Studies in Second Language Acquisition, 2018, **40**, 171–199.
doi:10.1017/S0272263116000437

LEXICAL AND SYNTACTIC CONGRUENCY IN L2 PREDICTIVE GENDER PROCESSING

Holger Hopp
Technische Universität Braunschweig

Natalia Lemmerth
Universität Mannheim

This article investigates how lexical and syntactic differences in L1 and L2 grammatical gender affect L2 predictive gender processing. In a visual-world eye-tracking experiment, 24 L1 Russian adult learners and 15 native speakers of German were tested. Both Russian and German have three gender classes. Yet, they differ in lexical congruency, that is, whether a noun ("house") is assigned to the same or a different gender class. Further, gender is syntactically realized on postnominal suffixes in Russian but on prenominal articles in German. For adjectives, both Russian and German mark gender on suffixes. In predictive gender processing, we find interactions of proficiency and congruency for gender-marked articles. Advanced L2 learners show nativelike gender prediction throughout. High-intermediate learners display asymmetries according to syntactic and lexical congruency. Predictive gender processing obtains for all nouns in the (syntactically congruent) adjective condition, yet only for lexically congruent nouns in the (syntactically incongruent) article condition.

We would like to thank the audiences at ISBPAC 2016 in Kaiserslautern and BUCLD 41 for helpful comments. Further, we are grateful to three anonymous reviewers and the editors of SSLA for providing constructive criticism which significantly helped us in improving the paper. All errors remain our responsibility.

Correspondence concerning this article should be addressed to Holger Hopp, English Linguistics, Technische Universität Braunschweig, Bienroder Weg 80, 38106 Braunschweig, Germany. E-mail: h.hopp@tu-bs.de

© Cambridge University Press 2016

https://doi.org/10.1017/S0272263116000437 Published online by Cambridge University Press

One of the central experiences of adult second language (L2) learners is that the first language (L1) interferes in speaking or comprehending the L2. On the one hand, the L1 can facilitate L2 acquisition and L2 processing, especially when L1 and L2 share a grammatical property. On the other hand, if the L1 differs from the L2 in how it realizes a phenomenon or if the L1 does not have an L2 property at all, the L1 can exert a negative effect on L2 acquisition and L2 processing. L1 effects in L2 morphosyntax have been central components of models of L2 acquisition (e.g., MacWhinney, 2008; Schwartz & Sprouse, 1996). In contrast, effects of the L1 morphosyntax have been more controversial in research on the real-time processing of the L2. Some models attribute a prominent role to L1 effects (e.g., the Full Transfer/Full Access/Full Parse Hypothesis; Dekydtspotter, Schwartz, & Sprouse, 2006), and yet other models suggest that there is no transfer in real-time processing (e.g., the Shallow Structure Hypothesis; Clahsen & Felser, 2006). In this article, we contribute novel data and present more fine-grained effects of L1 interference at the lexical and the syntactic level in L2 processing.

Much of the evidence of L1 effects in L2 processing comes from studies on grammatical gender. Grammatical gender, that is, the classification of nouns into different gender classes (Corbett, 1991), presents persistent problems to adult second language learners up to and including the highest proficiency levels (Franceschina, 2005). Learners have difficulty with gender assignment, that is, the target lexical classification of words into gender categories (e.g., "casa" [house] → feminine) and with gender agreement, that is, the systematic morphosyntactic realization of gender marking across elements inside a noun phrase (e.g., "l**a**$_{\text{FEM}}$ cas**a**$_{\text{FEM}}$ blanc**a**$_{\text{FEM}}$" [the house white]). The extent of these problems appears to be largely dependent on whether the L1 also has grammatical gender (for overview, see Franceschina, 2005; Grüter, Lew-Williams, & Fernald, 2012).

L1 effects on gender processing have been shown in neurophysiological ERP studies on gender agreement violations (e.g., "la cas**a**$_{\text{FEM}}$ blanc**o**$_{\text{MASC}}$"), in which late L2 learners whose L1 does not have grammatical gender are much less sensitive to gender violations than L2 speakers with L1s that do realize gender (e.g., Sabourin & Stowe, 2008). Similar findings have been obtained in visual-world eye tracking, where gender acts as a cue on the article for predicting the following noun ("**la**$_{\text{FEM}}$ → casa$_{\text{FEM}}$"): L2 learners whose L1s have gender are generally better able to use gender as a predictive cue, whereas L2 learners with L1s that do not instantiate gender show no or much attenuated predictive use of gender (e.g., Dussias, Valdés Kroff, Guzzardo Tamargo, & Gerfen, 2013; Grüter et al., 2012; Hopp, 2013; Lew-Williams & Fernald, 2010). Predictive processing, that is, the preactivation of upcoming linguistic input before it is encountered in the input (Kuperberg & Jaeger, 2016), has been identified as a particular area of difficulty in adult L2 processing (e.g., Martin et al., 2013). Some approaches hold that prediction ability is generally

https://doi.org/10.1017/S0272263116000437 Published online by Cambridge University Press

attenuated in an L2 (e.g., Grüter & Rohde's "RAGE" hypothesis; Grüter, Rohde, & Schafer, 2016) or specifically compromised in the processing of syntactic dependencies (e.g., Kaan, Dallas, & Wijnen, 2010; for discussion, see Kaan, 2014; Phillips & Ehrenhofer, 2015). In this context, we study whether and under which conditions the L1 can help overcome the particular problems with predictive processing of grammatical gender in an L2.

Current models of L1 transfer or L1 interference in L2 gender processing conceptualize L1 effects in gender processing as the recruitment of L1 properties in L2 comprehension and production. Formal models of L1 transfer hold that grammatical features of the L1, for example an abstract gender feature, transfer at the initial state of adult L2 acquisition and thus facilitate the acquisition and processing of L2 gender (e.g., Full Transfer/Full Access/Full Parse Model; Dekydtspotter et al., 2006). In cases in which the L1 does not realize grammatical gender, adult L2 learners may show persistent difficulties in L2 gender acquisition (e.g., Representational Deficit Hypothesis; Hawkins & Casillas, 2008; Tsimpli & Dimitrakopoulou, 2007), or they need to embark on a stepwise process of (re-)assembling abstract gender features in the course of L2 learning (e.g., Feature Reassembly Hypothesis; Lardiere, 2009). In functional models of L1 transfer, for example the Competition Model, speakers of L1s that possess grammatical gender transfer gender and can readily use it as a cue in processing the L2 (e.g., MacWhinney, 2008). Overall, these models conceive of L1 effects of grammatical gender largely as all-or-nothing effects in that the existence of grammatical gender in the L1 determines the degree of targetlikeness of gender processing in an L2. However, more recent studies on L2 gender processing suggest that L1 effects may be more varied than previously thought.

For instance, even adult L2 speakers whose L1s do not have grammatical gender (e.g., English) are able to acquire targetlike gender processing in the L2 (e.g., Alemán Bañón, Fiorentino, & Gabriele, 2014; Dussias et al., 2013; Foucart & Frenck-Mestre, 2011; Gillon-Dowens, Vergara, Barber, & Carreiras, 2010; Hopp, 2013, 2016; Tokowicz & MacWhinney, 2005). At the same time, L2 learners whose L1s do have grammatical gender do not invariably come to show targetlike L2 gender processing in spite of having high levels of L2 proficiency (e.g., Meulman, Stowe, Sprenger, Bresser & Schmid, 2014; Loerts, 2012; Sabourin, & Stowe, 2008). These findings underscore that the scope and the ways in which the L1 can facilitate or hinder grammatical gender processing in an L2 need to be refined.

In the present study, we therefore focus on adult L2 learners whose L1 has grammatical gender, and we specifically test whether lexical and syntactic differences in how the L1 and L2 realize grammatical gender affect L2 processing of gender agreement. We investigate Russian as the L1 and German as the L2. Although both Russian and German have grammatical gender and both categorize nouns into one of three gender classes (masculine, feminine, neuter), they differ in lexical congruency,

https://doi.org/10.1017/S0272263116000437 Published online by Cambridge University Press

that is, whether a particular noun (e.g., "house") is assigned to the same or a different gender class across languages. Further, they differ syntactically in that gender is expressed on postnominal suffixes in Russian but on prenominal determiners in German. These cross-linguistic differences allow us to test the relative impacts of L1-L2 lexical and syntactic contrasts on L2 gender processing and give insights into the levels at which the L1 is recruited in L2 grammatical processing.

LEXICAL AND SYNTACTIC CONGRUENCY IN GENDER MARKING

Even if both the L1 and the L2 have grammatical gender, they can differ in the amount of overlap in how they realize grammatical gender at the lexical and at the morphosyntactic level. In this section, we discuss effects of L1-L2 overlap, first, in terms of lexical congruency and, second, in terms of syntactic congruency in gender realization.

In terms of lexical congruency, L1 and L2 gender systems can overlap (a) in the number of gender distinctions they make and (b) in the classification of individual nouns into gender classes. For (a), Klassen (2016) reports that Spanish-German bilinguals show different responses in lexical decision and naming to the third type of gender of German, neuter, vis-à-vis masculine and feminine that have counterparts in the Spanish lexicon. Hence, the different ways in which L1 and L2 divide the lexicon across gender classes affect L2 gender processing. For (b), the classification of nouns to individual gender classes, studies have found robust effects of gender congruency. In picture naming, L2 speakers are faster in naming pictures if the translation equivalent of the L2 word in the L1 belongs to the same gender class (e.g., fem-fem; Bordag & Pechmann, 2007; Paolieri et al., 2010a). Similarly, bilinguals translate gender-congruent words faster than incongruent words (Paolieri et al., 2010a; Salamoura & Williams, 2007), and they benefit from gender congruency in accuracy in lexical decision tasks (e.g., Lemhöfer, Spalek, & Schriefers, 2008). These findings indicate that lexical gender representations are shared or accessed across languages.

For gender agreement processing, Weber and Paris (2004) report that French-German L2 speakers activate the lexical gender of the L1 in L2 gender processing. In a visual-world eye-tracking experiment, participants saw two objects that shared phonological onset and gender in L2 German (e.g., "Perle$_{FEM}$/Perücke$_{FEM}$" [pearl/wig] or "Kassette$_{FEM}$/Kanone$_{FEM}$" [tape/cannon]). In spoken word recognition, phonological overlap reliably leads to competition, in that listeners make initial looks to both objects (e.g., Allopenna, Magnuson, & Tanenhaus, 1998). In addition, the labels of the objects shared phonological onset in L1 French ("perle/perruque," "cassette/canon"), yet either matched the German word in gender ("perruque$_{FEM}$") or did not ("canon$_{MASC}$"). Even though

https://doi.org/10.1017/S0272263116000437 Published online by Cambridge University Press

the experiment was in L2 German only, participants showed less competition from the phonological competitor if the gender was incongruent with French than if it overlapped in gender. Similarly, Italian-Spanish L2 speakers fixate target objects in a visual-world task sooner for congruent than incongruent nouns (Morales, Paolieri, Dussias, Valdés Kroff, & Gerfen, 2015). Moreover, L2 speakers access gender representations of the L1 translation equivalents in L2 gender agreement processing. For instance, German-Dutch learners were found to outperform Romance-Dutch speakers in an ERP experiment on gender agreement (Sabourin & Stowe, 2008). Although both German and Romance languages realize grammatical gender, German and Dutch overlap to a large degree in gender assignment of nouns (see also Lemhöfer, Schriefers, & Indefrey, 2014), while Romance languages and Dutch do less so (Sabourin & Stowe, 2008). In conjunction, these studies suggest that lexical overlap in gender assignment facilitates gender agreement processing in an L2.

Yet other studies find that congruency does not play a role in grammatical gender processing in Polish-Dutch bilinguals (Loerts, 2012). These differences may partially reflect different degrees of transparency in gender marking. Whereas both Italian and Spanish have phonologically highly transparent gender markers for the vast majority of words ("-a" for feminine, "-o" for masculine), gender marking in Dutch or German is relatively opaque. In consequence, effects of gender congruency may be enhanced if phonological congruency or transparency additionally supports across-language correspondences.

Further, syntactic congruency, that is, overlap in the morphosyntax of L1 and L2 gender realization, may affect the degree of targetlike L2 gender processing. Foucart and Frenck-Mestre (2011) tested L1 English and L1 German advanced learners of French on gender agreement violations in determiner-noun (DN), determiner-adjective-noun (DAN) orders, and determiner-noun-adjective (DNA) orders. Although the L1 German group demonstrated nativelike ERP responses to gender violations in the DN orders, they were not sensitive to gender agreement errors in the DAN and the DNA orders. In this respect, they mirrored the performance of the English group whose L1 does not have gender. Foucart and Frenck-Mestre argue that the null finding for the German group on DNA orders may be related to the fact that German syntax does not license DNA orders, as it does not have postnominal adjectives, and does not mark grammatical gender on adjectives in predicative constructions in which the adjective appears postnominally ("The book is red."). Syntactic congruency in terms of the positions in which L1 and L2 realize inflectional markings may thus modulate the degree to which the availability of gender in the L1 facilitates L2 gender processing. Similarly, Tokowicz and MacWhinney (2005) report that L1 English adult learners of Spanish respond to tense violations and gender violations in DN sequences with

https://doi.org/10.1017/S0272263116000437 Published online by Cambridge University Press

a targetlike P600 ERP component, yet they do not show sensitivity to DN number agreement violations. The authors relate the lack of L2 sensitivity to nominal number agreement to the fact that English does not realize number agreement between definite determiners and nouns. In this study, the mismatch in syntactic congruency for number agreement between English and Spanish resulted in less targetlike performance than the processing of gender agreement, which is unique to Spanish (yet, see Alemán Bañón et al., 2014). Differences in the syntactic ways in which inflectional markings are realized thus appear to affect L2 processing.

In sum, previous research has found that both lexical and syntactic congruency between L1 and L2 in gender realization affect L2 processing. So far, each of these aspects has been studied independently, and it is an open question as to whether and, if yes, how lexical and syntactic congruency interact in L2 gender processing. In the present study, we therefore investigate how lexical and syntactic congruency may have additive or interactive effects in L2 processing. We study gender processing in L2 German by L1 adult speakers of Russian.

GENDER IN GERMAN

German has a three-way gender distinction (masculine, feminine, neuter), and gender marking is realized through gender agreement on determiners and adjectives within complex noun phrases. Nouns are opaque with respect to gender assignment. Although there are some phonological and semantic generalizations (e.g., Köpcke & Zubin, 1996), exceptions abound and the generalizations are rarely the subject of instruction in foreign language teaching. For instance, even the most general rule that trochaic bisyllabic words ending in schwa are feminine (e.g., "Hose" [trousers], "Blume" [flower]) has many (high-frequency) exceptions (e.g., "Hase$_{\text{MASC}}$" [hare], "Auge$_{\text{NEUT}}$" [eye]). Given the lack of robust correspondences between morphophonological form and gender class, learners need to store the gender for each noun individually as part of its lemma entry.

Critically, the surface realization of German gender is characterized by a large degree of syncretism because determiners and prenominal attributive adjectives also mark definiteness (definite vs. indefinite), case (nominative, accusative, dative, genitive), and number (singular vs. plural). Table 1 charts the determiner and adjectival forms for indefinite and definite articles in nominative case with singular nouns.

Of importance for the present study, Table 1 shows that definite articles clearly disambiguate between all the genders, while adjectives are three-way ambiguous for gender in definite noun phrases. In contrast, for indefinites, articles are ambiguous between masculine and neuter, while adjectival endings disambiguate between masculine and neuter.

https://doi.org/10.1017/S0272263116000437 Published online by Cambridge University Press

Table 1. German article forms and adjective endings in nominative case and singular

			Article	Adjective ending
definite	singular	masculine	der	-e
		feminine	die	-e
		neuter	das	-e
indefinite	singular	masculine	ein	-er
		feminine	eine	-e
		neuter	ein	-es

GENDER IN RUSSIAN

Russian also has three genders (masculine, feminine, neuter) that are marked on nouns and attributive adjectives as suffixes (Corbett, 1991). Russian does not have articles, and, for example, definiteness is contextually determined. Russian nouns and adjectives also mark case (nominative, genitive, dative, accusative, instrumental, locative, partitive, prepositional, vocative) and number (singular vs. plural).[1] Gender forms differ across declensional classes or have variants depending on phonological properties of nouns and/or adjectives; yet, in each class, the three genders are clearly differentiated (Cubberley, 2002, p. 110). Russian gender markers are partially phonologically transparent as per noun endings (e.g., "-o" is mostly neuter, "-a" marks feminine, e.g., Corbett, 1991, p. 62). Table 2 lists Russian gender forms for nouns and adjectives in nominative case in different declensional classes.

In sum, German and Russian share the same number and types of grammatical genders. In both languages, most nouns are either masculine or feminine, and fewer nouns are classified as neuter. In addition, adjectives mark gender through gender agreement on suffixes in both languages. German and Russian differ in that German marks gender on prenominal articles and determiners, while Russian marks gender on nominal suffixes. Importantly, gender assignment in both German and Russian is nontransparent, and the two languages are typologically unrelated.

Table 2. Russian gender suffixes on nouns and adjective endings in nominative case and singular. Variants in each cell denote different declensional classes

			Noun	Adjective ending
(in-)definite	singular	masculine	-a, Ø, -b	-ый
		feminine	-a, -b	-ая
		neuter	-o, -Я	-ое

https://doi.org/10.1017/S0272263116000437 Published online by Cambridge University Press

RESEARCH QUESTIONS AND HYPOTHESES

Against the backdrop of the cross-linguistic differences and previous research findings summarized in the preceding sections this study investigates whether L1 Russian learners use German gender marking predictively in on-line comprehension. We test predictive processing of grammatical gender being marked either on articles or on adjectives. Although Russian marks grammatical gender, it does not have articles. However, Russian marks gender on prenominal attributive adjectives, and Russian native speakers use adjectival gender marking for predicting the upcoming noun (Sekerina, 2015). The syntactic contexts in which gender is marked have been shown to affect the degree of L1 facilitation in L2 gender processing (Foucart & Frenck-Mestre, 2011; Tokowicz & MacWhinney, 2005). Therefore, the present study considers predictive gender processing on articles and on adjectives. We hypothesize that L2 gender processing is moderated by syntactic gender congruency in that gender-marked articles lead to less prediction than gender-marked adjectives (Hypothesis 1).

Further, lexical gender congruency affects predictive gender processing for (typologically related) languages with (partially) transparent gender marking like French, Italian, and Spanish (Morales et al., 2015; Weber & Paris, 2004). We hypothesize that lexical gender congruency extends to a typologically distinct language pairing and to languages with opaque gender assignment like Russian and German in that gender prediction is magnified for lexically congruent versus incongruent nouns (Hypothesis 2).

Finally, previous studies by Hopp (2013) for L1 English learners of German show that general proficiency and target lexical knowledge of gender assignment moderate predictive gender processing. Similarly, Dussias et al. (2013) report that rising proficiency leads to more predictive gender processing in Italian and English learners of Spanish. Therefore, the present study includes effects of proficiency on predictive gender processing. We hypothesize that more highly proficient L2 learners demonstrate more targetlike predictive gender processing, that is, lower effects of syntactic or lexical congruency with the L1 (Hypothesis 3).

THE STUDY

Participants

Twenty-four L1 Russian adult learners of German (mean age: 26.4 years, sd = 5.8 years; 21 females) and 15 native speakers of German (mean age: 27.3 years, sd = 6.1 years; 10 females) were recruited for the experiment. All L2 speakers of German had started learning German after the age of

https://doi.org/10.1017/S0272263116000437 Published online by Cambridge University Press

10 (mean: 14.8 years, sd = 5.38) and were residents in Germany at the time of testing. None of the German native speakers had any knowledge of Russian. Table 3 gives an overview of the participants.

The L2 participants completed two proficiency measures: First, they filled in a standardized 30-item written placement test (Goethe Institut, 2010) focusing on lexico-grammatical knowledge. On average, they scored 24.9 points (range: 19–29), which placed them into the upper-intermediate to advanced range. Second, they completed the unspeeded lexical decision task LexTALE (Lemhöfer & Broersma, 2012): on average, they scored 72% (range: 51–89%). The scores of the two tests were strongly correlated (r = .755). In consequence, we computed one composite proficiency score consisting of the mean of the proportional scores from each test.

Materials and Methods

For the linguistic stimuli, we crossed the factors Condition (article vs. adjective), Congruency (congruent vs. incongruent), and Type (different gender vs. same gender). In total, 100 pictureable nouns were selected from medium to high frequency bands, 50 having congruent genders in Russian and German, and 50 having incongruent genders. All nouns had a corpus frequency of at least 1,843 tokens in the 25-billion-word COSMAS II corpus of contemporary written German (COSMAS II, 2008; date accessed 21 March 2015). A one-way ANOVA with the factor Congruency showed that there was no statistical difference in corpus frequency between congruent and incongruent nouns ($F(1,98) = 1.083$, $p = .299$). None of the items was a cognate word, so that lexical overlap across languages would be restricted to gender. In each set of 50 items, 30 bore gender marking on the article as in (1a), 10 for each gender (i.e., masculine, feminine, neuter). In (1a), which we refer to as the "article condition," only the articles showed gender marking because attributive adjectives with definite articles are ambiguous between all genders.

Table 3. Participant information

	L1 German (n = 15)		L2 German (n = 24)	
	Mean (sd)	Range	Mean (sd)	Range
Age	27.3 (6.1)	19–38	26.4 (5.8)	18–38
Proficiency Score (out of 100)	91.1 (5.1)	80–98	77.6 (10.3)	57–93
Length of Exposure (years)			11.9 (7.5)	2–29
Length of Residence (years)			7.5 (7.5)	0.5–23

https://doi.org/10.1017/S0272263116000437 Published online by Cambridge University Press

Twenty items had gender marking on the adjective as in (1b), that is, the adjective condition. In (1b), only the adjective bore gender marking because indefinite articles are ambiguous between masculine and neuter. In contrast, the article for indefinite feminine nouns carries overt gender marking ("eine") and would disambiguate before the adjective. In consequence, we only used masculine and neuter nouns in the adjective condition. Of the 20 items in the adjective condition, 10 were masculine and 10 were neuter. Of the 10 nouns for each gender, half (i.e., 5) were assigned to the different gender condition and the other half to the same gender condition. Table 4 illustrates the design of the materials.

(1) a. Wo ist der/die/das gelbe [N]? (article condition)
Where is the$_{MASC/FEM/NEUT}$ yellow [N]?
b. Wo ist ein kleiner/s gelber/s [N]? (adjective condition)
Where is a small$_{MASC/NEUT}$ yellow [N]

In the adjective condition, two adjectives were used to make sure that the gender information at the end of the first adjective did not occur right before the onset of the noun. The first adjective was always a size adjective ("klein" [small] or "groß" [large]), and the second adjective was a color adjective.[2]

For the visual displays, pictures of the 100 nouns and 145 additional pictureable objects were selected. The pictures were pretested on a group of five native German and three native Russian speakers to check that they were clearly recognizable and were named consistently. Five objects were replaced, and the new objects were retested for naming consistency. All 245 objects were then colored in red, green, yellow, and blue using open-source picture editors.

Sixty displays were created as follows for the article condition. For the different gender trials, one object with a clearly identifiable color was the target; there were two identically colored competitor objects that bore one of the other two genders each, and the fourth object in the display was a differently colored distracter of a gender different from the target. Hence, gender on the article would unambiguously refer to

Table 4. Design of the materials by condition. Number of items in parentheses

Condition	Type	Congruency	Gender
article (60)	different (30)	congruent (15)	masculine (10)
	same (30)	incongruent (15)	feminine (10)
			neuter (10)
adjective (40)	different (20)	congruent (10)	masculine (10)
	same (20)	incongruent (10)	neuter (10)

https://doi.org/10.1017/S0272263116000437 Published online by Cambridge University Press

the target object. For the same gender trials, displays were designed along the same lines (Figure 1); however, the three color-matched objects all had the same gender, and the differently colored distractor had a different gender. All objects were similar in size and placed in identically sized quadrants on the display.

For the adjective condition, 40 displays were created that contained four identically colored objects, three of the same size and one distractor of a different size (Figure 2). For the different gender condition, only one object had the target gender. For the same gender trials, the three same-sized objects all had the same gender and the differently sized object bore a different gender.

In this way, the article and adjective conditions presented the same amount of ambiguity: Three objects were identical in color and size in each display. In the different gender (difference) trials, gender marking on the article (1a) or the adjective (1b) would disambiguate to one of the three objects; in the same trials, neither the articles nor the adjectives would disambiguate to any of the three objects, and only the noun disambiguated reference. In each display, nouns with phonologically different onsets were used to avoid temporary competition between nouns (e.g., Allopenna et al., 1998). In the same gender trials, then, noun onset would lead to clear disambiguation.

Given the limited availability of pictureable objects, some of the 245 objects had to be reused. It was made sure that objects would only be

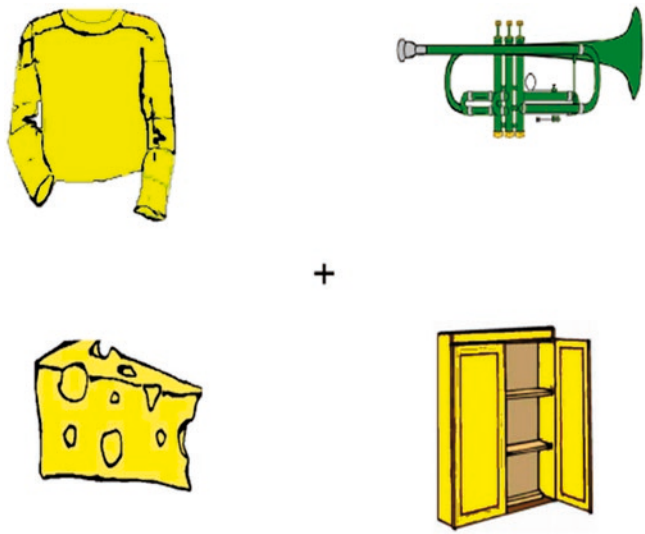


Figure 1. Display for determiner condition: same condition (all identically-coloured objects have same gender; differently-coloured distractor has different gender).

https://doi.org/10.1017/S0272263116000437 Published online by Cambridge University Press

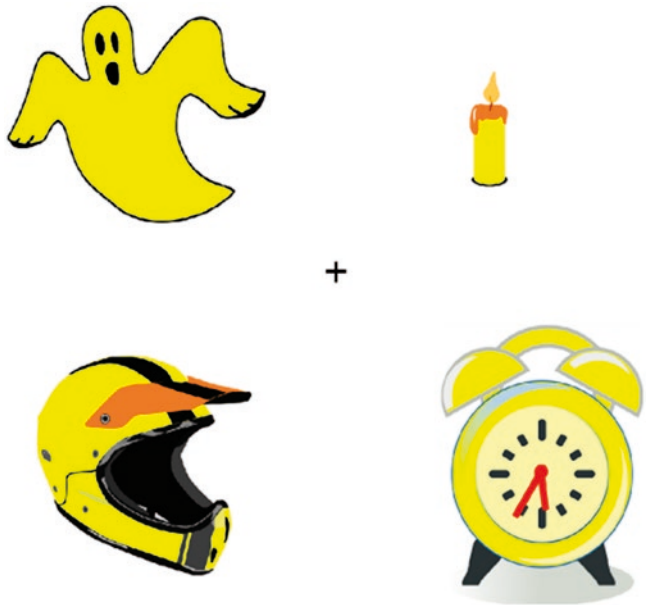


Figure 2. Display for adjective condition: difference condition (all identically-sized objects have different gender; differently-sized distractor has different gender compared to other objects).

targets once: If they were targets in one condition (difference or same), they would be reused as distractors in the other condition. Hence, no participant heard the same label twice as a target. We created three lists that rotated target gender (masculine, feminine, and neuter), object color, and object position.

Procedure

All sentences were recorded by a female native speaker of standard German at a slow-to-moderate pace with neutral intonation. The mean length of the article was 255 ms (sd = 22.0), 717 ms (sd = 134.9) for the adjectives and 642 ms (sd = 159.9) for the nouns. For the article condition, the onset of the disambiguating noun was set 1200 ms after the onset of the article. In the adjective condition, noun onset was set at 1150 ms after the offset of the first adjective. The duration of the final segment on the adjective (/r/ vs. /s/) that carried gender marking was around 50 ms. Hence, the duration between the provision of gender information and the onset of the disambiguating noun was comparable across the article or the adjective conditions, that is, 1200 ms.

https://doi.org/10.1017/S0272263116000437 Published online by Cambridge University Press

The procedure was identical to the experiments in Hopp (2013, 2016) in that the experiment combined a production and a comprehension task. Participants were tested individually in a lab where they sat in front of a 19-inch screen at a distance of approximately 70 cm. Participants were instructed, and they completed four practice trials before a calibration began. The calibration aimed for visual acuity below 0.5 degrees and was repeated during the experiment if necessary. Throughout the experiment, an SMI RED eye tracker recorded gaze position at 60 Hz.

Production Task. In the production task that immediately preceded the comprehension task, participants saw still images of the four-picture displays. They were asked to name the four objects in each display, including their color (see Figures 1 and 2). This way, we could (a) establish whether participants recognized the objects and knew their labels, and (b) assess which gender L2 participants assigned to the labels of the objects. In their descriptions of the objects, the participants used noun phrases containing articles, color adjectives, and nouns. Participants' responses were audio-recorded and transcribed. Responses were coded for gender on the basis of the gender form used on the article ("der$_{\text{MASC}}$," "die$_{\text{FEM}}$," "das$_{\text{NEUT}}$") for definite NPs as in example (1a) or, if the article used was indefinite, the adjective (e.g., "gelb-er," "gelb-e," "gelb-es" – yellow$_{\text{MASC/FEM/NEUT}}$). Overall accuracy was coded as percentage of target responses.

Comprehension Task. After the production part, the screen changed immediately to show the identical displays including a fixation cross. A sound signal alerted participants to fixate on the cross. The participants' gaze was directed to the fixation cross prior to each trial to avoid baseline effects of participants already looking at the target item before the onset of the critical noun phrase (Barr, Gann, & Pierce, 2011). The auditory presentation of the sentences began 800 ms after the sound signal. In total, the experiment took about 35 minutes for the natives and between 40 and 50 minutes for the L2 participants. Participants received five euros as compensation.

Analysis

In the comprehension task, we probed whether participants use gender marking on the article to predict upcoming nouns. We analyzed the reaction time, that is, the first fixation on the target picture after article onset or adjective offset in the adjective condition. Gaze position after article onset was coded every 20 ms for 3,000 ms. All trials in which the

https://doi.org/10.1017/S0272263116000437 Published online by Cambridge University Press

first fixation to the target object occurred within 200 ms after article onset or adjective offset were excluded from analyses because such early fixations could not possibly have been guided by linguistic information (e.g., Huettig, Rommers, & Meyer, 2011). In all, this affected 4.8% of the trials.

For the remaining trials, we computed reaction times for the subset of trials based on the production accuracy of gender assignment by the L2ers, that is, the subjective gender assignment of the participants. Following Hopp (2013) and Lemhöfer et al. (2014), we thus made sure that we only analyzed (a) trials in which the genders assigned by the participants in previous production could potentially act as an informative cue for the subsequent noun (difference trials) or (b) trials in which the gender marking would not single out one referent as per the participant's gender assignment (same trials).[3] Specifically, we classified the comprehension trials of the L2ers as (a) difference trials, if the L2er had assigned the target lexical gender to the target item and other (target) genders to the other items. Hence, these were trials in which gender marking on articles or adjectives would uniquely identify the target item according to the learner's subjective gender assignment in the production task. Correspondingly, trials were defined as (b) same trials, if the participant used the target gender in labeling the target object in the same trials and used the same gender for at least one other object in the display. In these trials, gender on the article or the adjective did not uniquely identify the target, and participants needed to wait to hear the onset of the nouns to disambiguate reference. For all the experimental trials in the article condition, the difference trials amounted to 478 out of 674 difference trials overall (70.9%; 67.1% for incongruent and 77.8% for congruent trials). The same trials were 552 out of 643 same trials overall (86.1%; 55.6% for incongruent and 73.6% for congruent trials). For the adjective condition, the difference trials encompassed 288 of 440 trials (65.5%; 59.8 for incongruent and 71.0% for congruent trials), and the same trials were 361 out of 434 trials (83.2%; 79.1% for incongruent and 87.1% for congruent trials). All other trials were treated as uninformative trials and excluded from the analysis.

Results

Table 5 lists the noun naming accuracy and gender assignment accuracy for the 400 objects used in the experiment. The L1 participants named all objects with the target genders. In total, the L2 participants did not accurately name 2.4% of the nouns. A mixed logistic regression for Noun Labeling Accuracy with Gender Congruency as fixed factors and participants and items as crossed random factors showed that

https://doi.org/10.1017/S0272263116000437 Published online by Cambridge University Press

Table 5. Accuracy of noun labeling and gender assignment in L2 group (n = 24)

	All nouns	Congruent nouns	Incongruent nouns
Noun labels	97.6% (sd = 3.4; range: 85–100%)	97.4% (sd = 3.6; range: 85–100%)	97.7% (sd = 3.3; range: 85–100%)
Gender assignment	87.6% (sd = 9.2; range: 63–100%)	91.7% (sd = 7.3; range: 70–100%)	83.5% (sd = 11.8; range: 56–100%)

there was no difference in noun naming accuracy between congruent (97.4%) and incongruent (97.7%) nouns (β = 0.097, SE = 0.149, t = 0.656, p = .512). The nouns that had incorrectly been named were excluded from the calculation of correct gender assignment. Overall, rate of target gender assignment correlated significantly with proficiency (r = .846).

Table 5 shows that the L2 participants were better at producing the target genders for congruent (91.7%) than for incongruent (83.5%) nouns. A mixed logistic regression for Gender Assignment with Congruency as a fixed factor and subject and item as crossed random factors showed a significant main effect of Congruency (β = 0.832, SE = 0.078, t = 10.658, p < .001).

We now turn to the comprehension results. Figures 3a and 3b graph the mean reaction times for the L1, and Figures 4a and 4b show the results for the L2 group in the respective conditions across congruent and incongruent conditions. All reaction times and standard deviations are listed in the Appendix.

For analyzing the reaction times, we used linear mixed regression analyses with subjects and items as crossed random factors. Analyses were conducted in IBM SPSS 23. Predictor variables were contrast coded, and the reaction times and proficiency scores were centered around their group means. In the first step, we defined Type (difference vs. same), Group (L1 vs. L2), Congruency (congruent vs. incongruent), Condition (article vs. adjective), and Proficiency (continuous variable) as fixed factors. Following current practice in psycholinguistics (Barr, Levy, Scheepers, & Tily, 2013), we initially constructed a maximal model that included random intercepts and slopes for all fixed effects and their interactions. When this model failed to converge, we gradually simplified the random effects structure by first removing interaction terms and then main effects. We report the results corresponding to the model with the maximal random effects structure that converged.

In the omnibus analysis, we found a highly significant main effect of Type (β = -210.59, SE = 23.28, t = -9.047, p < .001), which was qualified by an interaction of Type and Group (β = 144.55, SE = 42.23, t = 3.423, p = .001). Further, the effect of Type was qualified by an interaction of Type, Condition, and Proficiency (β = -14.21, SE = 6.12, t = -2.323, p = .020) and

https://doi.org/10.1017/S0272263116000437 Published online by Cambridge University Press

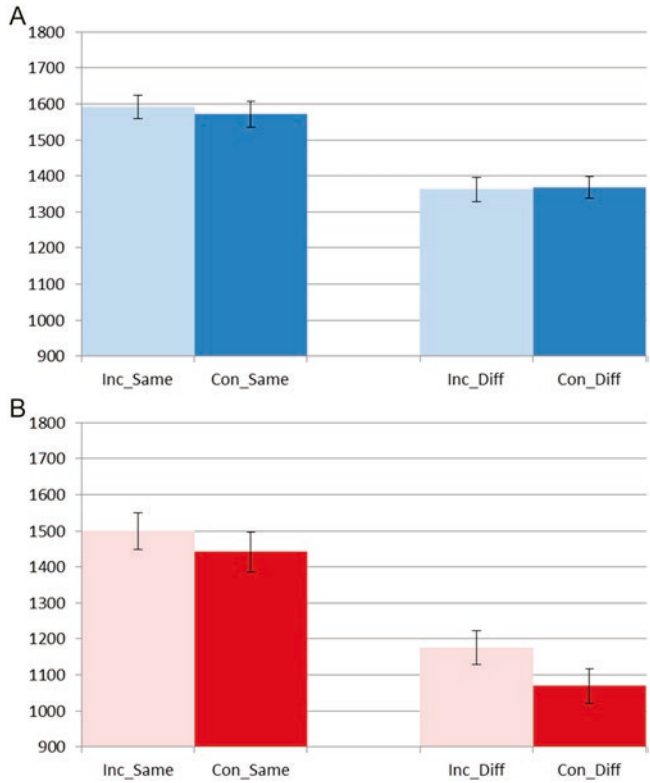


Figure 3. (a) Mean reaction times (in ms after article onset) for article condition in L1 group (n = 15). (b) Mean reaction times (in ms after adjective offset) for adjective condition in L1 group (n = 15). Inc = Incongruent, Con = Congruent, Same = Same Condition, Diff = Difference Condition. Error bars show standard error of the mean.

further (marginally) by Condition, Congruency, Group, and Proficiency (β = 23.59, SE = 12.21, t = 1.932, p = .053) as well as Type, Condition, Congruency, and Group (β = -252.60, SE = 148.95, t = 1.696, p = .090).[4] In light of the interactions with Group and Condition, we next analyze the findings separately by group to interpret the interactions.

For the L1 group, there was a main effect of Type (β = -281.30, SE = 40.44, t = -6.956, p < .001) and an interaction of Type and Condition (β = 126.18, SE = 58.45, t = 2.159, p = .032). There were no significant main effects of Congruency or any interactions with it. Given the interactions of Type and Condition, we calculated separate comparisons for articles and adjectives. For the article condition, we found a highly significant effect of Type (β = -219.39, SE = 33.96, t = -6.460, p < .001), with the gender effect, that is, the mean difference between difference and same trials, being 215 ms. The factor Type also became significant for adjectives (gender

https://doi.org/10.1017/S0272263116000437 Published online by Cambridge University Press

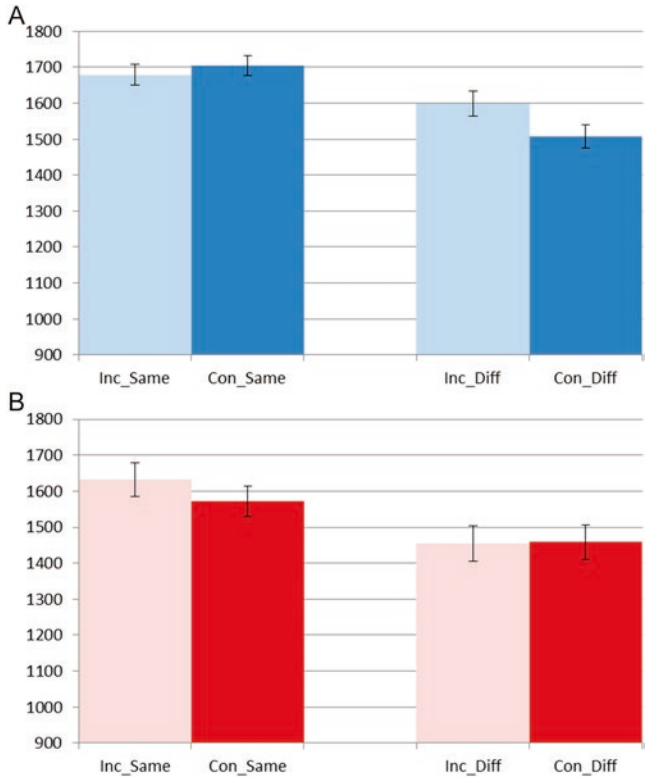


Figure 4. (a) Mean reaction times (in ms after article onset) for article condition in L2 group (n = 24). (b) Mean reaction times (in ms after adjective offset) for adjective condition in L2 group (n = 24). Inc = Incongruent, Con = Congruent, Same = Same Condition, Diff = Difference Condition. Error bars show standard error of the mean.

effect: 349 ms; β = -345.05, SE = 51.27, t = -6.731, p < .001). There were no significant interactions with Congruency (all ts < 1). Hence, native speakers use gender marking predictively on both articles and adjectives, with the effect being larger for adjectives than for articles.

For the L2 group, there was no main effect of Type (β = -59.05, SE = 84.62, t = -0.689, p = .485) and the interaction of Type and Condition also failed to reach significance (β = 198.79, SE = 173.63, t = 1.145, p = .260). Given our hypotheses about differences between articles and adjectives, and to ensure comparability with the analysis of the natives, we calculated separate analyses by Condition even in the absence of a significant interaction term. For the article condition, the L2 group showed no main effect of Type (β = 42.87, SE = 95.13, t = 0.451, p = .652), yet we observed a (marginally) significant interaction of Type and Proficiency (β = -113.08, SE = 57.54, t = 1.960, p = .050). For the adjective condition,

https://doi.org/10.1017/S0272263116000437 Published online by Cambridge University Press

there was a significant effect of Type (β = -146.61, SE = 43.32, t = -3.385, p = .001) and no significant further effects or interactions (all ts < 1.2). These findings show that the L2 group used gender on adjectives, which are syntactically congruent between German and Russian, predictively, irrespective of lexical gender congruency or proficiency level. In contrast, predictive use of gender on articles, which are syntactically incongruent between German and Russian, was affected by the proficiency level of the L2 participants.

To illustrate the effects of proficiency further, we performed a median split of the L2 group according to the proficiency score and computed separate analyses for each group. For information, Table 6 lists the mean gender assignment accuracy in the production task for the two proficiency groups.

As Table 6 shows, the advanced proficiency group had more target-like gender assignment than the high-intermediate group. A mixed logistic regression for Gender Accuracy yielded a main effect of Proficiency (β = 1.206, SE = 0.360, t = 3.356, p = .001) and Congruency (β = 0.712, SE = 0.128, t = 5.554, p < .001), yet no interaction between them (t < 1).

The mean reaction times for the high-intermediate proficiency group are shown in Figures 5a and 5b, and the mean reaction times for the advanced group are given in Figures 6a and 6b.

For the high-intermediate group, the model for articles showed no effect of Type (gender effect: 79 ms; β = -67.71, SE = 46.52, t = -1.456, p = .168) or Congruency (β = -49.94, SE = 45.0, t = -1.110, p = .268), yet a marginally significant interaction of the two (β = -169.25, SE = 90.14, t = 1.878, p = .061). Pairwise comparisons indicated that there was a significant effect of Type for the congruent items (gender effect: 173 ms; β = -160.34, SE = 77.66, t = -2.065, p = .041), yet not for the incongruent items (gender effect: -21 ms; β = 11.29, SE = 68.58, t = 0.165, p = .869). For adjectives, we found a significant effect of Type (gender effect: 163 ms; β = -152.23, SE = 70.29, t = -2.166, p = .031), yet no effects of or interactions with Congruency (all ts < 1). These results show that, at high-intermediate proficiency level, lexical congruency affects predictive gender processing in that gender prediction is restricted to lexically congruent nouns in syntactically incongruent (article) contexts.

Table 6. Accuracy of gender assignment by proficiency group (n = 12 each)

	All nouns	Congruent nouns	Incongruent nouns
High-Intermediate Proficiency	83.4% (sd =9.7; range: 63–95%)	87.9% (sd = 8.3; range: 70–98%)	76.7% (sd = 12.1; range: 56–92%)
Advanced Proficiency	92.3% (sd = 4.8; range: 84–100%)	95.4% (sd = 3.4; range: 89–100%)	90.3% (sd = 6.5; range: 78–100%)

https://doi.org/10.1017/S0272263116000437 Published online by Cambridge University Press

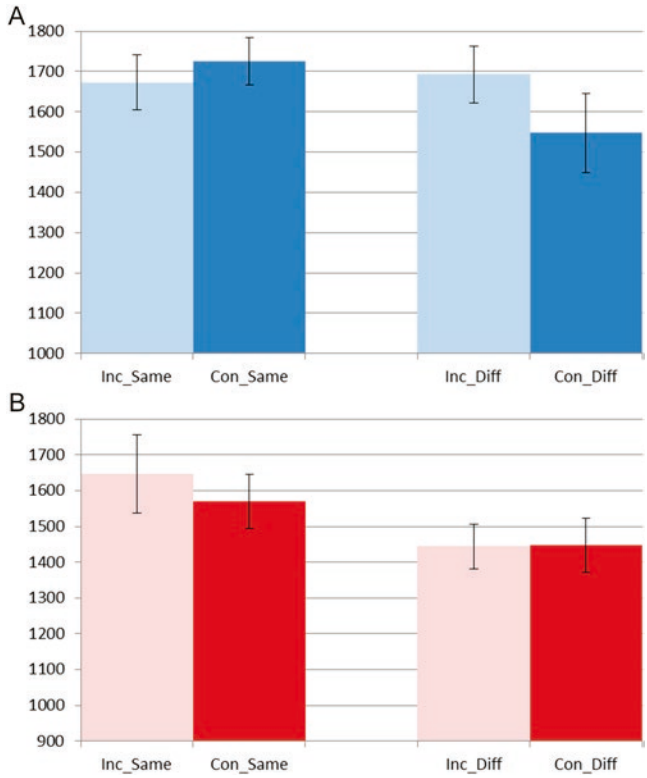


Figure 5. (a) Mean reaction times (in ms after article onset) for article condition in high-intermediate-proficiency L2 group (n = 12). (b) Mean reaction times (in ms after adjective offset) for adjective condition in high-intermediate-proficiency L2 group (n = 12). Inc = Incongruent, Con = Congruent, Same = Same Condition, Diff = Difference Condition. Error bars show standard error of the mean.

For the advanced group, the analysis for articles returned a significant effect of Type (gender effect: 188 ms; β = -184.61, SE = 37.40, t = -4.936, $p < .001$), yet no further effects or interactions (all ts < 1). Similarly, for adjectives, the only effect that reached significance was Type (gender effect: 142 ms; β = -141.92, SE = 57.46, t = -2.407, $p = .030$), and no interactions neared significance (all ts < 1). At higher proficiency levels, then, predictive gender processing obtains across syntactic and lexical congruency.

DISCUSSION

The present experiment addressed the extent to which particular syntactic and lexical properties of grammatical gender in the L1 affect L2

https://doi.org/10.1017/S0272263116000437 Published online by Cambridge University Press

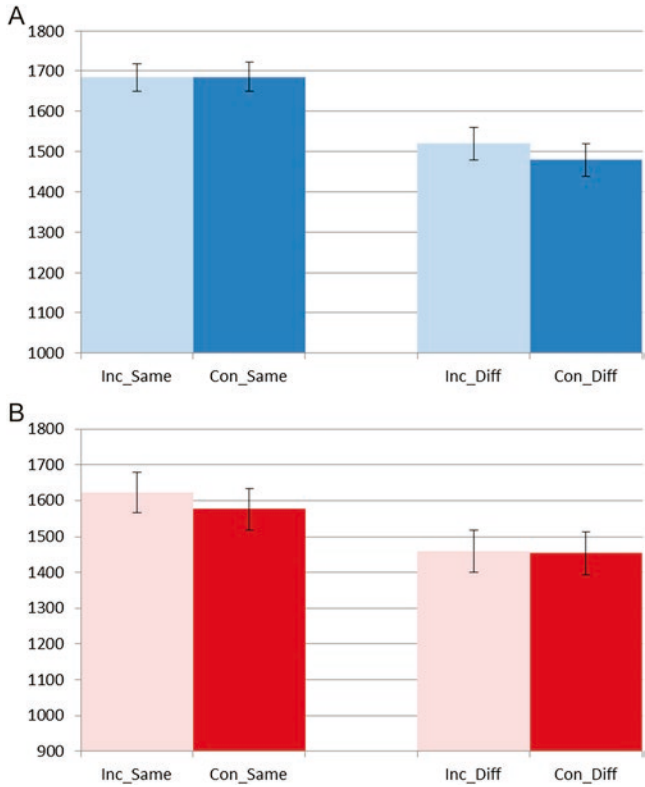


Figure 6. (a) Mean reaction times (in ms after article onset) for article condition in advanced L2 group (n = 12). (b) Mean reaction times (in ms after adjective offset) for adjective condition in advanced L2 group (n = 12). Inc = Incongruent, Con = Congruent, Same = Same Condition, Diff = Difference Condition. Error bars show standard error of the mean.

gender processing. In a visual-world experiment on predictive gender processing in German, we found that L1 Russian adult learners as a group showed predictive gender agreement processing in the adjective condition, albeit at lower levels than German natives (Figures 4a and 4b). At the same time, effects of proficiency and congruency indicated that both learner factors and linguistic factors moderate L2 predictive gender processing.

First, the separate analyses according to proficiency demonstrated that more advanced Russian learners of German come to show target L2 predictive gender processing irrespective of lexical congruency or syntactic overlap between L1 and L2 gender marking. For both gender marking on articles and on adjectives, the advanced learners exhibited predictive gender processing (Figures 6a and 6b). Confirming Hypothesis 3, these proficiency effects corroborate previous research on L1 English learners of Spanish (Dussias et al., 2013) and German (Hopp, 2013).

https://doi.org/10.1017/S0272263116000437 Published online by Cambridge University Press

Further, the present findings complement those of previous studies by documenting that fully targetlike prediction according to gender agreement also holds for highly proficient adult L2 learners of an opaque TL gender system like German whose L1 is typologically distinct and employs partially different morphosyntactic means of gender marking.

Second, beyond learner characteristics like proficiency, the present study also finds that linguistic properties of the L1 affect predictive L2 gender processing at lower proficiency levels. The high-intermediate proficiency group showed gender prediction for gender marking on adjectives, yet not for marking on articles (Figures 5a and 5b). In line with Hypothesis 1, this discrepancy suggests that identity between L1 and L2 in how gender is morphosyntactically realized, that is, on adjective endings, facilitates predictive L2 gender processing, while syntactically incongruent realizations of gender (articles vs. noun suffixes) attenuate L2 gender prediction (see also Foucart & Frenck-Mestre, 2011; Tokowicz & MacWhinney, 2005).

In this regard, the present findings contrast with most previous studies that compare different types of gender agreement. Typically, adjective-noun agreement is associated with greater difficulty than article-noun agreement in adult L2 processing (e.g., Foucart & Frenck-Mestre, 2011). Such an asymmetry likely reflects the facts (a) that article-noun agreement is more frequent in the input than adjective-noun agreement, (b) that learners typically learn grammatical gender by memorizing sequences of definite articles and nouns (e.g., "die Blume" – the_{FEM} $\text{flower}_{\text{FEM}}$), and (c) that article-noun agreement is restricted to a closed class of articles while adjectival agreement extends across lexical adjectives.[5] Against this backdrop, the data from the high-intermediate proficiency learners in this study suggest that syntactic congruency of L1 gender marking determines L2 gender processing to a greater extent than quantitative properties of the L2 input.

At present, it remains an open question whether it is the lack of articles in Russian or the difference in the morphosyntax of gender between Russian and German that imposes burdens on L2 predictive processing, or whether it is the nonpredictive aspect of the postnominal realization of gender in Russian that attenuates prediction according to gender marking on articles in the L2 German of L1 Russian learners. Because gender on postnominal suffixes cannot be used predictively in Russian, the lack of the predictive function of nominal case marking may impede the use of L2 gender for prediction. Future studies will need to test whether congruency in morphosyntax per se or the possibility of L1 inflectional marking of acting as a predictive cue in similar domains affects L2 predictive processing. In this respect, it is worth noting that L2 English learners of German at comparable intermediate proficiency levels to the high-intermediate Russian learners in this study were also found not to use gender (Hopp, 2013, 2016) or case (Hopp, 2015) on

https://doi.org/10.1017/S0272263116000437 Published online by Cambridge University Press

German articles to predict upcoming referents. Because English and German both have articles, it cannot solely be the lack of articles in Russian that accounts for the difficulties in using L2 German articles for gender prediction. Instead, it seems likely that facilitatory effects of the L1 hail both from congruent syntactic options as well as from analogous realization of inflectional morphology in the L1. These findings on gender prediction suggest that problems with prediction are not general features of L2 syntactic processing, as claimed by some proposals (e.g., Grüter et al., 2016; Kaan et al., 2010); rather, these problems are restricted to contexts in which the L1 does not afford predictive processing (see also Foucart, Martin, Moreno, & Costa, 2014).

Third, on top of syntactic congruency effects, the intermediate group only used gender predictively in particular lexical contexts. The high-intermediate proficiency group only showed predictive gender processing for lexically congruent nouns in cases in which L1 and L2 realize gender in morphosyntactically different ways, that is, on prenominal articles in German and on postnominal suffixes in Russian (Figure 5a). In contrast, gender marking on adjectives was used predictively irrespective of whether the nouns were lexically congruent in their gender assignment between Russian and German. Hence, Hypothesis 2 is partially confirmed for the high-intermediate L2 group in the present study. Finding such differences complements effects of lexical congruency in Italian-Spanish (Morales et al., 2015) and French-German learners (Weber & Paris, 2004). It thus bolsters approaches stipulating that lexical gender nodes are interconnected across languages (e.g., Paolieri et al., 2010b; Salamoura & Williams, 2007). Such effects further surface for typologically distinct languages with opaque gender systems. In this regard, they strongly suggest that cross-linguistic activation of gender in syntactic processing proceeds on the basis of shared lexical gender nodes rather than form-based phonological congruency (Italian-Spanish; Morales et al., 2015) or correspondences in the gender classification of individual lexical items (Dutch-German; Sabourin & Stowe, 2008).

The present study thus finds that both syntactic congruency and lexical congruency facilitate L2 prediction according to gender at intermediate proficiency levels. Critically, these factors interact in that L2 predictive gender processing obtains only for lexically congruent nouns in the syntactically incongruent (article) condition. In contrast, there is no effect of lexical congruency in the syntactically congruent (adjective) condition. Overall, the present study thus indicates that L1 and L2 representations are highly interactive in L2 gender processing at both the morphosyntactic and the lexical level at intermediate proficiency levels (see also Hopp, 2010; Klassen, in press; Sabourin & Stowe, 2008).

We now consider these findings in the context of bilingual processing and acquisition approaches. Upon hearing gender-marked articles or adjectives in comprehension, listeners need to activate the gender node

https://doi.org/10.1017/S0272263116000437 Published online by Cambridge University Press

associated with the gender marking on the article or the adjective. Subsequently, they construct a syntactic agreement relation between the gender marker and the noun labels of the objects in the display. In formal linguistic and psycholinguistic models of nominal concord, these agreement computations are construed as feature checking or feature valuation between head nouns that bear an interpretable gender feature and articles or adjectives that have uninterpretable gender features (e.g., Carstens, 2000; for psycholinguistic models, see, e.g., Franck, Vigliocco, Antón-Méndez, Collina, & Frauenfelder, 2008).

Nonselective models of bilingual lexical access (e.g., Dijkstra & Van Heuven, 2002) assume an integrated (e.g., Klassen, 2016; Salamoura & Williams, 2007) or an interconnected (Paolieri et al., 2010b) gender system. Because the task in the present study requires participants to select the label of the objects in the display, the objects automatically activate both the L2 and the L1 lexical representations including their gender nodes. With lexically congruent nouns, the gender is associated with the same object in both languages and the gender form on the article or adjective relates to a unique referent through agreement in the difference trials. In contrast, the gender activates different objects for incongruent nouns according to German or Russian gender, respectively, so that the agreement relation computed on the gender form is not unique or weaker compared to lexically congruent trials (see also Morales et al., 2015). As proficiency rises, L2 learners appear to become better at inhibiting the activation of or competition from the L1 gender, as shown in the production accuracy of congruent versus incongruent nouns (Table 6) and the lack of congruency effects in comprehension in the advanced group (Figure 6).

In addition, Russian learners of German need to acquire articles and configure the feature assembly of uninterpretable features including gender, so that gender agreement between articles and nouns can be computed. Many production studies testify to the frequent omission of articles by adult Russian learners of English (e.g., Schönenberger, 2014) or the nontarget use of articles according to the features realized on them, for example definiteness and specificity (e.g., Ionin, Ko, & Wexler, 2004). The present study finds that these problems spill over into comprehension and gender marking in that articles are used less reliably for predictive gender agreement at intermediate proficiency levels. Critically, lexical and syntactic congruency interact in that the joint activation of the same referent according to gender in both Russian and German facilitates predictive processing of gender-marked articles because a unique agreement relation to the target referent can be computed.

These interactions of lexical and syntactic L1 effects cannot easily be accommodated in existing model of cross-linguistic influence in grammatical gender in adult L2 acquisition and L2 processing. Formal models that encode grammatical gender in the presence of a single feature

https://doi.org/10.1017/S0272263116000437 Published online by Cambridge University Press

(e.g., [+gen]) can account neither for effects of lexical congruency nor for effects of syntactic congruency (e.g., Hawkins & Casillas, 2008; Schwartz & Sprouse, 1996; Tsimpli & Dimitrakopoulou, 2007). The Feature Reassembly Model (Lardiere, 2009), which conceives of L1 transfer and restructuring as pertaining to sets of features on lexical and functional items, can potentially capture more fine-grained L1 effects in terms of lexical and syntactic congruency. It is yet to be worked out how exactly lexical and syntactic features are encoded and how they are reassembled as proficiency rises (for some discussion, see Lardiere, 2000). Within functional models, the notion of gender as a cue would need to be specified, so that the particular lexical and syntactic aspects of gender as a cue in the L1 can be delineated that affect L2 processing. At any rate, the present findings underline that L1 effects in the L2 processing of gender warrant a more fine-grained analysis. Rather than restricting the scope of cross-linguistic facilitation in L2 processing to the mutual realization of a grammatical feature per se in both L1 and L2, studies of cross-linguistic influence need to incorporate discussion of how differences in the realization of these features impact L2 processing.

Although the present data are cross-sectional in nature, they suggest a possible developmental trajectory in L2 predictive gender processing. At first, predictive gender processing may be specific to syntactically congruent contexts, that is, adjective-noun combinations. Later, gender prediction emerges in further syntactic contexts, yet it initially remains specific to lexically congruent nouns in article-noun contexts that facilitate gender processing by virtue of gender coactivation. At the most advanced proficiency levels, predictive gender processing generalizes across syntactic and lexical contexts in the L2. This pattern of development from L1-governed and lexically based processing to grammaticalized processing of gender seems to resonate with findings on the development of agreement processing in research using ERPs. Cross-sectional studies testing L2 learners of Spanish at different proficiency levels (e.g., Gabriele, Fiorentino, & Alemán Bañón, 2013) as well as longitudinal studies of subject-verb agreement in beginning to intermediate L2 learners of French (e.g., Osterhout, McLaughlin, Pitkanen, Frenck-Mestre, & Molinaro, 2006) or learners of an artificial language with grammatical gender (Morgan-Short, Sanz, Steinhauer, & Ullman, 2010) report developmental changes in the L2 processing of inflection (see also Steinhauer, White, & Drury, 2009). Learners initially respond to (gender) agreement violations by showing an N400 component, which is typically taken to index lexical-semantic processing (e.g., Kutas & Hillyard, 1983). At higher proficiency levels, L2 learners respond to agreement violations by evincing a P600 (McLaughlin et al., 2010), which is commonly associated with syntactic and more abstract processing (e.g., Osterhout & Holcomb, 1992). Although individual differences

https://doi.org/10.1017/S0272263116000437 Published online by Cambridge University Press

between learners may cloud the developmental pattern from more lexically to more structure-based processing (e.g., Tanner, McLaughlin, Herschensohn, & Osterhout, 2013), advances in L2 proficiency and exposure appear to lead to the grammaticalization of L2 morphosyntactic information in L2 processing. The present behavioral findings are broadly compatible with evidence from neurophysiological studies suggesting that learners go through a developmental trajectory from more lexically based to syntactic processing of gender agreement. In future research, it will be important to explore how data from visual-world eye tracking can usefully complement neurophysiological data in mapping L2 development.

Needless to say, the conclusions of the present study are constrained by the relatively small sample size and the fact that only one language combination was studied. To generalize how lexical and syntactic overlap between the L1 and the L2 affects L2 processing, future research needs to test larger sets of learners across different L1-L2 combinations that systematically differ in lexical and syntactic congruency. Work along these lines will delineate the relative impacts of L1 properties at different levels on L2 processing.

In sum, the study finds that both lexical and syntactic properties of the L1 interactively affect predictive L2 gender processing. These findings underscore that cross-linguistic influence in L2 gender processing obtains at various lexical and syntactic levels, and they suggest that L2 development of gender proceeds from lexically based to grammar-based processing. It will be fruitful to explore such interactions further, so that future models of L1 influence in L2 sentence processing can capture the specifics of cross-linguistic influence in a comprehensive account of L2 processing.

Received 23 January 2016
Accepted 10 October 2016
Final Version Received 18 October 2016

NOTES

1. In addition, Russian marks dual number in some contexts. Further, gender forms are further distinguished by animacy in accusative case (see, e.g., Cubberley, 2002, for discussion). These further markings are not relevant in the context of the current study.

2. In the adjective condition, we used a second adjective to pad out the region between the first gender-marked adjective and the noun, so that the article and adjective conditions would be comparable. In consequence, gender marking surfaced twice before the onset of the noun. Given that the second instance occurred right before the onset of the noun and it did not provide any cue different from the first gender marker, we hold that its effects on prediction are negligible.

3. Two anonymous reviewers ask why gender assignment was not tested after the eye-tracking experiment and wonder how the production task may have affected the eye-tracking results. We did not test gender assignment after the experiment because

https://doi.org/10.1017/S0272263116000437 Published online by Cambridge University Press

participants listened to the target gender markings in the eye-tracking task. Hence, there would have been no way of telling whether the participants' gender assignment after the eye tracking corresponded to their gender representation before hearing the target gender assignment in the eye-tracking experiment. Instead, we classified participants' gender assignment according to the rationale that the gender recruited in the participant's production would be active during immediately subsequent comprehension. Future studies would need to test whether the predictive use of gender in comprehension patterns differently if there is no preceding production part.

4. Following the suggestion by an anonymous reviewer, we checked whether these effects were affected by the type of gender of the target item (masculine, feminine, neuter). There were no significant main effects of gender of the target item or any interactions with this factor.

5. An anonymous reviewer points out that "definite articles (not determiners in general!) form a special class in L2 learning, because they are often acquired as chunks together with the noun" and wonders how the L2 learners would deal with gender marking on other determiners, for example, demonstratives. This is an empirical question worth addressing.

REFERENCES

Alemán Bañón, J., Fiorentino, R., & Gabriele, A. (2014). Morphosyntactic processing in advanced second language (L2) learners: An event-related potential investigation of the effects of L1-L2 similarity and structural distance. *Second Language Research*, *30*, 275–306.

Allopenna, P. D., Magnuson, J. S., & Tanenhaus, M. K. (1998). Tracking the time course of spoken word recognition using eye movements: Evidence for continuous mapping models. *Journal of Memory and Language*, *38*, 419–439.

Barr, D. J., Gann, T. M., & Pierce, R. S. (2011). Anticipatory baseline effects and information integration in visual world studies. *Acta Psychologica*, *137*, 201–207.

Barr, D. J., Levy, R., Scheepers, C., & Tily, H. (2013). Random-effects structure for confirmatory hypothesis testing: Keep it maximal. *Journal of Memory and Language*, *68*, 255–278.

Bordag, D., & Pechmann, T. (2007). Factors influencing L2 gender processing. *Bilingualism: Language and Cognition*, *10*, 299–314.

Carstens, V. (2000). Concord in minimalist theory. *Linguistic Inquiry*, *31*, 319–355.

Clahsen, H., & Felser, C. (2006). Grammatical processing in language learners. *Applied Psycholinguistics*, *27*, 3–42.

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

COSMAS II (2008). Corpus Search, Management and Analysis System. Retrieved 20 March 2015 from http://www.ids-mannheim.de/cosmas2.

Cubberley, P. (2002). Russian: A linguistic introduction. Cambridge, UK: Cambridge University Press.

Dekydtspotter, L., Schwartz, B. D., & Sprouse, R. A. (2006). The comparative fallacy in L2 processing research. In M. Grantham O'Brien, C. Shea, & J. Archibald (Eds.), *Proceedings of the 8th generative approaches to second language acquisition conference (GASLA 2006): The Banff conference* (pp. 33–40). Somerville, MA: Cascadilla Press.

Dijkstra, T., & van Heuven, W. J. B. (2002). The architecture of the bilingual word recognition system: From identification to decision. *Bilingualism: Language and Cognition*, *5*, 175–197.

Dussias, P. E., Valdés Kroff, J. R., Guzzardo Tamargo, R. E., & Gerfen, C. (2013). When gender and looking go hand in hand: Grammatical gender processing in L2 Spanish. *Studies in Second Language Acquisition*, *35*, 353–387.

Foucart, A., & Frenck-Mestre, C. (2011). Grammatical gender processing in L2: Electrophysiological evidence of the effect of L1–L2 syntactic similarity. *Bilingualism: Language and Cognition*, *14*, 379–399.

Foucart, A., Martin, C. D., Moreno, E. M., & Costa, A. (2014). Can bilinguals see it coming? Word anticipation in L2 sentence reading. *Journal of Experimental Psychology: Learning, Memory, and Cognition*, *40*, 1461–1469.

Franceschina, F. (2005). Fossilized second language grammars: The acquisition of grammatical gender. Amsterdam, The Netherlands: John Benjamins.

https://doi.org/10.1017/S0272263116000437 Published online by Cambridge University Press

Franck, J., Vigliocco, G., Antón-Méndez, I., Collina, S., & Frauenfelder, U. (2008). The interplay of syntax and form in sentence production: A cross-linguistic study of form effects on agreement. *Language and Cognitive Processes*, *23*, 329–374.
Gabriele, A., Fiorentino, R., & Alemán Bañón, J. (2013). Examining second language development using event-related potentials: A cross-sectional study on the processing of gender and number agreement. *Linguistic Approaches to Bilingualism*, *3*, 213–232.
Gillon Dowens, M., Vergara, M., Barber, H. A., & Carreiras, M. (2010). Morphosyntactic processing in late second-language learners. *Journal of Cognitive Neuroscience*, *22*, 1870–1887.
Goethe Institut (2010). German placement test. Retrieved 21 July 2015 from http://www.goethe.de/cgi-bin/einstufungstest/einstufungstest.pl.
Grüter, T., Lew-Williams, C., & Fernald, A. (2012). Grammatical gender in L2: A production or a real-time processing problem? *Second Language Research*, *28*, 191–215.
Grüter, T., Rohde, H., & Schafer, A. J. (2016). Coreference and discourse coherence in L2: The roles of grammatical aspect and referential form. *Linguistic Approaches to Bilingualism*. doi: 10.1075/lab.15011.gru.
Hawkins, R., & Casillas, G. (2008). Explaining frequency of verb morphology in early L2 speech. *Lingua*, *118*, 595–612.
Hopp, H. (2010). Ultimate attainment in L2 inflectional morphology: Performance similarities between non-native and native speakers. *Lingua*, *120*, 901–931.
Hopp, H. (2013). Grammatical gender in adult L2 acquisition: Relations between lexical and syntactic variability. *Second Language Research*, *29*, 33–56.
Hopp, H. (2015). Semantics and morphosyntax in L2 predictive sentence processing. *International Review of Applied Linguistics*, *53*, 277–306.
Hopp, H. (2016). Learning (not) to predict: Grammatical gender agreement in non-native processing. *Second Language Research*, *32*, 277–307.
Huettig, F., Rommers, J., & Meyer, A. S. (2011). Using the visual world paradigm to study language processing: A review and critical evaluation. *Acta Psychologica*, *137*, 151–171.
Ionin, T., Ko, H., & Wexler, K. (2004). Article semantics in L2 acquisition: The role of specificity. *Language Acquisition*, *12*, 3–69.
Kaan, E. (2014). Predictive sentence processing in L2 and L1: What is different? *Linguistic Approaches to Bilingualism*, *4*, 257–282.
Kaan, E., Dallas, A. C., & Wijnen, F. (2010). Syntactic predictions in second-language sentence processing. In J.-W. Zwart & M. de Vries (Eds.), *Structure preserved. Festschrift in the honor of Jan Koster* (pp. 207–213). Amsterdam, The Netherlands: John Benjamins.
Klassen, R. (2016). The representation of asymmetric grammatical gender systems in the bilingual mental lexicon. *Probus, 28*(1), 9–28.
Köpcke, M., & Zubin, D. (1996). Prinzipien für die Genuszuweisung im Deutschen. In E. Lang & G. Zifonun (Eds.), *Deutsch typologisch. Jahrbuch des Instituts für Deutsche Sprache 1995* (pp. 473–491). Berlin: de Gruyter.
Kuperberg, G., & Jaeger, T. F. (2016). What do we mean by prediction in language comprehension? *Language, Cognition and Neuroscience*, *31*, 32–59.
Kutas, M., & Hillyard, S. A. (1983). Event-related brain potentials to grammatical errors and semantic anomalies. *Memory and Cognition*, *11*, 539–550.
Lardiere, D. (2000). Mapping features to forms in second language acquisition. In J. Archibald (Ed.), *Second language acquisition and linguistic theory* (pp. 102–129). Cambridge, MA: Blackwell.
Lardiere, D. (2009). Some thoughts on a contrastive analysis of features in second language acquisition. *Second Language Research*, *25*, 173–227.
Lemhöfer, K., & Broersma, M. (2012). Introducing LexTALE: A quick and valid Lexical Test for Advanced Learners of English. *Behavior Research Methods*, *44*, 325–343.
Lemhöfer, K., Schriefers, H., & Indefrey, P. (2014). Idiosyncratic grammars: Syntactic processing in second language comprehensions uses subjective feature representations. *Journal of Cognitive Neuroscience*, *26*, 1428–1444.
Lemhöfer, K., Spalek, K., & Schriefers, H. (2008). Cross-language effects of grammatical gender in bilingual word recognition and production. *Journal of Memory and Language*, *59*, 312–330.
Lew-Williams, C. & Fernald, A. (2010). Real-time processing of gender-marked es by native and non-native Spanish speakers. *Journal of Memory and Language*, *63*, 447–464.
Loerts, H. (2012). *Uncommon gender: Eyes and brains, native and second language learners and grammatical gender* (Doctoral dissertation). Rijksuniversiteit Groningen: Grodil Press.

https://doi.org/10.1017/S0272263116000437 Published online by Cambridge University Press

MacWhinney, B. (2008). A unified model. In P. Robinson & N. Ellis (Eds.), *Handbook of cognitive linguistics and second language acquisition* (pp. 341–372). Mahwah, NJ: Lawrence Erlbaum Associates.

Martin, C. D., Thierry, G., Kuipers, J-R., Boutonnet, B., Foucart, A., & Costa, A. (2013). Bilinguals reading in their second language do not predict upcoming words as native readers do. *Journal of Memory and Language*, *69*, 574–588.

McLaughlin, J., Tanner, D., Pitkänen, I., Frenck-Mestre, C., Inoue, K., & Valentine, G. (2010). Brain potentials reveal discrete stages of L2 grammatical learning. *Language Learning*, *60*, 123–150.

Meulman, N., Stowe, L. A., Sprenger, S. A., Bresser, M., & Schmid, M. S. (2014). An ERP study on L2 syntax processing: When do learners fail? *Frontiers in Psychology*, *5*, 1072. doi: 10.3389/fpsyg.2014.01072.

Morales, L., Paolieri, D., Dussias, P. E., Valdés Kroff, J. R., & Gerfen, C. (2015). The gender congruency effect during bilingual spoken-word recognition. *Bilingualism: Language and Cognition*, *19*, 294–310. doi: 10.1017/S1366728915000176.

Morgan-Short, K., Sanz, C., Steinhauer, K., & Ullman, M. T. (2010). Second language acquisition of gender agreement in explicit and implicit training conditions: An event-related potential study. *Language Learning*, *60*, 154–193.

Osterhout, L., & Holcomb, P. J. (1992). Event-related brain potentials elicited by syntactic anomaly. *Journal of Memory and Language*, *31*, 785–806.

Osterhout, L., McLaughlin, J., Pitkanen, I., Frenck-Mestre, C., & Molinaro, N. (2006). Novice learners, longitudinal designs, and event-related potentials: A paradigm for exploring the neurocognition of second-language processing. *Language Learning*, *56*, 199–230.

Paolieri, D., Cubelli, R., Macizo, P., Bajo, M. T., Lotto, L., & Job, R. (2010a). Grammatical gender processing in Italian and Spanish bilinguals. *The Quarterly Journal of Experimental Psychology*, *63*, 1631–1645.

Paolieri, D., Lotto, L., Morales, L., Bajo, M. T., Cubelli, R., & Job, R. (2010b). Grammatical gender processing in Romance languages: Evidence from bare noun production in Italian and Spanish. *European Journal of Cognitive Psychology*, *22*, 335–347. doi: 10.1080/09541440902916803.

Phillips, C., & Ehrenhofer, L. (2015). The role of language processing in language acquisition. *Linguistic Approaches to Bilingualism*, *5*, 409–453.

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

Salamoura, A., & Williams, J. N. (2007). The representation of grammatical gender in the bilingual lexicon: Evidence from Greek and German. *Bilingualism: Language and Cognition*, *10*, 257–275.

Schönenberger, M. (2014). Article use in L2 English by L1 Russian and L1 German speakers. *Zeitschrift für Sprachwissenschaft*, *33*, 77–105.

Schwartz, B. D., & Sprouse, R. A. (1996). L2 cognitive states and the fulltransfer/fullaccess model. *Second Language Research*, *12*, 40–72.

Sekerina, I. (2015). *Is grammatical agreement predictive in heritage language bilingualism? Paper presented at the workshop "Grammar in the Mental Lexicon,"* University of Potsdam, August 1.

Steinhauer, K., White, E. J., & Drury, J. E. (2009). Temporal dynamics of late second language acquisition: Evidence from event-related brain potentials. *Second Language Research*, *25*, 13–41.

Tanner, D., McLaughlin, J., Herschensohn, J., & Osterhout, L. (2013). Individual differences reveal stages of L2 grammatical acquisition: ERP evidence. *Bilingualism: Language and Cognition*, *16*, 367–382.

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*, 173–204.

Tsimpli, I. M., & Dimitrakopoulou, M. (2007). The Interpretability Hypothesis: Evidence from wh-interrogatives in second language acquisition. *Second Language Research*, *23*, 215–242.

Weber, A., & Paris, G. (2004). The origin of the linguistic gender effect in spoken-word recognition: Evidence from non-native listening. *Proceedings of the Twenty-Sixth Annual Meeting of the Cognitive Science Society*, 1446–1451.

https://doi.org/10.1017/S0272263116000437 Published online by Cambridge University Press

APPENDIX

Table A1. Mean reaction times by group and condition in ms (standard deviation)

			L1 (n = 15)	L2 (n = 24)	L2 High-Intermediate (n = 12)	L2 Advanced (n = 12)
articles	same	congruent	1571 (257)	1704 (184)	1725 (224)	1686 (144)
		incongruent	1591 (241)	1679 (213)	1673 (238)	1684 (194)
	different	congruent	1369 (220)	1507 (259)	1548 (334)	1479 (164)
		incongruent	1362 (231)	1599 (255)	1693 (273)	1520 (185)
adjectives	same	congruent	1442 (292)	1573 (246)	1570 (280)	1576 (220)
		incongruent	1499 (237)	1632 (486)	1647 (572)	1623 (379)
	different	congruent	1069 (478)	1459 (330)	1448 (374)	1454 (296)
		incongruent	1176 (256)	1454 (338)	1444 (256)	1460 (335)

https://doi.org/10.1017/S0272263116000437 Published online by Cambridge University Press