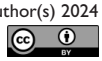
Original Article

The International Journal of Bilingualism

International Journal of Bilingualism
2025, Vol. 29(5) 1474–1488



DOI: 10.1177/13670069241288318
journals.sagepub.com/home/ijb


The gender congruency effect in sentence context depends on gender transparency and L2 proficiency: A self-paced reading study with Polish–German bilinguals

Kamil Długosz and Anna Olszewska
Adam Mickiewicz University, Poznań

Abstract

Aims and objectives: The gender congruency effect, and so the facilitated retrieval of grammatical gender for translation equivalents of the same gender, has been documented in the recognition of isolated words in L2. Our study investigates whether this effect extends to word recognition in sentence context and whether it is modulated by gender transparency and L2 proficiency.

Design/methodology/approach: Sixty late unbalanced Polish–German bilinguals completed a self-paced reading task with noun phrases manipulated for lexical gender congruency with L1 (congruent, incongruent) and noun-internal gender transparency in L2 (transparent, non-transparent). L2 proficiency was assessed through the Dialang test.

Data and analysis: Reading times on the noun were analysed by means of linear mixed-effects modelling.

Findings/conclusion: Results showed a gender congruency effect that was limited to transparent nouns and which was greater at lower L2 proficiency levels.

Originality: The study is the first to investigate the gender congruency effect in the recognition of nouns embedded in grammatical sentences.

Significance/implications: The study shows that the presence of a contextualising sentence does not eliminate the cross-language effects of grammatical gender. Furthermore, it provides first evidence that individual noun-internal gender cues modulate the gender congruency effect. The findings are discussed in terms of the relationship between gender transparency at the micro-level and the degree of activation of gender values.

Keywords

Gender congruency effect, gender transparency, sentence context, self-paced reading, L2 German

Corresponding author:
Kamil Długosz, Institute of Applied Linguistics, Faculty of Modern Languages and Literatures, Adam Mickiewicz University, Collegium Novum, Aleja Niepodległości 4, PL 61-874 Poznań, Poland.
Email: kamdlu@amu.edu.pl

Introduction

Strong evidence exists that the two gender systems of a bilingual interact in language processing. In L2 comprehension and production, bilinguals typically process nouns that share gender between L1 and L2 faster than nouns that differ in gender between L1 and L2 (e.g., Bordag & Pechmann, 2007; Lemhöfer et al., 2008; Morales et al., 2016; Paolieri et al., 2010; see Sá-Leite et al., 2020 for a detailed discussion). The effect of gender congruency between L1 and L2 has thus far only been documented in the processing of isolated nouns and noun phrases. Therefore, it is still an open question whether it would also emerge under more ecologically valid conditions, that is, in the processing of nouns embedded in grammatical sentences. Furthermore, although the effect of gender congruency has been replicated in many studies, its nature still remains a matter of controversy due to various L2 learner characteristics and language-specific idiosyncrasies. For example, languages vary with respect to gender transparency, that is, the extent to which the gender of a noun can be predicted by its morpho-phonological properties. Gender transparency affects the way gender is processed in L2 (e.g., Bordag & Pechmann, 2007; Caffarra et al., 2017), which, in turn, may modulate the size of lexical gender congruency effects. Although evidence exists that the effect of L1 on gender processing in L2 is more easily obtained in languages that are more transparent overall, such as Spanish (e.g., Sá-Leite et al., 2020), the role of individual gender cues, such as noun termination, in modulating the cross-language effects of gender is far from understood. Finally, although L2 proficiency is recognised as an important modulator of cross-language activation in general (e.g., Schwartz, 2015), its role in L2 gender processing remains understudied. Until very recently, L2 proficiency had never been included as a factor in the analyses of any study on gender congruency effects (Sá-Leite et al., 2023).

To shed light on these issues, our study addresses the question of whether the gender congruency effect also emerges when reading nouns embedded in grammatical sentences in L2, and whether it is modulated by the transparency of individual nouns and L2 proficiency. We report a self-paced reading experiment with late unbalanced Polish–German bilinguals. The results reveal a complex interplay between the factors in question.

The effect of gender congruency in bilingual processing and its modulators

Interactions between gender systems are typically captured through the facilitated processing of nouns that share gender between L1 and L2 relative to nouns that have a different gender in L1 and L2, a finding referred to as the ‘gender congruency effect’ (e.g., Sá-Leite et al., 2020). Until now, most evidence for gender congruency effects has been obtained in production studies (e.g., Bordag & Pechmann, 2007; Lemhöfer et al., 2008; Paolieri et al., 2010; Sá-Leite et al., 2023). In turn, gender congruency effects in bilingual comprehension have been addressed in only a few studies, using primed lexical decisions (Lemhöfer et al., 2008), translation recognition (Paolieri et al., 2020), gender decisions (Długosz, 2023), and visual-world eye tracking (Morales et al., 2016).

Languages in which the nouns’ gender can be derived from their morpho-phonological properties are claimed to have transparent gender assignment systems. For example, Spanish exhibits a great degree of transparency because the majority of nouns have a suffix that is a reliable gender cue (‘-o’ for masculine and ‘-a’ for feminine). Importantly, the presence of gender cues has been found to affect the way gender is processed during L2 comprehension (e.g., Bordag et al., 2006; Bordag & Pechmann, 2007; Caffarra et al., 2017). The processing of transparent nouns is typically facilitated compared to non-transparent nouns, which is captured through shorter response latencies and lower error rates across different types of tasks.

To account for the differences in the processing of transparent versus non-transparent nouns, Gollan and Frost (2001) proposed a model of language comprehension according to which gender information can be accessed via a lexical, and a form-based route. The two routes can be employed whenever there is a need to retrieve the gender of a noun. The lexical route derives gender as an abstract representation from the mental lexicon, whereas the form-based route derives gender from its correlation with morpho-phonological cues (see also Caffarra et al., 2014, 2017). In their recent proposal on gender selection in language comprehension, Sá-Leite and Lago (2024) argue that the employment of the form-based route depends on the reliability of individual gender cues. Reliable gender cues, such as '-o' and '-a' in Spanish, are claimed to be maintained at a higher level of activation than gender cues that are less reliable and less frequent, which affects the process of gender selection.

Importantly to our study, cross-language interactions at the level of gender vary as a function of the overall degree of transparency of a given language (see Sá-Leite et al., 2020 for discussion). More precisely, gender congruency effects seem to be more easily obtained in languages with transparent gender assignment systems. As Sá-Leite et al. (2020, p. 684) put it, the availability of gender cues increases the level of activation of gender information. The higher its activation, the greater the competition between L2 and L1 translation equivalents, and hence the size of gender congruency effects. Although gender congruency effects have also been observed in languages with less transparent gender assignment systems, such as German (Bordag & Pechmann, 2008) and Dutch (Lemhöfer et al., 2008), they seem to be somewhat weaker, because they are absent in the analysis of variance by items (Lemhöfer et al., 2008), or emerge in translation but not in picture naming (Bordag & Pechmann, 2008).

Apart from treating gender transparency globally, a more detailed perspective has also been adopted. Since transparent languages also have nouns whose gender cannot be predicted by their form, and, at the same time, even non-transparent languages can exhibit some gender cues, the effects of gender transparency have also been analysed with respect to morpho-phonological regularities within a language itself. For example, Bordag and Pechmann (2007) studied the role of gender cues in German, a language with a much less transparent gender assignment system than Romance languages. L1 Czech learners of L2 German completed a picture naming task in which the gender transparency of target nouns was manipulated. Participants were slower and made more errors with non-transparent nouns than with transparent nouns. Although the authors were also interested in gender congruency effects, they did not analyse them in interaction with transparency. We believe that even in less transparent languages, such as German, the level of activation of gender values can be higher for transparent nouns than for non-transparent nouns, even if the reliability of a given gender cue is variable. Therefore, gender transparency might also interact with gender congruency within one and the same language. To our knowledge, such an interaction has not yet been explored with respect to language comprehension.

Although L2 proficiency is acknowledged as an important modulator of cross-language interactions (e.g., Schwartz, 2015), its role in shaping the degree of gender congruency effects has not been systematically studied. Nevertheless, there are indications that gender congruency effects are stronger at lower L2 proficiency levels (Bordag & Pechmann, 2007; Paolieri et al., 2010). In addition, Sá-Leite et al. (2023) have provided evidence that the size of gender congruency effects may depend on language balance. In forward and backward translation tasks with early and late Portuguese–German bilinguals, the authors found a gender congruency effect in Portuguese that was stronger with increasing language imbalance, operationalised as the difference in proficiency between German and Portuguese.

Research on the impact of L2 proficiency on cross-language activation beyond gender suggests that when the target language is the weaker language, greater proficiency in that language

results in decreased effects of cross-language activation (Schwartz, 2015, p. 345). Therefore, greater cross-language activation in L2 is expected at lower L2 proficiency levels. Indeed, this is predicted by most models of bilingual lexical processing and acquisition (e.g., Dijkstra et al., 2019; Ecke, 2015; Grainger et al., 2010). For example, the Parasitic Model of vocabulary acquisition assumes that learners unconsciously detect and use similarities between new and already known word forms:

> Once learners detect a new word's similarity in (phonological or orthographic) form and/or meaning with an already known item, they will (by default) assume equivalence and create connections between the two forms using the stable one as a host representation to anchor the new form in the lexicon. (Ecke, 2022, p. 234)

Importantly, establishing connections between assumed formal equivalents leads to the adoption of the so-called syntactic frame of the host's lexical entry that includes gender information. During the initial stages of learning, many instances of cross-linguistic influence are predicted because new lexical items are parasitic upon the established L1 hosts until they receive sufficient activation to detach from the host.

More globally, the presence of gender congruency effects supports the assumption of non-selective lexical access in bilinguals. The non-selectivity of lexical access in L2 has also been demonstrated in sentence context, which is discussed in the following section.

Lexical co-activation in sentence context

As the large body of evidence on the non-selectivity of lexical access has been based on word recognition in isolation, which does not correspond to the way language is usually processed, a number of studies were interested in whether the findings could be extended to more ecologically valid conditions, namely sentential contexts (see Lauro & Schwartz, 2017 for review). Studies exploring cross-language activation of lexical items embedded in sentence contexts predominantly report cognate facilitation effects, that is, facilitated processing of words that share form and meaning between L1 and L2, as compared to non-cognates that lack such overlap (e.g., Bultena et al., 2014; Duyck et al., 2007), and interlingual homograph interference effects, that is, an interaction between words that share form in L1 and L2, but differ in meaning (e.g., Elston-Güttler et al., 2005), or both (e.g., Schwartz & Kroll, 2006).

Beyond yielding more generalisable results, research investigating bilingual word recognition in context has shown that the effects obtained can be modulated both by language and by semantic cues. In terms of the former, the effects are mostly shown among bilinguals reading in their L2, but non-selectivity of lexical access has also been reported in studies performed in L1 (e.g., Van Assche et al., 2009) and in mixed-language conditions (e.g., Dijkstra et al., 2015). Although language cues on their own are not sufficient to suppress the non-target language and induce selectivity, the observed effects can vary in size, being smaller when observed in L1, but more robust in mixed conditions (e.g., Dijkstra et al., 2015). As for the semantic cues provided by the preceding context, a recurring finding is that the sole presentation of a word embedded in a sentence, where the target is plausible, but not anticipated (i.e., a low-constraint sentence) does not largely modulate the observation of cross-language interactions. Semantically high-constraint sentences, however, can constrain lexical access and tend to diminish the effects (e.g., Schwartz & Kroll, 2006).

When it comes to gender congruency effects in sentence context, they were addressed in a recent study by Klassen et al. (2022), although from a different perspective than ours. The authors were interested in the interplay between lexical (gender) and syntactic (word order) overlap between L1

and L2. They carried out a self-paced reading experiment with L1 Spanish L2 German learners, in which L1-L2 lexical gender congruency and adjective–noun word order were manipulated. An interaction between lexical gender congruency and grammaticality was found, in that participants were only sensitive to word order violations in sentences with gender-congruent nouns. However, Klassen et al. (2022) did not find a main effect of lexical gender congruency, which means that participants did not read the gender-congruent nouns faster.

Research questions

The main goal of our study is to investigate whether the gender congruency effect observed in previous studies on the processing of isolated nouns and noun phrases extends to sentence contexts and whether it is modulated by gender transparency and L2 proficiency. We focus on the Polish–German language pair, in which the gender systems are symmetrical. Despite being typologically distant languages, both Polish and German have a three-way gender system with masculine, feminine, and neuter. They also exhibit a similar distribution of the three gender values, with masculine gender being the most frequent one, followed by feminine, and the least frequent neuter (e.g., Bauch, 1971; Stefańczyk, 2007). While the Polish gender system is highly transparent (e.g., Stefańczyk, 2007), the transparency of the German gender system is claimed to be rather low (e.g., Kupisch et al., 2022). The reliability of gender cues in German is highly variable. Most of them are probabilistic, and only some of them are deterministic in nature (e.g., the suffix '-heit', *Freiheit*$_{F}$; Hohlfeld, 2006). Gender is expressed overtly in phrases with singular definite articles, which clearly disambiguate between the three genders ('der$_{M}$', 'die$_{F}$', 'das$_{N}$' in the nominative case and 'den$_{M}$', 'die$_{F}$', 'das$_{N}$' in the accusative case). Indefinite articles for masculine and neuter genders are identical in the nominative case ('ein', as opposed to 'eine' for feminine gender), but not in the accusative case ('einen' for masculine gender and 'ein' for neuter gender).

Attempting to elucidate the nature of cross-language gender effects in sentence context by means of the self-paced reading paradigm, we address the following research questions:

1. Does the gender congruency effect in L2 extend to noun recognition in grammatical sentences?
2. Is the gender congruency effect, if present, modulated by gender transparency?
3. Is the gender congruency effect, if present, modulated by L2 proficiency?

The data and scripts used in this study are available online at the following link: https://osf.io/396hw/.

Method

Participants

The original sample consisted of 66 late unbalanced Polish–German bilinguals. Six participants had to be excluded because they reported significant knowledge of other languages with grammatical gender (Ukrainian, Spanish, or French). Thus, the final sample comprised 60 participants (50 female, $M_{age} = 19.6$ years, $SD = 1.4$). All of them were students of Applied Linguistics at a Polish university. They started learning L2 German after the age of 6 in school settings in Poland and then continued at university ($M_{AOA\ of\ L2\ German} = 11.4$ years, $SD = 3.5$). At the time of testing, participants' L2 German proficiency ranged from intermediate to advanced level.

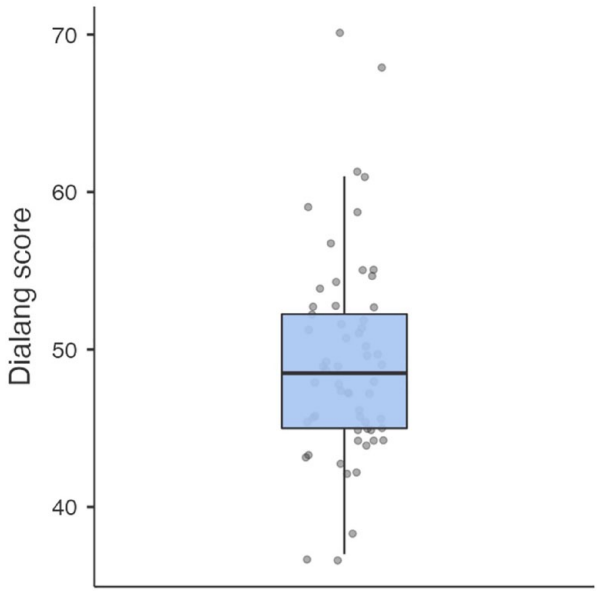


Figure 1. Participants' scores on the German placement test, Dialang.

Participants constituted a homogeneous group because they were taught by the same teachers and were using the same learning materials. The participants' background information was gathered via a short questionnaire. All participants gave informed consent and received course credits for their participation.

Proficiency

Proficiency in L2 German was measured through the German placement test, Dialang (Alderson, 2005). The test consists of 50 real words and 25 pseudo-words that are presented as a list and need to be classified as either existing or non-existing German words. We used an adaptation in which the items are presented on the screen one at a time (see Di Pisa et al., 2022). Participants were asked to press the key 'J' (=*ja*, 'yes') if they thought the word existed or 'N' (=*nein*, 'no') if they thought the word did not exist. The test score was the sum of all correct answers. The maximum possible score was thus 75. Participants scores ranged from 37 to 70 ($M = 49.3$, $SD = 6.4$). Figure 1 presents participants' scores.

Materials

The self-paced reading task included 32 grammatical, low-constraint sentences (see the OSF repository). We opted for low-constraint sentences because previous research has shown that sentence context can eliminate cross-language activation if sufficient semantic cues are provided to suppress the non-response language (e.g., Lauro & Schwartz, 2017). Sentences always contained (1) the subject 'ich', (2) the auxiliary 'haben', (3) a definite article in the accusative case, (4) a trigger noun, (5) a prepositional phrase, and (6) a past participle (see Example 1).

(1)	*Ich habe das Angebot ohne Zögern angenommen.* I have thеN offerN without hesitation accepted 'I have accepted the offer without hesitation'.

All the sentences were created using the perfect tense. This enabled us to eliminate the influence of context, because the semantic information conveyed by the verb was only available to participants at the end of a sentence.

We manipulated the gender congruency and gender transparency of the trigger nouns. We crossed these factors, which resulted in four conditions: congruent and transparent, congruent and non-transparent, incongruent and transparent, and incongruent and non-transparent. Congruent nouns always had the same gender in German and Polish, whereas incongruent nouns had a different gender in German than in Polish. For transparent nouns, we decided to use nouns ending in a schwa '-e', which is a cue of feminine gender (*Tasche*, 'bag'). The choice was motivated by two considerations. First, L2 learners strongly rely on this cue when assigning gender to German nouns (e.g., Walter & MacWhinney, 2015). Second, this cue is frequent in the input, as L2 learners encounter a large number of nouns ending in a schwa from the very beginning. For non-transparent nouns, we opted for nouns ending in a consonant (we did not use nouns with informative endings, such as '-er' or '-el'). The rationale behind this was that the consonantal ending is associated with the masculine (*Tisch*, 'table'), neuter (*Kreuz*, 'cross'), and even feminine (*Milch*, 'milk') gender in German, which means that it does not provide a reliable cue to any of the three genders. Importantly, the contrast between the schwa and the consonantal ending has produced the intended transparency effect in previous studies (Bordag et al., 2006; Bordag & Pechmann, 2007).

Half of the non-transparent nouns were masculine and half neuter. All transparent nouns were feminine. In order to balance the number of nouns of a given gender, we added 16 sentences, 8 with masculine and 8 with neuter nouns. They followed the same syntactic structures as the target items with the exception of the subject. We used the first-person plural pronoun 'wir' in lieu of the first-person singular pronoun 'ich' to minimise repetition effects.

Each condition included eight nouns that were one to three syllables long. We controlled for their length in letters, logarithmic frequency in German according to SUBTLEX-DE (Brysbaert et al., 2011), concreteness (Köper & Schulte im Walde, 2016), and formal similarity to Polish based on the normalised Levenshtein distance (NLD) as taken from the NIM database (Guasch et al., 2013). Since participants also had knowledge of English, we controlled for formal similarity to English, again using the NLD. Table 1 presents the characteristics of the trigger nouns in each of the four conditions.

The Kruskal–Wallis H tests did not reveal any differences between the four conditions with respect to length in letters, $H(3)=0.27$, $p=.965$; logarithmic frequency in German, $H(3)=0.09$, $p=.992$; concreteness, $H(3)=6.01$, $p=.111$; NLD (Polish), $H(3)=0.85$, $p=.836$; and NLD (English), $H(3)=2.27$, $p=.516$.

In addition, we avoided German nouns with more than one Polish translation, nouns that could also be verbs (*Essen*, 'food/to eat'), nouns that were emotionally loaded, and polysemic nouns.

Procedure

Participants read the items for comprehension in a linear, non-cumulative procedure with word-by-word segmentation. They used the space key to proceed through each segment at their own pace. To ensure that participants were paying attention, a binary yes/no comprehension question targeting the prepositional phrase was presented after half of the items. Participants responded by

Table 1. Characteristics of the trigger nouns.

Condition	Length (no. of letters)	Logarithmic frequency	Concreteness	NLD (Polish)	NLD (English)
Congruent transparent	5.25 (1.04)	2.89 (0.46)	7.58 (0.46)	0.16 (0.17)	0.35 (0.35)
Congruent non-transparent	5.50 (2.20)	2.85 (0.39)	6.65 (0.98)	0.15 (0.12)	0.28 (0.35)
Incongruent transparent	5.25 (0.46)	2.86 (0.73)	5.87 (1.91)	0.10 (0.12)	0.20 (0.15)
Incongruent non-transparent	5.50 (1.31)	2.81 (0.18)	6.91 (1.78)	0.16 (0.20)	0.13 (0.19)

pressing the keys 'j' ('yes') or 'n' ('no') on their keyboard. The task began with written instructions in German and a practice block. Participants were informed that the task targeted reading comprehension in German and no feedback was provided during the experiment. The order of the items was fully randomised.

In order to control for participants' lexical knowledge of gender in German, we conducted an untimed gender assignment task. This was important because any effect of gender congruency can only emerge if target-like lexical gender knowledge is in place. Participants were presented with all the trigger nouns used in the self-paced reading task one at a time and were asked to indicate their gender by choosing one of the three definite articles: 'der', 'die', or 'das'. Participants could also choose the option 'I don't know'.

All tasks were hosted on PsyToolkit – web-based software for programming and running reaction-time experiments (Stoet, 2010, 2017). Each participant completed all the tasks in a single testing session, including the self-paced reading task, the German placement test, Dialang, the control gender assignment task, and the background questionnaire, in that order.

Results

Control gender assignment task

Participants did not assign gender to 70 trigger nouns (3.6%). They achieved an overall accuracy of 87.5%. To gain insights into participants' performance on the control task, we analysed their accuracy using the statistical software R (R Core Team, 2024) with lmerTest (Kuznetsova et al., 2017) and lme4 (Bates et al., 2015) packages. Gender assignment accuracy was entered into a generalised linear mixed-effects model with a binomial distribution and logit link function. The model was fitted by maximum likelihood with the Laplace approximation technique. As fixed effects, the model included congruency, transparency, and L2 proficiency. Congruency and transparency were coded using sum contrast coding (–0.5, +0.5). L2 proficiency was centred and transformed into *z*-scores. Participant and item were entered as random effects (intercepts). The model output is available in the OSF repository.

The analysis showed significant effects of congruency, $estimate = -0.85$, $SE = 0.25$, $z = -2.00$, $p = .046$; transparency, $estimate = 2.02$, $SE = 0.43$, $z = 4.67$, $p < .001$; and L2 proficiency, $estimate = 0.42$, $SE = 0.13$, $z = 3.12$, $p = .002$. Participants performed better with congruent (96.2%, $SE = 0.01$) than incongruent nouns (91.5%, $SE = 0.02$). Accuracy was also higher with transparent (97.8%, $SE = 0.01$) than non-transparent nouns (85.7%, $SE = 0.04$). Finally, accuracy improved with advancing L2 proficiency.

The results of the control task thus show that congruency and transparency have a significant positive effect on participants' gender assignment accuracy. We take this as evidence that our manipulation of the trigger nouns with respect to these two variables was successful.

Table 2. Mean reading times and standard errors.

Congruency	Transparency	Mean RTs	*SE*
Congruent	Non-transparent	705	20.5
Incongruent	Non-transparent	657	18.9
Congruent	Transparent	586	14.0
Incongruent	Transparent	685	20.1

Self-paced reading task

Accuracy on comprehension questions was always higher than 80% ($M = 92.6\%$, $SD = 0.05$), meaning that participants paid attention when reading the target items. Therefore, no participant was excluded. To clean the reading time data on the trigger noun, we applied the absolute cutoff method (Jegerski, 2014). We excluded all reading times below 150 ms and above 3,000 ms from the analysis (0.8%). Furthermore, we removed all items with unknown or unassigned gender based on the results of the control gender assignment task (15.6%). Table 2 presents mean reading times and standard errors.

To analyse the reading times on the trigger noun, we employed linear mixed-effects modelling using the statistical software R (R Core Team, 2024) with lmerTest (Kuznetsova et al., 2017) and lme4 (Bates et al., 2015) packages. We log-transformed the reading times to adjust for the skewness of their distribution. As fixed effects, the model included congruency, transparency, L2 proficiency, and their interactions. Congruency and transparency were coded using sum contrast coding (–0.5, +0.5). L2 proficiency was centred and transformed into *z*-scores. The maximal model that converged and did not return a singular fit (Barr et al., 2013) included random intercepts for participant and item as well as a by-item random slope for L2 proficiency. The emmeans (Lenth, 2024) package was used to obtain pairwise contrasts, with *p* values adjusted using the false discovery rate (FDR) method. The model output is available in the OSF repository.

The results showed non-significant effects of congruency, $estimate = 0.04$, $SE = 0.03$, $t = 1.32$, $p = .199$, and L2 proficiency, $estimate = -0.01$, $SE = 0.04$, $t = -0.22$, $p = .829$, as well as non-significant interactions between congruency and L2 proficiency, $estimate = -0.00$, $SE = 0.02$, $t = -0.06$, $p = .957$, and between transparency and L2 proficiency, $estimate = -0.00$, $SE = 0.02$, $t = -0.14$, $p = .887$. The effect of transparency was significant, $estimate = -0.07$, $SE = 0.03$, $t = -2.31$, $p = .028$, with faster reading times for transparent than non-transparent nouns. The interaction between congruency and transparency was significant, $estimate = 0.15$, $SE = 0.06$, $t = 2.51$, $p = .018$, so was the interaction between congruency, transparency, and L2 proficiency, $estimate = -0.09$, $SE = 0.04$, $t = -2.39$, $p = .024$.

The pairwise contrasts for the interaction between congruency and transparency revealed that participants read congruent transparent nouns faster than congruent non-transparent nouns, $estimate = 0.15$, $SE = 0.04$, $t = 3.42$, $p = .012$; incongruent transparent nouns, $estimate = -0.11$, $SE = 0.04$, $t = -2.73$, $p = .033$; and incongruent non-transparent nouns, $estimate = 0.11$, $SE = 0.04$, $t = 2.55$, $p = .033$. Differences between the other conditions were not significant (all $ps > .500$). The interaction between congruency and transparency is illustrated in Figure 2.

The pairwise contrasts for the interaction between congruency, transparency, and L2 proficiency (at three levels: mean – 1 *SD*, mean, mean + 1 *SD*) revealed that L2 proficiency modulated the effect of congruency for transparent nouns. With L2 proficiency at the level of 1 *SD* below the mean, the difference between congruent and incongruent transparent nouns was significant,

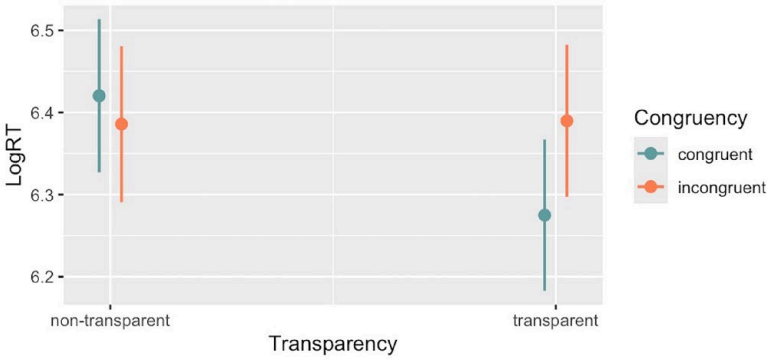


Figure 2. Plot of interaction between congruency and transparency.

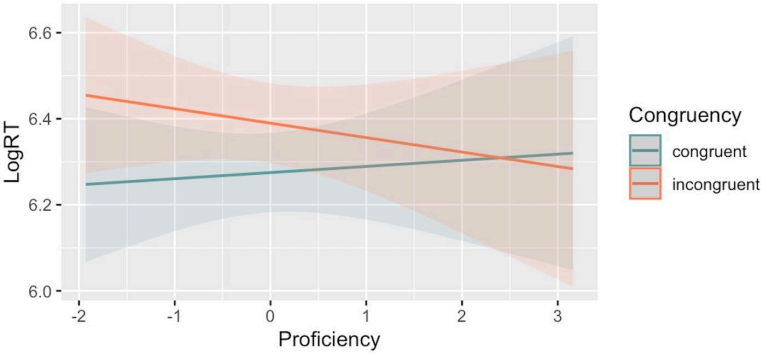


Figure 3. Plot of interaction between congruency and L2 proficiency for transparent nouns.

estimate = −0.16, *SE* = 0.05, *t* = −3.18, *p* = .004. The difference between congruent and incongruent transparent nouns was also significant, with L2 proficiency at mean level, *estimate* = −0.11, *SE* = 0.04, *t* = −2.73, *p* = .011, but disappeared with L2 proficiency at the level of 1 *SD* above the mean, *estimate* = −0.07, *SE* = 0.05, *t* = −1.38, *p* = .178. This indicates that participants read congruent transparent nouns faster than incongruent transparent nouns, but the effect of congruency decreased with advancing L2 proficiency (see Figure 3). A summary of the post hoc analyses is available in the OSF repository.

Discussion

The results showed that participants read gender-congruent nouns in their L2 German faster than gender-incongruent nouns, but only if they were transparent. For non-transparent nouns, no gender congruency effect was observed. Therefore, gender congruency did have a measurable effect on the recognition of words embedded in grammatical sentences, but it was constrained by the presence of noun-internal gender cues.

Why would the gender congruency effect be limited to transparent nouns? Following the argument put forward by Sá-Leite et al. (2023), the degree of activation of gender values should be inherently low in languages with non-transparent gender assignment systems. This would lead to

slippery gender congruency effects. In turn, the activation level of gender values should be increased in languages with transparent gender assignment systems, which would make gender competition, and hence gender congruency effects more easily observed. The results of the study can be explained by this argument. The availability of gender cues amplified the level of activation of gender values in L2 German, which, in turn, gave rise to stronger competition between L2 and L1 translation equivalents and hence to a gender congruency effect. However, our results indicate that it is not the global transparency of a given language that determines the degree of activation of gender values, but rather the transparency of individual nouns in that language. This is in line with the recent proposal by Sá-Leite and Lago (2024). Therefore, the slippery gender congruency effects observed in previous studies involving non-transparent languages (e.g., Bordag & Pechmann, 2008; Lemhöfer et al., 2008) could have resulted from the fact that the factor of noun-internal gender transparency was not included in their design.

As a matter of fact, the activation of gender in our study could have been increased in yet another way. Since the trigger nouns in the self-paced reading task were preceded by a gender-marked definite article, their gender could always be anticipated. Unfortunately, we are not able to assess to what extent the presence of an agreement context contributed to the observed gender congruency effect. Note, however, that nouns are typically accompanied by articles in German sentences.

The lack of a gender congruency effect for non-transparent nouns could also be accounted for in terms of task-specific factors. Most of the previous studies that found a gender congruency effect employed picture naming or translation. In these tasks, processing proceeds top-down, from semantics to form, since the target word is not available (in contrast with a self-paced reading task, for example). This consumes more time and, therefore, leaves more space for the activation of non-target language to be observed (Lauro & Schwartz, 2017, p. 220).

In addition, response-time evidence in favour of gender congruency effects in L2 comprehension only comes from two studies. In the study by Paolieri et al. (2020), Catalan–Spanish bilinguals recognised translation pairs faster in Spanish if they were gender congruent than if they were not. However, the majority of nouns used in the experiment were transparent, and, more generally, Spanish has a highly transparent gender assignment system. In the study by Lemhöfer et al. (2008), a gender congruency effect was found in a primed lexical decision task with German–Dutch bilinguals, but it was limited to cognates. Unlike Spanish, Dutch is rather a non-transparent language. Therefore, it seems that a higher level of cross-language activation is required for L1 effects to emerge during L2 gender retrieval in the absence of gender cues. In the study by Lemhöfer et al. (2008), cross-language activation was increased by the inclusion of L2 nouns with form-similar translations in L1.

Interestingly, we found that the interaction between gender congruency and transparency was modulated by L2 proficiency. The gender congruency effect observed for transparent nouns was largest at low L2 proficiency levels and decreased in size with advancing L2 proficiency. This finding confirms previous studies that suggested a possible negative relationship between the magnitude of gender congruency effects and L2 proficiency (Bordag & Pechmann, 2007; Paolieri et al., 2010). It is also in line with most models of bilingual lexical processing and acquisition (e.g., Dijkstra et al., 2019; Ecke, 2015; Grainger et al., 2010). Specifically, our data can be accommodated by the Parasitic Model of vocabulary acquisition (e.g., Ecke, 2015). Weaker gender congruency effects in more proficient L2 learners follow from the assumption that parasitic connections between new lexical items in L2 and the established L1 hosts diminish as L2 proficiency increases. Since more proficient L2 learners are more balanced bilinguals, the results are also in line with the study by Sá-Leite et al. (2023). More generally, our study

provides additional evidence that greater L2 proficiency is linked to decreased effects of cross-language activation (e.g., Schwartz, 2015).

Our findings add further strength to the growing body of evidence that the two gender systems of a bilingual interact in L2 comprehension. We show that the gender congruency effect is not limited to the processing of isolated nouns and extends to sentence contexts, similar to the cognate facilitation effect (e.g., Bultena et al., 2014; Duyck et al., 2007). Importantly, we provide the first evidence for gender congruency effects in L2 based on the Polish–German language pair. Therefore, we increase the range of language pairs in which the effects emerge, thus providing evidence for their universal character.

Our study is not free of limitations. First of all, we only included one gender cue and hence one gender value in the transparent condition. Therefore, we were not able to disentangle the effect of transparency from the potential effect of gender. Although the schwa-ending is a strong cue to feminine gender (Bordag & Pechmann, 2007; Walter & MacWhinney, 2015), more research is needed to find out whether the effect observed in this study is specific to this cue, or generalises to other gender cues in German. Second, our study was limited to low-constraint sentence contexts, which limits the generalisation of the findings. Future studies should explore whether the presence of semantic and pragmatic cues in a sentence could restrict gender retrieval to L2. Finally, although the self-paced reading paradigm has been successfully used to explore lexical co-activation in sentence context (e.g., Bultena et al., 2014), it would be desirable to complement its findings with more sensitive measures, such as eye-tracking, to mitigate the influence of task-specific factors on gender processing in L2.

The study contributes to our understanding of bilingual gender processing in numerous ways. It is the first to demonstrate a gender congruency effect in the recognition of nouns embedded in grammatical sentences. We consider this an important contribution, because bilinguals typically read words embedded in meaningful sentences. Our study assesses the generalisability of previous studies on isolated word recognition and constitutes a step towards more ecologically valid research on cross-language effects of gender. Furthermore, we shed new light on the relationship between gender congruency and gender transparency in L2. More specifically, we show that the effects of gender congruency depend not only on global transparency, but also on the transparency of individual nouns, even in a language that is considered non-transparent in global terms, such as German. By showing that the activation of gender information in L1 during L2 processing depends on individual gender cues in L2, we open up new avenues for research on bilingual gender processing.

Data availability statement

The materials, scripts, and data that support the findings of this study are openly available in OSF at https://osf.io/396hw/.

Declaration of conflicting interests

The author(s) declared no potential conflicts of interest with respect to the research, authorship, and/or publication of this article.

Funding

The author(s) disclosed receipt of the following financial support for the research, authorship, and/or publication of this article: This work was supported by the National Science Centre, Poland (Narodowe Centrum Nauki) (grant number 2023/49/B/HS2/00173).

ORCID iD

Kamil Długosz https://orcid.org/0000-0002-9600-6452

Supplemental material

Supplemental material for this article is available online on OSF at: https://osf.io/396hw/.

References

Alderson, J. C. (2005). *Diagnosing foreign language proficiency: The interface between learning and assessment*. Continuum.

Bates, D., Mächler, M., Bolker, B., & Walker, S. (2015). Fitting linear mixed-effects models using lme4. *Journal of Statistical Software*, *67*(1), 1–48. https://doi.org/10.18637/jss.v067.i01

Bauch, H. J. (1971). Zum Informationsgehalt der Kategorie Genus im Deutschen, Englischen und Polnischen [On the information content of the gender category in German, English, and Polish]. *Wissenschaftliche Zeitschrift der Universität Rostock*, *20*(6), 411–419.

Barr, D. J., Levy, R., Scheepers, C., & Tily, H. J. (2013). Random effects structure for confirmatory hypothesis testing: Keep it maximal. *Journal of Memory and Language*, *68*(3), 255–278. https://doi.org/10.1016/j.jml.2012.11.001

Bordag, D., Opitz, A., & Pechmann, T. (2006). Gender processing in first and second languages: The role of noun termination. *Journal of Experimental Psychology: Learning, Memory, and Cognition*, *32*(5), 1090–1101. https://doi.org/10.1037/0278-7393.32.5.1090

Bordag, D., & Pechmann, T. (2007). Factors influencing L2 gender processing. *Bilingualism: Language and Cognition*, *10*(3), 299–314. https://doi.org/10.1017/S1366728907003082

Bordag, D., & Pechmann, T. (2008). Grammatical gender in translation. *Second Language Research*, *24*(2), 139–166.

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

Bultena, S., Dijkstra, T., & van Hell, J. G. (2014). Cognate effects in sentence context depend on word class, L2 proficiency, and task. *Quarterly Journal of Experimental Psychology*, *67*(6), 1214–1241. https://doi.org/10.1080/17470218.2013.853090

Caffarra, S., Barber, H., Molinaro, N., & Carreiras, M. (2017). When the end matters: Influence of gender cues during agreement computation in bilinguals. *Language, Cognition and Neuroscience*, *32*(9), 1069–1085. https://doi.org/10.1080/23273798.2017.1283426

Caffarra, S., Janssen, N., & Barber, H. A. (2014). Two sides of gender: ERP evidence for the presence of two routes during gender agreement processing. *Neuropsychologia*, *63*, 124–134. https://doi.org/10.1016/j.neuropsychologia.2014.08.016

Dijkstra, T., van Hell, J. G., & Brenders, P. (2015). Sentence context effects in bilingual word recognition: Cognate status, sentence language, and semantic constraint. *Bilingualism: Language and Cognition*, *18*(4), 597–613.

Dijkstra, T., Wahl, A., Buytenhuijs, F., Van Halem, N., Al-Jibouri, Z., De Korte, M., & Rekké, S. (2019). Multilink: A computational model for bilingual word recognition and word translation. *Bilingualism: Language and Cognition*, *22*(4), 657–679. https://doi.org/10.1017/S1366728914000388

Di Pisa, G., Kubota, M., Rothman, J., & Marinis, T. (2022). Effects of markedness in gender processing in Italian as a heritage language: A speed accuracy tradeoff. *Frontiers in Psychology*, *13*, Article 965885. https://doi.org/10.3389/fpsyg.2022.965885

Długosz, K. (2023). The effect of L2 German on grammatical gender access in L1 Polish: Proficiency matters. *Linguistics Vanguard*, *9*(1), 87–97. https://doi.org/10.1515/lingvan-2022-0068

Duyck, W., Van Assche, E., Drieghe, D., & Hartsuiker, R. J. (2007). Visual word recognition by bilinguals in a sentence context: Evidence for nonselective lexical access. *Journal of Experimental Psychology: Learning, Memory, and Cognition*, *33*(4), 663–679. https://doi.org/10.1037/0278-7393.33.4.663

Ecke, P. (2015). Parasitic vocabulary acquisition, cross-linguistic influence, and lexical retrieval in multilinguals. *Bilingualism: Language and Cognition*, *18*(2), 145–162. https://doi.org/10.1017/S1366728913000722

Ecke, P. (2022). Grammatical gender and article use in beginning learners of German. In D. Ayoun (Ed.), *The acquisition of gender: Crosslinguistic perspectives* (pp. 127–156). John Benjamins.

Elston-Güttler, K. E., Gunter, T. C., & Kotz, S. A. (2005). Zooming into L2: Global language context and adjustment affect processing of interlingual homographs in sentences. *Cognitive Brain Research*, *25*(1), 57–70. https://doi.org/10.1016/j.cogbrainres.2005.04.007

Gollan, T. H., & Frost, R. (2001). Two routes to grammatical gender: Evidence from Hebrew. *Journal of Psycholinguistic Research*, *30*, 627–651. https://doi.org/10.1023/a:1014235223566

Grainger, J., Midgley, K., & Holcomb, P. J. (2010). Re-thinking the bilingual interactive-activation model from a developmental perspective (BIA-d). In M. Kail, & M. Hickmann (Eds.), *Language acquisition across linguistic and cognitive systems* (pp. 267–283). John Benjamins.

Guasch, M., Boada, R., Ferré, P., & Sánchez-Casas, R. (2013). NIM: A Web-based Swiss army knife to select stimuli for psycholinguistic studies. *Behavior Research Methods*, *45*(3), 765–771. https://doi.org/10.3758/s13428-012-0296-8

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

Jegerski, J. (2014). Self-paced reading. In J. Jegerski, & B. VanPatten (Eds.), *Research methods in second language psycholinguistics* (pp. 20–49). Routledge.

Klassen, R., Kolb, N., Hopp, H., & Westergaard, M. (2022). Interactions between lexical and syntactic L1-L2 overlap: Effects of gender congruency on L2 sentence processing in L1 Spanish-L2 German speakers. *Applied Psycholinguistics*, *43*(6), 1221–1256. https://doi.org/10.1017/S0142716422000236

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.

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

Kuznetsova, A., Brockhoff, P. B., & Christensen, R. H. B. (2017). lmerTest package: Tests in linear mixed effects models. *Journal of Statistical Software*, *82*(13), 1–26. https://doi.org/10.18637/jss.v082.i13

Lauro, J., & Schwartz, A. I. (2017). Bilingual non-selective lexical access in sentence contexts: A meta-analytic review. *Journal of Memory and Language*, *92*, 217–233. https://doi.org/10.1016/j.jml.2016.06.010

Lemhöfer, K., Spalek, K., & Schriefers, H. (2008). Cross-language effects of grammatical gender in bilingual word recognition and production. *Journal of Memory and Language*, *59*, 312–330. https://doi.org/10.1016/j.jml.2008.06.005

Lenth, R. (2024). *emmeans: Estimated marginal means, aka least-squares means* (R Package Version 1.10.1). https://CRAN.R-project.org/package=emmeans

Morales, L., Paolieri, D., Dussias, P. E., Valdés Kroff, J. R., Gerfen, C., & Bajo, T. (2016). The gender congruency effect during bilingual spoken-word recognition. *Bilingualism: Language and Cognition*, *19*(2), 294–310. https://doi.org/10.1017/S1366728915000176

Paolieri, D., Cubelli, R., Macizo, P., Bajo, T., Lotto, L., & Job, R. (2010). Grammatical gender processing in Italian and Spanish bilinguals. *Quarterly Journal of Experimental Psychology*, *63*(8), 1631–1645. https://doi.org/10.1080/17470210903511210

Paolieri, D., Demestre, J., Guasch, M., Bajo, T., & Ferré, P. (2020). The gender congruency effect in Catalan–Spanish bilinguals: Behavioral and electrophysiological evidence. *Bilingualism: Language and Cognition*, *23*(5), 1045–1055. https://doi.org/10.1017/S1366728920000073

R Core Team. (2024). *R: A language and environment for statistical computing*. https://www.R-project.org/

Sá-Leite, A. R., Flores, C., Eira, C., Haro, J., & Comesaña, M. (2023). Language balance rather than age of acquisition: A study on the cross-linguistic gender congruency effect in Portuguese–German bilinguals. *Bilingualism: Language and Cognition*, *26*(5), 1079–1092. https://doi.org/10.1017/S1366728923000378

Sá-Leite, A. R., & Lago, S. (2024). The role of word form in gender processing during lexical access: A theoretical review and novel proposal in language comprehension. *Psychonomic Bulletin & Review*. Advance online publication. https://doi.org/10.3758/s13423-023-02426-8

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

Schwartz, A. I. (2015). Bilingual lexical access during written sentence comprehension. In J. W. Schwieter (Ed.), *The Cambridge handbook of bilingual processing* (pp. 325–365). Cambridge University Press.

Schwartz, A. I., & Kroll, J. F. (2006). Bilingual lexical activation in sentence context. *Journal of Memory and Language*, *55*(2), 197–212. https://doi.org/10.1016/j.jml.2006.03.004

Stefańczyk, W. T. (2007). *Kategoria rodzaju i przypadka polskiego rzeczownika: próba synchronicznej analizy morfologicznej* [The category of gender and case of the Polish noun: An attempt at a synchronous morphological analysis]. Wydawnictwo Uniwersytetu Jagiellońskiego.

Stoet, G. (2010). PsyToolkit: A software package for programming psychological experiments using Linux. *Behavior Research Methods*, *42*, 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

Van Assche, E., Duyck, W., Hartsuiker, R. J., & Diependaele, K. (2009). Does bilingualism change native-language reading? Cognate effects in a sentence context. *Psychological Science*, *20*(8), 923–927. https://doi.org/10.1111/j.1467-9280.2009.02389.x

Walter, D., & MacWhinney, B. (2015). US German majors' knowledge of grammatical gender. *Die Unterrichtspraxis/Teaching German*, *48*(1), 25–40. https://doi.org/10.1111/tger.10179

Author biographies

Kamil Długosz is an Assistant Professor at Adam Mickiewicz University, Poznań, Poland. His research focuses on bi- and multilingual acquisition and processing.

Anna Olszewska in a PhD student at Adam Mickiewicz University, Poznań, Poland. She is currently investigating the role of regressive cross-linguistic influence in third language acquisition.