S. Maryam Fatemi

The influence of grammatical gender on the conceptualization of the world: A systematic literature review

Zusammenfassung

Der Einfluss des grammatikalischen Geschlechts auf die Konzeptualisierung der Welt. Eine systematische Literaturübersicht

Viele aktuelle Studien zu grammatikalischem Geschlecht und sprachlicher Relativität vernachlässigen das Neutrum und reduzieren das Thema auf binäre Geschlechter. Diese Studie unternimmt eine kritische Untersuchung der Forschung des vergangenen Jahrzehnts anhand des PRISMA-Protokolls. Die Ergebnisse zeigen einen deutlichen rückläufigen Trend in der Anzahl der durchgeführten Studien in den letzten vier Jahren und eine Fokussierung auf die Sprachen Deutsch und Englisch. Zudem werden Sprachen mit drei Genera oft ungenau dargestellt, indem das Neutrum ausgeschlossen wird, wodurch Repräsentativität begrenzt wird. Die Auswirkungen betreffen sowohl die Forschung zur sprachlichen Relativität als auch genderrelevante Forschungsbereiche. Die mangelnde Berücksichtigung des Neutrums und der rückläufige Trend in der Anzahl an Studien werfen Bedenken hinsichtlich der Vollständigkeit der aktuellen Forschungsansätze auf und unterstreichen die Notwendigkeit für inklusivere und verfeinerte Methoden in der Zukunft.

Schlüsselwörter
Genus, Linguistische Relativität, Binäres Geschlecht, PRISMA-Protokoll

Summary

Many recent studies on grammatical gender and linguistic relativity overlook the neuter, simplifying the issue to a binary gender system. This study critically and systematically reviews research from the past decade, following the PRISMA protocol. The findings reveal a significant downward trend in the number of studies conducted over the past four years and a bias toward German and English. Additionally, languages with three-gender systems are often inaccurately represented by excluding the neuter, limiting generalizability. These implications affect both research on linguistic relativity and gender-related disciplines. The lack of attention to neuter gender and the recent reduction in the number of studies raise concerns about the comprehensiveness of current research approaches and highlight the need for more inclusive and refined methodologies moving forward.

Keywords
grammatical gender, linguistic relativity, binary gender, PRISMA protocol

1 Introduction

There is a plethora of literature on gender in linguistics, including studies on grammatical gender (Boroditsky/Schmidt/Phillips 2003; Günthner/Hüpper/Spieß 2012; Hellinger/Bußmann 2002). It is commonly accepted that three grammatical genders exist: masculine, feminine, and neutral (Corbett 1991). However, most studies on gram-

 https://doi.org/10.3224/gender.v16i3.12

matical gender published within the last decade ignored the existence of the neutral gender. By doing so, researchers are in danger of reducing the richness of grammatical gender to binary male/female social and biological gender categories. Through a systematic literature review, the present study demonstrates that the above-mentioned binary treatment of grammatical gender is caused to a large extent by the methodologies of the reviewed studies.

In linguistic studies, the category *genus* refers to grammatical gender, which classifies nouns and pronouns as masculine, feminine, or neuter, while *sex* refers to biological sex (Kotthoff/Nübling 2018). Considering Ferdinand de Saussure's semiotic theory (Lagopoulos/Boklund-Lagopoulou 2020), the complexity of studying gender linguistically can be described thus: One challenge is the difficulty in clearly separating the signified[1] concepts associated with genus and sex. While signifiers are clear and functionally represented through pronouns and declensions, indicating genus markers, they also carry cultural and biological associations (signifieds). These connotations can influence the attributes assigned to grammatical genders, as seen with nouns like the sun (*die Sonne*) and the moon (*der Mond*), which reflect broader cultural meanings. This blending of sociocultural and biological connotations complicates the analysis and must be acknowledged as a significant aspect of linguistic studies on genus.

Therefore, gender (genus/sex) and language are closely linked in various ways studied across the disciplines of semiotics, pragmatics, and sociolinguistics (Hellinger 1985; Gorny 1995; Trömel-Plötz 1997; Hellinger/Bußmann 2002; Leaper/Ayres 2007; Newman et al. 2008; Palomares 2008; Günthner/Hüpper/Spieß 2012; Pusch 2014; Acke 2019). Linguistic studies on genus and sex begin with the observation that genus and sex are performed differently across various aspects and levels of communication. In daily social interactions, different characteristics and attributes are ascribed to different biological sexes, treated differently, and often follow stereotypical patterns (Günthner/Hüpper/Spieß 2012: 4).

Furthermore, studies extend beyond the above disciplines to the investigation of the influence of grammatical gender on conceptualization. This research provides evidence that genus plays a crucial role within language systems, shaping the perception of objects (Boroditsky/Schmidt/Phillips 2003; Cook 2016). Genus initially serves as a linguistic classification but gains contextual meaning through behavioral patterns, socialization, and cultural factors. This contextualization influences how speakers perceive and interact with linguistic gender categories, affecting cognitive frameworks and cultural interpretations (Boroditsky/Schmidt/Phillips 2003; Cook 2016). Since the emergence of the *linguistic relativity hypothesis* or *Sapir-Whorf hypothesis*, researchers have shifted their focus from the lexeme level to a meta-level of mental representations, exploring how grammatical categories like genus structure thought processes and influence cognitive scaffolds. The linguistic relativity theory goes back to the 19th-century linguistic-philosophical work of Wilhelm von Humboldt. Humboldt's views were further developed in 20th-century empirical research (Mertins 2018). Indeed, advocates of the linguistic relativity principle claim that the grammatical structure of the languages we speak influ-

1 Adapted from Ferdinand de Saussure's distinction between *signifier* (the form a sign takes) and *signified* (the concept it represents). In this context, grammatical gender is the signifier and biological gender is the signified.

ences the way we see, conceptualize, categorize, and speak about the world (Boroditsky/Schmidt/Phillips 2003; Lucy 1992; Mertins 2018; Slobin 1996; Whorf 2007 [1956]). Thus, genus as a grammatical category expressed differently among languages is a popular aspect in studying the impact of language on worldview.

Genus assigns nouns to specific classes, determined by both semantics and morphological and phonological principles (Alvanoudi 2015). Semantic principles involve direct mapping of biological gender/sex onto genus. For example, in Dravidian languages like Olari, nouns representing male humans are categorized as belonging to one class, and nouns representing female humans and other non-male entities belong to another group (Corbett 1991). On the contrary, morphological and phonological principles found in Russian and other Slavic languages extend beyond semantic distinctions. In Russian, genus is determined not only by semantics but also by morphological factors (Corbett 1991). That is, the genus of the words depends on the concluding consonants of the words (Fatemi 2023). However, genus assignment, though partly rule-based, is mostly arbitrary (Corbett 1991).

Languages form three groups based on this grammatical feature: *genderless languages* (e.g., Farsi, Finnish, Estonian), *natural gender languages* (e.g., English, Swedish, Norwegian), and *grammatical gender languages* (e.g., Spanish, German, French) (Prewitt-Freilino et al. 2011). Various studies have raised the question of whether such differences in the grammatical coding of nouns lead to language-specific conceptualization of the objects. For instance, the expression for *sun* being feminine in German (*die Sonne*), masculine in Spanish (*el sol*), and genderless in English could lead to different mental representations of the word for *sun* among German, Spanish, and English speakers, although all three languages refer to the same biologically sexless object. To this point, studies have found diverging results on the influence of this grammatical feature on our worldview. That is, the outcomes have shown that individuals speaking gendered languages may subconsciously perceive the world with feminine or masculine connotations (e.g., Bender/Beller/Klauer 2016a, 2016b; Kousta et al. 2008). While some studies did not find any influence of grammatical gender on conceptualization (Montefinese/Ambrosini/Roivainen2019), a closer examination reveals substantial flaws in their methodology.

This study demonstrates that while the influence of grammatical gender on cognition is recognized, the inconsistent findings as found in the existing research on the influence of genus on cognition are due to significant methodological limitations in those studies. These limitations include failing to account for all genera in languages such as German, particularly the neuter. The author argues that the neuter as a grammatical category (signifier) also carries meaning. For example, in Polish, dziewczę (a diminutive form of dziewczyna – a girl) is neutral (Maciuszek/Mateusz/Świątkowska 2019), and in German, das Mädchen (the girl) is also neutral. This shift in grammatical category can alter conceptualization by presenting diminutive forms through the neuter gender, rather than their expected or typical semantic associations with femininity. Another flaw is replicating previous research designs without accounting for their limitations. Such methodological shortcomings can lead to a defective understanding of the relationship between genus and cognition.

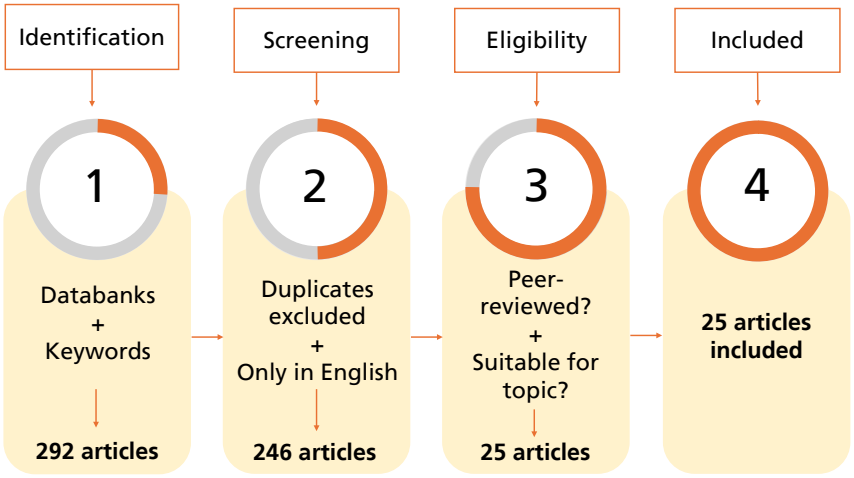
To demonstrate the flawed methodologies, the current study looks critically at 1) the examined languages linguistically in terms of the underlying genus system and

2) research designs used to show the relation between genus and mental representations of the objects. This knowledge can guide future studies to better answer the question of whether genus truly impacts cognition.

2 Method

The initial search was based on the four stages of identification, screening, eligibility, and inclusion outlined in the PRISMA guidelines (Preferred Reporting Items for Systematic Reviews and Meta-Analyses) (Moher et al. 2015). The keywords *genus and linguistic relativity, grammatical gender and Sapir-Whorf hypothesis, grammatical gender and cognition*, and *grammatical gender and conceptualization* for peer-review in English published between 2013 and 2023 yielded a total of 292 articles. Articles were selected from Google Scholar, Web of Science, Scopus, Eric, and Science Direct. 46 articles addressed the topic of sex from political or societal perspectives and were removed. After compiling the list of articles, the abstracts were analyzed to identify papers specifically examining the impact of gender on the conceptualization level.[2] This left the researcher with a total of 25 articles for the final analysis. Figure 1 shows the PRISMA guidelines for paper selection:

Figure 1: Flow diagram of searched, screened, and selected studies



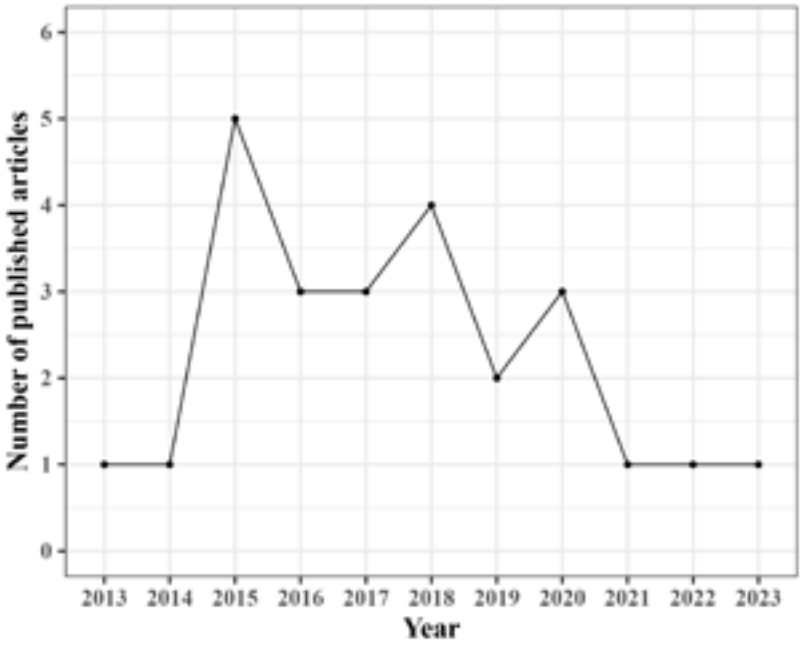
Source: Adapted from and inspired by Newmann et al. (2019).

2 In conducting the literature search, duplicate articles were identified and removed through a comprehensive process, ensuring that each unique study was accepted only once across the various databases.

3 Results and discussion

Examination of articles on linguistic relativity with a focus on genus published between 2013 and 2023 yielded the following distribution:

Figure 2: Distribution of published articles within the last ten years



Source: Data compiled from the articles included in this systematic review of research studies published between 2013 and 2023.

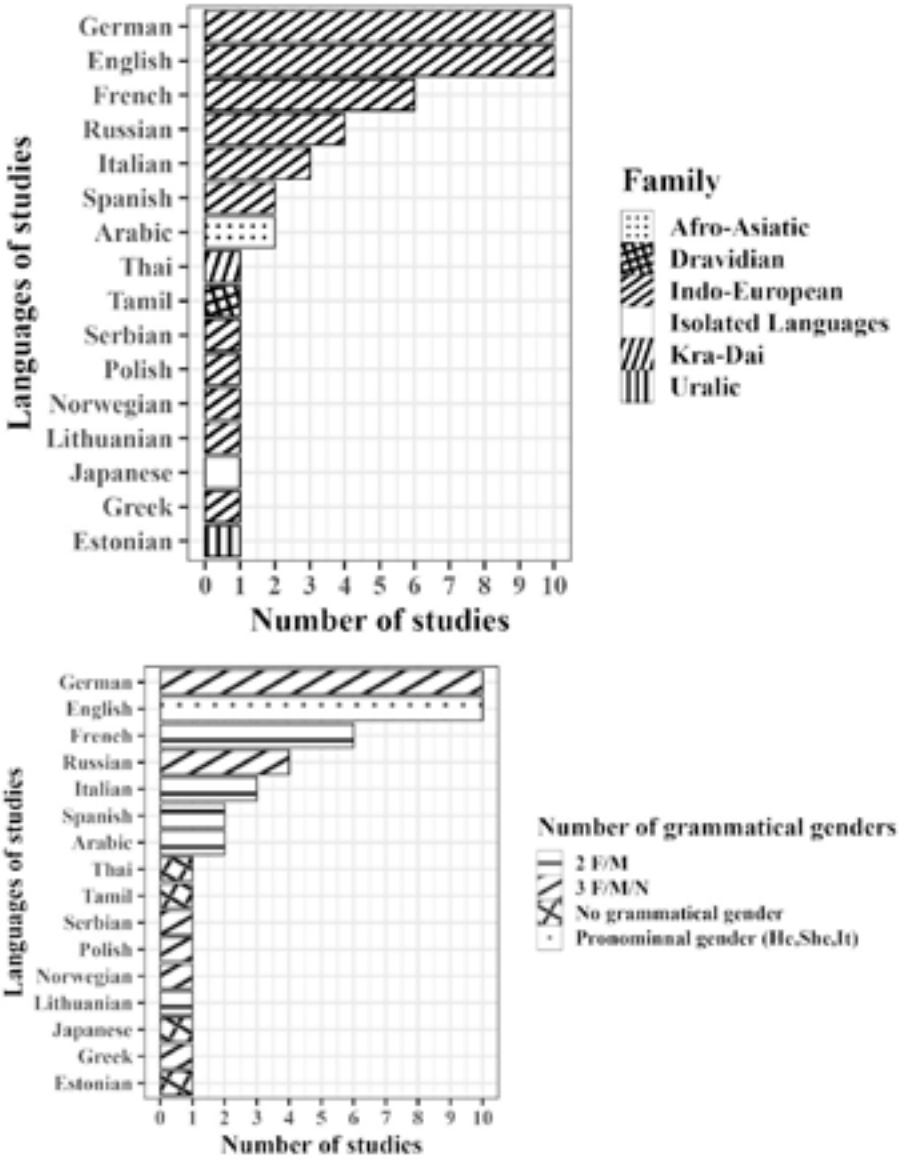
There was a notable increase in published articles in 2015, a decline in 2016–2017, and a resurgence in 2018. This pattern from 2015 to 2018 is also seen from 2018 to 2020, though with fewer publications. Overall, a downward trend from 2015 to 2023 indicates limitations in current methodologies for demonstrating the influence of genus on worldview. This trend, coupled with the continued use of ineffective or conflicting research designs, highlights the need for researchers to refine their approaches and avoid methodological repetition. Reassessing research designs is crucial for advancing understanding in this field. Key information from these 25 articles is summarized in Appendix A.

3.1 Which languages were examined?

This review identifies languages studied within the linguistic relativity framework over the past decade and examines their genus assignments. Understanding these variations is crucial for assessing the generalizability of findings on the conceptualization of reality under linguistic relativity and illustrating the variations in genus distinctions across lan-

guages. Figure 3 illustrates the languages studied within the linguistic relativity framework, their respective language families, the number of genus distinctions they exhibit, and the number of studies conducted on each language in the last decade.

Figure 3: Overview of languages of studies: Distribution by language families, article counts, and grammatical gender



Source: Data compiled from the articles included in this systematic review of research studies published between 2013 and 2023.

As shown in figure 3, these 25 studies cover only 16 languages. According to Hammarström/Haspelmath/Bank (2022), these languages—German, English, French, Italian, Spanish, Russian, Norwegian, Polish, Serbian, Lithuanian, Greek, Arabic, Tamil, Estonian, Thai, and Japanese—belong to just six of 100 language families, with limited representation from others. 15 are from Indo-European, Afro-Asiatic, Dravidian, Uralic, and Kra-Dai families, and Japanese is an isolate, each reflecting different genus systems (Fatemi 2023). Over the past decade, German and English have been the most studied Indo-European languages in this field, each examined in ten studies on the effect of gender on cognition (e.g., Imai et al. 2013; Bender/Beller/Klauer 2016a, 2016b; Cook 2016; Pavlidou/Alvanoudi 2019; Sato/Anthanasopoulos 2018; Gygax et al. 2021). German's high research interest is attributed to its complex grammatical structure and prominence in Europe (Kotthoff/Nübling 2018). French and Russian follow, with six and four studies respectively (e.g., Lambelet 2015; Speed/Majid 2019; Thongniam/Prasithrathsint 2020).

Nonetheless, figure 3 illustrates that few publications study linguistic relativity in other languages. For instance, merely five articles have studied the relativity hypothesis in Italian and Arabic combined (Vernich/Argus/Kamandulyté-Merfeldiené 2017; Montefinese/Ambrosini/Roivainen 2018; Bin Dawood/Sen/Wu 2020; AlSabbagh 2023; Vernich 2017). For Estonian, Greek, Japanese, Lithuanian, Norwegian, Polish, Serbian, Spanish, Tamil and Thai there is only one article per language (Imai et al. 2013; Vuksanović/Bjekić/Radivojević 2014; Beller et al. 2015; Sedlmeier/Tipandjan/Jänchen 2015; Pavlidou/Alvanoudi 2019; Vernich 2017; Vernich/Argus/Kamandulytė-Merfeldienė 2017; Incera et al. 2018; Maciuszek/Polak/Świątkowska 2019; Thongniam/Prasithrathsint 2020). These findings underscore the existing gap in the literature and emphasize the need for further investigation in this linguistic context.

Regarding genus distinctions, the languages mentioned in figure 3 vary substantially. German, Polish, Russian, Greek, Norwegian, and Serbian feature a three-genus system (masculine, feminine, neuter), while Italian, Spanish, French, Lithuanian, and Arabic use a two-genus system (feminine, masculine). English alone belongs to the category of natural gender languages, where pronouns like *he* and *she* indicate the sex of the referent entity. Additionally, Estonian, Japanese, Thai, and Tamil are classified as genderless languages, lacking grammatical gender markers for all human or nonhuman nouns (Corbett 1991).

This review aims to determine if studies on languages with a three-genus system adequately addressed the neuter. Neglecting the neuter genus biases results toward a binary understanding, reinforcing social norms. Of 25 reviewed studies, 16 examined languages with a three-genus system; however, ten excluded neuter gender analysis in German, Polish, Norwegian, and Serbian (Beller et al. 2015; Bender/Beller/Klauer 2016a, 2016b; Gygax et al. 2021; Imai et al. 2013; Maciuszek/Polak/Świątkowska 2019; Montefinese/Ambrosini/Roivainen 2018; Speed/Majid 2019; Vernich 2017; Vuksanović/Bjekić/Radivojević 2014). Only six articles included neuter genus stimuli in German, Greek, and Russian (Pavlidou/Alvanoudi 2019; Bender/Beller/Klauer 2018; Cook 2016; Nicoladis/Da Costa/Foursha-Stevenson 2015; Sedlmeier/Tipandjan/Jänchen 2015; Vernich 2017), with Bender/Beller/Klauer (2018) stressing the need to include the neuter and consider the full spectrum of genus categories for a comprehensive understanding

of language's influence on cognition. However, their study restricted participants to binary gender decisions, a critical limitation.

This study found that, over the past decade, research on linguistic relativity and grammatical gender has been limited to only 16 languages. Findings from this sample may not generalize to a broader range of world languages (Hammarström/Haspelmath/Bank 2022; Fatemi 2023). Recent investigations have primarily emphasized Indo-European languages, particularly German and English. However, studies on languages with three genera often dismiss the existence of the third neuter genus and focus on associations between feminine and masculine genders within a binary sex framework instead. Consequently, gender studies within the relativity framework may lack complete inclusiveness and generalizability, potentially leading to an incomplete understanding of how language reflects and shapes societal perceptions of gender.

These findings have significant implications for gender studies. Firstly, excluding the neuter in studies on languages with three genera may lead to an incomplete understanding of gender systems and narrow interpretations of gender representation in language. Secondly, restricting gender studies to binary frameworks may overlook the richness and complexity of genus expressions across diverse linguistic systems.

3.2 Research design employed

This section will focus on participant selection and data collection methods, adhering to the PRISMA protocol.

The 25 studies can be categorized into two groups: 14 focusing on monolingual speakers and eleven on bilingual speakers. The latter investigated how genera in both first (L1) and second (L2) languages influence conceptualization.

Eleven of the examined articles explored bilingualism's impact on cognition, focusing on native speakers of various languages with mostly English as their L2 (Beller et al. 2015; Lambelet 2015; Nicoladis/Da Costa/Foursha-Stevenson 2015; Samuel/Roehr-Brackin/Roberson 2015; Cook 2016; Vernich 2017; Sato/Athanasopoulos 2018; Speed/Majid 2019; Sato/Casaponsa/Anthanasopoulos 2020; Bin Dawood/Sen/Wu 2020; White/Cunningham/Zampini 2022). In these studies, proficiency levels in the second language were assessed using the Oxford Placement Test, a pre-self-rated questionnaire, or LEAP-Q (Marian/Blumenfeld/Kauschanslkaya 2007). Among these studies, Lambelet (2015), Cook (2016), and Vernich (2017) stand out for investigating the influence of learning a second language with a distinct genus system on bilingual processing and categorization. Lambelet examined 282 Erasmus students with diverse linguistic backgrounds learning French as a second language, including speakers whose first languages lack grammatical gender. Participants' French proficiency ranged from A1 to C1 levels, assessed using CEFR descriptors. Cook examined 32 advanced-level Russian speakers learning English, all of whom were graduate students not enrolled in ESL classes during the experiment. They were compared with 24 English native speakers who had no significant experience learning a gendered language, serving as monolingual controls. Vernich recruited four groups of Lithuanian speakers (n=128) majoring in philology at the University of Kaunas (VDU) and Vilnius University (VU). The study focused on English and additional languages like Italian, Russian, and German.

White/Cunningham/Zampini (2022) and Speed/Majid (2019) conducted investigations into the influence of genus on odor perception. Investigating the hypothesis that genus may exert automatic and implicit effects in this domain, White/Cunningham/Zampini explored how bilingualism in a gendered language influences gender/sex assignments to odorants. Speed/Majid further investigated grammatical gender's impact on odor perception by exploring how German-English, German-Dutch, and French-English participants memorized descriptions that either matched or mismatched the odor's genus in a description task.

The final two studies that recruited bilingual participants adopted a somewhat distinct approach (Beller et al. 2015; Nicoladis/Da Costa/Foursha-Stevenson 2015). Beller et al. investigated 107 Norwegian participants, examining two language varieties spoken in Norway, denoted as Nynorsk and Bokmål, which feature distinct genus-marking systems. The study aimed to explore how two linguistic varieties spoken within a single cultural context impact the dynamic interplay between language, cognition, and culture. Nicoladis/Da Costa/Foursha-Stevenson investigated how Russian, a gendered language, affects toy classification among 20 Canadian-born preschool bilinguals aged three years, two months to six years, three months in a Russian-English context. These children, exposed to Russian by at least one parent, had their language proficiency assessed by their parents. The study also compared their results with those of 14 English monolingual adults raised in Canada to explore cultural influences on toy classification. However, criticism has been raised about comparing children and adults due to their differing developmental stages (De Houwer 2009: 19–51).

From the 14 studies investigating the influence of genus on cognition in monolingual contexts, two groups emerge: the first, represented by two articles (Montefinese/Ambrosini/Roivainen 2018; Gygax et al. 2021), utilized alternative research approaches without direct human participant involvement; the second, encompassing twelve studies, involved monolingual speakers of various languages (Imai et al. 2013; Vuksanović/Bjekić/Radivojević 2014; Samuel/Roehr-Brackin/Roberson 2015; Sedlmeier/Tipandjan/Jänchen 2015; Bender/Beller/Klauer 2016a, 2016b; Pavlidou/Alvanoudi 2019; Vernich/Argus/Kamandulyté-Merfeldiené 2017; Incera et al. 2018; Maciuszek/Polak/Świątkowska 2019; Thongniam/Prasithrathsint 2020; AlSabbagh 2023).

Montefinese/Ambrosini/Roivainen (2018) collected normative data from Schmidtke and colleagues' (2014) German affective norms (as cited in Montefinese/Ambrosini/Roivainen 2018: 6) and 1.121 Italian words from the Italian adaptations of ANEW (Fairfield et al. 2017; Montefinese et al. 2013a, 2014b, as cited in Montefinese/Ambrosini/Roivainen 2018: 6) in Italian and German, while Gygax et al. (2021) conducted experiments on French, German, and English to explore semantic ambiguities of masculine forms, aligning their findings with previous research.

Five out of the eleven studies on monolingual participants recruited German participants, either solely or with speakers of Greek, Japanese, and Tamil (Imai et al. 2013; Sedlmeier/Tipandjan/Jänchen 2015; Bender/Beller/Klauer 2016a, 2016b; Pavlidou/Alvanoudi 2019).

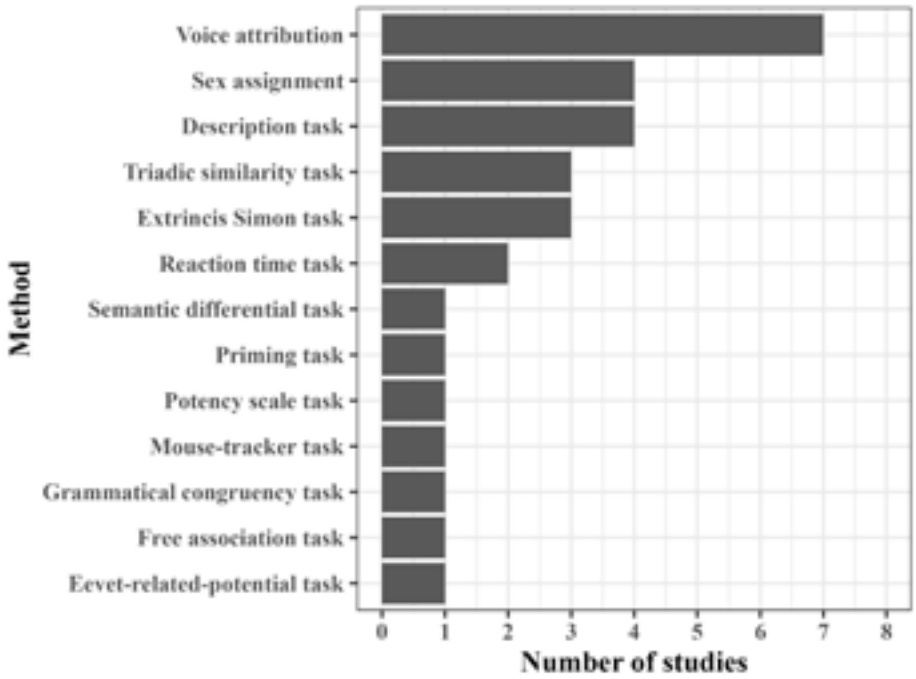
Vernich/Argus/Kamandulyté-Merfeldiené (2017) opted for Estonian as a less-studied language within the relativity framework, comparing it to Lithuanian and Italian. Moreover, Incera et al. (2018) examined how genus in Spanish influences conceptual

representations of musical instruments compared to English, involving 26 Spanish and 26 English monolinguals. In another study, Thongniam/Prasithrathsint (2020) investigated the influence of genus on cognition with a smaller cohort of twelve Russian and Thai participants each.

The last three articles (Vuksanović/Bjekić/Radivojević 2014; Maciuszek/Polak/Świątkowska 2019; AlSabbagh 2023) examined the impact of genus on the cognition of native speakers in Serbian, Polish, and Arabic, respectively.

Next, this study categorized the methodologies of the studies, as outlined in figure 4. After examining the tasks in the 25 studies, one was excluded due to focusing on literature review rather than using psycholinguistic methods (Gygax et al. 2021).

Figure 4: Summary of methods employed in reviewed articles



Source*:* Data compiled from the articles included in this systematic review of research studies published between 2013 and 2023.

As shown in figure 4, voice attribution was the predominant method used, either alone or with other approaches. Beller et al. (2015), Bender/Beller/Klauer (2016a), Bin Dawood/Sen/Wu (2020), Lambelet (2015), Maciuszek/Polak/Świątkowska (2019), Vernich (2017), and Vernich/Argus/Kamandulyté-Merfeldiené (2017) adopted this method, instructing participants to attribute a female or male voice to objects. Notably, three of these studies examined languages with three genera. However, this method warrants criticism as it explicitly prompts binary gender assignments, overlooking the neuter genus in languages like German, Polish, and Russian. This reveals a bias towards a

gender-binary framework, neglecting the nuances of languages with multiple genus categories and potentially undermining the study's purpose.

Description tasks are the second most used method, appearing in four studies either alone or together with other methods (Sedlmeier/Tipandjan/Jänchen 2015; Speed/Majid 2019; Vuksanović/Bjekić/Radivojević 2014; White/Cunningham/Zampini 2022). In these tasks, participants provide the first three adjectives that come to mind to describe an object or odor. A separate group then determines if these adjectives are perceived as feminine or masculine (Boroditsky/Schmidt/Phillips 2003: 69). Critically, this method neglects the neuter genus in studies involving languages like Serbian and German (Sedlmeier/Tipandjan/Jänchen 2015; Vuksanović/Bjekić/Radivojević 2014). This exclusion skews the analysis toward a binary feminine-masculine framework, undermining the primary goal of linguistic relativity: to understand how grammatical features shape reality (Whorf 2007 [1956]).

Sex assignment, the Extrinsic Affective Simon Task (EAST), and the Triad Similarity Task are the third most common methods, each used in three studies (Imai et al. 2013; Nicoladis/Da Costa/Foursha-Stevenson 2015; Sedlmeier/Tipandjan/Jänchen 2015; Pavlidou/Alvanoudi 2019; Bender/Beller/Klauer 2016a, 2016b, 2018; Maciuszek/Polak/Świątkowska 2019; Thongniam/Prasithrathsint 2020). Sex assignment, where participants determine the sex of the stimulus, faces criticism for potentially yielding unrealistic results, especially in German, Greek, and Russian contexts (Pavlidou/Alvanoudi 2019; Imai et al. 2013; Nicoladis/Da Costa/Foursha-Stevenson 2015). In EAST, stimuli on an LCD screen include colors like black, blue, or green and are accompanied by a definite article relevant to the language (e.g., *der, die,* or *das* in German). Each color scheme represents a specific category, such as male/female in black and distractors in blue and green, positioned in the bottom corners of the screen. Participants categorize the stimuli by pressing a key for correct and incorrect judgments. Figure 5 shows an example of an EAST adapted from Bender/Beller/Klauer (2016a).

Figure 5: Four examples of Extrinsic Simon Task

Example 1 Basic Category CONGRUENT ANIMATEAS Color: black	Example 2 Reference Category CONGRUENT ANIMATEAS Color: green or **blue**	Example 3 Target Category GENERIC ANIMATEAS Color: green or **blue**	Example 4 Target Category NON- ANIMATEAS Color: green or **blue**
Tante female male green **blue**	Onkel female male green **blue**	**Ziege** female male green **blue**	**Löffel** female male green **blue**
Decision: biological sex X	Decision: color X	Decision: color X	Decision: color X

Source: Adapted from Bender/Beller/Klauer (2016a).

Triadic Similarity Judgment involves groups of three words where two share the same genus and the third differs. Participants judge the similarity of each triad (Maciuszek/Polak/Świątkowska 2019; Sedlmeier/Tipandjan/Jänchen 2015; Thongniam/Prasithrathsint 2020).

Among studies using EAST and Triadic Similarity Judgment, only Bender/Beller/Klauer (2018), Sedlmeier/Tipandjan/Jänchen (2015), and Thongniam/Prasithrathsint (2020) included neuter stimuli. Notably, Bender/Beller/Klauer (2018) is the only recent study to intentionally balance neuter stimuli with feminine and masculine ones. Bender/Beller/Klauer (2018) used grammatically neuter generic animate words in German (e.g., *Krokodil*[3], *Kätzchen*[4]) for their EAST task but instructed participants to categorize stimuli as only female/male or blue/green. This approach limits participants by framing stimuli strictly within a binary sex context, despite including neuter genus. Sedlmeier/Tipandjan/Jänchen (2015) and Thongniam/Prasithrathsint (2020) included neuter items, but fewer than feminine and masculine items. In these studies, participants categorized neuter stimuli based on perceptions of femininity and masculinity.

In addition, seven approaches were used in only one study each, while the reaction time task appeared in two studies (Cook 2016; Maciuszek/Polak/Świątkowska 2019). These approaches include the potency scale, semantic differential, grammatical congruency, event-related potential, free association, mouse tracking, and priming (AlSabbagh 2023; Incera et al. 2018; Montefinese/Ambrosini/Roivainen 2018; Samuel/Roehr-Brackin/Roberson 2015; Sato/Casaponsa/Anthanasopoulos 2020; Sato/Athanasopoulos 2018; Sedlmeier/Tipandjan/Jänchen 2015). Although used less frequently, these methods highlight the diverse approaches to exploring linguistic relativity and grammatical gender in psycholinguistics. Researchers using the potency scale and free-association approach included the neuter gender. Sedlmeier/Tipandjan/Jänchen (2015) employed a potency scale but did not maintain an equal number of stimuli for each genus in German. In contrast, Cook (2016) used a free-association task to examine the neuter third-person singular pronoun in Russian and English, comparing bilingual speakers and monolingual English controls and found no significant differences in processing costs under neutral conditions.

As pointed out, this systematic literature review examined peer-reviewed articles published in English between 2013 and 2023, focusing on grammatical gender's influence on sex representation and bias in gendered languages. Using the PRISMA protocol (Moher et al. 2015), the review elucidates participant selection criteria and methodologies employed across the articles.

The impact of genus on conceptual representation ties into the broader debate on language's influence on thought. Linguistic relativity suggests that grammar shapes perspectives. Researchers studying this look for evidence of genus affecting cognitive processes like categorization, memory etc. (Vernich 2017; Maciuszek/Polak/Świątkowska 2019).

The analysis of the 25 studies found that researchers selected either bilingual or monolingual participants, with eleven studies focusing on bilinguals. It is crucial to differentiate between bilingualism and foreign language learning. Bilingualism involves

3 English translation: *crocodile*.
4 English translation: *kitten*.

growing up with two languages, acquiring a second through migration, or learning it institutionally with significant exposure to both. In contrast, foreign language learning occurs in educational settings or for personal interest (Delucchi/Mertins 2018). Bilingualism can be categorized by usage frequency, acquisition type, and proficiency (Delucchi/Mertins 2018). Oversights in participant recruitment, such as those in Cook (2016), Lambelet (2015), and Vernich (2017), illustrate this issue. Lambelet (2015) included Erasmus students with varying French proficiency, Cook (2016) grouped graduate students as Russian-English bilinguals despite differing proficiency levels, and Vernich (2017) examined Lithuanian philology students with varied language exposure. Such recruitment practices lead to inaccurate results, as genuine bilingualism requires substantial contact and experience with both languages.

As shown in figure 4, most studies used similar methods, like voice attribution, description, EAST, sex assignment, and triadic similarity tasks, with voice attribution being the most common. However, there is a critical issue with methods that explicitly instruct participants to assign female or male voices. Such methods reveal the experiment's objective, which may bias results and prevent interpreting them as unconscious effects of grammatical gender. Similar concerns apply to methods like sex assignment and description, where tasks involve a binary gender spectrum, affecting performance or the assignment of binary genders to adjectives. Notably, among the 25 studies reviewed, only seven used entirely distinct tasks.

Of the 16 studies examining languages with three genera (German, Polish, Russian, and Serbian), only four included the neuter in their analysis (Bender/Beller/Klauer 2018; Cook 2016; Sedlmeier/Tipandjan/Jänchen 2015; Thongniam/Prasithrathsint 2020). Notably, Bender/Beller/Klauer (2018) uniquely included an equal number of neuter stimuli in German, matching the quantities of feminine and masculine stimuli. However, the EAST task in these studies required participants to categorize neutral items as either feminine or masculine. Cook (2016) is notable for testing the neutral third-person singular pronoun in both Russian and English.

Overall, the studies from the past decade are limited to a narrow set of languages and use tasks that fail to capture the unconscious effects of genus on conceptualization. Consequently, their findings lack generalizability to most languages and do not support the relativity hypothesis.

The current study results have serious implications for both theory and practice. Theoretically, it is crucial to broaden the range of languages studied and refine current methods while developing new approaches beyond those identified in the literature. Practically, researchers should be cautious when interpreting results from established procedures and avoid generalizing binary genus concepts to languages with different gender systems including those marking neuter genus.

Conclusion

This review examined the impact of genus on cognition within the relativity framework. It revealed a downward trend in the number of studies conducted over the past four years and a focus on few languages, with a bias toward German and English. Studies

often misrepresented three-genus systems by excluding the neuter or forcing binary sex categorization. Future research should improve theoretical and methodological approaches by broadening linguistic diversity and refining existing methods.

References

Acke, Hanna (2019). Sprachwandel durch feministische Sprachkritik. *Zeitschrift für Literaturwissenschaft und Linguistik*, *49*(2), 303–320. https://doi.org/10.1007/s41244-019-00135-1

AlSabbagh, Zainab Abdulaziz (2023). Grammatical Gender & Linguistic Relativity: Does the Grammatical Gender Native Arabic Speakers Assign Neutral Nouns Affect How They Perceive Them? *European Journal of Foreign Language Teaching*, *7*(1). https://doi.org/10.46827/ejfl. v7i1.4673

Alvanoudi, Angeliki (2015). *Grammatical Gender in Interaction: Cultural and Cognitive Aspects*. Leiden: Brill.

Beller, Sieghard; Brattebø, Karen; Lavik, Kristina; Reigstad, Rakel & Bender, Andrea (2015). Culture or Language: What Drives Effects of Grammatical Gender? *Cognitive Linguistics*, *26*(2), 331–359. https://doi.org/10.1515/cog-2014-0021

Bender, Andrea; Beller, Sieghard & Klauer, Karl (2016a). Crossing Grammar and Biology for Gender Categorization: Investigating the Gender Congruency Effect in Generic Nouns for Animates. *Journal of Cognitive Psychology*, *28*(5), 530–558. https://doi.org/10.1080/20445911.2016.1148042

Bender, Andrea; Beller, Sieghard & Klauer, Karl (2016b). Lady Liberty and Godfather Death as Candidates for Linguistic Relativity? Scrutinizing the Gender Congruency Effect on Personified Allegories with Explicit and Implicit Measures. *Quarterly Journal of Experimental Psychology*, *69*(1), 48–64. https:// doi.org/10.1080/17470218.2015.1021701

Bender, Andrea; Beller, Sieghard & Klauer, Karl (2018). Gender Congruency from a Neutral Point of View: The Roles of Gender Classes and Conceptual Connotations. *Journal of Experimental Psychology: Learning, Memory, and Cognition*, *44*(10), 1580–1608. https://doi.org/10.1037/xlm0000534

Bin Dawood, Abdullah; Sen, Maya & Wu, Yan Jing (2020). Linguistic Relativity: Object Categorisation Differences between Arabic and English Speakers. *Jurnal Psikologi Malaysia*, *34*(4), 14–32.

Boroditsky, Lera; Schmidt, Lauren & Phillips, Webb (2003). Sex, Syntax and Semantics. In Dedre Gentner & Susan Goldin-Meadow (eds.), *Language in Mind: Advances in the Study of Language and Thought* (pp. 61–79). Boston: Boston Review.

Cook, Svetlana (2016). Gender Matters: From L1 Grammar to L2 Semantics. *Bilingualism: Language and Cognition, 21*(1), 13–31. https:// doi.org/10.1017/s1366728916000766

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

De Houwer, Annick (2009). *Bilingual First Language Acquisition*. Bristol: Multilingual Matters.

Delucchi, Renate & Mertins, Barbara (2018). Psycholinguistische Grundlagen der Inklusion: Schwerpunkt Bilingualismus. In Stephan Hußmann & Barbara Welzl (eds.), *DoProfiL – Das Dortmunder Profil für inklusionsorientierte Lehrerinnen- und Lehrerbildung* (pp. 161–178). Münster: Waxmann.

Fatemi, Maryam (2023). Synchronic Overview of Grammatical Gender Marking in Indo-European Languages. In Barbara Mertins & Renate Delucchi Danhier (eds.), *Diversity in Cognition* (pp. 13–60). Berlin: Peter Lang.

Gorny, Hildegard (1995). Feministische Sprachkritik. In Georg Stötzel & Martin Wengeler (eds.), *Kontroverse Begriffe: Geschichte des öffentlichen Sprachgebrauchs*

in der Bundesrepublik Deutschland (pp. 517–562). Berlin: De Gruyter. https://doi.org/10.1515/9783110881660.517

Günthner, Susanne; Hüpper, Dagmar & Spieß, Constanze (2012). *Genderlinguistik: Sprachliche Konstruktionen von Geschlechtsidentität*. Berlin: De Gruyter.

Gygax, Pascal; Sato, Sayaka; Öttl, Anton & Gabriel, Ute (2021). The Masculine Form in Grammatically Gendered Languages and its Multiple Interpretations: A Challenge for Our Cognitive System. *Language Sciences, 83*, 101328. https://doi.org/10.1016/j.langsci.2020.101328

Hammarström, Harald; Forkel, Robert; Haspelmath, Martin & Bank, Sebastian (2022). *Glottolog 4.6.* Leipzig: Max Planck Institute for Evolutionary Anthropology. https://doi.org/10.5281/zenodo.6578297

Hellinger, Marlis (ed.). (1985). *Sprachwandel und Feministische Sprachpolitik: Internationale Perspektiven*. Opladen: Westdeutscher Verlag. https://doi.org/10.1007/978-3-322-83937-4

Hellinger, Marlis & Bußmann, Hadumod (2002). *Gender Across Languages: The Linguistic Representation of Women and Men*. Amsterdam: Benjamins. https://doi.org/10.1075/impact.11

Imai, Mutsumi; Schalk, Lennart; Saalbach, Henrik & Okada, Hiroyuki (2013). All Giraffes Have Female-specific Properties: Influence of Grammatical Gender on Deductive Reasoning About Sex-specific Properties in German Speakers. *Cognitive Science, 38*(3), 514–536. https://doi.org/10.1111/cogs.12074

Incera, Sara; McLennan, Conor; Stronsick, Lisa & Zetzer, Emily (2018). Is Tuba Masculine or Feminine? The Timing of Grammatical gender. *Mind & Language, 34*(5), 667–680. https://doi.org/10.1111/mila.12223

Kotthoff, Helga & Nübling, Damaris (2018). *Genderlinguistik: Eine Einführung in Sprache, Gespräch und Geschlecht*. Tübingen: Narr Francke Attempto.

Kousta, Stavroula-Thaleia; Vinson, David & Vigliocco, Gabriella (2008). Investigating Linguistic Relativity Through Bilingualism: The Case of Grammatical Gender. *Journal of Experimental Psychology: Learning, Memory, and Cognition*, *34*(4), 843–858. https://doi.org/10.1037/0278-7393.34.4.843

Lambelet, Amelia (2015). Second Grammatical Gender System and Grammatical Gender-linked Connotations in Adult Emergent Bilinguals with French as a Second Language. *International Journal of Bilingualism, 20*(1), 62–75. https://doi.org/10.1177/1367006915576832

Leaper, Campbell & Ayres, Melanie (2007). A Meta-analytic Review of Gender Variations in Adults' Language Use: Talkativeness, Affiliative Speech, and Assertive Speech. *Personality and Social Psychology Review*, *11*(4), 328–363. https://doi.org/10.1177/1088868307302221

Lagopoulos, Alexandros & Boklund-Lagopoulou, Karin (2020). *Theory and Methodology of Semiotics: The Tradition of Ferdinand de Saussure*. Berlin: De Gruyter.

Lucy, John Arthur (1992). *Language Diversity and Thought: A Reformulation of the Linguistic Relativity Hypothesis*. Cambridge: Cambridge University Press.

Maciuszek, Józef; Polak, Mateusz & Świątkowska, Natalia (2019). Grammatical Gender Influences Semantic Categorization and Implicit Cognition in Polish. *Frontiers in Psychology, 10*. https://doi.org/10.3389/fpsyg.2019.02208

Marian, Viorica; Blumenfeld, Henrike & Kaushanskaya, Margarita (2007). The Language Experience and Proficiency Questionnaire (LEAP-Q): Assessing Language Profiles in Bilinguals and Multilinguals. *Journal of Speech, Language, and Hearing Research*, *50*(4), 940–967. https://doi.org/10.1044/1092-4388(2007/067)

Mertins, Barbara (2018). *Sprache und Kognition: Ereigniskonzeptualisierung im Deutschen und Tschechischen*. Berlin: De Gruyter.

Moher, David; Shamseer, Larissa; Clarke, Mike; Ghersi, Davina; Liberati, Alessandro; Petticrew, Mark; Shekelle, Paul; Stewart, Lesley & PRISMA-P Group (2015). Preferred Reporting Items for Systematic Review and Meta-analysis Protocols (PRISMA-P) 2015 Statement. *Systematic Reviews*, *4*(1). https://doi.org/10.1186/2046-4053-4-1

Montefinese, Maria; Ambrosini, Ettore & Roivainen, Eka (2018). No Grammatical Gender Effect on Affective Ratings: Evidence from Italian and German Languages. *Cognition and Emotion, 33*(4), 848–854. https://doi.org/10.1080/02699931.2018.1483322

Newman, Alexander; Round, Heather; Wang, Shuanglong & Mount, Matthew (2019). Innovation Climate: A Systematic Review of the Literature and Agenda for Future Research. *Journal of Occupational and Organizational Psychology*, *93*(1), 73–109. https://doi.org/10.1111/joop.12283

Newman, Matthew; Groom, Carla; Handelman, Lori & Pennebaker, James (2008). Gender Differences in Language Use: An Analysis of 14,000 Text Samples. *Discourse Processes*, *45*(3), 211–236. https://doi.org/10.1080/01638530802073712

Nicoladis, Elena; Da Costa, Natalie & Foursha-Stevenson, Cassandra (2015). Discourse Relativity in Russian-English Bilingual Preschoolers' Classification of Objects by Gender. *International Journal of Bilingualism, 20*(1), 17–29. https://doi.org/10.1177/1367006915576826

Palomares, Nicholas (2008). Explaining Gender-based Language Use: Effects of Gender Identity Salience on References to Emotion and Tentative Language in Intra- and Intergroup Contexts. *Human Communication Research*, *34*(2), 263–286. https://doi.org/10.1111/j.1468-2958.2008.00321.x

Pavlidou, Theodossia-Soula & Alvanoudi, Angeliki (2019). Conceptualizing the World as "Female" or "Male": Further Remarks on Grammatical Gender and Speakers' Cognition. *Selected Papers of ISTAL, 23*, 317–332. https://doi.org/10.26262/istal.v23I0.7351

Prewitt-Freilino, Jennifer; Caswell, T. Andrew & Laakso, Emmi (2011). The Gendering of Language: A Comparison of Gender Equality in Countries with Gendered, Natural Gender, and Genderless Languages. *Sex Roles*, *66*(3–4), 268–281. https://doi.org/10.1007/s11199-011-0083-5

Pusch, Luise (2014). *Gerecht und Geschlecht: Neue Sprachkritische Glossen*. Göttingen: Wallstein.

Samuel, Steven; Roehr-Brackin, Karen & Roberson, Debi (2015). "She says, he says": Does the Sex of an Instructor Interact with the Grammatical Gender of Targets in a Perspective-taking Task? *International Journal of Bilingualism, 20*(1), 40–61. https://doi.org/10.1177/1367006915576831

Sato, Sayaka & Athanasopoulos, Panos (2018). Grammatical Gender Affects Gender Perception: Evidence for the Structural-feedback Hypothesis. *Cognition, 176*, 220–231. https://doi.org/10.1016/j.cognition.2018.03.014

Sato, Sayaka; Casaponsa, Aina & Athanasopoulos, Panos (2020). Flexing gender perception: Brain potentials reveal the cognitive permeability of grammatical information. *Cognitive Science, 44*(9). https://doi.org/10.1111/cogs.12884

Sedlmeier, Peter; Tipandjan, Arun & Jänchen, Anastasia (2015). How Persistent are Grammatical Gender Effects? The Case of German and Tamil. *Journal of Psycholinguistic Research, 45*(2), 317–336. https://doi.org/10.1007/s10936-015-9350-x

Slobin, Dan Isaac (1996). From "Thought and Language" to "Thinking for Speaking". In John Gumperz & Stephen Levinson (eds.), *Rethinking Linguistic Relativity* (pp. 70–96). Cambridge: Cambridge University Press.

Speed, Laura & Majid, Asifa (2019). Linguistic Features of Fragrances: The Role of Grammatical Gender and gender associations. *Attention, Perception, & Psychophysics, 81*(6), 2063–2077. https://doi.org/10.3758/s13414-019-01729-0

Thongniam, Kusuma & Prasithrathsint, Amara (2020). The Influence of Grammatical Gender on Russian and Thai Speakers' Cognition. *Manusya: Journal of Humanities, 23*(1), 40–59. https://doi.org/10.1163/26659077-02301003

Trömel-Plötz, Senta (ed.). (1997). *Gewalt durch Sprache: Die Vergewaltigung von Frauen in Gesprächen*. Frankfurt/Main: Fischer.

Vernich, Luca (2017). Does Learning a Foreign Language Affect Object Categorization in Native Speakers of a Language with Grammatical Gender? The Case of Lithuanian Speak-

ers Learning Three Languages with Different Types of Gender Systems (Italian, Russian and German). *International Journal of Bilingualism, 23*(2), 417–436. https://doi.org/10.1177/1367006917728593

Vernich, Luca; Argus, Reili & Kamandulytė-Merfeldienė, Laura (2017). Extending Research on the Influence of Grammatical Gender on Object Classification: A Cross-linguistic Study Comparing Estonian, Italian and Lithuanian Native speakers. Eesti Rakenduslingvistika Ühingu Aastaraamat. *Estonian Papers in Applied Linguistics, 13*, 223–240. https://doi.org/10.5128/erya13.14

Vuksanović, Jasmina; Bjekić, Jovana & Radivojević, Natalija (2014). Grammatical Gender and Mental Representation of Object: The Case of Musical Instruments. *Journal of Psycholinguistic Research, 44*(4), 383–397. https://doi.org/10.1007/s10936-014-9293-7

White, Theresa; Cunningham, Caitlin & Zampini, Mary (2022). Is That "Mr." or "Ms." Lemon? An Investigation of Grammatical and Semantic Gender on the Perception of Household Odorants. *Brain Sciences, 12*(10), 1313. https://doi.org/10.3390/brainsci12101313

Whorf, Benjamin Lee (2007 [1956]). *Language, Thought, and Reality: Selected Writings of Benjamin Lee Whorf*. Cambridge: MIT Press.

Author's details

S. Maryam Fatemi, MA, PhD student and research assistant, TU Dortmund, Institute for Diversity Studies. Research focus: Language and Cognition.
Contact: Institute for Diversity Studies, TU Dortmund, Emil-Figge-Straße 51, 44227 Dormund
Email: seyedehmaryam.fatemi@tu-dortmund.de

Appendix A Summary of the reviewed studies

Year	Authors	Language	Participant Number and Sex	Participant Selection Method	Population Type	Task Type
2013	Imai, Schalk, Saalbach & Okada	German, Japanese	57	Random sampling	Students from Zurich and Tokyo areas	SAT[5]
2014	Vuksanović, Bjekić & Radivojević	Serbian	136 female, 131 male	Random sampling	Students	Description Task
2015	Beller, Brattebø, Lavik, Reigstad & Bender	Nynorsk, Bokmål (Two varieties of Norwegian)	64 female, 43 male	Random sampling	Volunteers from Bergen	VAT[6]
2015	Lambelet	L2 French, L1 various languages	89 female, 193 male	Self evaluation	Erasmus students in Switzerland	VAT
2015	Nicoladis, Da Costa & Foursha-Stevenson	Russian, English	11 female, 9 male	Parent evaluation	Russian-English bilingual kids	SAT
2015	Samuel, Roehr-Brackin & Roberson	L2 English, L1 various languages	43 female, 21 male	Random sampling	Students from Essex	Grammatical Congruency Task
2015	Sedlmeier, Tipandjan & Jänchen	German, Tamil	117 female, 112 male	Random sampling	Students from Chemnitz	Potency Scale Task, Triadic Similarity Task and Description Task
2016a	Bender, Beller & Klauer	German	93 female, 84 male	Random sampling	Volunteers from Freiburg	EAST[7] and VAT
2016b	Bender, Beller & Klauer	German	66 female, 52 male	Random sampling	Volunteers from Freiburg	EAST and SAT
2016	Cook	Russian, English	21 female, 22 male	Self evaluation	Russian students and English monolinguals	Reaction Time Task
2017	Vernich	Lithuanian, Italian, German, Russian	128	Self evaluation	Students from Kaunas and Vilnius	VAT
2017	Vernich, Argus & Kamandulytė-Merfeldienė	Estonian, Italian	60 female, 60 male	Random sampling	Students from Milan, Kaunas and Tallin	VAT
2018	Bender, Beller & Klauer	German	50 female, 26 male	Random sampling	Students from Freiburg	EAST

5 Sex Assignment Task
6 Voice Attribution Task
7 Extrinsic Affective Simon Task

Year	Authors	Language	Participant Number and Sex	Participant Selection Method	Population Type	Task Type
2018	Incera, McLennan, Stronsick & Zetzer	English, Spanish	52	Random sampling	Volunteers from Cleveland and Santander	MouseTracker Task
2018	Montefinese, Ambrosini & Roivainen	Italian, German	—	—	—	Semantic Differential Task
2018	Sato & Athanaso-poulos	French, English	33 female, 26 male	Specific placement test	French-English bilinguals and native English monolinguals	Priming Task
2019	Maciuszek, Polak & Świątkowska	Polish	272 female, 100 male	Random sampling	Students form Krakow	Triadic Similarity Task, Reaction Time Task (modified Implicit Association Test) and VAT
2019	Pavlidou & Alvanoudi	German, Greek	63 female, 63 male	Random sampling	Students from Berlin and Thessaloniki	SAT
2019	Speed & Majid	German, French	83 female, 20 male	Self evaluation	Volunteers from Nijmegen and Lyon	Description Task
2020	Bin Dawood, Sen & Wu	Arabic, English	49 female, 71 male	Self evaluation	Students from Riyadh and Marquette	VAT
2020	Sato, Casaponsa & Anthahas-opoulos	French, English	27 female, 13 male	Self evaluation	Students from Lancaster	Event-related Potentials Task
2020	Thongniam & Prasithrathsint	Russian, Thai	12 female, 12 male	Random sampling	Students from Bangkok	Triadic Similarity Task
2021	Gygax, Sato, Öttl & Gabriel	French, German, English	—	—	—	Memory-based Approaches to Reading
2022	White, Cunningham & Zampini	French, English	20 female, 12 male	Self evaluation	Volunteers from Montreal	Description Task
2023	AlSabbagh	Arabic, English	25 female, 25 male	Random sampling	Students from Bahrain	Free Association Task

Source: Data compiled from the articles included in this systematic review of research studies published between 2013 and 2023.