**Research Article**

# Sensitivity to grammatical gender in L1 Polish learners of L2 German: The contribution of semantic, lexical, and structural congruency

Kamil Długosz*
Adam Mickiewicz University, Poznań

Jakub Przybył
Adam Mickiewicz University, Poznań

**Abstract English:** Are adult L2 learners sensitive to grammatical gender during L2 processing? Most previous research has examined learners of L2s like Spanish with L1 English, a language lacking grammatical gender. Less is known about how sensitivity to grammatical gender develops in L2 learners with a gendered L1. This study investigates how L1 Polish learners of L2 German process grammatical gender within the noun phrase, focusing on the role of congruency factors and L2 proficiency. Polish and German both have a three-gender system (masculine, feminine, neuter) but differ in lexical gender congruency. Furthermore, while both languages mark gender on adjectives (structurally congruent), only German marks it on prenominal articles (structurally incongruent).
Two primed lexical decision tasks were conducted. Experiment 1 tested correctly vs. incorrectly gender-marked adjectives and manipulated the semantic congruency of adjective-noun pairs. Experiment 2 tested correctly vs. incorrectly inflected articles and manipulated lexical gender congruency. The results showed clear sensitivity to structurally congruent gender marking (adjectives) but not to incongruent marking (articles), regardless of semantic congruency, lexical gender congruency, or L2 proficiency. These findings suggest that structural congruency with L1 plays a primary role in shaping L2 learners' sensitivity to grammatical gender, while lexical and semantic factors are less influential.

**Keywords:** grammatical gender processing, gender agreement processing, grammatical gender priming, structural congruency, semantic congruency, gender congruency, L2 German, cross-linguistic influence

## 1 Introduction

The question of whether adult L2 learners are able to develop target-like competence in grammatical morphemes and in agreement processes involving these morphemes has been of considerable importance in L2 acquisition research (for an extensive overview, see Jiang, 2018). In the 2000s, researchers have shifted their focus from the accuracy of grammatical-morpheme realisation to language processing measures, such as response times or brain responses, with the aim of exploring L2 learners' sensitivity to morphosyntactic features. A great deal of studies in this area have been concerned with grammatical gender, a morphosyntactic feature known as agreement between the noun and its modifiers, such as

determiners and adjectives (Hockett, 1958, p. 231). A particularly productive approach has been to examine L2 learners' sensitivity to grammatical gender violations, a phenomenon frequently observed among native speakers (e.g., Jiang, 2018). The degree to which late L2 learners are sensitive to grammatical gender marking, as well as the factors influencing this sensitivity, continues to be a topic of debate. Some studies provide evidence of target-like sensitivity to gender marking in L2 learners (e.g., Sagarra & Herschensohn, 2010; Alemán Bañón et al., 2017), although this sensitivity may be weaker compared to that of native speakers (e.g., Garrido-Pozú, 2022). Conversely, other studies report no sensitivity effects (e.g., Sabourin, 2003; Foucart & Frenck-Mestre, 2011).

In the present study, we contribute to this debate by focusing on sensitivity to grammatical gender marking in late L1 Polish L2 German learners through the lens of priming effects. The priming paradigm, such as the primed lexical decision task, offers several advantages for studying grammatical gender processing in L2 learners compared to other methods. These advantages stem from the task's ability to tap into automatic and implicit aspects of linguistic processing. Unlike tasks that explicitly require grammatical judgments (e.g., acceptability judgments), priming tasks prevent access to metalinguistic knowledge, providing a more direct measure of underlying grammatical competence (Jiang, 2013: 98, p. 105).

Previous priming studies with late L2 learners yielded mixed results. While Foote (2014) and Romanowa and Gor (2017) found clear grammatical gender priming effects in late L2 learners, Guillelmon and Grosjean (2001) and Scherag et al. (2004) failed to do so. However, all these studies involved advanced L2 learners whose L1 was English, a non-gendered language. Therefore, it remains unclear how L2 proficiency and the presence of gender in L1 modulate gender priming effects in L2. To rectify this situation, we present priming data from intermediate to advanced learners of L2 German, who are native speakers of Polish, a language that encodes gender. Our study investigates factors identified as critical in prior research on grammatical gender processing: (i) structural congruency, i.e., the similarity in gender marking within the noun phrase between L1 and L2; (ii) semantic congruency, i.e., the degree of semantic relatedness between the noun and its modifier; and (iii) lexical gender congruency, i.e., the overlap in gender assignment between L1 and L2.

The paper is organised as follows: Section 2 gives an overview of previous research on gender agreement processing in L2, focusing on the priming paradigm and congruency factors. Section 3 presents how gender is marked in Polish and German. The methodology is presented in Section 4. Section 5 presents the results, and Section 6 discusses the findings.

## 2 Theoretical background

### 2.1 Gender agreement in adult L2 learners

Mastering gender agreement has been observed to be notoriously difficult in adult L2 acquisition (e.g., Carroll, 1989; Franceschina, 2005). In language production, L2 learners frequently fail to use correct gender-marked determiners and adjectives, even after prolonged language exposure (e.g., Dewaele & Veronique, 2001; Unsworth, 2008; Grüter et al., 2012). It has been argued that non-native language learners are not capable of acquiring morphosyntactic features that differ from those in their L1 (Hawkins & Chan, 1997; Tsimpli & Dimitrakopoulou, 2007). This means that if syntactic features triggering morphological agreement are instantiated in L1, they are accessible to non-native language learners. This view has been challenged by studies showing that late L2 learners have persistent difficulties with gender agreement even if

their L1 grammar represents the gender feature (e.g., Meisel, 2009). Therefore, other accounts have been developed to explain L2 learners' variability with respect to gender. Computational approaches, for example, underscore that non-native grammars are not deficient and associate difficulties with establishing gender agreement with difficulties with the overt realisation of morphology (Prévost & White, 2000), default inflection (White et al., 2004), or variability in lexical gender assignment (e.g., Grüter et al., 2012; Hopp, 2013). Nevertheless, having gender in L1 has been found to benefit L2 learners in terms of gender assignment (e.g., Sabourin et al., 2006) and gender agreement (e.g., Mačiukaité, 2008; Ellis et al., 2012).

Early L2 research on gender, however, particularly relied on accuracy data. To minimise the involvement of participants' metalinguistic knowledge and escape the problem of setting an acquisition criterion, researchers have begun to consider language processing measures. Reaction times and brain responses shed new light on the acquirability of grammatical gender by exploring gender agreement processes in real-time language use. This line of research has shown that native speakers and early L2 learners benefit from gender-marked words in real-time language processing. For example, native speakers of German show shorter response times when making lexical decisions on nouns primed by adjectives that agree vs. disagree in gender with the noun (e.g., Scherag et al., 2004). In what follows, we discuss previous research that used the priming paradigm to investigate gender processing in L2.

## 2.2 Gender priming in adult L2 learners

The evidence regarding sensitivity to gender marking in L2 is inconsistent. In their seminal paper, Guillelmon and Grosjean (2001) report a study of how early and late L1 English L2 French learners with an advanced L2 proficiency react to gender marking in L2 processing. Participants heard noun phrases consisting of a determiner, an adjective, and a noun, and they were asked to repeat the noun. The gender marking on the determiner was either congruent or neutral, or incongruent or neutral. Early L2 learners demonstrated clear facilitation and inhibition effects, similar to their monolingual French counterparts. Late L2 learners, in turn, showed no sensitivity to gender marking, whether congruent or incongruent.

Similar findings were obtained by Scherag et al. (2004), who tested advanced L1 English L2 German learners and German monolinguals, among others, by means of a primed lexical decision task. Auditory primes were German adjectives correctly or incorrectly marked for gender and semantically congruent or incongruent with the target noun. While German monolinguals gained from semantically and morphosyntactically congruent primes, late L2 learners showed no priming effects for morphosyntactically congruent primes when the adjective was not semantically associated with the noun. While Guillelmon and Grosjean (2001) highlight a lack of sensitivity to gender marking in L2 learners, Scherag et al. (2004) demonstrate that L2 learners can develop sensitivity to gender marking, but only when provided with meaningful adjective-noun pairs.

In contrast to these studies, Foote (2014) demonstrated that advanced L1 English L2 Spanish learners were sensitive to gender marking in their L2. In her partial replication of Guillelmon and Grosjean (2001), participants were presented auditorily with noun phrases consisting of a determiner, an adjective, and a noun, and they were asked to repeat the noun. Comparable gender priming effects were observed in L2 learners and native speakers. Similar to Foote (2014), Romanova and Gor (2017) found that advanced L1 English L2 Russian learners performed like native speakers of Russian in a primed lexical decision task. Both groups made lexical decisions faster when the noun was preceded by a correctly inflected adjective than by

an incorrectly inflected adjective. However, the study also included a neutral-prime condition, which enabled the authors to distinguish between facilitation and inhibition effects. While gender priming effects in native speakers consisted of facilitation and inhibition, L2 learners only showed inhibition. Therefore, although the study points to native-like priming effects in L2, the underlying mechanisms seem to differ between native speakers and L2 learners.

To our knowledge, the four studies reviewed above are the only ones that explore sensitivity to gender marking through the lens of gender priming. Although they provide valuable insights into whether advanced L2 learners with an ungendered L1 can capitalise on gender information in language comprehension, they do not consider L2 learners at lower proficiency levels whose L1 encodes gender. The primary aim of our study is to address this research gap.

Considering gender priming in German, the L2 of our participants, priming effects have been observed across all three genders in native speakers (e.g., Hillert & Bates, 1996). Similarly, the eye-tracking study by Hopp (2013) did not find any asymmetries between different genders in predictive gender agreement processing in late L2 learners. Beyond priming, L2 learners generally tend to overgeneralise a default gender in both production and comprehension but this tendency is particularly prevalent among learners with L1 English (e.g., Klassen, 2016).

## 2.3 Structural and lexical congruency in adult L2 gender processing

Contemporary research on L2 gender processing focuses on how L2 learners deploy their grammatical knowledge depending on language-internal factors that characterise the gender systems in L1 and L2. The central question is whether similarities between the two gender systems facilitate L2 gender acquisition and processing, and if so, what types of similarities—such as structural patterns, lexical classifications (e.g., number of gender categories), or morphological forms—promote or hinder L2 grammatical development and processing (e.g. Johannessen et al., 2024).

Starting with lexical similarities, a substantial body of research has demonstrated that the L1 and L2 gender systems interact during L2 processing. Specifically, when the gender of a noun in L2 aligns with the gender of its translation equivalent in L1, the processing of that noun is facilitated. This phenomenon is commonly referred to as the 'gender congruency effect' (e.g., Sá-Leite et al., 2020). The impact of lexical gender congruency on gender agreement processing in L2 has been investigated in a few studies, which, at the same time, focused on structural similarity.

The study by Hopp and Lemmerth (2018) revealed an interaction between lexical and structural congruency in predictive gender processing. Using a visual-world eye-tracking experiment, they compared L1 Russian L2 German learners to German monolinguals, focusing on their ability to predict upcoming nouns based on gender cues encoded on articles (present in German but absent in Russian) or adjectives (present in both languages). Advanced L2 learners exhibited native-like gender prediction across all conditions. In contrast, high-intermediate learners showed asymmetries influenced by structural and lexical congruency. Specifically, predictive gender processing was observed for all nouns in the (structurally congruent) adjective condition, but only for lexically congruent nouns in the (structurally incongruent) article condition. These findings highlight the combined role of syntactic and lexical congruency in facilitating predictive gender processing in L2 learners.

Johannessen et al. (2024) explored predictive gender processing in L2 Norwegian among learners with L1 Greek, Russian, and Turkish. The study focused on the presence vs. absence of the gender feature in L1 (Norwegian, Greek, and Russian have gender, whereas Turkish is a

genderless language), lexical gender congruency, structural congruency in gender marking, and L2 proficiency. Using a visual-world eye-tracking experiment, the study found that high-proficiency learners with L1 Greek and L1 Russian demonstrated stronger predictive processing effects compared to high-proficiency learners with L1 Turkish. Interestingly, neither lexical nor structural congruency influenced the results. This suggests that the presence of grammatical gender in L1 and L2 proficiency are more critical factors in facilitating predictive gender processing than L1-L2 similarities.

The role of structural similarity has been examined in numerous neurolinguistic studies, often focusing on the P600 effect—a distinct type of event-related potential (ERP) observed in EEG research, traditionally associated with grammar processing, including the detection of morphosyntactic violations. Sabourin and Stowe (2008) found different ERP signatures for sentences violating gender agreement in determiner phrases, depending on participants' L1. L1 German L2 Dutch learners showed a P600 effects (albeit weaker and with a later peak compared to native speakers), whereas L2 Dutch learners with a Romance L1 did not. The authors concluded that L2 learners may show sensitivity to gender errors when their L1 and L2 grammars are very similar.

Further evidence showing that similarities in gender marking between L1 and L2 may be a precondition for target-like gender processing comes from the study by Foucart and Frenck-Mestre (2011). Advanced L1 German L2 French learners and French native speakers read sentences while ERPs were recorded. Gender agreement violations in the noun phrase were manipulated in three experiments: (1) between the definite article and the noun, (2) between the postnominal adjective and the noun, and (3) between the prenominal adjective and the noun. A comparable P600 effect was observed for L2 learners and native speakers only when gender was marked similarly in German and French. Even advanced L2 learners were not able to process gender agreement in a target-like manner when gender marking was structurally incongruent between L1 and L2.

Numerous studies have identified L2 proficiency as a key factor influencing sensitivity to gender marking in L2. Research examining the detection of gender agreement violations has demonstrated that as L2 learners advance in their language development, they gradually become more sensitive to gender errors (Sagarra & Herschensohn, 2010; Długosz, 2023). Similarly, visual-world eye-tracking studies focusing on predictive processing have revealed that advanced L2 learners, but not intermediate ones, are capable of using gender cues predictively (Lew-Williams & Fernald, 2010; Grüter et al., 2012; Dussias et al., 2013).

In summary, while native speakers exhibit a high degree of sensitivity to gender marking, most studies indicate that L2 learners' gender processing depends on various factors, including the presence or absence of gender in their L1, structural similarities in gender marking between L1 and L2, lexical gender congruency, and L2 proficiency. However, contradictory findings—particularly regarding the role of structural similarities—highlight the need for further research to determine whether gender marking in L1 constrains gender processing in L2.

## 3 Linguistic background

German has three gender values: masculine, feminine, and neuter. Most nouns are feminine (46%), followed by masculine (34%), and neuter (20%) (Duden, 2023). Gender assignment is assumed to be arbitrary to a large extent, with noun-internal cues providing only limited predictability (e.g., Kupisch et al., 2022). Most of these cues are probabilistic, while only a few are deterministic (e.g., Hohlfeld, 2006). Gender is marked on different types of determiners,

including prenominal definite and indefinite articles, and adjectives. The following examples are provided in the nominative case:

(1)  *ein*           *klein-er*        *Tisch*
     aSG.M          small-SG.M       tableSG.M
     'a small table'
(2)  *ein*           *klein-es*        *Bier*
     aSG.N          small-SG.N       beerSG.N
     'a small beer'
(3)  *eine*          *klein-e*         *Lampe*
     aSG.F          small-SG.F       lampSG.F
     'a small lamp'

The examples show that indefinite articles do not disambiguate between masculine and neuter, whereas adjectival suffixes do. Nouns with a definite reading are marked overtly for definiteness by a prenominal free determiner, which clearly indicates gender in the nominative case ('derM', 'dieF', 'dasN').

Similar to German, Polish distinguishes between masculine, feminine, and neuter, but standard Polish grammar forms (e.g., Grzegorczykowa et al., 1999) tend to split masculine gender into three classes depending on animacy and virility in the accusative case (masculine virile, masculine animate/non-virile, masculine inanimate). According to Stefańczyk (2007, p. 48), masculine includes approximately 50% of all Polish nouns and is thus the most frequent gender category, followed by feminine with a frequency of approximately 40%. Neuter is least frequent in Polish as it only covers 10% of nouns. Gender assignment is highly predictable by the morphological shape of the noun in the nominative singular. Nouns ending in a consonant are typically masculine, nouns ending in '-a' and '-(o)ść' are feminine, and nouns ending in '-o', 'e/ę', and '-um' are neuter (Stefańczyk, 2007).

Gender is marked within the noun phrase. Polish has no articles, but gender is marked on demonstratives and prenominal attributive adjectives, among others. They clearly disambiguate between masculine, feminine, and neuter:

(4)  *mał-y*         *stół*
     small-SG.M     tableSG.M
     'small table'
(5)  *mał-e*         *piwo*
     small-SG.N     beerSG.N
     'small beer'
(6)  *mał-a*         *lampa*
     small-SG.F     lampSG.F
     'small lamp'

Thus, a structural parallel between German and Polish lies in the fact that gender-marked adjectives in the nominative case clearly disambiguate gender and appear prenominally in both languages. However, unlike German, which marks gender on prenominal articles, Polish does not, as it is an articleless language.

# 4 The study

## 4.1 Research questions

To explore sensitivity to grammatical gender marking in late L2 learners, we devised two primed lexical decision tasks. These tasks also allowed us to complement previous findings concerning L2 learners with non-gendered L1s by testing late L1 Polish L2 German learners at different proficiency levels. Sensitivity to grammatical gender marking was operationalised as a slowdown effect in response to a gender agreement violation, reflecting the process of accessing gender information from the preceding noun modifier. We posed the following research questions:

RQ1. Are late L1 Polish L2 German learners sensitive to grammatical gender marking on prenominal modifiers?

RQ2. If sensitivity is present, does it depend on structural congruency in gender marking between L1 and L2?

RQ3. If sensitivity is present, does it depend on semantic and lexical gender congruency?

RQ4. If sensitivity is present, does it vary with L2 proficiency?

In line with the studies reviewed in the previous section, we predict that L1 Polish L2 German learners will exhibit sensitivity to grammatical gender marking during L2 processing, as their L1 encodes grammatical gender (e.g., Hawkins & Chan, 1997; Tsimpli & Dimitrakopoulou, 2007). However, considering the growing evidence that gender agreement processing in L2 is influenced by various language-internal congruency factors (e.g., Foucart & Frenck-Mestre, 2011; Hopp & Lemmerth, 2018), we entertain the possibility that participants' sensitivity to gender marking may depend on structural similarities between L1 and L2. If this holds true, participants are expected to demonstrate greater or exclusive sensitivity to gender agreement violations in adjective-noun pairs, as these structurally align with L1 Polish. Similarly, if lexical and semantic factors contribute to gender agreement processing in L2, we would expect to observe a grammaticality effect modulated by lexical gender congruency (e.g., Hopp & Lemmerth, 2018) or semantic congruency (Scherag et al., 2004). Finally, participants' sensitivity to gender agreement violations is predicted to increase with higher levels of L2 proficiency (e.g., Sagarra & Herschensohn, 2010; Długosz, 2023).

## 4.2 Experiment 1

### 4.2.1 Participants

We recruited 34 Polish learners/speakers of German at a Polish university, including six males ($M_{age}$ = 28.5 years, $SD$ = 3.4). All participants began learning German after the age of six in school settings ($M_{age\ of\ onset}$ = 10.2 years, $SD$ = 2.1). The sample included students at various stages of their studies ($n$ = 22) as well as German teachers ($n$ = 12) to ensure a broad range of German proficiency (from intermediate to advanced levels). Proficiency in German was

assessed using the LexTALE test (Lemhöfer & Broersma, 2012), which yielded a mean score of 74.6% ($SD$ = 16.4, range = 43–98). Informed consent was obtained from all participants, and they were compensated with bookstore gift cards for their participation in the study.

## 4.2.2 Materials

The stimuli comprised 48 real words (nouns) and 48 nonce words, which served as targets and were paired with adjectives functioning as primes. The grammaticality and semantic congruency of the adjective-noun pairs were manipulated, resulting in four conditions: grammatical and semantically congruent (e.g., *guter Tag*, 'good$_M$ day$_M$'), grammatical and semantically incongruent (e.g., *schlankes Glück*, 'slim$_N$ happiness$_N$'), ungrammatical and semantically congruent (e.g., *lange Brief*, 'long$_F$ letter$_M$'), and ungrammatical and semantically incongruent (e.g., *faule Licht*, 'lazy$_F$ light$_N$'). Each condition included 12 German nouns, ranging from three to nine letters in length. The three genders—masculine, feminine, and neuter—were equally represented across the conditions. All nouns were inanimate[1] and matched the gender of their Polish translation equivalents.

The semantic congruency of adjective-noun pairs was independently evaluated by 51 non-target participants from the same population. Participants rated the semantic congruency of each pair on a scale from 1 ('not congruent at all') to 4 ('completely congruent'). The mean ratings for congruent pairs were 3.9 ($SD$ = 0.2, range = 3.7–4.0), while incongruent pairs scored 1.5 ($SD$ = 0.4, range = 1.1–1.7), with the difference being statistically significant (Mann-Whitney $U$ test, $p$ < .001). Importantly, grammatical ($M$ = 2.7, $SD$ = 1.3) and ungrammatical pairs ($M$ = 2.7, $SD$ = 1.2) did not differ significantly (Mann-Whitney $U$ test, $p$ = .975).

The nouns in the four conditions were controlled for relevant lexical variables, including word length (in letters), frequency according to SUBTLEX-DE (Brysbaert et al., 2011), concreteness, arousal, valence, and imageability (Köper & Schulte im Walde, 2016), as well as formal similarity to Polish (calculated using normalised Levenshtein distance). Kruskal-Wallis tests showed no significant differences between the four conditions ($p$s > .160). In addition, the adjectives in the four conditions were controlled for length in letters ($p$ = .133), frequency ($p$ = .199) and normalised Levenshtein distance ($p$ = .160).

The nonce words were generated using UniPseudo (New et al. 2024), based on its built-in word database and matched in length with the real words. Only nonce words without real orthographic neighbours, as determined by GermanPOND (Marian et al., 2012), were selected. To mimic the structure of adjective-noun pairs involving real words, nonce words were paired with randomly selected gender-marked adjectives that differed from those used in the real adjective-noun pairs. The stimuli are provided in Appendix A.

## 4.2.3 Procedure

Each testing session included a lexical decision task, a control gender decision task, a control translation task, the LexTALE test, and a background questionnaire, in that order. The entire experiment was conducted using the PsyToolkit platform (Stoet, 2010, 2017).

In the lexical decision task, each trial began with a fixation cross presented for 500 ms, followed by the adjective as a prime, displayed for 500 ms. After the prime, a letter string

---

1 We avoided animate nouns, as previous research indicates that gender agreement with animate nouns is cognitively more demanding than with inanimate nouns (e.g., Sagarra & Herschensohn, 2011).

appeared on the screen for 3,000 ms or until the participant responded. The intertrial interval was set at 1,000 ms. Participants were asked to press the 'j' (= *ja*, 'yes') key if the letter string was a real German word, and the 'n' (= *nein*, 'no') key if it was not. The stimuli were displayed in white uppercase Arial font against a black background at the centre of a computer screen. The order of stimulus presentation was randomised for each participant. Prior to the task, participants received instructions in German and completed a separate practice block featuring unrelated items.

The two control tasks were untimed. In the gender decision task, participants assigned grammatical genders to all nouns used in the experiment. In the translation task, they translated all German adjectives into Polish. The results of these control tasks were used to refine reaction time data.

### 4.2.4 Data processing and analysis

All calculations and data analyses were conducted using Jamovi (The Jamovi Project, 2023), including the GAMLj module (Gallucci, 2019). The data were analysed using linear mixed-effects models (LMEM) on log-transformed RTs. Categorical variables were dummy coded, and LexTALE scores were standardized as *z*-scores. Random intercepts for participants and items were included in all models to account for variability in baseline performance across individuals and words, respectively. This approach ensures the analysis captures generalisable effects while controlling for idiosyncratic differences in participant response tendencies and item characteristics. All models converged successfully without warnings. Omnibus tests for the fixed effects were conducted to evaluate the overall significance of the factors included in the analysis. The results of these tests are reported in the main text, while the fixed-effects parameter estimates are provided in the tables. Degrees of freedom were calculated using the Satterthwaite method.

Responses exceeding the 3,000 ms response window were excluded from the analysis (1.0%). Additionally, RTs corresponding to the items with unknown gender or unclear adjective meaning were excluded from the analysis (6.7%), as were RTs associated with incorrect responses (1.5%). For the remaining RTs, an absolute cut-off method was applied, excluding RTs shorter than 300 ms and longer than 2,500 ms (1.0%), in accordance with the recommendations of Jiang (2013, p. 70).

## 4.2.5 Results

Figure 1 presents the predicted RTs as a function of grammaticality and semantic congruency.[2]

**Figure 1**

*Predicted RTs as a function of grammaticality and semantic congruency.*
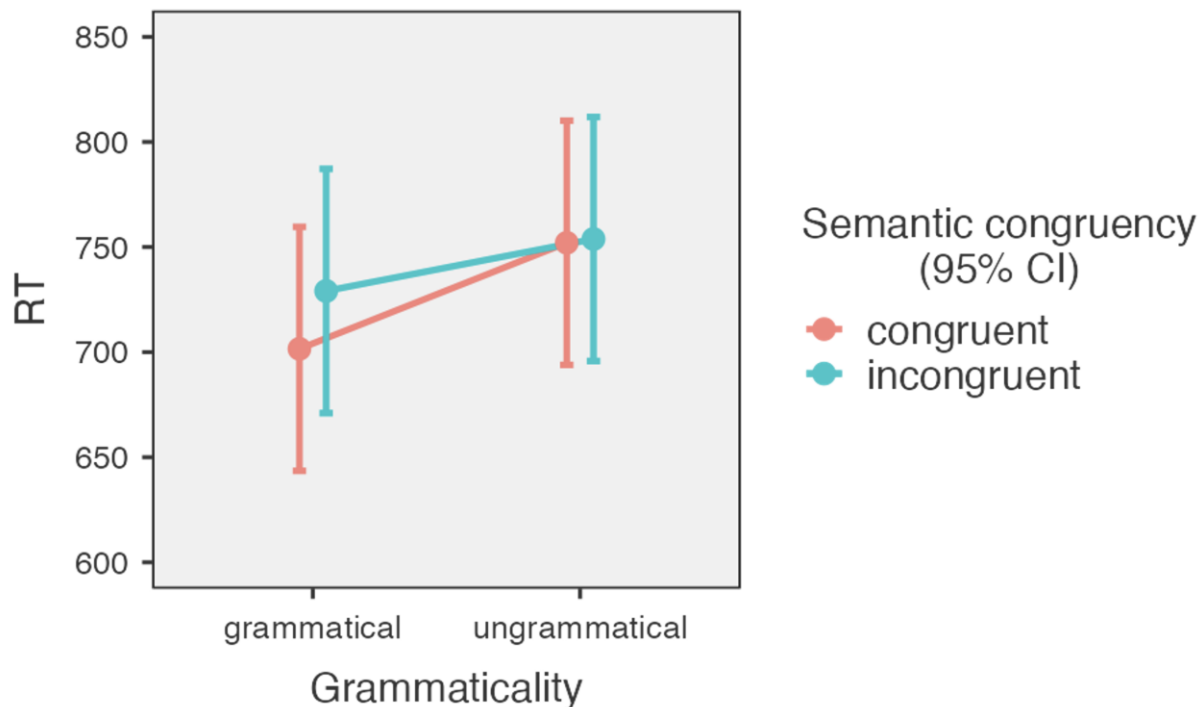


As fixed effects, the model included grammaticality (grammatical, ungrammatical), semantic congruency (congruent, incongruent), gender (masculine, feminine, neuter), and proficiency (standardised LexTALE results as *z*-scores). To investigate whether semantic congruency influences grammaticality effects, the interaction between grammaticality and semantic congruency was included in the model. Additionally, to examine whether grammaticality effects are modulated by German proficiency, the interaction between grammaticality and proficiency was incorporated into the analysis. Fixed effects parameter estimates are included in Table 1.

The results revealed a statistically significant main effect for grammaticality ($F(1, 40.6) = 5.58$, $p = .023$). In contrast, no significant effects were found for semantic congruency ($F(1, 40.6) = 2.10$, $p = .155$), proficiency ($F(1, 32.0) = .09$, $p = .762$), or gender ($F(2, 40.6) = 1.20$, $p = .312$). Regarding interaction effects, the combination of grammaticality and semantic congruency did not yield a significant interaction ($F(1, 40.6) = 1.02$, $p = .318$). Similarly, the interaction between grammaticality and proficiency was not significant ($F(1, 1408.5) = 0.00$, $p = .966$).

These findings indicate that participants were sensitive to gender marking on adjectives; however, this sensitivity was not influenced by semantic congruency or language proficiency.

---

2 For nonce words, the mean RT was 909 ms, which was significantly longer than the RTs for both grammatical items ($p < .001$) and ungrammatical items ($p < .001$), as determined by the Mann-Whitney *U* test.

**Table 1**

*Fixed effects parameter estimates (Experiment 1)*

| Parameter | Effect | Estimate | SE | 95% CI lower | 95% CI upper | df | t | p |
|---|---|---|---|---|---|---|---|---|
| (Intercept) | Intercept | 6.5496 | 0.0338 | 6.4834 | 6.6159 | 35.4 | 193.816 | < .001 |
| Grammaticality | Ungrammatical − Grammatical | 0.0467 | 0.0198 | 0.0079 | 0.0856 | 40.6 | 2.362 | 0.023 |
| Semantic congruency | Congruent − Incongruent | 0.0287 | 0.0198 | −0.0101 | 0.0676 | 40.6 | 1.451 | 0.155 |
| Proficiency | Proficiency | 0.0099 | 0.0326 | −0.0540 | 0.0739 | 32.0 | 0.305 | 0.762 |
| Gender 1 | Masculine − Feminine | 0.0298 | 0.0243 | −0.0176 | 0.0774 | 40.5 | 1.232 | 0.225 |
| Gender 2 | Neuter − Feminine | −0.0049 | 0.0242 | −0.0523 | 0.0425 | 40.2 | −0.203 | 0.840 |
| Grammaticality × Semantic congruency | Ungrammatical − Grammatical × Incongruent − Congruent | −0.0400 | 0.0396 | −0.1177 | 0.0376 | 40.6 | −1.010 | 0.318 |
| Grammaticality × Proficiency | Ungrammatical − Grammatical × Proficiency | 0.0005 | 0.0122 | −0.0234 | 0.0245 | 1408.5 | 0.043 | 0.966 |

# 4.3 Experiment 2

## 4.3.1 Participants

We recruited 32 Polish learners/speakers of German at a Polish university different from those in Experiment 1, including five males ($M_{age}$ = 26.4 years, $SD$ = 9.6). All participants began learning German after the age of six in school settings ($M_{age\ of\ onset}$ = 12.2 years, $SD$ = 4.2). The sample included students at various stages of their studies ($n$ = 23) as well as German teachers ($n$ = 9) to ensure a broad range of German proficiency (from intermediate to advanced levels). Proficiency in German was assessed using the LexTALE test (Lemhöfer & Broersma, 2012), which yielded a mean score of 75.5% ($SD$ = 11.4, range = 61–99). Informed consent was obtained from all participants, and they were compensated with bookstore gift cards for their participation in the study.

## 4.3.2 Materials

The stimuli comprised 80 real words (nouns) and 80 nonce words, which served as targets and were paired with definite articles functioning as primes. The grammaticality and lexical gender congruency (i.e., overlap in grammatical gender between L1 and L2 nouns) of the article-noun pairs were manipulated, resulting in four conditions: grammatical and gender-congruent (e.g., *die Flasche*, 'theF bottleF'), grammatical and gender-incongruent (e.g., *der Löffel*, 'theM spoonM'), ungrammatical and gender-congruent (e.g., *die Tisch*, 'theF tableM'), and ungrammatical and gender-incongruent (e.g., *die Gesicht*, 'theF faceN'). Each condition included 20 German nouns, varying in length from three to thirteen letters. The number of stimuli was greater than in Experiment 1, as the cross-linguistic effects of gender congruency are typically smaller compared to more robust effects of grammatical and semantic violations

(e.g., Sá-Leite et al., 2020). All nouns were inanimate, except for two: *Schmetterling* ('butterfly') and *Schwein* ('pig').

The three grammatical genders—masculine, feminine, and neuter—were not equally represented across the conditions due to the challenges associated with manipulating gender congruency for neuter nouns. Consequently, each condition comprised 8 masculine, 8 feminine, and 4 neuter nouns.

The nouns in the four conditions were controlled for relevant lexical variables, including word length (in letters), frequency according to SUBTLEX-DE (Brysbaert et al., 2011), concreteness, arousal, valence, and imageability (Köper & Schulte im Walde, 2016), as well as formal similarity to Polish (calculated using normalised Levenshtein distance). Kruskal-Wallis tests showed no significant differences across the four conditions ($ps > .470$).

The nonce words were generated using UniPseudo (New et al., 2024), based on its built-in word database and matched in length with the real words. Only nonce words without real orthographic neighbours, as determined by GermanPOND (Marian et al., 2012), were selected. The stimuli are provided in Appendix B.

### 4.3.3 Procedure

Each testing session included a lexical decision task, a control gender decision task, the LexTALE test, and a background questionnaire, in that order. The entire experiment was conducted using the PsyToolkit platform (Stoet, 2010, 2017). The procedure of the lexical decision task used in Experiment 2 was identical to that of Experiment 1, except that the primes were changed from adjectives (Experiment 1) to definite articles (Experiment 2). The control gender decision task was untimed.

### 4.3.4 Data processing and analysis

Responses exceeding the 3,000 ms response window were excluded from the analysis (0.7%). Additionally, RTs corresponding to the items with unknown gender were excluded from the analysis (22.9%), as were RTs associated with incorrect responses (18.8%). For the remaining RTs, an absolute cut-off method was applied, excluding RTs shorter than 300 ms and longer than 2,500 ms (0.4%), in accordance with the recommendations of Jiang (2013, p. 70). The analytical procedures were identical to those used in Experiment 1.

### 4.3.5 Results

Figure 2 presents the predicted RTs as a function of grammaticality and lexical gender congruency.[3]

---

3 For nonce words, the mean RT was 948 ms, which was significantly longer than the RTs for both grammatical items ($p < .001$) and ungrammatical items ($p < .001$), as determined by the Mann-Whitney *U* test.

**Figure 2**

*Predicted RTs as a function of grammaticality and lexical gender congruency.*
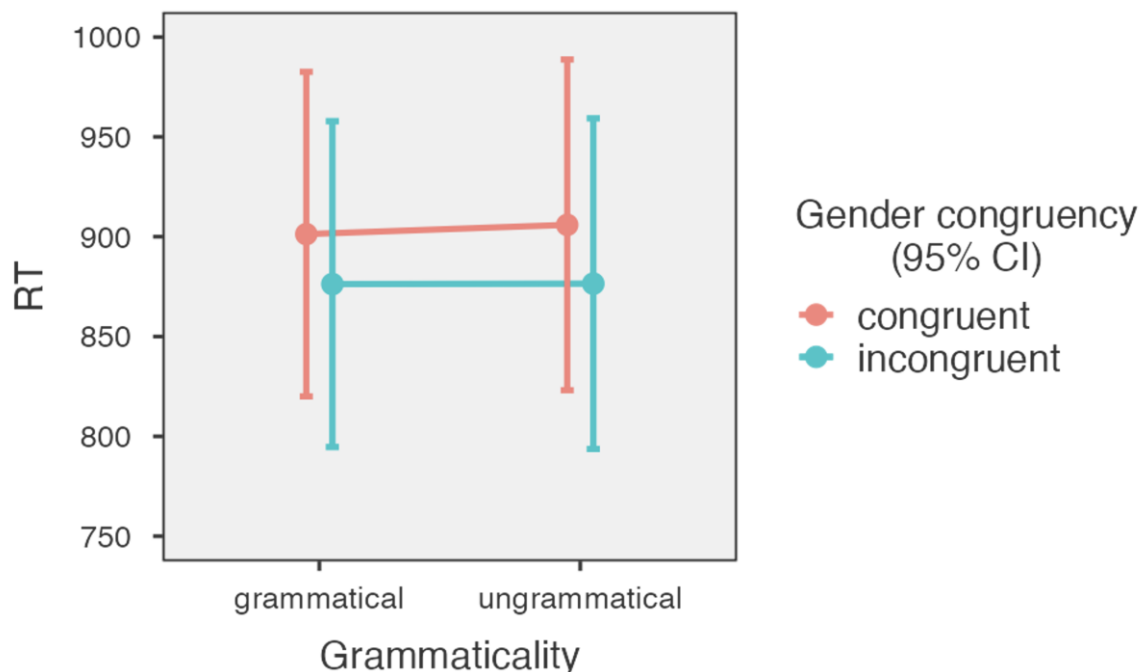


As fixed effects, the model included grammaticality (grammatical, ungrammatical), lexical gender congruency (congruent, incongruent), gender (masculine, feminine, neuter), and proficiency (standardised LexTALE results as *z*-scores). To explore whether lexical gender congruency influences grammaticality effects, the interaction between grammaticality and lexical gender congruency was included. Additionally, to determine whether grammaticality effects are modulated by German proficiency, the interaction between grammaticality and proficiency was analysed. Finally, given that lexical gender congruency effects may depend on L2 proficiency (e.g., Długosz & Olszewska, 2024), the interaction between lexical gender congruency and proficiency was also included in the model. Fixed effects parameter estimates are included in Table 2.

The results revealed no significant main effects or interactions across the variables analysed. Specifically, the effects of grammaticality ($F(1, 74.5) = 0.01$, $p = .940$), lexical gender congruency ($F(1, 71.8) = 1.40$, $p = .240$), gender ($F(2, 70.5) = 0.14$, $p = .871$), and proficiency ($F(1, 29.9) = 0.18$, $p = .672$) were not significant.

Likewise, interactions between these variables did not yield significant results. The interactions of grammaticality and lexical gender congruency ($F(1, 71.9) = 0.10$, $p = .755$), grammaticality and proficiency ($F(1, 1584.9) = 0.34$, $p = .559$), and lexical gender congruency and proficiency ($F(1, 1572.9) = 0.03$, $p = .861$) all failed to reach significance.

These findings indicate that participants were insensitive to gender marking on definite articles, regardless of lexical gender congruency and language proficiency.

**Table 2**

*Fixed effects parameter estimates (Experiment 2)*

| Parameter | Effect | *Estimate* | *SE* | 95% CI lower | 95% CI upper | *df* | *t* | *p* |
|---|---|---|---|---|---|---|---|---|
| (Intercept) | (Intercept) | 6.7320 | 0.0389 | 6.6550 | 6.8084 | 32.6 | 172.848 | < .001 |
| Grammaticality | Ungrammatical − Grammatical | 0.0015 | 0.0202 | -0.0380 | 0.0410 | 74.5 | 0.076 | 0.940 |
| Lexical gender congruency | Incongruent − Congruent | -0.0236 | 0.0200 | -0.0628 | 0.0155 | 71.8 | -1.184 | 0.240 |
| Gender 1 | Masculine − Feminine | -0.0029 | 0.0223 | -0.0467 | 0.0409 | 70.9 | -0.130 | 0.896 |
| Gender 2 | Neuter − Feminine | 0.0112 | 0.0272 | -0.0420 | 0.0644 | 69.8 | 0.412 | 0.681 |
| Proficiency | Proficiency | -0.0159 | 0.0372 | -0.0889 | 0.0571 | 29.9 | -0.428 | 0.672 |
| Grammaticality × Lexical gender congruency | Ungrammatical − Grammatical × Incongruent − Congruent | -0.0125 | 0.0400 | -0.0909 | 0.0658 | 71.9 | -0.313 | 0.755 |
| Grammaticality × Proficiency | Ungrammatical − Grammatical × Proficiency | 0.0077 | 0.0132 | -0.0181 | 0.0336 | 1584.9 | 0.585 | 0.559 |
| Lexical gender congruency × Proficiency | Incongruent − Congruent × Proficiency | 0.0022 | 0.0125 | -0.0222 | 0.0266 | 1572.9 | 0.1754 | 0.861 |

## *4.3.6 Native speakers of German*

In order to determine whether our priming task taps into implicit gender processing and whether the null effect observed in the L2 group can indeed be interpreted as a lack of gender priming, we tested 22 native speakers of German residing in Germany, including two males ($M_{age} = 23.8$ years, $SD = 5.0$). The procedures were the same as those for the L2 group, except for the control gender decision task and the LexTALE test. Participants were recruited through personal contacts and sent a link to the experiment, which they completed in their free time.

Responses exceeding the 3,000 ms response window were excluded from the analysis (0.6%), as were RTs associated with incorrect responses (3.2%). Among the remaining RTs, none were shorter than 300 ms or longer than 2,500 ms. As fixed effects, the model included grammaticality (grammatical, ungrammatical), gender (masculine, feminine, neuter), and their interaction.

The results revealed a main effect of grammaticality ($F(1, 73.1) = 12.6$, $p < .001$). In turn, the effect of gender was not significant ($F(2, 73.1) = 0.76$, $p = .472$), nor was the interaction between grammaticality and gender ($F(2, 73.1) = 0.07$, $p = .934$). Figure 3 presents the predicted RTs as a function of grammaticality, while Table 3 shows the fixed effects parameter estimates.

**Table 3**

*Fixed effects parameter estimates (Experiment 2, native speakers of German)*

| Parameter | Effect | Estimate | SE | 95% CI lower | 95% CI upper | df | t | p |
|-----------|--------|----------|-----|--------------|--------------|-----|-----|-----|
| (Intercept) | (Intercept) | 6.4769 | 0.0415 | 6.3956 | 6.5583 | 21.7 | 156.017 | < .001 |
| Grammaticality | Ungrammatical − Grammatical | 0.0515 | 0.0145 | 0.0231 | 0.0801 | 73.1 | 3.546 | <.001 |
| Gender 1 | Masculine − Feminine | 0.0036 | 0.0154 | -0.0266 | 0.0339 | 73.5 | 0.235 | 0.815 |
| Gender 2 | Neuter − Feminine | -0.0189 | 0.0189 | -0.0560 | 0.0180 | 72.9 | -0.006 | 0.318 |
| Grammaticality × Gender 1 | Ungrammatical − Grammatical × Masculine − Feminine | 0.0019 | 0.0309 | -0.0586 | 0.0625 | 73.5 | 0.064 | 0.949 |
| Grammaticality × Gender 2 | Ungrammatical − Grammatical × Neuter − Feminine | 0.0135 | 0.0378 | -0.0604 | 0.0875 | 72.9 | 0.3587 | 0.721 |

**Figure 3**

*Predicted RTs as a function of grammaticality (native speakers of German).*
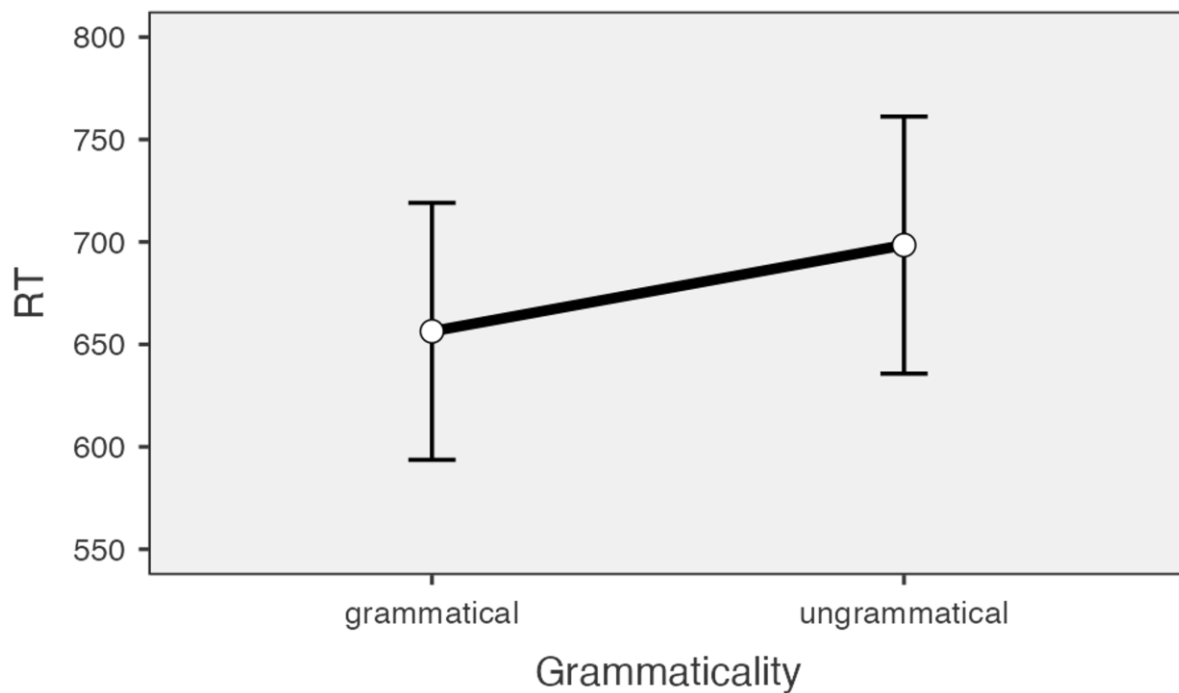


The results from the native group demonstrate clear gender priming results, indicating that our method was effective and that the L2 learners' lack of sensitivity to gender marking on the article was a genuine finding.

# 5 Discussion

This study aimed to examine whether adult L2 German learners with a gendered L1, Polish, demonstrate sensitivity to grammatical gender during L2 processing. Specifically, we investigated how this sensitivity is influenced by structural, lexical, and semantic congruency, as well as L2 proficiency. To address these questions, two primed lexical decision tasks were conducted, exploring the extent to which these factors contribute to grammaticality effects arising from contrasts between gender agreement and disagreement.

Our findings revealed that late L1 Polish L2 German learners took longer to respond when nouns were preceded by incorrectly inflected adjectives compared to correctly inflected ones. This slowdown effect was not modulated by semantic congruency between the noun and the adjective, indicating that participants processed gender disagreement more slowly than agreement, even when the adjective-noun pairs lacked meaning. Furthermore, the slowdown effect was not modulated by L2 proficiency, suggesting no differences in gender agreement processing between intermediate and highly advanced L2 learners. In contrast, no slowdown effects were observed in the processing of article-noun pairs, even when the nouns were gender-congruent with their L1 translation equivalents, and even among the most proficient participants.

These findings indicate that sensitivity to gender information in L2 is shaped by how gender is marked in L1, supporting previous research on the central role of L1-L2 structural similarities (e.g., Sabourin & Stowe, 2008; Foucart & Frenck-Mestre, 2011; Hopp & Lemmerth, 2018). Participants were sensitive to gender agreement violations when processing adjective-noun phrases, which share structural similarities in gender marking, but not when processing article-noun phrases, which do not. Importantly, this target-like sensitivity in adjective-nouns phrases was unaffected by L2 proficiency, suggesting that structural similarities in gender marking between L1 and L2 enable L2 learners to process gender agreement in a target-like manner even at intermediate proficiency levels.

Conversely, our findings show that even if the gender feature is present in the learners' L1, gender agreement processing in L2 is still constrained by the similarities and differences in how gender is marked between L1 and L2. Although Polish, like German, is a gendered language with a three-way gender system, our participants did not exhibit sensitivity to gender violations involving article-based gender marking. This is likely because such marking is absent in Polish, an articleless language. Notably, this pattern was observed even among the most proficient participants.

The impact of semantic congruency between adjectives and nouns on gender agreement processing in L2 has not been thoroughly investigated. To our knowledge, only one study has addressed this issue. Using a primed lexical decision task, Scherag et al. (2004) found that late L1 English L2 German learners were sensitive to gender marking on adjectives only when the adjectives were semantically associated with nouns. This contrasts with the findings of our study, where such sensitivity was observed even in the absence of semantic associations. It is likely that the gendered nature of our participants' L1, Polish, provided an advantage in processing meaningless adjective-noun pairs. This suggests that the presence of gender in L1 may enable L2 learners to compute gender agreement independently of semantic context. Further research is needed to confirm whether the differences between our findings and those of Scherag et al. (2004) are indeed attributable to the participants' L1 background.

With regard to another factor of interest, namely lexical gender congruency, we found that participants' performance was not influenced by L1-L2 gender correspondences. This contrasts with previous studies that have reported lexical gender congruency effects in various

populations (e.g., Sá-Leite et al., 2020) and specifically in L1 Polish L2 German learners (Długosz & Olszewska, 2024). It is important to note that in Experiment 2, no grammaticality effects were observed, meaning that any potential modulating effect of lexical gender congruency could not be detected. Future research is needed to determine whether a gender congruency effect might emerge in the processing of adjective-noun pairs, for which we did observe grammaticality effects. Note, however, that in Experiment 1, all nouns were incongruent in gender with their L1 Polish translation equivalents, which did not eliminate the grammaticality effects. In addition, we believe that the necessary inclusion of ungrammatical article-noun pairs may have prevented the occurrence of a lexical gender congruency effect, as processing gender disagreement may override or interfere with lexical processing and increase cognitive load.

In addition, we found no differences in the processing of masculine, feminine, and neuter nouns. This suggests that late L1 Polish L2 German learners are sensitive to gender marking on adjectives to a similar extent across all genders, thus confirming previous findings regarding L1 and L2 German (e.g., Hillert & Bates, 1996; Hopp, 2013).

To our knowledge, this study is the first to examine gender priming effects in L2 among learners with a gendered L1. By demonstrating target-like gender priming effects in late L2 learners irrespective of proficiency, as well as differing effects based on the type of gender marking, our study has important implications for L2 acquisition theory and teaching practice. First, adult L2 learners are able to make use of morphosyntactic information in L2 processing when the realisation of the gender feature in L2 aligns with that in their L1. Therefore, target-like competence in grammatical morphemes and in agreement processes involving these morphemes is achievable for adult L2 learners. Our findings further show that late L2 learners whose L1 encodes gender in a structurally similar way can benefit from gender-marked words in L2 comprehension, potentially giving them an advantage over learners with ungendered L1s or L1s with different gender marking systems. In the context of increasingly multilingual classrooms, these insights are particularly important for educators to consider when assessing learners' grammatical competence.

Given the nature of the method employed in this study—specifically, the combination of a priming paradigm with a timed lexical decision task, which taps into implicit gender agreement processes (as participants' attention is not focused on grammaticality and they have limited time to access metalinguistic knowledge)—we can draw meaningful conclusions regarding the acquisition of gender agreement in L2. Our findings suggest that the development of implicit knowledge of gender agreement, i.e., knowledge utilised without conscious awareness (Suzuki, 2017), may be shaped by structural similarities and differences between L1 and L2. This implies that the mere presence of a morphosyntactic feature in L2 may not guarantee its successful acquisition. What appears to be more crucial is how this feature is realised in both L1 and L2, highlighting the critical role of structural alignment between the two languages.

In conclusion, our study provides evidence that structural similarity plays a pivotal role in gender agreement processing in L2. When the structural realisation of gender in L2 aligns with that in L1, L2 learners show sensitivity to gender marking regardless of their L2 proficiency and lexical factors, such as semantic associations between the noun and its modifier or lexical gender congruency. This is underscored by the grammaticality effect observed even when all nouns were gender-incongruent with their L1 counterparts.

Our study has some limitations. First, we focused on L2 learners at intermediate to advanced L2 proficiency levels. In future studies, it would be interesting to explore whether the grammaticality effect observed in our study extends to beginner L2 learners. Second, the inclusion of a group of L2 learners with an ungendered L1 would have provided valuable

insights, allowing us to distinguish the effects of structural similarity from the presence or absence of gender in L1. Third, due to organisational constraints, native speakers of German were only tested in Experiment 2, which remains a limitation of our study. Lastly, a larger sample size would enhance the reliability and generalisability of our findings.

## Acknowledgment

## References

Alemán Bañón, J., Miller, D., & Rothman, J. (2017). Morphological variability in second language learners: An examination of electrophysiological and production data. *Journal of Experimental Psychology: Learning, Memory, and Cognition*, *43*(10), 1509–1536. https://doi.org/10.1037/xlm0000394

Brysbaert, M., Buchmeier, M., Conrad, M., Jacobs, A. M., Bölte, J., & Böhl, A. (2011). The word frequency effect: a review of recent developments and implications for the choice of frequency estimates in German. *Experimental Psychology*, *58*(5), 412–424. https://doi.org/10.1027/1618-3169/a000123

Carroll, S. E. (1989). Language acquisition studies and a feasible theory of grammar. Canadian *Journal of Linguistics*, *34*(4), 399–418. https://doi.org/10.1017/S0008413100024300

Dewaele, J.-M., & Véronique, D. (2001). Gender assignment and gender agreement in advanced French interlanguage: A cross-sectional study. *Bilingualism: Language and Cognition*, *4*(3), 275–297. https://doi.org/10.1017/S136672890100044X

Długosz, K. (2023). Processing gender agreement in an additional language: The more languages the better? *Second Language Research*, *39*(4), 997–1026. https://doi.org/10.1177/02676583221113333

Długosz, K. & Olszewska, A. (2024). The gender congruency effect in sentence context depends on gender transparency and L2 proficiency: A self-paced reading study with Polish–German bilinguals. *International Journal of Bilingualism*. Online First. https://doi.org/10.1177/13670069241288318

Duden (2023, December 15). Die Verteilung der Artikel (Genusangabe) im Rechtschreibduden. https://duden.de/sprachwissen/sprachratgeber/Die-Verteilung-der-Artikel-Genusangabe-im-Rechtschreibduden

Dussias, P. E., Valdés Kroff, J. R., Guzzardo Tamargo, R. E., & Gerfen, C. (2013). When gender and looking go hand in hand: Grammatical gender processing in L2 Spanish. *Studies in Second Language Acquisition*, *35*(2), 353–387. https://doi.org/10.1017/S0272263112000915

Ellis, C., Conradie, S., & Huddlestone, K. (2012). The acquisition of grammatical gender in L2 German by learners with Afrikaans, English or Italian as their L1. *Stellenbosch Papers in Linguistics*, *41*, 17–27. https://doi.org/10.5774/41-0-131

Foote, R. (2014). Age of acquisition and sensitivity to gender in Spanish word recognition. *Language Acquisition*, *21*(4), 365–385. https://doi.org/10.1080/10489223.2014.892948

Foucart, A., & Frenck-Mestre, C. (2011). Grammatical gender processing in L2: Electrophysiological evidence of the effect of L1-L2 syntactic similarity. *Bilingualism: Language and Cognition*, *14*(3), 379–399. https://doi.org/10.1017/S136672891000012X

Franceschina, F. (2005). *Fossilized second language grammars: The acquisition of grammatical gender*. John Benjamins.

Gallucci, M. (2019). jAMM: jamovi Advanced Mediation Models [jamovi module]. https://jamovi-amm.github.io/

Garrido-Pozu, J. J. (2022). Predictive processing of grammatical gender: Using gender cues to facilitate processing in Spanish. *Lingua*, *278*, 103416. https://doi.org/10.1016/j.lingua.2022.103416

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

Grzegorczykowa, R., Laskowski, R., & Wróbel, H. (1999). *Gramatyka współczesnego języka polskiego. Morfologia*. PWN.

Guillelmon, D., & Grosjean, F. (2001). The gender marking effect in spoken word recognition: The case of bilinguals. *Memory & Cognition*, *29*(3), 503–511. https://doi.org/10.3758/BF03196401

Hawkins, R., & Chan, C. (1997). The partial accessibility of Universal Grammar in second language acquisition: The failed functional features hypothesis. *Second Language Research*, *13*(2), 187–226. https://doi.org/10.1191/026765897671476153

Hillert, D., & Bates, E. (1996). *Morphological constraints on lexical access: Gender priming in German*. Technical Report 9601, University of California, San Diego, Center for Research in Language.

Hockett, C. F. (1958). *A course in modern linguistics*. MacMillan.

Hohlfeld, A. (2006). Accessing grammatical gender in German: The impact of gender-marking regularities. *Applied Psycholinguistics*, *27*(2), 127–142. https://doi.org/10.1017/S0142716406060218

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

Hopp, H., & Lemmerth, N. (2018). Lexical and syntactic congruency in L2 predictive gender processing. *Studies in Second Language Acquisition*, *40*(1), 171–199. https://doi.org/10.1017/S0272263116000437

Jiang, N. (2013). *Conducting reaction time research in second language studies*. Routledge.

Jiang, N. (2018). *Second language processing: An introduction*. Routledge.

Johannessen, J. B., Lundquist, B., Rodina, Y., Tengesdal, E., Kaldhol, N. H., Türker, E., & Fyndanis, V. (2024). Cross-linguistic effects in grammatical gender assignment and predictive processing in L1 Greek, L1 Russian, and L1 Turkish speakers of Norwegian as a second language. *Second Language Research*. Online First. https://doi.org/10.1177/02676583241227709

Klassen, R. (2016). When masculine as default supercedes L1 transfer: Bilingual speakers of languages with asymmetric gender systems. In A. Alba de la Fuente, E. Valenzuela & C. Martínez Sanz (Eds.), *Language acquisition beyond parameters: Studies in honour of Juana M. Liceras* (pp. 73–98). John Benjamins.

Köper, M., & Schulte im Walde, S. (2016). Automatically generated affective norms of abstractness, arousal, imageability, and valence for 350,000 German lemmas. In N. Calzolari, K. Choukri, T. Declerck, S. Goggi, M. Grobelnik, B. Maegaard, J. Mariani, H. Mazo, A. Moreno, J. Odijk, & S. Piperidis (Eds.), *Proceedings of the Tenth International Conference on Language Resources and Evaluation* (LREC'16) (pp. 2595–2598). European Language Resources Association. https://aclanthology.org/L16-1413.pdf

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

Lemhöfer, K., & Broersma, M. (2012). Introducing LexTALE: A quick and valid lexical test for advanced learners of English. *Behavior Research Methods*, *44*(2), 325–343. https://doi.org/10.3758/s13428-011-0146-0

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

Mačiukaitė, S. (2008). Differences in adult L2 grammars of Lithuanian: The case of gender agreement. *Eesti Rakenduslingvistika Ühingu Aastaraama*, *4*, 87–103. http://dx.doi.org/10.5128/ERYa4.06

Marian, V., Bartolotti, J., Chabal, S., & Shook, A. (2012). CLEARPOND: Cross-linguistic easy-access resource for phonological and orthographic neighborhood densities. *PLoS ONE*, *7*(8), e43230. https://doi.org/10.1371/journal.pone.0043230

Meisel, J. M. (2009). Second language acquisition in early childhood. *Zeitschrift für Sprachwissenschaft*, *28*(1), 5–34. https://doi.org/10.1515/ZFSW.2009.002

New, B., Bourgin, J., Barra, J., & Pallier, C. (2024). UniPseudo: A universal pseudoword generator. *Quarterly Journal of Experimental Psychology*, *77*(2), 278–286. https://doi.org/10.1177/17470218231164373

Prévost, P., & White, L. (2000). Missing Surface Inflection or Impairment in second language acquisition? Evidence from tense and agreement. *Second Language Research*, *16*(2), 103–133. https://doi.org/10.1191/026765800677556046

Romanova, R., & Gor, K. (2017). Processing of gender and number agreement in Russian as a second language: The devil is in the details. *Studies in Second Language Acquisition*, *39*(1), 97–128. https://doi.org/10.1017/S0272263116000012

Sabourin, L. (2003). *Grammatical gender and second language processing: An ERP study* [Unpublished doctoral thesis]. Rijksuniversiteit te Groningen.

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

Sabourin, L., & Stowe, L. A. (2008). Second language processing: When are first and second languages processed similarly? *Second Language Research*, *24*(3), 397–430. https://doi.org/10.1177/0267658308090186

Sagarra, N., & Herschensohn, J. (2010). The role of proficiency and working memory in gender and number agreement processing in L1 and L2 Spanish. *Lingua*, *120*(8), 2022–2039. https://doi.org/10.1016/j.lingua.2010.02.004

Sagarra, N., & Herschensohn, J. (2011). Proficiency and animacy effects on L2 gender processing during comprehension. *Language Learning*, *61*(1), 80–116. https://doi.org/10.1111/j.1467-9922.2010.00588.x

Sá-Leite, A. R., Luna, K., Fraga, I., & Comesaña, M. (2020). The gender congruency effect across languages in bilinguals: A meta-analysis. *Psychonomic Bulletin & Review*, *27*, 677–693. https://doi.org/10.3758/s13423-019-01702-w

Scherag, A., Demuth, L., Rösler, F., Neville, H. J., & Röder, B. (2004). The effects of late acquisition of L2 and the consequences of immigration on L1 for semantic and morpho-syntactic language aspects. *Cognition*, *93*(3), B97–B108. https://doi.org/10.1016/j.cognition.2004.02.003

Stefańczyk, W. T. (2007). *Kategoria rodzaju i przypadka polskiego rzeczownika: próba synchronicznej analizy morfologicznej*. Kraków: Wydawnictwo UJ.

Stoet, G. (2010). PsyToolkit: A software package for programming psychological experiments using Linux. *Behavior Research Methods*, *42*(4), 1096–1104. https://doi.org/10.3758/BRM.42.4.1096

Stoet, G. (2017). PsyToolkit: A novel web-based method for running online questionnaires and reaction-time experiments. *Teaching of Psychology*, *44*(1), 24–31. https://doi.org/10.1177/0098628316677643

Suzuki, Y. (2017). Validity of new measures of implicit knowledge: Distinguishing implicit knowledge from automatized explicit knowledge. *Applied Psycholinguistics*, *38*(5), 1–33. https://doi.org/10.1017/S014271641700011X

The Jamovi Project. (2023). Jamovi (Version 2.3) [Computer Software]. https://www.jamovi.org

Tsimpli, I., & Dimitrakopoulou, M. (2007). The Interpretability Hypothesis: Evidence from wh-interrogatives in second language acquisition. *Second Language Research*, *23*(2), 215–242. https://doi.org/10.1177/0267658307076546

Unsworth, S. (2008). Age and input in the acquisition of grammatical gender in Dutch. *Second Language Research*, *24*(3), 365–395. https://doi.org/10.1177/0267658308090185

White, L., Valenzuela, E., Kozlowska-Macgregor, M., & Leung, Y. I. (2004). Gender and number agreement in nonnative Spanish. *Applied Psycholinguistics*, *25*(1), 105–133. https://doi.org/10.1017/S0142716404001067

# Appendix A

*Adjective-noun pairs used in Experiment 1*

| Prime | Target | Grammaticality | Semantic congruency |
|---|---|---|---|
| NEUER | ANFANG | grammatical | congruent |
| GUTER | TAG | grammatical | congruent |
| WEIßER | BUS | grammatical | congruent |
| NEUER | COMPUTER | grammatical | congruent |
| ROHES | EI | grammatical | congruent |
| OFFENES | FENSTER | grammatical | congruent |
| GUTES | KINO | grammatical | congruent |
| EIGENES | BÜRO | grammatical | congruent |
| LANGE | STRAßE | grammatical | congruent |
| LEERE | FLASCHE | grammatical | congruent |
| LANGE | PAUSE | grammatical | congruent |
| KLEINE | LAMPE | grammatical | congruent |
| NEUES | BAHNHOF | ungrammatical | congruent |
| LANGE | BRIEF | ungrammatical | congruent |
| LANGES | URLAUB | ungrammatical | congruent |
| SCHRIFTLICHE | TEST | ungrammatical | congruent |
| BREITER | BETT | ungrammatical | congruent |
| BUNTE | GLAS | ungrammatical | congruent |
| KULTURELLER | ZENTRUM | ungrammatical | congruent |
| SCHNELLE | TEMPO | ungrammatical | congruent |
| RECHTER | HAND | ungrammatical | congruent |
| ERSTES | HILFE | ungrammatical | congruent |
| GRÜNER | KARTE | ungrammatical | congruent |
| GANZES | KLASSE | ungrammatical | congruent |
| TRAURIGER | HUNGER | grammatical | incongruent |
| NETTER | TISCH | grammatical | incongruent |
| TOTER | FILM | grammatical | incongruent |
| HELLER | APPETIT | grammatical | incongruent |
| KRANKES | BIER | grammatical | incongruent |
| SCHLANKES | GLÜCK | grammatical | incongruent |
| KLUGES | KONTO | grammatical | incongruent |
| JUNGES | HOBBY | grammatical | incongruent |
| DICKE | ARBEIT | grammatical | incongruent |
| NASSE | BITTE | grammatical | incongruent |
| GESUNDE | MASCHINE | grammatical | incongruent |
| SAURE | GRUPPE | grammatical | incongruent |
| KALTES | FEHLER | ungrammatical | incongruent |
| FRECHE | SCHIRM | ungrammatical | incongruent |
| WARMES | PLAN | ungrammatical | incongruent |
| LECKERE | TERMIN | ungrammatical | incongruent |
| GEFÄHRLICHER | FOTO | ungrammatical | incongruent |
| FAULE | LICHT | ungrammatical | incongruent |

| | | | |
|---|---|---|---|
| NAIVER | RADIO | ungrammatical | incongruent |
| DUMME | STUDIO | ungrammatical | incongruent |
| SCHWACHER | ROSE | ungrammatical | incongruent |
| SONNIGES | PUPPE | ungrammatical | incongruent |
| SÜßER | FIRMA | ungrammatical | incongruent |
| FETTES | SCHULE | ungrammatical | incongruent |

# Appendix B

*Article-noun pairs used in Experiment 2*

| Prime | Target | Grammaticality | Gender congruency |
|-------|--------|----------------|-------------------|
| DIE | BÜRSTE | grammatical | congruent |
| DIE | WÜSTE | grammatical | congruent |
| DIE | SÄGE | grammatical | congruent |
| DIE | KERZE | grammatical | congruent |
| DIE | WELLE | grammatical | congruent |
| DIE | PUPPE | grammatical | congruent |
| DIE | PFEIFE | grammatical | congruent |
| DIE | KIRSCHE | grammatical | congruent |
| DER | SCHATTEN | grammatical | congruent |
| DER | KNOPF | grammatical | congruent |
| DER | SCHLÜSSEL | grammatical | congruent |
| DER | MOND | grammatical | congruent |
| DER | TISCH | grammatical | congruent |
| DER | GÜRTEL | grammatical | congruent |
| DER | REGEN | grammatical | congruent |
| DER | TELLER | grammatical | congruent |
| DAS | OHR | grammatical | congruent |
| DAS | FLEISCH | grammatical | congruent |
| DAS | KNIE | grammatical | congruent |
| DAS | GEFÄNGNIS | grammatical | congruent |
| DER | PFÜTZE | ungrammatical | congruent |
| DER | NARBE | ungrammatical | congruent |
| DER | FLASCHE | ungrammatical | congruent |
| DER | MÜNZE | ungrammatical | congruent |
| DER | TASCHE | ungrammatical | congruent |
| DER | FAUST | ungrammatical | congruent |
| DAS | WURST | ungrammatical | congruent |
| DAS | SCHUBLADE | ungrammatical | congruent |
| DIE | WALD | ungrammatical | congruent |
| DIE | SCHMETTERLING | ungrammatical | congruent |
| DIE | TOPF | ungrammatical | congruent |
| DIE | SCHNEE | ungrammatical | congruent |
| DIE | KAMM | ungrammatical | congruent |
| DIE | DAUMEN | ungrammatical | congruent |
| DAS | KORB | ungrammatical | congruent |
| DAS | TEPPICH | ungrammatical | congruent |
| DER | AUGE | ungrammatical | congruent |
| DER | BETT | ungrammatical | congruent |
| DIE | EI | ungrammatical | congruent |
| DIE | FRÜHSTÜCK | ungrammatical | congruent |
| DIE | DUSCHE | grammatical | incongruent |
| DIE | NASE | grammatical | incongruent |

| | | | |
|---|---|---|---|
| DIE | ECKE | grammatical | incongruent |
| DIE | ZIGARETTE | grammatical | incongruent |
| DIE | BRÜCKE | grammatical | incongruent |
| DIE | MAUER | grammatical | incongruent |
| DIE | MILCH | grammatical | incongruent |
| DIE | FEDER | grammatical | incongruent |
| DER | KNOCHEN | grammatical | incongruent |
| DER | FLUSS | grammatical | incongruent |
| DER | AUFZUG | grammatical | incongruent |
| DER | KÜHLSCHRANK | grammatical | incongruent |
| DER | LÖFFEL | grammatical | incongruent |
| DER | PFIRSICH | grammatical | incongruent |
| DER | STUHL | grammatical | incongruent |
| DER | DURST | grammatical | incongruent |
| DAS | MESSER | grammatical | incongruent |
| DAS | HAUS | grammatical | incongruent |
| DAS | NETZ | grammatical | incongruent |
| DAS | SCHWEIN | grammatical | incongruent |
| DER | TROMMEL | ungrammatical | incongruent |
| DER | BLUME | ungrammatical | incongruent |
| DER | KIRCHE | ungrammatical | incongruent |
| DER | TOMATE | ungrammatical | incongruent |
| DER | KETTE | ungrammatical | incongruent |
| DER | GURKE | ungrammatical | incongruent |
| DER | FAHRRAD | ungrammatical | incongruent |
| DER | GEHIRN | ungrammatical | incongruent |
| DIE | GESICHT | ungrammatical | incongruent |
| DIE | HEMD | ungrammatical | incongruent |
| DIE | TURM | ungrammatical | incongruent |
| DIE | KÄFIG | ungrammatical | incongruent |
| DIE | WÜRFEL | ungrammatical | incongruent |
| DIE | BERG | ungrammatical | incongruent |
| DIE | VORHANG | ungrammatical | incongruent |
| DIE | FUSS | ungrammatical | incongruent |
| DAS | FLÜGEL | ungrammatical | incongruent |
| DAS | SPIEGEL | ungrammatical | incongruent |
| DAS | STADT | ungrammatical | incongruent |
| DAS | BUTTER | ungrammatical | incongruent |