


ORIGINAL ARTICLE

Sensitivity to microvariation in bilingual acquisition: morphophonological gender cues in Russian heritage language

Natalia Mitrofanova^{1*} , Olga Urek¹, Yulia Rodina¹ and Marit Westergaard^{1,2}

¹Department of Language and Culture, UiT, The Arctic University of Norway, Tromsø, Norway and

²Department of Language and Literature, NTNU, Norwegian University of Science and Technology, Trondheim, Norway

*Corresponding author. Email: natalia.mitrofanova@uit.no

(Received 13 November 2019; revised 17 November 2020; accepted 15 February 2021; first published online 27 September 2021)

Abstract

Previous research on the acquisition of grammatical gender has shown that this property is acquired early in transparent gender systems such as Russian. However, it is not clear to what extent children are sensitive to the assignment cues and to what extent they simply memorize correspondences between frequent lexical items. Furthermore, we do not know if bilingual children are different from monolingual children in this respect. This article reports on a study investigating bilingual children's sensitivity to gender assignment cues in Russian. A group of 64 bilingual German–Russian children living in Germany participated in the study, as well as 107 monolingual controls in Russia. The elicitation experiments used both real and nonce words, as well as noun phrases with mismatched cues (where the morphophonological shape of the noun cued one gender and the agreement on the modifying adjective another). The results show that both bilinguals and monolinguals are highly sensitive to cues, both to the frequent transparent cues and to more fine-grained gender regularities in situations where there is ambiguity. There is also an age effect, showing that younger children pay more attention to the cue on the noun itself, thus displaying a preference for regular patterns, while older children are more sensitive to gender agreement on other targets.

Keywords: grammatical gender; transparent cues; ambiguous cues; nonce words; mismatched cues; Russian; heritage language

Introduction

In this paper, we investigate monolingual and bilingual children's sensitivity to microvariation in the input, more specifically to morphophonological cues for grammatical gender assignment in Russian. The study is framed within the micro-cue model of language acquisition (Westergaard, 2009a, b, 2014), which argues that children are sensitive to fine-grained linguistic distinctions from early

© The Author(s), 2021. Published by Cambridge University Press. This is an Open Access article, distributed under the terms of the Creative Commons Attribution licence (<https://creativecommons.org/licenses/by/4.0/>), which permits unrestricted re-use, distribution, and reproduction in any medium, provided the original work is properly cited.

on. This has been attested in work on L1 acquisition, including research on grammatical gender in Russian (Rodina, 2008; Rodina & Westergaard, 2012). However, it remains an open question whether this sensitivity may also be found in bilingual contexts. On the one hand, previous research on the acquisition of grammatical gender by bilingual children has shown that, given sufficient input, their gender system may be similar to that of monolinguals (Rodina & Westergaard, 2017). On the other hand, this same research has generally used experiments with existing nouns, and it is thus unclear whether the bilinguals are in fact sensitive to cues or whether they have simply memorized the agreement patterns associated with frequent lexical items in the input.

For this reason, we have used both existing nouns and nonce words in the current study. Furthermore, inspired by Karmiloff-Smith's (1979) seminal paper, we have included a nonce word experiment with mismatched cues, where the morphophonological form of the nonce noun and the agreement marker on the adjective indicate different genders. The three experiments were carried out with German–Russian bilinguals (aged 3–10) as well as a group of monolingual Russian children (aged 3–7).

Our findings show that both monolinguals and bilinguals rely on morphophonological cues to assign gender to both real and nonce nouns, with a somewhat higher accuracy on the former, arguably due to retrieval being more efficient than computation. For the experiment with mismatched cues, both monolingual and bilingual children rely more on noun-internal cues from early on, with the importance of noun-external gender agreement increasing with age.

Background

Theoretical model

This study is couched within the micro-cue model of Westergaard (2009a, b, 2014), which means that we consider the Russian gender system from an acquisition perspective. The model argues that children are highly sensitive to variation in the input, paying attention to fine-grained linguistic distinctions from early on. This means that children do not learn by setting (macro-) parameters, as has been assumed in traditional generative literature. Instead, children are argued to be equipped with an innate endowment that enables them to parse the input and build syntactic structure in a stepwise fashion, based on positive evidence in the primary linguistic data. Thus, children are conservative learners, and the acquisition process is also affected by a principle of economy. The model has so far mainly been based on data from syntactic phenomena, such as verb movement, object shift, possessive movement (see, e.g., Westergaard, 2014), showing that young monolingual children do not initially make major generalizations, but typically make the relevant distinctions between different clause types, verb types, subject types, or other linguistic categories from their earliest possible production. For example, when acquiring verb-second (V2) word order in *wh*-questions in Norwegian (which is variable in the adult language in most dialects), the children do not apply the V2 rule across the board, but make the relevant fine-grained distinctions, for example, providing categorical V2 when the *wh*-element is disyllabic or longer and variable V2 (dependent on information structure) when the *wh*-element is monosyllabic, as in the adult

language. Despite this, children clearly do make generalizations in the acquisition process in their attempt to figure out the system behind the input variation that they are exposed to. That is, the result of language acquisition is not just a collection of item-based information, but a rule-based system from early on (see, e.g., Westergaard, 2009b). The important aspect of the micro-cue model is that children's generalizations are small, not affecting major categories (e.g., all verbs or all nouns), but take place across a small class or subcategory (one micro-cue at a time).

However, it is also well known that there are cases where children *over*-generalize, especially in morphology, for example, the often-cited examples of overextensions of regular past tense in English (*comed* instead of *came*; see, e.g., Pinker, 1999) or overgeneralization of common gender to neuter nouns in Dutch (*de paard* instead of *het paard* "the horse"; see Blom, Poliřenská & Weerman, 2008). So far, the micro-cue model has only been used to account for one morphological phenomenon, grammatical gender in Russian: Rodina (2008) and Rodina and Westergaard (2012) show that, while Russian-speaking monolinguals acquire the transparent gender system of morphophonological cues early (see below), they also occasionally overgeneralize these patterns to ambiguous cases, for example, to *papa*-type ("daddy") nouns, which have the shape of feminine nouns but take semantic (masculine) gender agreement; see example (1) from Rodina and Westergaard (2012:1092).

- (1) *deduřka oranřev-aja na tareločke* (Vera 3;9)
 granddad_(M) orange-F on plate
 'The orange granddad is on the plate.'
 Target: *deduřka oranřev-yj na tareločke*

Importantly, Rodina and Westergaard (2012) show that there is a limit to this type of overgeneralization, in that when the semantic rule is learned, it is used differently across relevant contexts; that is, across different noun classes (hybrid nouns, double-gender nouns, etc.) and male versus female reference. For example, while semantic agreement for *papa*-type nouns, hybrid nouns with male reference and double-gender nouns with female reference is almost at ceiling (92.4%, 96.9%, and 87.4%, respectively), it is only occasionally used with female names with a typically masculine ending (*-ik*, *-ok*) and hybrid nouns with female reference (14.3% and 18.8%, respectively). Rodina and Westergaard (2012) thus conclude that L1 Russian children are sensitive to fine-grained distinctions in the input, not generalizing semantic agreement across the board, but only within subclasses of nouns.

Gender in Russian from an acquisition perspective

Gender is a category sorting nouns into different classes. We assume that the particular gender value of each noun is a feature on the noun stem, and this feature is expressed externally on elements agreeing with the noun (see the traditional definition in Hockett, 1958). In Russian, these elements are adjectives, participles, verbs in the past tense, demonstratives, possessives, many quantifiers, and certain numerals. All the agreeing elements except for verbs are declinable (have different forms depending on the case).

Table 1. Declension of inanimate nouns with adjectives

	I: M/N	II: F	III: F (Pal)
Nominative	Bolsh- oj stol-Ø/bolsh- oje okn- o	Bolsh- aja sten- a	dver'-Ø
Accusative	-oj -Ø /-oje -o	-uju -u	-Ø
Genitive	-ogo -a	-oj -y	-i
Dative	-omu -u	-oj -e	-i
Instrumental	-ym -om	-oj -oj	-ju
Locative	-om -e	-oj -e	-i

Apart from a small number of the so-called double-gender nouns and pluralia tantum, each noun in Russian belongs to either masculine (M), feminine (F), or neuter (N) gender. The numerical distributions reported in various sources show that masculine is the most frequent and neuter the least frequent class (Lazova, 1974; Mučnik, 1971; Zaliznyak, 1977; Ilola & Mustajoki, 1989; Corbett, 1991). According to Corbett (1991:78), of a total of 33,952 nouns derived from Russian dictionaries, only 13% are neuter, 46% masculine, and 41% feminine. Masculine is considered to be the linguistic default, as it is the least marked form (see below) and is also typically assigned to borrowings (Corbett, 1991:78).

There are three main declension classes in Russian (Shvedova, 1980; Zaliznyak, 1977), and it is traditionally assumed that there is a reliable correlation between gender and declension class (Zaliznyak, 1967; Corbett, 1991). Thus, masculine and neuter nouns typically belong to declension class I,¹ feminine nouns ending in *-a* belong to declension class II, while feminines ending in a palatal belong to declension class III. Since we take a cue-based acquisition approach in this study, we focus on the distinctions that children need to be sensitive to in order to acquire the gender system. The declension paradigms of inanimate nouns with adjectives are illustrated in Table 1 for the examples in (2).

- (2) a. bolš-[ó] stol²
 big-M table(M)
 'a big table'
 b. bolš-[ó(j)ə] okn[ó]
 big-N window(N)
 'a big window'
 c. bolš-[á(j)ə] sten[á]
 big-F wall(F)
 'a big wall'
 d. bolš-[á(j)ə] dver'
 big-F door(F)
 'a big door'

Note that adjectival declension paradigms in masculine and neuter singular are the same, except for the forms in the nominative and accusative, while there is a separate adjectival paradigm for feminine singular (see Shvedova, 1980; Zaliznyak, 1977; Halle & Matushansky, 2006).

There are generally two types of information available in the input that children can use when acquiring grammatical gender:

- (a) noun-external information (on the agreeing elements);
- (b) noun-internal information (phonological and/or semantic regularities that are typical of nouns belonging to a particular gender).

It has been shown in previous literature (Gagliardi & Lidz, 2014; Gagliardi, 2012; Culbertson, Gagliardi & Smith, 2017; Culbertson et al., 2019) that young children are highly sensitive to noun-internal phonological information when acquiring gender (i.e., probabilities based on the phonological form of the noun), and it plays a more important role than other types of information (semantic cues or noun-external gender agreement information).

Such phonological regularities exist in Russian too. In this study, we focus on the phonological form of nouns in the nominative singular, which is the most frequent form (see, e.g., Slioussar & Samoilova, 2015). For inanimate masculine and all neuter nouns, the nominative is syncretic with the accusative, which is the second most frequent form. As noted by Corbett (1991:35), the gender of the majority of Russian nouns can be predicted from the nominative singular. In what follows we present some of the phonological “cues” (i.e., regularities based on the phonological form of the noun) in the nominative singular that are available in Russian.

Transparent gender cues: -C, [-á], and [-ó]

The majority of nouns that end in a non-palatal consonant (henceforth -C) are masculine, for example, *stol* “table” in (2a). Stressed [-á] is a cue for feminine, for example, *sten-[-á]* ‘wall’ in (2c) and stressed [-ó] is a cue for neuter, for example, *okn-[-ó]* “window” in (2b). Nouns in [-ó] are less productive than the nouns in -C and [-á] (cf. Lazova 1974 cited in Corbett 1982:206, Table IV). These cues also have notable exceptions, viz. nouns that are assigned a gender based on the semantic/sex-based properties of the noun referent, including *papa*-type (“daddy”) nouns which are masculine, as well as hybrids like *doktor* “doctor” and double-gender nouns like *zadira* “tease”, which can be either feminine or masculine.

Ambiguous masculine/feminine: -C'

Nominative singular is not a reliable predictor of gender for a subset of nouns ending in a palatalized consonant (henceforth -C'). Such nouns can be masculine (*korabl'* “ship”) or feminine (*krovat'* “bed”). Thus, other types of information, such as declension class (see Table 1) and agreement on noun-external elements, must be used to acquire the grammatical gender of these nouns.

Recent research has suggested that there exist more fine-grained phonological regularities in the nominative singular that might be predictive of the gender of individual nouns ending in a palatal. Nessel (2003) proposes several generalizations tested against a corpus consisting of all nouns ending in palatals from Zaliznjak (1977), which are largely confirmed by data from the Russian National Corpus (RNC) reported in Slioussar (2018). Thus, different palatalized consonants may predict either masculine or feminine: for example, nouns ending in [b', p', v', f, m', t', s']

tend to be feminine, and nouns ending in [r', l'] tend to be masculine (Slioussar, 2018)³. In other words, this suggests that the palatalized consonants may represent more fine-grained cues for learners than the -C' cue.

Ambiguous feminine/neuter: -[ə]

Predicting gender based on the nominative singular is also problematic when nouns end in an unstressed vowel. The unstressed -a and -o endings both sound like schwa (-[ə]) and create ambiguity between feminine and neuter. The gender of these nouns can be predicted from their declensions (cf., Table 1). In contrast, adjectival agreement is often unable to disambiguate between these nouns, because in the majority of cases, the adjectival endings are also unstressed and thus ambiguous, illustrated in (3). Only prenominal modifiers with stress on the final syllable can resolve this ambiguity, for example, *golub*-[á(j)ə] *kníg*-[ə] ‘a light blue book’ versus *golub*-[ó(j)ə] *óblak*-[ə] ‘light blue cloud’.

- (3) a. *bél*-[ə(j)ə] *kníg*[ə]
 white-F book(F)
 ‘a white book’
 b. *bél*-[ə(j)ə] *óblak*[ə]
 white-N cloud(N)
 ‘a white cloud’

The cues in the nominative are presumably not the only ones that children pay attention to when acquiring grammatical gender in Russian. As mentioned above, there is a reliable correlation between the declension class that the noun belongs to and its gender. It is thus likely that the knowledge of the declension class of the noun facilitates the acquisition of gender (see Corbett, 1991; Tarasenkova, 2010). Furthermore, the gender of a noun can be predicted by the specific diminutive suffix this noun combines with, which has been argued to promote the acquisition of gender (see Kempe, Brooks, Mironova & Fedorova, 2003; Montrul, de la Fuente, Davidson & Foote, 2013; Janssen, 2014). In this paper, we abstract away from these types of cues, and only focus on children’s sensitivity to the form of the noun in the nominative. We leave questions concerning the sensitivity to other types of cues and their contribution to the acquisition of grammatical gender in Russian for future research.

Previous research on the acquisition of grammatical gender

Grammatical gender assignment in Russian has been investigated in monolingual children as well as in child and adult heritage speakers (Gvozdev, 1961; Ceitlin, 2009; Rodina, 2008; Rodina & Westergaard, 2012, 2017; Mitrofanova *et al.*, 2018; Polinsky, 2008; Schwartz *et al.*, 2015), and transparency of the morphophonological gender cues has been shown to play a major role. Furthermore, the presence or absence of predictable gender cues in one of the languages of a bilingual has been argued to have an effect on the other language in terms of acceleration or delay in the acquisition process (e.g., Eichler, Jansen & Müller, 2013; Egger, Hulk & Tsimpli, 2018; Kaltsa, Tsimpli & Argyri, 2019; Kupisch, Müller & Cantone, 2002). Both

monolinguals and bilinguals have been found to generally use target-consistent gender marking with masculine and feminine transparent nouns in Russian, showing sensitivity to the transparent *-C* and *-á* cues. Ambiguous nouns are more problematic, and monolinguals have been found to use two strategies: masculine is overgeneralized to feminine ambiguous and feminine to neuter ambiguous nouns (Gvozdev, 1961; Ceitlin, 2009). Similar patterns are observed in child and adult heritage speakers (Rodina & Westergaard, 2012, 2017; Mitrofanova et al., 2018; Polinsky, 2008; Schwartz et al., 2015). Polinsky (2008) has found that less proficient heritage speakers may develop a two-gender system of masculine and feminine, where neuters are reanalyzed as feminines, while Rodina and Westergaard (2017) and Mitrofanova et al. (2018) provide results showing that certain bilingual children use only masculine across all noun classes, meaning that they have no gender at all. It has thus been argued that some heritage Russian speakers are not sensitive to morphophonological cues for grammatical gender assignment.

With the exception of Rodina (2008) and Rodina and Westergaard (2012), research focusing specifically on learner sensitivity to gender cues in Russian is limited. The impact of morphophonological properties has been shown to be stronger for younger than for older children. Tarasenkova (2010) shows that Russian 3–5-year-olds, who experience problems with ambiguous neuters in *-[ə]*, are able to make generalizations about the gender of novel nouns based on the instrumental singular form ending in *-om*, which disambiguates between neuter and feminine (cf., *-om* vs. *-oj* endings in Table 1). Furthermore, she shows that children are significantly more successful at assigning gender to novel nouns if they can rely on morphophonological gender cues rather than on adjectival agreement.

Such findings support similar investigations on other languages including the seminal work by Karmiloff-Smith (1979). Using nonce nouns, she showed that French-speaking children are highly sensitive to the phonological information encoded on the noun suffix to determine the gender of a noun. For young children, neither syntactic information (i.e., the indefinite article provided by the experimenter) nor semantic information (i.e., the sex of persons depicted in the drawings) is as predictive in eliciting gender agreement as the form of the noun. Only gradually are “the phonological procedures $\langle \dots \rangle$ (in some cases from 6 years, but more frequently at roughly 9 years) replaced by the natural gender clues and by the more foolproof syntactic ones” (Karmiloff-Smith, 1979:167). Similarly, using artificial language experiments, Culbertson et al. (2019) have shown that children over-rely on phonological cues when assigning gender, and they argue that this is due to a general phonological bias as well as the early availability of phonological cues in the input. Evidence from L2 acquisition may also be relevant here: In a gender processing study comparing L1 speakers to L2 learners of German, Bordag et al. (2006) found that the latter group had shorter reaction times and lower error rates with transparent nouns than with nouns that had an ambiguous ending. However, there was no such effect in the L1 group, and the authors argue that this provides evidence that L1 and L2 processing is essentially the same, but that with more exposure and higher proficiency in the language, the connection between a particular noun and the appropriate gender node becomes so strong (in the L1 group) that the connections necessary to compute gender based on the phonological cue become weaker.

Table 2. Morphophonological gender cues in Russian

Masculine transparent	Feminine transparent	Neuter transparent	Masculine ambiguous	Feminine ambiguous	Neuter ambiguous
-C	-á	-o	-C'	-C'	-/ə/
<i>dom</i> “house”	<i>lisá</i> “fox”	<i>vedró</i> “bucket”	<i>korabl'</i> “ship”	<i>kost'</i> “bone”	<i>myl'/ə/</i> “soap”

To date, it is still unclear whether child speakers of heritage Russian are sensitive to morphophonological gender cues and use the same strategies to resolve gender mismatches. A recent investigation by Mitrofanova *et al.* (2018) shows that Norwegian–Russian bilinguals are generally sensitive to the gender cues presented in Table 2, but concludes that there is considerable variation dependent on a combination of background variables and proficiency measures. Thus, we need much further research in this area.

Research questions and predictions

While the research reviewed in the previous section has provided some answers to important questions related to young children’s sensitivity to variation in the input, there are still many unexplored aspects of the acquisition of grammatical gender. With the exception of a few studies, most of the previous research on the acquisition of gender has investigated existing nouns. It is therefore still unclear to what extent children acquire rules for gender assignment and to what extent they simply use memorized lexical knowledge (as they are clearly able to do in languages with nontransparent gender systems, such as Norwegian; see, e.g., Rodina & Westergaard 2013, 2015). There are also many unanswered questions related to the types of cues children are most sensitive to as well as the behavior of bilinguals in this respect. In this study, we therefore compare monolingual Russian and bilingual German–Russian-speaking children’s performance on three tasks, focusing on real and nonce words with both transparent and ambiguous cues as well as nonce words with mismatching cues.

Based on the structure of Russian and the results of previous research, we ask the following research questions (1a–4a) and make corresponding predictions (1b–4b):

1. a. To what extent are monolingual and bilingual children sensitive to gender cues in Russian and to what extent do they rely on memorized lexical knowledge?
- b. Following considerable previous research on L1 acquisition, we expect young children to acquire rule-based patterns early (e.g., Westergaard, 2009b; Yang, 2016) and thus be sensitive to the transparent correspondences for gender assignment in Russian. We also expect to replicate the findings in Mitrofanova *et al.* (2018), which show that bilingual children also display sensitivity to the assignment cues, although to a lesser extent than the monolinguals. Thus, we predict that the bilinguals are more prone to use the masculine default. Furthermore, we expect both monolinguals and bilinguals to perform

- better with real than nonce nouns, indicating a role also for memorized lexical knowledge in the acquisition of gender.
2. a. Do young children make fine-grained linguistic distinctions in gender assignment from early on (cf., the micro-cue model, Westergaard, 2009a, b, 2014)? More specifically, are monolingual and bilingual children sensitive to ambiguous gender cues in Russian (i.e., stem-stressed nouns ending in a vowel (FN) or palatalized consonants (MF), or do they default to masculine when exposed to nonce words with such cues?
 - b. Following the micro-cue model, L1 children should be sensitive to fine-grained variation, at least in syntax. Previous studies of morphology have provided mixed results (cf., the above section). For Russian gender, previous research on L1 acquisition (e.g., Gvozdev, 1961; Ceitlin, 2009) has shown overgeneralization of F to ambiguous neuters and overgeneralization of M to ambiguous feminines. There is also an effect of the masculine default, especially in bilingual speakers with reduced exposure to Russian (Rodina & Westergaard, 2017). We therefore expect such overgeneralizations to be predominant in the ambiguous gender conditions, mainly with the bilinguals, but possibly also with the monolingual children.
 3. a. Which cues are Russian-speaking monolingual and bilingual children most sensitive to for gender assignment, nominal endings or gender agreement? Does this change with age?
 - b. Following Karmiloff-Smith (1979), we predict younger children to pay attention to the shape of the noun when assigning gender to nonce words, while older children should be more sensitive to gender agreement with other targets.
 4. a. Are there only quantitative or also qualitative differences between monolinguals and bilinguals? That is, are there signs of a reduced gender system, for example, a two-gender system in heritage speaker data (cf., findings from bilingual adults in Polinsky, 2008) or no gender system at all (cf., findings from bilingual children in Rodina & Westergaard, 2017)?
 - b. Based on previous research with bilingual children (e.g., Rodina & Westergaard, 2017, Mitrofanova et al., 2018), we predict that the bilinguals will either have an intact three-gender system or a gender system that is heavily affected by the masculine default.

Methodology and participants⁴

Methodology

Real and Nonce word tasks

Experiments 1 and 2 investigated whether children make use of formal cues in assigning grammatical gender to familiar and novel nouns. In Experiment 1, we elicited adjectival gender agreement with 30 familiar Russian nouns falling into 6 conditions: feminine, masculine, and neuter nouns with transparent as well as ambiguous gender cues (cf., the description of Russian gender above). In Experiment 2, we elicited adjectival gender agreement with 25 novel nouns

constructed to conform to Russian phonotactics and falling into 5 conditions: nouns with transparent feminine, masculine, and neuter gender cues, as well as stem-stressed schwa-final nouns (ambiguous FN) and nouns ending in palatal consonants (ambiguous FM). In order to avoid neighborhood density effects, only novel nouns without nominal phonological neighbors were selected. The list of experimental stimuli is given in Tables 3 and 4.

The elicitation procedure used in Experiments 1 and 2 was an adapted version of the picture-based elicitation task in Rodina and Westergaard (2013, 2017). In both experiments, children were presented with a set of pictures on a laptop screen, each depicting a pair of identical but differently colored objects. In Experiment 1, the objects were familiar household items and animals. In Experiment 2, the objects were novel items selected from the Novel Object and Unusual Name Database (NOUN; Horst & Hout, 2016), illustrated in Figure 1. In both experiments, the object was first named by the experimenter (in such a way as not to provide any agreement cues to the gender of the target noun). The child was then asked to name the colors of the items on the screen, producing adjectival gender agreement. The experimenter then pressed a key causing one of the objects to disappear, and the child was prompted to name the object that disappeared. The script for the elicitation dialogue is given in (5).

(5) *Elicitation dialogue in Experiments 1 and 2 (in Russian)*

Experimenter: “That’s what we call *dom*(M)(real)/*punip*(M)(novel). What color are they?”

Child: “A **blue**-M *dom*(M)/*punip*(M) and a **golden**-M *dom*(M)/*punip*(M).”

Experimenter: “And now, what disappeared?”

Child: “The **blue**-M *dom*(M)/*punip*(M).”

Mixed cues task

Experiment 3 investigated the relative sensitivity to formal (morphophonological) and syntactic (adjectival agreement) gender cues in cases where these cues provide conflicting information. Transparent novel nouns from Experiment 2 (see Table 5) were paired with color adjectives that either matched or mismatched with them in gender. This resulted in nine conditions, illustrated in (6).

(6) *Experimental stimuli in Experiment 3*

Adjective – nonce noun match

MM: golub-[*ó(j)*]-M puníp(M)

FF: golub-[*á(j)ə*]-F kluv[*á*](F)

NN: golub-[*ó(j)ə*]-N garp[*ó*](N)

Adjective – nonce noun mismatch

FM: golub-[*á(j)ə*]-F puníp(M)

NM: golub-[*ó(j)ə*]-N puníp(M)

MF: golub-[*ó(j)*]-M kluv[*á*](F)

NF: golub-[*ó(j)ə*]-N kluv[*á*](F)

MN: golub-[*ó(j)*]-M garp[*ó*](N)

FN: golub-[*á(j)ə*]-F garp[*ó*](N)

Table 3. Experimental stimuli in the Real word task (Experiment 1)

Russian noun endings in nominative singular/gender		Russian noun	Gloss
Transparent cues	Hard consonant (masculine)	tort _M	cake
		most _M	bridge
		stakan _M	glass
		poezd _M	train
		dom _M	house
	-a (feminine)	mašina _F	car
		ljaguška _F	frog
		čška _F	cup
		zmeja _F	snake
		liša _F	fox
	-o (neuter)	moloko _N	milk
		kol'tso _N	ring
		vedro _N	bucket
		krylo _N	wing
		pal'to _N	coat
Ambiguous cues	-/ə/ (neuter)	kreslo _N	armchair
		mylo _N	soap
		platje _N	dress
		odejalo _N	blanket
		sito _N	sieve
	Palatalized consonant (masculine)	fonar' _M	streetlight
		ru' _M	steering wheel
		jakor' _M	anchor
		korabl' _M	ship
		gus' _M	goose
	Palatalized consonant (feminine)	ten' _F	shadow
		tsep' _F	chain
		kost' _F	bone
		medal' _F	medal
		peč' _F	oven

In this experiment, the child was again presented with a picture of identical but differently colored novel objects. The experimenter named one of the objects along with its color, thus providing both the formal cue on the noun and the syntactic – matching or mismatching – cue on the gendered adjective. The child was then

Table 4. Experimental stimuli in the Nonce word task (Experiment 2)

Transparent cues	Hard consonant	punip
		gapuk
		vipan
		kabol
		gamut
	-a	bul'ga
		kluva
		prisa
		tranga
		punta
	-o	garpo
		tivlo
		pruno
		glamo
		kluzo
Ambiguous cues	Palatalized consonant	prosh
		dron'
		knov'
		dryst'
		klyan'
	-/ə/	pruz -/ə/
		smik -/ə/
		klir -/ə/
		gryp -/ə/
		mukt -/ə/

Note: Stressed syllables are highlighted in bold.

prompted to name the second object and its color. The experimenter then pressed the key causing one of the objects to disappear, and the child was prompted to name the object that disappeared.

(7) *Elicitation dialogue in Experiment 3 (in Russian)*

Experimenter: "This is golub-[á(j)ə]-F puníp(M). What is that?"
 Child: "Zolot-[ó(j)]-M puníp(M) OR zolot -[á(j)ə]-F puníp(M)"
 Experimenter: "What disappeared?"
 Child: "Golub-[ó(j)]-M puníp(M) OR golub-[á(j)ə]-F puníp(M)"

Table 5. Experimental stimuli in the Nonce word task (Experiment 2)

Transparent cues	Hard consonant	
		punip
		gapuk
		vipan
		kabol
		gamut
	-a	bul' ga
		klu va
		pr isa
		tr anga
		pun ta
	-o	gar po
		tiv lo
		prun o
		gl amo
		kluz o

Note: Stressed syllables are highlighted in bold.



Figure 1. Example of an experimental novel object.

Participants

Tables 6–8 provide an overview of the participants' demographics. We recruited a group of 64 German–Russian bilingual children (aged 3–10), who were tested in Berlin, Stuttgart, and Singen. All the bilingual children involved in this study were either born in Germany or arrived there before the age of three. The majority of children ($n = 46$) came from families with two Russian-speaking parents, while fewer than one-third ($n = 18$) were growing up in families with one Russian-speaking and one German-speaking parent. A total of 107 monolingual Russian-speaking children (aged 3–6) were tested as controls in Moscow and Ivanovo (these monolingual control data have been taken from Mitrofanova et al., 2018, as the two first tasks were identical). The participants in the mixed cues experiment were tested in

Table 6. Participants in Experiment 1: Real words

Group	N	Age range, mean age
Monolinguals	20	4–6-year-olds (mean 5.3)
Bilinguals	64	3–10-year-olds (mean 6.16)

Table 7. Participants in Experiment 2: Nonce words

Group	N	Age range, mean age
Monolinguals	87	3–7-year-olds (mean 5.2)
Bilinguals	64	3–10-year-olds (mean 6.16)

Table 8. Participants in Experiment 3: Mixed cues

Group	N	Age range, mean age, and <i>SD</i>
Monolingual	38	3–7-year-olds (mean 4.89)
Bilingual	64	3–10-year-olds (mean 6.16)

Moscow, specifically for this study. The monolinguals were included in the study as a baseline for learner behavior with the nonce nouns in Experiments 2 and 3. Previous research shows that gender is in place in Russian-speaking monolinguals by the age of 6–7, even with nouns that have nontransparent gender cues (Gvozdev, 1961; Ceitlin, 2005, 2009). Heritage language children have been found to behave like younger monolinguals, rather than age-matched monolingual peers (Schwartz *et al.*, 2015; Unsworth, 2014). Therefore, the age range was broader for the bilingual heritage Russian participants (3–10 years) than for the monolinguals (3–7 years).

Results

Experiment 1: Real words

Figure 2 presents the children's responses across the six experimental conditions (Table 3) in the Real word task for the two participant groups: Russian monolingual children and Russian–German bilingual children. The accuracy rates of the monolinguals reveal that gender assignment was at the ceiling in M-, F-, and N-transparent as well as M-ambiguous conditions. Somewhat less accurate performance was observed in F- and N-ambiguous conditions, where the accuracy rate was 84% and 85%, respectively. Bilinguals scored at ceiling on the M-transparent and M-Palatal conditions (98% and 92%, respectively), followed by transparent F and N conditions (85% and 74%, respectively), while ambiguous F and N conditions were the least accurate (67% and 66%, respectively).

As evident from Figure 2, monolinguals overused M in the F-Palatal condition (16% of the time) and F in the N-Unstressed condition (15% of the time). These errors are in line with previous findings (Gvozdev 1961, Ceitlin 2009) and are

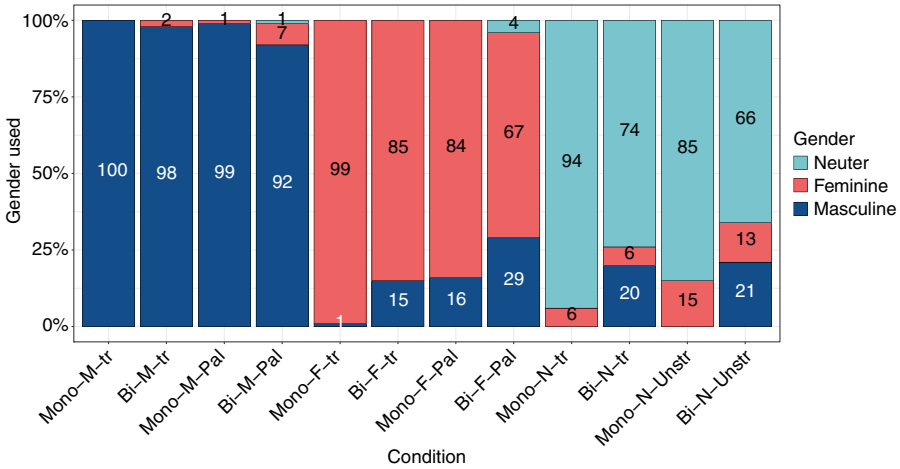


Figure 2. Gender use on the Real word experiment by condition and group. **Conditions:** M-tr = Masculine transparent nouns; M-Pal = Masculine nouns ending in a palatal consonant; F-tr = Feminine transparent nouns; F-Pal = Feminine nouns ending in a palatal consonant; N-tr = Neuter transparent nouns; N-U = Neuter nouns ending in an unstressed vowel. **Groups:** Mono = Monolingual children; Bi = Bilingual children.

compatible with cue-driven gender assignment (the final palatal consonant is compatible with M, while the final shwa often serves as a cue for F). Bilinguals exhibit an additional overgeneralization pattern, with a substantial portion of M responses in the F-transparent (15% of the cases) and Neuter transparent and ambiguous conditions (20% and 21%, respectively). This type of overgeneralization can be regarded as a default strategy and is substantially more pronounced in bilinguals than in monolinguals.

We employed a generalized linear mixed-effects logistic regression analysis to statistically model the effects of Group, Condition, and Age on the accuracy of gender assignment. ANOVA model comparison was used to choose the best model. To compare the groups within conditions, we conducted post hoc pairwise comparisons with the help of the R package *emmeans* (Lenth et al., 2020). All models were fit using R version 4.0.3 (release 2020-10-10), and categorical variables were dummy-coded unless specified otherwise. The winning model for the Real word task predicted the binary variable accuracy as a sum of two interactions: Condition and Group, and Group and Age (continuous). Participants and items were included as random intercepts. The model revealed a significant effect of Age ($p = .01$) – suggesting that both groups of children get significantly more accurate with age – and Condition (M-Pal and M-transparent, $p = .005$ and $p < .001$, respectively). The output of the model is presented in Table A1 in the appendix. Post hoc pairwise comparisons of the groups within conditions showed that the Bilinguals performed significantly less accurately than the Monolinguals on two conditions: Feminine transparent ($p = .03$) and Neuter ambiguous ($p = .01$, see Table A2 in the appendix).

To disentangle the participants’ sensitivity to the signal from a possible response bias (i.e., *over-* or *underuse* of a certain gender), we applied an alternative analysis

couched within Signal Detection Theory⁵ (SDT; Peterson *et al.*, 1954; Pallier, 2002; Macmillan & Creelman, 2004). SDT is widely used in speech perception studies, but has also been applied in other areas of linguistics, such as grammaticality judgement studies (see Huang & Ferreira, 2020). When applied to the Nonce word data, this approach allows us to quantify both the sensitivity to the morphophonological gender cues and the participants' response bias within one analysis. In the Real word experiment, the sensitivity to signal measure does not directly quantify the participants' sensitivity to gender cues, since the gender of real words may be stored and retrieved from the lexicon. The sensitivity to signal measure in the Real word experiment is thus not directly comparable to that of the Nonce word experiment. In the Real word task, our primary aim in applying the SDT analysis was to assess the participants' response bias. To do so, we calculated the number of *hits*, *misses*, *false alarms*, and *correct rejections* based on data from three transparent gender conditions. For example, for calculating sensitivity to the F gender, the categorization of responses is the following: *hits* are F responses in the F condition; *misses* are M or N responses in the F condition; *false alarms* are F responses in M and N conditions; and *correct rejections* are all non-F responses in non-F (i.e., M and N) conditions. Table A3 in the appendix summarizes the number of *hits*, *misses*, *false alarms*, and *correct rejections* for the bilingual and monolingual participants.

Based on Table A3, we calculated two indices: *d*-prime (sensitivity) and *c* (bias) for the two groups of participants in each condition⁶. To calculate the variance and confidence intervals of *d*-prime and *c*, we followed the method proposed in Gurevitch & Galanter (1967; see also Huang & Ferreira, 2020). Table A4 in the appendix summarizes the mean, variance, *SD*, and 95% confidence interval for *d*-prime (discriminability/sensitivity) and *c* (bias) for the two groups of participants and three transparent gender cues. A *c* value equal to 0 indicates that the participants show no bias in their responses. A positive *c* value indicates a “no”-bias, that is, the participants tend to *underproduce* a specific gender. A negative *c* value is evidence of a “yes”-bias, meaning that a particular gender is *overused* in contexts without the corresponding gender signal. Our analysis reveals that monolinguals have no bias with respect to any of the genders (the 95% confidence interval for *c* for all genders includes 0), while the bilinguals show a “yes-bias” toward masculine (overuse of masculine), and a “no-bias” for the feminine and neuter genders (underuse of feminine and neuter).

Experiment 2: Nonce words

As mentioned above, the Nonce word experiment had three transparent and two ambiguous conditions. Figure 3 compares the results for the M-, F-, and N-transparent conditions for the two participant groups across these conditions in the Nonce and Real word tasks. In the three transparent conditions, children from both groups assign gender more “accurately” (i.e., in accordance with the respective morphophonological cues) to real words than to nonce words. To compare the group performance on the two tasks statistically, we fit a set of generalized linear mixed-effects models to predict gender assignment in accordance with the three transparent gender cues based on the effects of Group (Monolingual vs. Bilingual), Age, Condition (M, F, or N), and Task (Real vs. Nonce). ANOVA model

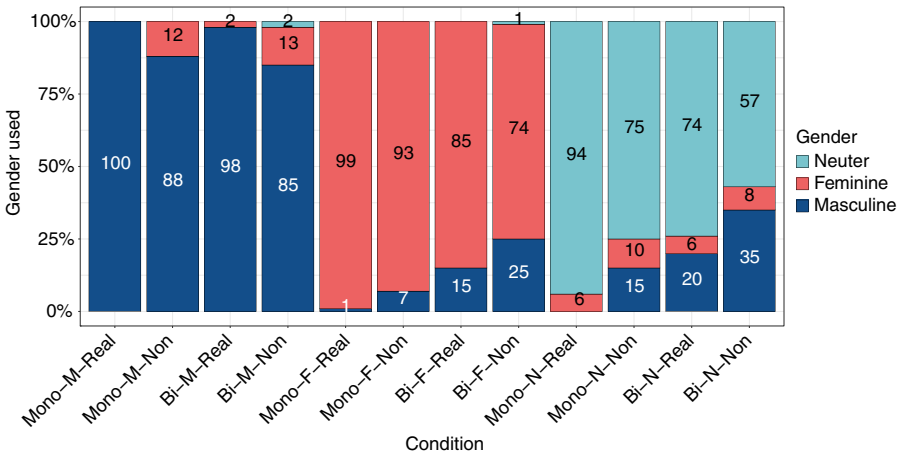


Figure 3. Gender assignment in transparent masculine, feminine, and neuter conditions on the Real and Nonce word tasks by the two groups of participants: Monolinguals versus Bilinguals.

comparison was applied to choose the winning model. The best model included three two-way interactions: Group by Age, Group by Condition, and Group by Task. Participants and items were included as random effects. The model revealed significant effects of Group ($p = .03$), Age ($p = .004$), Condition (M and N, $p < .001$ in both cases), and Task ($p < .001$), as well as a significant interaction of Group and Condition (M, $p < .001$), and Group and Task ($p = .01$; see Table B1 in the appendix). No other effects were significant. Post hoc pairwise comparisons of tasks within groups (controlled for Age and Condition) revealed that both Bilinguals and Monolinguals scored significantly higher on the Real word task than on the Nonce word task (see Table B2 in the Appendix). To sum up, the results of the analysis indicate that both groups perform more target-like on the Real word than the Nonce word task, both groups become more target-like with age, and for both groups, Masculine is significantly easier than Feminine, while Neuter is significantly more challenging than Feminine.

Figure 4 illustrates the use of different genders in the Nonce word experiment in all conditions. As we see, the most common overgeneralization pattern observed in the bilingual groups is the overuse of masculine in all non-masculine conditions (as in the Real word task). We employed generalized linear mixed-effects logistic regression analysis to estimate the effects of Group, Age, and Condition on the probability of a Masculine response. The best model included a three-way interaction of these predictors. Participants and items were included as random intercepts. The results of the model revealed a significant effect of Condition (FN, $p = .02$), a significant interaction of Group and Condition (FN, $p < .001$, N, $p = .03$), as well as significant interactions of Age and Condition (FM, $p = .004$, FN, $p = .008$, and M, $p < .001$) and Group, Age, and Condition (FN, $p < .001$, N, $p = .01$). Post hoc pairwise comparisons indicate that Bilinguals were significantly more likely than the Monolinguals to use Masculine in all conditions except the transparent M condition ($p < .001$ for all contrasts except M). The output of the model and the post hoc pairwise comparisons are presented in Tables B3 and B4 in the appendix.

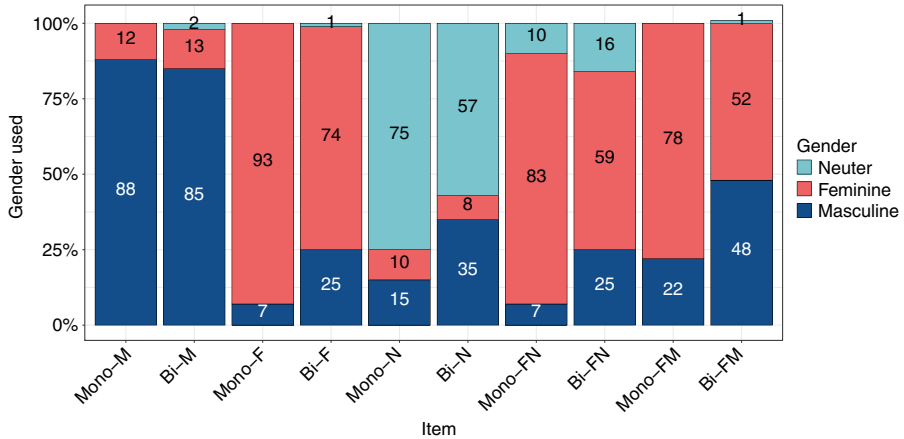


Figure 4. Gender assignment on the Nonce word task by the two groups of participants (Mono = Monolinguals; Bi = Bilinguals) and condition (M = Masculine transparent; F = Feminine transparent; N = Neuter transparent; FN = Feminine/Neuter ambiguous; FM = Feminine/Masculine ambiguous).

Sensitivity to gender cues

We again applied SDT analysis to assess the participants' sensitivity to gender cues, as well as their response bias. We calculated the number and proportion of *hits*, *misses*, *false alarms*, and *correct rejections* for the bilingual and monolingual participants in the three transparent conditions of the Nonce word experiment (see Table B5 in the appendix). Masculine gender has the highest rate of *false alarms* (i.e., cases where a masculine response is used in the absence of a masculine cue), with the proportion being higher for bilinguals (30%) than monolinguals (10%). Note that this proportion is substantially higher than that observed in the Real word experiment: 18% false alarms for masculine in the bilingual group and 1% in the monolingual group (see Table A3 in the appendix). The rate of false alarms for feminine and neuter is lower (i.e., these genders are rarely used in the absence of the corresponding gender cue). Table B6 in the appendix summarizes the mean, variance, *SD*, and 95% confidence interval for *d*-prime (sensitivity) and *c* (bias) for the two groups of participants and three transparent gender cues. Recall that a *d*-prime value close to 0 indicates no sensitivity, while a higher value indicates higher sensitivity. As evident from Table B6, the 95% confidence intervals for *d*-prime values for all genders in both groups are positive and do not include 0. This indicates that monolingual and bilingual children show significant sensitivity to the three transparent gender cues.

A *c* value close to 0 indicates that the participants show no bias in their responses. In the context of our experiment, this means that the participants neither oversupply a specific gender in conditions without this gender cue, nor do they undersupply it in contexts with this cue. A positive *c* value indicates a "no"-bias and a negative *c* value is evidence of a "yes"-bias. Our analysis reveals that monolinguals have no bias in F and M conditions (the 95% confidence interval for *c* in these conditions

includes 0), while they have a “no-bias” (underproduction) in the N condition. The bilinguals show a “yes-bias” in the M condition and a “no-bias” in the F and N conditions (see Table B6).

To compare the sensitivity to the three gender cues between the monolinguals and bilinguals, we calculated individual *d*-prime values for individual participants in each gender. The distributions of *d*-prime values were significantly different from normal, as confirmed by the application of the Shapiro–Wilk normality test ($p < .001$ for all genders and groups). We therefore applied a nonparametric two-sample Mann–Whitney–Wilcoxon test to determine whether the two groups were sampled from populations with identical distributions. The results indicate that the values of *d*-prime for the monolingual group were significantly different from those in the bilinguals in the masculine ($W = 1977$, $p < .001$), feminine ($W = 2130$, $p < .01$), and neuter ($W = 1948$, $p < .001$).

Ambiguous conditions. To estimate sensitivity to gender cues in the two ambiguous conditions, we compared the use of gender agreement in these conditions to the use of these genders in the transparent conditions (the FM condition was compared to the F and M conditions, while the FN condition was compared to the F and N conditions). Based on the group responses in these conditions, we calculated two sensitivity values. The *d*-prime_1 value estimated the sensitivity to the presence of the F cue (for the FM condition, we compared the distribution of responses in the FM and M condition; for the FN condition, we compared the proportions of responses in the FN and N conditions). The *d*-prime_2 value reflected the sensitivity to the presence of the non-F cue. In the FM condition, the *d*-prime_2 value measured the sensitivity to the M cue (we achieved this by comparing the proportions of responses in the FM and F conditions). In the FN condition, the *d*-prime_2 value estimated sensitivity to the N cue (this was achieved by comparing the proportions of responses in the FN condition to the F condition). We consider the two ambiguous conditions individually.

FM condition

Table B7 summarizes the proportions of masculine responses in the ambiguous FM condition as compared to the two control conditions, F and M. Based on this table, we calculated two *d*-prime values to assess sensitivity to the presence of F and M cues. As evident from Table B8, both sensitivity values (*d*-prime_1 and *d*-prime_2) are positive and do not include 0 in their confidence intervals. This suggests that both monolinguals and bilinguals are sensitive to the presence of both M and F gender cues in the FM condition. Finally, we calculated individual *d*-prime_1 and *d*-prime_2 for each participant and compared values using the two-sample Mann–Whitney–Wilcoxon test. The results revealed significant differences between the two groups in their sensitivity to F (*d*-prime_1) in the FM condition ($W = 1468$, $p < .001$). This indicates that monolinguals have a higher sensitivity to the F cue in the FM condition than the bilinguals. At the same time, there was no statistical difference between the groups in their sensitivity to the M cue in the FM condition ($W = 3285$, $p = 0.12$).

FN condition

Table B9 summarizes the proportions of neuter responses in the ambiguous FN condition as compared to the transparent N and F conditions, and Table B10 presents the statistics for the two d-prime values. Similar to the FM condition, monolingual and bilingual children showed sensitivity to both F and N in the ambiguous FN condition; however, the sensitivity to F is higher in both groups. The two-sample Mann–Whitney–Wilcoxon test run on the two d-prime values calculated for individual participants revealed a significant difference between the groups in their sensitivity to the F cue in the FN condition ($W = 2037$, $p = .002$). However, there was no significant difference in the sensitivity to the N cue between the groups ($W = 3119$, $p = .33$).

To sum up, the results of the d-prime analysis revealed that both monolingual and bilingual children show sensitivity to gender cues in the Nonce word experiment. Monolinguals show a significantly higher sensitivity to gender cues in the three nonambiguous conditions, and a higher sensitivity to feminine gender cues in the FM and FN conditions than the bilinguals.

Follow-up: gender assignment with nouns ending in palatals

As evident from Figure 4, the preferred response in the FM condition was feminine, not the masculine default that we had predicted. If we zoom in on individual items of the palatal condition (Figure 5), we can see that (a) Neuter was almost never assigned; (b) Bilinguals generally assign more Masculine than the Monolinguals; and (c) the relative proportions of assigned Masculine/Feminine is very similar across the two groups (with most M assigned for the nonce word *prosh*, and a stepwise decrease of M for *dron*, *knov*, *dryst*, and *klian*). The results indicate that, although Bilinguals generally tend to assign more Masculine than the Monolinguals (arguably, due to defaulting), their sensitivity to fine-grained cues associated with individual nonce items is qualitatively similar to that of the Monolinguals.

In order to investigate this further, we conducted a follow-up experiment with a new group of monolingual Russian children ($n = 85$, aged 3–7). This experiment was identical to Experiment 2, but included 13 items ending in a palatalized consonant. Again, the items conformed to Russian phonotactics and did not have frequent nominal neighbors. The distribution of responses (proportion of F vs. M) is shown in Figure 6.

As we can see, there is considerable variability among palatal-final items with respect to the proportion of masculine agreement they trigger, which seems to be linked to the phonological nature of the word-final palatal (e.g., palatal liquids appear more likely to trigger masculine agreement, while final obstruents tend toward feminine). A possible source of this pattern could be the distribution of palatal-final nouns in the Russian lexicon, where nouns of a particular shape might be more likely to be feminine or masculine. To confirm this intuition, we extracted all noun lemmas ending in a postalveolar/palatal consonant, along with their grammatical gender specification from the Russian National Corpus. The resulting list consisted of 6,309 unique entries, of which 2,127 were masculine. The list was tagged for lemma frequency information, derived from the Frequency Dictionary of Russian (Sharoff, 2002), and all lemmas with the frequency of less than 1 ipm were

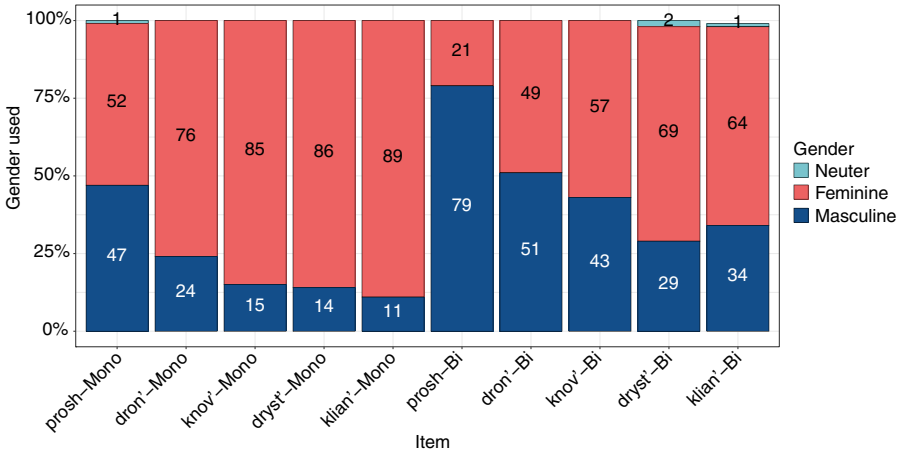


Figure 5. Gender assignment patterns with nonce nouns ending in a palatal by two groups of participants (Russian monolinguals; Russian-German bilinguals).

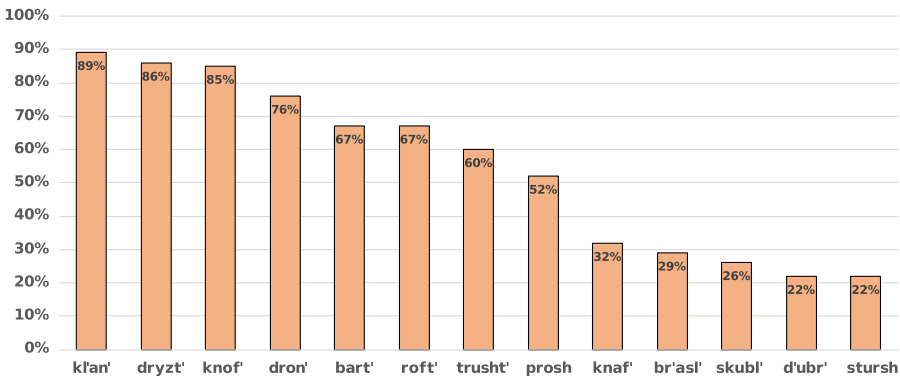


Figure 6. Percentage of feminine agreement with an extended set of nonce nouns ending in a palatal by monolingual Russian children (n = 85, 3–7-year-olds).

eliminated. This resulted in a final list of 1,560 relatively frequent nouns tagged for gender, frequency, and final consonant, of which 603 were masculine. Figure 7 shows the proportion of masculine and feminine noun lemmas for each word-final palatal/postalveolar consonant.

As evident from Figure 7, the relative proportion of masculine and feminine nouns varies considerably among the palatal-final nouns of different shapes. For example, nouns ending in [r'] are masculine in 82% of cases, while nouns ending in [r] are predominantly feminine. Comparing Figures 6 and 7, we see that the responses provided by the children match the corpus distribution quite closely. In fact, for 9 items out of 13, the majority of children assign the gender that is more frequently associated with nouns of this shape in the corpus.

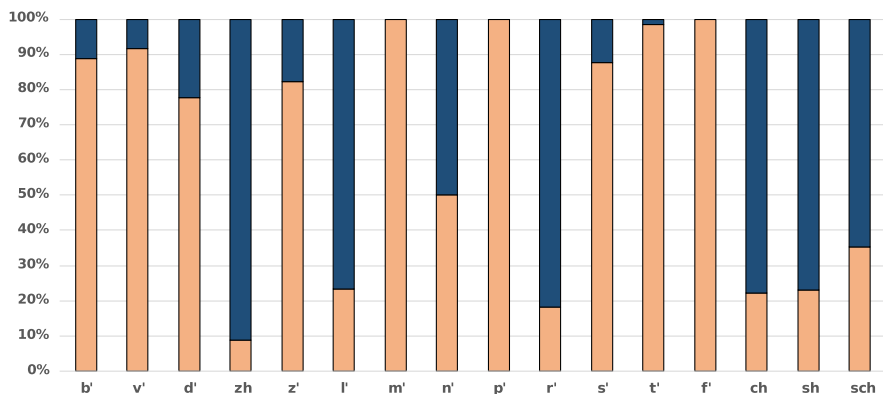


Figure 7. Corpus distribution of masculine and feminine noun lemmas by final consonant.

Let us therefore look more closely at the four items where the children's judgments seem to go against the frequency information: *knaf'* is judged as masculine by 60% of the children, while the only noun ending in [f'] available in our corpus is feminine (*verf', dockyard*). On the other hand, *knov'* (pronounced as [knof'] due to word-final devoicing) is judged as feminine by over 70% of the children. While the treatment of *knov'* aligns nicely with the corpus data (92% of v'-final nouns are feminine), the *knaf'/knov'* asymmetry is interesting and might suggest that the children make use of other formal properties (e.g., final nucleus) when they calculate gender probabilities (62% of the nouns that have [a] as a final nucleus are masculine). Another intriguing possibility is that assigning masculine to *knaf'* could be a result of a defaulting strategy. Since *verf'* is a rare noun, it can be assumed that it is not part of the active vocabulary of (most) children. Therefore, when they encounter a noun ending in [f'], they assign the default gender to it, as they have no previous experience with nouns of this shape. At present, we have no data to convincingly argue for this possibility and therefore leave it for future research (Urek, [submitted](#); Urek et al., [forthcoming/2021](#)).

The results on the remaining three mismatched items also suggest that the children might be using other formal cues in addition to the final consonant: *dron'* and *kljan'* are judged as predominantly feminine, although the corpus data show a 50/50 gender distribution for n'-final nouns. Conversely, sh-final nouns are mostly masculine in the corpus, while the children's judgments on *prosh* are divided almost equally between the two genders. In order to test the role of other formal properties for gender assignment, we further tagged the nouns in our corpus for the number of syllables, initial segment, final nucleus, final segment, final coda, stress, and morphological complexity. The resulting list was then used to train the J48 decision tree algorithm (Weka software, Witten & Frank, 2005; see also Gagliardi, 2012 for gender classification application) to assign gender based on the given phonological properties. After the training, the algorithm achieved about 90% accuracy on real palatal-final nouns, which shows that the gender of palatal-final nouns in Russian is strongly predictable from their phonological shape. When run on the set of 13 novel nouns we had used as stimuli in our follow-up experiment, the algorithm produced a

categorical classification that very closely matched the classification produced by the children – that is, nouns that most children judged to be feminine were also assigned feminine by the algorithm. In fact, the children and J48 mismatched on only two items – *knaf* (judged as feminine by J48) and *prosh* (judged as masculine). Pending further research (e.g., Urek, [submitted](#); Urek et al., [forthcoming/2021](#)), we take these results as a preliminary indication that gender assignment judgments exhibited by children are determined – at least to some extent – by the relative probability of feminine and masculine items of a given phonological shape in the lexicon of Russian.

Experiment 3: mixed cues

Recall that in one-third of the trials of this experiment, the children heard a prompt where the gender agreement cue on the adjective corresponded to the noun-internal phonological gender cue (“Matched cues”), and in two-thirds of the trials, the cue on the adjective did not match the noun-internal phonological cue (“Mismatched cues”). In her response, the child could either follow (a) the adjectival cue in the prompt, (b) the phonological noun-internal cue, or (c) neither. We excluded cases where the children used plural agreement, failed to repeat the noun, or when the response was not identifiable (6% of data in the bilingual dataset and < 1% of data in the monolingual dataset). Figure 8 summarizes gender agreement patterns in Match and Mismatch conditions in the two participant groups: In the Match condition, both groups of children overwhelmingly used gender agreement forms that corresponded to the cue on the noun and the adjective, although bilinguals used substantially more agreement forms that did not match either the adjectival or the nominal cue (between 11% and 12% of the cases).

Figure 9 plots the proportions of gender agreement patterns per participant in the Match cues condition by Age. As evident from the plot, the proportion of responses that match the adjectival/nominal cues increases with age in both groups. To model the results statistically, we fit a generalized linear mixed-effects model where the probability of following the cues was predicted based on a three-way interaction of Group (Monolingual vs. Bilingual), Condition (FF, MM, or NN), and Age (continuous). Participants and items were included as random intercepts. The model did not reveal any significant effects. The output of the model is summarized in the Appendix (Table C1).

Figures 10 and 11 plot the proportions of cases (per participant) where the participants followed the adjectival cue (Figure 10) or the nominal cue (Figure 11) in the Mismatch cues condition by Age. The proportion of responses that match the adjectival cue increases, and the proportion of responses that match the nominal cue decreases with age in both groups. To model the results statistically, we fit two generalized linear mixed-effects model where the probability of following the adjectival or the nominal cue was predicted based on a three-way interaction of Group (Monolingual vs. Bilingual), Condition (FF, MM, or NN), and Age (continuous). Participants and items were included as random intercepts. The first model predicted the probability of following the adjectival cue and revealed a significant effect of Age ($p = .03$), Cue (MF, $p = .02$, and MN, $p = .01$), and a significant interaction of Group and Cue (MF, $p = .007$; see Table C2 in the Appendix). The second model predicted the probability of following the nominal cue. The results revealed

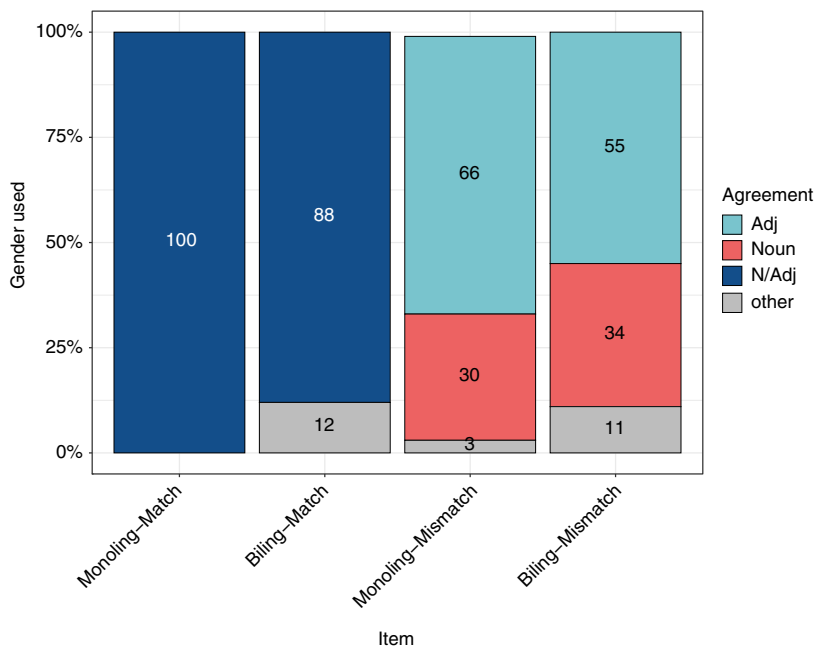


Figure 8. Gender agreement patterns in Match and Mismatch conditions for Monolinguals and Bilinguals.

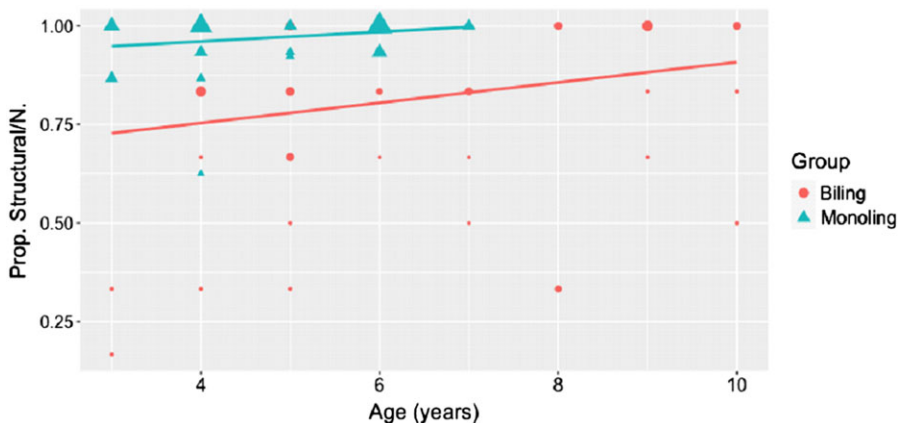


Figure 9. Matching cues. Proportion of responses (per participant) that followed the adjectival and nominal cues in the prompt by Age.

significant effects of Age ($p < .001$) and Cue (MF, $p = .01$, MN, $p = .004$, and NF, $p = .01$), significant interactions of Group and Cue (MF, $p = .006$) and Age and Cue (MF, $p = .008$, MN, $p = .006$, and NF, $p = .006$), and a significant three-way interaction of Group, Age, and Cue (MF, $p = .03$; see Table C3 in the Appendix).

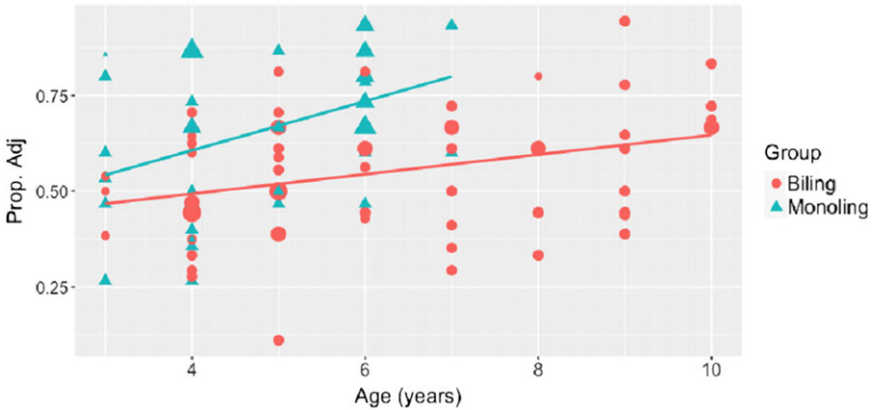


Figure 10. Mismatching cues. Proportion of responses (per participant) that followed the adjectival cue by Age.

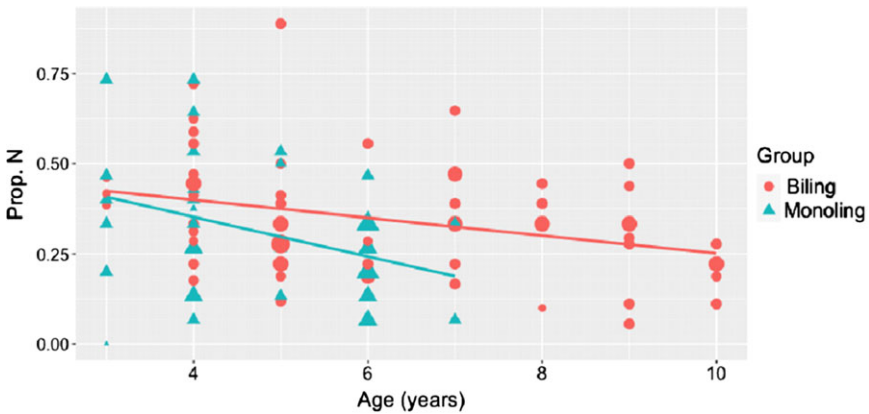


Figure 11. Mismatching cues. Proportion of responses (per participant) that followed the nominal cue by Age.

To sum up, the models confirmed that the importance of the syntactic cue (the adjectival gender cue) significantly increased, and the importance of the nominal cue significantly decreased with Age for both monolingual and bilingual participants. The fixed effect of Group and the interaction of Group and Age were not significant, which is in line with the hypothesis that Bilinguals and Monolinguals generally follow the same developmental path, gradually becoming more and more adult-like in recognizing the adjectival agreement cue as the most reliable predictor of the noun’s grammatical gender.

Discussion

We now return to our four research questions, which we repeat before the discussion of each of them.

RQ1: To what extent are monolingual and bilingual children sensitive to gender cues in Russian and to what extent do they rely on memorized lexical knowledge?

The answer to this question is that both monolinguals and bilinguals clearly rely on morphophonological cues to assign gender, as evident from significant sensitivity coefficients (d-prime values) for both groups of learners. Bilinguals default to masculine more than monolinguals, as indicated by a significant bias toward masculine in all non-masculine conditions in both Real and Nonce word tasks. When comparing the children's performance on Real and Nonce word tasks, we observed that there is a significantly higher accuracy on real than nonce words in both groups, indicating that lexical knowledge also plays a role, in addition to the cue-driven assignment. This resonates with the processing account in Bordag *et al.* (2006) discussed above, which argues that with increased experience and proficiency in the language, the connections necessary to activate gender selection of familiar words become weaker, as the direct connection between the gender node and the specific nouns becomes stronger and more automated. We thus propose the following model for the children's behavior: when presented with a real word and requested to provide an adjective agreeing in gender, a monolingual child will typically not perform online computation of the noun's gender based on its formal properties. Rather, the child extracts gender information from the relevant lexical entry and chooses an adjectival gender exponent based on this information. The difference between monolingual and bilingual children on this task is due to the former group having developed stronger connections between familiar words and their respective gender nodes, while the latter group must to a larger extent rely on the phonological cues.

When the children need to produce a gendered adjective agreeing with a novel noun, the task is quite different. In this case, both monolinguals and bilinguals need to resort to computation, as there is no stored connection between the nonce word and a specific gender node. We assume that in this case the child first performs an online analysis of the formal properties of the noun. Since the form of the noun in the nominative provides cues to the noun's gender in Russian, and the relevant morphophonological information is contained at the right edge of words, the child presumably pays attention predominantly to word-final segments. When a gender cue on the given noun is identified (e.g., stressed [-á] or a palatal consonant), the child can categorize a novel noun with the nouns that have a similar ending (i.e., gender cue) in her mental lexicon, and form a "gender hypothesis" for the novel noun. Given the young age of our participants, the connections between the phonological cues and the respective gender nodes are also weaker in bilinguals (unlike the L2 adults investigated in Bordag *et al.* 2006), accounting for their lower accuracy compared to the monolinguals also on this task.

In case of highly reliable near-deterministic cues such as stressed [-á], this gender hypothesis is near-deterministic as well (the only type of non-cue-driven response that we expect in this case should be in the direction of the grammatical default, i.e., masculine in Russian). In case of ambiguous, probabilistic cues, such as palatal consonants, the gender hypothesis the child forms will also be probabilistic in nature. Thus, if say, 70% of palatal-final nouns contained in her current lexicon are feminine, she will assign feminine to palatal-final novel nouns 70% of the time. The level of granularity that is relevant for children in forming probabilistic gender

hypotheses is an empirical question. Based on data from monolingual Russian adults, Urek ([submitted](#)) argues that at least three scenarios are possible: (a) children are generally sensitive to the fact that palatal-final nouns in Russian can be either feminine or masculine, and as a result novel nouns of this shape get assigned masculine or feminine gender at chance; (b) children are sensitive to the fact that feminines constitute about 70% of all palatal-final lemmas, and therefore assign feminine to palatal-final nouns 70% of the time; (c) children are sensitive to probability distributions associated with each individual word-final palatal and mimic those in gender assignment. We leave further investigation of this question for future research.

RQ2: Do young children make fine-grained linguistic distinctions in gender assignment from early on (cf., the micro-cue model, Westergaard, 2009a, b, 2014)? More specifically, are monolingual and bilingual children sensitive to ambiguous gender cues in Russian (i.e., stem-stressed nouns ending in a vowel (FN) or palatalized consonants (MF)) or do they default to masculine when exposed to nonce words with such cues?

The answer to RQ2 constitutes one of the most interesting findings of our study, viz. that when children are exposed to nonce words with ambiguous cues (either FN ending in a schwa or MF ending in a palatalized consonant), they do NOT default to masculine across the board. In the FN condition, both monolinguals and bilinguals predominantly use F gender, which may reflect overgeneralization to the most frequent gender compatible with the available cue. In the Real word task, overgeneralization to feminine is also more pronounced with neuter nouns ending in a schwa than with neuter nouns ending in a stressed [-o] in both groups of participants. Thus, both monolingual and bilingual children seem to make a distinction between a final schwa (which can be a cue for feminine) and a final [-o] (which is incompatible with feminine). This experimental finding corresponds to previous findings from spontaneous production mentioned above (Gvozdev, 1961; Popova, 1973; Polinsky, 2008).

Surprisingly, both participant groups also show a preference for the feminine in the ambiguous M/F condition, which was unexpected, given previous findings from spontaneous production reporting nontarget-consistent use of *masculine* with existing feminine words ending in palatals (cf., Gvozdev, 1961; Ceitlin, 2009). Our follow-up study of this phenomenon based on Russian corpus data revealed that some palatalized endings are characteristic of F and others of M (cf., above). On closer inspection, the test items in our Nonce word experiment (M/F condition) predominantly contained endings characteristic of F. This finding indicates that both monolinguals and bilinguals are sensitive to even finer gender cues than what has been reported in the literature based on production. As mentioned under RQ1, further research is needed to investigate this intriguing phenomenon, and some is already under way (Urek, [submitted](#)).

RQ3: Which cues are Russian-speaking monolingual and bilingual children most sensitive to for gender assignment, nominal endings or gender agreement? Does this change with age?

Our results from Experiment 3 (mismatched cues) mirror the findings in Karmiloff-Smith (1979); see also Gagliardi & Lidz, (2014); Culbertson et al. (2019) among others. The probability of following the gender agreement on the adjective increases

significantly with age for both monolinguals and bilinguals, while the probability of following the cue on the noun decreases. At the same time, the effect of group is significant too, with the monolinguals being more likely than the bilinguals to follow the agreement cue on the adjective and disregard the cue on the noun. Thus, children with less exposure are less likely to disregard the phonological cue on the noun and tend to change/regularize the gender agreement on the adjective to produce a “harmonic” (i.e., matched) adjective–noun combination. This finding provides further evidence that young children are sensitive to gender cues. With more linguistic experience, both monolingual and bilingual children become more likely to produce a mismatched adjective–noun combination, indicating that they have learned that there are exceptions and to pay more attention to the syntactic cue (which is of course the exponent of gender and as such 100% reliable). This means that they gradually become able to disregard nominal cues if there is a mismatch between the adjectival and the nominal cue.

RQ4: Are there only quantitative or also qualitative differences between monolinguals and bilinguals? That is, are there signs of a reduced gender system, for example, a two-gender system in the bilingual data (cf., findings from heritage language adults in Polinsky, 2008) or no gender system at all (cf., findings from bilingual children in Rodina & Westergaard, 2017)?

As mentioned above, the bilinguals used the M default strategy more than the monolinguals, which is likely due to overall less input in the heritage language. However, we did not find any children who produced only M and thus seemed to have no gender at all (as was found for Norwegian–Russian bilinguals in Rodina & Westergaard, 2017). The reason for this is presumably that the Russian heritage children in Germany have a higher proficiency in the heritage language, mainly due to a later onset of acquisition of the majority language (cf., Rodina *et al.*, 2020). Furthermore, we did not find signs of a robust two-gender system (M and F) based on the contrasts “final C-nouns” versus “final V-nouns”, as was found for adult heritage speakers in the USA (Polinsky, 2008). Although the neuter condition is the most error-prone condition in all experiments and for all participants, the errors appearing with neuter nouns, especially nouns ending in a stressed [-ó], mainly constitute overuse of the masculine, not the feminine. Possible explanations for this are that masculine agreement is syntactically unmarked/underspecified, and/or that neuter is attracted to masculine rather than to feminine due to a substantial paradigm overlap between M and N in the adjectival and nominal declension (see above).

Summary and outlook

This paper has investigated monolingual and bilingual children’s sensitivity to morphophonological cues for grammatical gender, both major rules (transparent cues) and more fine-grained distinctions in ambiguous contexts. In order to answer our research questions, we carried out three elicitation experiments on a group of 64 Russian heritage children growing up in Germany (aged 3–10) as well as 107 monolingual controls living in Russia (aged 3–7). We focused on the three transparent cues for Russian gender assignment, masculines ending in a non-palatalized consonant (-C), feminines ending in a stressed [-á], and neuters ending in a stressed [-ó], as well as two ambiguous cues, nouns ending in a schwa, which may be either

feminine or neuter (FN), and nouns ending in a palatalized consonant, which are either masculine or feminine (MF). Experiment 1 focused on existing nouns expected to be familiar to the children, Experiment 2 on nonce nouns with the same cues (three transparent, two ambiguous), while Experiment 3 consisted of noun phrases with mismatched cues, where the ending of the noun cued one (transparent) gender, while the gender on an agreeing adjective was a different gender. Our findings show that both monolingual and bilingual children are sensitive to the transparency of the gender cues in the Real word task, with all groups of children performing significantly more accurately with nouns having transparent gender cues than nouns without such cues. Furthermore, younger bilinguals performed significantly less accurately than the monolinguals and older bilinguals and defaulted to masculine significantly more than the other groups of children. Results from the Nonce word task show that all children are sensitive to the gender cues, as indicated by a significant sensitivity to cue indices in all conditions for both monolinguals and bilinguals, obtained via a *d*-prime analysis. The bilinguals showed a significantly higher bias toward masculine and a significantly lower sensitivity to the three transparent gender cues than their monolingual peers. However, surprisingly, overuse of masculine cannot explain the children's behavior in the two ambiguous conditions, which contained FN and MF cues, for neither monolingual nor bilingual children. Both groups of children display a preference for the feminine, not only for the FN cue, but also for the MF cue. In the former case, this is presumably due to a generally higher frequency of feminine nouns ending in a schwa, while in the latter case, the result indicates that children are sensitive to even more fine-grained cues related to the type of palatal consonant, stress, and quality of the final vowel.

In the mismatch experiment, we attested an age effect with respect to the type of cue children are more sensitive to, with the likelihood to follow the syntactic cue on the adjective significantly increasing with age in both monolingual and bilingual children, echoing findings from French monolinguals in Karmiloff-Smith (1979), as well as results reported in Gagliardi & Lidz (2014) and Culbertson et al. (2019).

The results of the current study lend some support to approaches to acquisition such as the micro-cue model (Westergaard, 2009a, b, 2014), in that they show that children are sensitive to fine distinctions in the input from relatively early on, not only in syntax, but also for complex morphosyntactic phenomena such as grammatical gender: While there are certain overgeneralizations in early child data, the sensitivity to cues for gender assignment is visible in the data from early on, not only for transparent, but also for ambiguous cues, and not only for monolinguals, but also for bilinguals, who typically have comparatively less input. Our findings also open up new questions and should lead to further research, for example, to what extent the sensitivity to fine-grained cues found in our follow-up study on nouns ending in a palatal consonant may hold in a larger study including different populations and a larger variety of items.

Acknowledgments. This research was supported by a grant from the Research Council of Norway for the project Micro-Variation in Multilingual Acquisition and Attrition Situations (MiMS), project number 250857. The paper was completed while the authors were working on the international research project MultiGender at the Centre for Advanced Study at the Norwegian Academy of Science and Letters in Oslo during the academic year 2019–2020. We thank Tanja Kupisch and Natalia Gagarina for their help with participant recruitment, and Anna Afanasieva for her help in transcribing the data.

Notes

- 1 We follow Shvedova (1980) for the numeration of the declension classes.
- 2 Here and further, Russian examples are given following the transliteration convention adopted in Corbett (1991). Where relevant, International Phonetic Alphabet (IPA) transcriptions are also provided in square brackets.
- 3 The corpus analysis in Slioussar (2018) became available only after we designed and conducted our study. It is noteworthy that the generalizations in Slioussar (2018) overlap with the results of our acquisition study.
- 4 The project was registered and approved by the Norwegian Social Science Data Service (NSD, <http://www.nsd.uib.no>). Data collection was conducted in accordance with NSD's ethical principles. Written informed consent was obtained from parents of all the participants prior to testing.
- 5 We thank the anonymous reviewer who suggested this analysis.
- 6 We added 0.5 to each rate value to circumvent the problem of extreme values (0 and 1), see Hautus (1995).

References

- Blom, E., Polišká, D., & Weerman, F. (2008). Articles, adjectives and age of onset: The acquisition of Dutch grammatical gender. *Second language research*, 24(3), 297–331.
- Bobrova, M. (2013). Genus ved homophone substantiv [Gender in homophonous nouns]. Masters thesis, University of Oslo.
- Bordag, D., Opitz, A., Pechmann, T. (2006). Gender processing in first and second languages: The role of noun termination. *Journal of Experimental Psychology: Learning, Memory, and Cognition*, 32(5), 1090–1101.
- Ceitlin, S. N. (2005). Kategorija roda v detskoj rechi [The category of gender in child language]. In A. V. Bondarenko & S. A. Shubik (eds.), *Problemy funkcionalnoj grammatiki [Issues in functional grammar]* (pp. 346–375). St Petersburg, Russia: Rossijskaja Akademija Nauk.
- Ceitlin, S. N. (2009). *Očerki po slovoobrazovaniju i formoobrazovaniju v detskoj rechi. [On Inflection and Derivation in Child Language]*. Znak.
- Corbett, G. G. (1982). Gender in Russian: An account of gender specification and its relationship to declension. *Russian linguistics*, 197–232.
- Corbett, G. G. (1991). *Gender*. Cambridge University Press. <https://doi.org/10.1017/CBO9781139166119>
- Culbertson, J., Gagliardi, A., & Smith, K. (2017). Competition between phonological and semantic cues in noun class learning. *Journal of Memory and Language*, 92, 343–358.
- Culbertson, J., Jarvinen, H., Haggarty, F., & Smith, K. (2019). Children's sensitivity to phonological and semantic cues during noun class learning: Evidence for a phonological bias. *Language*, 95(2), 268–293.
- Egger, E., Hulk, A., & Tsimpli, I. M. (2018). Crosslinguistic influence in the discovery of gender: the case of Greek–Dutch bilingual children. *Bilingualism: Language and Cognition*, 21(4), 694–709.
- Eichler, N., Jansen, V., & Müller, N. (2013). Gender acquisition in bilingual children: French–German, Italian–German, Spanish–German and Italian–French. *International Journal of Bilingualism*, 17(5), 550–572.
- Gagarina, N., Klop, D., Kunnari, S., Tantele, K., Välimaa, T., Balčiūnienė, I., Bohnacker, U., & Walters, J. (2012). *Multilingual Assessment Instrument for Narratives (MAIN)*. ZAS papers in linguistics, 56. ZAS. Retrieved from <http://www.zas.gwz-berlin.de/zaspil56.html>
- Gagliardi, A., & Lidz, J. (2014). Statistical insensitivity in the acquisition of Tsez noun classes. *Language* 90, 1–32.
- Gagliardi, A. C. (2012). *Input and intake in language acquisition*. Doctoral dissertation, University of Maryland.
- Gourevitch, V., and Galanter, E. (1967). A significance test for one parameter isosensitivity functions. *Psychometrika* 32, 25–33. <https://doi.org/10.1007/BF02289402>.
- Gvozdev, A. N. (1961). *Formirovanie u rebenka grammatičeskogo stroja russkogo jazyka. [Language Development of a Russian Child]*. APN RSFSR.
- Halle, M., & Matushansky, O. (2006). The morphophonology of Russian adjectival inflection. *Linguistic Inquiry* 37, 351–404.
- Hautus, M. J. (1995). Corrections for extreme proportions and their biasing effects on estimated values of d'. *Behavior Research Methods, Instruments, & Computers*, 27, 46–51. <https://doi.org/10.3758/BF03203619>.
- Hockett, C. F. (1958). *A course in modern linguistics*. New York: Holt/Rinehart and Winston.

- Horst, J. S., & Hout, M. C.** (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.
- Huang Y., & Ferreira F.** (2020). The application of signal detection theory to acceptability judgments. *Frontiers in Psychology*, *11*. <https://doi.org/10.3389/fpsyg.2020.00073>.
- Ilola, E., & Mustajoki, A.** (1989). *Report on Russian morphology as it appears in Zaliznyak's grammatical dictionary*. Helsinki University Press.
- Janssen, B. E.** (2014). Frequency effects on the acquisition of Polish and Russian gender morphology. In E. Fortuin, P. Houtzagers, J. Kalsbeek & S. Dekker (ed.), *Dutch Contributions to the Fifteenth International Congress of Slavists*, Minsk, August 20-27, 2013 (Studies in General and Slavic Linguistics, 40) (pp. 109–126). Rodopi.
- Kaltsa, M., Tsimpli, I. M., & Argyri, F.** (2019). The development of gender assignment and agreement in English-Greek and German-Greek bilingual children. *Linguistic Approaches to Bilingualism*, *9*(2), 253–288.
- Karmiloff-Smith, A.** (1979). *A functional approach to child language: A study of determiners and reference*. Cambridge: Cambridge University Press.
- Kempe, V., Brooks, P. J., Mironova, N., & Fedorova, O.** (2003). Diminutivization supports gender acquisition in Russian children. *Journal of Child Language*, *30*(2), 471–485.
- Köpcke, K.-M., & Zubin, D. A.** (1983). Die kognitive Organisation der Genuszuweisung zu den einsilbigen Nomen der Deutschen Gegenwartssprache. *Zeitschrift für Germanistische Linguistik*, *11*(2), 166–182.
- Kupisch, T., Mitrofanova, N., & Westergaard, M.** (in prep.). The role of formal gender cues in the acquisition of German as an L1 and early L2: Experiments with real and nonce words.
- Kupisch, T., Müller, N., & Cantone, K. F.** (2002). Gender in monolingual and bilingual first language acquisition: Comparing Italian and French. *Lingue e linguaggio*, *1*(1), 107–150.
- Lazova, M. V.** (ed.) (1974). *Obratnyj slovar' russkogo jazyka: Okolo 125 000 slov.* [Reverse dictionary of Russian: Approximately 125 000 words.] Sovetskaja Enciklopedija, Moskva.
- Lenth, R., Singman, H., Love, J., Buerkner, P., & Herve, M.** (2020). R package emmeans: Estimated Marginal means, aka Least-Squares Means. R repository: <https://cran.r-project.org/web/packages/emmeans/index.html>
- Macmillan, N., & Creelman, C.** (2004). *Detection theory: A user's guide*. Psychology Press.
- Mitrofanova, N., Rodina, Y., Urek, O., & Westergaard, M.** (2018). Bilinguals' sensitivity to grammatical gender cues in Russian: the role of cumulative input, proficiency, and dominance. *Frontiers in Psychology*, <https://doi.org/10.3389/fpsyg.2018.01894>.
- Montrul, S., de la Fuente, I., Davidson, J., & Foote, R.** (2013). The role of experience in the acquisition and production of diminutives and gender in Spanish: Evidence from L2 learners and heritage speakers. *Second Language Research*, *29*(1), 87–118.
- Mučnik, I. P.**, (1971). Grammatičeskie kategorii glagola i imeni v sovremenom russkom literaturnom jazyke [The grammatical categories of verb and noun in Russian]. Nauka, Moskva, 177–244.
- Nesset, T.** (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.
- Pallier, C.** (2002). Computing discriminability and bias with the R software. Available online: <https://www.pallier.org/pdfs/aprime.pdf>
- Peterson, W. W. T. G., Birdsall, T., & Fox, W.** (1954). The theory of signal detectability. *Transactions of the I.R.E. Professional Group on Information Theory*, *4*, 171–212. <https://doi.org/10.1109/TIT.1954.1057460>
- Pinker, S.** (1999). *Words and Rules: The Ingredients of Language*. Weidenfeld and Nicolson.
- Polinsky, M.** (2008). Gender under incomplete acquisition: heritage speakers' knowledge of noun categorization. *Heritage Language Journal*, *6*, 40–71.
- Popova, M. I.** (1973). **Grammatical elements of language in the speech of pre-preschool children.** In C. A. Ferguson & D. I. Slobin (eds.), *Studies of child language development* (pp. 269–280). New York: Holt, Rinehart and Winston.
- Radford, A.** (1992). The acquisition of the morphosyntax of finite verbs in English. In J. M. Meisel (ed.), *The Acquisition of Verb Placement: Functional Categories and V2 Phenomena in Language Acquisition* (pp. 23–62). Kluwer Academic Publishers.
- Rodina, Y.** (2008). *Semantics and morphology: The acquisition of grammatical gender in Russian*. PhD dissertation, University of Tromsø.
- Rodina, Y.** (2017). Narrative abilities of preschool bilingual Norwegian-Russian children. *International Journal of Bilingualism*, *21*, 617–635.

- Rodina, Y., & Westergaard, M.** (2012). A cue-based approach to the acquisition of grammatical gender in Russian. *Journal of child language*, **39**(5), 1077–1106.
- Rodina, Y., & Westergaard, M.** (2013). The acquisition of gender and declension in a non-transparent system: monolinguals and bilinguals. *Studia Linguistica*, **67**(1), 47–67.
- Rodina, Y., & Westergaard, M.** (2015). Grammatical gender in Norwegian: Language acquisition and language change. *Journal of Germanic Linguistics*, **27**(2), 145–187.
- Rodina, Y., & Westergaard, M.** (2017). Grammatical gender in bilingual Norwegian-Russian Acquisition: the role of input and transparency. *Bilingualism: Language and Cognition*, **20**, 197–214. <https://doi.org/10.1017/S1366728915000668>
- Rowland, C. F.** (2007). Explaining errors in children's questions. *Cognition*, **104**(1), 106–134.
- Schwartz, M., Minkov, M., Dieser, E., Protassova, E., Moin, V., & Polinsky, M.** (2015). Acquisition of Russian gender agreement by monolingual and bilingual children. *International Journal of Bilingualism*, **19**(6), 726–752.
- Sharoff, S.** (2002). Meaning as use: exploitation of aligned corpora for the contrastive study of lexical semantics. *Proceeding of Language Resources and Evaluation Conference (LREC02)*. May, 2002, Las Palmas, Spain.
- Shvedova, N., ed.** (1980). *Russkaya grammatika [Russian grammar]*. Nauka, Moscow.
- Slioussar, N., & Samoilova, M.** (2014). A database to estimate frequencies of different grammatical features and inflectional affixes in Russian nouns. In *The 9th International Conference on the Mental Lexicon*. Niagara-on-the-Lake: Brock University and McMaster University. 104–105.
- Slioussar, N. A.** (2018). Gender, declension and stem-final consonants: an experimental study of gender agreement in Russian. *Computational Linguistics and Intellectual Technologies. Papers from the Annual International Conference "Dialogue"*, 688–700.
- Snyder, W.** (2007). *Child language: The parametric approach*. Oxford University Press.
- Taraskenkova, O.** (2010). *Acquisition of Agreement and Morphological Features Within the Noun Phrase in Russian*. Doctoral Dissertation. University of Connecticut. <https://opencommons.uconn.edu/dissertations/AAI3451400>
- Trosterud, T.** (2001). Genustilordning i norsk er regelstyrt [Gender assignment in Norwegian is rule-based]. *Norsk lingvistisk tidsskrift*, **19**, 29–57.
- Unsworth, S.** (2014). **Comparing the role of input in bilingual acquisition across domains.** In T. Grüter & J. Paradis (eds.), *Input and experience in bilingual development* (pp. 181–201). Amsterdam: John Benjamins.
- Urek, O.** (Submitted). Categorization of ambiguously cued novel nouns: evidence from grammatical gender assignment in Russian.
- Urek, O., Lohndal, T., & Westergaard, M.** (Forthcoming/2021). En splyv eller et splyv? Tilordning av grammatisk genus til nonord-substantiv i norsk. *Norsk lingvistisk tidsskrift*.
- Wegener, H.** (1995). *Die Nominalflexion des Deutschen - verstanden als Lerngegenstand*. Niemeyer.
- Westergaard, M.** (2009a). *The Acquisition of Word Order: Micro-cues, information structure, and economy*. John Benjamins Publishing.
- Westergaard, M.** (2009b). Usage-based vs. Rule-based Learning: The Acquisition of Word Order in Wh-Questions in English and Norwegian. *Journal of Child Language*, **36**(5), 1023–1051.
- Westergaard, M.** (2014). Linguistic variation and micro-cues in first language acquisition. *Linguistic Variation*, **14**(1), 26–45.
- Witten, I. H., & Frank, E.** (2005). *Data mining: Practical machine learning tools and techniques*. Morgan Kaufmann.
- Yang, C.** (2016). *The price of linguistic productivity: How children learn to break the rules of language*. MIT press.
- Zaliznjak, A. A.** (1967). *Russkoje imennoje slovoizmenenije [Russian nominal inflection system]*. Moscow.
- Zaliznjak, A. A.** (1977). *Grammaticheskij slovarj russkogo jazyka [Grammatical dictionary of Russian]*. Moscow.

Appendix. Details of the statistical analysis

Table A1. Real words (Accuracy predicted by Condition and Age Group and their interaction)

```

Generalized linear mixed model fit by maximum likelihood (Laplace Approximation) ['glmerMod']
Family: binomial ( logit )
Formula: Accuracy ~ Group*Age +Condition * Group + (1 | Code) + (1 | Item)
Bilinguals & F-Pal condition as Intercept

Fixed effects:
              Estimate Std. Error z value Pr(>|z|)
(Intercept)   -1.7346    1.6279   -1.066 0.286618
GroupMono      1.6672    3.8049    0.438 0.661265
Age            0.6384    0.2614    2.442 0.014610 *
ConditionF-tr  0.9396    0.6769    1.388 0.165098
ConditionM-Pal 2.0015    0.7162    2.795 0.005196 **
ConditionM-tr  2.8935    0.8056    3.592 0.000328 ***
ConditionN-tr -0.9443    0.6720   -1.405 0.159973
ConditionN-Unstr -0.9116    0.6586   -1.384 0.166334
GroupMono:Age -0.1334    0.7636   -0.175 0.861321
GroupMono:ConditionF-tr 2.1271    1.3422    1.585 0.113019
GroupMono:ConditionM-Pal 1.1026    1.3453    0.820 0.412419
GroupMono:ConditionM-tr 16.3512  3651.1896  0.004 0.996427
GroupMono:ConditionN-tr 19.6434  3719.2087  0.005 0.995786
GroupMono:ConditionN-Unstr 1.2744    0.7695    1.656 0.097687 .
---
Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
    
```

Table A2. Real words: Post hoc pairwise comparisons of groups within conditions

```

> contrast(ml.emm, "pairwise")
Condition = F-Pal:
contrast              estimate          SE df z.ratio p.value
Biling 5.3221850613155 - Mono 5.3221850613155   -0.957    0.893 Inf -1.072 0.2835

Condition = F-tr:
contrast              estimate          SE df z.ratio p.value
Biling 5.3221850613155 - Mono 5.3221850613155   -3.084    1.462 Inf -2.110 0.0349

Condition = M-Pal:
contrast              estimate          SE df z.ratio p.value
Biling 5.3221850613155 - Mono 5.3221850613155   -2.060    1.461 Inf -1.410 0.1585

Condition = M-tr:
contrast              estimate          SE df z.ratio p.value
Biling 5.3221850613155 - Mono 5.3221850613155  -17.308 3651.190 Inf -0.005 0.9962

Condition = N-tr:
contrast              estimate          SE df z.ratio p.value
Biling 5.3221850613155 - Mono 5.3221850613155  -20.601 3719.209 Inf -0.006 0.9956

Condition = N-Unstr:
contrast              estimate          SE df z.ratio p.value
Biling 5.3221850613155 - Mono 5.3221850613155   -2.232    0.911 Inf -2.449 0.0143

Results are given on the log odds ratio (not the response) scale.
    
```

Table A3. Proportion (and number) of responses in each of the four categories defined by the signal detection analysis for monolinguals and bilinguals in the masculine, feminine, and neuter **transparent conditions** (Real words experiment)

	Hits	Misses	False Alarms	Correct rejections
Bilinguals Masculine	0.98 (336)	0.02 (8)	0.18 (115)	0.88 (539)
Monolinguals Masculine	0.99 (83)	0.00 (0)	0.01 (1)	0.99 (138)
Bilinguals Feminine	0.85 (281)	0.15 (50)	0.03 (26)	0.92 (641)
Monolinguals Feminine	0.98 (78)	0.02 (1)	0.00 (0)	0.99 (143)
Bilinguals Neuter	0.74 (240)	0.26 (83)	0.00 (0)	0.99 (675)
Monolinguals Neuter	0.94 (57)	0.06 (3)	0.00 (0)	0.99 (162)

Table A4. Mean, variance (*SD*), and 95% confidence interval for sensitivity (*d'*) and bias (*c*) with respect to M, F, and N by monolinguals and bilinguals in **transparent conditions** (Real word experiment)

	<i>d'</i>			<i>c</i>		
	Mean	var (<i>SD</i>)	95% conf. int.	mean	var (<i>SD</i>)	95% conf. int.
Bilinguals Masculine	2.9	0.02 (0.1)	2.64 to 3.15	-0.52	0.004 (0.1)	-0.78 to -0.26
Monolinguals Masculine	4.82	0.31 (0.6)	3.73 to 5.9	-0.11	0.08 (0.28)	-1.2 to 0.98
Bilinguals Feminine	2.78	0.02 (0.1)	2.51 to 3.05	0.36	0.05 (0.07)	0.09 to 0.64
Monolinguals Feminine	4.78	0.4 (0.7)	3.45 to 6.11	0.31	0.12 (0.34)	-1.02 to 1.64
Bilinguals Neuter	3.83	0.35 (0.6)	2.66 to 5	1.26	0.08 (0.3)	0.1 to 2.43
Monolinguals Neuter	5.14	0.7 (0.8)	3.52 to 6.76	0.17	0.17 (0.4)	-1.45 to 1.79

Table B1. Transparent conditions Real versus Nonce words (Accuracy predicted by the interactions of Group and Age, Condition and Group, and Group and Task)

```

Generalized linear mixed model fit by maximum likelihood (Laplace Approximation)
[‘glmerMod’] Family: binomial ( logit )

Formula: Accuracy ~ Group * Age + gender * Group + Group * task + (1 | Code) + (1 | Item)
Data: datal
Control: glmerControl(optimizer = "bobyqa")

Fixed effects:
            Estimate Std. Error z value Pr(>|z|)
(Intercept) -0.17748    0.61347  -0.289  0.77235
GroupMono   2.14860    0.98673   2.177  0.02944 *
Age          0.26966    0.09251   2.915  0.00356 **
genderM      1.10047    0.28904   3.807  0.00014 ***
genderN     -1.13577    0.26930  -4.217  2.47e-05 ***
taskReal     1.25368    0.30801   4.070  4.70e-05 ***
GroupMono:Age -0.05039    0.17145  -0.294  0.76884
GroupMono:genderM -1.69042    0.35603  -4.748  2.05e-06 ***
GroupMono:genderN -0.58190    0.32300  -1.802  0.07161 .
GroupMono:taskReal 2.79341    1.12074   2.492  0.01269 *
---
Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
    
```

Table B2. Post hoc pairwise comparisons of tasks within groups: Transparent cues

```

Group = Biling, Age = 5.56:
contrast      estimate      SE  df z.ratio p.value
Nonce - Real   -1.25  0.308 Inf -4.070 <.0001

Group = Mono, Age = 5.56:
contrast      estimate      SE  df z.ratio p.value
Nonce - Real   -4.05  1.100 Inf -3.680 0.0002

Results are averaged over the levels of: gender
Results are given on the log odds ratio (not the response) scale.
    
```

Table B3. Nonce words (Probability of M agreement predicted by the interaction of Group, Age, and Condition)

```

Generalized linear mixed model fit by maximum likelihood (Laplace Approximation)
['glmerMod']
Family: binomial ( logit )
Formula: M ~ Group * Age*Condition + (1 | Code) + (1 | Item)
Bilinguals & F condition as Intercept

Fixed effects:
              Estimate Std. Error z value Pr(>|z|)
(Intercept)   -0.2256    0.9807  -0.230  0.818093
GroupMono     -0.2966    1.6679  -0.178  0.858847
Age           -0.2388    0.1494  -1.599  0.109924
ConditionFM   -0.8409    0.8437  -0.997  0.318928
ConditionFN    2.1806    0.9504   2.294  0.021772 *
ConditionM    1.0938    1.0033   1.090  0.275615
ConditionN    1.5251    0.8744   1.744  0.081129 .
GroupMono:Age -0.4477    0.3168  -1.413  0.157596
GroupMono:ConditionFM  1.9497    1.4272   1.366  0.171898
GroupMono:ConditionFN -6.2782    1.6892  -3.717  0.000202 ***
GroupMono:ConditionM  1.8464    1.7296   1.068  0.285731
GroupMono:ConditionN -3.2512    1.4640  -2.221  0.026362 *
Age:ConditionFM  0.3525    0.1221   2.886  0.003897 **
Age:ConditionFN -0.3927    0.1472  -2.667  0.007647 **
Age:ConditionM  0.5725    0.1587   3.607  0.000309 ***
Age:ConditionN -0.1115    0.1258  -0.886  0.375403
GroupMono:Age:ConditionFM -0.2368    0.2793  -0.848  0.396551
GroupMono:Age:ConditionFN  1.2293    0.3269   3.761  0.000169 ***
GroupMono:Age:ConditionM  0.2010    0.3347   0.600  0.548185
GroupMono:Age:ConditionN  0.7276    0.2857   2.546  0.010882 *
---
Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
    
```

Table B4. Nonce words: Post hoc pairwise comparisons of groups within conditions

Condition = F:						
contrast			estimate	SE	df	z.ratio p.value
Biling	5.61532364597094	- Mono 5.61532364597094	2.811	0.529	Inf	5.318 <.0001
Condition = FM:						
contrast			estimate	SE	df	z.ratio p.value
Biling	5.61532364597094	- Mono 5.61532364597094	2.191	0.406	Inf	5.397 <.0001
Condition = FN:						
contrast			estimate	SE	df	z.ratio p.value
Biling	5.61532364597094	- Mono 5.61532364597094	2.186	0.445	Inf	4.915 <.0001
Condition = M:						
contrast			estimate	SE	df	z.ratio p.value
Biling	5.61532364597094	- Mono 5.61532364597094	-0.164	0.428	Inf	-0.383 0.7014
Condition = N:						
contrast			estimate	SE	df	z.ratio p.value
Biling	5.61532364597094	- Mono 5.61532364597094	1.976	0.400	Inf	4.936 <.0001
Results are given on the log odds ratio (not the response) scale.						

Table B5. Proportion (and number) of responses in each of the four categories defined by the signal detection analysis for monolinguals and bilinguals in the masculine, feminine, and neuter **transparent conditions** (Nonce word experiment)

	Hits	Misses	False Alarms	Correct rejections
Bilinguals Masculine	0.85 (282)	0.15 (48)	0.3 (198)	0.7 (465)
Monolinguals Masculine	0.88 (372)	0.12 (52)	0.1 (92)	0.9 (754)
Bilinguals Feminine	0.74 (245)	0.26 (86)	0.1 (69)	0.9 (593)
Monolinguals Feminine	0.93 (399)	0.07 (28)	0.1 (92)	0.9 (751)
Bilinguals Neuter	0.57 (190)	0.43 (142)	0.01 (9)	0.99 (652)
Monolinguals Neuter	0.75 (314)	0.25 (105)	0.002 (1)	0.998 (850)

Table B6. Mean, variance (*SD*), and 95% confidence interval for sensitivity (*d'*) and bias (*c*) with respect to M, F, and N by monolinguals and bilinguals in **transparent conditions** (Nonce word experiment)

	<i>d'</i>			<i>c</i>		
	Mean	var (<i>SD</i>)	95% conf. int.	Mean	<i>SD</i>	95% conf. int.
Bilinguals Masculine	1.58	0.01(0.1)	1.39 to 1.76	-0.26	0.002(0.05)	-0.45 to -0.08
Monolinguals Masculine	2.38	0.01(0.1)	2.2 to 2.58	0.04	0.002(0.05)	-0.16 to 0.23
Bilinguals Feminine	1.89	0.01(0.1)	1.69 to 2.11	0.31	0.003(0.05)	0.1 to 0.52
Monolinguals Feminine	2.73	0.01(0.1)	2.53 to 2.94	-0.13	0.003(0.05)	-0.34 to 0.07
Bilinguals Neuter	2.37	0.03(0.2)	2.01 to 2.73	1.00	0.01(0.1)	0.64 to 1.37
Monolinguals Neuter	3.59	0.13(0.4)	2.87 to 4.31	1.12	0.03(0.2)	0.41 to 1.84

Table B7. Proportion (and number) of masculine responses in the **ambiguous FM condition** as compared to M and F conditions (Nonce word experiment)

	M	FM	F
Bilingual	0.87 (282)	0.48 (160)	0.25 (82)
Monolingual	0.88 (372)	0.22 (94)	0.07 (28)

Table B8. Mean, variance, and 95% confidence interval for sensitivity (d') to the presence of the F cue (comparison between the M and FM conditions = d' -prime_1) and to the presence of the M cue (comparison between the FM and F conditions = d' -prime_2) in the FM condition (Nonce word experiment)

	d-prime_1 (M to FM)			d-prime_2 (FM to F)		
	Mean	var (SD)	95% conf. int.	Mean	var	95% conf. int.
Bilingual	1.17	0.01 (0.1)	0.95 to 1.39	0.62	0.01 (0.1)	0.42 to 0.82
Monolingual	1.94	0.01 (0.1)	1.73 to 2.14	0.74	0.01 (0.1)	0.52 to 0.97

Table B9. Proportion (and number) of neuter responses in the **ambiguous FN condition** as compared to N and F conditions (Nonce word experiment)

	N	FN	F
Bilingual	0.88 (190)	0.21 (53)	0.02 (4)
Monolingual	0.88 (314)	0.11 (44)	0.00 (0)

Table B10. Mean, variance, and 95% confidence interval for sensitivity (d') to the presence of the F cue (comparison between the N and FN conditions = d' -prime_1) and to the presence of the N cue (comparison between the FN and F conditions = d' -prime_2) in the FN condition (Nonce word experiment)

	d-prime_1 (M to FM)			d-prime_2 (FM to F)		
	Mean	var (SD)	95% conf. int.	Mean	var	95% conf. int.
Bilingual	1.97	0.02 (0.1)	1.69 to 2.24	1.30	0.04 (0.2)	0.89 to 1.71
Monolingual	2.41	0.01 (0.1)	2.17 to 2.65	1.80	0.19 (0.4)	0.95 to 2.66

Table C1. Mixed cues Matching Conditions (Probability of following the prompted adjectival agreement pattern as predicted by the interaction of Group, Age, and Condition)

```

Generalized linear mixed model fit by maximum likelihood (Laplace Approximation)
[glmerMod]
Family: binomial ( logit )
Formula: Adj ~ Group * Age* Condition + (1 | Code) + (1 | Item)
Control: glmerControl (optimizer = "bobyqa")
Bilinguals & FF condition as Intercept

Fixed effects:

```

	Estimate	Std. Error	z value	Pr(> z)
(Intercept)	1.37208	1.11015	1.236	0.2165
GroupMonoling	-2.56702	2.51197	-1.022	0.3068
Age	0.11446	0.17065	0.671	0.5024
Cue_totMM	4.81555	2.67803	1.798	0.0721
Cue_totNN	-1.05284	1.21100	-0.869	0.3846
GroupMonoling:Age	1.06732	0.57735	1.849	0.0645
GroupMonoling:Cue_totMM	9.39681	9.19330	1.022	0.3067
GroupMonoling:Cue_totNN	4.96903	2.82390	1.760	0.0785
Age:Cue_totMM	-0.27859	0.37250	-0.748	0.4545
Age:Cue_totNN	0.09738	0.18880	0.516	0.6060
GroupMonoling:Age:Cue_totMM	-2.16389	1.56563	-1.382	0.1669
GroupMonoling:Age:Cue_totNN	-1.01385	0.63465	-1.597	0.1102

```

---
Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1

```

Table C2. Mixed cues Mismatching Conditions (Probability of following the prompted adjectival agreement pattern as predicted by the interaction of Group, Age, and Condition)

```

Generalized linear mixed model fit by maximum likelihood (Laplace Approximation)
['glmerMod']
Family: binomial ( logit )
Formula: Adj ~ Group * Age* Condition + (1 | Code) + (1 | Item)
Control: glmerControl(optimizer = "bobyqa")
Bilinguals & FM condition as Intercept

Fixed effects:

```

	Estimate	Std. Error	z value	Pr(> z)
(Intercept)	-1.024228	0.558062	-1.835	0.0665
GroupMonoling	2.081944	1.189206	1.751	0.0800
Age	0.184387	0.084512	2.182	0.0291 *
Cue_totFN	-0.404793	0.732965	-0.552	0.5808
Cue_totMF	1.661179	0.731757	2.270	0.0232 *
Cue_totMN	1.829526	0.735738	2.487	0.0129 *
Cue_totNF	0.480040	0.721882	0.665	0.5061
Cue_totNM	-0.616830	0.745537	-0.827	0.4080
GroupMonoling:Age	-0.120824	0.221923	-0.544	0.5861
GroupMonoling:Cue_totFN	-1.834815	1.503533	-1.220	0.2223
GroupMonoling:Cue_totMF	-4.057708	1.499157	-2.707	0.0068 **
GroupMonoling:Cue_totMN	-2.611808	1.616980	-1.615	0.1063
GroupMonoling:Cue_totNF	-3.078703	1.612850	-1.909	0.0563
GroupMonoling:Cue_totNM	-2.289059	1.624908	-1.409	0.1589
Age:Cue_totFN	0.065738	0.111828	0.588	0.5566
Age:Cue_totMF	-0.164228	0.110364	-1.488	0.1367
Age:Cue_totMN	-0.199652	0.111230	-1.795	0.0727
Age:Cue_totNF	-0.090593	0.109007	-0.831	0.4059
Age:Cue_totNM	-0.007428	0.110972	-0.067	0.9466
GroupMonoling:Age:Cue_totFN	0.339273	0.290640	1.167	0.2431
GroupMonoling:Age:Cue_totMF	0.519514	0.284300	1.827	0.0676
GroupMonoling:Age:Cue_totMN	0.281131	0.309634	0.908	0.3639
GroupMonoling:Age:Cue_totNF	0.521806	0.313731	1.663	0.0963
GroupMonoling:Age:Cue_totNM	0.228580	0.305858	0.747	0.4549

```

---
Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1

```

Table C3. Mixed cues Mismatching Conditions (Probability of following the prompted **nominal** agreement pattern as predicted by the interaction of Group, Age, and Condition)

```

Generalized linear mixed model fit by maximum likelihood (Laplace Approximation)
['glmerMod']
Family: binomial ( logit )
Formula: Nominal ~ Group * Age* Condition + (1 | Code) + (1 | Item)
Control: glmerControl(optimizer = "bobyqa")
Bilinguals & FM condition as Intercept

Fixed effects:
                Estimate Std. Error z value Pr(>|z|)
(Intercept)      1.19169    0.59356   2.008 0.044676 *
GroupMonoling    -2.36447    1.24052  -1.906 0.056645 .
Age              -0.32480    0.09684  -3.354 0.000796 ***
Cue_totFN       -0.81937    0.76742  -1.068 0.285661
Cue_totMF       -1.95575    0.75517  -2.590 0.009603 **
Cue_totMN       -2.25636    0.77474  -2.912 0.003586 **
Cue_totNF       -1.90533    0.76038  -2.506 0.012218 *
Cue_totNM        0.18389    0.76376   0.241 0.809729
GroupMonoling:Age  0.24617    0.23858   1.032 0.302171
GroupMonoling:Cue_totFN  2.93057    1.57297   1.863 0.062452 .
GroupMonoling:Cue_totMF  4.27190    1.54274   2.769 0.005622 **
GroupMonoling:Cue_totMN  3.17791    1.67442   1.898 0.057706 .
GroupMonoling:Cue_totNF  3.25677    1.70194   1.914 0.055676 .
GroupMonoling:Cue_totNM  2.46772    1.64678   1.499 0.133999
Age:Cue_totFN    0.13861    0.12395   1.118 0.263454
Age:Cue_totMF    0.32001    0.11996   2.668 0.007638 **
Age:Cue_totMN    0.34050    0.12276   2.774 0.005540 **
Age:Cue_totNF    0.33370    0.12032   2.774 0.005545 **
Age:Cue_totNM    0.06874    0.12187   0.564 0.572748
GroupMonoling:Age:Cue_totFN -0.49643    0.30765  -1.614 0.106608
GroupMonoling:Age:Cue_totMF -0.63199    0.29748  -2.124 0.033634 *
GroupMonoling:Age:Cue_totMN -0.40091    0.32272  -1.242 0.214130
GroupMonoling:Age:Cue_totNF -0.54164    0.33299  -1.627 0.103825
GroupMonoling:Age:Cue_totNM -0.23955    0.31474  -0.761 0.446581
---
Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1

```

Cite this article: Mitrofanova, N., Urek, O., Rodina, Y., and Westergaard, M. (2022). Sensitivity to microvariation in bilingual acquisition: morphophonological gender cues in Russian heritage language. *Applied Psycholinguistics* 43, 41–79. <https://doi.org/10.1017/S0142716421000382>