

Processing of gender and number agreement in late Spanish bilinguals

International Journal of Bilingualism
0(0) 1–21

© The Author(s) 2012

Reprints and permission:

sagepub.co.uk/journalsPermissions.nav

DOI: 10.1177/1367006912453810

ljb.sagepub.com



Nuria Sagarra

Rutgers University, USA

Julia Herschensohn

University of Washington, USA

Abstract

Drawing from Sagarra & Herschensohn (2010, 2011), we evaluate current approaches to grammatical representation and processing of L2 gender and number agreement in Spanish determiner phrases (DPs), some advocating incomplete acquisition and computation of L2 grammatical features and others claiming native-like representation and computation of new grammatical values. Testing Spanish monolinguals, beginning and intermediate L2 learners, we evaluate comprehension data with a moving window paradigm and a grammaticality judgment task to investigate whether late learners of ungendered L1s can gain native-like behavioral patterns of sensitivity to grammatical gender and number agreement violations. We also consider the additional factors of proficiency level (beginners vs. intermediates) and processing cost (animate vs. inanimate nouns; gender vs. number agreement). Moving window and grammaticality judgment data reveal that intermediates, but not beginners, show qualitatively similar reactions to monolinguals (gender and number concord/discord distinctions), confirming the importance of proficiency while suggesting native-like processing by L2 learners. They also show that both grammatical (gender, number) and semantic (animacy) features differentially impact concord processing for both natives and L2 learners, and that working memory plays a role in developing L2 processing skills.

Keywords

Adjective agreement, animacy, functional categories, gender/number, L2 Spanish, working memory

I Introduction

Universal Grammar (UG) approaches to second language acquisition (L2A) consider the representation of the learner's abstract grammar (interlanguage), often in comparison to the native

Corresponding author:

Nuria Sagarra, Rutgers University, 57 U.S. Highway 1 New Brunswick, NJ 08901, USA

Email: nuria.sagarra@gmail.com

speaker's grammar, whereas processing studies examine the manner in which the grammatical knowledge is put to practice in real-time computation of language use (comprehension/production). Assuming that grammatical representation and computational procedures develop in tandem, we evaluate current approaches to representation and processing that cover a range of perspectives. Some models of representation predict incomplete acquisition by adults of grammatical features absent in their L1, owing to a critical period effect (Franceschina, 2005; Hawkins & Franceschina, 2004; Tsimpli & Dimitrakopoulou, 2007), while certain approaches to computation claim adult learners' online representations lack structural depth and grammatical detail, with preference for shallow over deep processing (Clahsen & Felser, 2006) in complex syntactic constructions such as relative clauses. In contrast, other models argue that adult learners can acquire representation and computation of new grammatical values that may be qualitatively comparable to native grammars. They hold that production errors are the result of default inflection (Bruhn de Garavito & White, 2002; Leung, 2005; White, Valenzuela, Kozłowska-MacGregor & Leung, 2004), difficulties in mapping grammatical features to Phonetic Form (Lardiere, 2007), or L1 transfer and performance factors (Hopp, 2007; Frenck-Mestre, Foucart, Carrasco & Herschensohn, 2009). UG representational approaches assume that some grammatical properties are universal and others may vary (parametrically); likewise, processing studies have shown some universal tendencies and other language-specific preferences (cf. Clahsen & Felser, 2006; Sabourin & Stowe, 2008). Determining exactly what is to be learned and what might be given is a major task for scholars investigating these areas.

In this study, we investigate whether adult Anglophone learners of L2 Spanish can show native-like behavioral patterns of sensitivity to gender and number agreement violations, and whether linguistic factors such as animacy and/or individual differences in working memory affect L2 acquisition. In Spanish, adjectives and determiners agree with the interpretable gender and number features on the head noun, a concord that facilitates processing (faster reading times [RT]) in native speakers. In contrast, English has no nominal number or gender and no concord on adjectives, making it an ideal case to test the deficit and accessibility models. Because linguistic competence includes both knowledge of grammar and the ability to implement it in real-time processing (Foucart, 2008; Hopp, 2007; Jin, Áfarli, & van Dommelen, 2007; Juffs & Harrington, 1995, 1996; Marinis, Roberts, Felser, & Clahsen, 2005), we employed the online—moving window paradigm—and offline techniques—grammaticality judgments—to assess representation and computation. An anonymous reviewer notes that grammaticality judgments presuppose, and will be influenced by, participants' processing of the experimental items (and thus involve computation), while online reading times may be affected by the nature of the representations that readers compute. It is impossible to clearly dissociate the two empirically, as we acknowledge.

2 Syntactic representation of L2 gender and number agreement

Following a Minimalist framework (Chomsky, 1995, 2002) for Spanish, grammatical (*un*interpretable) features of determiners and adjectives motivate agreement (concord) with interpretable counterparts on the head noun (Carstens, 2000, 2003), and determine syntactic differences between Spanish and English (Bosque & Picallo, 1996, Mallen, 1997). Given this perspective, adult Anglophone learners would need to gain interpretable gender for Spanish nouns, [*ugender*] on determiners (English has limited number agreement on determiners like *this*, *these*, *that*, *those*), and [*ugender*]/[*unumber*] on adjectives. Representational deficit approaches (e.g. Franceschina 2001a, 2001b; Hawkins & Franceschina, 2004) posit that L2 grammatical features different from

those of the L1 cannot be acquired after the critical period, whereas for other UG accounts, there would be no critical period deficit (L2A is similar for children and adults) and *u*features are available in principle (e.g. Bruhn de Garavito & White, 2002; White et al., 2004).

These proposals have been evaluated in terms of offline (cited earlier) and online data. Monolinguals and early bilinguals of a gender concord language have demonstrated congruency/incongruency effects by showing faster recognition of nouns with congruent rather than incongruent gender marking on determiners and adjectives (for French and Spanish monolinguals, see Antón-Méndez, Nicol, & Garrett, 2002; Grosjean, Pommergues, Cornu, Guillelmon, & Besson, 1994; Jacobowicz & Faussart, 1998; for French bilinguals, Guillelmon and Grosjean, 2001). Some online studies report that Anglophones learning a gendered L2 as adults do not show sensitivity to determiner-noun gender discord (Guillelmon & Grosjean, 2001), and Lew-Williams and Fernald's (2007) eyetracking data suggest that such learners do not rely on determiners to process gender agreement like natives do. In contrast, other studies insist that learner sensitivity to determiner-noun and noun-adjective gender and number agreement violations is evident from behavioral measures (gender: Alarcón 2006; Herschensohn & Frenck-Mestre 2005; Keating, 2009; Sagarra & Herschensohn, 2008; number: Frenck-Mestre, Osterhout, McLaughlin & Foucart, 2008) and neurocognitive methodologies (gender: Rossi, Gugler, Friederici, & Hahne, 2006; Tokowicz & MacWhinney, 2005; number: Osterhout, Poliakov, Inoue, McLaughlin, Valentine, Pitkanen, Frenck-Mestre & Herschensohn, 2008; Rossi et al., 2006). The majority of these studies employ online or offline techniques to explore grammatical gender agreement or investigate gender and number agreement separately. We use both online (moving windows) and offline (grammaticality judgments) techniques to compare (a) gender agreement with animate and inanimate nouns, and (b) gender and number agreement with inanimate nouns.

3 Processing of gender agreement with animate and inanimate nouns

In Spanish, the gender of animate nouns correlates with animate sex whereas that of inanimate nouns is semantically arbitrary. Scholars disagree as to whether gender agreement with animate and inanimate nouns is processed differently. Some studies show no differences (Spanish monolinguals: Antón-Méndez et al., 2002; Barber, Salillas, & Carreiras, 2004). Others indicate that gender agreement with animate nouns is easier to process (fewer agreement production errors in nouns and shorter RTs at adjectives) than gender agreement with inanimate nouns, because the gender of animate nouns corresponds to biological sex (Spanish monolinguals: Alarcón, 2009; Antón-Méndez, 1999; Vigliocco & Franck, 1999; L2 Spanish learners: Alarcón, 2009; Fernández-García, 1999; Finnemann, 1992). Finally, the data of other studies suggest that gender agreement with animate nouns is more difficult to process (more agreement production errors in nouns and longer RTs at adjectives) than gender agreement with inanimate nouns, because animate nouns require the processor to choose between two options (in line with both lexical accounts to gender: *esposo/esposa* 'spouse_{masc, sing}', and syntactic accounts to gender: *-o/-a*) vs. single gendered inanimate nouns (see next section for more on lexical and syntactic accounts) (Spanish monolinguals: Igoa, García-Albea, & Sánchez-Casas, 1999; L2 Spanish learners: Alarcón, 2009; Bruhn de Garavito & White, 2002, with low proficiency learners).

Of the aforementioned studies, only Alarcón (2009) employs an online task to compare gender agreement with animate and inanimate nouns in adult English–Spanish learners of different proficiency levels. Her participants read complex NPs such as *el arquitecto del museo es...* 'the

architect of the museum is...’ and chose between a masculine or feminine adjective (*orgullosa/orgullosa* ‘proud_{fem, sing/masc, sing}’). For noun animacy, she found shorter RTs at incongruent adjectives with animate head nouns (*arquitecto*) but longer RTs at adjectives with animate attractor nouns (*museo*). These data confirm that the complement nouns (see also Bock, 1995) and the proximity between a noun/pronoun and its modifiers affect processing (see also Abu-Rabia, 2003; Dekydtspotter, Donaldson, Edmonds, Liljestrand Fultz & Petrush, 2008; Keating, 2009). For example, Keating reported that increased structural distance between a noun and an adjective decreased English–Spanish learners’ sensitivity to gender agreement violations. Alarcón’s study lacks a [+proximity] condition, and Keating’s study excludes gender agreement with animate nouns. The present study focuses on gender concord with animate and inanimate nouns and looks only at contiguous adjectives, obviating the structural distance issue.

4 Processing of gender and number agreement with inanimate nouns

Research on the processing of grammatical gender and number agreement is as inconclusive as that on the processing of animate and inanimate gender agreement. Thus, some monolingual studies show that gender and number are processed similarly—the P 600 effect in event-related potentials (ERPs) for both and no differences in grammatical priming with word pairs—(Monolinguals: Colé & Segui, 1994; Lukatela, Kostic, Todorovic, Carello, & Turvey, 1987; Osterhout & Mobley, 1995). In contrast, other studies suggest that gender agreement is more difficult to process (more agreement production errors, longer RTs in modifiers, and longer ERP latencies) than number agreement (Monolinguals: Antón-Méndez et al., 2002; Barber & Carreiras, 2003, 2005; De Vicenzi & Di Domenico, 1999; Faussart, Jakubowicz, & Costes, 1999, Nicol & O’Donnell, 1999; Vigliocco Hartsuiker, Jarema, & Kolk, 1996; L2 learners: Bruhn de Garavito & White, 2002; Franceschina, 2001a, 2002; White et al., 2004). This finding is in line with lexical (e.g., Igoa et al., 1999) and syntactic (e.g., Sicuro Correa, Almeida, & Sobrino Porto, 2004) accounts to gender. If gender is a stem feature accessed from the full word form and number is affixal (lexical accounts), gender discord is more taxing than number discord because the former requires the processor to return to the initial lexical identification stage to assess whether the right lexical entry had been selected, whereas the latter only checks the final syntactic stage. In turn, if both gender and number are affixal (syntactic accounts), gender disagreement is more cognitively demanding than number disagreement because the former is less predictable than the latter (there are more Spanish nouns violating the *-o/-a* rule to denote masculine/feminine gender than Spanish nouns violating the *-s/-es* rule to mark plurality) (see e.g., Ellis, 2001, and Hernández, Hoffman, & Kotz, 2007, for evidence that grammatical items with more irregularities consume more attentional resources than those with fewer irregularities).

Particularly relevant for our study, Gillon-Dowens, Barber, Vergara, & Carreiras (in press) reported that highly-proficient English–Spanish learners behave like Spanish natives for determiner-noun number discord but they lacked the LAN-P600 effect shown in Spanish monolinguals for gender discord. They conclude that cognitive factors such as working memory could be responsible for the gender-number difference in the L2 learners, because the distinction was found in learners with (English) and without (Chinese) determiner-noun number agreement (meaning the difference could not be due to transfer), and because previous research establishes a correspondence between working memory and both the LAN effect (Kluender & Kutas, 1993) and the processing of L2 grammatical information (see next section).

5 Working memory and L2 processing

If there are differences between animate and inanimate gender agreement and between grammatical gender and number agreement, and the differences can be due to cognitive demands, it is important to investigate the relationship between working memory and these L2 agreement mechanisms. Single-resource theories of attention (e.g., Baddeley, 2003; Just & Carpenter, 1992) claim that we have a limited capacity to process and store incoming information during complex cognitive tasks, such as processing L2 gender agreement (see Williams, 2011, for a comprehensive review on working memory and L2A). If task demands exceed a person's limited capacity, processing will slow down (longer RTs) and/or storage will decrease (lower comprehension). Because acquiring another language as an adult imposes an additional processing load (Hasegawa, Carpenter, & Just, 2002), it is not surprising that working memory has been linked to L2 syntactic processing (e.g., Havik, Roberts, van Hout, Schreuder, & Haverkort, 2009; cf. Juffs, 2004), L2 verbal and nominal agreement (Sagarra, 2007a, 2007b), and L2A in general (Ellis & Sinclair, 1996; Miyake & Friedman, 1998). Sawyer and Ranta note that "if attention at any moment is limited to working memory capacity, then there must logically be a close relationship between amount of learning and size of working memory" (2001, p. 342). Especially important to the present research project, Sagarra (2007b) asked English–Spanish learners to read Spanish sentences with noun-adjective concord/discord and answer comprehension questions, and found that those with larger working memory were more sensitive to gender agreement violations (longer RTs at adjectives and lower comprehension). Considering the vast number of studies showing a correspondence between working memory and L2 processing as well as the L1 and L2 behavioral and neurocognitive evidence pointing to cognitive demands as the explanation of concord processing differences, we decided to include working memory as a variable in our study.

6 The study

As previously mentioned, many L2 studies are offline and focus on proficient learners. Heeding White's (2003) admonition to use varied tests, the studies cited in this article (Sagarra & Herschensohn, 2010, 2011) employ an online technique, a non-cumulative self-paced moving window test, and an offline methodology, a grammaticality judgment task, to investigate two questions: (1) whether beginning and intermediate adult learners with an ungendered L1 can gain native-like computation of L2 Spanish adjectives and, if they do, whether L2 proficiency and working memory affect the acquisition process; and (2) whether they process (a) animate and inanimate gender agreement and (b) inanimate gender and number agreement differently, and if they do, whether working memory modulates the increased difficulty of one over the other. We also explore the question of whether learners whose RTs are qualitatively similar (i.e., longer RTs to discord than concord) to those of native Hispanophones have gained grammatical features of gender and number on L2 Spanish adjectives. Adopting the idea of possible acquisition of L2 computation and representation, we hypothesize that beginning learners will be insensitive to agreement violations but that for intermediates, sensitivity to number violations will emerge before sensitivity to gender violations. This developmental path also indicates the importance of experience with the language and proficiency. We also expect the intermediates to process number better than gender agreement violations, and we expect to see animacy as a cognitively complex feature, as for monolinguals. Finally, we expect greater working memory to facilitate sensitivity to gender disagreement and processing of grammatical features that are more cognitively taxing (e.g., gender concord may consume more attentional resources than number concord). In this article we present the

results of two earlier studies that appeared separately in order to reconsider their results together. So we recap our experiments on L2 and monolingual noun-adjective concord/discord with respect to number, gender, animacy (and WM) to see if L2ers process like monolinguals, and if so what that could mean.

7 Experiment I

7.1 Participants

To investigate the first two questions, Sagarra & Herschensohn (2011) asked 63 Spanish monolinguals and 69 beginning and 64 intermediate English–Spanish learners to read sentences in Spanish containing noun-adjective gender concord/discord with animate or inanimate nouns. For the moving window test, participants answered comprehension questions after each sentence and had to be at least 60% accurate, and for the grammaticality judgment test, they indicated whether each sentence was grammatically correct. Participants could not have lived in a foreign-speaking country or a bilingual Spanish province for more than one month. The Spanish monolinguals grew up in areas of southern Spain without post-vocalic /-s/ lenition (something relevant for the investigation of number concord, Experiment 2) and spoke no L2 apart from English. The learners needed to have studied Spanish postpuberty, have no knowledge of other L2s, score within 3 SDs from the mean of a Spanish proficiency test, and obtain a perfect score on a grammar test and a vocabulary test to ensure that slower RTs were not due to poor L2 knowledge. The intermediate learners were enrolled in their 7th or 8th semester of study of Spanish and the beginning learners in their 3rd semester.

7.2 Materials and procedure

Participants completed seven tests in two one-hour sessions one week apart: a language background questionnaire (one for Spanish monolinguals and one for Spanish learners), a Spanish proficiency test (learners only), and a moving window test in session 1; then a grammaticality judgment test, a Spanish grammar and vocabulary tests (learners only), and a working memory test in session 2. The Spanish monolinguals' English self-ratings, part of their language background questionnaire, revealed that their L2 functional proficiency was too low to affect L1 processing (the means for the four skills ranged between 3.13 and 4.21, in a Likert scale in which 1 = minimum ability, and 10 = native proficiency) (see Bonnet, 2002, for correlations between L2 proficiency and self ratings). The Spanish proficiency test, an excerpt of the *Diploma de Español como Lengua Extranjera* (intermediate level), revealed that the intermediates—scores between 15 and 25—showed significantly higher L2 knowledge ($M = 20.65$, $SD = 3.00$) than the beginners—scores between 0 and 10—($M = 7.04$, $SD = 2.89$) (the results of an independent-samples t test were $t(125) = -22.667$, $p < .01$; Levene's $F = .000$, $p < .05$).

Because complex tasks can hinder sensitivity to the already redundant and non-salient adjectival morphology, the moving window and grammaticality judgment tests were written and self-paced (see Montrul, Foote, & Perpiñán, 2008, and Sabourin, 2003, for evidence that written and self-paced tasks are easier than oral and timed ones). Each test contained grammatical and ungrammatical practice sentences (in line with the experimental and filler sentences), 40 experimental sentences (10 per condition), and 70 fillers. Sentences were pseudo-randomized using a Latin square design that divided the sentences into blocks to avoid the appearance of two ungrammatical sentences close to each other. Sentences were grammatically and lexically controlled for L2 level

(grammar, syntax, and vocabulary adequate to third semester Spanish learners) and length (9–15 words long). Experimental sentences followed the same syntactic structure and had four conditions: [+/-concord, +/-animate]. For example, *El padre quiere el esposo/trabajo perfecto/*perfecta para su hija* ‘The father wants the_{masc, sing} husband_{masc, sing}/job_{masc, sing} perfect_{masc, sing}/*perfect_{fem, sing} for his daughter.’ Finally, the moving window and the grammaticality judgment tests used the same pool of experimental nouns and adjectives but a given noun-adjective combination only appeared once to avoid practice effects.

The target noun phrase was formed by a masculine singular noun that was countable in English and Spanish (to avoid interpretive issues with mass nouns) and a descriptive adjective, both with transparent gender. We focused on contiguous noun-adjective concord because distance between a noun and an adjective affects sensitivity to grammatical gender discord in L2 Spanish (e.g., Keating, 2009) and we excluded nouns that formed part of gender-inflected pairs (e.g., *puerto-puerta* ‘seaport_{masc, sing} door_{fem, sing}’) to avoid possible priming with the opposite gender. Also, the unmarked form of gender (masculine: Harris, 1991; transparent: Antón-Méndez, 1999) and number (singular: Eberhard, 1997) was chosen for several reasons. First, for logistical reasons (participants were already reading 114 sentences per test), additional conditions such as masculine/feminine, transparent/opaque gender, singular/plural would have forced us to decrease the number of sentences per condition, thus sacrificing the statistical power of the experiment. Furthermore, considering that the beginning learners only had 3 semesters of instruction, it made sense to investigate an area where it was anticipated that grammaticalization first appears (McCarthy’s 2008 findings that L2 learners are more accurate with masculine singular than feminine plural forms in comprehension and production indicates that grammatical representation happens earlier in default/unmarked forms). In addition, there is L1 and L2 Spanish evidence that a noun’s gender or number does not affect RTs of congruent/incongruent determiners and adjectives in L1 or L2 Spanish (e.g., Alarcón, 2009; Antón-Méndez et al., 2002; Keating, 2009).

To further demonstrate that longer RTs at the adjective should be interpreted as sensitivity to a grammatical violation (e.g., *trabajo *perfecta* ‘job_{masc, sing} perfect_{*fem, sing}’) rather than as a mere reaction to the marked form of gender (feminine) or number (plural), 30 of the fillers contained feminine singular ($k = 10$), masculine plural ($k = 10$), and feminine plural ($k = 10$) noun-adjective well-formed combinations. These gave rise to five conditions: (1) masculine singular noun with masculine singular adjective (see Table 1 for descriptive statistics); (2) feminine singular noun with feminine singular adjective (beginners: $M = 855.58$, $SD = 242.18$; intermediates: $M = 691.12$, $SD = 192.25$; Spanish monolinguals: $M = 455.28$, $SD = 97.87$); (3) masculine plural noun with masculine plural adjective (beginners: $M = 852.19$, $SD = 237.00$; intermediates: $M = 683.61$, $SD = 189.78$; Spanish monolinguals: $M = 448.34$, $SD = 87.07$); (4) masculine singular noun with feminine singular adjective (see Table 1 for descriptive statistics); and (5) masculine singular noun with masculine plural adjective (see Table 1 for descriptive statistics). A repeated-measures ANOVA with a 5 (Condition) \times 3 (Group) factorial design on the mean RTs at the adjective revealed a significant main effect for Condition ($F(4,756) = 31.793$, $p < .01$) and Group ($F(4,756) = 31.793$, $p < .01$), and a significant interaction of Condition \times Group ($F(4,756) = 31.793$, $p < .01$). Bonferroni posthoc tests indicated longer mean RTs in gender and number discord than gender/number concord in intermediates and Spanish monolinguals, independently of the noun’s gender and number—default vs. marked—(no concord–discord differences were obtained for the beginning learners). These results demonstrate that it is sensitivity to agreement violations rather than markedness that causes longer mean RTs at adjectives in sentences with agreement violations.

The sentences were presented with E-Prime on a Dell 19” monitor with 16-font Arial black characters on a white background. Letters were presented in standard upper and lower case format.

Table 1. Mean RTs at the noun (N-I) and at the adjective (N) in milliseconds, in the moving window test.

	Beginners		Intermediates		Spanish monolinguals	
	M	SD	M	SD	M	SD
Word before adjective						
<i>Animate nouns:</i>						
Gender/number concord	1034	451	867	393	504	217
Gender discord	1021	376	832	393	506	221
<i>Inanimate nouns:</i>						
Gender/number concord	986	342	891	315	469	168
Gender discord	986	315	835	291	493	175
Number discord	957	368	819	288	495	175
Adjective						
<i>Animate nouns:</i>						
Gender/number concord	930	344	744	225	500	158
Gender discord	934	311	9143	381	614	264
<i>Inanimate nouns:</i>						
Gender/number concord	901	333	709	207	465	115
Gender discord	911	310	8243	277	555	190
Number discord	952	360	856	259	546	167
Word after adjective						
<i>Animate nouns:</i>						
Gender/number concord	547	133	480	107	420	94
Gender discord	546	135	569	157	477	106
<i>Inanimate nouns:</i>						
Gender/number concord	551	144	446	85	536	138
Gender discord	531	121	536	138	450	111
Number discord	521	139	515	120	434	81

$n = 69$ for beginners, $n = 64$ for intermediates, and $n = 63$ for Spanish monolinguals. $K = 10$.

For each sentence of the moving window test, participants looked at a fixation sign for 500 ms, saw dashes indicating where the words would appear, pressed the space bar key to read the first word silently, then pressed the space bar key to make the first word disappear and the second one appear, and repeated this process until the last word of the sentence disappeared and a comprehension question was displayed; then participants read the question and answered by pressing a 'yes' or a 'no' button. Comprehension questions ensured that possible reading latencies were not a result of lack of understanding, excluded the adjective, and did not assess gender/number marking or concord. The moving window test generated two scores: (1) mean RTs, the mean of all word RTs between 200 and 2000 ms within a condition (the cutoff is based on Rayner & Pollatsek's 1989 findings that English monolinguals need between 225 and 300 ms to process single words), and (2) accuracy on comprehension questions, with 1 point per correct answer, and taking into account that statistics excluded sentences with incorrect answers to the comprehension questions to minimize obtaining RTs due to lack of understanding.

For the grammaticality judgment test, participants read Spanish sentences silently, decided whether they were correct, circled the source of the error in incorrect sentences, and indicated

how confident they were about their answer on a 5-point Likert scale. We used confidence rating because the use of multiple tests is necessary to discern between performance (accuracy score) and awareness (confidence score. Tunney, 2005). The grammaticality judgment test gave rise to two scores: (1) accuracy, defined as 1 point for identifying correct sentences as correct or incorrect ones as incorrect including circling the right error source, and (2) confidence ratings, based on a 5-point continuous score ranging from 1 (not sure at all) to 5 (completely sure). Only confidence judgments for sentences with accurate grammaticality judgments were part of the statistics to be able to compare accuracy and confidence ratings to draw conclusions based on two judgment tests.

After the grammaticality judgment test, participants completed three additional tests. The grammar test required them to identify the gender and number of a list of Spanish nouns ($\frac{1}{4}$ masculine singular (all experimental nouns), $\frac{1}{4}$ feminine singular, $\frac{1}{4}$ masculine plural, $\frac{1}{4}$ feminine plural) and ensured that longer RTs were not the result of lack of grammatical knowledge. For example, they read *trabajo* ‘job_{masc, sing}’ and they had to circle one of these options: masculine singular, feminine singular, masculine plural, feminine plural. The vocabulary test controlled for familiarity with the meaning of the target nouns and adjectives by asking learners to match them with their English translation. Finally, the working memory test consisted of a reading span test for which participants read sets of 2 to 6 sentences silently at a fast pace, decided whether each sentence was plausible, and produced the last word of each sentence at the end of each set. Participants read sentences in their native language following research suggesting that working memory is language-independent (Osaka & Osaka, 1992; Xue, Dong, Jin, & Chen, 2004) and taking into consideration the low proficiency level of the beginning learners.

7.3 Results

7.3.1 Moving window test. Table 1 shows the means and standard deviations for the RTs at the noun, adjective, and adjective + 1 for Experiments 1 and 2. The mean RTs for the word before the adjective seem similar across conditions for the three groups but the mean RTs at the adjective and the word after the adjective for intermediates and Spanish monolinguals tend to be longer in the discord conditions.

A series of repeated-measures ANOVAs with a 2 (gender agreement) \times 2 (noun animacy) \times 3 (group) factorial design were carried out on the mean RTs at the word preceding the adjective, the adjective, and the word following the adjective (to control for delayed processing). We separated gender agreement from noun animacy to be able to determine the effects of the two variables separately (running a 4 \times 3 ANOVA would not inform us about the source of a main significant effect). Mean RTs at the word before the adjective revealed no differences across conditions within a group, confirming a solid baseline before participants reached the adjective. The mean RTs at the adjective and the word after the adjective revealed a main effect for agreement (adjective: $F(1, 193) = 61.968, p < .01$; adjective + 1: $F(1, 193) = 59.870, p < .01$), animacy (adjective: $F(1, 193) = 11.236, p < .01$; adjective + 1: $F(1, 193) = 19.203, p < .01$), and group (adjective: $F(2, 193) = 47.797, p < .01$; adjective + 1: $F(2, 193) = 24.608, p < .01$), as well as a significant interaction of agreement \times group (adjective: $F(2, 193) = 10.198, p < .01$; adjective + 1: $F(2, 193) = 25.005, p < .01$). Pairwise contrast comparisons showed longer RTs at the adjective and the word following the adjective in intermediates and Spanish monolinguals (a) in sentences with gender disagreement than agreement, regardless of noun animacy (all $p < .01$) and (b) in sentences with animate rather than inanimate nouns, regardless of gender agreement (the animacy effect was evident in the adjective for both groups, except for intermediates who showed a delayed animacy effect in sentences

Table 2. Accuracy on the comprehension questions of the moving window test.

	Beginners		Intermediates		Spanish monolinguals	
	M	SD	M	SD	M	SD
<i>Animate nouns:</i>						
Gender/number concord	8.58	1.01	8.45	1.37	9.13	.75
Gender discord	8.26	1.21	8.40	1.27	8.43	1.09
<i>Inanimate nouns:</i>						
Gender/number concord	8.67	1.23	8.80	1.28	9.49	.64
Gender discord	8.57	1.04	8.72	1.29	8.84	.99
Number discord	8.32	1.18	8.75	1.20	9.44	.64

$n = 69$ for beginners, $n = 64$ for intermediates, and $n = 63$ for Spanish monolinguals. $K = 10$.

with gender disagreement). The main effect of group was caused by beginners reading slower than intermediates and the latter in turn slower than Spanish monolinguals, an expected finding.

The descriptive statistics for the total number of correct responses to the comprehension questions for Experiments 1 and 2 are displayed in Table 2. For all groups, the means tend to be lower in sentences with animate gender concord/discord than with inanimate gender concord/discord.

A $2 \times 2 \times 3$ repeated-measures ANOVA confirmed a significant main effect for animacy ($F(1,193) = 17.299, p < .01$), owing to better understanding of sentences with animate than inanimate nouns for all groups. In addition, there was a significant interaction of animacy \times group ($F(2,193) = .620, p < .05$) and animacy \times agreement \times group ($F(2,193) = .283, p < .05$), because Spanish monolinguals understood sentences better than intermediates and intermediates better than beginners, across conditions (all $p < .05$).

7.3.2 Grammaticality judgment test. The means and standard deviations for the grammaticality and confidence judgments for Experiments 1 and 2 are shown in Table 3. The means show a clear ceiling effect for Spanish monolinguals across conditions, and a tendency for L2 learners to be more accurate correctly identifying animate/inanimate gender concord than discord.

Two additional $2 \times 2 \times 3$ repeated-measures ANOVAs were conducted: one on the total number of correct grammaticality judgments and one on the mean of confidence judgments on sentences with accurate grammaticality judgments. The ANOVAs showed a significant main effect for agreement (grammaticality judgments: $F(1, 165) = 145.210, p < .01$; confidence judgments: $F(1,165) = 32.113, p < .01$), animacy (grammaticality judgments: $F(1, 165) = 15.579, p < .01$; non-significant for confidence judgments), and group (grammaticality judgments: $F(2, 165) = 128.298, p < .01$; confidence judgments: $F(2,165) = 126.058, p < .01$), as well as a significant interaction of agreement \times group (grammaticality judgments: $F(2, 165) = 67.654, p < .01$; confidence judgments: $F(2,165) = 21.560, p < .01$) and animacy \times group (grammaticality judgments: $F(2, 165) = 6.833, p < .01$; non-significant for confidence judgments). Thus, participants did not feel more or less confident about their answers depending on noun animacy. Pairwise contrast comparisons showed no differences among conditions for Spanish monolinguals because of ceiling effects. However, beginners and intermediates were more accurate with gender agreement than disagreement (all $p < .01$) because they classified sentences as correct when unsure. In addition, intermediates were more accurate in sentences with inanimate than animate nouns ($p < .01$), in line with the moving window

Table 3. Mean and standard deviations of the grammaticality judgment test.

	Beginners		Intermediates		Spanish monolinguals	
	<i>M</i>	<i>SD</i>	<i>M</i>	<i>SD</i>	<i>M</i>	<i>SD</i>
Grammaticality judgments (<i>k</i> = 10)						
<i>Animate nouns:</i>						
Gender/number concord	8.52	1.66	7.34	2.08	10.00	.00
Gender discord	2.62	3.82	5.64	3.03	9.97	.18
<i>Inanimate nouns:</i>						
Gender/number concord	8.73	1.51	8.26	1.32	10.00	.00
Gender discord	2.92	4.11	6.47	3.40	10.00	.00
Number discord	5.94	2.51	7.98	1.54	10.00	.00
Confidence judgments (<i>k</i> = 5) (1 = not sure at all; 5 = completely sure)						
<i>Animate nouns:</i>						
Gender/number concord	3.53	.82	3.56	.71	5.00	.00
Gender discord	3.61	.78	3.96	.69	5.00	.00
<i>Inanimate nouns:</i>						
Gender/number concord	3.51	.85	3.54	.67	5.00	.00
Gender discord	3.56	.91	3.97	.72	5.00	.00
Number discord	3.97	.82	4.42	.60	5.00	.00

n = 52 for beginners, *n* = 53 for intermediates, and *n* = 63 for Spanish monolinguals. The sample size varied slightly because some L2 learners did not complete the grammaticality judgment test.

findings suggesting that animate nouns are more cognitively taxing than inanimate nouns. Finally, the group effect resulted from Spanish monolinguals being more accurate and confident than the L2 learners and intermediates being more accurate and confident than beginners. Most importantly, in support of non-deficit approaches, intermediates were more accurate in sentences with gender disagreement than beginners.

7.3.3 Working memory test. The descriptive statistics per group were: beginners: *M* = 44.99, *SD* = 12.80; intermediates: *M* = 56.64, *SD* = 12.78; Spanish monolinguals: *M* = 42.94, *SD* = 11.88. A one-way ANOVA showed significant differences ($F(2,193) = 22.443$, $p < .01$) and Bonferroni posthoc tests revealed that the intermediate group had a higher working memory level than the other two groups. Then, a series of correlations were carried out between reading span scores and the mean proportional increase in RTs on ungrammatical adjectives compared to grammatical ones, following Waters and Caplan (1996). The findings indicated that, for the intermediate group, reading span correlated positively with the mean proportional increase in RTs (a) on adjectives that disagreed compared to those that agreed in gender with animate nouns ($r = .461$, $p < .05$), and (b) on those that disagreed compared to those that agreed in gender with inanimate nouns ($r = .303$, $p < .05$). The rest of correlations were non-significant (see Discussion section for an explanation of these findings).

7.4 Discussion

Taken together, the results from Experiment 1 demonstrate that while beginning L2 Spanish learners are insensitive to adjective concord/discord and animacy, intermediate adult learners with an

ungendered L1 display more target-like patterns than beginning learners, both in terms of grammatical features (gender concord–discord) and semantic features (animate/inanimate), namely longer RTs and lower accuracy in grammaticality judgments in gender discord than concord and in animate than inanimate nouns. These data indicate that adults can acquire target-like computation of grammatical features absent in their L1 in a way that is qualitatively similar to monolinguals (see also Bruhn de Garavito & White, 2002; White et al., 2004). The findings also suggest that language experience affects the computation of concord/discord in Spanish L2 adjectives, and might potentially affect the representation: early stages of L2A may be limited to transfer of L1 features and are clearly not sensitive to L2 grammatical concord, whereas sensitivity to grammatical features seems to develop over time.

The results also reveal that animate nouns are more difficult to process than inanimate nouns for intermediate learners as well as for native speakers. This can be due to the processor having to return to the lexical identification stage in animate but not inanimate nouns (lexical accounts posit that *esposo* ‘spouse_{masc, sing}’ primes *esposa* ‘spouse_{fem, sing}’ whereas *mesa* ‘table_{fem, sing}’ does not prime **meso*), or to the processor requiring more time to choose between masculine and feminine suffixes in animate nouns than in single gendered inanimate nouns. In either case, the learners show the same pattern (albeit slower) as monolinguals. We now turn to our second experiment which looks at a different grammatical feature, number (a feature that exists in L1 English on nouns and determiners, although not adjectives), and compare its processing to that of gender on inanimate nouns.

Finally, our findings suggest that individual differences in working memory can modulate the processing of L2 gender agreement with animate and inanimate nouns: intermediates with larger working memory capacity displayed greater differences between discord and concord conditions (i.e., greater sensitivity to gender agreement violations) than those with smaller working memory capacity. The lack of significant correlations for beginners can be explained in terms of developmental readiness à la Pienemann (see Pienemann & Kessler, 2011). The grammar test demonstrated that the beginners knew the gender and number of the target nouns and the grammaticality judgment test showed that they had declarative knowledge of gender agreement, but the moving window test confirmed that they did not have procedural knowledge about gender agreement—meaning they had not really acquired gender agreement—(see Paradis, 2009; Morgan-Short, & Ullman, 2011 for more information on declarative and procedural knowledge). The explicit and offline nature of grammaticality judgment tasks accounts for the lack of working memory effects in both L2 groups. With regard to the Spanish monolingual group, the lack of relationship between working memory and mean RTs or grammaticality judgments was expected for being native speakers: discord–concord RT differences were large (all $p < .01$) and grammaticality judgments were at ceiling.

8 Experiment 2

8.1 Participants, method and procedures

In a related study that examined similarities and differences of gender and number processing in L2 Spanish, Sagarra and Herschensohn (2010) examine two grammatical features: number, existent as an uninterpretable feature on English determiners on UG accounts, and gender, non-existent in L1. Having determined in Experiment 1 that animacy plays a role in intermediate L2 processing similar to that of monolinguals, we aimed to see if there was a difference between the grammatical features of gender and number for inanimate nouns. We report the findings of another moving

window and grammaticality judgment experiment comparing beginning and intermediate Spanish adult learners' and Spanish monolinguals' responses to inanimate gender and number discord, and we also consider the role of working memory. The participants, method and procedure of Experiment 2 mirrored those of Experiment 1 (same participants for both studies), and once again canonical masculine singular transparent adjective concord/discord contrasted with feminine and plural yielding identical results to those obtained in Experiment 1. Experiment 2 included four practice sentences (same as in Experiment 1) 30 experimental sentences, and 80 fillers (10 of which were experimental sentences for Experiment 1). There were three conditions: (+ gender/number concord, - gender concord, - number concord). For example, *El ingeniero presenta el prototipo famoso/*famosa/*famosos en la conferencia* 'The engineer presents the_{masc, sing} prototype_{masc, sing} famous_{masc, sing}/*famous_{fem, sing}/*famous_{masc, pl} at the conference'. Sentences including both gender and number discord were excluded because adult learners conceive these errors as semantic rather than grammatical violations (Rossi et al., 2006).

8.2 Results

8.2.1 Moving window test. Table 1 displays the means and standard deviations for the RTs at the noun, adjective, and adjective + 1, and shows similarity of the mean RTs at the word before the adjective across conditions, as well as a tendency for intermediates and Spanish monolinguals to need more time to process adjectives and the word after the adjectives in sentences with grammatical violations.

The inferential statistics consisted of three 3 (condition) × 3 (group) repeated-measures ANOVAs: one for the word preceding the adjective, one for the adjective, and one for the word following the adjective. As expected, the first ANOVA showed no differences among conditions. The remaining ANOVAs revealed a significant main effect for condition (adjective: $F(2,386) = 16.608, p < .01$; adjective + 1: $F(2,386) = 11.887, p < .01$) and group (adjective: $F(2,193) = 54.995, p < .01$; adjective + 1: $F(2,193) = 27.129, p < .01$), as well as a significant interaction of condition × group (adjective: $F(4,386) = 2.407, p < .05$; adjective + 1: $F(4, 386) = 7.923, p < .01$). Bonferroni posthoc comparisons showed no differences among conditions for beginners, but intermediates and Spanish monolinguals displayed longer RTs with gender or number discord than gender/number concord (all $p < .01$). No differences were obtained between gender and number concord/discord. Finally, between-group comparisons revealed that Spanish monolinguals read faster than intermediates and intermediates were in turn faster than beginners.

Table 2 displays the descriptive statistics for the total number of correct responses to the comprehension questions. The means indicate a tendency for Spanish monolinguals to understand sentences better with gender than number disagreement. In line with the RT findings, a 3 × 3 repeated-measures ANOVA revealed a significant main effect for Condition ($F(2,386) = 4.282, p < .05$) and Group ($F(2,193) = 16.258, p < .01$), as well as a significant interaction of Condition × Group ($F(4,386) = 4.075, p < .01$). Bonferroni posthoc tests showed that Spanish monolinguals understood sentences better with gender/number concord than gender discord, as well as those with number than gender discord (all $p < .01$). Finally, between-group comparisons revealed that Spanish monolinguals understood sentences better than intermediates and intermediates in turn better than beginners.

8.2.2 Grammaticality judgment test. The means and standard deviations for the grammaticality and confidence judgments are shown in Table 3, and suggest that Spanish monolinguals were very accurate for all conditions, and that intermediates seemed to be more accurate with sentences with

gender/number concord than discord. Additional ANOVAs were conducted on the total number of correct grammaticality judgments and the mean of confidence judgments on sentences with accurate grammaticality judgments. The ANOVAs showed a significant main effect for condition (grammaticality judgments: $F(2,330) = 95.218, p < .01$; confidence judgments: $F(2,330) = 47.665, p < .01$) and group (grammaticality judgments: $F(2,165) = 137.259, p < .01$; confidence judgments: $F(2,165) = 125.199, p < .01$), and a significant interaction of condition \times group (grammaticality judgments: $F(4,330) = 45.113, p < .01$; confidence judgments: $F(4,330) = 16.993, p < .01$). As in Experiment 1, Bonferroni posthoc comparisons revealed (a) no differences across conditions in Spanish monolinguals because of ceiling effects, (b) higher accuracy with gender/number concord than gender or number discord in beginners and intermediates (all at least $p < .05$), and (c) higher accuracy and confidence ratings in Spanish monolinguals than intermediates and in intermediates than beginners (all at least $p < .05$). As for gender–number comparisons, all learners were more accurate and confident with number discord than gender discord and more accurate at identifying number than gender errors (all at least $p < .05$) (ceiling effects explain lack of significant differences in Spanish monolinguals).

8.2.3 Working memory test. Section 7.3.3 presents the descriptive statistics for the working memory test per group and explains how we analyzed the relationship between reading span and sensitivity to gender violations. Following Experiment 1, several correlations were conducted between reading span scores and the mean proportional increase in RTs on ungrammatical adjectives compared to grammatical ones. The findings indicated that, for the intermediate group, reading span correlated positively with the mean proportional increase in RTs on adjectives that disagreed compared to those that agreed in gender with inanimate nouns ($r = .303, p < .05$)—already reported in Experiment 1. The rest of comparisons were non-significant. We discuss these findings in the next section.

8.3 Discussion

Experiment 2 indicates that Anglophone learners with number but no L1 gender can gain sensitivity to gender and number agreement in L2 Spanish. Such sensitivity, indicating grammatical computation comparable to that of Spanish monolinguals, may reveal acquisition of L2 grammatical features of gender and number on adjectives. As in Experiment 1, moving window and grammaticality judgment data show that intermediates and Spanish monolinguals, but not beginners, are sensitive to gender and number agreement violations on inanimate nouns. With regard to differences between gender and number, all learners are more accurate and confident in their grammaticality judgments about sentences with number discord than gender discord. Taking together these results and the findings that larger working memory span positively correlated with sensitivity to gender agreement (significant correlation between reading span and the mean proportional increase in RTs on gender incongruent compared to gender congruent adjectives) but not with number agreement in intermediates, we suggest that gender disagreement may be more difficult to process than number disagreement (see section 7.4 for discussion of the other working memory findings).

9 General discussion

In this section, we reconsider our two research questions in terms of previous research on adult L2 grammatical representation and processing.

9.1 Question 1: Can adults of an ungendered L1 gain native-like computation of L2 grammatical features of gender and number? If they can, do language proficiency and working memory affect the acquisition process?

The UG approaches to representation we have outlined contrast the idea that L2 grammatical features ([*ugen*], [*unum*]) and concord/discord sensitivity are unavailable if nonexistent in the L1 (e.g. Hawkins & Franceschina, 2004) with the idea that these features and concord sensitivity are possible in the L2 (e.g., Hopp, 2007). Results from monolinguals and beginners from our study were not surprising: Spanish monolinguals show predictable native sensitivity to gender and number agreement violations (longer RTs), whereas beginners are not sensitive to any agreement violations (equivalent RTs and accuracy/confidence in grammaticality judgments for concord/discord).

The intermediate group differentiates the two approaches, because it is qualitatively similar to the Spanish natives in showing longer latencies to both gender and number disagreement than to agreement conditions on the adjective, and it shows native-like responses to animacy effects (shorter RTs and higher accuracy in sentences with inanimate than animate nouns). The qualitative similarities of the intermediates with the monolinguals show that they are processing adjectival concord and discord in a similar manner, potentially an indication of a parallel means of computation. While similar behavior may result from quite distinct native vs. L2 processing algorithm, it is also possible that “L2 sentence processing may in fact obey the same parsing principles as native sentence processing, but processing may break down to varying degrees as slower computations time out in the limits of available resources” (Dekydtspotter et al., 2008, p. 460). It is unclear exactly what processing can tell us about representation, but we may infer that implicit grammatical knowledge plays a role in computation. For example, Hawkins and Franceschina (2004) essentially claim that concord RTs by monolinguals, early and late bilinguals (Guillelmon & Grosjean, 2001) show critical period effects pointing to a representational deficit in late L2. If their opinion is correct, the results from our study seem to indicate that L2 learners start to show target-like processing and representation although beginners do not. If so, the emerging L2 grammar reflects a sensitivity to adjective concord that suggests that these learners have gained [*ugender*] and [*unumber*] features on Spanish adjectives. We conclude that intermediate learners are gaining L2-like computational strategies and could well have grammaticalized L2 gender and number features for grammatical features absent in their L1. We also conclude that higher L2 proficiency and working memory span facilitates sensitivity to gender agreement mismatches.

The emerging grammatical concord for adjective gender and number is unexpected under the view holding that intermediates should be incapable after puberty of gaining grammatical features nonexistent in their native language. Our results show, however, that intermediates can gain this sensitivity in online processing (i.e., implicit command of gender/number agreement on novel noun-adjective pairs, not explicitly memorized morphological rules). Beginners—even though they are able to use adequate cognitive strategies to produce adjective concord on classroom tests—behaviorally do not have automatized procedural knowledge (this is why there were no working memory effects in this group). In contrast, intermediates are gaining procedural knowledge of gender agreement. Like Spanish monolinguals, they are sensitive to gender disagreement, but it still consumes a great deal of attentional resources (this is why intermediates with larger working memory span are more sensitive to gender agreement violations than those with smaller working memory span). Based on the findings discussed in this section, we conclude that the intermediates are qualitatively similar to monolinguals and are not restricted to beginner cognitive strategies (cf. Hawkins & Franceschina, 2004).

9.2 Question 2: Do Anglophones process (a) natural and grammatical gender agreement and (b) grammatical and number agreement differently in L2 Spanish? If they do, does working memory modulate the increased difficulty of one over the other?

As mentioned earlier, we have a limited working memory capacity to process language, and both lexical and syntactic accounts of Spanish gender suggest that animate nouns and grammatical gender agreement are more cognitively demanding than inanimate nouns and number agreement, respectively. It is generally accepted (a) that gender mismatches with animate nouns force the processor to choose between two options vs. single-gendered inanimate nouns (lexical accounts: *esposo/esposa* ‘spouse-m/f’ vs. *libro* ‘book-m;’ syntactic accounts: *-o/-a* vs. *-o* or *-a*), and (b) that grammatical gender is more cognitively taxing than number agreement (lexical accounts: the processor needs to return to the lexical identification stage in gender discord but only to the final steps of syntactic recognition for number discord; syntactic accounts: it is more difficult to process gender than number concord because number marking has a single morpheme *-s*, whereas gender marking is less regular than number marking).

Experiment 1 reveals that animate nouns are more difficult to process than inanimate nouns: intermediates and Spanish monolinguals show longer RTs and lower accuracy in grammaticality judgments in sentences with animate than inanimate nouns, and all groups are more accurate in comprehension questions about sentences with inanimate than animate nouns. These findings demonstrate that the learners’ emerging competence is due to neither L1 transfer nor universal grammatical properties, but rather the computational aspect of language in use. Along the same lines, Experiment 2 shows that inanimate gender agreement is more cognitively demanding than number agreement: Spanish monolinguals are more accurate in comprehension questions about sentences with number than gender disagreement; all learners are more accurate and confident in their grammaticality judgments regarding number than gender discord; and intermediates with greater working memory capacity are more sensitive to gender disagreement (longer RTs and higher grammaticality judgment accuracy), but there are no working memory effects on the computation of number disagreement. Summarizing, the two experiments suggest that grammatical features consuming more attentional resources inhibit the computation of gender agreement mechanisms. Working memory is correlated with gender discord but not noun animacy because the former consumes significantly more resources than the latter (Table 1 shows that, for intermediates, the difference between RTs at the adjective for discord and concord conditions was greater than the difference between RTs at the adjective for animate and inanimate nouns), and working memory effects are only visible when performing *complex* cognitive tasks.

When we look specifically at the results of the two experiments, they give more insight to L2 computation. Like monolinguals, the intermediates compute number concord faster than gender, and process inanimate faster than animate nouns. This pattern could not be related to a facilitating effect of L1 English interpretable gender on animates (since it is the opposite result), nor to any sort of transfer since English has no gender concord. We conclude that the intermediates are beginning to develop L2 computation and presumably also grammatical representation.

10 Conclusion

Moving window and grammaticality judgment data showed that adult Anglophone L2 Spanish intermediates and Spanish monolinguals were sensitive to noun-adjective gender and number agreement violations and to noun animacy, but beginners were not. While the results

of the beginners and the Spanish monolinguals are expected, the intermediates' data shows the emergence of target-like processing and suggests the possible development of [*ugender*] and [*unnumber*] on adjectives. The findings also show that both grammatical (gender vs. number agreement) and semantic (noun animacy) features differentially impact concord processing for both natives and L2 learners, and that working memory modulates computation of gender agreement in learners who are developmentally ready.

References

- Abu-Rabia, S. (2003). The influence of working memory on reading and creative writing processes in a second language. *Educational Psychology, 23*, 209–222.
- Alarcón, I. (2006). *The second language acquisition of Spanish gender agreement: The effects of linguistic variables on accuracy*. Munich, Germany: Lincom Europa.
- Alarcón, I. (2009). The processing of gender agreement in L1 and L2 Spanish: Evidence from reaction-time data. *Hispania, 92*, 814–828.
- Antón-Méndez, M. I. (1999). *Gender and number agreement processing in Spanish*. Unpublished doctoral dissertation, University of Arizona.
- Antón-Méndez, M. I., Nicol, J. L., & Garrett, M. (2002). The relation between gender and number agreement processing. *Syntax, 5*, 1–25.
- Baddeley, A. (2003). Working memory and language: An overview. *Journal of Communication Disorders, 36*, 189–208.
- Barber, H. & Carreiras, M. (2003). Integrating gender and number information in Spanish word pairs: An ERP study. *Cortex, 39*, 465–482.
- Barber, H. & Carreiras, M. (2005). Grammatical gender and number agreement in Spanish: An ERP comparison. *Journal of Cognitive Neuroscience, 17*, 137–153.
- Barber, H., Salillas, E., & Carreiras, M. (2004). Gender or genders agreement? In M. Carreiras & C. Clifton (eds.), *The on-line study of sentence comprehension. Eyetracking, ERPs, and beyond* (pp. 309–327). New York: Psychology Press.
- Bock, K. (1995). Producing agreement. *Current Directions in Psychological Science, 4*, 56–61.
- Bonnet, G. (2002). *The assessment of pupils' skills in English in eight European countries*. Paris, France: European Network of Policy Makers for the Evaluation of Education Systems.
- Bosque, I. & Picallo, C. (1996). Postnominal adjectives in Spanish DPs. *Journal of Linguistics, 32*, 349–385.
- Bruhn De Garavito, J. & White, L. (2002). L2 acquisition of Spanish DPs: The status of grammatical features. In A.-T. Pérez-Lerous, & J. Licerias (Eds), *The acquisition of Spanish morphosyntax: The L1/L2 connection* (pp. 153–178). Dordrecht, the Netherlands: Kluwer.
- Carstens, V. (2000). Concord in minimalist theory. *Linguistic Inquiry, 31*, 319–355.
- Carstens, V. (2003). Rethinking complementizer agreement: Agree with a case-checked goal. *Linguistic Inquiry, 34*, 393–412.
- Chomsky, N. (1995). *The minimalist program*. Cambridge, MA: MIT Press.
- Chomsky, N. (2002). *On Nature and Language*. Cambridge, UK: Cambridge University Press.
- Clahsen, H., & Felser, C. (2006). Grammatical processing in language learners. *Applied Psycholinguistics, 27*, 3–42.
- Colé, P., & Ségui, J. (1994). Grammatical incongruency and vocabulary types. *Memory and Cognition, 22*, 387–394.
- Dekydspotter, L. Donaldson, B., Edmonds, A. C., Liljestrands Fultz, A., & Petrush, R. A. (2008). Syntactic and prosodic computations in the resolution of relative clause attachment ambiguity by English–French learners. *Studies in Second Language Acquisition, 30*(4), 453–480.
- De Vicenzi, M., & Di Domenico, E. (1999). A distinction among phi-features: The role of gender and number in the retrieval of pronoun antecedents. *Rivista di Linguistica, 11*, 41–74.
- Eberhard, K. M. (1997). The marked effect of number on subject–verb agreement. *Journal of Memory and Language, 36*, 147–164.

- Ellis, N. (2001). Memory for language. In P. Robinson (Ed.), *Cognition and second language instruction* (pp. 33–68). Cambridge Applied Linguistics. Cambridge, UK: Cambridge University Press.
- Ellis, N., & Sinclair, S. (1996). Working memory in the acquisition of vocabulary and syntax: Putting language in good order. *The Quarterly Journal of Experimental Psychology*, 49A, 234–250.
- Faussart, C., Jakubowicz, C., & Costes, M. (1999). Gender and number processing in spoken French and Spanish. *Rivista di Linguistica*, 11, 75–101.
- Fernández-García, M. (1999). Patterns of gender agreement in the speech of second language learners. In J. Gutiérrez-Rexach, & F. Martínez-Gil (Eds.), *Advances in Hispanic linguistics: Papers from the 2nd Hispanic linguistics symposium* (pp. 3–15). Somerville, MA: Cascadilla Press.
- Finnemann, M. D. (1992). Learning agreement in the noun phrase: The strategies of three first-year Spanish students. *IRAL*, 30, 121–136.
- Foucart, A. (2008). *Syntactic gender processing in French as a first and a second language*. Unpublished doctoral dissertation. Université Aix-Marseille, France and University of Edinburgh, UK.
- Franceschina, F. (2001a). Morphological or syntactic deficits in near-native speakers? An assessment of some current proposals. *Second Language Research*, 17, 213–247.
- Franceschina, F. (2001b). Against an L2 morphological deficit as an explanation for the differences between native and non-native grammars. In S. H. Foster-Cohen, & A. Nizgorodcew (Eds.), *EUROSLA yearbook*, 1 (pp. 143–158). Amsterdam, the Netherlands: John Benjamins.
- Franceschina, F. (2002). Case and phi-feature agreement in advanced L2 Spanish grammars. In S. H. Foster-Cohen, R. Ruthenberg, & M. L. Poschen (Eds.), *EUROSLA yearbook*, 2 (pp. 71–86). Amsterdam, the Netherlands: John Benjamins.
- Franceschina, F. (2005). *Fossilized second language grammars: The acquisition of syntactic gender*. Amsterdam, the Netherlands and Philadelphia, PA: John Benjamins.
- Frenck-Mestre, C., Foucart, A., Carrasco, H., & Herschensohn, J. (2009). Processing of grammatical gender in French as a first and second language: Evidence from ERPs. *EUROSLA Yearbook*, 76–106.
- Frenck-Mestre, C., Osterhout, L., McLaughlin, J., & Foucart, A. (2008). The effect of phonological realization of inflectional morphology on verbal agreement in French: Evidence from ERPs. *Acta Psychologica*, 128, 528–536.
- Gillon-Dowens, M., Barber, H., Vergara, M., & Carreiras, M. (In press). Morpho-syntactic processing in late L2 learners: ERP evidence. *Journal of Cognitive Neuroscience*.
- Grosjean, F., Pommergues, J.-Y., Cornu, E., Guillelmon, D., & Besson, C. (1994). Gender-marking effect in spoken word recognition. *Perception and Psychophysics*, 56, 590–598.
- Guillelmon, D., & Grosjean, F. (2001). The gender marking effect in spoken word recognition: the case of bilinguals. *Memory and Cognition*, 29, 503–511.
- Harris, J. (1991). The exponence of gender in Spanish. *Linguistic Inquiry*, 22, 27–62.
- Hasegawa, M., Carpenter, P. A., & Just, M. A. (2002). An fMRI study of bilingual sentence comprehension and workload. *Neuroimage*, 14, 647–660.
- Havik, E., Roberts, L., Van Hout, R., Schreuder, R., & Haverkort, M. (2009). Processing subject–object ambiguities in the L2: A self-paced reading study with German L2 learners of Dutch. *Language Learning*, 59, 73–112.
- Hawkins, R., & Franceschina, F. (2004). Explaining the acquisition and non-acquisition of determiner-noun gender concord in French and Spanish. In P. Prévost, & J. Paradis (Eds.), *The acquisition of French in different contexts* (pp. 175–206). Amsterdam, the Netherlands and Philadelphia, PA: John Benjamins.
- Hernández, A., Hoffman, J., & Kotz, S. (2007). Age of acquisition modulates neural activity for both regular and irregular syntactic functions. *Neuroimage*, 36, 912–923.
- Herschensohn, J., & Frenck-Mestre, J. (2005). Failed [ugender] features in L2 French: What ERPs reveal about L1 influence. Paper presented at *EUROSLA*. Dubrovnik, Croatia.
- Hopp, H. (2007). *Ultimate attainment at the interfaces in second language acquisition: Grammar and processing*. Groningen, the Netherlands: Grodil Press.
- Igoa, J. M., García-Albea, J. E., & Sánchez-Casas, R. (1999). Gender–number dissociations in sentence production in Spanish. *Rivista di Linguistica*, 11, 163–196.

- Jacobowicz, C., & Faussart, C. (1998). Gender agreement in the processing of spoken French. *Journal of Psycholinguistic Research*, 27, 597–617.
- Jin, F., Āfarli, T. A., & Van Dommelen, W. A. (2007). Endstate L2 competence: Does perception parallel production? In A. Belikova, A. et al. (Eds.), *Proceedings of the 2nd conference on Generative Approaches to Language Acquisition North America* (pp. 200–211). Somerville, MA: Cascadilla Press.
- Juffs, A. (2004). Representation, processing, and working memory in a second language. *Transactions of the Philological Society*, 102, 199–225.
- Juffs, A., & Harrington, M. (1995). Parsing effects in L2 sentence processing: Subject and object asymmetries in wh-extraction. *Studies in Second Language Acquisition*, 17, 483–512.
- Juffs, A., & Harrington, M. (1996). Garden path sentences and error data in second language sentence processing. *Language Learning*, 46, 283–326.
- Just, M., & Carpenter, P. (1992). A capacity theory of comprehension: Individual differences in working memory. *Psychological Review*, 99, 122–149.
- Keating, G. D. (2009). Sensitivity to violations of gender agreement in native and nonnative Spanish: An eye-movement investigation. *Language Learning*, 59, 503–535.
- Kluender, R., & Kutas, M. (1993). Subjacency as a processing phenomenon. *Language and Cognitive Processes*, 8, 573–633.
- Lardiere, D. (2007). *Ultimate attainment in second language acquisition: A case study*. Mahwah, NJ: Lawrence Erlbaum.
- Leung, Y-K. (2005). L2 and L3 initial state: A comparative study of the acquisition of French DPs by Vietnamese monolinguals and Cantonese–English bilinguals. *Bilingualism: Language and Cognition*, 8, 39–61.
- Lew-Williams, C., & Fernald, A. (2007). How first and second language learners use predictive cues in online sentence interpretation in Spanish and English. In H. Caunt-Nulton, S. Kulatilake, & I. Woo (Eds.) *BUCLD 31: Proceedings of the 31st Annual Boston University Conference on Language Development* (pp. 382–393). Somerville, MA: Cascadilla Press.
- Lukatela, G., Kostic, A., Todorovic, D., Carello, C., & Turvey, M. T. (1987). Type and number of violations and the grammatical congruency effect in lexical decision. *Psychological Research* 49, 37–43.
- Mallen, E. (1997). A minimalist approach to concord in noun phrases. *Theoretical Linguistics*, 23, 49–77.
- Marinis, T., Roberts, L., Felser, C., & Clahsen, H. (2005). Gaps in second language sentence processing. *Studies in Second Language Acquisition*, 27, 53–78.
- McCarthy, C. (2008). Morphological variability in the comprehension of agreement: An argument for representation over computation. *Second Language Research*, 24, 459–486.
- Michael, E., & Gollan, T. (2005). Being and becoming bilingual: Individual differences and consequences for language production. In J. F. Kroll, & A. M. B. de Groot (Eds.), *Handbook of bilingualism: Psycholinguistic approaches* (pp. 389–407). New York: Oxford University Press.
- Miyake, A., & Friedman, N. (1998). Individual differences in second language proficiency: Working memory as language aptitude. In A. F. Healy, & L. E. Bourne (Eds.), *Foreign language learning: Psycholinguistic studies on training and retention* (pp. 339–364). Mahwah, NJ: Lawrence Erlbaum.
- Montrul, S., Foote, R., & Perpiñán, S. (2008). Gender agreement in adult second language learners and Spanish heritage speakers: The effects of age and context of acquisition. *Language Learning*, 58, 503–553.
- Morgan-Short, K., & Ullman, M. (2011). The neurocognition of second language. In S. Gass, & A. Mackey (Eds.), *The Routledge Handbook of second language acquisition* (pp. 282–300). London, UK and New York: Routledge.
- Nicol, J., & O'Donnell, M. (1999). Pronominal feature distinctions in English. *Rivista di Linguistica*, 11, 133–163.
- Osaka, M., & Osaka, N. (1992). Language-independent working memory as measured by Japanese and English reading span tests. *Bulletin of the Psychonomic Society*, 30, 287–289.
- Osterhout, L., & Mobley, L. A. (1995). Event-related potentials elicited by failure to agree. *Journal of Memory and Language*, 34, 739–773.

- Osterhout, L., Poliakov, A., Inoue, K., McLaughlin, J., Valentine, G., Pitkanen, I., Frenck-Mestre, C., & Herschensohn, J. (2008). Second-language learning and changes in the brain. *Journal of Neurolinguistics, 21*, 509–521.
- Paradis, M. (2009). *Declarative and procedural determinants of second languages*. Amsterdam, the Netherlands: John Benjamins.
- Pienemann, M., & Keßler, J.-U. (2011). Processability theory. In S. Gass, & A. Mackey (Eds.), *The Routledge handbook of second language acquisition* (pp. 228–246). London, UK and New York: Routledge.
- Rayner, K., & Pollatsek, A. (1989). *The psychology of reading*. New York: Prentice Hall.
- Rossi, S., Gugler, M. F., Friederici, A. D., & Hahne, A. (2006). The impact of proficiency on syntactic second-language processing of German and Italian: Evidence from event-related potentials. *Journal of Cognitive Neuroscience, 18*, 2030–2048.
- Sabourin, L. (2003). Grammatical gender and second language processing. *Groningen Dissertations in Linguistics*, volume 42. Groningen, the Netherlands: CLCG.
- Sabourin, L., & Stowe, L. A. (2008). Second language processing: When are first and second languages processed similarly? *Second Language Research, 24*, 397–430.
- Sagarra, N. (2007a). Working memory and L2 processing of redundant grammatical forms. In Z. Han (Ed.), *Understanding second language process* (pp. 142–159). Clevedon, UK: Multilingual Matters.
- Sagarra, N. (2007b). Online processing of gender agreement in low proficient English–Spanish late bilinguals. In J. Camacho, N. Flores-Ferrán, L. Sánchez, V. Déprez, & M.-J. Cabrera (Eds.), *Current issues in linguistic theory series* (pp. 240–253). Amsterdam, the Netherlands: John Benjamins, 240–253.
- Sagarra, N., & Herschensohn, J. (2008). Processing gender in L2 Spanish. In Chan, H., Jacob, H., & Kapia, E. (Eds.), *32 BUCLD Proceedings, of the 32nd annual Boston University Conference on Language Development* (pp. 427–437). Somerville, MA: Cascadilla Press.
- Sagarra, N., & Herschensohn, J. (2010). The role of proficiency and working memory in gender and number agreement processing in L1 and L2 Spanish. *Lingua, 120*, 2022–2039.
- Sagarra, N., & Herschensohn, J. (2011). Proficiency and animacy effects on L2 gender agreement processes during comprehension. *Language Learning, 61*, 80–116.
- Sawyer, M., & Ranta, L. (2001). Aptitude, individual differences and L2 instruction. In P. Robinson (Ed.), *Cognition and second language instruction* (pp. 319–353). Cambridge, UK: Cambridge University Press.
- Sicuro Correa, L. M., Almeida, D. A., & Sobrino Porto, R. (2004). On the representation of Portuguese gender-inflected words in the mental lexicon. *Brain and Language, 90*, 63–73.
- 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.
- Tunney, R. J. (2005). Sources of confidence in implicit cognition. *Psychonomic Bulletin & Review, 12*, 367–373.
- Viglioccio, G., & Franck, J. (1999). When sex and syntax go hand in hand: Gender agreement in language production. *Journal of Memory and Language, 40*, 455–478.
- Viglioccio, G., Hartsuiker, R. J., Jarema, G., & Kolk, H. H. J. (1996) How many labels on the bottles? *Language and Cognitive Processes, 11*, 407–421.
- Waters, G. S., & Caplan, D. (1996). The measurement of verbal working memory capacity and its relationship to reading comprehension. *Quarterly Journal of Experimental Psychology, 49*, 51–79.
- White, L. (2003). *Second language acquisition and Universal Grammar*. Cambridge, UK: Cambridge University Press.
- White, L., Valenzuela, E., Kozłowska-Macgregor, M., & Leung, Y.-K. (2004). Gender and number agreement in non-native Spanish. *Applied Psycholinguistics, 25*, 105–133.
- Williams, J. N. (2011). Working memory and SLA. In A. Mackey, & S. Gass (Eds.), *The Routledge Handbook of Second Language Acquisition* (pp. 427–441). London, UK and New York: Routledge.
- Xue, G., Dong, Q., Jin, Z., & Chen, Ch. (2004). Mapping of verbal working memory in nonfluent Chinese–English bilinguals with functional MRI. *Neuroimage, 22*, 1–10.

About the authors

Nuria Sagarra is Associate Professor of Linguistics in the Department of Spanish and Portuguese at Rutgers University. Her research straddles the domains of cognitive science, linguistics, and second language acquisition, seeking to identify what factors explain adults' difficulty learning morphosyntax in a foreign language, with the aim of informing linguistic and cognitive models, as well as instructional practices. She investigates these topics using self-paced reading, eyetracking, and more recently, event-related potentials. She is the recipient of a National Science Foundation grant, and has published in notable journals, including *Applied Psycholinguistics*, *Bilingualism: Language and Cognition*, *Language Learning*, *Lingua*, and *Studies in Second Language Acquisition*.

Julia Herschensohn is Professor and Chair of Linguistics at the University of Washington. The range of her work spans areas of generative syntax, second language acquisition theory, and applied linguistics. Her main areas of specialization are theoretical morphosyntax, nonnative language learning, age effects, and bilingual processing. She has published three single-author books and (co)-edited four books dealing with Romance linguistics and second-language acquisition. She has recently published in journals such as *Studies in Second Language Acquisition*, *Second Language Research* and *Journal of French Language Studies*.