraurud 1999 on toponyms in Swedish, Hickey 1999 on the gender of names of hotels, cinemas and cars in German; De Clercq 2008 on the gender of common nouns derived from proper names. Also tree-denoting nouns are feminine in Latin, irrespective of the declension class. I will not attempt to address these cases here.

as well as with the expected neuter of *gosudarstvo* ‘state’) and for *Somali*, which appears in RNC only with masculine and neuter adjectives (contrary to RDK).

- (6) a. Bol’šoj Zimbabwe (Velikij Zimbabwe)
big.MSG Zimbabwe great.MSG Zimbabwe
- b. Bol’šoe Zimbabwe
big.NSG Zimbabwe
Great Zimbabwe

For indeclinable river names I have found only the expected feminine agreement, though two of my informants when asked about the Alazani River conceded that maybe other options are possible and one actually preferred masculine. The generalization is therefore strong where it comes to rivers, but less so for states. Similar variability is observed for semantic agreement with human-denoting nouns: they do not allow agreement by the gender of the denotatum all to the same extent. Importantly, the alternative pattern (neuter for inanimates) is semantic as well, although it does not seem to be available as a default (or else no hesitations would be expected).

My tentative conclusion is therefore that the sortal-determined gender of indeclinables is linguistic to the same extent as semantic agreement is in general.

3.2. The syntax of gender features

Known toponyms like *Peru* are easily used by any competent speaker of Russian in a context that does not require gender agreement, though some of these speakers would falter when gender agreement is required (cf. Doleschal 1996:125). Yet any usual take on the syntax of nouns (and proper names) is that gender is a core lexical property of nouns that is available from the lexicon and merely used in syntax. Even if we concede that unknown nouns are not in the lexicon of a given speaker and thus their gender may also be unknown, we do not expect a difference in behavior in function of overt agreement, especially in approaches that locate gender on the N itself and even more so, on *n* (Kihm 2005, Lowenstamm 2007, Percus 2011, Kramer 2012, 2015, 2016, etc.). Indeed, if gender or noun class is part of what defines a noun in languages that have gender, we expect no differences between environments where gender agreement occurs and those where it does not.

It could be objected that these toponyms are not nouns *sensu stricto*. Indeed, the consensus is that proper names are DPs, and it could be hypothesized that indeclinable proper names are deficient syntactic objects like book or movie titles (see Testelets 2010, Kholodilova 2013, and Trubitsina 2017 for their behavior with respect to agreement and case-marking). There are, however, two objections to this approach. Firstly, book or movie titles that are nominal (and not coordinate structures) do not exhibit any special behavior with respect to agreement, nor do declinable unfamiliar toponyms, and so the special nature of indeclinable toponyms stems from their morphology. Secondly, the same agreement issues arise with indeclinable common nouns, even if the denotation is known. Thus the indeclinable *kolibri* ‘hummingbird’ has been attested with both the expected masculine (for animates) and the unexpected feminine (for the feminine hypernym *ptica* ‘bird’) agreement. Likewise, *suluguni* (a type of Georgian cheese, *syr* ‘cheese’ is masculine) can be used productively with full understanding in non-agreement contexts and may result in hesitation or in masculine or neuter in agreement contexts; other examples are too numerous to list.

Moreover, the issue is not limited to Russian. The gender of city names in French is no less thorny a question and grammars diverge on it wildly, as do individual speakers or texts. Two major empirical explorations of the matter, Hasselrot 1943/1944 and Edwardsson 1968, note a high degree of confusion and optionality in gender resolution for city names, with the same city name potentially yielding both feminine and masculine agreement within the same

sentence, as in (7), from Edwardsson 1968:271. Informal observations, both my own and discussed on the Internet, indicate that French speakers mostly try to avoid gender agreement altogether for city names.

- (7) **Fondée** en 1869, Kemi est **devenu** un centre industriel important. E271
founded.FSG in 1869, Kemi is become.MSG a center industrial important
Founded in 1869, Kemi became an important industrial center.

Even proper names that have an integrated gender-marked definite article may show gender disagreement (8):

- (8) a. le Caire, nostalgique et **orgueilleuse** E279
the.MSG Cairo, nostalgic and proud.FSG
Cairo, nostalgic and proud
- b. **Le** La Rochelle de mon enfance a disparu. Internet
the.MSG the.FSG Rochelle of my childhood has disappeared
The La Rochelle of my childhood has disappeared.

Some empirical generalizations are attested. Thus for the quantifier *tout* ‘all’ and adjectives *vieux* ‘old’ (in the sense of the older part of the city), *petit* ‘small’ and *grand* ‘big’ only masculine agreement is possible (Hasselrot 1943/1944:215, Edwardsson 1968). Conversely, the definite postnominal epithet (e.g., *Alger la blanche* ‘Algiers the White’) is claimed to permit feminine adjectives only, while elsewhere feminine agreement is preferred (Edwardsson 1968), rationalized to the feminine sortal *ville* ‘city, town’ (9). Indeed, Hasselrot 1943/1944 indicates (with no actual numbers) that for villages (a masculine noun in French) masculine is preferred.

- (9) a. Predicate past participle (including depictives and appositives):
Grand Larousse Encyclopédique (661): masculine 130 (20%), feminine 531 (80%)
Grand Larousse (2158): masculine 885 (41%), feminine 1273 (59%)
- b. Predicate adjectives (including those in apposition):
No corpus indicated (90): masculine 18 (20%), feminine 72 (80%).
- c. Nonrestrictive adjectives (*épithets de nature*):
No corpus indicated (44): masculine 13 (30%), feminine 31 (70%)
- d. Indefinite article (Hasselrot 1943/1944: masculine preferred)
No corpus indicated (90): masculine 60 (66%), feminine 30 (33%)

It is important to emphasize that in French the issue is not limited to a subset of toponyms, all city names exhibit this behavior, so the sortal-oriented agreement that we have observed in Russian for a morphologically defined group of inanimate proper names, in French is attested for the entire lexical-semantic class of city and village names, with no problems arising in non-agreement contexts.

Once again, it could be suggested that proper names do not form part of the lexicon. This argument, however, is very weak. Firstly, under most approaches they end up as DPs, so the issue is just postponed. Secondly, this assumption does not explain why agreement contexts should be different from non-agreement ones if the same syntactic object is eventually built.

4. SUMMARY AND CONCLUSIONS

We have shown that inanimate proper names can be underspecified for gender, which can then be computed from their denotation. Russian indeclinable toponyms are assigned gender based on their hypernym (sortal), though sometimes neuter or even masculine can surface in violation of this generalization. French city and village names, on the other hand, suffer from the competition between the general masculine for inanimates and semantic gender based on

their sortal, which means that they are mostly avoided altogether in agreement contexts. This suggests that gender is not an obligatory property of proper names, and this conclusion is also supported by the behavior of Russian indeclinable common nouns.

These data can be taken as evidence for the purely syntactic computation of gender: it is not part of either N or *n* but is instead computed on the head that enters into gender agreement with the relevant (extended) NP. This view has several advantages. Firstly, it can be extended to gender agreement with coordinated NPs, which are assigned gender on the basis of mixed formal and semantic criteria (see Lyskawa 2021 for an overview and references) and can also show ineffability in agreement contexts (see, e.g., Adamson and Anagnostopoulou 2021 in addition to Lyskawa 2021:275). Secondly, syntactic computation of gender for the purposes of agreement is essential for treating semantic agreement, which is triggered by semantic rather than formal properties of an NP and can target different levels of an NP (cf. (i) in fn. 1, which illustrates NP-internal mixed agreement). While it could be hypothesized that proper names should not be treated the same as common nouns, evidence from coordinated NPs and semantic agreement with animate nouns suggest that the issue is not limited to proper names.

Thirdly, the hypothesis that nouns and specifically proper names may enter syntax under- or unspecified for phi-features and gender is computed during syntactic derivation might pave the way for analyzing the role that phi-congruence plays in case-assignment to proper names introduced by a sortal (Matushansky 2021):

- (10) a. na ulice Jakimanka/Jakimanke ✓phi-congruent
 in street.FSG.LOC Yakimanka.FSG.NOM/LOC
 on the Yakimanka street
- b. na ulice Balčug/*Balčuge *phi-congruent
 in street.FSG.LOC Balčug.MSG.NOM/LOC
 on the Balčug street

As the contrast in (10) shows, a masculine proper name appears in the default nominative case when preceded by the feminine sortal *ulica* ‘street’, while a feminine proper name also has the option of agreeing in case with the sortal. Assuming that case-agreement is parasitic on agreement in gender (Matushansky 2021) requires the assumption that proper names can enter syntactic computation without specified gender features.

A separate issue is that of what semantically determined inanimate gender tells us about the semantics of gender features in general. The fact that a formal feature of the sortal is used in semantic analogy strongly indicates that it is not as semantically neutral as usually assumed.

Summarizing, gender-deficient proper names, both in Russian and in French, cast doubt on the standard perception of gender as a property of a noun given in the lexicon or established for *n*: if such were the case, we would see no difference in behavior in agreement and non-agreement contexts. Syntactic computation of gender might, on the other hand, offer support for an exoskeletal model, like that argued for in Borer 2005a, b.

5. REFERENCES

- Adamson, Luke, and Elena Anagnostopoulou. 2021. Interpretability and gender features in coordination: Evidence from Greek. Paper presented at *Agreement in multivaluation constructions*, Frankfurt University (Online), May 19-20, 2021.
- Baranova, Anna S. 2016. Отражение вариативности родовой характеристики несклоняемых существительных в современных словарях [Variation in gender characteristics of indeclinable nouns in contemporary dictionaries]. *Вестник Санкт-Петербургского Университета. Язык и литература* [Vestnik of Saint Petersburg University. Language and Literature] 13, 45-57.

- Borer, Hagit. 2005a. *Structuring Sense 1*. Oxford: Oxford University Press.
- Borer, Hagit. 2005b. *Structuring Sense 2*. Oxford: Oxford University Press.
- 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.
- 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.
- Doleschal, Ursula. 1996. Gender assignment revisited. In *Gender in Grammar and Cognition: I: Approaches to Gender*, ed. by Unterbeck Barbara, Rissanen Matti, Nevalainen Terttu and Saari Mirja, 117-166: De Gruyter Mouton.
- Edwardsson, Roland. 1968. Le genre des noms de villes en français. *Studia Neophilologica* 40, 265-316.
- 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.
- Graudina, Ljudmila Karlovna, Viktor Aleksandrovič Ickovič, and Lia Pavlovna Katlinskaja. 1976. *Грамматическая правильность русской речи. Стилистический словарь вариантов [Grammatical correctness of the Russian speech. A stylistic dictionary of variants]*. Moscow: Nauka.
- Hasselrot, Bengt. 1943/1944. Le genre des noms de villes en français. *Studia Neophilologica* 16, 201-223.
- Herschensohn, Julia. 1978. Deep and surface nominalized adjectives in French. *Linguistic Inquiry* 9, 135-137.
- Hickey, Raymond. 1999. The phonology of gender in Modern German. In *Gender in Grammar and Condition*, ed. by Barbara Unterbeck and Matti Rissanen, 621-663. Berlin: Mouton de Gruyter.
- Iomdin, L.L. 1980. О русских существительных так называемого обščего рода. *Izvestija Akademii Nauk SSSR. Serija literatury i jazyka* 39, 456-461.
- Kholodilova, Maria. 2013. Свойства вершины у элементов названий в русском языке [Head properties in elements of names in Russian]. Paper presented at *Русский язык: конструкционные и лексико-семантические подходы [Russian language: constructive and lexical-semantic approaches]*, Saint-Petersbourg, 12–14.09.2013.
- 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.
- Kramer, Ruth. 2015. *The Morphosyntax of Gender*. Oxford: Oxford University Press.
- Kramer, Ruth. 2016. The location of gender features in the syntax. *Language and Linguistics Compass* 10, 661-677.

- Lowenstamm, Jean. 2007. On little *n*, $\sqrt{\quad}$, and types of nouns. In *The Sounds of Silence: Empty Elements in Syntax and Phonology*, ed. by Jutta Hartmann, Veronika Hegedüs and Henk Van Riemsdijk. Amsterdam: Elsevier.
- Lyskawa, Paulina. 2021. Coordination without Grammar-Internal Feature Resolution. Doctoral dissertation, University of Mary.
- Matushansky, Ora. 2021. *Die Stadt Leipzig* and other fun places in Russian. In *FDSL 14. Leipzig* (Online).
- Murphy, Dianna L. 2000. The gender of inanimate indeclinable common nouns in Modern Russian. Doctoral dissertation, Ohio State University.
- Percus, Orin. 2011. Gender features and interpretation: a case study. *Morphology* 21, 167-196.
- Rozental, D. E., E. V. Džandžakova, and N. P. Kabanova. 1998. *Spravočnik po pravopisaniju, proiznošeniju, literaturnomu redaktirovaniju*. Moscow: ČeRo.
- Testelefs, Yakov. 2010. Case deficient elements and the direct case condition in Russian. *Язык и речевая деятельность [Language and language behavior]* 9, 126-143.
- Trubitsina, Maria. 2017. Согласование в конструкциях с названиями в русском языке [Agreement in constructions with names in Russian]. *Труды Института русского языка им. В. В. Виноградова [Proceedings of the V.V. Vinogradov Russian Language Institute]* 13, 295-303.
- Unbegaun, Boris Ottokar. 1947. Les substantifs indéclinables en russe. *Revue des Études Slaves*, 130-145.
- Wang, Qiang. 2014. Gender Assignment of Russian Indeclinable Nouns, MA thesis.