Morphology (2023) 34:1–32
https://doi.org/10.1007/s11525-023-09418-3

Processing of grammatical gender agreement morphemes in Polish: evidence from the Visual World Paradigm

Zuzanna Fuchs[1]

Received: 7 January 2022 / Accepted: 25 September 2023 / Published online: 13 October 2023

Abstract
This paper presents a psycholinguistic study of the processing of grammatical gender agreement morphemes in Polish, which has three gender categories (masculine, feminine, neuter), as well as what language-internal factors impact this processing. Results from an eye-tracking study using the Visual World Paradigm show that, during real-time language comprehension, adult monolingual speakers of Polish use cues from gender agreement on a prenominal adjective to anticipate the upcoming noun. An exploration of language-internal factors affecting this anticipatory processing finds this effect in all three genders, suggesting that encountering the relevant nominative-case agreement morpheme during language comprehension leads to automatic activation of a gender node in the mental lexicon, consistent with the literature on other languages with grammatical gender. These results hold true for the neuter agreement morpheme, despite the fact that this morpheme also instantiates default gender agreement in the language and is syncretic with the nominative plural agreement morpheme in all three genders. Further investigation finds that, while agreement morphemes for each gender prompt anticipatory processing, the reliability of a masculine agreement morpheme as a cue to gender is reduced in the presence of a neuter distractor, and vice versa. This raises questions regarding phonological proximity between the realized suffix and the suffix that would cue the distractor, with implications for the acquisition and processing of gender agreement morphology in Polish.

Keywords Grammatical gender · Polish · Psycholinguistics · Eye-tracking

1 Introduction

Grammatical gender is a categorizing feature that groups nouns in a given language into classes based on how they interact with the inflectional properties of other words

✉ Z. Fuchs
zfuchs@usc.edu

[1] Department of Linguistics, University of Southern California, Los Angeles, CA, USA

in the clause (Corbett, 1991; Hockett, 1958; Kramer, 2015). In languages with grammatical gender, the number of categories varies – anywhere from two categories (ex. Spanish: masculine and feminine; Swedish: common and neuter), to three (ex. German: masculine, feminine, neuter), to twelve or more in some Bantu languages. In some languages, the gender category that a noun is assigned to can be arbitrary; for instance, in a language like Spanish, for a noun referring to an inanimate item such as a table or pen, there is no real-world property that would determine whether it should be treated as masculine or feminine. And yet despite the arbitrariness of this categorization, this information is acquired relatively early by children, who in most languages with gender systems develop gender categories around age 2-3 (ex. Spanish, cf. Mariscal, 2009), although there is variation in this according to the transparency of the gender system. In adulthood, monolingual speakers make relatively few errors in the production and comprehension of grammatical gender agreement[1] in naturalistic settings (for an overview of work on naturalistic and experimentally elicited agreement errors, see Wang & Schiller, 2019 and Schriefers & Jescheniak, 1999).

The arbitrary and yet robust nature of grammatical gender in some languages has prompted a number of questions both in the literature on the morphosyntax of gender as well as the literature on the processing of gender. In the morphosyntactic literature, two main questions concern the structural representation of grammatical gender: where in the syntax this feature is located (ex. Fuchs et al., 2015; Fuchs & van der Wal, 2022; Kramer, 2015; Scontras et al., 2018; Steriopolo & Wiltschko, 2010), and what the internal makeup of the feature is, i.e. what is the relationship between the possible values of the feature (ex. Adamson & Šereikaite, 2019; Harley & Ritter, 2002; Kramer, 2015; Wechsler, 2008).

In the psycholinguistics literature, one prominent question concerns the processing of gender agreement features during real-time language comprehension. In particular, research has asked whether, when the listener encounters an agreement morpheme in the speech stream, they are able to access the abstract gender feature on the morpheme and integrate it into their word recognition to anticipate properties of the upcoming noun. The general answer appears to be yes, as evidenced by faster looks to target items in the Visual World Paradigm (Tanenhaus et al., 1995) when the target noun is preceded by an article or adjective that matches only the target noun in gender (Aumeistere et al., 2022; Cholewa et al., 2019; Dussias et al., 2013; Fuchs, 2021; Grüter et al., 2012; Hopp & Lemmerth, 2016; Lemmerth & Hopp, 2019; Loerts et al., 2013; Lew-Williams & Fernald, 2007, 2010; Mornati et al., 2023; Sekerina, 2015; a.o.). In a display with only two images and an auditory stimulus with a prenominal adjective marked for gender agreement with the noun, the logic is as follows. If the two images on the screen are of the same grammatical gender (a "match" condition), then gender marking on the prenominal determiner or adjective in the auditory stimulus does not provide any additional cues to identifying the target image; the first disambiguating cue to the target item is the phonological onset of the lexical item in the auditory prompt. If, however, the two images on the screen are of different

[1] Here and throughout, "agreement" is used atheoretically to indicate matching in features between a noun and either an element in the verbal domain or in the nominal domain, the latter sometimes referred to as "concord".

genders (a "mismatch" condition), then the gender agreement on the prenominal adjective provides a cue that can be used by the participant to uniquely identify the target item earlier than in the match condition.

The mechanism driving these anticipatory effects is thought to pertain to the automatic spread of pre-activation in the mental lexicon. Assuming that there are abstract gender nodes in the mental lexicon, each of which is connected to all noun lemmas in the corresponding gender category in the language (Roelofs, 1992; Levelt et al., 1999), and assuming an interactive activation model of the lexicon (McClelland & Rumelhart, 1981; Rumelhart & McClelland, 1982),[2] lexical retrieval is suggested to proceed as follows. Participants hear a gender cue on the agreeing element, which automatically activates the corresponding gender node. This spreads activation to the set of nouns of that gender, which facilitates word recognition once the target noun is encountered in the input.

But as Kaan and Grüter (2021) point out in an overview of studies in this area, facilitative (or predictive) use of gender information is not all-or-nothing. Within this literature, the focus has largely been on exploring language-external (or rather, experience-based) factors that affect an individual's access to and integration of gender information from a morphological cue during real-time language comprehension. Much of this concerns factors in the processing of gender agreement by second language learners, such as the role of exposure, language proficiency, and language transfer in accounting for variability in using gender cues predictively. Within the literature on monolinguals, experience has also been shown to play a role. When presented with instances of ungrammatical agreement in an experimental context, participants stop using gender cues predictively, suggesting recent experience can affect their perception of gender cues as reliable (Brothers et al., 2019; Heyselaar et al., 2020; Hopp, 2016). Additionally, variation in these effects within monolingual populations may be in part dependent on working memory (Huettig & Janse, 2016; Ito et al., 2018).

Less emphasis has been placed on understanding language-internal factors that impact these effects in language processing. In a study of Dutch, which has a two-gender system, with gender categories common and neuter, Loerts et al. (2013) found that participants used only common gender cues on adjectives and definite articles for prediction, and Johnson (2005) found similar results for Dutch-speaking children. Loerts et al. (2013) observe that all nouns in Dutch become neuter when diminutivized; as a result, neuter pre-nominal agreement may not be a reliable cue to properties of the upcoming noun. However, Brouwer et al. (2017) found that Dutch adults could use both genders facilitatively, while children could only use neuter facilitatively, in contrast with previous work; the authors suggest differences in age range (for the children) and experimental design may be the cause of discrepancies. Huettig and Guerra (2019) did not report differences between use of common and neuter gender cues on Dutch articles. In Spanish, also a two-gender system, asymmetries in the use of gender cues on articles have not been noted for monolingual speakers but have

[2]While the discussion and proposal do not necessarily hinge on the exact model of the mental lexicon assumed, for concreteness the assumptions are made clear here and interpretation of results proceeds in Sect. 5 with reference to these assumptions.

been observed in second-language speakers of Spanish whose first language is Italian, as well as for low-proficiency second-language speakers of Spanish whose first language is English (Dussias et al., 2013).

Language-internal asymmetries in the use of gender cues in three-gender systems have to-date been discussed mostly in passing. In a study on the predictive use of definite articles in German, which index gender agreement (masc. *der*, fem. *die*, neut. *das*), Hopp (2013) noted that masculine articles led to slower anticipatory processing than did neuter or feminine. Hopp does not discuss this asymmetry further, though it is a surprising result given that *die* is the definite plural article in all genders, which would imply – in the spirit of the Dutch findings in Loerts et al. (2013) – that *die* might also be perceived as an unreliable cue. Hopp and Lemmerth (2016) investigated predictive use of gender morphology on German determiners and adjectives, but reported only an overall analysis without comparison between gender cues; they did, however, report that cues on adjectives led to overall faster looks to the target than did cues on articles. Cholewa et al. (2019) found that German-speaking children can also use gender facilitatively but did not report differences between the three genders. In studies of use of adjectival agreement in Russian, another three-gender language, Sekerina (2015) and Aumeistere et al. (2022) used experimental items only in masculine and feminine. In a study on the use of adjectival agreement in Polish, Fuchs (2022) found that heritage speakers of Polish use agreement morphology to anticipate properties of the noun, as did control speakers, and while the author noted group-internal asymmetries in use of the three gender cues, these were not discussed further.[3] In an experimental study on the anticipatory use of gender-inflected articles in Norwegian, Lundquist and Vangsnes (2018) did in fact focus on language-internal asymmetries in gender use, because some Norwegian dialects have shifted from being a three-gender system to a two-gender system, in which the distinction between masculine and feminine nouns has disappeared. The authors found that speakers of a three-gender dialect of Norwegian could use articles in all three genders to anticipate the upcoming noun, but speakers of a two-gender dialect could not use masculine articles anticipatorily when the competing noun was feminine.

This sets the stage for a more systematic approach to understanding the language-internal properties that may impact the processing of gender morphology to anticipate properties of an upcoming noun. This matter is particularly compelling in a three-gender system, in which the reliability of a gender agreement morpheme could be dependent not only on the gender of the target noun – as it would be in a two-gender system – but also on the gender of a competing candidate noun. Specifically in a mismatch condition, in which the prenominal gender cue is expected to be reliable, the gender of the competing candidate item can vary (as opposed to in a two-gender system, in which the competing candidate item will always be of the same gender), and this could play a role in the processing of the agreement morpheme that is realized in the speech stream.

Polish offers a neat empirical domain for such an exploration. In addition to being a three-gender system, Polish has an adjectival agreement paradigm in which, at

[3]The results presented in the present manuscript in Sect. 4 are also included in Fuchs (2022) as the results for the control group, for comparison with the heritage bilingual group, who are the focus of that paper.

least in the nominative singular, agreement in each gender is equally morphologically specified (see Alemán Bañón & Rothman (2016) for discussion of known effects of morphological specification on processing) and unambiguous between genders. Moreover, unlike in Russian, unstressed vowels in Polish do not undergo substantive reduction that would lead to opacity of the gender cue (Crosswhite, 2001; Janssen, 2016; Rojczyk, 2019). These properties allow for a direct comparison of the reliability of gender agreement morphemes in the three genders both as a function of the agreement morpheme itself and as a function of the gender category of the competing candidate item.

To pursue this matter, Sect. 2 will first introduce the relevant properties of Polish grammatical gender and adjectival agreement that will be the empirical domain for the experimental study presented in Sect. 3. Analysis and results will be presented in Sect. 4, and the implications of these findings as well as the open questions that they invite will be discussed in Sect. 5. Section 6 concludes.

2 Gender in Polish

2.1 Three-gender system

Polish is generally classified as having three genders, like many other members of the Slavic language family, and six cases (excluding the vocative, which is generally thought to be extra-syntactic, cf. Luczynski, 2002). In the citation form in Polish (nominative singular) there are three evident "global" gender categories: masculine, feminine, and neuter. There are morphological correlates typically associated with each gender that occur on nouns: *-a* for feminine (1-a); *-o*, *-e*, or *-ę* for neuter (1-b), and consonants for masculine (1-c) (see also Table 1).

(1) a. ta dziewczyna 'this girl(F)'
ta kobieta 'this woman(F)'
ta koszula 'this shirt(F)'
ta książka 'this book(F)'
b. to jajko 'this egg(N)'
to okno 'this window(N)'
to wiadro 'this bucket(N)'
to dziewczę 'this maiden(N)'
c. ten chłopiec 'this boy(M)'
ten pies 'this dog(M)'
ten stół 'this table(M)'
ten wazon 'this vase(M)'

However, these correspondences between the morphophonology of nouns – particularly their endings, or word markers, in the sense of Harris (1991) – and the genders they are assigned to is not one-to-one, with exceptions in each gender category:

(2) a. ta sól 'this salt(F)'
ta pięść 'this fist(F)'
ta rzecz 'this thing(F)'

Table 1 Gender correlates and agreement markers in nominative singular in Polish, presented both in Polish spelling and in phonological transcription

	M		F		N	
Correlates on noun	*-C*		*-a*	/a/	*-o*	/o/
Suffix on attributive adjective	*-i* or *-y*	*/i/* or */ɨ/*	*-a*	/a/	*-ie* or *-e*	/ɛ/
Suffix on past tense verbs	*-Ø*		*-a*	/a/	*-o*	/o/

b. to coś ‘this something(N)’
to menu ‘this menu(N)’
to museum ‘this muzeum(N)’
c. ten mężczyzna ‘this man(M)’
ten artysta ‘this artist(M)’
ten sędzia ‘this judge(M)’

Given the lack of a one-to-one correspondence between gender and the morphophonology of nouns in that gender category, the most reliable indicator of the gender of nouns in Polish, as in many gendered languages, is the behavior of associated words (Hockett, 1958). In Polish, gender is marked on elements such as demonstratives (as in (1) and (2) above), attributive and predicative adjectives, relative pronouns, and/or verbs in the past tense and future imperfective tense (Swan, 2015). This is illustrated for demonstratives, attributive adjectives, and verbs in (3).

(3) a. Ten star-y wazon był w kuchni.
dem.M.SG old-M.SG vase(M) be.PST.3SG.M in kitchen
‘That old vase was in the kitchen.’
b. Ta star-a książka był-a w kuchni.
dem.F.SG old-F.SG book(F) be.PST-3SG.F in kitchen
‘That old book was in the kitchen.’
c. To star-e wiadro był-o w kuchni.
dem.N.SG old-N.SG bucket(N) be.PST-3SG.N in kitchen
‘That old bucket was in the kitchen.’

The above example illustrates inflection on agreeing elements with nouns in the nominative singular; a simplified system of nominal and adjectival markers in the nominative singular is schematized for illustration in Table 1. As mentioned, unlike in Russian, these gender suffixes do not undergo vowel reduction, despite typically occurring in unstressed position (Crosswhite, 2001; Rojczyk, 2019) and are thus considered to be consistently transparent for gender (cf. Janssen, 2016).

Nouns and adjectives in Polish inflect not only for gender but also for number and case. Thus, Table 1 is only a fraction of the paradigm. Thus far the discussion has been restricted to the nominative singular, and although the experimental study will also focus on the nominative singular as well, at this point it is useful to address how this fits into the broader paradigm. Table 2 provides an illustration of the inflectional paradigm by providing the paradigm for *stary* ‘old’ in the three global genders (masculine, feminine, neuter), in singular and plural, and in the six cases. As can be seen in Table 2, the inflectional paradigms of the three genders are entirely syncretic in

Table 2 An illustration of the inflectional paradigm of three global (inanimate) genders in the singular and plural of six cases for the adjective *stary* 'old'

	Singular			Plural		
	M	F	N	M	F	N
Nominative	stary	stara	stare	stare	stare	stare
Genitive	starego	starej	starego	starych	starych	starych
Dative	staremu	starej	staremu	starym	starym	starym
Accusative	stary	starą	stare	stare	stare	stare
Instrumental	starym	starą	starym	starymi	starymi	starymi
Locative	starym	starej	starym	starych	starych	starych

Table 3 An illustration of the inflectional paradigm of subcategories of masculine nouns for the adjective *stary* 'old'

	Singular		Plural			
	Animate	Inanimate	Personal	De-virilized	Animate	Inanimate
Nominative	stary	stary	starzi	stare	stare	stare
Genitive	starego	starego	starych	starych	starych	starych
Dative	staremu	staremu	starym	starym	starym	starym
Accusative	starego	stary	starych	starych	stare	stare
Instrumental	starym	starym	starymi	starymi	starymi	starymi
Locative	starym	starym	starych	starych	starych	starych

the plural. Note however that Table 2 illustrates adjectives agreeing with inanimate nouns only.

When animacy distinctions are considered, the masculine category itself is not uniform, with minimally different inflectional paradigms – including in the plural – for a few subgroups, as illustrated in Table 3 (the feminine and neuter do not have similar subdivisions). For masculine nouns, the inflectional paradigm of inanimate nouns (the empirical domain of the experimental study presented below) is distinguished from that of animate nouns in the singular accusative (animate *-ego* vs inanimate *-y* on agreeing adjectives). In other words, in the masculine animate paradigm, the accusative is syncretic with the genitive rather than with the nominative. The masculine personal category contains those nouns whose referents are male and human, such as *mężczyzna* 'man'; in the adjectival inflectional paradigm they are set apart by a unique form in the nominative plural and by the accusative plural suffix *-ych*, syncretic with the masculine genitive and locative plural. The masculine de-virilized category contains nouns that are referentially masculine personal but have a pejorative connotation, such as *brudas* 'dirty person.' They share the accusative plural agreement suffix *-ych* with masculine personal nouns but diverge in the nominative plural (cf. *starzy mężczyźni* 'old men, nom. pl.' versus *stare brudasy* 'old dirty ones, nom. pl.').

While the inclusion of animacy distinctions is important in a broad introduction to Polish agreement morphology, the empirical domain of the present study excludes

animate nouns, investigating gender agreement morphology with inanimate nouns only. There is independent motivation for this choice, from several perspectives. First, work on the morphosyntax of grammatical gender argues for a distinction between the gender of inanimate nouns (grammatical gender) and the gender of animate nouns (conceptual gender), and some syntactic research suggests these may even be represented as separate features in the mental grammar (Bobaljik & Zocca, 2011; Kramer 2009, 2014; Steriopolo & Wiltschko, 2010; although see Kramer, 2015, for arguments in favor of not separating animate and inanimate gender features). Furthermore, inclusion of animate nouns in an experimental paradigm would mean that properties of the real-world referent might be used by participants to recognize the gender category of the upcoming noun, whereas the goal of this study is to isolate and observe use of the abstract grammatical gender feature. Additionally, while animacy clearly plays a role in determining membership in what Corbett (1983) considers to be subgenders of masculine in Polish, the exact status of the animacy feature in the modern Polish grammar warrants further investigation. Bogusławski (1986) and Kosta (2003) note that words for things such as dances and mushrooms, among others, also determine the masculine animate agreement paradigm in Table 3 despite being semantically inanimate. Additionally, Fuchs (2014) demonstrates that, regardless of their semantic animacy, recent loanwords from English to Polish are inflected in the accusative with the nominal case suffix *-a*, which is usually restricted to animate masculine nouns (with the exceptions discussed by Bogusławski, 1986, and Kosta, 2003). These data raise questions as to the status of animacy distinctions in the Polish gender system.

In addition to the considerations noted above, narrowing the domain to only the three global genders – masculine, feminine, and neuter – is consistent with existing experimental work on Polish (ex. Brehmer & Rothweiler, 2012). Additionally, only the singular number will be considered; given the syncretism in declension paradigms across the genders in the plural, incorporating plurality would not be informative in this study.

2.2 Default gender

Given that the study below pertains to the use of gender agreement morphology during processing, it is important to discuss default gender agreement in the language. Determining which gender value is "default" requires identifying which gender occurs when no other rules apply (Corbett & Fraser, 1999; Haspelmath, 2006). The notion of default has been used to refer to various concepts in the literature (Haspelmath, 2006), and environments in which a default will be applied may be language-specific. It is therefore important to make explicit that, given the nature of the study below, the discussion of gender default as presented here pertains specifically to what gender agreement morphology is observed in Polish when there is no overt controller of agreement or when a controller of agreement lacks gender specification.

In the absence of a controller of agreement in (4), both the verb and the predicate adjective occur with neuter gender morphology. Similarly, in the absence of gender specification, as when the controller of agreement does not have an inherent gender feature (ex. a clausal subject, as in (5)), this also determines neuter agreement on the verb. Metalinguistic use of a word – which should not carry an inherent gender feature – will also trigger neuter agreement on an attributive adjective (6).

(4) Był-o zimn-o.
be.PST-3SG.N cold-N.SG
'(It) was cold (outside).'

(5) [Że Jaś nie przeczytał lektur-y] był-o
COMP Jaś(M) NEG read.PST.3SG.M book(F)-GEN.SG be.PST-3SG.N
jasn-e.
clear-3SG.N
'That Jaś had not read the schoolbook was clear.'

(6) Rozmazan-e "nie" sprawił-o, że znak był
smudged-N.SG no cause-PST.3SG.N COMP sign(M).SG be.PST.3SG.M
nie-czyteln-y
not-legible-M.SG
'The smudged "no" caused the sign to be illegible.'

Crosslinguistically, there are a number of other environments that can be considered for evidence of default gender agreement, but these are not applicable in Polish for various reasons.[4] Adjectival agreement with coordinated noun phrases can sometimes constitute a test case, as it is an instance of conflicting gender cues and may be an environment in which a language deploys a default agreement strategy. However, in Polish, closest conjunct agreement is applied in this environment (Willim, 2012). Verbal agreement with coordinated noun phrases, which might also be used to test for default agreement in some languages, also does not provide insight into default gender agreement in Polish because coordinated noun phrases are always plural; recall that the inflectional paradigms for plurals in all three gender categories are syncretic, modulo animacy distinctions for the masculine discussed above. When both noun phrases are animate, verbal agreement with coordinated noun phrases suggests masculine as the default for gender agreement; however, as discussed, animacy (and thus conceptual gender) is outside of the scope of the present study.

2.3 Research questions

The goal of this study is to investigate how gender information on agreement morphology in Polish is processed. Specifically, by using eye-tracking methodology within the Visual World Paradigm (Tanenhaus et al., 1995), this study will measure how Polish speakers use gender agreement marking on prenominal adjectives to facilitate lexical retrieval of the subsequent noun during real-time language processing.

[4] An anonymous reviewer asks about agreement with loanwords as a diagnostic for default gender. Indeed, this diagnostic has been used in various languages as additional evidence for default gender agreement in the presence of an overt controller of agreement that lacks a gender feature. However, the behavior of loanwords in Polish with respect to gender agreement appears to be undergoing a change: while many older loanwords determined masculine (inanimate) agreement (not including those that ended in *-a* or *-o*, ex. *ta opera* 'this opera(F)' or *to radio* 'this radio(N)'), current loanwords appear to generally pattern with masculine animate nouns, regardless of semantic animacy (Fuchs, 2014). This does not align with the diagnostics for agreement with an overt controller lacking a gender feature in (5) and (6) in this section, and is itself an area for future research. With these considerations in mind, it would be an unreliable diagnostic on which to base an understanding of default gender agreement given the availability of other diagnostics.

As discussed in Sect. 1, this question has been addressed for two-gender languages such as Spanish (Lew-Williams & Fernald, 2007, 2010; Grüter et al., 2012; Dussias et al., 2013; Fuchs, 2021) and Dutch (Brouwer et al., 2017; Johnson, 2005; Loerts et al., 2013; Huettig & Janse, 2016), as well as for three-gender languages such as German (Cholewa et al., 2019; Hopp, 2013; Hopp & Lemmerth, 2016) and Russian (Sekerina, 2015; Aumeistere et al., 2022). However, the focus of most of these studies, with the exception of work on Dutch (Brouwer et al., 2017; Johnson, 2005; Loerts et al., 2013) and Norwegian (Lundquist & Vangsnes, 2018), was variation in these effects based on language-external factors such as first- vs second-language acquisition, age, and working memory. Language-internal distinctions either were not pursued or were mentioned as secondary to the research aims.

The present study therefore investigates possible language-internal factors that may affect the facilitative use of gender for lexical retrieval in a three-gender system. By narrowing the empirical domain to agreement targets in Polish that are inanimate and occur in the nominative singular, the experimental design described below will ensure that gender agreement morphemes are equally morphologically specified (cf. Alemán Bañón & Rothman, 2016), and as discussed before, these are not phonologically reduced and thus not opaque to gender. Given this, the first question is whether agreement morphemes for all three genders can serve as a reliable cue to gender.

Research Question 1 Can Polish speakers use pre-nominal gender cues in each of the three gender categories to facilitate lexical retrieval of the subsequent noun?

Given that monolingual speakers of languages such as German and Russian have been shown to be able to use grammatical gender on adjectives to facilitate lexical retrieval (Aumeistere et al., 2022; Hopp, 2013; Hopp & Lemmerth, 2016; Lemmerth & Hopp, 2019; Sekerina, 2015), it is expected that speakers of Polish will be able to do so as well, at least for some of the gender features. The expectations specifically for neuter are not clear, and there is reason to believe the neuter may generate less strong expectations. As discussed in Sect. 2.2, default gender agreement with inanimate nouns in Polish is neuter. Additionally, the nominative singular agreement morpheme in neuter is syncretic with the plural nominative morpheme in all three genders (cf. Table 2). The results for Polish neuter could therefore pattern with the lack of anticipatory looks observed for Dutch neuter (Johnson, 2005; Loerts et al., 2013; but see Brouwer et al., 2017 for results in the opposite direction), or they could pattern with the results for German *die*, which, while similarly syncretic with the plural form of the definite article for nouns of all genders, still led to anticipatory looks (Hopp, 2013).

Unique to working with a three-gender system to answer questions regarding language-internal variation in the use of gender to facilitate lexical retrieval is the opportunity to investigate whether this effect is dependent on the grammatical gender of the other candidate noun. In mismatch conditions in general we expect the agreement morpheme to be a reliable cue for anticipatory processing, but in a three-gender system there are two grammatical genders that may be instantiated by a distractor in the display. It is possible that manipulating this factor could reveal variation in the reliability of the agreement morpheme as a cue to gender, dependent on the gender

of the other candidate lexical item in the given context. This motivates the second research question:

Research Question 2 Is Polish speakers' use of gender cues to facilitate lexical retrieval of a target noun modulated by the gender of a competing candidate noun?

Previous studies have not reported analyzing results to this degree.[5] Should the present study find such asymmetries between the looks to the target based on the gender of the distractor, these asymmetries would have to be unpacked further to determine what may be driving them.

3 Methods

3.1 Materials & design

The design of the study follows much of the previous literature on the facilitative use of grammatical gender (Fuchs, 2021, 2022; Grüter et al., 2012; Hopp & Lemmerth, 2016; Lemmerth & Hopp, 2019; Loerts et al., 2013; Lew-Williams & Fernald, 2007, 2010; Sekerina, 2015). The prenominal words that host the gender cue in this study are prenominal adjectives (Aumeistere et al., 2022; Hopp & Lemmerth, 2016; Lemmerth & Hopp, 2019; Loerts et al., 2013; Sekerina, 2015; a.o.). These were color adjectives, as in the absence of additional context these were most felicitous.

3.1.1 Image selection

36 identifiable images of inanimate objects were selected as target items[6] for the study: 12 representing masculine items, 12 feminine, and 12 neuter. The corresponding nouns all had at least two syllables in order to allow for looking time following the onset of the noun in the auditory prompt. All lexical items also had as their first phoneme a consonant, so as to create a clear boundary between the gender information on the preceding adjective (realized as a suffixal vowel) and the onset of the lexical item itself.

The images were split into three color groups: green, red, and blue. These colors were selected because the corresponding adjectives in Polish are of approximately equal number of syllables (Table 4). For each gender, images were equally distributed among the colors: 4 images were green, 4 red, 4 blue.

[5] Here the exception is the findings in Lundquist and Vangsnes (2018): Norwegian speakers of dialects in which the three-gender system is disappearing can use masculine articles predictively only when the distractor is neuter, but not when it is feminine. This is because masculine articles can precede both traditionally masculine and traditionally feminine nouns in these dialects.

[6] See Table A.1 in Appendix A for a full list of experimental nouns, their English translations, and their frequencies in the Frequency Dictionary Corpus subpart of the National Corpus of Polish (*Narodowy Korpus Języka Polskiego*, NKJP).

Table 4 Color adjectives used in the study, with inflectional suffixes for each gender

		M	F	N
czerwon-	red	*-y*	*-a*	*-e*
zielon-	green	*-y*	*-a*	*-e*
niebiesk-	blue	*-i*	*-a*	*-ie*

Table 5 Experimental conditions for each gender

		Distractor gender		
		M	F	N
Target gender	M	match	mismatchF	mismatchN
	F	mismatchM	match	mismatchN
	N	mismatchM	mismatchF	match

3.1.2 Visual displays

The selected images were combined into 108 total pairings for the visual displays. Each display consisted of two images (matched for color) with a fixation cross in the center. In each pairing, the lexical items corresponding to the two images in the display always had a different first phoneme, in order to ensure that the onset of the noun in the auditory prompt was clearly the moment of disambiguation between candidate images in match conditions (as well as in mismatch conditions, in case the prenominal gender cue is not attended to). One image was the target item that the corresponding auditory prompt would ask the participant to direct their gaze to, and the other image was the distractor. Because there are three genders in Polish and given the goals of Research Question 2, there were three possible conditions per target gender, as presented in Table 5.

Each image appeared as the target item three times, once in a match condition and once in each of the two possible mismatch conditions for its gender. To illustrate, the image for *książka* 'book(F)' appeared once with a feminine distractor (match, Fig. 1a), once with a masculine distractor (mismatchM, Fig. 1b), and once with a neuter distractor (mismatchN, Fig. 1c). The result was 36 match stimuli, 24 mismatchM stimuli, 24 mismatchF stimuli, and 24 mismatchN stimuli.

Following previous work, the dependent measure in this study is the time of first fixation on the target item following the onset of the gender information on the adjectival suffix. If participants are able to access and deploy gender information on agreement suffixes on prenominal adjectives in order to facilitate lexical retrieval of the target item, then times of first fixation should be faster in mismatch trials, in which the gender information is a unique cue to one of the items in the visual display, than in match trials, in which the first available disambiguating information is the first phoneme of the target lexical item. Additionally, if the ability to use a particular gender feature to facilitate lexical retrieval is modulated by the gender of the distractor (per Research Question 2), then looking times to the target item of that gender may be slower in one mismatch condition than the other.

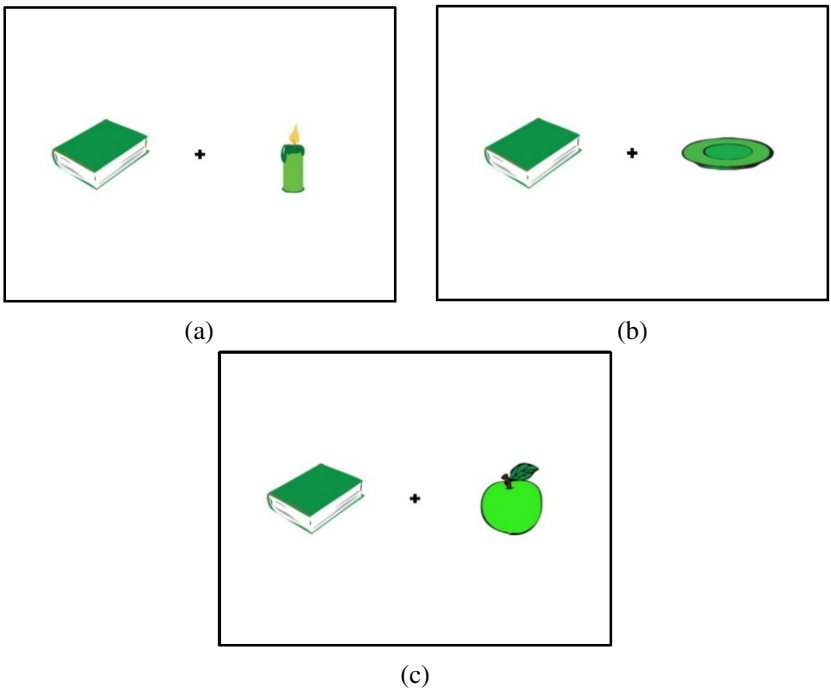

Fig. 1 (a) Sample match display with F target: *książka* 'book(F)' and *świeczka* 'candle(F)'. (b) Sample mismatchM display with F target: *książka* 'book(F)' and *talerz* 'plate(M)'. (c) Sample mismatchN display with F target: *książka* 'book(F)' and *jabłko* 'apple(N)' (Color figure online.)

3.1.3 Auditory stimuli

Each target lexical item was placed in a carrier phrase as schematized in (7). This was modeled on the experimental designs in previous work (Fuchs, 2021, 2022; Grüter et al., 2012; Hopp & Lemmerth, 2016; Lemmerth & Hopp, 2019; Loerts et al., 2013; Lew-Williams & Fernald, 2007, 2010; Sekerina, 2015).

(7) Gdzie jest COLOR-M.NOM/F.NOM/N.NOM NOUN?
where is color-M.NOM/F.NOM/N.NOM noun
'Where is (the) green/red/blue noun?

All sentences were recorded by a male native speaker of Polish immigrated to the United States less than a year prior to the time of recording. The final auditory stimuli were created by splicing together parts of the recordings, such that across all trials the onset of the gender information always occurred at the same time and the lexical item occurred at the same time.[7] The auditory stimuli began with 800 ms of looking time and a tone that indicated the end of looking time, and then 200 ms of

[7]The method of splicing was carefully manipulated to ensure that the speech still sounded like natural human speech, although some artificiality was introduced by the slightly slower rate of speech that the native speaker who recorded the stimuli was instructed to adapt (cf. Huettig & Guerra, 2019). This was to

silence before the onset of the carrier phrase. A single token of *gdzie jest* 'where is' was spliced with single tokens of each of the nine adjectives (three color adjectives each inflected for three genders, cf. Table 4), such that the duration from the onset of the carrier phrase to the onset of the gender suffix was consistently 1150 ms.[8] Using a single token of each gender-inflected adjective eliminated co-articulation effects that a target noun could have on the quality of the preceding vowel (i.e. the adjectival suffix); this strategy therefore avoids unintended phonological cues to the target item. Recordings were manipulated to ensure that the total time from the onset of the gender cue to the onset of the noun was 480 ms in all auditory stimuli. The average duration of nouns was 700 ms.

Corresponding visual displays and auditory stimuli were presented together, and each trial – consisting of a single visual display and a single auditory stimulus – lasted 6 seconds. Between each trial there was a one-second interlude during which only a fixation cross was visible on the screen.

3.2 Participants

Twenty-three adult native speakers of Polish (average age 31.8 years, sd = 8.7) recently immigrated to the United States participated in the study. The inclusion criteria required that participants had lived in Poland for at least the first 18 years of their life; on average participants in this study lived 25.0 years in Poland (sd = 8.7). Participants completed an abbreviated version of the Language Experience and Proficiency Questionnaire (LEAP-Q; Kaushanskaya et al., 2020; Marian et al., 2007) that was translated into Polish for the purposes of this study.[9]

3.3 Procedure

Participants were tested individually in a lab. They first filled out the LEAP-Q and then completed an oral lexical identification task, the results of which were used in data cleaning, as discussed in Sect. 4. During this task, participants viewed each of the 36 images used as target items during the study and were asked to orally label the image using a color adjective and a noun (Lieberman et al., 2018; Fuchs, 2021). The task was self-paced, and participants were allowed to skip an item if they were not able to identify an image or recall the appropriate lexical item.

Following this, participants completed the eye-tracking comprehension task. They sat facing a 53.5-cm screen, approximately 75 cm away from it. A chin-support apparatus ensured minimal head movement during the task. Participants received instructions in written and oral form and completed four practice trials. Subsequently,

ensure phonological breaks between words that would make splicing easier, and to allow for looking time between the gender cue and the onset of the lexical item.

[8] An anonymous reviewer correctly points out that some of the suffixed vowels could have caused coarticulation on the preceding phonemes, leading to cues to gender before the gender suffix. This is indeed a possibility. In particular, front vowels (in masculine and neuter) preceding /k/ in *niebiesk-i*/*-ie* 'blue-M/-N' could result in the palatalization of the preceding consonant.

[9] This translated version of the LEAP-Q is now available at https://bilingualism.northwestern.edu/leapq/.

an SMR Eyelink 1000 was calibrated, with the goal of achieving visual acuity below 0.5 degrees. Gaze position was recorded at 2000 Hz. The comprehension task was split into two parts of 54 trials each. After the first half, participants were allowed a break of self-determined duration. Calibration of the Eyelink 1000 was repeated before the second half of the trials. In total, the study took approximately 45 minutes, and participants were compensated for their time.

4 Analysis & results

Before analyzing the data, the results of the oral production task were used for cleaning. The purpose of the oral identification task was to determine what lexical item a given participant associated with each target image and to control for any dialectal or less common equivalents. For instance, for the image of a box, the intended lexical item was *pudełko* 'box(N)', but a few participants labeled it using the noun *paczka* 'package(F)'. Given that, for these participants, the image represented a lexical item with feminine grammatical gender (rather than the intended neuter), it would be undesirable to include in the analysis these participants' times of first fixation for trials in which the box was either the target or the distractor. Therefore, if in the oral production task a participant labeled an image using a label that was not the same as the one intended in the experimental materials, any trial containing that image was removed from the analysis for that participant. This process removed approximately 6% of the data.

For the remaining trials, times of first fixation were gathered for each participant for each trial in the SR Research Eyelink software and then analyzed in R (R Core Team, 2021). Time of first fixation was defined as the earliest fixation in the interest area of the target item later than 3250 ms from the start of the trial (i.e. after the onset of the gender suffix on the pre-nominal color adjective). Times of first fixation were trimmed to within two standard deviations of the mean; this process removed approximately 4% of the trials for the dataset.

The time of first fixation past the onset of gender information on the color adjective was compared across match and mismatch conditions. Since this is a 3×3 design – for each of three target genders there is a match condition and two mismatch conditions – the results will be presented according to the gender of the target item. In the analyses presented below, the condition variable was Helmert-contrast coded. Helmert coding compares each level of a categorical variable to previous levels of the variable. In this particular context, this allows testing for two effects of interest to the research questions. The first comparison tests for a difference between the match condition and the mean of the two mismatch conditions, thereby testing whether participants were able to fixate on targets in mismatch trials faster than on targets in match trials. For ease of exposition, this effect will be labeled CONDITION-MATCH and reflects whether participants are overall able to use grammatical gender to facilitate lexical retrieval (Research Question 1). The second comparison tests for a difference between mean looking times in one mismatch condition versus the other mismatch condition, thereby testing for an effect of the gender of the distractor on

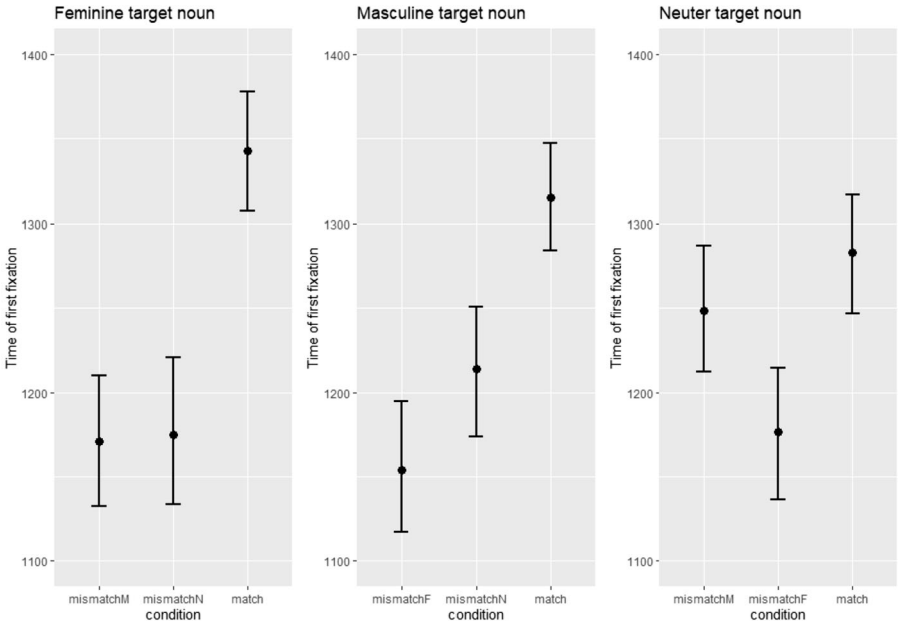

Fig. 2 Average time of first fixation on the target item after the onset of the gender cue on the pre-nominal adjective

looking times in mismatch conditions. For ease of exposition, this effect will be labeled CONDITION-DISTRACTOR in the presentation of the analysis below and reflects whether the gender of the distractor in a mismatch condition modulates the time of first fixation on the target item (Research Question 2).

4.1 Feminine target noun conditions

The mean first fixation times for trials with feminine targets are presented in Fig. 2. A mixed effects linear model was fitted to the data, predicting time of first fixation by CONDITION and TRIAL, with random intercepts and slopes for CONDITION grouped by PARTICIPANT.[10] The output of the model is provided in Table 6. The model found no significant effect of CONDITION-DISTRACTOR (mismatchM mean = 1170 ms, sd = 318 ms; mismatchN mean = 1175 ms, sd = 341 ms), but did find a significant effect of CONDITION-MATCH (match mean = 1343 ms, sd = 286 ms; mismatch mean = 1173, sd = 329 ms) ($\beta = 57.65$, SE = 10.83, $p < 0.01$). Participants' mean time of first fixation was overall faster on mismatch conditions than on the match condition, as illustrated in Fig. 2. The model also found a significant effect of TRIAL ($\beta = -1.33$, SE = 0.31, $p < 0.01$): the average time of first fixation was faster in later trials. A model with an additional INTERACTION between condition and trial did not find a significant effect of interaction. Here and

[10] A model with additional random intercepts grouped by item was also fitted to the data but resulted in singular fit.

Table 6 Fixed effects for linear mixed effects model predicting time of first fixation on feminine target items by condition and trial number

	Time of first fixation		
Condition-distractor	−1.35	(14.30)	
Condition-match	57.65	(10.83)	**
Trial	−1.33	(0.31)	**
Constant	1,127.48	(38.54)	**
Observations		731	
Log Likelihood		−5,113.20	
Akaike Inf. Crit.		10,242.40	
Bayesian Inf. Crit.		10,279.15	

Note: * p < 0.05, ** p < 0.01

Model: Time ∼ Condition + Trial + (1 + Condition | Participant)

Table 7 Fixed effects for linear mixed effects model predicting time of first fixation on masculine target items by condition and trial number

	Time of first fixation		
Condition-distractor	29.42	(11.46)	*
Condition-match	44.38	(8.83)	**
Trial	−1.40	(0.26)	**
Constant	1,127.60	(41.23)	**
Observations		769	
Log Likelihood		−5,276.46	
Akaike Inf. Crit.		10,568.91	
Bayesian Inf. Crit.		10,606.07	

Note: * p < 0.05, ** p < 0.01

Model: Time ∼ Condition + Trial + (1 + Condition | Participant)

throughout, models were fitted parsimoniously, and predictors that did not improve the fit of the model in model comparisons were removed.

4.2 Masculine target noun conditions

The mean first fixation times for conditions in which the target noun was masculine are plotted in Fig. 2. A mixed effects linear model was fitted to the data, predicting time of first fixation by CONDITION and TRIAL, with random intercepts and slopes for CONDITION grouped by PARTICIPANT. The output of the model is provided in Table 7. The model found a significant effect of CONDITION-DISTRACTOR (mismatchF mean = 1154 ms, sd = 314 ms; mismatchN mean = 1214 ms, sd = 303 ms) ($\beta = 29.42$, SE = 22.46, $p < 0.05$). Participants were overall slower on the mismatchN conditions than the mismatchF conditions. The model also found a significant effect of CONDITION-MATCH (match mean = 1315 ms, sd = 273 ms; mismatch mean = 1184 ms, sd = 310 ms) ($\beta = 44.38$, SE = 8.83, $p < 0.01$). This indicates that participants' first fixations were earlier in the mismatch conditions than in the match condition. The model also found a significant effect of TRIAL ($\beta = -1.40$, SE = 0.26, $p < 0.01$), suggesting participants' looking times increased in speed over time. A model with an additional INTERACTION between condition and trial did not find a significant effect of interaction.

Table 8 Fixed effects for linear mixed effects model predicting time of first fixation on neuter target items by condition and trial number

	Time of first fixation		
Condition-distractor	−37.66	(11.22)	**
Condition-match	23.65	(6.68)	**
Trial	−1.35	(0.26)	**
Constant	1,129.91	(44.70)	**
Observations		748	
Log Likelihood		−5,124.98	
Akaike Inf. Crit.		10,265.97	
Bayesian Inf. Crit.		10,302.90	

Note: * $p < 0.05$, ** $p < 0.01$

Model: Time ∼ Condition + Trial + (1 + Condition | Participant)

4.3 Neuter target noun conditions

The mean first fixation times on neuter target items are presented in Fig. 2. A mixed effects model predicting time of first fixation by CONDITION and TRIAL was fitted to the data, with random intercepts and slopes for CONDITION grouped by PARTICIPANT. The output of this model is presented in Table 8. The model found a significant effect of CONDITION-DISTRACTOR ($\beta = -37.66$, SE = 11.22, $p < 0.01$), indicating that participants were faster on mismatchF trials (mean = 1076 ms, sd = 316 ms) than on mismatchM trials (mean = 1148 ms, sd = 306 ms). The model also found a significant effect of CONDITION-MATCH ($\beta = 23.65$, SE = 6.68, $p < 0.01$), indicating that participants were overall slower on match trials (mean = 1182 ms, sd = 289) than on mismatch trials (mean = 1113 ms, sd = 312 ms). The model also found a significant effect of TRIAL ($\beta = -1.35$, SE = 0.26, $p < 0.01$), indicating that fixation times increased in speed over time. A model with an additional INTERACTION between condition and trial did not find a significant effect of interaction.

5 Discussion

5.1 Facilitative use of grammatical gender in Polish

I first discuss the results with respect to Research Question 1, repeated here: Can Polish speakers use pre-nominal gender cues in each of the three gender categories to facilitate lexical retrieval of the subsequent noun? The study found looks to target items to be faster in mismatch condition trials than in match condition trials in all genders. These results demonstrate that, during real-time language processing, Polish speakers do access abstract gender features on prenominal agreement morphemes; these features are integrated into the word recognition process through the automatic activation of a gender node in the mental lexicon that subsequently spreads pre-activation to all connected lemmas. This allows participants to anticipate properties of the subsequent noun.

The results for the neuter are particularly notable. As mentioned, neuter agreement is default in Polish, and thus even in the singular it does not uniquely cue neuter

nouns. Second, the neuter gender agreement morpheme in the nominative singular is syncretic with the nominative plural and accusative plural agreement morphemes for all three genders for inanimate nouns. Thus, the adjectival suffix *-e* can in principle cue nouns of any gender – much like the definite article *die* in German cues feminine in the singular but all genders in the plural (cf. Hopp, 2013) and how neuter in Dutch can cue any noun in the diminutive (cf. Loerts et al., 2013). And yet in this study the suffix served as a reliable cue, aligning more with the results for German *die* (though see further discussion in Sect. 5.2). This could be at least partly driven by two aspects of the experimental method. The first is related to processing: in the auditory prompt, the adjective was preceded by the verb *jest* 'is', which is singular. If participants held this singular agreement in memory and integrated it into recognition of the feature content of the adjectival suffix, this would explain at least why any unreliability stemming from syncretism with the plural paradigm might be obscured in the current findings. Additionally, all target items in visual displays were single items. Thus, participants could plausibly not attend to the singular verb in the prompt but still adapt to the experimental design and anticipate only singular nouns, again eliminating ambiguity in the featural content of the suffix *-e*. This would be consistent with studies in the Visual World Paradigm that have shown that recent experience within a study can affect participants' perception of gender cues as more or less reliable (Brothers et al., 2019; Heyselaar et al., 2020; Hopp, 2016). Follow-up studies with singular and plural targets may investigate this matter further.

Having established that all three gender morphemes are themselves reliable cues to gender that participants do attend to and integrate into word recognition during processing, we can transition to the second research question, which asks whether this reliability is dependent on the properties of a competing candidate. The remainder of this section is therefore dedicated to implications of the results for Research Question 2 and the open questions they invite.

5.2 Asymmetries between masculine and neuter

Overall results from the study suggest that looks to masculine targets were modulated by the gender of the distractor, as were looks to neuter targets. Specifically, masculine gender agreement morphemes appeared to be less reliable cues to the upcoming noun when the distractor was neuter than when the distractor was feminine, and analogously for neuter gender morphemes. In this subsection, these asymmetries will be unpacked to determine whether acoustic or perceptual properties of the morphemes – which in the case of each gender are, in the nominative singular, realized as a vowel – may be driving these effects.[11]

Polish has six oral vowels (Jassem, 2003), represented schematically in the vowel space in Fig. 3. The vowels that occur in this study as gender agreement suffixes on the relevant adjectives are /a/ (fem. *-a*), /ɛ/ (neut. *-e*), and /i/ and /ɨ/ (masc. *-i* and *-y*, respectively) (cf. Table 4). While studies of perceptual distance between these vowels

[11] Thank you to an anonymous reviewer and the editor for raising this consideration.

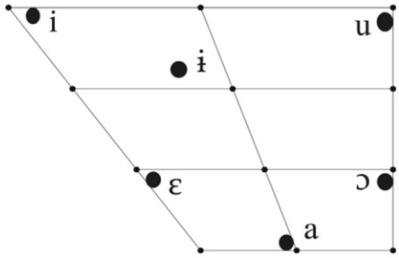

Fig. 3 Polish vowel system (Rojczyk, 2019, p. 83)

Table 9 Phonological realization of inflectional gender suffixes on the three color adjectives used in the study

		M	F	N
czerwon-	red	-/ɨ/	-/a/	-/ɛ/
zielon-	green	-/ɨ/	-/a/	-/ɛ/
niebiesk-	blue	-/i/	-/a/	-/ɛ/

in Polish have – to this author's knowledge – not been pursued to date, production studies that investigate acoustic distance between these vowels suggest that, based on F1 and F2 values, /a/ (fem.) may be acoustically further from /ɛ/ (neut.) than /ɛ/ (neut.) is from /ɨ/ (masc.) (Jassem et al., 1976; Rojczyk, 2019). Although Polish vowels do not undergo reduction (Crosswhite, 2001; Rojczyk, 2019), there are some differences between stressed and unstressed vowels in terms of formant frequencies, particularly for /a/ (Rojczyk, 2019). Nevertheless, the possible asymmetry in acoustic proximity of /ɛ/ (neut.) to /ɨ/ (masc.) persists even in unstressed vowels.

The question at hand is whether this may play a role in what was observed in the present study, namely why looks to neuter targets were slowed in the presence of masculine (but not feminine) distractors and why looks to masculine targets were slowed in the presence of neuter (but again not feminine) distractors.

An experimental manipulation initially introduced for reasons incidental to the crucial design factors, as discussed in Sect. 3.1, allows us to probe this question. Recall from Table 4 that there were three colors adjectives used as hosts for the gender cue. Standard phonological processes in Polish yield different suffixal vowels on *niebiesk-i/-a/-ie* 'blue' as compared to the other two color adjectives, as schematized in Table 9.

This motivates a posthoc analysis wherein observations from trials in which the gender cue was on the adjective *niebiesk-i/-a/-ie* 'blue' are analyzed separately from trials in which the gender cue was on the other two adjectives. If acoustic proximity or perceptual distance between morphemes realizing gender agreement for masculine versus neuter is at play in the asymmetries reported in Sect. 4.2 & 4.3, then we might expect these asymmetries to be absent in the *niebiesk-i/-a/-ie* 'blue' data. This is in fact what the results show (Tables B.1–B.3 in Appendix B). When a masculine target is cued by *niebieski* 'blue', there is no effect of the distractor; the same is true for a neuter target. Conversely, when a masculine target is cued by *czerwony* 'red' or *zielony* 'green', there is an effect of the distractor, and similarly for neuter targets cued by agreement on these color adjectives (see analysis and results in Tables C.1–C.3 in Appendix C).

This raises questions regarding the reliability of some Polish gender morphology as a cue to grammatical gender. Polish vowels are typically considered to be "transparent". A contrast is often drawn within Slavic between languages like Russian, in which vowels in unstressed position are reduced, and languages like Polish, in which such reduction is relatively small if at all present (Crosswhite, 2001; Rojczyk, 2019). The logic is that, given the lack of vowel reduction in Polish, morphemes indexing an adjective's agreement in gender with a head noun that are realized as a vowel should be devoid of ambiguity. This has implications for a number of other subfields, such as language acquisition; for instance, Janssen (2016) directly compared the acquisition of gender morphology between Polish-speaking and Russian-speaking children, finding that Polish-speaking children are faster to learn certain gender distinctions.

However, the results presented here suggest that the reliability of a gender agreement cue may be reduced depending on properties of the competing noun being considered. Take the case of the adjective *zielon-y* 'green-M' (cf. Appendix C). When a masculine target noun is cued by this adjective in the presence of a feminine distractor, participants fixate on masculine targets faster than in a match condition. But when cued by this same adjective in the presence of a neuter distractor, participants are significantly slower to do so. This suggests that the cue on *zielon-y* 'green-M' is reliable enough in the presence of a feminine distractor, which would have been preceded by *zielon-a* 'green-F' instead, but not reliable in the presence of a neuter distractor, which would have been preceded by *zielon-e* 'green-N.'

Further work is needed to determine whether the slowness of these mismatch conditions (masculine target, neuter distractor; and vice versa) reflects the lack of activation of any gender node in the mental lexicon upon encountering the relevant morpheme in the input (i.e. no anticipatory processing) or rather activating both masculine and neuter gender nodes (i.e. anticipatory processing of two genders). In the former scenario, upon encountering the relevant ambiguous morpheme, participants do not have enough information to activate any gender node, so there is no pre-activation of lexical items and thus no anticipation. In the latter scenario, participants perceive the morpheme as consistent with both masculine and neuter gender morphology, activating both nodes and thus pre-activating lexical items connected to both nodes. The present results cannot adjudicate between these two possibilities. Instead, future work would need to consider whether there are any anticipatory looks prompted by an adjective *zielon-y* 'green-M' when there are three items in the display – masculine, feminine, and neuter: do participants show a preference for both masculine and neuter, to the exclusion of feminine (anticipatory processing of two genders), or do they not show any preferential looks until the onset of the noun (no anticipatory processing)?

If it is indeed the case that the asymmetries reported above are driven by perceptual similarity between vowels, further work ought to investigate what distances in the vowel space are small enough to lead to the kind of processing patterns observed here – possibly in the vein of work that uses discrimination tasks to estimate perceptual distance. This would have implications for questions of gender processing and gender acquisition in Polish. Agreement morphology is thought to be central to the acquisition of gender, given that the morphophonology of the noun itself can often

be opaque or even misleading in terms of cues to gender category (see, for example, discussion of the role of gender-inflected articles in the acquisition of gender in Spanish, ex. Lleó (1998) and Mariscal (2009) and citations therein). Thus, a deeper understanding of the reliability of cues as they are realized in the speech stream that serves as the input to the child's grammar may provide a more nuanced understanding of the acquisition of the Polish gender system.

While the asymmetries observed in the present study leave open several questions, as described in this section, the crucial takeaway from the results is that Polish speakers do indeed access gender information on agreeing adjectives in real time, but that the ability to do so may be reduced as a function of the gender of another candidate noun. Specifically, the possibility considered here is that the reliability of the morpheme as a cue may be dependent on its phonological similarity to the cue that would prompt the distractor. In light of this, research on the processing or acquisition of Polish gender may benefit from further investigation of perceptual distance between vowels involved in the realization of agreement in gender and in other agreement features in Polish.

5.3 A lack of asymmetry: is feminine agreement morphology special?

In the results detailed in the main analysis (Sect. 4) as well as the post hoc analysis (Appendix B and Appendix C), the feminine appears to stand on its own: it is not involved in the asymmetries discussed above. As far as the data suggests, the processing of a feminine gender morpheme is independent of the gender of the distractor, and the presence of feminine distractors does not affect the processing of masculine or neuter morphemes.

Under one approach, this special status of the feminine is reducible to the same mechanism that drives the results above, namely the possible perceptual distance between vowels. Under such an account, the word-final unstressed /a/ that realizes feminine gender agreement in the nominative singular is always "far enough" from neuter /ɛ/ and masculine /ɨ/ to be perceptually unambiguous, leading to the lack of asymmetries in feature activation and subsequently word recognition. To reiterate, perceptual distance for Polish vowels has, to this author's knowledge, not been explored, although acoustic distance (per Fig. 3) could be consistent with such an interpretation of the results.

Under an alternate account, the feminine is set apart from the masculine and neuter in morphosyntactic terms – a notion for which there is some converging evidence from other domains. From the perspective of formal linguistic work, in the inflectional paradigms illustrated in Table 2, whereas the masculine and neuter demonstrate syncretism in four of six syntactic cases, the feminine gender suffix is distinct from the others in all six cases. Further evidence comes from language acquisition: while Polish children acquire initial gender distinctions by around age 2;0 (Smoczyńska, 1985), studies of children's production of inflectional endings suggest that the distinction between feminine and masculine is acquired before the neuter, although these studies focus on inflectional endings on nouns rather than on agreeing elements (Dąbrowska, 2006). A study by Janssen (2016) found that though monolingual Polish-speaking children (ages 3;8-6;6) rarely produce agreement errors on

adjectives, when they do, the most common error is to use neuter agreement where masculine agreement would be expected.

In light of this evidence from other domains, the special status of feminine in the present study constitutes additional evidence that the feminine category in the Polish gender feature may be representationally distinct from the masculine and the neuter. In spirit, this would be similar to Halle and Matushansky's (2006) take on Russian gender, in which they formalize this distinction as the feminine gender category being represented as [+FEMININE], while the masculine and neuter are grouped together featurally as [-FEMININE], with masculine further specified as [+MASCULINE]. Evidence from the present study as well as from the discussion above is tentatively compatible with such a representation of the internal makeup of the Polish gender feature, while traditional approaches to Polish gender assume an equal representation of each of the three global gender categories (ex. Corbett, 1983).

Given the confounding role of phonetic distinctness in explaining the asymmetries presented in the study above, the results cannot conclusively argue for an interpretation that rests strictly in the morphosyntactic representation of Polish gender category features as split into feminine and non-feminine, with a subordinate distinction between masculine and neuter. However, there are conceptual reasons to consider the hierarchical analysis of Polish gender, and the results from the present study contribute to converging evidence from various domains – highlighted above – that such an analysis should be explored. If such a proposal is on the right track, reflexes of this mental representation should be evident in experimental paradigms that investigate language processing while minimizing the effect of phonological perception. One particularly promising avenue for further research might be in the processing of agreement and/or agreement errors in a self-paced reading paradigm. Given that agreement morphemes in Polish are equally morphologically specified and – with careful selection of adjective lexical items – of equal length, one could expect agreement errors to be more surprising (leading to longer reading times) when a feminine adjective agrees with a masculine or neuter noun, than when a masculine adjective agrees with a neuter noun or vice versa.

6 Conclusion

An eye-tracking study in the Visual World Paradigm measured whether adult speakers of Polish can use gender information on pre-nominal adjectives to facilitate the lexical retrieval of the subsequent noun during online processing. The research questions focused on language-internal variation in this kind of anticipatory processing, specifically whether the reliability of a gender cue is dependent on the gender cue itself and/or on properties of the competing candidate noun.

The study found that participants were able to access gender information on adjectives in real time for all three Polish gender values. That this is true for gender agreement in the neuter is particularly notable given that default gender agreement in Polish is realized as neuter, and the particular agreement morpheme (*-e*) for the neuter

in nominative singular is also syncretic with agreement morphology for the nominative plural in all genders. In other words, in Polish, *-e* cues not only neuter in the singular but also nouns of any gender in the plural and gender-less nominal elements. Results for this study suggest that participants nevertheless perceived it as a reliable cue to neuter, in line with results for German *die* (Hopp, 2013), which cues not only singular feminine nouns but also plural nouns in all genders. In this study, this effect could be an outcome of participants integrating singular agreement from the earlier verbal element in the auditory prompt, or it could be the product of their experience in the study, in which all target items were in the singular – other work has shown that even recent experience can affect perception of a gender cue as (un)reliable (Brothers et al., 2019; Heyselaar et al., 2020; Hopp, 2016).

Analysis of the results with respect to whether effects were modulated by the gender of the distractor showed that access to feminine target items proceeded independently of the gender of the distractor. However, access to masculine target items and neuter target items was modulated by the gender of the distractor: looks to a masculine target item were slower in the presence of a neuter distractor (relative to a feminine distractor), and looks to a neuter target item were slower in the presence of a masculine distractor (but proceeded as expected in the presence of a feminine distractor). Post hoc analyses revealed that this was driven by conditions in which the masculine agreement morpheme was realized as /ɨ/ (*-y*), which is acoustically and possibly perceptually closer to the neuter agreement morpheme realized as /ɛ/ (*-e*) than is the masculine agreement morpheme realized as /i/ (*-i*). For trials in which masculine agreement on the adjective was realized as the latter, looks to masculine or neuter targets in mismatch conditions were not modulated by the gender of the distractor. The effects of phonological proximity between vowels on processing of gender agreement in Polish may be unexpected given that Polish vowels do not undergo substantive reduction in unstressed position (Rojczyk, 2019) and thus Polish gender agreement morphology is typically considered to be transparent for gender (cf. Janssen, 2016). These results therefore leave open questions with respect to perceptual distance between Polish vowels and how this impacts the acquisition and/or processing of Polish gender agreement. Further work should explore this aspect of Polish agreement morphemes, as well as whether the stand-alone status of the feminine holds up in methodologies that eliminate a reliance on phonological properties of inflectional suffixes.

Appendix A: Experimental items

This section provides the list of experimental nouns used in the study, as described in Sect. 3.1.1. Items are listed here with their corresponding gender in Polish and with their English translation, along with their frequencies in parts-per-million (ppm) in the Frequency Dictionary Corpus subpart (0.5M segments) of the National Corpus of Polish (*Narodowy Korpus Języka Polskiego*, NKJP). The numbers here combine frequencies for diminutive and non-diminutive forms of the words, since the corpus considers them to belong to the same lemmas.

Table A.1 List of experimental nouns used in the present study

Polish	Gender	English	NKJP-FDC(ppm)
koło	neut.	wheel	266
pudło	neut.	box	12
mydło	neut.	soap	18
wiadro	neut.	bucket	6
lustro	neut.	mirror	16
drzewo	neut.	tree	120
jabłko	neut.	apple	12
krzesło	neut.	chair	30
łóżko	neut.	bed	60
gniazdo	neut.	nest	28
pióro	neut.	feather	22
jajko	neut.	egg	64
rower	masc.	bike	14
grzebień	masc.	comb	0
namiot	masc.	tent	10
pasek	masc.	belt	38
zegar	masc.	clock	22
łańcuch	masc.	chain	18
talerz	masc.	plate	24
widelec	masc.	fork	2
dzbanek	masc.	kettle	4
szalik	masc.	scarf	4
młotek	masc.	hammer	10
samolot	masc.	plane	118
strzała	fem.	arrow	14
koszula	fem.	shirt	6
linijka	fem.	ruler	4
ręka	fem.	arm	200
świeczka	fem.	candle	8
książka	fem.	book	173
łopata	fem.	shovel	10
drabina	fem.	ladder	2
huśtawka	fem.	swing	2
truskawka	fem.	strawberry	4
lalka	fem.	doll	8
sukienka	fem.	dress	4

Appendix B

In this appendix, I present the results of a post hoc analysis, similar to the main analysis presented in Sect. 4.1-4.3, but applied only to the subset of the data in which

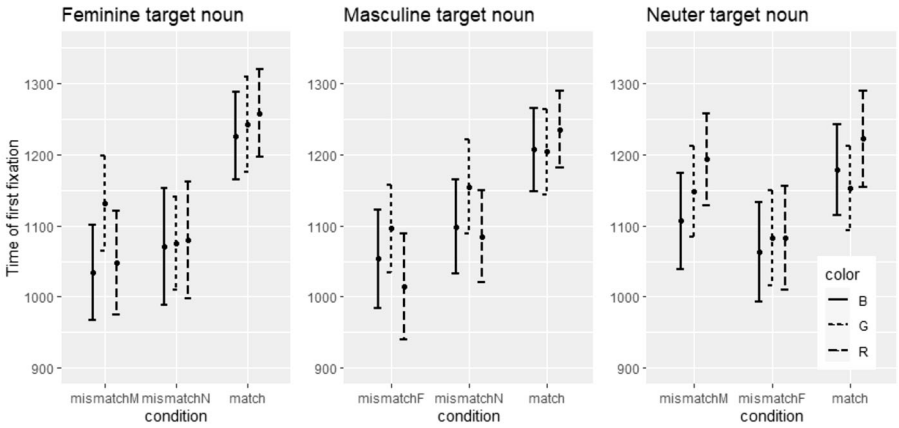

Fig. B.1 Experimental results for each target gender, split by color adjective whose suffix was the gender agreement morpheme

Table B.1 Fixed effects for linear mixed effects model fit to subset with only blue-color conditions, predicting time of first fixation on feminine target items by condition and trial number

Note: * p < 0.05, ** p < 0.01

Model: Time ~ Condition + Trial + (1 + Condition | Participant)

	Time of first fixation		
Condition-distractor	18.37	(21.63)	
Condition-match	56.83	(15.26)	**
Trial	−1.31	(0.53)	*
Constant	1121.84	(43.01)	**
Observations		236	
Log Likelihood		−1641.32	
Akaike Inf. Crit.		3298.64	
Bayesian Inf. Crit.		3326.35	

the gender cue was a suffix on the adjective *niebiesk-il-a/-ie* 'blue', as motivated in Sect. 5.2.

This subset of the data contains 735 observations. The structure of the models for each target gender condition is the same as in the corresponding subsection, referenced each time. It should be noted that these models have less power to detect effects, so null results should be interpreted with caution. A visual representation of the results split by color (also relevant for Appendix C) appears in Fig. B.1.

For feminine target nouns, the model found no significant effect of CONDITION-DISTRACTOR, but did find a significant effect of CONDITION-MATCH (β = 56.83, SE = 15.26, p = 0.001), indicating looks to targets were faster in mismatch condition trials (mean = 1051 ms, sd = 329 ms) than in match condition trials (mean = 1243 ms, sd = 235 ms). The model also found a significant effect of TRIAL ($\beta = -1.31$, SE = 0.54, p = 0.015).

For masculine target nouns, the model did not find a significant effect of CONDITION-DISTRACTOR. There was a significant effect of CONDITION-MATCH (β = 42.37, SE = 12.03, p = 0.0017), indicating that looks were faster to targets in mismatch condition trials (mean = 1075 ms, sd = 307 ms) than in match

Table B.2 Fixed effects for linear mixed effects model fit to subset with only blue-color conditions, predicting time of first fixation on masculine target items by condition and trial number

Note: * p < 0.05, ** p < 0.01

Model: Time ~ Condition + Trial + (1 + Condition | Participant)

	Time of first fixation		
Condition-distractor	21.40	(18.26)	
Condition-match	42.37	(12.02)	**
Trial	−1.11	(0.46)	*
Constant	1120.83	(44.16)	**
Observations		254	
Log Likelihood		−1741.51	
Akaike Inf. Crit.		3499.01	
Bayesian Inf. Crit.		3527.31	

Table B.3 Fixed effects for linear mixed effects model fit to subset with only blue-color conditions, predicting time of first fixation on neuter target items by condition and trial number

Note: * p < 0.05, ** p < 0.01

Model: Time ~ Condition + Trial + (1 | Participant)

	Time of first fixation		
Condition-distractor	23.21	(17.58)	
Condition-match	31.70	(10.23)	**
Trial	−1.58	(0.49)	**
Constant	1123.31	(45.39)	**
Observations		245	
Log Likelihood		−1686.46	
Akaike Inf. Crit.		3384.91	
Bayesian Inf. Crit.		3405.92	

condition trials (mean = 1206 ms, sd = 281 ms). The model also found a significant effect of TRIAL ($\beta = -1.11$, SE = 0.46, p = 0.017).

For neuter target items, the model[12] did not find a significant effect of CONDITION-DISTRACTOR. The model did find a significant effect of CONDITION-MATCH ($\beta = 31.52$, SE = 10.45, p = 0.005), indicating that participants were overall slower on match trials (mean = 1178 ms, sd = 286 ms) than on mismatch trials (mean = 1085 ms, sd = 312 ms). The model also found a significant effect of TRIAL ($\beta = -1.58$, SE = 0.49, p = 0.0016).

To synthesize this part of the post hoc analysis, results suggest that for trials in which the gender marking was on the adjective *niebiesk-i/-a/-ie* 'blue', looks to targets in mismatch condition trials are not modulated by the gender of the distractor.

Appendix C

This Appendix presents the post hoc analysis for the subset of observations for which the gender cue was hosted on the color adjective *czerwon-y/-a/-e* 'red' or *zielon-y/-a/-e* 'green'. This subset of the data contains 1513 observations. A visual representation of the results is in Fig. B.1.

[12]This model differs from the model fit to all data in Section 4.3 in its random effect structure. A model with a random slope for Condition grouped by Participant had a singular fit.

Table C.1 Fixed effects for linear mixed effects model fit to subset with green- and red-color conditions, predicting time of first fixation on feminine target items by condition and trial number

	Time of first fixation		
Condition-distractor	−9.01	(14.48)	
Condition-match	55.76	(8.38)	**
Trial	−1.23	(0.39)	**
Constant	1131.15	(39.60)	**
Observations		495	
Log Likelihood		−3475.06	
Akaike Inf. Crit.		6962.13	
Bayesian Inf. Crit.		6987.36	

Note: * p < 0.05, ** p < 0.01

Model: Time ∼ Condition + Trial + (1 | Participant)

Table C.2 Fixed effects for linear mixed effects model fit to subset with green- and red-color conditions, predicting time of first fixation on masculine target items by condition and trial number

	Time of first fixation		
Condition-distractor	34.37	(14.52)	*
Condition-match	42.03	(10.77)	**
Trial	−1.55	(0.32)	**
Constant	1130.21	(39.90)	**
Observations		515	
Log Likelihood		−3535.64	
Akaike Inf. Crit.		7087.28	
Bayesian Inf. Crit.		7121.23	

Note: * p < 0.05, ** p < 0.01

Model: Time ∼ Condition + Trial + (1 + Condition | Participant)

Table C.3 Fixed effects for linear mixed effects model fit to subset with only blue-color conditions, predicting time of first fixation on neuter target items by condition and trial number

	Time of first fixation		
Condition-distractor	44.72	(12.31)	**
Condition-match	19.32	(7.54)	*
Trial	−1.24	(0.32)	**
Constant	1132.97	(45.28)	**
Observations		503	
Log Likelihood		−3449.67	
Akaike Inf. Crit.		6915.34	
Bayesian Inf. Crit.		6949.10	

Note: * p < 0.05, ** p < 0.01

Model: Time ∼ Condition + Trial + (1 + Condition | Participant)

For feminine targets in conditions in which the color adjective was green or red, the model found a significant effect of CONDITION-MATCH ($\beta = 55.76$, SE = 8.38, $p < 0.001$) but no significant effect of CONDITION-DISTRACTOR. Additionally, the model found a significant effect of TRIAL ($\beta = -1.23$, SE = 0.39, $p = 0.002$).

For masculine targets in conditions in which the color adjective was green or red, the model found a significant effect of CONDITION-MATCH ($\beta = 42.03$, SE = 10.77, $p < 0.001$) as well as a significant effect of CONDITION-DISTRACTOR ($\beta = 34.37$, SE = 14.52, $p = 0.02$). The model also found a significant effect of TRIAL ($\beta = -1.55$, SE = 0.32, $p < 0.001$).

For conditions with neuter target items in which the color adjective was green or red, the model found a significant effect of CONDITION-MATCH ($\beta = 19.32$, SE $= 7.54$, p $= 0.02$) as well as a significant effect of CONDITION-DISTRACTOR ($\beta = 44.72$, SE $= 12.31$, p < 0.001). The model also found a significant effect of TRIAL ($\beta = -1.24$, SE $= 0.32$, p < 0.001).

To synthesize this part of the post hoc analysis, results suggest that for trials in which the gender marking was on the adjective *czerwon-y/-a/-e* 'red' or *zielon-y/-a/-e* 'green', looks to the target in mismatch condition trials are modulated by the gender of the distractor.

Acknowledgements This work was supported in part by a Graduate Research Grant from the Institute for Quantitative Social Sciences at Harvard University, and logistical support was provided by the Harvard Center for Brain Sciences. For very helpful feedback and discussion I am grateful to Maria Polinsky, Jonathan Bobaljik, Holger Hopp, Kevin Ryan, and Irina Sekerina, as well as to the audience at the 94th Annual Meeting of the Linguistic Society of America. All errors are my own.

Funding Open access funding provided by SCELC, Statewide California Electronic Library Consortium. Partial financial support was received from the Institute for Quantitative Social Sciences at Harvard University.

Data Availability Anonymous data is publicly available on OSF at https://osf.io/ue6n8/?view_only=786fb76f78ca4d0881e9c0f2af0b937a.

Code Availability R code used for data analysis is publicly available on OSF at https://osf.io/ue6n8/?view_only=786fb76f78ca4d0881e9c0f2af0b937a.

Declarations

Competing Interests The author has no conflicts of interest or competing interests to declare that are relevant to the content of this article.

References

Adamson, L., & Šereikaite, M. (2019). Gender representation and defaults in Lithuanian. *Glossa*, *4*(1). https://doi.org/10.5334/GJGL.860.

Alemán Bañón, J., & Rothman, J. (2016). The role of morphological markedness in the processing of number and gender agreement in Spanish: An event-related potential investigation. *Language, Cognition and Neuroscience*, *31*(10), 1273–1298. https://doi.org/10.1080/23273798.2016.1218032.

Aumeistere, A., Bultena, S., & Brouwer, S. (2022). Wisdom comes with age? The role of grammatical gender in predictive processing in Russian children and adults. *Appl. Psycholinguist.*, *43*(4), 867–887. https://doi.org/10.1017/S0142716422000170.

Bobaljik, J. D., & Zocca, C. L. (2011). Gender markedness: The anatomy of a counter-example. *Morphology*, *21*(2), 141–166. https://doi.org/10.1007/s11525-010-9156-3.

Bogusławski, A. (1986). Accusative singular operations on Polish nouns. In R. D. Brecht & J. S. Levine (Eds.), *Case in Slavic. Studies dedicated to the memory of Roman Jakobson* (pp. 152–169). Columbus: Slavica Publishers.

Brehmer, B., & Rothweiler, M. (2012). The acquisition of gender agreement marking in Polish. In K. Braunmuller & C. Gabriel (Eds.), *Multilingual individuals and multilingual societies* (pp. 81–100). Amsterdam: Benjamins.

Brothers, T., Dave, S., Hoversten, L. J., Traxler, M. J., & Swaab, T. Y. (2019). Flexible predictions during listening comprehension: Speaker reliability affects anticipatory processes. *Neuropsychologia*, *135*, 107225. https://doi.org/10.1016/j.neuropsychologia.2019.107225.

Brouwer, S., Sprenger, S., & Unsworth, S. (2017). Processing grammatical gender in Dutch: Evidence from eye movements. *Journal of Experimental Child Psychology*, *159*, 50–65. https://doi.org/10.1016/j.jecp.2017.01.007.

Cholewa, J., Neitzel, I., Bürsgens, A., & Günther, T. (2019). Online-processing of grammatical gender in noun-phrase decoding: An eye-tracking study with monolingual German 3rd and 4th graders. *Frontiers in Psychology*, *10*, 2586. https://doi.org/10.3389/fpsyg.2019.02586.

Corbett, G. (1983) The number of genders in Polish. *Papers and studies in contrastive linguistics*, *16*, 83–89.

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

Corbett, G., & Fraser, N. (1999). Default genders. In B. Unterbeck & M. Rissanen (Eds.), *Gender in grammar and cognition* (pp. 55–97). Berlin: de Gruyter.

Crosswhite, K. (2001). *Vowel reduction in optimality theory*. New York: Routledge.

Dąbrowska, E. (2006). Low-level schemas or general rules? The role of diminutives in the acquisition of Polish case inflections. *Language Sciences*, *28*, 120–135. https://doi.org/10.1016/j.langsci.2005.02.005.

Dussias, P., Valdés Kroff, J. R., Guzzardo Tamargo, R. E., & Gerfen, C. (2013). When gender and looking go hand in hand. *Studies in Second Language Acquisition*, *35*(2), 353–387. https://doi.org/10.1017/S0272263112000915.

Fuchs, Z. (2014). Gender and analogical extension: From animacy to borrowings in Polish. *New Insights into Slavic Linguistics*, *3*, 115–127.

Fuchs, Z. (2021). Facilitative use of grammatical gender in Heritage Spanish. *Linguistic Approaches to Bilingualism*. https://doi.org/10.1075/lab.20024.fuc.

Fuchs, Z. (2022). Eyetracking evidence for heritage speakers' access to abstract syntactic agreement features in real-time processing. In F. Bayram, M. Kubota, & S. Soares (Eds.), *The next phase in heritage language studies: Methodological considerations and advancements [special issue]. Frontiers of Psychology* (p. 13). https://doi.org/10.3389/fpsyg.2022.960376.

Fuchs, Z., Polinsky, M., & Scontras, G. (2015). The differential representation of number and gender in Spanish. *The Linguistic Review*, *32*(4), 703–737. https://doi.org/10.1515/tlr-2015-0008.

Fuchs, Z., & van der Wal, J. (2022). The locus of parametric variation in Bantu gender and nominal derivation. *Linguist. Var.*, *22*(2), 268–324. https://doi.org/10.1075/lv.20007.fuc.

Grüter, T., Lew-Williams, C., & Fernald, A. (2012). Grammatical gender in L2: A production or a real-time processing problem? *Second Lang. Res.*, *28*(2), 191–215. https://doi.org/10.1177/0267658312437990.

Halle, M., & Matushansky, O. (2006). The morphophonology of Russian adjectival inflection. *Linguistic Inquiry*, *37*(3), 351–404.

Harley, H., & Ritter, E. (2002). Person and number in pronouns: A feature-geometric analysis. *Language*, *78*(3), 482–526.

Harris, J. W. (1991). The exponence of gender in Spanish. *Linguistic Inquiry*, *22*(1), 27–62.

Haspelmath, M. (2006). Against markedness (and what to replace it with). *Journal of Linguistics*, *42*(1), 25–70.

Heyselaar, E., Peeters, D., & Hagoort, P. (2020). Do we predict upcoming speech content in naturalistic environments? *Language, Cognition and Neuroscience*, *36*(4), 1–22. https://doi.org/10.1080/23273798.2020.1859568.

Hockett, C. (1958). *A course in modern linguistics*. New York: Macmillan.

Hopp, H. (2013). Grammatical gender in adult L2 acquisition: Relations between lexical and syntactic variability. *Second Lang. Res.*, *29*(1), 33–56. https://doi.org/10.1177/0267658312461803.

Hopp, H. (2016). Learning (not) to predict: Grammatical gender processing in second language acquisition. *Second Lang. Res.*, *32*(2), 277–307. https://doi.org/10.1177/0267658315624960.

Hopp, H., & Lemmerth, N. (2016). Lexical and syntactic congruency in L2 predictive gender processing. *Studies in Second Language Acquisition*, *40*(1), 171–199. https://doi.org/10.1017/S0272263116000437.

Huettig, F., & Guerra, E. (2019). Effects of speech rate, preview time of visual context, and participant instructions reveal strong limits on prediction in language processing. *Brain Research*, *1706*, 196–208. https://doi.org/10.1016/j.brainres.2018.11.013.

Huettig, F., & Janse, E. (2016). Individual differences in working memory and processing speed predict anticipatory spoken language processing in the visual world. *Language, Cognition and Neuroscience*, *31*(1), 80–93. https://doi.org/10.1080/23273798.2015.1047459.

Ito, A., Corley, M., & Pickering, M. J. (2018). A cognitive load delays predictive eye movements similarly during L1 and L2 comprehension. *Bilingualism: Language and Cognition*, *21*(2), 251–264. https://doi.org/10.1017/S1366728917000050.

Janssen, B. (2016). The acquisition of gender and case in Polish and Russian: A study of monolingual and bilingual children. *Uitgeverij Pegasus*. https://dare.uva.nl/search?identifier=cb49f518-8561-4268-8535-fc72e4439f27.

Jassem, W. (2003). Illustration of the IPA: Polish. *J. Int. Phon. Assoc.*, *33*, 103–107.

Jassem, W., Krzyśko, W., & Dyczkowski, A. (1976). Identification of isolated Polish vowels. *Speech Analysis and Synthesis*, *4*, 106–133.

Johnson, E. (2005). Grammatical gender and early word recognition. In E. A. Dutch Brugos, M. R. Clark Cotton, & S. Ha (Eds.), *Proceedings of the 29th annual Boston University Conference on Language Development* (Vol. 1, pp. 320–330). Somerset, MA: Cascadilla Press.

Kaan, E., & Grüter, T. (2021). Prediction in second language processing and learning. In E. Kaan & T. Grüter (Eds.), *Prediction in second language processing and learning* (pp. 1–24). Amsterdam: Benjamins.

Kaushanskaya, M., Blumenfeld, H. K., & Marian, V. (2020). The language experience and proficiency questionnaire (LEAP-Q): Ten years later. *Bilingualism*, *23*(5), 945–950. https://doi.org/10.1017/S1366728919000038.

Kosta, P. (2003). The new animacy category in slavic languages: Open questions of syntax, semantics and morphology. *Germanoslavica: Zeitschrift für germano-slawische Studien*, *9*, 179–198.

Kramer, R. (2009). *Definite Markers, Phi-features, and Agreement: A Morphosyntactic Investigation of the Amharic DP*. PhD thesis, University of California, Santa Cruz.

Kramer, R. (2014). Gender in Amharic: A morphosyntactic approach to natural and grammatical gender. *Language Sciences*, *43*, 102–115. https://doi.org/10.1016/J.LANGSCI.2013.10.004.

Kramer, R. (2015). *The morphosyntax of gender*. London: Oxford University Press. https://doi.org/10.1093/acprof:oso/9780199679935.001.0001.

Lemmerth, N., & Hopp, H. (2019). Gender processing in simultaneous and successive bilingual children: Cross-linguistic lexical and syntactic influences. *Language Acquisition*, *26*(1), 21–45. https://doi.org/10.1080/10489223.2017.1391815.

Levelt, W., Roelofs, A., & Meyer, A. (1999). A theory of lexical access in speech production. *Behavioral and Brain Sciences*, *22*, 1–75.

Lew-Williams, C., & Fernald, A. (2007). Young children learning Spanish make rapid use of grammatical gender in spoken word recognition. *Psychological Science*, *18*(3), 193–198.

Lew-Williams, C., & Fernald, A. (2010). Real-time processing of gender-marked articles by native and non-native Spanish speakers. *Journal of Memory and Language*, *63*(4), 447–464. https://doi.org/10.1016/j.jml.2010.07.003.

Lieberman, A., Borovsky, A., & Mayberry, R. (2018). Prediction in a visual language: Real-time sentence processing in American sign language across development. *Language, Cognition & Neuroscience*, *33*, 387–401. https://doi.org/10.1080/23273798.2017.1411961.

Lleó, C. (1998). Proto-articles in the acquisition of Spanish: Interface between phonology and morphology. In R. Fabri, A. Ortmann, & T. Parodi (Eds.), *Modelle der Flexion: 18. Jahrestagung der Deutschen Gesellschaft für Sprachwissenschaft* (pp. 175–195). Tubingen: Niemeyer.

Loerts, H., Wieling, M., & Schmid, M. S. (2013). Neuter is not common in Dutch: Eye movements reveal asymmetrical gender processing. *Journal of Psycholinguistic Research*, *42*, 551–570. https://doi.org/10.1007/s10936-012-9234-2.

Luczynski, E. (2002). Zagadnienie kompetencji fleksyjnej dzieci w wieku przedszkolnym [na przykładzie rzeczownika] (The issue of inflectional competence of preschool children [case of noun]). *Język Polski*, *1*, 43–50.

Lundquist, B., & Vangsnes, Ø. (2018). Language separation in bidialectal speakers: Evidence from eye tracking. *Frontiers in Psychology*, *9*. https://doi.org/10.3389/fpsyg.2018.01394.

Marian, V., Blumenfeld, H. K., & Kaushanskaya, M. (2007). The language experience and proficiency questionnaire (LEAP-Q): Assessing language profiles in bilinguals and multilinguals. *Journal of Speech, Language, and Hearing Research*, *50*(4), 940.

Mariscal, S. (2009). Early acquisition of gender agreement in the Spanish noun phrase: Starting small. *Journal of Child Language*, *36*, 143–171. https://doi.org/10.1017/S0305000908008908.

McClelland, J. L., & Rumelhart, D. E. (1981). An interactive activation model of context effects in letter perception: I. An account of basic findings. *Psychological Review*, *88*(5), 375–407. https://doi.org/10.1037/0033-295X.88.5.375.

Mornati, G., Riva, V., Vismara, E., Molteni, M., & Cantiani, C. (2023). Infants aged 12 months use the gender feature in determiners to anticipate upcoming words: An eye-tracking study. *Journal of Child Language*, *50*, 841–859. https://doi.org/10.1017/S030500092200006X.

R Core Team (2021). *R: A language and environment for statistical computing*. Vienna, Austria: R Foundation for Statistical Computing. https://www.R-project.org/.

Roelofs, A. (1992). A spreading-activation theory of lemma retrieval in speaking. *Cognition*, *42*, 107–142.

Rojczyk, A. (2019). Quality and duration of unstressed vowels in Polish. *Lingua*, *217*, 80–89. https://doi.org/10.1016/j.lingua.2018.10.012.

Rumelhart, D. E., & McClelland, J. L. (1982). An interactive activation model of context effects in letter perception: II. The contextual enhancement effect and some tests and extensions of the model. *Psychological Review*, *89*(1), 60–94. https://doi.org/10.1037/0033-295X.89.1.60.

Schriefers, H., & Jescheniak, J. D. (1999). Representation and processing of grammatical gender in language production: A review. *Journal of Psycholinguistic Research*, *28*, 575–600. https://doi.org/10.3389/fpsyg.2014.00589.

Scontras, G., Polinsky, M., & Fuchs, Z. (2018). In support of representational economy: Agreement in heritage Spanish. *Glossa: A Journal of General Linguistics*, *3*(1), 1. https://doi.org/10.5334/gjgl.164.

Sekerina, I. A. (2015). Predictions, fast and slow. *Linguistic Approaches to Bilingualism*, *5*(4), 532–536. https://doi.org/10.1075/lab.5.4.16sek.

Smoczyńska, M. (1985). The acquisition of Polish. In D. I. Slobin (Ed.), *The crosslinguistic study of language acquisition* (Vol. 1, pp. 595–686). Hillsdale: Erlbaum.

Steriopolo, O., & Wiltschko, M. (2010). The distributed gender hypothesis. In G. Zybatow, P. Dudchuk, S. Minor, & E. Pshehotskaya (Eds.), *Formal Studies in Slavic Linguistics*. (pp. 155–172). Oxford: Peter Lang.

Swan, O. (2015). Polish gender, subgender, and quasi-gender. *Journal of Slavic Linguistics*, *23*(1), 83–122. https://doi.org/10.1353/jsl.2015.0001.

Tanenhaus, M. K., Spivey-Knowlton, M. J., Eberhard, K. M., & Sedivy, J. C. (1995). Integration of visual and linguistic information in spoken language comprehension. *New Series*, *268*(5217).

Wang, M., & Schiller, N. O. (2019). A review on grammatical gender agreement in speech production. *Frontiers in Psychology*. [Online], 92754–2754.

Wechsler, S. (2008). Elsewhere in gender resolution. In K. Hanson & S. Inkelas (Eds.), *The nature of the word: Essays in honor of Paul Kiparsky* (pp. 567–586). Cambridge: MIT Press.

Willim, E. (2012). Concord in Polish coordinate NPs as agree. In M. Ziková & M. Dočekal (Eds.), *Slavic Languages in Formal Grammar: Proceedings of FDSL 8.5, Brno 2010*, Oxford: Peter Lang.