Essay

Micro-variation and Predictability in Gender Assignment for Russian Nouns Ending in Palatalized Consonants

Natalia Mitrofanova, Olga Urek & Marit Westergaard

Department of Language and Culture, UiT The Arctic University of Norway

natalia.mitrofanova@uit.no, olga.urek@uit.no, marit.westergaard@uit.no

1. Introduction

Nouns ending in palatalized consonants in Russian pose a challenge to learners and linguists alike, in that these nouns may be either feminine or masculine. In this short paper, we highlight the similarity between Tore Nesset's work using Russian corpus data and our own experimental work investigating gender assignment to nonce nouns by Russian children and adults. We conclude with proposing some possible future research directions that could build on the cumulative results of Nesset's and our studies.

2. Background

In a number of empirical studies, our Tromsø research group on language acquisition and multilingualism – which has had several names over the past decade, e.g. LAVA and AcqVA (both acronyms for (Language) Acquisition Variation & Attrition) and currently Center for Language, Brain & Learning (C-LaBL) – we have documented that both monolingual and bilingual Russian-speaking children seem to learn the rules for gender assignment in Russian quite early, not only in cases where assignment is rule-governed, but also in cases where the gender cues are ambiguous (e.g. Rodina & Westergaard 2012, 2015, Mitrofanova et al. 2018, and Rodina et al. 2020). The Russian gender system is generally considered to be rule-based, in that the majority of nouns can be accounted for by the following three general rules: Nouns ending in consonants are typically masculine (e.g. *dom* 'house'),

Slovo 66:1 (2026), 118–130. DOI https://doi.org/10.33063/slovo.v66i1.1227
CC BY 4.0

nouns ending in *-a* are typically feminine (e.g. *stená* 'wall'), and nouns ending in *-o* are typically neuter (e.g. *oknó* 'window'); see e.g. Hockett (1958), Corbett (1991). However, there are two classes of nouns in Russian where the final segment is an ambiguous gender cue; these are nouns ending in an unstressed vowel (schwa), which make them (in speech) ambiguous between feminine and neuter (e.g. *kníga* 'book', which is feminine, and *óblako* 'cloud', which is neuter) and nouns ending in palatalized consonants, which may be either masculine or feminine (e.g. *den'* 'day', which is masculine and *ten'* 'shadow', which is feminine).[1]

In our above-mentioned gender studies on monolingual and bilingual Russian-speaking children, the experimental stimuli were always existing nouns that the children were expected to know. This meant that it was unclear whether the children had actually acquired the gender rules of Russian or whether they had simply learned the gender of individual nouns in an item-based fashion (which must be a possible learning trajectory, given that there are gendered languages where there are few, if any, reliable gender rules, e.g. Norwegian; see e.g. Rodina & Westergaard 2013, 2021, Westergaard & Rodina 2016, Busterud et al. 2019, Lohndal & Westergaard 2021, cf. Urek et al. 2022). In order to investigate this question, Mitrofanova et al. (2021) carried out a study where monolingual and German-Russian bilingual children were exposed to novel nouns combined with pictures of inanimate items taken from the Novel Object and Unusual Name Database (NOUN, Horst & Hout 2016), e.g. *punip*, *bul'ga* and *glamo*, for the three general gender assignment rules. In addition, we included five items in the two ambiguous rule categories, i.e., nouns ending in unstressed vowels (e.g. *pruz-/schwa/*) and nouns ending in palatalized consonants (e.g. *dron'* and *knov'*). Our prediction was that, if the children had acquired the three general rules, they would prefer feminine in the nouns ending in unstressed vowels (as feminine is more frequent than neuter, and also tends to be acquired earlier , also with existing nouns; see e.g. Gvozdev 1961 or Ceitlin 2009). We further predicted that they would prefer masculine for nouns ending in palatalized consonants, given that nouns ending in plain consonants are typically masculine, and masculine is generally considered to be the default gender in Russian (Corbett 1991). In the next section, we focus on this latter subclass of Russian nouns.

[1] In addition, there are other exceptional noun classes in Russian where semantic rules override morphology, e.g. nouns ending in *-a* referring to males, which are always masculine (e.g. *papa* 'daddy') or hybrid nouns of the so-called 'common gender', which may have either masculine or feminine agreement depending on the biological gender of the referent (e.g. *vrač* 'doctor').

3. What do Russian Children Do with Nouns Ending in Palatalized Consonants?

The results of the Mitrofanova et al (2021) study showed that both the monolingual (N=87, mean age 5.2) and bilingual (N=64, mean age 6.16) Russian-speaking children generally obeyed the three general rules for gender assignment, producing masculine gender for novel nouns ending in consonants, feminine for novel nouns ending in (stressed) -*a*, and neuter for novel nouns ending in (stressed) -*o*. As predicted, they also defaulted to feminine with nouns ending in an unstressed vowel. But, contrary to our prediction and quite surprising to us, the children generally preferred feminine gender instead of defaulting to masculine with the novel nouns ending in palatalized consonants. If we had read Tore Nesset's (2003) article on the microvariation in gender assignment with Russian nouns ending in palatalized consonants, we would not have been so surprised. More about that below.

In order to check that this result was reliable, we carried out a follow-up study with only monolingual Russian children (N=85, age 3-7) with 13 novel items, ending in different palatalized consonants. This time, the results showed a clear distinction between different types of palatalized consonants, with a preference for masculine with some items, and a preference for feminine with other items. This is shown in the following figure (adapted from Mitrofanova et al. 2021: 61).

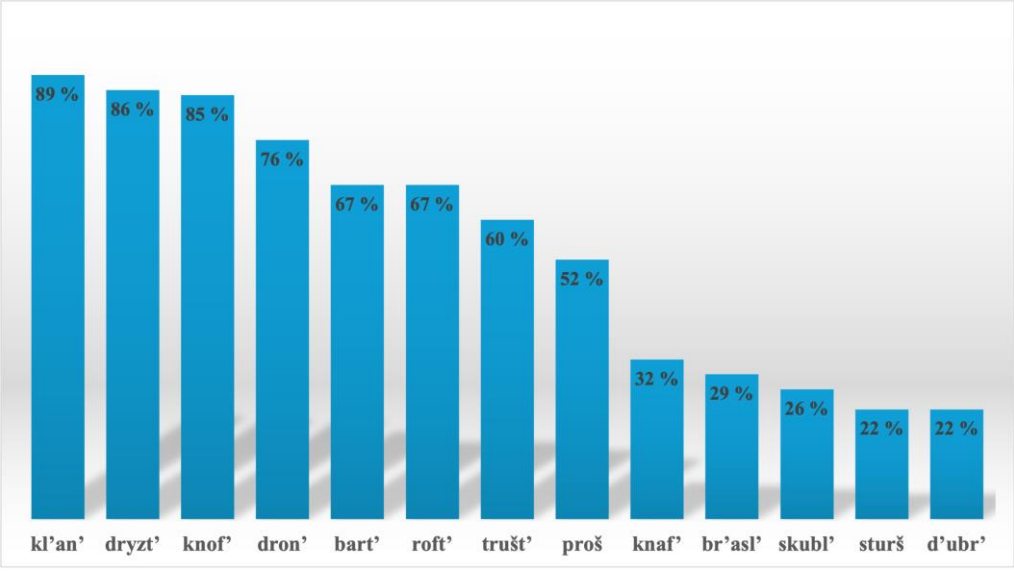


Figure 1. Percentage of feminine agreement with an extended set of nonce nouns ending in a palatal by monolingual Russian children (n = 85, 3–7-year-olds).

The five nonce nouns included in the original Mitrofanova et al. (2021) study were *proš, dron', knof', dryzt'*, and *kl'an'*, and as we can see from the figure above, four of these are clearly preferred with feminine gender (76-89%), and the fifth (*proš*) has a more or less equal distribution of masculine and feminine. Thus, the follow-up study confirmed that the children's choices were not random.

Furthermore, the distribution in the figure suggests that the preferred gender is linked to the type of final palatal segment: palatal stops tend to be associated with feminine, whereas palatal liquids more often trigger masculine agreement.

Another follow-up investigation of the 1,560 most frequent palatal-final nouns in the Russian National Corpus (of which approximately 40% are masculine) confirmed that the distribution of masculine and feminine gender with these nouns in the Russian language is dependent on the nature of the final segment, in combination with a few other factors such as stress and final nucleus. Importantly, the distribution of the two genders in the Russian language very closely matches the preference of the children.

This means that not only had the children in the Mitrofanova et al. (2021) study acquired the main gender cues in Russian; they were also sensitive to much finer-grained cues for gender assignment than the authors had anticipated. This finding lends support to the Micro-Cue Model of language acquisition (e.g. Westergaard 2009, 2014), which argues that young children do not learn language by setting large-scale parameters, but instead are sensitive to micro-variation in the input from early on.

4. What do Russian Adults Do?

Urek (submitted) ran a large-scale online experiment with 267 monolingual Russian-speaking adults, who were asked to classify orthographically presented novel palatal-final monosyllables as either masculine or feminine.



Figure 2. Example of stimulus presentation (from Urek, submitted)

The dataset comprised 1,648 novel monosyllabic items, created by combining 12 simplex non-shibilant palatal codas, 8 nuclei, and 20 simplex onsets, divided across 24 lists. Each participant evaluated the items in their assigned list twice, making it possible to compute test-retest reliability. The underlying assumption was that reasonably strong test-retest

consistency in gender judgments would reflect sensitivity to noun-internal cues to grammatical gender. In addition to the gender task, participants completed three background measures: a receptive vocabulary test, a digit span task, and a pattern detection task.

The results indicated that around 75% of the sample showed test-retest reliability above chance. This suggests that at least three quarters of participants consistently used noun-internal phonological information when assigning gender to novel nouns. These participants, labeled "analyzers", also scored higher than those whose reliability was at chance ("guessers") on all three background tasks. Urek (ibid.) concludes that one possible interpretation is that people with stronger pattern-detection abilities and larger receptive vocabularies are more likely to pick up on distributional regularities in their linguistic input.

Urek (ibid) also reports that even though responses show an overall preference for the feminine gender, the probability of assigning feminine gender to palatal-final nonce words can be consistently predicted from the phonological characteristics of the onset, nucleus, and coda. This pattern mirrors earlier experimental findings (Slioussar 2018) as well as our own previous results with monolingual and bilingual children (Mitrofanova et al., 2018, 2021) and reinforces the idea that probabilistic rule-based mechanisms guide the classification of nouns with ambiguous marking (Karmiloff-Smith 1979; Pinker 1984; see also Meylan et al. 2017 and McCauley & Christiansen 2019 for a discussion on the role of rule-based mechanisms in grammatical acquisition).

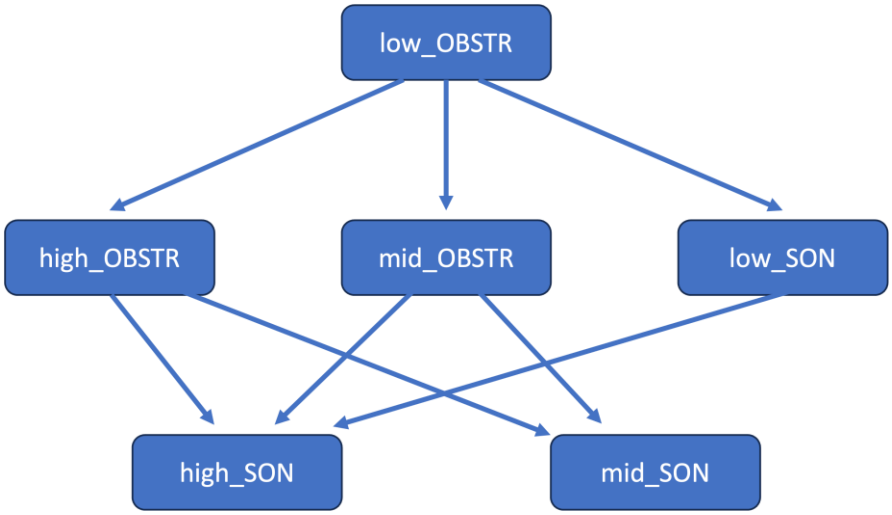


Figure 3. Likelihood of feminine response by rhyme (from Urek, submitted)

Crucially, differences between specific onsets, nuclei, and codas can be captured by broader phonological categories: obstruent codas, low vowels in the nucleus and continuant

onsets are significantly more likely to elicit feminine responses than sonorant codas, high nuclei, and non-continuant onsets, which aligns very closely with the corpus results reported by Nesset (see Fig 4. from Nesset (2003:308)). Moreover, when nuclei and codas are considered together as rhymes, their influences combine in an additive way. This pattern is illustrated in the Hasse diagram in Fig. 3 (fom Urek et al., submitted), where arrows represent statistically significant – and transitive – relationships. For instance, a rhyme in which a low vowel is followed by an obstruent (e.g., **-at**') is more likely to elicit a feminine response than a rhyme in which a low vowel is followed by a sonorant (e.g., **-an**'). By transitivity, both of these rhymes are more likely to trigger feminine agreement than rhymes in which a sonorant is preceded by a high or mid vowel (e.g., **-un**' or **-en**'). Note that because there is no arrow between **-un**' and **-en**', these two rhymes do not differ significantly in their likelihood of triggering feminine agreement. Although analyzers differentiate more finely among onset, nucleus, and coda types than guessers, both groups show the same overall direction of effects, and the trends seen for analyzers are never reversed in the guessers' data.

5. A Comparison of our Approach with Nesset (2003)

In this section, we examine both the differences and converging findings from Mitrofanova et al. (2021), Urek (submitted), and Nesset (2003). These studies employed distinct experimental designs, including tasks where children and adults were asked to assign gender to novel nouns (pseudowords) ending in a palatal consonant, as well as analyses of existing corpus data to assess generalizations based on the gender of existing nouns.

In Mitrofanova et al (2021), all included nouns were simplex monosyllabic nonce nouns. We avoided nouns that could be analysed as having a specific morphological suffix and used pictures of novel objects to accompany the audio prompts, ensuring that all nouns were interpreted as inanimate count nouns, thus isolating the effect of phonological (micro-)cues only. The study also reports on a corpus analysis of monosyllabic nouns ending in palatal consonants.

In Urek (submitted), all nouns were monosyllabic simplex nouns, while overlap with existing morphological cues (suffixes) was avoided. The stimuli were presented orthographically (hence words ending in /ž/, /š/, /č'/, /š'č'/ where the gender can be determined based on the presence/absence of the special soft sign were excluded). The stimuli were presented without accompanying pictures, and no semantic restrictions were put on the interpretation of the novel nouns.

Our focus on (morpho)phonological cues was motivated by the phonology-first proposal in language acquisition literature (see Culbertson et al. 2019; Gagliardi & Lidz 2014)

Corpus data with existing nouns and experiments with pseudowords/novel nouns suggest that children rely more on the phonological cues than semantic cues and even deterministic agreement cues (in case these are in mismatch to the probabilistic phonological cues). See examples from child corpora like *le maman* with masculine article used with a female referent 'mother' (Karmiloff-Smith 1979) or feminine adjective agreement with masculine nouns ending in –a like *papa* and *deduška* in Russian (Rodina & Westergaard 2012). Nesset (2003) took a more holistic approach than our more focussed studies. He analysed available corpus data identifying a hierarchy of morphological, semantic and phonological cues as predictors of grammatical gender and declension class in Russian. When discussing phonological cues for morphologically simplex inanimate nouns, Nesset proposed the following hierarchy of cues (see Fig. 4 from Nesset 2003: 309):

Final labial consonants favor feminine gender
Final post-alveolar consonants favor masculine gender
Alveolar consonants have a more complex hierarchy:
Final obstruents favor feminine gender
Nasal sonorants tend to favor feminine gender
Oral sonorants: gender depends on the sonority of the penultimate segment:
Consonant or high vowel: favor masculine gender.
Non-high vowels: favor feminine gender

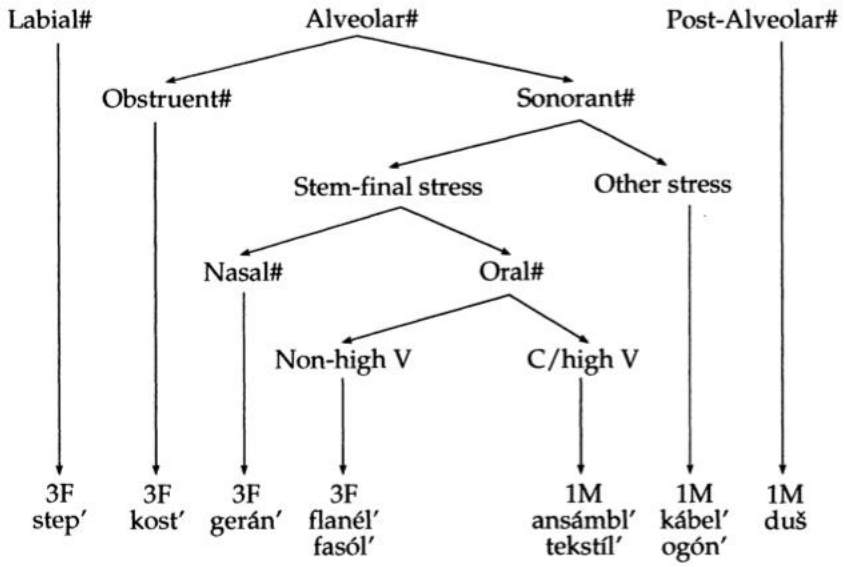


Figure 4. Overview of phonological generalizations, Nesset (2003).

Finally, another study that contributes both corpus analysis of existing nouns and experimental evidence on the gender and declension class of nouns ending in palatal consonants in Russian is Slioussar (2018). The most relevant experimental study reported in the paper investigated how adult Russian speakers assigned gender to nonce nouns ending in -b', -d', -s', -n', -l', and -r'. The results revealed a significant preference for masculine gender in nouns ending in -l' and -r', whereas nouns ending in -b', -d', -s', and -n' showed a feminine preference. Overall, the findings align with generalizations reported by Nesset (2003) and our own studies. Specifically, Slioussar (2018) observed that the majority of nouns ending in labials (p', b', v', f', m'), as well as alveolar obstruents (d', t', and s'), are feminine, while most nouns ending in r' and l' are masculine.

Taken together, all reviewed studies highlight that speakers of Russian (child and adult participants alike, monolingual as well as heritage bilingual children) are sensitive to microvariation in the input, using specific phonological properties as **(micro-)cues** to assign gender, in cases where traditional descriptions of Russian suggest ambiguity.

6. Future Directions

Below we propose some future research direction that in our view could be promising in advancing the understanding of gender assignment in nouns ending in palatal consonants.

First, it would be valuable to investigate the interplay between morphological and purely phonological cues. For example, the suffix -ost' is highly productive and strongly associated with the feminine gender in Russian. An experimental study could test whether human intuitions about novel nouns ending in -t' (but not -ost') align more closely with the overall proportions of all nouns ending in -t' in the corpus or with the proportions of nouns ending in -t' excluding those with the -ost' suffix. This would shed light on the relative weight of morphological versus phonological cues in gender assignment.

Second, future studies could incorporate semantic cues into experimental designs, building on Nesset's hierarchy of cues, which prioritizes morphological cues over semantic and phonological ones. By systematically manipulating semantic information, researchers could better understand how these cues interact with phonological and morphological factors in shaping gender assignment.

Third, a comparative approach could be taken to examine gender assignment in novel nouns between humans and AI systems. In a preliminary query conducted in January 2026 (see Appendix), we asked four AI systems (ChatGPT, ChatUiT, Gemini, and Claude) to determine the grammatical gender of 11 pseudo-nouns from the child survey reported on in Mitrofanova et al. (2021), as if they existed in Russian. The study originally has 13 nouns (see above), but to avoid orthographic cues, two nouns ending in -š were excluded from

the prompts. The systems' responses varied significantly. ChatGPT emerged as an outlier, assigning masculine gender to all 11 pseudo-nouns. ChatUiT's responses aligned with human judgments in 7 out of 11 cases, though it diverged in four instances: two nouns were identified as masculine by the system but feminine by humans, and two were identified as feminine by the system but masculine by humans. Gemini and Claude performed similarly, with both systems matching human judgments in 9 out of 11 cases. In both instances, the mismatched responses involved nouns assigned feminine gender by the systems but masculine by humans. Overall, Gemini's responses most closely approximated human intuitions, while ChatGPT deviated the most. However, all AI systems demonstrated some level of divergence from human judgments, highlighting the complexity of gender assignment and the limitations of current AI models in replicating human linguistic intuitions.

Finally, a cross-linguistic perspective could provide valuable insights by comparing gender assignment for nouns ending in palatal consonants across Slavic languages, including Ukrainian, Polish, and Czech. This could involve investigating microvariation within Slavic languages and examining how monolingual and bilingual speakers navigate co-existing systems of micro-cues. Such research could reveal both language-specific patterns and broader typological trends, as well as the potential influence of bilingualism and crosslinguistic influence on gender assignment processes.

By pursuing these avenues, future research could gain a more nuanced understanding of the factors influencing gender assignment in Slavic languages, while also shedding light on broader questions about the interaction of morphology, phonology, and semantics in mono- and bilingual language acquisition and processing.

7. Concluding Remarks

This chapter focuses on gender assignment in a special class of Russian nouns ending in palatal consonants, a class traditionally described as having ambiguous gender cues (compatible both with feminine and masculine gender). We compare findings from several experimental studies with Russian-speaking children and adults to the analysis proposed by Nesset (2003). Experimental tasks involving nonce words suggest that both children and adults are sensitive to very fine-grained phonological cues, with gender preferences varying (non-randomly) by the type of final consonant. These findings align with Tore Nesset's (2003) analysis pertaining to phonology, highlighting the role of fine-grained phonological **micro-cues** in gender assignment.

References

Busterud, Guro, Terje Lohndal, Yulia Rodina & Marit Westergaard. 2019. The loss of feminine gender in Norwegian: A dialect comparison. *Journal of Comparative Germanic Linguistics* 22(2). 141–167. https://doi.org/10.1007/s10828-019-09108-7.

Ceitlin, S. N. 2009. *Očerki po slovoobrazovaniju i formoobrazovaniju v detskoj reči. [On Inflection and Derivation in Child Language]*. Moscow: Znak.

Corbett, Greville G. 1991. *Gender*. Cambridge: Cambridge University Press. https://doi.org/10.1017/CBO9781139166119.

Gvozdev, A. N. 1961. *Formirovanie u rebenka grammatičeskogo stroja russkogo jazyka. [Language Development of a Russian Child]*. Moscow: APN RSFSR.

Hockett, Charles F. 1958. *A course in modern linguistics*. New York: MacMillan.

Horst, Jessica S. & Michael C. Hout. 2016. The Novel Object and Unusual Name (NOUN) Database: a collection of novel images for use in experimental research. *Behavior Research Methods* 48(4). 1393–1409. https://doi.org/10.3758/s13428-015-0647-3.

Lohndal, Terje & Marit Westergaard. 2021. Grammatical gender: Acquisition, attrition, and change. *Journal of Germanic Linguistics* 33(1). 95–121. https://doi.org/10.1017/S1470542720000057.

McCauley, S. M. & M.H. Christiansen. 2019. Language learning as language use: A cross-linguistic model of child language development. *Psychological Review* 126(1). 1–51. https://doi.org/10.1037/rev0000126.

Meylan, S. C., M.C. Frank, B.C. Roy & R. Levy. 2017. The Emergence of an Abstract Grammatical Category in Children's Early Speech. *Psychological Science* 28(2). 181–192. https://doi.org/10.1177/0956797616688341.

Mitrofanova, Natalia, Yulia Rodina, Olga Urek & Marit Westergaard. 2021. Sensitivity to microvariation in bilingual acquisition: morphophonological gender cues in Russian heritage language. *Applied Psycholinguistics* 43(1). 41–79. https://doi.org/10.1017/S0142716421000382.

Mitrofanova, Natalia, Yulia Rodina, Olga Urek & Marit Westergaard. 2018. Bilinguals' sensitivity to grammatical gender cues in Russian: the role of cumulative input, proficiency, and dominance. *Frontiers in Psychology* 9. https://doi.org/10.3389/fpsyg.2018.01894.

Nesset, Tore. 2003. The Assignment of Gender and Declension to Russian Nouns in Soft Consonants: Predictability and Rule Interaction. *Journal of Slavic Linguistics* 11(2). 287–322.

Rodina, Yulia & Marit Westergaard. 2021. Grammatical gender and declension class in language change: A study of the loss of feminine gender in Norwegian. *Journal of Germanic Linguistics* 33(3). 235–263. https://doi.org/10.1017/S1470542719000217.

Rodina, Yulia, Tanja Kupisch, Natalia Meir, Natalia Mitrofanova, Olga Urek & Marit Westergaard. 2020. Internal and External Factors in Heritage Language Acquisition:

Evidence from Heritage Russian in Israel, Germany, Norway, Latvia and the UK. *Frontiers in Education* 5. https://doi.org/10.3389/feduc.2020.00020.
Rodina, Yulia & Marit Westergaard. 2015. Grammatical gender in bilingual Norwegian-Russian acquisition: The role of input and transparency. *Bilingualism: Language and Cognition* 20(1). 197–214. https://doi.org/10.1017/S1366728915000668.
Rodina, Yulia & Marit Westergaard. 2013. The acquisition of gender and declension in a non-transparent system: monolinguals and bilinguals. *Studia Linguistica* 67(1). 47–67.
Rodina, Yulia & Marit Westergaard. 2012. A cue-based approach to the acquisition of grammatical gender in Russian. *Journal of Child Language* 39(5). 1077–1106. https://doi.org/10.1017/S0305000911000419.
Slioussar, N. (2018). Gender, declension and stem-final consonants: An experimental study of gender agreement in Russian. *Computational Linguistics and Intellectual Technologies* 17. 688–700.
Urek, Olga, Terje Lohndal & Marit Westergaard. 2022. En splyv eller et splyv? Tilordning av grammatisk genus til pseudosubstantiv i norsk. *Norsk Lingvistisk Tidsskrift* 40(1). 7–26.
Urek, Olga. Submitted. Categorization of novel nouns with ambiguous gender cues: experimental evidence from Russian.
Westergaard, Marit. 2009. *The Acquisition of Word Order: Micro-cues, information structure, and economy*. Amsterdam & Philadelphia: John Benjamins Publishing.
Westergaard, Marit. 2014. Linguistic variation and micro-cues in first language acquisition. *Linguistic Variation* 14(1). 26–45. https://doi.org/10.1075/lv.14.1.02wes.
Westergaard, Marit & Yulia Rodina. 2016. Hvor mange genus er det i Tromsødialekten? *Maal og Minne* 2. 159-189.

Appendix

AI responses to the query about the nouns клянь, дрысть, кнофь, дронь, барть, рофть, трушть, дюбрь, кнафь, брясль, скубль (red marks misalignment with human responses), January 2026:

ChatUiT

Клянь — скорее мужской род, по аналогии со словами "день", "огонь".
Дрысть — скорее женский род, по аналогии со словами "грусть", "честь".
Кнофь — скорее женский род, по аналогии со словами "кровь", "любовь".
Дронь — скорее мужской род, по аналогии со словами "конь", "огонь".
Барть — скорее женский род, по аналогии со словами "дверь", "твердь".
Рофть — скорее женский род, по аналогии со словами "мощь", "печь".
Трушть — скорее женский род, по аналогии со словами "честь", "грусть".
Дюбрь — скорее мужской род, по аналогии со словами "январь", "сентябрь".
Кнафь — скорее женский род, по аналогии со словами "кровь", "твердь".
Брясль — скорее женский род, по аналогии со словами "мысль", "пыль".
Скубль — скорее мужской род, по аналогии со словами "уголь", "шкворень".

ChatGPT

Все приведённые формы:
клянь, дрысть, кнофь, дронь, барть, рофть, трушть, дюбрь, кнафь, брясль, скубль
односложные или квазикорневые, без суффиксов, характерных для ж. р. (-ость, -есть, -ань, -ынь и т. п.), фонетически напоминают мужские существительные типа гвоздь, дождь, рыцарь (основа).
По умолчанию они будут интерпретированы как существительные мужского рода.

Gemini

Слово	Род
Клянь	Женский
Дрысть	Женский
Кнофь	Женский
Дронь	Женский
Барть	Мужской/Женский
Рофть	Женский
Трушть	Женский
Дюбрь	Мужской
Кнафь	Женский
Брясль	Женский
Скубль	Мужской

Claude

Женский род (наиболее вероятный):
клянь → по аналогии с длань, рань, дань — ж.р.
дрысть → суффикс -сть почти всегда женский (кость, власть, страсть) — ж.р.
кнофь → как морковь, бровь, кровь — ж.р.
дронь → как бронь, вонь, благовонь — ж.р.
барть → существует архаичное слово борть (дупло для пчёл) — ж.р.
рофть → по аналогии с редкими словами на -фть — ж.р.
трушть → по модели -сть/-шть — ж.р.
дюбрь → как дробь, скорбь, хворь — ж.р.
кнафь → как гарь, тварь, дрянь — ж.р.
Мужской род (менее типичный):
брясль → по аналогии с мысль (ж.р.), но скорее с кисель — м.р.
скубль → как рубль, костыль, ковыль — м.р.