TU Dortmund University

Faculty of Cultural Studies

Grammatical Gender and Cognition

Cumulative Dissertation submitted in fulfillment of the requirements for the degree of
Doctor of Philosophy (Dr. phil.)
by

S.Maryam Fatemi, M.A.

First Examiner: Prof. Dr. Christian Zimmer
Second Examiner: Dr. Benedetta Bassetti
Date of Defence: 28.08.2026
Dortmund

To my father, in loving memory, whose pride continues to guide me.
To my mother, whose unwavering love has been my strength throughout this journey.

Acknowledgments

This dissertation is not only the result of intensive scientific work, but also the outcome of many meaningful encounters, valuable collaborations, and the support of remarkable individuals. I feel deeply grateful for the people I have met along this path and consider myself truly fortunate for their support. My doctoral journey has also taught me an important lesson about supervision: the right guidance should create opportunities for learning and growth, foster independence, and support researchers in bringing their work to completion rather than creating barriers along the way.

I would also like to express my heartfelt appreciation to my colleagues. Katrin, your support meant more to me than I can fully express. During moments of exhaustion and doubt, you encouraged me to keep going and reminded me how far I had already come. You gave me strength when I felt I had none left, and our conversations—ranging from academic to personal—made this journey not only manageable, but truly meaningful. Renate, I am especially grateful for your motivating words during uncertain phases and for your thoughtful suggestions, which helped me move forward in a focused and efficient way. Joël, thank you for your constructive feedback, which greatly supported the scientific development of my work. Linda, I sincerely thank you for our many valuable conversations and your continuous academic support—your kindness and encouragement meant a great deal to me. All of you have been very special to me, and I am deeply grateful for your support and friendship.

I would also like to thank the students who supported this project, in particular Sarah and Nicolas, for their immense help in reviewing my drafts and assisting with formatting. My thanks also go to Tim Schwarz for his openness and support in initiating the first pilot study, as well as to all members of Team Mertins for their engagement and collaborative spirit.

I am also sincerely grateful to Carmen van Meegen from the Statistical Consultation at TU Dortmund for her invaluable statistical advice, which significantly strengthened the methodological rigor of my work. I would also like to thank Dr. Tetyana Müller-Lyaskovets from the writing group at TU Dortmund. Her workshops opened up a new perspective for me and helped me discover my appreciation for writing, while her guidance in developing clarity and precision has greatly enhanced my academic writing skills.

I would like to express my sincere gratitude to all members of the Iranian community in Germany who took part in this study without any expectation of financial compensation. Your willingness to contribute to academic research and your genuine interest in supporting this work were deeply meaningful and reflect the strong intellectual and academic spirit of the Iranian diaspora.

My heartfelt thanks also go to the participants in Iran, who contributed to this study despite challenging circumstances, including unstable internet connections and periods of uncertainty in the country. Your commitment, resilience, and willingness to participate under such conditions were invaluable and played a crucial role in making this research possible.

My sincere thanks go to my first reviewer and examiner, Prof. Dr. Christian Zimmer, for the time and care he devoted to reading my dissertation and for his thoughtful and constructive evaluation of my work. I would also like to thank my second reviewer and examiner, Prof. Dr. Benedetta Bassetti, for taking the time to review and evaluate my dissertation

I would also like to thank my sister, Haleh, for always being there for me. Your continuous love, patience, and encouraging words—whenever we were able to stay in touch, despite the difficulties in communication—gave me strength throughout this demanding time. Your trust in me meant more than words can express.

Last but not least, my deepest gratitude goes to my husband, Armin Razavi. Throughout all these years, you have been by my side as my strongest support and greatest source of motivation. Your patience, understanding, and unwavering encouragement are beyond what I could have ever imagined. You stood by me through long days and nights of work, when I was often at my desk or in the lab, with little time for anything else. Yet, you never stopped supporting me and always made me feel cared for. Without you, this dissertation would not have been possible. In many ways, this achievement belongs to you as well.

Abstract

This cumulative dissertation is situated within the fields of cognitive linguistics and psycholinguistics and investigates the relationship between grammatical gender and cognition within the framework of linguistic relativity. In doing so, the dissertation integrates typological, theoretical, and experimental approaches in its research methodology. This dissertation focuses on grammatical gender as a morphosyntactic category and examines its role in cognition within the framework of linguistic relativity across multiple cognitive levels—conceptual representation, lexical access, and real-time processing—in both monolingual and bilingual contexts. Throughout the five articles comprising the dissertation, it addresses a central question in cognitive science and linguistics—namely, to what extent grammatical structure influences cognitive processing— while extending the body of research on language–cognition interactions through a systematic investigation of grammatical gender.

The findings of this dissertation provide evidence that grammatical structure shapes cognition in both behavioral measures and real-time language processing by guiding attention and interpretation, while not fully determining conceptual representations. Instead, its influence is dynamic and modulated by linguistic context and speaker experience.

The first study investigated how grammatical gender systems are distributed and documented across languages and whether current typological knowledge provides a reliable basis for research on linguistic relativity. Using a synchronic typological analysis of 413 Indo-European languages based on large-scale linguistic databases, the study systematically examined the presence of grammatical gender systems. The findings revealed substantial linguistic documentation gaps and a strong bias toward well-described European languages, with many linguistic systems insufficiently classified. These results indicate that assumptions about grammatical gender and its cognitive effects are often based on Eurocentric and WEIRD-biased, and thus incomplete, datasets. This is important for researchers in typology and cognitive linguistics, as broader and more representative data are necessary for generalizable conclusions.

The second study investigated the operationalization of grammatical gender in empirical research on linguistic relativity. Using a PRISMA-guided systematic review of studies published between 2013 and 2023, it analyzed language selection, methodological approaches, and the treatment of gender categories. The findings revealed a strong focus on a limited set of

languages, primarily German and English, alongside a tendency to reduce grammatical gender to binary distinctions. Crucially, the neuter category is often excluded, even in languages that encode it. This suggests that existing evidence on grammatical gender and cognition is methodologically constrained. The study highlights the need for more comprehensive designs, which are essential for researchers seeking to accurately assess how grammatical structure influences cognition.

Building on this foundation, the third study critically examined research on grammatical gender in bilingual and multilingual contexts. Through a critical review of studies on bilingual speakers, the analysis evaluated definitions of bilingualism, control of language experience, and methodological approaches. The findings revealed that inconsistencies largely stem from heterogeneous participant profiles and a reliance on offline behavioral measures. These limitations obscure the extent to which grammatical gender influences cognition in bilinguals. This study directly informs the central research question by emphasizing that sensitive, real-time methods are required. This is especially important for psycholinguistic studies of bilingual language processing, which aim to capture dynamic effects in real-time processing.

The empirical studies presented in this dissertation directly address the empirical and methodological gaps identified in the preceding studies, namely the limited range of language comparisons, the exclusion of the neuter category, and the reliance on simplified experimental designs. A behavioral experiment comparing German, a language with a three-gender system, and Farsi, a genderless language, demonstrated that grammatical gender influences conceptualization in German speakers. These effects are particularly evident at early stages of lexical retrieval, suggesting an automatic link between grammatical gender and conceptual processing. In contrast, Farsi speakers showed no systematic gender-related patterns, indicating that such effects are grounded in language-specific grammatical structures. Importantly, the inclusion of the neuter category reveals that it functions as a cognitively meaningful class rather than a residual grammatical form. The findings demonstrate that grammatical gender shapes conceptualization in a language-specific manner, emerging early in processing and extending to the neuter category, thereby providing a more comprehensive test of linguistic relativity.

The final study extended this investigation to bilingual speakers and employed eye-tracking within the Visual World Paradigm to examine the role of grammatical gender in real-time language processing. The findings showed that grammatical gender guides visual attention, with gender-congruent information capturing and sustaining gaze in German monolingual

speakers. Farsi–German bilingual speakers, whose degree of bilingualism was systematically assessed, thereby addressing a key gap in research on linguistic relativity and bilingualism, exhibited similar but attenuated effects when processing German, while no such effects were observed when they performed the task in Farsi. These results demonstrate that grammatical gender can be integrated into processing even when acquired later in life. However, its influence is modulated by the linguistic context in use.

Taken together, the findings in all five articles support a dynamic and process-based account of linguistic relativity. Grammatical gender does not impose fixed constraints on conceptual representations, but instead, it influences the way information is attended to, structured, and processed by speakers during language use. Driven by typological breadth, theoretical clarity, and experimental precision, this dissertation contributes to a more nuanced understanding of how linguistic structure interacts with cognition across languages and speakers. In addition, it highlights the importance of methodological diversity for advancing research in cognitive science and psycholinguistics.

This dissertation contributes to the field by integrating typological, theoretical, and experimental approaches to investigate grammatical gender across cognitive levels. By incorporating the neuter as a meaningful category, combining behavioral and eye-tracking methods, and examining an understudied language pairing in monolingual and bilingual contexts, it provides new evidence for the dynamic and context-dependent nature of language–cognition interactions.

Table of Contents

Chapter 1 Introduction 1
1.1 Aims of the Dissertation 2
1.2 Research Questions 2
1.3 Overview of Thesis Structure 3
1.4 Contribution to the Field 5
Chapter 2 Cognition and Theoretical Framework 7
2.1 Defining Cognition from Cognitive Linguistics Perspectives 8
2.2 Language Production and Grammatical Gender: Levelt's Model 12
2.3 Linguistic Relativity 15
2.3.1 Historical Roots of the Language–Thought Relationship 16
2.3.2 Humboldt's Weltansicht and Weltanschauung 16
2.3.3 From Humboldt to Sapir and Whorf 17
2.3.4 The Political and Historical Contingencies of Early Interpretations: Monolingualism as National Ideal 19
2.3.5 Empirical Studies on Language and Cognition 20
2.3.6 Critical Assessment of Linguistic Relativity 22
2.4 Conclusion 23
Chapter 3 Grammatical Gender in Cross-Linguistic Perspective and Its Role in Linguistic Relativity 24
3.1 Introduction 24
3.2 Defining Grammatical Gender and Cross-Linguistic Variation 25
3.2.1 A Semantic Core of Animacy and Humanness 26
3.2.2 Morphological, Phonological, and Arbitrary Assignment 27
3.2.3 Cross-Linguistic Diversity in Gender Systems 28
3.2.4 Agreement and Functional Motivations 29
3.2.5 Interim Summary 30
3.3 Approaches to Grammatical Gender Across Traditions 30
3.3.1 Psycholinguistic Perspectives 31
3.3.2 Sociolinguistic Perspectives 32
3.3.3 Linguistic Relativity Perspective 33
3.4 Conclusion 35
Chapter 4 Methodological Framework 36
4.1 Typological and Review-Based Methods 37
4.2 Theoretical and Critical Review Approaches 39
4.3 Behavioral Experimental Methods 40
4.4 Eye-Tracking and the Visual World Paradigm 42
4.5 Language Background Questionnaires 44
4.6 Integration of Methodological Approaches 46

Chapter 5 Synchronic overview of grammatical gender marking in Indo- European languages 48

Chapter 6 The influence of grammatical gender on the conceptualization of the world: A systematic literature review 96

Chapter 7 Language-Specific Patterns in Multilingual Contexts: The Case of Grammatical Gender 115

Chapter 8 The influence of grammatical gender on Cognition: The Case of German and Farsi 145

Chapter 9 Grammatical Gender in a Grammatically Genderless Mind: Eye-Tracking Evidence from Farsi-German Bilinguals 177

Chapter 10 Discussion and Conclusion 235

10.1 Theoretical implications 237

10.1.1 Core Position: Soft, Dynamic Relativity 238

10.1.2 Grammatical Gender Across Cognitive Levels: Representation and Processing 239

10.1.3 Psycholinguistic Processing and the Automatic Activation of Grammatical Gender 241

10.1.4 Bilingual Cognition and Context-Dependent Language Processing 242

10.2 Methodological Contributions 243

10.3 Final Conclusion and Outlook 245

11 References 248

Chapter 1 Introduction

Language is more than a means of communication. It is a cognitive system that offers a window into the mind, providing insights into the nature, structure, and organization of human thought (Evans & Green, 2018). A long-standing question in psycholinguistics and cognitive linguistics is whether—and to what extent—linguistic structures, particularly grammatical categories, influence non-linguistic cognition, such as categorization, attention, memory, and decision-making. One such category is *grammatical gender*, which systematically assigns nouns to distinct classes (e.g., masculine, feminine, neuter) (Corbett, 1991; Dahl, 2000). Grammatical gender is found in many of the world's languages, although the exact proportion of gendered versus genderless languages remains unknown (Corbett, 2015).

The central question is whether grammatical gender, as a structural linguistic feature, leaves a trace on our mental representations. Do speakers of gendered languages conceptualize objects differently from those whose languages lack this feature? For instance, native speakers of German or Spanish may exhibit grammatical gender-congruent patterns in object categorization, while speakers of Chinese or Farsi—languages without grammatical gender—do not. This question is rooted in the framework of the *linguistic relativity hypothesis*, which posits that the structure of a language can influence or shape the way its speakers think (Whorf, 1956). In this context, the present dissertation critically examines whether grammatical structures such as gender have cognitive consequences, particularly in shaping patterns of object representation and categorization.

The question of whether grammatical gender shapes conceptual representations becomes even more complex in the case of multilingual speakers who navigate between systems with and without grammatical gender. Do the differing grammatical gender systems in the languages of bilingual and multilinguals lead to the development of hybrid or language-specific conceptual patterns for nouns and entities? Do their cognitive strategies regarding grammatical gender shift depending on the language context?

While previous research has provided some evidence that grammatical gender affects cognition in monolingual speakers (e.g. Boroditsky et al., 2003; Konishi, 1993; Sera et al., 2002), the extent, robustness, and generalizability of these effects remain contested. Moreover, much of the literature has focused on speakers of Western European languages, often within WEIRD (Western, Educated, Industrialized, Rich, and Democratic) populations. However,

even within this context, several Western European languages with rich grammatical gender systems—such as Czech—remain underrepresented in empirical research. As a result, little is known about the cognitive implications of grammatical gender in these languages, or more broadly, in bilingual and multilingual contexts.

1.1 Aims of the Dissertation

This dissertation pursues a multi-layered research agenda that combines typological mapping, systematic review, theoretical analysis, and experimental investigation. Specifically, it aims to:

1. Provide a synchronic overview of grammatical gender across Indo-European languages.
2. Systematically review empirical research on grammatical gender and cognition between 2013 and 2023, identifying methodological patterns and recurring limitations.
3. Analyze multilingual perspectives on grammatical gender, with particular attention to underrepresented languages such as Czech.
4. Experimentally test whether grammatical gender influences semantic property assignment in monolingual German speakers, using Farsi as a baseline.
5. Investigate real-time processing in Farsi–German bilinguals using eye-tracking methodology.

Together, these aims seek to clarify not only whether grammatical gender influences cognition, but how, when, and under what linguistic conditions such effects emerge.

1.2 Research Questions

This cumulative dissertation investigates the relationship between grammatical gender and cognition from cross-linguistic and multilingual perspectives. It is based on five interrelated studies. Together, these studies address the following central research questions:

- **RQ1:** What is the status of Indo-European languages with respect to the presence or absence of grammatical gender in their morphosyntactic systems?
- **RQ2:** What methodological approaches have been employed in psycholinguistic studies on grammatical gender and cognition between 2013 and 2023? Are there recurring limitations or design flaws that may account for inconsistencies in findings across studies?

- **RQ3:** From a theoretical perspective, how has grammatical gender been examined in multilingual contexts, and to what extent have underrepresented European languages—specifically Czech—been included in this line of research?
- **RQ4:** Does grammatical gender influence the assignment of semantic properties to objects and animals in monolingual speakers of a gendered language (e.g., German), and how do speakers of a genderless language (e.g., Farsi) differ in this regard?
- **RQ5:** Does grammatical gender modulate real-time cognitive processing in Farsi–German bilinguals, as measured through an eye-tracking visual world paradigm?

1.3 Overview of Thesis Structure

This dissertation is organized into ten chapters, moving from theoretical foundations to empirical investigations and, finally, to broader implications for linguistic relativity research. The following sections provide a detailed overview of each chapter and its contribution to the overall research aims:

- **Chapter 1: Introduction**
  This chapter outlines the background, rationale, and research questions of the dissertation. It introduces the central themes of linguistic relativity and grammatical gender, highlighting the significance of these topics for understanding language-cognition interfaces across monolingual and multilingual contexts.
- **Chapter 2: Cognition and the Linguistic Relativity Hypothesis**
  Given the dissertation's focus on the influence of grammatical structures on cognition, this chapter begins by defining cognition from multiple disciplinary perspectives to ensure theoretical precision, methodological clarity, and empirical interpretability. It then introduces the theoretical framework of linguistic relativity as it has been developed and applied in psycholinguistic research. Finally, it critically examines competing theories that challenge or refine the Whorfian hypothesis, thereby delineating the theoretical boundaries within which the present inquiry is situated.
- **Chapter 3: Grammatical Gender: Definitions, Typology, and Cross-Linguistic Variation**
  This chapter provides a comprehensive definition of grammatical gender as a morphosyntactic category and traces its typological distribution across the world's languages. It also includes a review of linguistic studies on gender, highlighting the

semantic, cognitive, and social dimensions that have shaped scholarly debates in this domain.

- **Chapter 4: Methodological Approaches in Language and Cognition Research**
  This chapter provides an overview of methodological strategies commonly employed in psycholinguistic research, with a particular focus on studies investigating the relationship between grammatical gender and cognition. It outlines the rationale for method selection in the present dissertation and critically examines key methodological considerations, including issues of internal and external validity, participant sampling, and task design. Special attention is given to behavioral and eye-tracking paradigms, both of which are employed in the empirical studies that follow.

The core of the dissertation is composed of five interconnected studies, each addressing a distinct but related aspect of the overall research questions:[1]

- **Study 1: Synchronic Overview of Grammatical Gender in Indo-European Languages**
  This foundational study maps grammatical gender marking across 413 Indo-European languages. It uncovers substantial WEIRD and Eurocentric biases in the literature, revealing that 215 Indo-European languages remain undescribed linguistically and that 61 lack grammatical gender altogether.
- **Study 2: Systematic Literature Review of the Influence of Grammatical Gender on Cognition (2013–2023)**
  This study critically evaluates 43 empirical articles to identify conceptual gaps, methodological inconsistencies, and limited language coverage in the existing literature. Emphasis is placed on experimental design, language sampling, and task selection, with the goal of establishing a more rigorous methodological agenda for future research.
- **Study 3: Language-Specific Patterns in a Multilingual Context: The Case of Grammatical Gender**
  Drawing on a theoretical analysis, this study explores the representation of grammatical gender in multilingual speakers, with a focus on Czech as an underrepresented Indo-European language. It highlights the lack of inclusion of richly

[1] Articles 1 to 5 (chapters 5-9) are reproduced in their original published format, including their original page numbering. The page numbers of the present dissertation are consistently displayed at the bottom centre of each page, whereas original page numbers of the published articles are retained in their respective positions.

gendered languages in cognitive and bilingualism research and argues for greater methodological standardization in participant profiling and study design.

- **Study 4: The Influence of Grammatical Gender on Cognition: The Case of German and Farsi**
  Using a behavioral experimental design, this study investigates whether grammatical gender influences how monolingual speakers of German describe objects and animals. Farsi, a typologically unrelated genderless language, is used as a comparison baseline. The findings emphasize the often-overlooked role of neuter gender in German and advocate for the inclusion of non-binary gender systems in cross-linguistic research.
- **Study 5: Grammatical Gender in a Grammatically Genderless Mind: Eye-Tracking Evidence from Farsi-German Bilinguals**
  This study uses an eye-tracking visual world paradigm to examine real-time noun-adjective processing in Farsi–German bilinguals who acquired German in adolescence. It distinguishes between different cognitive domains—perception, categorization, conceptualization, and worldview—thereby refining key constructs within the linguistic relativity framework.

Finally, chapter 10 synthesizes findings from the five studies and relates them to the overarching theoretical framework. It discusses implications for psycholinguistic theory, language typology, and bilingual cognition. Methodological contributions and directions for future research are also outlined.

1.4 Contribution to the Field

This dissertation contributes to the field in five interrelated ways.

Theoretical Advancement

Rather than treating linguistic relativity as a binary question of presence versus absence of cognitive effects, this dissertation refines the conceptual distinctions underlying the debate. It differentiates between perception, categorization, conceptualization, and worldview, thereby clarifying what is meant by "cognitive influence." By distinguishing between stable conceptual representations and language-specific online processing, it offers a system-sensitive account of grammatical gender effects that moves beyond strong determinism and simple null-effect claims.

Empirical Diversification

The dissertation expands the empirical base of grammatical gender research by incorporating underrepresented languages such as Czech and Farsi and by examining bilingual populations rarely included in cognitive studies. It demonstrates that current generalizations about grammatical gender are derived from structurally limited language samples and argues for broader typological inclusion in future research.

Methodological Clarification

Through a systematic review and the design of original experimental studies, this dissertation identifies methodological sources of inconsistency in the literature. It shows that reported effects are often task-dependent and sensitive to participant profiling, timing measures, and operational definitions. By integrating behavioral and real-time eye-tracking methods, it provides a more nuanced framework for testing linguistic relativity in controlled settings.

Extension to Multilingual Contexts

A central contribution of this work lies in extending linguistic relativity into multilingual contexts in a testable and theoretically precise way. By examining Farsi–German bilinguals, the dissertation demonstrates how cognitive processing adapts dynamically to the active grammatical system. This approach reframes linguistic relativity as context-sensitive and cognitively flexible rather than monolithic.

Interdisciplinary Relevance

Beyond psycholinguistics, the findings inform cognitive science, language typology, and second language acquisition research. By clarifying how grammatical systems interact with processing mechanisms, the dissertation contributes to broader discussions of cognitive plasticity and language-cognition interfaces.

Chapter 2 Cognition and Theoretical Framework

In the context of linguistic relativity, the question of whether and how grammatical features influence cognitive processes remains a central and contested issue (Lucy, 1992; Slobin, 1996; Wolff & Holmes, 2011). Terms such as *cognition*, *thought*, *worldview*, and *conceptualization* are often used interchangeably in the literature (Athanasopoulos, 2009; Flecken et al., 2015), frequently without precise definitional boundaries. This terminological ambiguity is especially pronounced in early research on linguistic relativity, where cognition was rarely defined explicitly and often functioned as an assumed rather than articulated construct. Only in more recent decades—particularly within cognitive science and cognitive linguistics—have scholars begun to clarify what is meant by cognition, not necessarily as the direct object of inquiry, but as a foundational concept shaping theoretical and methodological choices (Evans & Green, 2018). This chapter addresses this gap by adopting a delineated, multidimensional understanding of cognition informed by current approaches in psycholinguistics and cognitive science. Establishing a clear definition is essential for articulating the theoretical framework of this dissertation and for enabling a consistent and empirically grounded investigation into the cognitive implications of grammatical structure.

Building on the need for a more precise and consistent understanding of cognition in linguistic relativity research, this chapter lays the theoretical groundwork for examining how grammatical gender may influence cognitive processes. It begins by defining cognition and delineating its various levels, drawing on insights from cognitive science, psycholinguistics, neurocognition, and cognitive linguistics. These distinctions are crucial for identifying the specific domains—such as perception, categorization, conceptualization, and language production—through which grammatical features such as gender may interact with cognitive functioning.

The chapter then briefly turns to three influential frameworks—Cognitive Semantics, Langacker's Cognitive Grammar, and Levelt's model of language production—to clarify how each conceptualizes cognition and where grammatical gender may be situated within these paradigms. Although these approaches differ in scope and orientation, they offer complementary perspectives on how linguistic structures are mentally represented and processed. Cognitive Semantics and Cognitive Grammar, both rooted in the broader tradition of cognitive linguistics, argue that language reflects and shapes conceptual structures, aligning closely with the assumptions of linguistic relativity. In contrast, Levelt's model adopts a more

modular and procedural account of language production, in which conceptualization is a distinct stage in a sequence of cognitive operations. By juxtaposing these models, the chapter highlights differing assumptions about the relationship between language and thought, while also clarifying the conceptual pathways through which grammatical gender may influence cognitive processes. The final sections of the chapter revisit the linguistic relativity hypothesis in greater detail and critically examine alternative theoretical accounts, thereby establishing the conceptual landscape for the empirical investigations that follow.

2.1 Defining Cognition from Cognitive Linguistics Perspectives

Cognition is a foundational concept that spans multiple disciplines, each emphasizing different facets of mental activity. In cognitive science and psychology, cognition broadly refers to the mental processes involved in acquiring, organizing, and using knowledge, including perception, attention, memory, categorization, reasoning, problem-solving, and language (Anderson, 2010; Eysenck & Keane, 2020). To better understand the structure and complexity of these processes, cognitive science often distinguishes between different layers of cognition. Low-level cognition involves basic perceptual and sensorimotor operations such as attention and visual tracking (Mandler, 2007). Mid-level cognition includes mechanisms like learning, memory, and categorization (Anderson, 2010; Eysenck & Keane, 2020). At the top of this hierarchy, high-level cognition encompasses abstract and symbolic capacities, such as reasoning, planning, and language use (Baars & Gage, 2010). These distinctions facilitate more precise theoretical modeling and empirical investigation by allowing researchers to isolate specific processes while also examining their interactions.

Within cognitive linguistics, cognition is not simply a background mechanism supporting language—it is considered the very source of linguistic structure. Like cognitive science, cognitive linguistics defines cognition as a range of mental processes, including perception, memory, categorization, and attention (Evans & Green, 2018). However, it differs in its theoretical orientation: language is viewed not as an autonomous rule-based system, but as a reflection of general cognitive processes rooted in embodied experience and conceptualization. A central premise of this approach is that linguistic knowledge emerges through patterns of usage shaped by experience, rather than being governed by a specialized mental grammar detached from other cognitive functions.

Cognitive linguistics may be broadly divided into two subfields: *cognitive semantics*, which investigates the relationship between language, conceptual structure, and embodied experience; and *cognitive approaches to grammar*, which focus on the symbolic units that pair form with meaning (Evans & Green, 2018). Importantly, this division rests on the assumption that meaning precedes grammar—that is, an account of how meaning is mentally represented (cognitive semantics) must underpin any theory of grammatical structure. Langacker explicitly argues for this primacy of meaning in language structure (Langacker, 1987, pp.12-13) and notes that grammar should be understood as a meaningful system rather than as a purely formal apparatus (Langacker, 1991, p.2).

Cognitive semantics, a strand within the broader field of cognitive linguistics, offers further insight into how cognition may be understood in relation to language. Developed in the late 1970s and 1980s by scholars such as George Lakoff, Mark Johnson, Leonard Talmy, and Ronald Langacker (Johnson & Lakoff, 1980; Langacker, 1987; Talmy, 1983), this approach is grounded in the premise that linguistic meaning does not derive from direct reference to an objective external world but is constructed through embodied experience. According to this view, linguistic expressions are tightly linked to mental representations that arise from our perceptual, motor, and cultural interactions with the world. Language is therefore not a neutral conduit for labeling pre-existing entities, but an active process of conceptualization—of constructing meaning in real time through use.

This dynamic view of meaning construction leads cognitive semanticists to adopt what is often termed an encyclopedic model of semantics: words are not isolated carriers of fixed meanings but serve as access points to vast networks of conceptual knowledge (Evans & Green, 2018). Meaning, in this framework, emerges not merely from linguistic structure but from the interplay between linguistic form and embodied conceptual structure. As such, cognition is understood not only as the background mechanism that supports language, but as the very medium through which linguistic meaning arises (Langacker, 2008, p.3).

Extending these principles into the grammatical domain, Cognitive Grammar—developed by Ronald Langacker (2008, 1987, 1991) — offers a unified account of linguistic structure grounded in cognitive processes. Cognitive Grammar rejects the notion that grammar constitutes a self-contained, formal system separate from meaning. Instead, it views all grammatical constructions as symbolic pairings of form and meaning, encompassing not only lexical items but also grammatical morphemes and abstract syntactic patterns. These symbolic

units are learned through repeated language use, with recurring patterns becoming entrenched in memory as schematic constructions (Langacker, 1987, pp.59-60). In this sense, grammar is not an autonomous computational module but a structured inventory of symbolic conventions shaped by experience.

This framework also maintains that grammatical organization reflects the speaker's conceptual system: the ways in which speakers construe, categorize, and prioritize aspects of experience are encoded not only lexically but also grammatically. Syntax, therefore, is not a system external to meaning but one that emerges from—and is constrained by—the same conceptual structures that underpin lexical semantics (Langacker, 1991, pp, 12-13). As such, grammar is not merely compatible with cognition; it is itself a cognitive phenomenon.

This meaning-based view of grammar stands in sharp contrast to formalist approaches, which treat grammar as an autonomous computational system and regard semantics as a secondary byproduct. For instance, in Chomskyan generative grammar (Chomsky, 1957, 1965), syntax is modeled as a self-contained rule system that operates independently of meaning. The sentence *Colorless green ideas sleep furiously*, while semantically incoherent, is considered syntactically well-formed under this framework (Chomsky, 1957, p. 15). Such examples highlight the formalist emphasis on grammatical structure in isolation from conceptual content—a position fundamentally at odds with the cognitive linguistic view adopted in this dissertation.

Taken together, cognitive semantics and Cognitive Grammar offer a coherent, cognitively grounded understanding of language that places cognition—broadly construed to include perception, memory, categorization, and conceptualization—at the heart of both meaning and grammar. This theoretical orientation provides a crucial point of contrast to much of the literature on linguistic relativity, in which key terms such as cognition, conceptualization, worldview, and thought are frequently used interchangeably and without explicit definition (Athanasopoulos, 2009; Flecken et al., 2015). Such terminological ambiguity is not a minor theoretical oversight: it carries significant consequences for the empirical validity and interpretability of experimental findings. If researchers fail to define clearly what cognitive domain their work addresses—whether perceptual salience, conceptual categorization, or language production—they risk conflating distinct constructs and drawing conclusions that are not theoretically sound.

The schematic in Figure 1, adapted from Evans and Green (2018), illustrates the layered relationship between perception, conceptualization, and linguistic expression. In this model, a *percept* is grounded in sensory and experiential reality; a *concept* is the mental abstraction that organises and interprets that percept; and linguistic meaning arises when speakers select and structure aspects of that concept in order to communicate about their experiences. Cognitive linguistics stresses that this last step is not automatic: speakers must choose lexical items and grammatical constructions that construe the concept in particular ways, highlighting some features while backgrounding others. Finally, the linguistic *form* realises this construal as a symbolic expression. This framework reflects the core assumption that language emerges from embodied experience and mental representation, and it provides terminological clarity for the present study by distinguishing perception, conceptualisation, and the linguistic construal that mediates between them.

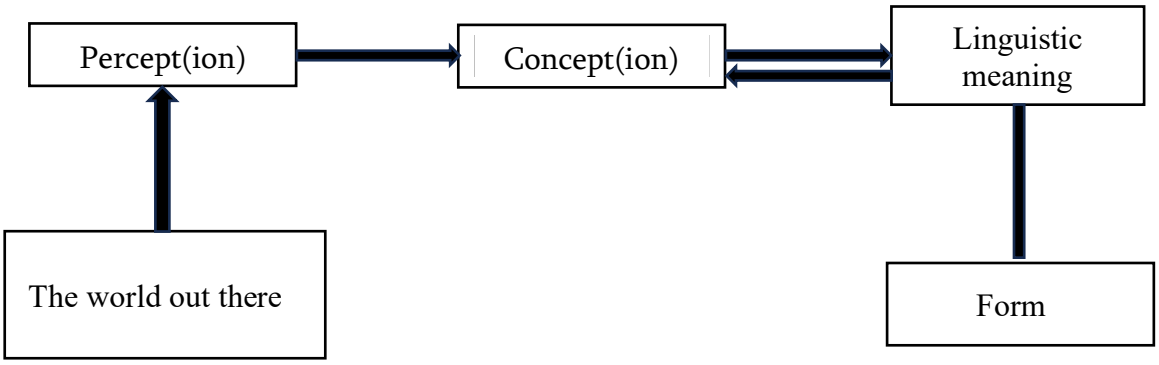


Figure 1 Levels of representation, adapted from Evans & Green (2018), p.7

Crucially, this schematic serves to disambiguate two frequently conflated notions in linguistic relativity research: perceptualization and conceptualization (Feinmann, 2020). In the context of this dissertation, perceptualization refers to the sensory grounding of experience, whereas conceptualization denotes the mental structuring of that experience into concepts that are linguistically expressible (Evans & Green, 2018; Feinmann, 2020; Langacker, 1987; Sloutsky, 2010). When this study investigates whether grammatical gender influences conceptualization, it refers to this intermediate level—concepts—not raw percepts nor the linguistic surface alone. In this way, Figure 1 offers a conceptual map of where grammatical structures such as gender might exert influence and how such influence might be detected or misidentified depending on the methodology employed.

As later chapters will demonstrate, different experimental tasks tap into different layers of this representational hierarchy. Some may reveal effects at the level of perceptual salience, others at the level of conceptual categorization or lexical retrieval. Without clearly distinguishing these levels, there is a risk of conflating linguistic influence on meaning construction with influence on perception or worldview—an error that has compromised theoretical clarity in previous research. The schematic therefore functions not only as a theoretical guidepost but also as a methodological anchor for the empirical investigations to follow.

2.2 Language Production and Grammatical Gender: Levelt's Model

While cognitive linguistics emphasizes the symbolic and experiential grounding of linguistic structure, psycholinguistic models of language production offer a complementary perspective by illuminating how language is processed during real-time communication. One of the most influential frameworks in this domain is the model of speech production developed by Willem Levelt (Levelt, 1989; Levelt et al., 1999). This model is especially relevant for the present dissertation, as it offers a procedural account of how grammatical features—including grammatical gender—are activated and selected during language use.

According to Levelt's model, language production unfolds in three major stages:

1. **Conceptualization**: the speaker forms a preverbal message based on communicative intent and background knowledge.
2. **Formulation**: the message is translated into linguistic form through two sub-processes:
    - **Lemma retrieval**: selecting abstract lexical entries that contain syntactic features such as grammatical gender, number, and word class.
    - **Morpho-phonological encoding**: retrieving phonological word forms and applying inflectional morphology.
3. **Articulation**: the motor execution of the planned utterance.

Based on this model, grammatical gender is particularly salient during the lemma retrieval sub-stage in formulation, where speakers retrieve not only the semantic content of a word but also its associated grammatical features (e.g., gender, number, syntactic category). Even in contexts where gender is not overtly marked on the surface—such as in elliptical constructions or experimental tasks like picture naming, where participants observe visual stimuli and are prompted to produce noun phrases—research suggests that gender is still automatically

activated during speech production (Schriefers et al., 2002). This makes Levelt's model particularly pertinent for understanding how grammatical gender may influence lexical access, response latencies, and conceptual processing, even in tasks that do not require explicit gender marking.

In German, for example, gender is encoded not only in articles (*der*, *die*, *das*) and adjectives but also plays a key role in nominalization. A single adjective like *schön* ('beautiful') can be nominalized differently depending on the grammatical gender of the implied referent, yielding *der Schöne* ('the handsome one'—masculine) or *die Schöne* ('the beautiful one'—feminine). These gendered forms typically evoke distinct conceptual interpretations, underscoring how grammatical structures can shape meaning even in the absence of explicit nouns. This morphosyntactic difference is not merely grammatical; it correlates with different mental representations. While nominalization is not the focus of this dissertation, it illustrates how grammatical gender can influence conceptualization, aligning well with the procedural stages described by Levelt.

The broader theoretical perspective adopted in this dissertation is that grammatical structures—particularly grammatical gender—may function not merely as formal linguistic categories but also as cues that interact with speakers' conceptual representations during language use. For speakers of languages that encode gender morphosyntactically, the hypothesis is that during the conceptualization stage of language production, nouns may become associated with specific grammatical gender features. This linkage shapes the mental representation of referents in such a way that speakers are cognitively predisposed to produce grammatical gender-congruent constructions—for example, choosing *die Schöne* rather than *der* or *das Schöne* in contexts where the noun conceptually aligns with feminine gender. Thus, grammatical gender is not only retrieved at the lemma level during formulation but is also deeply embedded in the conceptual framework guiding language production and interpretation.

From this perspective, Levelt's model provides a useful framework for examining how grammatical gender may influence lexical access and conceptual processing. If grammatical gender forms part of the lexical information stored at the lemma level, it is plausible that gender features become activated automatically during language production, even in tasks where explicit gender marking is not required. Such activation could influence processes such as adjective selection or referential interpretation during speech production and comprehension.

The empirical studies presented in the later chapters of this dissertation are designed to examine these possibilities by comparing speakers of German—a language with grammatical gender—and speakers of Farsi, a language that lacks grammatical gender. This contrast provides an opportunity to explore whether the presence or absence of gender marking in a language system influences how speakers access lexical representations and process referential information during language use.

Although Levelt's model was originally developed to describe language production, its procedural architecture and staged processing framework have significantly influenced theories of language comprehension as well. Increasingly, language production and comprehension are regarded as overlapping processes that rely on shared cognitive and neural mechanisms, especially during early lexical and syntactic stages (Pickering & Garrod, 2013). This bidirectional view is supported by eye-tracking research demonstrating that listeners rapidly integrate grammatical cues—such as gender-marked articles and adjectives—to anticipate upcoming referents in real time (e.g. Lew-Williams & Fernald, 2007). Notably, this predictive use of grammatical gender emerges early in development: even toddlers aged 34–42 months show sensitivity to gender cues during online comprehension, suggesting that grammatical gender plays a guiding role from an early age (Lew-Williams & Fernald, 2007). Expanding on this, Pickering and Garrod's (2013) interactive alignment model argues that comprehension involves covert simulations of production, further blurring the boundaries between these two domains.

At the same time, researchers have pointed to limitations in Levelt's serial, feed-forward architecture when applied to the dynamic and predictive nature of real-time comprehension. For instance, although Levy (2008) does not directly critique Levelt, his expectation-based framework emphasizes the probabilistic and context-sensitive features of comprehension, which implicitly challenge the assumptions of a rigid, modular design. Likewise, Brouwer, Fitz, and Hoeks (2012) highlight the role of discourse-level inferencing, arguing that such flexible, top-down processes are not easily accommodated within Levelt's modular and linear framework. In the production domain, alternative models such as the cascading activation framework (e.g. Rapp & Goldrick, 2000) suggest that lexical, syntactic, and phonological information may be processed in overlapping and interactive stages.

Empirical studies further question key assumptions of Levelt's model: for example, Mahon et al. (2007) demonstrated that lexical selection may not involve competitive inhibition between

semantically related words, as previously assumed, but rather facilitation, challenging the idea that lexical access operates via strict competition at the lemma level. In addition, Hartsuiker, Pickering, and Veltkamp (2004) showed that syntactic structures can be primed both within and across languages, indicating that syntactic processing is more flexible and interactive than the modular, feed-forward design proposed by Levelt. Together, these studies support interactive and parallel accounts of lexical and syntactic retrieval, where semantic and syntactic information are activated dynamically and contextually, reflecting a more nuanced and integrated process than Levelt's original architecture suggests. Electrophysiological evidence also points to rapid, automatic integration of extralinguistic information into syntactic processing, as demonstrated by Alekseeva et al. (2022). Their ERP findings reveal early brain responses sensitive to mismatches between speaker voice gender and grammatical gender, suggesting an interactive processing mechanism that integrates linguistic and extralinguistic cues in real time. Such results further problematize the serial, feed-forward assumptions of Levelt's speech production model, highlighting the complex, dynamic nature of language processing that transcends strictly modular stages.

Nonetheless, Levelt's model remains a valuable theoretical scaffold—particularly its distinction between lemma retrieval and morpho-phonological encoding. In the present dissertation, this framework provides a conceptual basis for examining how grammatical gender may interact with both expressive and receptive aspects of language processing, as explored in the empirical chapters that follow.

2.3 Linguistic Relativity

Having established the cognitive foundations of language processing—from conceptualization to formulation and articulation, as well as comprehension to production—this chapter now turns to the central theoretical lens of the present study: linguistic relativity. Also known as the Sapir–Whorf Hypothesis, linguistic relativity posits that the structure of a language can influence how its speakers perceive, categorize, and interact with the world (Whorf, 1956)[2]. This framework provides the conceptual grounding for investigating how

[2] The volume *Language, Thought, and Reality: Selected Writings of Benjamin Lee Whorf (Whorf, 1956*) is a posthumous collection of Whorf's previously unpublished manuscripts, lecture notes, and papers. These materials were compiled and edited by his former colleague and friend, John B. Carroll, and published in 1956, fifteen years after Whorf's death in 1941.

grammatical gender may subtly shape patterns of thought in speakers of gendered languages, and how these patterns may persist, shift, or attenuate in bilingual contexts.

2.3.1 Historical Roots of the Language–Thought Relationship

The idea that language influences human thought predates the 20th century and is deeply embedded in 18th- and 19th-century philosophy. Early Romantic thinkers such as Johann Gottfried von Herder (1744–1803) and Johann Georg Hamann (1730–1788) rejected Enlightenment universalism, arguing instead that each speech community possesses a distinctive mode of thought, expressed and maintained through its language (Pavlenko, 2014, p.2). In his *Treatise on the Origin of Language* (1772/2002), Herder maintained that language not only reflects but also structures thought, with linguistic differences corresponding to the lived realities and cultural needs of their speakers rather than implying cognitive hierarchies.

Among these early theorists, Wilhelm von Humboldt (1767–1835) offered one of the most influential and conceptually precise accounts of the relationship between language and cognition. He described each language as a self-contained system that "draws a circle" around its speakers (Humboldt, 1836/1963, p. 294), shaping how reality is apprehended and represented. At the same time, Humboldt was deeply concerned with how these linguistic boundaries might be transcended—whether through the refinement of one's native language, through literature (and poetry in particular), or through the acquisition of foreign languages (Humboldt, 1836/1963). These avenues, in his view, allowed individuals to expand their conceptual repertoire beyond the constraints of a single linguistic system. While Humboldt did not explicitly formulate them in these terms, his framework anticipates two ideas that remain central in contemporary cognitive science:

1. The difference between implicit, grammar-shaped patterns of thought and explicit, culturally mediated beliefs.
2. The potential influence of bilingualism on conceptualization and mental representation.

2.3.2 Humboldt's Weltansicht and Weltanschauung

Wilhelm von Humboldt's philosophical reflections on language and thought laid the groundwork for what later scholars conceptualized as the distinction between *Weltansicht* and *Weltanschauung*. While Humboldt himself used these German terms in his writings (Humboldt,

1836/1963), their precise theoretical differentiation was more fully elaborated by subsequent linguists and philosophers (Duranti, 1997; Pavlenko, 2014).

Weltansicht — often translated as "view of the world" or "worldview" — refers to the largely unconscious, habitual ways in which language structures the perception and categorization of experience. It reflects the fundamental capacity of the mind to process the world through language, shaping implicit conceptual patterns embedded within grammatical and lexical systems. These patterns guide everyday language use and cognition without requiring active, conscious reflection.

Weltanschauung, by contrast, denotes an explicit, reflective worldview or ideology shaped by cultural, social, and historical contexts. Unlike *Weltansicht*, *Weltanschauung* is not determined solely by language but constitutes a more complex system of beliefs and values that individuals hold and articulate.

Underhill (2009) emphasizes that for Humboldt, *Weltansicht* represents "the cornerstone of his linguistic philosophy" (p. 55) — the way speakers habitually follow linguistic patterns to negotiate and structure reality beneath the level of conscious thought.

Grammatical gender, a core focus of the present dissertation, exemplifies a *Weltansicht*-level category: it operates unconsciously to influence how speakers implicitly classify and conceptualize objects and concepts. This distinction is essential for investigating whether bilingual speakers whose other language lacks grammatical gender (such as Farsi) exhibit transfer or attenuation of these implicit patterns when processing adjectives in German (Chapter 8).

Finally, Humboldt's insight into language learning as "a new standpoint toward one's worldview" (Humboldt, 1836/1963, p. 294) directly connects to contemporary research on bilingualism and cross-linguistic influence, suggesting that while explicit worldviews may shift with new linguistic contexts, the implicit *Weltansicht* patterns may persist or interact in complex ways.

2.3.3 From Humboldt to Sapir and Whorf

Building on the foundational insights of Wilhelm von Humboldt, the early 20th century witnessed the emergence of a more explicitly articulated theory of language–thought relations in the work of Franz Boas, Edward Sapir, and his student and collaborator Benjamin Lee Whorf.

Boas (1858–1942), often regarded as the father of American anthropology, emphasized the interrelatedness of language, culture, and cognition, advocating for the study of linguistic diversity as essential to understanding human experience. His empirical work and methodological rigor influenced Sapir, who further developed the idea that language is not merely a tool for expressing thought but a framework that shapes habitual patterns of our worldview (Sapir, 1921). Whorf (1956), extending Sapir's perspective, argued that the grammatical and lexical categories embedded in a language influence how its speakers habitually attend to, segment, and interpret reality.

Notably, the term "*Sapir-Whorf hypothesis*" was first introduced by Hoijer (1956) to characterize the idea that language functions not only as a means of communication but also as a device through which experience itself is defined and structured.

Whorf's analyses—most famously of Hopi and Standard Average European—led him to propose that linguistic structure exerts a probabilistic, though not absolute, influence on cognition (Whorf, 1956). Crucially, both Sapir and Whorf conceived of linguistic relativity as a tendency rather than a deterministic law. This view resonates with Humboldt's metaphor of language "drawing a circle" around its speakers, shaping but not imprisoning the mind.

Despite their profound conceptual contributions, Sapir and Whorf did not formulate precise, operational hypotheses amenable to systematic empirical testing, a gap that has led to the critique of their framework as more heuristic than scientific (Lucy, 1997). The premature deaths of both scholars—Sapir in 1939 at the age of 55 and Whorf in 1941 at the age of 44—undoubtedly curtailed the development of a unified, rigorous research program, leaving their work open to reinterpretation and, at times, oversimplification (Pavlenko, 2014).

One notable oversimplification was the emergence of the so-called strong version of linguistic relativity, or linguistic determinism, which posits that language rigidly determines thought and precludes alternative conceptualizations. Although often attributed to Sapir and Whorf, neither explicitly endorsed this deterministic view. Instead, the strong interpretation largely emerged as a mid-20th-century construct, popularized by both proponents and critics (e.g. Brown, 1958; Brown & Lenneberg, 1954) seeking a falsifiable proposition. The weaker and now more widely accepted position—that language influences but does not dictate cognition—aligns more closely with Sapir's and Whorf's original writings, which recognized

cognitive flexibility, the possibility of bilingualism, and potential for conceptual change (Lucy, 1992; Pavlenko, 2014).

This nuanced understanding of linguistic relativity provides a conceptual bridge to the present study's focus on grammatical gender. In languages such as German, grammatical gender may operate at the level of Humboldt's *Weltansicht*, subtly influencing implicit classification and conceptualization. The key question—central to the empirical work of this dissertation—is whether and how such implicit patterns persist, shift, or attenuate when speakers are bilingual in a language that lacks grammatical gender, such as Farsi.

2.3.4 The Political and Historical Contingencies of Early Interpretations: Monolingualism as National Ideal

Although early interpretations of Humboldt's, Sapir's, and Whorf's ideas often framed language as deterministically shaping thought, this monolingual misreading was neither accidental nor purely scholarly (e.g., Pavlenko, 2014; Weinreich, 1953). Instead, it was deeply influenced by the political and social context of the early 20th century, a period when monolingualism was widely regarded as the cultural and national norm (Anastasi & Cordova, 1953; Bossard, 1945; Pavlenko, 2014). Emerging nation-states across the globe promoted linguistic homogeneity as a foundation for identity and unity, casting multilingualism as a threat to social cohesion and intellectual clarity (B. Anderson, 1991; Jespersen, 1922; Weinreich, 1953). This ideological backdrop shaped scholarly perspectives, encouraging interpretations of linguistic relativity that emphasized language boundaries as rigid cognitive limits. Consequently, the complex and nuanced views of language's role in thought—originally conceived as more flexible and dynamic—were often overshadowed by politically motivated assumptions favoring monolingualism as the ideal.

In the specific context of early 20th-century America, this misreading was further intensified by a political and historical climate that strongly promoted monolingualism as a marker of national loyalty and cultural assimilation. During and after World War I, rising anti-immigrant sentiment precipitated widespread suppression of languages other than English, manifested through laws banning foreign-language instruction in schools and the removal of German-language materials from public life (Fitz-Gerald, 1918; Luebke, 1980; Pavlenko, 2002). President Roosevelt's 1919 declaration that "we have room for but one language here, and that is the English language … and we have room for but one sole loyalty" (Roosevelt. 1919, as cited in Pavlenko, 2002, p. 184) encapsulates this political push toward linguistic

assimilation as a necessary element of patriotism and unity. Educational and academic institutions further reinforced this stance by regarding bilingualism as detrimental, relegating multilingualism to the realm of elite privilege rather than a societal norm (Pavlenko, 2014).

Notably, key figures in linguistic anthropology and linguistic relativity, such as Franz Boas, Edward Sapir, and Benjamin Lee Whorf, themselves advocated for linguistic pluralism. They emphasized the cognitive and cultural value of diverse languages and viewed multilingualism as a corrective to the "tyrannical hold" of any single linguistic framework (Sapir, 1929; Whorf, 1956). The irony of the political climate in the United States during this period becomes apparent when contrasted with Soviet multilingual policies, which—despite the regime's totalitarian reputation—were comparatively liberal and supportive of minority languages. This contrast illustrates that monolingualism as the dominant ideological paradigm in the U.S. was not an inevitable consequence of linguistic science but rather a product of specific political and social forces (Pavlenko, 2014).

Thus, the widespread monolingual interpretation of Humboldt's, Sapir's, and Whorf's ideas must be understood as historically and politically contingent rather than purely scholarly or linguistic in nature. Recognizing this context is essential for a fuller and more accurate appreciation of their work, as well as for contemporary investigations into how language shapes thought in multilingual individuals and societies. Researchers interested in bilingualism and multilingualism, aware of the complexity of these phenomena, have played a central role in these discussions, not only as proponents supporting Whorfian claims but also as critical voices challenging and refining the hypothesis. Recognizing that these early interpretations were shaped as much by sociopolitical forces as by scholarly debate is essential for accurately situating contemporary research. This historical awareness also underlines the need to examine linguistic relativity through empirical evidence rather than ideological assumptions. In the following section, I therefore turn to a range of experimental and observational studies that have investigated whether, and under what conditions, the structures of language influence cognitive processes.

2.3.5 Empirical Studies on Language and Cognition

In recent decades, the linguistic-relativity debate has been recast in empirical terms. Researchers have moved from anecdotal reports to controlled experiments that ask whether language influences perception, categorization, and reasoning in specific domains. A rich body

of work now shows that linguistic categories can bias cognitive processing, but these effects tend to be probabilistic and task-dependent rather than deterministic.

Evidence from colour perception is a case in point. Speakers discriminate hues more quickly when their language distinguishes them lexically (e.g., light versus dark blue) than when they fall under a single term (e.g. Brown et al., 2011; Franklin et al., 2005; Witzel & Gegenfurtner, 2013). Similar cross-linguistic differences have been documented for spatial frames of reference, object relations, and object properties (Imai & Gentner, 1997; Park & Ziegler, 2014). Other studies extend the inquiry to abstract domains such as quantity, time and motion: languages with obligatory number markers affect speakers' reasoning about quantity (Athanasopoulos, 2006); metaphors for time correlate with spatial conceptualisation of temporal sequences (Boroditsky, 2001; Casasanto et al., 2004); and speakers of satellite- and verb-framed languages attend to manner or path when describing motion events (Slobin, 1996, 2004). These findings collectively suggest that language guides attention and construal in domains where grammatical or lexical distinctions are salient.

Importantly, some of these results have spurred methodological debates. Studies on colour perception, for example, raise questions about the role of perceptual salience versus linguistic encoding (A. M. Brown et al., 2011; Franklin et al., 2005). Scholars have therefore argued that grammatical categories may offer a more reliable test bed than purely lexical categories (Montefinese et al., 2019). Dan Slobin's (1996, 2004) thinking-for-speaking hypothesis proposes that speakers habitually attend to those aspects of a scene that must be encoded grammatically; cross-linguistic work on motion events supports this view by showing that English and Spanish speakers highlight different components of a motion depending on their language's encoding requirements. Mertins's (2018) seeing-for-speaking framework builds on this insight by demonstrating, through eye-tracking, that grammatical aspect and word order modulate visual attention during language production.

Beyond these domains, research on grammatical gender—reviewed in detail in Chapter 3—illustrates how morphological systems can bias semantic categorisation and even implicit cognition. Experiments in languages like Polish show that participants' similarity judgments and reaction times are influenced by the grammatical gender of nouns (Maciuszek et al., 2019). Nevertheless, gender effects vary with the structure of a language's gender system and with task demand (Maciuszek et al., 2019), underscoring the need for careful cross-linguistic comparison.

Overall, the empirical literature supports a moderate, usage-based version of linguistic relativity. Language can bias attention and conceptualisation, but its influence depends on the domain, the linguistic distinctions involved, and the speaker's experience (Mertins & Delucchi Danhier, 2023). This thesis builds on that foundation by investigating a grammatical category—gender—that has been understudied relative to other domains, and by applying both behavioural and eye-tracking methods to monolinguals and bilinguals.

2.3.6 Critical Assessment of Linguistic Relativity

The empirical findings reviewed in the previous section suggest that strong deterministic formulations of the Sapir–Whorf hypothesis are difficult to support. Language does not generate entirely new cognitive categories; instead, it appears to bias the salience of certain features or distinctions under specific circumstances. Studies of colour naming, spatial frames of reference, and motion description reveal that speakers of different languages sometimes attend to different aspects of the same perceptual scene (e.g., Casasanto et al., 2004; Slobin, 1996, 2004). Yet these effects are far from ubiquitous, and their magnitude depends on factors such as task demands, attentional focus, and developmental stage. Such variability underscores the need for carefully designed cross-linguistic experiments and nuanced interpretation of results.

Complementary theoretical perspectives place important constraints on how far language can shape cognition. Chomsky's Universal Grammar (Chomsky, 1965, 1993) posits that the capacity for language rests on an innate, species-specific architecture. This view treats grammatical diversity as surface variation on a shared cognitive foundation, implying that linguistic differences cannot wholly explain differences in thought. Similarly, Pinker (1995) and Fodor (1975) emphasize that mental representations and conceptual structures exist independently of linguistic form, implying that cognitive capacities can operate without language. Piaget (1952) further highlights universal developmental stages that shape reasoning before linguistic influence. Related arguments from cognitive psychology propose that conceptual representations and developmental trajectories exist independently of language. Infants distinguish between objects and substances before they learn to speak (Soja et al., 1991), and early conceptual structures are elaborated—but not created—through linguistic exposure (Imai & Gentner, 1997). Cross-cultural evidence further shows that languages lacking categories such as grammatical gender or tense support the same cognitive functions as

languages that mark them (Mertins & Delucchi Danhier, 2023). Collectively, these perspectives remind us that any linguistic influence operates against a backdrop of shared human cognition.

Viewing linguistic relativity through this lens shifts the focus from whether language determines thought to how language interacts with pre-existing cognitive systems. Research on bilinguals provides a rich testing ground for this interaction. Speakers who command two languages with contrasting grammatical systems often adjust their cognitive strategies depending on the language being used (Athanasopoulos, 2007; Grosjean, 1998). Such adjustments are probabilistic and modulated by factors like frequency of exposure, pragmatic relevance and social context (Boroditsky et al., 2011; Cook et al., 2006). Proposed mechanisms include the ease with which certain distinctions can be coded linguistically, the transfer of analogical patterns across languages, and the structuring of discourse (Imai & Gentner, 1997; Lucy, 2014; Slobin, 1996). Effects involving grammatical gender illustrate this conditionality: speakers are more likely to attribute masculine or feminine qualities to objects when their language grammatically encodes gender but the strength and direction of the effect vary with the language's gender system and the experimental task (Maciuszek et al., 2019; Samuel et al., 2019).

Taken together, these lines of evidence support a soft-relativity position in which language influences cognition without overriding universal cognitive capacities. In this view, grammatical and lexical distinctions guide attention and categorisation in a probabilistic manner while fundamental cognitive structures remain stable across linguistic communities (Athanasopoulos & Bylund, 2023; Boroditsky et al., 2011). Critical assessment requires integrating structural, functional, and interactional dimensions of language, while recognizing the overarching influence of universal cognitive constraints (Lucy, 2014; Slobin, 1996). In this sense, the "opposing" theories are not in conflict with linguistic relativity per se; rather, they highlight that language operates on top of pre-existing cognitive capacities, and its effects are contingent, probabilistic, and context-sensitive.

2.4 Conclusion

This chapter has laid the conceptual groundwork for the dissertation by clarifying the layered nature of cognition and synthesising insights from cognitive linguistics and psycholinguistics. By tracing the historical and political contexts of the linguistic-relativity debate, it underscored that linguistic categories guide attention and categorisation in

probabilistic, task-dependent ways. Evidence from colour perception, spatial reasoning, motion description, and other domains points toward a soft-relativity view in which language biases, but does not determine, cognitive processes. Within this broader landscape, grammatical gender emerged as an especially intriguing case. Widely attested yet typologically diverse, gender systems provide a lens for examining how morphosyntactic categories may influence conceptual organisation across both monolingual and bilingual speakers.

Chapter 3 Grammatical Gender in Cross-Linguistic Perspective and Its Role in Linguistic Relativity

This chapter focuses specifically on grammatical gender, outlining its formal definitions, typological variation, and the central debates in gender research. Building on the cognitive framework developed in the previous chapter, it situates grammatical gender within a cross-linguistic perspective and highlights why languages beyond the usual Western European sample—including those without grammatical gender systems—are essential for a comprehensive account of language–cognition interactions. Establishing this typological foundation is necessary before evaluating claims about linguistic relativity, since differences in gender systems must first be described with structural precision before their potential cognitive implications can be assessed. This discussion also prepares the ground for the empirical core of the dissertation (Chapters 5–9), which investigates the cognitive consequences of grammatical gender through behavioural and eye-tracking experiments while addressing the WEIRD and Eurocentric biases identified in previous research.

3.1 Introduction

Grammatical gender[3] is a morphosyntactic system of noun classification that systematically controls agreement across elements of the clause. Because it combines semantic motivation, structural obligatoriness, and cross-linguistic variation, it has become a central phenomenon in typology, psycholinguistics, sociolinguistics, and linguistic relativity research. Gender systems differ in the number of classes they distinguish, the principles governing

[3] Throughout this dissertation, the term gender refers to grammatical gender unless otherwise specified. When sex-based distinctions are relevant, the term biological gender is used for clarity.

assignment, and the morphological domains in which agreement is realized across the clause (Corbett, 1991; Walter et al., 2021).

This chapter situates grammatical gender within a cross-linguistic typological framework before examining how it has been treated across major linguistic traditions. Establishing this typological foundation is essential for evaluating claims about linguistic relativity: differences in gender systems must first be described with structural precision before their potential cognitive consequences can be assessed.

The chapter proceeds by outlining principles of gender assignment, patterns of cross-linguistic variation, and the functional role of agreement, before turning to psycholinguistic, sociolinguistic, and linguistic relativity perspectives.

3.2 Defining Grammatical Gender and Cross-Linguistic Variation

Grammatical gender is a morphosyntactic system for classifying nouns (Corbett, 1991). In typological research, it is defined as the division of nouns into two or more classes whose membership is reflected in the agreement behaviour of associated elements such as determiners, adjectives, pronouns, or verbs (Corbett, 1991; Kramer, 2015, 2020). Crucially, gender is identified not simply by semantic distinctions, but by systematic agreement patterns that are grammatically encoded across the clause.

Typological research on grammatical gender aims not only to document these classificatory systems across languages but also to identify cross-linguistic regularities, map the range of structural variation, and explain their functional motivations (Corbett, 2013; Seifart, 2010). This comparative perspective provides the structural basis for understanding how gender systems differ in organization, transparency, and distribution.

Languages typically mark gender via agreement on determiners, adjectives, verbs, or other targets, such that the gender of the noun triggers specific agreement patterns (Corbett, 1991; Greenberg, 1978; Hockett, 1958; Kramer, 2015; Nicholas, 1992, inter alia). In the Romance language Spanish, for example, masculine and feminine nouns trigger distinct adjectival and pronominal forms (e.g., *niño alto* vs. *niña alta*). German, by contrast, exhibits a three-gender system (masculine, feminine, neuter), in which agreement distinctions are realized across articles, adjectives, and pronouns (e.g., *der alte Tisch*, *die alte Lampe*, *das alte Buch*). Other

languages, such as Farsi or Mandarin Chinese, lack grammatical gender altogether and do not systematically encode gender distinctions through agreement morphology.

Grammatical gender, as an obligatory agreement-triggering feature, differs from other nominal classification phenomena. Declension classes, for example, may organize nouns morphologically without systematically controlling agreement. Similarly, classifier systems, as found in languages such as Mandarin Chinese, categorize nouns but do not trigger agreement across the clause (Corbett, 1991). This distinction is particularly important because the presence of systematic agreement ensures that gender features are repeatedly activated in language use, potentially increasing their cognitive salience.

A more detailed examination of Indo-European gender systems is provided in Chapter 5. The aim of the present section is not to duplicate those analyses, but to establish a shared typological background that frames the subsequent discussion of assignment principles, cross-linguistic diversity, and functional motivations.

3.2.1 A Semantic Core of Animacy and Humanness

Cross-linguistic research consistently demonstrates that grammatical gender systems typically include a semantic core, whereby a subset of nouns is assigned gender on the basis of animacy, humanness, or biological sex (Corbett, 1991; Greenberg, 1978; Hockett, 1958; Kramer, 2015; Nicholas, 1992). Kramer (2015, 2020) identifies the semantic core generalization—the tendency for grammatical gender assignment to align, at least in part, with animacy, humanness, or biological sex—as a foundational typological property that differentiates grammatical gender from other φ-features such as number. In formal linguistic frameworks, φ-features (phi-features) refer to morphosyntactic agreement features associated with nominals, typically including person, number, and gender, which are encoded in syntactic structure and reflected in agreement morphology (Chomsky, 2000; Harley & Ritter, 2002).

While Kramer's account is firmly situated within the morphosyntactic domain, it is reasonable to hypothesize, as an interpretive extension, that the semantic grounding inherent in this core may function as a form of cognitive anchoring, providing speakers with a conceptual basis for certain aspects of gender assignment. Although this cognitive interpretation is not explicitly advanced in Kramer's work, it opens the possibility that gender categories—by aligning with salient semantic distinctions—could influence the way speakers mentally organize their world.

Nevertheless, beyond this semantic nucleus, the relationship between a noun's meaning and its grammatical gender often becomes opaque. In many languages—such as French and German—grammatical gender frequently diverges from the semantic properties of the referent. For example, in French, *la table* ("table") is grammatically feminine despite denoting an inanimate object, and abstract nouns such as *la masculinité* ("masculinity") are also grammatically feminine despite their semantically masculine content (Chen & Faitaki, 2025). Such mismatches highlight the partially arbitrary and lexically determined nature of gender assignment outside the semantic core, even in languages where gender plays a central role in morphosyntax.

This duality between semantic motivation and arbitrariness has made grammatical gender a focal point for linguistic relativity research, since it raises the question of whether such systems can bias conceptualization even when their assignments are semantically opaque.

3.2.2 Morphological, Phonological, and Arbitrary Assignment

Beyond the semantic core, gender assignment can depend on morphological, phonological, or arbitrary criteria. Morphological assignment links gender to particular derivational affixes or inflectional classes. In many Bantu languages, such as Swahili, noun class prefixes correlate with agreement patterns: nouns beginning with *m-* (Class 1) trigger singular agreement (*mtoto mzuri* "good child"), while their plural counterparts with *wa-* (Class 2) trigger plural agreement (*watoto wazuri* "good children") (Katamba, 2003).

Phonological assignment is rarer, but in Miya (a Chadic language), nouns ending in /-i/ are typically feminine, whereas those ending in /-a/ are typically masculine (Schuh, 1989).

Importantly, not all languages rely on morphological or phonological cues: some assign gender semantically for a subset of nouns and arbitrarily for the rest. The arbitrariness of gender assignment explains why even closely related languages may assign different genders to cognates; for example, the word for "sun" is feminine in German (*die Sonne*) but masculine in Spanish (*el sol*), despite both descending from Proto-Indo-European *sóh₂wl̥* (Corbett, 1991).

Moreover, a few languages assign gender solely on semantic grounds—for example, in Tamil, masculine gender is assigned to male humans, feminine gender to female humans, and neuter gender to all inanimate objects and most animals—whereas most languages exhibit a mixture of semantic, morphological, and arbitrary assignment (Corbett, 1991; Lehmann, 1993).

This opacity highlights an important challenge for linguistic relativity claims. If cognitive effects of grammatical gender are observed only for nouns whose gender assignment aligns with biological sex or animacy, such effects could plausibly be attributed to semantic associations rather than to grammar per se. However, if speakers exhibit systematic cognitive biases even for nouns whose gender assignment is semantically opaque or historically arbitrary, then these effects cannot be reduced to lexical meaning alone. Instead, they would suggest that the grammatical system itself—through its obligatory agreement patterns and repeated activation during language use—may contribute to shaping conceptual representations.

From this perspective, arbitrariness becomes theoretically revealing rather than problematic. It provides a critical testing ground: precisely because certain gender assignments lack transparent semantic motivation, any consistent cognitive patterns associated with them would point toward the influence of morphosyntactic structure rather than conceptual content. For linguistic relativity research, this distinction is crucial, as it separates effects driven by shared semantic knowledge from those potentially arising from the structural properties of language.

3.2.3 Cross-Linguistic Diversity in Gender Systems

Grammatical gender is widely attested but varies greatly across languages. Most languages with gender have between two and four distinct classes, though some exhibit far more. Common systems include masculine–feminine, masculine–feminine–neuter, animate–inanimate, and common–neuter distinctions (Corbett, 1991). Masculine–feminine systems predominate in Romance, Baltic, and many Afroasiatic languages, whereas three-gender systems (masculine–feminine–neuter) are typical of Germanic, Slavic, and Indo-Aryan languages. Some Australian and Papuan languages distinguish animate from inanimate nouns (Dixon, 1972; Olsson, 2021), and certain Dravidian languages contrast common and neuter (Steever, 1998).

Typologically, grammatical gender appears in approximately half of the world's languages (Corbett, 2013; Gygax et al., 2008; McCarthy, 2020). However, as will be discussed in detail in Chapter 5, this general estimate requires caution: within the Indo-European family alone, my analysis demonstrates—for the first time to my knowledge—that for more than half of the documented languages, no conclusive linguistic evidence is available to determine the presence or absence of grammatical gender. This highlights that even well-studied language families may

contain significant gaps in our typological knowledge, which complicates attempts to make global generalisations.

Furthermore, the presence or absence of grammatical gender does not correlate neatly with language family or geographic area. Even within the Indo-European family, genderless languages such as Farsi coexist with richly gendered languages such as German and Russian (Corbett, 1991; Nicholas, 1992).

This cross-linguistic variation underscores why comparative work is indispensable: cognitive effects of grammatical gender cannot be assumed to be universal but must be tested across languages with different typological profiles. As I demonstrate in Chapter 5, even within Indo-European—arguably the best-documented family—typological evidence about the presence or absence of gender remains incomplete. This gap complicates attempts to make global generalisations and highlights why relativity research must be grounded in more systematic typological foundations.

3.2.4 Agreement and Functional Motivations

Grammatical gender is generally treated as a lexical property of nouns that systematically affects agreement on other constituents. Agreement may appear on determiners, pronouns, adjectives, numerals, verbs, and even adpositions. In addition to marking semantic distinctions, grammatical gender serves important discourse functions: it aids in disambiguation, clarifies antecedent relationships, and helps distinguish homophones. Gender marking can also facilitate reference tracking in narratives and reduce repetition (Corbett, 1991; Kramer, 2015).

These functional benefits likely contribute to the persistence of gender systems despite their partial arbitrariness. From a cognitive perspective, agreement is especially significant because it ensures that gender categories are activated automatically during both comprehension and production. This property has made gender agreement a central focus in psycholinguistic research, where it provides insight into the online processing of grammatical features. For example, ERP studies in German and Spanish reveal rapid neural responses to gender-agreement violations (Barber & Carreiras, 2005; Molinaro et al., 2011), and eye-tracking experiments show that speakers use gender-marked articles to anticipate upcoming referents (Lew-Williams & Fernald, 2007).

It is important to distinguish this line of research from studies on linguistic relativity: while psycholinguistic studies demonstrate that gender features are accessed and monitored automatically during language processing, they do not in themselves address whether such grammatical structures influence non-linguistic conceptualisation. Relativity research takes this question further, asking whether gender systems bias the way speakers perceive, categorise, or reason about entities in the world.

3.2.5 Interim Summary

In sum, grammatical gender classifies nouns into agreement-triggering categories based on a mixture of semantic, morphological, phonological, and arbitrary principles. The number and organization of gender classes vary widely across languages, and this typological diversity is essential for interpreting any claims about the cognitive effects of gender. Precisely because grammatical gender combines semantically motivated cores with substantial arbitrariness, it offers a uniquely informative testing ground for exploring the potential relationship between language structure and cognition.

The next section builds on this typological foundation by examining how grammatical gender has been treated in different linguistic frameworks, providing a broader disciplinary context for the empirical and relativity-focused analyses developed in Chapters 5–9.

3.3 Approaches to Grammatical Gender Across Traditions

Research on grammatical gender spans multiple linguistic traditions, each oriented toward distinct but complementary questions. Psycholinguistic studies investigate how gender features are represented and processed during comprehension and production. Sociolinguistic research examines how gender systems interact with identity, ideology, and social practice. Linguistic relativity research explores whether the structural properties of gender systems influence non-linguistic cognition. Taken together, these traditions provide a comprehensive picture of grammatical gender as a phenomenon at the intersection of grammar, processing, cognition, and society.

The aim of this dissertation is situated primarily in the third strand—linguistic relativity—but it is important to situate that inquiry against the backdrop of psycholinguistic and sociolinguistic perspectives. The following sections therefore review these traditions in turn,

highlighting what they contribute to our understanding of grammatical gender and how they frame the relativity questions developed in Chapters 5–9.

3.3.1 Psycholinguistic Perspectives

Psycholinguistic research has consistently shown that grammatical gender is automatically activated during language processing. In comprehension, studies using eye-tracking and event-related potentials (ERPs) demonstrate that listeners and readers exploit gender cues predictively, anticipating upcoming nouns on the basis of agreement markers (Lew-Williams & Fernald, 2007; Wicha et al., 2004). ERP studies further reveal robust sensitivity to mismatches between expected and actual gender features, suggesting rapid integration of gender information during sentence processing (Barber & Carreiras, 2005; Molinaro et al., 2011).

In production, the gender-congruency effect has been widely documented: speakers retrieve and produce words more quickly when distractors share the gender of the target noun, indicating that gender features are activated at the lemma level even when they are not strictly required for task completion (Schriefers et al., 2002). A recent meta-analysis confirmed the robustness of this effect, while showing that its size is modulated by the locus of agreement (determiners vs. adjectives), the transparency of gender marking, and the structural context (Sá-Leite et al., 2022).

Acquisition research shows that children are sensitive to gender cues early in development. For instance, Lew-Williams and Fernald (2007) found that Spanish-speaking children exploit gender-marked articles to anticipate nouns before mastering irregular cases. Bilingual studies further demonstrate that grammatical gender can be co-activated across languages, with facilitation or interference depending on typological distance and proficiency (Grüter et al., 2012; Hopp, 2013).

Taken together, psycholinguistic studies reveal two key insights: (i) grammatical gender is activated automatically in real-time comprehension and production, and (ii) the behavioral footprint of this activation is shaped by factors such as transparency, proficiency, and developmental profile. While these studies illuminate the processing dimension of gender, they do not directly address whether grammatical gender influences conceptualization beyond language. It is precisely this broader question that motivates the relativity perspective developed in this dissertation.

3.3.2 Sociolinguistic Perspectives

Whereas psycholinguistic research emphasizes the automatic activation of grammatical gender in processing, sociolinguistic studies foreground its role as a site of social meaning, ideology, and identity construction (Cameron, 1995; Eckert & McConnell-Ginet, 2013). From this perspective, grammatical gender is not only a morphosyntactic category but also a resource for negotiating power relations and enacting social belonging.

A central focus has been the so-called *generic masculine* in languages such as German, French, and Spanish. In German, for example, plural masculine forms such as *die Studenten* ('the students') have traditionally been used to refer to mixed-gender groups, theoretically including both male and female individuals. Similarly, occupational titles such as *der Arzt* ('the doctor') have historically been used in generic contexts intended to refer to people of any gender (Stahlberg et al., 2007). Although such forms are traditionally described as encompassing both male and female referents, experimental evidence demonstrates that masculine generics are not cognitively neutral: they systematically bias mental representations toward male exemplars (Gabriel et al., 2008; Stahlberg et al., 2007). These findings converge with psycholinguistic research on default interpretations, while sociolinguistic studies extend the analysis by showing how such linguistic defaults may reproduce androcentric norms and maintain symbolic asymmetries in language use.

In response to these biases, many communities have developed inclusive language reforms. German innovations such as the *Binnen-I* (e.g., *StudentInnen*) or the gender asterisk (*Student**innen, the introduction of the gender-neutral pronoun *hen* in Swedish, and the increasing acceptance of singular *they* in English illustrate how speakers actively contest inherited gender systems (Hellinger & Bußmann, 2015). Empirical studies confirm that such inclusive forms can reduce male bias in role-noun interpretation tasks (Gygax et al., 2009), thereby demonstrating how linguistic practices can intervene in cognitive defaults.

Variation further highlights the social embedding of gender. Heritage speakers and bilinguals frequently simplify gender systems (Montrul, 2015). While these reductions are sometimes stigmatized, they can also function as positive markers of identity and group affiliation. Similarly, the use of inclusive innovations indexes alignment with particular political or ideological stances. In this way, sociolinguistic research demonstrates that grammatical gender is deeply enmeshed with issues of ideology, identity, and belonging,

underscoring that its significance extends beyond morphosyntax and into the realm of social practice. These findings are also highly relevant for linguistic relativity research. They suggest that experimental tasks designed to probe conceptual effects of grammatical gender are not ideologically neutral but are embedded in broader social discourses about gender and inclusion. Participants' responses may therefore reflect not only structural properties of the grammatical system but also metalinguistic awareness, sociopolitical positioning, or sensitivity to ongoing language reforms. From this perspective, sociocultural context may function as a moderating variable in the manifestation of gender-related cognitive effects.

This possibility has important methodological implications. Studies that rely on explicit judgment tasks or reflective descriptions may be particularly susceptible to ideological influences or social desirability effects. For this reason, the empirical chapters of this dissertation incorporate both controlled behavioral paradigms and implicit real-time measures, allowing for a clearer separation between structural grammatical activation and conscious sociopolitical reflection.

Taken together, sociolinguistic studies show that grammatical gender is not only a structural category but also a resource for negotiating ideology, identity, and social meaning. While this strand of research emphasizes the social consequences of gendered systems, the next section turns to linguistic relativity research, where the central concern is whether grammatical gender exerts influence beyond language use itself—shaping how speakers conceptualize and represent the world.

3.3.3 Linguistic Relativity Perspective

Grammatical gender has long served as a central testing ground for theories of linguistic relativity. Because it is obligatory in some languages but entirely absent in others, it provides a particularly clear case for examining whether structural properties of language influence habitual thought. Early research suggested that speakers of gendered languages may conceptualize objects in ways that reflect their grammatical classification (e.g. Boroditsky et al., 2003). Yet subsequent work produced mixed results, with effects often contingent on task demands, language pairings, and methodological choices (Bender et al., 2011; Flaherty, 2001; Sera et al., 2002). This variability has fueled debates over whether grammatical gender genuinely shapes conceptualization or whether previously reported effects are artifacts of experimental design.

A central difficulty in this line of research is that grammatical gender is not a straightforward mapping of biological sex. Many inanimate nouns carry masculine, feminine, or neuter grammatical genders in languages such as German without reference to natural categories, and even some animate nouns exhibit grammatical genders that contradict their biological sex (e.g., *das Mädchen* "girl" in German is neuter). For example, asking German speakers to explain why *das Mädchen* ("girl") is grammatically neuter would elicit conscious reflection on the language system, rather than revealing whether gender categories shape their everyday conceptualization of animacy or sex. Accordingly, scholars have increasingly argued that more sensitive methodologies and broader typological sampling are needed to assess relativity claims (Athanasopoulos & Casaponsa, 2020; Lucy, 2016).

In this dissertation, I contribute to this debate in two main ways. First, it offers a systematic review of the methodological landscape of grammatical gender studies in the linguistic relativity tradition (Chapter 6, Article 2). By comparing paradigms ranging from explicit voice-assignment tasks to implicit measures such as electroencephalography (EEG), this review not only identifies the strengths and limitations of existing approaches but also reveals systematic gaps in the field's methodological repertoire. In particular, it highlights three issues: (i) an over-reliance on explicit tasks, (ii) the underrepresentation of gender systems beyond the masculine–feminine binary, and (iii) the scarcity of evidence from genderless languages and bilingual contexts. By bringing these shortcomings into focus, the review provides a framework for evaluating how experimental design conditions the detection of relativity effects and sets the stage for the empirical studies presented in later chapters.

The subsequent empirical studies (Chapters 5, 7–9) were designed with these gaps in mind. Chapter 5 addresses the problem of typological coverage by situating Indo-European languages within a broader genealogical perspective and showing that even within this well-studied family, evidence for the presence or absence of gender is far less complete than typological surveys suggest. Chapter 7 shifts focus to speakers of Czech, another fully fledged gendered language, thereby providing the critical counterpoint that is largely missing from the literature. Chapter 8 responds to the field's binary bias by systematically examining the German neuter, a category often overlooked in relativity research but whose conceptual consequences may diverge from those of masculine and feminine. In addition, it incorporates Farsi, another underrepresented language in the linguistic relativity framework, thereby enabling a direct comparison with German. Finally, Chapter 9 extends this analysis to Farsi–German bilinguals, examining how exposure to a new gender system interacts with an otherwise genderless

conceptual repertoire. Crucially, it employs an eye-tracking paradigm, a method not previously applied in grammatical gender research within the linguistic relativity framework, thereby providing both methodological innovation and novel empirical evidence. Together, these studies not only diversify the typological and methodological bases of relativity research but also test, across multiple populations and paradigms, whether and how grammatical gender shapes conceptualization.

Taken as a whole, the dissertation thus moves beyond documenting isolated relativity effects to reframing how grammatical gender should be studied within this tradition. By integrating methodological critique with typological expansion and bilingual evidence, it argues that the influence of grammatical gender on cognition is neither universal nor illusory but is contingent on linguistic structure, experimental design, and speaker experience.

3.4 Conclusion

This chapter has examined grammatical gender from multiple perspectives, situating it both as a structural feature of language and as a focal point of theoretical debate. After outlining patterns of cross-linguistic variation and describing the semantic, morphological, and phonological dimensions of gender assignment, the chapter reviewed how different linguistic traditions—psycholinguistic, sociolinguistic, and linguistic relativity approaches—have addressed the role of gender systems in language and cognition.

Of particular relevance to the present dissertation is the way grammatical gender serves as a natural testing ground for the framework of linguistic relativity. Because gender systems encode categories that are sometimes semantically motivated yet often arbitrary, while remaining deeply embedded in grammatical structure, they provide an opportunity to examine whether language-specific patterns may influence cognitive representations.

Taken together, the discussion in this chapter highlights two central insights. First, grammatical gender cannot be understood solely as a system of formal agreement; it also interacts with semantic interpretation and social meaning in language use. Second, existing research offers both empirical indications and theoretical motivation for examining whether speakers' representations of grammatical gender may shape cognitive processes, particularly in multilingual contexts where gender systems diverge across languages.

Building on these insights, the next chapter introduces the methodological framework that guides the empirical studies in this dissertation. By combining controlled experimental paradigms with cross-linguistic comparisons, the studies aim to assess whether grammatical gender exerts measurable effects on cognition and to clarify the structural and contextual conditions under which such effects may emerge.

Chapter 4 Methodological Framework

This chapter outlines the methodological framework underpinning the empirical and theoretical investigations presented in this dissertation. The novelty of this work lies in its integrative, multi-method design, which combines complementary approaches to examine the relationship between grammatical gender and cognition across different levels of analysis. In addition, to the best of our knowledge, this dissertation presents the first eye-tracking study to investigate grammatical gender within the framework of linguistic relativity**.** Rather than relying on a single methodological tradition, the dissertation brings together typological analysis, systematic literature review, theoretical–critical inquiry, behavioral experimentation, eye-tracking techniques, and questionnaire-based measures. This combined approach enables the research questions to be addressed from both cross-linguistic and experimental perspectives, thereby providing a more comprehensive and nuanced understanding of the cognitive implications of grammatical gender.

Each of these approaches addresses a different level of analysis and contributes distinct types of evidence, with questionnaire-based measures ensuring precise participant characterization and control over individual differences. Typological and review-based methods provide a descriptive and conceptual foundation by mapping cross-linguistic variation and identifying methodological patterns in existing research. Behavioral and eye-tracking experiments, in contrast, allow specific hypotheses about the influence of grammatical gender on conceptual and online processing to be tested under controlled conditions.

Table 4.1 provides an overview of the methodological approaches employed across the five studies included in this cumulative dissertation. Each study contributes a distinct methodological perspective while collectively forming an integrated research program.

Table 4.1 *Overview of methodological approaches used across the five studies in this cumulative dissertation*

Study	Research Type	Method	Data Source	Purpose
Study 1	Typological survey	Cross-linguistic database analysis	Glottolog (413 Indo-European languages)	Identify typological diversity and sampling bias
Study 2	Systematic literature review	PRISMA-guided SLR	Peer-reviewed studies (2013–2023)	Identify methodological patterns in gender–cognition research
Study 3	Theoretical-critical analysis	Conceptual analysis	Linguistic relativity literature	Evaluate theoretical assumptions about gender systems
Study 4	Behavioral experiment	Property-judgment paradigm	German vs. Farsi monolingual participants	Test implicit gender-concept associations
Study 5	Eye-tracking experiment	Visual World Paradigm	German monolinguals and Farsi-German bilinguals	Examine real-time processing of grammatical gender
Study 4 & 5	Participant profiling	Language background questionnaires	Self-report data	Classify participants and control for linguistic and socio-demographic variables

The following sections introduce these methodological approaches from a scientific perspective, outlining their role in linguistic and psycholinguistic research before explaining their application within the individual studies of this dissertation.

4.1 Typological and Review-Based Methods

The first two studies of this dissertation employ typological and systematic review methods in order to establish the broader empirical and methodological context for the experimental investigations that follow. In linguistic relativity research, claims about the cognitive effects of linguistic structures depend not only on how these structures are distributed across languages but also on how they have been examined in previous studies. For this reason, it is necessary to consider both the typological diversity of grammatical gender systems and the methodological

properties and potential biases of existing research before testing specific hypotheses experimentally.

Typological research plays an important role in linguistic inquiry because it documents how linguistic structures vary across languages and language families (Comrie, 1989; Croft, 2003). In the case of grammatical gender, typological studies have demonstrated substantial cross-linguistic variation in the number of gender categories, the principles governing gender assignment, and the domains in which agreement is expressed (Corbett, 1991). Such variation is particularly relevant for linguistic relativity research, which often relies on comparisons between languages that differ in the presence or structure of gender systems. However, earlier research has frequently focused on a relatively limited set of languages, primarily within Western European contexts.

Study 1 adopts a typological survey design based on the Glottolog database (Hammarström et al., 2021). The study examines documented Indo-European languages in order to assess the distribution of grammatical gender systems within this language family. By systematically classifying languages according to the presence and number of gender categories, the analysis provides a broader typological perspective on gender systems while also highlighting gaps in available linguistic documentation.

Complementing this typological perspective, Study 2 employs a systematic literature review (SLR) to examine recent experimental work on grammatical gender and cognition. Systematic reviews have become an important methodological tool for synthesizing research findings and identifying patterns across studies (Borenstein et al., 2009; Norris & Ortega, 2006). In this dissertation, the review follows the PRISMA framework (PRISMA-P Group et al., 2015), which provides transparent procedures for identifying, screening, and reporting relevant studies. Using clearly defined search strategies and inclusion criteria, the review surveys experimental studies published between 2013 and 2023 that investigate the relationship between grammatical gender and cognitive processing, with particular attention to language coverage and methodological variation.

Taken together, these two approaches provide the descriptive and methodological foundation for the experimental work presented later in the dissertation. The typological analysis situates grammatical gender within a broader cross-linguistic perspective, while the

systematic review identifies key empirical trends and gaps that motivate the design of the experimental studies.

4.2 Theoretical and Critical Review Approaches

In addition to typological surveys and systematic literature reviews, theoretical and critical review approaches constitute an important component of linguistic and psycholinguistic research. While systematic reviews aim to provide a structured synthesis of empirical studies based on predefined search strategies, theoretical reviews focus on interpreting and integrating research findings in order to clarify conceptual relationships and identify broader theoretical implications (Grant & Booth, 2009; Torraco, 2016). Such approaches are particularly valuable in research domains where empirical findings are heterogeneous or where underlying theoretical assumptions require closer examination.

Within the field of linguistic relativity, theoretical analyses have played a central role in evaluating how linguistic structures may interact with cognitive processes. Research on grammatical gender has produced a diverse body of findings, with some studies reporting gender-related conceptual biases and others observing weak or inconsistent effects (Bassetti & Nicoladis, 2016; Samuel et al., 2019). These differences have often been attributed to variation in experimental paradigms, language pairings, and participant populations. Theoretical reviews provide a framework for assessing how such methodological and linguistic factors shape the interpretation of empirical results.

Theoretical approaches are particularly relevant when examining bilingual and multilingual populations, where multiple linguistic systems interact within the same cognitive architecture. In such contexts, questions arise regarding whether conceptual representations are language-specific or shared across languages (Athanasopoulos & Casaponsa, 2020). This perspective is essential for interpreting potential cross-linguistic influences and for understanding how grammatical gender effects may vary across language contexts.

Within this framework, Study 3 adopts a theoretical–critical perspective to examine how grammatical gender has been conceptualized within the linguistic relativity debate, with particular attention to multilingual contexts. The analysis focuses on studies published between 2013 and 2023, which are examined from a conceptual perspective to evaluate how grammatical gender effects have been interpreted in recent research. Rather than providing a descriptive synthesis, the study critically evaluates underlying assumptions, identifies

inconsistencies in the interpretation of findings, and highlights conceptual gaps that inform the design of the subsequent experimental studies. By situating these findings within a broader conceptual and typological context, the study contributes to the theoretical foundation for the experimental investigations presented in the subsequent studies of this dissertation.

4.3 Behavioral Experimental Methods

Behavioral experimental paradigms are widely used in psycholinguistic research to investigate how linguistic structures may influence cognitive representations and conceptual processing (Mertins, 2016; Traxler, 2012). Unlike typological and review-based approaches, which examine structural patterns across languages or synthesize findings from previous research, behavioral experiments allow researchers to test specific hypotheses under controlled conditions. By systematically manipulating linguistic stimuli and measuring participants' responses, such paradigms provide empirical evidence about the relationship between linguistic structure and conceptual processing.

Within research on linguistic relativity, behavioral methods have frequently been employed to examine whether grammatical structures influence speakers' conceptual representations of objects and their associated properties (e.g., Bassetti & Nicoladis, 2016; Mertins, 2016; Samuel et al., 2019). One widely used approach involves property attribution or judgment tasks, in which participants generate or evaluate descriptive properties of objects whose grammatical gender differs across languages, while a second group of participants subsequently rates these attributes along a gendered scale. These paradigms are designed to capture conceptual associations that may reflect the influence of grammatical gender without explicitly directing participants' attention to grammatical categories. Previous studies have employed similar tasks to investigate whether speakers attribute gender-related characteristics to objects (Boroditsky et al., 2003; Imai et al., 2014; Saalbach et al., 2012; Samuel et al., 2019; Sera et al., 2002).

Property-judgment paradigms are particularly suitable for investigating implicit conceptual associations because they allow participants to respond in a relatively natural task environment while minimizing explicit reflection on grammatical gender. By eliciting descriptive properties of visually presented objects, such tasks enable the examination of whether speakers systematically associate gendered, stereotypical attributes (e.g., masculine vs. feminine connotations) with objects whose grammatical gender differs across languages. As a result,

these paradigms provide an indirect measure of how linguistic structure may shape conceptualization (Boroditsky et al., 2003; Sedlmeier et al., 2015; Sera et al., 2002).

When applying property-judgment paradigms to grammatical gender, several methodological considerations are crucial. First, it is essential to represent the full gender system of the language under investigation. In the case of German, this entails the inclusion of masculine, feminine, and neuter nouns, as reducing the system to a binary distinction may obscure important structural effects. Second, the measurement of gender associations requires a sufficiently sensitive rating scale. For this reason, a 7-point Likert scale capturing a full continuum—from clearly feminine to clearly masculine, including a gender-neutral midpoint—provides a more fine-grained assessment of conceptual associations than binary categorizations, and more accurately reflects the gradient nature of gendered representations in real-world cognition. Third, careful construction of the stimulus set is essential. This includes the systematic selection of critical items, appropriate control items, and filler stimuli in order to minimize strategic responding and to ensure that observed effects can be attributed to the manipulation of grammatical gender rather than to extraneous factors.

In the present dissertation, Study 4 employs a property-judgment paradigm to examine whether grammatical gender influences conceptual associations in speakers of languages with different gender systems. Participants generated descriptive adjectives for a set of objects, which were subsequently evaluated in terms of their gender-related semantic associations using a 7-point Likert scale ranging from very feminine to very masculine, including a neutral midpoint. By comparing responses from German monolingual speakers, whose language encodes a three-gender system, with responses from Farsi monolingual speakers, whose language lacks grammatical gender, the study investigates whether grammatical gender is reflected in implicit semantic judgments. Through this behavioral paradigm, the study provides an experimental test of the relationship between grammatical gender and conceptualization under controlled conditions.

The resulting data were analyzed using linear mixed-effects models to account for both participant- and item-level variability (Fahrmeir et al., 2013). The models included random intercepts for participants and items, as well as by-subject and by-item random slopes where supported by the data, following Barr et al. (2013). This approach allows for robust estimation in repeated-measures designs and captures both within-subject and within-item variability. All analyses were conducted in R (R Core Team & others, 2023).

4.4 Eye-Tracking and the Visual World Paradigm

Eye-tracking has become a central method in psycholinguistic research for examining how language processing unfolds over time and how linguistic information guides visual attention. By recording participants' eye movements while they process spoken or written language, the method provides a continuous measure of moment-to-moment cognitive processing during language comprehension (Holmqvist & Andersson, 2017; Huettig et al., 2011; Rayner, 1998). Because eye movements are closely linked to attentional allocation, the locations and durations of fixations can reveal which elements of a visual scene participants attend to while interpreting linguistic input.

Modern eye-tracking systems typically use infrared illumination to detect both the position of the pupil and the reflection of light from the cornea. The relative position between these signals allows the system to calculate gaze direction with high spatial and temporal precision (Holmqvist & Andersson, 2017). The resulting data consist of sequences of eye movements that are commonly classified into fixations and saccades. Fixations represent relatively stable periods during which visual information is processed, whereas saccades are rapid movements between fixation points during which visual perception is largely suppressed (Rayner, 1998).

As an online method, eye-tracking provides mediated access to underlying cognitive processes as they unfold during language comprehension. These processes are typically rapid, automatized, and largely inaccessible to conscious introspection, as participants are required to respond to experimental stimuli with minimal temporal delay (Mertins, 2016). Consequently, eye-tracking enables the investigation of implicit aspects of language processing that are difficult to capture using offline behavioral measures[4].

The analysis of eye-tracking data typically involves several stages, beginning with the preprocessing of raw gaze data. During this step, artifacts such as blinks or temporary loss of pupil detection are identified and corrected where possible. The data are then segmented into relevant eye-movement events to enable subsequent analysis of gaze behavior (Holmqvist & Andersson, 2017). In psycholinguistic research, several metrics are commonly analyzed,

[4] Offline methods are characterized by their indirect access to cognitive processes, as they capture the outcomes of conscious and reflective decision-making rather than real-time processing. Responses in such tasks are typically produced with a temporal delay following stimulus presentation, allowing for strategic and controlled processing. Typical examples include questionnaires, rating tasks, and object-naming paradigms, which are often employed in studies involving special populations, such as individuals with aphasia (Mertins, 2016).

including fixation duration (often referred to as dwell time), number of fixations, latency to first fixation, and fixation proportion. These measures are typically calculated within predefined Areas of Interest (AOIs), which represent specific regions of the visual display that are relevant to the experimental hypotheses. The definition of AOIs is a critical step in the analysis, as it directly determines how visual attention is quantified and interpreted. Careful and theoretically motivated delineation of AOIs is therefore essential to ensure that observed gaze patterns can be reliably linked to the cognitive processes under investigation.

A widely used experimental paradigm in psycholinguistic eye-tracking research is the Visual World Paradigm (VWP). In this paradigm, participants view a visual display containing multiple objects while simultaneously listening to spoken language (Huettig et al., 2011; Tanenhaus et al., 1995). Eye movements are recorded continuously as the auditory input unfolds, allowing researchers to track how linguistic information guides visual attention over time. Because listeners tend to fixate on objects that are semantically, phonologically, or syntactically related to the spoken input, gaze patterns provide insight into the incremental processes underlying language comprehension.

The Visual World Paradigm has been widely applied to investigate a range of psycholinguistic phenomena, including lexical access, predictive processing, and bilingual language processing (Huettig et al., 2011; Ito & Knoeferle, 2022; Knoeferle & Guerra, 2016). Importantly, eye-movement measures such as fixation counts, dwell time, and latency to first fixation allow researchers to detect subtle processing biases that may emerge before explicit behavioral responses. Recent methodological work has further highlighted the value of eye-tracking for examining the interaction between linguistic structure and cognitive processing in real time (Ito & Knoeferle, 2023).

In the present dissertation, Study 5 employs an eye-tracking experiment based on the Visual World Paradigm to investigate whether grammatical gender influences online language processing in bilingual speakers. The Visual World Paradigm was selected because it provides a highly sensitive method for capturing real-time and implicit processing, making it particularly well suited for detecting subtle effects of grammatical gender on visual attention as language unfolds. Participants viewed visual displays containing multiple objects while listening to an auditory stimulus. By analyzing fixation patterns and dwell times across predefined Areas of Interest, the study examines whether grammatical gender information affects anticipatory gaze behavior and lexical activation during language comprehension. This online methodology

complements the behavioral paradigm used in Study 4 by enabling the investigation of the temporal dynamics of grammatical gender processing.

The resulting eye-tracking data were analyzed using linear mixed-effects models following the same analytical framework as in the behavioral study, allowing the examination of gaze behavior while accounting for both participant- and item-level variability (Fahrmeir et al., 2013). Random intercepts for participants and items were included, as well as by-subject and by-item random slopes were supported by the data (Barr et al., 2013). This approach is particularly suited to eye-tracking data, as it accounts for both participant- and item-level variability in repeated measures, thereby providing a more accurate and reliable analysis of gaze behavior. All analyses were conducted in R (R Core Team & others, 2023).

4.5 Language Background Questionnaires

Language background questionnaires are a widely used methodological tool in psycholinguistic and bilingualism research for systematically assessing participants' linguistic experience, language proficiency, and patterns of language use (Anderson et al., 2018; Li et al., 2014; Marian et al., 2007). In addition to linguistic variables, such instruments typically collect demographic and socio-economic information, enabling a more comprehensive characterization of participants and facilitating control over potential confounding factors.

In research on linguistic relativity and cross-linguistic influence, the use of language background questionnaires is particularly crucial. Differences in language experience—such as age of acquisition, frequency of use, and degree of exposure—may directly influence cognitive processing and conceptual representations (Bialystok et al., 2012; Marian et al., 2007). At the same time, socio-economic variables, including educational background, may also affect linguistic behavior and cognitive performance (Bialystok et al., 2012). A careful assessment of both linguistic and demographic factors is therefore essential to ensure that observed effects can be attributed to structural properties of language rather than to uncontrolled individual differences.

In the present dissertation, language background questionnaires were administered in both experimental studies (Studies 4 and 5) to ensure precise participant classification and to control for potential confounding variables. Two complementary questionnaires were employed, each tailored to the specific requirements of the respective study.

For the behavioral experiment (Study 4), an online questionnaire developed at the *psycholinguistics laboratories* at TU Dortmund was used to establish strict monolingual profiles while also collecting relevant demographic information. Participants provided detailed information about their linguistic history, including languages acquired in early childhood, patterns of language use within the family environment, and exposure to other languages through residence abroad or regular interaction. The questionnaire further included items on age, gender, country of birth, and educational background, allowing for the consideration of socio-economic factors. Only participants who reported a predominantly monolingual upbringing were included. Additional criteria ensured that participants had not experienced sustained exposure to another language during early development and had not lived outside their home country (Germany or Iran) for extended periods. While some participants reported classroom-based foreign language instruction, such exposure was not considered to constitute functional bilingualism, as it primarily reflects explicit, declarative knowledge rather than implicit, automatized processing (Bialystok, 2001; Grosjean, 2008).

For the eye-tracking experiment (Study 5), a second questionnaire was used to obtain a more fine-grained and quantitative assessment of bilingual experience. Although several instruments have been developed to assess bilingualism—such as the Language Experience and Proficiency Questionnaire (Marian et al., 2007), the Language History Questionnaire (Li et al., 2020), and the Language and Social Background Questionnaire (Anderson et al., 2018) —there is no unified consensus regarding item selection, scaling procedures, or the integration of variables into a comprehensive measure of bilingualism. Existing instruments often assess different dimensions of bilingual experience separately rather than providing a single composite index that captures the graded nature of bilingualism.

For this reason, the present study employed a questionnaire-based index developed at the *psycholinguistics laboratories* at TU Dortmund, which allows for the calculation of a continuous measure of bilingualism and thus enables a more fine-grained characterization of participants' linguistic profiles. This index, adapted for an adult population, incorporates multiple dimensions of language experience, including age of acquisition, self-rated proficiency, frequency of use across contexts, and length of residence in German-speaking environments. Based on these measures, a composite bilingualism score was calculated, allowing for a graded characterization of bilingual experience rather than a simple categorical classification. This approach makes it possible to capture individual differences in bilingualism

more precisely and to relate them to patterns observed in the eye-tracking data (Odermann & Mertins, 2026).

Despite their advantages, questionnaire-based methods also have certain limitations. As self-report instruments, they may be subject to biases such as inaccuracies in self-assessment or socially desirable responding (Hlawatsch & Krickl, 2019). Moreover, as offline measures, questionnaires do not provide direct access to real-time cognitive processes (Mertins, 2016). Nevertheless, their systematic and standardized nature makes them an essential tool for controlling participant-related variables in experimental research.

4.6 Integration of Methodological Approaches

The methodological framework of this dissertation is based on the assumption that the relationship between grammatical gender and cognition cannot be adequately addressed through a single research method. Instead, the cumulative design combines complementary approaches that operate at different levels of analysis and provide distinct forms of evidence.

This integrative approach is motivated by the need to capture both the structural properties of grammatical gender systems and their cognitive consequences across different contexts of language use. In particular, combining behavioral and eye-tracking methods allows the investigation of grammatical gender effects at both the conceptual level and in real-time language processing. By incorporating an online eye-tracking paradigm into the study of linguistic relativity, the present work extends previous research, which has largely relied on offline behavioral measures, and enables the detection of fine-grained and time-sensitive processing effects that may not be observable in offline measures.

The typological survey and systematic literature review establish the descriptive and methodological background of the research by identifying cross-linguistic variation and recurrent patterns in earlier studies. The theoretical–critical review extends this perspective by examining how empirical findings have been interpreted within the broader linguistic relativity debate, particularly in multilingual contexts. Together, these approaches provide the conceptual and empirical foundation for the experimental investigations.

The behavioral and eye-tracking studies build on this foundation by testing specific hypotheses under controlled conditions. Whereas the behavioral paradigm captures implicit conceptual associations through property judgments, the eye-tracking experiment allows the

temporal dynamics of grammatical gender processing to be examined in real time. The inclusion of language background questionnaires in both experimental studies further strengthens the design by enabling precise participant classification and by controlling for linguistic and socio-demographic variables that may affect cognitive processing.

Taken together, these methodological components form an integrated research program that combines breadth and depth: breadth through typological and review-based approaches, and depth through controlled experimental and online measures. This multi-method design provides a comprehensive basis for examining the role of grammatical gender in cognition across different levels of analysis. The following chapters (Chapters 5–9) present the five studies of this cumulative dissertation in detail, each applying these methodological approaches to address specific aspects of the research problem.

Maryam Fatemi

Synchronic overview of grammatical gender marking in Indo-European languages

Abstract: In line with the theories of *linguistic relativity* (Whorf 1956), *thinking for speaking* (Slobin 1996), and *seeing for speaking* (Mertins 2018), which all argue that the grammatical feature of the language we speak influences the way we see the world and conceptualize it, grammatical gender makes a relevant category to explore such effect on our thoughts. The morphosyntactic feature of gender is linguistically marked in high number of languages in diverse forms. These diverse forms of representing gender of animate and inanimate entities in languages are based on semantic, morphological, and phonological principles. Different from the languages which express gender according to the aforementioned principles, the instance of a language like Farsi, with no grammatical gender-marking system, has depicted that gender information of the nouns can also be optionally represented by lexical description in postnominal positions. Such diverse forms of representing gender in various languages have made this grammatical feature a valid and relevant category for studying diversity in cognition. Looking at the literature on the influence of grammatical gender on thought, the issues of the WEIRD bias as well as Eurocentric bias have been observed. The results of such biased studies on only European languages can explain neither the linguistic variation nor can be taken as the basis for showing the possible cognitive diversity triggered through such variations. The goal of this article is to explicitly address the issues of the WEIRD bias as well as the Eurocentric bias by initially critically reviewing 43 empirical studies on grammatical gender in linguistic relativity and secondly exploring the current state of 413 Indo-European languages concerning +/- grammatical gender features, including the many languages without any expression of grammatical gender. The results of the critical review on 43 studies proves additional evidence for Eurocentric and WEIRD bias in the linguistic relativity framework, since in these studies only 20 languages were considered, among which fifteen Indo-European languages mainly spoken in Europe. The findings of this study on the state of 413 Indo-European languages concerning +/- grammatical gender feature firstly convey that 72 % of the Indo-European languages are spoken outside Europe. Secondly, it was found that only 198 languages have linguistic proof for the absence or presence of gender-marking systems. Furthermore, the results have revealed that most of these languages, 96 in number, express gender in a twofold system of masculine and feminine. Finally, out of all Indo-European languages, 215 languages have not been linguistically studied regarding their expression of grammatical gender.

Keywords: linguistic relativity, thinking for speaking, seeing for speaking, grammatical gender, Indo-European languages, WEIRD bias, Eurocentric bias

1. Introduction

Eurocentrism, which is the tendency to view the European ideas and social standards as the basis for evaluative judgments of other practices, has been widely criticized (Joseph et al. 1990). Eurocentric bias can be observed not only in social relations and politics, but also in research, specifically including linguistic research.

Following Eurocentrism, the term "WEIRD" societies was coined (Henrich, Heine & Norenzayan 2010; Henrich 2020), which refers to the *Western Educated Industrialized Rich Democratic* societies. Most scientific studies in various fields are concentrated on WEIRD people. In linguistics, psycholinguistics, and corpus linguistics, the majority of studies have looked merely into the languages spoken in Europe, which make up approximately 1 % of the 7,099 existing languages spoken in the world. These facts lead to the question of how linguistic variations can be explained, when the existing research body has only taken 1 % of all languages into consideration. This article addresses explicitly the issue of the WEIRD bias as well as the Eurocentric bias by exploring the category of grammatical gender in 413 Indo-European languages, including the many languages without any expression of grammatical gender.

In a high number of languages, the morphosyntactic feature of gender is linguistically marked. Alvanoudi (2000, 2015) defines this feature as a grammaticized agreement system, which correlates with specific semantic characteristics of the noun such as sex, animacy or humanness. These features are realized outside the noun itself. That is, the agreement is applied through morphosyntactic marking of the words that the gendered noun is morphosyntactically related to (Boroditsky et al. 2003a; Corbett 1991). The term *gender* has been interchangeably used with the term *noun class* in literature. Alvanoudi (2000) has proposed that the term *noun class* should be used as the cover term for both *gender* and *noun class* and the term *gender* should be utilized for the systems including two or three genders which always have at least feminine and masculine. Alvanoudi (2000) argues in his typological approach that nouns are assigned to specific classes not only based on *semantic* but also on *morphological* and *phonological* principles.

Semantic principles refer to the systems in which the meaning of a noun determines its gender. For instance, in Dravidian languages such as Kolami, Parji and Ollari nouns denoting male humans belong to one class and nouns denoting female humans and any other non-male entities belong to a different class (Corbett 1991). An opposite pattern can be observed in Diyari, an Australian Indigenous language. Diyari differentiates between nouns denoting female human/animates

and the rest, which comprise male human/animates and all inanimates (Dixon 1972). Dyirbal, which is another Australian Indigenous language spoken in North Queensland, categorizes the nouns yet differently. There are four different classes: (i) *bayi* referring to males and non-human animates, (ii) *balan* referring to females, fire, water, battle, and most birds, (iii) *balam* referring to non-protein food, and (iv) *bala* for the rest (Dixon 1972).

A completely different take on the classification of nouns can be seen in Manambu, a language belonging to the Ndu family spoken in Papua New Guinea. The nouns are categorized into feminine and masculine based on their size and shape. For example, adult humans are categorized based on their sex, however, bigger or older children are classified as masculine, and smaller or younger children are referred to as feminine. Furthermore, longer, and larger inanimate objects are categorized as masculine, while smaller and/or rounder objects are categorized as feminine (Aikhenvald 2012).

In the Indo-European languages such as Danish, Icelandic, Spanish, Polish, German, Italian, Serbian/Croatian/Bosnian, Portuguese, Swedish, and Greek, the gender assignment is mainly motivated semantically when it refers to a human and some animals. That is, nouns in these languages denoting female humans and some animals are usually grammatically feminine, and nouns denoting male humans and some animals are usually grammatically masculine (Hellinger & Bussmann 2001, 2002, 2003). Therefore, there is a grammatical congruency between the gender of the nouns and the articles and/or the adjectives which precede the nouns (the same agreement rules exist in animate and inanimate). For instance, the German nouns in Example 1 show the sex of the person both lexically and grammatically:

Example 1:

Masculine-Definite
(a) Deutsch

Der	*groß-e*	*Mann*
The (M.D.A.)	tall-NOM.SG.	man
'The	tall	man'

Feminine-Definite
Deutsch

Die	*groß-e*	*Frau*
The (F.D.A.)	tall-NOM.SG.	woman
'The	tall	woman'

Masculine-Indefinite
(b) Deutsch

Ein	*groß-er*	*Mann*
A (M.ND.A.)	tall-NOM.SG.	man
'A	tall	man'

Feminine-Indefinite
Deutsch

Eine	*groß-e*	*Frau*
A (F.ND.A.)	tall-NOM.SG.	woman
'A	tall	woman'

Masculine-animate
(c) Deutsch

Der	*schnell-e*	*Tiger*
The (M.D.A.)	fast-NOM.SG.	(male) tiger
'The	fast	tiger'

Feminine-animate
Deutsch

Die	*schnell-e*	*Tigerin*
A (F.D.A.)	fast-NOM.SG.	(female) tiger
'The	fast	tiger'

At the same time, in the Indo-European languages, gender assignment in inanimate nouns such as in food (e.g., egg, bread) or furniture (e.g., chair, table) is generally considered to be semantically arbitrary. The peculiarity of gender assignment does not end with assigning a feminine, masculine and/or neutral gender to inanimate objects. But rather, in some of the languages of this family, speakers are required to not only mark objects as gendered through modifying articles, pronouns, and adjectives according to the gender of the noun, but also to add affixes to verbs to agree with the gender of the noun in specific tenses. For instance, in Czech, which belongs to the Slavonic group of languages in the Indo-European family, the verbal suffix *-y* is used in the past tense only when the referred noun is plural feminine animate, children, and/or all inanimate. On the other hand, the verbal suffix *-i* marks the past tense for the plural masculine animate. Moreover, this distinctive assignment of grammatical gender on verbs is observable only in written language and as mentioned in the past tense.

Knowing this grammatical information in Czech, which is a pro-drop language, is important in order to be able to infer the nature of the subject in the sentence, especially the gender of the subject.

Example 2:

(a) Czech

Včera	*šl-i*	*do*	*kina.*
Yesterday	went-he	to	cinema
'Yesterday	he went	to	the cinema'

(b) Czech

Včera	*šl-y*	*do*	*kina.*
Yesterday	went-they	to	cinema
'Yesterday	they went	to	the cinema'

In (2a), the verbal suffix *-i* marks that the subject of this sentence must be masculine, whereas in (2b) the verbal suffix *-y* can only refer to the animate feminine referent, and/or to children based on the context. However, from the pragmatic point of view, the inanimate referent is not conveyed here, as the inanimate referent cannot go to the cinema. A lack of this grammatical information in Czech leads to false interpretation of the meaning of sentences.

Furthermore, morphological, and phonological principles play an important role in gender assignment. For instance, in Russian not only does the semantic distinction of sex specify the gender of the noun (e.g., *mat* "mother" is gendered as feminine and *otec* "father" is gendered as masculine), but also in other cases nominative, accusative, genitive, dative, instrumental and prepositional nouns are classified as masculine, feminine and neuter according to morphological principles. Nouns of declension 1, which usually end in a consonant are masculine (e.g., zakon "law"), nouns in declension 2 and 3 which end in -a and- я are feminine (e.g., škola "school"), and nouns of declension 4, which end in -o, -e, or -ë are neuter (e.g., vino "wine") (Corbett 1991). In Katcha, a Kordofanian language spoken in Sudan, nouns beginning with m- are feminine, except for the nouns which have a male referent (Heine 1982).

Different from the languages which express gender semantically, morphologically or phonologically, which has consequences for the syntactic structure, Farsi, another Indo-European language mainly written and spoken in Iran, is a genderless language. In this language, nouns are not classified into gendered classes. Consequently, there is no gender marking related to case

marking. Furthermore, pronouns in third person singular are not gendered either. Example 3 gives the translations of the above Examples 1. It should be mentioned that Farsi is written from right to left.

Example 3:

Masculine
(3a) Farsi

-	بزرگ /bozorg/	مرد /Mærd/
-	tall	man
'The	tall	man'

Feminine
Farsi

	بزرگ /bozorg/	زن /zæn/
	tall	woman
'The	tall	woman'

Masculine/Feminine-animate
(3b) Farsi

	بزرگ /bozorg/	ببر /bæbr/
	big	tiger
'the	big	tiger'

As shown in Example 3, in Farsi, nouns (whether animate and/or inanimate) can be directly used to refer to definiteness. In other words, nouns are definite by default. Indefiniteness is expressed through adding either the suffix ی *(-i)* to the noun, or the article یک /*Yek*/, similar to the English *a*/*an* used before the singular nouns, which doesn't attribute any gendered meaning to the noun. However, referring to the gender of an animal or an inanimate object is possible but optional through descriptively adding the translations of the lexemes (feminine and masculine) only in the postnominal position. It is important to stress that the sequence of adjective and noun in Farsi is "nouns + adjective". Due to the lack of grammatical gender marking, personal pronouns in Farsi only provide information about person and number. Example 4 gives the translation of the English sentence "He came" in Farsi.

Example 4:

Farsi

است. /ast/	مرد/mard/	یک /yek/	او/u/	آمد . /amad/	او /u/
Is	man	M.ND.A.	he/she	came	he/she
'He	came.'				

Conforming to the above explained facts about the grammatically genderless system of nouns in Farsi, including all animate/inanimate nouns (e.g., *cat, dog, door, window, desk*), nouns referring to jobs (e.g., *teacher, engineer*) are genderless as well. However, in accordance with the Farsi syntax, it is possible to add gender information to the noun. As described above, such information can be further added to the noun phrase in the postnominal position. For instance, the genderless word *teacher* /ʔɑmuzeˈgɑ̊ɻ/ can refer both to a feminine teacher or a masculine teacher. If a Farsi speaker would like to give information about the gender of a teacher, they can do so by adding a gendered lexeme in the postnominal position.

Grammatical gender has been studied in different linguistic fields such as psycholinguistics (Levelt, Roelofs & Meyer 1999; Friedmann & Biran 2003), cognitive linguistics (Evans & Green 2006; Lakoff 1987) or linguistic relativity (Lucy 1997). The framework of the present article is the field of linguistic relativity. Linguistic relativity is a theory which can be traced back to the linguistic-philosophical work of Wilhelm von Humboldt in the beginning of 19th century, where he described language as the "formative organ of thought". The principles of his theory could be equated with the later Sapir-Whorf hypothesis, which was proposed and supported with first empirical evidence in the 20th century. According to the Sapir-Whorf hypothesis[1] the grammatical features of the language we speak influences the way we see the world, conceptualize, categorize, and speak about it. In other words, speakers of different languages apply different language-specific perspectives on one and the same event existing in the external world (Boroditsky et al. 2003a; Lucy 1992; Mertins 2018; Slobin 1996; Whorf 1956, 2012). Studies focusing on linguistic relativity and grammatical gender have experimentally tested the effects of grammatical gender on the conceptualization and interpretation of nouns by speakers of different languages (Dilkina, McClelland & Boroditsky 2007; Schiefke & Stefanowitsch 2004; Landor 2014; Haertlé 2017). These studies have shown that grammatical gender can

1 This principle is later called *thinking for speaking* by Slobin (1996), and *seeing for speaking* by Mertins (2018).

lead to different kinds of conceptualizations. In other words, it affects cognitive processes and speakers' categorizations of the external world. However, most of these studies have concentrated on languages spoken in Europe, in which gender is grammatically marked. As a consequence of the WEIRD bias, this view on grammatical gender must be skewed.

In order to get a more impartial view of how gender influences thought, a detailed, systematic study is necessary that focuses on all alive languages from the Indo-European family, including languages without grammatical gender marking. Additionally, the strong focus of the research so far on preferences from English or German speakers cannot be taken as a basis for other, especially less studied languages. To broaden the focus in this way will help to identify possible differences in the conceptualization of objects in gendered compared with non-gendered languages.

The present study raises the following question: What is the current state of Indo-European languages concerning the presence or absence of grammatical gender features? Based on the existing linguistic and psycholinguistic research, one could hypothesize that most of the Indo-European languages are gendered. This is not the case, but rather a false impression given that the majority of the existing studies have investigated languages marking grammatical gender. The goal of this study is to show the linguistic bias of the current research in the field of linguistic relativity on the one hand, and to provide an overview of the expression of grammatical gender in all spoken L1 Indo-European languages on the other. In line with John Lucy (1997), only the study of languages in cross-linguistic contexts enables us to gain a full insight into the features, which would show support for the linguistic relativity hypothesis.

The structure of this paper is as follows: Section 2 encompasses a meta-analysis of existing studies on grammatical gender within the frame of linguistic relativity by critically reviewing a large-scale study. In Section 3 follows the overview of how Indo-European languages depict grammatical gender in their systems. Section 4 contains a presentation of the findings from the previous section. In Section 5 the overall results from Sections 2 and 3 are discussed. The article closes with a summary and a conclusion.

2. Criticism of a review on 43 empirical studies on grammatical gender in linguistic relativity

Many studies have examined the influence of grammatical gender on thought (Boroditsky et al. 2003b; Chen & Su 2011; Flaherty 2001; Konishi 1993; Sera, Berge & Pintado 1994; Vigliocco et al. 2005). A systematic and large-scale review of 43 empirical studies (from 1993 to 2018) was carried out by Samuel, Cole and Eacott (2019), in which all studies investigated the topic of the influence of grammatical gender on thought, testing 5,895 participants in total. The authors of this review classified these 43 articles based on the methods each study implemented and their findings in terms of +/- support for the hypothesis. In addition, six other parameters were part of the review:

1. The first parameter is *saliency of gender or sex* referring to the task of *voice assignment*. In this task the participants are instructed to explicitly assign a biological sex to the objects by assigning/attributing a female or male voice, when seeing a word or a picture.
2. The second parameter is *language.* This parameter controls the extent to which the tasks use language for finding evidence of the effect of language on thought as opposed to a non-verbal or a non-linguistic task such as application of eye-tracking methods. Therefore, the tasks were classified into two groups: Studies with high recruitment of language in their tasks and studies with low recruitment of language in their tasks.
3. The third parameter is *gendered language*. The underlying hypothesis here is that if someone's language creates a gender-bias conceptualization of the objects, such a conceptualization and gender-bias world view should be observed not only in the context of that particular gendered language, but also in the context of a non-gendered language. Accordingly, the studies were classified in terms of whether the participants were tested in a gendered or in a non-gendered language.
4. The next, and for this paper most relevant parameter, is the *number of genders*. It is important to pay attention that all of the 43 studies investigated only gendered languages, and excluded non-gendered languages. Consequently, the authors of this systematic review classified the studies in terms of whether the participants spoke a two- or three-gendered language and considered English as a non-gendered language.

5. The fifth parameter is *the effect with animate and inanimate targets,* which deals with the question of whether the studies reviewed only showed the influence of grammatical gender on the conceptualization of animate targets.
6. The final parameter is the *age* of the participants, i.e., stronger effects in support of linguistic relativity have been found among adults than children. Based on the results of these studies, weaker effects at younger ages are possibly consistent with less experience of younger children with grammatical gender. Furthermore, formal and institutional instructions play an important role in the formation of meta-linguistic knowledge about grammatical gender.

The methods implemented across the different articles for testing the effect of grammatical gender on the conceptualization of objects were as heterogeneous as *property judgment*, *voice choice*, *sex assignment*, *extrinsic affective Simon task*, *priming*, *similarity judgment*, *association*, and *object-name memory association*.

Along with the applied methods and the parameters listed above, the results of these 43 articles have been classified into three groups: *support* (studies showing consistent effects of grammatical gender on conceptualization), *mixed support* (studies showing partial confirmation of the influence of grammatical gender on thought, e.g., an effect of one particular gender only), and *no support* (studies showing no evidence for any effect of grammatical gender on thought).

Table 1 lists these 43 studies structured in separate columns according to the information about the *languages* being tested, the applied *method*, the percentage of the *overall studies applying a specific method*, and the results of the overall meta-analysis.

Table 1: Summary of 43 empirical studies on grammatical gender and linguistic relativity

Studies	Flaherty 2001; Imai et al. 2014; Saalbach et al. 2012; Haertlé 2017; Konishi 1993; Sedlmeier,Tipandaj, & Jänchen 2016; Semenuks et al. 2017; Landor 2014; Mickan, Schiefke & Stefanowitsch 2014; Montefinese et al. 2019	Almutrafi 2015; Athanasopoulos & Boutonnet 2016; Beller et al. 2015; Bender er al. 2016a; Bender et al. 2018; Haertlé 2017; Kurniski et al. 2016; Lambelet 2016: Ramos & Roberson 2011; Sera et al. 1994; Sera et al. 2002; Vernich 2017; Vernich, Argus, & Kamandulyte-Merfeldiene 2017; Bassetti 2007; Kurinski & Sera 2011; Forbes et al. 2008;	Pavlidou & Alvanoudi 2013; Pavlidou & Alvanoudi 2018; Belacchi & Cubelli 2012; Bender et al. 2016b; Flaherty 2001; Nicoladis 2019; Nicoladis et al. 2016; Nicoladis & Foursha Stevenson 2012; Sera et al. 1994;	Bender et al. 2016a, 2016b, 2018	Bender et al. 2011; Sato & Athanasopoulos 2018; Bender et al. 2011; Degani 2007; Mickan et al. 2014; Samuel et al. 2016
Languages	Spanish, English, German, Japanese, Polish, French, Tamil, Romanian, Finno-Ugric (the traditional name of Ulric Languages), Semitic, Italian	Arabic, English, French, Spanish, Norwegian, German, Polish, Hungarian, Portuguese, Estonian, Italian, Lithuanian, Italian, Russian	Greek, German, Italian, English, Spanish, Russian, French	German	German, French, English, Spanish, Romanian, Lithuanian, Italian, Bulgarian, Latvian, Ukrainian
Task type	Property judgment	Voice choice	Sex assignment	EAST	Priming
Percentage of the overall studies	37 %	28 %	12 %	9 %	8 %
support	3 %	64 %	66 %	11 %	34 %
Mixed support	23 %	22 %	26 %	56 %	-
No support	74 %	14 %	8 %	33 %	66 %

Based on the data presented in Table 1, it can be concluded that the absence/presence of a grammatical gender effect on conceptualization is highly related to the task/method used in a study. More specifically, when the tasks *voice choice* or *sex assignment* were used, in which the participants were explicitly asked to assign a biological gendered voice or a name when naming objects, the influence of grammatical gender was stronger. Following the argumentation of Samuel et al. (2019), these two tasks form the cornerstone for proving the linguistic relativity theory (in terms of *support*: voice choice 64 %; sex assignment 66 %). By contrast, the tasks *association*, and *property judgement* stand at the other end of the continuum, providing evidence for the absence (no support) of such effects.

In conclusion, it is important to consider the fact that all of the 43 empirical studies mentioned above investigated only fifteen Indo-European languages, all spoken in Europe, and five non-Indo-European languages. These fifteen languages cover merely 3.5 % of the 413 languages spoken within the Indo-European family. Such linguistic focus confirms yet again the Eurocentric bias often found in linguistic studies. In order to reduce this bias and to gain a more comprehensive as well as evidence-based view on the cross-linguistic occurrence of grammatical gender, the inclusion of other Indo-European languages into the analysis is a must. In line with this argumentation, in the following a detailed scrutiny of 413 Indo-European languages will be presented and analyzed in terms of +/- grammatical gender.

3. Indo-European languages

The backbone for the analyzes in the present paper was the language database *Glottolog* version 4.6 (Hammarström et al. 2022), which is a comprehensive reference information for the world's languages, with focus on the lesser-known languages. According to this database, the Indo-European language family represents one of the overall existing 427 language families. The Indo-European family consists of 581 languages, from which only 413 of the languages are alive and spoken, with the rest being extinct. In line with *Glottolog*, these 413 languages, including languages which are moribund,[2] threatened,[3] or endangered,[4] were considered for the purpose of the current study.[5] Following Hammarström et al. (2022), the Indo-European language family was classified as following:

2 This term refers to languages which are severely endangered and are near to death.
3 This term refers languages which are still spoken, but threatened to language change or lost.
4 This term refers the languages which are massively endangered but not yet extinct.
5 The classification does not include any extinct languages, such as the Anatolian group.

Classical Indo-European

- Albanian
- Armenic
- Balto-Slavic
- Celtic
- Germanic
- Graeco-Phrygian
- Indo-Iranian
- Italic

Figure 1 presents a map of the geographical distribution of all Indo-European languages, including the extinct groups.

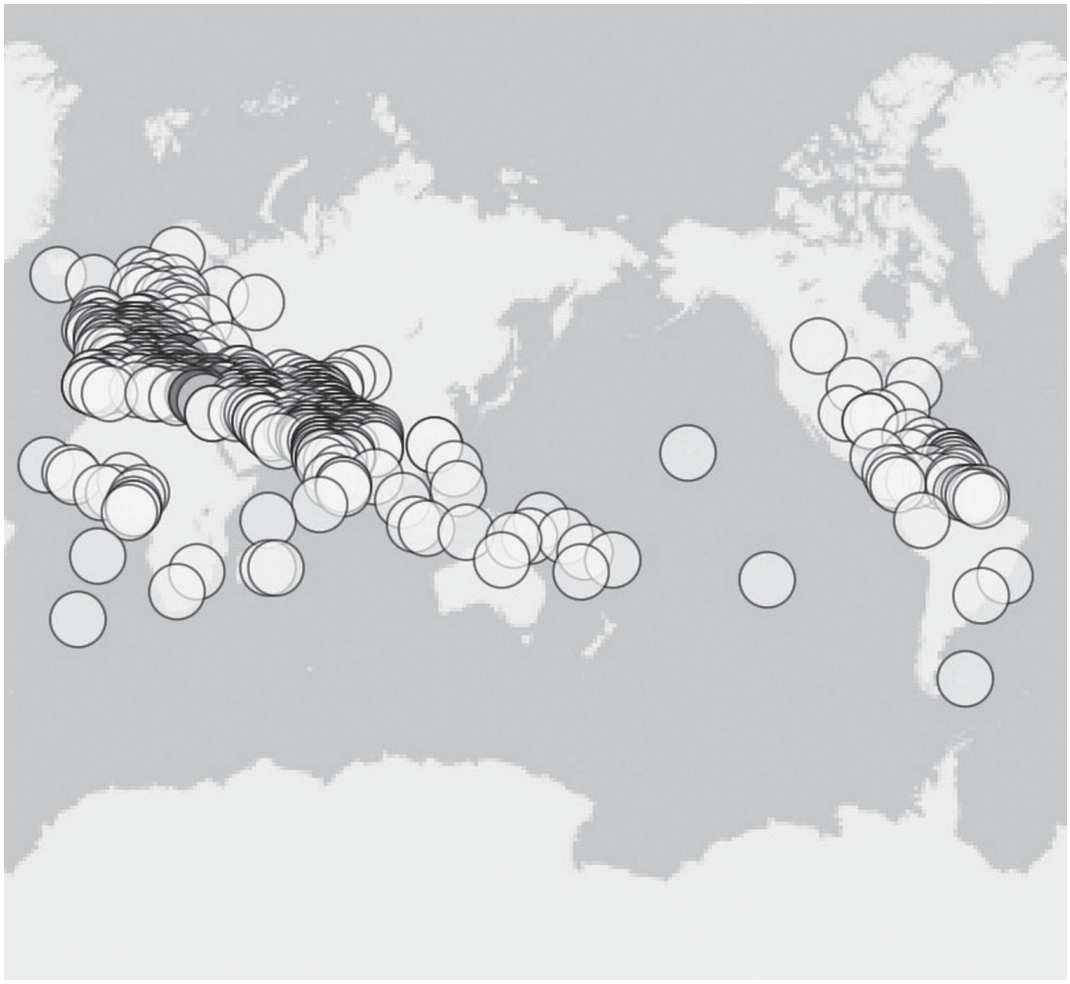


Figure 1: A map of the main Indo-European languages

As shown in Figure 1, contrary to what the name of this family of languages might imply, the languages of this group are not merely spoken in Europe and Asia but are rather spoken in a broader scope of continents such as in Africa, Australia, and even Central America. However, the fact that the density of the geographical location of these languages is mostly in the Middle East and Europe cannot be neglected.

All language groups listed above divide into subgroups called *sub-branch languages*. Each of these *sub-branch languages* also includes various dialects. For instance, Swedish is a sub-branch language belonging to the *Germanic* group and it branches further into four different dialects: *Götamal*, *old Swedish*, *Standard Swedish* and *Sveamal.* In this article, dialects in any sub-branch language are not part of the analysis.

The expression of gender as a grammatical feature is found not only in the Indo-European language group, but also in many other language families, for instance in the Afroasiatic languages, Atlantic Languages, Australian Languages, Austroasiatic Languages, Dravidian Languages, or Languages of the Americas. Despite this diversity, grammatical gender is far from constituting a universal feature (Ibrahim 2014). Even within the Indo-European language group, the marking of grammatical gender is not the default: "The inner distribution of grammatical gender marking is not homogenous within the entire Indo-European family" (Allassonnière-Tang & Dunn 2020). In order to find out whether or not a language from the Indo-European family belongs to +/- gender systems, we will examine this group by analyzing each sub-branch language in the next section.

4. Results

Having introduced the sub-branch groups of Indo-European languages, it is important to know how many languages are spoken in each group. Figure 1 shows the distribution of languages for each one of the eight groups presented in the above classification.

In line with Figure 2, the languages spoken in each sub-branch are not evenly distributed since most languages are found in the Indo-Aryan, Iranian, and Italic sub-branches.

The languages in each sub-branch were divided in terms of +/- expression of grammatical gender into five categories: (1) NGG – languages without grammatical gender; (2) 1GG- languages with one grammatical gender; (3) 2GG – languages with two grammatical genders; (4) 3GG – languages with three grammatical genders; and (5) No Info – no information found. It is

important to stress that in this paper the 1GG group includes languages with pronominal genders, such as English. There are other linguistic approaches, in which English is considered a non-gendered language (Sabourin, Stowe & de Haan 2006). The summary of the analyzes carried out for the purpose of this article are presented in Figure 3.

In Figure 3, the different sub-branch groups are plotted on the x-axis whereas the y-axis shows the number of languages belonging to each category, marked with a different color.[6]

Overall, only seven languages of all the analyzed Indo-European languages belong to the 1GG category. There are three languages in the Iranian, one in the Indo-Aryan, one in the Graeco-Phrygian, and two in the Germanic sub-branch group. As for the category 2GG, there are 96 languages in total, from which 39 languages were identified for the Indo-Aryan, 33 in the Italic, 13 in the Iranian, five in the Celtic, four in the Germanic, and two languages in the Balto-Slavic sub-branches. Regarding the 3GG systems, a total of 34 languages were identified belonging to this category. The majority, eleven languages, were found in the Balto-Slavic sub-branch, followed by seven languages in the Indo-Aryan, four languages in the Germanic as well as in the Italic sub-branch. In the Graeco-Phrygian group there were four languages, and finally Albanian with three languages. The last category No GG was found in 61 languages in total, with 44 languages in the Iranian group, followed by the Indo-Aryan with 10 languages and the Italic sub-branch with two languages. The sub-branch Armenic with a total of two languages also belongs to No GG. Finally, there was one language in the Germanic, and one language in the Graeco-Phrygian sub-branch, which belong to the No GG category.

6 As for 215 languages no linguistic information regarding their grammatical gender marking system is available, the category of No Info has not been presented in Figure 3.

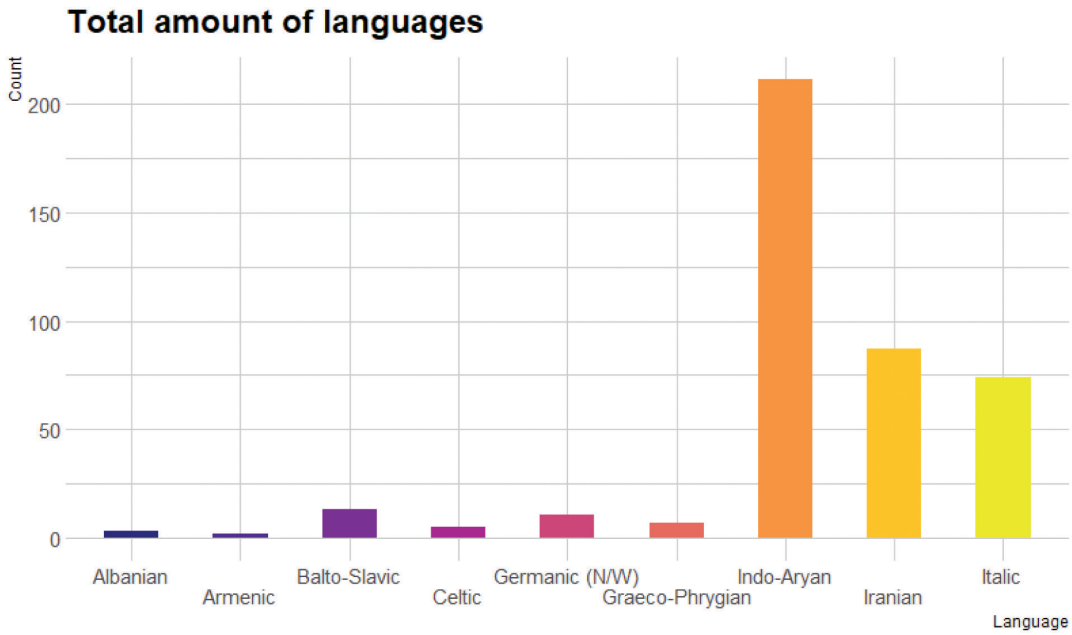


Figure 2: The number of languages in each sub-branch of Indo-European family

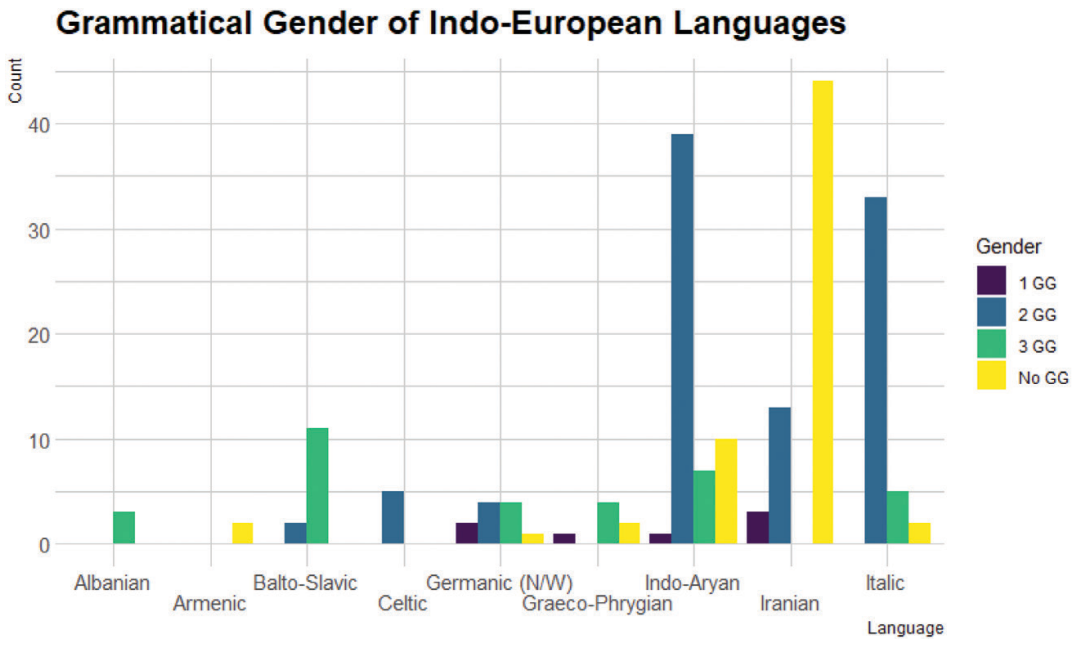


Figure 3: Grammatical gender of each sub-branch in Indo-European family

Looking at the results, it can be further observed that the sub-branches of Germanic and Indo-Aryan include more diverse grammatical gender systems, since both sub-branches have languages belonging to all four categories. Although Germanic languages are considered as languages with gender systems, it is important to highlight that the only Germanic language without a grammatical gender marking system is Afrikaans. This language is indeed the youngest fully standardized member of the West Germanic branch, spoken in South Africa. It is also considered as a daughter of Dutch and is characterized by extreme simplifications of Dutch morpho-syntax, including missing Dutch's gender system (Brown & Ogilvie 2009).

On the other hand, Armenic, Albanian, and Celtic sub-branches show the least diversity in terms of grammatical gender, since the two languages in the Armenic sub-branch only include No GG systems and all three languages from the Albanian sub-branch express grammatical gender in a threefold system, i.e., in the 3GG category. Finally, the alive languages belonging to the Celtic sub-branch assign gender solely in a twofold system of feminine and masculine.

It is also important to note that languages categorized as 1GG, 2GG, 3GG, and NGG made up only 47.9 % (198 languages) of the total number of languages spoken in the Indo-European family. This means that for more than half of the spoken languages in the Indo-European family, no linguistic research is available. Out of all 413 languages spoken in this family, only seven of them express gender as one category, mostly expressed as a pronominal feature; 96 languages have a twofold gender system, 34 express the category gender in a threefold system, and 61 languages have no grammatical genders at all.

In addition, taking the gendered Indo-European languages, it can be observed that most of them, 96 in total, have a twofold gender system. This means that among the Indo-European languages, most of the languages show a strong correlation between biological sex and grammatical gender for all animates (humans and animals), when compared to the three-gendered languages such as German.

Taking the three principles of semantic, morphological and phonological gender assignment into consideration, the results of the present study further confirm the fact mentioned in section 1 of the current article, that gender assignment in the Indo-European languages is mainly motivated semantically when referring to humans and some animals.

Given the high number of languages considered for this analysis, the detailed results for every sub-branch language are presented in appendices A1, A2, A3 and A4.

5. Discussion

Taken the Eurocentric and WEIRD biases in research in general and specifically in linguistic and psycholinguistic studies, this article has found additional evidence for these biases in the linguistic relativity framework, specifically in research on grammatical gender. In the systematic review of Samuel et.al. (2019), 43 empirical studies with the focus on grammatical gender marking on different languages were reviewed. Having critically re-assessed these 43 articles, it was found that only twenty languages were taken into account, among which fifteen Indo-European and five non-Indo-European languages. These fifteen Indo-European are mainly spoken in Europe.

These results inspired me to study all languages alive in the Indo-European family, with the focus on their underlying grammatical gender marking systems. The findings of the present study firstly convey that 72 % of the Indo-European languages are spoken outside Europe. These languages belong to the sub-branch languages of Indo-Aryan and Iranian languages, which all together comprise 298 languages of the total of 413 alive languages in this language family. Secondly, it was found that only 198 languages have linguistic proof for the absence or presence of a gender-marking systems. These languages belong to the sub-branches of Indo-Aryan, Italic, Iranian, Celtic, Germanic, Balto-Slavic, Albanian, Armenic, and Graeco-Phrygian groups. Furthermore, the results have revealed that most of these languages express gender in a twofold system of masculine and feminine. Therefore, it can be claimed that most of the grammatical gender of nouns referring to animals and humans in these twofold gendered languages are congruent with their biological sex. This confirms that the majority of the gendered Indo-European languages classify nouns based on the semantic principle. However, it should not be overlooked that the classification of inanimate nouns among these twofold gendered languages are semantically arbitrary. The reason is that these nouns are, by definition, inanimate and have therefore no biological sex, i.e., gender is not an intrinsic characteristic of those nouns (Bender et al. 2018). Finally, out of all Indo-European languages, 215 languages have not been linguistically studied regarding their expression of grammatical gender.

Moreover, having presented the categories of grammatical gender marking systems among Indo-European languages, it is also important to include the precise nature of grammatical gender marking systems. The categories of No GG, 1GG, 2GG and 3GG can be explained as forming a continuum, on each of their extreme ends including the categories of No GG and 3 GG. It is important to mention that *sex* and *gender* are distinguished from the sociological perspective.

Sex is a biological term, which refers to the chromosomes, hormonal profiles, and internal and external sex organs, since *gender* is a cultural term and denotes the characteristics that societies or cultures delineate along the spectrum of masculinity and femininity (Butler 2002). Taking German as an example for a 3 GG language, some nouns referring to human/animals can be claimed to carry a grammatical gender inconsistent with their social gender, due to the existence of the third neuter gender. For instance, the lexeme *girl* in German is expressed with neutral gender *das Mädchen*, although by virtue of being probably socially feminine, the neutral grammatical gender would not accord with the social gender. No GG languages, on the other hand, do not in any way mark gender grammatically. However, it can be argued that despite the absence of this grammatical feature, nouns can be associated with specific genders culturally. As mentioned above, 1GG languages refer to the languages with pronominal genders. Such languages pronominally refer to humans and animals consistent with their social gender. 2GG languages in the Indo-European family form two groups. One group divides the nouns into feminine and masculine such as Spanish, French, and Italian. This group forms the most frequent type of twofold gender marking systems among the Indo-European languages. The second group, which is less common, take a different approach in their classification of the nouns: They classify nouns into *gendered* and *neuter* categories. Dutch and Swedish are two instances of such Indo-European languages with a different take on the classifications of nouns. Looking at Samuel et al. (2019), it should also be emphasized that such languages have also been excluded in their review.

In this discussion, the focus is also on the areas which need addressing in future research on grammatical gender in linguistic relativity. Although the systematic review by Samuel et al. (2019) provided evidence for the effect of grammatical gender on conceptualization and thought, testing only fifteen languages of the 413 Indo-European languages does not provide comprehensive results. The generalization of these findings based on only fifteen languages to the complete family can be considered as debatable. Following John Lucy (1997), it is only through cross-linguistic studying of languages that we can gain full insights into the features, which show support for linguistic relativity. Thus, more languages need to be studied to have a better understanding of language-specific effects of grammar on the conceptualizations of the world. As the results of the current data on the Indo-European languages have proven, there exists a diversity in the marking of grammatical gender within the Indo-European family. Consequently, studies must be carried out including languages which mark gender semantically, morphologically and phonologically, but also including languages which do not express gender grammatically at all. An example of a language with no

gender marking system is Farsi, in which gender information can be presented through a lexical description and/or by context. As mentioned above, although nouns in Farsi do not mark gender, conveying such information through the lexical insertion of the words *masculine* and *feminine* in postnominal positions is possible and does not violate the underlying structural rules. Therefore, the concept of gender/sex does exist in Farsi, albeit not grammatically marked. Comparing the results of empirical studies on such non-gendered languages with gendered languages is therefore essential to fill the gap in study of the effects of linguistic diversity on cognition.

Another potential point of relevance of taking more non-gendered languages into consideration, besides finding further support for linguistic relativity, is the acquisitional perspective, especially in migration contexts. In line with the main topic of the current study, it is also relevant to pose the following questions: How do native speakers of non-gendered languages acquire and master gendered languages, such as German, French, or English? How do second language speakers of these languages adjust to the extreme gendered classifications of every noun? Having briefly introduced Farsi as one example of a non-gendered language, one could relate the above raised questions with the consistent increase in migration among Iranians. According to a study at Stanford University in 2020, the total number of Iranian migrants grew from about 130,000 in 1970 to 3.1 million in 2020 (Azadi et al. 2020). Among the countries which host the largest number of Iranian migrants are the United States (32 %), Canada (14 %), Germany (11 %), United Kingdom (6 %) and Sweden (5 %). In accordance with this diaspora of Iranian speakers to countries in which gendered languages are the dominant language of the society in every aspect, it is relevant to study the linguistic behavior of Farsi speakers in acquiring such languages.

This study is not free of limitations. The biggest limitation is that there is no comprehensive resource for all spoken languages in the Indo-European family that could serve as a basis for the analysis. Although there exists reliable linguistic works such as *Encyclopaedia of Languages of the World* (Brown & Ogilvie 2009), they are not up to date because they do not include information on language change and are therefore inaccurate. As pointed out before, the list of languages in the appendix display a large number of languages belonging to the category of *No Info*. This means it is difficult to draw conclusions on the results. Nonetheless, the results prove the fact that further research is necessary to fill in the gaps considering many other understudied languages.

6. Conclusions

Based on the analyzes of the current article, diverse forms of representing gender of animate and inanimate entities (semantically, morphologically, phonologically, and by lexical description) in various languages have made this grammatical feature a valid and relevant category for studying diversity in cognition. In line with the linguistic relativity theory, grammatical gender is a relevant feature for finding evidence for the effects of language on conceptualization. Nevertheless, the findings of this study clearly manifest the existence of the Eurocentric and the WEIRD bias in studies on grammatical gender within the linguistic relativity principle. The current study showed that the review by Samuel et al. (2019) is clearly affected by both biases. On the basis of the present study, it can be concluded that among all Indo-European languages a long list of understudied languages exists. Therefore, a more comprehensive and evidence-based research on all, especially understudied Indo-European languages is necessary in order to corroborate the findings on the effects of grammatical gender on thought. This article has further paved the way for future studies by providing the existing linguistic information about all sufficiently documented alive languages from the Indo-European family on the +/- expression of grammatical gender.

References

Adib, D. (1972). *Further data on khuniite from the Tschah Khuni Mine, Anarak, Iran.*

Aikhenvald, A. Y. (2000). *Classifiers: A typology of noun categorization devices*. Oxford: Oxford University Press.

Aikhenvald, A. Y. (2012). Round women and long men: Shape and size in gender choice in Papua New Guinea and beyond. *Anthropological Linguistics*, *54*(1), 33–86.

Alam, S. (2011). Word formation in Surjapuri. *Language in India, 11*(3), 487–497.

Alimova, K. (2020). Dari word-forming transposotion phenomenon. *International Journal of Pharmaceutical Research, 12*, 516–521.

Allassonnière-Tang, M. & Dunn, M. (2020). The evolutionary trends of grammatical gender in Indo-Aryan languages. *Language Dynamics and Change*, *11*(2), 211–240.

Almutrafi, F. (2015). *Language and cognition: Effects of grammatical gender on the categorisation of objects* (Unpublished doctoral thesis). Newcastle University, Newcastle upon Tyne, UK.

Alvanoudi, A. (2015). *Grammatical gender in interaction: Cultural and cognitive aspects*. *Brill's studies in language, cognition and culture: v. 9*. Brill, Leiden, The Netherland.

Andonova, E., D'amico, S., & Decescovi, A. (2004). Gender and lexical access in Bulgarian. *Perception & Psychophysics, 66*(3), 496–507.

Anonby, E. J. (2003). Update on Luri: How many languages? *Journal of the Royal Asiatic Society of Great Britain & Ireland*, *13*(2), 171–197. Retrieved (03/2023) from https://doi.org/10.1017/S1356186303003067

Arora, A. (2020, December 6). *Historical phonology and other observations on historical phonology and other observations on Kholosi*. Georgetown University, Washington, United States.

As-Sultan, Ab. H. J. (2010). Causative in Shabaki. *Acta Linguistica, 4* (2), 79–88.

Athanasopoulus, P. (2009). Cognitive representation of colour in bilinguals: The case of Greek blues. *Bilingualism: Language and Cognition*, *12*(1), 83–95.

Athanasopoulos, P. & Boutonnet, B. (2016). Learning grammatical gender in a second language changes categorization of inanimate objects: Replications and new evidence from English learners of L2 French. In R. Alonso (Ed.), *Cross-linguistic influence in second language acquisition* (pp. 173–192). Bristol, UK: Multilingual Matters.

Audring, J. (2008). Gender assignment and gender agreement: Evidence from pronominal gender languages. *Morphology*, *18*(2), 93–116. Retrieved (03/2023) from https://doi.org/10.1007/s11525-009-9124-y

Azadi, P., Mirramezani, M., & Mesgaran, M. B. (2020). Migration and brain drain from Iran. *Stanford Iran*, *2040*, 1–30.

Baart, J. (1999). *Sketch of Kalam Kohistani grammar: Studies in languages of Northern Pakistan*. Quaid-i-Azam University and Summer Institute of Linguistics, Islamabad, Pakistan.

Bano, S., Deo, A., & Butt, M. (2019). Achievement predicates and tense paradigms in Hazaragi. *Proceedings of the LFG'19 Conference*, 394–414.

Barber, H. & Carreiras, M. (2005). Grammatical gender and number agreement in Spanish: An ERP comparison. *Journal of Cognitive Neuroscience, 17*(1), 137–153.

Barker, M. S., Nelson, N. L., & Robinson, G. A. (2020). Idea formulation for spoken language production: The interface of cognition and language. *Journal of the International Neuropsychological Society: JINS*, *26*(2), 226–240.

Bartens, A. & Sandström, N. (2008). Explaining Kabuverdianu nominal plural formation. *Language Complexity: Typology, Contact and Change*, 94, 305–320, John Benjamins, Amsterdam, The Netherlands.

Bashir, E., Conners, T. J., & Hefright, B. (2019). *A descriptive grammar of Hindko, Panjabi, and Saraiki*. *Mouton-CASL Grammar Series*: *volume 4*. De Gruyter Mouton.

Bassetti, B. (2007). Bilingualism and thought: Grammatical gender and concepts of objects in Italian–German bilingual children. *International Journal of Bilingualism*, *11*, 251–273.

Batzarov, Z. (2000). *Orbis Latinus: Descriptive grammar of the Occitan language*. https://www.orbilat.com/Languages/Occitan/Grammar/index.html

Bavanpouri, M., Mirzaei Al-Husseini, S. M., & Fathi Iranshahi, T. (2021). Investigation of meaning and phonological processes of Arabic words involved in Laki language [In Persian]. *Journal of Language Teaching, Literature & Linguistics, 4*(1), 29–48.

Baxter, A. N. (1988). *A grammar of kristang (malacca creole portuguese)*. The Australian National University.

Belacchi, C. & Cubelli, R. (2012). Implicit knowledge of grammatical gender in preschool children. *Journal of Psycholinguistic Research*, *41*, 295–310.

Beller, S., Brattebø, K. F., Lavik, K. O., Reigstad, R. D., & Bender, A. (2015). Culture or language: What drives effects of grammatical gender? *Cognitive Linguistics*, *26*, 331–359.

Bender, A., Beller, S., & Klauer, K. C. (2011). Grammatical gender in German: A case for linguistic relativity? *Quarterly Journal of Experimental Psychology*, *64*, 1821–1835.

Bender, A., Beller, S., & Klauer, K. C. (2016a). Crossing grammar and biology for gender categorisations: Investigating the gender congruency effect in generic nouns for animates. *Journal of Cognitive Psychology*, *28*, 530–558.

Bender, A., Beller, S., & Klauer, K. C. (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*, 48–64.

Bender, A., Beller, S., & Klauer, K. C. (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*, 1580–1608.

Benkato, A. (2015). Recent work in Yaghnobi studies. *Iran and the Caucasus, 19*(3), 283–294.

Bernstein, J. (1991). DPs in French and Walloon: Evidence for parametric variation in nominal head movement. *Probus*, *3*(2). Retrieved (03/2023) from https://doi.org/10.1515/prbs.1991.3.2.101

Bhattacharya, S., Choudhury, M., Sarkar, S., & Basu, A. (2005). Inflectional morphology synthesis for Bengali noun, pronoun and verb systems.

Proceedings of the National Conference on Computer Processing of Bangla (NCCPB 05), 34–43.

Borijan, H. (2004). Review: Northern Talysh by Wolfgang Schulze. *Iran & the Caucasus, 8*(1), 165–168.

Borjian, H. (2008). *The Komisenian dialect of aftar. Advance online publication.* Retrieved (03/2023) from https://doi.org/10.7916/D8N02H9X

Borjian, H. (2014). What is Judeo-Median – and how does it differ from Judeo-Persian? *Journal of Jewish Languages, 2*(2), 117–142.

Borjian, H (2019). *The Caspian language of Šahmirzād. Journal of the American Oriental Society, 2*(139), 361–380.

Borjian, H. [Habib] (2019). The Mazandarani dialect of Kalijān Rostāq. *Iranian Studies, 52*(3–4), 551–573. Retrieved (03/2023) from https://doi.org/10.1080/00210862.2019.1611371

Borjian, H. (2021). *Essays on three Iranian language groups: Taleqani, Biabanaki, Komisenian*. Bristol, CT (USA), ISD LLC.

Boroditsky, L. (2012). How the languages we speak shape the ways we think. In M. Spivey, K. McRae, & M. Joanisse (Eds.), *The Cambridge handbook of psycholinguistics* (pp. 615–632). Cambridge University Press, Cambridge, United Kingdom. Retrieved from https://doi.org/10.1017/CBO9781139029377.042.

Boroditsky, L. & Schmidt, L. A. (2000). Sex, syntax, and semantics. In L. R. Gleitman & A. K. Joshi (Eds.), *Proceedings of the twenty second annual conference of the cognitive science society* (pp. 42–46). Mahwah, NJ: Erlbaum.

Boroditsky, L., Schmidt, L. A., & Phillips, W. (2003a). Sex, syntax, and semantics. *Language in Mind: Advances in the Study of Language and Thought* (22), 61–79.

Boroditsky, L., Schmidt, L. A., & Phillips, W. (2003b). Can quirks of grammar affect the way you think? Grammatical gender and object concepts. *Proceedings of the Annual Meeting of the Cognitive Science Society* (25), 928–933.

Braine, M. D. (2014). What is learned in acquiring word classes – A step toward an acquisition theory. In Mechanisms of Language Acquisition, *The 20th annual Carnegie Mellon symposium on cognition.*

Brown, K. & Ogilvie, S. (2009). *Concise encyclopedia of languages of the world.* Elsevier Ltd, Oxford, United Kingdom.

Bussmann, H. & Hellinger, M. (Eds.). (2001, 2002, 2003). *Gender across languages: The linguistic representation of women and men*. J. Benjamins, Amsterdam, The Netherlands.

Butler, J. (2002). *Gender trouble*. Routledge, New York, United States.

Chambers, F. M. (1985). *An introduction to old Provençal versification* (Vol. 167). American Philosophical Society, Philadelphia, United States.

Chen, J.-Y. & Su, J.-J. (2011). Differential sensitivity to the gender of a person by English and Chinese speakers. *Journal of Psycholinguistic Research*, *40*(3), 195–203. Retrieved (03/2023) from https://doi.org/10.1007/s10936-010-9164-9

Chockalingam, K. (2016). *Languages of Tamil Nadu Lambadi an Indo-Aryan dialect*. Messrs Alamu Printing, Part XII, Vol-IX, India.

Comrie, B. (Ed.). (1987). *The world's major languages*. Croom Helm, New York, United States.

Corbett, G. (1991). *Gender: Textbooks in linguistics*. Cambridge University Press, Cambridge.

Dash, B. (2014). Kurmali noun inflection: Aspectual inquiry into gender and number. *International Journal of Linguistics* (3), 1–14. Retrieved (03/2023) from https://doi.org/10.13140/RG.2.1.4993.8004.

David, A. B. (2013). Descriptive grammar of Pashto and its dialects. *Descriptive Grammar of Pashto and Its Dialects (Vol.1)*. Walter de Gruyter, Berlin, Germany.

Degani, T. (2007). *The semantic role of gender: Grammatical and biological gender match effects in English and Spanish* (Unpublished master's thesis). University of Pittsburgh, Pittsburgh, PA.

Dhakal, D. N. (2011). *The Darai language*. UGS Press, Tribhuvan, Kirtipur, Kathmandu.

Dhanunjay, B. (2010). *Pangwali grammar write-up*. Retrieved (03/2023) from https://www.pangi.in/en/Pangwali-Linguistic-papers

Dilkina, K., McClelland, J. L., & Boroditsky, L. (2007). How language affects thought in a connectionist model. *Proceedings of the Annual Meeting of the Cognitive Science Society, 29,* 215–220.

Dixon, R. M. W. (1972). *The Dyirbal language of North Queensland*. London: CambridgeUniversity Press.

Dovudov, G., Pomikálek, J., Suchomel, V., & Smerk, P. (2011). Building a 50M Corpus of Tajik Language (RASLAN), 89–95.

Dryer, M. S. & Haspelmath, M. (2013). *The world atlas of language structures online*. Max Planck Institute for Evolutionary Anthropology. Retrieved (03/2023) from https://wals.info/feature/30A#2/26.7/148.9

Dwivedi, A. V. (2013). *A grammar of Bhadarwahi*. *Languages of the world. Materials: Vol. 495*. LINCOM Europa, Munich, Germany.

Enguita-Utrilla, J. M. (2008). Spanish and historical dialect in Aragon. *International Journal of the Sociology of Language*, *2008*(193/194). Retrieved (03/2023) from https://doi.org/10.1515/IJSL.2008.048

Evans, V. & Green, M. (2006). *Cognitive linguistics: An introduction*. Edinburgh University Press, Edinburgh, Scotland.

Everett, C. (2013). *Linguistic relativity*. De Gruyter Mouton, Berlin, Germany.

Fareed, N. (2016). Morphology of Mewati language. *Proceedings of the Conference on Language and Technology*, 25–32.

Firestone, C. & Scholl, B. J. (2016). Cognition does not affect perception: Evaluating the evidence for "top-down" effects. *The Behavioral and Brain Sciences*, *39*, e229.

Flaherty, M. (2001). How a language gender system creeps into perception. *Journal of Cross-Cultural Psychology*, *32*(1), 18–31.

Forbes, J. N., Poulin-Dubois, D., Rivero, M. R., & Sera, M. D. (2008). Grammatical gender affects bilinguals' conceptual gender: Implications for linguistic relativity and decision making. *Open Applied Linguistics Journal*, *1*, 68–76.

Friedmann, N. & Biran, M. (2003). When is gender accessed? A study of paraphasias in Hebrew anomia. *Cortex*, *39*(3), 441–463. Retrieved (03/2023) from https://doi.org/10.1016/S0010-9452(08)70258-2

Gambhir, S. K. & Moag, R. F. (1981). Fiji Hindi. A basic course and reference grammar. *Journal of the American Oriental Society*, *101*(4), 506. Retrieved (03/2023) from https://doi.org/10.2307/601362

Gnanadesikan, A. E. (2017). *Dhivehi: The language of the Maldives. Mouton-CASL grammar series: Vol. 3*. De Gruyter Mouton, Berlin, Germany.

Greenbaum, S. (1996). *English grammar*. Oxford University, Oxford, United Kingdom.

Gusain, L. (1999). *A discriptive grammar of Bagri* [PhD Dissertation]. New Delhi: Jawaharlal Nehru University.

Gusain, L., Ph.D. (2002). *Language in India: Eendangered language: A case study of Sansiboli*. Retrieved (03/2023) from http://www.languageinindia.com/dec2002/sansi.html

Haertlé, I. (2017). Does grammatical gender influence perception? A study of Polish and French speakers. *Psychology of Language and Communication*, *21*(1), 386–407.

Haig, G. & Khan, G. (Eds.). (2018). *The languages and linguistics of Western Asia: An areal perspective: (Vol. 6)*. Walter de Gruyter GmbH & Co KG, Berlin/ Boston.

Hammarström, H., Forkel, R., Haspelmath, M., & Bank, S. (2022). *glottolog/glottolog: Glottolog database 4.6.* Retrieved (03/2023) from https://doi.org/10.5281/zenodo.6578297

Heine, B. (1982). African noun class systems. In H. Seiler & C. Lehmann (Eds.), *Apprehension: Das sprachliche Erfassen von Gegenständen, i: Bereich und Ordnung der Phänomene* (pp. 189–216). Tübingen: Narr.

Henrich, J., Heine, S. J., & Norenzayan, A. (2010). Most people are not WEIRD. *Nature*, *466*(7302), 29. Retrieved (03/2023) from https://doi.org/10.1038/466029a

Hickmann, M. (2000). Linguistic relativity and linguistic determinism: Some new directions. *Linguistics,* 38(2). Retrieved (03/2023) from https://doi.org/10.1515/ling.38.2.409

Hirvonen, M. & Igareda, P. (2017). Imaginary places–Verbalization of setting in the English, German and Spanish audio descriptions of Slumdog Millionaire. Minor *t*ranslating *m*ajor-*m*ajor *t*ranslating *m*inor-*m*inor *t*ranslating *m*inor, *341*, Diavlos Publications, Athens, Greece.

Ibrahim, M. H. (2014). *Grammatical gender*. De Gruyter Mouton, The Hague, Netherlands. Imai, M., Schalk, L., Saalbach, H., & Okada, H. (2014). All giraffes have female-specific properties: Influence of grammatical gender on deductive reasoning about sex-specific properties in German speakers. *Cognitive Science, 38*, 514–536.

Ivanovska, B. (2015). *The category of definiteness in language pair German-Macedonian*. Anchor Academic Publishing, Hamburg, Germany.

Ivanow, W. (1926). Two dialects spoken in the Central Persian desert. *Journal of the Royal Asiatic Society, 58*(3), 405–431.

Jenner, H. (1904). *A handbook of the Cornish language: Chiefly in its latest stages, with some account of its history and literature*. D. Nutt.

Joseph, G. G., Reddy, V., & Searle-Chatterjee, M. (1990). Eurocentrism in the social sciences. *Race & Class*, *31*(4), 1–26. Retrieved (03/2023) from https://doi.org/10.1177/030639689003100401

Kalnaca, A. (2014). *A typological perspective on Latvian grammar*. DE GRUYTER OPEN. Retrieved (03/2023) from https://doi.org/10.2478/9783110411317

Karacan, H. & Khalid, H. S. (2016). Adjectives in Kurdish language: Comparison between dialects. *International Journal of Kurdish Studies, 2*(2), 15–23.

Karatsareas, P. (2011). *A study of Cappadocian Greek nominal morphology from a diachronic and dialectological perspective* [PhD Dissertation]. Wolfson College, Cambridge.

Karim, S. O. (2021). *The synchrony and diachrony of New Western Iranian nominal morphosyntax* [Doctoral dissertation, The Ohio State University)].

RIS. Retrieved (03/2023) from https://search.proquest.com/openview/92c24f6c9f4944f07333190c95228fe8/1?pq-origsite=gscholar&cbl=18750&diss=y

Karimi Doostan, G. H. & Tajalli, T. (2015). Definite and indefinite articles in Ardekani dialect of fars within the minimalist framework. *Journal of Researchers in Linguistics, 7*(2), 85–100.

Katz, J. & Pitzanti, G. (2019). The phonetics and phonology of lenition: A Campidanese Sardinian case study. *Laboratory Phonology: Journal of the Association for Laboratory Phonology, 10*(1), 16. Retrieved (03/2023) from https://doi.org/10.5334/labphon.18

Kaushanskaya, M. & Smith, S. (2016). Do grammatical-gender distinctions learned in the second language influence native-language lexical processing? *International Journal of Bilingualism*, *20*, 30–39.

Kiran, S. & Peterson, J. (Preprint 2021). *Sadani/Sadri: Новые индоарийские языки.(New Indo-Aryan languages)* [Redcall. LEE Kulikov, I Oranskaya, A Yu Rusakov, 367–379].

Konishi, T. (1993). The semantics of grammatical gender: A cross-cultural study. *Journal of Psycholinguistic Research*, *22*(5), 519–534. Retrieved (03/2023) from https://doi.org/10.1007/BF01068252

Korangy, A. & Mahmoodi-Bakhtiari, B. (2019). *Essays on typology of Iranian languages*. De Gruyter Mouton, Berlin/ Boston.

Korn, A. (2017). Notes on the nominal system of Bashkardi. *Transactions of the Philological Society, 115*(1), 79–97.

Koul, O. N. & Wali, K. (2006). *Modern Kashmiri grammar*. Dunwoody Press, Springfield, United States.

Kumar, P. & Grace, G. V. (2017). *A write-up of Bhariati grammar*. ILL, Dehradun, India.

Kurinski, E., Jambor, E., & Sera, M. D. (2016). Spanish grammatical gender: Its effects on categorization in native Hungarian speakers. *International Journal of Bilingualism*, *20*, 76–93.

Kurinski, E. & Sera, M. D. (2011). Does learning Spanish grammatical gender change English-speaking adults' categorization of inanimate objects? *Bilingualism: Language and Cognition*, *14*, 203–220.

Kushal, G. & Padra, B. (2021). Agreement in Sambalpuri: A lesser-known language. *Online International Interdisciplinary Research Journal* (11), 49–61.

Lakoff, G. (1987). *Women, fire and dangerous things: What categories reveal about the mind.* The University of Chicago Press.

Lambelet, A. (2016). Second grammatical gender system and grammatical gender-linked connotations in adult emergent bilinguals with French as a second language. *International Journal of Bilingualism*, *20*, 62–75.

Lamm, W. (2001). *Scottish Gaelic*. LINCOM Europa, Munich, Germany.

Landor, R. V. (2014). *Grammatical categories and cognition across five languages: The case of grammatical gender and its potential effects on the conceptualisation of objects*. DataCite, Griffith University, Brisbane, Australia.

Lautin, J. (2016). *Grammatical gender in Hindukush languages: An aerial-typological study* [Bachelor's thesis]. Stockholm University, Stokholm, Sweden, a research project supported by the Swedish Research Council.

Learn Haitian: Haitian Gender. (n.d.). Retrieved (03/2023) from http://learn101.org/haitian_gender.php

Levelt, W. J., Roelofs, A., & Meyer, A. S. (1999). A theory of lexical access in speech production. *The Behavioral and Brain Sciences*, *22*(1), 1–38; discussion 38–75. Retrieved (03/2023) from https://doi.org/10.1017/S0140525X99001776

Liljegren, H. (2008). *Towards a grammatical description of Palula: An Indo-Aryan language of the Hindu Kush*, Arkitektkopia AB, Stockholm, Sweden.

Liljegren, H. (2019). Gender typology and gender (in) stability in Hindu Kush Indo-Aryan languages. *Grammatical Gender and Linguistic Complexity* (1), 279–328.

Lipski, J. M. (2001). Chabacano/Spanish and the Philippine linguistic identity: Shedding light on the Chabacano language: Learning from general linguistic and similar cases (= *Estudios de Sociolingüística 2.2*), 119–164.

Lucy, J. A. (1992). *Language diversity and thought: A reformulation of the linguistic relativity hypothesis. Studies in social and cultural foundations of language: Vol. 0012*. Cambridge University Press, New York, United States.

Lucy, J. A. (1997). Linguistic relativity. *Annual Review of Anthropology*, *Vol. 0012,* 291–312.

Lucy, J. A. (2020, September 18). *Influence of language on thought: Relativity, acquisition, and standardization.* Department of Anthropology. University of New Mexico, Albuquerque, NM, via Zoom. Retrieved (03/2023) from https://www.youtube.com/watch?v=cMGwu396N50&t=267s

Maciuszek, J., Polak, M., & Świa Tkowska, N. (2019). Grammatical gender influences semantic categorization and implicit cognition in Polish. *Frontiers in Psychology*, *10*, 2208.

MacKinnon, C. R. (1974). *The phonology and morphology of Dezfuli-Shushtari: A study in West Persian dialectology*. University of California: Los Angeles.

Maffei, A. F. X. (1882). *A Konkani grammar*. Basel Mission & Tract Depository.

Martinez, I. M. & Shatz, M. (1996). Linguistic influences on categorization in preschool children: A crosslinguistic study. *Journal of Child Language*, *23*(3), 529–545. Retrieved (03/2023) from https://doi.org/10.1017/s030500090000893x

Martins, C. & Santos, I. (2015). Standardising mirandese: Processes and challenges. *Sociolinguistica, 29*(1), 141–156.

Matras, Y. (2005, cop. 2002). *Romani: A linguistic introduction* (1st ed.). Cambridge University Press. Retrieved (03/2023) from https://doi.org/10.1017/cbo9780511486791

Matras, Y. (2012). *A grammar of Domari*. *Mouton Grammar Library [MGL]: Vol. 59*. De Gruyter Mouton: De Gruyter. Retrieved (03/2023) from https://doi.org/10.1515/9783110291421

Maurer, P. (2008). A first step towards the analysis of tone in Santomense. *Roots of Creole Structures*, Vol. xvii, 253–261.

Maurer-Cecchini, P. (2021). *A grammar of Tuatschin: A Sursilvan Romansh dialect*. Language Science Press, Berlin, Germany.

Mehta, J. (2018, November 2). *Here's everything you need to know about Mewari language*. Retrieved (03/2023) from https://www.udaipurblog.com/mewari-language.html

Melissaropoulou, D. (2019). Morphological innovations in Propontis Tsakonian. *He Morphology of Asia Minor Greek*, Vol.13, 284–314.

Mertins, B. (2018). *Sprache und Kognition*. DE GRUYTER, Berlin, Deutschland.

Michaelis, S. M., Maurer, P., Haspelmath, M., & Huber, M. (Ed.). *The survey of pidgin and creole language structures.* Oxford University Press, USA.

Mickan, A., Schiefke, M., & Stefanowitsch, A. (2014). Key is a llave is a Schlussel: A failure to replicate an experiment from Boroditsky et al. 2003. *Yearbook of the German Cognitive Linguistics Association*, *2*, 39–50.

Mirdehghan, M. & Yusefi, S. (2013). Case and case marking in Vafsi. *Language Studies, 3*(5), 85–105.

Montefinese, M., Ambrosini, E., & Roivainen, E. (2019). No grammatical gender effect on affective ratings: Evidence from Italian and German languages. *Cognition and Emotion*, *33*, 848–854. Retrieved (03/2023) from https://doi.org/10.1080/02699931.2018.1483322

Moreton, R. L. (2001). *Mississippi Gulf Coast French: Phonology and morphology*. Tulane University, Language, Linguistics.

Newmark, L., Hubbard, P., & Prifti, P. (1982). *Standard Albanian: A reference grammar for students*. Stanford University Press.

Nicoladis, E. (2019). [Unpublished data].

Nicoladis, E., Da Costa, N., & Foursha-Stevenson, C. (2016). Discourse relativity in Russian-English bilingual preschoolers' classification of objects by gender. *International Journal of Bilingualism, 20*, 17–29.

Nicoladis, E. & Foursha-Stevenson, C. (2012). Language and culture effects on gender classification of objects. *Journal of Cross-Cultural Psychology, 43*, 1095–1109.

Nourzaei, M., Jahani, C., Anonby, E., & Ahangar, A. A. (2015). *Koroshi: A corpus-based grammatical description*. Acta Universitatis Upsaliensis, Uppsala, Sweden.

Oberlies, T. (2011). *Pali: A grammar of the language of the Theravada Tipitaka. With a concordance to Pischel's Grammatik der Prakrit-Sprachen*. Walter de Gruyter, Berlin/ New York.

O'Donovan, J. (1845). *A grammar of the Irish language.*

Ojha, A. K., Behera, P., Singh, S., & Jha, G. N. (2015). Training & evaluation of POS taggers in Indo-Aryan languages: A case of Hindi, Odia and Bhojpuri. *The proceedings of 7th language & technology conference: Human language technologies as a challenge for computer science and linguistics*, 524–529.

Palmer, T. (2016). *Verbal morphology and grammatical aspect in Sarikoli* [Master Thesis]. University of North Dakota.

Parsieva, L. K. & Gatsalova, L. B. (2014). Linguistic and cultural specificity of the derivational category of grammatical persons' nominations in Ossetian language. *Journal of Language and Literature,* 5(4), 217–219.

Pavlidou, T.-S. & Alvanoudi, A. (2013). Grammatical gender and cognition. In N. Lavidas, T. Alexio, & A. Sougari (Eds.), *Major trends in theoretical and applied linguistics* (Vol. 2, pp. 109–124). London, UK: Versita.

Pavlidou, T.-S. & Alvanoudi, A. (2018). Conceptualizing the world as "female" or "male": Further remarks on grammatical gender and speakers' cognition. In N. Topintzi, N. Lavidas, & M. Moumtzi (Eds.), *Selected papers on theoretical and applied linguistics from ISTAL23*. Thessaloniki, Greece: School of English, Aristotle University of Thessaloniki.

Phillips, M. P. (2012). *Dialect continuum in the Bhil tribal belt grammatical aspects* [PhD Thesis]. University of London, London, England.

Phillips, W. & Boroditsky, L. (2003, August). *Can quirks of grammar affect the way you think? Grammatical gender and object concepts*. Paper presented at the 25th annual meeting of the Cognitive Science Society, Boston, MA.

Press, I. (1986). *A grammar of modern Breton*. *Mouton grammar library: Vol. 2*. Mouton de Gruyter

Pylyshyn, Z. W. (1984). *Computation and cognition*. MIT Press.

Rahimian, G. (1978). *Spoken Bakhtiari.* SIT Study abroad, SIT Graduate Institute, Brattleboro, Vermont.

Rajapurohit, B. B. (2012). *Grammar of Shina language and vocabulary: (Based on the dialect spoken around Dras).* Central Institute of Indian Languages, Mysore, India.

Ramos, S. & Roberson, D. (2011). What constrains grammatical gender effects on semantic judgements? Evidence from Portuguese. *Journal of Cognitive Psychology, 23*(1), 102–111.

Rani Das, A. (October 2017). *A comparative study of Bangla and Sylheti grammar* [PhD Dissertation]. University of Naples Federico II, Napoli.

Rojavin, M. (2010,). The semantic category of gender in Russian and Ukrainian. *The Slavic and East European Journal, 54*(3), 503–526.

Saalbach, H., Imai, M., & Schalk, L. (2012). Grammatical gender and inferences about biological properties in German-speaking children. *Cognitive Science, 36*, 1251–1267.

Sabourin, L., Stowe, L. A., & de Haan, G. J. (2006). Transfer effects in learning a second language grammatical gender system. *Second Language Research*, *22*(1), 1–29. Retrieved (03/2023) from https://doi.org/10.1191/0267658306 sr259oa

Saksenā, B. (1971). *Evolution of Awadhi: (a branch of Hindi)* (2nd ed.). Motilal Banarsidass.

Samuel, S., Cole, G., & Eacott, M. J. (2019). Grammatical gender and linguistic relativity: A systematic review. *Psychonomic Bulletin & Review*, *26*(6), 1767–1786. Retrieved (03/2023) from https://doi.org/10.3758/s13423-019-01652-3

Sande, H. (2013). Nouchi as a distinct language: The morphological evidence. *Annual Conference on African Linguistics* (44), 7–10.

Sato, S. & Athanasopoulos, P. (2018). Grammatical gender affects gender perception: Evidence for the structural-feedback hypothesis. *Cognition*, *176*, 220–231.

Sedlmeier, P., Tipandjan, A., & Jänchen, A. (2016). How persistent are grammatical gender effects? The case of German and Tamil. *Journal of Psycholinguistic Research*, *45*, 317–336.

Semenuks, A., Phillips, W., Dalca, I., Kim, C., & Boroditsky, L. (2017). Effects of grammatical gender on object description. In G. Gunzelmann, A. Howes, T. Tenbrink, & E. Davelaar (Eds.), *Proceedings of the 39th annual meeting of the Cognitive Science Society* (pp. 1060–1065). Austin, TX: Cognitive Science Society.

Sera, M. D., Berge, C. A., & Del Pintado, J. C. (1994). Grammatical and conceptual forces in the attribution of gender by English and Spanish speakers. *Cognitive Development*, *9*(3), 261–292. Retrieved (03/2023) from https://doi.org/10.1016/0885-2014(94)90007-8

Sera, M. D., Elieff, C., Forbes, J., Burch, M. C., Rodríguez, W., & Dubois, D. P. (2002). When language affects cognition and when it does not: An analysis of grammatical gender and classification. *Journal of Experimental Psychology: General*, *131*, 377–397.

Siddique, K. (2006, June). *Memoni language project.* The Truth (As-Sidq). Retrieved (03/2023) from http://www.as-sidq.org/memoni/

Slobin, D. I. (1996). From "thought and language" to "thinking for speaking". In D. I. Slobin (Ed.), *Rethinking linguistic relativity* (pp. 70–96). Cambridge University Press, Cambridge, United Kingdom.

Slobin, D. I. (Ed.). (1996). *Rethinking linguistic relativity*. Cambridge University Press, Cambridge, United Kingdom.

Smith, I. R. (1977). *Sri Lanke Creole Portuguese phonology*. Cornell University.

Southworth, S. C. (1970). Review: A grammar of old Marathi by Alfred Master. *Linguistic Society of America, 46*(2, Part 1), 507–513.

Stemberger, J. P. (2006). Proper names in Slovene: Implications for defaults in inflectional morphology. *Proceedings of the Annual Meeting of the Cognitive Science Society, 28* (8), 2186–2191.

Taherkhani, N. & Ourang, M. (2013). A study of derivational morphemes in Lari & Tati as two endangered Iranian languages: An analytical-contrastive examination with Persian. *Journal of American Science, 9*(11s), 38–45.

Toker, H. (2014). *A practical guide to Hindko grammar*. Trafford Publishing, Mounton & Co.B.V, Hungary.

Trofimov, A. (2019). On the place of Parachi and Ormuri among the Iranian languages according to the data of annotated Swadesh lists. *Journal of Language Relationship, 16*(3–4), 277–292.

Trumpp, E. (1872). *Grammar of the Sindhi language: Compared with the Sanskrit-Prakrit and the cognate Indian vernaculars*. Trübner.

Truppi, C. (2019). Copulas in contact: Kriyol, Upper Guinea Creoles, and their substrate. *Journal of Ibero-Romance Creoles, 9*(1), 85–114.

Tucker, G. R., Lambert, W. E., & Rigault, A. A. (2019). The French speaker's skill with grammatical gender. *The French Speaker's Skill with Grammatical Gender*. Mounton & Co.B.V, Hungary.

Turner, R. L. (1932). The language of the Ashkun Kafirs. By G. Morgenstierne.: Extract from Norsk Tidsshrift for Sprogvidenskap, Bindii, 1929. pp. 192–289. *Journal of the Royal Asiatic Society, 6a4*(1), 173–175.

Uddin, N. & Uddin, J. (2019). A step towards Torwali machine translation: an analysis of morphosyntactic challenges in a low-resource language. *Proceedings of the 2nd Workshop on Technologies for MT of Low Resource Languages*, 6–10.

Vahman, F., Vahman, F., & Asatrian, G. (2002). *Notes on the language and ethnography of the Zoroastrians of Yazd* (Vol. 85). Kgl. Danske Videnskabernes Selskab.

van der Wal Anonby, C. A. (2015). *A grammar of Kumzari: A mixed Perso-Arabian language of Oman.* Leiden University, Doctoral dissertation.

Vernich, L. (2017). Does learning a foreign language affect object categorization in native speakers of a language with grammatical gender? The case of Lithuanian speakers learning three languages with different types of gender systems (Italian, Russian and German). *International Journal of Bilingualism*, *23*, 417–436.

Vernich, L., Argus, R., & Kamandulytė-Merfeldienė, L. (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*, *13*, 223–240.

Vigliocco, G., Vinson, D. P., Paganelli, F., & Dworzynski, K. (2005). Grammatical gender effects on cognition: Implications for language learning and language use. *Journal of Experimental Psychology. General*, *134*(4), 501–520. Retrieved (03/2023) from https://doi.org/10.1037/0096-3445.134.4.501

Villeneuve, A. J. (2011). *A sociolinguistic study of Vimeu French* [PhD Dissertation]. Indiana University, Indiana.

Visser, W. (2011). Historical gender change in West Frisian. *Morphology*, *21*(1), 31–56.

Wam, C. (2019). *Grammatical gender in Afrikaans–simplification in action.* North-West University, Potchefstroom South Africa. Retrieved (03/2023) from https://www.ds.uzh.ch/dam/jcr:499d7d96-d307-410f-a924-ca927ecc3d6f/afrikaans_wanniecarstens.pdf

Wheeler, M., Yates, A., & Dols, N. (2002). *Catalan: A comprehensive grammar.* Routledge, England.

Whorf, B. (1956). *Language thought and reality*. MIT Press.

Whorf, B. L. (Ed.). (2012). *Language, thought, and reality, second edition: Selected writings of Benjamin Lee Whorf* (2nd ed.). The MIT Press.

Williams, S. J. (1986). *A Welsh grammar* (Repr). University of Wales Press.

Windfuhr, G. (2009). *The Iranian languages* (1st ed.). *Routledge language family Sseries*. Routledge, England.

Yadav, S. P. (July 2016). *A sociolinguistic survey of Chitoniya.* Linguistic Survey of Nepal, Department of linguisitcs, Tribhuvan University, Kathmandu, Nepal.

Appendix A.1

Name of the main branch	Name of L1 spoken languages	Number of grammatical gender
Albanian	Gheg and Tosk	Three genders: masculine, feminine, rare neutral
	Arvanitika	
	Arbëresh	
Armenic	Standard Armenic	No grammatical Gender
	East Armenic	
Baltic	Latvian/Lettish	Two genders: masculine and feminine
	Lithuanian	
Slavonic	Bulgarian	Three genders: masculine, feminine, rare neutral
	Macedonian	
	Serbo-Croat	
	Slovene	
	Czech	
	Slovak	
	Polish	
	Upper and lower Sorbian (Lusatian)	
	Russian	
	Ukrainian	
	Belorussian	

Name of the main branch	Name of L1 spoken languages	Number of grammatical gender
Celtic	Breton	Two genders: masculine and feminine
	Cornish	
	Welsh	

Name of the main branch	Name of L1 spoken languages	Number of grammatical gender
	Scotts Gaelic	
	Irish	
Northern Germanic	Norwegian	Three genders: masc., fem., neut.
	Modern Icelandic	Three genders: masc., fem., neut.
	Modern Danish	Two genders: common and neuter
	Modern Swedish	Two genders: common and neuter
Western Germanic	Modern German	Three genders: masc., fem., neut.
	Yiddish	Three genders: masc., fem., neut.
	Modern Dutch	Two genders: common and neuter
	Afrikaans	No grammatical Gender
	English	Pronominal gender
	Scots	Pronominal gender
	Frisian	Two genders: common (non-neuter) and neuter
Graeco-Phrygian	Apulia-Calabrian Greek/ Griko	No grammatical gender
	Cypriot Greek	No information found
	Romano-Greek	Three genders: masc., fem., neut.
	Cappadocian Greek	One gender: neutral
	Pontic	Three genders: masc., fem., neut.
	Modern Greek	Three genders: masc., fem., neut.
	Taskonian	Three genders: masc., fem., neut.

Appendix A.2

Name of the main branch	Name of L1 spoken languages	Number of grammatical gender
Indo-Aryan	Angika	No information found
	Bhojpuri	One gender: fem.
	Caribbean Hindustani	No information found
	Magahi	Two genders: masc., fem.
	Maithili	Two genders: masc., fem.
	Kudmali/Kurmali	Two genders: masc., fem.
	Oraon Sadri	No grammatical gender
	Sadri	No grammatical gender
	Panchpargania	No information found
	Pali	Three genders: masc., fem., neut
	Chitwania Tharu/ Chitoniya Tharu	No grammatical gender
	Dangaura Tharu	No information found
	Sonha	No information found
	Kathoriya Tharu	No information found
	Kochila Tharu	No information found
	Rana Tharu	Two genders: masc., fem.
	Buksa	No information found
	Majhi	No information found
	Musasa	No information found
	Kumhali	No information found
	Bote	No information found
	Darai	Two genders: masc., fem.
	Dewas-Done Danuwar	No information found
	Kochariya-East Danuwar	No information found
	Dhivehi	No grammatical gender
	Sinhala	Two genders: masc., fem.
	Veddah	No information found
	Domaaki	Two genders: masc., fem.
	Domari	Two genders: masc., fem.
	Archaic Angloromani	No information found
	Welsh Romani	No information found
	Kalo Finnish Romani	No information found
	Sinte-Manus Romani	No information found
	Balkan Romani	No information found

Name of the main branch	Name of L1 spoken languages	Number of grammatical gender
	Baltic Romani	No information found
	Caló	No information found
	Central Romani	Two genders: masc., fem.
	Vlax Romani	No information found
	Sauraseni Prakrit	No information found
	Bauria	No information found
	Bhili	No information found
	Chodri	No information found

Name of the main branch	Name of L1 spoken languages	Number of grammatical gender
Indo-Aryan	Dhodia-Kukna	No information found
	Dungra Bhil	No information found
	Gamit	No information found
	Adiwasi Garasia	No information found
	Rajput Garasia	No information found
	Malvi	No information found
	Mawchi	No information found
	Nimadi	No information found
	Pardhi	No information found
	Bhilali	No information found
	Nahali-Baglani	No information found
	Pauri Bareli	No information found
	Rathwi Bareli	No information found
	Palya Bareli	No information found
	Rathawi	No information found
	Vaagri Booli	No information found
	Noiri	No information found
	Vasavi	No information found
	Wagdi	Three genders: masc., fem., neut.
	Awadhi	Two genders: masc., fem.
	Bagheli	No information found
	Chhattisgarhi	Two genders: masc., fem.
	Powari	No information found
	Dubli	No information found
	Gujarati	Three genders: masc., fem., neut.
	Saurashtra	No information found
	Aer	No information found
	Kachi Koli	No information found

Name of the main branch	Name of L1 spoken languages	Number of grammatical gender
	Vaghri	No information found
	Wadiyara Koli	No information found
	Bagri	Two genders: masc., fem.
	Jandavra	No information found
	Dhatki	No information found
	Dhundari	No information found
	Hadothi	No information found
	Kabutra	No information found
	Kanjari	No information found
	Lambadi	No grammatical gender
	Merwari	No information found
	Gade Lohar	No information found
	Mewari	Two genders: masc., fem.
	Godwari	No information found
	Gujari	No information found
	Marwari	No information found

Name of the main branch	Name of L1 spoken languages	Number of grammatical gender
Indo-Aryan	Mewati	Two genders: masc., fem.
	Od	No information found
	Parya	No information found
	Shekhawati	No information found
	Bhaya	No information found
	Goaria	No information found
	Parkari Koli	No information found
	Dhanki	No information found
	Khandesi	No information found
	Middle Indo-Aryan	No information found
	Braj	Two genders: masc., fem.
	Bharia/Bhariati	Two genders: masc., fem.
	Bundeli	Two genders: masc., fem.
	Haryanvi	No information found
	Andaman Creole Hindi	No information found
	Fiji Hindi	No grammatical gender
	Hindi	Two genders: masc., fem.
	Urdu	Two genders: masc., fem.
	Kanauji	No information found
	Ghera	No information found

Name of the main branch	Name of L1 spoken languages	Number of grammatical gender
	Gurgula	No information found
	Bhatri	No information found
	Bhunjia	No information found
	Halbi	No information found
	Kamar	No information found
	Nahari	No information found
	Bengali	No grammatical gender
	Bishnupriya Manipuri	No information found
	Hajong	No information found
	Sylheti	Two genders: masc., fem.
	Kharia Thar	No information found
	Lodhi	No information found
	Mar Paharia of Dumka	No information found
	Rohingya	No information found
	Chakma	No information found
	Chittagonian	No information found
	Tangchangya	No information found
	Kurmukar	No information found
	Assamese	No grammatical gender
	Naga Pidgin	No information found
	Central-Eastern Kamta	No information found
	Rajbanshi	No information found
	Surjapuri	Two genders: masc., fem.
	Bodo Parja	No information found

Name of the main branch	Name of L1 spoken languages	Number of grammatical gender
Indo-Aryan	Kotia-Adivasi Oriya-Desiya	No information found
	Kupia	No information found
	Odia	No grammatical gender
	Reli	No information found
	Sambalpuri	No grammatical gender
	Garhwali	No information found
	Kumaoni	No information found
	Dotyali	No information found
	Jumli	No information found
	Nepali	Two genders: masc., fem.
	Jaunsari	No information found
	Bhadrawahi	Two genders: masc., fem.

Name of the main branch	Name of L1 spoken languages	Number of grammatical gender
	Churahi	No information found
	Bhattiyali	No information found
	Bilaspuri	No information found
	Chambeali	No information found
	Gaddi	No information found
	Pangwali	Three genders: masc., fem., neut
	Dogri	No information found
	Kangri	No information found
	Mandeali	No information found
	Hinduri	No information found
	Indo-Aryan Kinnauri	No information found
	Kullu Pahari	No information found
	Mahasu Pahari	No information found
	Chitral Kalasha	No information found
	Khowar	No information found
	Kashmiri	Two genders: masc., fem.
	Khah	No information found
	Gawri	Two genders: masc., fem.
	Torwali	Two genders: masc., fem.
	Indus Kohistani	No information found
	Bateri	No information found
	Mankiyali	No information found
	Chilisso	No information found
	Gowro	No information found
	Dameli	Two genders: masc., fem.
	Gawar-Bati	No information found
	Grangali	No information found
	Shumashti	No information found
	Northeast Pashayi	No information found
	Southeast Pashayi	No information found
	Northwest Pashayi	No information found
	Southwest Pashayi	No information found

Name of the main branch	Name of L1 spoken languages	Number of grammatical gender
Indo-Aryan	Brokskat	No information found
	Kohistani Shina	Two genders: masc., fem.
	Ushojo	No information found
	Kundal Shahi	Two genders: masc., fem.

Name of the main branch	Name of L1 spoken languages	Number of grammatical gender
	Shina	Two genders: masc., fem.
	Kalkoti	No information found
	Palula	Two genders: masc., fem.
	Savi	No information found
	Eastern Panjabi	No information found
	Sansi/Sansiboli	Two genders: masc., fem.
	Northern Hindko	Two genders: masc., fem.
	Southern Hindko	Two genders: masc., fem.
	Pahari Potwari	No information found
	Sirajic	No information found
	Saraiki	Two genders: masc., fem.
	Western Panjabi	No information found
	Kachchi	No information found
	Khetrani	No information found
	Jadgali	No information found
	Lasi	No information found
	Luwati	No information found
	Memoni	Two genders: masc., fem.
	Sindhi	Two genders: masc., fem.
	Sindhi Bhil	No information found
	Kholosi	Two genders: masc., fem.
	Katkari	No information found
	Goan Konkani/Konkani	Three genders: masc., fem., neut.
	Dungar Varli	No information found
	Marathi	Three genders: masc., fem., neut.
	Samvedi	No information found
	Varhadi-Nagpuri	No information found
	Konkan Marathi	No information found
	Phudagi	No information found
	Old Marathi	Three genders: masc., fem., neut.
	Chinali	No information found
	Lahul Lohar	No information found

Appendix A.3

Name of the main branch	Name of L1 spoken languages	Number of grammatical gender
Iranian	Gurani	No grammatical gender
	Bajelnai	No grammatical gender
	Shabaki	No grammatical gender
	Alamuti	No grammatical gender
	Kajali	No grammatical gender
	Karanic	No information found
	Shahrudi-Southern Talysh	No grammatical gender
	Kho´ni	moribund/pronominal gender (two genders: masculine, feminine)
	Maraghei	One gender, feminie
	North-Central Talysh	No grammatical gender
	Kabatei	Two genders: masc., fem.,
	Upper Taromi	No information found
	Harzani-Kilit	No information found
	Karingani-Kalasuri-Khoynarudi	No information found
	Alviri-Vidari	moribund/No grammatical gender
	Eshtehardi	No information found
	Rezajedri	No information found
	Takestani	No information found
	Ashtiani	No information found
	Vafsi	Two genders: masc., fem.
	Dimli	Two genders: masc., fem.
	Kirmanjki/Kurmanji	Two genders: masc., fem.
	Eastern-Baluchi	No grammatical gender
	Koroshi	No grammatical gender
	southern Baluchi	No grammatical gender
	Western Baluchi	No grammatical gender
	Gilaki	No grammatical gender
	Rudbari	No information found
	Mazandarani	No grammatical gender
	Shahmirzadi	No grammatical gender
	Kavir	No information found

Name of the main branch	Name of L1 spoken languages	Number of grammatical gender
	Gazic	No grammatical gender
	Judeo-Hamadani-Borujerdi	No grammatical gender
	Natanzic	No information found
	Soic	No information found
	Khunsaric	No information found
	Nayinic	No grammatical gender
	Zoroastrian Yazdi	No grammatical gender
	Sivandi	No information found
	Lasgerdi	No information found

Name of the main branch	Name of L1 spoken languages	Number of grammatical gender
Iranian	Sangisari	No information found
	Sorkhei-Aftari	No grammatical gender
	Central Kurdish	No grammatical gender
	Northern Kurdish	Two genders: masc., fem.
	Southern Kurdish	No grammatical gender
	Laki	No grammatical gender
	Semnani-Biyabuneki	No information found
	Digor Ossetian	No grammatical gender
	Iron Ossetian	No grammatical gender
	Yagnobi	No grammatical gender
	Ormuri	No grammatical gender
	Parachi	No grammatical gender
	Central Pashto	Two genders: masc., fem.
	Northern Pashto	Two genders: masc., fem.
	Southern Pashto	Two genders: masc., fem.
	Waneci	No information found
	Ishkashimi	No information found
	Sanglechi	No information found
	Sarikoli	No grammatical gender
	Shughni	No information found
	Yazgulyam	pronominal gender
	Bashkardi	No grammatical gender
	Fars Dialects	No grammatical gender
	Judeo-Tat	No grammatical gender

Name of the main branch	Name of L1 spoken languages	Number of grammatical gender
	Muslim Tat	No grammatical gender
	Aimaq	No information found
	Dari	No grammatical gender
	Dehwari	No information found
	Hazaragi	No grammatical gender
	Bukharic	No information found
	Tajik	No grammatical gender
	Judeo-Persian	No grammatical gender
	Western Farsi	No grammatical gender
	Kumzari	No grammatical gender
	Larestani	No grammatical gender
	Dezfuli-Shushtari	No grammatical gender/nearly extinct
	Bakhtiari	No grammatical gender
	Southern Luri	No grammatical gender
	Northern Luri	No grammatical gender
	Wakhi	No information found
	Munji	No information found
	Yidgha	No information found
	Ashkun	Two genders: masc., fem.

Name of the main branch	Name of L1 spoken languages	Number of grammatical gender
Iranian	Katë	Two genders: masc., fem.
	Nuristani Kalasha	Two genders: masc., fem.
	Tregami	Two genders: masc., fem.
	Prasun	Two genders: masc., fem.

Appendix A.4

Name of the main branch	Name of L1 spoken languages	Number of grammatical gender
Italic	Aromanian	Three genders: masc., fem., neut.
	Megleno Romanian	Three genders: masc., fem., neut.
	Romanian	Three genders: masc., fem., neut.
	Istro Romanian	Three genders: masc., fem., neut.
	Istriot	Two genders: masc., fem.
	Continental Southern Italian	Two genders: masc., fem.
	Italian	Two genders: masc., fem.
	Sicilian	Two genders: masc., fem.
	Emiliano	No information found
	Romagnol	No information found
	Judeo-Italian	Two genders: masc., fem.
	Ligurian	Two genders: masc., fem.
	Lombard	Two genders: masc., fem.
	Piemontese	Two genders: masc., fem.
	Sassarese Sardinian	Two genders: masc., fem.
	Venetian	Two genders: masc., fem.
	Friulian	Two genders: masc., fem.
	Ladin	Two genders: masc., fem.
	Cajun French	Two genders: masc., fem.
	Gulf Coast French	No information found
	Guianese Creole French	No information found
	Uaçá Creole French	No information found
	Haitian	Two genders: masc., fem.
	Guadeloupe-Martinique Creole French	Two genders: masc., fem.
	Saint Lucian Creole French	Two genders: masc., fem.
	Louisiana Creole French	Two genders: masc., fem.
	French	Two genders: masc., fem.
	Nouchi	Two genders: masc., fem.
	Morisyen	No information found
	Seselwa Creole French	No information found
	Réunion Creole French	No information found
	Tayo	No information found
	Picard	Two genders: masc., fem.,

Name of the main branch	Name of L1 spoken languages	Number of grammatical gender
	Arpitan	No information found
	Faeto and Celle San Vito Francoprovencal	No information found
	Walloon	Two genders: masc., fem.
	Romansh	Two genders: masc., fem.
	Catalan	Two genders: masc., fem.
	Gardiol	No information found
	Occitan	Two genders: masc., fem.

Name of the main branch	Name of L1 spoken languages	Number of grammatical gender
Italic	Old Provençal	Two genders: masc., fem.
	Asturian-Leonese-Cantabrian	No information found
	Mirandese	Two genders: masc., fem.
	Extremaduran	No information found
	Ladino	No information found
	Palenquero	No information found
	Peruvian Amazonian Spanish	No information found
	Chavacano	Two genders: masc., fem.
	Spanish	Two genders: masc., fem., three genders in pronouns masc., fem., neut.
	Fala	No information found
	Galician	No information found
	Cafundo	No information found
	Portuguese	Two genders: masc., fem.
	Malabar-Sri Lanka Portuguese	Two genders: masc., fem.
	Daman-Diu Portuguese	No information found
	Korlai Portuguese	No information found
	Angolar	No information found
	Annobonese	No information found
	Principense	No information found
	Sãotomense	Two genders: masc., fem.
	Macanese	No information found

Name of the main branch	Name of L1 spoken languages	Number of grammatical gender
	Malacca-Batavia Portuguese Creole	No grammatical gender
	Kabuverdianu	Two genders: masc., fem.
	Papiamento	No information found
	Upper Guinea Crioulo	No grammatical gender
	Minderico	No information found
	Aragonese	Three genders: masc., fem., neut
	Corsican	No information found
	Gallurese Sardinian	No information found
	Campidanese Sardinian	Two genders: masc., fem.
	Logudorese Sardinian	No information found
	South Lucanian	No information found

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-

Open Access © 2024 Autor*innen. Dieses Werk ist bei der Verlag Barbara Budrich GmbH erschienen und steht unter der Creative Commons Lizenz Attribution 4.0 International (CC BY 4.0).

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.

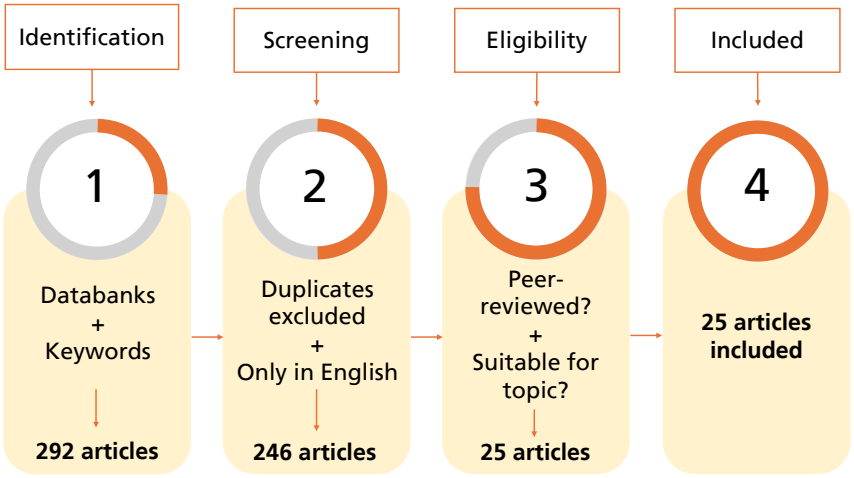
To demonstrate the flawed methodologies, the current study looks critically at 1) the examined languages linguistically in terms of the underlying genus system and

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



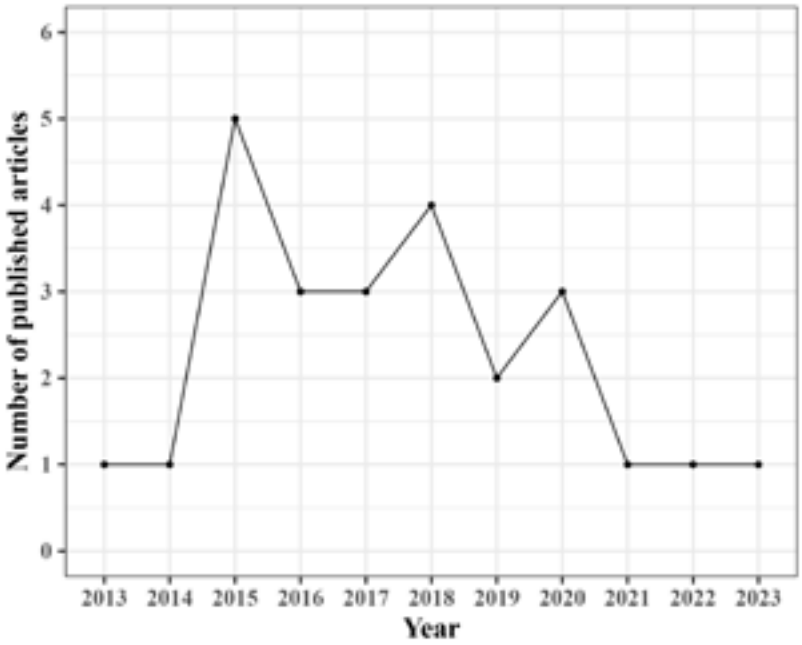
Source: Adapted from and inspired by Newmann et al. (2019).

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.

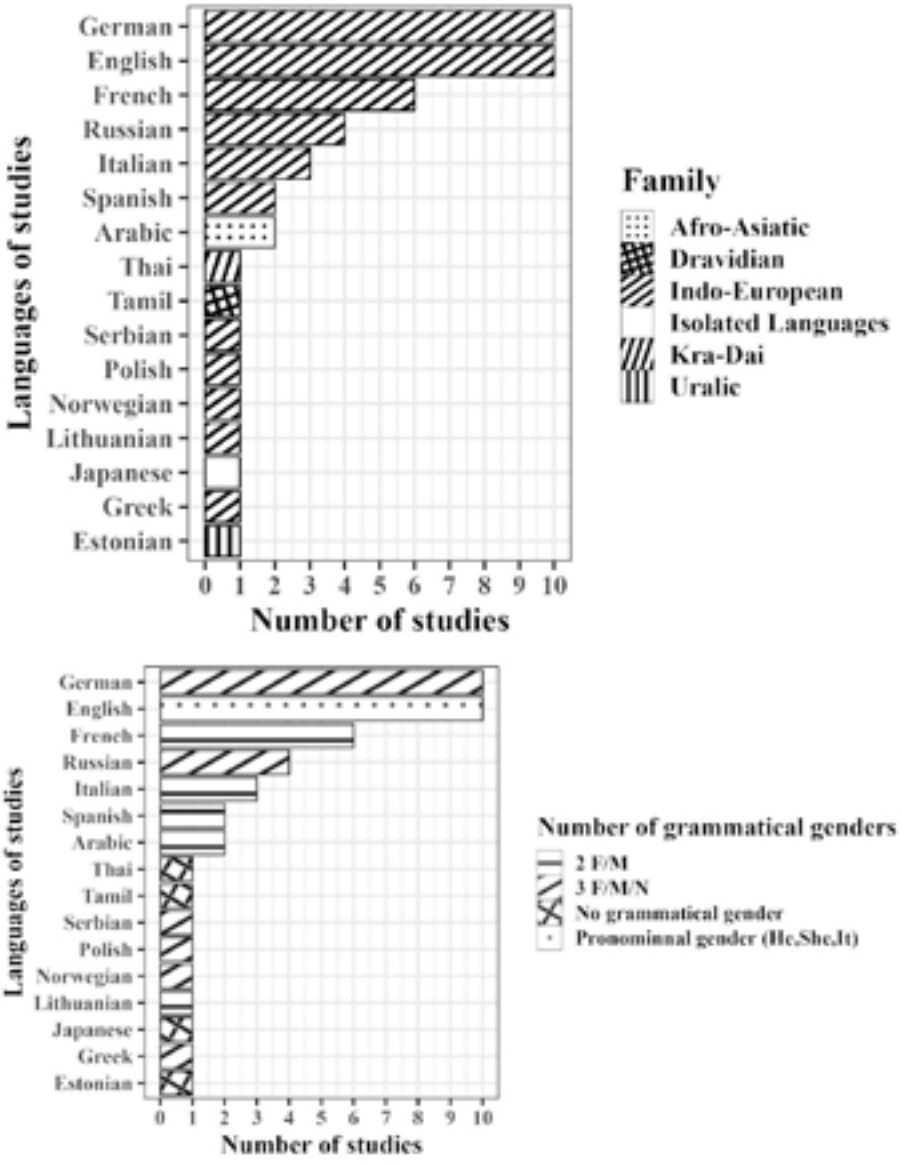
There was a notable increase in published articles in 2015, a decline in 2016–2017, and a resurgence in 2018. This pattern from 2015 to 2018 is also seen from 2018 to 2020, though with fewer publications. Overall, a downward trend from 2015 to 2023 indicates limitations in current methodologies for demonstrating 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).

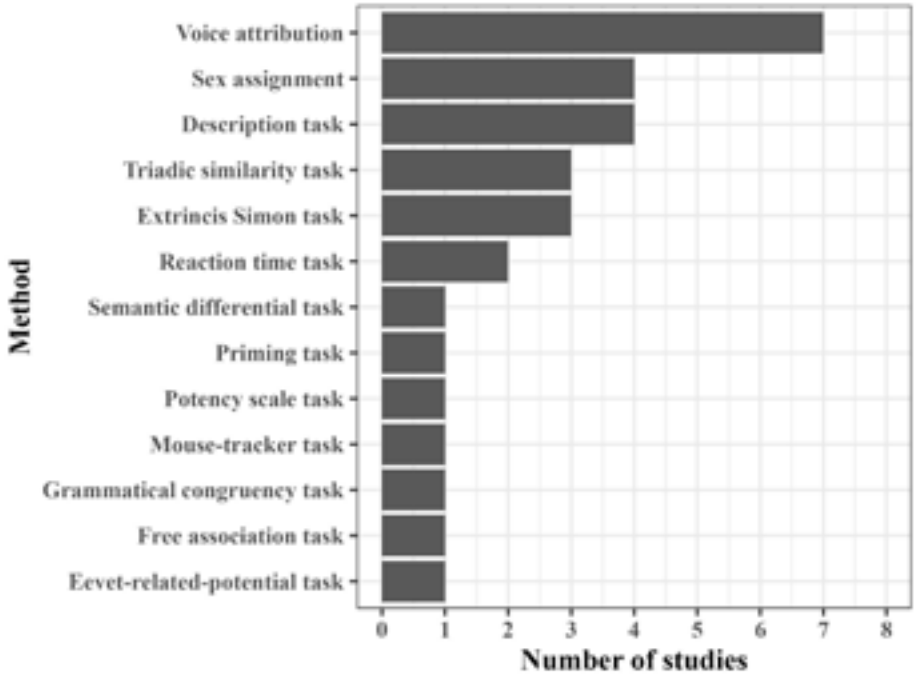
Vernich/Argus/Kamandulyté-Merfeldiené (2017) opted for Estonian as a less-studied language within the relativity framework, comparing it to Lithuanian and Italian. Moreover, Incera et al. (2018) examined how genus in Spanish influences conceptual

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.

Fatemi, Maryam, & Mertins, Barbara (2026). Language-specific patterns in multilingual contexts: The case of grammatical gender. In B. Mertins & K. Šormová (Eds.), *New empirically based perspectives on multilingualis* (pp. 119·148). Peter Lang.

Research Article

Maryam Fatemi*, Barbara Mertins and Vincent Schroeder

The influence of grammatical gender on cognition: the case of German and Farsi

https://doi.org/10.1515/opli-2025-0071
Received February 13, 2025; accepted September 29, 2025; published online November 20, 2025

Abstract: This study investigates whether and how grammatical gender influences cognition by comparing speakers of German, a language with three grammatical genders (masculine, feminine, neutral), and Farsi, a genderless language. Using object description tasks, we examined how grammatical gender affects the assignment of descriptive properties to objects. German monolinguals demonstrated significant effects of grammatical gender on object categorization. This effect was particularly evident in the first adjectives produced, indicating an immediate and automatic link between grammatical gender and conceptual representation. In contrast, Farsi monolinguals showed no such effects, consistent with Farsi's lack of grammatical gender, and their adjective choices were unaffected by gender category, adjective order, sex, or age. Notably, German speakers perceived objects with neutral grammatical gender as genuinely neutral, underscoring the cognitive reality of the neuter category. These findings highlight how the presence or absence of grammatical gender in a language shapes cognitive processing and support linguistic relativity. The study also emphasizes the importance of including neutral gender and of cross-linguistic comparisons in future research to deepen our understanding of language–thought interactions.

Keywords: grammatical gender; cognition; conceptual representation; linguistic relativity

1 Introduction

In the realm of linguistic diversity, some grammatical categories vary significantly across languages (Haspelmath et al. 2005). This variation is observed even across languages within the same language family, such as the Indo-European family, highlighting the nuanced structural evolution of grammar caused by historical and functional adaptations (Comrie 1989). Thus, different languages may vary in case marking (extensive case systems in Russian versus limited or no cases in Romance languages) or in verb conjugation patterns (complex aspectual marking in Slavic languages versus simpler tense systems in Germanic languages) (Comrie 1989). This study focuses on one of these grammatical categories: grammatical gender, present in some languages, yet notably absent in others (Corbett 1991).

Understanding these variations is crucial as they provide insights into how language shapes cognition (Boroditsky 2011), influences social interactions (Gumperz 1982), and reflects cultural values (Duranti 2011). Grammatical gender, in particular, can affect how speakers perceive and categorize the world, with implications for identity, communication, and gender biases (Boroditsky et al. 2003). Examining these variations enhances our understanding of linguistic diversity and its profound influence on thought and society.

***Corresponding author: Maryam Fatemi**, Technical University of Dortmund, Dortmund, Germany,
E-mail: seyedehmaryam.fatemi@tu-dortmund.de. https://orcid.org/0009-0009-5686-4186
Barbara Mertins and Vincent Schroeder, Technical University of Dortmund, Dortmund, Germany, E-mail: barbara.mertins@tu-dortmund.de (B. Mertins), vincent.schroeder@tu-dortmund.de (V. Schroeder). https://orcid.org/0000-0002-0271-0431 (B. Mertins). https://orcid.org/0009-0009-1578-4155 (V. Schroeder)

Open Access. © 2025 the author(s), published by De Gruyter. This work is licensed under the Creative Commons Attribution 4.0 International License.

The purpose of the present study is to explore whether grammatical gender influences cognition[1] at the level of conceptualization.[2] Previous studies have demonstrated differences in how speakers of languages with grammatical gender perceive and categorize objects compared to speakers of languages without grammatical gender (Sedlmeier et al. 2016; Sera et al. 2002). This study further investigates these differences in categorization. Specifically, it examines whether grammatical gender leaves traces on how speakers conceptualize objects. Do German and Farsi speakers have distinct associations regarding gender attributes assigned, for example, to the noun 'cloud'? Do German speakers categorize 'cloud' as feminine ('die Wolke'), while Farsi speakers attribute no gender to 'cloud' ('ابر' /ʔaˈbaɻ/)? In contrast, a speaker of French would perceive 'cloud' as masculine ('le nuage'), highlighting the differences among these languages: 'die Wolke' is feminine in German, 'le nuage' is masculine in French, and 'ابر' /ʔaˈbaɻ/ is genderless in Farsi.

To investigate the link between grammatical gender and cognition, this study turns to *the principle of linguistic relativity*, often referred to as *the Sapir-Whorf hypothesis*. This hypothesis posits that the grammatical features inherent in the languages we speak influence our conceptualization and worldview (Lucy 2016; Whorf 1956). Whorf (1956), followed by other scholars, argued that languages impose external impressions through their grammar (Gentner and Goldin 2003; Gumperz and Levinson 1991; Lucy 1992; Whorf 1956). This study explores how the grammatical gender of nouns may contribute to these categorical differences, particularly in the context of German and Farsi.

Grammatical gender, as one of the obligatory properties of nouns in some languages, was researched within the relativity framework as early as the 1990s (Konishi 1993; Sera et al. 1994). Yet, these studies yielded inconsistent outcomes. Some of them found an impact of grammatical gender on conceptualization – particularly in how objects are categorized and perceived – while others found no such influence, suggesting that cognitive effects may vary depending on context or methodology (Flaherty 2001; Imai et al. 2014; Landor 2014; Mickan et al. 2014; Montefinese et al. 2019).

The inconsistency in results is not tied solely to task variations and methodologies. There is a contentious debate within the psycholinguistic and sociolinguistic literature about associating grammatical gender with masculinity and femininity due to the perceived associations with biological gender or sex (Bender et al. 2011; Samuel et al. 2019). Yet, in many languages, the gender distinction does not coincide with a sex distinction. For instance, in grammatically gendered languages such as Dutch and Swedish, a *common gender* includes what other languages classify as *masculine* or *feminine* (Bender et al. 2011). The complexity of grammatical gender is particularly pronounced in German, where three genders must be assigned not only to animate but also to inanimate objects. For example, nouns such as *der Tisch* (the table, masculine), *die Lampe* (the lamp, feminine), and *das Buch* (the book, neuter) illustrate this intricacy. Consequently, much of the existing literature often fails to incorporate these non-binary grammatical gender forms in their experimental tasks (Bender et al. 2016a, 2016b; Imai et al. 2014). Instead, researchers frequently simplify the process by prompting participants to make binary gender judgments (male/female), thus overlooking the nuanced gender system of such languages such as German. To remedy this situation, researchers should revise the tasks employed to investigate the influence of grammatical gender on cognition, ensuring experiments encompass the full range of grammatical categories. Such an approach would facilitate a more comprehensive understanding of how these diverse grammatical forms impact cognition rather than limiting the analysis to binary gender judgments.

To enhance inclusivity in grammatical gender research, this study adopts Whorf's (1956) and Lucy's (2016) argument to investigate the influence of grammatical gender on conceptualization in native German speakers and compares the perspectives of German speakers' responses with those of Farsi speakers. Unlike three-folded gendered German, Farsi, another Indo-European language, is genderless. Thus, this study explores whether native speakers of Farsi and German conceptualize the world differently based on their language. This research

1 The term "cognition" encompasses a range of mental processes involved in acquiring knowledge and understanding through thought, experience, and senses. These processes take place at different levels, including perception, attention, memory, reasoning, and conceptualization, which involves forming and organizing concepts based on experiences (Langacker 1987).

2 This process is influenced by linguistic and cultural factors, shaping how individuals categorize and interpret their experiences (Boroditsky 2001).

trajectory allows for further scrutiny of the linguistic relativity hypothesis. The study seeks to answer the following research questions:

(1) Does the grammatical gender of the nouns influence the assignment of properties to objects and animals by German native speakers?
(2) Do Farsi native speakers differ from German native speakers in assigning properties to objects and animals?

To address these research questions, this paper is structured as follows. Section 2 provides the theoretical framework, beginning with an overview of linguistic relativity and its spectrum (2.1), followed by a discussion of grammatical gender as an obligatory feature in certain languages (2.2). Section 2.3 examines the gender systems of German and Farsi, while Section 2.4 reviews previous research on the influence of grammatical gender on cognition within the linguistic relativity framework. Section 3 outlines the methodology of the present study, detailing the experimental design and analytical approach. Section 4 presents the results, and Section 5 discusses their implications in relation to the research questions, concluding with a synthesis of the study's contributions and future research directions.

2 Theoretical framework

2.1 Linguistic relativity and its spectrum

Linguistic relativity, also known as the Sapir-Whorf hypothesis, is typically discussed in terms of strong and weak versions. The strong version holds that language determines thought and restricts cognitive abilities, while the weak version suggests that language influences thought and perception without fully constraining them (Whorf 1956). Few contemporary scholars endorse the strong determinist position (Athanasopoulos 2009).

At the opposite end of this spectrum lies the universalist perspective, which posits that thought is independent of language (Pinker 1994). Between these two extremes lie various intermediate positions. One such view is the *thinking-for-speaking* hypothesis, which argues that language guides our attention toward perceptual features that are more easily described linguistically, thereby shaping thought as we engage in speaking (Boroditsky 2011; Slobin 1996, 91). Mertins (2018) extended this idea with the *seeing-for-speaking* hypothesis, which emphasizes that visual perception precedes and is influenced by the grammatical encoding required by a given language.

The extent to which language shapes perception and cognition remains controversial. Many scholars agree that our perceptions and conceptualizations are influenced or shaped by language (Athanasopoulos 2006; Dolscheid et al. 2013; Lucy 2016). According to Lucy (2016, 498), it is now widely acknowledged that language-specific patterns can impact cognition, and the burden of proof lies with those who deny these effects. However, as Malt and Wolff (2010) argue, the relationship between language and cognition is complex: some cognitive processes – such as categorization, memory, attention, conceptualization, and pragmatics – may be more affected by linguistic structures than others. This suggests that language influences thought to varying degrees depending on the specific cognitive domain or task involved.

Numerous studies have demonstrated language effects across cognitive domains, including color, pitch, number, motion, and time (Athanasopoulos 2006; Athanasopoulos and Bylund 2013; Boroditsky 2001; Dolscheid et al. 2013; Roberson et al. 2008; Stutterheim and Nüse 2003). For example, Athanasopoulos (2009) showed that bilingual Greek-English speakers shift their cognitive color representations based on language dominance and cultural context. Similarly, Dolscheid et al. (2013) found that Dutch and Farsi speakers, who use different metaphors for pitch (high/low versus thin/thick), represented pitch differently in non-verbal tasks.

Athanasopoulos (2006) found that Japanese-English bilinguals showed different numerical judgments depending on their L2 proficiency. While Japanese monolinguals, whose language lacks plural marking, focused on non-countable objects, advanced L2 speakers aligned more closely with English speakers, suggesting grammar can shape – and be reshaped – by language learning.

Stutterheim and Nüse (2003) demonstrated that speakers of English, German, and Algerian Arabic segment motion events differently due to language-specific patterns of perspective-taking. Boroditsky (2001) found that English and Mandarin speakers conceptualized time using horizontal and vertical spatial metaphors, respectively – differences that persisted even in non-verbal reasoning tasks.

Research on grammatical gender falls within this broader tradition of investigating how linguistic structures affect cognition (Almutrafi 2015; Athanasopoulos and Boutonnet 2016; Bassetti 2007; Beller et al. 2015; Bender et al. 2016a; Haertlé 2017). These studies explore how grammatical gender influences cognitive processes such as object perception, categorization, and gender associations.

For example, Almutrafi (2015) found that Arabic speakers, whose language includes grammatical gender, showed gender-based voice attribution in object judgments – effects absent in English speakers. Beller et al. (2015) found that voice assignment patterns in gendered languages were shaped both by grammatical gender and cultural context. Similarly, Bassetti (2007) observed that Italian-German bilinguals showed weaker gendered conceptualization when exposed to conflicting grammatical gender cues across their two languages.

Haertlé (2017) examined how grammatical gender shaped perceptions of natural versus artificial objects in Polish and French speakers. The study demonstrated that grammatical gender cues can lead to gendered assessments of objects. In contrast, Bender et al. (2016a) found that gender congruency effects were often item-specific rather than driven by grammatical gender.

These findings suggest that grammatical gender contributes to how people conceptualize and categorize objects, but that its effects are nuanced, shaped by context, culture, and task design.

2.2 Grammatical gender: an obligatory feature in some languages

Grammatical gender systematically classifies nouns into categories, typically based on attributes such as sex (e.g., masculine, feminine) or other inherent properties. According to Hockett (1960), grammatical gender categories are revealed through agreement patterns of related words, including adjectives, determiners, and verbs. This means that the grammatical gender of a noun governs the form of words that syntactically agree with it. For example, in German, adjectives and determiners must align with the noun's gender, as in *eine kleine Flasche* ('a small bottle'), where *eine* (a) and *kleine* (small) agree with the feminine noun *Flasche* (bottle). In languages with grammatical gender, such agreement is an obligatory structural feature distinguishing noun classes and enabling syntactic consistency.

While grammatical gender sometimes aligns with real-world sex distinctions, this alignment is often partial or absent (Corbett 1991). For instance, in German, *Mädchen* ('girl') has neuter gender despite referring to a feminine entity. Similarly, in Polish, the diminutive *dziewczę* ('girl') carries neuter gender, and some nouns for females may adopt masculine gender and vice versa (Maciuszek et al. 2019).

Languages such as Czech demonstrate the pervasive nature of grammatical gender, with adjectives always carrying gender-specific endings and lacking a base form. For example, *velký* ('big') appears as masculine, *velká* feminine, and *velké* neuter, requiring adjective endings to match the noun's gender (Tahal 2010). This highlights the obligatory nature of grammatical gender in such languages.

Exploring grammatical gender as an obligatory feature helps illuminate how language might influence our worldview. As Lucy (1992) argues, distinguishing obligatory categories from optional ones is essential to understanding how structural imperatives differ from usage preferences (Lucy 2016, 489).

Grammatical gender varies across languages in the number of categories and marking mechanisms. One group features a three-gender system, realized through pre-nominal articles (e.g., German der, die, das) or post-nominal suffixes (e.g., Polish). In Czech, three genders are also marked by post-verbal suffixes (-y, -i) that indicate gender and plurality in verb forms, crucial for grammatical agreement (Fatemi 2023; Tahal 2010).

The second group employs a binary gender system in which grammatical gender predominantly distinguishes between feminine and masculine forms. Languages such as Spanish, French, and Portuguese mark gender distinctions through pre-nominal articles and post-nominal suffixes (Corbett 1991).

The third group includes natural gender languages such as English, where gender is primarily indicated through pronouns (he, she, it), aligning mostly with biological sex (Corbett 1991; Prewitt-Freilino et al. 2012). While some classify English as genderless due to the lack of grammatical gender marking on inanimate nouns (Forbes et al. 2008; Kousta et al. 2008), English enforces gender agreement in pronoun use (e.g., she for girl), reflecting a gender system (Corbett 1991).

In conclusion, grammatical gender is an obligatory feature across these three types of languages – three-gender, two-gender, and natural-gender systems – expressed through various linguistic mechanisms such as articles, suffixes, or pronouns. These markings are represented in morphosyntax, affecting sentence structure, verb tense, and inflection of related words (Corbett 1991). For speakers aiming for grammatical accuracy, awareness of their language's grammatical gender system is essential.

Not all languages conform to these patterns; for example, Farsi lacks grammatical gender distinctions entirely (Esmaili 2023). Such variation underscores the importance of exploring how structural differences in grammatical gender systems might influence cognition and conceptualization, setting the stage for the following examination of grammatical gender in German and noun phrase structure in Farsi.

2.3 Gender system in German and Farsi

This study focuses on a cross-linguistic comparison between German and Farsi in terms of grammatical gender. German features a grammatical gender system with three categories – feminine, masculine, and neuter – governed by both semantic and formal rules (Dudenredaktion 2005; Duke 2005). In contrast, Farsi lacks grammatical gender distinctions entirely, placing the two languages at opposite ends of the grammatical gender spectrum (Fatemi 2023). This divergence makes them ideal for critically examining the linguistic relativity hypothesis, which posits that grammatical structures influence cognition (Lucy 2016, 487–490).

In German, grammatical gender is pervasive and affects nouns, articles, and adjectives across the four grammatical cases – nominative, accusative, dative, and genitive – resulting in dynamic inflectional patterns. Gender assignment in German relies on semantic categories (e.g., *der Mann* 'the man' is masculine; *die Frau* 'the woman' is feminine) and formal markers, such as suffixes (e.g., nouns ending in -er are usually masculine, while -e indicates feminine). Tables 1 and 2 summarize these semantic and morphological criteria (Dudenredaktion 2005).

Nevertheless, a significant number of exceptions persist in nouns whose gender cannot be deduced through the aforementioned principles. For instance, diminutive nouns in German are obligatorily assigned the neuter gender, regardless of the natural biological gender of the referent. Examples include "*das Mädchen*" (the girl) and "*das Büblein*" (the little boy). These cases highlight the precedence of grammatical rules over natural gender in such instances. Similarly, nouns such as "*der Käse*" (the cheese) and "*die Jeans*" (the jeans) deviate from formal assignment criteria. Overall, these exceptions underscore the nuanced and occasionally non-intuitive nature of the German gender system.

Considering the frequency with which grammatical gender is realized in basic syntactic structures in German, it becomes evident that this feature is inherently present in all linguistic expressions, continuously

Table 1: Grammatical gender classes in German: semantic criteria overview.

Grammatical gender class/gender class	Semantic criteria
Masculine	Seasons, months, days of the week, star constellations, winds, forms of precipitation, types of soil and rock, currencies, proper names for male persons, mountains, and cars
Feminine	Individual flowers and trees, some abstract concepts, nominalized numbers, fonts, names denoting female persons, ships and airplanes (especially those named after places), motorcycles
Neutral	Most types of metal, chemical elements and medications, diminutives (including proper names), nominalized parts of speech and letters (das A), collectives and pejorative terms beginning with Ge-, place names in general, most names of countries, hotels, cafes, and cinemas

Table 2: Grammatical gender classes in German: formal criteria overview.

Grammatical gender class/gender class	Formal criteria
Masculine	-ich, -ig, -ling, -s, -and, -ant, -är, -ast, -eur / -ör, -ient, -ier, -iker, -ikus, -ismus, -ist, -or
Feminine	-e, -ei, -itl, -heit, -keit, -schüft, -ung, -a, -ade, -age, -aille, -aise / -äse, -ance, -äne, -anˆ -ation, -elk, ette, -euse, -ie, -{i) entˆ -(i)ere, -ik, -tile, -ine, ion / -ation, -isse, -(i)tät, -itis, -ive, -ose, -sis / -se, ur, -ürr
Neutral	-chen, -lein, -le, -icht, -tel, -tum, -eau / -o, -ett, in, -ing, - (, i) um, -ma, -ment

influencing syntactic structures at every level of language use. This feature exhibits a dynamic nature across the four grammatical cases – nominative, accusative, dative, and genitive. In each case, the articles and potentially accompanying adjectives are modified according to the specific grammatical gender of the noun, highlighting the inherent variability of grammatical gender across different linguistic contexts. Furthermore, the definite and indefinite forms of the noun introduce an additional layer of complexity. The articles, both definite and indefinite, undergo declination in different forms across these four cases in German, which further illustrates the syntactic and morphological role of grammatical gender as an integral feature of the German language (Dudenredaktion 2005). Table 3 illustrates inflection patterns for feminine nouns, showing how articles and adjectives change across cases. These patterns differ for masculine and neuter nouns, emphasizing grammatical gender's impact on morphosyntactic structure.

As depicted above, grammatical gender is a fundamental and obligatory feature of German. The language exhibits a robust obligatory gender system, where morphological cues embedded in nouns, adjectives, and articles enforce consistent agreement across syntactic contexts. This pervasive feature significantly shapes the structure of the language.

Farsi,[3] on the other hand, is a gender-neutral language with a simpler nominal system (Sagot and Walther 2010) and no grammatical gender. While it lacks morphological case marking (Lazard and Paper 1959),

Table 3: Declination of feminine nouns with definite and indefinite articles across German cases.

Cases	Adjective + definite/indefinite noun (feminine)		
Nominative	*Die*	*schön-e*	*Tochter*
	The (F.NOM. ART.)	nice-F.SG	daughter
	eine	*schön-e*	*Tochter*
	a (F. INDF. ART.)	nice-F.SG	daughter
Accusative	*Die*	*schön-e*	*Tochter*
	The (F.ACC. ART.)	nice-F.SG	daughter
	eine	*schön-e*	*Tochter*
	a (F. INDF.ACC ART.)	nice-F.SG	daughter
Dative	*Der*	*schön-en*	*Tochter*
	The (F.DAT. ART)	nice-F.DAT.SG	daughter
	einer	*schön-en*	*Tochter*
	a (F. INDF.DAT. ART.)	nice-F.SG	daughter
Genitive	*Der*	*schön-en*	*Tochter*
	The (F.GEN. ART)	nice-F.GEN.SG	daughter
	einer	*schön-en*	*Tochter*
	a (F. INDF.GEN. ART.)	nice-F.GEN.SG	daughter

3 In this article, the term Farsi is used to refer collectively to Persian, Parsi, and Dari for the sake of clarity and accessibility. Farsi is an inflectional language with a subject-object-verb (SOV) word order and is written from right to left. It belongs to the Western Iranian branch of the Indo-European language family and is predominantly spoken in Iran.

Table 4: Case system in Farsi.

Cases	Adjective + definite/indefinite noun (feminine)			
Nominative	آمد/amæd come-PST-3rd s The beautiful girl came.	زیبا/ziba Beautiful/ATTR	دخترِ/Dokhtær-e girl	
Accusative	بوسید/bus-id kiss-PST-3rd s The father kissed the daughter.	را/ra postposition “râ” ACC-marker	دختر/dokhtær girl/daughter. ACC	پدر/pedær father
Dative/ Completive	کمك کرد/komak-kærd help-did-3rd The father helped the daughter.	دختر/dokhtær girl/daughter/IOBJ	به/be preposition “be” to	پدر/pedær father
Predicative	است/æst To be-3rd SG The weather is cold.	سرد/særd ADJ.PRED cold	هوا/hæva weather	
Genitive	است/æst to be-3rd SG Maryam’s mother is beautiful.	زیبا/ziba beautiful-adj	مریم/Maryam Maryam	/e EZ/GEN-marker	مادر/madær mother
Attributive	آمد/amæd come-PST-3rd s The devoted mother came.	فداکار/fædakar devoted/adj	/e EZ/GEN-marker	مادر/madær mother
Vocative	خدا/Xoda God Oh god!		اي/ey Oh/EXCLM	
Adverbial	است/æst to be-3rd SG The bird is on the tree.	درخت/derækht tree/LOC	روي/ruye on/PREP	پرنده/pærænde bird
Object completive	نامم ي/mi-nam-æm PRS. PREFIX.to name. 1ST SG I name him Armin.	آرمین/Armin OBJ.CMPL	را/ræ postposition “râ”. ACC-marker	او/u PRO.3rd SG	من/mæn I
Appositive case	بود/bud to be-PS.3rd SG Yesterday, Armin, the carpenter, was here	اینجا/indʒa here.ADV	نجار/Nædʒar carpenter.APPOS	آرمین/Armin	دیروز/Diruz yesterday

researchers note ten semantic and syntactic case roles in Farsi (Ahamadi-Givi 2004; Khayyampour 1998). These include nominative, accusative, dative, genitive, attributive, predicative, adverbial, object-completive, vocative, and appositive, as shown in Table 4.

In Farsi, case relationships are expressed through prepositions, postpositions, and syntax rather than inflected endings. The nominative case identifies the subject of a sentence, typically positioned at the beginning. The accusative case is marked by the postposition [ra], essential for direct objects of transitive verbs such as *busidæn* (to kiss). The dative case employs prepositions such as [bɛ] to indicate indirect objects, while the predicative case is formed with the copula verb *budan* (to be). Genitive and attributive relationships are expressed by using the Ezafe marker [-e], which links nouns or adjectives (e.g., *ketab-e man* “my book”). Notably, this marker is pronounced but remains unwritten in standard Farsi. Other roles, such as vocative, adverbial, and object-completive, rely on word order and context rather than explicit markers. Similarly, the appositive case clarifies a noun and is inferred from sentence semantics. These features demonstrate Farsi’s reliance on syntactic and auxiliary elements to establish grammatical relationships and its avoiding fixed morphological rules.

By contrasting grammatically gendered German with genderless Farsi, we explore how grammatical gender may influence cognition. If German speakers show a significant influence of grammatical gender on their conceptualization of nouns, while Farsi speakers exhibit divergent patterns, we claim that these results lend

support to the linguistic relativity hypothesis. Thus, this analysis highlights the cognitive implications of grammatical gender and the broader relationship between language and thought.

2.4 Navigating gendered constructs: insights from linguistic relativity studies

2.4.1 Previous research

Research into the influence of grammatical gender on cognition spans over a century, with studies dating back to 1915, predating the Sapir-Whorf hypothesis (Konishi 1993). Konishi (1993) discusses how grammatical gender may guide speakers to anthropomorphize words based on the grammatical gender assigned to objects in their languages. For instance, a seminal 1915 Russian study found that participants consistently personified Monday, Tuesday, and Thursday as masculine, aligning with their grammatical gender, while Wednesday, Friday, and Saturday were personified as feminine (Konishi 1993).

Following the Sapir-Whorf hypothesis, numerous studies have investigated how grammatical gender influences cognitive processes. Sera et al. (1994) used classification tasks to examine whether Spanish speakers assign gendered attributes to objects based on grammatical gender. They found Spanish speakers tended to attribute masculine or feminine qualities consistent with grammatical gender in Spanish, whereas English speakers, whose language lacks grammatical gender, associated masculine traits with artificial objects and feminine traits with natural ones. Subsequent research extended these findings to languages such as Arabic, French, and Norwegian using voice-assignment tasks (Almutrafi 2015; Athanasopoulos and Boutonnet 2016; Beller et al. 2015). Although these studies support the hypothesis that grammatical gender shapes conceptualization, the use of explicit gender-judgment tasks raises questions about whether the effects reflect unconscious linguistic influences or conscious reasoning.

Boroditsky et al. (2003) conducted a widely cited study in which German and Spanish speakers described objects using adjectives. They found that participants tended to choose adjectives aligned with the grammatical gender of the object in their native language – for example, describing the German masculine word *Schlüssel* ('key') as 'hard' or 'heavy,' and the Spanish feminine word *llave* ('key') as 'lovely' or 'intricate.' However, this study excluded nouns that are grammatically neuter in German and did not include statistical analyses. In a follow-up study by Phillips and Boroditsky (2003), participants rated the similarity between pairs consisting of a human figure (male or female) and an inanimate object or animal with masculine or feminine grammatical gender. Although conducted in English, similarity judgments were influenced by grammatical gender congruency in German and Spanish, suggesting that native grammatical gender systems shape conceptual associations beyond native language use.

Some scholars argue that grammatical gender effects are evident only in languages with two grammatical genders, such as Arabic, Italian, French, and Spanish (Sera et al. 2002; Vigliocco et al. 2005). They used categorization tasks and found no effects in languages with more than two grammatical genders, such as German. However, contrary evidence exists from studies on Polish, a three-gender language, which demonstrated that grammatical gender effects can occur in three-gendered languages (Haertlé 2017; Maciuszek et al. 2019). This suggests that restricting the influence of grammatical gender on cognition to two-gender languages is not logically sound, highlighting the need for further research across diverse grammatical gender systems.

More recent replication attempts (Elpers et al. 2022) with improved methodologies and larger samples have yielded mixed results, finding no effect of grammatical gender on similarity judgments between German and Spanish speakers. These mixed findings highlight the complexity of comparing languages with different gender systems (three genders in German versus two in Spanish).

Electrophysiological studies using electroencephalogram (EEG)[4] provide evidence for implicit, unconscious activation of grammatical gender during language processing (Boutonnet et al. 2012; Sato et al. 2020).

4 EEG, widely utilized in psycholinguistics, measures electrical brain activity and serves as an effective tool to capture automatic and unconscious neural responses during language processing.

Event-related potential components (ERP) such as the N400[5] and Left-Anterior Negativity (LAN)[6] indicate morphosyntactic processing even when gender is not explicitly judged, supporting the view that grammatical gender influences cognition beyond conscious awareness.

Behavioral paradigms such as the Extrinsic Affective Simon Test (EAST),[7] which measures implicit associations between stimuli and affective responses, have been used to assess gender congruency effects (Bender et al. 2016a, 2018). These studies found stronger grammatical gender effects for animate objects whose biological sex matches grammatical gender, while effects for inanimate or grammatically neuter nouns were weaker or item-specific. Importantly, these tasks probe implicit processing, unlike explicit voice assignment tasks that involve conscious judgments.

Cross-linguistic comparisons with genderless languages, such as Tamil (Sedlmeier et al. 2016), emphasize the unique role of grammatical gender in shaping cognition. These studies found significant gender effects in German but not in Tamil, suggesting that grammatical gender influences cognition in a language-specific rather than universal manner.

Despite extensive research, significant gaps remain. In particular, the neutral grammatical gender has been underrepresented in previous studies, with most research focusing on masculine and feminine categories, overlooking the complexity of languages such as German with three genders. Additionally, the variety of methods – ranging from explicit to implicit tasks – and the inconsistent findings call for more integrative approaches. A summary of the key studies reviewed in this section is presented in Table 5, providing an overview of the methodologies, language systems, gender categories, and principal findings.

Table 5: Summary of key studies on grammatical gender and cognition.

Study	Language(s)	Methodology/task	Gender focus	Main findings
Konishi (1993)	Russian	Anthropomorphizing days of week	Masculine, feminine	Grammatical gender influences personification of days
Sera et al. (1994)	Spanish, English	Classification, voice assignment	Masculine, feminine	Spanish speakers assigned traits consistent with grammatical gender; English speakers associated masculinity with artificial and femininity with natural objects.
Boroditsky et al. (2003)	German, Spanish	Adjective description, similarity ratings	Masculine, feminine	Gender affects object description and similarity judgments
Phillips and Boroditsky (2003)	German, Spanish	Similarity judgement	Masculine, feminine	Gender affects similarity ratings
Elpers et al. (2022)	German, Spanish, English	Similarity judgments	Masculine, feminine	No significant gender effects on similarity ratings
Boutonnet et al. (2012)	Spanish-English	EEG semantic categorization	Masculine, feminine	Implicit activation of grammatical gender in bilinguals
Sato and Athanasopoulos (2018)	French-English	ERP gender-related judgments	Masculine, feminine	Grammatical gender modulates early neural processing
Bender et al. (2016a, 2018)	German	EAST, voice assignment	Masculine, feminine, neuter	Gender congruency effects vary by animacy and gender type
Sedlmeier et al. (2016)	German, Tamil	Potency scales, adjectivet production & similarity judgement	Masculine, feminine, neuter	Effects found in German but not in genderless Tamil

5 The N400 wave is a negative-going potential peaking around 400 ms post-stimulus, reflecting semantic integration processes (Kutas and Hillyard 1984).

6 The Left-Anterior Negativity is an ERP marker of morphosyntactic processing.

7 Reaction time–based psychological paradigm used to measure implicit associations between concepts – such as grammatical gender and emotional valence – without requiring participants to explicitly judge the relevant features.

This review shows that grammatical gender can influence how speakers assign properties to objects, especially in languages such as German and Spanish. However, findings vary, particularly when comparing languages with different gender systems. A notable gap remains: the limited attention to the neutral grammatical gender and the need for methodologies that balance implicit and explicit measures. The present study addresses these gaps by including neuter nouns in German and contrasting responses with Farsi, a genderless language, for a comprehensive cross-linguistic examination.

3 Methodology

In this paper, we aim at filling some of the research gaps that we detected in the current state of the art. First, the study explores whether the grammatical gender of nouns in German influences their conceptual meaning and the mental representation of objects among L1 German monolingual speakers. Second, the study examines whether Farsi native speakers differ from German native speakers in assigning properties to objects and animals. To answer these questions, a two-phased study was conducted. During the first phase, German and Farsi native speakers completed a property judgment task in which they were asked to describe a set of nouns using the first three adjectives that came to mind. The adjectives generated from these responses were then used in the second phase, where a different group of native German and Farsi speakers rated the adjectives on a Likert scale. The resulting data were analyzed using a Linear Mixed Model, with the mean of the adjective ratings serving as the dependent variable. Post-hoc tests were conducted to further assess the significance of the model's findings.

3.1 Participants

The study involved two phases with two separate groups of participants. In the first phase, 74 German monolinguals[8] and 65 Farsi monolinguals participated. The German monolinguals were recruited from the Dortmund area in Germany (45 females, age: M = 27, SD = 8.3, range: 18–60), while the Farsi monolinguals were recruited from Mashhad, Iran (35 females, age: M = 34.6, SD = 8.2, range: 20–59). In this phase, the participants were asked to describe objects as part of the experimental task.

In the second phase, 106 German monolinguals and 60 Farsi monolinguals were recruited from Dortmund and Mashhad, respectively. The German group consisted of 84 females and 21 males (one participant did not disclose their gender; age: M = 23.7, SD = 3.7, range: 18–49). The Farsi group included 37 females and 23 males (age: M = 35.9, SD = 8.9, range: 16–55). The participants rated the adjectives on a 7-point Likert scale.

In both phases, the participants exhibited comparable age ranges, with standard deviations indicating similar variability in age. Gender distributions were skewed toward females across both groups but remained consistent between phases and regions.

To assess socioeconomic status (SES), educational attainment was used as a proxy. In the first phase, the majority of German participants had completed the Abitur[9] (45.95 %) or held a bachelor's degree (41.89 %), whereas most Farsi participants held a bachelor's degree (65 %). In the second phase, 75.24 % of German participants had completed the Abitur, and 23.81 % had earned bachelor's degrees. Farsi participants in this phase showed a similar pattern to those in the first, with 65 % holding bachelor's degrees. While the overall educational levels were relatively high across both groups, it should be noted that in the second phase, a higher proportion of Farsi participants held bachelor's degrees compared to the German group (65 % vs. 23.81 %). Most German participants were university students who received one ECTS credit as an incentive. The remaining German

8 Speakers were considered monolingual if they had exclusively communicated in either German or Farsi during their preschool years. Additionally, participants were required not to have lived outside Germany or Iran for more than six months up to the age of 22, which was regarded as the typical age for completing a bachelor's degree. This approach ensured that the linguistic competencies of both groups were predominantly influenced by their first language and their extended residence in the country where their first language is spoken.

9 The German Abitur is equivalent to a high school diploma or A-levels in English-speaking countries.

Table 6: Participants' demographic information.

Phase	Group	Total N	Female N	Male N	Age range	Age mean (SD)
Phase 1	German L1	74	45	29	18–60	27 (8.3)
	Farsi L1	65	35	30	20–59	34.6 (8.2)
Phase 2	German L1	106	84	21	18–49	23.7 (3.7)
	Farsi L1	60	37	23	16–55	35.9 (8.9)

participants volunteered without financial compensation. All Farsi participants volunteered and did not receive any financial compensation. Table 6 provides a summary of participant characteristics.

3.2 Material/stimuli selection

The study utilized 30 stimuli, 30 control items, and 45 fillers. The stimuli and five control items were selected from the German version of the Peabody Picture Vocabulary Test (PPVT). These items consisted of everyday vocabulary, represented through pictures of both animate and inanimate objects. A balanced number of feminine, masculine, and neutral stimuli and control items was included, with each category containing ten words.

Out of the 228 PPVT words, 94 refer to objects. Among these, 36 are feminine, 38 are masculine, and 20 are neuter in grammatical gender. The frequencies of these 94 words were obtained using the COSMAS II query interface, accessing the W-Archive of the written language (W-public) – a publicly available subset of the German Reference Corpus (DeReKo) that includes written texts from a variety of genres and includes recent acquisitions. To categorize word frequencies, we used a frequency class table (see Appendix C) following the method outlined on the Ruhr University Bochum's corpus linguistics page (homepage.ruhr-uni-bochum.de/Stephen.Berman/Korpuslinguistik, R Core Team 2023). This method involves transforming raw frequency values into discrete frequency classes using a logarithmic scale, which allows for meaningful comparisons between words across a large frequency range. While DeReKo includes raw frequency data, this classification provides a more interpretable metric for psycholinguistic analysis. The purpose of applying this frequency classification was to ensure that the selected feminine, masculine, and neuter nouns fell within comparable frequency bands, minimizing the potential confounding effect of frequency on participants' responses. From the original pool, we selected the ten nouns from each grammatical gender category with the highest frequency class (i.e., the most frequent items according to the classification system). All remaining analyses, including statistical modeling and visualizations, were conducted using R (R Core Team, 2023).

The control items were lexical items denoting human referents whose grammatical and biological genders aligned with feminine and masculine cases, such as "Mutter" (mother), "Braut" (bride), "Vater" (father), and "Koch" (cook). For neutral grammatical gender, words associated with a specific sex in real life but assigned neutral grammatical gender in German were chosen, such as "Mädchen" (girl), "Herrchen" (master), and "Fräulein" (miss). Additionally, this category included words whose overarching terms have neutral grammatical gender in German but whose real-world references possess a specific biological sex, such as "Geschlecht" (gender), "Opfer" (victim), or "Baby" (baby).

In this study, the linguistic terms of objects and their corresponding images were incorporated. Most stimulus images were adapted from the PPVT, except for the image corresponding to the word "Huhn" (hen), which was sourced from Pixabay's free image collection. Among the five control items – "Archäologin" (female archaeologist), "Athlet" (male athlete), "Zimmermann" (male carpenter), "Ritter" (male knight), and "Fußgängerin" (female pedestrian) – both the terms and their images were adopted from the PPVT. The remaining images were either sourced from Pixabay's free image collection or hand-drawn by a colleague.

The 45 fillers consisted of clock images, with participants tasked with telling the time. These clocks were adapted from an ongoing study conducted by a PhD colleague. The number of fillers was chosen to prevent participants from deducing the study's main objectives while avoiding an overwhelming number of irrelevant

tasks. All images were presented in black and white to avoid triggering gender-specific color associations (e.g., pink as feminine or blue as masculine) that could bias participants' judgments. A comprehensive list of all stimuli images and their corresponding linguistic terms is provided in Tables A1, A2 and A3 (Appendix B).

3.3 Survey procedure

The primary aim of this study was to investigate the influence of grammatical gender on cognition using an online questionnaire administered via LimeSurvey. While all German-speaking participants completed the survey online, a subset of Farsi-speaking participants was unable to do so due to internet shutdowns in Iran. As a result, these participants completed the survey in a pen-and-paper format.

The survey was designed to take no longer than 20 min and included randomized stimuli, control items, and fillers. At the end of the survey, participants completed a linguistic competence questionnaire to gather information on their language acquisition history and to confirm their monolingual status. Demographic information, such as age, sex, and education level, was also collected. The survey was translated into Farsi for Farsi-speaking participants, including all items and the demographic questionnaire.

This study was reviewed and approved by the Joint Ethics Committee of Faculties 9, 11–17 at TU Dortmund University (Approval Number: GEKTUDO_2024-76). The research adhered to ethical guidelines for scientific investigations and was deemed ethically unobjectionable.

At the beginning of the survey, all participants received instructions in both German and Farsi, briefly outlining the tasks. In the first phase, participants were asked to describe images and provide spontaneous responses. A second group of participants then rated the adjectives generated in the first phase. Detailed instructions for both phases, including the image design and their translations, can be found in Appendix A.

4 Results

The initial phase of the study yielded 6,660 adjectives from German monolinguals and 5,850 adjectives from Farsi monolinguals. These adjectives were carefully reviewed based on specific inclusion criteria. First, any spelling errors were corrected. Second, German adjectives derived from Participle 1 were excluded, as these are formed by adding "-d" to the infinitive verb, thus focusing on adjectives independent of verb origins. Additionally, adjectives based on verbs or nouns, rather than pure adjectives, were excluded to ensure consistency. Only adjectives mentioned by participants at least four times were retained to ensure sufficient frequency representation. As a result, 326 German adjectives and 157 Farsi adjectives were included in the final analysis. These adjectives were then presented to a separate group of participants (106 native German speakers and 60 native Farsi speakers) who were unaware of the origin of the adjectives and asked to rate them on a 7-point Likert scale from 1 = very feminine to 7 = very masculine. The primary objective of this second phase was to obtain a quantitative representation of the perceived gender connotation for each adjective. The mean score of each adjective across all raters was then calculated and used as a continuous variable in the subsequent statistical modeling as the response variable. This fine-grained approach was inspired by the methodology of Sedlmeier et al. (2016), though our study employed a significantly larger sample of raters (n = 166 versus n = 6), thus enhancing the reliability and generalizability of the findings.

The analysis of the 326 German adjectives revealed that some were used to describe objects across all three grammatical gender categories, while others were predominantly associated with a specific gender. To facilitate the rating task in the second phase of the study, the 326 adjectives were distributed into six groups, each containing a manageable number of adjectives for evaluation. A total of 194 adjectives were used across all grammatical gender categories (feminine, masculine, and neutral), while the remaining adjectives were specifically linked to a single gender. Each group contained a balanced selection of these adjectives, with a maximum of 60 adjectives per group, ensuring that participants could complete the task without excessive cognitive load. The grouping method aimed to evenly distribute adjectives used across all genders while maintaining the gender-

specific distinctions. A detailed breakdown of the adjective distribution across groups in the second phase of the study is provided in Table 7.

Among the 157 Farsi adjectives, only زیبا (/ʒiˈbɑ/; 'beautiful, nice') was used consistently across all objects. Consequently, this adjective was included in all four groups of Farsi adjectives. Group A and Group B each contained 39 adjectives, Group C included 40 adjectives, and Group D comprised 39 adjectives. These adjectives were subsequently evaluated by the second cohort of participants in Farsi. Table 8 provides detailed breakdown of the adjective distribution across groups.

After the adjective ratings were completed, we calculated the average score for each adjective. Our Likert scale ranged from 1 to 7, indicating the possible minimum and maximum average scores. Strategically equidistant boundaries were introduced to define logical limits for further evaluation. Ratings of [1,3] were labeled as feminine, (3,5] as neutral, and (5,7] as masculine.[10]

The 194 adjectives produced by participants to describe objects of all grammatical genders in German were specifically examined, as these adjectives were suspected to reveal particular tendencies in participant responses. It was found that 81.93 % of these adjectives were categorized as neutral, i.e., with an average score between 3 and 5. In contrast, 10.82 % were assigned as feminine [1,3] and 8.25 % as masculine (5,7].

In the analysis, these average scores of each adjective were employed alongside the key variables of the participants from the first study, such as age, gender, and education level. These variables were integrated into our statistical analysis to account for potential influences on adjective ratings. To provide a comprehensive overview, Figure 1 presents the descriptive analysis of the data for both German and Farsi, utilizing the average scores for each adjective incorporated into our data set.

Figure 1 displays the average scores separated for German and Farsi, categorized according to the grammatical gender system in German. The groupings to feminine, masculine, and neutral for Farsi were adapted from German grammar due to Farsi's genderless system. We observe a downward trend in the median in both German

Table 7: Distribution of adjectives across gender categories and groups in the second phase of the study for German speakers.

Group	Total adjectives	Adjectives used across all gender	Gender-specific adjectives
A	60	29	31
B	59	29	30
C	59	30	29
D	59	30	29
E	45	39	8
F	44	37	8
Total	326	194	

Table 8: Distribution of adjectives across object categories and groups in the second phase of the study for the Farsi speakers.

Group	Total adjectives	Adjectives used across all objects	Object-specific adjectives
A	39	1 (زیبا/ʒiˈbɑ/ - "beautiful, nice")	38
B	39	1 (زیبا/ʒiˈbɑ/ - "beautiful, nice")	38
C	40	1 (زیبا/ʒiˈbɑ/ - "beautiful, nice")	39
D	39	1 (زیبا/ʒiˈbɑ/ - "beautiful, nice")	38
Total	157	1	

10 In this context, the notations [1,3] represent an interval that includes both endpoints, 1 and 3. Conversely (3,5] denotes an interval that begins just after 3 (excluding 3) and includes 5. Similarly, (5,7] indicates an interval starting just after 5 (excluding 5) and including 7.

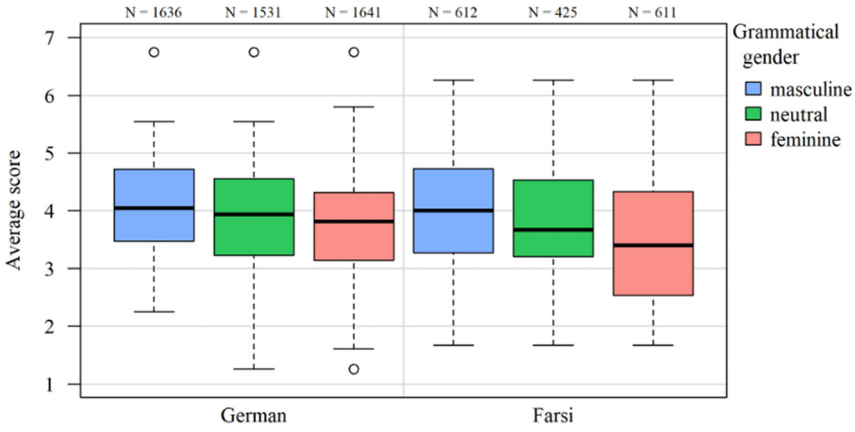


Figure 1: Boxplots of the mean scores separated according to the grammatical gender of the objects for German and Farsi.

and Farsi, progressing from masculine (German median = 4.05, Farsi median = 4.00) to neutral (German median = 3.94, Farsi median = 3.67) and ending with feminine (German median = 3.81, Farsi median = 3.40). What is particularly evident here is the variability in the results. The range of the average score for Farsi speakers is identical across all three groups: feminine, masculine, and neutral. In contrast, the German data indicates less variability for the masculine grammatical gender. However, for feminine and neutral genders in German, the variability is nearly identical. Lexeme-specific variability in both the German and Farsi datasets is further examined in Figures 2, 2.4 and 2.5 (Appendix D), which presents boxplots of the mean adjective scores, grouped by participant gender, the grammatical gender of the stimulus nouns, and the order of the adjectives (1–3), as provided by the participants. Although this visualization enables a more fine-grained examination of how individual lexical items vary across conditions, it does not pertain directly to the core hypothesis of this study. To determine whether the observed differences in average scores are statistically significant, we applied a linear mixed-effects model, following the framework outlined by Fahrmeir et al. (2013).

The following paragraph addresses our research question concerning the potential influence of grammatical gender on German speakers, as determined by the grammatical gender inherent in the stimulus nouns. An extensive analysis was conducted using a linear mixed model (Fahrmeir et al. 2013). The dependent variable was the mean rating of the adjectives on a Likert scale, representing the perceived gender connotation. For each participant and each noun, up to three adjective ratings were collected, capturing how masculine, feminine, or neutral the noun was perceived to be. These scores served as the basis for evaluating whether grammatical gender systematically influences adjective production.

The model incorporated both fixed and random effects to account for participant variability and object characteristics. The independent variables included the grammatical gender with values masculine, feminine, and neutral (in German der, die, das), as well as the order of the adjectives (first, second, third) produced by each participant in describing each object. Furthermore, we considered the interaction between the grammatical gender and the order of the adjectives to take into account that the effect of the grammatical gender could be modified by the order of the adjectives. Importantly, the first adjective is understood to reflect the most immediate and spontaneous cognitive association, while the second and third adjectives tend to represent more deliberate and reflective judgments. Additionally, participants' gender and age (continuous variables) were included as independent variables in our analysis, given that this study investigates associations between grammatical and conceptual gender. It is plausible that participants' own biological sex, as well as generational differences reflected in age, could influence how they perceive and attribute gendered characteristics to objects. Including these variables allowed us to control for potential confounding effects in the analysis. Notably, the reference levels for the three categorical variables (Grammatical gender, order of the adjectives, the gender of the participant) in our model were set to *masculine* (in German *der*) for the grammatical gender, to *first* for the ordering of the adjectives being produced, and to *male* for participants' gender. In the analysis, a p-value < 0.05 was considered statistically significant.

The results of the linear mixed model for German monolinguals are summarized in Table 9. The study found no significant difference in the average score between female and male participants (b = −0.024, SE = 0.031, t (68.64) = −0.761, p = 0.449). Similarly, participant age did not have a significant effect (b = 0.002, SE = 0.002, t (78.38) = 1.277, p = 0.205). However, our analysis using a linear mixed model revealed a significant negative main effect for feminine grammatical gender (b = −0.434, SE = 0.181, t(29.28) = −2.405, p = 0.023), indicating nouns with feminine grammatical gender were described as more feminine than nouns with masculine grammatical gender. Conversely, no significant difference was found in the average score between the neutral and masculine grammatical genders (b = −0.243, SE = 0.181, t(29.38) = −1.346, p = 0.189). Regarding the adjective order, the average score for the second-mentioned adjectives was significantly lower than the first-mentioned adjectives (b = −0.135, SE = 0.045, t(4718.80) = −2.983, p = 0.003) and the same was true for third-listed adjectives (b = −0.114, SE = 0.046, t(4721.49) = −2.446, p = 0.014). This suggests that the second and third adjectives are more feminine than the first one, with a stronger effect for the second adjective. The study found significant interaction effects between the grammatical genders and the order of the adjectives. Notably, the average score for feminine grammatical gender described by the second adjective significantly increased (b = 0.240, SE = 0.064, t(4720.44) = 3.739, p < 0.001) as did that of the third adjective (b = 0.145, SE = 0.065, t(4720.23) = 2.214, p = 0.027), suggesting a relative shift toward more masculine-associated language with later adjectives in feminine noun contexts. Furthermore, for nouns with neutral grammatical gender, the average score associated with the second adjective significantly increased (b = 0.162, SE = 0.065, t (4719.37) = 2.477, p = 0.013), as did the average score for the third adjective (b = 0.136, SE = 0.067, t (4722.56) = 2.041, p = 0.041). Since all estimated coefficients for the interaction terms between grammatical gender and adjective order were statistically significantly different from zero, we conducted post-hoc tests to further explore the differences in the average scores between the grammatical gender categories within the order of adjectives.

To pinpoint the specific differences in the average score between the grammatical genders within the order of the above-mentioned adjectives, we conducted post-hoc tests, i.e. pairwise comparisons using z-tests. These were corrected with Holm's sequential Bonferroni (1979) procedure. An overview of the post-hoc results is given in Table 10. We found a significant difference between feminine and masculine grammatical gender for the first adjectives (d = −0.434, z = −2.405, SE = 0.181, p = 0.048), showing more feminine adjectives were used for feminine grammatical gender compared to masculine. This significant difference is only present in the first adjectives and no other pairs of grammatical genders showed any significant values.

The results of the linear mixed model for Farsi monolinguals are presented in Table 11. To critically examine the hypothesis of linguistic relativity, this study artificially imposed German grammatical gender categories onto the Farsi dataset. The rationale behind this approach was that, since Farsi is a genderless language, any grammatical gender effects found in German should not be replicated in Farsi. In this context, grammatical gender was

Table 9: Results of the linear mixed model for German monolinguals.

	b	SE	df	*t*-value	*p*-value
Intercept	4.085	0.141	43.114	28.886	< 0.001
GG = feminine	−0.434	0.181	29.279	−2.405	0.023
GG = neutral	−0.243	0.181	29.384	−1.346	0.189
Adjective order = second	−0.135	0.045	4718.795	−2.983	0.003
Adjective order = third	−0.114	0.046	4721.493	−2.446	0.014
Participant gender = female	−0.024	0.031	68.637	−0.761	0.449
Age (in years)	0.002	0.002	78.375	1.277	0.205
GG = feminine × adjective order = second	0.240	0.064	4720.435	3.739	< 0.001
GG = neutral × adjective order = second	0.162	0.065	4719.305	2.477	0.013
GG = feminine × adjective order = third	0.145	0.065	4720.276	2.214	0.027
GG = neutral × adjective order = third	0.136	0.067	4722.555	2.041	0.041

The linear mixed model is based on a total of 4808 observations from 74 participants. Both the participant's ID and the 30 objects to be described with an adjective were considered as random intercepts. b, estimated regression coefficient; SE, standard error; df, degrees of freedom; GG, grammatical gender. Marginal R^2, 0.0251, and Conditional R^2, 0.2354.

Table 10: Results of the post-hoc tests for German monolinguals.

Adjective order	Contrasts between GGs	*d*	SE	*z*-value	*p*-value
First	Masculine versus feminine	0.434	0.181	2.405	0.048
	Masculine versus neutral	0.243	0.181	1.346	0.356
	Feminine versus neutral	−0.191	0.181	−1.056	0.356
Second	Masculine versus feminine	0.194	0.181	1.071	0.852
	Masculine versus neutral	0.082	0.181	0.451	1.000
	Feminine versus neutral	−0.112	0.181	−0.619	1.000
Third	Masculine versus feminine	0.290	0.181	1.597	0.331
	Masculine versus neutral	0.107	0.182	0.590	0.629
	Feminine versus neutral	−0.183	0.181	−1.006	0.629

d, estimated mean difference; SE, standard error; GGs, grammatical genders.

included in the Farsi condition solely for purposes of cross-linguistic comparison, with German functioning as the experimental condition and Farsi as the control. To maintain methodological consistency and ensure comparability, the same stimulus groupings used in the German condition, based on grammatical gender (feminine, masculine, and neuter), were applied to the Farsi data. The objective was to investigate whether Farsi speakers, despite the absence of grammatical gender in their language, would nevertheless attribute similar gendered connotations to the stimuli as German speakers. Furthermore, to ensure analytical parity, the same fixed and random effects structure was retained across both language groups.

The study found no significant difference in the average score between female and male participants (b = 0.039, SE = 0.048, t(1612.723) = 0.805, p = 0.421). Similarly, participant age did not have a significant effect (b < 0.001, SE = 0.003, t(1612.782) = −0.147, p = 0.883). Consistent with the German data, feminine grammatical gender showed a significant negative effect in comparison to masculine grammatical gender, which also served as the reference category here (b = *−0.659*, SE = 0.272, t(31.817) = −2.421, p = 0.021). As with the German data, no significant effect was observed for the neutral grammatical gender. However, unlike the German data, the order of the adjectives did not have a significant effect in the Farsi dataset. Additionally, the interaction between the order of the adjectives and the grammatical gender was not significant.

To compare the German and Farsi data, we conducted the post-hoc tests following the same procedure used for the German data to maintain comparability in the analysis. These tests were adjusted using Holm's sequential Bonferroni (1979) correction as well to control for multiple comparisons. An overview of the post-hoc results is presented in Table 12. Pairwise comparisons were performed between the imposed grammatical gender

Table 11: Results of the linear mixed model for Farsi monolinguals.

	b	SE	df	*t*-value	*p*-value
(Intercept)	4.065	0.225	58.773	18.095	< 0.001
GG = feminine	−0.659	0.272	31.817	−2.421	0.021
GG = neutral	−0.145	0.277	33.909	−0.525	0.603
Adjective order = second	−0.161	0.093	1613.023	−1.727	0.084
Adjective order = third	0.006	0.096	1613.346	0.065	0.948
Participant gender = female	0.039	0.048	1612.723	0.805	0.421
Age (in years)	< 0.001	0.003	1612.782	−0.147	0.883
GG = feminine × adjective order = second	0.163	0.132	1612.360	1.234	0.217
GG = neutral × adjective order = second	0.151	0.149	1613.321	1.016	0.310
GG = feminine × adjective order = third	0.033	0.135	1612.502	0.240	0.810
GG = neutral × adjective order = third	−0.153	0.149	1614.416	−1.023	0.306

The linear mixed model is based on a total of 1648 observations from 65 participants. Both the participant's ID and the 30 objects to be described with an adjective were considered as random intercepts. *b*, estimated regression coefficient; SE, standard error; df, degrees of freedom. Marginal R^2, 0.0562 and conditional R^2, 0.3012.

Table 12: Results of the post-hoc tests for Farsi monolinguals.

Adjective order	Contrasts between GGs	*d*	SE	*z*-value	*p*-value
First	Masculine versus feminine	0.659	0.272	31.013	0.065
	Masculine versus neutral	0.145	0.277	33.054	0.603
	Feminine versus neutral	−0.514	0.278	33.347	0.146
Second	Masculine versus feminine	0.496	0.273	31.429	0.237
	Masculine versus neutral	−0.006	0.277	33.261	0.983
	Feminine versus neutral	−0.502	0.278	33.417	0.237
Third	Masculine versus feminine	0.627	0.274	31.878	0.087
	Masculine versus neutral	0.298	0.277	33.159	0.488
	Feminine versus neutral	−0.329	0.277	33.221	0.488

d, estimated mean difference; SE, standard error; GGs, grammatical genders.

categories on the Farsi data for each level of adjective order separately. Unlike the German data, no significant differences were found between grammatical gender categories across the different adjective orders in the Farsi dataset.

These results underscore the influential role of grammatical gender in shaping the conceptualization and categorization of objects among German speakers, as reflected in the adjectives they produce – particularly those mentioned first. The significant interaction between grammatical gender and adjective order further suggests that grammatical gender influences cognitive processing and guides the sequence of adjective selection. In contrast, the Farsi data did not exhibit a comparable pattern, showing no significant interaction between adjective order and the imposed grammatical gender. Moreover, post-hoc analyses revealed no significant differences between any grammatical gender categories in the Farsi condition, reinforcing the absence of grammatical gender effects in this genderless language.

5 Discussion

5.1 The discussion organized by research questions

5.1.1 Does the grammatical gender of the nouns influence the assignment of properties to objects and animals by German native speakers?

The findings indicate that the grammatical gender of nouns may subtly influence how monolingual German speakers assign descriptive properties to objects and animals. In the German dataset, significant main effects emerged for adjective order (i.e., first, second, and third position), as well as a significant interaction between grammatical gender and adjective order (see Table 9). Notably, this interaction was driven primarily by feminine stimuli, for which the effect of adjective order was strongest. Post-hoc pairwise comparisons confirmed that this effect was most robust for the first adjective produced, revealing a statistically significant contrast between nouns with masculine and feminine grammatical gender (see Table 10). Specifically, the first adjective selected by participants consistently aligned with the grammatical gender of the referent noun, suggesting that grammatical gender can shape immediate, automatic conceptual associations. This pattern aligns with the notion of a primacy effect (Asch 1946), whereby early responses are more susceptible to implicit linguistic structures, preceding more reflective or individualized judgments.

In contrast, second and third adjectives showed no significant gender-based differences, which may reflect participants' increased reliance on personal associations or object familiarity as the task progressed. This pattern is in line with prior research demonstrating that grammatical gender activates early and unconsciously during language processing, even in the absence of explicit gender judgments (Boutonnet et al. 2012; Sato et al. 2020). Importantly, our use of a free description task, as opposed to explicit gender labeling or voice assignment, further

strengthens the conclusion that participants were not consciously focused on grammatical gender, yet it still guided their responses in systematic ways.

A particularly novel contribution of our study lies in the inclusion and treatment of neuter grammatical gender as an independent category. Although the post-hoc tests for neuter-gender nouns yielded no statistically significant differences, this outcome aligns with our hypothesis. Adjectives in the range (3,5] were defined as neutral traits, while [1,3] and (5,7] were associated with feminine and masculine traits, respectively. We anticipated – and confirmed – that neutral nouns would not differ meaningfully from either masculine or feminine nouns in gendered associations. The lack of significant contrast supports the interpretation that German speakers categorize neuter-gendered objects as conceptually neutral, rather than interpreting them as implicitly masculine or feminine by default. The lack of significance is thus meaningful: it reflects a psychologically real and conceptually distinct category. In this way, the grammatical structure of German appears to support a three-way gender conceptual framework, rather than forcing binary categorization. Moreover, participants' age and sex did not significantly influence the results, which suggests that these patterns are not driven by demographic variation, but likely reflect structural effects of the German language system itself.

This interpretation strengthens earlier findings from studies on three-gender languages such as Polish and Norwegian, which show that grammatical gender effects are not limited to binary gender systems (Beller et al. 2015; Boroditsky et al. 2003; Haertlé 2017; Maciuszek et al. 2019; Sedlmeier et al. 2016).

Compared to earlier studies that used explicit gender cues, such as assigning voices or requiring gender categorization (e.g., Almutrafi 2015; Athanasopoulos and Boutonnet 2016), our open-ended description task offered a more naturalistic context for uncovering the effects of grammatical gender. Importantly, we observed these effects without drawing participants' attention to gender, which enhances the ecological validity of our findings. That the gender alignment appeared primarily in the first adjective strongly suggests automatic activation of grammatical gender.

We also improved upon previous designs by including a larger and more diverse participant pool in the second phase of the study. Participants evaluated adjectives on a 7-point Likert scale, enabling a more fine-grained, continuous measurement of gender associations than binary tasks typically allow (e.g., Boroditsky et al. 2003; Sedlmeier et al. 2016). This methodological expansion increased both reliability and interpretive richness.

Most crucially, our treatment of neutral grammatical gender as a fully integrated category, rather than excluding it or collapsing it into binary groupings – offers a key theoretical contribution. Studies that do include neuter terms often force binary judgments (e.g., Bender et al. 2018), but our approach treated neuter gender as conceptually independent. The finding that neuter nouns elicited neutral trait ratings supports the view that German speakers treat them not as ambiguous or default categories but as conceptually distinct. This opens a path for cognitive-linguistic research to move beyond binary models of gender representation.

In theoretical terms, our findings suggest that grammatical gender systems do more than highlight male/female distinctions – they can also encode a third, non-binary conceptual category. This expands current understandings of linguistic relativity by demonstrating that gender effects are not inherently dichotomous and may instead reflect the specific structural affordances of a language. It challenges the assumption that gendered cognition always maps onto binary categories and supports a more typologically inclusive perspective.

5.1.2 Do Farsi native speakers differ from German native speakers in assigning properties to objects and animals?

The analysis demonstrates that Farsi monolingual speakers did not exhibit systematic differences in the adjectives they assigned based on the grammatical gender categories imposed by German. A significant effect emerged for feminine nouns in the model ($p = 0.021$). However, neutral nouns, adjective order, and the interaction between grammatical gender and adjective order were not significant. Post-hoc analyses of the Farsi data (Table 12), using the same procedure as for German, did not reveal any significant pairwise differences between grammatical gender conditions across adjective positions.

Although a main effect appeared in the overall model, this may reflect variability introduced during noun selection. However, the selection process was not guided by any intention to bias conceptual gender; as described in the Materials section, stimuli were chosen based on pre-established criteria unrelated to gender connotation. No consistent or interpretable differences emerged between gender groups. We interpret this as evidence that the gender structure of the German stimuli does not exert a systematic influence on adjective selection in Farsi. This contrasts with the German data, where grammatical gender significantly influenced the assignment of properties, particularly in the first adjective position. In Farsi, no significant main effects were found for participant sex or age, nor for the order of adjective responses. Furthermore, no interaction was observed between the imposed grammatical gender categories and adjective order. These findings support the hypothesis that grammatical gender, as a structural feature of language, exerts an influence on conceptualization, one that is absent when such a feature is not grammatically encoded.

Importantly, although German grammatical gender categories were artificially overlaid onto the Farsi dataset to ensure analytical comparability, the absence of corresponding effects underscores the role of grammatical gender as a genuine cognitive cue for German speakers. The imposed categories did not trigger differentiated responses among Farsi speakers, which reinforces the idea that these effects are not driven by universal semantic associations, but rather by the grammatical structures embedded in the speaker's native language. This lack of a grammatical gender effect in Farsi aligns with earlier findings in other genderless languages such as Tamil (Sedlmeier et al. 2016), further corroborating the structuralist interpretation of linguistic relativity.

The difference between the two language groups offers empirical support for the idea that speakers of languages with grammatical gender (such as German) develop automatized cognitive associations that subtly shape perception, even when gender is not explicitly attended to. In contrast, speakers of genderless languages (such as Farsi) are unlikely to spontaneously map objects to gendered traits in the absence of such cues. The fact that the Farsi participants' ratings remained stable across the adjective positions further supports the conclusion that their responses were not influenced by linguistic structure but were instead guided by individual semantic associations or subjective interpretations of the visual stimuli.

Moreover, while the German results showed evidence of a primacy effect (Asch 1946), where the first adjective aligned with grammatical gender more strongly, this pattern was absent in the Farsi data. This strengthens the claim that grammatical gender in German functions as a cognitive cue that becomes automatically activated early in the response process. In Farsi, where no such grammatical scaffolding exists, the response pattern lacked this early bias.

Together, these findings provide cross-linguistic evidence for the principle of linguistic relativity: structural features of language can shape habitual thought in ways that are both automatic and language-specific. The contrast between German and Farsi offers a compelling demonstration of how the presence or absence of grammatical gender affects conceptualization at a subtle level, especially in contexts where participants are not explicitly prompted to think in gendered terms. By placing these two languages at opposite ends of the grammatical gender spectrum (Fatemi 2023), the study reveals how deep-seated structural features can influence even simple descriptive tasks.

6 Conclusions

This study underscores the critical importance of developing robust methodological frameworks to investigate how grammatical gender influences cognition. By systematically including all grammatical gender categories – especially the often-overlooked neutral gender in German – our findings provide nuanced insights into the ways linguistic structures shape thought. Furthermore, expanding research to include understudied languages like Farsi enriches our understanding of linguistic relativity across diverse linguistic and cultural contexts.

Several limitations warrant consideration. While high-frequency words were carefully selected to maintain relevance, excluding animate nouns might have further refined the stimuli. Additionally, the shift between online and pen-and-paper data collection due to logistical constraints was managed carefully to preserve data integrity, although such procedural adjustments should be noted in interpreting the results.

Our findings contribute both theoretically and empirically. Theoretically, they reinforce and extend linguistic relativity by demonstrating how the presence or absence of grammatical gender influences conceptualization, highlighting a three-way gender framework in German and the absence of such effects in genderless Farsi. Empirically, these insights offer valuable implications for gender studies, language acquisition, and cognitive psychology, informing efforts to design gender-sensitive research, improve language teaching strategies, and deepen cross-cultural understanding of cognition.

Overall, this research advances interdisciplinary inquiry into the cognitive effects of grammatical structures, providing a foundation for future studies to explore the complex interactions between language, thought, and culture.

Acknowledgment: The consulting service provided by the Statistical Consulting and Analysis Department at the Centre for Higher Education at the TU Dortmund University was instrumental in the completion of this study. We also like to thank three anonymous reviewers for their constructive comments.
Author contributions: All authors have accepted responsibility for the content of this manuscript and approved its submission. MF and BM designed the experiments; MF conducted the experimental work. The initial statistical model was developed by MF in collaboration with the Statistical Consulting and Analysis Department at the Centre for Higher Education at TU Dortmund University. VS assisted with the statistical modeling incorporated into the revised manuscript following reviewer feedback. MF prepared the final manuscript with getting consultation from all co-authors.
Funding Information: Authors state no funding involved.
Conflict of interest: The authors declare no conflict of interest.
Data availability statement: The datasets generated and/or analyzed during the current study are available in the Open Science Framework repository and can be accessed at https://osf.io/qmsbh.

Appendix A: Survey instructions and test design

The following instructions were provided to participants at the start of the test:

Phase 1: German instructions:

„Liebe Teilnehmerinnen und Teilnehmer, Herzlich willkommen! In dieser Studie sollen Sie kurz Fragen beantworten, die sich auf darunterstehende Bilder beziehen. Für manche Fragen müssen Sie die Bilder beschreiben, indem Sie spontan die ersten drei Adjektive aufschreiben, die Ihnen in den Sinn kommen. In anderen Fragen werden Sie gebeten, spontan die Uhrzeit anzugeben. Ihre Antworten werden streng vertraulich behandelt. Haben Sie herzlichen Dank für Ihre Zeit und Ihr Interesse! Um die Umfrage zu öffnen, akzeptieren Sie bitte unsere Datenschutzerklärung."

Translation:

Dear participants, welcome! In this study, you will be asked to briefly answer questions relating to the images below. For some questions, you will describe the images by spontaneously writing down the first three adjectives that come to mind. In other questions, you will be asked to spontaneously indicate the time. Your answers will be kept strictly confidential. Thank you very much for your time and interest! To begin the test, please click on the provided link to confirm your participation and proceed with the tasks.

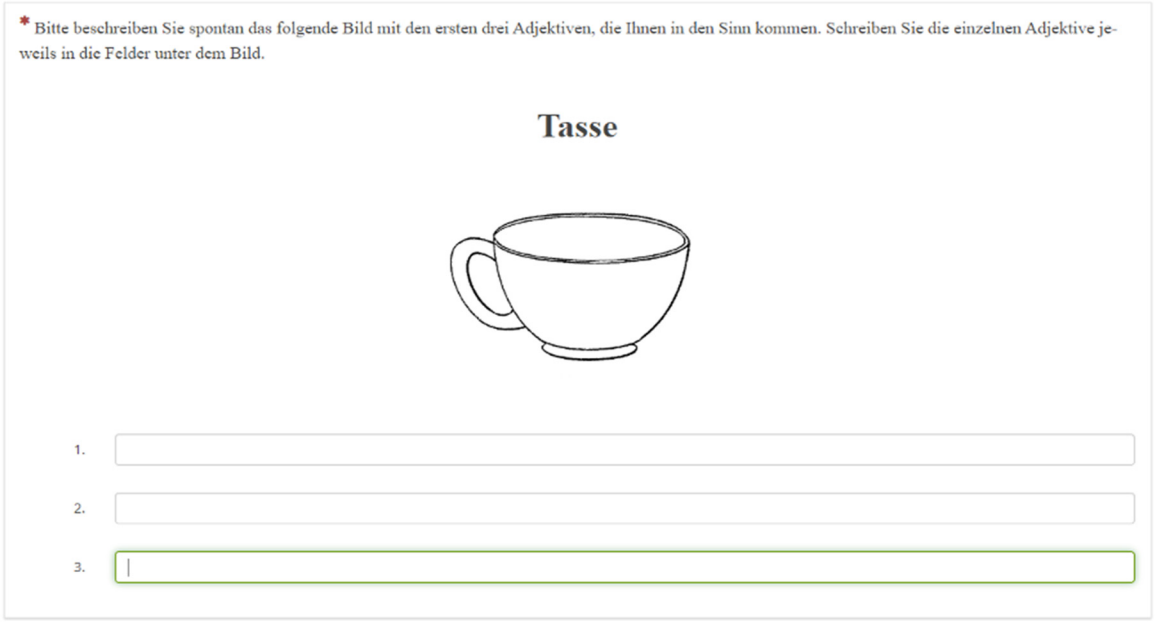
Test prompt (after instructions):

Participants were guided to the main task with the following prompt:

„Bitte beschreiben Sie spontan das folgende Bild mit den ersten drei Adjektiven, die Ihnen in den Sinn kommen. Schreiben Sie die einzelnen Adjektive jeweils in die Felder unter dem Bild.“

Translation:

Please describe the following image spontaneously with the first three adjectives that come to mind. Write each adjective in the fields below the image.



Phase 2: Gender assignment

“Liebe Teilnehmerinnen und Teilnehmer, in dieser Phase bitten wir Sie, den folgenden Adjektiven ein Geschlecht zuzuordnen. Bewerten Sie das Adjektiv auf einer Skala von 1 bis 7:1 = sehr feminin, 7 = sehr maskulin.”

Translation

Dear participants, in this phase, we ask you to assign a gender to the following adjectives. Please rate each adjective on a scale from 1 to 7: 1 = very feminine, 7 = very masculine.



Appendix B

See Tables A1–A3.

Table A1: List of feminine-stimuli item and images.

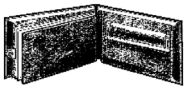
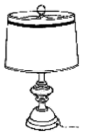
1	Ente	
2	Taube	
3	Fliege	
4	Schildkröte	
5	Blume	
6	Börse	
7	Tasse	
8	Lampe	
9	Gruppe	
10	Uniform	

Table A2: List of masculine-stimuli item and images.

1	Hund	
2	Zeh	
3	Pinguin	

Table A2: (continued)

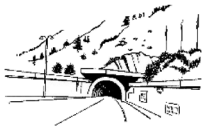
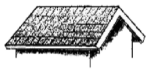
4	Ball	
5	Bus	
6	Fuß	
7	Tunnel	
8	Brunnen	
9	Zaun	
10	Globus	

Table A3: List of neutral-stimuli item and images.

1	Huhn	
2	Eichhörnchen	
3	Ziel	
4	Herz	
5	Feuer	
6	Netz	

Table A3: (continued)

7	Dach	
8	Geschenk	
9	Gepäck	
10	Getränk	

Appendix C

See Table A4.

Table A4: Frequency class table by the psycholinguistics laboratory of TU Dortmund University using Ruhr University Bochum's method.

Frequency class	Min	Max
0	230188366	325535509
1	115094183	230188365
2	57547092	115094182
3	28773546	57547091
4	14386773	28773545
5	7193387	14386772
6	3596694	7193386
7	1798347	3596693
8	899174	1798346
9	449587	899173
10	224794	449586
11	112397	224793
12	56199	112396
13	28100	56198
14	14050	28099
15	7025	14049
16	3513	7024
17	1757	3512
18	879	1756
19	440	878
20	220	439
21	110	219
22	55	109
23	28	54
24	14	27
25	7	13
26	4	6
27	2	3
28	1	1

Appendix D

Lexeme-specific variability in both the German and Farsi datasets (Figures 2–2.5).

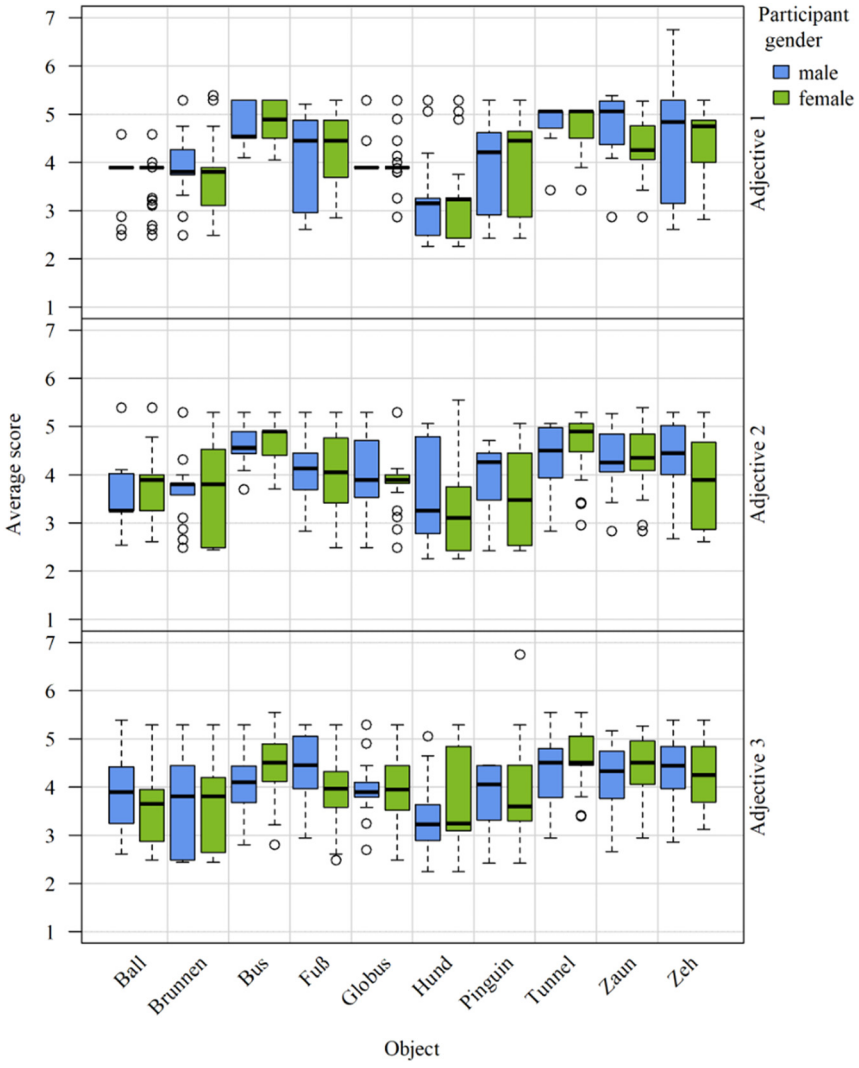


Figure 2: Boxplots of the mean adjective scores separated by participant gender and repetition for the M-stimuli objects (German participants).

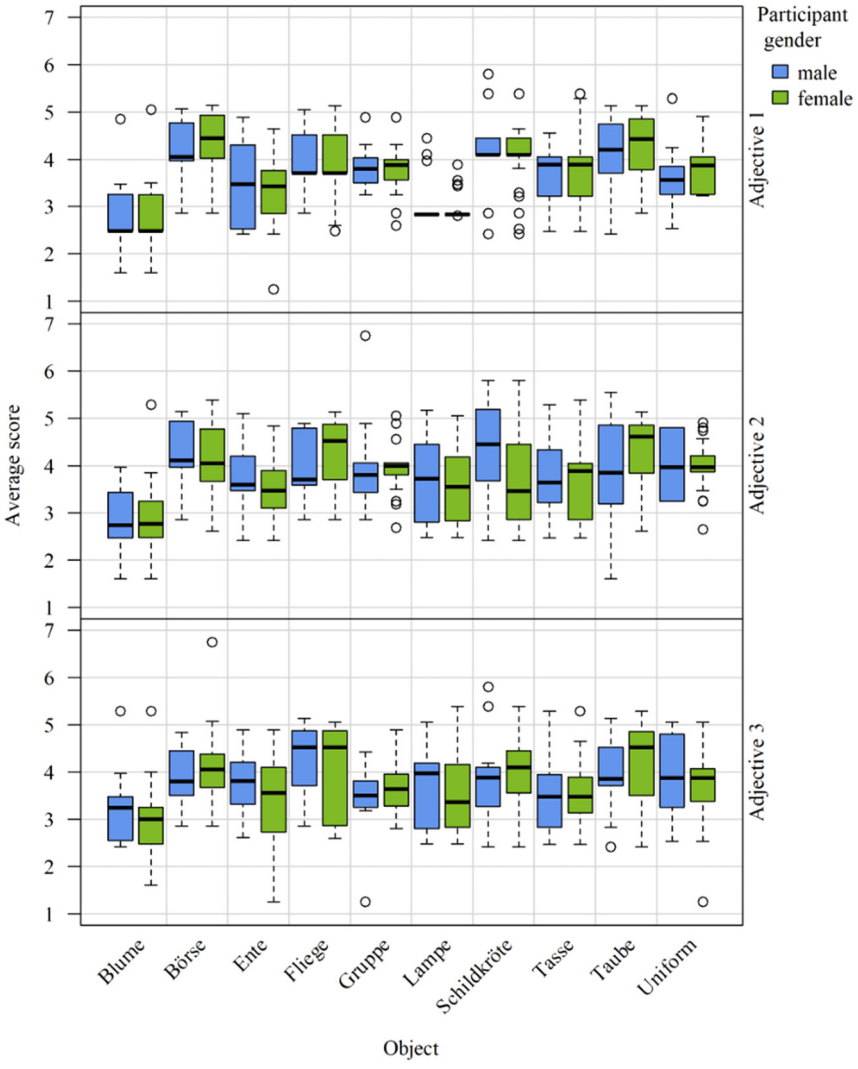


Figure 2.1: Boxplots of mean adjective scores by participant gender and repetition for F-stimuli objects (German participants).

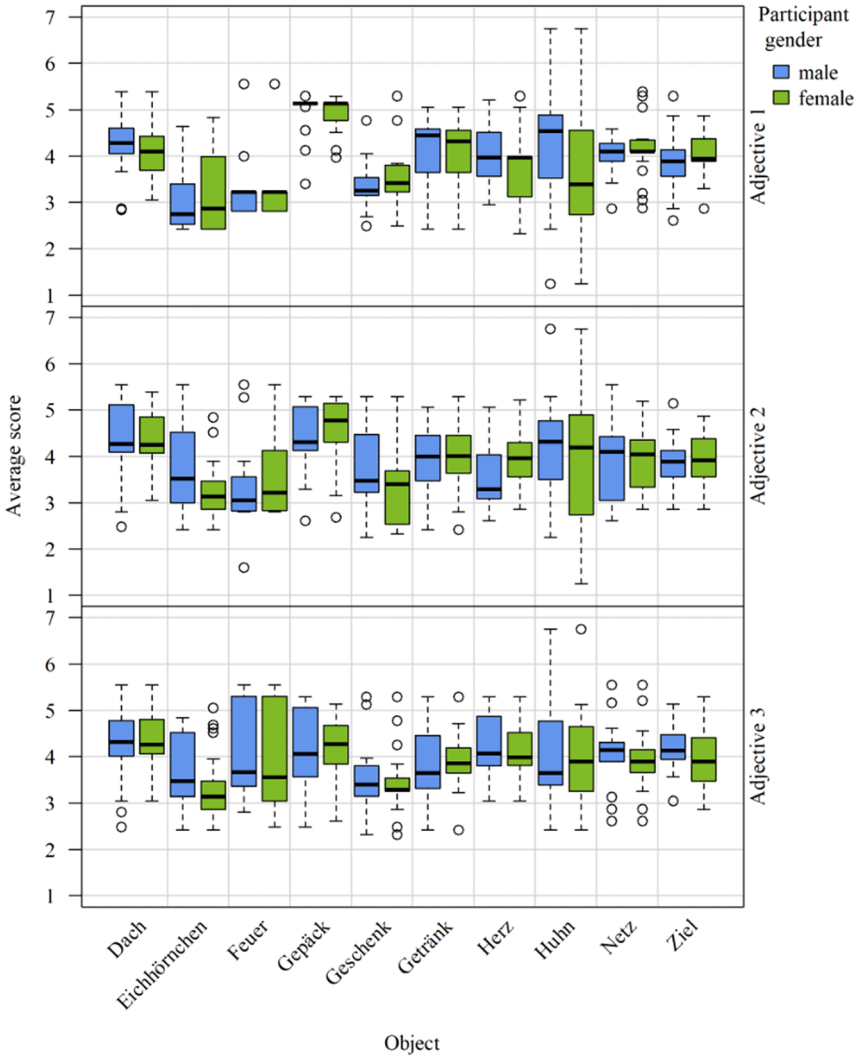


Figure 2.2: Boxplots of mean adjective scores by participant gender and repetition for N-stimuli objects (German participants).

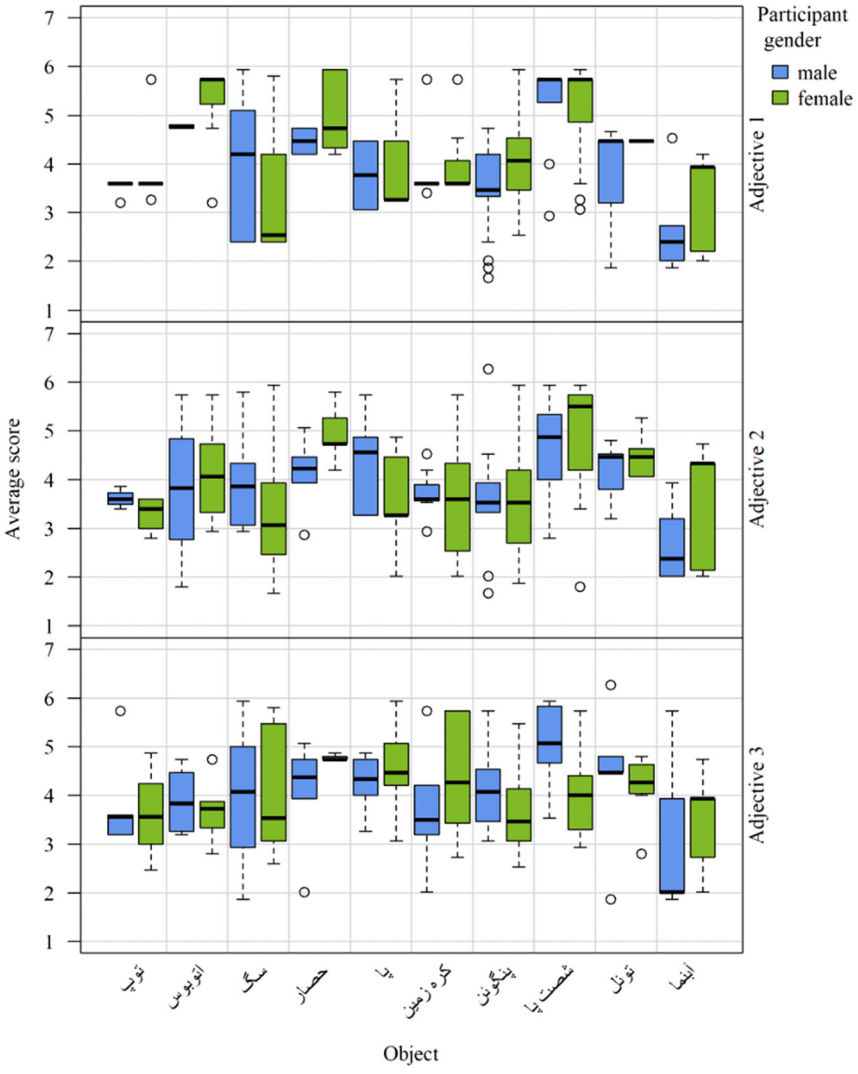


Figure 2.3: Boxplots of the mean adjective scores by participant gender and adjective order for the M-stimuli objects, based on German grammatical gender (Farsi participants). Word translations from left to right: ball, bus, dog, fence, foot, globe, penguin, toe, tunnel, fountain.

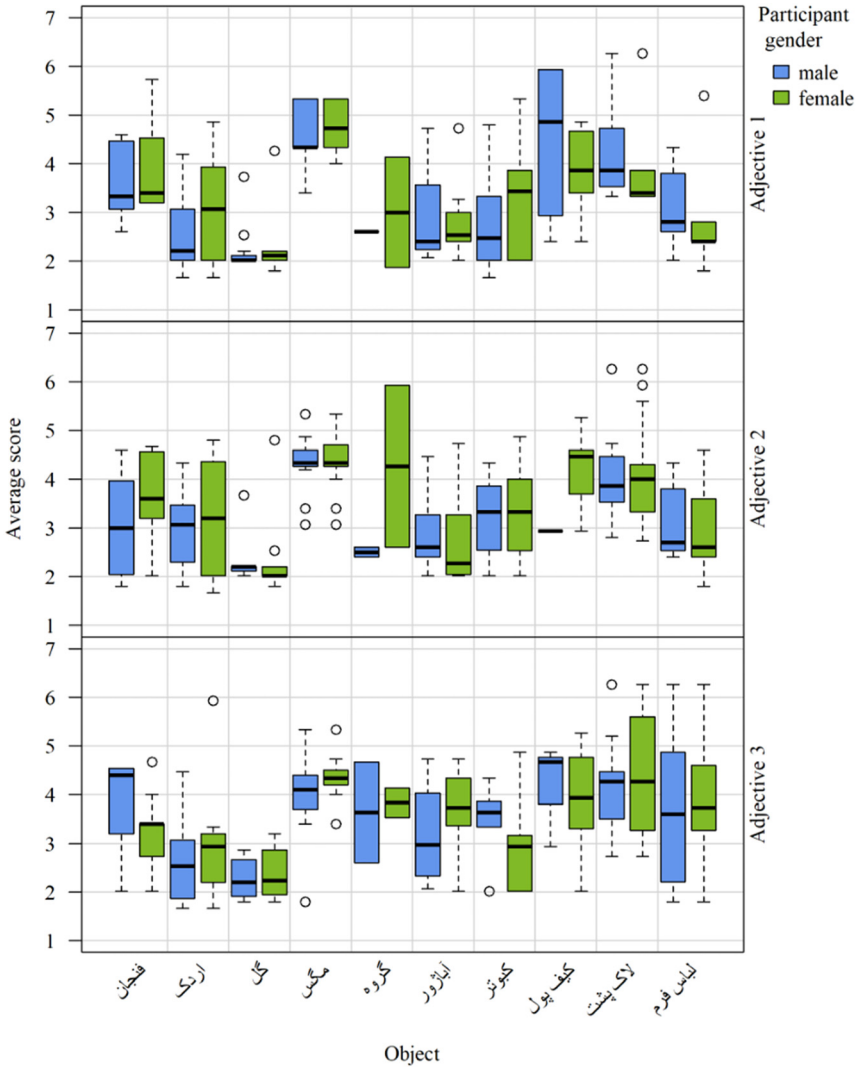


Figure 2.4: Boxplots of the mean adjective scores by participant gender and adjective order for the F-stimuli objects, based on German grammatical gender (Farsi participants). Word translations from left to right: cup, duck, flower, fly, group, Lamp, pigeon, wallet, turtle, uniform.

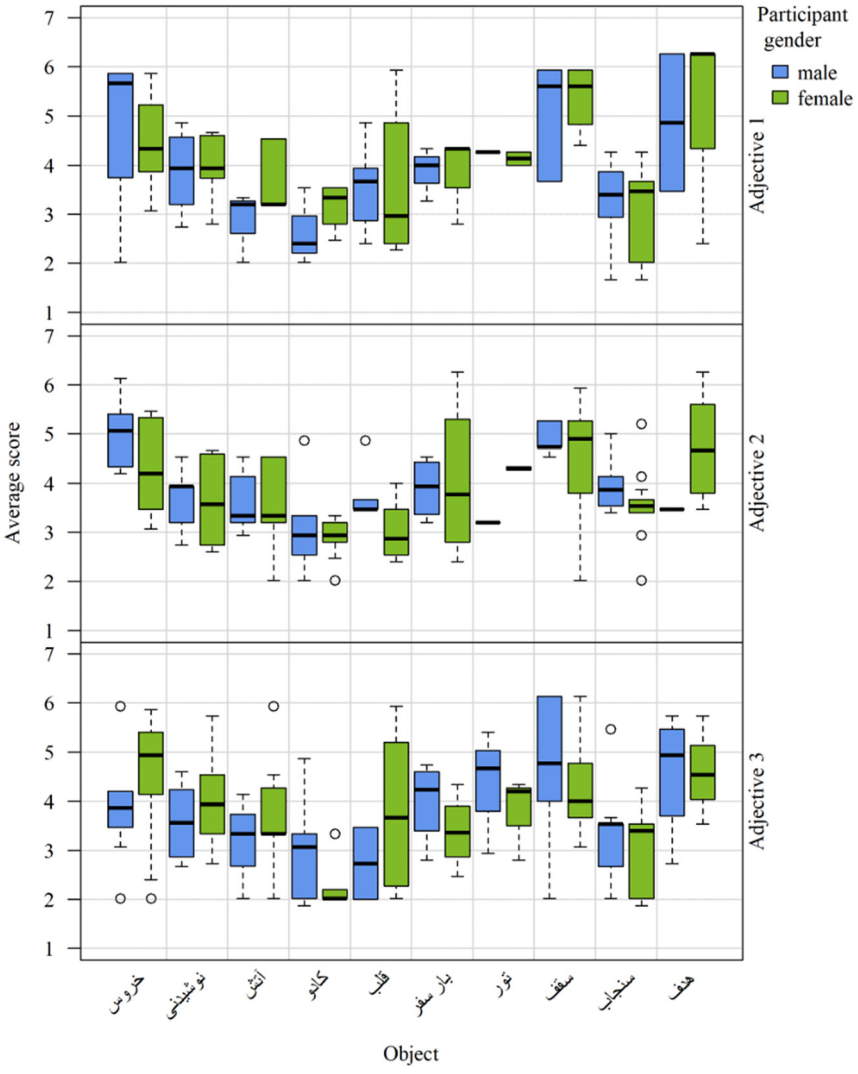


Figure 2.5: Boxplots of the mean adjective scores by participant gender and adjective order for the N-stimuli objects, based on German grammatical gender (Farsi participants). Word translations from left to right: rooster, drink, fire, gift, heart, luggage, net, roof, squirrel, goal.

References

Ahamadi-Givi, Hassan. 2004. *Comparative grammar of Turkish and Persian*. Tehran: Ghatreh (In Persian).

Almutrafi, Fatimah. 2015. *Language and cognition: Effects of grammatical gender on the categorization of objects*. Newcastle: Newcastle University PhD diss.

Asch, Solomon E. 1946. Forming impressions of personality. *Journal of Abnormal and Social Psychology* 41(3). 258–290.

Athanasopoulos, Panos. 2006. Effects of the grammatical representation of number on cognition in bilinguals. *Bilingualism: Language and Cognition* 9(1). 89–96.

Athanasopoulos, Panos. 2009. Cognitive representation of colour in bilinguals: The case of Greek blues. *Bilingualism: Language and Cognition* 12(1). 83–95.

Athanasopoulos, Panos & Bastien Boutonnet. 2016. Learning grammatical gender in a second language changes categorization of inanimate objects: Replications and new evidence from English learners of L2 French. In R. Alonso Alonso (ed.), *Crosslinguistic influence in second language acquisition*, 173–192. *Second Language Acquisition* S.

Athanasopoulos, Panos & Emanuel Bylund. 2013. Does grammatical aspect affect motion event cognition? A cross-linguistic comparison of English and Swedish speakers. *Cognitive Science* 37(2). 286–309.

Bassetti, Benedetta. 2007. Bilingualism and thought: Grammatical gender and concepts of objects in Italian-German bilingual children. *International Journal of Bilingualism* 11(3). 251–273.

Beller, Sieghard, Karen Fadnes Brattebø, Kristina Osland Lavik, Rakel Drønen Reigstad & Andrea Bender. 2015. Culture or language: What drives effects of grammatical gender? *Cognitive Linguistics* 26(2). 331–359.

Bender, Andrea, Sieghard Beller & Karl Christoph Klauer. 2011. Grammatical gender in German: A case for linguistic relativity? *Quarterly Journal of Experimental Psychology* 64(9). 1821–1835.

Bender, Aandrea, Sieghard Beller & Karl Christoph Klauer. 2016a. Crossing grammar and biology for gender categorizations: Investigating the gender congruency effect in generic nouns for animates. *Journal of Cognitive Psychology* 28(5). 530–558.

Bender, Aandrea, Sieghard Beller & Karl Christoph Klauer. 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.

Bender, Andrea, Sieghard Beller & Karl C. Klauer. 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). https://doi.org/10.1037/xlm0000534.

Boroditsky, Lera. 2001. Does language shape thought?: Mandarin and English speakers' conceptions of time. *Cognitive Psychology* 43(1). 1–22.

Boroditsky, Lera. 2011. How language shapes thought. *Scientific American* 304(2). 62–65.

Boroditsky, Lera, Laurren A. Schmidt & Webb Phillips. 2003. Sex, syntax, and semantics. In Dedre Gentner & Susan Goldin-Meadow (eds.), *Language in mind*, 61–80. The MIT Press.

Boutonnet, Bastien, Panos Athanasopoulos & Guilaume Thierry. 2012. Unconscious effects of grammatical gender during object categorization. *Brain Research* 1479. 72–79.

Comrie, Bernard. 1989. *Language universals and linguistic typology: Syntax and morphology*. Chicago: University of Chicago Press.

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

Dolscheid, Sarah, Shakila Shayan, Asif Majid & Daniel Casasanto. 2013. The thickness of musical pitch: Psychophysical evidence for linguistic relativity. *Psychological Science* 24(5). 613–621.

Dudenredaktion. 2005. *Duden: Die Grammatik 8*. Aufl. Mannheim, Wien, Zürich: Dudenverlag.

Duke, Janet. 2005. Gender systems and grammaticalization: Examples from German and Germanic. In Torsten Leuschner, Tanja Mortelmans & Sarah De Groodt (eds.), *Grammatikalisierung im Deutschen*, 31–58. Berlin, New York: De Gruyter.

Duranti, Alessandro. 2011. Linguistic anthropology: Language as a non-neutral medium. *The Cambridge handbook of sociolinguistics*. 28–46. https://doi.org/10.1017/CBO9780511810190.

Elpers, Nan, Greg Jensen & Kevin J. Holmes. 2022. Does grammatical gender affect object concepts? Registered replication of Phillips and Boroditsky 2003. *Journal of Memory and Language* 127. 104357.

Esmaili, Mohammad Mehdi. 2023. A Study of gender agreement in Abyānaٔī Dialect. *Elm-e Zabân Language Science* 245. https://doi.org/10.22054/LS.2023.73653.1579.

Fahrmeir, Ludwig, Thomas Kneib, Stefan Lang & Brian Marx. 2013. *Regression: Models, methods and applications*. Berlin, Heidelberg: Springer Berlin Heidelberg.

Fatemi, Maryam. 2023. Synchronic overview of grammatical gender marking in Indo-European languages. In Barbara Mertins & Renate Delucchi Danhier (eds.), *Diversity in cognition*, 13–60. Berlin: Peter Lang Group AG.

Flaherty, Mary. 2001. How a Language gender System creeps into perception. *Journal of Cross-Cultural Psychology* 32(1). 18–31.

Forbes, James N., Diane Poulin-Dubois, Magda R. Rivero & Maria D. Sera. 2008. Grammatical gender affects bilinguals' conceptual gender: Implications for linguistic relativity and decision making. *The Open Applied Linguistics Journal* 1(1). 68–76.

Gentner, Dedre & Meadow Goldin. 2003. *Language in mind: Advances in the study of language and thought*. Cambridge: MIT Press.

Gumperz, John. 1982. *Discourse strategies*. Cambridge: Cambridge University Press.

Gumperz, John J. & Stephan C. Levinson. 1991. Rethinking linguistic relativity. *Current Anthropology* 32(5). S. 613–623.

Haertlé, Izabella. 2017. Does grammatical gender influence perception? A study of Polish and French speakers. *Psychology of Language and Communication* 21(1). 386–407.
Haspelmath, Martin, Matthew S. Dryer, David Gil & Bernard Comrie. 2005. *The world atlas of language structures*. Oxford University Press. Available at: https://katalog.ub.tu-dortmund.de/id/ir01388a:ubd.lobid:990133519940206441.
Hockett, Charles F. 1960. *A course in modern linguistics*. New York: Macmillan.
Holm, Sture. 1979. A simple sequentially rejective multiple Test procedure. *Scandinavian Journal of Statistics* 6(2). 65–70.
Imai, Muttsumi, Lennart Schalk, Henrik Saalbach & Hiroyuki Okada. 2014. 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.
Khayyampour, Abbas Reza. 1998. *Persian grammar 10th edition*. Tabriz: Tehran Press (In Persian).
Konishi, Toshi. 1993. The semantics of grammatical gender: A cross-cultural study. *Journal of Psycholinguistic Research* 22(5). 519–534.
Kousta, Stavroula-Thaila, David P. Vinson & G. Vigliocco. 2008. Investigating linguistic relativity through bilingualism: The case of grammatical gender. *Journal of Experimental Psychology: Learning, Memory, and Cognition* 34(4). 843–858.
Kutas, Marta & Stevan A. Hillyard. 1984. Brain potentials during reading reflect word expectancy and semantic association. *Nature* 307(5947). 161–163.
Landor, Roland Viktor. 2014. *Grammatical categories and cognition across five languages: The case of grammatical gender and its potential effects on the conceptualization of objects*, 46–47. Brisbane, Australia: Griffith University PhD diss.
Langacker, Roland W. 1987. *Foundations of cognitive grammar: Volume I: Theoretical prerequisites*, Bd. 1. Stanford University Press. Available at: https://katalog.ub.tu-dortmund.de/id/ir01388a:ubd.lobid:990083560210206441.
Lazard, Gilbert & Herbert H. Paper. 1959. Grammaire du Persan Contemporain. *Journal of the American Oriental Society* 79(1). 31.
Lucy, John A. 1992. *Language diversity and thought: A reformulation of the linguistic relativity Hypothesis*. Cambridge: Cambridge University Press.
Lucy, John A. 2016. Recent advances in the Study of linguistic relativity in historical context: A critical assessment. *Language Learning* 66(3). 487–515.
Maciuszek, Josef, Mateuz Polak & Natalia Świątkowska. 2019. Grammatical gender influences semantic categorization and implicit cognition in polish. *Frontiers in Psychology* 10. 2208.
Malt, Barbara C. 2010. *Words and the mind: How words capture human experience*. New York: Oxford University Press.
Mertins, Barbara. 2018. *Sprache und Kognition: Ereigniskonzeptualisierung im Deutschen und Tschechischen*. Berlin, Germany: De Gruyter.
Mickan, Anne, Maren Schiefke & Anatol Stefanowitsch. 2014. Key is a llave is a Schlüssel: A failure to replicate an experiment from Boroditsky et al. 2003. *Yearbook of the German Cognitive Linguistics Association* 2(1). https://doi.org/10.1515/gcla-2014-0004.
Montefinese, Maria, Ettore Ambrosini & Eka Roivainen. 2019. No grammatical gender effect on affective ratings: Evidence from Italian and German languages. *Cognition & Emotion* 33(4). 848–854.
Phillips, Webb & Lera Boroditsky. 2003. Can quirks of grammar affect the way you think? Grammatical gender and object concepts. In *Proceedings of the 25th annual cognitive science society*, 928–933. Psychology Press. https://escholarship.org/uc/item/31t455gf.
Pinker, Steven. 1994. On language. *Journal of Cognitive Neuroscience* 6(1). 92–98.
Prewitt-Freilino, Jennifer L., Andrew T. Caswell & Emmi K. Laakso. 2012. 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.
Roberson, Debi, Hyensou Pak & Richard J. Hanley. 2008. Categorical perception of colour in the left and right visual field is verbally mediated: Evidence from Korean. *Cognition* 107(2). 752–762.
Sagot, Bonit & Geraldine Walther. 2010. A morphological lexicon for the Persian language. In *Proceedings of the 7th language resources and evaluation conference LREC'10*. https://shs.hal.science/halshs-00751629.
Samuel, Steven, Geoff Cole & M. J. Eacott. 2019. Grammatical gender and linguistic relativity: A systematic review. *Psychonomic Bulletin & Review* 26(6). 1767–1786.
Sato, Sayako & Panos Athanasopoulos. 2018. Grammatical gender affects gender perception: Evidence for the structural-feedback hypothesis. *Cognition* 176. 220–231.
Sato, Sayako, Aina Casaponsa & Panos Athanasopoulos. 2020. Flexing gender perception: Brain potentials reveal the cognitive permeability of grammatical information. *Cognitive Science* 44(9). e12884.
Sedlmeier, Peter, Arun Tipandjan & Anastasia Jänchen. 2016. How persistent are grammatical gender effects? The case of German and Tamil. *Journal of Psycholinguistic Research* 45(2). 317–336.
Sera, Maria D., Christian A. H. Berge & Javier Del Castillo Pintado. 1994. Grammatical and conceptual forces in the attribution of gender by English and Spanish speakers. *Cognitive Development* 9(3). 261–292.
Sera, M. D., Chryle Elieff, James Forbes, Melissa Clark Burch, Wanda Rodríguez & Diane Poulin Dubois. 2002. When language affects cognition and when it does not: An analysis of grammatical gender and classification. *Journal of Experimental Psychology: General* 131(3). 377–397.
Slobin, Dan. 1996. From 'thought and language' to 'thinking for speaking'. In John Joseph Gumperz & Stephen C. Levinson (eds.), *Rethinking linguistic relativity*. Cambridge University Press. https://psycnet.apa.org/record/1996-98701-002.
Stutterheim, Cchristina Von & Ralf Nüse. 2003. Processes of conceptualization in language production: Language-specific perspectives and event construal. *Linguistics* 41(5). https://doi.org/10.1515/ling.2003.028.
Tahal, Karel. 2010. *A grammar of Czech as a foreign language*. Factum CZ, sro. https://dl.melalinstitute.com/cheki/czech-grammar.pdf.
Vigliocco, Gabiella, David P. Vinson, Fedrica Paganelli & Katharina Dworzynski. 2005. Grammatical gender effects on cognition: Implications for language learning and language use. *Journal of Experimental Psychology: General* 134(4). 501–520.
Whorf, Benjamin L. 1956. In John B. Carroll (ed.), *Language, thought, and reality: Selected writings*. Cambridge, Massachusetts: Technology Press of Massachusetts.

Grammatical Gender in a Grammatically Genderless Mind: Eye-Tracking Evidence from Farsi-German Bilinguals

Maryam Fatemi, Renate Delucchi Danhier, Barbara Mertins

Abstract

This study extends research on linguistic relativity by examining whether German grammatical gender influences real-time language processing in monolingual and bilingual speakers. Three groups of 99 adults (German monolinguals, Farsi–German bilinguals tested in German, and a subset of Farsi–German bilinguals tested in Farsi; n = 46, 53, and 36, respectively) participated in a visual world eye-tracking experiment. Participants completed 30 trials in which they selected the adjective that best matched a spoken noun from four options (critical, two competitors, distractor). We also assessed participants' degree of bilingualism. Mixed-effects modelling of dwell time revealed that German monolinguals showed robust gender-congruency effects, with greater attention to gender-congruent adjectives. Farsi–German bilinguals exhibited similar but attenuated effects in German, but no such effects in Farsi. These findings suggest that grammatical gender guides visual attention during online processing, can be integrated in a later-acquired language, and is modulated by the linguistic context.

Keywords: grammatical gender; linguistic relativity; bilingualism; eye-tracking; Visual World Paradigm; real-time language processing

1. Introduction

The influence of language on thought, as proposed by the linguistic relativity principle (Whorf, 1956), has long been a central topic in cognitive science and linguistics. Early debates focused on whether linguistic structures shape non-linguistic cognition, often in opposition to universalist perspectives that emphasize the independence of conceptual representations from language (Pinker, 1994). More recent research has converged on a moderated view, suggesting that specific grammatical features—such as gender—can bias how speakers attend to, categorize, and conceptualize the world without determining cognition in an absolute sense (Boroditsky et al., 2003; Samuel et al., 2019; Sato & Athanasopoulos, 2018).

In languages such as German, nouns—including those referring to inanimate objects—are obligatorily marked for grammatical gender (masculine, feminine, or neuter), often without transparent semantic motivation (Corbett, 1991). For example, the German noun *Tasche* (‘bag’) is grammatically feminine despite lacking inherent associations with femininity. By contrast, languages such as Farsi do not encode grammatical gender. This typological contrast is particularly relevant for bilingual cognition, raising the question of whether grammatical gender can influence conceptualization across co-existing gendered and non-gendered linguistic systems.

Crucially, the present study adopts a dynamic view of bilingualism that moves beyond fixed L1–L2 distinctions. Following Grosjean’s (2001, 2008) language mode framework, both languages remain active, with relative activation varying as a function of context. Language processing is therefore context-dependent, with the relative influence of each language varying across situations (De Bot et al., 2007). While the role of age of acquisition remains debated in bilingualism research, and the critical period hypothesis has been increasingly challenged (e.g., Birdsong, 2018; Bylund & Norrman, 2023; Caldwell-Harris & MacWhinney, 2023; Delucchi Danhier & Mertins, 2018; Flege, 2018; Grotjahn, 2016;

Singleton, 2023; Vanhove, 2014), the present study characterizes participants in terms of length of exposure, proficiency, and language use—factors shown to modulate bilingual processing (Consonni et al., 2013)—rather than fixed L1–L2 categories. In this sample, German was acquired at a later stage of life (mean age of acquisition = 24.5 years); however, this characterization is used descriptively and does not imply that age of acquisition is the primary determinant of processing outcomes. Instead, variation in language experience, proficiency, and use is assumed to underlie differences in bilingual processing. The experimental conditions (German vs. Farsi) are interpreted as inducing different language contexts, which modulate the activation of linguistic representations during processing.

Most empirical evidence on grammatical gender and cognition comes from monolingual speakers, demonstrating that speakers of gendered languages often conceptualize objects in ways that align with grammatical gender categories (Boroditsky et al., 2003; Cubelli et al., 2011; Elpers et al., 2022; Sera et al., 2002). However, considerably less is known about how such effects operate in bilingual speakers whose languages differ in whether they encode grammatical gender. The present study bridges this gap by investigating whether grammatical gender influences online language processing in Farsi–German bilingual speakers. Furthermore, it employs a novel eye-tracking approach based on the Visual World Paradigm, enabling the examination of the subtle, implicit effects of grammatical gender as they unfold in real time.

Although there is a limited number of studies that employed eye-tracking methodologies to investigate grammatical gender processing, these have primarily focused on lexical prediction and agreement during online comprehension (Fuchs, 2022; Riordan et al., 2015) rather than on the influence of grammatical gender on conceptualization. As a result, eye-tracking research on grammatical gender within the linguistic relativity framework remains neglected, particularly in bilingual populations.

Within the emerging literature on bilingualism in linguistic relativity (Athanasopoulos & Bylund, 2013; Bassetti & Nicoladis, 2016; Samuel et al., 2019; Sato & Athanasopoulos, 2018; Sato et al., 2020), studies on bilingual speakers whose languages differ in grammatical gender marking remain relatively limited. Existing work has predominantly focused on bilinguals whose early-acquired language encodes grammatical gender (e.g., German) and whose later-acquired language does not (e.g., English), suggesting that gender-based effects may persist even when processing occurs in a language without grammatical gender (Bassetti & Nicoladis, 2016). In contrast, less is known about bilingual configurations in which one language lacks grammatical gender (e.g., Farsi) while the other encodes it (e.g., German). The present study addresses this gap by examining whether grammatical gender can influence online language processing in Farsi–German bilingual speakers, thereby contributing to a more comprehensive understanding of how linguistic structure shapes cognition across underexplored bilingual contexts.

Empirical findings on grammatical gender in bilingual speakers remain mixed: some studies report persistent gender-congruency effects, whereas others find reduced or absent effects. This variability appears to depend on factors such as language pairing, task design, and degree of immersion in the gender-marking language (Lambelet, 2016; Sato & Athanasopoulos, 2018). At the same time, bilingual experience itself is heterogeneous. Factors such as length of immersion, proficiency, and patterns of language use have been shown to modulate bilingual processing (Grosjean, 2010; Luk & Bialystok, 2013). Understanding how bilingual speakers process grammatical gender, therefore, requires careful consideration of these factors, particularly in relation to differences in exposure, including its length, intensity, and quality, which may shape how strongly gender-related information is integrated into online processing.

The present study provides the first eye-tracking evidence on grammatical gender processing in Farsi–German bilinguals across two language contexts. It examines whether

bilingual speakers with a relatively shorter length of exposure to German, whose other language (Farsi) does not encode grammatical gender, internalize and cognitively engage with the German gender system. Critically, bilingual processing is examined across two language contexts in order to test whether such effects are language-dependent or reflect more stable conceptual representation. In one condition, bilingual participants perform the task in German, allowing us to examine whether grammatical gender acquired through later language experience guides real-time visual attention during lexical processing. In a second condition, a subset of participants performs the same task in Farsi, enabling us to test whether such effects persist in the absence of grammatical gender marking. This experimental sequence directly follows from linguistic relativity accounts, which predict that language-specific structures may influence cognition in a context-dependent manner (Athanasopoulos & Bylund, 2013; Sato & Athanasopoulos, 2018).

To contextualize bilingual performance, the study also includes a group of German monolingual speakers. This comparison allows us to assess whether bilingual speakers engage grammatical gender in a manner comparable to native-like processing. Using the visual-world eye-tracking paradigm, the study examines how participants allocate visual attention to gender-congruent and gender-incongruent adjectives during spoken language processing. By combining cross-linguistic comparison using eye-tracking, the study provides direct evidence on whether grammatical gender effects in bilinguals are contingent on the currently active language context or reflect more stable patterns of conceptual representation.

Based on these objectives, the present study investigates how grammatical gender influences real-time visual attention in monolingual and bilingual speakers. Specifically, the study engages with the following research questions:

1. Do German monolingual speakers allocate greater visual attention (dwell time) to gender-congruent adjectives than to gender-incongruent competitors and distractors during spoken language processing?

2. Do late Farsi–German bilinguals allocate greater visual attention to gender-congruent adjectives when performing the task in German, indicating sensitivity to grammatical gender acquired through later language experience?

3. When the task is performed in Farsi, a language without grammatical gender, are these gender-related visual attention patterns reduced or absent?

4. Do German monolingual speakers and bilingual speakers tested in German differ in the extent to which grammatical gender guides visual attention during spoken language processing?

5. Do bilingual speakers show differences in gender-related visual attention patterns when performing the task in German compared to Farsi?

By answering these questions, the study contributes to understanding how structural features of language influence cognitive processing in bilingual speakers, particularly when the languages involved differ fundamentally in their grammatical architecture. In addition, it establishes a methodological precedent for using eye-tracking to investigate implicit, real-time effects of grammatical gender during language comprehension.

2. Theoretical framework

The linguistic relativity hypothesis (LRH) proposes that structural features of language can influence cognitive processes (Lucy, 1992; Whorf, 1956). Contemporary approaches adopt a moderated view, suggesting that language guides attention and habitual patterns of thought without determining cognition (Boroditsky et al., 2003; Levinson, 2003; Sato & Athanasopoulos, 2018). Grammatical gender has been one of the most extensively investigated domains in this context, providing a test case for examining how morphosyntactic categories may shape conceptualization (Bassetti & Nicoladis, 2016; Lambelet, 2016; Samuel et al., 2019).

Cross-linguistic variation in grammatical gender provides a useful framework for examining how speakers process and conceptualize gender-related information. This variation is particularly informative in bilinguals, who manage linguistic systems that differ in whether and how gender is encoded. The present study focuses on Farsi–German bilinguals, combining a genderless language (Farsi) with a three-gender system (German). This pairing allows us to directly investigate how gender-related representations emerge when one language encodes grammatical gender and the other does not, and how these systems interact within a bilingual cognitive system.

A key challenge in linguistic relativity research concerns the precise operationalization of cognition. Cognition is often used broadly to refer to constructs such as thought, perception, or worldview (Chen & Faitaki, 2024; Flecken et al., 2015). In empirical contexts, however, it must be defined in terms of measurable processes, including attention, categorization, perception, and conceptualization (Langacker, 1987). In the present study, cognition is examined through perceptual attention and categorization as reflected in real-time visual processing. Eye-tracking provides high temporal resolution and allows moment-by-moment measurement of visual attention during language comprehension, making it well suited to detecting subtle, context-dependent effects of language on processing (Rayner, 1998).

2.1 Linguistic relativity in a bilingual mind

In bilingual speakers, linguistic relativity is not assumed to operate uniformly but to be shaped by the interaction of co-existing languages (Bassetti & Filipović, 2022). Evidence shows that bilinguals can exhibit language-specific patterns of attention and categorization depending on the language in use, while also showing cross-linguistic influence across contexts (Athanasopoulos & Bylund, 2013; Bassetti & Filipović, 2022; Bassetti & Nicoladis,

2016). This reflects both flexibility and partial integration within the bilingual cognitive system.

These findings suggest that linguistic relativity in bilinguals is characterized by both flexibility and interaction: speakers may shift between language-specific cognitive patterns while also showing transfer or convergence across languages. Such variability is modulated by factors including proficiency, age of acquisition, and language use (Grosjean, 2001, 2008).

Theoretical accounts provide mechanisms for these effects. Slobin's (1996) *Thinking for Speaking* hypothesis proposes that linguistic structures shape cognition during speech planning, while Mertins' (2018) *Seeing for Speaking* framework extends this influence to visually grounded tasks. Together, these accounts emphasize that linguistic effects are task- and context-dependent, emerging during online processing rather than reflecting fixed conceptual restructuring.

When languages differ in fundamental grammatical features—such as the presence versus absence of grammatical gender—these dynamics become particularly salient. Gender-based effects may emerge when a gender-marking language is active, may be reduced in a non-gendered context, or may persist as cross-linguistic influence. Such patterns are further modulated by factors including proficiency, length of exposure, and language use (Grosjean, 2001, 2008). This variability highlights the importance of examining linguistic relativity in bilinguals using methods that capture real-time processing.

2.2 Grammatical gender and bilingualism

Grammatical gender is a morphosyntactic feature that classifies nouns into categories such as masculine, feminine, or neuter (Corbett, 1991). In gender-marking languages such as Spanish, German, and French, this system has been shown to influence how speakers attribute properties to inanimate objects. Behavioral studies using property attribution and voice assignment tasks demonstrate that speakers tend to associate nouns with gender-congruent

semantic features (e.g., masculine nouns with stereotypically “masculine” attributes) (Boroditsky et al., 2003; Sato & Athanasopoulos, 2018). In contrast, speakers of languages without grammatical gender, such as Farsi and Tamil, typically do not exhibit systematic gender-based conceptual biases in comparable tasks (Prewitt-Freilino et al., 2012).

Research on grammatical gender and cognition varies along two primary dimensions: task type and participant population. Explicit tasks—such as voice assignment or gender evaluation—require conscious judgments and have been widely used across languages including French, Spanish, German, and English (Boroditsky et al., 2003; Sera et al., 2002). In contrast, implicit tasks assess gender effects indirectly through measures such as adjective generation, semantic ratings, or object descriptions. For instance, in an adjective-generation task conducted in English, Romanian-German and Romanian-French speakers produced more gender-congruent adjectives depending on the grammatical gender of translation equivalents, demonstrating implicit effects of grammatical gender on conceptualization (Semenuks et al., 2017). Similarly, Maciuszek et al. (2019) reported that monolingual Polish speakers assigned gender-consistent properties to inanimate nouns in a semantic judgment task.

Methodologically, studies range from behavioral paradigms (e.g., semantic categorization and similarity judgments; Cubelli et al., 2011; Vigliocco et al., 2005) to neurocognitive approaches such as event-related potentials (ERPs). ERP studies with Slovenian monolinguals (Mikić Ljubi et al., 2022) and French–English bilinguals (Sato et al., 2020) indicate that grammatical gender modulates neural responses at early stages of processing, suggesting partially automatic activation of gender information. Notably, such neural effects are sometimes observed even when behavioral measures yield weak or inconsistent results.

Findings from monolingual populations remain mixed. While some studies report robust gender-congruency effects (e.g., Polish; Maciuszek et al., 2019), others—particularly in languages such as German and Italian—find weak, absent, or task-dependent effects

(Bender et al., 2016a, 2016b; Montefinese et al., 2019). This variability has been attributed to methodological factors, including task demands, stimulus characteristics, and properties of specific gender systems (Bassetti, 2007; Sato & Athanasopoulos, 2018).

Within bilingualism research, grammatical gender effects have been examined across three main language configurations. First, bilinguals who use two gender-marking languages (e.g., Spanish–German or Italian–German) provide insight into how distinct gender systems interact. Evidence from this group suggests that gender representations are flexible and context-dependent. For example, Lambelet (2016) found that learners of French as a later-acquired gender-marking language showed strong influence from the earlier-acquired language (e.g., German, Spanish, Russian), with limited and unstable effects of the later-acquired system. Complementary evidence from within-group cross-linguistic comparisons indicates that grammatical gender can influence tasks such as voice assignment. For instance, speakers of two Norwegian varieties differing in their gender systems (three vs. two genders) show gender effects that are relatively small compared to culturally driven associations (Beller et al., 2015).

Bidirectional influence between languages has also been documented. For example, Andonova et al. (2007) examined bilingual speakers of two gender-marking languages (e.g., Bulgarian–German or Spanish–German) using a picture-based gender assignment task, and found that gender from the later-acquired language was occasionally extended to nouns in the earlier-acquired language, indicating cross-linguistic interaction in grammatical gender representations. However, other studies report more limited and variable effects. For instance, Vernich (2019) compared Lithuanian speakers with different language backgrounds (Lithuanian–English, Lithuanian–Italian, Lithuanian–Russian, Lithuanian–German) using a voice attribution task, and found that gender effects varied across groups depending on the presence and type of grammatical gender in the additional language, as well as factors such as proficiency and exposure.

Second, studies have examined bilinguals whose earlier-acquired language encodes grammatical gender while the later-acquired language does not (e.g., German–English, Spanish–English). In this configuration, gender effects from the earlier-acquired language often persist during processing in the later-acquired language. For instance, Boroditsky et al. (2003) showed that bilinguals produced gender-consistent descriptions even when performing tasks in English. Using a visual priming paradigm, Sato & Athanasopoulos (2018) further demonstrated that grammatical gender can automatically influence perceptual judgments in bilinguals, even when it is not task-relevant. Converging evidence from ERP studies indicates that grammatical gender from the earlier-acquired language is activated during processing in the later-acquired language (Boutonnet et al., 2012). Developmental findings suggest that such effects become more stable with increasing experience (Nicoladis et al., 2016).

The third configuration—central to the present study—concerns bilinguals whose earlier-acquired language lacks grammatical gender, while the later-acquired language encodes it (e.g., Farsi–German bilinguals). In contrast to the extensively studied direction in which grammatical gender is acquired early, relatively little is known about whether grammatical gender acquired later in life can shape conceptual processing. This configuration provides a critical test case for linguistic relativity in bilingualism, as it allows examination of whether newly acquired grammatical categories can influence cognition in speakers whose other linguistic system does not encode such distinctions.

Bilinguals: genderless-gendered language pairings

Research on bilinguals who use a genderless language alongside a gender-marking language remains comparatively limited. This configuration is particularly informative, as it allows examination of whether grammatical gender—absent from one language but present in the other—can influence conceptual processing.

Existing accounts suggest that exposure to a gender-marking language may lead to sensitivity to grammatical gender distinctions, even in speakers whose other language does not encode such categories (Bassetti & Nicoladis, 2016). However, empirical evidence for such effects is scarce and less consistent than in bilingual populations involving two gender-marking languages or those in which grammatical gender is present in the earlier-acquired language. This gap highlights the need for studies that directly examine whether grammatical gender can shape conceptual processing in bilinguals for whom gender is not a feature of one of their linguistic systems, and under what conditions such effects may emerge.

Empirical evidence demonstrates that grammatical gender can influence conceptual processing in this configuration. Chen and Faitaki (2024) showed that bilingual speakers of English and French assigned gender-congruent voices to inanimate objects in line with French grammatical gender, despite French being acquired later. Importantly, this effect was found to be independent of L2 proficiency, suggesting that neither age of acquisition nor proficiency alone fully accounts for the observed patterns. Similarly, Kurinski and Sera (2011) showed that adult English speakers learning Spanish categorized inanimate objects in ways that increasingly aligned with Spanish grammatical gender in a classification task. This shift was associated with increased exposure to the gender-marking language, rather than with proficiency per se, as the magnitude of the effect did not vary systematically with proficiency.

While these studies provide some evidence that grammatical gender may be engaged during task performance in bilingual speakers, the available findings remain limited, both in scope and consistency, and are largely based on behavioral measures in a narrow range of language pairings.

The present study addresses this methodological gap by examining Farsi–German bilinguals, a population in which grammatical gender is encoded in one language but not the other. Using an eye-tracking paradigm, it investigates whether grammatical gender is engaged

during online comprehension, thereby providing a process-sensitive test of whether previously observed effects reflect stable representations or task-specific strategies.

3. Methodology

The study employed a visual-world eye-tracking paradigm to examine real-time sensitivity to grammatical gender during online processing. The design included three phases. In Phase 1, German monolingual speakers completed the task in German, characterizing gender-related processing in a monolingual context. In Phase 2, Farsi–German bilingual speakers completed the same task in German. In Phase 3, a subset of bilingual participants additionally completed the task in Farsi.

This design enables both between-group comparisons (monolingual vs. bilingual speakers in German) and within-group comparisons across languages (German vs. Farsi), allowing assessment of whether sensitivity to grammatical gender depends on the language context in which processing occurs.

This design tests linguistic relativity by examining whether sensitivity to grammatical gender varies across language contexts. If grammatical gender influences online processing, gender-congruent attention patterns should emerge in a gender-marking language but be reduced or absent in a language without grammatical gender. Combined between- and within-group comparisons distinguish language-specific from more stable effects. Importantly, an online, implicit task minimizes strategic or metalinguistic influences.

3.1 Participants

A total of 99 adults participated in the study. Of these, 46 were German monolinguals (33 females; age: $M = 24.8$, $SD = 3.6$, range = 21–40). The remaining 53 were Farsi–German bilinguals who completed the task in German.

The bilingual group consisted of 53 native speakers of Farsi (24 females, 29 males), aged 23–61 years ($M = 34.6$, $SD = 8.6$). All had acquired German after adolescence and

reported regular use of German in both every day and professional contexts. The mean age of acquisition of German was 24.6 years (*SD* = 4.8, range = 14–31), and participants had lived in Germany for an average of approximately 11 years (range = 7–41 years).

Self-reported German proficiency was assessed using the Common European Framework of Reference for Languages (CEFR). Most participants reported advanced proficiency: 34 at C1, 8 at C2, 9 at B2, and 2 at B1.

Participants also completed a detailed language background questionnaire developed at the *psycholinguistics laboratories* of TU Dortmund University, assessing language use, proficiency, and attitudes across multiple social, academic, and professional contexts. Based on these responses, a bilingualism index was calculated following Odermann and Mertins (2026). This index integrates language use, input, and self-rated proficiency in German and Farsi across contexts. The results indicated that 43 of the 53 bilingual participants had fully balanced bilingual profiles, whereas the remaining participants showed only minor asymmetries between the two languages.

A subset of the bilingual group (*n* = 36, 15 females; age: *M* = 35.8, *SD* = 5.7, range = 22–47) additionally completed the task in Farsi, allowing within-group comparison across the participants' first language (Farsi) and second language (German).

The study was approved by the Ethics Committee of TU Dortmund University. All participants provided informed consent before participation. German monolingual participants received course credit as part of their class performance, whereas the bilingual Farsi–German participants took part voluntarily without financial compensation.

3.2 Materials

The experiment included 30 spoken noun stimuli and 30 visually presented adjectives, arranged into 30 stimulus sets. In each trial, participants heard one noun while viewing four adjectives on the screen. Nouns were presented without definite articles to avoid explicit morphosyntactic cues to grammatical gender. The stimuli were generated using the

ElevenLabs text-to-speech system with a natural-sounding, neutral voice and were amplitude-normalized to ensure consistency in pronunciation, speech rate, and prosody across items.

The adjectives were drawn from a large pool compiled in a behavioral study by the first two authors (Fatemi et al., 2025), in which 102 native speakers of German rated 4,808 adjectives on a 7-point scale for gender connotation (1 = *very feminine*, 4 = *gender-neutral*, 7 = *very masculine*). Based on these ratings, 30 adjectives were selected for the present study: 10 feminine, 10 masculine, and 10 neutral.

Each trial contained one critical adjective, two semantic competitors, and one distractor. The critical adjective was grammatically congruent with the noun and strongly semantically associated with it. The competitor adjectives, drawn from the remaining gender categories, were semantically related but gender-incongruent. Specifically, for each grammatical gender condition, the competitors were selected from the two non-matching gender categories (e.g., feminine and neuter competitors for masculine targets). In the analyses, these competitor types are labeled according to their grammatical gender as Comp F (F = feminine), Comp M (M = masculine), and Comp N (N = neuter). The distractor was neither semantically related nor gender-congruent.

The auditory stimuli consisted of 30 nouns selected based on their collocational relationships with the adjectives using the German Reference Corpus (DeReKo) via the COSMAS II interface. In addition to identifying the adjective–noun combinations, likelihood ratios for the German collocations were extracted from COSMAS II; the full stimulus list and collocation information are reported in Appendix A.

Nouns were presented without definite articles to avoid explicit morphosyntactic cues to grammatical gender in German, where gender is typically marked on articles and agreement morphology. In contrast, Farsi does not encode grammatical gender, and nouns are not associated with gender-marking morphology. This ensured that any preference for gender-

congruent adjectives in German reflected implicit activation of grammatical gender, while no such cues were available in Farsi.

The positions of the four adjectives were counterbalanced across trials so that each adjective type appeared equally often in each screen position.

The bilingual participants tested in German received the same German materials. For the Farsi version, all nouns and adjectives were translated into Farsi by the first author. Four stimulus sets were excluded because the translations did not preserve the intended semantic relation or experimental manipulation: *Sprache* (“language”), *Mühe* (“effort”), *Harmonie* (“harmony”), and *Buch* (“book”). For example, the Farsi translation of *Sprache* (“/zabān/”) is ambiguous between “language” and “tongue,” and several critical adjectives could not be translated into a single semantically appropriate Farsi adjective.

Because no corpus-based collocation analysis comparable to COSMAS II / DeReKo was available for Farsi, the translated items were evaluated by 20 native speakers of Farsi residing in Iran, who rated the semantic naturalness of each noun–adjective combination on a 7-point Likert scale. Only items receiving high naturalness ratings were retained.

Stimulus presentation in the Farsi task was counterbalanced in the same way as in the German task. Because Farsi is written from right to left, the spatial arrangement of the adjectives followed the Farsi reading direction.

3.3 Procedure

Participants were tested individually in a quiet laboratory environment. Stimuli were presented on a computer monitor while eye movements were recorded with an SMI RED 60 eye tracker (60 Hz) connected to a standard PC.

Participants sat approximately 60 cm from the screen. Eye-tracking accuracy was established with a four-point calibration followed by validation. Calibration was accepted when the average error was below 1° of visual angle; otherwise, calibration was repeated.

After calibration, participants heard prerecorded task instructions generated with a synthetic voice. In German, the instruction was: “Ihnen werden vier Wörter auf dem Bildschirm angezeigt. Daraufhin werden Sie ein weiteres Wort hören. Wählen Sie von den angezeigten Wörtern eines aus, welches am besten zu dem von Ihnen gehörten Wort passt und nennen Sie es.“ (“You will see four words on the screen. Afterwards, you will hear another word. Please choose the word from the displayed options that best matches the word you heard and say it aloud.”) Two practice trials preceded the main experiment.

Each trial began with a central fixation cross to establish a common starting position. A display containing four written words then appeared on the screen (see Figure 1). The display had a resolution of 1920 × 1080 pixels and was divided into four quadrants, with one word centered in each quadrant. The critical item was counterbalanced across quadrants and randomized across trials.

Following standard procedures in visual world paradigm studies, the visual display appeared 200 ms before auditory onset, providing a brief preview period commonly used in visual-world studies to reduce strategic reading and control anticipatory eye movements (Huettig & McQueen, 2007; McQueen & Viebahn, 2007; Pier Salverda et al., 2011). The display remained visible for 3.000 ms after auditory onset, and eye-movement analyses were time-locked to the onset of the spoken noun.

Participants were instructed to select and name aloud the adjective that best described or was most appropriate for the spoken noun. Eye movements were recorded throughout the trial. Participants who completed both language versions did so within the same session, with a break of approximately 5-10 minutes between tasks.

To maintain attention and reduce predictability, 40 filler trials involving a color-detection task were included. In these trials, participants indicated whether colored squares on the screen were the same or different. Following the practice and filler trials, participants completed the experimental trials in randomized order.

Eye-movement data were recorded using the SMI Experiment Center and exported from SMI BeGaze for preprocessing. All statistical analyses were conducted in *R*. Data quality was assessed using the tracking ratio, defined as the proportion of valid gaze samples recorded during a trial. Trials with a tracking ratio below 90% were excluded.

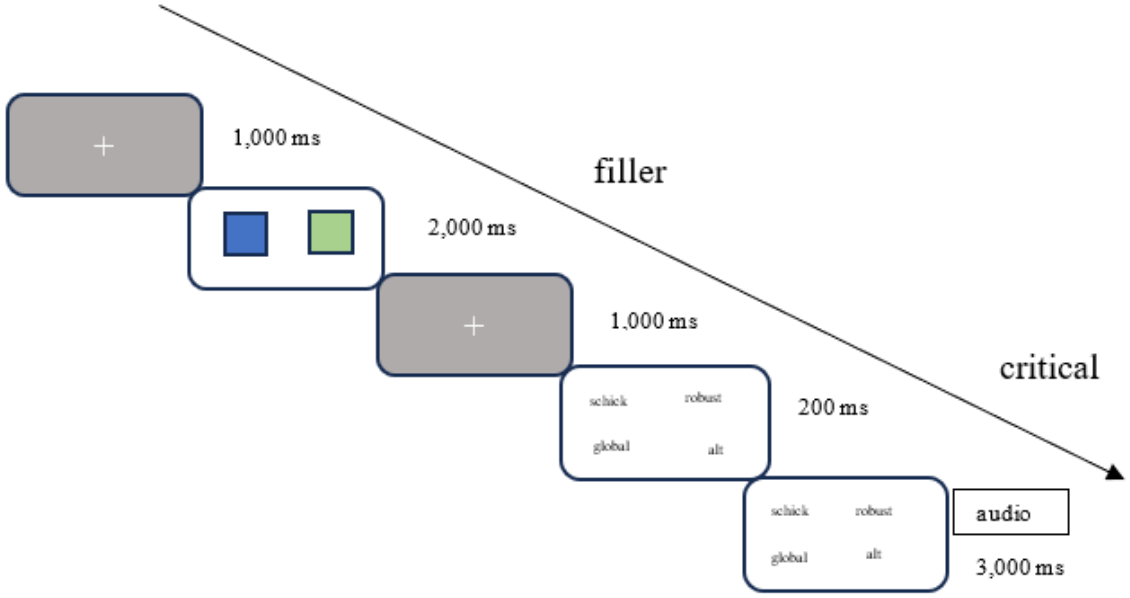


Figure 1 Overview of the experimental design and the sequence of the events within each trial

3.4 Eye-movement measures and statistical analysis

Areas of interest (AOIs) were defined around each of the four words displayed on the screen and categorized according to their experimental role (critical, competitors [F, M, N], and distractor). To ensure comparability across stimuli and languages, AOIs were standardized using the maximum width and height of the longest word in the stimulus set. The same AOI dimensions were applied to all German and Farsi items, ensuring that gaze differences were not attributable to word length.

Eye-movement data were extracted using SMI BeGaze and analyzed in *R*. The primary dependent variable was dwell time, defined as the total duration of fixations within an AOI during the analysis window. Dwell time was selected as the main measure because several

adjectives in each display represented plausible descriptions of the auditory noun, making cumulative attention more informative than early fixation differences alone.

In addition, fixation proportions were calculated in 100-ms time bins to examine the temporal dynamics of gaze allocation across AOIs, following standard practice in visual-world research. For each time bin, the proportion of fixations directed to each AOI was computed relative to all fixations within the display. Eye-movement analyses were time-locked to auditory onset, excluding the initial 200 ms following display onset to minimize center-bias effects.

Eye-tracking data were analyzed using linear mixed-effects models (LMMs). Because fixation-duration measures are typically positively skewed, dwell time was log-transformed prior to analysis to improve normality and meet model assumptions. All models included participants and items as random intercepts. Fixed effects included grammatical gender, AOI type, congruency condition, and participant group. Post hoc pairwise comparisons were conducted using estimated marginal means and adjusted for multiple comparisons. All statistical tests were evaluated using a significance level of $\alpha = .05$.

4. Results

The primary outcome measure was dwell time during the eye-tracking task, as it provides a robust index of overall attentional allocation in contexts where multiple response alternatives are semantically plausible (Huettig et al., 2011; Tanenhaus et al., 1995). In the present design, several adjectives per display constituted reasonable candidates for the spoken noun, making dwell time a more appropriate measure of graded preference than discrete fixation measures or time-course analyses alone (Huettig et al., 2011).

To examine how grammatical gender influenced visual attention, we analyzed dwell time using linear mixed-effects models with grammatical gender, AOI type, and participant group (German monolinguals, Farsi–German bilinguals tested in German, and Farsi–German bilinguals tested in Farsi) as predictors.

Post hoc pairwise comparisons were conducted using Holm's sequential Bonferroni procedure (Holm, 1979). In addition, Pearson correlation analyses examined the relationship between participants' behavioral responses and their dwell times on the critical and competitor AOIs. Finally, fixation proportions were calculated in 100-ms time bins to examine the temporal dynamics of gaze allocation across AOIs; these time-course analyses are reported in Appendix C (Figure C4).

4.1 German monolingual speakers

Research Question 1 addressed whether grammatical gender cues from the spoken noun guide visual attention toward the gender-congruent adjective, resulting in a pattern of greater dwell time on the critical adjective relative to competitors and the distractor.

Log-transformed dwell time was analyzed using a linear mixed-effects model with grammatical gender (*der, die, das;* masculine, feminine, neuter) and AOI (critical, competitor F, competitor M, competitor N, distractor) as fixed effects and their interaction. Following Barr et al. (2013), the model included random intercepts for participants and items (Table 1). The reference levels were *der* and critical, such that the intercept reflects dwell time on the gender-congruent adjective for masculine nouns.

The model revealed a significant intercept ($\beta = 6.23$, $SE = 0.063$, $t\,(118.9) = 99.63$, $p < .001$). No main effects of grammatical gender were observed (*die*: $\beta = -0.015$, $SE = 0.071$, $t\,(6180) = -0.22$, $p = .828$; *das*: $\beta = 0.070$, $SE = 0.071$, $t\,(6180) = 0.98$, $p = .326$). Several AOI contrasts differed significantly from the critical condition. Dwell time was reduced for competitor M ($\beta = -0.165$, $SE = 0.071$, $t\,(200.1) = -2.32$, $p = .020$), competitor N ($\beta = -0.226$, $SE = 0.071$, $t\,(153.0) = -3.18$, $p = .001$), and the distractor ($\beta = -0.356$, $SE = 0.071$, $t\,(201.7) = -4.99$, $p < .001$). The comparison with competitor F did not reach statistical significance, although the effect was numerically in the expected direction ($\beta = -0.137$, $SE = 0.071$, $t\,(240.3) = -1.92$, $p = .055$). A significant interaction was observed between neuter nouns (*das*)

and competitor F ($\beta = -0.348$, $SE = 0.101$, t (6179) = −3.46, $p = .001$), indicating reduced attention to feminine competitors when neuter nouns were processed.

Table 1. Linear mixed-effects model for log-transformed dwell time (RQ1: German monolinguals)

Term	Estimate β	*SE*	*df*	*t*	*p*
Intercept	6.230	.063	118.9	99.63	< .001
die	-.015	.071	6180	-.22	.828
das	.070	.071	6180	.98	.326
Comp F	-.137	.071	240.3	-1.92	.055
Comp M	-.165	.071	200.1	-2.32	.020
Comp N	-.226	.071	153.0	-3.18	.001
Distractor	-.356	.071	201.7	-4.99	< .001
das × Comp F	-.348	.101	6179	-3.46	.001
die × Comp M	.045	.101	6178	.45	.653
die × Comp N	-.098	.101	6179	-.97	.332
die × Distractor	.188	.101	6179	1.87	.061
das × Distractor	.022	.101	6179	.22	.827

Linear mixed-effects model for log-transformed dwell time. The model was based on 5,520 observations from 46 participants, with random intercepts for participants and items. Reference levels were der (grammatical gender) and critical (AOI). β = estimated regression coefficient; *SE* = standard error; *t* = test statistic. Marginal $R2 = .010$; conditional $R2 = .061$.

Bonferroni-corrected pairwise comparisons further showed that the critical adjective received greater attention than several incongruent alternatives. For masculine nouns, dwell time on the critical adjective exceeded competitor N ($z = 3.18$, $p = .004$) and the distractor ($z = 4.99$, $p < .001$). For feminine nouns, the critical adjective differed from competitor N ($z = 4.55$, $p < .001$). For neuter nouns, the critical adjective exceeded competitor F ($z = 6.81$, $p < .001$) and the distractor ($z = 4.69$, $p < .001$). The full set of Bonferroni-corrected comparisons is reported in Appendix B. The predicted dwell-time patterns across AOIs are illustrated in Figure C1 (Appendix C).

To examine the relationship between behavioral responses and visual attention, a Pearson correlation analysis was conducted at the participant level. The correlation between the proportion of critical responses and mean dwell time on the critical AOI was positive but not significant, r (44) = .13, p = .385, 95%, *CI*: [−.17, .41].

4.2 Farsi-German bilingual speakers tested in German

To test whether grammatical gender guides visual attention when bilingual participants performed the task in German, log-transformed dwell time was analyzed using a linear mixed-effects model with grammatical gender (*der, die, das*) and AOI type (critical, competitor F, competitor M, competitor N, distractor) as fixed effects and random intercepts for participants and items (Table 2). The reference levels were *der* and *critical*, such that the intercept reflects the mean logarithmized dwell time on the gender-congruent adjective for masculine nouns.

The model revealed a significant intercept (β = 6.38, *SE* = 0.028, t (354.1) = 229.66, p < .001). No main effects of grammatical gender were observed (*die*: β = −0.010, *SE* = 0.034, t (5972.7) = −0.31, p = .760; *das*: β = 0.031, *SE* = 0.033, t (5972.5) = 0.94, p = .350).

Several AOI contrasts differed significantly from the critical condition. Relative to the critical adjective, dwell time was reduced for competitor M (β = −0.100, *SE* = 0.033, t (5973.1) = −3.01, p = .003), competitor N (β = −0.147, *SE* = 0.034, t (5972.0)= −4.37, p < .001), and the distractor (β = −0.167, *SE* = 0.034, t (5971.9) = −5.00, p < .001). The comparison with competitor F did not reach significance (β = −0.058, *SE* = 0.033, t (5972.8) = −1.74, p = .083).

Among the interaction terms, a significant interaction was observed between neuter nouns (*das*) and competitor F (β = −0.156, *SE* = 0.047, t (5972.2) = −3.31, p < .001), indicating reduced attention to feminine competitors when neuter nouns were processed. In addition, the interaction between feminine nouns (*die*) and the distractor was significant (β = 0.104, *SE* = 0.047, t (5972.1) = 2.21, p = .027).

Table 2. Linear mixed-effects model for log-transformed dwell time (RQ2: Bilinguals tested in German).

Term	Estimate β	*SE*	*df*	*t*	*p*
Intercept	6.379	.028	354.1	229.66	< .001
die	-.010	.034	5972.7	-.31	.760
das	.031	.033	5972.5	.94	.350
Comp F	-.058	.033	5972.8	-1.74	.083
Comp M	-.100	.033	5973.1	-3.01	.003
Comp N	-.147	.034	5972.0	-4.37	< .001
Distractor	-.167	.034	5971.9	-5.00	< .001
die × Comp M	.022	.047	5972.4	.48	.633
die × Comp N	-.017	.047	5972.2	-.36	.716
die × Distractor	.104	.047	5972.1	2.21	.027
das × Comp F	-.156	.047	5972.2	-3.31	< .001
das × Distractor	.041	.047	5972.0	.87	.387

Linear mixed-effects model for log-transformed dwell time. The model was based on 6,337 observations from 53 participants, with random intercepts for participants and items. Reference levels were der (grammatical gender) and critical (AOI). β = estimated regression coefficient; *SE* = standard error; *t* = test statistic. Marginal R^2 = .003; conditional R^2 = .181.

Holm-corrected pairwise comparisons examined differences between the critical adjective and competing alternatives within each grammatical gender condition. For masculine nouns (*der*), dwell time on the critical adjective exceeded that on competitor N ($z = 4.37$, $p < .001$) and the distractor ($z = 5.00$, $p < .001$), whereas the comparison with competitor F was not significant ($p = .165$). For feminine nouns (*die*), the critical adjective attracted more dwell time than competitor N ($z = 4.92$, $p < .001$), while the comparisons with competitor M ($z = 2.32$, $p = .061$) and the distractor ($p = .117$) were not significant. For neuter nouns (*das*), dwell time on the critical adjective exceeded that on competitor F ($z = 6.45$, $p < .001$), competitor M ($z = 3.01$, $p = .008$), and the distractor ($z = 3.82$, $p < .001$). The full set of Holm-

corrected pairwise comparisons is reported in Appendix B (Table B2). The predicted dwell-time patterns across AOIs are illustrated in Figure C2 (Appendix C).

To examine the relationship between behavioral responses and visual attention, a Pearson correlation analysis was conducted at the participant level between the proportion of critical responses and mean dwell time on the critical AOI. The analysis revealed no significant association between behavioral preference and visual attention, r (51) = .13, p = .389, 95%, *CI*: [−.17, .41].

4.3 Farsi–German bilingual speakers tested in Farsi

Research Question 3 addressed whether gender-related visual attention patterns are reduced or absent when the task is performed in Farsi, such that no preference for gender-congruent adjectives is expected. To examine this, log-transformed dwell time was analyzed using a linear mixed-effects model with grammatical gender and AOI type as fixed effects and random intercepts for participants and items (Table 3), using the same model specification as in Hypotheses 1 and 2. The reference levels were der (grammatical gender) and critical (AOI), such that the intercept reflects the mean logarithmized dwell time on the gender-congruent adjective for masculine nouns.

The model revealed a significant intercept (β = 8.145, SE = 0.100, t (110.1) = 81.70, p < .001). No main effects of grammatical gender were observed (*die*: β = 0.183, SE = 0.135, t (305.2) = 1.35, p = .176; *das*: β = −0.060, SE = 0.126, t (305.2) = −0.48, p = .634). Similarly, no significant differences were observed between the critical item and the competing items or distractor (competitor F: β = −0.120, SE = 0.128, t (174.3) = −0.94, p = .349; competitor M: β = 0.013, SE = 0.127, t (700.4) = 0.11, p = .916; competitor N: β = −0.065, SE = 0.123, t (170.1) = −0.53, p = .596; distractor: β = −0.078, SE = 0.120, t (787.9) = −0.65, p = .518).

No significant interaction effects between grammatical gender and AOI type were observed (all ps > .12). A marginal trend was observed for the interaction between die and

competitor N ($\beta = -0.322$, $SE = 0.186$, t (3521.4) $= -1.73$, $p = .084$), but this effect did not reach statistical significance.

Table 3. Linear mixed-effects model for log-transformed dwell time (RQ3: Bilinguals tested in Farsi).

Term	Estimate β	*SE*	*df*	*t*	*p*
Intercept	8.145	.100	110.1	81.70	< .001
die	.183	.135	305.2	1.35	.176
das	-.060	.126	305.2	-.48	.634
Comp F	-.120	.128	174.3	-.94	.349
Comp M	.013	.127	700.4	.11	.916
Comp N	-.065	.123	170.1	-.53	.596
Distractor	-.078	.120	787.9	-.65	.518
das × Comp F	-.024	.174	3521.4	-.14	.888
die × Comp M	-.290	.191	3521.4	-1.52	.128
die × Comp N	-.322	.186	3521.4	-1.73	.084
die × Distractor	-.130	.186	3521.4	-.70	.486
das × Distractor	.089	.174	3521.4	.51	.607

Linear mixed-effects model for log-transformed dwell time. The model was based on 4,550 observations from 35 participants, with random intercepts for participants and items. Reference levels were der (grammatical gender) and critical (AOI). β = estimated regression coefficient; *SE* = standard error; t = test statistic. Marginal $R^2 = 0.270$; conditional R^2 was not available due to singularity in the random-effects structure.

Exploratory Bonferroni-corrected pairwise comparisons further confirmed the absence of systematic gender-congruency effects. Only one contrast reached significance: for feminine nouns, dwell time on the critical item exceeded that on competitor N ($z = -2.66$, $p = .023$). All remaining comparisons were not significant after correction (all ps $\geq .163$). The full set of Bonferroni-corrected comparisons is reported in Appendix B. The predicted dwell-time patterns across AOIs are illustrated in Figure C3 (Appendix C).

To examine the relationship between behavioral responses and visual attention, a Pearson correlation analysis was conducted at the participant level between the proportion of critical responses and mean dwell time on the critical AOI. The analysis revealed no significant association between behavioral preference and visual attention, $r = .13$, $p = .385$, 95%, *CI*: [−.17, .41].

4.4 Group comparison: monolingual vs. bilingual speakers tested in German

To examine whether visual attention differed between German monolinguals and bilingual speakers tested in German, a linear mixed-effects model was fitted with group (monolingual vs. bilingual_German) and AOI type (critical vs. competitor) as fixed effects. The intercept reflects the mean baseline log-dwell time for bilingual participants tested in German while fixating competitor objects, which served as the reference level in the model.

The model revealed a significant main effect of group, $\beta = 3.716$, $SE = 0.116$, t (8508) = 31.92, $p < .001$, indicating overall differences in visual attention between the two groups. A significant main effect of object type was also observed, $\beta = 1.301$, $SE = 0.145$, t (12133) = 8.96, $p < .001$, reflecting differences in attention between critical targets and competitors. These effects were qualified by a significant interaction between group and object type, β = −3.799, SE = 0.233, t (12132) = −16.29, $p < .001$. The full model results are presented in Table 4.

Table 4. Linear mixed-effects model for log dwell time monolingual vs. bilingual-German (RQ4)

Term	Estimate β	*SE*	*df*	*t*	*p*
Intercept	6.800	.100	72.8	67.87	< .001
Group (Bilingual – Monolingual)	3.716	.116	8508	31.92	< .001
Object type (Competitor – Critical)	1.301	.145	12133	8.96	< .001
Group × AOI type	-3.799	.233	12132	-16.29	< .001

Linear mixed-effects model for log-transformed dwell time. The model was based on 12,220 observations from 99 participants (53 bilingual speakers tested in German and 46 monolingual speakers) and 30 items, with random intercepts for participants and items. Marginal R^2 = .003; conditional R^2 = .181.

Bonferroni-corrected post hoc comparisons showed that bilingual speakers tested in German fixated significantly more on the critical target than on competitor objects, $z = -8.96$, $p < .001$. In contrast, monolingual participants showed greater dwell time on competitor objects than on the critical target $z = 13.70$, $p < .001$, suggesting a different pattern of visual attention in this group.

Between-group comparisons revealed a significant difference in the competitor condition, with monolingual speakers fixating competitors more than bilingual speakers, $z = -31.92$, $p < .001$. No significant difference between groups was observed for the critical target condition, $z = 0.40$, $p = .691$. The full set of Bonferroni-corrected comparisons is reported in Appendix B (Tables B4 and B5).

The pattern of group differences between targets and competitors is illustrated in Figure 2.

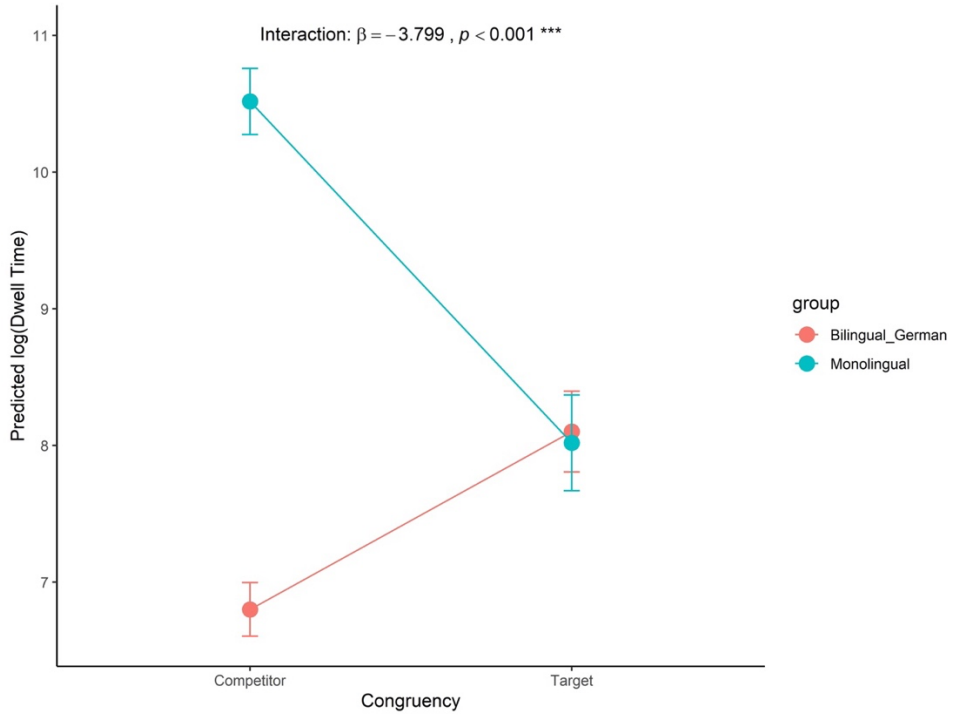


Figure 2 Estimated marginal means of log-transformed dwell time for critical targets and competitor objects in German monolingual and bilingual speakers tested in German. Error bars represent 95% confidence intervals.

Together, these results suggest that visual attention to competitor objects differed substantially between the two groups, whereas attention to the critical target did not vary reliably across groups.

4.5 Group comparison: bilingual speakers tested in German vs. Farsi

Research Question 5 addressed whether bilingual speakers differ in gender-related visual attention patterns when performing the task in German compared to Farsi. To examine this, a linear mixed-effects model was fitted with group (bilingual_German vs. bilingual_Farsi) and AOI type (Target vs. Competitor) as fixed effects. The intercept reflects the mean baseline log dwell time for bilingual participants tested in German when fixating the target object.

The model revealed a significant intercept, $\beta = 8.101$, $SE = 0.102$, $t\,(112.82) = 79.31$, $p < .001$. No significant main effect of group was observed, $\beta = 0.078$, $SE = 0.162$, $t\,(113.89) = 0.48$, $p = .631$, indicating that overall visual attention did not differ between bilingual

participants tested in German and those tested in Farsi. A significant main effect of AOI type was observed, $\beta = -0.134$, $SE = 0.052$, t (7362.19) $= -2.57$, $p = .010$, reflecting greater visual attention to the target object than to the competitor. The interaction between group and AOI type was not significant, $\beta = -0.014$, $SE = 0.086$, t (7361.65) $= -0.16$, $p = .871$. The full model results are presented in Table 5.

Table 5. Linear mixed-effects model for log-dwell time for monolingual and bilingual speakers tested in German (RQ5)

Term	Estimate β	*SE*	*df*	*t*	*p*
Intercept	8.101	.102	112.82	79.31	< .001
Bilingual_German – Bilingual_Farsi	.078	.162	113.89	.48	.631
Competitor	-.134	.052	7362.19	-2.57	.010
Group (Bilingual_Farsi) × Competitor	-.014	.086	7361.65	-.16	.871

Linear mixed-effects model for log-transformed dwell time. The model was based on 7,481 observations from 88 participants (53 bilingual speakers tested in German and 35 bilingual speakers tested in Farsi) and 30 items, with random intercepts for participants and items. Marginal $R^2 = .002$; conditional $R^2 = .138$.

Bonferroni-corrected post hoc *comparisons* within each group showed that both bilingual groups fixated the target object significantly more than the competitor. For bilingual speakers tested in German, the difference between target and competitor objects was significant, $z = 2.57$, $p = .010$. A similar pattern was observed for bilingual speakers tested in Farsi, $z = 2.15$, $p = .031$.

Between-group comparisons revealed no significant differences between bilingual speakers tested in German and those tested in Farsi for either the target condition, $z = -0.48$, $p = .630$, or the competitor condition, $z = -0.42$, $p = .678$. The full set of Bonferroni-corrected comparisons is reported in Appendix B (Tables B6 and B7).

The pattern of visual attention across groups and object types is illustrated in Figure 3.

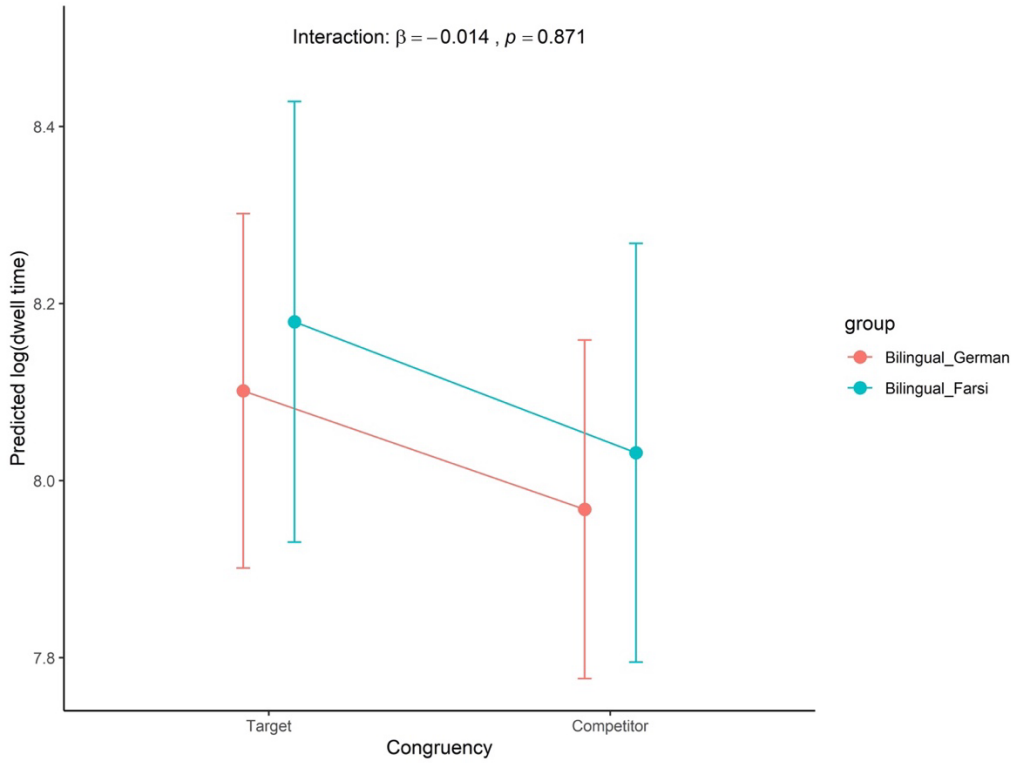


Figure 3 Estimated marginal means of log-transformed dwell time for target and competitor objects in bilingual speakers tested in German and Farsi. Error bars represent 95% confidence intervals.

4.6 Time-course analysis of fixation proportions

In addition to the dwell time analyses reported above, fixation proportions were calculated in 100-ms time bins to examine the temporal dynamics of gaze allocation, following common practice in Visual World Paradigm studies (e.g., Huettig et al., 2011; Tanenhaus et al., 1995). For each trial, fixation proportions were computed for the critical adjective, the three competitor adjectives**,** and the distractor across the entire analysis window.

Because several adjectives within each display constituted plausible responses to the auditory noun, the resulting time-course plots did not show clear divergence between gender-congruent and incongruent alternatives. Instead, fixation proportions remained relatively similar across AOIs throughout the time window. Consequently, these analyses did not reveal additional effects beyond the dwell time results reported in Sections 4.1–4.5**.**

For completeness, the full set of fixation-proportion time-course plots for each participant group and grammatical gender condition is provided in Appendix C (Figure C4)**.**

5. Discussion

This study investigated the linguistic relativity hypothesis by examining whether grammatical gender cues influence visual attention during lexical processing in German monolinguals and Farsi–German bilinguals using a Visual World eye-tracking paradigm. Specifically, we tested whether participants preferentially attended to adjectives whose stereotypical gender matched the grammatical gender of an auditory noun.

Three main findings emerged. First, German monolinguals showed clear gender-congruency effects in visual attention. Second, Farsi-German bilingual speakers tested in German exhibited a similar but weaker pattern. Third, the effect disappeared when Farsi-German bilinguals performed the task in Farsi, a language without grammatical gender. Taken together, these findings provide support for the linguistic relativity hypothesis by demonstrating that grammatical gender can modulate online visual attention during language processing, yet the findings indicate that the gender congruency effect depends on both the grammatical structure of the language being processed and the speaker's linguistic experience. Building on these findings, the present study provides the first eye-tracking investigation within the linguistic relativity framework to examine gender-congruency effects in visual attention across the full German grammatical gender system (*der, die, das*), while also comparing German monolingual speakers with Farsi–German bilinguals tested in both German and Farsi.

Time-course analyses of fixation proportions were additionally conducted to examine the temporal dynamics of gaze allocation; however, because multiple adjectives in each display constituted plausible responses to the auditory noun, fixation trajectories remained largely similar across alternatives and did not reveal effects beyond those captured by the dwell time analyses.

5.1 Grammatical gender effects in monolingual processing

German monolingual participants showed clear gender-congruency effects in visual attention. When processing nouns marked for grammatical gender, participants preferentially attended to adjectives whose stereotypical gender matched the grammatical gender of the auditory noun. This pattern indicates that grammatical gender information was rapidly integrated during comprehension and guided the allocation of visual attention. The findings align with previous research suggesting that grammatical gender cues can shape conceptual expectations during language processing (Boroditsky et al., 2003; Sato & Athanasopoulos, 2018). In this sense, the present results provide further evidence that linguistic structure can influence perceptual attention during real-time comprehension.

Importantly, the strength of the gender-congruency effect varied across competitors. The contrast between the critical adjective and the feminine competitor was weaker than the contrasts involving the other competitors. This asymmetry may reflect an interaction between grammatical gender and conceptual gender stereotypes associated with the adjectives. Previous research has shown that such stereotypes can influence how speakers interpret objects even when grammatical gender assignment itself is arbitrary (Boroditsky et al., 2003; Sato & Athanasopoulos, 2018). In addition, masculine forms in German are often interpreted generically, referring to individuals regardless of gender (Gygax et al., 2008; Stahlberg et al., 2001). Such generic interpretations may reduce the perceived mismatch between masculine nouns and feminine adjectives, thereby weakening the contrast between congruent and incongruent alternatives. However, item-level lexical properties may also have contributed to the observed variation in effect strength, a possibility that future studies should examine more systematically.

Finally, the correlation analysis revealed no significant association between the proportion of gender-congruent responses and mean dwell time on the critical adjective. This dissociation suggests that visual attention and explicit responses may reflect different stages

of processing. Eye movements provide a sensitive measure of early attentional processes during comprehension, whereas explicit responses likely reflect later decision processes (Huettig et al., 2011; Rayner, 1998). The absence of a strong correlation, therefore indicates that participants' gaze behavior was already sensitive to grammatical gender cues even when this sensitivity was not directly reflected in their overt responses.

5.2 Gender processing in bilingual speakers tested in German

The results for Farsi–German bilingual participants performing the task in German indicated a pattern consistent with Research Question 2. Similar to the monolingual group, bilingual participants showed increased dwell time on the gender-congruent adjective relative to several competing alternatives, particularly the neuter competitor and the distractor. This pattern indicates that grammatical gender cues were also used during bilingual processing in German, guiding visual attention toward linguistically compatible items. However, the effect was weaker and less consistent than that observed for monolingual speakers.

Importantly, the bilingual participants in the present study were highly proficient users of German, although the language was typically acquired after adolescence. The presence of gender-related attentional effects in this group, therefore, suggests that grammatical gender acquired later in life can become integrated into online language processing. At the same time, the weaker pattern relative to monolingual speakers may indicate that such representations remain less strongly entrenched when grammatical gender is acquired in adulthood and is not supported by the first language (Sabourin et al., 2006). With increasing experience and continued use of German, however, such representations may become more strongly entrenched, potentially leading to more native-like sensitivity to grammatical gender cues.

These findings are consistent with studies showing that grammatical gender acquired in a later-learned language can influence conceptual processing (Bassetti & Nicoladis, 2016;

Chen & Faitaki, 2024; Kurinski & Sera, 2011), while contrasting with research that has reported limited or inconsistent conceptual effects of grammatical gender in other experimental contexts (e.g., Bender et al., 2016a, 2016b). The present results extend this literature by showing that gender-related effects can emerge not only in explicit judgments but also in real-time visual attention during spoken language comprehension.

Consistent with the monolingual condition, the correlation analysis did not reveal a significant association between the proportion of gender-congruent responses and mean dwell time on the critical adjective. This pattern further supports the interpretation that eye movements capture early attentional processes that are not always directly reflected in explicit behavioral responses.

5.3 Absence of gender effects in Farsi processing

The results for bilingual participants tested in Farsi revealed a pattern consistent with Research Question 3. In contrast to the German conditions, no systematic gender-congruency pattern emerged in visual attention when the task was performed in Farsi. Dwell time did not differ reliably between the critical adjective and the competing alternatives, indicating that participants did not preferentially attend to adjectives whose stereotypical gender matched the grammatical gender of the corresponding German noun. This finding is consistent with the linguistic structure of Farsi, which lacks grammatical gender marking and therefore does not provide gender cues that could guide visual attention during language comprehension.

Although one marginal interaction in the main model and a single exploratory comparison suggested a possible difference between the critical adjective and the neuter competitor in the feminine condition, these isolated effects did not form a consistent pattern across conditions. One possible explanation for this isolated effect may relate to broader cultural conventions in Persian descriptive language. Persian literary tradition, particularly classical poetry, frequently employs aesthetic and affectionate descriptive expressions

associated with beauty and delicacy (Lewisohn, 2012; Yarshater, 1988). Such associations may have slightly increased attention toward adjectives with feminine connotations in isolated cases. However, because this pattern did not appear consistently across conditions, this interpretation should be treated cautiously.

Importantly, the disappearance of the gender-congruency effect cannot simply be attributed to differences between participant groups. While all bilingual participants completed the German version of the task, only a subset of these participants also took part in the Farsi condition. The absence of gender-based attentional patterns in the Farsi condition therefore most plausibly reflects the linguistic context in which the task was performed rather than differences in participant characteristics. More broadly, this pattern is consistent with accounts of bilingual language processing that emphasize context-dependent activation of linguistic systems (Grosjean, 2001, 2008).

The correlation analysis likewise did not reveal a significant association between behavioral responses and dwell time on the critical adjective in the Farsi condition. This pattern further supports the interpretation that grammatical gender cues were not guiding visual attention when the task was performed in a language that does not encode grammatical gender.

5.4 Monolingual–bilingual differences in visual attention

The fourth research question examined whether visual attention during lexical processing differed between German monolingual speakers and Farsi–German bilingual speakers performing the task in German. The results revealed a clear interaction between group and AOI (adjective) type. When competitor objects were collapsed into a single category and contrasted with the critical adjective, bilingual speakers showed greater dwell time on the gender-congruent adjective, whereas monolingual speakers directed relatively more attention toward competing alternatives. Importantly, the two groups did not differ

reliably in fixation on the critical target itself; rather, the difference emerged in how strongly competitor objects attracted visual attention.

This pattern suggests that both groups were sensitive to grammatical gender cues, but that they differed in how lexical competition was resolved during processing. Previous research has reported mixed findings regarding monolingual–bilingual differences in the use of linguistic cues during comprehension. Some studies have shown that bilingual speakers can demonstrate sensitivity to morphosyntactic cues during comprehension (Dussias & Sagarra, 2007; Hopp, 2013), whereas other work suggests that bilingual processing may involve different strategies for integrating linguistic cues and resolving competition (Bialystok, 2017; Kroll et al., 2006). The present findings contribute to this discussion by suggesting that grammatical gender information can guide visual attention in both groups. At the same time, differences may arise in the management of competing lexical alternatives during comprehension.

Such a pattern is consistent with accounts of bilingual language processing that emphasize adaptive mechanisms for regulating competition between linguistic representations (Kroll et al., 2006). Differences between monolingual and bilingual speakers may therefore reflect variation in how lexical competition is resolved rather than differences in the availability of grammatical information.

5.5 Language-dependent activation of grammatical gender

The fifth research question examined whether the grammatical gender bias observed in German would also appear when bilingual speakers performed the task in Farsi. Unlike the comparison reported in Section 5.4, this analysis focused on the bilingual group alone and compared their performance across two language contexts. The results showed that bilingual participants fixated on the critical adjective more than the competitor in both conditions. However, no reliable differences were observed between the German and Farsi versions of the

task, and the interaction between language condition and adjective type was not significant. This pattern suggests that the overall distribution of visual attention between target adjectives and competitors remained comparable across the two language contexts.

Previous research has shown that grammatical cues can guide attention during comprehension when they are present in the input language (Dussias & Sagarra, 2007; Hopp, 2013). At the same time, other studies have proposed that conceptual associations related to grammatical gender may transfer across languages in bilingual speakers (Bassetti & Nicoladis, 2016; Boroditsky et al., 2003). The present findings provide limited evidence for such cross-language transfer and instead suggest that attentional patterns were largely determined by the immediate linguistic context.

Taken together with the results reported for fourth research question, these findings indicate that differences in visual attention are more strongly related to patterns of lexical competition than to the language context itself. This interpretation is consistent with context-dependent accounts of bilingual processing in which linguistic structures influence attention primarily when they are activated by the language currently in use (Kroll et al., 2006).

5.6 Theoretical implications

The present findings contribute to ongoing debates concerning linguistic relativity by showing that grammatical gender cues affect perceptual attention during real-time language comprehension. German monolingual speakers demonstrated clear gender-congruency effects in visual attention, indicating that grammatical gender information was rapidly integrated during lexical processing. At the same time, the absence of comparable effects when bilingual speakers performed the task in Farsi suggests that such influences are not fixed conceptual properties but emerge dynamically during language use.

This pattern aligns with context-dependent accounts of linguistic relativity, which propose that linguistic structures guide attentional processes during communication without permanently restructuring underlying conceptual representations (Levinson, 2003; Lucy,

1992). The findings are also consistent with frameworks such as Thinking for Speaking (Slobin, 1996) and Seeing for Speaking (Mertins, 2018), which emphasize that language-specific structures shape how speakers attend to information during language use.

From the perspective of bilingual cognition, the results further indicate that grammatical gender representations acquired in a later-learned language can become integrated into online processing even when the first language lacks such a system. However, the weaker effects observed for bilingual speakers compared with monolinguals suggest that the strength of these representations may depend on language experience and the degree of entrenchment of grammatical structures. Together, the findings highlight the dynamic interaction between linguistic structure, bilingual experience, and attentional processes during comprehension.

The findings may also have implications for second language learning and cross-linguistic influence. Learners whose first language lacks grammatical gender often experience difficulties acquiring gender systems in a later-learned language (Sabourin et al., 2006; Shimanskaya & Slabakova, 2019). At the same time, conceptual associations linked to grammatical gender may interact with learners' intuitions about objects during acquisition of a later-learned language (Chen & Faitaki, 2025). The present results suggest that sensitivity to gender cues can nevertheless develop through extensive exposure to the second language.

5.7 Limitations

Several limitations of the present study should be acknowledged. First, although the adjective–noun combinations were selected using corpus-based likelihood ratios from the COSMAS corpus to ensure natural collocations in German, it was not possible to achieve perfectly balanced collocation strengths across all stimulus pairs. Differences in natural co-occurrence probabilities between items may therefore have influenced attentional patterns to some extent. In addition, the absence of a comprehensive and comparable corpus resource for Farsi limited the possibility of conducting equivalent corpus-based collocation analyses across

the two languages. Future research could address this issue by further refining stimulus selection and by developing more extensive corpus-based resources for cross-linguistic stimulus control.

Second, the study did not include a monolingual Farsi control group. Such a comparison would have allowed a more direct evaluation of whether the absence of gender effects in the Farsi condition reflects properties of the language itself or characteristics of the bilingual participants. Although the original research design envisaged collecting eye-tracking data from monolingual Farsi speakers, practical and contextual constraints prevented the implementation of this component. Future studies, including monolingual speakers of genderless languages, would therefore provide an important extension of the present findings.

5.8 Conclusion

The present study examined whether grammatical gender influences visual attention during language processing in German monolingual speakers and Farsi–German bilingual speakers. Using a Visual World eye-tracking paradigm, the findings showed that grammatical gender cues guided visual attention in German monolingual participants and, to a weaker extent, in bilingual speakers performing the task in German. In contrast, no systematic gender-congruency effects emerged when the task was performed in Farsi, a language without grammatical gender. These results suggest that grammatical gender can influence perceptual attention during real-time language comprehension, but that such effects depend on both the linguistic structure of the language being processed and the speaker's linguistic experience.

Appendix A: Stimulus Combinations: nouns, adjectives, and LLR scores

Table A1: Noun–adjective combinations for masculine nouns, including LLR scores

Noun	Critical ADJ	Competitor F	Competitor N	Distractor
Wagen	Gepanzert (4246)	Schick (506)	Alt (1205)	Lieb (0)
Roboter	Schwer (175)	Niedlich (111)	Schlau (148)	Zart (0)
Trip	Gefährlich (58)	Schön (82)	Toll (57)	Menschlich (0)
Tag	Hart (1065)	Sommerlich (182)	Gesellig (45)	Zerbrechlich (0)
Gesang	Rauchig (134)	Zart (102)	Toll (454)	Schlau (0)
Teppich	Groß (104)	Schön (41)	Alt (113)	Kindlich (0)
Bilderrahmen	Hölzern (74)	Schön (48)	Antik (43)	Lieb (0)
Wein	Kräftig (124)	Süß (1079)	Trocken (828)	Kurz (0)
Stiefel	Robust (82)	Schick (121)	Alt (88)	Liebevoll (0)
Preis	Stabil (160)	Schön (65)	Toll (243)	Blumig (0)

Table A2: Noun–adjective combinations for feminine nouns, including LLR scores

Noun	Critical ADJ	Competitor M	Competitor N	Distractor
Kleidung	Schick (1508)	Robust (790)	Alt (928)	Global (0)
Stimme	Schön (2662)	Rauchig (7131)	Menschlich (7750)	Gemischt (0)
Puppe	Niedlich (82)	Hölzern (75)	Alt (170)	Rauchig (0)
Sprache	Blumig (1475)	Schwer (1505)	Menschlich (1237)	Geöffnet (0)
Mühe	Lieb (11886)	Groß (48149)	Viel (110129)	Hölzern (0)
Figur	Zart (65)	Stabil (65)	Kindlich (51)	Gesellig (0)
Stimmung	Sommerlich (775)	Gefährlich (75)	Toll (138)	Gepanzert (0)
Vase	Zerbrechlich (66)	Groß (65)	Antik (405)	Hölzern (0)
Harmonie	Süß (56)	Groß (554)	Toll (54)	Schlau (0)

Umarmung	Liebevoll (1064)	Kräftig (132)	Kurz (122)	Stabil (0)

Table A3: Noun–adjective combinations for neutral nouns, including LLR scores

Nouns	Critical ADJ	Competitor M	Competitor F	Distractor
Sofa	Alt (1663)	Groß (159)	Schick (70)	Gefährlich (0)
Buch	Schlau (555)	Groß (298)	Schön (3701)	Kräftig (0)
Haar	Kurz (11195)	Kräftig (287)	Schön (498)	Gepanzert (80)
Gesicht	Kindlich (689)	Hart (83)	Zart (429)	Blumig (0)
Auto	Toll (1362)	Groß (766)	Schön (1380)	Zart (0)
Gefühl	Menschlich (395)	Gefährlich (55)	Süß (245)	Robust (0)
Bauwerk	Antik (207)	Hölzern (652)	Schön (235)	Niedlich (0)
Leben	Gesellig (1394)	Hart (1840)	Süß (5384)	Rauchig (0)
Gehäuse	Geöffnet (76)	Stabil (146)	Schick (57)	Liebevoll (0)
Papier	Trocken (75)	Stabil (52)	Schön (202)	Sommerlich (0)

Table A4: Visual Stimuli: adjectives and mean ratings on a gendered Likert scale

Masculine	Feminine	Neuter
Gepanzert (5.8)	Zart (1.6)	Alt (4.4)
Gefährlich (5.55)	Blumig (1.9)	Schlau (3.89)
Hart (5.34)	Lieb (2.25)	Kurz (4.2)
Rauchig (5.3)	Liebevoll (2.23)	Kindlich (4)
Groß (5.29)	Süß (2.42)	Toll (3.84)
Hölzern (5.27)	Sommerlich (2.44)	Menschlich (4)
Kräftig (5.2)	Zerbrechlich (2.46)	Antic (4)
Robust (5.19)	Schön (2.48)	Gesellig (4)
Stabil (5.17)	Niedlich (2.53)	Geöffnet (4)
Schwer (5.14)	Schick (2.53)	Trocken (4.37)

Appendix B: Bonferroni-corrected post-hoc comparisons

Table B1: Bonferroni post-hoc comparisons for AOI within grammatical gender (German monolinguals)

Term	Estimate β	*SE*	*z*	*p*
der: Critical − Comp_F	.137	.071	1.918	.165
der: Critical − Comp_N	.226	.071	3.180	.004
der: Critical − Distractor	.356	.071	4.993	< .001
die: Critical − Comp_M	.120	.071	1.686	.276
die: Critical − Comp_N	.324	.071	4.552	< .001
die: Critical − Distractor	.167	.071	2.348	.057
das: Critical − Comp_F	.485	.071	6.811	< .001
das: Critical − Comp_M	.165	.071	2.322	.061
das: Critical − Distractor	.334	.071	4.685	< .001

Table B2: Bonferroni post-hoc comparisons for AOI within grammatical gender (Bilinguals tested in German)

Contrast	Estimate β	*SE*	*z*	*p*
der: Critical – Comp_F	.058	.033	1.74	.165
der: Critical – Comp_N	.147	.034	4.37	< .001
der: Critical – Distractor	.167	.034	5.00	< .001
die: Critical – Comp_M	.077	.033	2.32	.061
die: Critical – Comp_N	.164	.033	4.92	< .001
die: Critical – Distractor	.063	.033	1.89	.117
das: Critical – Comp_F	.214	.033	6.45	< .001
das: Critical – Comp_M	.100	.033	3.01	.008
das: Critical – Distractor	.127	.033	3. 82	< .001

Table B3: Bonferroni post-hoc comparisons for AOI within grammatical gender (Bilinguals tested in Farsi)

Contrast	Estimate β	*SE*	*z*	*p*
der: Distractor − Critical	-.120	.128	-.94	1.000
der: Comp_N − Critical	-.065	.123	-.53	1.000
der: Distractor − Critical	-.078	.120	-.65	1.000
die: Comp_M − Critical	-.277	.144	-1.92	.163
die: Comp_N − Critical	-.388	.146	-2.66	.023
die: Distractor − Critical	-.208	.143	-1.45	.443
das: Comp_F − Critical	-.145	.134	-1.08	.841
das: Comp_M − Critical	.013	.127	.11	1.000
das: Distractor − Critical	.012	.127	.09	1.000

Table B4.1: Bonferroni post-hoc comparisons within groups (bilingual-German vs. monolingual)

Group	Contrast	Estimate β	*SE*	*z*	*p*
Bilingual_German	Competitor − Critical	-1.301	.145	-8.96	< .001
Monolingual	Competitor − Critical	2.498	.182	13.70	< .001

Table B4.2: Bonferroni post-hoc comparisons between groups (bilingual-German vs. monolingual)

Condition/AOI	Contrast	Estimate β	*SE*	*z*	*p*
Competitor	Bilingual_German − Monolingual	-3.716	.116	-31.92	< .001
Target	Bilingual_German − Monolingual	.082	.207	.40	.691

Table B5.1: Bonferroni post-hoc comparisons within groups (bilingual-German vs. bilingual-Farsi)

Group	Contrast	Estimate $\boldsymbol{\beta}$	*SE*	*z*	*p*
Bilingual_German	Target − Competitor	.134	.052	2.57	.010
Bilingual_Farsi	Target − Competitor	.148	.069	2.15	.031

Table B5.2: Bonferroni post-hoc comparisons between groups (H5: bilingual-German vs. bilingual-Farsi)

Congruency	Contrast	Estimate β	*SE*	*z*	*p*
Target	Bilingual_German − Bilingual_Farsi	-.078	.162	-.48	.630
Competitor	Bilingual_German − Bilingual_Farsi	-.064	.154	-.42	.678

Figure C1: Predicted log-transformed dwell time across areas of interest (AOIs) for German monolinguals as a function of grammatical gender. Error bars represent 95% confidence intervals.

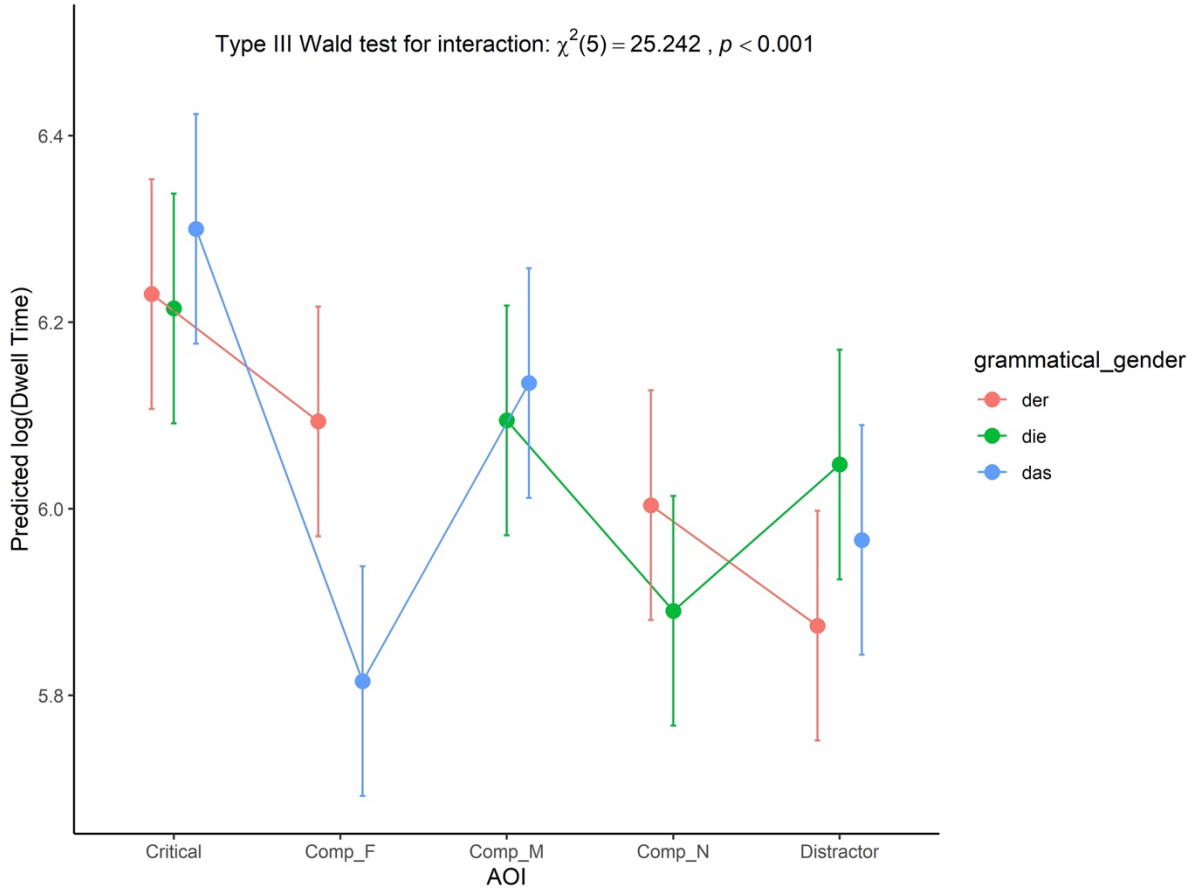


Figure C2: Predicted log-transformed dwell time across areas of interest (AOIs) for Farsi-German bilinguals tested in German as a function of grammatical gender. Error bars represent 95% confidence intervals.

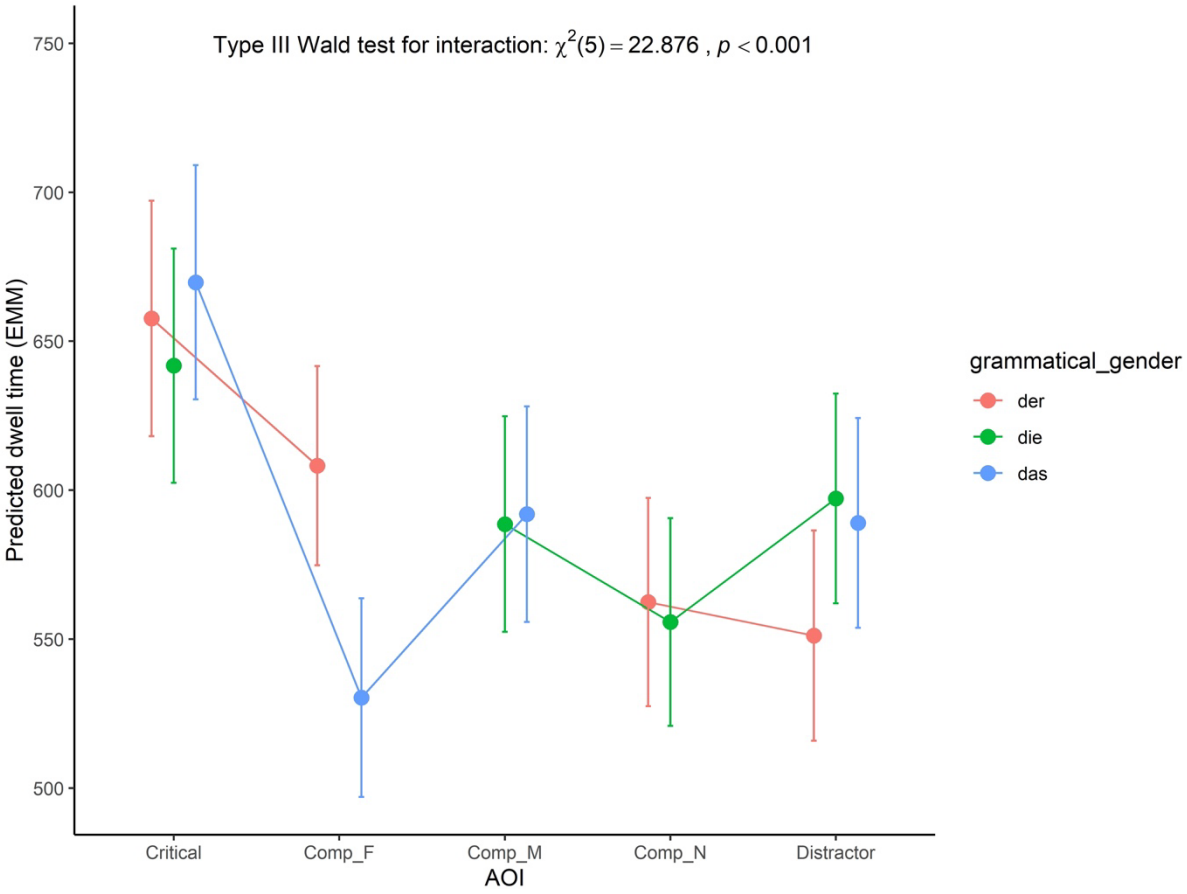


Figure C3: Predicted log-transformed dwell time across areas of interest (AOIs) for Farsi-German bilinguals tested in Farsi as a function of grammatical gender. Error bars represent 95% confidence intervals.

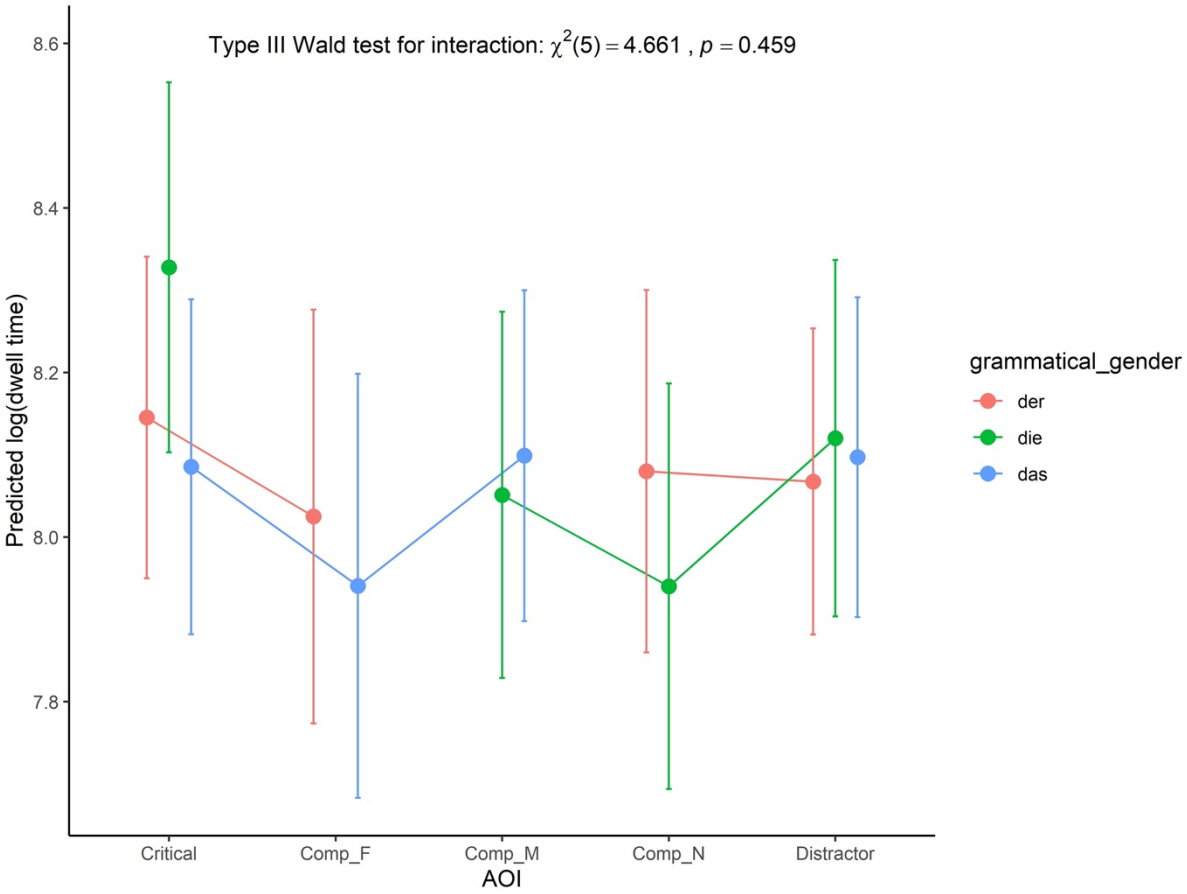
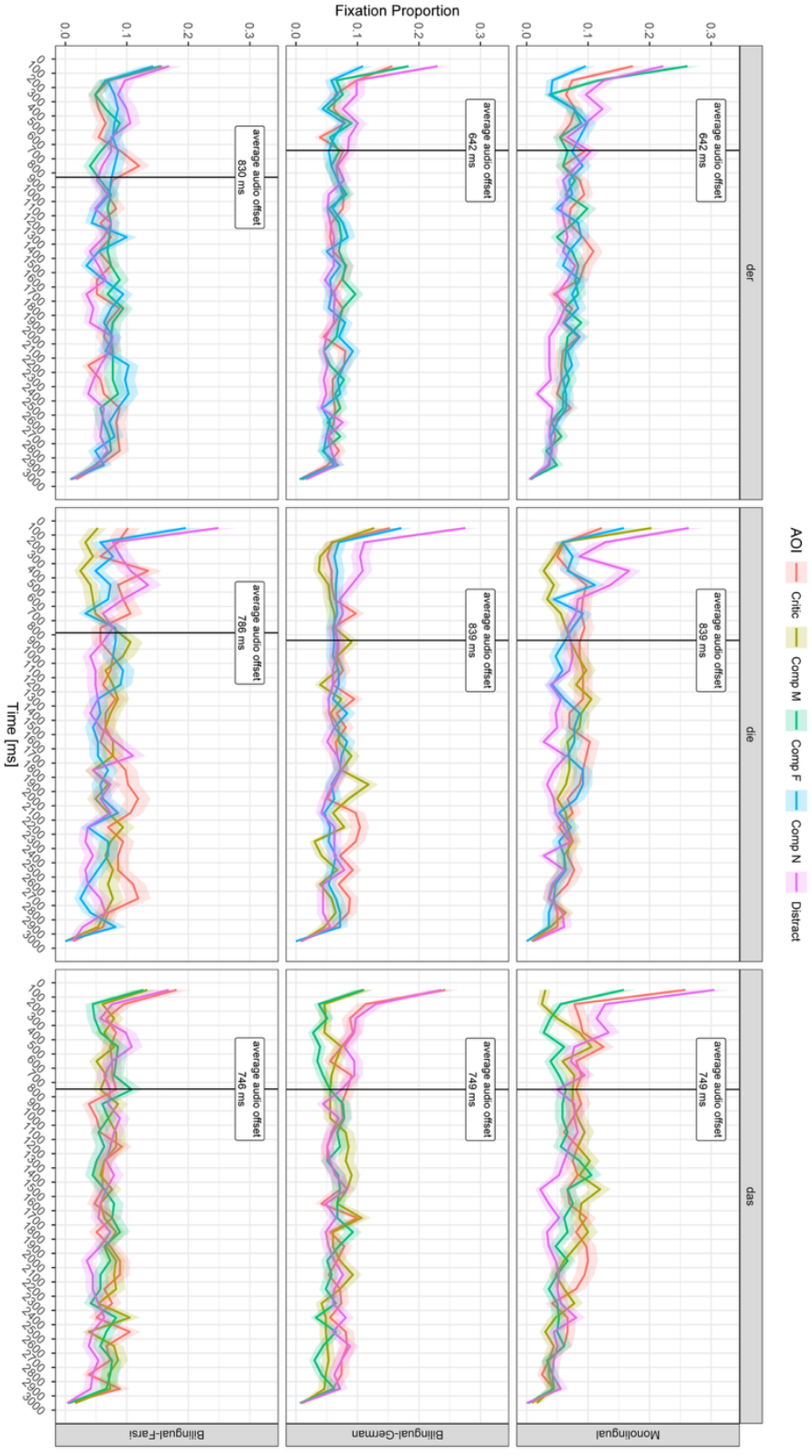


Figure C4: Fixation proportions across areas of interest (AOIs) over time for each grammatical gender condition and participant group. Vertical lines indicate the average offset of the auditory noun. Shaded areas represent ± 1 *SE*.



References

Andonova, E., Gosheva, A., Janyan, A., & Schaffai, J. S. (2007). Second language gender system affects first language gender classification. In I. Kecskes & L. Albertazzi (Eds.), *Cognitive aspects of bilingualism* (pp. 271-299). Springer.

Athanasopoulos, P., & Bylund, E. (2013). Cognitive restructuring. Psychophysical measurement of time perception in bilinguals. *Bilingualism. Language and Cognition, 26*(4), 809–818. https://doi.org/10.1017/S1366728922000876.

Barr, D. J., Levy, R., Scheepers, Ch., & 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.

Bassetti, B. (2007). Bilingualism and thought. Grammatical gender and concepts of objects in Italian-German bilingual children. *International Journal of Bilingualism, 11*(3), 251-273. https://doi.org/10.1177/13670069070110030101.

Bassetti, B., & Filipović, L. (2022). Researching language and cognition in bilinguals. *International Journal of Bilingualism, 26*(1), 3-12. https://doi.org/10.1177/13670069211022860.

Bassetti, B., & Nicoladis, E. (2016). Research on grammatical gender and thought in early and emergent bilinguals. *International Journal of Bilingualism, 20*(1), 3–16. https://doi.org/10.1177/1367006915576824.

Beller, S., Fadnes Brattebø, K., Osland Lavik, K., Drønen Reigstad, R., & Bender, A. (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, A., Beller, S., & Klauer, K. Ch. (2016a). Crossing grammar and biology for gender categorizations. Investigating the gender congruency effect in generic nouns for animates. *Journal of Cognitive Psychology, 28*(5), 530–558.

Bender, A., Beller, S., & Klauer, K. Ch. (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.

Bialystok, E. (2017). The bilingual adaptation. How minds accommodate experience. *Psychological Bulletin, 143*(3), 233–262. https://doi.org/10.1037/bul0000099.

Birdsong, D. (2018). Plasticity, variability and age in second language acquisition and bilingualism. *Frontiers in Psychology, 9*, e81. https://doi.org/10.3389/fpsyg.2018.00081.

Boroditsky, L., Schmidt, L. A., & Phillips, W. (2003). Sex, syntax, and semantics. In D. Gentner & S. Goldin-Meadow (Eds.), *Language in mind. Advances in the study of language and thought* (pp. 61-80). MIT Press. https://doi.org/10.7551/mitpress/4117.003.0010.

Boutonnet, B., Athanasopoulos, P., & Thierry, G. (2012). Unconscious effects of grammatical gender during object categorisation. *Brain Research, 1479*, 72–79. https://doi.org/10.1016/j.brainres.2012.08.044.

Bylund, E., & Norrman, G. (2023). The CPH is dead. Long live the critical period hypothesis. *Brain and Language, 246*, e105341. https://doi.org/10.1016/j.bandl.2023.105341.

Caldwell-Harris, C. L., & MacWhinney, B. (2023). Age effects in second language acquisition. Expanding the emergentist account. *Brain and Language, 241*, e105269. https://doi.org/10.1016/j.bandl.2023.105269.

Chen, Z., & Faitaki, F. (2025). Effects of the French grammatical gender system on bilingual adults' perception of objects. *Bilingualism. Language and Cognition, 28*(2), 401–414. https://doi.org/10.1017/S1366728924000464.

Consonni, M., Cafiero, R., Marin, D., Tettamanti, M., Iadanza, A., Fabbro, F., & Perani, D. (2013). Neural convergence for language comprehension and grammatical class production in highly proficient bilinguals is independent of age of acquisition. *Cortex, 49*(5), 1252-1258. https://doi.org/10.1016/j.cortex.2012.04.009.

Corbett, G. G. (1991). *Gender*. Cambridge University Press.

COSMAS I/II (Corpus Search, Management and Analysis System: 1991-2022 Leibniz-Institut für Deutsche Sprache, Mannheim). Retrieved on March 30, 2026, from http://www.ids-mannheim.de/cosmas2/.

Cubelli, R., Paolieri, D., Lotto, L., & Job, R. (2011). The effect of grammatical gender on object categorization. *Journal of Experimental Psychology. Learning, Memory, and Cognition, 37*(2), 449–460.

De Bot, K., Lowie, W., & Verspoor, M. (2007). A dynamic systems theory approach to second language acquisition. *Bilingualism. Language and Cognition, 10*(1), 7–21. https://doi.org./10.1017/S1366728906002732.

Delucchi Danhier, R., & Mertins, B. (2018). Psycholinguistische Grundlagen der Inklusion. Schwerpunkt Bilingualismus. In S. Hußmann & B. Welzel (Eds.), *DoProfiL-Das Dortmunder Profil für inklusionsorientierte Lehrerinnen- und Lehrerbildung* (pp. 161-178). Waxmann.

Dussias, P. E., & Sagarra, N. (2007). The effect of exposure on syntactic parsing in Spanish–English bilinguals. *Bilingualism. Language and Cognition, 10*(1), 101–116. https://doi.org/10.1017/S1366728906002847.

Elpers, N., Jensen, G., & Holmes, K. J. (2022). Does grammatical gender affect object concepts? Registered replication of Phillips and Boroditsky (2003). *Journal of Memory and Language, 127*, e104357. https://doi.org/10.1016/j.jml.2022.104357.

Fatemi, S. M., Mertins, B., & Schröder, V. (2025). The influence of grammatical gender on cognition. The case of German and Farsi. *Open Linguistics, 11*(1), e20250071. https://doi.org/10.1515/opli-2025-0071.

Flecken, M., Athanasopoulos, P., Kuipers, J. R., & Thierry, G. (2015). On the road to somewhere. Brain potentials reflect language effects on motion event perception. *Cognition, 141*, 41–51. https://doi.org/10.1016/j.cognition.2015.04.006.

Flege, J. E. (2018). It's input that matters most, not age. *Bilingualism. Language and Cognition, 21*(5), 919-920. https://doi.org/10.1017/S136672891800010X.

Fuchs, Z. (2022) Eyetracking evidence for heritage speakers' access to abstract syntactic agreement features in real-time processing. *Frontiers in Psychology, 13*, e960376. https://doi.org/10.3389/fpsyg.2022.960376.

Grosjean, F. (2001). The bilingual's language modes. In J. Nicol (Ed.), *One mind, two languages. Bilingual language processing* (pp. 1-22). Blackwell.

Grosjean, F. (2008). *Studying bilinguals*. Oxford University Press.

Grosjean, F. (2010). *Bilingual. Life and reality*. Harvard University Press. https://doi.org/10.4159/9780674056459.

Grotjahn, R. (2016). Alter. In E. Burwitz-Melzer, G. Mehlhorn, C. Riemer, K.-R. Bausch & H.-J. Krumm (Eds.), *Handbuch Fremdsprachenunterricht* (6th ed.; pp. 251–254). Narr Francke Attempto.

Gygax, P., Gabriel, U., Sarrasin, O., Oakhill, J., & Garnham, A. (2008). Generically intended, but specifically interpreted. When beauticians, musicians, and mechanics are all men. *Language and Cognitive Processes, 23*(3), 464–485.

Holm, St. (1979). A simple sequentially rejective multiple test procedure. *Scandinavian Journal of Statistics, 6*(2), 65–70.

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

Huettig, F., & McQueen, J. M. (2007). The tug of war between phonological, semantic and shape information in language-mediated visual search. *Journal of Memory and Language, 57*(4), 460–482. https://doi.org/10.1016/j.jml.2007.02.001.

Huettig, F., Rommers, J., & Meyer, A. S. (2011). Using the visual world paradigm to study language processing. A review and critical evaluation. *Acta Psychologica, 137*(2), 151–171. https://doi.org/10.1016/j.actpsy.2010.11.003.

Kroll, J. F., Bobb, S. C., & Wodniecka, Z. (2006). Language selectivity is the exception, not the rule. Arguments against a fixed locus of language selection in bilingual speech. *Bilingualism. Language and Cognition, 9*(2), 119–135. https://doi.org/10.1017/S1366728906002483.

Kurinski, E., & Sera, M. D. (2011). Does learning Spanish grammatical gender change English-speaking adults' categorization of inanimate objects? *Bilingualism. Language and Cognition, 14*(2), 203–220. https://doi.org/10.1017/S1366728910000179.

Lambelet, A. (2016). 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.

Langacker, R. W. (1987). *Foundations of cognitive grammar. Volume I: Theoretical prerequisites*. Stanford University Press.

Leibniz-Institut für Deutsche Sprache (n. d.). Deutsches Referenzkorpus / Archiv der Korpora geschriebener Gegenwartssprache 2025-I-RC1 (RC vom 31.01.2025). [German Reference Corpus / Archive of Corpora of Written Contemporary Language 2025-I-

RC1 (RC from January 31, 2025)]. Leibniz-Institut für Deutsche Sprache. Retrieved on March 30, 2026, from www.ids-mannheim.de/dereko.

Levinson, St. C. (2003). *Space in language and cognition. Explorations in cognitive diversity*. Cambridge University Press. https://doi.org/10.1017/CBO9780511613609.

Lewisohn, L. (Ed.). (2012). *The heritage of Sufism. Classical Persian Sufism from its origins to Rumi (700-1300)*. Oneworld Publications.

Lucy, J. A. (1992). *Language diversity and thought. A reformulation of the linguistic relativity hypothesis*. Cambridge University Press. https://doi.org/10.1017/CBO9780511620843.

Luk, G., & Bialystok, E. (2013). Bilingualism is not a categorical variable. Interaction between language proficiency and usage. *Journal of Cognitive Psychology, 25*(5), 605–621. https://doi.org/10.1080/20445911.2013.795574.

Maciuszek, J., Polak, M., & Świątkowska, N. (2019). Grammatical gender influences semantic categorization and implicit cognition in Polish. *Frontiers in Psychology, 10*, e2208. https://doi.org/10.3389/fpsyg.2019.02208.

McQueen, J. M., Viebahn, M. C. (2007). Tracking the recognition of spoken words by tracking looks to printed words. *Quarterly Journal of Experimental Psychology, 60*(5), 661–671. https://doi.org/10.1080/17470210601183890.

Mertins, B. (2018). *Sprache und Kognition. Ereigniskonzeptualisierung im Deutschen und Tschechischen*. De Gruyter. https://doi.org/10.1515/9783110615746.

Mikić Ljubi, J., Matkovič, A., Bon, J., & Kanjuo Mrčela, A. (2022). The effects of grammatical gender on the processing of occupational role names in Slovene. An event-related potential study. *Frontiers in Psychology, 13*, e1010708. https://doi.org/10.3389/fpsyg.2022.1010708.

Montefinese, M., Ambrosini, E., & Roivainen, E. (2019). No grammatical gender effect on affective ratings. Evidence from Italian and German languages. *Cognition and Emotion, 33*(4), 848-854.

Nicoladis, E., Da Costa, N. & Foursha-Stevenson, C. (2016). 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.

Odermann, K., & Mertins, B. (2026). *Semantisches Bewusstsein bei Vorschulkindern. Wie Bilingualismus das metalinguistische Wissen stärkt*. Zeitschrift für Erziehungswissenschaft. https://doi.org/10.1007/s11618-026-01379-7.

Pier Salverda, A., Brown, M., & Tanenhaus, M. K. (2011). A goal-based perspective on eye movements in visual world studies. *Acta Psychologica, 137*(2), 172-180. https://doi.org/10.1016/j.actpsy.2010.09.010.

Pinker, St. (1994). On language. *Journal of Cognitive Neuroscience, 6*(1), 92-98. https://doi.org/10.1162/jocn.1994.6.1.92.

Prewitt-Freilino, J. L., Caswell, A. T., & Laakso, E. K. (2012). The gendering of language. A comparison of gender equality in countries with gendered, natural gender, and genderless languages. *Sex Roles, 66*, 268–281. https://doi.org/10.1007/s11199-011-0083-5.

Rayner, K. (1998). Eye movements in reading and information processing. 20 years of research. *Psychological Bulletin, 124*(3), 372–422.

Riordan, B., Dye, M., & Jones, M. N. (2015). Grammatical number processing and anticipatory eye movements are not tightly coordinated in English spoken language comprehension. *Frontiers in Psychology, 6*, e590. https://doi.org/10.3389/fpsyg.2015.00590.

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

Samuel, S., Cole, G., & Eacott, M. J. (2019). Grammatical gender and linguistic relativity. A systematic review. *Psychonomic Bulletin & Review, 26*, 1767-1786. https://doi.org/10.3758/s13423-019-01652-3.

Sato, S., & Athanasopoulos, P. (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, S., Casaponsa, A. & Athanasopoulos, P. (2020). Flexing gender perception. Brain potentials reveal the cognitive permeability of grammatical information. *Cognitive Science, 44*, e12884. https://doi.org/10.1111/cogs.12884.

Semenuks, A., Phillips, W., Dalca, I., Kim, C., & Boroditsky, L. (2017). Effects of grammatical gender on object description. In G. Gunzelmann, A. Howes, T. Tenbrink & E. Davelaar (Eds.), *Proceedings of the 39th annual meeting of the Cognitive Science Society* (pp. 1060–1065). Cognitive Science Society.

Sera, M. D., Elieff, C., Forbes, J., Burch, M. C., Rodríguez, W., & Dubois, D. P. (2002). When language affects cognition and when it does not. An analysis of grammatical gender and classification. *Journal of Experimental Psychology. General, 131*(3), 377–397. https://doi.org/10.1037/0096-3445.131.3.377.

Shimanskaya, E., & Slabakova, R. (2019). L1–L2 differences in the L2 classroom. Anticipating anglophone learners' difficulties with French pronoun interpretation. *Language Teaching Research, 23*(2), 259-277. https://doi.org/10.1177/1362168817739650.

Singleton, D. (2023). Bilingualism and age. In C. A. Chapelle & D. Gabryś-Barker (Eds.), *The Encyclopedia of Applied Linguistics (Bilingualism and Multilingualism)*. https://doi.org/10.1002/9781405198431.wbeal0097.pub2.

Slobin, D. I. (1996). From "thought and language" to "thinking for speaking". In J. J. Gumperz & St. C. Levinson (Eds.), *Rethinking linguistic relativity* (pp. 70-96). Cambridge University Press.

Stahlberg, D., Sczesny, S., & Braun, F. (2001). Name your favorite musician. Effects of masculine generics and of their alternatives in German. *Journal of Language and Social Psychology, 20*(4), 464-469. https://doi.org/10.1177/0261927X01020004004.

Tanenhaus, M. K., Spivey-Knowlton, M. J., Eberhard, K. M., & Sedivy, J. C. (1995). Integration of visual and linguistic information in spoken language comprehension. *Science, 268*(5217), 1632–1634. https://doi.org/10.1126/science.7777863.

Vanhove, J. (2014). The Critical Period Hypothesis in Second Language Acquisition: A Statistical Critique and a Reanalysis. *PLOS ONE, 9*(7), e102922. https://doi.org/10.1371/journal.pone.0102922.

Vernich, L. (2019). Does learning a foreign language affect object categorization in native speakers of a language with grammatical gender? The case of Lithuanian speakers 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.

Vigliocco, G., Vinson, D. P., Paganelli, F., & Dworzynski, K. (2005). Grammatical gender effects on cognition. Implications for language learning and language use. *Journal of Experimental Psychology. General, 134*(4), 501–520. https://doi.org/10.1037/0096-3445.134.4.501.

Whorf, B. L. (1956 [2012]). *Language, thought, and reality; selected writings* (J. B. Carroll, St. C. Levinson & P. Lee, Eds.; 2nd ed.). The MIT Press.

Yarshater, E. (Ed.). (1988). *Persian literature*. Bibliotheca Persica / SUNY Press.

Chapter 10 Discussion and Conclusion

Across the five studies presented in this dissertation, converging evidence from typological analysis, systematic literature review, theoretical inquiry, behavioral experimentation, and eye-tracking data provides a comprehensive account of the relationship between grammatical gender and cognition. Taken together, the findings demonstrate that grammatical gender can systematically influence conceptual processing; however, this influence is neither universal nor uniform, but is constrained by linguistic structure, methodological design, and speaker-related factors.

The first two studies revealed fundamental limitations in the existing literature on linguistic relativity and grammatical gender. Study 1 demonstrated that grammatical gender systems across Indo-European languages are more diverse than commonly represented in empirical research, with a substantial proportion of languages—particularly within the Indo-Aryan and Iranian branches—remaining unstudied. Study 2 further showed that empirical investigations between 2013 and 2023 have been restricted to a narrow set of languages, predominantly German and English, and have relied heavily on methodological paradigms that impose binary gender frameworks, even in languages with more complex grammatical gender systems. This is particularly surprising, as German is widely studied, yet it is consistently reduced to a binary system, a simplification that warrants explicit discussion. Together, these findings point to systematic typological and methodological biases in the field, raising concerns about the generalizability of previous conclusions. These typological and methodological constraints are crucial for interpreting the present experimental findings, as they underscore the importance of extending research beyond well-studied language systems and simplified gender frameworks.

Building on this foundation, Study 3 critically evaluated the theoretical and empirical landscape of grammatical gender research in bilingual and multilingual contexts. The review identified substantial inconsistencies in previous findings, which were largely attributable to methodological shortcomings, including heterogeneous participant samples, insufficient control of proficiency and language experience, and limited use of online processing measures. Importantly, this study emphasized that grammatical gender effects are not restricted to explicit conceptual judgments, such as voice assignment or sex attribution, but may also emerge in implicit and early stages of processing, underscoring the need for more sensitive experimental paradigms capable of capturing real-time cognitive dynamics.

Taken together, Studies 1–3 highlight several critical gaps that guide the design of the empirical investigations in this dissertation. First, they point to the need to extend research beyond frequently studied languages by including understudied linguistic systems such as Farsi. Second, they demonstrate the importance of accounting for the full complexity of grammatical gender systems, rather than reducing them to binary categories, thereby motivating the explicit inclusion of neuter gender as an independent category. Third, they emphasize the necessity of systematically investigating bilingual speakers in order to better understand how grammatical gender operates across different language contexts. Finally, they underscore the importance of employing online methodologies capable of capturing real-time processing, as opposed to relying solely on offline or explicit measures.

The empirical studies presented in this dissertation directly address these gaps. Study 4 demonstrated that grammatical gender influences conceptualization in monolingual speakers of German, particularly at early stages of response generation. The finding that gender-congruent effects were strongest for the first adjective supports the interpretation that grammatical gender operates as an automatized cue during conceptual processing. Crucially, the inclusion of neuter gender as an independent category showed that grammatical gender systems are not inherently binary; rather, accounting for the full gender system—including neuter—captures a conceptually meaningful category and more accurately reflects the structure of languages such as German. In contrast, Farsi speakers—whose language does not encode grammatical gender—did not exhibit systematic gender-based patterns, indicating that such effects are grounded in language-specific grammatical structures and are not reducible to universal semantic associations.

Study 5 extended these findings by examining grammatical gender effects in real-time language processing using eye-tracking methodology. Importantly, the duration of visual attention to gender-congruent items (i.e., longer dwell times) emerged as a key indicator of grammatical gender effects, suggesting that gender-congruent information captures and maintains attentional resources during comprehension. German monolingual speakers displayed clear gender-congruency effects in visual attention, demonstrating that grammatical gender information is rapidly integrated and can guide perceptual attention. Farsi–German bilingual speakers exhibited similar but attenuated effects when processing German, suggesting that grammatical gender representations acquired in a language in adulthood, particularly under conditions of reduced immersion compared to German monolinguals, can become integrated into online processing, albeit with reduced strength. Importantly, these effects disappeared

when the Farsi-German participants performed the task in Farsi, even after a short break, providing strong evidence that the influence of grammatical gender is dynamically modulated by the linguistic context in use rather than reflecting stable, language-independent conceptual representations.

Taken together, these findings converge on a central conclusion: grammatical gender influences cognition in a systematic yet context-dependent manner. These effects are most clearly observed during language processing, particularly at early and implicit stages, as reflected in both behavioral primacy effects and eye-tracking measures of visual attention. Crucially, this pattern indicates that grammatical gender operates as an initial organizing cue during processing rather than as a post hoc interpretive factor. At the same time, the absence of comparable effects in a genderless language and the attenuation of effects in bilingual speakers indicate that these influences depend on the interaction between linguistic structure, language experience, and task demands. The present dissertation supports a dynamic view of linguistic relativity, according to which language influences cognition during real-time processing, but does not permanently alter underlying conceptual representations (e.g., Levinson, 2003; Lucy, 1992; Mertins, 2018; Slobin, 1996).

10.1 Theoretical implications

The theoretical implications of this dissertation are grounded in the framework developed in Chapter 2 and are evaluated in light of the conceptual distinctions outlined there. In particular, the findings are interpreted with respect to the proposed multi-level model of cognition, the distinction between conceptual representation and online processing, and the role of grammatical structure in shaping attentional and conceptual processes. Building on this foundation, the discussion first situates the results within a dynamic and constrained account of linguistic relativity, examining to what extent the empirical findings support a processing-based interpretation of language–cognition interaction. It then considers how grammatical gender operates across different cognitive levels, with particular attention to the distinction between stable conceptual representations and context-dependent activation during processing. Furthermore, the findings are interpreted within cognitive-linguistic approaches that treat grammar as a meaningful and usage-based system, as well as psycholinguistic models that assume the rapid and automatic activation of grammatical information. Finally, the discussion addresses bilingual cognition by examining how grammatical gender acquired in a language

with reduced exposure compared to monolingual speakers is integrated into processing and modulated by the immediate language context.

10.1.1 Core Position: Soft, Dynamic Relativity

The findings of this dissertation support a dynamic and constrained interpretation of linguistic relativity. Grammatical gender did not emerge as a deterministic force structuring thought; rather, its effects were systematic but context-dependent, shaped by linguistic structure, task demands, and speaker experience. This pattern aligns with contemporary "weak" accounts of linguistic relativity, according to which language biases attention and conceptualization without overriding general cognitive capacities (Athanasopoulos & Bylund, 2023; Lucy, 1992, 2014; Slobin, 1996, 2004; Wolff & Holmes, 2011).

Across the empirical studies presented in this dissertation, grammatical gender effects were most evident during online processing, particularly in early stages of production and in real-time visual attention. In Study 4, gender-congruent effects were strongest for the first adjective produced, indicating that grammatical gender influenced conceptualization at the earliest stage of response generation. In the eye-tracking study, visual attention was systematically guided by grammatical gender cues, as reflected in increased dwell time on gender-congruent adjectives. Crucially, the comparison across language contexts in the bilingual group (Study 5) showed that these effects were present when the task was performed in German but disappeared when the participants performed the task in Farsi. This within-participant modulation provides strong evidence that grammatical gender effects are contingent on the currently active linguistic system. Together, these findings suggest that grammatical gender primarily influences how information is attended to and processed during language use, rather than reshaping conceptual representations in a stable or language-independent manner. The modulation of these effects across linguistic contexts—most notably their attenuation in bilingual speakers and absence in Farsi—further supports a view in which language–cognition interactions are dynamically shaped by the currently active linguistic system.

These findings are consistent with the frameworks of *Thinking for Speaking* (Slobin, 1996, 2004) and *Seeing for Speaking* (Mertins, 2018). Specifically, the results of Study 4, in which gender-congruent effects emerged most strongly in the first adjective produced, align with the notion that language-specific structures guide conceptualization during speech planning. At the same time, the eye-tracking findings, which showed that visual attention was systematically

directed toward gender-congruent adjectives during comprehension, support the view that language influences perceptual attention in real time, as captured by the *Seeing for Speaking* framework. Together, these patterns suggest that grammatical gender shapes both how speakers organize information for production and how they allocate attention during language processing. The results extend these accounts by showing that grammatical gender—beyond lexical semantics—can guide both linguistic encoding and visual attention during comprehension. Eye-tracking evidence demonstrates that gender-congruent information captures and sustains visual attention, suggesting that morphosyntactic structure guides the allocation of attention toward linguistically compatible referents during comprehension.

Taken together, the findings support a model in which grammatical gender shapes cognition in a dynamic, probabilistic, and context-dependent manner, primarily during online processing rather than through permanent restructuring of conceptual knowledge. If grammatical gender were to restructure conceptual representations in a stable manner, one would expect consistent gender-congruent effects across linguistic contexts, including when bilingual Farsi–German speakers perform the task in Farsi, a language that does not encode grammatical gender. However, the absence of such effects in the Farsi condition, as well as their modulation across language contexts in bilingual speakers—emerging in German but not in Farsi—suggests that grammatical gender does not operate at the level of stable conceptual restructuring, but instead exerts its influence in a context-dependent manner during processing. Taken together, the findings suggest that attentional modulation constitutes a central mechanism through which grammatical gender exerts its influence on cognition.

10.1.2 Grammatical Gender Across Cognitive Levels: Representation and Processing

A central theoretical contribution of this dissertation lies in demonstrating that grammatical gender operates across multiple levels of cognition, as defined in Chapter 2. Adopting a multi-layered view of cognition—distinguishing between perceptual processes, conceptualization, and linguistic encoding (Anderson, 2010; Evans & Green, 2018; Eysenck & Keane, 2020) — the present findings show that grammatical gender effects emerge differently depending on the cognitive domain engaged.

In the behavioral study, grammatical gender influenced conceptual associations during lexical retrieval, particularly at early stages of response generation. More specifically, these early effects can be located at the level of lexical access, where grammatical gender information

is retrieved as part of lemma activation and interacts with conceptual selection processes. Within the framework outlined in Chapter 2, this effect can be located at the level of conceptualization, where linguistic structure interacts with the formation and activation of concepts (Evans & Green, 2018; Langacker, 1987, 1991). In contrast, the eye-tracking study revealed effects at the level of perceptual attention, with gender-congruent information guiding gaze allocation during comprehension. Importantly, these effects were observed in relation to written linguistic stimuli rather than non-linguistic visual input, indicating that grammatical gender influences the allocation of attention within the processing of linguistic information. This corresponds to lower-level attentional processes involved in online language use (Mandler, 2007), demonstrating that grammatical gender can influence how visual information is selected and processed in real time.

Crucially, these findings clarify the distinction between conceptual representation and processing, as outlined in Chapter 2. Rather than indicating that grammatical gender leads to stable changes in conceptual representations, the results suggest that it influences how concepts are accessed and activated during language use. If grammatical gender were to reshape conceptual representations in a stable manner, one would expect gender-congruent effects to persist across linguistic contexts, including when bilingual speakers perform the task in a language that does not encode grammatical gender. However, the absence of such effects in the Farsi condition indicates that these influences do not reflect stable changes in conceptual structure, but instead arise during language-specific processing. In this sense, the findings are more consistent with usage-based and activation-based accounts, in which conceptual representations remain available, but their activation is dynamically modulated by the linguistic context. This interpretation is further supported by the cognitive-linguistic view that meaning is not fixed but dynamically constructed in context, emerging through the interaction between linguistic structure and conceptual knowledge (Evans & Green, 2018; Langacker, 2008).

By linking behavioral and eye-tracking evidence to different levels of the cognitive system, the present findings provide a more differentiated account of how grammatical gender interacts with cognition. They show that linguistic effects are not uniform across cognitive domains, but are most pronounced at early and implicit stages of processing, thereby helping to reconcile previously inconsistent findings in the literature, which range from strong claims about the influence of grammatical gender on cognition (e.g., Boroditsky et al., 2003) to more cautious

or null findings emphasizing the role of task demands and methodological variation (e.g., Bender et al., 2011; Vigliocco et al., 2005).

10.1.3 Psycholinguistic Processing and the Automatic Activation of Grammatical Gender

The present findings contribute to psycholinguistic models of language processing by providing evidence for the rapid and automatic activation of grammatical gender, as outlined in Chapter 2. Within Levelt's (1989; Levelt et al., 1999) model of language production, grammatical gender is encoded at the lemma level and is assumed to be retrieved automatically during lexical access. The observed effects in early adjective production are consistent with this assumption, suggesting that gender information becomes available at initial stages of formulation and can influence lexical selection even when gender is not explicitly required by the task.

At the same time, the eye-tracking findings extend this account beyond production to real-time comprehension. The observed dwell-time patterns indicate that grammatical gender information is not only retrieved but also guides attentional allocation during language processing. This supports the view, discussed in Chapter 2, that production and comprehension rely on shared and overlapping mechanisms (Pickering & Garrod, 2013), and that grammatical features can be exploited predictively during comprehension, with such predictive use emerging early in language development (Lew-Williams & Fernald, 2007).

Importantly, the findings are not only compatible with Levelt's model but also align with more interactive and dynamic accounts of language processing. Evidence from both behavioral and eye-tracking data suggests that grammatical gender is activated rapidly and interacts with conceptual and perceptual processes in real time, rather than being processed in strictly sequential stages. This interpretation is consistent with models that assume cascading and interactive activation of linguistic information (Hartsuiker et al., 2004; Mahon et al., 2007; Rapp & Goldrick, 2000), as discussed in Chapter 2.

Taken together, these findings support the view that grammatical gender is an automatically activated feature of lexical representations that can influence both production and comprehension processes. At the same time, they highlight the importance of moving beyond strictly modular accounts such as Levelt's (1999) serial model of speech production, in which

processing proceeds in discrete, non-interacting stages, toward more dynamic models in which grammatical, conceptual, and attentional processes interact during language use.

10.1.4 Bilingual Cognition and Context-Dependent Language Processing

Building on the preceding discussion of processing mechanisms and cognitive levels, the findings from bilingual speakers provide particularly important insights into how grammatical gender operates within a dynamic bilingual system. Farsi–German bilingual speakers showed gender-congruent effects when processing German, although these effects were weaker than those observed in monolingual speakers. Crucially, no such effects were observed when the task was performed in Farsi, indicating that grammatical gender information from German did not transfer to the processing of Farsi lexical items.

This pattern supports context-dependent models of bilingual processing, in which linguistic representations are selectively activated depending on the language currently in use (Grosjean, 2001; Kroll et al., 2016). It also suggests that grammatical gender representations can become integrated into processing, but remain less strongly entrenched than in native speakers, likely due to reduced exposure and language experience.

Rather than supporting a categorical distinction between the presence and absence of grammatical gender effects, the findings point to graded sensitivity shaped by language experience, proficiency, and task context. This interpretation speaks directly to an ongoing debate in bilingualism research concerning the nature of cross-linguistic interaction and the extent to which bilingual language systems are dynamically integrated. While some approaches assume relatively stable and separable language systems (Francis, 2007; Ullman, 2001), others propose a fundamentally dynamic architecture in which both languages remain co-activated and continuously interact (e.g., Grosjean, 2001, 2008; de Bot et al., 2007; Kroll et al., 2016). Importantly, even within this dynamic perspective, there is no single unified account, models differ in how they conceptualize the degree and mechanisms of cross-linguistic interaction (e.g., Dijkstra & van Heuven, 2002; Kroll et al., 2016).

Non-selective access models (e.g., Dijkstra & Van Heuven, 2002) assume parallel activation of both languages and predict widespread cross-linguistic influence. In contrast, language mode frameworks (Grosjean, 2001, 2008) emphasize that the degree of co-activation is context-dependent, allowing one language to be highly active while the other is only weakly engaged. More recent usage-based and adaptive control accounts (De Bot et al., 2007; Kroll et

al., 2016) further conceptualize bilingual processing as a continuum shaped by experience, frequency, and interactional demands.

The present findings are most consistent with these graded, context-sensitive accounts. The emergence of gender effects in German indicates that grammatical gender representations can become integrated into the bilingual system through experience, whereas their absence in Farsi demonstrates that such representations are not obligatorily activated across languages. Instead, grammatical gender appears to function as a probabilistic, experience-dependent cue, whose influence depends on the current language context and the bilingual's history of use.

Taken together, these findings support a dynamic perspective on bilingual cognition and linguistic relativity, highlighting that the influence of grammatical structure is neither fixed nor categorical, but modulated by language context, experience, and patterns of use.

10.2 Methodological Contributions

A central methodological contribution of this dissertation lies in its integrative multi-method design, which addresses several persistent limitations in research on grammatical gender and linguistic relativity, including restricted language sampling, a predominance of binary gender frameworks, and a reliance on isolated experimental paradigms. By combining typological analysis, systematic review, theoretical synthesis, and experimental approaches—both behavioral and eye-tracking—the present work establishes a coherent research program that links broad cross-linguistic patterns to controlled investigations of cognitive processing. This integration allows for a more comprehensive and internally consistent account of how grammatical gender operates across different levels of analysis.

A second key contribution concerns the operationalization of grammatical gender as a multi-category system. Previous research has frequently reduced grammatical gender to binary distinctions, even in languages that encode more complex systems. By systematically incorporating masculine, feminine, and neuter categories into the experimental design, this dissertation demonstrates that such reductions obscure important aspects of linguistic structure. The findings show that neuter constitutes a cognitively meaningful category, thereby underscoring the necessity of aligning experimental design with typological reality and refining theoretical accounts of gender representation.

The dissertation further advances methodological practice through the integration of implicit and online measures of cognition. The use of a property-judgment paradigm (Study 4) enables the investigation of conceptual associations without explicitly directing participants' attention to grammatical gender, capturing more implicit aspects of processing. This is complemented by the application of eye-tracking in a Visual World Paradigm (Study 5), which provides access to real-time processing dynamics and early stages of language comprehension. By combining offline and online measures, the present work offers a fine-grained and temporally sensitive account of grammatical gender effects across cognitive levels.

Importantly, the use of eye-tracking in this domain represents a novel methodological contribution. While previous research has predominantly relied on explicit or offline tasks, the present dissertation demonstrates that eye-tracking can reveal subtle, time-sensitive effects of grammatical gender that remain undetected in traditional paradigms. This approach provides new insights into how linguistic structure guides visual attention during real-time language processing, thereby extending methodological possibilities in the study of linguistic relativity.

Finally, the dissertation contributes to methodological rigor through its cross-linguistic and participant design. By including both monolingual speakers of a gendered language and bilingual speakers whose languages differ in grammatical gender marking, the study enables a more precise examination of language-specific effects. The inclusion of a grammatically genderless language (Farsi) is particularly critical, as it provides a baseline for distinguishing language-driven effects from more general cognitive tendencies.

Taken together, these methodological contributions demonstrate that advancing the study of linguistic relativity requires not only theoretical precision but also methodological diversity, typological sensitivity, and access to real-time processing measures. The present dissertation shows that integrating these dimensions yields a more nuanced and empirically grounded understanding of how grammatical gender interacts with cognition.

Despite these methodological strengths, several limitations should be acknowledged. At the typological level, the analysis (Study 1) is constrained by incomplete documentation, as a substantial proportion of Indo-European languages lack clear evidence regarding grammatical gender systems. Similarly, the systematic and theoretical reviews (Studies 2 and 3) are limited by the availability and heterogeneity of existing research, including inconsistencies in the operationalization of bilingualism and potential publication biases.

At the experimental level, full control over collocational frequency and semantic associations in Studies 4 and 5 was not possible, particularly due to the absence of a comprehensive Farsi corpus, which may have introduced variability across items. In addition, the lack of a Farsi monolingual eye-tracking group limits the symmetry of cross-linguistic comparisons in the online domain. Furthermore, the focus on a specific bilingual configuration (Farsi–German) restricts the generalizability of the findings to other language pairings with different typological properties.

These limitations do not undermine the central conclusions of the dissertation, but rather delineate the scope of interpretation and point to important directions for future research, including broader typological sampling, more balanced cross-linguistic experimental designs, and the extension of online methodologies to additional language contexts.

10.3 Final Conclusion and Outlook

The present dissertation has shown that grammatical gender can influence cognition in systematic ways, particularly during real-time language processing. By integrating typological, theoretical, behavioral, and eye-tracking approaches, it provides a comprehensive account of how linguistic structure interacts with conceptualization and attention across different cognitive domains and speaker populations.

Beyond its empirical findings, the dissertation contributes to a more refined understanding of linguistic relativity. The results demonstrate that language does not impose fixed constraints on thought, but rather shapes how information is selected, structured, and attended to during language use. In this sense, grammatical gender operates as a probabilistic and context-sensitive cue that guides cognitive processing, rather than determining conceptual representations.

The findings further highlight the importance of methodological precision and cross-linguistic diversity for future research. Expanding the range of languages, refining experimental paradigms, and integrating online and offline measures will be essential for advancing the field. In particular, the present results open new avenues for investigating how grammatical systems interact with attention and conceptualization in multilingual contexts.

The findings also have important implications for applied linguistics, particularly in the domains of bilingualism and second language acquisition. They suggest that grammatical gender is not merely a formal property of language, but actively shapes attentional and

conceptual processing during language use. This has direct relevance for understanding why grammatical gender systems are persistently difficult to acquire, especially in contexts of limited exposure. The results indicate that successful acquisition may depend not only on explicit rule learning, but also on the development of implicit, automatized processing routines that integrate grammatical gender into real-time comprehension and production (Sabourin et al., 2006; Shimanskaya & Slabakova, 2019).

Beyond academic contexts, these findings have broader implications for language teaching and learning. They suggest that instructional approaches to grammatical gender should move beyond purely rule-based explanations and instead incorporate exposure-based and usage-oriented methods that promote implicit learning and real-time processing. Teaching practices that emphasize repeated exposure, contextualized input, and attention to form during comprehension may facilitate the integration of grammatical gender into online processing. Furthermore, the results highlight the importance of considering learners' linguistic backgrounds, as the absence of grammatical gender in a first language may influence how learners attend to and process gender distinctions in a second language. In this way, the present findings contribute to a more nuanced understanding of how grammatical structures shape language learning in multilingual settings.

More broadly, these findings underscore that grammatical structures are not merely formal features of language, but play an active role in shaping how speakers attend to and interpret the world during communication.

Future research can build on these findings in several directions. First, expanding the range of language pairings will be essential to determine the extent to which the observed patterns generalize across different typological configurations. In particular, comparisons involving two gendered languages, or languages with different types of gender systems, may further clarify the role of grammatical structure in shaping cognition.

Second, future studies should investigate the role of proficiency, exposure, and language use in modulating grammatical gender effects, ideally through longitudinal designs that track changes over time. Such approaches would provide a more detailed understanding of how grammatical gender becomes integrated into the bilingual processing system.

Third, methodological extensions—such as combining eye-tracking with neurocognitive measures (e.g., EEG)—offer promising avenues for gaining deeper insight into the temporal dynamics and neural correlates of grammatical gender processing.

Finally, future research may examine how interactions between linguistic relativity and cross-linguistic influence shape learners' sensitivity to grammatical gender, and whether such effects can be systematically leveraged in instructional contexts.

Taken together, this dissertation contributes to a growing body of research that moves beyond the question of whether language influences cognition. Instead, it clarifies how, when, and under which conditions such influences emerge. In doing so, it supports a broader shift in the field from static, representation-based accounts toward dynamic, process-oriented models of language–cognition interaction.

11 References

Albert, E. M., & Hoijer, H. (1956). Language in Culture: Proceedings of a Conference on the Interrelations of Language and other Aspects of Culture. *Philosophy and Phenomenological Research*, *16*(3), 426. https://doi.org/10.2307/2104733

Alekseeva, M., Myachykov, A., Bermudez-Margaretto, B., & Shtyrov, Y. (2022). Neurophysiological correlates of automatic integration of voice and gender information during grammatical processing. *Scientific Reports*, *12*(1), 13114. https://doi.org/10.1038/s41598-022-14478-2

Anastasi, A., & Cordova, F. A. (1953). Some effects of bilingualism upon the intelligence test performance of Puerto Rican children in New York City. *Journal of Educational Psychology*, *44*(1), 1–19. https://doi.org/10.1037/h0061639

Anderson, B. (1991). *Imagined communities: Reflections on the origin and spread of nationalism (Rev. Ed.)*.

Anderson, C. (2010). Presenting and Evaluating Qualitative Research. *American Journal of Pharmaceutical Education*, *74*(8), 141. https://doi.org/10.5688/aj7408141

Anderson, J. A. E., Mak, L., Keyvani Chahi, A., & Bialystok, E. (2018). The language and social background questionnaire: Assessing degree of bilingualism in a diverse population. *Behavior Research Methods*, *50*(1), 250–263. https://doi.org/10.3758/s13428-017-0867-9

Athanasopoulos, P. (2006). Effects of the grammatical representation of number on cognition in bilinguals. *Bilingualism: Language and Cognition*, *9*(1), 89–96. https://doi.org/10.1017/S1366728905002397

Athanasopoulos, P. (2007). Interaction between grammatical categories and cognition in bilinguals: The role of proficiency, cultural immersion, and language of instruction. *Language and Cognitive Processes*, *22*(5), 689–699. https://doi.org/10.1080/01690960601049347

Athanasopoulos, P. (2009). Cognitive representation of colour in bilinguals: The case of Greek blues. *Bilingualism: Language and Cognition*, *12*(1), 83–95. https://doi.org/10.1017/S136672890800388X

Athanasopoulos, P., & Bylund, E. (2023). Cognitive restructuring: Psychophysical measurement of time perception in bilinguals. *Bilingualism: Language and Cognition*, *26*(4), 809–818. https://doi.org/10.1017/S1366728922000876

Athanasopoulos, P., & Casaponsa, A. (2020). The Whorfian brain: Neuroscientific approaches to linguistic relativity. *Cognitive Neuropsychology*, *37*(5–6), 393–412. https://doi.org/10.1080/02643294.2020.1769050

Baars, B. J., & Gage, N. M. (2010). *Cognition, Brain, and Consciousness*. Elsevier. https://doi.org/10.1016/C2009-0-01556-6

Barber, H., & Carreiras, M. (2005). Grammatical Gender and Number Agreement in Spanish: An ERP Comparison. *Journal of Cognitive Neuroscience*, *17*(1), 137–153. https://doi.org/10.1162/0898929052880101

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

Bassetti, B., & Nicoladis, E. (2016). Research on grammatical gender and thought in early and emergent bilinguals. *International Journal of Bilingualism*, *20*(1), 3–16. https://doi.org/10.1177/1367006915576824

Bender, A., Beller, S., & Klauer, K. C. (2011). Grammatical Gender in German: A Case for Linguistic Relativity? *Quarterly Journal of Experimental Psychology*, *64*(9), 1821–1835. https://doi.org/10.1080/17470218.2011.582128

Bialystok, E. (2001). *Bilingualism in development: Language, literacy and cognition*. Cambridge University Press.

Bialystok, E., Craik, F. I. M., & Luk, G. (2012). Bilingualism: Consequences for mind and brain. *Trends in Cognitive Sciences*, *16*(4), 240–250. https://doi.org/10.1016/j.tics.2012.03.001

Borenstein, M., Hedges, L. V., Higgins, J. P. T., & Rothstein, H. R. (2009). *Introduction to Meta-Analysis* (1st ed.). Wiley. https://doi.org/10.1002/9780470743386

Boroditsky, L. (2001). Does Language Shape Thought?: Mandarin and English Speakers' Conceptions of Time. *Cognitive Psychology*, *43*(1), 1–22. https://doi.org/10.1006/cogp.2001.0748

Boroditsky, L., Fuhrman, O., & McCormick, K. (2011). Do English and Mandarin speakers think about time differently? *Cognition*, *118*(1), 123–129. https://doi.org/10.1016/j.cognition.2010.09.010

Boroditsky, L., Schmidt, L. A., & Phillips, W. (2003). Sex, syntax, and semantics. *Language in Mind: Advances in the Study of Language and Thought*, *22*(61–79), 1.

Bossard, J. H. S. (1945). The bilingual individual as a person—Linguistic identification with status. In *American Sociological Review* (Vol. 6, pp. 699–709).

Brouwer, H., Fitz, H., & Hoeks, J. (2012). Getting real about Semantic Illusions: Rethinking the functional role of the P600 in language comprehension. *Brain Research*, *1446*, 127–143. https://doi.org/10.1016/j.brainres.2012.01.055

Brown, A. M., Lindsey, D. T., & Guckes, K. M. (2011). Color names, color categories, and color-cued visual search: Sometimes, color perception is not categorical. *Journal of Vision*, *11*(12), 2–2. https://doi.org/10.1167/11.12.2

Brown, R. (1958). How shall a thing be called? *Psychological Review*, *65*(1), 14–21. https://doi.org/10.1037/h0041727

Brown, R. W., & Lenneberg, E. H. (1954). A study in language and cognition. *The Journal of Abnormal and Social Psychology*, *49*(3), 454–462. https://doi.org/10.1037/h0057814

Cameron, D. (1995). *Verbal Hygiene* (0 ed.). Routledge. https://doi.org/10.4324/9780203123898

Casasanto, D., Boroditsky, L., Phillips, W., Greene, J., Goswami, S., Bocanegra-Thiel, S., Santiago-Diaz, I., Fotakopoulou, O., Pita, R., & Gil, D. (2004). How Deep Are Effects of language on Thought? Time Estimation in Speakers of English, Indonesian, Greek and Spanish. *Kenneth Forbus, Dedre Gentner, Terry Reiger: Proceedings of the 26th Annual Cognitive Science Society, Lawrence Erlbaum (2004)*.

Chen, Z., & Faitaki, F. (2025). Effects of the French grammatical gender system on bilingual adults' perception of objects. *Bilingualism: Language and Cognition*, *28*(2), 401–414. https://doi.org/10.1017/S1366728924000464

Chomsky, N. (1957). *Syntactic structures (The Hague, Netherlands: Mouton)*.

Chomsky, N. (1965). *Aspects of the theory of syntax*. The MIT Press.

Chomsky, N. (1993). *Lectures on Government and Binding: The Pisa Lectures* (Vol. 9). De Gruyter Mouton. https://doi.org/doi:10.1515/9783110884166

Chomsky, N. (2000). Minimalist inquiries: The framework (MITOPL 15). *Step by Step: Essays on Minimalist Syntax in Honor of Howard Lasnik*, 89–155.

Comrie, B. (1989). *Language universals and linguistic typology: Syntax and morphology* (2nd ed). University of Chicago Press.

Cook, V., Bassetti, B., Kasai, C., Sasaki, M., & Takahashi, J. A. (2006). Do bilinguals have different concepts? The case of shape and material in Japanese L2 users of English. *International Journal of Bilingualism*, *10*(2), 137–152. https://doi.org/10.1177/13670069060100020201

Corbett, G. G. (1991). *Gender* (1st ed.). Cambridge University Press. https://doi.org/10.1017/CBO9781139166119

Corbett, G. G. (Ed.). (2013). *The Expression of Gender:* DE GRUYTER. https://doi.org/10.1515/9783110307337

Croft, W. (2003). *Typology and universals* (2nd ed). Cambridge University Press.

Dahl, Ö. (2000). Animacy and the notion of semantic gender. In B. Unterbeck, M. Rissanen, T. Nevalainen, & M. Saari (Eds.), *Gender in Grammar and Cognition* (pp. 99–116). DE GRUYTER MOUTON. https://doi.org/10.1515/9783110802603.99

De Bot, K., Lowie, W., & Verspoor, M. (2007). A Dynamic Systems Theory approach to second language acquisition. *Bilingualism: Language and Cognition*, *10*(1), 7–21. https://doi.org/10.1017/S1366728906002732

Dijkstra, T., & Van Heuven, W. J. B. (2002). The architecture of the bilingual word recognition system: From identification to decision. *Bilingualism: Language and Cognition*, *5*(3), 175–197. https://doi.org/10.1017/S1366728902003012

Dixon, R. M. (1972). *The Dyirbal language of north Queensland*. CUP Archive.

Duranti, A. (1997). *Linguistic Anthropology* (1st ed.). Cambridge University Press. https://doi.org/10.1017/CBO9780511810190

Eckert, P., & McConnell-Ginet, S. (2013). *Language and Gender* (2nd ed.). Cambridge University Press.

Evans, V., & Green, M. (2018). *Congnitive Linguistics an Introduction* (1st ed.). Routledge. https://doi.org/10.4324/9781315864327

Eysenck, M. W., & Keane, M. T. (2020). *Cognitive Psychology: A Student's Handbook* (8th ed.). Psychology Press. https://doi.org/10.4324/9781351058513

Feinmann, D. (2020). Language and Thought in the Motion Domain: Methodological Considerations and New Empirical Evidence. *Journal of Psycholinguistic Research*, *49*(1), 1–29. https://doi.org/10.1007/s10936-019-09668-5

Fitz-Gerald, J. (1918). *Modern Language Journal*, *3*(2), 49–62.

Flaherty, M. (2001). How a Language Gender System Creeps into Perception. *Journal of Cross-Cultural Psychology*, *32*(1), 18–31. https://doi.org/10.1177/0022022101032001005

Flecken, M., Athanasopoulos, P., Kuipers, J. R., & Thierry, G. (2015). On the road to somewhere: Brain potentials reflect language effects on motion event perception. *Cognition*, *141*, 41–51. https://doi.org/10.1016/j.cognition.2015.04.006

Fodor, J. A. (1975). *The language of thought*. Harvard Univ. Pr.

Francis, N. (2007). MICHEL PARADIS , *A neurolinguistic theory of bilingualism*. *Language in Society*, *36*(1), 105–108. https://doi.org/10.1017/S004740450721005X

Franklin, A., Clifford, A., Williamson, E., & Davies, I. (2005). Color term knowledge does not affect categorical perception of color in toddlers. *Journal of Experimental Child Psychology*, *90*(2), 114–141. https://doi.org/10.1016/j.jecp.2004.10.001

Gabriel, U., Gygax, P., Sarrasin, O., Garnham, A., & Oakhill, J. (2008). Au pairs are rarely male: Norms on the gender perception of role names across English, French, and German. *Behavior Research Methods*, *40*(1), 206–212. https://doi.org/10.3758/BRM.40.1.206

Grant, M. J., & Booth, A. (2009). A typology of reviews: An analysis of 14 review types and associated methodologies. *Health Information & Libraries Journal*, *26*(2), 91–108. https://doi.org/10.1111/j.1471-1842.2009.00848.x

Greenberg, J. H. (1978). How does a language acquire gender markers. *Universals of Human Language*, *3*, 47–82.

Grosjean, F. (1998). Studying bilinguals: Methodological and conceptual issues. *Bilingualism: Language and Cognition*, *1*(2), 131–149. https://doi.org/10.1017/S136672899800025X

Grosjean, F. (2001). The Bilingual's Language Modes. In J. Nicol (Ed.), *One Mind, Two Languages: Bilingual Language Processing* (pp. 1–25). Blackwell.

Grosjean, F. (2008). *Studying Bilinguals*. Oxford University Press.

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

Gygax, P., Gabriel, U., Sarrasin, O., Oakhill, J., & Garnham, A. (2008). Generically intended, but specifically interpreted: When beauticians, musicians, and mechanics are all men. *Language and Cognitive Processes*, *23*(3), 464–485. https://doi.org/10.1080/01690960701702035

Gygax, P., Gabriel, U., Sarrasin, O., Oakhill, J., & Garnham, A. (2009). Some grammatical rules are more difficult than others: The case of the generic interpretation of the masculine. *European Journal of Psychology of Education*, *24*(2), 235–246. https://doi.org/10.1007/BF03173014

Hammarström, H., Forkel, R., Haspelmath, M., & Bank, S. (2021). glottolog/glottolog: Glottolog database 4.4. *(No Title)*.

Harley, H., & Ritter, E. (2002). Person and Number in Pronouns: A Feature-Geometric Analysis. *Language*, *78*(3), 482–526. https://doi.org/10.1353/lan.2002.0158

Hartsuiker, R. J., Pickering, M. J., & Veltkamp, E. (2004). Is Syntax Separate or Shared Between Languages?: Cross-Linguistic Syntactic Priming in Spanish-English Bilinguals. *Psychological Science*, *15*(6), 409–414. https://doi.org/10.1111/j.0956-7976.2004.00693.x

Hellinger, M., & Bußmann, H. (2015). Gender across languages: The linguistic representation of women and men. In M. Hellinger & H. Motschenbacher (Eds.), *IMPACT: Studies in Language, Culture and Society* (Vol. 36, pp. 1–25). John Benjamins Publishing Company. https://doi.org/10.1075/impact.36.01hel

Herder, J. G. von. (2002). Treatise on the Origin of Language (1772). In M. N. E. Forster (Ed.), *Herder: Philosophical Writings* (pp. 65–164). Cambridge University Press.

Hlawatsch, A., & Krickl, T. (2019). Einstellungen zu Befragungen. In N. Baur & J. Blasius (Eds.), *Handbuch Methoden der empirischen Sozialforschung* (pp. 357–364). Springer Fachmedien Wiesbaden. https://doi.org/10.1007/978-3-658-21308-4_24

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

Holmqvist, K., & Andersson, R. (2017). *Eye tracking: A comprehensive guide to methods, paradigms, and measures* (2nd edition). Lund Eye-Tracking Research Institute.

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

Huettig, F., Rommers, J., & Meyer, A. S. (2011). Using the visual world paradigm to study language processing: A review and critical evaluation. *Acta Psychologica*, *137*(2), 151–171. https://doi.org/10.1016/j.actpsy.2010.11.003

Humboldt, W. (1963). *Humanist without portfolio*.

Imai, M., & Gentner, D. (1997). A cross-linguistic study of early word meaning: Universal ontology and linguistic influence. *Cognition*, *62*(2), 169–200. https://doi.org/10.1016/S0010-0277(96)00784-6

Imai, M., Schalk, L., Saalbach, H., & Okada, H. (2014). 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

Ito, A., & Knoeferle, P. (2022). Analysing data from the psycholinguistic visual-world paradigm: Comparison of different analysis methods. *Behavior Research Methods*, *55*(7), 3461–3493. https://doi.org/10.3758/s13428-022-01969-3

Jespersen, O. (1922). *Language: Its nature, development and origin*.

Johnson, M., & Lakoff, G. (1980). *Metaphors we live by* (Vol. 1). University of Chicago press Chicago.

Katamba, F. (2003). Bantu nominal morphology. *The Bantu Languages*, *103*, 120.

Knoeferle, P., & Guerra, E. (2016). Visually Situated Language Comprehension. *Language and Linguistics Compass*, *10*(2), 66–82. https://doi.org/10.1111/lnc3.12177

Konishi, T. (1993). The semantics of grammatical gender: A cross-cultural study. *Journal of Psycholinguistic Research*, *22*(5), 519–534. https://doi.org/10.1007/BF01068252

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

Kramer, R. (2020). Grammatical Gender: A Close Look at Gender Assignment Across Languages. *Annual Review of Linguistics*, *6*(1), 45–66. https://doi.org/10.1146/annurev-linguistics-011718-012450

Kroll, J. F., Fricke, M., & Dussias, P. E. (2016). Phonetic variation in bilingual speech: A lens for studying the production–comprehension link. *Journal of Memory and Language*, *89*, 110–137. https://doi.org/10.1016/j.jml.2015.10.001

Langacker, R. (2008). *Cognitive Grammar: A Basic Introduction* (1st ed.). Oxford University PressNew York. https://doi.org/10.1093/acprof:oso/9780195331967.001.0001

Langacker, R. W. (1987). *Foundations of cognitive grammar: Volume I: Theoretical prerequisites* (Vol. 1). Stanford university press.

Langacker, R. W. (1991). *Foundations of Cognitive Grammar: Volume II: Descriptive Application*. Stanford University Press. https://doi.org/10.1515/9780804764469

Lehmann, T. (1993). *A grammar of modern Tamil* (2nd ed.). Pondicherry Institute of Linguistics and Culture.

Levelt, W. J. M. (1989). *Speaking: From Intention to Articulation*. The MIT Press. https://doi.org/10.7551/mitpress/6393.001.0001

Levelt, W. J. M., Roelofs, A., & Meyer, A. S. (1999). A theory of lexical access in speech production. *Behavioral and Brain Sciences*, *22*(1), 1–38. https://doi.org/10.1017/S0140525X99001776

Levinson, S. C. & others. (2003). *Language and mind: Let's get the issues straight!.* MIT Press.

Levy, R. (2008). Expectation-based syntactic comprehension. *Cognition*, *106*(3), 1126–1177. https://doi.org/10.1016/j.cognition.2007.05.006

Lew-Williams, C., & Fernald, A. (2007). Young Children Learning Spanish Make Rapid Use of Grammatical Gender in Spoken Word Recognition. *Psychological Science*, *18*(3), 193–198. https://doi.org/10.1111/j.1467-9280.2007.01871.x

Li, P., Zhang, F., Tsai, E., & Puls, B. (2014). Language history questionnaire (LHQ 2.0): A new dynamic web-based research tool. *Bilingualism: Language and Cognition*, *17*(3), 673–680. https://doi.org/10.1017/S1366728913000606

Lucy, J. A. (1992). *Grammatical Categories and Cognition: A Case Study of the Linguistic Relativity Hypothesis* (1st ed.). Cambridge University Press. https://doi.org/10.1017/CBO9780511620713

Lucy, J. A. (1997). Linguistic Relativity. *Annual Review of Anthropology*, *26*(1), 291–312. https://doi.org/10.1146/annurev.anthro.26.1.291

Lucy, J. A. (2014). Methodological approaches in the study of linguistic relativity: Corpus method and cognitive theory. In L. Filipović & M. Pütz (Eds.), *Human Cognitive Processing* (Vol. 44, pp. 17–44). John Benjamins Publishing Company. https://doi.org/10.1075/hcp.44.01luc

Lucy, J. A. (2016). Recent Advances in the Study of Linguistic Relativity in Historical Context: A Critical Assessment. *Language Learning*, *66*(3), 487–515. https://doi.org/10.1111/lang.12195

Luebke, F. C. (1980). *Bonds of Loyalty: German-Americans and World War I*. Northern Illinois University Press.

Maciuszek, J., Polak, M., & Świątkowska, N. (2019). Grammatical Gender Influences Semantic Categorization and Implicit Cognition in Polish. *Frontiers in Psychology*, *10*, 2208. https://doi.org/10.3389/fpsyg.2019.02208

Mahon, B. Z., Costa, A., Peterson, R., Vargas, K. A., & Caramazza, A. (2007). Lexical selection is not by competition: A reinterpretation of semantic interference and facilitation effects in the picture-word interference paradigm. *Journal of Experimental Psychology: Learning, Memory, and Cognition*, *33*(3), 503–535. https://doi.org/10.1037/0278-7393.33.3.503

Mandler, J. M. (2007). *Foundations of Mind*. Oxford University Press. https://doi.org/10.1093/acprof:oso/9780195311839.001.0001

Marian, V., Blumenfeld, H. K., & Kaushanskaya, M. (2007). The Language Experience and Proficiency Questionnaire (LEAP-Q): Assessing Language Profiles in Bilinguals and

Multilinguals. *Journal of Speech, Language, and Hearing Research*, *50*(4), 940–967. https://doi.org/10.1044/1092-4388(2007/067)

Mertins, B. (2016). The use of experimental methods in linguistic research: Advantages, problems and possible pitfalls. *Slavic Languages in Psycholinguistics. Chances and Challenges for Experimental Research*, 15–33.

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

Mertins, B., & Delucchi Danhier, R. (Eds.). (2023). *Diversity in Cognition*. Peter Lang Verlag. https://doi.org/10.3726/b21118

Molinaro, N., Barber, H. A., & Carreiras, M. (2011). Grammatical agreement processing in reading: ERP findings and future directions. *Cortex*, *47*(8), 908–930. https://doi.org/10.1016/j.cortex.2011.02.019

Montrul, S. (2015). *The Acquisition of Heritage Languages* (1st ed.). Cambridge University Press. https://doi.org/10.1017/CBO9781139030502

Nicholas, J. (1992). Linguistic diversity in space and time. Chicago and London: The University of Chicago Press. *Journal of Linguistics*, *29*(2), 494–500. https://doi.org/10.1017/S0022226700000438

Norris, J. M., & Ortega, L. (Eds.). (2006). *Synthesizing Research on Language Learning and Teaching* (Vol. 13). John Benjamins Publishing Company. https://doi.org/10.1075/lllt.13

Odermann, K., & Mertins, B. (2026). Semantisches Bewusstsein bei Vorschulkindern: Wie Bilingualismus das metalinguistische Wissen stärkt. *Zeitschrift für Erziehungswissenschaft*. https://doi.org/10.1007/s11618-026-01379-7

Olsson, B. (2021). *A Grammar of Coastal Marind*. De Gruyter. https://doi.org/10.1515/9783110747065

Park, H. I., & Ziegler, N. (2014). Cognitive shift in the bilingual mind: Spatial concepts in Korean–English bilinguals. *Bilingualism: Language and Cognition*, *17*(2), 410–430. https://doi.org/10.1017/S1366728913000400

Pavlenko, A. (2002). We have room for but one language here: Language and national identity in the US at the turn of the 20th century. In *Multilingua* (Vol. 21, pp. 163–196).

Pavlenko, A. (2014). *The Bilingual Mind: And What it Tells Us about Language and Thought* (1st ed.). Cambridge University Press. https://doi.org/10.1017/CBO9781139021456

Piaget, J. (1952). *The origins of intelligence in children.* (M. Cook, Trans.). W W Norton & Co. https://doi.org/10.1037/11494-000

Pickering, M. J., & Garrod, S. (2013). An integrated theory of language production and comprehension. *Behavioral and Brain Sciences*, *36*(4), 329–347. https://doi.org/10.1017/S0140525X12001495

Pinker, S. (1995). *The language instinct* (1st HarperPerennial ed). HarperPerennial.

PRISMA-P Group, Moher, D., Shamseer, L., Clarke, M., Ghersi, D., Liberati, A., Petticrew, M., Shekelle, P., & Stewart, L. A. (2015). Preferred reporting items for systematic review and meta-analysis protocols (PRISMA-P) 2015 statement. *Systematic Reviews*, *4*(1), 1. https://doi.org/10.1186/2046-4053-4-1

R Core Team. & others. (2023). *R: A language and environment for statistical computing (version 4.3. 0)[Computer software]. Last visited 5 December 2023*.

Rapp, B., & Goldrick, M. (2000). Discreteness and interactivity in spoken word production. *Psychological Review*, *107*(3), 460–499. https://doi.org/10.1037/0033-295X.107.3.460

Rayner, K. (1998). Eye movements in reading and information processing: 20 years of research. *Psychological Bulletin*, *124*(3), 372–422. https://doi.org/10.1037/0033-2909.124.3.372

Saalbach, H., Imai, M., & Schalk, L. (2012). Grammatical Gender and Inferences About Biological Properties in German-Speaking Children. *Cognitive Science*, *36*(7), 1251–1267. https://doi.org/10.1111/j.1551-6709.2012.01251.x

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

Sá-Leite, A. R., Luna, K., Tomaz, Â., Fraga, I., & Comesaña, M. (2022). The mechanisms underlying grammatical gender selection in language production: A meta-analysis of the gender congruency effect. *Cognition*, *224*, 105060. https://doi.org/10.1016/j.cognition.2022.105060

Samuel, S., Cole, G., & Eacott, M. J. (2019). Grammatical gender and linguistic relativity: A systematic review. *Psychonomic Bulletin & Review*, *26*(6), 1767–1786. https://doi.org/10.3758/s13423-019-01652-3

Sapir, E. (1921). An introduction to the study of speech. *Language*, *1*(1), 15.

Sapir, E. (1929). The Status of Linguistics as a Science. *Language*, *5*(4), 207. https://doi.org/10.2307/409588

Schriefers, H., Jescheniak, J. D., & Hantsch, A. (2002). Determiner selection in noun phrase production. *Journal of Experimental Psychology: Learning, Memory, and Cognition*, *28*(5), 941–950. https://doi.org/10.1037/0278-7393.28.5.941

Schuh, R. G. (1989). Gender and number in Miya. In Z. Frajzyngier (Ed.), *Current Issues in Linguistic Theory* (Vol. 62, pp. 171–181). John Benjamins Publishing Company. https://doi.org/10.1075/cilt.62.09sch

Seifart, F. (2010). Nominal Classification. *Language and Linguistics Compass*, *4*(8), 719–736. https://doi.org/10.1111/j.1749-818X.2010.00194.x

Sera, M. D., Elieff, C., Forbes, J., Burch, M. C., Rodríguez, W., & Dubois, D. P. (2002). When language affects cognition and when it does not: An analysis of grammatical

gender and classification. *Journal of Experimental Psychology: General*, *131*(3), 377–397. https://doi.org/10.1037/0096-3445.131.3.377

Shimanskaya, E., & Slabakova, R. (2019). L1–L2 differences in the L2 classroom: Anticipating Anglophone learners' difficulties with French pronoun interpretation. *Language Teaching Research*, *23*(2), 259–277. https://doi.org/10.1177/1362168817739650

Slobin, D. I. (1996). From "Thought and Language" to "Thinking for Speaking." In J. J. Gumperz & S. C. Levinson (Eds.), *Rethinking Linguistic Relativity* (pp. 70–96). Cambridge University Press.

Slobin, D. I. (2004). The Many Ways to Search for a Frog: Linguistic Typology and the Expression of Motion Events. *Relating Events in Narrative: Vol. 2. Typological and Contextual Perspectives*, *2*, 219–275.

Sloutsky, V. M. (2010). From Perceptual Categories to Concepts: What Develops? *Cognitive Science*, *34*(7), 1244–1286. https://doi.org/10.1111/j.1551-6709.2010.01129.x

Soja, N. N., Carey, S., & Spelke, E. S. (1991). Ontological categories guide young children's inductions of word meaning: Object terms and substance terms. *Cognition*, *38*(2), 179–211. https://doi.org/10.1016/0010-0277(91)90051-5

Stahlberg, D., Braun, F., Irmen, L., & Sczesny, S. (2007). Representation of the sexes in language. *Social Communication*, 163–187.

Steever, S. B. (1998). *The Tamil auxiliary verb system: A layered-SCI approach*. Springer.

Talmy, L. (1983). How language structures space. In *Spatial orientation: Theory, research, and application* (pp. 225–282). Springer.

Tanenhaus, M. K., Spivey-Knowlton, M. J., Eberhard, K. M., & Sedivy, J. C. (1995). Integration of Visual and Linguistic Information in Spoken Language Comprehension. *Science*, *268*(5217), 1632–1634. https://doi.org/10.1126/science.7777863

Torraco, R. J. (2016). Writing Integrative Literature Reviews: Using the Past and Present to Explore the Future. *Human Resource Development Review*, *15*(4), 404–428. https://doi.org/10.1177/1534484316671606

Traxler, M. J. (2012). *Introduction to psycholinguistics: Understanding language science*.

Ullman, M. T. (2001). The Declarative/Procedural Model of Lexicon and Grammar. *Journal of Psycholinguistic Research*, *30*(1), 37–69. https://doi.org/10.1023/A:1005204207369

Underhill, J. (2009). *Humboldt, Worldview and Language*. Edinburgh University Press. https://doi.org/10.1515/9780748640225

Vigliocco, G., Vinson, D. P., Paganelli, F., & Dworzynski, K. (2005). Grammatical Gender Effects on Cognition: Implications for Language Learning and Language Use. *Journal of Experimental Psychology: General*, *134*(4), 501–520. https://doi.org/10.1037/0096-3445.134.4.501

Walter, A., Fritzsche, T., & Hoehle, B. (2021). Grammatical Gender Acquisition in German: Three-Year-Old Children Use Phonological Cues to Learn the Gender of Novel Nouns. In D. Dionne & L.-A. Vidal Covas (Eds.), *Proceedings of the 45th Annual Boston University Conference on Language Development* (Vol. 2, pp. 746–760). Cascadilla Press.

Weinreich, U. (1953). *Languages in Contact: Findings and Problems*. DE GRUYTER MOUTON. https://doi.org/10.1515/9783110802177

Whorf, B. L. (1956). *Language, thought, and reality: Selected writings of Benjamin Lee Whorf*. MIT press.

Wicha, N. Y. Y., Moreno, E. M., & Kutas, M. (2004). Anticipating Words and Their Gender: An Event-related Brain Potential Study of Semantic Integration, Gender Expectancy, and Gender Agreement in Spanish Sentence Reading. *Journal of Cognitive Neuroscience*, *16*(7), 1272–1288. https://doi.org/10.1162/0898929041920487

Witzel, C., & Gegenfurtner, K. R. (2013). Categorical sensitivity to color differences. *Journal of Vision*, *13*(7), 1–1. https://doi.org/10.1167/13.7.1

Wolff, P., & Holmes, K. J. (2011). Linguistic relativity. *WIREs Cognitive Science*, *2*(3), 253–265. https://doi.org/10.1002/wcs.104