

Boundary Crossing from the Start: 55 Years of Second Language Grammatical Gender Research in Review



Daniel R. Walter

Abstract In this chapter, I explore how the study of grammatical gender has been a prime example of the impact that boundary crossing can have on the understanding of a phenomenon by creating nuance through the incorporation and juxtaposition of multiple perspectives. In the first section, I start with a description of grammatical gender and its different instantiations (or absence) across languages, including cue patterns, agreement structures, and noun-class parallels. I then move on to the ways in which grammatical gender has been investigated and how this research has crossed theoretical and methodological boundaries. The third section focuses on grammatical gender in SLA and aims to deconstruct the boundary between theory and practice, with a focus on what functional approaches have taught us about instruction of grammatical gender. The fourth section focuses on the history of grammatical gender instruction and the various pedagogies that have been tested to support the teaching of grammatical gender, including innovative approaches based on functional and sociocultural approaches. In the final section, I reflect on the many boundaries crossed: theoretical, methodological, disciplinary, linguistic, developmental, typological and, maybe most importantly, temporal. I end with a discussion of the future boundaries to be crossed in this area of research.

Keywords SLA · Grammatical gender · History · Boundary crossing · Teaching

1 Introduction

In 1967, G. Richard Tucker completed his thesis for his PhD from McGill University under the title *The French speaker's skill with grammatical gender: An example of rule-governed behavior*. This thesis focused on first language (L1) French speakers

D. R. Walter (✉)

Humanities Division, Emory University, Oxford College, Oxford, GA, USA

e-mail: daniel.walter@emory.edu

© The Author(s), under exclusive license to Springer Nature
Switzerland AG 2023

D. Zhang, R. T. Miller (eds.), *Crossing Boundaries in Researching,
Understanding, and Improving Language Education*, Educational Linguistics
58, https://doi.org/10.1007/978-3-031-24078-2_3

and the rules that govern French grammatical gender marking. In his own words, Tucker (1967, p. 38) explains that “the study of gender may prove to be a relatively easily explicable example of rule acquisition in language, and the process of attempting to explain this acquisition may provide a useful model for more complex linguistic features (...)” (p. 1). One year later, Tucker published an article with colleagues Lambert, Rigault, and Segalowitz titled “A psychological investigation of French speakers’ skill with grammatical gender.” Here, Tucker uses a different vocabulary to talk about French grammatical gender:

In the case of first-language learning, it seems apparent that the French child is able to distinguish and utilize, by induction, from the recurring regularities in the language, those patterns of cues that mark gender. He appears to be very skilled in generalizing from these patterns to novel occurrences. (Tucker et al., 1968, p. 315).

Right from the start, we can see one of the first of many Tuckerian boundary crossings through the vocabulary used to describe a single phenomenon. In the first quote, the words “rule” and “govern” are reminiscent of a Chomskian perspective on language acquisition. In the second, the words “psychological”, “patterns”, “regularities”, and “cues” call to mind a different, still young field of psycholinguistics, which offered alternative ways of describing the nature of language.

This theoretical boundary crossing is not the only one made between these two studies. Other boundaries include the methodologies used to investigate the role of grammatical gender in linguistic systems and their uses, as well as the boundaries between pedagogical approaches. In addition, Tucker’s initial focus on first language acquisition (L1A) of French grammatical gender eventually led him to questions about the role of grammatical gender in second language acquisition (SLA), in addition to studies in languages beyond the borders of French. Over the course of his work, he leaped between typically separate fields to develop a more robust understanding of grammatical gender on the whole, as a phenomenon of study that does not leave itself easily confined within the borders of one language, one discipline, or one theoretical approach.

Over the last 55 years since these two publications, inquiries into grammatical gender in applied linguistics have changed from the Chomskian rule-governed understanding which peaked at the time of the young Dr. Tucker to a data-driven era that has embraced non-rule-based features and incorporated perspectives and methodologies from diverse academic disciplines. Tucker’s early work set the stage for a complex and nuanced understanding of grammatical gender and its acquisition in first and second languages. In this chapter, I explore how the study of grammatical gender has been a prime example of the impact that boundary crossing can have on the understanding of a phenomenon. First, I describe grammatical gender and its different instantiations (or absence) across languages, including cue patterns, agreement structures, and noun-class parallels. I will then move on to the ways in which grammatical gender research has been investigated and how this research has crossed theoretical and methodological boundaries. The third section focuses on grammatical gender in SLA and aims to deconstruct the boundary between theory and practice with a focus on what functional approaches have taught us about instruction of grammatical gender. The fourth section focuses on the history of grammatical gender instruction and the various pedagogies that have been tested to

support the teaching of grammatical gender, including innovative approaches based on functional and sociocultural approaches. In the final section, I reflect on the many boundaries crossed in the research on grammatical gender: theoretical, methodological, disciplinary, linguistic, developmental, typological and, maybe most importantly, temporal. I conclude with thoughts on the future boundaries to be crossed in this area of research.

2 Grammatical Gender across Linguistic Boundaries

The first questions to pose is, what is grammatical gender? It is integral to so many Indo-European languages, but at the same time the term *grammatical gender* fails to properly describe the phenomenon, especially since a similar phenomenon is observed in noun-class systems such as Swahili, which operate in structurally similar ways but differ in semantic overlapping. In fact, in order to really understand the nature and function of grammatical gender, I must cross quite a few linguistic, typological, and semantic boundaries.

2.1 On the Nature of Grammatical Gender

One of the first boundaries we come across when we look at the nature of grammatical gender is the semantic distinction between the terms we use and the objects they refer to. In languages with grammatical gender, all nouns are assigned a grammatical gender on top of and in addition to any naturally occurring semantic gender or sex, in the biological sense (from here on, sex). These grammatical genders and sexes can overlap, but do not have to, and in the grand scope of all nouns in a language, the vast majority of the time the grammatical gender has nothing to do with sex. The gender assigned to a particular noun is mainly a way to capture the similarity among nouns that behave in similar ways. And while the assignment of grammatical genders to nouns is usually aided by phonological, morphological, or semantic similarities among nouns, the designations *masculine* and *feminine* are rather arbitrary.

This previous point about assignment of grammatical gender is particularly important because gender assignment was previously thought to be random in some languages. For example, in German, it wasn't until studies by Zubin and Köpcke (Köpcke & Zubin, 1984, 1996; Zubin & Köpcke, 1981, 1986) that gender assignment was understood to be highly predictable. Rather than a reliance on memorization for all nouns, we now knew that people could rely on cues to assign grammatical gender to the vast majority of nouns. Remaining with German, there are, in fact, phonological, morphological, and semantic cues to grammatical gender. Phonological cues differ from morphological ones in that, for phonological cues, the sound itself is the cue and does not carry meaning. Meanwhile, for morphological cues, the sound is tied to a particular meaning. For example, most words in

German that end in /ə/ like *Katze* [cat] are feminine, but there is no meaning attached to the /ə/ itself; it is simply part of the sound of the word, i.e., the phonological representation of the word, and is, therefore, a phonological cue. On the other hand, the morphological cue /kart/, as in *Glücklichkeit* (happiness) has a morphological function – to turn adjectives into nouns (similar to the function of ‘-ness’ in *happiness*). There are also semantic cues in German. For example, all days of the week are masculine, all alcoholic beverages (except beer) are masculine, and all baby animals are neuter. So within one linguistic system, the cues to gender cut across linguistic categories. These types of grammatical gender cues, as well as their distribution, reliability, and competition with one another cross language boundaries. One language may only have phonological cues to gender, while another only has semantic, and a third has both.

This cross-linguistic difference is also true for the number of grammatical gender categories in a language. Whereas most Romance languages only have two levels, masculine and feminine, that is not true for Romanian, which retained its neuter category from Latin. The dropping of the third gender occurred in all other Romance languages as well as some other Indo-European languages such as the Germanic language Dutch (whose two resulting genders were renamed to *neuter* and *common* in formal linguistic descriptions) and English (which dropped almost all of its grammatical gender over time). This is interesting because closed-class words are less productive and harder to change over time.

2.2 *An Ill-Fitting Name*

Here, I would like to return to the term *gender* itself, because when we discuss typologically different languages, we see similar linguistic features that are not traditionally called *genders* that nevertheless act in very similar ways. So in order to truly understand what grammatical gender is, we again need to cross the boundary between languages or language groups. This effort to cross and examine multiple languages’ and language families’ boundaries leads to a better understanding of the role that semantics and morphology play in languages with grammatical gender.

First, when we compare language with genders to other languages with noun-classes, we can see that much of the confusion arises as a direct result of the ill-fitting term *gender*. This problem in linguistic terminology can be seen in at least four ways:

- the lack of overlap between grammatical gender and sex present in languages,
- the origins of the term *grammatical gender* and an attempt to encase all languages with this typological feature into the same linguistic box,
- the similar function that noun-class systems in non-Indo-European languages play,
- and the possible over-extension of biological sex connotations onto inanimate objects.

First, the overlap of sex with grammatical gender complicates and overemphasizes this relationship. In languages with grammatical gender, all nouns are assigned a gender. In many languages (including German, Spanish, French, Latin, and Greek, among others), sex often overlaps with grammatical gender for people and professions. For example, in German, *die Frau* and *der Mann*, “the woman” and “the man” respectively, agree in both grammatical gender and sex. However, one only needs to look to the word for “girl”, *das Mädchen*, to see the break between grammatical gender and sex. *Das Mädchen* in German (as is evident from the difference in the article-noun agreement pattern visible on the definite article) is not the same gender as *die Frau*. Instead, “girl” in German is neuter. This grammatical gender assignment has nothing to do with the semantic nature of the referent (although, as previously mentioned, other semantic cues for gender assignment do exist in German). Grammatical gender assignment here is related to the morphological marking, specifically the umlaut over the *a* and the suffix *-chen*, which makes nouns diminutive. All diminutive nouns in German are formed through this process,¹ and therefore all diminutive nouns in German are assigned neuter grammatical gender. If not even nouns with semantic cues to sex provide a reliable basis for grammatical gender assignment, then it is not a surprise that in the majority of cases, grammatical gender has nothing to do with sex.

The second problem with the term *gender* is that the origin of this term comes from the study of classical languages. The first known use of the term *gender* to describe noun classes was by the Greek philosopher Protagoras in the fifth century B.C., when he divided Ancient Greek nouns into three classes, *masculine*, *feminine*, and *inanimate* (Aikhenvald, 2004). Greek has a noun-class system typical of Indo-European languages. Over the course of centuries, various philosophers and others interested in the nature of language forced, for better and often for worse, the pre-established categories from Greece onto the grammatical gender systems of other Indo-European languages, seemingly regardless of their fit. Even within Indo-European languages, there can be important distinctions between grammatical gender systems that make this classical three gender system unfit. As previously mentioned, many Romance languages only have two genders.

Finally, the existence of languages with similarly functioning grammatical categories should add another layer of skepticism for using *gender* as the overarching term for this phenomenon. A perfectly reasonable replacement term, *noun class*, exists and can more accurately encompass the grammatical gender systems of all languages as well as other features present in languages like Swahili, which has 16 noun classes and uses categories such as *human*, *animal*, and *plant* as semantic categories to organize nouns into different noun classes.

How much of the linguistic understanding from the study of Greek structures has interfered with our ability to differentiate the functions and forms of grammatical gender systems from the way languages refer to sex? How much has the word *gender* tainted our ability, as researchers and speakers, to understand noun-class

¹Or through the suffix *-lein*, although *-chen* is more productive, especially in modern German.

structures in languages that we observe as having *gender* systems? This question about linguistic relativity has gained increasing scrutiny in recent research as researchers test the boundaries and implications of a weak Sapir-Whorf hypothesis, i.e., the idea that the language(s) we speak do not determine our ways of thinking but do play a role in highlighting certain features of our environment via entrenched grammatical processes (Kay & Kempton, 1984). One of the major claims of a weak Sapir-Whorf hypothesis is that grammar affects what we pay attention to as humans. According to Samuel et al. (2019), “labels or grammatical information hone attention to associated features, which in turn feed back down to lower-level processes in a feedback loop. These effects can be upregulated or downregulated by the salience of the relevant linguistic information in the task,” (p. 1782). Therefore, if a grammar is telling you to pay attention to the way nouns are being categorized by morphologically similar declensional paradigms, then that information is continually being fed back to the processor and assigned a salient role in the input.

When we assign a name like *gender* to these paradigms, which overlap with sex to minor degrees, it is easy to see how a constant influx of gender information about nouns could influence semantic connections. This point is argued by Phillips and Boroditsky (2003), who report that,

A series of studies found effects of grammatical gender on people’s perceptions of similarity between objects and people. This was true even though the tasks were performed in English (a language devoid of grammatical gender), even when the tasks were non-linguistic (e.g., rating similarities between unlabeled pictures), and even while subjects were engaged in a verbal interference task. Finally, results showed that crosslinguistic differences in thought can be produced just by grammatical differences and in the absence of other cultural factors. It is striking that even a fluke of grammar (the nearly arbitrary assignment of a noun to be masculine or feminine) can have an effect on how people think about things in the world. (p. 933)

Studies by Boroditsky (2001) and Boroditsky et al. (2003) seem to provide additional evidence for these effects.

However, these findings are not consistent. In a study by Kousta et al. (2008) that investigated semantic transfer among L1 Italian speakers of L2 English, the authors “found no evidence of transfer from Italian to English of the semantic effects of gender and interpret this lack of transfer as evidence for the constrained role grammatical gender has on bilingual cognition” (p. 854). Similarly, Bender et al. (2011), who conducted a study on L1 German speakers, did not find an effect of language on thought regarding grammatical gender. At this time, it is still unclear how much, if at all, grammatical gender influences a person’s semantic representation of inanimate objects. What is clear, however, is that more research is needed to understand whether grammatical gender crosses the boundary of morphology into semantics.

3 Grammatical Gender Research across Theoretical and Methodological Boundaries

Grammatical gender is a topic of great interest within a number of research perspectives, and although past efforts to describe grammatical gender from these different perspectives have been very fruitful with regard to our understanding of grammatical gender, we only benefit from these multiple perspectives when we allow ourselves to cross the boundaries of various research paradigms and theories.

From formal linguistic perspectives working within the parameters of Universal Grammar (UG), grammatical gender has been a doorway into the study of feature assignment across languages. One of the major questions is where and how grammatical gender information is assigned. In one analysis of grammatical gender in Romance and Bantu languages, Carstens (2010) used grammatical gender to argue that nouns can have “intrinsically valued but uninterpretable” (p. 28) features. On the other hand, Kramer (2014) argued that gender description in Amharic, a Semitic language, relies on both grammatical gender and sex. Rather than separating grammatical gender and sex, the author “developed a gender assignment system that is almost entirely based on sex as an interpretable feature on [nouns], and ‘masculine’ forms as a default for anything that does not have a [+FEM] on [the noun]” (Kramer, 2014, p. 11). Kramer (2014) argued that this reanalysis “is more successful than previous analyses in that all the Amharic facts are accounted for, there is no need for a discourse referent lexicon connection, and there is no more ‘calculation’ of gender from sex” (p. 11). Interestingly, within a UG framework there seems to be an important discussion related to boundary crossing – whether non-syntactic/morphological aspects of language, like semantics, can affect feature assignment.

Beyond formal linguistic approaches, developmental psychology is interested in how children acquire grammatical gender in both L1 and early L2 acquisition. For example, Blom et al. (2008) investigated acquisition of the common and neuter genders on both determiners and adjectives in Dutch among L1 Dutch speaking children, child Moroccan L2 learners of Dutch, and adult Moroccan L2 learners of Dutch. By comparing the three groups, the researchers found that “the vast majority of the children’s errors could be interpreted as use of the common form (i.e., *schwa*-adjective) in neuter contexts and were in this respect consistent with the errors in definite articles,” (Blom et al., 2008, p. 322). From this we learn that child learners’ developmental trajectories may differ from those of adults, and that crossing boundaries between L1A and SLA may lead to further insights about how particular grammatical features develop. In contrast with L1 learners, the adult L2 learners in this study often produced “bare adjectives” (Blom et al., 2008, p. 317); i.e., they did not apply any grammatical gender to attributive adjectives.² This could imply a difference in language acquisition, where children, having fewer defined L1 pathways, automatically process grammatical information in the input, even if the same information is absent in their L1, while adults rely on L1 pathways that would not be

²This result was also found by Walter (2020) for L1 English high school learners of L2 German.

looking for grammatical information in new places (i.e., grammatical gender marking on adjectives), assuming grammatical information is not located in the same linguistic environments in their L1. Unlike L1 learners who are forming an initial pathway to process grammatical gender, L2 learners may be required to either abandon or reform processing pathways to account for gender information. However, the authors point to their results and other studies which indicate that “in bilingual/child L2 development, definite articles do show a (prolonged) development, whereas attributive adjectives seem to fossilize” (Blom et al., 2008, p. 322). Thus, there seems to be a difference in either the relevance or frequency of grammatical gender marking on determiners as opposed to adjectives that is causing a divergence between the acquisitional trajectories of these two parts of speech. The previously described study by Blom et al. (2008) also crossed into the field of corpus linguistics; the source of the data they used to come to their findings was “based on Dutch adult- and child directed speech in three CHILDES corpora” (p. 303).

Researchers within usage-based approaches are interested in similar questions about acquisition as those posed by formal linguists and developmental psychologists, but their methods and theories differ substantially. In a study using an artificial language, Arnon and Ramscar (2012) asked whether the grain size and order of acquisition could affect the acquisition of a novel grammatical gender system. By modifying the grain size (whether there was a large or small boundary between articles and nouns) and the order in which participants saw the relevant grammatical gender structure, the authors showed that participants who saw items with smaller boundaries and/or had the sequence-first condition outperformed other groups on both the forced-choice and production tasks. They argued that these findings “[fit] nicely with usage-based models of language, which posit that grammatical relations emerge from a gradual process of abstraction over stored utterances” (p. 2116). This explanation and these findings contrast strongly with the positions of formal linguists, who have focused on the availability of transfer and access, rather than properties of the input and learning.

Crossing another disciplinary boundary, we find that psycholinguists are also interested in grammatical gender, and in looking at this phenomenon, they collect and interpret knowledge from various mental functions such as acquisition, processing/comprehension, and production. Beginning with acquisition, it is not the case that L1 speakers always pick up grammatical gender so effortlessly, as is so often the sentiment regarding L1 acquisition. For example, cue reliability and frequency have significant effects on the L1 acquisition of grammatical gender (e.g., in French: Matthews, 2010; in German: Mills, 2012). Psycholinguistic researchers are also interested in how the existence of grammatical gender affects processing. One way to investigate this is to measure the time it takes to access gender information using timed lexical decision tasks. Using this method for L1 Spanish number and gender processing, Dominguez et al. (1999) found surface frequency effects for access times by gender, within gender, and between singular- and plural-dominant forms. However, when the researchers compared within-group access times, they found differences between real word items but not between non-word items. This means that semantic information, like connotations between genders and sex, were applied

to real words and slowed down processing but that this additional processing step for semantic information was not carried out for non-words.

So why is there an additional processing step for real words as opposed to non-words? If we cross over into connectionist models, they help shed light on how these structures are processed and the number of routes through which grammatical gender information can be accessed. Studying L1 Hebrew speakers, Gollan and Frost (2001) proposed a connectionist model with multiple routes to gender using a gender decision and grammaticality judgment task. Based on the results, they postulated “a model containing two routes to grammatical gender: one that involves an abstract gender node, and another that is form-based and is assumed to play a greater role in recovery from agreement errors,” (Gollan & Frost, 2001, p. 627). The question of how one “gets” to gender information must then at least raise questions about whether semantic, phonetic, or morphological information is the faster way to access gender and whether these routes to gender are variable based on the way gender is assigned and distributed within a language.

Offline measures can also provide insight into variable routes to gender assignment. Hohlfeld (2006), investigating L1 German speakers’ assignment of gender to non-words, also found evidence for multiple routes to gender within the same language; one they called *lexical* and the other *rule-based*. The author argued that without some semantic referent, as is given with non-words, there must be some other way that German speakers are able to assign gender aside from the learned gender/noun pairing through usage. They argue that “gender assignment might be guided by either gender marking regularities alone or lexical information as well as gender cue information (postlexical checking)” (Hohlfeld, 2006, p. 139). The author points to studies that found similar results (e.g., Bates et al., 1995, 1996; Gollan & Frost, 2001); but they also note that these findings contrast with some on French (e.g., Desrochers & Paivio, 1990; Taft & Meunier, 1998).

The methods that psycholinguists use to investigate grammatical gender go beyond traditional behavioral ones. In recent years, researchers have begun to use functional magnetic resonance imaging (fMRI) and electroencephalography (EEG) to look at the brain during processing. For example, de Resende et al. (2019) used EEG to look at event-related potentials (ERPs) of distinct neurocognitive mechanisms for different grammatical gender conditions. In this study, L1 Brazilian Portuguese speakers “read sentences containing congruent and incongruent grammatical gender agreement between a determiner and a regular or an irregular form (condition 1) and between a regular or an irregular form and an adjective (condition 2),” (de Resende et al., 2019, p. 181). After analysis of the ERP results, the researchers found a “LAN/P600 effect for gender agreement violation involving regular and irregular forms in both conditions,” (de Resende et al., 2019, p. 181). The authors argued that these results “suggest that gender agreement between determiner and nouns recruits the same neurocognitive mechanisms regardless of the nouns’ form and that, depending on the grammatical class of the words involved in gender agreement, differences in ERP signals can emerge,” (de Resende et al., 2019, p. 181). This ERP effect on grammatical gender non-congruence has been fairly well established and can have variable effects based on the way the grammatical gender

system is distributed within a language. In another example, Caffarra et al. (2015) found that among L1 Italian speakers, transparent nouns, i.e., ones that had clear and reliable phonological cues to gender, “elicited an increased frontal negativity and a late posterior positivity compared to irregular nouns (350–950 ms), suggesting that the system is sensitive to gender-to-ending consistency from relatively early stages of processing” (p. 1019). One possible interpretation of this data might be that while one route, either lexical or morpho-phonological, might be faster in terms of access, the strength of the connections might be increased when both routes are available.

Psycholinguists can also cross methodological boundaries within the same study to provide a more detailed, triangulated explanation of grammatical gender. In a study of L1 Spanish speakers, Caffarra et al. (2014) used behavioral (both a gender agreement judgement task and a recognition memory task) and electrophysiological (EEG) methods to test the differences between transparent and opaque nouns. The findings from this study also support the idea of multiple routes to gender, which were found in previously mentioned studies espousing different theoretical approaches and methods. In their explanation, they compare the results of their multiple behavioral tests to their own ERP findings as well as previous behavioral work on multiple routes to gender (e.g., Bates et al., 1995, 1996; Gollan & Frost, 2001; Hernandez et al., 2004; Taft & Meunier, 1998).

Finally, we can also move across boundaries of linguistic modality to provide a more complete picture of grammatical gender processing. In speech, the intricacies of language processing are further complicated by conscious and unconscious mechanisms of cognitive planning as well as coordination with both articulation and paralinguistic behaviors, such as gesturing. This can often lead to differences in grammatical gender performance that contrast with established abilities of assigning and processing grammatical gender. In other words, production might not reflect comprehension, and it is important to understand why. In order to produce correct noun phrases in a language with grammatical gender, that information needs to be accessed in real time while producing the desired utterance (assuming it is not an unanalyzed chunked expression). In a review of the available research on grammatical gender in speech production (which the authors note is significantly smaller than that on grammatical gender comprehension and assignment), Wang and Schiller (2019) cover behavioral, electrophysiological, and fMRI studies that have investigated grammatical gender assignment during speech production. Based on their review, the authors conclude that, “It is generally agreed that grammatical gender is represented as a separate lexico-syntactic feature in the mental lexicon,” (Wang & Schiller, 2019, p. 5). However, despite this agreement, the authors note a number of areas of uncertainty, an important one being whether grammatical gender needs to be, or even is, accessed if grammatical gender information would have no impact on the final articulated form of the noun-phrase. That is, if the selection of one grammatical gender or another would not change the produced determiner, adjective, or noun, then it is not certain that grammatical information is actually accessed at all. This has importance regarding how and where grammatical information is linked to the lexicon. The authors continue: “Nevertheless, emerging evidence has shown

distinctive mechanisms underlying the selection of grammatical gender in Romance languages like Italian and Spanish, and Germanic languages like German and Dutch” (Wang & Schiller, 2019, p. 5). They found that this was confirmed in fMRI studies that

provide evidence for distinctive neural networks for the processing of grammatical gender and suggest that participants tend to adopt a more form-related route to access gender information in Romance languages where the gender-to-ending regularity modulates the gender effect. By contrast, participants tend to adopt a more lexically based route to access grammatical gender in Dutch and German where the noun’s morpho-phonological form is generally not strongly marked by gender. (Wang & Schiller, 2019, p. 5)

From a speech production standpoint, the findings seem to indicate that the strength of morpho/phonological cues can change the way grammatical gender is accessed. In other words, whether the anticipation of a word ending alone or the entire word is required to trigger gender assignment during oral production.

In a curious case on L1 Italian, Vigliocco et al. (1997) compared the effects of grammatical gender to “tip of the tongue (TOT) states,” (p. 314). According to the authors, “The TOT state reflects the failure to recall a word for which one has well-established knowledge” (Vigliocco et al., 1997, p. 314). Relevant to grammatical gender, a pertinent question is whether, in such a state, speakers still have access to the grammatical gender of the word or if it is blocked until the lexical entry itself is retrievable. The authors report that “speakers in a positive TOT state do have access to syntactic features of words for which they cannot yet generate a pronunciation code,” (Vigliocco et al., 1997, p. 316). This provides strong evidence for multiple routes to gender and the idea that grammatical gender information is, as stated by Wang and Schiller (2019), stored as a separate lexico-syntactic feature in the mental lexicon that is, by itself, retrievable. This finding regarding TOT states, however, returns us to the question of route to gender. I will leave this section with a question, rather than attempt any conclusion, as more research is necessary: If speakers in positive TOT states can still access grammatical gender without a lexical or morpho-phonological route, what route are they taking, and can a connectionist network explain this effect?

In this section, we have seen a number of ways that researchers think about grammatical gender, as well as ways in which they study this phenomenon, and in doing so, we crossed a number of theoretical and methodological boundaries. By moving between formal and psycholinguistic approaches to gender, we gained a deeper understanding of the structure of grammatical gender through its typological differences, its linguistic forms, and its mental representations. By contrasting findings from behavioral experiments and brain-imaging technology, we better understand how grammatical gender comprehension and production is represented by biological mental states. And by crossing between comprehension-based and production-based methods, we better understand not only how grammatical gender is processed and produced, but also the connections between processing and production generally and their broad implications for other linguistic features and human communication as a whole.

4 Grammatical Gender in SLA

Across the border between L1 and L2 research, I find both similar and unique questions about grammatical gender, and L1-L2 differences in grammatical gender provide a rich point of entry. As a base question, we would want to know how different languages themselves influence L2 grammatical gender acquisition. Linguistic systems can differ considerably in the variability, reliability, and frequency of gender cues. These differences can show the complexity that grammatical gender espouses when it is combined with other grammatical features like case, number, and gender. This can alter the relative ease of acquisition of grammatical gender. In two studies, Kempe and MacWhinney (1996, 1998) trained two groups of L1 English speakers either German or Russian as an L2 to see which of the grammatical gender systems would be more difficult to learn. German has three genders, one plural form, and four cases which combine to form 16 unique gender, number, and case combinations. Russian also has three genders, plural, and six cases, which makes 24 unique gender, number, and case combinations. One might expect German to be easier for the learners, since there are eight fewer gender, number, and case combinations to learn. However, the researchers showed that it is not simply a function of the number of combinations needed to learn, but rather an issue of cue reliability. Each gender, number, and case combination in Russian has a unique form, whereas in German, learners run into homophony/homography. The word *der* [the] illustrates this point nicely, in that it could refer to a masculine, singular noun in nominative case; a feminine, singular noun in dative or genitive case; or a plural noun in genitive case. So, when it comes to second language acquisition, the make-up of the grammatical gender system to be learned can play a significant role in ease of acquisition.

Beyond the structure of language to be learned, the type of knowledge that a person can bring from their first language(s), or transfer, also affects L2 acquisition of grammatical gender. Formal linguists, especially those interested in UG, have tested multiple theories to answer the question of transfer, including No Transfer/No Access (Epstein et al., 1996), Partial Transfer/Full Access (e.g., Minimal Trees Hypothesis, Vainikka & Young-Scholten, 1996; Failed Functional Features Hypothesis, Hawkins & Chan, 1997), and Full Transfer/Full Access (Schwartz & Sprouse, 1996). The comparison of these models had important repercussions for formal SLA, and a number of researchers attempted to find the hypothesis that best modelled learner data. In a study of L1 English learners of Arabic, Aljadani (2019) tested the Failed Functional Features Hypothesis against Full Transfer/Full Access using a grammaticality judgment task (GJT) of demonstrative pronouns. Participants were grouped by proficiency level and compared to a native speaker group. The results of the study showed that the more proficient learners did not vary significantly from the native-speaker group, whereas the less proficient group did. Based on their findings, the authors argued that since proficiency had an effect, it “could be in some way evident to the FT/FA hypothesis,” (Aljadani, 2019, p. 84). In another study using a GJT as well as a production task, Ayoun (2007) examined learnability

of grammatical gender among L1 English learners of French. Again, the authors were interested in comparing the expected results from the Failed Functional Features Hypothesis to the Full Transfer/Full Access model. Like Aljadani's (2019) study, Ayoun (2007) noted that "participants were more accurate as their level of proficiency increased," (p. 160). While both of these studies seem to provide evidence for transferability from and access to UG, from a formal perspective, it is still a messy process. This was the case in a study by Sabourin et al. (2006), which investigated learners of Dutch from multiple L1 backgrounds (German, Romance language, or English). While the results clearly bore out expected advantages (German providing a greater advantage due to its similarity with Dutch, then Romance languages because they have grammatical gender, and finally English, where grammatical gender is all but absent), the results varied by task. The authors noted that "the German group show[ed] less advantage of surface transfer for gender agreement than for gender assignment," (Sabourin et al., 2006, p. 26), which they interpret to mean that there are different effects for knowledge transfer than language use. As the authors state, "It seems that certain aspects of grammar (e.g., gender) can be learned to a high degree of accuracy but that using this knowledge remains a problem for L2 learners," (Sabourin et al., 2006, p. 26). And while proficiency played a role on some items,

an effect of L1 remains for the middle frequency items. This shows that agreement is still more difficult than assignment. Further, the English group that performs well on the assignment task only performs at chance on the agreement task. This suggests that at least for a group with no gender in their L1, gender agreement is very difficult and may be impossible to acquire. (Sabourin et al., 2006, p. 26)

The various levels of L1 overlap with the L2 grammatical system, while still influenced by proficiency and language ability, provide at least a gray-scale picture of what transfer can look like between languages, and that sharing a similar functional feature between languages can facilitate both assignment of the L2 feature, as well as, if not as much as, production of agreement patterns.

Other theoretical perspectives informed by cognitive science and developmental psychology in SLA have also taken an interest in issues of transfer and gender assignment. In a study by Carroll (1999) on beginning L2 learners of French, the researcher questioned how grammatical gender is applied at the onset of L2 acquisition. She found that learners rely on prior semantic and abstract knowledge, not just from the "objective patterns in the speech signal" (p. 38). This means that it is not just a matter of cue distribution and reliability within the input, which is a very strong driver in L1 grammatical gender acquisition. For L2 learners, the semantic relations already stored in the L1 lexicon can affect the way learners expect grammatical gender to work in an L2. She also argues that this provides "support for theories of linguistic cognition involving mediating structural representations, as well as learning theories in which conceptual information can guide grammatical development" (Carroll, 1999, p. 38). This would include Vygotskian sociocultural theory, which places an emphasis on language as a mediating tool.

SLA researchers taking cues from L1 research know that L1 readers make use of grammatical gender, but it was unclear whether L2 learners were capable of attaining L1-like processing. In order to test whether L2 learners do or are able to behave similarly, Dussias et al. (2013) compared Spanish L1 and L2 speakers' anticipatory eye movement during a reading task. They found that "late English Spanish learners revealed sensitivity to gender marking on Spanish articles similar to that found in native speakers, but this sensitivity was affected by the level of proficiency" (Dussias et al., 2013, p. 377). Thus, from a psycholinguistic perspective, we see new reading behaviors emerge through increased noticing and processing of grammatical gender that mirrors, or at least comes close to mirroring, L1 behaviors that are used in grammatical gender processing. In another study of L1 and L2 Spanish, Grüter et al. (2012) looked at advanced L2 Spanish learners with persistent production problems of grammatical gender. The authors asked whether these production problems were the result of production-specific performance or gender information retrieval in real time. To tease out these differences, participants took part in a sentence-picture matching task, an elicited production task, and a looking-while-listening task. Results showed that in the offline language task, namely the sentence-picture matching task, advanced L2 learners performed at the same level (i.e., at ceiling level) as the L1 speakers. However, for the other tasks, which were online (real-time) tasks, the L2 group did not perform as well. The authors argued that these performance differences between L1 and L2 speakers lie at the level of representation, where a more robust L1 system allows for easier real-time processing.

This question of representation of grammatical gender in an L2 system brings us back to issues of transfer, this time from a connectionist perspective. How would a connectionist model account for differences in L1/L2 grammatical gender processing? A study by Klassen (2016) looked at how gender congruency between an L1 and L2 can affect processing. To study this, the researcher asked L1 Spanish learners of German to complete a picture naming task. In Spanish, there are only two genders (masculine and feminine), while in German, there are three (masculine, feminine, and neuter). Between these two languages, then, there are some words that overlap with the same grammatical gender (*el sombrero* [the hat] is masculine in Spanish and its German counterpart *der Hut* is as well), some that are assigned the opposite gender (*la mesa* [the table] is feminine in Spanish but its German counterpart *der Tisch* is masculine), and some that are neuter in German and either masculine or feminine in Spanish (*la casa* [the house] is feminine in Spanish but its German counterpart *das Haus* is neuter). The outcomes of the picture naming task showed "faster responses in (...) L1–L2 gender-congruent nouns than for gender-incongruent ones" (Klassen, 2016, p. 24). From a connectionist perspective, this would indicate some connectivity at the lexical level based on the shared gender information which would be speeding up access. The author posits that this "show[s] that genders common to both languages are represented as L1–L2 shared gender nodes, much like what has been shown for bilinguals whose languages have symmetric gender systems" (Klassen, 2016, p. 24). In addition to this finding, the contrast between a three- and two-gender system is of note. This is because the information stored on neuter lexical entries is not "opposite" in the sense that it is

the other gender option in the L1; it is simply different information. Surprisingly, the results show that,

neuter nouns patterned similarly to L1–L2 gender-congruent nouns, illustrating that L1–L2 gender-incongruent nouns (masculine–feminine mismatches) are subject to significantly higher levels of interference in the production of bare nouns and DPs than both L1–L2 gender congruent and L2 neuter nouns. This finding suggests that nouns of different genders in the L1 and the L2 are not all subject to the same levels of interference: gender values present only in the L2 have a distinct representation that is significantly less affected by the activation of a different L1–L2 shared gender node. (Klassen, 2016, p. 24)

In sum, the level of interference from non-congruence is highly correlated to L1/L2 similarities and differences, in addition to issues of proficiency, as was investigated in other studies on processing described earlier.

The various borders crossed in SLA reflect many of the borders crossed broadly in linguistics as discussed in the previous section. Of particular note here, though, is that the relative size of the fields does not need to be similar for the benefits of boundary crossing to be prevalent. In some ways, SLA gains an immense amount of information by crossing into theories, approaches, and findings from formal linguistics, developmental psychology, L1A, and the like, but the reverse is also true. These other fields also gain an immense amount of insight by crossing into the boundaries of more niche fields to see how their general theories play out in particular contexts and either provide evidence to strengthen their claims or force them to reconcile their theories with new evidence from specific contexts.

5 L2 Grammatical Gender Instruction

Based on the previous review, it is clear that grammatical gender is complex and often difficult to learn, although this is relative based on the L2 grammatical gender system and any opportunities for transfer from previously learned L1/L2s. Many studies provide strong evidence for difficulties learning an L2 grammatical gender system. In one study on L2 German, Walter and MacWhinney (2015) surveyed students majoring in German at US universities in their final year of their studies. They found that even among students who were about to receive a bachelor's degree in German, control over grammatical gender assignment varied considerably, with many participants only knowing approximately 50 percent of the genders of the nouns they were asked about, and there was a fairly clear lack of understanding of cues that could have provided them with information about the correct gender, even if they did not recognize the word. Knowing that grammatical gender is difficult to learn means that it is important to find evidence-based methods that can support learning. Based on my experience and interactions with fellow teachers and researchers, most pedagogues emphasize explicit instruction, awareness raising, and other fairly straight-forward approaches based on research and practices with other grammatical features. But are there differences that come about from these proposed methodologies? Is there an effect of instruction on acquisition? Instructed

SLA, as an area focused primarily on this question, would want to know how the boundary between instructed and non-instructed learners differs in their attainment, understanding, and use of grammatical gender. And within the classroom, the boundaries between different teaching methods might also play a part in this development. For example, how do comprehension-based activities differ from and production-based activities in the way they change how learners acquire, process, and use grammatical gender? Keppen et al. (2021) asked this very question and found that, for beginning L2 learners at least, production-based activities were superior for helping students on comprehension and production tasks.

In a study that utilized a cognitive tutor based on psycholinguistic principles, Presson et al. (2014) provided learners with individualized feedback on hundreds of French nouns. In each trial, learners were informed whether they were correct or not, and if incorrect, they were provided with individualized feedback about their errors. In some cases, learners received explicit rules about gender assignment in French in addition to corrective feedback, while other learners only received corrective feedback or feature focusing. The group that received explicit instruction outperformed the other two groups, although they found no effect for more versus less frequent words.

In addition to developing technological solutions to overcome the deficit in L2 gender acquisition, some researchers have tried to utilize the (fairly weak, but existent) link between sex and gender in instruction. Most of these methods have been focused on providing semantically associated information, such as objects that are stereotypically or culturally more associated with a particular sex, via mnemonic devices. Color-coding, especially, has been the focal point of textual enhancement approaches because there is at least some link between many gender systems and sex, so traditional cultural connections between some colors (e.g., pink for feminine and blue for masculine in a Western context) seemed like possible avenues to highlight bare textual information for learners and provide some additional support for gender awareness during reading. This theory was tested in studies by Desrochers (e.g., Desrochers, 1982; Desrochers et al., 1989) in the mid-1980s and early 1990s, and has received some renewed attention in recent years, especially from a deeper understanding of connectionist frameworks, where instead of just one (color) link between the gender and sex, additional semantic connections are also made. These include, for example, the sex of the actors who use or are pictured with similarly gendered nouns, as well as voiced productions of words with overlapping gender/sex categorization, i.e., a male voice actor for masculine nouns and a female voice actor for feminine nouns. Two recent studies, one by Dias de Oliveira Santos (2015) and one by Arzt and Kost (2016), both tested the effects that different combinations of these mnemonic devices can have on gender acquisition. Each study compared a control group to different conditions of color/voice/image pairings. In both cases, the researchers did not find a significant immediate effect for this type of instruction, but Arzt and Kost (2016) did find a delayed effect for both forms of visual enhancement on retention of gender information.

5.1 Grammatical Gender as Part of a Complex Grammatical-Functional System

Based on all of the issues presented with instructing grammatical gender, it is clear that simply providing learners with which genders map to which nouns is insufficient. Even many of the teaching strategies that try to bring the form to the forefront, such as visual enhancements, mnemonics, explicit instruction, and input processing approaches, seem to have little effect on the long-term acquisition of grammatical gender. I argue that one of the major reasons for the lack of both uptake and understanding of grammatical gender is the relative dearth of approaches that link this grammatical feature to its functionality, and therefore relevance, within a complex grammatical system. The function that grammatical gender plays is variable based on language, but let's take the case of German grammatical gender as one example. As previously mentioned, German grammatical gender overlaps with case and number information and is often homophonic and homographic. This leads to extensive problems beyond simply mapping gender to nouns, as the cues needed to do so are unreliable. So, rather than understanding grammatical gender as something that needs to be learned for correctness, approaches such as sociocultural theory and systemic functional linguistics put the function of the feature at the forefront rather than the form. Each language's grammatical gender feature might play a different functional role, but in German, it is intertwined with the case marking system and therefore essential knowledge to unravel sentential role. Because role is marked by case and not syntactic order, word order is very flexible in German. Thus, the functions of movement and topicalization can be used to motivate control of noun phrase declension in German. For example, if you ask a German speaker who they saw yesterday, they can easily respond "Meinen Freund sah ich." [my friend saw I], just as easily as they can respond "Ich sah meinen Freund." [I saw my friend], with the first version even being preferable because the placement of the object in the first position, marked as such by the *-en* on the possessive adjective *mein*, emphasizes and highlights the answer to the question that was actually posed.

In Walter and van Compernelle (2017) the authors enact this through concept-based instruction (CBI) (Negueruela, 2003). CBI, as defined in van Compernelle and Henery (2014),

is an approach to L2 pedagogy that centers on promoting the internalization of categories of meaning as psychological tools that mediate L2 communication. CBI draws on Vygotsky's (1986) analysis of the development of scientific thinking in formal schooling and expands and adapts Gal'perin's (1989, 1992) systemic-theoretical instruction and Davydov's (2004) germ-cell model approach to teaching subjects such as mathematics (see, e.g., Stetsenko & Arievidtch, 2010). In L2 CBI, categories of meaning are presented as abstract, systematic concepts that "are semantically driven, recontextualizable, and agentive." (p. 72)

In Walter and van Compernelle (2017), the authors, working with the framework of CBI (also concept-based language instruction, C-BLI), taught grammatical gender and case through the concepts of movement and topicalization via animated slides

and in-class enrichment activities where students practiced creating their own topicalized sentences. After comparing pre- and post-tests on a subject-identification task with reflective questions, “the CBI enrichment program not only improved learners’ scores, but also helped them to develop conscious, declarative knowledge of how the language works” (Walter & van Compernelle, 2017, p. 81). This is in contrast with previous mnemonic studies which did not show gains on immediate posttests.

In an extension of this work, Walter (2020) united this teaching framework with a cognitive-tutor implementation similar to the one described in Presson et al. (2014). In this study, four high school second-year German classes were divided into two groups. Two classes received a CBI approach with post-lesson practice on hundreds of sentence-level and noun-phrase-level exemplars, and the other two classes received traditional, text-book style explicit information before beginning with the same exemplar training. The researcher found learning in both groups, but the types of knowledge gained diverged. The more traditional approach with the large amount of corrective feedback and practice resulted in small to moderate gains in both passive and productive tasks. The CBI program, on the other hand, had larger effects for a smaller number of categories, specifically the inclusion and accuracy of attributive adjective endings and the picture-sentence matching task. Also, reflective comments provided in the post-test, which asked specifics about the grammatical gender system of German, showed differences in the type of vocabulary used by learners, with the CBI group focusing much more on the function of grammar over linguistic terminology.

The different approaches in Instructed SLA derived from multiple SLA theories, as well as other pedagogical traditions, has led to an array of methods designed to best instruct what we know about grammatical gender, how it is learned and acquired, and how best to teach it. Currently, there are a number of competing methods. Interpreting each of the learning outcomes from a single perspective is unlikely to provide a true representation of the different effects of these methods, and therefore boundary crossing will likely be needed to understand no whether one approach is “better” than another, but how each method affects learning differently. It is highly likely that a new method or combination of methods works best. One that only becomes visible when one takes an approach that is informed by diligent border crossing and interdisciplinary cooperation.

6 Boundaries Crossed and Those Still to Cross

From this review of research on grammatical gender, it is clear that fully understanding this phenomenon requires a multidisciplinary approach that crosses a number of boundaries. It is important to cross language boundaries to understand how grammatical gender is instantiated in one language versus another, and how the differences in those systems affect the way grammatical gender is assigned and used. It is important to cross semantic boundaries between terms like grammatical gender and sex, and grammatical gender and noun class, and see how the use of

certain terminology can affect the way people perceive these linguistic structures. It is important to cross methodological boundaries because different methods and theories can provide us a more complete picture of grammatical gender. It is important to cross teleological boundaries to understand how language change can affect the structure and use of grammatical gender over phylogenetic, ontogenetic, and microgenetic timescales, as well as through historical and cultural processes. And finally, it is important to cross pedagogical boundaries in an attempt to make this difficult-to-acquire linguistic feature as learnable as possible.

In the spirit of this book and the Tuckerian impact of boundary crossing, it is essential that researchers from different theoretical and methodological backgrounds engage with those from other perspectives. I believe that by providing spaces for this to occur, scholars would better understand the reasons behind others' perspectives. It is also imperative that researchers allow themselves to become boundary crossers. This means that if another perspective provides a better model to support the data, that person must be willing to either accept that other model as more accurate or be willing to alter their model to become the more accurate model of the data. I can see this from Dr. Tucker's initial assessment of French linguistics described at the beginning of this chapter. Boundary crossing is a humbling experience. One presents oneself, an expert over here, as somewhat of a neophyte over there. In these moments of humility, one can gain a new appreciation for the work and perspective of others and find oneself in a position to listen and learn, rather than to profess and explain.

It is also critical that we allow for more boundary crossing between the theoretical/experimental study of linguistic features and pedagogical/instructional implementation (cf. how Sects. 1 and 2 crosscut in this volume). I see a particular need with studies on grammatical gender to understand how the terminology we use impacts learners' assumptions and development through a combination of work from research on linguistic relativism, psycholinguistics, and instructed SLA. This will require more direct communication between researchers and teachers, and, again, providing a space where teachers see themselves as integral to our understanding of how language acquisition happens in the classroom, and where researchers see themselves as sources for classroom innovation.

References

- Aikhenvald, A. (2004). *Gender and noun class*. Walter de Gruyter.
- Aljadani, A. S. (2019). The acquisition of grammatical gender in Arabic demonstratives by English native speakers. *International Journal of Contemporary Education*, 2(1), 78–86.
- Arnon, I., & Ramscar, M. (2012). Granularity and the acquisition of grammatical gender: How order-of-acquisition affects what gets learned. *Cognition*, 122(3), 292–305.
- Arzt, J., & Kost, C. (2016). Effect of different teaching techniques on the acquisition of grammatical gender by beginning German second language learners. *Die Unterrichtspraxis/Teaching German*, 49(1), 1–23.

- Ayoun, D. (2007). The second language acquisition of grammatical gender and agreement. *French Applied Linguistics*, 16, 130–170.
- Bates, E., Devescovi, A., Pizzamiglio, L., D'amico, S., & Hernandez, A. (1995). Gender and lexical access in Italian. *Perception & Psychophysics*, 57(6), 847–862.
- Bates, E., Devescovi, A., Hernandez, A., & Pizzamiglio, L. (1996). Gender priming in Italian. *Perception & Psychophysics*, 58(7), 992–1004.
- 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.
- Blom, E., Polisenska, D., & Weerman, F. (2008). Articles, adjectives and age of onset: The acquisition of Dutch grammatical gender. *Second Language Research*, 24(3), 297–331.
- Boroditsky, L. (2001). Does language shape thought? Mandarin and English speakers' conceptions of time. *Cognitive Psychology*, 43(1), 1–22.
- 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. 61–79). The MIT Press.
- Caffarra, S., Janssen, N., & Barber, H. A. (2014). Two sides of gender: ERP evidence for the presence of two routes during gender agreement processing. *Neuropsychologia*, 63, 124–134.
- Caffarra, S., Siyanova-Chanturia, A., Pesciarelli, F., Vespignani, F., & Cacciari, C. (2015). Is the noun ending a cue to grammatical gender processing? An ERP study on sentences in Italian. *Psychophysiology*, 52(8), 1019–1030.
- Carroll, S. E. (1999). Input and SLA: Adults' sensitivity to different sorts of cues to French gender. *Language Learning*, 49(1), 37–92.
- Carstens, V. (2010). Implications of grammatical gender for the theory of uninterpretable features. In M. T. Putnam (Ed.), *Exploring crash-proof grammars* (pp. 31–58). John Benjamins.
- Davydov, V. (2004). *Problems of developmental instruction: A theoretical and experimental psychological study*. (P. Moxay, Trans.). Akademyia Press.
- De Resende, N. C. A., Mota, M. B., & Seuren, P. (2019). The processing of grammatical gender agreement in Brazilian Portuguese: ERP evidence in favor of a single route. *Journal of Psycholinguistic Research*, 48(1), 181–198.
- Desrochers, A. (1982). *Effects of an imagery mnemonic on the acquisition and retention of French article-noun pairs*. Unpublished Doctoral Dissertation. University of Western Ontario.
- Desrochers, A., & Paivio, A. (1990). Le phonème initial des noms inanimés et son effet sur l'identification du genre grammatical. *Canadian Journal of Psychology/Revue canadienne de psychologie*, 44(1), 44–57.
- Desrochers, A., Gélinas, C., & Wieland, L. D. (1989). An application of the mnemonic key-word method to the acquisition of German nouns and their grammatical gender. *Journal of Educational Psychology*, 81(1), 25–32.
- Dias de Oliveira Santos, V. (2015). Can colors, voices, and images help learners acquire the grammatical gender of German nouns? *Language Teaching Research*, 19(4), 473–498.
- Dominguez, A., Cuetos, F., & Segui, J. (1999). The processing of grammatical gender and number in Spanish. *Journal of Psycholinguistic Research*, 28(5), 485–498.
- Dussias, P. E., Kroff, J. R. V., Tamargo, R. E. G., & Gerfen, C. (2013). When gender and looking go hand in hand: Grammatical gender processing in L2 Spanish. *Studies in Second Language Acquisition*, 35(2), 353–387.
- Epstein, S. D., Flynn, S., & Martohardjono, G. (1996). Universal grammar and second language acquisition: The null hypothesis. *Behavioral and Brain Sciences*, 19(4), 746–758.
- Gal'perin, P. I. (1989). Organization of mental activity and the effectiveness of learning. *Soviet Psychology*, 27(3), 65–82.
- Gal'perin, P. I. (1992). Stage-by-stage formation as a method of psychological investigation. *Journal of Russian & East European Psychology*, 30(4), 60–80.
- Gollan, T. H., & Frost, R. (2001). Two routes to grammatical gender: Evidence from Hebrew. *Journal of Psycholinguistic Research*, 30(6), 627–651.

- 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.
- Hawkins, R., & Chan, C. Y.-H. (1997). The partial availability of universal grammar in second language acquisition: The 'failed functional features hypothesis'. *Second Language Research*, 13(3), 187–226.
- Hernandez, A. E., Kotz, S. A., Hofmann, J., Valentin, V. V., Dapretto, M., & Bookheimer, S. Y. (2004). The neural correlates of grammatical gender decisions in Spanish. *Neuroreport*, 15(5), 863–866.
- Hohlfeld, A. (2006). Accessing grammatical gender in German: The impact of gender-marking regularities. *Applied PsychoLinguistics*, 27(2), 127–142.
- Kay, P., & Kempton, W. (1984). What is the Sapir-Whorf hypothesis? *American Anthropologist*, 86(1), 65–79.
- Kempe, V., & MacWhinney, B. (1996). The crosslinguistic assessment of foreign language vocabulary learning. *Applied PsychoLinguistics*, 17(2), 149–183.
- Kempe, V., & MacWhinney, B. (1998). The acquisition of case marking by adult learners of Russian and German. *Studies in Second Language Acquisition*, 20, 543–587.
- Keppenne, V., Hopman, E. W., & Jackson, C. N. (2021). Production-based training benefits the comprehension and production of grammatical gender in L2 German. *Applied PsychoLinguistics*, 42(4), 907–936.
- Klassen, R. (2016). The representation of asymmetric grammatical gender systems in the bilingual mental lexicon. *Probus*, 28(1), 9–28.
- Köpcke, K.-M., & Zubin, D. A. (1984). Sechs Prinzipien für die Genuszuweisung im Deutschen: Ein Beitrag zur natürlichen Klassifikation. *Linguistische Berichte*, 93, 26–50.
- Köpcke, K.-M., & Zubin, D. A. (1996). Prinzipien für die Genuszuweisung im Deutschen. In E. Lang & G. Zifonun (Hrsg.), *Deutsch typologisch. Jahrbuch des Instituts für Deutsche Sprache 1995* (pp. 473–491). Walter de Gruyter.
- Kousta, S.-T., Vinson, D. P., & Vigliocco, G. (2008). Investigating linguistic relativity through bilingualism: The case of grammatical gender. *Journal of Experimental Psychology: Learning, Memory, and Cognition*, 34(4), 843–858.
- Kramer, R. (2014). Gender in Amharic: A morphosyntactic approach to natural and grammatical gender. *Language Sciences*, 43, 102–115.
- Matthews, C. (2010). On the nature of phonological cues in the acquisition of French gender categories: Evidence from instance-based learning models [DP- internal information structure]. *Lingua*, 120(4), 879–900.
- Mills, A. E. (2012). *The acquisition of gender: A study of English and German*. Springer-Verlag.
- Negueruela, E. (2003). *Systemic-theoretical instruction and l2 development: A sociocultural approach to teaching-learning and researching l2 learning*. Unpublished Doctoral Dissertation. The Pennsylvania State University.
- Phillips, W., & Boroditsky, L. (2003). Can quirks of grammar affect the way you think? Grammatical gender and object concepts. In R. Alterman & D. Kirsh (Eds.), *Proceedings of the 25th annual meeting of the cognitive science society* (pp. 928–933). Cognitive Science Society.
- Presson, N., MacWhinney, B., & Tokowicz, N. (2014). Learning grammatical gender: The use of rules by novice learners. *Applied PsychoLinguistics*, 35(4), 709–737.
- 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.
- Samuel, S., Cole, G., & Eacott, M. J. (2019). Grammatical gender and linguistic relativity: A systematic review. *Psychonomic Bulletin & Review*, 26(6), 1767–1786.
- Schwartz, B. D., & Sprouse, R. A. (1996). L2 cognitive states and the full transfer/full access model. *Second Language Research*, 12(1), 40–72.
- Stetsenko, A., & Arievidt, I. M. (2010). Cultural-historical activity theory: Foundational worldview and major principles. In J. Martin & S. Kirschner (Eds.), *The sociocultural turn in psychology: The contextual emergence of mind and self* (pp. 231–253). Columbia University Press.

- Taft, M., & Meunier, F. (1998). Lexical representation of gender: A quasiregular domain. *Journal of Psycholinguistic Research*, 27(1), 23–45.
- Tucker, G. R. (1967). *The French speaker's skill with grammatical gender: An example of rule-governed behavior*. [Unpublished doctoral dissertation]. McGill University.
- Tucker, G. R., Lambert, W. E., Rigault, A., & Segalowitz, N. (1968). A psychological investigation of French speakers' skill with grammatical gender. *Journal of Verbal Learning and Verbal Behavior*, 7(2), 312–316.
- Vainikka, A., & Young-Scholten, M. (1996). Gradual development of L2 phrase structure. *Second Language Research*, 12(1), 7–39.
- van Compernelle, R. A., & Henery, A. (2014). Instructed concept appropriation and L2 pragmatic development in the classroom. *Language Learning*, 64(3), 549–578.
- Vigliocco, G., Antonini, T., & Garrett, M. F. (1997). Grammatical gender is on the tip of Italian tongues. *Psychological Science*, 8(4), 314–317.
- Vygotsky, L. (1986). *Thought and language*. The MIT Press.
- Walter, D. R. (2020). The acquisition of German declension in additive and concept-based approaches to instruction via computer-based cognitive tutors. *Language, Interaction and Acquisition*, 11(1), 130–162.
- Walter, D. R., & MacWhinney, B. (2015). US German majors' knowledge of grammatical gender. *Die Unterrichtspraxis/Teaching German*, 48(1), 25–40.
- Walter, D. R., & van Compernelle, R. A. (2017). Teaching German declension as meaning: A concept-based approach. *Innovation in Language Learning and Teaching*, 11(1), 68–85.
- Wang, M., & Schiller, N. O. (2019). A review on grammatical gender agreement in speech production. *Frontiers in Psychology*, 9, 27–54.
- Zubin, D. A., & Köpcke, K.-M. (1981). Gender: A less than arbitrary grammatical category. *Chicago Linguistic Society*, 17, 439–449.
- Zubin, D. A., & Köpcke, K.-M. (1986). Gender and folk taxonomy: The indexical relation between grammatical and lexical categorization. In C. Craig (Ed.), *Noun classes and categorization* (pp. 139–180). John Benjamins.