Kupisch, Tanja & Geiss, Miriam & Mitrofanova, Natalia & Westergaard, Marit. 2022. Structural and phonological cues for gender assignment in monolingual and bilingual children acquiring German. Experiments with real and nonce words. *Glossa: a journal of general linguistics* 7(1), pp. 1–37. DOI: https://doi.org/10.16995/glossa.5696

Structural and phonological cues for gender assignment in monolingual and bilingual children acquiring German. Experiments with real and nonce words

Tanja Kupisch, University of Konstanz, DE and UiT Tromsø, NO, Tanja.Kupisch@uni-konstanz.de
Miriam Geiss, University of Konstanz, DE, miriam.geiss@uni-konstanz.de
Natalia Mitrofanova, UiT Tromsø, NO, natalia.mitrofanova@uit.no
Marit Westergaard, UiT Tromsø, NO, marit.westergaard@uit.no

We investigate the acquisition of grammatical gender marking in German by monolingual children as well as German-Russian bilingual children who grow up in Germany as heritage speakers of Russian. We ask to what extent monolingual and bilingual children use phonological and/or structural cues to assign nominal gender, and to what extent they rely on lexical knowledge. To this end, we designed three experiments. Experiment 1 tests gender assignment with real nouns; Experiment 2 tests gender assignment to nonce nouns with the same set of noun-internal phonological cues as in Experiment 1, and in Experiment 3 we compare gender assignment to nonce nouns based on phonological vs. structural (agreement) cues when both types of cues are provided. Results show that children are significantly less successful when assigning gender to nonce nouns as compared to real nouns, which highlights the importance of lexical learning. At the same, we observe sensitivity to noun-internal phonological cues for both mono- and bilingual children. Bilingual children show similar patterns as monolingual children but different default strategies. For the bilingual children, we discuss the possibility of cue transfer from Russian to German. Finally, we observe that the role of structural (agreement) cues increases with age, while the role of noun-internal phonological cues decreases with age for the bilingual children, in line with previous findings from other languages.

1. Introduction

One central debate in the field of early bilingual language acquisition concerns the nature and the extent of cross-linguistic influence (CLI), i.e., whether a bilingual child's two languages develop autonomously or influence each other in a significant way. A recent meta-analysis by van Dijk et al. (2021) measured the effect of CLI in 750 simultaneous and early sequential bilingual children (from 26 different studies and 17 language combinations) and found "a significant small to moderate average effect size of CLI, indicating that CLI is part and parcel of bilingual development" (p. 1). However, further research is needed to establish in what way CLI is modulated by the linguistic distance and structural similarity between the languages involved. Moreover, it remains unclear if the effect of CLI is equally strong across different linguistic domains and features. In the current study we focus on the acquisition of grammatical gender in German by German-Russian bilingual children, considering the global effect of overall transparency of the gender systems in the two languages as well as the role of more local and fine-grained structural and phonological gender cues.

The number of studies on the acquisition of gender in bilingual children has increased massively over the past few years, and the results have often been interpreted as CLI, which may affect either gender discovery, i.e., the moment when children start marking gender on gender targets (e.g., articles or adjectives) or the rate of correct gender marking. The idea behind these studies is typically that a language with a more transparent gender marking system can influence a language with a less transparent system positively (resulting in acceleration) or, conversely, that a language with a less transparent gender marking system can affect a language with a more transparent system negatively (resulting in delays) (see Kupisch et al. 2002; Cornips & Hulk 2008; Hulk & van der Linden 2010; Eichler et al. 2013). For example, a language without a gender marking system, such as English, may have a delaying effect on the acquisition of gender in another language (e.g., Unsworth et al. 2014; Kaltsa et al. 2019). Conversely, a language with a transparent gender-marking system, such as Greek, may have an accelerating effect on the acquisition of gender in another language with a less-transparent gender-system (Egger et al. 2018; Kaltsa et al. 2019).

In most of these studies, it is implicitly or explicitly assumed that languages differ in terms of how transparent their gender marking systems are. For example, Dutch and Norwegian have relatively non-transparent systems, which take long to be acquired (up to age 7, Rodina & Westergaard 2015; Egger et al. 2018), while Spanish, Italian and Russian have transparent systems, which are acquired fast with only a few errors after age 3 – 4 (Perez-Pereira 1991; Ceitlin 2009). The parameters referred to are typically (i) the transparency of morphological, phonological or semantic cues on the noun itself, (ii) the transparency of agreement marking on elements which are in a syntactic relation with the noun, (iii) how often gender is relevant (i.e., how many targets show gender marking), and (iv) the number of genders. As for the latter, it

has been proposed that two-way systems with masculine and feminine can generally be acquired faster than systems with an additional neuter category (Eichler et al. 2013).

While most of these studies make plausible assumptions about transparency based on the grammatical gender systems of the investigated target languages, children's actual knowledge of the rules and regularities underlying these systems is rarely ever tested explicitly. Importantly, we would need to know what bilingual children's knowledge of the gender system consists of in order to make assumptions about the properties that can potentially be transferred from one language system into the other. If, for example, children have merely lexical knowledge, e.g., knowing that the Italian noun *sole* is masculine, then transferring this knowledge would not be facilitative because translation equivalents often have different genders.[1] If, on the other hand, children have internalized knowledge that noun endings can be used to predict nominal gender, they might apply this more abstract knowledge to the other language. Therefore, in addition to investigating whether children assign gender successfully to existing nouns, we need to investigate whether they are sensitive to more general cues (phonological regularities of nouns belonging to the same gender) and exactly which cues they are sensitive to.

If we assume that languages differ gradually with respect to the predictability of noun-internal cues for gender assignment, we may set up a scale of transparency, where languages such as Italian, Spanish and Russian would be at one end of the scale (with a few highly predictive rules that may account for the majority of nouns in the language), while languages like Dutch and Norwegian would be at the other end of the scale (see **Figure 1**).

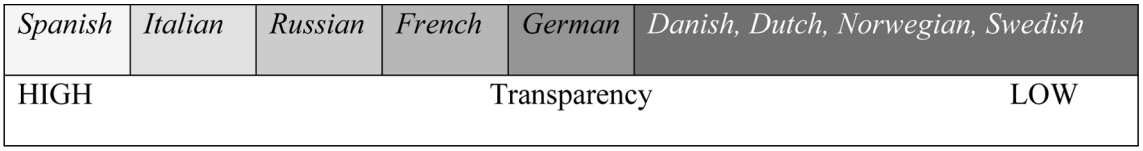

Figure 1: Predictability of gender marking noun-internal cues in selected languages.

Since children ultimately acquire all these gender systems, it should be possible to learn grammatical gender both by paying attention to cues and by rote-learning gender features of individual lexical items, with the former strategy being more relevant for Italian and Russian and the latter strategy for Dutch and Norwegian. Nevertheless, it has been shown in a number of studies that transparent systems are learned early, while less transparent ones lag behind (e.g., Kupisch et al. 2002; Rodina & Westergaard 2015; 2017; Egger at al. 2018). But what is the preferred strategy for children acquiring a language that is somewhere in between, such as German? Will the children be learning item by item for an extended period of time, or will they discover the statistical patterns for gender assignment from early on?

[1] This type of transfer would correspond to Sabourin, Stowe & de Haan's (2006) notion of "surface transfer".

The paper is structured as follows. In section 2, we introduce gender assignment and agreement patterns in German, motivating the assumption that the gender marking system in German is semi-transparent, and we also provide an overview of previous research relevant to the paper. In section 3 we formulate the research questions and predictions. Sections 4 – 6 present one Experiment each. We discuss the results and conclude in sections 7 and 8.

2. Background

2.1 Gender assignment and agreement in German

German has a three-way gender system with masculine, feminine and neuter. Gender is marked on determiners (definite and indefinite articles, demonstratives, possessives), attributive adjectives and relative pronouns. Masculine is the most frequent and neuter the least frequent gender, although the exact numbers vary depending on the type of corpus (e.g., spoken vs. written, child-directed vs. adult-directed, basic vocabulary vs. advanced vocabulary). Schiller & Caramazza (2003) report 43% masculine, 38% feminine and 19% neuter gender, based on lemmas in the CELEX database (4164 monomorphemic nouns). For child speech, Binanzer (2017) reports 42% masculine, 31% feminine and 27% neuter; for basic vocabulary Wegener (1995) finds 44% masculine, 33% feminine and 23% neuter nouns. In many accounts on German, masculine is considered to be the default for gender assignment in German. The reason is not only its comparatively higher frequency – masculine is only slightly more frequent than feminine – but rather the assigned gender when assignment rules are in competition or when no specific gender assignment rules apply (e.g., Steinmetz 2006). On the other hand, Opitz & Pechmann (2016) have provided evidence that masculine gender requires increased processing load, which would speak against its default status. According to Köpcke & Zubin (2009, pp. 148 – 150), it is in fact possible to construct an argument for all three genders as the default in German, depending on theoretical position. Based on examples of headless NPs such as ***das*** *Schöne* 'the beautiful' and reference to unknown objects (*Was ist denn* ***das?*** 'What's that?'), they claim that Neuter gender could be considered the 'system default', though admitting that such examples could also be genderfree. Corbett & Fraser (2000: 69f) argue that such examples could be construed as showing the *syntactic* gender default in German. The syntactic default is generally assumed to be Neuter, based on presentative or expletive constructions (***das*** *ist meine Mutter* 'that is my mother', *es scheint* ... 'it seems …'). Thus, there is no general consensus about the notion of default gender in German (see also Enger (2009) for a critical discussion of defaults, and Corteen (2019) for a discussion of defaults within Optimality Theory).

The most frequent elements on which gender marking is visible are articles, which are obligatory with singular count nouns. Gender interacts with number, definiteness and case, and there is considerable syncretism, which makes structural cues to nominal gender somewhat opaque. For example, gender is unambiguously marked on definite singular DPs in the nominative

case, although the adjective is underspecified with respect to gender. In indefinite DPs in the nominative case, there is syncretism between neuter and masculine articles, but the adjective is unambiguously marked for gender (see **Table 1**). In the plural, there is no gender marking.

	M	**F**	**N**
Singular definite	**der** schwarze Hund DEF.M.SG black dog	**die** schwarze Katze DEF.F.SG black cat	**das** schwarze Auto DEF.N.SG black car
Singular indefinite	**ein** schwarz**er** Hund INDEF.M.SG black-M.SG dog	**eine** schwarz**e** Katze INDEF.F.SG black- F.SG cat	**ein** schwarz**es** Auto INDEF.N.SG black-N.SG car
Plural definite	**die** schwarzen Hunde DEF.PL black-PL dog-PL	**die** schwarzen Katzen DEF.PL black-PL cat-PL	**die** schwarzen Autos DEF.PL black- PL car-PL

Table 1: Examples of gender marking in German (nominative case).

Besides structural cues to nominal gender (i.e., elements in syntactic agreement with the noun), there are morphological, phonological and semantic cues on the noun itself (Köpcke 1982; Köpcke & Zubin 1983). German generally follows the natural gender rule, in that nouns referring to male beings are masculine, and nouns referring to female beings are feminine.[2] Suffixes can be associated with particular genders, e.g., *–chen* and *–lein* are associated with neuter gender, even if there is a clash with natural gender, as illustrated by the examples *Mädchen*$_{N}$ 'girl' and *Männlein*$_{N}$ 'little man'. There are numerous phonological cues to gender assignment, but many of them reflect probabilistic tendencies rather than rules. Amongst the most reliable ones for each gender are the following: Disyllabic nouns ending in [ə] are associated with feminine gender 90% of the time (e.g., *Katze*$_{F}$ 'cat'), monosyllabic nouns starting with the onset cluster /ʃ/ + C_ are associated with masculine gender 81% of the time (e.g., *Strand*$_{M}$ 'beach'), and nouns ending in [ɛt] are associated with neuter 95% of the time (e.g., *Bett*$_{N}$ 'bed') (see, e.g., Köpcke 1982).

2.2 The acquisition of gender by German monolingual children

Gender is most frequently marked on articles, which is why first gender marking is closely related to first article use. Mills (1986: 63) reports first uses of definite articles in spontaneous child data around age 1;10 – 2;2, and Stern & Stern (1928: 44, 86) around 1;2 – 1;10. Articles are typically not gender marked at this point but reduced or underspecified forms, e.g., *de.* Such forms are often considered to be proto-articles whose function is to mark a syntactic position, while the acquisition of morphology lags behind (e.g., Bottari et al. 1993/1994). Sporadic use of gender-marked articles is reported from age 2;0 – 2;4 (Mills 1986), 1;9 (Bittner 2006), and 1;5

[2] There is more to be said about semantic regularities, but since we did not test them in this paper, we refrain from discussing them in more detail.

(Szagun 2007), though with substantial individual variation. ACQUISITION tends to be defined as the point when children cease to make gender-marking mistakes (90% accuracy). This happens around age 3;0 (Mills 1986; Szagun et al. 2007).

Based on a single case study, Bittner (2006) argued that children's early articles do not have a gender-marking function at all, but are used to mark syntactic function instead, with *der* referring to subjects and *den* to objects. *Der* in nominative position constitutes more than 90% of the forms used before age 2;4 regardless of the noun's gender; *die* is overused between 2;5 and 2;11, and *das* is hardly ever overused. According to Bittner the gender system only emerges at 2;11, but the dissociation between masculine and neuter only by the age of 4. However, since her data are primarily presented in terms of case and only non-target uses are quantified, the study leaves open what the rate of correct gender marking was. Further, the early acquisition of case in German is not generally confirmed (e.g., Tracy 1986).

Based on longitudinal data, Mills (1986: 64) observed that the feminine article *eine* was the most common early form, also with nouns of masculine or neuter gender. When indefinite articles became frequent, *ein* and *eine* fluctuated, sometimes in combination with the same noun. Mills speculated that the incorrect use of the indefinite article may be due to "the high frequency of the ending *–e* in the article and adjective paradigms […] leading to an overgeneralization of the *–e* ending on all prenominal units." Definites appeared later than indefinites. While Mills observed an association of monosyllabic nouns with *der* (p. 83) and of the schwa-ending with *die* (p. 85), errors suggested that not all assignment was cue based. Errors involved overuse of *die,* even with nouns containing masculine cues, e.g., *Bagger* (*-er* is a cue for masculine) and *Hund* (monosyllabic nouns tend to be masculine). Thus, overuse of *die* does not result from early acquisition of the feminine cue alone. Article omission was still frequent between the ages of two and three years. Since the children produced far more definite articles with feminine nouns than with nouns of the other two genders, Mills concluded that children are more confident with the feminine gender. Her account is largely in line with Stern & Stern (1928), who reported a distinction between masculine and feminine forms by 2;6 as well as gender errors resulting from overuse of *die*. Mills (1986) also investigated older children (ages 3;2 – 6;3) in an experimental situation, and results showed few incorrect productions of articles (5%), with none of the 6-year-olds producing any errors. There were comparatively fewer errors with feminine, and when errors were made with masculine and feminine, children rarely resorted to neuter. Feminine and neuter were rarely confused with one another.

Overall, the literature review indicates considerable individual variation, but some overall tendencies: (i) despite the challenges posed by the system, children's development of gender in German appears to be comparatively fast (mostly before age 3), (ii) feminine articles emerge early and neuter articles late, (iii) there is some sensitivity to phonological noun-internal cues,

and (iv) article omission or production of underspecified articles may be considered children's strategies to avoid morphological markings of gender and case.

2.3 The acquisition of gender by German bilingual children

Studies on the acquisition of German by early bilinguals can be distinguished in terms of whether German is acquired simultaneously as an L1 or sequentially as an early L2. Another relevant distinction is whether German is acquired along with a language that has gender and, if so, how transparent gender marking in that language is.

In Müller's (1999) study of three German-French simultaneous bilingual children, gender was marked from around 1;10 – 2;4. Initially, the children mostly omitted determiners, except for French *un* and German *ein*, which seemed to be used as numerals and irrespectively of the gender of the noun they preceded. Subsequently, they started using determiners productively and assigned gender based on phonological and semantic criteria. Gender assignment with definite articles was unproblematic, except for neuter, which occurred only from around age 2;4 – 2;8. Assignment and overgeneralization errors were determined by phonological and semantic properties. The children used feminine definite articles with nouns ending in schwa, even if these nouns were masculine, i.e., overusing the schwa-regularity. For instance, the children produced *die*$_{F}$ *Hase*$_{M}$ 'rabbit' or *die*$_{F}$ *Affe*$_{M}$ 'monkey'; alternatively they used the correct gender but adjusted the noun ending to make assignment more rule-conform, as in *der *Aff* 'the monkey' (monosyllabic nouns being associated with masculine). One of the children also overused *der* with monosyllabic nouns (*der*$_{M}$ *Haus*$_{N}$ 'house', *der*$_{M}$ *Geld*$_{N}$ 'money', *der*$_{M}$ *Bett*$_{N}$ 'bed'), indicating sensitivity to the tendency for monosyllabic nouns to be masculine. Neuter remained a problem until the age of 5, and neuter nouns were initially classified as masculine (Müller 1999: 393).

Eichler, Jansen & Müller (2013) investigated gender marking in naturalistic data from 17 simultaneous bilingual children acquiring a Romance language (French, Italian or Spanish) along with German or another Romance language. They compared gender errors across all languages. The children were followed longitudinally between 2 and 3;6 years and gender marking was investigated based on article production. The bilingual children were more successful in their Romance languages than in German (and children acquiring Italian and Spanish had advantages over children acquiring French), although most of the children grew up in Germany. Knowledge of cues was not explicitly discussed, but the children were aware of the schwa-rule, as evident from overgeneralizations such as *die*$_{F}$ *Junge*$_{M}$ 'the boy' (p. 565). Neuter was acquired late.

Also based on naturalistic data, Wegener (1995) studied the acquisition of gender by L1-Turkish, Polish, and Russian children acquiring German as an early L2. The children started out using bare nominals; their first articles had the function of marking indefinite (nonspecific)

as opposed to definite reference. *Ein* was the most frequent indefinite article and *die* the most frequent definite one, arguably due to the high token frequency of these forms in the input. Subsequently, the children used articles for case-marking or argument-marking functions: *r*-forms for subjects, *s*-forms for direct objects and *e*-forms for plurals, somewhat similar to what is claimed by Bittner (2006) for monolingual German. The children eventually used articles successfully, according to Wegener primarily due to lexical learning (p. 15), but they also showed sensitivity to formal rules. For example, after only 13 – 15 months of exposure the L1-Polish and L1-Russian children assigned masculine gender to monosyllabic nouns and nouns in *–er* and feminine to nouns in schwa (pp. 16 – 17). The L1-Turkish children in grade 4 performed at chance, possibly due to the absence of gender in Turkish.

Dieser (2009) presented longitudinal naturalistic data (ages 1 – 6;6) and cross-sectional semi-naturalistic and experimental data (ages 3 – 10) from German-Russian children. She investigated both languages and compared simultaneous and sequential bilinguals. All learners, independently of type of bilingualism, passed through similar acquisition sequences in both languages, starting with a two-way gender system, feminine and masculine for Russian, feminine and non-feminine for German (see also Hopp 2011). Gender differentiation started with indefinite articles, before being extended to definite articles. Dieser concluded that monolinguals and bilinguals make use of lexical and rule-based learning in both languages, with a higher amount of rule-based learning in Russian. The German rules included the schwa-rule.

Ruberg (2013) compared German monolingual children (ages 3 – 5) and sequential bilingual children with different heritage languages (Polish, Turkish, Russian; ages 4;0 – 6;3). Based on real and nonce words, his results show that by the age of 4;0 monolingual children use correct definite and indefinite articles. The sequential bilingual children did not show target-like gender marking after 30 months of contact with German. When making assignment errors, both the monolingual and the sequential bilingual children defaulted to masculine, in line with Steinmetz's (2006) claim that masculine is the default gender in German. Ruberg's results further showed that both monolingual and bilingual children made use of morphophonological cues, assigning gender more often correctly if the nouns ended in schwa (associated with feminine) or were monosyllabic (associated with masculine). When the noun did not exhibit a prominent gender cue, children mainly used the default strategy. As in Wegener's study, the L1 Turkish children did not attend to cues to the same extent as the two other bilingual groups.

Finally, Lemmerth & Hopp (2018) investigated lexical and syntactic CLI in simultaneous and sequential bilingual German-Russian children (mean age 8), focusing on German. Elicited production data showed that the children assigned gender correctly independently of whether the nouns have congruent or different genders in the children's two languages. By means of eye-tracking, the authors further investigated whether the children used articles and adjectives to anticipate an upcoming noun. The simultaneous bilingual children behaved like monolingual

controls, while the sequential bilinguals made use of gender marking only when the gender of the noun was equivalent in the two languages. Lemmerth & Hopp (2018) speculated that sequential bilinguals assign gender in the L2 by first accessing the lexicon of their L1, while there is no such influence in simultaneous bilinguals. This latter claim raises the interesting question on what basis the children drew a parallel between feminine, masculine and neuter gender in German and Russian. Unless the children had been taught German and Russian grammar explicitly, we believe that there is no reason for them to match specific genders one-to-one in the two languages.

In summary, although there is substantial variation in the acquisition of German gender by bilinguals, there are many indications that bilingual children show sensitivity to the same cues as monolingual children. The variation seems to be determined by the language combination, specifically the presence/absence and transparency of gender marking-rules in the other language, but it remains unclear what exactly can be transferred across languages. **Table 2** summarizes some relevant studies on German.

2.4 Grammatical gender in Russian and its acquisition

As mentioned in the Introduction, gender in Russian is mostly transparent (Corbett 1982; 1991), with three rules that predict gender assignment to the majority of nouns in the language: In the nominative singular, masculine nouns typically end in a consonant (e.g., *dom*$_M$ 'house'), most feminine nouns end in *-a* (e.g., *lisá*$_F$ 'fox'), and most neuter nouns end in *-o* (e.g., *vedró*$_N$ 'bucket'). There are also some phonological gender cues that are ambiguous, most notably affecting nouns that end in a palatalized consonant, which may be either masculine or feminine (e.g., *gus'*$_M$ 'goose' or *kost´*$_F$ 'bone'), as well as nouns ending in schwa, which may be either feminine or neuter (e.g., *parta*$_F$ [partə] 'desk' or *sito*$_N$ [sitə] 'sieve'). The majority of nouns ending in a schwa are feminine, leading to early overgeneralizations of schwa-final neuters to feminine (Ceitlin 2009; Mitrofanova et al. 2018). Russian has no articles, but gender is marked as agreement on adjectives, demonstratives, quantifiers, some numerals and verbs in the past tense. According to Corbett (1991), masculine and feminine genders are almost equally frequent (46% and 41% respectively), with masculine being considered the default for gender assignment, while neuter is by far less frequent (13%).

Table 3 summarizes (morpho-)phonological gender cues in Russian, comparing them to those we introduced above for German. We remind the reader that there are more possible noun endings in both German and Russian, which are often predictive of gender, but we restrict ourselves to those relevant for the present study.

The current study is a continuation of similar studies on gender assignment in bilingual populations acquiring two gender languages that differ with respect to transparency. Rodina & Westergaard (2017) elicited real words to investigate the gender system of Norwegian-Russian

	Age range	Focus	Results
Mills (1986)	1;8 – 2;6	L1	– naturalistic data: gender errors below 90% by age 3;0; early overuse of *eine* and *die*; some sensitivity to phonological rules
	5 – 9;0		– experimental data: (frequent) phonological rules are applied
Szagun (2007)	1;4 – 3;8	L1	– naturalistic data – gender errors below 90% by age 3;0; some sensitivity to phonological rules
Müller (1999)	1;10 – 5;0	2L1	– naturalistic data – sensitivity to semantic and phonological rules; delay with neuter
Eichler, Jansen & Müller (2013)	2 – 3;6	2L1	– naturalistic data – higher error rates in German than in the Romance languages; sensitivity to schwa-rule; delay with neuter
Wegener (1995)	unclear	eL2	– naturalistic data – evidence for lexical learning along with sensitivity to formal rules
Dieser (2009)	1;6 – 10;0	2L1, eL2	– naturalistic and (semi-)experimental data – evidence for lexical and rule-based learning, including the schwa-rule – late split of masculine and neuter categories
Ruberg (2013)	4;0 – 6;3	eL2	– elicited production data – use of phonological cues and defaulting to masculine
Lemmerth & Hopp (2018)	7;2 – 9;6	2L1, eL2	– elicited production and – use of articles and adjectives to anticipate gender; congruency effects in eL2 bilinguals

Table 2: Some studies on the acquisition of German (2L1 = simultaneous bilinguals; eL2 = early sequential bilinguals).

bilinguals and found that transparency plays a major role, as gender was acquired faster in Russian than in Norwegian, even though Russian was the children's heritage language. There was no difference between the bilinguals and monolingual controls in the majority language (Norwegian), but the children's performance in Russian was dependent on the amount of input.

	M	F	N
German	monosyllabic /ʃ/ + C_CC Strand ‘beach’	Disyllabic ending in [ə] Katze ‘cat’	-[ɛt] Bett ‘bed’
Russian	C *dom* ‘house’	-[a]/-[ə] *lisá*$_F$ ‘fox’ *parta*$_F$ [partə] ‘desk’	-o *vedró*$_N$ ‘bucket’

Table 3: Comparison of (prominent) phonological rules in German and Russian.

Unlike Polinsky (2008), who found that some (adult) heritage speakers of Russian in the US were developing a reduced gender system with only masculine and feminine, Rodina & Westergaard found that the children generally had a three-gender system and that only the children with the least amount of input defaulted to masculine, thus seemingly developing a Russian variety without gender. Mitrofanova, Rodina, Urek & Westergaard’s (2018) follow-up study used nonce words and showed that the heritage children were generally sensitive to the transparent phonological cues in Russian, and defaulting to masculine decraesed with a higher Cumulative Length of Exposure to Russian.

The German-Russian children in the current study have also been investigated in their heritage language Russian and compared with monolingual controls (Mitrofanova et al. 2021), with the same three experiments as in the present study: real words, nonce words, and nonce words with mismatched cues (i.e., noun endings and adjectival morphology cueing for different genders). The findings show sensitivity to both transparent and ambiguous cues in Russian for both monolingual and bilingual children, but bilingual children used the masculine default to a considerably larger extent. For the ambiguous cues, all children displayed a preference for feminine with nouns ending in a schwa and sensitivity to more fine-grained cues for the nouns ending in palatals, such as the type of palatal consonant. In the mismatch experiment, the importance of the syntactic (agreement) cue increased significantly with age, while that of the noun-internal phonological cue decreased, echoing findings from French by Karmiloff-Smith (1979).

3 Research questions and predictions

Our literature review shows early sensitivity to gender based on naturalistic corpora, but there are several hints that not all regularities have been acquired by age 3, and that the acquisition of assignment regularities is still ongoing during the primary school years. This is plausible given the magnitude of regularities in German, their relatively low extension and the assumption that children postulate rules and regularities only when they have collected sufficient evidence in the input. Since bilingual children tend to develop their vocabularies more slowly than monolinguals

(see Montanari et al., 2018 and Klassert et al., 2014, for German-Russian children), they may need longer to generalize over a class of nouns and postulate a rule. Alternatively, given their smaller vocabularies, they may be more inclined to rely on formal cues from earlier on, especially if they have another language in which formal cues are more prominent. This motivates our first research question and corresponding prediction.

> **RQ1:** In a language with only semi-transparent gender cues/statistical tendencies, such as German, are children sensitive to phonological gender cues? When exposed to new nouns, do they default to one gender, and if so, which one?
> **Prediction:** Since monolingual German children have been found to produce few non-target-consistent gender forms in spontaneous speech after age 3, we predict that both mono- and bilingual children will show some sensitivity to certain phonological gender cues, despite the relatively low reliability of such cues in German. We also expect this sensitivity to be dependent on the relative reliability of the cue in the language and thus that the cue *-e* for F nouns would be the one where the children show the highest cue sensitivity. At the same time, given the generally low reliability of phonological cues in German, we also predict that the children to a large extent will use a default gender with nonce nouns. According to previous research both within gender typology and acquisition, M gender is the default in Russian and at least some children default to masculine also in German. We thus expect both mono- and bilingual children to some extent to default to M gender (cf. section 2.1 and 2.2).

As outlined in the introduction, many studies have argued for positive or negative CLI from one language to another leading to acceleration or delay. Here, we ask whether such cross-linguistic effects, if we find them, can be linked to transfer of experience with phonological cues:

> **RQ2:** Are there quantitative and/or qualitative differences between monolinguals and bilinguals, and, if so, can they be related to CLI from Russian?
> **Prediction:** Given the high reliability of (morpho-)phonological cues in Russian, we expect there to be a facilitative effect of Russian in the results of the bilinguals. That is, we predict that the bilinguals show a generally higher sensitivity to phonological cues in German compared to monolingual children. Thus, while the monolinguals might be more accurate than the bilinguals with existing words, we expect the bilinguals to score relatively higher with nonce words.

Finally, since we know from previous research that older children are more sensitive to structural cues, we speculate that the same might hold for German-learning children. However, they may show an even stronger tendency to rely on structural cues, since the phonological cues in German are weak. This motivates our third question.

RQ3: Which strategy do German-speaking mono- and bilinguals use for gender assignment when exposed to conflicting cues – do they pay more attention to the phonological shape of the noun or structural cues on other targets, and does this change with age?
Prediction: Given findings from other languages (L1 French, 2L1 Russian, see above), we predict that young children are more sensitive to the phonological cue on the noun early on, even when it is less reliable than structural cues. In accordance with previous research, we further predict that children will pay more attention to structural cues with increasing age.

4. Experiment 1 (real words)

The purpose of Experiment 1 was to establish whether bilingual children assign the target gender to nouns that reflect gender assignment regularities and whether they differ from monolinguals. Since we used specific cues for gender assignment in nonce words (Experiment 2), we wanted to establish whether the children knew the gender of real nouns that contained precisely these cues.

4.1 Participants

53 bilingual children (ages 3 – 11, mean 6,2) were recruited from a Russian language school in Singen, Hohentwiel (South Germany), a Russian daycare centre in Stuttgart, and four in Berlin. All children were heritage speakers of Russian growing up in Germany. These children spoke no other languages than German and Russian. 18 children came from German-Russian families (ages 3 – 11; mean 6.22; sd = 2.42), i.e., families in which one parent spoke Russian and the other German (henceforth 'mixed' families) and 35 from families in which both parents spoke Russian (ages 4 – 10, mean 6,22; sd = 1.92). Family type coincides with type of bilingualism, i.e., children from mixed families were simultaneous bilinguals and children from only Russian-speaking families were sequential bilinguals. Age of onset to German varied (range 0 – 6 years, mean 2.1). For most children first exposure to German was at the age of 3 years when entering kindergarten.[3] Additionally, we tested 24 monolingual German children (ages 3 – 9; mean 4.92; sd = 1.43).

4.2 Method

Words in Experiment 1 (see Appendix 2) were based on selected gender assignment tendencies, as specified in Köpcke (1982), Köpcke & Zubin (1983), and Wegener (1995), cf. section 2.1 Disyllabic nouns ending in schwa were used for feminine, mono- and disyllabic nouns in /ɛt/

[3] The relatively large age range and varying AoO also means that total length of exposure varied. However, *amount of exposure* to German also varied greatly so that in some cases children with shorter exposure were more fluent than children with longer exposure. Thus, neither AoO nor length of exposure were necessarily indicative of the children's proficiencies, especially amongst the younger children. We therefore included all children, but those whose proficiency was too low were filtered out if they could not respond to the task (see description of experimental procedures below).

were used for neuter, and monosyllabic nouns starting with /ʃ/ + C__ and ending in __Nasal + C targeted masculine gender. The selected nouns were inanimate and concrete. There were 18 nouns in total, six per gender category. Since we targeted specific phonological cues and since we needed inanimate nouns that could be depicted, certain drawbacks could not be avoided. First, it was impossible to balance congruent and incongruent nouns within the three gender categories. All masculine nouns that fulfilled our selection criteria happened to have congruent translation equivalents in Russian, and most neuter nouns happened to be incongruent. For the feminine category in schwa, which is generally more frequent, congruent and incongruent nouns were balanced. Overall, 50% of the nouns were congruent. Second, the frequency of the noun in each gender category differed. Nouns in the feminine category tend to be the most frequent in child corpora, while most nouns in the neuter and masculine categories were of low frequency (see **Table 1** in Appendix 2).

The experiment was framed in a story of Martians coming for their first visit to Earth. In Experiment 1, the Martians took pictures of objects they saw during their visit, but since they had no idea what these objects were called, the children had to tell the Martians the names of the objects. The children were seated in front of a PowerPoint presentation on a screen and their responses were audiotaped with an Olympus LS-11. The children saw two objects of different sizes and colours, and they were asked to name them. Then one object disappeared and the children had to guess which one, again using an article, an adjective and a noun. An example of the procedure is given in (1). Each trial targeted the elicitation of **two indefinite DPs** and one **definite DP** (relevant DPs in bold). The experimenters were native speakers of German and were carefully instructed on how to carry out the experiment. Preceding the experiment, there was a training session with three items, one of each gender.

(1) **Example procedure for Experiment 1**

Pictures of a blue and red candle appear simultaneously on the screen

Experimenter: Was siehst du?
(What do you see?)

Expected answer: **Eine$_F$ blaue$_F$ (Kerze)** und **eine$_F$ rote$_F$ Kerze.**
(a blue (candle) and a red candle)

The red candle disappears

Experimenter: Was ist verschwunden?
(what disappeared?)

Expected answer: **Die$_F$ rote Kerze.**
(the red candle)

4.3 Analysis

For a first overview, we classified all utterances in terms of whether the **elicited** gender target (definite or indefinite article or adjective) was realized or not. The gender targets were **indefinite articles** and **adjectives** (indefinite condition) or a **definite article** (definite condition). If the elicited target was provided by the child, the realization was coded as M, F or N. If the child did not provide the elicited gender target, we distinguished **replacements** of indefinite by definite articles and vice versa, **article omissions** and **non-classifiable responses**. The responses were not classifiable if the noun was changed or if no gender target was provided. An overview of the responses for the monolingual and bilingual groups is provided in **Table 4**. Note that masculine and neuter responses cannot be distinguished in the indefinite condition due to syncretism.

		Elicited gender target provided			**Elicited gender target not provided**		
		F	**M**	**N**	**replacement def./indef.**	**omission**	**non-classifiable**
INDEF	MON	122 (28.3%)	**219 (50.8%)**		8 (1.9%)	52 (12.1%)	30 (7%)
	BIL	305 (26.5%)	**396 (34.4%)**		34 (3%)	227 (19.7%)	190 (16.5%)
ADJ	MON	**139 (32.3%)**	106 (24.6%)	116 (26.9%)	0	25 (5.8%)	45 (8.8%)
	BIL	**397 (34.5%)**	215 (18.7%)	283 (24.6%)	0	4 (0.3%)	253 (22%)
DEF	MON	**139 (32.3%)**	109 (25.3%)	127 (29.5%)	19 (4.4%)	14 (3.2%)	23 (5.3%)
	BIL	163 (14.1%)	122 (10.6%)	162 (14.1%)	296 (25.7%)	**374 (32.5%)**	35 (3%)

Table 4: Real word experiment, overview of responses (predominant answer pattern in bold).

For further analysis, gender assignment was coded based on the definite article, but some adjustments were necessary because the children did not always produce the elicited target.

- If no article was produced, the adjective in the indefinite condition was used as an indicator of assigned gender, but only if it was clearly marked as masculine (e.g., *roter*) or as neuter (e.g., *rotes*). Potentially feminine marked adjectives (e.g., *rote*) were not considered, because the form can represent all genders in definite DPs (cf. **Table 1**). Further, **Table 2**

suggests that the bilingual children still have problems with definiteness markings.[4] Note that this procedure potentially biased the analysis *against* feminine marked DPs.

- If neither an article (or other determiner) nor an adjective marked as neuter or masculine was produced, we refrained from coding any assigned gender.
- Responses in which the children used a different noun than the intended one were excluded, unless the noun contained the same gender assignment cue as the target word. For instance, *Flagge*$_{F}$ instead of *Fahne*$_{F}$ 'flag' was included because both contain the schwa-cue but *Bodensee*$_{M}$ 'Lake of Constance' instead of *Strand*$_{M}$ 'beach' was excluded.
- If children could not name the depicted item themselves, the answer was not included. The rationale for this is that an unknown word is like a nonce word for the child, and we wanted to restrict the results of experiment 1 to words that the children knew.
- We included non-target-like productions in which the children added an 's' or 'r' to the adjective to make the adjective rhyme with the definite determiner in the neuter and masculine conditions, e.g., *das*$_{N}$ **rote-s Bett*$_{N}$ ('the red bed'), *der*$_{M}$ **grosse-r Schrank*$_{M}$ ('the big closet'). Although the adjective form is not target-like, it can be taken to indicate the assigned gender.
- We coded inconsistently marked gender across the three elicited responses and agreement mismatches within the DP, e.g., *eine*$_{F}$ *grosses*$_{N}$ *Bett*$_{N}$ ('a big bed'). In both cases, it is not clear which gender the children have assigned, and these responses were not included when determining which gender the children had assigned.

Finally, we excluded data from nine children who completed less than 50% of the task or provided responses without determiners or any gender-marked adjective more than 50% of the time, two from mixed families and several from Russian families (see **Table 3**).

4.4 Results

The remaining dataset consisted of 44 bilingual German-Russian children, 16 from German-Russian families (age 3 – 11; mean 6,62; sd = 2.40) and 28 from Russian families (ages 4 – 10, mean 6.51; sd = 1,78). **Table 5** shows that the children from Russian families omitted determiners most often. Gender switches across DPs were infrequent in all groups.

As a next step, we excluded children who defaulted to one particular gender. Defaulting was defined as using one gender 75% to 100% of the time. There were no monolingual defaulters, but among the bilinguals five defaulted to feminine and one to neuter. Discarding defaulters left us with 23 monolinguals and 38 bilinguals (23 from RR families and 15 from mixed families).

[4] In other words, we cannot be sure whether a bare noun produced in the indefinite condition was intended to be definite, in which case the feminine-looking adjective form could potentially represent all genders.

	Monolinguals (GG)	German-Russian family (GR)	Russian family (RR)
Responses without a determiner	2%	17%	27%
Gender switch	5%	2%	3%

Table 5: Children producing no gender marked forms and inconsistent gender marking.

The results of the Real word task are illustrated in **Figure 2** for the feminine condition (left), the masculine condition (middle) and the neuter condition (right) for each of the three groups of children. Arrows indicate where we find significant differences.

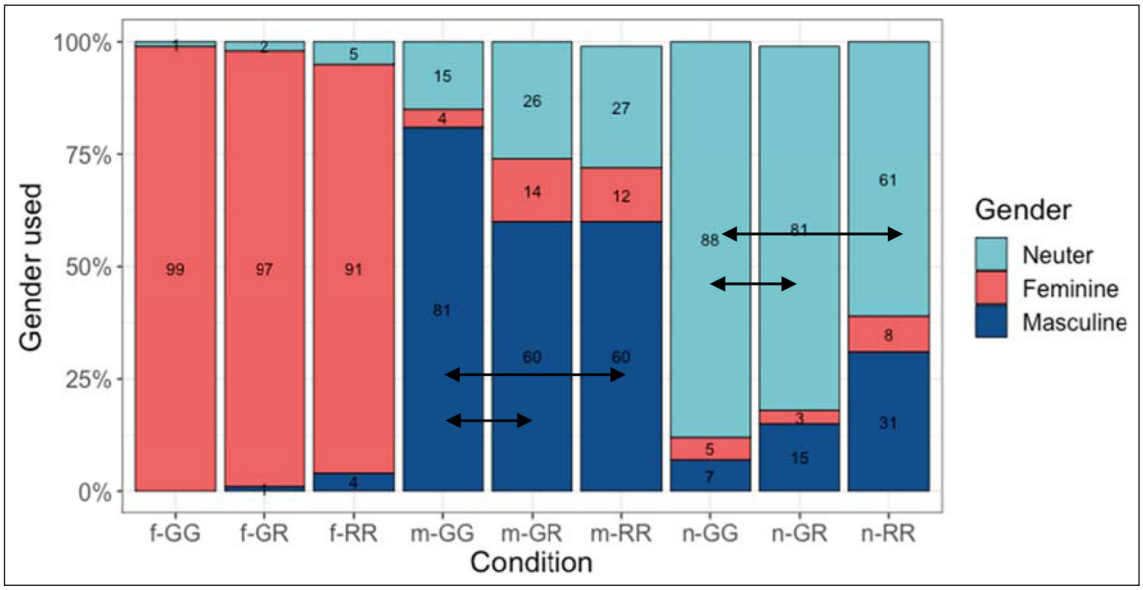

Figure 2: Gender assignment with real nouns in German: monolingual children (GG), children from German-Russian (GR) and children from Russian-Russian (RR) families. Significant contrasts between the groups are marked by arrows.

We applied generalized linear mixed effects logistic regression analysis to statistically model the probability of an accurate response on the real word task depending on the three main predictors (Family type, Condition and Age) and their interactions. Model comparison (Anova) indicated that Age (continuous) as a main effect and the interaction of Family and Age could be dropped. The final model included the main effects of Family type and Condition, their interaction, as well as an interaction of Age and Condition, and a three-way interaction of Condition, Family and Age. Participants and Items were included as random intercepts. All models were fit using R version 4.0.3 (release 2020-10-10). The analysis revealed significant interactions of Age and Condition (masculine $p = .03$, and neuter $p = .02$), indicating that all

groups of participants become significantly more accurate with age with respect to masculine and neuter gender as compared to the feminine. The output of the final model is summarized in Table A1 in the Appendix. For a more detailed comparison of the groups within conditions, we conducted post-hoc pairwise comparisons using the R package *emmeans* (Lenth et al. 2019). These revealed that the monolingual children performed significantly better than the children from Russian (RR) and mixed families (GR) on the masculine ($p < .001$ in both cases) and neuter conditions ($p = .003$ and $p < .001$, respectively). No other contrasts were significant (see Tables A2 in Appendix 1).

Overall, the results indicate that children in all groups use all three genders productively (above 60% accuracy with all genders), although nouns in the feminine condition seem to be the least difficult.

5. Experiment 2 (Nonce words)

The purpose of Experiment 2 was to see how children performed with nonce nouns that contained the same gender cues as the real nouns in Experiment 1. The Nonce word task preceded the Real word task and was carried out on a different day.

5.1 Participants

The participants were the same as in Experiment 1, but more children took this task. The initial dataset consisted of 60 bilingual German-Russian children, 20 from German-Russian families (age 3 − 11; mean 6.67; sd = 2.47) and 40 from Russian families (age 4 − 10 years, mean 6.50; sd = 1.89). We tested the same 23 monolingual German children as in Experiment 1.

5.2 Method

Inspired by Karmiloff-Smith's (1979) Experiment 8, we created a task with 24 nonce words (8 per gender category), e.g., *Schromp*$_{(M)}$, *Fruxe*$_{(F),}$ *Knett*$_{(N)}$, which followed the same assignment rules as those in Experiment 1 (see Appendix 2). We ensured that these nonce nouns formed no minimal pairs with existing words in German and piloted them with monolingual German adults. The adults assigned the predicted gender between 70% and 100% of the time (71% M; 80% N; 88% F).

As in Experiment 1, the children were told a story about Martians, but this time they had brought several objects that were non-existent on earth. When leaving their spacecraft to explore the area, a thief broke into their spacecraft, and the children were asked to help him carry the bag with stolen items. The children were given a small bag into which they could place the items (pictured on small cards). The items were introduced to the children by means of a PowerPoint presentation (with drawings of imaginary objects taken from the Horst & Hout 2016 noun database). They were non-animate objects of different shapes. When the items were

introduced, special care was taken not to reveal the targeted gender. Again, two indefinite DPs and one definite DP were elicited (see (2), relevant DPs in bold).

(2) **Example procedure for Experiment 2**

Situation: Children are shown a yellow Knett and a blue Knett.

Experimenter:	Was du hier siehst heißt Knett. Was ist das?
	(What you see here is called $Knett_{(N)}$. What do you see?)
Expected answer:	**Ein gelbes (Knett)** und **ein blaues Knett.**
	(a_N $yellow_N$ (Knett) and a_N $blue_N$ Knett)

Situation: Animated appearance of the thief. The blue Knett disappears, while the yellow Knett remains visible.

Experimenter:	Was klaut der Dieb?
	(What's the thief stealing?)
Expected answer:	**Das blaue Knett.**
	(the_N blue Knett)

There were three practice items whose gender was predictable based on suffixes (one associated with each gender) that were not used in the experiment. If necessary, the practice session was repeated until the child understood the task.

5.3 Analysis

Tables 6 provides an overview of the responses for the monolingual and bilingual groups. Again, masculine and neuter responses cannot be distinguished in the indefinite condition due to syncretism.

The coding of responses followed the same algorithm as in Experiment 1 (see above) with some adjustments for nonce nouns. We discarded the following types of items:

- Cases in which the child did not (re)produce the target word, unless the cue was kept constant (e.g., $Klapett_{(N)}$ for $Lapett_{(N)}$ was included, but $Fretts_{(?)}$ for $Fruxe_{(F)}$ was excluded);
- Instances in which it was impossible to determine the assigned gender, (i) because the gender was incongruent between the determiner and the noun, e.g., *$eine_F$ $blaues_N$ *Plett*, (ii) because the gender was underspecified, e.g., *blau Plett* (no article, no inflection on the adjective), (iii) because the children switched genders across the elicited DPs.

As before, we included only children who completed more than 50% of the task. The resulting dataset consisted of 23 monolingual German-speaking children, 44 bilingual German-Russian children: 18 from German-Russian families and 26 from Russian families.

		Elicited gender target provided			Elicited gender target not provided		
		F	**M**	**N**	**replacement def./indef.**	**omission**	**non-classifiable**
INDEF	MON	77 (13.4%)	**376 (65.3%)**		2 (0.3%)	33 (5.7%)	88 (15.3%)
	BIL	319 (18.2%)	545 (31.1%)		78 (4.5%)	192 (11%)	**616 (35.3%)**
ADJ	MON	110 (19.1%)	34 (5.9%)	**382 (66.3%)**	0	6 (1%)	44 (7.6%)
	BIL	**531 (30.3%)**	187 (10.7%)	525 (30%)	0	0	509 (19.1%)
DEF	MON	92 (16%)	13 (2.3%)	**345 (59.9%)**	35 (6.1%)	39 (6.8%)	52 (9%)
	BIL	252 (14.4%)	117 (6.7%)	326 (18.6%)	275 (15.7%)	**578 (33%)**	204 (11.6%)

Table 6: Nonce word experiment, overview of responses (predominant answer pattern in bold).

5.4 Results

Similar to the Real word task, the groups differed significantly in the probability of providing an article-less response: children from Russian families did so 32% of the time, while children from German-Russian families did so only 15%. Again, the children rarely switched (only 4% of the time with no differences between children from the two family types). The monolinguals failed to provide an article or another gender-marked element only 6% of the time, while gender switches constituted 1% of the data. When coding the data, we proceeded as follows, in analogy to Experiment 1.

- The assigned gender was coded based on the definite article.
- If no article was produced in the definite condition, we used the indefinite condition as an indicator of the assigned gender: adjectives marked as masculine (e.g., *roter*) or neuter (e.g., *rotes*) were taken to be indicators of the assigned gender. Recall that in the absence of an article, feminine marked adjectives (e.g., *rote*) were not considered, because the form can represent all genders in definite DPs (cf. **Table 1** and footnote 4).
- We included non-target-like productions in which the children added an 's' or 'r' to the adjective to make the adjective rhyme with the definite determiner in the neuter and masculine conditions, e.g., *das*$_{N}$ **rote-s Plett*$_{N}$ ('the red bed').

- If neither an article (or other determiner) nor an adjective marked as neuter or masculine was produced, we refrained from coding any assigned gender.
- We excluded cases in which gender was marked inconsistently across the three elicited responses and cases of agreement mismatches within the DP, e.g., *eine*$_F$ *grosses*$_N$ *Plett*$_N$ ('a big Plett').

As a result of the first cleaning, we excluded cases where the child did not respond, produced a noun that was different from the prompted noun, a bare noun and a bare adjective, an incomprehensible response or one that couldn't be classified as either M, F or N. The remaining data were classified into M, F and N responses following the procedure described in 5.3 above. **Figure 3** illustrates the distribution of responses in the Nonce word experiment.

To statistically model the probability of following the phonological gender cue in the Nonce word task, we fit a series of generalized linear mixed effects logistic regressions with different sets of predictors. The best model included three main effects: Condition, Family type, and Age (continuous), two two-way interactions (Condition and Family, and Condition and Age), and a three-way interaction of Condition, Family and Age. An Anova model comparison suggested that the interaction of Family and Age could be dropped. Participants and Items were included as random intercepts. The model revealed a significant effect of Family (RR: $p = .007$), Condition (Masculine: $p = .008$, Neuter: $p < .001$) and Age ($p = .002$), as well as a significant interaction

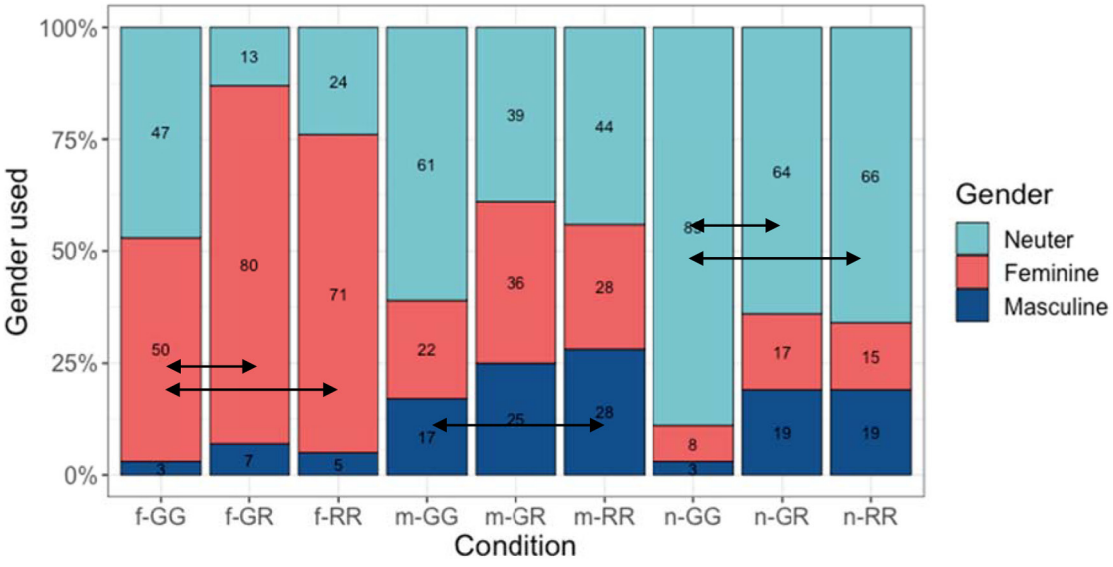

Figure 3: Gender assignment with nonce nouns in German: monolingual children (GG), children from German-Russian (GR) and children from Russian-Russian (RR) families. Significant contrasts between groups are marked by arrows.

of Condition (N) and Family (DR: $p = .03$; RR: $p < .001$), a significant interaction of Condition (N) and Age ($p < .001$) and four significant three-way interactions of Condition (F), Age and Family (DR: $p = .004$; RR: $p < .001$), and Condition (N), Age and Family (DR: $p = .003$; RR: $p = .001$). The output of the model is summarized in Table B1 in the Appendix.

Post-hoc pairwise comparisons of groups within conditions revealed three significant differences. First, the GR and RR children were significantly more likely than the GG children to use Feminine gender in the Feminine condition ($p < .001$). GR children were marginally but not significantly different from the RR children ($p = .07$). Second, RR children were significantly more likely to use Masculine in the Masculine condition than GG children ($p = .01$). GR children were marginally different from the GG children ($p = .07$). Third, GG children were significantly more likely to use Neuter in the Neuter condition than both the GR and the RR children ($p < .001$ in both cases). No other differences were significant. The results of the post-hoc comparisons are summarized in Table B2 in the Appendix.

It should be noted that many children chose one particular gender (almost) throughout the entire task. We had a closer look at these defaulters (classified as children who used one particular gender 75% of the time or more) and non-defaulters separately. Among the 23 monolinguals, 11 children used neuter 75% or more of the time, and one child (3 years old) used feminine 100% of the time. The bilingual data varies more in terms of default strategies. Among the 44 bilinguals, 13 defaulted to neuter and 8 to feminine. The latter are distributed equally over German-Russian and Russian families. Interestingly, defaulting correlates with age, as younger children default less to neuter (all 3-year-olds are non-neuter-defaulters, while most 6-year-olds and all 7-year-olds are neuter-defaulters). **Table 7** provides an overview of defaulting strategies comparing mono- and bilinguals. The distribution of (non-)defaulters and their demographics is presented in **Table 8**, indicating that as many as half of the children in each group chose a default strategy in their responses. Interestingly, the vast majority of monolingual defaulters used Neuter gender

	F-defaulters (>75% F)	**N-defaulters (>75% N)**	**Non-defaulters**
Monolinguals	1 (4%)	11 (48%)	11 (48%)
Bilinguals	8 (18%)	13 (30%)	23 (52%)

Table 7: Default strategies in mono- and bilinguals.

GG (monolinguals)	**GR (mixed families)**	**RR (minority families)**
N = 11 (age 3−7, mean: **4.55**; sd = 1.15)	N = 9 (age 4−10, mean: **6.85**, sd = 2.10)	N = 14 (age 4−10, mean: **7.3**, sd = 1.63)

Table 8: Distribution of non-defaulters in Experiment 2.

across the board, while bilingual defaulters fall into two groups: children that used Feminine across the board and participants that defaulted to Neuter. There were no Masculine defaulters in the dataset.

Figure 4 presents a comparison of responses by condition in the monolinguals and the two groups of bilinguals (excluding defaulters).

We used a similar analysis as described above to assess the effect of the main predictors (Family type, Condition and Age) on the probability of following the gender cue provided in the Nonce word for a subset of children who did not default to one gender throughout the experiment. The winning model included the same predictors as the winning model for the whole Nonce word dataset, i.e., the three main effects: Condition, Family type, and Age, two two-way interactions (Condition and Family, and Condition and Age), and a three-way interaction of Condition, Family and Age. Participants and Items were included as random intercepts. The model revealed a significant effect of Condition (Masculine: $p = .001$ and Neuter: $p = .004$), as well as a significant interaction of Condition (Neuter) and Age ($p = .002$) and two significant three-way interaction of Condition (Neuter), Age and Family (GR: $p = .04$ and RR: $p = .02$). The output of the model is summarized in Table B3 in the Appendix.

Post-hoc pairwise comparisons of groups within conditions revealed the following significant differences. First, the GR and RR children were significantly more likely than the GG children to use Feminine gender in the Feminine condition ($p < .001$ in both cases). Second, RR children

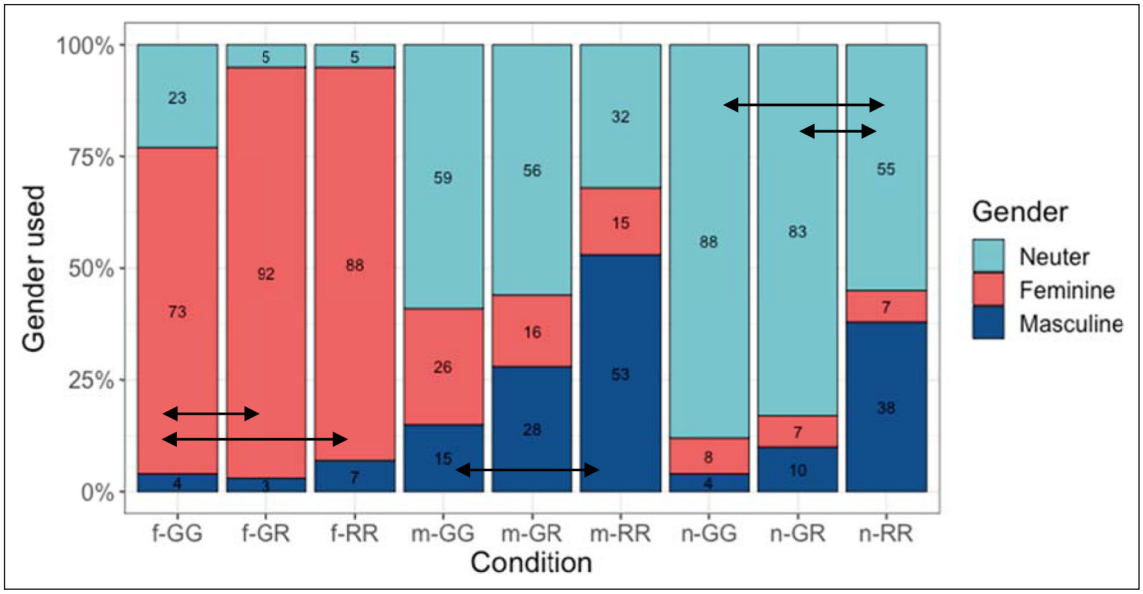

Figure 4: Gender response per condition: Monolinguals (GR) vs. Bilinguals from German-Russian (GR) and Russian (RR) families (excluding defaulters). Significant contrasts between the groups are marked by arrows.

were significantly more likely to use Masculine in the Masculine condition than the GG children ($p = .03$). Third, the GG and GR children were significantly more likely to use Neuter in the Neuter condition than the RR children ($p = .005$ and $p = .03$, respectively). No other differences were significant. The results of the post-hoc comparisons are summarized in Table B4 in the Appendix.

The results indicate that after the exclusion of the defaulters both bilingual groups exhibit a significantly higher sensitivity to the feminine gender cue than the monolinguals, and the RR children exhibit a higher sensitivity to the masculine cue. In the neuter condition, the monolinguals and children from mixed families produce more neuter responses than children from Russian families. The children from Russian families use more masculine in the neuter condition, and it is possible that they are delayed in separating the masculine vs. neuter categories, as has been reported for monolinguals in the literature (cf. section 2.2).

6. Experiment 3 (mixed cues)

The purpose of Experiment 3 was to see whether children pay more attention to phonological or to structural (agreement) cues when assigning gender to nonce nouns. Since sensitivity to structural cues as opposed to phonological cues is known to increase with age in languages with relatively transparent assignment systems, such as French (e.g., Karmiloff-Smith 1979) and Russian (Rodina 2008; Rodina & Westergaard, 2012), the same might hold for German. Unlike in Experiment 2, we provided a gender cue when introducing the nonce noun. Experiment 3 followed Experiment 2 but was carried out on a different day.

5.1 Participants

We tested 20 monolinguals (ages 3 – 9, mean 5) and 38 bilinguals (age 4 – 10, mean 7), 11 from GR families and 27 from RR families. The bilingual children were a subset of the children who participated in Experiments 1 and 2.

6.2 Method

The experiment was inspired by Karmiloff-Smith's (1979) Experiment 9 and involved 18 stimuli. The nouns were taken from a larger pool of items than in Experiment 2. A third of the stimuli (N = 6) contained DPs in which the gender cue on the noun was consistent (matched) with that on the article and adjective (e.g., *eine*$_{\text{F}}$ *blaue*$_{\text{F}}$ *Fruxe*$_{\text{(F)}}$); the remainder (N = 12) contained DPs with conflicting (mismatching) gender cues on article + adjective vs. noun ending (e.g., *eine*$_{\text{F}}$ *blaue*$_{\text{F}}$ *Glett*$_{\text{(N)}}$). The stimuli were balanced for gender (see Appendix 2 for an overview of the stimuli).

The children were told that the Martians were on another trip exploring Earth, and while they were gone, the thief broke into their spacecraft again. The experimenter invited the children

to look at what he stole this time. This time, one DP with cues was provided and only two additional DPs were elicited. Note that in the mixed cue condition, the child may repeat the gender markings present in the adult utterance, although they contain a mismatch between the structural and the phonological cue, or they may change the article and adjective (i.e., the structural cues) to match them with the noun ending (i.e., the phonological cue).

(3) **Example procedure in Experiment 3, children see a blue Glett**

Experimenter: Das auf meiner Seite ist **eine**$_{F}$ **blaue**$_{F}$ **Glett**$_{(N)}$. Und auf deiner Seite?
(This here on my side is a blue Glett. And on your side?)

Exp. Answer: **Eine**$_{F}$ **rote**$_{F}$ **Glett**$_{(N)}$**/ein**$_{M/N}$ **rotes**$_{N}$ **Glett**$_{(N)}$.
(a red Glett/a red Glett)

Experimenter: Und was ist weg?
(And what is gone?)

Exp. Answer: **Die**$_{F}$ **rote Glett**$_{(N)}$**/das**$_{N}$ **rote Glett**$_{N}$.
(the red Glett/the red Glett)

6.3 Analysis

We followed a similar cleaning procedure as in Experiments 1 and 2 and discarded responses where the child did not produce the target word or where it was impossible to determine the assigned gender due to incongruent gender markings or underspecified gender. Unlike in Experiment 2, where no structural gender cues were provided, we did not observe any strong defaulting patterns. Only one bilingual child used feminine 100% of the time, and there were no neuter or masculine defaulters. Five further children (all from Russian families) were excluded from the analysis, because they produced less than 50% codable responses. The resulting dataset included 20 German monolingual children, 11 children from German-Russian families and 21 children from Russian families. No significant differences were observed between the bilingual children from German-Russian and Russian families, which motivated analyzing their results together. Unlike in the other experiments, children switched genders more often between their two responses. Monolinguals switched genders 37% of the time, children from German-Russian families 25% of the time and children from Russian families 28% of the time. As much as 80% (72/89) of all switches in the monolingual group were from M or F to N (consistent with the defaulting strategy). Bilinguals switched to N less than 50% of the time. The higher number of switches in this experiment was expected because the children were provided with conflicting cues.

6.4 Results

When the cues on the two targets were consistent with those on the noun (match condition), both mono- and bilingual children mostly produced a definite DP with the same gender as the one provided in the prompt, in fact as much as 85% of the time. **Figure 5** summarizes the proportions of responses where children used the same gender as the one provided by the experimenter in the Match condition.[5]

As evident from **Figure 5**, both groups of children exhibited high sensitivity to the prompted gender cues. At the same time, younger children (especially bilinguals) produced a higher proportion of responses that did not follow the gender cue provided by the experimenter.

The main research question of Experiment 3 was whether we would observe any effect of group or age in the probability of following either the structural cue (on the determiner and adjective) or the phonological gender cue (on the noun ending) provided in the Mismatch condition. **Figures 6** and **7** illustrate the proportions of responses that followed the adjectival or the phonological cue by individual participants from the two groups.

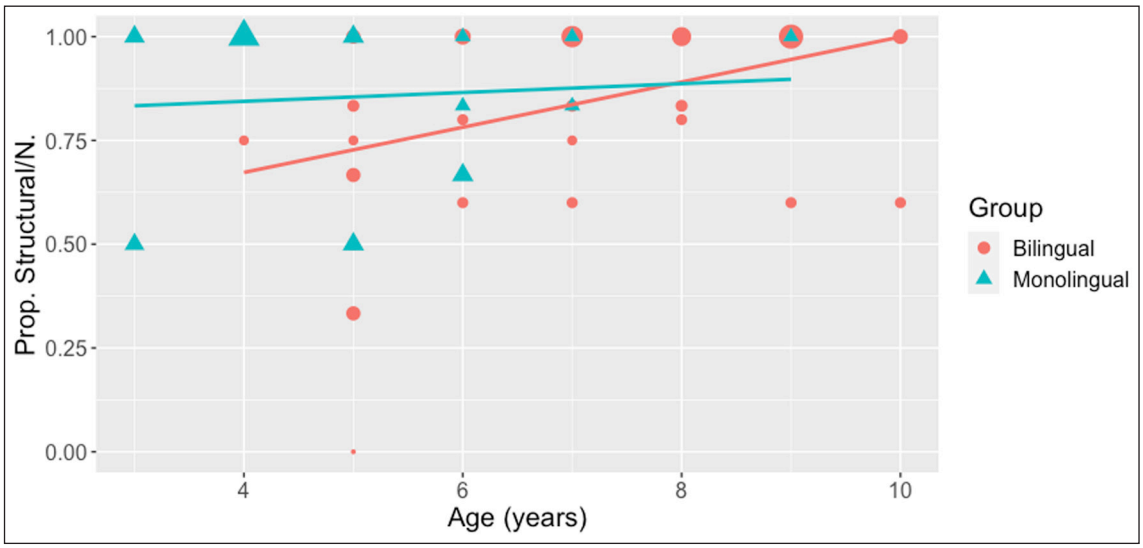

Figure 5: Proportion of responses following the prompted (structural) gender cues in the Match condition by age.

Figures 6 and **7** indicate that monolinguals have a higher probability of following the structural gender cue in their responses than the bilinguals. In both groups, the probability of

[5] The highest number of non-cued gender responses occurred in the masculine condition (37N, 75F = 28%) and the lowest in the feminine condition (4M, 19N = 6%) (neuter condition: 22M, 35F = 15%). When the children assigned a gender that was not cued at all, they used all three genders, but bilingual children tended to use more feminine and monolingual children more neuter, mirroring their defaulting-preferences in Experiment 2.

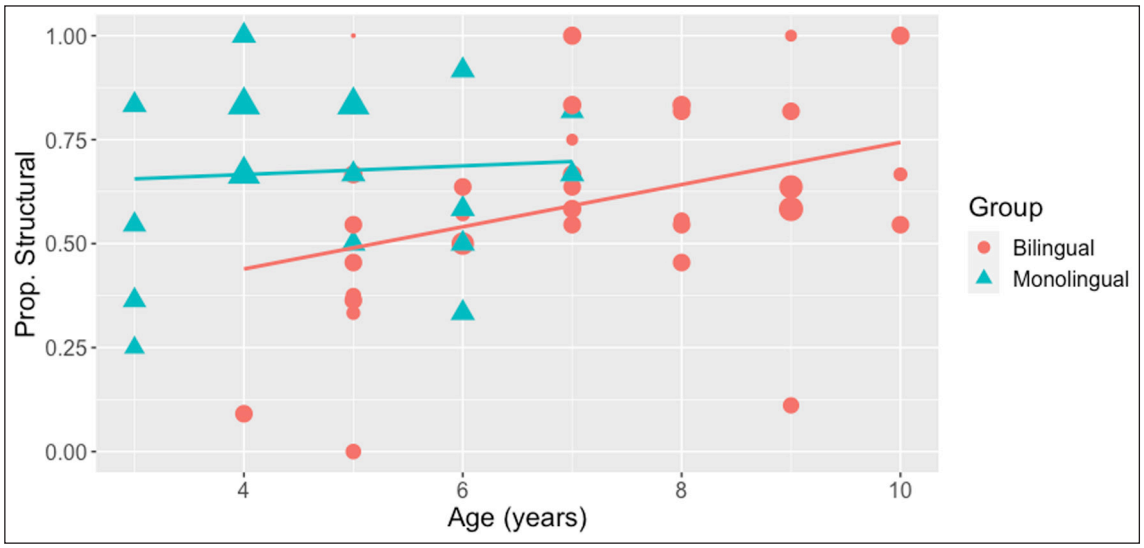

Figure 6: Proportion of responses following the structural cue in the Mismatch condition by age.

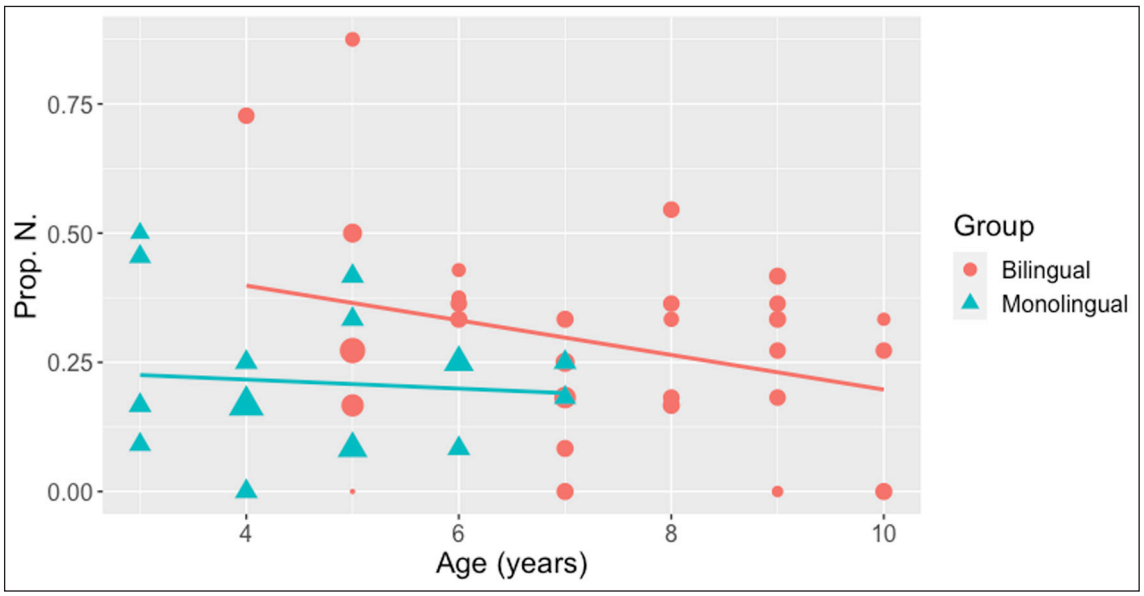

Figure 7: Proportion of responses following the phonological cue in the Mismatch condition by age.

following the structural (and not the phonological cue) increases with age, and this increase is more visible in the bilingual group.

To analyze the results statistically, we applied generalized linear mixed effects logistic regression analysis where the probability of following the structural cue or the phonological

cue was modelled based on the effects of Group (Bilingual vs Monolingual), Condition (Match vs Mismatch) and Age (continuous), as well as their interactions. An Anova model comparison was applied to choose the best model for each dependent variable. The final model predicting the probability of following the **structural cue** included three main fixed effects, an interaction of Group and Age, and a three-way interaction of all predictors. Participants and Items were included as random intercepts. The model revealed a marginally significant effect of group ($p = .05$), and a significant effect of age ($p < .001$), indicating that both groups become more likely to follow the structural gender cue with age, and that the monolinguals are marginally more likely to follow this cue than the bilinguals. No other effects were significant. The output of the model is summarized in Table C1 in the Appendix. Post-hoc pairwise comparisons of groups within conditions indicated that the monolinguals were significantly different from the bilinguals in the mismatch condition ($p = .02$), while there were no differences in the match condition (see Table C2 in the Appendix).

The model that predicted the probability of following the **phonological cue** included three main fixed effects, two interactions: Group and Condition, and Group and Age, and a three-way interaction of all predictors. Participants and Items were included as random intercepts. The model revealed a significant effect of Group ($p = .03$), and a significant effect of Age ($p < .001$), indicating that both groups become less likely to follow the phonological gender cue with age, and that the bilinguals were generally more likely to follow the phonological gender cue than the monolinguals. Furthermore, the model revealed a significant interaction of Group and Condition ($p = .007$), and a significant three-way-interaction of Group, Condition and Age ($p < .001$). No other effects were significant. The output of the model is summarized in Table C3 in the Appendix. Post-hoc pairwise comparisons of groups within conditions indicated that the monolinguals were again significantly different from the bilinguals in the mismatch condition ($p = .02$), while there were no differences in the match condition.

7. Discussion

We have investigated whether German-learning children are sensitive to phonological gender cues and which default strategies they use (**RQ1**). We expected both mono- and bilingual children to be sensitive to phonological gender cues and we expected this sensitivity to be dependent on cue reliability; we further expected defaulting to masculine. We explored potential quantitative and/or qualitative differences between monolinguals and bilinguals and asked whether those could be related to CLI from Russian (**RQ2**). We expected a facilitative effect of Russian and higher sensitivity to phonological cues in bilinguals compared to monolinguals, especially with nonce words. We finally investigated the assignment strategies German-speaking mono- and bilinguals prefer by exposing them to conflicting cues, i.e., structural vs. phonological (**RQ3**).

We expected a higher sensitivity to phonological cues in younger children and to structural cues in older children.

7.1 Sensitivity to formal cues in German vs. lexical learning

Experiment 1 showed that with familiar nouns, mono- and bilingual children acquiring German assign gender successfully in all three gender categories. Both groups of children were most accurate with the feminine nouns and least accurate with the masculine nouns. Given that most neuter nouns in our experiment had a very low frequency too, this suggests that factors beyond the frequency of noun items are at play. We hypothesize that form-based regularities connected to the phonological endings of the nouns are implicated: children showed at-ceiling accuracy with the feminine schwa-final nouns, where cue reliability is the highest, followed by the neuter nouns ending in *-ett*. The reliability of the masculine cue was very low, and accordingly, the children's accuracy on this condition was the lowest. There were quantitative differences between the monolingual and the bilingual children, which can be expected given that the bilinguals, especially those from Russian families, had had less exposure to German.

Despite a general sensitivity to cues, the comparison between Experiments 1 and 2 shows that both groups are considerably less successful when assigning gender to novel words where only a phonological cue on the noun is provided. This finding is in line with the proposal that there exist two routes of grammatical gender processing: the lexical route, which retrieves grammatical gender features of the noun directly from the lexicon, and a form-based route that relies on form-based properties of the grammatical gender classes (see Caffarra et al. 2014). While the gender features of real words are stored in the lexicon and are deterministic, the form-based cues only reflect probabilistic tendencies and may differ in their reliability. The results of our Nonce word task may further reflect the underlying differences in reliability between different types of form-based gender cues. The feminine gender cue (final *-e*) was by far the most reliable cue for the participants, followed by the neuter cue, and finally the masculine cue. This mirrors the results for the adults when we piloted the nonce words.

Overall, phonological gender cues seem to have a relatively low reliability in German. In both groups, many children resort to a default gender in Experiment 2, choosing one particular gender 75% of the time or more. At the same time, children who do not default show a certain sensitivity to the respective gender cues. Those monolingual children who do not default to neuter assign neuter to nouns with the neuter cue, use feminine gender most often when the phonological cue is feminine, and masculine gender when the phonological cue is masculine. This means that defaulting and sensitivity to cues and regularities do not exclude one another, as has also been found by Ruberg (2013; see Enger 2009 for similar observations in adult systems). The picture for the bilingual children is less clear because some children default to neuter and

others to feminine. More specifically, children from Russian families with less German input are more likely to default to feminine than children from mixed families with more exposure to German (the latter, by defaulting to neuter, behave more akin to German monolinguals). The non-defaulters use all three genders, with feminine being most frequent in the feminine condition, masculine in the masculine condition, and neuter in the neuter condition; the division between masculine vs. neuter is less clear, as shown in many previous studies on German (e.g., Stern & Stern 1928; Mills 1986; Müller 1999; Dieser 2009; Hopp 2011).

Finally, a comparison of Experiment 2 and the matching cue condition in Experiment 3 shows that mono- and bilingual children are more successful at assigning gender if there is a structural cue on a gender target (article or adjective) in addition to a phonological cue on the noun. Indeed, the defaulting patterns observed in Experiment 2 were not replicated in Experiment 3, where the children received structural cues in addition to phonological cues. This may be taken to indicate that children are sensitive to the relevant cues, but that phonological cues in German are weaker than in other languages.

7.2 Differences between monolinguals and bilinguals

As outlined above, both mono- and bilingual children have been found to be sensitive to phonological and structural gender cues in German. Based on previous research (Kaltsa et al. 2019; Egger et al. 2018; van Dijk et al. 2021) we hypothesized that bilingual German-Russian children may experience 'acceleration' (or facilitation) in the acquisition of gender in German due CLI from their heritage language Russian. More specifically, bilinguals were expected to be more sensitive to gender regularities in German, due to the high transparency of gender assignment in Russian. This was indeed borne out, as the bilinguals performed better than the monolinguals on both the feminine and the masculine cue in the nonce word task, i.e., nouns ending in schwa and nouns ending in a consonant (Experiment 2, see **Figure 4**). However, they did not perform better than the monolinguals on the neuter cue (which also ends in a consonant). This could be because they treat all consonant-final nouns as belonging to the same noun class (as in Russian) and thus use more masculine gender in this context. Alternatively (or additionally), the high performance of the monolinguals on the neuter cue is not due to sensitivity to the cue, but rather the tendency of the monolinguals to default to neuter.

One major difference between monolinguals and bilinguals concerns their defaulting strategies (N in monolinguals and F or N in bilinguals). First of all, the defaulting patterns observed in our study are somewhat unexpected given previous claims about masculine being the default for gender assignment in German (e.g., Steinmetz 2006; Ruberg 2013) and neuter being late acquired (Müller 1999; Eichler et al. 2013). Interestingly, among the monolingual children, the neuter defaulters represent the *older* children. It thus seems as if the F and M cues on the nouns were to some degree available to younger children, but as they grow older, the phonological

cues get overridden by a more powerful neuter default. There are three possible explanations for this: (i) children identify the nonce objects with things, and the word for thing (*Ding*) is neuter in German (furthermore, according to Zubin & Köpcke 1986, nouns with a 'vague' reference, e.g., *Objekt* 'object', *Zeug* 'stuff', and nouns denoting superordinates, e.g., *Instrument* 'instrument', *Möbel* 'furniture', are generally Neuter in German, although some of these nouns may not be frequent in colloquial and child speech), (ii) the children were primed by the relative pronoun and question word *was,* because it ends in *–s,* which is a typical ending for neuter gender targets,[6] viz. the definite article *das* and inflected adjectives, such as *rotes* 'red', or (iii) children default to neuter gender because it is in fact the default gender in German (cf. Köpcke & Zubin 2009, see section 2.1). Alternatively, the children are affected by the syntactic default, a possibility which gains plausibility given the challenges pertaining to German gender: if children postulate phonological rules, they must discover that there are many exceptions to them, while there are no exceptions to syntactic rules of gender agreement. On the other hand, there is no evidence, so far, that syntactic defaults are transferred to function as morphological default. Clearly, more research is needed to investigate the issue of grammatical gender default in German. We leave this question for future research.

We further suspect that the bilinguals' defaulting to feminine has different reasons than the monolinguals' defaulting to neuter, and that it could be motivated phonologically. The feminine cue is the strongest for the bilinguals, since it is the category with the highest percentage of target-like responses. This makes sense because the phonological cue for feminine is not only present in noun endings, but also on articles and adjectives, e.g., *ein**e** klein**e** Kerz*e 'a small candle', and the adjective form reappears with all genders in definite DPs (*der*$_{M}$ *gross**e** X, die*$_{F}$ *gross**e** X, das*$_{N}$ *gross**e** X* 'the big X'). Thus, children hear determiners, adjectives and nouns ending in schwa very frequently (see also Mills 1986), and by using more feminine overall, children might prime themselves. That is, the more they hear these forms, the more they produce them. An alternative (or additional) explanation is the existence of a similar cue in Russian, viz. nouns ending in schwa, which is mainly feminine (cf. section 2.3). While we think it is unlikely that the bilinguals link feminine gender in Russian to feminine gender in German, they may simply be more sensitive to this phonological cue in German, knowing that a similar cue is important (and highly predictive) in their heritage language.

Another difference between monolinguals and bilinguals is that the bilingual children seem to be more sensitive to the masculine cue. This is particularly evident in the group of children from Russian families. Given that consonantal endings in Russian are associated with masculine gender, it is not unlikely that children have used their knowledge of Russian cues and applied it to German, where our masculine nouns ended in a consonant. In fact, some children do not

[6] The relative pronoun *was* occurs in the absence of an antecedent with specified gender and number features.

seem to distinguish clearly between masculine and neuter, and this can be expected because the neuter nouns in Experiment 2 ended in a consonant too. Note that transfer of cues, as we have speculated here, would not imply that the bilingual children make a connection between masculine, feminine, and neuter gender in German and the corresponding genders in Russian. Rather, since we know that they are highly sensitive to the consonantal noun-ending cue in Russian, they may simply assume that ending in a consonant is a transparent gender cue also in German, and thus expect that **all nouns in -C will belong to the same gender**. This means that it is not important **which gender** this is. The bilingual children may select the M gender forms simply because they are more frequent than N in German. Under this scenario, the 'transfer' of a cue would be based on the phonological feature and unrelated to a lexical feature.

In conclusion, mono- and bilingual children are sensitive to cues, but monolingual children default to neuter, while bilingual children default to neuter or feminine. It is possible that this is merely an effect of exposure and age, since the bilingual children have had less cumulative exposure to German, and it might be that the feminine defaulters would have picked a different strategy, had they been older. Admittedly, our study covers a large age range and the bilingual group is heterogeneous in terms of the learning profiles in German, specifically age of onset and exposure, which makes further research with larger participant groups desirable. At the same time, we were interested in the developmental patterns implicated in the acquisition of gender in German, thus looking at a wider age range was necessary in our study.

7.3 Phonological vs. structural cues

The results of Experiment 3 show that all groups of children rely more on structural cues than on phonological cues in a situation of cue mismatch. For the bilingual children we observed a significant developmental trend, with their reliance on phonological cues significantly decreasing and reliance on syntactic cues significantly increasing with age, in line with previous findings from languages with more transparent phonological gender rules (Karmiloff-Smith 1979, Lidz & Gagliardi 2015). The reason why we see relatively little sensitivity to phonological cues (unlike what Karmiloff-Smith (1979) has shown for French) might indeed be that these are less reliable in German, and that children start paying attention to structural cues from an earlier age, i.e., once they discover that there are numerous exceptions to phonological regularities.

8. Conclusions

We have investigated sensitivity to phonological and structural cues to gender marking in monolingual German and Russian-German bilingual children. The results show that both groups of children use phonological cues on the nouns to assign gender, but they are significantly better at assigning gender to existing and familiar nouns exhibiting the same cues. This suggests that

cue sensitivity takes some time to develop, but it may also be the case that retrieval (of a familiar noun and its gender) is easier and faster than computation (assigning gender to a noun that is not in the mental lexicon). In addition, many children default to one specific gender, with mono- and bilinguals displaying different defaulting strategies. Children are more successful at assigning gender and refrain from defaulting when structural cues are provided, suggesting that using phonological and structural cues is an available and important strategy even in the acquisition of German, where phonological cues are less reliable than in many other languages. There is a trend for the bilingual children, in particular those with considerable exposure to Russian, to rely more on phonological cues than the monolinguals. This was especially visible in the masculine condition of Experiment 2, which tested gender assignment with nonce nouns, and also to some extent in the feminine. This finding suggests that bilingual children may transfer their knowledge of cues for grammatical gender from one language to another, at least if the other language has more reliable phonological cues, which was the case in our study. This indicates that the bilingual children in our study did not transfer the lexical gender of individual nouns but show general sensitivity to phonological cues on a more abstract level. Thus, our study contributes to the literature on CLI in bilingualism. Another interesting finding is that neither monolingual nor bilingual children defaulted to masculine gender in German, although masculine is the suggested default in the theoretical literature. This raises new questions about the definition of defaults, which should be investigated in future research.

Additional Files

The additional files for this article can be found as follows:

- **Appendix 1.** Details of the Statistical Analysis. DOI: https://doi.org/10.16995/glossa.5696.s1
- **Appendix 2.** Experimental Stimuli. DOI: https://doi.org/10.16995/glossa.5696.s2

Acknowledgements

We wish to thanks Simone Waitz and Sarah Zander for their invaluable help during data collection and coding. This research was supported by a grant from the Research Council of Norway for the project Microvariation in Multilingual Acquisition & Attrition Situations (MiMS), project code 250857. The paper was completed while most authors were fellows at the Centre for Advanced Study at the Norwegian Academy of Science and Letters in Oslo during the academic year 2019–2020, working on the research project MultiGender.

Competing Interests

The authors have no competing interests to declare.

References

Binanzer, Anja. 2017. *Genus–Kongruenz und Klassifikation: Evidenzen aus dem Zweitspracherwerb des Deutschen* (Bd. 17). Berlin/Boston: Walter de Gruyter GmbH & Co KG. DOI: https://doi.org/10.1515/9783110548549

Bittner, Dagmar. 2006. Case before gender in the acquisition of German. *Folia Linguistica* 40(1 – 2). 115–134. DOI: https://doi.org/10.1515/flin.40.1part2.115

Bottari, Piero & Cipriani, Paola & Chilosi, Anna Maria. 1993/1994. Protosyntactic devices in the *acquisition* of Italian free morphology. Language Acquisition 3(4). 327–369. DOI: https://doi.org/10.1207/s15327817la0304_1

Caffarra, Sendy & Janssen, Niels & Barber, Horacio A. 2014. Two sides of gender: ERP evidence for the presence of two routes during gender agreement processing. *Neuropsychologia* 63. 124–34. DOI: https://doi.org/10.1016/j.neuropsychologia.2014.08.016

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

Corbett, Greville. 1982. Gender in Russian: An account of gender specification and its relationship to declension. *Russian Linguistics* 6. 197–232. DOI: https://doi.org/10.1007/BF03545848

Corbett, Greville. 1991. *Gender.* Cambridge: Cambridge University Press.

Corbett, Greville & Fraser, Norman. 2000. Default genders. In Unterbeck, Barbara & Rissanen, Matti (eds.), *Gender in Grammar and Cognition*, 55–97. Berlin: De Gruyter. DOI: https://doi.org/10.1515/9783110802603.55

Cornips, Leonie & Hulk, Aafke. 2008. Factors of success and failure in the acquisition of grammatical gender in Dutch. *Second Language Research* 24(3). 267–295. DOI: https://doi.org/10.1177/0267658308090182

Corteen, Emma C. 2019. *The assignment of grammatical gender in German: Testing Optimal Gender Assignment Theory*. University of Cambridge dissertation.

Dieser, Elena. 2009. *Genuserwerb im Russischen und Deutschen: Korpusgestützte Studie zu ein-und zweisprachigen Kindern und Erwachsenen.* München/Berlin: Sagner. DOI: https://doi.org/10.3726/b12030

Egger, Evelyn & Hulk, Aafke & Tsimpli, Ianthi Maria. 2018. Crosslinguistic influence in the discovery of gender: the case of Greek–Dutch bilingual children. *Bilingualism: Language and Cognition* 21(4). 694–709. DOI: https://doi.org/10.1017/S1366728917000207

Eichler, Nadine & Jansen, Veronika & Müller, Natascha. 2013. Gender acquisition in bilingual children: French–German, Italian–German, Spanish–German and Italian–French. *International Journal of Bilingualism* 17(5). 550–572. DOI: https://doi.org/10.1177/1367006911435719

Enger, Hans-Olav. 2009. The role of core and non-core semantic rules in gender assignment. *Lingua* 119(9). 1281–1299. DOI: https://doi.org/10.1016/j.lingua.2009.02.004

Hopp, Holger. 2011. Internal and external factors in the child L2 development of the German determiner Phrase. *Linguistic Approaches to Bilingualism* 1(3). 238–264. DOI: https://doi.org/10.1075/lab.1.3.02hop

Horst, Jessica S. & Hout, Michael C. 2016. The Novel object and unusual name (NOUN) database: a collection of novel images for use in experimental research. *Behavioral Research Method* 48(4). 1393–1409. DOI: https://doi.org/10.3758/s13428-015-0647-3

Hulk, Aafke & van der Linden, Elisabeth 2010. How vulnerable is gender? In Guijarro-Fuentes, Pedro & Domiínguez, Laura (eds.), *New Directions in Language Acquisition: Romance Languages in the Generative Perspective,* 107–134. Cambridge Scholars Publishing.

Kaltsa, Maria & Tsimpli, Ianthi Maria & Argyri, Froso. 2019. The development of gender assignment and agreement in English-Greek and German-Greek bilingual children. *Linguistic Approaches to Bilingualism* 9(2). 253–288. DOI: https://doi.org/10.1075/lab.16033.kal

Karmiloff-Smith, Annette. 1979. *A functional approach to child language: A study of determiners and reference.* Cambridge: Cambridge University Press.

Klassert, Annegret & Gagarina, Natalia & Kauschke, Christina. 2014. Object and action naming in Russian-and German-speaking monolingual and bilingual children. *Bilingualism: Language and Cognition* 17(1). 73–88. DOI: https://doi.org/10.1017/S1366728913000096

Köpcke, Klaus-Michael. 1982. *Zum Genussystem der deutschen Gegenwartssprache*. Tübingen: Niemeyer. DOI: https://doi.org/10.1515/9783111676562

Köpcke, Klaus-Michael & Zubin, David. 1983. Die kognitive Organisation der Genuszuweisung zu den einsilbigen Nomen der deutschen Gegenwartssprache. *Zeitschrift für germanistische Linguistik* 11. 166–182. DOI: https://doi.org/10.1515/zfgl.1983.11.2.166

Köpcke, Klaus-Michael & Zubin, David. 2009. Genus. In Hentschel, Elke & Vogel, Petra M. (eds.) *Deutsche Morphologie*,132–154. Berlin: de Gruyter.

Kupisch, Tanja & Müller, Natascha & Cantone, Katja Francesca. 2002. Gender in monolingual and bilingual first language acquisition: Comparing Italian and French. *Lingue e Linguaggio* 1. 107–150.

Lemmerth, Natalia & Hopp, Holger. 2018. Gender processing in simultaneous and successive bilingual children: Cross-linguistic lexical and syntactic influences. *Language Acquisition* 26(1). 21–45. DOI: https://doi.org/10.1080/10489223.2017.1391815

Lenth, Russell & Singman, Henrik & Love, Jonathon & Buerkner, Paul & Herve, Maxime. 2019. R package *emmeans*: Estimated Marginal means, aka Least-Squares Means. R repository: https://cran.r-project.org/web/packages/emmeans/index.html

Lidz, Jeffrey & Gagliardi, Annie. 2015. How Nature Meets Nurture: Universal Grammar and Statistical Learning. *Annual Review of Linguistics* 1(1). 333–353. DOI: https://doi.org/10.1146/annurev-linguist-030514-125236

Mills, Anne E. 1986. *The Acquisition of Gender. A Study of English and German*. Berlin/Heidelberg: Springer. DOI: https://doi.org/10.1515/ling.1986.24.1.31

Mitrofanova, Natalia & Rodina, Yulia, & Urek, Olga & Westergaard, Marit. 2018. Bilinguals' sensitivity to grammatical gender cues in Russian: the role of cumulative input, proficiency, and dominance. *Frontiers in Psychology* 9. 1894. DOI: https://doi.org/10.3389/fpsyg.2018.01894

Mitrofanova, Natalia & Urek, Olga & Rodina, Yulia & Westergaard, Marit. 2021. Sensitivity to microvariation in bilingual acquisition: morphophonological gender cues in Russian heritage language. *Applied Psycholinguistics*, 1–39. DOI: https://doi.org/10.1017/S0142716421000382

Montanari, Elke G. & Abel, Roman & Graßer, Barbara & Tschudinovski, Lilia. 2018. Do bilinguals create two different sets of vocabulary for two domains? *Linguistic Approaches to Bilingualism* 8(4). 502–522. DOI: https://doi.org/10.1075/lab.16021.mon

Müller, Natascha. 1999. Gender and number in acquisition. In Unterbeck, Barbara & Rissanen, Matti (eds.), *Gender in grammar and cognition: Approaches to Gender*, 351–399. The Hague: Mouton de Gruyter. DOI: https://doi.org/10.1515/9783110802603.351

Opitz, Andreas & Thomas, Pechmann. 2016. Gender features in German. Evidence for underspecification. *The Mental Lexicon* 11(2). 216–241. DOI: https://doi.org/10.1075/ml.11.2.03opi

Pérez-Pereira, Miguel. 1991. The acquisition of gender: what Spanish children tell us. *Journal of Child Language* 18(3). 571–590. DOI: https://doi.org/10.1017/S0305000900011259

Polinsky, Maria. 2008. Gender under incomplete acquisition: heritage speakers' knowledge of noun categorization. *Heritage Language Journal* 6. 40–71. DOI: https://doi.org/10.46538/hlj.6.1.3

Rodina, Yulia. 2008. *Semantics and Morphology: The Acquisition of Grammatical Gender in Russian*. University of Tromsø dissertation.

Rodina, Yulia & Westergaard, Marit. 2012. A cue-based approach to the acquisition of grammatical gender in Russian. *Journal of Child Language* 39. 1077–1106. DOI: https://doi.org/10.1017/S0305000911000419

Rodina, Yulia & Westergaard, Marit. 2015. Grammatical gender in Norwegian: Language acquisition and language change. *Journal of Germanic Linguistics* 27(2). 145–187. DOI: https://doi.org/10.1017/S1470542714000245

Rodina, Yulia & Westergaard, Marit. 2017. Grammatical gender in bilingual Norwegian-Russian acquisition: The role of input and transparency. *Bilingualism: Language and Cognition* 20(1). 197–214. DOI: https://doi.org/10.1017/S1366728915000668

Ruberg, Tobias. 2013. *Der Genuserwerb ein- und mehrsprachiger Kinder*. Hamburg: Dr. Kovac.

Sabourin, Laura & Stowe, Laurie A. & de Haan, Ger J. 2006. Transfer effects in learning a second language grammatical gender system. *Second Language Research* 22(1). 1–29. DOI: https://doi.org/10.1191/0267658306sr259oa

Schiller, Niels O. & Caramazza, Alfonso. 2003. Grammatical feature selection in noun phrase production: Evidence from German and Dutch. *Journal of Memory and Language* 48(1). 169–194. DOI: https://doi.org/10.1016/S0749-596X(02)00508-9

Steinmetz, Donald. 2006. Gender shifts in Germanic and Slavic: semantic motivation for neuter? *Lingua* 116(9). 1418–1440. DOI: https://doi.org/10.1016/j.lingua.2004.06.014

Stern, Clara & Stern, William. 1928. *Die Kindersprache: Eine psychologische und sprachtheoretische Untersuchung*. Leipzig: Barth.

Szagun, Gisela & Stumper, Barbara & Sondag, Nina & Franik, Melanie. 2007. The acquisition of gender marking by young German-speaking children: Evidence for learning guided by phonological regularities. *Journal of Child Language* 34(3). 445–471. DOI: https://doi.org/10.1017/S0305000906007951

Tracy, Rosemarie. 1986. The acquisition of case morphology in German. *Linguistics* 24. 47–78. DOI: https://doi.org/10.1515/ling.1986.24.1.47

Unsworth, Sharon & Argyri, Froso & Cornips, Leonie & Hulk, Aafke & Sorace, Antonella & Tsimpli, Ianthi. 2014. On the role of age of onset and input in early child bilingualism in Greek and Dutch. *Applied Psycholinguistics* 35(4). 765–805. DOI: https://doi.org/10.1017/S0142716412000574

van Dijk, Chantal & vam Wonderen, Elise & Koutamanis, Elly & Kootstra, Gerrit Jan & Dijkstra, Ton & Unsworth, Sharon. 2021. Cross-linguistic influence in simultaneous and early sequential bilingual children: A meta-analysis. *Journal of Child Language*. DOI: https://doi.org/10.1017/S0305000921000337

Wegener, Heide. 1995. Das Genus im DaZ-Erwerb. Beobachtungen an Kindern aus Polen, Russland und in der Türkei. In Handwerker, Brigitte (ed.), *Fremde Sprache Deutsch. Grammatische Beschreibung, Erwerbsverläufe, Lehrmethodik*, 1 – 24. Tübingen: Narr.

Zubin, David & Köpcke, Klaus-Michael. 1986. Gender and folk taxonomy: The indexical relation between grammatical and lexical categorization. In Craig, Colette G. (ed.), *Noun classes and categorization*, 139–180. John Benjamins. DOI: https://doi.org/10.1075/tsl.7.12zub