Lingua xxx (2010) xxx–xxx

Contents lists available at ScienceDirect

Lingua

journal homepage: www.elsevier.com/locate/lingua

The role of proficiency and working memory in gender and number agreement processing in L1 and L2 Spanish

Nuria Sagarra [a,*], Julia Herschensohn [b]

[a] The Pennsylvania State University, Department of Spanish, Italian and Portuguese, 231 Burrowes Building, University Park, PA 16802, United States
[b] University of Washington, United States

ARTICLE INFO

Article history:
Received 29 August 2009
Received in revised form 11 February 2010
Accepted 11 February 2010
Available online xxx

Keywords:
Grammatical gender
Grammatical number
L2 Spanish
Adjective agreement
Language processing
Working memory

ABSTRACT

Most adult learners cannot attain native competence in a second language (L2). Some approaches maintain that L2 learners cannot access features unavailable in L1 after puberty (Hawkins and Franceschina, 2004) and that they process only superficial structures (Clahsen and Felser, 2006), due to a maturationally constrained critical period for L2 acquisition. In contrast, other approaches hold that late learners may acquire all L2 grammatical features (White et al., 2004) and that they may gain complex structural processing (Hopp, 2007). This study aims to test these models and to investigate the role of language proficiency and working memory on adult L2 acquisition. Beginning and intermediate adult English-speaking learners of Spanish and Spanish monolinguals completed a self-paced reading (online) and a grammaticality judgment task (offline) containing sentences with noun–adjective gender/number agreement/disagreement. The results revealed that all participants were highly accurate in the offline task, but only intermediates and Spanish monolinguals showed sensitivity to gender and number violations in the online task. In addition, intermediates with higher working memory were more accurate on some comprehension questions. These findings indicate that adult learners can develop processing patterns qualitatively similar to those of native speakers and that proficiency and working memory influence their acquisition.

1. Introduction

Universal Grammar (UG) approaches to second language acquisition (L2A) are representational in that they examine the nature of the interlanguage grammar compared to the native grammar of the target language (cf. Herschensohn, 2000; Hawkins, 2001; White, 2003). Computational approaches complement these studies by investigating the implementation of the grammar in real-time production and perception (cf. Doughty and Long, 2003; Dörnyei, 2009). During acquisition of the native language (L1), the child must develop computational procedures in tandem with the growing grammatical competence to facilitate split-second processing in adult speech and comprehension. In both of these realms, recent research has fallen into two general approaches that predict opposing possibilities for ultimate attainment of grammatical features in adult interlanguage grammars: deficit and accessibility approaches. Representational deficit approaches (e.g., Hawkins and Chan, 1997; Franceschina, 2001, 2002, 2005; Hawkins and Franceschina, 2004; Tsimpli and Dimitrakopoulou, 2007) posit that adult second language (L2) learners cannot gain grammatical (uninterpretable) features of functional categories (e.g., tense, gender, number) absent in the L1 due to impaired underlying syntactic competence. Along the same lines, certain computational deficit scholars (e.g., Clahsen and Felser, 2006) argue that adult L2 learners' parsing is quite distinct from that

* Corresponding author. Tel.: +1 814 865 4252; fax: +1 814 863 7944.
E-mail address: sagarra@psu.edu (N. Sagarra).

0024-3841/$ – see front matter © 2010 Elsevier B.V. All rights reserved.
doi:10.1016/j.lingua.2010.02.004

of native speakers (who rely heavily on abstract hierarchical syntactic structure) in that it lacks grammatical detail and favors shallow over deep processing (i.e., lexical, semantic and pragmatic cues as opposed to syntactic cues). In contrast to deficit approaches, representational accessibility approaches such as Full Transfer Full Access (Schwartz and Sprouse, 1996) predict possible accessibility of new L2 grammatical features (e.g., White et al., 2004; Leung, 2005) and link L2 morphological errors to performance/mapping problems in morphological realization. Similarly, computational accessibility approaches (e.g., Hopp, 2007; Frenck-Mestre et al., 2009) hold that late learners are capable of gaining native-like grammatical representation and processing.

The L2A of grammatical gender and number agreement has been a fruitful research arena for both representational and computational models in the last decade (e.g., Friederici et al., 1999; Bartning, 2000; White et al., 2004; Hawkins and Franceschina, 2004; Granfeldt, 2005; Ayoun, 2007; Bordag, 2007; Bordag and Pechmann, 2007; Blom et al., 2008; Acuña-Fariña, 2009), a topic of particular interest when the L1 and L2 grammars differ on this point. Recent studies have looked at L2 learners' abilities to implement determiner phrase (DP) agreement of grammatical gender and number in languages that require it (e.g., French, Spanish, Dutch, German) by learners whose L1 does or does not require concord (e.g., Liceras et al., 2008). Most studies on gender and number agreement in L2 Spanish focus exclusively on production data (Franceschina, 2001; Hawkins and Franceschina, 2004; Bruhn de Garavito and White, 2002; Foote, 2008), some combine perception and production tasks (Franceschina, 2002; McCarthy, 2008; Montrul et al., 2008; White et al., 2004), and only few employ online tasks to measure real-time processing during comprehension (Alarcón, 2006, 2009; Gillon-Dowens et al., 2008, 2009; Keating, 2009; Sagarra, 2007; Tokowicz and MacWhinney, 2005).

The present study includes both offline (grammaticality judgments) and online (non-cumulative self-paced reading) tasks to address the dual nature of linguistic competence: knowledge of grammar and the ability to implement it in real-time computation (Hopp, 2007; Jin et al., 2007; Foucart, 2008; Juffs and Harrington, 1995, 1996; Juffs, 2004; Marinis et al., 2005). Online tasks are necessary to directly evaluate possible deficits in underlying competence, the main factor responsible for limited attainment in older L2 learners, according to certain models (Sato and Felser, 2007). This combination of offline and online measures also allows us to complement previous research entirely based on grammaticality judgments (which are inadequate as single measures, according to Birdsong, 1989; Sorace, 1996), as well as to use multiple tasks to tap linguistic competence (White, 2003). We use these two types of tasks to investigate whether adult English-speaking learners of Spanish can gain gender and number agreement properties for adjectives in Spanish L2, and whether they process gender and number agreement differently.

Offline and online studies with Spanish monolinguals and adult learners of L2 Spanish suggest that processing grammatical gender is cognitively more difficult than number, and that these representational differences influence agreement mechanisms (e.g., Spanish monolinguals: Antón-Méndez et al., 2002; Barber and Carreiras, 2003, 2005; Faussart et al., 1999; Spanish late learners: Bruhn de Garavito and White, 2002; Gillon-Dowens et al., 2004, 2009; McCarthy, 2008; Tokowicz and MacWhinney, 2005; White et al., 2004). Most of these studies focus on native speakers of English—a language with number agreement on nouns (N), on some determiners (D) such as *this*/*these*, *that*/*those* but no number agreement on adjectives (A) and no gender agreement on D or A—and are thus unable to determine whether learners' increased difficulty with gender is due to L1 transfer or processing demands. Gillón-Dowens et al.'s (2008, in press) results show that native speakers of languages with and without D–N number agreement show the same gender–number differences, a result that led them to speculate that such differences are not the result of L1 transfer but of cognitive demands linked to working memory (i.e., an individual's limited capacity to process and store information during complex cognitive tasks, Baddeley, 2003, 2007; Just and Carpenter, 1992). While Gillón and colleagues did not measure working memory, their hypothesis is in line with lexical and syntactic accounts of gender, which claim that gender is cognitively more complex than number (e.g., Antón-Méndez et al., 2002), as well as with Sagarra's (2007) findings that English–Spanish learners with higher working memory span are more sensitive to noun–adjective (N–A) gender agreement violations than those with lower memory span. The inclusion of working memory in our study will shed light into the role of this variable in the processing of determiner phrase (DP) agreement processes.

In sum, the current investigation explores the role of age of acquisition (whether late learners can show native-like processing patterns), language proficiency (whether higher proficiency learners show more native-like patterns than lower proficiency ones), and processing demands (whether late learners process gender agreement differently from number agreement and, if they do, whether the increased difficulty of one over the other is influenced by working memory). Despite much research on adult L2A of gender and number agreement, the relative importance and interaction of these factors is still unclear and online studies examining these issues with N–A agreement are scarce. To address these questions, we examine how Spanish monolinguals and beginning and intermediate English–Spanish learners process N–A gender and number agreement, by means of online and offline techniques (non-cumulative self-paced reading and grammaticality judgments).

2. Background

2.1. Gender and number agreement in Spanish DPs

In Spanish, nouns carry number and gender features (Carroll, 1989; Dewaele and Véronique, 2000, 2001). In Spanish, nouns are marked as either masculine or feminine, with a natural criterion for assigning gender to animate nouns, based on biological sex, and an arbitrary criterion for inanimate nouns (Corbett, 1991). Gender is generally realized

Please cite this article in press as: Sagarra, N., Herschensohn, J., The role of proficiency and working memory in gender and number agreement processing in L1 and L2 Spanish. Lingua (2010), doi:10.1016/j.lingua.2010.02.004

morphosyntactically with the inflectional morphemes /-o/ for masculine (M) nouns (*libro* 'book-M') and /-a/ for feminine (F) nouns (*blusa* 'blouse-F') (transparent gender) (Green, 1988). Some exceptions include M nouns ending in /-a/ (*día* 'day-M'), F nouns ending in /-o/ (*radio* 'radio-F'), and M and F nouns ending in /-o/ and /-a/ (*el/la testigo* 'the-M/F witness-M/F'), /-e/ (*puente* 'bridge-M,' *fuente* 'fountain-F'), or consonant (*cartel* 'poster-M,' *pared* 'wall-F'). In addition to being [+/− feminine], Spanish nouns are [+/− plural]. Plurality (P) signals quantity of the semantic referent and is formed by adding /-s/ to singular (S) count nouns (*libros* 'books-MP'), /-es/ to nouns ending in a consonant (*paredes* 'walls-FP'), and seldom /Ø/ (*atlas* 'atlas-MP').

According to Corbett (1991), the existence of gender and number in a language is indicated by agreement. In Spanish, determiners, adjectives, pronouns and past participles vary their gender and number according to the noun to which they refer, a procedure called agreement or concord (Carroll, 1999; Zagona, 2002) (*el libro blanco* 'the-MS white-MS book-MS,' *los libros blancos* 'the-MP white-MP books-MP,' *la blusa blanca* 'the-FS white-FS blouses-FS,' *las blusas blancas* 'the-FP white-FP blouses-FP'). Thus, agreement is a syntactic mechanism that reflects both semantic features (biological sex of animate referents and quantity of animate and inanimate referents) and grammatical co-referential relations (through morpho-phonological markers) necessary to comprehend sentences. In contrast, English has no gender marking on nouns (nouns like *waiter*/*waitress* bear no gender feature and denote a female in the same way that the proper name Alice conventionally does), limited gender–number agreement between nouns and determiners (*this-these*, *that-those*) and no gender–number agreement between nouns and adjectives. Therefore, English–Spanish learners must gain the gender features for inanimate nouns and agreement on determiners and adjectives in the L2.

Gender concord is not simply a redundant agreement phenomenon in gendered languages, but rather is a facilitating factor for native speakers to access lexical items and concord relationships with greater speed, a factor known as congruency. The gender congruency effect refers to the facilitation that gender-marked items other than nouns have on processing: congruent gender concord speeds up and incongruent concord slows down processing of the noun and its modifiers. The gender congruency effect is widely documented with monolinguals both with behavioral data (Antón-Méndez et al., 2002; Cacciari and Padovani, 2007; Colé and Ségui, 1994; Schriefers, 1993; Schriefers and Teruel, 2000; Vainio et al., 2003; but see Miozzo et al., 2000, for behavioral evidence that the gender congruency effect is obtained in Romance but not Germanic languages) and neurocognitive data (Barber and Carreiras, 2005; Foucart, 2008; Gillon-Dowens et al., 2004; Gunter et al., 2000; Hagoort and Brown, 1999). However, it is not clear whether it is also present in adult L2 learners of ungendered L1s or not.

2.2. Syntactic representation of L2 gender and number agreement

UG approaches to L2A presuppose that learners build an interlanguage grammar using at least the available input and the innate language predisposition (UG) that guides L1A and defines universal properties of languages in general. Given a minimalist framework (Chomsky, 2000, 2001, 2002), grammatical competence in a mature native speaker comprises lexical categories (e.g., verbs, nouns) and functional categories (e.g., auxiliaries, determiners) instantiating grammatical features (e.g., gender, number) that can be semantically interpretable or grammatical (uninterpretable). Interpretable features (e.g., [+/− plural]) are semantically necessary and may be linked to grammatical uninterpretable [*u*features] (e.g., [*u*number]) on syntactically linked items such as subject–verb agreement or noun adjective concord.[1] In minimalist terms, the operation Agree matches interpretable and *u*features, deleting the latter. In the case of gender and number agreement, [+/− feminine], [+/− plural] interpretable features of N check and delete the grammatical [*u*gender], [*u*number] Det, Adj features in close enough proximity, hypothetically in nominal projections above N (Carstens, 2000, 2003). The question of whether adult L2 learners can acquire *u*features absent in their L1 differentiates deficit from accessibility approaches to adult L2A.

2.2.1. Deficit approaches

As mentioned earlier, deficit approaches posit that *u*features can only be acquired during the critical period (cf. Tsimpli and Roussou, 1991). These L1 *u*features remain available through adulthood, but new L2 values that differ from L1 cannot be acquired (unlike interpretable features which are available through adulthood). Thus, adult Anglophone learners of L2 Romance languages like Spanish can gain interpretable number on nouns, interpretable gender on L2 animate nouns, and (presumably) [*u*number] for number on D (given English agreement such as *this-these*), but not [*u*gender] on D, A because it is not instantiated in English. Hawkins, Franceschina and others have argued that several elicited production and reaction time (RT) studies support this claim. In terms of processing, the Shallow Structure Hypothesis (Clahsen and Felser, 2006) holds that adult L2 learners' parsing relies less on morphosyntactic cues than on lexical ones (thus favoring "shallow" over deep processing). Silva and Clahsen (2008) demonstrate a difference between native Anglophones' sensitivity to priming of inflectional and derivational morphology and L2 learners' lack of sensitivity to the same (cf. Neubauer and Clahsen, 2009).

Elicited production studies with English–French (Hawkins, 1998) and English–Spanish (Franceschina, 2001) late learners reveal accuracy rates on gender and number agreement on D and A above 90% but more gender than number agreement errors and overuse of the masculine form, facts they claim corroborate the establishment of gender on the basis of phonology and the consequent unavailability of [*u*gender]. Nevertheless, the learners did not produce perfect number concord with D,

[1] We adopt the accepted label "interpretable," although strictly speaking a given feature may not always impart the same "interpretation" depending on the context (cf. Carstens, 2000 on gender).

Please cite this article in press as: Sagarra, N., Herschensohn, J., The role of proficiency and working memory in gender and number agreement processing in L1 and L2 Spanish. Lingua (2010), doi:10.1016/j.lingua.2010.02.004

which the transfer of L1 [*u*number] would predict. Later on, Franceschina (2001) and Hawkins and Franceschina (2004) reported that adult L2 learners of Spanish of a gendered (Italian) or ungendered (English) L1 only made few errors with gender or number agreement of D and A, and, of the errors they made, Italian speakers made fewer gender agreement errors than English speakers, favoring representational deficit approaches. Franceschina (2002) explored the same question using perception tasks and concluded that all participants were accurate at interpreting number agreement and that Spanish monolinguals were better than the L1 ungendered group but the same as the L1 gendered group, bolstering the representational deficit claim that [*u*gender] features are only acquirable if present in the L1.

These studies have used offline data that can fall short at assessing underlying competence (Jin et al., 2007; Juffs and Harrington, 1995, 1996; Marinis et al., 2005; Hopp, 2007). As Sato and Felser (2007) note, online comprehension studies can help to empirically differentiate representational and computational accounts to adult L2A. Monolinguals with gender concord have demonstrated congruency/incongruency effects by showing faster recognition of nouns with congruent than incongruent gender marking on determiners and adjectives (for French and Spanish monolinguals, see Antón-Méndez et al., 2002; Grosjean et al., 1994; Jacubowicz and Faussart, 1998). For bilinguals, Guillelmon and Grosjean's (2001) auditory word repetition task revealed that monolinguals and early (5 year old) but not late (24 year old) English–French bilinguals demonstrated congruency/incongruency effects (see Foote, 2008, for similar findings with early and late English–Spanish learners). These data suggest that early bilinguals behave more similarly to monolinguals than their late counterparts.

2.2.2. Feature accessibility approaches

Claims in support of the deficit hypothesis have been challenged by studies suggesting that late bilinguals can achieve native-like grammatical competence in the L2. These approaches often maintain that adult L2 learners initially transfer L1 morphosyntactic settings (cf. Schwartz and Sprouse, 1996) but eventually may gain L2 *u*features through gradual restructuring (e.g., White et al., 2004; Leung, 2005). There is no critical period functional deficit to examine since L2A is similar for children and adults. Syntactic competence is not directly reflected by mastery of morphological inflection, because mistakes relate to matching difficulties between syntactic terminal nodes and surface morphology. This line of research attributes L2 morphology errors to problems with morphological mapping rather than impaired underlying syntactic competence, a tradition that draws on missing surface inflection (Haznedar and Schwartz, 1997; Lardiere, 1998, 2000; Prévost and White, 2000). Thus, adult English speakers of L2 Romance languages like Spanish initially transfer L1 properties and may reset nominal features to the L2 values, eventually gaining both [*u*gender] and [*u*number] on D and A (Gess and Herschensohn, 2001; Herschensohn, 2001; Bruhn de Garavito and White, 2002; Prévost, 2004; White et al., 2004; Herschensohn and Arteaga-Capen, 2007). While number is morphologically marked on English nouns and vaguely on English determiners (e.g., *this-these*), there is no [*u*num] feature on English adjectives, which show no concord in gender or number.

2.3. Processing gender and number concord

2.3.1. L2 Spanish noun concord effects

Offline studies indicate that adult learners of different proficiency levels are highly accurate at perceiving and producing D–N and N–A gender and number agreement, independently of whether or not these features have been instantiated in their L1 (e.g., Fernández-García, 1999; Bruhn de Garavito and White, 2002; McCarthy, 2008; White et al., 2004, for perception and production data with French–Spanish and English–Spanish learners). These studies also reveal that beginning and intermediate learners of gendered and ungendered L1s are more accurate at number than gender agreement. These findings are in line with results from L2 processing studies.

For example, lexical decision and eyetracking studies reveal that native speakers and advanced late learners are sensitive to both grammatical gender disagreement (e.g., monolinguals; Antón-Méndez et al., 2002; Colé et al., 2003; advanced late learners: Alarcón, 2009; Foucart, 2008; Herschensohn and Frenck-Mestre, 2005; Lew-Williams and Fernald, 2007) and number disagreement in noun phrases (e.g., monolinguals: Hartsuiker et al., 2001; Antón-Méndez et al., 2002). Such sensitivity has also been reported in electrophysiological studies and neuroimaging studies with respect to both grammatical gender disagreement (e.g., monolinguals: Barber and Carreiras, 2005; Carreiras et al., 2010; Hammer et al., 2007; advanced late learners: Gillon-Dowens et al., 2009; Rossi et al., 2006; Sabourin et al., 2006; Tokowicz and MacWhinney, 2005) and number disagreement (e.g., monolinguals: Carreiras et al., 2010; Barber and Carreiras, 2005; advanced late learners: Rossi et al., 2006; Osterhout et al., 2008; Gillon-Dowens et al., 2009).

It is important to note that advanced learners' sensitivity to gender/number agreement violations is absent in low proficient learners, suggesting that sensitivity to such discord partly depends on their L2 proficiency (e.g., RT data: Alarcón, 2009; Keating, 2009; Sagarra, 2007; event-related potential (ERP) data: Osterhout et al., 2008; Tokowicz and MacWhinney, 2005). For example, the latter study revealed that beginning late learners were sensitive (P600 effect) to violations of L2 features that are formed similarly in the L1 but not of features that are nonexistent or formed differently in the L1. This study also showed a clear divergence between implicit (brain responses to grammatical violations) and explicit measures (grammaticality judgments) of L2 learning. In summary, offline and online studies have argued that three important factors to learners' sensitivity to grammatical concord of gender and number are age of acquisition, L1 feature inventory, and L2 proficiency level.

Please cite this article in press as: Sagarra, N., Herschensohn, J., The role of proficiency and working memory in gender and number agreement processing in L1 and L2 Spanish. Lingua (2010), doi:10.1016/j.lingua.2010.02.004

2.3.2. Cognitive differences in gender and number agreement processing

Empirical evidence on whether gender and number agreement are represented differently at the lexical level and, if they are, how and when these representational differences affect agreement mechanisms during L1 and L2 syntactic processing is mixed. Some monolingual studies claim that gender and number are processed similarly. For example, Colé and Segui (1994) and Lukatela et al. (1987) found no differences between gender and number agreement in grammatical priming with word pairs. Osterhout and Mobley's (1995) ERP data are compatible with these findings: gender and number agreement violations produced the same effects (i.e., P600 effect, related to grammatical anomalies), and both were different from semantic violations (semantic violations produced a N400 effect, which is linked to semantic anomalies).

In contrast, other studies reveal that monolinguals have more difficulty processing gender than number agreement (De Vincenzi, 1999; De Vincenzi and Di Domenico, 1999). For example, speech-error studies show that monolinguals make more grammatical gender than number errors (Antón-Méndez, 1996; Vigliocco et al., 1996; Nicol and O'Donnell, 1999; Antón-Méndez et al., 2002). In turn, lexical decision studies reveal longer RTs when words disagree in gender than in number (Faussart et al., 1999). Finally, Spanish ERP data show that gender agreement violations produce longer latencies than number agreement violations (Barber and Carreiras, 2003, 2005). The results of these studies support both lexical and syntactic accounts to gender. If gender is a stem inherent feature that is accessed from the full word form and number is a morphological feature that combines with the stem of the word (lexical accounts to gender) (e.g., Domínguez et al., 1999; Igoa et al., 1999), gender disagreement should be more cognitively taxing than number disagreement because gender failure forces the processor to go back to the lexical identification stage in order to check if the right entry had been chosen vs. number failure that only requires the system to check the final processes of syntactic recognition without having to return to the initial processes of lexical access. Similarly, even if gender and number are affixal (e.g., Sereno and Jongman, 1997; Sicuro Correa et al., 2004), gender disagreement should be more cognitively demanding than number disagreement because irregular grammatical items are more difficult to process than regular grammatical ones (Hernández et al., 2007) (as mentioned earlier, number marking is consistently marked with *–s* or *–es*, whereas gender marking is only partially predictable because many nouns violate the rule of using *–o* for masculine and *–a* for feminine).

2.3.3. Working memory

In line with these studies, research examining how native speakers of gendered and ungendered languages process and produce gender and number agreement in L2 Spanish suggests that gender is cognitively more difficult than number. Offline studies show that beginning and intermediate French–Spanish and English–Spanish learners are more accurate at perceiving and producing number than gender agreement (Bruhn de Garavito and White, 2002; Franceschina, 2002; McCarthy, 2008; White et al., 2004). As for online studies, Gillon-Dowens et al.'s (2008, 2009) ERP data reveal that highly proficient English–Spanish and Chinese–Spanish learners behave like Spanish monolinguals (biphasic LAN-P600 pattern indicating the presence of both automatic and repair processes) for D–N number disagreement but the learners lack the P600 effect for gender disagreement. Because these gender–number differences are present in learners of an L1 with (English) and without (Chinese) D–N number agreement, they conclude that such differences are not due to L1 transfer but to cognitive factors, such as individual differences in working memory. Their interpretation of the findings follows previous research suggesting a link between the LAN effect and some kind of verbal working memory (Kluender and Kutas, 1993).

Working memory, the cognitive resources needed to temporarily store and process information during complex cognitive actions (King and Just, 1991; Geva and Ryan, 1993; Ellis and Sinclair, 1996; Baddeley, 2003, 2007), is limited (Just and Carpenter, 1992; Vos et al., 2001), and tasks that deplete a person's working memory capacity can result in less storage and slower processing (cf. MacDonald et al., 1992; Fiebach et al., 2002; Mackey et al., 2002). Current models of L2 acquisition agree that learning a language as an adult is a daunting task that consumes a great amount of cognitive resources (Hasegawa et al., 2002; Walter, 2004) and that these processing demands affect knowledge and implementation of linguistic information, such as gender agreement (e.g., Lardiere, 2007; Hopp, 2007). There is mounting evidence that working memory is associated with L2 lexical retrieval and morphosyntactic processing (e.g., Havik et al., 2009; Michael and Gollan, 2005; Miyake and Friedman, 1998; Miyake et al., 1994; Sagarra, 2007) and evidence has begun to accrue that working memory is linked to L2 gender agreement processing even at low proficiency levels. For example, Sagarra (2007) found that English–Spanish beginning learners were insensitive to gender agreement violations but that those with higher working memory capacity showed sensitivity to the violations. Heeding previous empirical evidence indicating that late learners of ungendered L1s show (a) greater difficulty with gender than number agreement regardless of the presence or absence of D–N number agreement in their native language (Gillon-Dowens et al., 2008, 2009) and (b) increased sensitivity to gender agreement violations with higher working memory (Sagarra, 2007), we investigate whether working memory affects the processing of gender and number agreement by learners of different proficiency levels.

3. The study

3.1. Research questions

As shown in the literature review, research on adult acquisition of gender and number agreement in L2 Spanish is abundant but it is still unclear whether adults can gain grammatical features absent in their L1 and whether greater difficulties in processing gender than number agreement are modulated by cognitive demands. The present study aims to

Please cite this article in press as: Sagarra, N., Herschensohn, J., The role of proficiency and working memory in gender and number agreement processing in L1 and L2 Spanish. Lingua (2010), doi:10.1016/j.lingua.2010.02.004

shed light on the role of age of acquisition (whether late learners can show native-like processing patterns), language proficiency (whether higher proficiency learners show more native-like patterns than lower proficiency ones), and processing demands (whether working memory influences the processing of gender and number agreement and whether it modulates possible differences between gender and number agreement) on the processing of N–A grammatical gender and number agreement by Spanish monolinguals and beginning and intermediate English–Spanish learners. The study uses online and offline techniques (non-cumulative self-paced reading and grammaticality judgments, respectively) to explore these issues. We investigate N–A rather than D–N agreement because research on the former is scarce (cf. Hernández-Pina, 1984, on child acquisition) and we examine beginning and intermediate learners because studies on the representation and computation of L2 gender and number agreement in low proficiency levels are few.

Lastly, we combine online (non-cumulative self-paced reading) and offline (grammaticality judgments) methodologies because linguistic competence entails both grammatical knowledge and implicit grammatical computation: comprehension and production data cannot speak to grammatical representation—since it is impossible to know where competence ends and performance begins—whereas behavioral data permit the direct comparison of L2 learners and monolinguals to ascertain computational procedures. Differences between online and offline data obtained in previous studies (De Mulder, 2006; Montrul et al., 2008) confirm the need to include both techniques within the same sample pool. Furthermore, learners' superior performance on self-paced over timed tasks (Sabourin, 2003) and on written over oral tasks (Montrul et al., 2008) led us to decide to use self-paced and written tasks (the higher processing demands of timed and oral tasks could hinder sensitivity to adjectival morphology).

The specific research questions of the study are

1. *Based on online and offline data, do Spanish monolinguals, and intermediate and beginning L2 learners show gender and number congruency effects*? We predict that beginning learners will not be sensitive to gender or number agreement violations (no differences between agreement and disagreement conditions) but that intermediate learners and Spanish monolinguals will show sensitivity to both types of agreement violations (longer RTs on the self-paced reading task and lower accuracy on the grammaticality judgment task). This prediction follows accessibility approaches to adult L2A, which assume that gender features can be acquired in adulthood independently of the L1 (e.g., representational models: White et al., 2004; Leung, 2005; computational accounts: Hopp, 2007; Foucart and Frenck-Mestre, 2004). The lack of gender congruency effects in beginners may relate to transfer of L1 values, to less optimal lexical mastery, or to cognitive factors such as individual differences in working memory. The latter is examined in the second research question.
2. *Based on online and offline data, do Spanish monolinguals, and intermediate and beginning L2 learners process gender agreement differently from number agreement? And if they do, is the increased difficulty of one over the other influenced by working memory?* We predict that beginners will show no differences, but that intermediates and Spanish monolinguals will spend more time processing gender agreement violations than number agreement violations. This hypothesis is based on previous studies with Spanish monolinguals (Antón-Méndez et al., 2002; Barber and Carreiras, 2003, 2005; Faussart et al., 1999) and Spanish late bilinguals (Bruhn de Garavito and White, 2002; Gillon-Dowens et al., 2008, 2009; McCarthy, 2008; Tokowicz and MacWhinney, 2005; White et al., 2004) and is in line with lexical and syntactic accounts of gender, which claim that computation of gender agreement is cognitively more demanding than number agreement. Furthermore, we expect L2 learners with higher working memory to be more sensitive to N–A agreement violations, particularly gender ones, following research showing that working memory modulates L2 lexical retrieval and morphosyntactic processing (e.g., Havik et al., 2009; Michael and Gollan, 2005; Miyake and Friedman, 1998), as well as processing of grammatical gender agreement in L2 Spanish (Sagarra, 2007).

3.2. Participants

A pool of 196 university students received 10 euros (63 Spanish monolinguals) or extra credit (69 beginning and 64 intermediate L2 learners) for participating in the study. To be included, participants could not have spent more than one month in a foreign language country or a bilingual Spanish province, and they needed to score above 60% in the comprehension questions of the self-paced reading task.[2] The Spanish monolingual group consisted of Spaniards with university education in areas other than linguistics. A language background questionnaire indicated that they were born and raised in Andalucia, they did not speak any other foreign languages, and they did not study English outside the required courses in middle school and high school. Studies indicating that Spanish highschoolers have one of the lowest L2 English proficiency levels of the European Union (Bonnet, 2002; CEDEFOP, 2004) and the monolinguals' self-ratings of their functional proficiency in English as low suggest that their proficiency level in this language was too low to have any effect on L1 processing.

The L2 groups comprised English native speakers enrolled in a third-semester (beginners) or seventh- or eighth-semester (intermediates) Spanish course at a North American university. A language background questionnaire for L2 learners determined that they began learning Spanish post-puberty, once their L1 syntax had been established (e.g., Guasti, 2002; Crain and Lillo-Martin, 1999; Hawkins and Franceschina, 2004; Herschensohn, 2007) and that they had no knowledge of

[2] While a 60% cutoff may not seem conservative at first glance, it is important to take into consideration (a) that many psycholinguistic studies follow the "above chance cutoff" (i.e., 51%), (b) that even the lowest mean was actually high (8.32 over 10 for the 69 beginners: see Table 2), and (c) that we included a comprehension question after every sentence (whereas many studies include comprehension questions randomly).

Please cite this article in press as: Sagarra, N., Herschensohn, J., The role of proficiency and working memory in gender and number agreement processing in L1 and L2 Spanish. Lingua (2010), doi:10.1016/j.lingua.2010.02.004

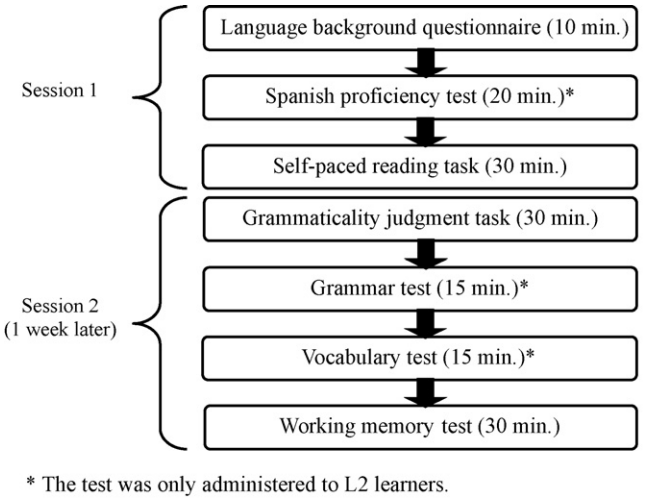

Fig. 1. Materials and procedure.

other languages. In addition to the language history questionnaire, L2 learners completed a Spanish proficiency test, with those scoring below or above 3 standard deviations from the mean being excluded from the study. Finally, L2 learners needed to obtain a perfect score in the diagnostic gender/number agreement and vocabulary tests to ensure that lack of morphological or semantic knowledge of the target nouns and adjectives did not produce longer latencies that could bias our results.

3.3. Materials and procedure

As shown in Fig. 1, participants performed seven tests in two sessions one week apart. The language background questionnaire, the Spanish proficiency test, and the self-paced reading task were conducted in the first session, and the grammaticality judgment task, the Spanish vocabulary test, the Spanish grammar test, and the working memory test in the second session. The Spanish vocabulary and grammar tests were administered later to ensure that the explicitness of these two tests did not bias the outcome of the self-paced reading or the grammaticality judgment task. Spanish monolinguals completed all the tests except the Spanish proficiency test, the Spanish vocabulary test, and the Spanish grammar test.

3.3.1. Language background questionnaire

Two versions of the language background questionnaire were administered: one for the L2 learners and one for the Spanish monolinguals. The questionnaire revealed that none of the participants had lived in a foreign-speaking country for more than one month and that the L2 learners received a similar amount of L2 instruction (three semesters for beginners and eight for intermediates). In addition, the questionnaire for Spanish monolinguals included self-ratings of English proficiency in each of the four skills using a Likert scale (1 = minimum ability; 10 = native proficiency) (see Bonnet, 2002, for evidence that self evaluations correlate with L2 proficiency). The questionnaire showed that Spanish monolinguals had a low functional proficiency in English (the means for the four skills ranged between 3.13 and 4.21 out of 10).

3.3.2. Spanish proficiency test

After the language background questionnaire, the L2 learners completed the Spanish proficiency test, which consisted of the grammar section of the *Diploma de Español como Lengua Extranjera* for intermediate learners. In this test, learners have to read isolated sentences (30 points) and a passage in Spanish (20 points) and fill in the blanks with one of four options provided per item. Correct answers received 1 point and incorrect answers 0 points. An independent-samples *t* test indicated that the intermediates ($M = 20.65$, $SD = 3.00$) were more proficient in Spanish than the beginners ($M = 7.04$, $SD = 2.89$): $t\ (125) = -22.667$, $p < .01$ (Levene's $F = .000$, $p > .05$).

3.3.3. Self-paced reading and grammaticality judgment tasks

For each task, participants read four practice sentences (half grammatical, half ungrammatical), 30 experimental sentences (10 per condition), and 80 filler sentences (70 well-formed, 10 with gender disagreement with animate nouns reported in another study). The conditions for the experimental sentences were: (1) gender and number agreement, (2) gender agreement violation (feminine for masculine), and (3) number agreement violation (plural for singular). Sentences combining gender and number agreement violations were excluded because Rossi et al. (2006) found that late L2 learners treat these errors as semantic rather than grammatical violations, regardless of proficiency level.

Please cite this article in press as: Sagarra, N., Herschensohn, J., The role of proficiency and working memory in gender and number agreement processing in L1 and L2 Spanish. Lingua (2010), doi:10.1016/j.lingua.2010.02.004

(1)	*El*	*ingeniero*	*presenta*	*el*	*prototipo*	*famoso*	*en*	*la*	*conferencia.*
	The	engineer	presents	the	prototype-MS	famous-MS	at	the	conference.
(2)	*El	*ingeniero*	*presenta*	*el*	*prototipo*	*famosa*	*en*	*la*	*conferencia.*
	*The	engineer	presents	the	prototype-MS	famous-FS	at	the	conference.
(3)	*El	*ingeniero*	*presenta*	*el*	*prototipo*	*famosos*	*en*	*la*	*conferencia.*
	*The	engineer	presents	the	prototype-MS	famous-MP	at	the	conference.

The following measures were taken to divert participants from discovering the goal of the study and focusing on the target structure. First, approximately two thirds of the sentences were fillers, and all sentences (experimental and fillers) were similar in length (9–15 words), syntactic structure, and lexicon (cognates and vocabulary typical of basic Spanish textbooks; see Appendix A for complete list of target nouns and adjectives). Second, sentences were randomized to avoid having two experimental sentences appear consecutively. Third, participants were not told that they would encounter ungrammatical sentences in the self-paced reading task. Finally, the same nouns and adjectives were used in both tasks for comparability purposes but in different noun–adjective combinations to avoid practice effects.

The target nouns consisted of 2–4 syllable masculine singular countable nouns with transparent gender and the target adjectives were 2–4 syllables long and also had transparent gender endings. The NP1 nouns were masculine singular ending in /-o/ to avoid possible gender or number biases on the NP2 nouns. Gender-inflected pairs such as *puerto-puerta* 'seaport$_{masc}$, door $_{fem}$' and *fruto-fruta* 'fruit$_{masc, fem}$' were excluded to avoid conceptual and grammatical relationships with the feminine noun (the processing of the masculine noun could activate the feminine noun resulting in shorter RTs at the feminine adjective that could hinder gender congruency effects). Also, Spanish uncountable nouns and Spanish countable nouns that are uncountable in English were not considered either, in order to ensure trial homogeneity (all nouns were countable) and avoid comprehension or L1 interference issues related to mass nouns lacking the number feature.

We used the unmarked form of gender (masculine: Harris, 1991; transparent: Antón-Méndez, 1999) and number (singular: Eberhard, 1997) due to several reasons. First, we wanted to develop a solid baseline to explore gender/number concord/discord that excludes variables irrelevant to our research (e.g., masculine/feminine, transparent/opaque, singular/plural, marked/unmarked), which would significantly increase the number of sentences, making the experiment logistically unfeasible. Second, previous research indicates that a noun's gender or number does not influence RTs of congruent/incongruent D/A in native speakers and late learners of Romance languages (e.g., Alarcón, 2009; Antón-Méndez et al., 2002; De Mulder, 2006; Keating, 2009). Still, to ensure that longer RTs at the adjective were due to sensitivity to disagreement rather than reaction to the marked form of gender and number, we ran statistical analyses on the fillers with N–A gender and number agreement based on feminine singular (FS) (k = 10), masculine plural (MP) (k = 10), and feminine plural (FP) nouns (k = 10) with transparent gender (see Results for more information). In the non-default form (F, P) latencies were a function of discord, not non-default status. Another consideration was that we are examining how processing of gender and number agreement takes place at early stages of acquisition, so we focused on the area where grammaticalization might first appear, in the transparent unmarked case. McCarthy (2008) has shown that L2 learners (manipulating Spanish gender and number) are more accurate for default (M, S) than marked (F, P) forms in both comprehension and production, a tendency suggesting that grammatical representation may be established earlier in default morphology. An anonymous reviewer questions our use of the default forms. Assuming that representation and computation necessarily interface (i.e., underlying grammatical competence and online processing), we seek to see this in the unmarked case since it should be diagnostic, the first place we predict to see evidence of grammatical sensitivity. We would also predict less sensitivity in the marked F, P if we were to run a similar test, but that is beyond the scope of the present study. Finally, we concentrated on contiguous N–A concord/discord, following previous studies indicating that longer structural distance between the noun and the modifying adjective results in decreased sensitivity to grammatical gender discord (L2 French: Myles, 1995; L2 Spanish: Keating, 2009; see Almor et al., 2001, for similar findings with subject–verb agreement violations in English monolinguals).

For the moving window task, participants read Spanish sentences silently on a computer screen, word-by-word, and answered yes-no comprehension questions (half with yes, half with no target answers) after each sentence. As illustrated in Fig. 2, each sentence began with a 500-ms fixation marker '+' that appeared at the center of the screen, followed by dashes

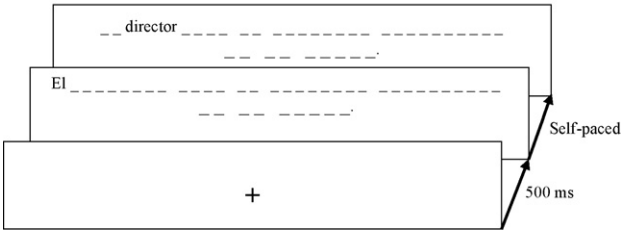

Fig. 2. Example of a trial for the self-paced reading task.

Please cite this article in press as: Sagarra, N., Herschensohn, J., The role of proficiency and working memory in gender and number agreement processing in L1 and L2 Spanish. Lingua (2010), doi:10.1016/j.lingua.2010.02.004

(dashes helped make reading more natural). Each dash represented a letter, and words were separated with spaces to present the visible characters normally available during natural reading. When participants pressed the space bar key, the first word of the sentence appeared, replacing the dashes for that word. By pressing the space bar key each time, the previous word was removed and the subsequent word was revealed. Pressing the key on the last word of the sentence prompted a comprehension question about the sentence they had just read, and participants indicated the answer by pressing a "yes" or a "no" button. Comprehension questions were included to ensure that low proficiency learners attended to the meaning of the sentences and that longer RTs at the adjective were due to sensitivity to agreement violations instead of lack of understanding. To avoid biasing the participants' attention to the adjective, the questions excluded the adjective and did not evaluate their knowledge of gender/number marking or agreement. Examples of yes and no questions for sentences (1)–(3) above are: *¿El ingeniero presenta el prototipo*? "Does the engineer present the prototype?" (Answer: Yes) and *¿El físico presenta el prototipo*? "Does the physicist present the prototype?" (Answer: No). The instructions emphasized the importance of accuracy in responding to the comprehension questions, and participants received 1 point for correct responses and 0 for incorrect ones.

The grammaticality judgment task was administered one week after the self-paced reading task and required participants to classify sentences as correct or incorrect, to identify the source of the error in incorrect sentences with a circle on the incorrect word(s), and to rate how confident they were about their answer in a 5-point Likert scale. Only confidence ratings of sentences with correct grammaticality judgments were used for statistical analyses, and the ratings were included because previous studies have shown that multiple judgment tasks are more informative than single judgment tasks (Schütze, 1996) and that confidence ratings made on a continuous scale are closely related to grammaticality (Tunney, 2005). Furthermore, the combination of grammaticality and confidence judgments allowed us to differentiate between performance (accuracy score and source of the error score) and awareness (confidence score) (Tunney and Shanks, 2003).

3.3.4. Spanish vocabulary and grammar tests

Following the grammaticality judgment test, L2 learners completed a vocabulary and a grammar test to control for familiarity with the meaning of the target nouns and adjectives and the target structure. For the vocabulary test, participants matched the nouns and adjectives in Spanish to their translation in English. For the grammar test, they identified the gender and number of a series of nouns that were counterbalanced for gender and number. Masculine singular nouns were based on the nouns used for the experimental sentences of the self-paced and grammaticality judgment task, and the rest corresponded to the nouns of filler sentences containing correct gender and number agreement with feminine singular nouns, masculine plural nouns and feminine plural nouns.

3.3.5. Working memory test

Following Waters and Caplan's (1996) reading span test, participants read sets of plausible and implausible sentences on a computer screen at a fast pace, one by one, indicated whether that sentence was plausible by pressing a 'yes' or a 'no' button, and wrote down the final word of each sentence at the end of each set of sentences. The test was given in the participants' L1 because working memory seems to be language-independent (Osaka and Osaka, 1992; Xue et al., 2004) and because deficits in L2 knowledge could affect the results of a test conducted in the target language.

3.3.6. Scoring

Participants received 1 point per correct answer and 0 per incorrect answer for all the measures except the mean RTs of the self-paced reading task and the confidence ratings of the grammaticality judgment task. The self-paced reading task produced two scores: mean RTs at the adjectives and accuracy on comprehension questions. Mean RTs, the mean of all word RTs within a condition, corresponded to the time spent between the appearance of a word on the screen and the press of a spacebar key. Word RTs faster than 200 ms and slower than 2000 ms were excluded because Anglophone college students need between 225 and 300 ms to process single words (Rayner and Pollatsek, 1989). As mentioned earlier, accuracy on comprehension questions was based on a binary score of 1 point for correct answers and 0 for incorrect answers and participants needed to be accurate above 60% to be included in the study. Only sentences with correct responses to the comprehension questions were taken into consideration for statistical analyses to minimize the possibility of reading latencies due to lack of understanding. The grammaticality judgment task also generated two scores: accuracy on identifying sentences as correct or incorrect and a confidence rating. For the accuracy score, participants received 1 point for identifying correct sentences as correct or identifying incorrect ones as incorrect and identifying the error accurately. The confidence rating score was based on a 5-point continuous score ranging from 1 (not sure at all) to 5 (completely sure). Finally, for the working memory test, participants received one point per sentence with both correct plausibility judgment and accurate recall of the final word.

3.4. Results

The study followed a mixed design with one within-subjects variable, condition (gender/number agreement, gender violation, number violation), and two between-subjects variables, namely group (beginners, intermediates, monolinguals) and working memory (high and low working memory).

Please cite this article in press as: Sagarra, N., Herschensohn, J., The role of proficiency and working memory in gender and number agreement processing in L1 and L2 Spanish. Lingua (2010), doi:10.1016/j.lingua.2010.02.004

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

	Gender/number agreement		Gender violation		Number violation	
	M	*SD*	*M*	*SD*	*M*	*SD*
Mean RTs at the word immediately preceding the adjective ($N-1$)						
Beginners	1033.97	450.69	1020.69	375.66	956.92	368.42
Intermediates	890.84	314.95	835.41	291.43	819.15	287.70
Spanish monolinguals	469.35	168.09	492.75	175.23	495.15	174.81
Mean RTs at the adjective (N)						
Beginners	901.32	332.83	911.36	310.13	952.47	359.71
Intermediates	708.83	207.43	823.82	276.62	855.64	258.98
Spanish monolinguals	465.34	115.06	554.96	189.89	546.39	166.88
Mean RTs at the word immediately following the adjective ($N+1$)						
Beginners	550.52	144.00	530.54	120.70	521.27	138.85
Intermediates	445.70	85.48	536.43	137.73	514.80	120.42
Spanish monolinguals	385.97	70.99	449.71	111.25	433.78	80.65

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

3.4.1. Self-paced reading task

The statistical analyses for the self-paced reading task included four repeated-measures ANOVAs with a 3 (agreement: gender/number agreement, gender violation, number violation) × 3 (group: beginners, intermediate, Spanish monolinguals) factorial design: one for the word immediately preceding the adjective, one for the adjective, one for the word immediately following the adjective, and one for accuracy on the comprehension questions. Table 1 displays descriptive statistics for the word preceding the adjective, the adjective, and the word following the adjective and Table 2 the descriptive statistics for accuracy on comprehension questions. Mean RTs at the word preceding the adjective were analyzed to ensure that latencies on the adjective were exclusively due to the variables under investigation. In turn, mean RTs at the word following the adjective were examined to measure possible residual effects in processing the adjective. Late latencies are important because the onset of L2 processing is delayed (e.g., Weber-Fox and Neville, 1996) and because differences between gender and number agreement have only been found in late measures (Barber and Carreiras, 2005). Finally, accuracy on comprehension questions were included to guarantee that participants understood the sentences they read and Table 2 shows that all participants had a good level of understanding (the lowest mean was 8.29 over 10).

The repeated-measures ANOVA conducted on the word immediately preceding the adjective showed no main effect for condition ($F(2,386) = 2.910$, $p > .05$) and no interaction of condition × group ($F(4,386) = 1.813$, $p > .05$), but there was a main significant effect for group ($F(2,193) = 84.992$, $p < .01$). Bonferroni post hoc tests revealed that beginners were slower readers than intermediates and that beginners and intermediates were in turn slower than Spanish monolinguals (all $p < .01$). These results indicate that higher L2 proficiency results in faster reading (see Clahsen and Felser, 2006, for similar findings).

The repeated-measures ANOVAs carried out for the mean RTs at the adjective, the word following the adjective, and accuracy on comprehension questions revealed a significant main effect for condition (adjective: $F(2,386) = 16.608$, $p < .01$; adjective + 1: $F(2,386) = 11.887$, $p < .01$; comprehension questions: $F(2,386) = 4.282$, $p < .05$) and group (adjective: $F(2,193) = 54.995$, $p < .01$; adjective + 1: $F(2,193) = 27.129$, $p < .01$; comprehension questions: $F(2,193) = 16.258$, $p < .01$), as well as a significant interaction of condition × group (adjective: $F(4,386) = 2.407$, $p < .05$; adjective + 1: $F(4, 386) = 7.923$, $p < .01$; comprehension questions: $F(4,386) = 4.075$, $p < .01$). Bonferroni post hoc tests showed no significant differences in RTs or accuracy among conditions for beginners (all $p > .05$). However, intermediates and Spanish monolinguals showed longer RTs in sentences with gender or number violations than with gender/number agreement (all $p < .01$), and Spanish monolinguals were in turn more accurate answering questions about sentences with gender/number agreement than gender disagreement ($p < .01$). In addition, no RT or accuracy differences were found between gender and number violations, except for Spanish monolinguals who were more accurate answering questions about sentences with number than gender violations ($p < .01$). Finally, as expected, between-group comparisons showed longer RTs and higher accuracy in Spanish monolinguals than intermediates and beginners, and in intermediates than beginners in sentences with gender/number agreement (all $p < .05$).

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

	Gender/number agreement		Gender violation		Number violation	
	M	*SD*	*M*	*SD*	*M*	*SD*
Beginners	8.67	1.23	8.57	1.05	8.32	1.18
Intermediates	8.80	1.28	8.72	1.19	8.75	1.20
Spanish monolinguals	9.49	.64	8.84	.99	9.44	.64

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

Please cite this article in press as: Sagarra, N., Herschensohn, J., The role of proficiency and working memory in gender and number agreement processing in L1 and L2 Spanish. Lingua (2010), doi:10.1016/j.lingua.2010.02.004

Because native speakers are by definition from a different population, it could be claimed that the significant main and interaction effects obtained at the adjective, adjective + 1, and comprehension questions were caused by this group. Three additional 3 (condition) × 2 (group: beginners, intermediates) repeated-measures ANOVAs revealed identical findings to those obtained in the 3 × 3 ANOVAs reported above for RTs at the adjective and the word following the adjective, namely, a significant main effect for condition (adjective: $F(2,262) = 8.911$, $p < .01$; adjective + 1: $F(2,262) = 3.820$, $p < .05$) and group (adjective: $F(1,131) = 2.288$, $p < .01$; adjective + 1: $F(1,131) = 4.716$, $p < .05$), as well as a significant interaction of condition × group (adjective: $F(2,262) = 3.001$, $p < .05$; adjective + 1: $F(2,262) = 11.175$, $p < .01$). Bonferroni post hoc tests showed the same significant differences as those reported in the 3 × 3 ANOVAs, demonstrating that the differences between the two L2 groups were not caused by the Spanish monolingual group. The only discrepancy between the 3 × 3 and the 3 × 2 ANOVAs was the expected lack of significant differences between beginners and intermediates in the comprehension questions: condition ($F(2,262) = 1.286$, $p > .05$), group: $F(2,262) = .923$, $p > 05$, and condition × group: $F(4,386) = 4.075$, $p > .05$).

As mentioned earlier, to ensure that longer RTs at the adjective were due to sensitivity to agreement violations rather than markedness, we conducted an additional 5 (condition) × 3 (group) repeated-measures ANOVA on the mean RTs at the adjective in the following conditions: gender/number agreement with unmarked gender/number (MS-MS experimental sentences), gender/number agreement with marked gender and number (FS-FS, MP-MP filler sentences), and gender/number disagreement with marked gender/number (MS-*FS, MS-*MP experimental sentences). Means and standard deviations can be found in Table 1, except for filler sentences with FS-FS and MP-MP: beginners FS-FS: $M = 855.58(242.18)$, beginners MP-MP: $M = 852.19(237.00)$, intermediates FS-FS: $M = 691.12(192.25)$; intermediates MP-MP: $M = 683.61$ (189.78), Spanish monolinguals FS-FS: $M = 455.28(97.87)$, and Spanish monolinguals MP-MP: $M = 448.34(87.07)$. The results showed 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 × group ($F(4,756) = 31.793$, $p < .01$). Bonferroni post hoc tests revealed no significant differences for beginners but longer RTs in sentences with gender/number disagreement (MS-*FS, MS-*MP) than those with gender/number agreement, regardless of gender/number markedness (MS-MS, FS-FS, MP-MP). These findings suggest that longer RTs obtained in sentences with gender/number violations are due to sensitivity to grammatical incongruencies rather than to markedness.

3.4.2. Grammaticality judgment task

As mentioned earlier, the grammaticality judgment task generated two scores: accuracy on identifying sentences as correct or incorrect and a confidence rating. The means and standard deviations for each score divided into group and condition can be found in Table 3.

For each score, we conducted a repeated-measures ANOVA with the same 3 × 3 factorial design. The ANOVAs showed a significant main effect for condition (accurate grammaticality judgments: $F(2,330) = 95.218$, $p < .01$; confidence rating: $F(2,330) = 47.665$, $p < .01$) and group (accurate grammaticality judgments: $F(2,165) = 137.259$, $p < .01$; confidence rating: $F(2,165) = 125.199$, $p < .01$), and a significant interaction between condition and group (accurate grammaticality judgments: $F(4,330) = 45.113$, $p < .01$; confidence rating: $F(4,330) = 16.993$, $p < .01$). In line with previous research (e.g., Herschensohn, 2000), Bonferroni post hoc tests revealed that L2 learners were more accurate at classifying sentences with gender/number agreement as correct than those with gender or number disagreement as incorrect (all at least $p < .05$), suggesting that, when in doubt, low proficient learners tend to assume that L2 sentences are grammatical. The sole exception, non-significant differences between number agreement and disagreement in the intermediate group, can be explained by ceiling effects. As for confidence ratings, beginners felt equally confident about their answers to sentences with gender/number agreement and gender disagreement. With regard to differences between gender and number violations, L2 learners were more confident and more accurate at classifying sentences with number disagreement than gender disagreement, and more accurate at identifying number errors than gender errors (all at least $p < .05$). There were no differences across conditions in the Spanish monolingual group due to ceiling effects. Finally, between-group comparisons indicated that beginners and

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

	Gender/number agreement		Gender violation		Number violation	
	M	*SD*	*M*	*SD*	*M*	*SD*
Accuracy identifying sentences as correct or incorrect ($k = 10$)						
Beginners	8.73	1.51	2.92	4.11	5.94	2.51
Intermediates	8.26	1.32	6.47	3.40	7.98	1.54
Spanish monolinguals	10.00	.00	10.00	.00	10.00	.00
Confidence rating ($k = 5$) (1 = not sure at all; 5 = completely sure)						
Beginners	3.51	.85	3.56	.91	3.97	.82
Intermediates	3.54	.67	3.97	.72	4.42	.60
Spanish monolinguals	5.00	.00	5.00	.00	5.00	.00

Note: $n = 52$ 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.

Please cite this article in press as: Sagarra, N., Herschensohn, J., The role of proficiency and working memory in gender and number agreement processing in L1 and L2 Spanish. Lingua (2010), doi:10.1016/j.lingua.2010.02.004

intermediates were equally accurate at classifying sentences with gender/number agreement and gender disagreement and equally confident in sentences with gender/number agreement. However, intermediates were more accurate than beginners at classifying sentences with number disagreement and identifying gender and number errors, and they were more confident than beginners in sentences with gender or number disagreement (all at least $p < .05$). As expected, Spanish monolinguals were more accurate and confident than L2 learners across conditions (all $p < .01$).

3.4.3. Working memory

To assess whether working memory affects processing of gender/number agreement/disagreement, multiple bivariate correlations were carried out for each variable and group. The results revealed that working memory positively correlated with sentences with gender agreement violations in the intermediate group (RTs at adjectives: $r = .334$, $p < .01$; accuracy on grammaticality judgments: $r = .290$, $p < .05$). These findings indicate that intermediate learners with higher working memory capacity were more sensitive to gender disagreement than those with lower memory. No significant correlations were obtained for Spanish monolinguals due to ceiling effects. With regard to beginners, the lack of significant effects can be explained by the low working memory mean of the group. In effect, a one-way ANOVA showed that the intermediate group had a higher working memory level ($M = 56.64$, $SD = 12.78$) than beginners ($M = 44.99$, $SD = 12.80$) and Spanish monolinguals ($M = 42.94$, $SD = 11.88$) ($F(2, 193) = 22.443$, $p < .01$).

4. Discussion

4.1. Research question 1

The first research question of the study investigated whether Spanish monolinguals, and beginning and intermediate adult L2 learners show sensitivity to gender and number agreement violations, based on online and offline data. The RT and grammaticality judgment data showed that beginners were not sensitive to gender or number agreement violations but intermediates and Spanish monolinguals showed sensitivity to both violation types (i.e., longer RTs for discord than concord). The results of Spanish monolinguals (at ceiling) and beginning learners (at floor) are in keeping with previous research. In representational terms, beginners' insensitivity could be expected in the early stages of L2A, as both UG approaches discussed earlier suggest. In contrast to the beginners, the intermediates are not completely insensitive to gender and number discord on the Spanish adjective, indicating a developing procedural sensitivity to gender and number agreement violations that should reflect the developing interlanguage grammar. In line with models such as Full Transfer Full Access, this could be indicative of the restructuring of the values of grammatical features of adjectives for L2 Spanish by Anglophone learners (cf. Herschensohn, 2000, 2001; Bruhn de Garavito and White, 2002; White et al., 2004).

A possible objection to this interpretation of the results is that longer RTs in incongruent sentences might be due to a reaction to the marked form of gender (feminine) and number (plural) in the adjective. There are two reasons why we consider this option unlikely. First, if non-default forms were responsible for reading latencies in feminine and plural adjectives in incongruent sentences, they should also be expected to produce reading latencies in feminine and plural adjectives in congruent sentences. However, statistical analyses comparing adjectives presented in the non-default form (feminine for gender, plural for number) in both congruent and incongruent sentences revealed latencies to be a function of discord, not non-default status. Indeed, the analyses showed longer RTs in FS and MP adjectives that disagreed with the noun than with FS and MP adjectives that agreed with it for both intermediate learners and Spanish monolinguals. Second, previous studies have failed to find any relation between a noun's gender or number and processing of agreement/disagreement in Romance languages as L1 or L2 (e.g., Alarcón, 2009; Antón-Méndez et al., 2002; De Mulder, 2006; Keating, 2009).

4.2. Research question 2

The second research question examined whether Spanish monolinguals, and intermediate and beginning L2 learners process gender agreement differently from number agreement, based on online and offline data, and if they do, whether the increased difficulty of one over the other was influenced by working memory. The data of the self-paced reading task (RTs and accuracy on comprehension questions) revealed no significant differences between gender and number concord/discord for any of the three groups. In the case of beginners and Spanish monolinguals, the lack of differences can be explained in terms of floor and ceiling effects, respectively. As for intermediate learners, the similarity between gender and number agreement could be due to several factors that we discuss in detail in the General Discussion section. These findings contradict models claiming that gender agreement is cognitively more demanding than number agreement (Faussart et al., 1999).

In contrast, the data of the grammaticality judgment task indicated that beginners and intermediates were more accurate at identifying number agreement errors (beginners were accurate 59.4% of the time and intermediates 79.8% of the time) than gender agreement errors (beginners were accurate 26.2% of the time and intermediates 56.4% of the time). Furthermore, beginners and intermediates were more confident in sentences with number discord than gender discord. These findings suggest that learners process number disagreement more easily than gender disagreement and are in line with models

Please cite this article in press as: Sagarra, N., Herschensohn, J., The role of proficiency and working memory in gender and number agreement processing in L1 and L2 Spanish. Lingua (2010), doi:10.1016/j.lingua.2010.02.004

Table 4
Summary of results for experiments 1 (moving window) and 2 (grammaticality judgment).

	Moving window test (ms. on adj.)	Grammaticality judgment test	
		Correct–incorrect	Confidence rating
Gender agreement vs. gender violation			
Beginners	agree = violation	agree > violation	agree = violation
Intermediates	agree < violation	agree > violation	agree < violation
Monolinguals	agree < violation	agree = violation	agree = violation
Number agreement vs. number violation			
Beginners	agree = violation	agree > violation	agree < violation
Intermediates	agree < violation	agree = violation	agree < violation
Monolinguals	agree < violation	agree = violation	agree = violation
Gender violation vs. number violation			
Beginners	gender = number	gender < number	gender < number
Intermediates	gender = number	gender < number	gender < number
Monolinguals	gender = number	gender = number	gender = number

claiming that processing gender agreement is cognitively more demanding than number disagreement. As for the Spanish monolingual group, as expected, ceiling effects produced no differences between gender and number violations for any of the measures of the grammaticality judgment test. Finally, we found that working memory was indeed significant at the intermediate level and also mediated the differences between gender and number concord processing, a topic we explore below.

4.3. General discussion

Given native sensitivity to concord violations (which induce longer RTs in processing), this study has examined two research questions: whether learners whose L1 lacks gender/number concord with adjectives can acquire these features in the L2, and if so, whether the two features are distinct or similar with respect to processing and interpretation. We also consider the roles of proficiency level and working memory to the computation of gender and number concord. These research questions address two orthogonal theoretical issues, the accessibility of L2 grammatical features to non-advanced learners and the significance of the gender–number distinction. Table 4 summarizes the results for the self-paced reading and grammaticality judgment tasks.

As expected, the native controls show significantly longer RTs to sentences with gender/number violations than to those with concord, and as we predicted, beginners in our study did not show sensitivity to gender or number agreement violations in the self-paced reading task (where reaction time was equivalent for concord/discord) or the grammaticality judgment task (where subjects were at floor). Assuming initial transfer of L1 English grammatical features (Hawkins, 2001; White, 2003), these beginning learners are not expected to transfer grammatical features of [*u*gender] or [*u*number] for adjectives because English only has an interpretable feature for number on nouns. Given the fact that concord is explicitly taught and that the learners could presumably learn it, they should (if teaching resulted in accurate representation and processing procedures) be 100% accurate. The fact that they aren't indicates that positive and negative evidence for concord is not sufficient for beginning L2 Spanish learners, and that more L2 exposure is required. In effect, learners with more L2 experience (the intermediates) are beginning to respond like the native speakers.

It is the intermediate group whose responses serve to differentiate between deficit and accessibility approaches. This group shows significantly longer RTs on violations of both gender and number concord than on non-violations (agreement) on the adjectives in the self-paced reading test. These results indicate a growing sensitivity in the intermediates to gender and number concord/discord in Spanish adjectives. A strictly shallow structure deficit view of processing is unsupported in that the intermediates—in contrast to the beginners—do not appear to be using lexical, semantic and pragmatic cues as opposed to morphosyntactic cues. Unlike the L2 learners (whose L2 target environment stay is measured in months) in Silva and Clahsen's (2008) study, our subjects are not totally indifferent to inflectional morphology. Silva and Clahsen conclude that further studies with more advanced learners are needed to ascertain the extent of the morphological insensitivity. While we do not want to exaggerate the native-like processing of our intermediate learners, their incipient sensitivity to discord points to development of target-like computation.[3] Indeed, they are more accurate and have faster RTs with inanimate (no semantic-pragmatic cue) than animate nouns (Sagarra and Herschensohn, 2011), an indication that they are developing the grammatical features [*u*gender] and [*u*number]. This development favors accessibility over deficit approaches to representation (cf. Hawkins, 2009) since the learners gain L2 uninterpretable as well as interpretable gender and number, even though grammatical features are unnecessary for interpretation (Chomsky, 2002). An anonymous reviewer notes that the fact that the intermediate learners show emergent sensitivity to a feature that is neither in the L1, nor needed at LF, is encouraging. The sentences of the task are

[3] A further difference is the task, a 30–60 ms priming task for Silva and Clahsen, as opposed to much later reaction time in our experiment.

Please cite this article in press as: Sagarra, N., Herschensohn, J., The role of proficiency and working memory in gender and number agreement processing in L1 and L2 Spanish. Lingua (2010), doi:10.1016/j.lingua.2010.02.004

novel in that they aren't sentences that the subjects have had likelihood to encounter (and have memorized), so the sequences are all new combinations, yet the intermediates and monolinguals distinguish between discord and concord. While their interlanguage grammars are still developing, and while their processing skills are slower than natives', they nevertheless manifest sensitivity to noun–adjective agreement in an online task. These results imply that L2 learners can gain grammatical features on L2 adjectives after puberty.

The second research question tested the relationship between two types of agreement, gender and number, examining possible differences between the two. As we noted in the initial section, there are differing views of the correct analysis of gender and number representation and processing in Spanish. According to Domínguez et al. (1999), psycholinguistic models diverge in their view of gender as a process achieved through a whole-word representation (*niña* 'girl') or across the stem of the word (*niñ-a* 'girl'). Furthermore, there are differences in processing both gender and number in terms of grammatical category, between D and A (cf. Centeno et al., 2007). In any case, the facts support a distinction between gender and number storage and processing in monolingual Spanish. We predicted that intermediates would process number more efficiently than gender for two reasons: their L1 English grammaticalizes number but not gender on nouns and determiners, and several studies show that Spanish monolinguals process number more efficiently than gender.

Our hypothesis that there would be no differences for beginners due to floor effects was confirmed. In contrast, our prediction that gender agreement violations would produce longer RTs at the adjective than number agreement violations for the intermediate learners and the Spanish monolinguals was not supported. Several factors could cause the lack of significant differences between gender and number agreement in the intermediate group. For example, delayed processing effects not visible in the mean RTs at the adjectives could explain the similarity between gender and number discord, but this possibility is refuted by the lack of significant differences obtained in the statistical analyses conducted at the word immediately following the adjective. Another possible explanation could be that the self-paced reading task is sensitive to the identification of grammatical incongruencies (concord-discord comparisons) but not sensitive enough to the processing of grammatical violations that consume more or less attentional resources (gender discord-number discord comparison), an issue we address in terms of working memory. Recall that working memory positively correlated with intermediates' ability to notice gender disagreement but not number disagreement, suggesting that gender disagreement is cognitively more taxing than number disagreement. Finally, the intermediate learners could have been using their metalinguistic knowledge to read the sentences of the self-paced reading task. Since the two structures are covered in class, they would be equally sensitive to the two types of disagreement. The data from the grammaticality judgment test let us address this option. In effect, the results of the grammaticality judgment test revealed that judgments on number are more accurate than those on gender, and confidence is higher on number than gender for all groups. These results together with the working memory findings corroborate lexical and syntactic models claiming that processing gender agreement is cognitively more demanding than number disagreement. They may also relate to the transfer of interpretable number on N from L1 English.

Finally, Spanish monolinguals showed ceiling effects both for the self-paced reading task and the grammaticality judgment task, indicating that our instrument is not sensitive enough to capture the distinction between gender and number agreement reported in the literature (see Barber and Carreiras, 2005, for ERP evidence of differences in the processing of gender and number disagreement in native Spanish). As for working memory effects in the Spanish monolingual and beginner groups, there were no significant correlations due to ceiling effects in Spanish monolinguals and the low working memory mean of beginners. In contrast, the intermediate group—the one that shows the emerging ability to compute gender concord—is also the bellwether for working memory. At this intermediate level, working memory can be a significant factor whereas it is essentially irrelevant at ceiling.

5. Conclusion

This study of processing of L2 Spanish gender and number agreement and disagreement on adjectives confirms monolingual patterns at intermediate L2 levels and shows insensitivity to concord/discord at beginning L2 stages. In terms of representational and computational accounts, our data for intermediate learners—whose results both in the self-paced reading and grammaticality judgment tasks demonstrate an emerging sensitivity to adjective concord/discord in L2 Spanish—favor models which allow eventual acquisition of grammatical features of [*u*gender] and [*u*number] on adjective. Furthermore, our results also confirm a distinction between gender and number, manifested both in the working memory of intermediate learners (higher working memory yields higher sensitivity to gender but not number agreement violations) and in the grammaticality judgments of beginning and intermediate learners (more accuracy and confidence in sentences with number than gender agreement violations). We interpret these working memory and grammaticality judgment findings as indicative of the differential representation and processing of the two features similar to that of Spanish monolinguals. Taken together, the results of this study suggest that adult learners with a certain proficiency level can demonstrate grammatical knowledge and implementation that is qualitatively comparable to that of native speakers; that gender agreement is cognitively more taxing than number disagreement; and that working memory facilitates gaining sensitivity to adjective morphology. Future research may tease out further details of why gender agreement consumes more attentional resources than number agreement and how learner cognitive individual differences such as working memory modulate agreement processes.

Please cite this article in press as: Sagarra, N., Herschensohn, J., The role of proficiency and working memory in gender and number agreement processing in L1 and L2 Spanish. Lingua (2010), doi:10.1016/j.lingua.2010.02.004

Appendix A

List of the nouns and adjectives used in the moving window and the grammaticality judgment tasks.

Inanimate nouns ($k = 40$): *acuario*, *acueducto*, *anillo*, *aparato*, *archivo*, *armario*, *auto*, *concierto*, *contrato*, *cuestionario*, *cuchillo*, *desayuno*, *dibujo*, *dinosaurio*, *documento*, *edificio*, *escritorio*, *gráfico*, *helado*, *instrumento*, *método*, *momento*, *monólogo*, *museo*, *negocio*, *panfleto*, *partido*, *periódico*, *pescado*, *prototipo*, *proyecto*, *regalo*, *sombrero*, *teléfono*, *torneo*, *trabajo*, *vehículo*, *verano*, *vestido*, *zapato*.

Adjectives ($k = 40$): *agresivo*, *bajo*, *bueno*, *caro*, *cómico*, *complicado*, *creativo*, *curioso*, *dinámico, divertido*, *fabuloso*, *famoso*, *fantástico*, *favorito*, *feo*, *limpio*, *loco*, *lógico*, *malicioso*, *malo*, *mediano*, *moderno*, *nuevo*, *ocupado*, *ortodoxo*, *pacífico*, *pequeño*, *perfecto*, *preferido*, *rápido*, *rico*, *romántico*, *serio*, *silencioso*, *simpático*, *sucio*, *típico*, *trágico*, *tranquilo*, *viejo*.

References

Acuña-Fariña, J.C., 2009. The linguistics and psycholinguistics of agreement: a tutorial overview. Lingua 119, 389–424.
Alarcón, I., 2006. The Second Language Acquisition of Spanish Gender Agreement: The Effects of Linguistic Variables on Accuracy. LINCOM Studies in Language Acquisition, Munich, Germany.
Alarcón, I., 2009. The processing of gender agreement in L1 and L2 Spanish: evidence from reaction-time data. Hispania 92 (4), 814–828.
Almor, A., MacDonald, M.C., Kempler, D., Andersen, E.S., Tyler, L.K., 2001. Comprehension of long distance number agreement in probable Alzheimer's disease. Language and Cognitive Processes 16, 35–63.
Antón-Méndez, M.I., 1996. Clitics and Attraction Errors: An Experimental Study of Language Production. Unpublished manuscript, University of Arizona, Tucson.
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), 1–25.
Ayoun, D., 2007. The second language acquisition of grammatical gender and agreement. In: Ayoun, D. (Ed.), French Applied Linguistics. John Benjamins, Amsterdam, pp. 130–170.
Baddeley, A.D., 2003. Working memory and language: an overview. Journal of Communication Disorders 36, 189–208.
Baddeley, A.D., 2007. Working Memory, Thought, and Action. Oxford University Press, Oxford.
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 (1), 137–153.
Bartning, I., 2000. Gender agreement in L2 French: pre-advanced vs. advanced learners. Studia Linguistica 54, 225–237.
Birdsong, D., 1989. Metalinguistic Performance and Interlanguage Competence. Springer, New York.
Blom, E., Polišenská, D., Unsworth, S., 2008. The acquisition of grammatical gender in Dutch. Second Language Research 24, 259–265.
Bonnet, G., 2002. The assessment of pupils' skills in English in eight European countries. Paris, European Network of policy makers for the evaluation of education systems.
Bordag, D., 2007. Psycholinguistische Aspekte der Interferenzerscheinungen in der Flexionsmorphjologie des Tschechischen als Fremdsprache. G. Olms Verlag, Hildesheim.
Bordag, D., Pechmann, T., 2007. Factors influencing L2 gender processing. Bilingualism: Language and Cognition 10 (3), 299–314.
Bruhn de Garavito, J., White, L., 2002. L2 acquisition of Spanish DPs: the status of grammatical features. In: Pérez-Leroux, A.-T., Liceras, J. (Eds.), The Acquisition of Spanish Morphosyntax: The L1/L2 Connection. Kluwer, Dordrecht, pp. 153–178.
Cacciari, C., Padovani, R., 2007. Further evidence of gender stereotype priming in language: semantic facilitation and inhibition in Italian role nouns. Applied Psycholinguistics 28, 277–294.
Carreiras, M., Carr, L., Barber, H.A., Hernández, A., 2010. Where syntax meets match: right intraparietal sulcus activation in response to grammatical number agreement violations. Neuroimage 49, 1741–1749.
Carroll, S., 1989. Second language acquisition and the computational paradigm. Language Learning 39, 535–594.
Carroll, S., 1999. Putting 'input' in its proper place. Second Language Research 15, 337–388.
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.
CEDEFOP, 2004. Lifelong Learning: Citizens Views. European Centre for Development of Vocational Training, Thessaloniki.
Centeno, J.G., Anderson, R., Obler, L., 2007. Communication Disorders in Spanish Speakers. Theoretical, Research and Clinical Aspects. Multilingual Matters, Clevedon, UK.
Chomsky, N., 2000. Minimalist inquiries: the framework. In: Martin, R., Michaels, D., Uriagereka, J. (Eds.), Step by Dtep: Essays on Minimalist Syntax in Honor of Howard Lasnik. MIT Press, Cambridge, pp. 89–155.
Chomsky, N., 2001. Derivation by phase. In: Kenstowicz, M. (Ed.), Ken Hale: A Life in Language. MIT Press, Cambridge, pp. 1–52.
Chomsky, N., 2002. On Nature and Language. Cambridge University Press, Cambridge.
Clahsen, H., Felser, C., 2006. Grammatical processing in language learners. Applied Psycholinguistics 27, 3–42.
Colé, P., Pynte, J., Andriamamonjy, P., 2003. Effect of grammatical gender on visual word recognition: evidence from lexical decision and eye-movement experiments. Perception and Psychophysics 65, 407–419.
Colé, P., Ségui, J., 1994. Grammatical incongruency and vocabulary types. Memory and Cognition 22, 387–394.
Corbett, G.G., 1991. Gender. Cambridge University Press, Cambridge.
Crain, S., Lillo-Martin, D., 1999. An Introduction to Linguistic Theory and Language Acquisition. Blackwell Publishers, Malden, MA.
De Mulder, H., 2006. Second Language Acquisition of Gender: Italian Gender in Native Speakers of Dutch and English. Unpublished M.A. Thesis. Utrecht University.
De Vincenzi, M., 1999. Differences between the morphology of gender and number: evidence from establishing coreferences. Journal of Psycholinguistic Research 28 (5), 537–553.
De Vincenzi, 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 (1), 41–74.
Dewaele, J.-M., Véronique, D., 2000. Relating gender errors to morphosyntax and lexicon in advanced French interlanguage. Studia Linguistica 54, 212–224.
Dewaele, J.-M., Véronique, D., 2001. Gender assignment and gender agreement in advanced French interlanguage: a cross-sectional study. Bilingualism: Language and Cognition 4, 275–297.
Dörnyei, Z., 2009. The Psychology of Second Language Acquisition. Oxford UP, Oxford.
Domínguez, A., Cuetos, F., Ségui, J., 1999. The processing of grammatical gender and number in Spanish. Journal of Psycholinguistic Research 28 (5), 485–498.
Doughty, C., Long, M., 2003. Handbook of Second Language Acquisition. Blackwell, Malden, MA/Oxford.
Eberhard, K.M., 1997. The marked effect of number on subject–verb agreement. Journal of Memory and Language 36, 147–164.
Ellis, N.C., Sinclair, S., 1996. Working memory in the acquisition of vocabulary and syntax. Quarterly Journal of Experimental Psychology 49, 234–250.
Faussart, C., Jakubowicz, C., Costes, M., 1999. Gender and number processing in spoken French and Spanish. Rivista di Linguistica 11, 75–101.

Please cite this article in press as: Sagarra, N., Herschensohn, J., The role of proficiency and working memory in gender and number agreement processing in L1 and L2 Spanish. Lingua (2010), doi:10.1016/j.lingua.2010.02.004

Fernández-García, M., 1999. Patterns of gender agreement in the speech of second language learners. In: Gutiérrez-Rexach, J.,Martínez-Gil, F. (Eds.),Advances in Hispanic Linguistics: Papers from the 2nd Hispanic Linguistics Symposium. Cascadilla Press, Somerville, MA, pp. 3–15.
Fiebach, C.J., Schlesewsky, M., Friederici, A.D., 2002. Separating syntactic memory costs and syntactic integration costs during parsing: the processing of German wh-questions. Journal of Memory and Language 47, 250–272.
Foote, R., 2008. Sensitivity to gender in early and late language learners. In: Hispanic Linguistics Symposium, Laval University, Quebec, Canada.
Foucart, A., 2008. Syntactic Gender Processing in French as a First and a Second Language. Unpublished Doctoral Dissertation. Université Aix-Marseille and University of Edinburgh.
Foucart, A., Frenck-Mestre, C., 2004. Processing of grammatical gender information in French as first and second language: evidence from ERPs. In: Architecture and Mechanisms of Language Processing (AMLaP), Aix-en-Provence, France.
Franceschina, F., 2001. Morphological or syntactic deficits in near-native speakers? An assessment of some current proposals. Second Language Research 17, 213–247.
Franceschina, F., 2002. Case and phi-feature agreement in advanced L2 Spanish grammars. In: Foster-Cohen, S.H., Ruthenberg, T., Poschen, M.L. (Eds.), EUROSLA Yearbook, 2. John Benjamins, Amsterdam, pp. 71–86.
Franceschina, F., 2005. Fossilized Second Language Grammars: The Acquisition of Grammatical Gender. John Benjamins, Amsterdam/Philadelphia.
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. In: EUROSLA Yearbook (76–106), John Benjamins, Amsterdam/Philadelphia.
Friederici, A.D., Garrett, M.F., Jacobsen, T., 1999. Editorial and introduction to Part I of the special issue Processing of Grammatical Gender. Journal of Psycholinguistic Research 28, 455–456.
Gess, R., Herschensohn, J., 2001. Shifting the DP parameter: a study of Anglophone French L2ers. In: Camps, J., Wiltshire, C. (Eds.), Romance Syntax, Semantics and their L2 Acquisition. John Benjamins, Philadelphia, pp. 105–119.
Geva, E., Ryan, E.B., 1993. Linguistic and cognitive correlates of academic skills in first and second Language. Language Learning 43 (1), 5–42.
Gillon-Dowens, M., Barber, H., Guo, J., Guo, T., Carreiras, M., 2008. ERP correlates of automatic syntactic processing in highly proficient L1 Chinese-L2 Spanish late learners. In: Paper presented at the 17th Annual Cognitive Neuroscience Society Meeting, San Francisco.
Gillon-Dowens, M., Barber, H., Vergara, M., Carreiras, M., 2004. Does practice make perfect? An ERP study of morphosyntactic processing in highly proficient English–Spanish late bilinguals. In: Poster presented at AMLaP, Aix-en-Provence, France.
Gillon-Dowens, M., Barber, H., Vergara, M., Carreiras, M., 2009. Morpho-syntactic processing in late L2 learners—ERP evidence, Journal of Cognitive Neuroscience (Epub ahead or print).
Granfeldt, J., 2005. The development of gender attribution and gender concord in French: a comparison of bilingual first and second language learners. In: Dewaele, J. (Ed.), Focus on French as a Foreign Language: Multidisciplinary Approaches. Multilingual Matters, pp. 164–190.
Green, J.N., 1988. Spanish. In: Harris, M., Vincent, N. (Eds.), The Romance Languages. Oxford University Press, New York, pp. 79–130.
Grosjean, F., Dommergues, J.-Y., Cornu, E., Guillelmon, D., Besson, C., 1994. Gender-marking effect in spoken word recognition. Perception and Psychophysics 56, 590–598.
Guasti, M.T., 2002. Language Acquisition: The Growth of Grammar. MIT Press, Cambridge, MA.
Guillelmon, D., Grosjean, F., 2001. The gender marking effect in spoken word recognition: the case of bilinguals. Memory and Cognition 29, 503–511.
Gunter, T.C., Friederici, A.D., Schriefers, H., 2000. Syntactic gender and semantic expectancy: ERPs reveal early autonomy and late interaction. Journal of Cognitive Neuroscience 12, 556–568.
Hagoort, P., Brown, C.M., 1999. Gender electrified: ERP evidence on the syntactic nature of gender processing. Journal of Psycholinguistic Research 28, 715–728.
Hammer, A., Goebel, R., Schwarzbach, J., Münte, T.F., Jansma, B.M., 2007. When sex meets syntactic gender on a neural basis during pronoun processing. Brain Research 1146, 185–198.
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 15, 647–660.
Hartsuiker, R.J., Antón-Méndez, I., Van Zee, M., 2001. Object attraction in subject–verb agreement construction. Journal of Memory and Language 45, 546–572.
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 (1), 73–112.
Hawkins, R., 1998. Explaining the difficulty of French gender attribution for speakers of English. In: Paper presented at the 8th Annual EUROSLA Conference, Paris.
Hawkins, R., 2001. Second Language Syntax: A Generative Introduction. Blackwell, Cambridge.
Hawkins, R., 2009. Second language acquisition of morphosyntax. In: Ritchie, W., Baktia, T. (Eds.), The New Handbook of Second Language Acquisition. Emerald, Leeds, UK, pp. 211–233.
Hawkins, R., Chan, C.Y., 1997. The partial availability of Universal Grammar in second language acquisition: the 'failed functional features hypothesis'. Second Language Research 13, 187–226.
Hawkins, R., Franceschina, F., 2004. Explaining the acquisition and non-acquisition of determiner-noun gender concord in French and Spanish. In: Prévost, P., Paradis, J. (Eds.), The Acquisition of French in Different Contexts. John Benjamins, Amsterdam, pp. 175–206.
Haznedar, B., Schwartz, B.D., 1997. Are there optional infinitives in child L2 acquisition? In: Hughes, E., Hughes, M., Greenhill, A. (Eds.), Proceedings of the 21st Annual Boston University Conference on Language Development. Cascadilla Press, Somerville, MA, pp. 257–268.
Hernández, A., Hofmann, J., Kotz, S., 2007. Age of acquisition modulates neural activity for both regular and irregular syntactic functions. Neuroimage 36 (3), 912–923.
Hernández-Pina, F., 1984. Teorías Psicosociolingüísticas y su Aplicación a la Adquisición del Español como Lengua Materna. Siglo XXI de España Editores, S.A, Madrid.
Herschensohn, J., 2000. The Second Time Around: Minimalism and L2 Acquisition. John Benjamins, Amsterdam/Philadelphia.
Herschensohn, J., 2001. Missing inflection in L2 French: accidental infinitives and other verbal deficits. Second Language Research 17, 273–305.
Herschensohn, J., 2007. Language Development and Age. Cambridge University Press, Cambridge.
Herschensohn, J., Arteaga-Capen, D., 2007. Parameters and processing: gender agreement in L2 French. In: Paper presented at the 17th EUROSLA Conference, Newcastle.
Herschensohn, J., Frenck-Mestre, 2005. Failed [ugender] features in L2 French: what ERPs reveal about L1 influence. In: Paper presented at the 15th EUROSLA Conference, Dubrovnik.
Hopp, H., 2007. Ultimate Attainment at the Interfaces in Second Language Acquisition: Grammar and Processing. PhD Dissertation, Rijksuniversiteit Groningen.
Igoa, J.M., García-Albea, J.E., Sánchez-Casas, R., 1999. Gender–number dissociations in sentence production in Spanish. Rivista di Linguistica 11 (1), 163–196.
Jacubowicz, C., Faussart, C., 1998. Gender agreement in the processing of spoken French. Journal of Psycholinguistic Research 27 (6), 597–617.
Jin, F., Åfarli, T.A., van Dommelen, W.A., et al., 2007. Endstate L2 competence: does perception parallel production? In: Belikova, A., et al. (Eds.), Proceedings of the 2nd Conference on Generative Approaches to Language Acquisition North America. Cascadilla Proceedings Project, Somerville, MA, pp. 200–211.
Juffs, A., 2004. Representation, processing, and working memory in a second language. Transactions of the Philological Society 102 (2), 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., Harrrington, M., 1996. Garden path sentences and error data in second language sentence processing. Language Learning 46, 283–326.
Just, M.A., Carpenter, P.A., 1992. A capacity theory of comprehension: individual differences in working memory. Psychological Review 99 (1), 122–149.

Please cite this article in press as: Sagarra, N., Herschensohn, J., The role of proficiency and working memory in gender and number agreement processing in L1 and L2 Spanish. Lingua (2010), doi:10.1016/j.lingua.2010.02.004

Keating, G.D., 2009. Sensitivity to violations of gender agreement in native and non-native Spanish: an eye-movement investigation. Language Learning 59 (3), 503–535.
King, J.W., Just, A.M., 1991. Individual differences in syntactic processing: the role of working memory. Journal of Memory and Language 30, 580–602.
Kluender, R., Kutas, M., 1993. Subjacency as a processing phenomenon. Language and Cognitive Processes 8 (4), 573–633.
Lardiere, D., 1998. Case and tense in the fossilized steady state. Second Language Research 14, 1–26.
Lardiere, D., 2000. Mapping features to forms in second language acquisition. In: Archibald, J. (Ed.), Second Language Acquisition and Linguistic Theory. Blackwell, Oxford, pp. 102–129.
Lardiere, D., 2007. Ultimate Attainment in Second Language Acquisition: A Case Study. L. Erlbaum, Mahwah, NJ.
Leung, Y.-K.I., 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: Caunt-Nulton, H., Kulatilake, S., Woo, I. (Eds.), BUCLD 31: Proceedings of the 31st Annual Boston University Conference on Language Development. Cascadilla Press, Somerville, MA, pp. 382–393.
Liceras, J.M., Fernández Fuentes, R., Perales, S., Pérez-Tattam, R., Spradlin, K.T., 2008. Gender and gender agreement in bilingual native and non-native grammar. A review from child and adult functional–lexical mixings. Lingua 118, 827–851.
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.
MacDonald, M.C., Just, M.A., Carpenter, P.A., 1992. Working memory constraints on the processing of syntactic ambiguity. Cognitive Psychology 24, 56–98.
Mackey, A., Philp, J., Egi, T., Fuji, A., Tatsumi, T., 2002. Individual differences in working memory, noticing of interactional feedback and L2 development. In: Robinson, P. (Ed.), Individual Differences and Instructed Language Learning. John Benjamins, pp. 181–209.
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 (4), 459–486.
Michael, E., Gollan, T., 2005. Being and becoming bilingual: individual differences and consequences for language production. In: Kroll, J.F., de Groot, A.M.B. (Eds.), Handbook of Bilingualism: Psycholinguistic Approaches. Oxford University Press, New York, pp. 389–407.
Miozzo, M., Costa, A., Caramazza, A., 2000. The absence of a gender congruency effect in Romance languages: a matter of stimulus onset asynchrony? Journal of Experimental Psychology: Learning, Memory, and Cognition 28, 388–391.
Miyake, A., Friedman, N.F., 1998. Individual differences in second language proficiency: working memory as language aptitude. In: Healy, A.F., Bounrne, L.E. (Eds.), Foreign Language Learning: Psycholinguistic Studies on Training and Retention. Erlbaum, Mahwah, NJ, pp. 339–364.
Miyake, A., Just, M.A., Carpenter, P.A., 1994. Working memory constraints on the resolution of lexical ambiguity: maintaining multiple interpretations in neutral contexts. Journal of Memory & Language 33, 125–202.
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 (3), 503–553.
Myles, F., 1995. Interaction between linguistic theory and language processing in SLA. Second Language Research 235–266.
Neubauer, K., Clahsen, H., 2009. Decomposition of inflected words in a second language. Studies in Second Language Acquisition 31, 403–435.
Nicol, J., O'Donnell, M., 1999. Pronominal feature distinctions in English. Rivista di Linguistica 11 (1), 133–163 11.
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.
Prévost, P., 2004. The semantic and aspectual properties of child L2 root infinitives. In: Prévost, P., Paradis, J. (Eds.), The Acquisition of French. John Benjamins, Amsterdam/Philadelphia, pp. 305–332.
Prévost, P., White, L., 2000. Truncation and missing inflection in second language acquisition. In: Friedmann, M.A., Rizzi, L. (Eds.), The Acquisition of Syntax. Longman, London, pp. 202–235.
Rayner, K., Pollatsek, A., 1989. The Psychology of Reading. Prentice Hall, New York.
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 (12), 2030–2048.
Sabourin, L., 2003. Grammatical Gender and Second Language Processing: An ERP Study. Ph.D. Dissertation. University of Groningen, The Netherlands.
Sabourin, L., Stowe, A.L.A., de Haan, G.J., 2006. Transfer effects in learning a second language grammatical gender system. Second Language Research 22, 1–29.
Sagarra, N., 2007. Online processing of gender agreement in low proficient English–Spanish late bilinguals. In: Cabrera, M.J., Camacho, J., Déprez, V., Flores, N., Sánchez, L. (Eds.), Current Issues in Linguistic Theory Series. John Benjamins, Amsterdam, pp. 240–253.
Sagarra, N., Herschensohn, J., 2011. Proficiency and animacy effects on L2 gender agreement processes during comprehension. Language Learning 61 (1).
Sato, M., Felser, C., 2007. Sensitivity to Semantic and Morphosyntactic Violations in L2 Sentence Processing: Evidence from Speeded Grammaticality Judgments. Ms. University of Essex.
Schriefers, H., 1993. Syntactic processes in the production of noun phrases. Journal of Experimental Psychology: Learning, Memory, and Cognition 19, 841–850.
Schriefers, H., Teruel, E., 2000. Grammatical gender in noun phrase production: the gender interference effect in German. Journal of Experimental Psychology: Learning, Memory, and Cognition 26, 1368–1377.
Schwartz, B.D., Sprouse, R., 1996. L2 cognitive states and the Full Transfer/Full Access model. Second Language Research 12, 40–72.
Schütze, C.T., 1996. The Empirical Base of Linguistics: Grammaticality Judgments and Linguistics Methodology. The University of Chicago, Chicago.
Sereno, J.A., Jongman, A., 1997. Processing of English inflectional morphology. Memory and Cognition 25 (4), 425–437.
Sicuro Correa, L.M., Almeida, D.A., de, A., Sobrino Porto, R., 2004. On the representation of Portuguese gender-inflected words in the mental lexicón. Brain and Language 90 (1–3), 63–73.
Silva, R., Clahsen, H., 2008. Morphologically complex words in L1 and L2 processing: evidence from masked priming experiments in English. Bilingualism: Language and Cognition 11, 245–260.
Sorace, A., 1996. The use of acceptability judgements in second language acquisition research. In: Ritchie, W., Bhatia, T.K. (Eds.), Handbook of Second Language Acquisition. Academic Press, San Diego, pp. 375–409.
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.
Tsimpli, I.M., Roussou, A., 1991. Parameter resetting in L2? University College London Working Papers in Linguistics 3, 149–169.
Tunney, R.J., 2005. Sources of confidence in implicit cognition. Psychonomic Bulletin & Review 12, 367–373.
Tunney, R.J., Shanks, D.R., 2003. Subjective measures of awareness and implicit cognition. Memory & Cognition 31, 1060–1071.
Vainio, S., Hyönä, J., Pajunen, A., 2003. Facilitatory and inhibitory effects of grammatical agreement: evidence from readers' eye fixation patterns. Brain and Language 85, 197–202.
Vigliocco, G., Hartsuiker, R.J., Jarema, G., Kolk, H.H.J., 1996. How many labels on the bottles? Language and Cognitive Processes 11, 407–421.

Please cite this article in press as: Sagarra, N., Herschensohn, J., The role of proficiency and working memory in gender and number agreement processing in L1 and L2 Spanish. Lingua (2010), doi:10.1016/j.lingua.2010.02.004

Vos, S.H., Gunter, T.C., Schriefers, H., Friederici, A.D., 2001. Syntactic parsing and working memory: the effects of syntactic complexity, reading span and concurrent load. Language and Cognitive Processes 16, 65–103.
Walter, C., 2004. Transfer of reading comprehension skills to L2 is linked to mental representations of text and to L2 working memory. Applied Linguistics 25 (3), 315–339.
Waters, G.S., Caplan, D., 1996. The measurement of verbal working memory capacity and its relation to reading comprehension. Quarterly Journal of Experimental Psychology 49, 51–79.
Weber-Fox, C., Neville, H.J., 1996. Maturational constraints on functional specializations for language processing: ERP and behavioral evidence in bilingual speakers. Journal of Cognitive Neuroscience 8, 231–256.
White, L., 2003. Second Language Acquisition and Universal Grammar. Cambridge UP, Cambridge.
White, L., Valenzuela, E., Kozlowska-Macgregor, M., Leung, Y.-K., 2004. Gender and number agreement in nonnative Spanish. Applied Psycholinguistics 25, 105–133.
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.
Zagona, K., 2002. The Syntax of Spanish. Cambridge University Press, Cambridge/New York.

Please cite this article in press as: Sagarra, N., Herschensohn, J., The role of proficiency and working memory in gender and number agreement processing in L1 and L2 Spanish. Lingua (2010), doi:10.1016/j.lingua.2010.02.004