Proficiency and Animacy Effects on L2 Gender Agreement Processes During Comprehension

Nuria Sagarra

Pennsylvania State University

Julia Herschensohn

University of Washington

This study examines whether adult second language (L2) learners of an ungendered first language (L1) are sensitive to gender congruency (grammatical feature absent in the L1) and noun animacy (semantic feature present in the L1) when processing L2 gender concord and whether L2 proficiency level determines such sensitivity. To address these questions, 63 Spanish monolinguals and 69 beginning and 64 intermediate Anglophone late learners of L2 Spanish completed a moving-window and a grammaticality judgment task with sentences with gender concord and discord with animate and inanimate nouns. The moving-window data reveal longer reading times in sentences with gender discord than concord and in those with animate than inanimate nouns in intermediates and Spanish monolinguals but not in beginners. Similarly, grammaticality judgments show that intermediates are more accurate in sentences with inanimate than animate nouns and are better than beginners in sentences with gender agreement violations. These results suggest that intermediate learners display targetlike patterns that are more qualitatively similar to those of natives than beginners, both in terms of semantic and grammatical features. In addition, these findings indicate that agreement with animate nouns is cognitively more demanding than with inanimate nouns both for intermediates and Spanish monolinguals, in line with lexical and syntactic accounts of gender.

1

Introduction

Native speakers' knowledge of the language—their underlying competence or representation of the grammar—is complemented by their ability to process and produce speech rapidly and efficiently in real time (cf. McCarthy, 2008). Native procedures develop along with grammatical competence during first language (L1) acquisition (Cutler, 1996; Jusczyck, 1997; Kuhl, 2004), eventually permitting the mature speaker automatized and rapid comprehension and production. Nonnative (L2) adults must likewise develop representational knowledge and procedural routines, but they start their task already possessing the grammar and processing strategies of their L1. Numerous studies show that L2 adults benefit from L1-L2 morphological similarity (see Hopp, 2007; Sabourin, 2003; Sabourin, Stowe, & de Haan, 2006; Sabourin & Stowe, 2008; Tokowicz & MacWhinney, 2005, for examples with gender marking similarities). However, research has not provided clear evidence that L2 adults can acquire L2 properties absent in their L1, and, if they can, what factors explain their persistent difficulty in this domain (see Bruhn de Garavito & White, 2002; Carroll, 1989; Dewaele & Véronique, 2001; Hawkins & Franceschina, 2004; White, Valenzuela, Kozlowska-Macgregor, & Leung, 2004, for examples with gender agreement).

Various proposals have been put forward to account for adult L2 learners' poor performance in morphological representation and computation. In terms of grammatical competence (see Snape, Leung, & Sharwood Smith, 2009, for recent studies), the representational deficit or failed functional features hypothesis (Franceschina 2001b, 2005; Hawkins & Chan, 1997, Hawkins & Franceschina, 2004) maintains that grammatical features such as gender are unavailable to adult L2 learners after a critical age (Hawkins and Franceschina, 2004, stated that it might be 9 years; see Herschensohn, 2007, for definitions and discussion of the notion of a critical period). In contrast, other accounts see no necessary grammatical deficit for adult learners and argue that morphological realization problems are due to various other factors, such as default inflection (Bruhn de Garavito & White, 2002; White, 2003b; White et al., 2004) or difficulties in mapping grammatical features to Phonetic Form (Lardiere, 2000) or (re)assembling features in the L2 (Lardiere 2007, 2008). In terms

of computational accounts, the fundamental identity hypothesis (Hopp, 2007) holds that late learners are capable of gaining nativelike grammatical representation and processing and that L2 adults' errors are attributed to L1 transfer or performance factors (cf. McDonald, 2006). In sum, some representational and computational accounts predict incomplete L2 acquisition in adulthood, whereas others argue that L2 learners can acquire representation and computation that is qualitatively comparable to native grammars. Despite theoretical differences on these points, most L2 scholars agree that the native language (L1) influences representation and processing of the L2, especially initially (cf. Hawkins & Franceschina, 2004; Schwartz & Sprouse, 1996); that learners show more target like abilities with increasing proficiency (cf. White et al., 2004; Vainikka & Young-Scholten, 2007); and that there are similarities between native speakers and L2 learners in computation and representation (cf. Clahsen & Felser, 2006; Hopp, 2007).

An area that has been explored in terms of both representation and processing is L2 gender concord (e.g. Blom, Polišenská, & Unsworth, 2008; Hawkins & Franceschina, 2004; White et al., 2004), a topic of particular interest when the L1 and L2 grammars differ on this point. For example, Anglophones learning L2 Spanish must gain grammatical gender (lexical knowledge), concord (grammatical knowledge), and facility in implementing the knowledge in real-time language use during comprehension and production. On the one hand, most studies on this topic have employed offline techniques to measure underlying grammatical representation, such as accuracy and grammaticality judgments for comprehension and oral errors for production (cf. Bruhn de Garavito & White, 2002; White et al., 2004). On the other hand, the majority of psycholinguistic studies on this topic have used online techniques to measure implicit grammatical computation in real time, such as moving window paradigms, eye-tracking, and event-related potentials (ERPs) (Alarcón, 2009; Gillón Dowens, Barber, Guo, Guo, & Carreiras, 2008; Gillón Dowens, Vergara, & Carreiras, 2004; Keating, 2009; Sagarra, 2007; Tokowicz & MacWhinney, 2005). However, linguistic competence embraces both grammatical knowledge and implicit grammatical computation, with the latter entailing the implementation of grammatical representation in real time (Hopp, 2007; Foucart, 2008; Jin, Åfarli, & van Dommelen, 2007; Juffs & Harrington, 1995, 1996; Marinis, Roberts, Felser, & Clahsen, 2005).

Heeding linguistic competence's dual nature and White's admonition about the need to use multiple tasks because "no methodology allows one to tap linguistic competence directly: in all cases, performance factors will be involved" (2003b, p. 17), we evaluate comprehension data in a moving-window reading

task and a grammaticality judgment task to investigate whether English-Spanish adult learners are sensitive to gender agreement violations in L2 Spanish. Because metalinguistic knowledge is often invoked when making grammaticality judgments, it is important to use grammaticality judgments in combination with online measures within the same population (cf. Blom, Marinis, Vasic, & Chondrogiannni, 2009; Ellis, 2005). This use of online and offline measures allows us to both account for the limitation of a single measure and to complement previous studies exclusively based on grammaticality judgments (which have been deemed inadequate as single measures; cf. Birdsong, 1989; Sorace, 1996). We examine the grammatical phenomenon of adjective agreement (concord) in Spanish, considering the role of proficiency level in nonadvanced learners. Based on behavioral (e.g., White et al., 2004; White, Valenzuela, Kozlowska-Macgregor, Leung, & Ben-Ayed, 2001) and neurocognitive (e.g., Gillón Dowens et al., 2008) studies, we predict that proficiency can help with the attainment of targetlike patterns of representation and processing of gender agreement in L2 Spanish.

Finally, we investigate whether intermediates also show monolingual-type behavioral patterns when processing gender agreement with animate nouns and we explore the relationship between natural and grammatical gender agreement, whether noun animacy affects agreement processes and, if it does, whether it facilitates (due to gender correlating with biological sex and L1 transfer) or it inhibits them (due to animate nouns' greater cognitive demands). The latter approach (i.e., natural gender agreement is more difficult to process than grammatical gender agreement) follows single-resource theories of attention (e.g., Baddeley, 2003; Just & Carpenter, 1992), holding that we have a limited capacity to process and store incoming information during complex cognitive tasks. 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 and that these processing demands affect knowledge and implementation of linguistic information, such as gender agreement (e.g., Hopp, 2007; Lardiere, 2007). Furthermore, the tenets of full listing and full parsing models (see Barber & Carreiras, 2005; Domínguez, Cuetos, & Segui, 1999; Igoa, García-Albea, & Sánchez-Casas, 1999, for a review) predict that transparent animate nouns are more difficult to access than transparent inanimate nouns: If gender is a stem inherent feature that is accessed from the full word form in Spanish, the processor will need more time to retrieve *esposo* "spouse-m" than *libro* "book-m" because *esposo* will, in turn, activate the referent *esposa* "spouse-f"; similarly, if gender is a morphological feature that combines with the stem

of the word in Spanish, the system will require more time to decide between the masculine and feminine suffixes than single-gendered inanimate nouns. Assuming that we have a limited capacity to process language, that linguistic cognitive demands affect L2 acquisition, and that accessing animate nouns is cognitively more demanding than accessing inanimate nouns, we predict that adult L2 learners of Spanish will have more difficulty processing gender agreement with animate than inanimate nouns.

Background

Gender Concord in Spanish DPs

Although Spanish has a binary gender system (masculine or feminine) for nouns, determiners, adjectives, and pronouns (Roca, 1989), only nouns have intrinsic gender (Carroll, 1989; DeWaele & Véronique, 2000, 2001). For animate nouns, gender may be assigned conceptually and allow for determining the biological sex of the referent (natural or semantic gender; but not always; cf. Harris, 1991) (Ibrahim, 1973); however, for both animate and inanimate nouns, gender must be assigned morphosyntactically. Gender is overtly marked on most animate and inanimate nouns with the morphological inflections /–o/ for masculine (*cuaderno* "notebook$_{masc}$") and /–a/ for feminine (*casa* "house$_{fem}$") (transparent gender) (Green, 1988). However, some nouns carry opaque gender, including masculine and feminine nouns ending in a consonant or in /–e/ (*señal* "sign$_{fem}$"; *diente* "tooth$_{masc}$"), masculine nouns ending in /–a/ or /–ema/ (*mapa* "map$_{masc}$"; *tema* "topic$_{masc}$"), feminine nouns ending in /–o/ (*mano* "hand$_{fem}$"), and gender-ambivalent nouns (*burócrata* "bureaucrat$_{masc/fem}$").

Gender of determiners, adjectives, and pronouns varies according to the gender of the noun they modify or substitute (*el cuaderno pequeño* "the$_{masc}$ small$_{masc}$ notebook$_{masc}$"; *la mesa pequeña* "the$_{fem}$ small$_{fem}$ table$_{fem}$") (Carroll, 1999; Zagona, 2002). Adjectives are overtly marked for gender (e.g., *pequeño/a* "small$_{masc,fem}$"), except for those that can be used invariably for masculine or feminine (e.g., *inteligente* "intelligent$_{masc,fem}$"). Like Spanish, English nouns are semantically characterized as [+/− animate], and animate nouns may be semantically [+/− feminine], a lexical feature that can be grammaticalized in English pronouns. In contrast with Spanish, English nouns do not carry syntactic gender features and do not agree with determiners or adjectives. Therefore, Anglophone learners of L2 Spanish must gain the gender feature for inanimate nouns and concord on determiners and adjectives.

UG Approaches to L2 Gender Agreement

We briefly review Universal Grammar (UG) premises, as a substantial amount of recent research on the topic of L2 gender has used this framework. UG approaches view grammatical competence as the realization of lexical categories (e.g., verbs, nouns, adjectives) through their implementation by functional categories (e.g., auxiliaries, determiners, complementizers). Properties of lexical and functional categories (e.g., tense, gender) are known as *features*, which are labeled as "interpretable" or "uninterpretable" (Chomsky, 1995). Interpretable features are lexical properties that are semantically necessary for interpretation such as [+/− past] and [+/− feminine]; uninterpretable features are grammatical properties whose implementation varies according to the language and that may result in agreement such as gender concord of adjectives with the head noun in Spanish. The operation Agree (Chomsky, 1995) matches interpretable and uninterpretable features, deleting the uninterpretable grammatical ones. In the case of nouns, the interpretable features of number and gender check and delete uninterpretable gender and number features of determiners and adjectives within the noun phrase (Carstens, 2000).[1] Although adjectives and determiners in Spanish are morphologically marked for gender and number, those characteristics are simply a reflection of the features of the noun (the interpretation of the entire phrase is determined by its head).

Representational deficit (RD) approaches view surface morphology errors as diagnostic of syntactic deficits and view parameterized L2 functional features as not acquirable after a critical period (cf. Tsimpli & Papadopoulou, 2009; Tsimpli and Roussou, 1991). L1 values of uninterpretable functional features remain available to adults, but L2 parametric values absent in the L1 cannot be acquired; inflectional errors in later L2 acquisition are therefore due to a syntactic deficit in underlying competence of L2 adult learners. Accordingly, Anglophone learners of L2 Spanish are incapable of gaining grammatical gender and agreement on determiners and adjectives through syntactic means. Hawkins (1998) and Franceschina (2001b) interpreted near-native late English-French and English-Spanish bilinguals' adoption of default masculine determiners and adjectives in elicited production data as confirmation that they had established gender on the basis of phonology and that they had not yet established uninterpretable features for the determiners and adjectives, suggesting that successful acquisition of L2 gender depends on the presence of gender features in the L1. In studies comparing L2 learners of Spanish with gendered (Italian) and ungendered (English) L1s immersed in a Spanish-speaking environment, Franceschina (2001a, 2002) and Hawkins & Franceschina (2004) considered the presence of fewer agreement errors in judgment and production

tasks in the [+ gen] L1 group as further evidence of a representational deficit. Given these assumptions, Anglophone L2 learners should not show targetlike behavior in a gender concord task.

The work of other UG scholars (Bruhn de Garavito & White, 2002; Herschensohn, 2001; Prévost, 2003; White et al., 2004) sees no necessary postcritical period syntactic deficit but rather views L2 morphosyntactic imperfections as related to mapping of morphology to phonetic form (e.g., Lardiere, 2000), default and null inflection (e.g., Prévost & White, 2000), or L1 prosodic interference (e.g., Goad & White, 2006). This research draws on full transfer full access (FTFA; Schwartz & Sprouse, 1996) and missing surface inflection accounts (Haznedar & Schwartz, 1997; Lardiere, 1998, 2000; Prévost & White, 2000). Learners initially transfer L1 morphosyntactic settings but eventually may reset parameters to L2 values through gradual restructuring induced by an inability to parse primary linguistic data. Syntactic competence is not directly reflected by mastery of morphological inflection because mistakes relate to matching difficulties between syntactic terminal nodes and surface morphology. There is no critical period functional deficit to examine because L2 acquisition is similar for children and adults. Accordingly, Anglophone learners of L2 Spanish initially transfer L1 properties (no gender on determiners or adjectives and no agreement) but may reset nominal features to L2 values, eventually gaining gender on determiners and adjectives with agreement. For example, Bruhn de Garavito and White (2002) reported that beginning and intermediate French-Spanish late learners were highly accurate at producing determiner-noun gender and number agreement, and White et al. (2004) found that intermediate and advanced French-Spanish and English-Spanish late learners were as accurate as native Spanish speakers on perceiving and producing gender and number agreement in Spanish DPs (for further evidence that learners produce the same errors independently of their L1, see Bartning, 2000; Dewaele & Véronique, 2000, 2001; Fernández-García, 1999).

Universal Grammar approaches may differ in their view of the abstract grammatical features of the final state L2 grammar, but they agree on other points. Learners are influenced in their L2 by properties of the L1, both lexical and functional (interpretable and uninterpretable features), and L2 interpretable features are acquirable. However, "the nature of ultimate attainment in a second language is partly, but not wholly, L1-determined" (Hawkins & Hattori, 2006, p. 295). From different theoretical perspectives, L2 grammatical features may be acquirable (e.g., FTFA) or they may not be by adults (e.g., RD), but in any case, L2 learners can approximate appropriate responses through a variety of means (e.g., alternative resources available through UG or input frequency).

Indeed, currently many UG scholars see non-UG resources as means of bootstrapping the L2 grammar (cf. Herschensohn, 2007, 2009) and suggest that "the processes of the language faculty (broadly conceived) are not exclusive to language" (Bley-Vroman, 2009, 185), for L1s and L2s draw on similar resources in acquisition and processing. The UG debate over whether there is a representational deficit or not is peripheral to the main goals of this article, which examines online and offline responses to Spanish gender nonagreement, comparing natives with L2 learners of two proficiency levels.

Grammatical Processing in L2

Two recent approaches to L2 processing point out differences and similarities between native language and L2 grammar computation (Clahsen & Felser, 2006; Hopp, 2007). Clahsen and Felser's (2006) shallow structure hypothesis—although admitting both shallow and deep processing for native speakers—proposes that late L2 learners prefer shallow processing, which relies more on lexical, semantic, and pragmatic cues as opposed to those related to a syntactic hierarchy. Comparing adult native processing with that of native children and adult L2 learners, the authors argued for a "continuity hypothesis" for native adults and children but for less detailed ("shallow" or deficient) grammatical processing by late L2 learners. For example, in a task of speeded production of German participles (Clahsen, Hadler, & Weyers, 2004), both children and adult natives showed a distinction between regular and irregular verbs. The children, however, displayed higher overregularization and longer production latencies, characteristics attributed to slower and less accurate lexical access by the younger natives. Likewise, in ambiguity resolution, monolingual children used the same structure-based, least effort processing principles as monolingual adults, whereas adult L2 learners favored lexical and pragmatic cues over syntactic ones to disambiguate sentences. Based on these findings, they concluded that adult L2 learners do not show nativelike processing of gender concord even after long periods of immersion. It is unclear how this relates to syntactic hierarchy in terms of adjacent nouns and adjectives.

Hopp's (2007) fundamental identity hypothesis argues that late bilinguals are capable of gaining grammatical representation and processing that is qualitatively similar to that of native speakers. According to Hopp, nonnative differences from native processing are attributed to L1 transfer or performance factors. Hopp (2006, 2007) conducted a series of experiments in native and L2 German on scrambling in embedded clauses, a syntactic movement that allows a reordering of constituents (e.g., SOV to OSV). The 2006 study—using two experiments to compare L1 English and L1 Dutch learners of L2 German

with natives—found nativelike parsing routines for the near-natives (the most advanced proficiency level) of both L1s. Hopp concluded that

> [t]he present findings do not support the hypothesis that L2 processing is "shallower" than native-language processing such that L2 speakers are sensitive to lexical-semantic and pragmatic information, yet make only very restricted or no use of syntactic or phrase-structural information in parsing. (2006, p. 391)

These approaches to L2 computation indicate that processing measures such as reaction time (RT) or eye movement records can assess real-time implementation of grammatical knowledge, thus providing a window on linguistic competence of both native speakers and L2 learners.

The Gender Congruency Effect in Adult L2 Learners

The gender congruency effect refers to the result that gender-marked items other than nouns have on processing: For native speakers of gendered languages, congruent gender concord speeds up and incongruent concord slows down processing of the noun and its modifiers. The redundant gender feature on adjectives is not then totally superfluous but rather is facilitating to rapid comprehension. The gender congruency effect is widely documented with monolinguals both with behavioral data (Antón-Méndez, Nicol, & Garrett, 2002; Cacciari and Padovani, 2007; Colé & Segui, 1994; Schriefers, 1993; Schriefers & Teruel, 2000; Vainio, Hyönä, & Pajunen, 2003; but see Miozzo, Costa, & Caramazza, 2000, for behavioral evidence that the gender congruency effect is obtained in Romance but not Germanic languages) and neurocognitive data (Barber & Carreiras, 2005; Foucart, 2008; Gillón-Dowens et al., 2004; Gunter, Friederici, & Schriefers, 2000; Hagoort & Brown, 1999). However, it is not clear whether it is also present in adult L2 learners of ungendered L1s.

For example, Guillelmon and Grosjean (2001) found no gender congruency effect in the way English-French advanced late bilinguals processed determiner-noun concord and discord in a gating experiment (in which a determiner + noun was incrementally revealed). However, a gender congruency effect was present in English-French early bilinguals and French monolinguals (cf. Grosjean, Pommergues, Cornu, Guillelmon, & Besson, 1994): Both groups showed faster recognition of nouns preceded by a correct gendered determiner and slower recognition for ungendered determiners and slowest for wrong gendered determiners. In the same line, Lew-Williams and Fernald's (2007) eyetracking study revealed that late L2 learners do not use the morphosyntactic (gender)

information of determiners to expedite the processing of gender agreement like natives do.

Other online studies demonstrate the gender congruency effect in late L2 learners even when their L1 lacks this grammatical feature, supporting similarity to nativelike L2. For instance, Foucart's (2008) eye movement records and ERPs showed the congruency effect in advanced English-French late learners, and such an effect has been reported in numerous behavioral and neurocognitive studies with other types of advanced learners whose L1 lacks gender concord (e.g., Alarcón, 2006; De Mulder, 2006; Foucart & Frenck-Mestre, 2004; Gillon-Dowens et al., 2004; Herschensohn & Frenck-Mestre, 2005; Keating, 2009; Sabourin et al., 2006). Although highly proficient learners are sensitive to gender agreement violations, no gender congruency effect has been found in low-proficient learners (behavioral data: Keating, 2009; Sagarra, 2007; neurocognitive data: Osterhout, Poliakov, Inoue, McLaughlin, Valentine, et al., 2008; Tokowicz & MacWhinney, 2005). In the same line, Alarcón's (2009) RT data indicate that sensitivity to gender agreement violations increases with proficiency level. These studies indicate the importance of proficiency to online computation in a nativelike fashion and suggest that processing studies need to consider the level of mastery as an important variable. One would expect that low-proficiency L2 learners would show less nativelike processing than those with high proficiency, but most of the studies have not used low-proficiency learners. To address this limitation, the current study includes low- and higher proficiency late L2 learners of Spanish. Our beginning and intermediate L2 Spanish learners provide a counterpoint to studies of very advanced learners (e.g. Hopp, 2007) and give us the opportunity to consider their gender concord in pseudolongitudinal terms.

The Noun Animacy Effect in Adult L2 Learners
The online studies described above focused on gender agreement with inanimate nouns. Gender in inanimate nouns is assigned arbitrarily and contrasts with animate gender, where gender is connected to biological sex. There is mounting neurocognitive evidence that monolinguals are sensitive to gender agreement violations (P600 effect) with both animate nouns (English: Osterhout, Bersick, & McLaughlin, 1997; Osterhout & Mobley, 1995; German: Schmitt, Lamers, & Münte, 2002; Hebrew: Deutsch & Bentin, 2001; Spanish: Barber, Salillas, & Carreiras, 2004; Corral, Barber, & Carreiras, 2008; Demestre, Meltzer, García-Albea, & Vigil, 1999) and inanimate nouns (Dutch: Hagoort & Brown, 1999; German: Gunter et al., 2000; Spanish: Barber et al., 2004). The absence of the N400 effect (semantic anomaly) in Spanish suggests that gender concord is

processed syntactically. Nevertheless, a range of studies indicates that animate and inanimate nouns may show differential effects in computation. Behavioral studies on gender agreement in Spanish monolinguals provide further evidence of this claim. For example, Igoa et al. (1999) asked Spanish monolinguals to perform a sentence completion task with complex noun phrases (using morpheme exchanges such as *unos gatos de la niña → una niña de los gatos* "some-m-pl cats-m-pl of a-f-sg girl-f-sg") and found that the subjects produced more errors with animate than inanimate nouns. In contrast, other studies with Spanish monolinguals argue that animate nouns are easier to process than inanimate nouns because their gender correlates with biological sex. For example, behavioral studies show both shorter RTs at adjectives (Alarcón, 2009) and fewer agreement production errors (Antón-Méndez, 1999; Vigliocco & Franck, 1999) in sentences with animate than inanimate head nouns. Finally, there are Spanish monolingual data showing no differences in the processing of agreement with animate and inanimate head nouns (Alarcón, in press; Antón-Méndez et al., 2002; Barber et al., 2004; Corral et al., 2008). In sum, Spanish monolingual data on the effects of noun animacy on agreement are inconclusive.

With regard to L2 Spanish data, some studies show more agreement production errors with animate nouns (see Bruhn de Garavito & White's, 2002, results with low-proficiency English-French learners), whereas others report fewer agreement production errors with animate nouns (see Finneman's, 1992, and Fernández-García's, 1999, findings with English-Spanish learners). Particularly relevant for the present research, the only study investigating the processing of agreement with animate and inanimate nouns in L2 Spanish with an online technique, Alarcón (in press), indicates shorter RTs at adjectives preceded by animate head nouns but longer RTs at adjectives preceded by animate attractor nouns (attractor noun refers to the second noun in the NP), suggesting that animate nouns proximal to the adjective are easier to process than inanimate nouns (see Keating, 2009, for further evidence that proximity between noun and modifiers affects L2 processing). The absence of both a [+ proximity] condition in Alarcón's study and a [+ animate] gender agreement in Keating's study begs the question of whether L2 learners treat [+/− animate] agreement within the DP alike. Our study addresses this issue and combines online and offline techniques.

The Study

As discussed in the literature review, previous research has not given definitive answers to the question of whether late L2 learners of ungendered L1s maintain

long-term deficits in grammatical representation and computation. To assess the development of grammatical gender concord in L2 Spanish, the current study uses online (moving-window task) and offline (grammaticality judgment task) data to examine the grammatical feature of noun-adjective gender concord in Spanish native speakers and beginning and intermediate English-Spanish late learners. Our study asks the following research questions:

1. Do Spanish monolinguals and intermediate and beginning L2 learners show gender congruency effects with animate and inanimate nouns, based on online and offline data?
2. Do Spanish monolinguals and intermediate and beginning L2 learners process gender agreement with animate and inanimate nouns differently, based on online and offline data? If they do, does noun animacy facilitate or inhibit agreement processes?

Our predictions for our first research question, whether L2 computation shows qualitative similarity to native behavioral patterns of concord/discord or whether it reflects L1 lack of gender concord, are that Spanish monolinguals and intermediate learners will be sensitive to gender agreement violations (longer RTs and lower accuracy on the grammaticality judgments), whereas beginning learners will not be (no differences between gender concord and discord conditions). This prediction follows targetlike computation accounts, which assume that gender features can be acquired after puberty independently of the L1 (Hopp, 2007; Foucart & Frenck-Mestre, 2004).

Our second research question examines whether noun animacy affects adjectival computation for Spanish monolinguals and the two groups of L2 learners. If it does, a related question is whether noun animacy facilitates processing due to the correlation between grammatical and biological sex and, in the case of L2 learners, also L1 transfer or whether it hinders processing due to greater cognitive demands. We hypothesize that Spanish monolinguals and intermediates will show a similar behavioral pattern—namely longer RTs and lower accuracy on the grammaticality judgments at sentences with animate than inanimate nouns, because animate nouns are cognitively more taxing than inanimate nouns. We base our prediction on previous research with Spanish monolinguals (e.g., Barber et al., 2004; Corral et al., 2008; Demestre et al., 1999; Igoa et al., 1999) and Anglophone and Francophone late learners of Spanish (e.g., Alarcón, in press, with attractor nouns; Bruhn de Garavito & White, 2002). In a similar vein, we expect that the accuracy of the grammaticality judgment will converge

with that of monolinguals for intermediates but not beginners and that monolinguals and intermediates will be less accurate in sentences with animate than inanimate nouns.

We employed both online (moving-window task) and offline (grammaticality judgment task) techniques in order to account for the dual nature of linguistic competence (grammatical knowledge and implicit grammatical computation): Grammatical representation is difficult to infer from comprehension and production—as it is impossible to know where competence ends and performance begins—whereas behavioral results do allow us to compare L2 learners and monolinguals directly to ascertain computational procedures.[2] This can also explain why previous research on L2 gender agreement has shown differences between RT and grammaticality judgment data (De Mulder, 2006; Montrul, Foote, & Perpiñán, 2008). Finally, we chose beginning and intermediate learners because studies on the representation and computation of L2 gender agreement in low-proficiency levels are few, and we examined noun-adjective gender concord because this type of agreement has been studied much less (cf. Hernández-Pina, 1984, on child acquisition) than noun-determiner concord.

Method

Participants

The participants were university students (69 beginning and 64 intermediate Spanish L2 learners) and 63 Spanish monolinguals who received extra credit or 10 euros, respectively, for participating in the study. The L2 learners were English speakers with no knowledge of languages other than Spanish who began learning Spanish after age 14 and were currently enrolled in third- (beginners) or seventh/eighth- (intermediates) semester Spanish courses in an American university. In order to be included in the study, the L2 learners had to score below or above three standard deviations from the mean of the language proficiency test and obtain a perfect score on the gender and vocabulary tests. The Spanish monolinguals had always lived in a Spanish monolingual community and had only studied one foreign language—English—and their proficiency level in this language was too low to have any effect on L1 processing. Spanish monolingual data were collected in southern Spain. Finally, all participants needed to complete all tests, score above 60% in the comprehension questions of the moving-window task, and not have lived in a foreign-language environment abroad for more than 1 month.

Materials and Procedure

Participants completed two screening tests (a language history questionnaire and, in the case of L2 learners, also a Spanish proficiency test) and a moving-window task in 1 hr. One week later, they performed a grammaticality judgment task and two additional screening tests (a vocabulary and gender agreement test) in 1 hr. The grammaticality judgment task and the vocabulary and gender tests were presented after the moving-window task so that the explicitness of the former did not bias the implicit processing of the latter.

Screening Tests

The language history questionnaire revealed that none of the participants had lived in a foreign-language environment abroad for more than 1 month and that the L2 learners received a similar amount of L2 instruction (beginners: three semesters; intermediates: eight semesters). The questionnaire also showed that the Spanish monolinguals had limited English knowledge, judging by their lack of exposure to English outside of the mandatory 3-hr/week courses offered in middle school and high school (Spain has one of the lowest L2 English proficiency levels of the European Union; Bonnet, 2002; CEDEFOP, 2004) and by their low self-ratings (reading: $M = 4.21$, $SD = 1.18$; listening: $M = 3.13$, $SD = 1.25$; writing: $M = 3.84$, $SD = 1.15$; speaking: $M = 3.40$, $SD = 1.36$, all out of 10 points).

For the Spanish proficiency test, L2 learners completed the grammar section of the Diploma de Español como Lengua Extranjera (intermediate level) and a *t*-test for independent samples showed that the intermediates ($M = 20.65$, $SD = 3.00$) were more proficient than the beginners ($M = 7.04$, $SD = 2.89$): $t(125) = -22.667$, $p < .01$ (Levene's $F = 0.000$, $p > .05$). Participants completed a vocabulary and a grammar test to control for familiarity with the meaning of the target nouns and adjectives and with the inherent gender of the target nouns. For the vocabulary test, L2 learners matched the target nouns and adjectives with their corresponding translation in English to control for familiarity with the target meaning. Finally, for the grammar test, the learners indicated the gender and number of a list of Spanish nouns without a determiner. One fourth of the nouns were masculine singular nouns that appeared in the experimental sentences and the rest (one fourth masculine plural nouns, one fourth feminine singular nouns, one fourth feminine plural nouns) were nouns that appeared in the filler sentences with correct gender and number agreement (see the next section for an explanation of the use of masculine singular nouns).

Moving Window Task and Grammaticality Judgment Task

These two tasks were written and self-paced because adult learners perform better on written than oral tasks (Montrul et al., 2008) and on self-paced than timed tasks (Sabourin, 2003). The use of complex tasks could consume attentional resources to the point of hindering sensitivity to redundant and unsalient morphological information in adjectives. For each task, participants completed practice sentences containing grammatical and ungrammatical sentences (in line with the experimental and filler sentences), 40 experimental sentences (10 per condition), and 70 filler sentences (60 well formed, 10 with number disagreement). All sentences were controlled for length and L2 level (vocabulary and grammar adequate for students in their third semester of study of Spanish). The sentences were randomized to avoid having two experimental sentences appear consecutively and participants were not told that some of the sentences were ungrammatical not to bias the results. Approximately two thirds of the sentences were fillers following standard quantitative psycholinguistic experiments to make the goal of the study less obvious. Four conditions were created for each experimental sentence: two for animate gender (agreement, disagreement) (1) and two for inanimate gender (agreement, disagreement) (2).

(1) *El abuelo   adora al nieto   nuevo/*nueva   de la familia.*
    The grandfather adores the grandson$_{masc}$ new$_{masc}$/*new$_{fem}$ of the family.
    "The grandfather adores the family's new grandson."

(2) *El abuelo   adora el auto   nuevo/*nueva   de la familia.*
    The grandfather adores the car$_{masc}$ new$_{masc}$/*new$_{fem}$ of the family.
    "The grandfather adores the family's new car."

   Experimental sentences followed the same syntactic structure, and the target NP contained two- to four-syllable masculine singular countable nouns with transparent gender plus a two- to four-syllable descriptive adjective.[3] Gender transparent masculine nouns were chosen for several reasons, as we used the unmarked forms of nouns (masculine: Harris, 1991; transparent: Antón-Méndez, 1999; singular: Eberhard, 1997) exclusively. First, we noted that previous research has shown that a noun's gender does not influence RTs or grammaticality judgments of determiners or adjectives in monolinguals or late bilinguals (Italian: De Mulder, 2006; Spanish: Alarcón, in press; Antón-Méndez et al., 2002; Keating, 2009). In our experiment, we wanted to create a solid and straightforward baseline to investigate gender agreement/disagreement. Additional comparisons such as masculine/feminine, transparent/opaque, or singular/plural could have biased our results and decreased k, yielding poor statistical power. Furthermore, additional comparisons would have increased the total number of

sentences participants had to read, making the experiment logistically unfeasible. Finally, because we are examining the low-proficiency end of the learner spectrum, we wanted to look at an area in which grammaticalization might first appear. Because the sole use of transparent masculine singular nouns prevents us from discerning between sensitivity to lexical idiosyncracies (opaque gender), markedness (feminine), and disagreement, 30 of the well-formed fillers had noun-adjective combinations with feminine singular, masculine plural, and feminine plural nouns with transparent gender (10 per noun type). Statistical analyses based on the moving-window data revealed longer RTs at adjectives in sentences with gender agreement violations (intermediates, Spanish monolinguals) than in those with gender and number agreement regardless of the gender and number of the noun, indicating that any reading latencies in the experimental sentences would be due to sensitivity to ungrammaticality rather than markedness.

In the moving-window task, participants read Spanish sentences silently on a computer screen, word-by-word, and answered yes-no comprehension questions after each sentence (target yes-no answers were equally distributed across sentences). For each sentence, participants looked at a 500-ms fixation sign ($+$) and saw dashes indicating where the words would appear, as illustrated in Figure 1 (the dashes helped make reading more natural); they pressed a space bar to read the first word of the sentence; each subsequent time they pressed the spacebar key, the next word appeared and the prior word disappeared until the end of the sentence; then they pressed the space bar to read a comprehension question based on the sentence they had just read and answered by pressing a "yes" or a "no" key. Half of the sentences (experimental and fillers) contained a yes question and half contained a no question. Examples of yes and no questions for sentences (1) and (2) above are *¿El abuelo adora el auto/al nieto nuevo?* "Does the grandfather adore the new car/grandson?" (answer: Yes) and *¿El hombre detesta el auto/al nieto nuevo?* "Does the grandfather detest the new car/grandson?" (answer: No). Therefore, the questions were based on the content of the sentence (not including the adjective) and hence did not evaluate the learner's knowledge of gender marking or agreement. This decision was made to avoid biasing participants' attention to the structure under investigation.

In the grammaticality judgment task, participants read sentences in Spanish, decided whether each sentence was correct or incorrect (if incorrect, they circled the incorrect word[s]) and indicated their self-confidence in their answers (see Schütze, 1996, for a discussion of the benefits of using multiple judgment tasks). Confidence judgments were based on a 5-point Likert scale ranging from 1 (not sure at all) to 5 (completely sure). For comparability purposes, the
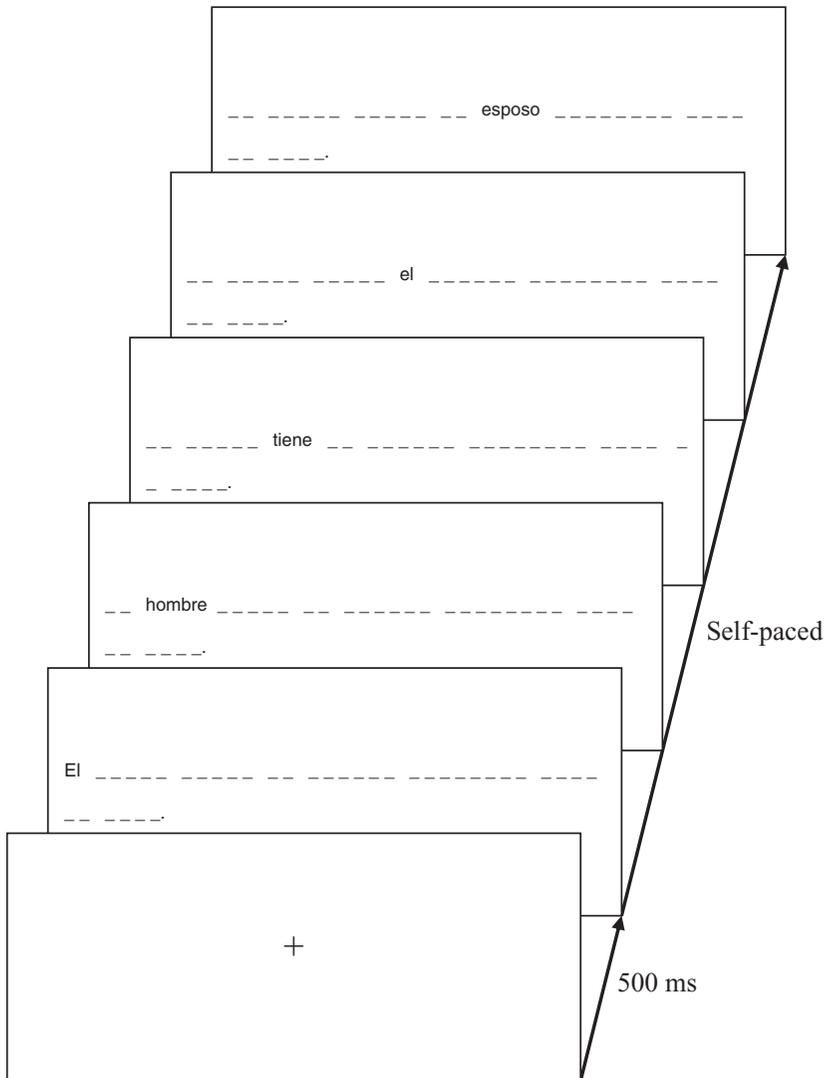
Figure 1 Example of a trial in the moving-window task.

moving-window and grammaticality judgment tasks contained the same number and type of randomized sentences, but the content and noun-adjective combinations were different to avoid possible practice effects that would increase sensitivity to gender agreement violations in the grammaticality judgment task.

Scoring

The Spanish proficiency, gender agreement, and vocabulary tests received 1 point per correct answer. The moving-window task generated two scores: mean RTs on the adjectives and accuracy on comprehension questions. Mean RTs were the mean of all word RTs per condition. Word RTs consisted of the time spent between the appearance of a word on the screen and the press of a spacebar key, which had to be between 200 and 2000 ms (according to Rayner & Pollatsek, 1989, Anglophone college students need between 225 and 300 ms to process single words). For the second moving-window score (accuracy on comprehension questions), 1 point was assigned to correct answers and 0 to incorrect ones, and participants needed to respond accurately above chance on all questions to be included in the study. It is important to note that statistical analyses were based only on sentences with correct responses to comprehension questions in order to explore processing of gender agreement during reading for comprehension. Finally, the grammaticality judgment task produced two scores: accuracy on classifying sentences as correct or incorrect and a confidence rating. The grammaticality judgment accuracy score was based on a binary score of 1 point for correct answers (i.e., identifying correct sentences as correct or identifying incorrect sentences as incorrect and identifying the source of the error accurately) and 0 for incorrect ones, for a total of 10 points (1 point per correct sentence). The confidence rating score of the grammaticality judgment task followed a 5-point continuous score ranging from 1 (not sure at all) to 5 (completely sure) and only confidence ratings of sentences with correct grammaticality judgments were included in the statistical analyses.

Results

Moving-Window Task

The means and standard deviations for RT and accuracy scores on the comprehension questions are shown in Tables 1 and 2, respectively. The high means of the total number of correct responses on comprehension questions across conditions and groups shown in Table 2 confirm that all participants understood the meaning of the sentences and that longer RTs at the adjective were not due to lack of understanding. Statistical analyses included four repeated-measures ANOVAs with a 2 (agreement: gender agreement vs. gender disagreement) × 2 (animacy: animate vs. inanimate nouns) × 3 (group: beginners, intermediates, 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 comprehension questions. Mean reading

Table 1 Mean RTs at the word preceding the adjective, at the adjective, and at the word following the adjective

| | [+ animacy] | | | | [− animacy] | | | |
| | [+ agreement] | | [− agreement] | | [+ agreement] | | [− agreement] | |
	M	SD	M	SD	M	SD	M	SD
Word before adj								
Beginners	1033.97	450.69	1020.69	375.66	986.32	342.45	986.05	315.49
Intermediates	866.78	392.79	832.08	392.79	890.84	314.95	835.41	291.43
Monolinguals	503.72	216.58	505.54	220.83	469.35	168.09	492.75	175.23
Adjective								
Beginners	930.49	344.14	943.14	311.37	901.32	332.83	911.36	310.13
Intermediates	744.49	225.15	913.74	381.13	708.83	207.43	823.82	276.62
Monolinguals	499.52	157.87	614.47	264.41	465.34	115.06	554.96	189.89
Word after adj								
Beginners	546.83	133.04	546.18	134.76	550.52	144.00	530.54	120.70
Intermediates	479.76	106.78	569.45	157.26	445.70	85.48	536.43	137.73
Monolinguals	420.08	94.38	476.71	106.34	385.97	70.99	449.71	111.25

Note. n = 69 for beginners, n = 64 for intermediates, and n = 63 for Spanish monolinguals.

Table 2 Accuracy on comprehension questions

	[+ animacy]				[− animacy]			
	[+ agreement]		[− agreement]		[+ agreement]		[− agreement]	
	M	SD	M	SD	M	SD	M	SD
Beginners	8.58	1.01	8.26	1.21	8.67	1.23	8.57	1.04
Intermediates	8.45	1.37	8.40	1.27	8.80	1.28	8.72	1.29
Monolinguals	9.13	0 .75	8.43	1.09	9.49	0.64	8.84	0.99

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

times at the word immediately preceding the adjective were analyzed to ensure that latencies on the adjective were exclusively due to the variables under investigation. In turn, mean reading times at the word immediately following the adjective were investigated to measure possible residual effects in processing the adjective.

Mean RTs at the Word Preceding the Adjective
The ANOVA conducted on the word preceding the adjective showed a lack of main effect for agreement, $F(1, 193) = 1.499, p > .05$, and animacy, $F(1, 193) = 2.284, p > .05$, and a lack of interaction of Agreement × Animacy $F(1, 193) = 0.226, p > .05$), agreement × Group, $F(2, 193) = 2.132, p > .05$, Animacy × Group, $F(2, 193) = 0.951, p > .05$, and Agreement × Animacy × Group, $F(2, 193) = 0.073, p > .05$. These results indicate that mean RTs on the word immediately preceding the adjective were equal across conditions and that latencies on the adjective were caused by the variables under investigation. The significant main effect obtained for Group, $F(2, 193) = 64.075, p < .01$ (beginners were slower than intermediates and both were, in turn, slower than Spanish monolinguals), is explained by L2 learners processing the target language more slowly than adult native speakers (Clahsen & Felser, 2006) due to a lack of L2 automaticity (Segalowitz, 2003).

Mean RTs at the Adjective and at the Word Following the Adjective
The ANOVAs conducted on the adjective and the word following the adjective (a preposition) revealed a significant main effect for Agreement [adjective: $F(1, 193) = 61.968, p < .01$; preposition: $F(1, 193) = 59.870, p < .01$], Animacy [adjective: $F(1, 193) = 11.236, p < .01$; preposition: $F(1, 193) = 19.203$,

$p < .01$], and Group [adjective: $F(2, 193) = 47.797$, $p < .01$; preposition: $F(2, 193) = 24.608$, $p < .01$], as well as a significant interaction of Agreement × Group [adjective: $F(2, 193) = 10.198, p < .01$; preposition: $F(2, 193) = 25.005, p < .01$]. However, there was no interaction of Agreement × Animacy [adjective: $F(1, 193) = 0.863, p > .05$; preposition: $F(1, 193) = 0.094, p > .05$], Animacy × Group [adjective: $F(2, 193) = 0.337, p > .05$; preposition: $F(2, 193) = 1.616, p > .05$], or Agreement × Animacy × Group [adjective: $F(2, 193) = 0.901, p > .05$; preposition: $F(2, 193) = 0.441, p > .05$].

Multiple-contrast comparisons revealed both a gender congruency effect and a noun animacy effect in intermediates and Spanish monolinguals but not in beginners. The gender congruency effect in intermediates and Spanish monolinguals was evident from the longer RTs at the adjective and the preposition in sentences with gender violations than with gender agreement, regardless of noun animacy (all $p < .01$). Similarly, the noun animacy effect in intermediates was confirmed by the longer RTs at the adjective in grammatical sentences with animate than inanimate nouns ($p < .05$ and this effect dissipated in the preposition, where $p > .05$), as well as by longer RTs at the preposition in ungrammatical sentences with animate than inanimate nouns ($p < .01$ and no differences were found for the adjective). The delayed effect of noun animacy in ungrammatical sentences (effect found in the preposition) vis-à-vis grammatical sentences (effect found in the adjective) in the intermediate group can be explained by the accumulation of two effects: (a) Processing ungrammatical sentences takes longer than grammatical sentences and (b) processing semantic gender takes longer than grammatical gender. As for the noun animacy effect in Spanish monolinguals, it was proven by the longer RTs at the adjective and the preposition in grammatical sentences with animate than inanimate nouns (all $p < .01$), as well as by longer RTs at the adjective in ungrammatical sentences with animate than inanimate nouns ($p > .01$ and this effect dissipated in the preposition, where $p > .05$). Furthermore, the presence of an animacy effect regardless of +/− agreement explains the lack of interaction of Agreement × Animacy, and the difference in the way beginners vis-à-vis intermediates and Spanish monolinguals reacted to agreement and animacy effects explains the significant interaction of Agreement × Group. Finally, the group effect was due to beginners reading more slowly than intermediates and both, in turn, doing slower reading than Spanish monolinguals across conditions (all $p < .01$), except nonsignificant differences between beginners and intermediates in sentences with gender violations, because the intermediates were sensitive to gender violations (i.e., longer RTs) but beginners were not.

Accuracy on Comprehension Questions

The ANOVA performed on the total number of correct responses on the comprehension questions showed no significant main effects for Agreement, $F(1, 193) = 23.299$, $p > .05$, or Group, $F(2, 193) = 7.270$, $p > .05$, and no significant interaction of Agreement × Animacy, $F(1, 193) = .339$, $p < .05$, or Agreement × Group, $F(2, 193) = 7.884$, $p > .05$. However, there was a significant main effect for Animacy, $F(1, 193) = 17.200$, $p < .01$, because all participants were more accurate responding to questions about sentences with inanimate than animate nouns (for both gender agreement and violation conditions). Finally, the interactions of Animacy × Group, $F(2, 193) = .620$, $p < .05$, and Agreement × Animacy × Group, $F(2, 193) = .283$, $p < .05$, were significant as a result of a slight trend for monolinguals to be more accurate than intermediates and the latter more accurate than beginners across conditions.

Grammaticality Judgment Task

The means and standard deviations for grammaticality and confidence judgments are shown in Table 3. Statistical analyses included two 2 × 2 × 3 repeated-measures ANOVAs: one for accuracy on grammaticality judgments and one for confidence judgments. As mentioned earlier, only confidence judgments for sentences with accurate grammaticality judgments were included for statistical analyses. This is important because it allows us to compare the results

Table 3 Mean and standard deviations of the grammaticality judgment task

	[+ animacy]				[− animacy]			
	[+ agreement]		[− agreement]		[+ agreement]		[− agreement]	
	M	SD	M	SD	M	SD	M	SD
Grammaticality judgments ($k = 10$)								
Beginners	8.52	1.66	2.62	3.82	8.73	1.51	2.92	4.11
Intermediates	7.34	2.08	5.64	3.03	8.26	1.32	6.47	3.40
Monolinguals	10.00	0.00	9.97	0.18	10.00	0.00	10.00	0.00
Confidence judgments ($k = 5$) (1 = not sure at all; 5 = completely sure)								
Beginners	3.53	0.82	3.61	0.78	3.51	0.85	3.56	0.91
Intermediates	3.56	0.71	3.96	0.69	3.54	0.67	3.97	0.72
Monolinguals	5.00	0.00	5.00	0.00	5.00	0.00	5.00	0.00

Note. $n = 52$ for beginners, $n = 53$ for intermediates, and $n = 63$ for Spanish monolinguals. The sample pool size of the grammaticality judgment task is lower than that of the moving-window task because not all participants completed the grammaticality judgment task.

of the grammaticality judgments and the confidence ratings and make overall conclusions based on two judgment measures.

Grammaticality Judgments

The ANOVA performed on the total number of correct responses to grammatical and ungrammatical sentences with animate or inanimate nouns showed a statistically significant main effect for Agreement, $F(1, 165) = 145.210$, $p < .01$, Animacy, $F(1, 165) = 15.579$, $p < .01$, and Group, $F(2, 165) = 128.298$, $p < .01$, as well as a statistically significant interaction of Agreement × Group, $F(2, 165) = 67.654$, $p < .01$, and Animacy × Group, $F(2, 165) = 6.833$, $p < .01$. However, there was no significant interaction of Agreement × Animacy, $F(1, 165) = .003$, $p > .05$, or Agreement × Animacy × Group, $F(2, 165) = .068$, $p > .05$. Multiple-contrast comparisons showed that beginners and intermediates were more accurate at classifying sentences with gender agreement than with gender violations (all $p < .01$) because they tended to classify sentences as correct when unsure about the correctness of the sentence. Spanish monolinguals did not show a gender congruency effect due to ceiling effects. Multiple-contrast comparisons also revealed no animacy effects for beginners and Spanish monolinguals, this being owed to ceiling effects for the latter group, but intermediates were more accurate in sentences with inanimate than animate nouns ($p < .01$). These findings are in line with those of the moving-window task (longer RTs and lower accuracy in comprehension questions at sentences with animate than inanimate nouns) and support our hypothesis that animate nouns are more difficult to process than inanimate nouns. Finally, the group effect resulted from Spanish monolinguals being more accurate than the two L2 groups, beginners being more accurate than intermediates in sentences with gender agreement (when unsure, beginners chose correct over incorrect) and, most importantly, intermediates being more accurate than beginners in sentences with gender disagreement (all $p < .01$). The latter findings are in line with the RT data of the moving-window task and lend support to nondeficit approaches.

Confidence Judgments

The ANOVA conducted on the average of confidence judgments on sentences with accurate grammaticality judgments showed a significant main effect for Agreement, $F(1, 165) = 32.113$, $p < .01$, and Group, $F(2, 165) = 126.058$, $p < .01$, and a significant interaction of Agreement × Group, $F(2, 165) = 21.560$, $p < .01$. There was no main effect for Animacy, $F(1, 165) = .609$, $p > .05$, and no interaction of Agreement × Animacy, $F(1, 165) = 0.003$, $p > .05$,

Animacy × Group, $F(2, 165) = .163$, $p > .05$, or Agreement × Animacy × Group, $F(2, 165) = 0.108$, $p > .05$. Multiple-contrast comparisons revealed that intermediates were more confident about their correct grammaticality judgments to sentences with gender disagreement than agreement (all $p < .01$), and there were no differences between gender agreement and disagreement in beginners (lack of sensitivity to disagreement) and Spanish monolinguals (ceiling effects). Between-group comparisons showed that Spanish monolinguals were more confident about their correct grammaticality judgments than beginners and intermediates, and the latter were, in turn, more confident about their correct grammaticality judgments than beginners (all $p < .01$).

Discussion

In this section, we examine the results of the moving-window and the grammaticality judgment tasks to determine whether Spanish monolinguals and beginning and intermediate English-Spanish late bilinguals are sensitive to gender congruency and noun animacy when processing noun-adjective gender agreement in Spanish DPs. Overall, the results show similarities between late L2 learners and Spanish monolinguals, but not until a certain level of proficiency is reached (intermediate, not beginner). In effect, RT and grammaticality judgment data reveal that beginners are insensitive to the gender concord/discord distinction, whereas intermediates show qualitatively similar reactions to monolinguals (longer RTs and lower accuracy in the grammaticality judgments in the gender discord than the gender concord condition). The same pattern is seen for animacy, where beginners show no difference between animate and inanimate nouns, whereas intermediates and monolinguals both show significant differences in RTs. In particular, intermediates and monolinguals showed longer RTs and lower accuracy in the grammaticality judgment in sentences with animate than inanimate nouns, suggesting that animate nouns are cognitively more demanding than inanimate nouns. Next, we discuss the findings for the following research questions of the study in more detail:

1. Do Spanish monolinguals and intermediate and beginning L2 learners show gender congruency effects with animate and inanimate nouns, based on online and offline tasks?
2. Do Spanish monolinguals, and intermediate and beginning L2 learners process gender agreement with animate nouns differently from inanimate nouns, based on online and offline tasks? If they do, does noun animacy facilitate or inhibit agreement processes?

Gender Congruency Effects

The first research question of the study explored whether L2 learners of an ungendered L1 can gain L2 gender concord features similar to that of monolinguals. We predicted that learners may develop representation and computation that are qualitatively similar to that of native speakers—that is, that they would perform on comprehension and production in a qualitatively if not quantitatively similar manner. We assumed that L2 proficiency is a factor because higher proficiency learners have been found to outperform lower ones on a range of tasks. Our prediction that beginning learners would show no sensitivity to gender agreement violations but intermediates and Spanish monolinguals would was supported. The moving-window data showed that intermediates and Spanish monolinguals, but not beginners, showed sensitivity to gender agreement violations with both animate and inanimate nouns (i.e., they spent more time reading adjectives that disagreed with a noun than those that agreed with it). In turn, the grammaticality judgment data revealed that intermediates were more accurate and confident than beginners in sentences with gender disagreement (Spanish monolinguals were equally accurate and confident in sentences with gender concord and discord due to ceiling effects). Although intermediates were far less accurate on grammaticality judgments for ungrammatical than grammatical sentences, their trajectory is precisely what one would expect of progressive acquisition of L2 grammar, because learners develop greater sensitivity to ungrammaticality with advancing proficiency (cf. Herschensohn, 2000, pp. 129–131, although the monolinguals tended to be more accurate than the intermediates on ungrammatical sentences). Taken as a whole, these findings indicate that adult L2 learners of a certain proficiency level show processing and performance patterns qualitatively similar to those of native speakers with regard to concord/discord of gendered adjectives. The patterns of the intermediate L2 learners suggest that adult learners can gain an ability to respond to grammatical features absent in the L1 in a manner that is qualitatively similar to monolinguals (e.g., Bruhn de Garavito & White, 2002; White, 2003a; White et al., 2004).

These results also suggest that proficiency level—the difference between beginning- and intermediate-level Anglophone learners—does indeed make a difference for the computation of concord/discord in Spanish L2 adjectives. Evaluating the theoretical issues in terms of the current results, we conclude that L2 differences in representation and processing are quantitative rather than qualitative. This investigation focuses on less advanced learners of L2 Spanish and demonstrates the gradual development of targetlike processing strategies regarding noun-adjective agreement. Rather than viewing L2 errors

as an indication of how different L2 learners are from monolinguals, we find that they manifest responses that are qualitatively similar at a remarkably early stage. It is precisely the distinction between the beginners—who show no sensitivity to concord/discord—and the intermediates—who register qualitatively similar responses to native speakers—that provides evidence for the developing L2 competence in both computation and representation.

Noun Animacy Effects

As for the question of whether L2 grammar resembles L1 grammar in terms of noun animacy effects on gender concord in Spanish adjectives, it is once again the distinction between beginners and intermediates that provides evidence for targetlike responses in the intermediate proficiency level. Unlike intermediates and Spanish monolinguals, beginners demonstrated no sensitivity to animacy, indicating (a) that the animacy gender feature in L1 English is not evident at this stage (cf. the idea of initial L1 transfer held by both deficit and nondeficit representational accounts) and (b) that the beginners' responses are different from those of intermediates and Spanish monolinguals, who showed behavioral differences between sentences with animate and inanimate nouns in RTs. A similar pattern was found in grammaticality judgments: Intermediates but not beginners showed differences in accuracy in sentences with animate versus inanimate nouns and ceiling effects prevented Spanish monolinguals from showing such differences.[4] These results support our prediction and mirror those obtained for the gender concord/discord distinction and indicate that adult L2 learners with a certain proficiency level (intermediates) show computational and representational patterns qualitatively similar to those of native speakers with regard to noun animacy, again supporting the idea that adult L2 learners can acquire targetlike computation of grammatical features absent in the L1.

As for the directionality of the noun animacy effects, intermediates and Spanish monolinguals showed longer RTs and lower accuracy on comprehension questions and grammaticality judgments in sentences with animate than inanimate nouns; nonsignificant differences in monolinguals' grammaticality judgments were due to ceiling effects. These findings indicate that the processing of animate nouns is cognitively more demanding than that of inanimate nouns and are compatible with previous research with Spanish monolinguals (e.g., Barber et al., 2004; Corral et al., 2008; Demestre et al., 1999; Igoa et al., 1999) and Anglophone and Francophone late learners of Spanish (e.g., Alarcón, in press, with attractor nouns; Bruhn de Garavito & White, 2002). In sentences with animate nouns, intermediates' longer RTs at the adjective in sentences with gender agreement and longer RTs at the word following the adjective in

sentences with gender agreement violations (delayed effects) provides further evidence that animate nouns are more difficult to process than inanimate nouns. The delayed effects in sentences with gender agreement violations would be due to an accumulation of the delay of processing discord plus the delay of processing animacy. As mentioned earlier, beginners' RTs or grammaticality judgments did not show sensitivity to noun animacy. However, it is important to note that their RTs followed the pattern of the other participants in the comprehension questions: lower accuracy rates in sentences with animate than inanimate grammatical gender concord.

As mentioned earlier, one of the goals of the present study was to investigate whether noun animacy affects agreement and, if it does, whether it has a facilitatory or an inhibitory effect on processing rate and accuracy. Our data show that the processing of agreement with animate nouns is cognitively more demanding than with inanimate nouns in intermediate but not beginning learners, confirming a role for proficiency level. Accounts of lexical access (see Barber & Carreiras, 2005; Domínguez et al., 1999; Igoa et al., 1999, for a review) propose differing explanations of why animate nouns are more difficult to access than inanimate nouns. On the one hand, lexical accounts claim that gender information is stored in the corresponding lexical entry and accessed from the full word and that gender failure forces the processor to go back to the lexical identification stage in order to check if the right entry had been chosen, a process cognitively more demanding in animate nouns (the processing of *esposo* "spouse-m" activates the related *esposa* "spouse-f") than inanimate nouns (*mesa* "table-f" does not prime **meso*). On the other hand, if gender is affixal, the processor needs more time to choose between the masculine affix *–o* and the feminine affix *–a* in animate nouns than to access single-gendered inanimate nouns. To discern between these accounts, future research could examine the RTs in word pairs with nontransparent invariable animate nouns (*hombre/mujer* "man/woman") and animate nouns with different meaning but the same stem (*puerto/puerta* "seaport/door"). Finally, to determine whether the differences between natural and grammatical gender are due to the animacy effect or to the fact that animate nouns are variable, future studies should examine the processing of gender agreement with animate nouns that can only have one gender (e.g., *monja* "nun-f") and those that have a dominant gender (e.g., *cirujano* "surgeon-m" or *enfermera* "nurse-f").

In sum, the RT and grammaticality judgment data of the intermediate group confirm growing sensitivity to gender congruency and to noun animacy, indicating targetlike sensitivity to L2 grammatical concord nonexistent in the L1. The sensitivity to gender agreement documented by online processing tasks

clearly relates to morphosyntactic knowledge, not lexical memorization (cf. Clahsen & Felser, 2006). An anonymous reviewer pointed out that L2 learners may develop sensitivity to transparent gender concord as a function of input frequency. We agree with the reviewer that transparent concord and input frequency undoubtedly play a role in bootstrapping the acquisition of concord sensitivity. The low-accuracy responses demonstrating insensitivity to adjective concord/discord of the beginners confirms their lack of targetlike competence in computation and representation at this level and no evidence of transfer of L1 English animate gender. In contrast, the intermediates demonstrate a clear emergence of targetlike computation—including sensitivity to noun animacy and to adjective concord/discord—indicating the importance of proficiency. We describe the difference between beginners and intermediates as one of proficiency, but the distinction also correlates with amount of language experience (input). Whereas the intermediates are quantitatively slower in RTs and less accurate in grammaticality judgments than the Spanish monolinguals, they nevertheless demonstrate that targetlike computation (and presumably representation) is evident well before ultimate achievement. In addition to the role of L2 proficiency, our results suggest that natural gender concord/discord is more difficult to process than grammatical gender concord/discord, as both intermediates and Spanish monolinguals show longer RTs in adjectives and lower accuracy on comprehension questions following animate than inanimate nouns. These findings suggest that gender concord with transparent animate nouns consumes more cognitive resources than processing of grammatical gender concord with inanimates.

Conclusion

This study has revealed insensitivity to gender concord/discord and agreement with animate/inanimate nouns at early L2 stages (the beginners) and confirm monolingual patterns in intermediate learners both for gender congruency effects (longer RTs and higher accuracy in gender concord than discord) and noun animacy effects (longer RTs and lower accuracy in adjectives following animate than inanimate nouns, with no differences in accuracy in monolinguals due to ceiling effects). These data show evidence of representation and computation of adjective gender concord (and discord) in Spanish DPs in higher proficiency learners of an ungendered L1. In effect, late learners with higher L2 proficiency and longer exposure display more monolingual-like patterns than those who exhibit less developed L2 proficiency. Finally, we consider that intermediates' and monolinguals' slower RTs to animate over inanimate grammatical gender

agreement may relate to animate nouns consuming more cognitive resources than inanimate nouns.

Although our study has found similar responses using online and offline formats using transparent default nouns, it has also raised several ideas for future research. With such a baseline, further studies could look at gender-opaque nouns that do not provide morphological cues to adjective form; comparisons between feminine and masculine concord/discord; comparisons between gender and number concord/discord; and comparisons between singular and plural concord/discord. With respect to the animacy criterion, areas of future exploration could include nontransparent invariable animate nouns (*hombre/mujer* "man/woman") and animate nouns with different meaning but the same stem (*puerto/puerta* "seaport/door"). Such studies could fruitfully expand the scope of the present study. Overall, the results of the present study are innovative in showing a convergence between online and offline responses of Spanish monolinguals and intermediate Spanish L2 learners to concord/discord of transparent noun phrases. The study has also demonstrated that instruction alone is insufficient to induce accurate grammaticality judgments or RT sensitivity to adjective discord for beginning Spanish L2 learners but that increased proficiency (with concomitant increased L2 experience) with the language does. Intermediate L2 Spanish learners' profiles for RT to concord/discord and to grammaticality judgment pattern with those of native speakers, indicating a developing competence and processing ability in L2 Spanish. Future research could further explore these early developmental stages and might even find profitable pedagogical applications for instructed learners.

<div align="right">Revised version accepted 21 July 2009</div>

Notes

1 An anonymous reviewer notes that "interpretable" is a problematic term for the gender of nouns (cf. White et al. 2004), a point with which we agree. Carstens (2000, p. 328) also pointed out that a better label is [inherent]; in this discussion we simply adopt commonly accepted UG terminology.
2 An anonymous reviewer pointed out that behavioral identity does not mean that the underlying processes are the same; gender agreement "could be carried out to success in a variety of ways."
3 Spanish uncountable nouns and Spanish countable nouns that are uncountable in English were excluded to ensure trial homogeneity (all nouns were countable) and avoid L1 interference. In addition, gender-inflected pairs such as *fruto-fruta* "fruit$_{masc,fem}$" and *puerto-puerta* "seaport$_{masc}$, door$_{fem}$" were excluded (cf. Camara,

1970) to avoid conceptual and grammatical relationships related 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 mask gender congruency effects).

4 Response times at the nouns (not the adjectives) were similar between animate and inanimate nouns because first fixations on words are devoted to extract meaning and only later (after participants have already pressed the button to see the next word) does the processing of gender-number take place. By then, because they have already pressed the button, the adjective is on the screen and receives a longer RT from the residual processing of the nouns' gender and number.

References

Alarcón, I. (2006). *The second language acquisition of Spanish gender agreement: The effects of linguistic variables on accuracy*. Munich: Lincom Europa.

Alarcón, I. (2009). The processing of gender agreement in L1 and L2 Spanish: Evidence from reaction-time data. *Hispania*, *92*(4), 814–828.

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

Antón-Méndez, M. I., Nicol, J. L., & Garrett, M. (2002). The relation between gender and number agreement processing. *Syntax, 5*, 1–25.

Baddeley, A. (2003). Working memory and language: An overview. *Journal of Communication Disorders, 36*, 189–208.

Barber, H., & Carreiras, M. (2005). Grammatical gender and number agreement in Spanish: An ERP comparison. *Journal of Cognitive Neuroscience, 17*, 137–153.

Barber, H., Salillas, E., & Carreiras, M. (2004). Gender or genders agreement? In M. Carreiras & C. Clifton (Eds.), *The on-line study of sentence comprehension. Eyetracking, ERPs and beyond* (pp. 309–327). New York: Psychology Press.

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*. New York: Springer.

Bley-Vroman, R. (2009). The evolving context of the Fundamental Difference Hypothesis. *Studies in Second Language Research, 31*, 175–198.

Blom, E., Marinis, T., Vasic, N., & Chondrogianni, V. (2009, July). *Bilingualism and SLI: Production and processing*. Paper presented at the VII International Symposium on Bilingualism. Universiteit Utrecht, Utrecht.

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.

Bruhn de Garavito, J., & White, L. (2002). L2 acquisition of Spanish DPs: The status of grammatical features. In A.-T. Pérez-Leroux & J. Liceras (Eds.), *The acquisition of Spanish morphosyntax: The L1/L2 connection* (pp. 153–178). Dordrecht: Kluwer.

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.

Câmara, J. M., Jr. (1970). *Estructura da lingua portuguesa*. Vozes, Petrópolis.

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.

CEDEFOP. (2004). *Lifelong learning: Citizens views*. Thessaloniki: European Centre for Development of Vocational Training.

Chomsky, N. (1995). *The minimalist program*. Cambridge, MA: MIT Press.

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

Clahsen, H., Hadler, M., & Weyerts, H. (2004). Speeded production of inflected words in children and adults. *Journal of Child Language, 31*, 683–712.

Colé, P., & Ségui, J. (1994). Grammatical incongruency and vocabulary types. *Memory and Cognition, 22*, 387–394.

Corral, J., Barber, H., & Carreiras, M. (2008, July). *EEG time-frequency analyses of syntactic and semantic gender agreement violations*. Paper presented at the IV Reunión de la Red Temática de Neurociencia Cognitiva. Santiago de Compostela, Spain.

Cutler, A. (1996). Prosody and the word boundary problem. In J. L. Morgan & K. Demuth (Eds.), *Signal to speech: Bootstrapping from speech to grammar in early acquisition* (pp. 87–99). Mahwah, NJ: Erlbaum.

Demestre, J., Meltzer, S., García-Albea, J. E., & Vigil, A. (1999). Identifying the null subject: Evidence from event-related brain potentials. *Journal of Psycholinguistic Research, 28*, 293–312.

De Mulder, H. (2006). *Second language acquisition of gender: Italian gender in native speakers of Dutch and English*. Unpublished M.A. thesis. Utrecht University, Utrecht.

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

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.

Domínguez, A., Cuetos, F., & Segui, J. (1999). The processing of grammatical gender and number in Spanish. *Journal of Psycholinguistic Research, 28*, 485–498.

Eberhard, K. M. (1997). The marked effect of number on subject-verb agreement. *Journal of Memory and Language, 36*, 147–164.

Ellis, R. (2005). Measuring implicit and explicit knowledge of a second language: A psychometric account. *Studies in Second Language Acquisition, 27*, 141–172.

Fernández-García, M. (1999). Patterns of gender agreement in the speech of second language learners. In J. Gutiérrez-Rexach & F. Martínez-Gil (Eds.), *Advances in Hispanic linguistics: Papers from the 2nd Hispanic linguistics symposium* (pp. 3–15). Somerville, MA: Cascadilla Press.

Finnemann, M. D. (1992). Learning agreement in the noun phrase: The strategies of three first-year Spanish students. *IRAL, 30*, 121–136.

Foucart, A. (2008). *Syntactic gender processing in French as a first and a second language*. Unpublished doctoral dissertation. Université Aix-Marseille and University of Edinburgh.

Foucart, A., & Frenck-Mestre, C. (in press). Grammatical gender processing in L2: ERP evidence. *Bilingualism, Language and Cognition*.

Franceschina, F. (2001a). Against an L2 morphological deficit as an explanation for the differences between native and non-native grammars. In S. H. Foster-Cohen & A. Nizegorodcew (Eds.), *EUROSLA yearbook, 1* (pp. 143–158). Amsterdam: Benjamins.

Franceschina, F. (2001b). 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 agreeement in advanced L2 Spanish grammars. In S. H. Foster-Cohen, R. Ruthenberg, & M. L. Poschen (Eds.), *EUROSLA yearbook, 2* (pp. 71–86). Amsterdam: Benjamins.

Franceschina, F. (2005). *Fossilized second language grammars: The acquisition of syntactic gender*. Amsterdam: Benjamins.

Gillon-Dowens, M., Barber, H., Guo, J., Guo, T., & Carreiras, M. (2008, April). *ERP correlates of automatic syntactic processing in highly proficient L1 Chinese-L2 Spanish late learners*. Paper presented at the Cognitive Neuroscience Society Annual Meeting. University of California, Davis.

Gillon-Dowens, M., Barber, H., Vergara, M., & Carreiras, M. (2004, September). *Does practice make perfect? An ERP study of morphosyntactic processing in highly proficient English-Spanish late bilinguals*. Poster presented at AMLaP, Aix-en-Provence, France.

Goad, H., & White, L. (2006). Ultimate attainment in interlanguage grammars: A prosodic approach. *Second Language Research, 22*, 243–268.

Green, J. N. (1988). Spanish. In M. Harris, & N. Vincent (Eds.), *The Romance languages* (pp. 79–130). New York: Oxford University Press.

Grosjean, F., Pommergues, J.-Y., Cornu, E., Guillelmon, D., & Besson, C. (1994). Gender-marking effect in spoken word recognition. *Perception and Psychophysics, 56*, 590–598.

Guillelmon, D., & Grosjean, F. (2001). The gender marking effect in spoken word recognition: The case of bilinguals. *Memory and Cognition, 29*, 503–511.

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.

Harris, J. (1991). The form classes of Spanish substantives. In G. Booij & J. van Marle (Eds.), *Yearbook of morphology* (pp. 65–88). Amsterdam: Kluwer Academic.

Hawkins, R. (1998, September). *Explaining the difficulty of French gender attribution for speakers of English*. Paper presented at EUROSLA, Paris.

Hawkins, R. (2001). *Second language syntax*. Malden, MA: Blackwell.

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 P. Prévost and J. Paradis (Eds.), *The acquisition of French in different contexts* (pp. 175–206). Amsterdam: Benjamins.

Hawkins, R., & Hattori, H. (2006). Interpretation of English multiple wh-questions by Japanese speakers: A missing uninterpretable feature account. *Second Language Research, 22*, 269–301.

Haznedar, B., & Schwartz, B.D. (1997). Are there optional infinitives in child L2 acquisition? In E. Hughes, M. Hughes, & A. Greenhill (Eds.), *Proceedings of the 21st Annual Boston University Conference on Language Development* (pp. 257–268). Somerville, MA: Cascadilla Press.

Hernández-Pina, F. (1984). *Teorías psicosociolingüísticas y su aplicación a la adquisición del español como lengua materna*. Madrid: Siglo XXI.

Herschensohn, J. (2000). *The second time around: Minimalism and L2 acquisition*. Amsterdam: Benjamins.

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: Cambridge University Press.

Herschensohn, J. (2009). Fundamental and gradient differences in language development. *Studies in Second Language Acquisition, 31*, 259–289.

Herschensohn, J., & Frenck-Mestre, J. (2005, September). *Failed [ugender] features in L2 French: What ERPs reveal about L1 influence*. Paper presented at *EUROSLA*, The University of Zagreb, Croatia.

Hopp, H. (2006). Syntactic features and reanalysis in near-native processing. *Second Language Research, 22*, 369–397.

Hopp, H. (2007). *Ultimate attainment at the interfaces in second language acquisition: Grammar and processing*. Groningen: Grodil Press.

Ibrahim, M. H. (1973). *Syntactic gender*. The Hague: Mouton.

Igoa, J. M., García-Albea, J. E., & Sánchez-Casas, R. (1999). Gender-number dissociations in sentence production in Spanish. *Rivista di Linguistica, 11*, 163–196.

Jin, F., Åfarli, T. A., & van Dommelen, W. A. (2007). Endstate L2 competence: Does perception parallel production? In A. Belikova, A. et al. (Eds.), *Proceedings of the 2nd conference on generative approaches to language acquisition North America* (pp. 200–211). Somerville, MA: Cascadilla Press.

Juffs, A., & Harrington, M. (1995). Parsing effects in L2 sentence processing: Subject and object asymmetries in wh-extraction. *Studies in Second Language Acquisition, 17*, 483–512.

Juffs, A., & Harrington, M. (1996). Garden path sentences and error data in second language sentence processing. *Language Learning, 46*, 283–326.

Jusczyk, P. W. (1997). *The discovery of spoken language*. Cambridge, MA: MIT Press.

Just, M., & Carpenter, P. (1992). A capacity theory of comprehension: Individual differences in working memory. *Psychological Review, 99*, 122–149.

Keating, G. D. (2009). Sensitivity to violations of gender agreement in native and nonnative Spanish: An eye-movement investigation. *Language Learning, 59*(*3*), 503–535.

Kuhl, P. (2004). Early language acquisition: Cracking the code. In M.S. Gazzaniga (Ed.), *The cognitive neurosciences* (3rd ed., pp. 831–843). Cambridge, MA: MIT Press.

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 J. Archibald (Ed.), *Second language acquisition and linguistic theory* (pp. 102–129). Oxford: Blackwell.

Lardiere, D. (2007). *Ultimate attainment in second language acquisition: A case study*. Mahwah, NJ: Erlbaum.

Lardiere, D. (2008). Feature-assembly in second language acquisition. In J. Liceras, H. Zobl, & H. Goodluck (Eds.), *The role of formal features in second language acquisition* (pp. 106–140). Mahwah, NJ: Erlbaum.

Lew-Williams, C., & Fernald, A. (2007). Young children learning Spanish make rapid use of grammatical gender in spoken word recognition. *Psychological Science, 33*, 193–198.

Marinis, T., Roberts, L., Felser, C., & Clahsen, H. (2005). Gaps in second language sentence processing. *Studies in Second Language Acquisition, 27*, 53–78.

McCarthy, C. (2008). Morphological variability in the comprehension of agreement: An argument for representation over computation. *Second Language Research, 24*, 459–486.

McDonald, J. L. (2006). Beyond the critical period: Processing-based explanations for poor grammaticality judgment performance by late second language learners. *Journal of Memory and Language, 55*, 381–401.

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.

Montrul, S., Foote, R., & Perpiñán, S. (2008). Gender agreement in adult second language learners and Spanish heritage speakers: The effects of age and context of acquisition. *Language Learning, 58*, 503–553.

Osterhout, L., Bersick, M., & McLaughlin, J. (1997). Brain potentials reflect violations of gender stereotypes. *Memory and Cognition, 25*, 273–285.

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., et al. (2008). Second-language learning and changes in the brain. *Journal of Neurolinguistics, 21*, 509–521.

Prévost, P. (2003). Truncation and missing inflection in initial child L2 German. *Studies in Second Language Acquisition, 25*, 65–97.

Prévost, P., & White, L. (2000). Missing surface inflection or impairment in second language acquisition? Evidence from tense and agreement. *Second Language Research, 16*, 103–134.

Rayner, K., & Pollatsek, A. (1989). *The psychology of reading*. New York: Prentice Hall.

Roca, I. M. (1989). The organization of syntactic gender. *Transactions of the Philological Society, 87*, 1–32.

Sabourin, L. (2003). *Grammatical gender and second language processing*. Groningen Dissertations in Linguistics Volume 42. Groningen: CLCG.

Sabourin, L., Stowe, L. A., & de Haan, G. J. (2006). Transfer effects in learning a second language grammatical gender system. *Second Language Research, 22*, 1–29.

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

Sagarra, N. (2007). Online processing of gender agreement in low proficient English-Spanish late bilinguals. In J. Camacho, N. Flores-Ferrán, L. Sánchez, V. Déprez, & M.-J. Cabrera (Eds.), *Current issues in linguistic theory series* (pp. 240–253). Amsterdam: Benjamins.

Schmitt, B. M., Lamers, M., & Münte, T. (2002). Electrophysiological estimates of biological and syntactic gender violation during pronoun processing. *Cognitive Brain Research, 14*, 333–346.

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.

Schütze, C. T. (1996). *The empirical base of linguistics: Grammaticality judgments and linguistics methodology*. Chicago: The University of Chicago Press.

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

Segalowitz, N. (2003). Automaticity and second language learning. In C. Doughty & M. Long (Eds.), *The handbook of second language acquisition* (pp. 382–408). Malden, MA: Blackwell.

Snape, N., Leung, Y.-K., & Sharwood Smith, M. (Eds.). (2009). *Representational deficits in SLA. Studies in honor of Roger Hawkins*. Amsterdam: Benjamins.

Sorace, A. (1996). The use of acceptability judgements in second language acquisition research. In W. Ritchie & T. K. Bhatia (Eds.), *Handbook of second language acquisition* (pp. 375–409). San Diego, CA: Academic Press.

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., & Papadopoulou, D. (2009). Aspect and the interpretation of motion verbs in L2 Greek. In N. Snape, Y.-K. Leung, & M. Sharwood Smith (Eds.), *Representational deficits in SLA. Studies in honor of Roger Hawkins* (pp. 187–227). Amsterdam: Benjamins.

Tsimpli, I.-M., & Roussou, A. (1991). Parameter resetting in L2? *University College London Working Papers in Linguistics, 3*, 149–169.

Vainikka, A., & Young-Scholten, M. (2007). Minimalism vs. organic syntax. In S. Karimi, V. Samiian, & W. K. Wilkins (Eds.), *Phrasal and clausal architecture: Syntactic derivation and interpretation* (pp. 319–338). Amsterdam: Benjamins.

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.

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

White, L. (2003a). On the nature of interlanguage representations: Universal Grammar in second language acquisition. In C. Doughty & M. Long (Eds.), *The handbook of second language acquisition* (pp. 19–42). Malden, MA: Blackwell.

White, L. (2003b). *Second language acquisition and Universal Grammar*. Cambridge: Cambridge University Press.

White, L., Valenzuela, E., Kozlowska-Macgregor, M., & Leung, Y.-K. (2004). Gender and number agreement in non-native Spanish. *Applied Psycholinguistics, 25*, 105–133.

White, L., Valenzuela, E., Kozlowska-Macgregor, M., Leung, Y.-K., & Ben-Ayed, H. B. (2001, November). *The status of abstract features in interlanguage: Gender and number in L2 Spanish*. Paper presented at the Annual Boston University Conference on Language Development, Somerville, MA.

Zagona, K. (2002). *The syntax of Spanish*. New York: Cambridge University Press.

Appendix

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.*

Animate nouns ($k = 40$): *abuelo, ahijado, amigo, arquitecto, banquero, barbero, biólogo, camarero, candidato, científico, compañero, consejero, cuñado, diputado, empleado, enfermo, esposo, físico, hermano, ingeniero, médico, mesero, ministro, misionero, modelo, muchacho, músico, nieto, notario, novio, piloto, plomero, político, químico, secretario, sargento, sobrinito, técnico, torero,vecino.*

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.*