Bilingualism: Language and Cognition 21 (1), 2018, 13–31 © Cambridge University Press 2016 doi:10.1017/S1366728916000766

Gender matters: From L1 grammar to L2 semantics*

SVETLANA V. COOK
University of Maryland

(Received: March 29, 2015; final revision received: June 7, 2016; accepted: June 20, 2016; first published online 9 September 2016)

The study investigates the effects of grammatical gender on bilingual processing. Native speakers of Russian (a gendered language) learning English and monolingual English controls performed a self-paced reading task in English (a non-gendered language). As predicted, bilingual speakers showed delayed latencies to gendered pronouns (he or she) that were incongruent with the noun's grammatical gender in Russian, indicating that first language (L1) grammatical gender assignment can be interpreted as biological gender in nonnative (L2) processing. The L1 gender bias was only found in sentences containing animate, but not inanimate, nouns. These results speak against the syntactic mechanism being solely responsible for gender biases, but rather support a semantic transfer account due to coactivation of linguistic and conceptual features as proposed in the sex and gender hypothesis (SAGH, Vigliocco, Vinson, Paganelli & Dworzynski, 2005). Overall, the study provides clear evidence for the L1 grammatical gender bias in bilingual processing, albeit constrained by animacy.

Keywords: non-selectivity, grammatical gender, linguistic relativity, Russian, self-paced reading

Introduction

Grammatical gender is one of the most profoundly linguistically entrenched categories. Indeed, grammatical (or formal) gender assignment is language-specific, arbitrary, and in many languages orthographically and phonologically opaque (Corbett, 1991; Comrie, 1999). It is not surprising that the acquisition of accurate gender representations presents a major challenge for second language learners. The fact that gender assignment varies so widely across languages has provided fruitful grounds for the investigation of issues that relate to bilingual representation and processing and for the linguistic relativity debate more generally (Whorf, 1956). The question of how formal gender knowledge affects representations and processing in bilinguals is the topic of the present study. More specifically, the study explores how profound the effects of native language (L1) grammatical gender representations can be on the nonnative language (L2) in bilingual processing and whether any of those effects interact with semantic features, such as animacy.

The prevailing view on bilingual lexical access has been of non-selectivity (e.g., Bilingual Interactive Activation model, BIA, Dijkstra & Van Heuven, 1998; Grainger & Dijkstra, 1992; BIA+, Dijkstra & Van Heuven, 2002), i.e., bilingual processing typically reveals activation of both languages regardless of the contextual, environmental, and task constraints (Brysbaert, 1998; Dijkstra, Van Heuven & Grainger, 1998; Spivey & Marian, 1999; Von Studnitz & Green, 2002; Dijkstra & Van Hell, 2003; Marian, Spivey & Hirsch, 2003; for review, see Kroll, Van Hell, Tokowicz & Green, 2010). Furthermore, a large body of research suggests that the direction of influence between the two (or more) languages can be predicted by language dominance, with L1 (typically, dominant) language exerting greater influence over the nonnative languages. As a result of this influence, linguistic knowledge relating to one of the languages is 'transferred' to a nonnative language system. If non-selectivity of bilingual processing assumes that L1 gender information is activated regardless of the language in use, will bilinguals show evidence of L1 gender biases in L2, and if so, can the bilingual brain tell the difference between the grammatical information, only relevant for L1, and semantic, when it comes to biological gender?

L1 gender biases

So far, the support for dissociation of formal gender information (grammatical gender) and biological gender information (semantic gender) in bilinguals has been inconsistent. Many L1 studies have found that the effects of grammatical gender are so profound that they can affect the judgment of concepts expressed in their native

* The author is deeply indebted to Nan Jiang, Kira Gor, and Robert DeKeyser for guidance at various stages of the study, Jonathan Cook for inspiration for the study, Carrie Bonilla for her comments on an earlier version of the manuscript, and to three anonymous reviewers for their insightful suggestions and comments. The remaining mistakes are my own. The study was supported in part by the School of Languages, Literatures and Cultures, and the College of Arts and Humanities at the University of Maryland.

Address for Correspondence:
Svetlana V. Cook, University of Maryland, National Foreign Language Center, 5245 Greenbelt Rd, Severn Building 810, College Park, MD 20742, USA
svcook@umd.edu

language (e.g., Belacchi & Cubelli, 2012; Konishi, 1993; Martinez & Shatz, 1996; Sera, Berge & Pintado, 1994). It is not surprising that the already existing L1 biases in categorization would consequently affect the acquisition of L2. Several studies on bilingual gender processing have shown that when the L2 translation equivalent is accessed, L1 gender becomes automatically activated, producing effects on categorization and judgment (Boroditsky & Schmidt, 2000; Phillips & Boroditsky, 2003; Boutonnet, Athanasopoulos & Thierry, 2012; Kousta, Vinson & Vigliocco, 2008; Kurinski & Sera, 2011; for a recent review, see Bassetti & Nicoladis, 2015). An often-quoted set of studies by Boroditsky and colleagues (Boroditsky, Schmidt & Phillips, 2003) has found evidence that in a group of English–Spanish and English–German bilingual speakers, the grammatical category of the language specific noun affected the semantic representation of the object, even though the participants were not aware of gender focus in the experimental tasks. Bilingual speakers were biased by the grammatical gender when assigning adjectives to pictured objects in a way that objects denoted by feminine and masculine nouns received more descriptive adjectives associated with feminine and masculine properties respectively (the 'masculinity' or 'femininity' of the adjectives were rated by independent raters). The studies of such sort, however, have received some criticism due to the fact that the focus of the studies was on the connotative aspect of lexical meaning, which is subjective and highly variable between individuals, rather than denotative meaning, which is assumed to be shared by all speakers of the same language (e.g., Cubelli, Paolieri, Lotto & Job, 2011). Regardless of the limitations, these studies have succeeded in illustrating that the influence of grammatical gender can extend beyond the morpho-syntactic domain and can affect the way bilinguals use the L2, even if the L2 does not have a formal gender system.

Recent studies, for the most part, have moved away from the experimental manipulations involving connotative judgments, opening the gender bias discussion to different conclusions. Using naming, classification and categorization tasks, more recent research has suggested that grammatical gender biases can be fully accounted for at a linguistic rather than at a conceptual level (e.g., Cubelli, Lotto, Paolieri, Girelli & Job, 2005; Cubelli et al., 2011; Bassetti, 2007). As, for example, suggested by Vigliocco and colleagues (Vigliocco et al., 2005; Vigliocco, Vinson, Arciuli & Barber, 2008), the locus of the gender effects should not be attributed to the conceptual representations, which are language-independent, but rather to the lexical-semantic representations, which are language-specific (see also Kousta et al., 2008; Bender, Beller & Klauer, 2011).

Recent neurophysiological evidence has provided, perhaps, the most convincing evidence in favor of the linguistic nature of gender biases. Boutonnet et al. (2012) report results of a picture categorization task with an event related potential (ERP) component. They show an automatic and spontaneous access to gender information in Spanish–English bilinguals, but not in English monolinguals, while performing a non-linguistic task. Access to grammatical gender was qualified by Left Anterior Negativity (LAN), typically implicated in morpho-syntactic processing (Friederici, Rueschemeyer, Hahne & Biebach, 2003). Boutonnet and colleagues conclude that "object conceptual retrieval and categorization are unconsciously affected by language-specific syntactic information, such as grammatical gender, even when such information is task-irrelevant" (Boutonnet et al., 2012: 76).

However, the current consensus on the gender biases in bilinguals is not universally shared. Contrasting evidence has been reported by Kousta et al. (2008), who found that fluent Italian–English bilinguals perform similarly to Italian and English monolinguals in each of the languages. While there were cross-language differences, bilinguals showed no Italian gender biases while performing a task in English. The authors have concluded that bilinguals in their study possess adequate linguistic representations for each respective language without L1 affecting the underlying conceptual structures, therefore, claiming that the notion of gender biases in bilinguals is unsubstantiated.

It should be mentioned, however, that the majority of the current evidence on gender biases comes from either naming studies (Cubelli et al., 2005; Kousta et al., 2008; Paolieri, Lotto, Leoncini, Cubelli & Job, 2011; Morales, Paolieri, Cubelli & Bajo, 2014) or categorization studies (Belacchi & Cubelli, 2012; Cubelli et al., 2011, Vigliocco et al., 2008; Boutonnet et al., 2012), which use the category membership assumption, or a congruency hypothesis, as a methodological premise. For example, in a naming experiment using a picture-interference task the prediction is that bare noun production times will be slower when target and distractor nouns share the same grammatical gender (or are congruent) than when they have different genders. Similarly, categorization tasks make predictions based on whether a set of pictures or words belongs to the same gender or different gender categories (e.g., Vigliocco et al., 2008; Cubelli et al., 2011). There is still an ongoing debate whether the conclusions about gender biases based on congruency assumption can, in fact, be used as evidence of altered (or more precisely, gendered-biased) mental representations. Some suggest that gender congruency effects can be, in essence, reduced to the category membership similarity. For instance, two words that are of the same gender are similar in the same way as two items of clothing are similar to each other by virtue of being included in the same category of "clothes" (e.g., Kousta et al., 2008, Eberhard, Scheutz & Heilman, 2005; but

see Pérez-Pereira, 1991, for conflicting evidence). Most importantly, the categorical approach assumes that, while grammatical gender membership does facilitate lexical access (as shown by gender congruency manipulation in picture/word interference tasks), it does not alter the underlying conceptual representations of words (Kousta et al., 2008). If this is the case, then gender-congruency effects can be seen as manifestation of a more general categorical membership, rather than as sharing the same linguistic gender-marked feature.

Effects of L1 on acquisition of L2 gender

A vast number of L2 grammar studies in an attempt to establish to what extend gender information can be acquired (or not) by speakers of a non-gendered languages have employed methods utilizing a more extended linguistic context (Coughlin & Tremblay, 2013; Dussias, Kroff, Tamargo & Gerfen, 2013; Foucart & Frenck-Mestre, 2012; Grüter, Lew-Williams & Fernald, 2012; Keating, 2009; Sabourin & Stowe, 2008; Sabourin, Stowe & de Haan, 2006; Sagarra & Herschensohn, 2011; Tokowicz & MacWhinney, 2005). The amount of literature to date is overwhelming; therefore, we will only review those studies that explore whether an abstract gender system in L1 can have transfer effects in L2.

The evidence to attest to the success of gender acquisition so far has been mixed. For example, Cornips, van der Hoek & Verwer (2006; see also Cornips & Hulk, 2008) reported no differences in success of acquisition of Dutch between the groups of bilinguals with Moroccan and Turkish languages as their L1, although Moroccan has an abstract gender system and Turkish does not, suggesting that the fact that the grammatical gender is present or absent in their L1 is not a decisive factor in the success of acquisition. Thus, no transfer of L1 gender representations to L2 occur. A similar finding was reported by Vatz (2009), when three L1 groups of different language background (L1 Spanish, L1 Dutch, and L1 English) failed to demonstrate any stable differences in gender assignment or gender agreement in L2 French. On the contrary, several other studies (Sabourin, Stowe & de Haan, 2006; Sabourin & Stowe, 2008) found support for an opposing claim that L2 acquisition of grammatical gender is indeed affected by the existence of the grammatical gender distinction in L1 and is further mediated by similarities between the two languages in terms of their gender morphology. Furthermore, many recent studies have found that the performance of L2 leaners on the offline tasks were target-like, while it is not the case for tasks targeting online processing (Coughlin & Tremblay, 2013; Dussias et al., 2013; Grüter et al., 2012). The failure of L2 leaners to produce a reliable neurophysiological correlate of syntactic processing (P600 effect)(Sabourin, Stowe & De Haan, 2008; Foucart & Frenck-Mestre, 2012; cf. Tokowicz & MacWhinney, 2005) or comparable to L1 early fixation behavior and gaze duration (Keating, 2009; Dussias et al., 2013) suggest that gender as an abstract category is not likely to be fully incorporated into the L2 processing mechanism required for native-like analysis of syntax. This evidence implies that, while the representational aspect of gender (i.e., gender assignment) can be acquired in L2, the processing (or computational) ability to process gender-marked syntax (i.e., gender agreement) might not be acquirable to a native-like level of performance. This finding is further supported by the role of L2 proficiency, reported in studies for English learners of Spanish (Keating, 2009; Sagarra & Herschensohn, 2011; Dussias et al., 2013) and English learners of French (Foucart & Frenck-Mestre, 2012; Coughlin & Tremblay, 2013), who report trends toward native-like performance as the proficiency increases.

The above-mentioned finding can be accommodated within the shallow structure hypothesis (SSH, Clahsen & Felser, 2006), suggesting that the differences in L2 processing of syntactic dependencies, including gender agreement, should be attributed to shallow processing of syntax. That is, nonnative speakers do not fully analyze the available morphological and syntactic cues the way native speaker do. Specifically, SSH suggests that L2 speakers do not possess the appropriate grammar to parse the syntactic structures, therefore, the parser is qualitatively different from that of a native speaker. However, an alternative explanation is possible. Without assuming any fundamental differences in grammar, the diverging evidence for representational and processing results in gender agreement can be attributed to the 'broken' parser itself. The hypothesis about this L2 deficit states that L2 learners lack the processing heuristic to parse new L2 syntactic structures, for which mechanisms are not yet available or sufficiently automatized. As Lardiere suggests, grammatical gender requires an additional 'layer' of mapping, or syntactic computation (Lardiere, 2000). This processing account of a syntactic deficit, espoused by Pienemann (1998), Sorace & Filiaci (2006) and Truscott & Sharwood-Smith (2004) among others, would also account for non target-like performance in online processing tasks still allowing intact representational level of abstract gender category at a lexical level. A similar idea is formulated in the morphological incongruency hypothesis (Jiang, 2004, 2007; Jiang, Hu, Chrabaszcz & Ye, 2015), which arose to account for a consistent difficulty of Chinese and Japanese learners of English as a second language (ESL) to acquire English plural morphemes. Based on the self-paced reading data from Chinese and Japanese ESL learners, who do not encode plurality morphologically in their L1, and Russian ESL learners, who do, this hypothesis also expands on the idea that L2 speakers utilize L1

https://doi.org/10.1017/S1366728916000766 Published online by Cambridge University Press

processing routines in their L2. More specifically, the processing routines dealing with morphologically poor L1 cannot accommodate routines necessary for the analysis of English plural morphology (Jiang et al., 2015).

What the reviewed literature on gender acquisition effects suggests is that the transfer of the grammatical gender information to L2 could lie outside of the semantic/conceptual domain and could be a consequence of the syntactic computational processes in use during L2 processing. Several recent gender congruency studies have attributed effects of gender to the syntactic level of representation even when the representations had to be accessed outside of the sentential context (Bender et al., 2011; Boutonnet et al., 2012; Cubelli et al., 2005; Cubelli et al., 2011). This possibility of a transfer from L1 to L2 grammar presents an alternative explanation capable of accounting for the gender biases found in L2, which is distinct from the semantic (or conceptual) motivation for these biases. In the context of the current study we will refer to this possible transfer as a 'syntactic transfer'. Unlike semantic transfer, syntactic transfer route assumes intact target-like lexical representations, including adequate language-specific semantic content, with gender biases arising as a consequence of a formal computational procedure, that requires activation of syntactic properties of lexical entries for agreement or coreference. The potential constraints on the extent of the possible syntactic transfer effects imposed by semantic features, in particular those related to animacy, will be discussed in the next section.

Gender and animacy in Russian and English

Grammatical gender assignment in general, and in Russian in particular, performs a purely formal function, especially evident in nouns that refer to inanimate objects without biological gender specifications (Corbett, 1991). Gender assignment, in turn, is further realized through two morpho-syntactic functions: first, it governs the selection of the declensional paradigm (Corbett, 1982); and second, a noun's gender determines proper agreement with verbs, adjectives, pronouns, and determiners (Aronoff, 1994). For example, in the case of noun and adjective agreement, the adjective defining a masculine noun should have a masculine gender inflection *–ый/ий* (*красивый /синий,* pretty.Masc/blue.Masc) and the adjective defining a feminine noun should have a feminine gender inflection *–ая/яя* (*красивая/синяя,* pretty.Fem/blue.Fem). The design of the study takes advantage of the fact that Russian language has three grammatical genders – masculine, feminine, and neuter. Most importantly, two of them – masculine and feminine – have a semantic core, which means that nouns that have a female referent (e.g., *дочь* 'daughter' or *львица* 'lioness') are assigned to feminine gender class and nouns with a male referent (e.g., *сын* 'son' or *лев* 'lion') are assigned to masculine gender (Corbett & Fraser, 2000). An important feature of most gendered languages, including Russian, is that in the overwhelming majority of cases if a noun denotes someone male or female, critically, an animate referent with an inherent biological gender, the grammatical gender of that noun will fully overlap with the semantic gender. For example, *бабушка* 'grandmother' is female and *брат* 'brother' is male, just as *одноклассник* 'male classmate' is marked for masculine gender and *одноклассница* 'female classmate' is marked for feminine gender. It is apparent that in such cases grammatical gender might be interpreted as semantic gender, and vice versa. Formal gender category, however, also extends to inanimate objects, whose referents do not possess any sex-differentiable characteristic. For example, *нож* 'knife' has a masculine and *вилка* 'fork' has a feminine gender assignment. As was mentioned earlier, gender assignment in most Russian nouns that do not have a biologically motivated gender assignment (also, exclusively inanimate) is arbitrary.

Alternatively, English belongs to a category of languages that are characterized by the sematic gender category (Corbett, 1991). In languages of this type there are no overt grammatical markings on ether nouns or other parts of speech that agree with nouns. The distinction between the gendered nouns is in some cases lexicalized and corresponds to the naturalistic sex distinction between males and females (e.g., *boy/girl, waiter/waitress, lion/lioness*, etc.). At the same time, English has retained some of the features from a formal gender system, which existed in the English language historically, such as sex-differentiated personal (*he/she*), possessive (*him/her*), and reflexive pronouns (*himself/herself*). Additionally, English differentiates between animate (*he/she*) and inanimate referents (*it*) in its pronominal system.

Among the first attempts to examine the relationship between grammatical and semantic gender were the studies by a group of authors (Vinson & Vigliocco, 2002; Vigliocco et al., 2005) and were motivated by the following hypothesis: gender effects arise as a consequence of a close correspondence between sex in humans and grammatical gender, which is further extended to other nouns, those that possess both the linguistic (gender of nouns) and conceptual (sex) features. Some preliminary evidence was found only for the weak version of the sex and gender hypothesis (SAGH), which states that when the language allows easy mapping of gender to animate referents that can be differentiated in terms of natural gender, the grammatical gender can be misinterpreted as semantic, which was confirmed to be true only for animate, but not inanimate referents.

The present study attempts to extend this finding with English–Russian bilinguals. Since these two types of gender largely coincide in Russian for animate

https://doi.org/10.1017/S1366728916000766 Published online by Cambridge University Press

referents (nouns denoting female referents are of feminine grammatical gender and nouns denoting male referents are masculine) (Corbett, 1991), we turn to English in an attempt to differentiate between the two in English pronoun agreement constructions. Since English does not have a formal gender feature, we will explore whether L1 grammatical gender assignment can manifest itself as a semantic category in L2 English without confounding the realization of grammatical with the semantic gender present in Russian.

Additionally, we take advantage of the differences between the two languages in their conceptualization of animacy. Both Russian and English use a general biological principle to assign semantic gender to human referents; however, the two languages exhibit important differences in how animacy is manifested in reference to animals. In Russian animal nouns (except for nouns for most domestic animals, such as *cows, sheep, goats,* etc.[1]) are epicenes and refer to both males and females members of the species while being marked for gender grammatically, as all Russian nouns do without exception. For example, *воробей* 'sparrow' is marked for masculine and *чайка* 'seagull' is marked for feminine gender regardless of the natural gender of the bird.[2] As for English, the traditional grammar prescribes using a neutral inanimate pronoun *it* when referred to the animals while the use of personal pronouns is also allowed, typically with pets, domestic animals, or towards other animals with an intent to express interest or empathy (Payne & Huddleston, 2002). To sum up, the use of gendered pronouns (feminine vs. masculine) with animal referent is compulsory for Russian and optional for English. Critically, the optionality of a gendered pronoun use in English with animate referents, and animal epicenes in particular, makes it possible for Russian bilinguals to show evidence of transfer the gender assignment from Russian to English without violating conventions of the English grammar. Therefore, any divergence in the behavior between monolingual English speakers and Russian–English bilinguals should be attributed to the differences in their gender representations stemming from L1 transfer and not ungrammaticality in the use of the English pronouns.

Present study

In order to overcome some of the methodological challenges shared by the previous studies of gender biases, here we describe an alternative method of investigating gender biases in bilinguals, which is a self-paced reading paradigm. Self-paced reading methodology has proven to be an effective method of assessing language processing in monolingual and bilingual speakers alike, because it is sensitive to any disruptions in normal processing, interfering with comprehension (for review of methodology, see Keating & Jegerski, 2014). In self-paced reading tasks the processing disruptions are introduced as an experimental manipulation. Test sentences are presented on a computer monitor one region at a time, from the beginning of the sentence to the end, imitating conventional left-to-right reading flow. The participants are instructed to press a key to advance to the next region as soon as they finish reading what is presented on the monitor. Unlike conventional reading, where there are opportunities to go back to the previously read material, self-paced reading tasks only allow rereading of the material in the region currently available on the screen. The reaction times to reading of each region are recorded and are primary data for the analysis. This methodology allows comparing reading times in sentences without a violation to sentences with a violation. Out of the two, the latter is expected to reveal any delays in processing associated with the experimental manipulation. Such processing delays are typically linked to the difficulty of a cognitive operation or to attempts to resolve a grammatical or semantic ambiguity and have been documented in extensive number of psychological and linguistic studies (e.g., Juffs & Harrington, 1995).

An additional benefit of using a self-paced reading task is in its ability to assess L2 processing and any possible influence of L1 implicitly without focusing the task specifically on gender and thus allowing us to avoid any strategic or deliberate behavior. It taps into the true online activation patterns of the available linguistic knowledge.

It should also be mentioned that many gender processing studies have previously explored the cataphoric (or predictive) qualities of agreement, such as between the gender-marked article or an adjective and the noun (e.g., Dussias et al., 2013; Sagarra & Herschensohn, 2011; Alarcón, 2010). The present study takes a different approach and explores anaphoric referential properties of the gender marked constituents, in particular, those between a gender-marked pronoun in the subordinate, or coordinate, clause 'looking back' to a noun in the main clause. This type of relationship between the constituents brings an extra dimension to the computations demands of the listener/reader, since in addition to the processing of morphology, there is a need to establish a structural coreference between the referent and the pronoun, which implies a necessity to maintain in the working memory (or keep available for recall) a mental representation of the referent for an extended period of time (Tyler & Marslen-Wilson, 1977; van Gompel & Liversedge, 2003). The type of referent–pronoun relationship explored

[1] We thank an anonymous reviewer for a suggestion to incorporate this clarification.

[2] The distinction between the male and female bird can be specified lexically if such a need arises by adding *самка* (*she*-bird) or *самец* (*he*-bird) to the name of the species.

https://doi.org/10.1017/S1366728916000766 Published online by Cambridge University Press

in the present study was motivated by practical as well as by theoretical reasons. English does not allow for gender marking outside of the pronominal system, which is a major practical constraint. At the same time, this practical limitation has a positive consequence and allows for a stronger theoretical claim: coreferential links between constituents in the different clauses in L2 grammar system are rather fragile. Furthermore, the strength of the coreferential links is affected by the structural distance, which is a primary consequence of high cognitive demands of computing the necessary structural and semantic information (Sorace & Filiaci, 2006; Pienemann, 1998). Such susceptibility to structural distance between the constituents has been also shown for gender coreference (Sabourin, 2003; Keating, 2009). On the flip side of this vulnerability, when the newly acquired processing routines have not been sufficiently entrenched, they will give way to L1 processing heuristics to affect processing as a more automatized and more readily available mechanism. As a result, if there are any syntactic effects of gender, they will be most likely to be manifested in structurally distant coreferential constructions, like the one explored in the present study.

We propose that grammatical gender representations in L1 Russian, being associated with the semantic gender, on the one hand, and with the computational heuristics, on the other, have the potential of being transferred to L2 English. In agreement with the non-selectivity account of bilingual processing, it is plausible that in order to deal with the constraints of the gender system of their L1 (in terms of declensional paradigm selection and morphosyntactic agreement), speakers of Russian automatically activate the grammatical gender component of the lexical representation. However, due to the fact that semantic and grammatical gender in Russian, for the most part, completely overlap for animate referents, a semantically-motivated gender concept is activated as well, whether relevant or not, purely by reasons of habitual coactivation. It happens regardless of the fact that the activation of the semantic gender feature becomes an inherent part of L1 processing and happens regardless of the fact that the grammatical gender concept is of no consequence for syntactic processing in non-gendered L2. It is quite possible that this mechanism is transferred to L2, regardless of whether the grammatical gender distinction exists or not in this language. These predictions align with the sex and gender hypothesis (SAGH) outlined earlier (Vigliocco et al., 2005).

What still remains to be seen is which version of the hypothesis will receive support – the strong version, which predicts gender bias effects for both animate and inanimate referents, or the weak version, which extends the predictions to animate referents only. Some preliminary evidence suggests that animate nouns are distinct from inanimate nouns in how they are processed even by the native speakers of the language. In particular, animate nouns show processing benefits during comprehension (Alarcón, 2009, 2010; Desrochers & Brabant, 1995; cf. Sagarra & Herschensohn, 2008), larger mismatch effects in sentences with violations (Deustch, Bentin & Katz, 1999), are more resistant to the attractions errors due to the adjacent location (Deutsch & Bentin 2001), and evoke fewer agreement production errors (Antón-Méndéz, 1999; Vigliocco & Franck, 1999). Advanced L2 speakers also show the same tendencies for the processing preference of the animate nouns (Alarcón, 2009, 2010; Spinner & Thomas, 2014). The superior status of animate nouns have been associated with the fact that the semantic component of gender in animate referents associated with biological gender reduces the processing demands for establishing agreement and coreference with other parts of the sentence (Deutsch et al., 1999). Alarcón (2009, 2010; see also Belacchi & Cubelli, 2012) further suggests that in pronoun resolution native speakers take equal advantage of morphological and semantic cues, while nonnative speakers favor semantic cues rather than morphological cues, with the latter being less available for automatic access. Indeed, animacy has been confirmed to have a privileged status for transfer to nonnative processing, which is accessible outside of the L1 linguistic system (semantic core hypothesis, Spinner & Thomas, 2014). If the gender biases in our data can only be observed with animate noun referents, it would confirm prior findings that the semantic gender component allows for an easier, more transparent mapping between biological gender and formal gender, leading to the misinterpretation of the grammatical gender as semantic.

Method

Participants

Two groups of participants took part in the study. The main population of interest was the bilingual group of Russian native speakers learning English ($n = 23$). Only advanced learners of English were selected for participation. The majority were current graduate students or recently graduated professionals, who have demonstrated their ability to use English in an academic setting. None of the students were enrolled in English as a second language (ESL) or other remedial-level English language classes at the time of the experiment. A screening questionnaire was filled out by all participants reporting their English language proficiency (for the bilingual group only) with self-reported ratings of their English language skills in reading, comprehension, writing and their command of grammar on a Likert scale from 0 (minimal knowledge) to 5 (native-like command). Only participants who reported their proficiency to be at 4 (good command) and above

https://doi.org/10.1017/S1366728916000766 Published online by Cambridge University Press

Table 1. *Participants' demographic and language proficiency data**.

Group	Male	Female	All
Monolingual			
Number of participants *(n)*	15	8	23
Age (*Mean*, in years)	26.4	22	24.2
Bilingual			
Number of participants *(n)*	7	13	20
Age (*Mean*, in years)	34.71	31.08	32.9
Self-assessed listening proficiency	4.57	4.54	4.55
Self-assessed reading proficiency	4.57	4.62	4.59
Self-assessed writing proficiency	4.57	4.62	4.59
Self-assessed knowledge of grammar	4.29	4.54	4.41

*Due to poor performance on the experiment four participants were dropped from the study. The details of the decision are discussed in more detail in the Analysis section of the paper. The table reflects the data for the participants retained in the study ($n = 43$).

(native-like) in all four skill areas were selected for participation (see Table 1 for details).

The second group included native English speakers ($n = 24$) who served as a monolingual control group. The candidates for participation were also carefully screened to make sure that none of them had any significant experience learning a gendered language (e.g., Spanish, Portuguese, French, German, Arabic, Hebrew, Dutch, or any Slavic languages). Only learners of Japanese, Korean and Chinese were considered eligible for participation due to the nature of the experiment.

Participants were not informed about the subject matter of the experiment in order to avoid strategic processing and prevent them from applying their explicit knowledge of the subject. After the completion of the experiment, the participants were debriefed and the administrator provided detailed information about the experiment and its purpose.

Materials

The experiment consisted of 84 sentences total with 24 critical and 60 filler sentences, presented to the participants in three counterbalanced lists. The 24 critical items were the sentences of two types based on the animacy of the referent: twelve critical base sentences had animate referents and another twelve sentences had inanimate referents (see Table 2 for examples of each type). Sentences with animate referents included epicenes, which were names of common animals, insects, and birds. Inanimate nouns were names of common household items. Special care was taken to avoid cognates or words that have a semantically motivated gender distinction as the critical noun, such as *cow* (Fem.)–*bull* (Masc.), or *rooster* (Masc.)–*hen* (Fem.). The critical manipulation involves three pronouns that differ in how (or whether) they mark gender and animacy – *he* (animate/Masc.), *she* (animate/Fem.) and *it* (inanimate/no gender). Each base sentence appeared on the presentation list in one of the three conditions – congruent, incongruent and neutral – depending on the pronoun used. The use of pronouns was balanced across conditions and across three presentation lists. The lexical frequency of the Russian translation equivalents of the English noun was controlled across types (19.89 instances per million with animate and 19.5 instances per million with inanimate referents). Filler sentences were sentences with varying sentence structure to distract the participants from the structural constitution of the sentences in the critical conditions.

All sentences were uniform in term of syntax with the pronoun appearing as the 8th word in each sentence. Each sentence was split into 6 regions, with the reference noun being in the first region and the pronoun – in the fourth region (see Figure 1). The comprehension questions were used to ensure that the participants are reading for understanding and do not simply advance through the sentence without proper processing of semantics and grammar. The distribution of the 'yes'/'no' answers to comprehension questions was balanced within as well as across the presentation lists.

Procedure

The participants were asked to read grammatical and ungrammatical sentences one section at a time and their reading time was measured at the critical regions where possible access of gender information takes place. The experiment was computer-delivered (DMDX, Forster & Forster, 2003), and the reaction time (RT) was recorded as a critical measure for further analysis. Each participant was tested individually in a quiet room on Dell® Latitude/D820 computers with Logitech® Precision USB game pads. The stimulus text was displayed in the center of

https://doi.org/10.1017/S1366728916000766 Published online by Cambridge University Press

Table 2. *Examples of critical conditions.*

	Conditions		
English grammar	Preferred grammatical acceptability	Optional grammatical acceptability	
Russian grammar	Neutral	Congruent	Incongruent
target noun: ЧЕРЕПАХА (fem.) – *turtle*			
Animate referent	The *turtle* crawled under the rock where *it* felt safe again.	The *turtle* crawled under the rock where *she* felt safe again.	The *turtle* crawled under the rock where *he* felt safe again.
target noun: ГАЗЕТА (fem.) – *newspaper*			
Inanimate referent	The *newspaper* sat on the counter when *it* caught my attention.	The *newspaper* sat on the counter when *she* caught my attention.	The *newspaper* sat on the counter when *he* caught my attention.

The **turtle** crawled under the rock where **it (he/she)** felt safe again.

1 2 3 4 5 6

Figure 1. Presentation sequence of a stimulus sentence.

the computer screen in black font (Times New Roman, size 14, bold) on a white background. The first frame of each item consisted of a set of dashes, equal to the number of symbols, including spaces, in the sentence the participant was about to read. With each press of a designated button on the gamepad the first reading segment became visible to the reader. By pressing the button the participant advanced through the sentence in a left-to-right fashion in such a way that only one segment of the sentence was available for reading, while the preceding and the following segments remained masked. The time-out for each segment was set to 5000 ms. After each sentence the participants were asked to make a 'yes'/'no' judgment in response to a question about the sentence by pressing an appropriate button on the gamepad to check for comprehension. No time-out was in place for the comprehension check. The next trial was advanced automatically after the response to the comprehension question was made.

Predictions

The following predictions were made. We hypothesize that the grammatical gender of the epicenes can affect bilingual performance in L2 in establishing coreference between an antecedent (which is an epicene in the animate condition or an inanimate object in the inanimate condition) with a gender-marked pronoun in the subordinate/coordinate clause. If the L1 speakers of Russian transfer gender information to L2, then differences are to be found between the performance of the bilingual group and the monolingual English controls as a function of the experimental condition – neutral, congruent, or incongruent. The neutral condition uses the pronoun *it*, which is the default grammatical option. The other two conditions – congruent and incongruent – are both possible grammatically, but dispreferred from the perspective of acceptable English grammar. However, for the bilingual speakers there is a possibility that the congruent condition will also be interpreted as an acceptable grammatical sentence, because it adheres to the grammatical rules of the L1. In the congruent condition, the Russian translation of the English noun will be congruent with the feminine or masculine interpretation of the pronoun as appropriate. For example, in the sentence in Figure 1, the Russian translation equivalent of *turtle* is *черепаха*, which is feminine, so when the pronoun *she* is used, the sentence is in the congruent condition; however, with the masculine pronoun *he*, the sentence becomes incongruent. The difference in reaction times for the Russian speakers is hypothesized to be in the incongruent condition, where they would tend to slow down in order to process the incongruency between the English pronoun and the Russian gender assignment of the antecendent.

In relation to the animacy manipulation, the predictions are as follows. If bilingual speakers attribute specific semantic properties to animate English nouns, as assumed by the semantic transfer account, then we will find evidence of grammatical gender being misanalysed as semantic. This will likely to be the case for animate referents only, who possess a biological gender, and masculine or feminine formal gender assignment have a potential of being interpreted as a natural gender. Inanimate nouns do not lend themselves to a semantic interpretation of gender and, therefore, will be resistant to this type of transfer. This outcome will provide support for the weak version

https://doi.org/10.1017/S1366728916000766 Published online by Cambridge University Press

of SAGH. Conversely, if both animate and inanimate nouns show evidence of transfer, then the data will support formal syntactic transfer account, suggesting that computational L1 routines activate formal grammatical category during L2 processing, giving rise to gender congruency effect irrespective of the animacy of the noun. Evidence in favor of the syntactic transfer account will provide support for the strong version of SAGH.

Results

Before the analysis, all data points that were below 50 ms and 2 SD above each subject's mean reaction time (RT) were eliminated from further analyses as outliers. In addition, data of individual participants were removed if they demonstrated less than 80% accuracy on the comprehension questions as indicative of insufficient English proficiency or failure to follow the instructions for completion of the experiment. Thus, data from four subjects – one monolingual control and three bilingual participants – were removed from further analysis. This trimming procedure affected 8.2% of the data.

We used R (R Core Team, 2012), *lme4* package (Bates, Maechler & Bolker, 2012), and *lmerTes*t package (Kuznetsova et al., 2013) to perform a linear mixed-effects model[3] analysis of the reaction time (RT) to the reading regions of interest. The main analyses focus, in particular, on two regions – the *critical* region, where the pronoun appeared, and the immediate *spillover* region (Rayner & Duffy, 1986), or region immediately following the pronoun, to account for any delayed processing effects (regions 4 and 5 respectively, see Figure 1). The model tested whether RTs varied as a function of gender assignment of the noun in Russian between the two populations of interest: monolingual speakers of English and L1 Russian bilingual speakers of English. The analyses for the two reading regions were carried out separately.

There were three main factors of primary interest: Animacy, Condition, and Group. Separate models were fitted to each of the Animacy conditions (Animate and Inanimate). The decision to analyze the two sets of data separately was motivated by two considerations. First, from the perspective of the experimental design, the direct comparison across the two Animacy conditions has little justification. The base sentences in animate and inanimate conditions were different; therefore, we can expect a great deal of variability in the overall responses between the sentences in the two conditions attributable to the lexical variation between the sentences rather than to an experimental manipulation related to animacy per se. More adequate conclusions can be drawn based on experimental manipulation of gendered personal pronouns within each of the animacy blocks, because the same base sentences are used in all three conditions and, therefore, the influence of any confounding factors is minimized.

Second, mixed-effects design, while parsimonious in many respects, does not offer a straightforward and comprehensive way of interpreting complex interactions. We also recognize that, while splitting the analyses in this way can facilitate interpretation of the effects, it does not provide a direct test of the most critical aspect of the data namely the three-way interaction between Animacy, Condition, and Group.[4] To address this issue, here we adopt a model-comparison approach based on the χ^2 likelihood ratio test, which is similar to testing a significance of a three-way interaction in ANOVA (Faraway, 2005). The χ^2 tests allow us to compare the data fit of each model to the observed data (of a model with a three-way interaction to the model without it) to see if the interaction in question accounts for the additional variability found in the data. A χ^2 likelihood ratio test confirmed that the model with a three-way interaction between Animacy, Condition, and Groups fit the data significantly better than a model that differs only by leaving out this interaction (χ^2 (15) = 7.31, $p < .05$ for Critical, and χ^2 (15) = 8.89, $p < .05$ for Spillover region). Thus, the results of the test provide grounds for comparing the results of the two separate models (Animate and Inanimate) with an implication of significant interaction of the Animacy conditions with the other fixed effects.

First, we will discuss the analyses of the animate condition, then the analyses of the inanimate condition.

Sentences with animate referents

To each reading region (Critical and Spillover) we fitted a full model with two main experimental factors (Condition and Group), the interaction term between the main factors, and the two random factors (Subject and Item). The data was treatment-coded for the fixed effects (Condition: Congruent, Incongruent, and Neutral; and Group: Monolingual and Bilingual), with the first level of each variable being the reference level for the model.

[3] Linear mixed-effects (LME) models offer several important advantages over traditional ANOVA models. Unlike traditional ANOVA with a by-subject and a by-item analyses, which treat random effects of participants and trials as independent, LME models allow to conduct an analysis of variance within items and participants in the same comprehensive analysis, allowing a crossed-design of random effects (Baayen, Davidson & Bates, 2008; Barr, Levy, Scheepers & Tily, 2013). Most importantly, LME modelling does not require averaging across subjects or items; therefore, LME approach allows for the model to be fitted to the actual data with its true distribution. As a consequence, compared to traditional by-subject and by-item analyses, LMEs can account for a greater proportion of variance in the data (Baayen et al., 2008). In addition, LMEs reduce Type 1 error (Clark, 1973) and perform better in handling of unbalanced data sets and categorical data (Dixon, 2008; Jaeger, 2008).

[4] We thank an anonymous reviewer for correctly insisting on this point.

https://doi.org/10.1017/S1366728916000766 Published online by Cambridge University Press

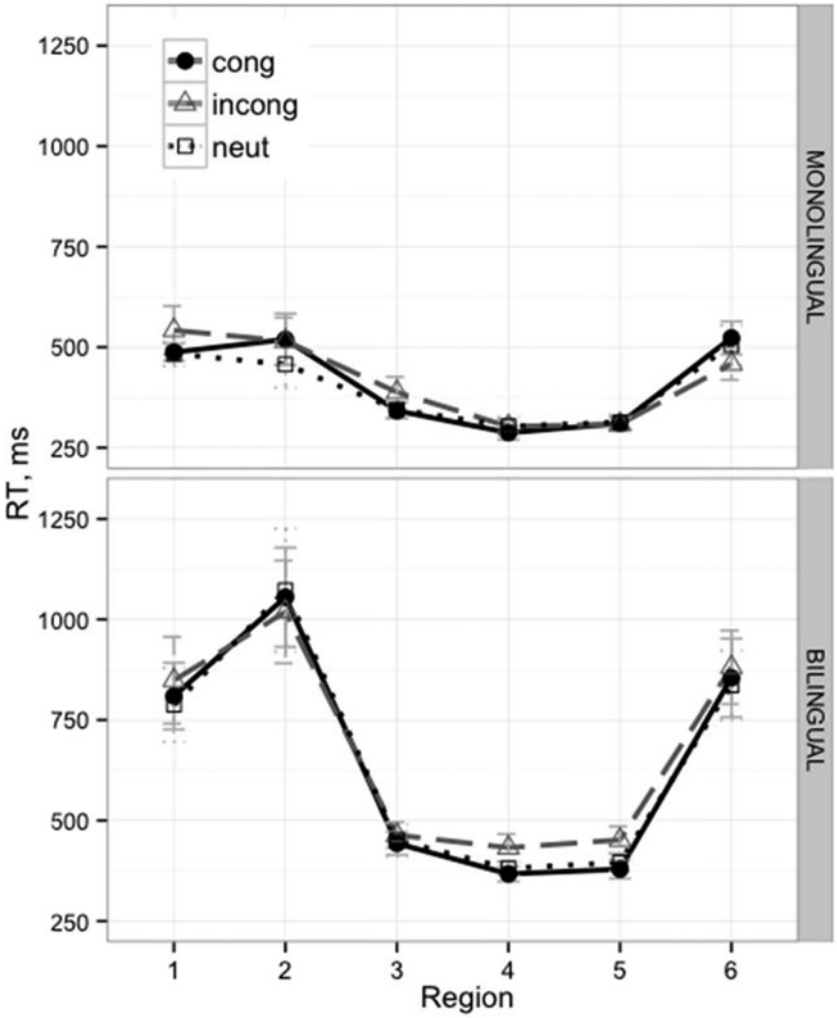

Figure 2. Mean reaction times (RTs) for sentences with animate referents (cong - congruent, incong - incongruent, neut - neutral).

Analogous to modeling by-subject and by-item ANOVA effects, random intercepts (i.e., allowing the overall RT intercept to vary by item and by subject) were included in all models. For a graphic representation of the results, see Figure 2.

For the critical region (Table 3) the analysis confirmed that the groups performed differently in the Congruent condition with slower reading times in the Bilingual group ($\beta = 79.55, SE = 30.47, t = 2.61, p < .01$). There were no statistically significant differences between the conditions in the Monolingual group, where the Congruent condition was not different from the Incongruent ($\beta = 15.87$, $SE = 21.72$, $t = 0.73$, $p = .465$) or the Neutral condition ($\beta = 19.64$, $SE = 0.90$, $t = -7.67$, $p = .371$). This was not the case for the Bilingual group. When Bilingual group was used as a reference group we found that RTs to a gender-incongruent pronoun were significantly slower than to a gender-congruent pronoun ($\beta = 65.00, SE = 22.21, t = 2.93, p < .05$); however, there was no difference between the Congruent and Neutral conditions ($\beta = 13.72$, $SE = 22.36$, $t = 0.61$, $p = .540$). To compare Neutral and Incongruent condition we conducted two additional analyses, this time with Neutral condition as a reference level of the Condition factor. Monolingual participants still showed no differences, but the Bilingual group showed a statistically significant effect ($\beta = 51.28$, $SE = 22.35$, $t = 2.30$, $p < .01$), with gender-incongruent pronouns delaying the RTs compared to a Neutral condition. The last analysis also revealed a marginally significant interaction term ($\beta = 55.01$, $SE = 31.31$, $t = 1.76$, $p = .079$), indicating that the magnitude of difference in RTs between the Neutral and the Incongruent condition were different in the two language groups, with bilingual participants displaying a greater RT difference between the two conditions by approximately 55 ms.

https://doi.org/10.1017/S1366728916000766 Published online by Cambridge University Press

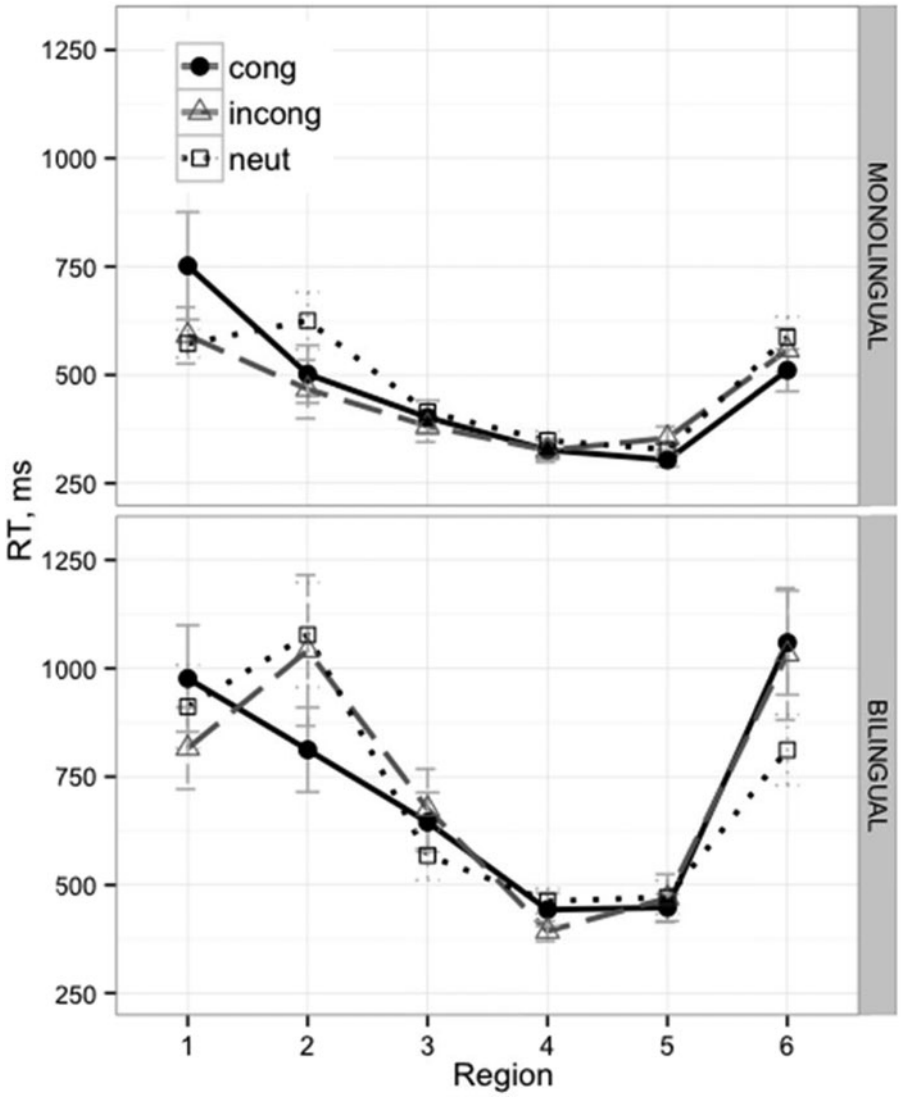

Figure 3. Mean reaction times (RTs) for sentences with inanimate referents (cong - congruent, incong - incongruent, neut - neutral).

The results of the analyses of the spillover region were similar to the results in the critical region (Table 4). Here we also found an effect of group for the Congruent level of the Condition factor ($\beta = 68.66$, $SE = 29.84$, $t = 2.3$, $p < .05$), indicating an advantage of the monolingual group. Similarly to the critical region results, only the Bilingual group showed an effect of a gender-incongruent pronoun, which delayed RTs by approximately 72 ms compared to a gender-congruent pronoun ($\beta = 72.38$, $SE = 22.8$, $t = 3.17$, $p < .01$). The model also revealed a significant interaction between Group and Condition, which is no longer marginal ($\beta = 73.65$, $SE = 31.91$, $t = 2.31$, $p < .05$). It is not uncommon for processing effects to have a stronger manifestation not at the point of the violation, or at the critical region, but rather at the spillover region, which can also be seen in our results (Pearlmutter, Garnsey & Bock, 1999; Sharkey & Sharkey, 1987). The spillover phenomenon can be associated with a delayed sensitivity to violation due to sentence-level content integration effects (Rayner & Duffy, 1986) and is particularly robust for L2 readers, due to the fact that the lack of native-like automaticity in L2 reading and the impaired access to lexical and grammatical representations delay the reaction to ungrammaticality, thus postponing it in the course of sentence processing until the spillover region is reached (Jiang, 2007).

The effect of incongruency was also found when comparing Incongruent and Neutral conditions in the Bilingual group ($\beta = 54.14$ $SE = 22.95$, $t = 2.36$, $p < .05$), which was not the case for the Monolingual group ($\beta = –6.12$, $SE = 22.51$, $t = –0.27$, $p = .786$). However, the effect of the interaction term (Condition x Group) in the Neutral condition was no longer robust ($\beta = 60.26$, $SE = 32.15$, $t = 1.87$, $p = .062$).

https://doi.org/10.1017/S1366728916000766 Published online by Cambridge University Press

Table 3. *Animate noun condition: Model estimates for critical region ((a) Congruent condition as a reference level; (b) Neutral condition as a reference level).*

	Group							
	Monolingual				Bilingual			
	β	*SE*	*t* value	*p* value	β	*SE*	*t* value	*p* value
(a)								
Congruent (intercept)	287.37	23.29	12.34	>0.001***	366.93	23.73	15.46	>0.001***
Group	79.55	30.47	2.61	0.011*	−79.55	30.47	−2.61	0.011*
Incongruent	15.87	21.72	0.73	0.465	65	22.21	2.93	0.004**
Neutral	19.64	21.92	0.90	0.371	13.72	22.36	0.61	0.540
Group × Incongruent	49.13	31.08	1.58	0.115	−49.13	31.08	−1.58	0.115
Group × Neutral	−5.92	31.31	−0.19	0.850	5.92	31.31	0.19	0.850
(b)								
Neutral (intercept)	307.02	23.48	13.08	>0.001***	380.65	23.86	15.95	>0.001***
Group	73.63	30.71	2.40	0.018*	−73.63	30.71	−2.40	0.018*
Congruent	−19.64	21.92	−0.90	0.371	−13.72	22.36	−0.61	0.540
Incongruent	−3.77	21.92	−0.17	0.864	51.28	22.35	2.30	0.022*
Group × Congruent	5.92	31.31	0.19	0.850	−5.92	31.31	−0.19	0.850
Group × Incongruent	55.05	31.31	1.76	0.079	−55.05	31.31	−1.76	0.079

Table 4. *Animate noun condition: Model estimates for spill over region ((a) Congruent condition as a reference level; (b) Neutral condition as a reference level).*

	Group							
	Monolingual				Bilingual			
Model	β	*SE*	*t* value	*p* value	β	*SE*	*t* value	*p* value
(a)								
Congruent (intercept)	310.21	23.57	13.16	>0.001***	378.87	23.98	15.46	>0.001***
Group	68.66	29.84	2.30	0.023*	−68.66	29.83	−2.61	0.023*
Incongruent	−1.28	22.30	−0.06	0.954	72.38	22.8	2.93	0.002**
Neutral	4.84	22.50	0.22	0.830	18.24	22.95	0.61	0.427
Group x Incongruent	73.65	31.91	2.31	0.021*	−73.65	31.91	−1.58	0.021*
Group x Neutral	13.40	32.15	0.42	0.677	−13.4	32.15	0.19	0.677
(b)								
Neutral (intercept)	315.05	23.76	13.26	>0.001***	397.11	24.12	16.46	>0.001***
Group	82.06	30.09	2.73	0.007**	−82.06	30.09	−2.73	0.007**
Congruent	−4.84	22.50	−0.22	0.830	−18.24	22.95	−0.79	0.427
Incongruent	−6.12	22.51	−0.27	0.786	54.14	22.95	2.36	0.019*
Group x Congruent	−13.40	32.15	−0.42	0.677	13.40	32.15	0.42	0.677
Group x Incongruent	60.26	32.15	1.87	0.062	−60.26	32.15	−1.87	0.062

Sentences with inanimate referents

The analyses of the data from sentences with the inanimate referent nouns followed the same coding scheme and were fitted with the same model as the data with the animate nouns (Tables 5 and 6).

For the critical region the analysis showed that bilingual readers were slower than monolingual readers across all conditions (β = 109.12, *SE* = 36.07, *t* = 2.61, *p* < .01). Monolingual group showed no sensitivity to the experimental conditions. The RTs in both Incongruent and Neutral conditions were no different from the Congruent

https://doi.org/10.1017/S1366728916000766 Published online by Cambridge University Press

Table 5. *Inanimate noun condition: Model estimates for critical region ((a) Congruent condition as a reference level; (b) Neutral condition as a reference level).*

	Group							
	Monolingual				Bilingual			
Model	β	*SE*	*t* value	*p* value	β	*SE*	*t* value	*p* value
(a)								
Congruent (intercept)	317.20	31.96	9.93	>0.001***	426.32	32.9	12.96	>0.001***
Group	109.12	36.07	3.03	0.003**	−109.12	36.07	−3.03	0.003**
Incongruent	19.11	34.53	0.55	0.580	−41.2	35.26	−1.17	0.243
Neutral	31.33	42.89	0.73	0.474	36.5	43.67	0.84	0.413
Group × Incongruent	−60.31	43.77	−1.38	0.169	60.31	43.77	1.38	0.169
Group × Neutral	5.17	40.37	0.13	0.898	−5.17	40.37	−0.13	0.898
(b)								
Neutral (intercept)	348.53	35.04	9.95	>0.001***	462.82	35.40	13.07	>0.001***
Group	114.28	34.16	3.35	0.001**	−114.28	34.16	−3.35	0.001**
Congruent	−31.33	42.89	−0.73	0.474	−36.50	43.67	−0.84	0.413
Incongruent	−12.22	42.85	−0.29	0.779	−77.69	42.97	−1.81	0.086
Group × Congruent	−5.17	40.37	−0.13	0.898	5.17	40.37	0.13	0.898
Group × Incongruent	−65.48	39.93	−1.64	0.102	65.48	39.93	1.64	0.102

Table 6. *Inanimate noun condition: Model estimates for spillover region ((a) Congruent condition as a reference level; (b) Neutral condition as a reference level).*

	Group							
	Monolingual				Bilingual			
Model	β	*SE*	*t* value	*p* value	β	*SE*	*t* value	*p* value
(a)								
Congruent (intercept)	314.92	41.48	7.59	>0.001***	479.18	42.83	11.19	>0.001***
Group	164.26	49.99	3.29	0.001**	−164.26	49.99	−3.29	0.001**
Incongruent	31.44	47.26	0.67	0.506	−13.61	48.25	−0.28	0.778
Neutral	13.07	53.60	0.24	0.810	−7.51	54.79	−0.14	0.892
Group × Incongruent	−45.06	61.00	−0.74	0.461	45.06	61.00	0.74	0.461
Group × Neutral	−20.58	56.35	−0.37	0.715	20.58	56.35	0.37	0.715
(b)								
Neutral (intercept)	327.99	43.78	7.49	>0.001***	471.68	44.35	10.64	>0.001***
Group	143.68	47.33	3.04	0.003**	−143.69	47.33	−3.04	0.003**
Congruent	−13.07	53.60	−0.24	0.810	7.51	54.79	0.14	0.892
Incongruent	18.37	53.57	0.34	0.735	−6.10	53.76	−0.11	0.911
Group × Congruent	20.58	56.35	0.37	0.715	−20.58	56.35	−0.37	0.715
Group × Incongruent	−24.48	55.73	−0.44	0.661	24.48	55.73	0.44	0.661

condition ($\beta = 19.11$, $SE = 34.53$, $t = 0.55$, $p = .580$, and $\beta = 31.33$, $SE = 42.89$, $t = 0.73$, $p = .473$, for Incongruent and Neutral conditions, respectively). The Bilingual group also showed no differences in RTs compared to the Congruent condition (β = –41.20, SE = 35.26, t = –1.17, $p = .240$; $\beta = 36.50$, $SE = 43.67$, $t = 0.84$, $p = .410$, for Incongruent and Neutral conditions, respectively). When Neutral condition was used as a reference level, there were still no statistically significant differences between conditions in either of the groups.

https://doi.org/10.1017/S1366728916000766 Published online by Cambridge University Press

The spillover region has replicated the lack of incongruency effect reported in the critical region. Only the effect of group was still found to be statistically significant, indicating slower reading speed in bilinguals compared to monolinguals ($\beta = 164.26$, $SE = 49.99$, $t = -3.29$, $p < .001$; $\beta = 143.68$, $SE = 47.33$, $t = 3.04$, $p < .01$, with Congruent and Neutral conditions, respectively, as a reference level). For graphic representation of the data, see Figure 3.

Discussion

In the animate condition, the results show a robust effect of Russian gender incongruency in L2 English reading by bilingual speakers of Russian. We found clear evidence of extensive cognitive effort associated with attempts in resolving the inconsistency between the gender assignment of the Russian equivalent of the preceding noun and the gender-marked personal pronoun. As our experiment suggests, when bilingual speakers encounter a pronoun inconsistent in gender to a preceding gendered noun (albeit in non-gendered L2), the processing falters. If gender information had not been accessed, both congruent and incongruent pronouns would have been processed in similar ways, which is not what our results show. On the contrary, our results suggest that the activation of L1 gender is, in fact, automatic and unconscious and that L1 is active even while performing an exclusively L2-focused task, which does not require, implicitly or explicitly, access to L1 lexical or syntactic knowledge.

In the inanimate subset we found no effects of gender in either of the groups. In contrast to the results of the sentences with animate noun referents, where we found clear evidence of the bilingual group's sensitivity to gender incongruency, with the inanimate noun referents we observed no such effects.

General discussion

The main research question of our study has been to investigate whether L1 grammatical gender can be transferred to a non-gendered L2 and whether the transfer has a semantic core, arising from the interaction of formal and biological gender of the referent as predicted by SAGH, or whether the transfer has a purely syntactic motivation and produces gender biases as a consequence of L1 processing heuristics.

The study found that Russian–English bilinguals, but not monolingual English speakers, have shown evidence of L1 gender biases while processing L2 English sentences: bilingual speakers have experienced processing costs involved in resolving the agreement incongruency between grammatical gender assignment of Russian nouns in processing of gendered English personal pronouns. The finding was constrained by animacy in a way that, while the gender bias was observed for nouns with the animate antecedents, no such bias was found in sentences with inanimate antecedents.

The present study has also set out to further explore via which route L1 gender information gets accessed during the L2-specific tasks. There are two possible routes that seem plausible: a semantic transfer route and a syntactic transfer route. Evidence for one or the other route hinges upon the bilingual group's performance on the sentences with inanimate referents.

Let's consider the syntactic processing route first. In line with this proposal gender information is accessed automatically regardless of the noun's semantic properties. Given that the mechanisms involved in the transfer are exclusively syntactic, animacy being a semantic feature plays no role in how semantic content is being integrated (and evaluated) during processing. The syntactic transfer route also assumes that gender effects with inanimate objects can still be observed in the absence of the semantic association with formal gender, but only in the case when activation of the formal gender is independent of the semantic content and is accessed on its own right as a part of the processing heuristics. As a result, both animate and inanimate nouns should show evidence of gender biases in L2 processing.

Our data do not support this proposal. Inanimate nouns showed no processing delays associated with gender incongruency. We interpreted the finding to mean that L1 gender information had not been accessed when the L2 noun was processed to bias the outcome of the pronoun coreference. There is, however, still a possibility, which should be entertained, that both types of nouns – animate and inanimate – received activation of the L1 gender feature, which gets passed down to the point of disambiguation of the pronoun, but animate and inanimate pronouns behave differently when establishing a coreference back to the referent. Since activation and interpretation of semantic information during sentence processing is delayed compared to activation and processing of the syntactic and morphological information (van Gompel & Liversedge, 2003; Kazanina, Lau, Lieberman, Yoshida & Phillips, 2007; Tyler & Marslen-Wilson, 1977), then the lack of gender effect can be attributed to the semantic coreference check, and not as evidence of prior gender activation, when the noun was accessed. When the pronoun is assessed for content integration, it is evaluated for semantic plausibility of an already established syntactic coreference (e.g., Deutsch et al., 1999; Hagoort, 2003). Recall that English personal pronouns *he* and *she* can only mark gender in animate referents, while *it* marks inanimate referents and is not gender-specific. In Russian, however, the masculine- and feminine-marked personal pronouns *он* and *она* can be used to refer to animate as well as to inanimate referents, performing primarily

https://doi.org/10.1017/S1366728916000766 Published online by Cambridge University Press

the morphosyntactic function of gender agreement. It can be argued that the lack of incongruency effect with inanimate nouns can be explained by the nature of the English pronominal system. If we assume that the noun access triggered activation of an inanimate feature, then animate pronouns would show no congruency effects, not because the gender feature was not activated for inanimate referents but rather due to the fact that animacy feature is blocked for interpretation at the pronoun. Consequently, both animate pronouns – *he* and *she* – will display similar behavior. At the same time, the use of English *he* and *she* by Russian–English bilinguals is not likely to be constrained by animacy, similarly to Russian use of *он* and *она*, and from the point of view of Russian grammar English sentences should be perceived perfectly well-constructed (e.g., The train has just arrived. *He was 2 hours late. – *Поезд только что прибыл. Он опоздал на 2 часа*.). If Russian–English bilinguals are governed by Russian syntactic processing heuristics and gender is selected for activation regardless of the semantics of the noun, gender biases will be able to emerge with inanimate referents to the same extent as with the animate referents. This is not what we found. In fact, neither of the group showed any animacy or gender dissociation in the inanimate conditions and displayed a similar pattern. Originally we hypothesized that L1 speakers of English would find the two conditions – with animate pronouns *he* and *she*, on the one hand, and inanimate pronoun *it*, on the other – to be of a various degree of acceptability. For animals English grammar favors the use of an inanimate pronoun unless the attitude towards the animal is marked by affection or other strong emotion (Payne & Huddleston, 2002). As our study suggest, this was not the case. The monolingual group did not prefer the default grammatical condition to the sentences with optional grammaticality. We found that there was no difference between the reaction times of the English speakers to pronouns in the neutral condition (*it*) and to pronouns in the two other conditions (*he* or *she*) in either animate or inanimate condition. The original predictions were not supported by the study. A potential explanation for the lack of sensitivity to an ungrammaticality could be along the same lines as the lack of the delay in the congruent condition in the Russian group. Similarly to the L2 group, English native speakers could interpret the referent of the animate noun as a sexuated entity, whether a 'he' or a 'she'. This interpretation does not violate the rules of the English grammar and is plausible in terms of their world knowledge. It should be noted, however, gender specifications attributed to animal nouns vary quite extensively among monolingual individuals even in respect to the same referent (e.g., Marcoux, 1973). For example, in a monolingual experiment a *canary* was attributed male gender by 23 respondents and female gender by 7, while 69 still preferred an inanimate pronoun; at the same time, a *parakeet* was treated as female by 42 and male by 40 respondents, while 14 respondents kept to an inanimate reference choice. The data in Marcoux's study can easily accommodate the results we report here. To the speaker of English it is not incomprehensible that a *squirrel* can be referred to as a 'she', but it can also be a 'he' with a similar outcome. For the L1 Russian speaker only one – congruent – option is acceptable, since it does not violate the agreement, constrained by the L1 gender system.

As far as the bilingual group is concerned, there is no evidence to substantiate a claim that bilinguals utilized L2-specific syntactic processing routines for inanimate, but not for animate referents. As the results of the animate condition indicate, bilingual participants treated congruent and neutral condition as synonymous. This important result allows us to argue that in relation to the use of inanimate pronoun bilinguals were governed by the English syntactic constrains, which allows for a congruent interpretation of the anaphoric dependency only within the English grammar system. In Russian *it* corresponds to a neuter-marked gender *оно*, interpretation of which as grammatical is only possible with a neuter-marked referent. Therefore, we argue that bilinguals performed within the constraints of the target language grammar in both animate and inanimate conditions and did not fall back on L1 syntactic processing routines. We interpret this result to mean that the lack of incongruency effect, observed in the bilingual group with the inanimate referents, is not due to the semantic checking mechanisms at the coreference check but rather due to lack of activation of the gender feature at the point of accessing the inanimate antecedent noun. This point discounts the syntactic transfer as a possible source of gender biases.

In sum, the current study found a bilingual gender effect with animate, but not inanimate, nouns, thus indicating that gender biases are constrained by the semantic properties of the nouns, which lie at the core of the transfer (e.g., Andonova, D'Amico, Devescovi & Bates, 2004; Sera, Elieff, Forbes, Burch, Rodriguez & Dubouis, 2002; Belacchi & Cubelli, 2012; Bender et al., 2011). The role of animacy in gender transfer, previously outlined in the weak version of SAGH (Vigliocco et al., 2005; see also Sera et al., 2002), has found full support in our study. Although the results of our experiment suggest that activation of L1 gender information is non-selective in terms of language and is accessible during L2 use, it is affected by the semantic, or real-life, interpretability of the gender feature more generally. The outcome of our study is also in conformity with findings by Alarcón (2009, 2010), who suggested that being a salient clue semantic gender reinforces the formal gender representation. Echoing the main claims of SAGH, stronger associations between the formal and semantic gender representations lead to a stronger activation of

https://doi.org/10.1017/S1366728916000766 Published online by Cambridge University Press

the semantic gender when formal gender is accessed. As a consequence, the stronger association between the semantic and formal gender leads to formal gender misanalysis as semantic gender, causing the bilingual processing confusion. Similar semantic-syntactic blends resistant to the dissociation of syntactic meaning from the semantic meaning have been reported for child language acquisition that persists into adolescence (e.g., Skeide, Brauer & Friederici, 2014). While it is not implausible that inanimate nouns could develop a similar to animate nouns association between semantic and formal gender, it is likely to be extremely weak, because semantic, or biological gender, does not typically apply to inanimate objects; therefore, no semantic content reinforces the association link with the formal gender assignment.

Some of the recent developments in our understanding of human cognition provide additional evidence in favor of the semantic-syntactic blend hypothesis. It has been proposed that speakers of different languages can end up with different patterns of associations between labels and external objects, which emerge as a product of operations over the mental representations (Lupyan, 2012). Most importantly, the distinction between different levels of representations (e.g., linguistic and nonlinguistic) should be abandoned, since most representations are "a hybrid of visuo-linguistic experience" (Lupyan, 2012: 4). In line with SAGH, the hybrid approach can be extended to mean that by virtue of repeated activation of gendered nouns, grammatical gender becomes an inherent part of the noun's representation, possibly blending with its semantic content.

Additional support for this view comes from a computer simulation study. Dilkina, McClellan & Boroditsky (2007) entered linguistic (e.g., grammatical gender, syntactic markings) and non-linguistic (what it looks like, what it sounds like, etc.) information about an entity into a connectionist model to explore whether the system's sensitivity to variation would be able to replicate grammatical gender effects. The results mirrored the reports on human subjects data from Boroditsky et al. (2003) in a way that the network's description of an object was feminine- or masculine-biased based on the assigned grammatical gender. By relying on coherent covariation the network learns to exhibit partial sensitivity in cases when such sensitivity should not be found, as, for example, a biological gender bias in inanimate objects, which are devoid of biological gender. One of the conclusions in Dilkina et al. (2007) echoes those of Vigliocco et al. (2005) and Lupyan (2012), suggesting that gender effects in non-linguistic tasks are possibly mediated by an on-line mechanism, in which a coactivation of linguistic and non-linguistic information can interact.

The present study has important implications for the debate of language processing in bilinguals. The reported experiment has found full support for the non-selectivity hypothesis, stating that L1 representations are in fact active while processing an L2, introducing an L1 bias and allowing speakers of Russian to attribute gender characteristics to the non-gendered L2 nouns. What is most important about our findings is that the information being transferred is not directly related to biological gender (or semantic, in Corbett's terms), but is a part of the linguistic knowledge. The gender feature is a formal attribute of the language that was acquired with and through the native language. As our results show, the monolingual group has no gender biases, evidencing that the source of the gender biases in the bilinguals lies in the grammatical system of the L1 Russian. This is the case where purely linguistic information obtained through the exposure to one language exposure (L1 Russian, in our case) can affect the perception of the linguistic information in a different linguistic system (L2 English).

Our findings are in line with the meaning-activated approach instantiated in Levelt's well-known model of speech production (Levelt, 1989). Levelt suggests that since different languages can encode the same meaning at different linguistic levels, depending on a language the same preverbal message will require a different combination of grammatical, semantic and morphological categories to express this message. This approach echoes Slobin's *thinking for speaking* hypothesis, which suggests that we are trained to think by our native language in certain patterns. These 'thinking patterns' are aligned to the grammatical and lexical structures that find their surface realization in various forms of the language (Slobin, 1997). For the speaker of Russian any noun has a gender assignment, which is an integral part of its lexical entry. The activation of gender information has to be a part of the processing routine, because without it being available, proper agreement of the noun with other parts of speech cannot be realized. As such, activation of grammatical gender is likely to be automatic and sustainable for at least short periods of time.

Conclusion

Overall, the study has provided clear evidence for the L1 grammatical gender bias effects in bilingual processing. We have shown that L1 Russian speakers operate in L2 English with mental representations that incorporate grammatical gender. While the exact mechanism of the transfer is yet to be fully understood, our study has succeeded in eliminating one of the potential explanations, such as a syntactic transfer route. The fact that we observed gender bias effects with animate referents only speaks against a global syntactic transfer, insensitive to semantics. On the contrary, gender bias in animate nouns points in the direction of the semantic transfer of gender information to L2, modulated by the animacy feature. In line with SAGH (Vigliocco et al., 2005), our data seem

https://doi.org/10.1017/S1366728916000766 Published online by Cambridge University Press

to suggest that the gender bias effects are a product of a coactivation of the linguistic gender information and conceptual information related to biological sex and the interaction between them. Only under those circumstances can grammatical gender be misinterpreted as semantic gender.

References

Alarcón, I. (2009). The processing of gender agreement in L1 and L2 Spanish: Evidence from reaction time data. *Hispania*, 814–828.

Alarcón, I. V. (2010). Gender assignment and agreement in L2 Spanish: The effects of morphological marking, animacy, and gender. *Studies in Hispanic and Lusophone Linguistics*, *3*(*2*), 267–300.

Andonova, E., D'Amico, S., Devescovi, A., & Bates, E. (2004). Gender and lexical access in Bulgarian. *Perception & psychophysics*, *66*(*3*), 496–507.

Antón-Méndez, M. I. (1999). *Gender and number agreement processing in Spanish*. Unpublished PhD dissertation, University of Arizona.

Aronoff, M. (1994). *Morphology by itself: Stems and inflectional classes* (No. 22). MIT Press.

Baayen, R. H., Davidson, D. J., & Bates, D. M. (2008). Mixed-effects modeling with crossed random effects for subjects and items. *Journal of memory and language*, *59*(*4*), 390–412.

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.

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.

Bassetti, B., & Nicoladis, E. (2015). Research on grammatical gender and thought in early and emergent bilinguals. *International Journal of Bilingualism*, 1367006915576824.

Bates, D., Maechler, M., & Bolker, B. (2012). lme4: Linear mixed-effects models using S4 classes.

Belacchi, C., & Cubelli, R. (2012). Implicit knowledge of grammatical gender in preschool children. *Journal of psycholinguistic research*, *41*(*4*), 295–310.

Bender, A., Beller, S., & Klauer, K. C. (2011). Grammatical gender in German: A case for linguistic relativity? *The Quarterly Journal of Experimental Psychology*, *64*(*9*), 1821–1835.

Boroditsky, L., & Schmidt, L. (2000). Sex, syntax and semantics. *Proceedings of the Twenty-Second Annual Meeting of the Cognitive Science Society.* Philadelphia, PA.

Boroditsky, L., Schmidt, L., & Phillips, W. (2003). Sex, Syntax, and Semantics. In Gentner & Goldin-Meadow (Eds.), *Language in Mind: Advances in the study of Language and Cognition*, 61–79.

Boutonnet, B., Athanasopoulos, P., & Thierry, G. (2012). Unconscious effects of grammatical gender during object categorisation. *Brain research*, *1479*, 72–79.

Brysbaert, M. (1998). Word recognition in bilinguals: Evidence against the existence of two separate lexicons. *Psychologica Belgica*, *38* (*3-4*), 163–175.

Clahsen, H., & Felser, C. (2006). Continuity and shallow structures in language processing. *Applied Psycholinguistics*, *27*(*01*), 107–126.

Clark, H. H. (1973). The language-as-fixed-effect fallacy: A critique of language statistics in psychological research. *Journal of verbal learning and verbal behavior*, *12*(*4*), 335–359.

Comrie, B. (1999). Grammatical gender systems: a linguist's assessment. *Journal of Psycholinguistic Research*, *28*(*5*), 457–466.

Corbett, G. G., & Fraser, N. M. (2000). Gender assignment: a typology and a model. *Systems of nominal classification*, *4*, 293–325.

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

Corbett, G. G. (1982). Gender in Russian: An account of gender specification and its relationship to declension. *Russian Linguistics*, *6*, 197–232.

Cornips, L., & Hulk, A. (2008). Factors of success and failure in the acquisition of grammatical gender in Dutch. *Second Language Research*, *24*(*3*), 267–295.

Cornips, L., van der Hoek, M., & Verwer, R. (2006). The acquisition of grammatical gender in bilingual child acquisition of Dutch (by older Moroccan and Turkish children). *Linguistics in the Netherlands*, *40*, 5.

Coughlin, C. E., & Tremblay, A. (2013). Proficiency and working memory based explanations for nonnative speakers' sensitivity to agreement in sentence processing. *Applied Psycholinguistics*, *34*(*03*), 615–646.

Cubelli, R., Lotto, L., Paolieri, D., Girelli, M., & Job, R. (2005). Grammatical gender is selected in bare noun production: Evidence from the picture–word interference paradigm. *Journal of Memory and Language*, *53*(*1*), 42–59.

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.

Desrochers, A., & Brabant, M. (1995). Interaction entre facteurs phonologiques et sémantiques dans une épreuve de catégorisation lexicale. *Revue Canadienne de Psychologie Expérimentale, 49*, 240–262.

Deutsch, A., & Bentin, S. (2001). Syntactic and semantic factors in processing gender agreement in Hebrew: Evidence from ERPs and eye movements. *Journal of Memory and Language*, *45*(*2*), 200–224.

Deutsch, A., Bentin, S., & Katz, L. (1999). Semantic influence on processing gender agreement: Evidence from Hebrew. *Journal of Psycholinguistic Research*, *28*(*5*), 515–535.

Dijkstra, T., & Van Heuven, W. J. (1998). The BIA model and bilingual word recognition. *Localist connectionist approaches to human cognition*, 189–225.

Dijkstra, T., Van Heuven, W. J., & Grainger, J. (1998). Simulating cross-language competition with the bilingual interactive activation model. *Psychologica Belgica*, *38*(*3-4*), 177–196.

Dijkstra, T., & Van Heuven, W. J. (2002). The architecture of the bilingual word recognition system: From identification

to decision. *Bilingualism: Language and Cognition*, *5*(*03*), 175–197.

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

Dixon, P. (2008). Models of accuracy in repeated-measures designs. *Journal of Memory and Language*, *59*(*4*), 447–456.

Dussias, P. E., Kroff, J. R. V., Tamargo, R. E. G., & Gerfen, C. (2013). When gender and looking go hand in hand. *Studies in Second Language Acquisition*, *35*(*02*), 353–387.

Eberhard, K. M., Scheutz, M., & Heilman, M. (2005). An empirical and computational test of linguistic relativity. In *Proceedings of the 27th annual conference of the Cognitive Science Society* (pp. 618–623).

Faraway, J. J. (2005). *Extending the linear model with R: generalized linear, mixed effects and nonparametric regression models*. CRC press.

Forster, K. I., & Forster, J. C. (2003). DMDX: A Windows display program with millisecond accuracy. *Behavior Research Methods, Instruments, & Computers*, *35*(*1*), 116–124.

Foucart, A., & Frenck-Mestre, C. (2012). Can late L2 learners acquire new grammatical features? Evidence from ERPs and eye-tracking. *Journal of Memory and Language*, *66*(*1*), 226–248.

Friederici, A. D., Rueschemeyer, S. A., Hahne, A., & Fiebach, C. J. (2003). The role of left inferior frontal and superior temporal cortex in sentence comprehension: localizing syntactic and semantic processes. *Cerebral cortex. 13*(*2*), 170–177.

Grainger, J., & Dijkstra, T. (1992). On the representation and use of language information in bilinguals. *Advances in psychology*, *83*, 207–220.

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.

Hagoort, P. (2003). Interplay between syntax and semantics during sentence comprehension: ERP effects of combining syntactic and semantic violations. *Cognitive Neuroscience, Journal of*, *15*(*6*), 883–899.

Jaeger, T. F. (2008). Categorical data analysis: Away from ANOVAs (transformation or not) and towards logit mixed models. *Journal of memory and language*, *59*(*4*), 434–446.

Jiang, N. (2004). Morphological insensitivity in second language processing. *Applied Psycholinguistics*, *25*, 603–634.

Jiang, N. (2007). Selective integration of linguistic knowledge in adult second language learning. *Language learning*, *57*(*1*), 1–33.

Jiang, N., Hu, G., Chrabaszcz, A., & Ye, L. (2015). The activation of grammaticalized meaning in L2 processing: Toward an explanation of the morphological congruency effect. *International Journal of Bilingualism*, 1367006915603823.

Juffs, A., & Harrington, M. (1995). Parsing effects in second language sentence processing. *Studies in second language acquisition*, *17*(*4*), 483–516.

Kazanina, N., Lau, E. F., Lieberman, M., Yoshida, M., & Phillips, C. (2007). The effect of syntactic constraints on the processing of backwards anaphora. *Journal of Memory and Language*, *56*(*3*), 384–409.

Keating, G. D. (2009). Sensitivity to Violations of Gender Agreement in Native and Nonnative Spanish: An Eye-Movement Investigation. *Language Learning*, *59*(*3*), 503–535.

Keating, G. D., & Jegerski, J. (2014). Experimental designs in sentence processing research. *Studies in Second Language Acquisition*, 1–32.

Konishi, T. (1993). The semantics of grammatical gender: A cross-cultural study. *Journal of psycholinguistic research*, *22*(*5*), 519–534.

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.

Kroll, J. F., Van Hell, J. G., Tokowicz, N., & Green, D. W. (2010). The Revised Hierarchical Model: A critical review and assessment. *Bilingualism: Language and Cognition*, *13*(*03*), 373–381.

Kurinski, E., & Sera, M. D. (2011). Does learning Spanish grammatical gender change English-speaking adults' categorization of inanimate objects? *Bilingualism: Language and Cognition*, *14*(*02*), 203–220.

Kuznetsova, A., Bruun Brockhoff, P., & Haubo Bojesen Christensen, R. (2013). lmerTest: Tests for Random and Fixed Effects for Linear Mixed Effect Modles (lmer Objects of lmer4 Package). R package version 2.0-0. Available online at: http://cran.r-prject.org/package=lmerTest

Lardiere, D. (2000). Mapping features to forms in second language acquisition. *Second language acquisition and linguistic theory*, 102–129.

Levelt, W.J.M. (1989). *Speaking: From intention to articulation*. Cambridge, MA: MIT Press.

Lupyan, G. (2012). Linguistically modulated perception and cognition: the label-feedback hypothesis. *Frontiers in psychology*, *3*(*54*).

Marcoux, D. R. (1973). Deviation in English gender. *American Speech*, *48*(*1/2*), 98–107.

Marian, V., Spivey, M., & Hirsch, J. (2003). Shared and separate systems in bilingual language processing: Converging evidence from eyetracking and brain imaging. *Brain and language*, *86*(*1*), 70–82.

Martinez, I. M., & Shatz, M. (1996). Linguistic influences on categorization in preschool children: A crosslinguistic study. *Journal of Child Language*, *23*(*3*), 529–545.

Morales, L., Paolieri, D., Cubelli, R., & Bajo, M. T. (2014). Transfer of Spanish grammatical gender to English: Evidence from immersed and non-immersed bilinguals. *Bilingualism: Language and Cognition*, 1–9.

Paolieri, D., Lotto, L., Leoncini, D., Cubelli, R., & Job, R. (2011). Differential effects of grammatical gender and gender inflection in bare noun production. *British Journal of Psychology*, *102*(*1*), 19–36.

Payne, J., & Huddleston, R. D. (2002). Noun and noun phrases. In R. D. Huddleston and G. K. Pullum, *Cambridge Grammar of the English Language*. Cambridge: Cambridge University Press.

Pearlmutter, N. J., Garnsey, S. M., & Bock, K. (1999). Agreement processes in sentence comprehension. *Journal of Memory and language*, *41*(*3*), 427–456.

Pérez-Pereira, M. (1991). The acquisition of gender: What Spanish children tell us. *Journal of Child Language*, *18*(*03*), 571–590.

Phillips, W., & Boroditsky, L. (2003). Can quirks of grammar affect the way you think? Grammatical gender and object concepts. Proceedings of the 25th Annual Meeting of the Cognitive Science Society. Boston, MA.

Pienemann, M. (1998). *Language processing and second language development: Processability theory* (Vol. 15). John Benjamins Publishing.

R Core Team (2012). R: A language and environment for statistical computing. Vienna: R foundation for Statistical Computing.

Rayner, K., & Duffy, S. A. (1986). Lexical complexity and fixation times in reading: Effects of word frequency, verb complexity, and lexical ambiguity. *Memory & Cognition*, *14*(*3*), 191–201.

Sabourin, L. L. (2003). *Grammatical gender and second language processing: An ERP study*. Unpublished PhD dissertation, Rijksuniversiteit Groningen.

Sabourin, L., & Stowe, L. A. (2008). Second language processing: when are first and second languages processed similarly? *Second Language Research*, *24*(*3*), 397–430.

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.

Sagarra, N., & Herschensohn, J. (2011). Proficiency and animacy effects on L2 gender agreement processes during comprehension. *Language Learning*, *61*(*1*), 80–116.

Sera, M. D., Elieff, C., Forbes, J., Burch, M. C., Rodriguez, 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.

Sera, M. D., Berge, C. A., & Pintado, J. D. C. (1994). Grammatical and conceptual forces in the attribution of gender by English and Spanish speakers. *Cognitive Development*, *9*(*3*), 261–292.

Sharkey, N. E., & Sharkey, A. J. (1987). What is the point of integration? The loci of knowledge-based facilitation in sentence processing. *Journal of Memory and Language*, *26*(*3*), 255–276.

Slobin, D. I. (1997). Mind, code, and text. In J. Bybee, J. Haiman & S. A. Thompson (eds.), *Essays on Language Function and Language Type*. Amsterdam: John Benjamins Publishing Company. 437–467

Skeide, M. A., Brauer, J., & Friederici, A. D. (2014). Syntax gradually segregates from semantics in the developing brain. *NeuroImage*, *100*, 106–111.

Sorace, A., & Filiaci, F. (2006). Anaphora resolution in near-native speakers of Italian. *Second Language Research*, *22*(*3*), 339–368.

Spinner, P., & Thomas, J. A. (2014). L2 learners' sensitivity to semantic and morphophonological information on Swahili nouns. *International Review of Applied Linguistics in Language Teaching*, *52*(*3*), 283–311.

Spivey, M. J., & Marian, V. (1999). Cross talk between native and second languages: Partial activation of an irrelevant lexicon. *Psychological science*, *10*(*3*), 281–284.

Tokowicz, N., & MacWhinney, B. (2005). Implicit and explicit measures of sensitivity to violations in second language grammar: An event-related potential investigation. *Studies in second language acquisition*, *27*(*2*), 173.

Truscott, J., & Sharwood-Smith, M. (2004). Acquisition by processing: A modular perspective on language development. *Bilingualism: Language and Cognition, 7*, 1–20.

Tyler, L. K., & Marslen-Wilson, W. D. (1977). The on-line effects of semantic context on syntactic processing. *Journal of Verbal Learning and Verbal Behavior*, *16*(*6*), 683–692.

Van Gompel, R. P., & Liversedge, S. P. (2003). The influence of morphological information on cataphoric pronoun assignment. *Journal of Experimental Psychology: Learning, Memory, and Cognition*, *29*(*1*), 128.

Vatz, K. L. (2009). *Grammatical gender representation and processing in advanced second language learners of French*. Unpublished PhD dissertation, University of Maryland.

Vigliocco, G., & Franck, J. (1999). When sex and syntax go hand in hand: Gender agreement in language production. *Journal of Memory and Language*, *40*(*4*), 455–478.

Vigliocco, G., Vinson, D. P., Arciuli, J., & Barber, H. (2008). The role of grammatical class on word recognition. *Brain and language*, *105*(*3*), 175–184.

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.

Vinson, D. P., & Vigliocco, G. (2002). A semantic analysis of grammatical class impairments: semantic representations of object nouns, action nouns and action verbs. *Journal of Neurolinguistics*, *15*(*3*), 317–351.

Von Studnitz, R. E., & Green, D. W. (2002). Interlingual homograph interference in German–English bilinguals: Its modulation and locus of control. *Bilingualism: Language and Cognition*, *5*(*01*), 1–23.

Whorf, B. L. (1956). Language, thought, and reality: Selected writings of Benjamin Lee Whorf. In J. B. Carroll (Ed.). New York: John Wiley & Sons.