Bilingualism: Language and Cognition 14 (3), 2011, 332–350 © Cambridge University Press 2011 doi:10.1017/S1366728910000222

Spanish gender agreement under complete and incomplete acquisition: Early and late bilinguals' linguistic behavior within the noun phrase*

IRMA V. ALARCÓN
Wake Forest University

(Received: August 13, 2009; final revision received: March 13, 2010; accepted: April 2, 2010; First published online 19 January 2011)

The present study explores knowledge of Spanish grammatical gender in both comprehension and production by heritage language speakers and second language (L2) learners, with native Spanish speakers as a baseline. Most L2 research has tended to interpret morphosyntactic variability in interlanguage production, such as errors in gender agreement, as a lack of native-like representation in the learner's grammar because of maturational constraints. From this perspective, adult English-speaking learners of Spanish are incapable of acquiring gender fully, whereas heritage Spanish speakers, who have been exposed to the language from birth, can attain complete gender acquisition. However, results of two tasks, one measuring written comprehension and the other oral production, show that advanced proficiency L2 learners, as well as advanced proficiency heritage speakers, have gender in their underlying grammars, and that the errors in oral production that L2 learners occasionally produce are due to difficulties in the surface manifestations of the abstract features of gender, i.e., the "mapping problem" (Lardiere, 2007).

Keywords: bilingualism, gender agreement, heritage Spanish speakers, second language acquisition

Leading researchers, most notably Valdés (2005, 2006), argue that few connections have been made between research on heritage language and second language (L2) acquisition. But since heritage speakers, or early bilinguals (see Valdés, 2000), have become a rapidly growing presence in the United States, Ortega's (2005) emphasis on real world needs suggests that we take a closer look at the language development and (re)acquisition of these language users, and at their potential contributions to our understanding of L2 acquisition processes. Since heritage language research is in its nascent stages, systematic study of a particular grammatical feature, such as gender, is undoubtedly a contribution to a field in which almost everything is "uncharted territory" (Polinsky, 2008, p. 41). Furthermore, grammatical gender is an interesting linguistic category for analysis because it provides a window on both lexical access and syntactic processing (Carroll, 1989; Corbett, 1991). For proficient speakers of a language, "grammatical gender has become knowledge that is stored as part of each noun's grammatical description in the mental lexicon" (Schriefers & Jescheniak, 1999, p. 577). This knowledge allows native speakers to produce correct gender agreements, suggesting gender is an example of rule-governed behavior (Tucker, Lambert & Rigault, 1977). However, for English-speaking learners, whose system lacks grammatical gender, mastery of gender is one of the most frustrating and difficult aspects of acquiring Spanish as an L2. Thus, heritage learners, who have acquired Spanish in early childhood at home, might provide a link between native speaker competence and adult L2 learners, or late bilinguals, who have learned Spanish after puberty in a formal classroom setting.

A relevant issue in L2 acquisition research is whether adult L2 learners are able to fully develop an implicit grammar of the target language, particularly of L2 grammatical features that are absent in their native (L1) language. In the present context, the interesting question is whether L2 learners are able to acquire competence of Spanish grammatical gender qualitatively and quantitatively comparable to that of native speakers. To address this question, the initial approach is to compare L2 with L1 Spanish acquisition. For native Spanish speakers gender agreement is acquired easily and early in life (e.g., Hernández-Pina, 1984; López-Ornat, 1997). In contrast, adult L2 learners exhibit persistent problems with gender (e.g., Fernández-García, 1999; Finnemann, 1992), and their acquisition appears to be incomplete.

* The author would like to thank the editor Carmen Silva-Corvalán as well as the anonymous reviewers for their many helpful suggestions, which greatly improved the article.

Address for correspondence:
Wake Forest University, Department of Romance Languages, P.O. Box 7566, Winston-Salem, NC 27109, USA
alarcoi@wfu.edu

Moreover, the emerging research on heritage languages shows that heritage speakers also display interrupted or incomplete knowledge of gender (e.g., Montrul, Foote & Perpiñán, 2008, for Spanish, and Polinsky, 2008, for Russian heritage speakers). This raises the questions of how to account for the divergence between heritage and native speaker gender acquisition, and whether this divergence extends to heritage and L2 learners of various proficiency levels.

From a theoretical perspective, there are two radically different proposals accounting for morphosyntactic variability in developing grammars. First are claims, such as the Fundamental Difference Hypothesis (Bley-Vroman, 1989) and the Failed Functional Features Hypothesis (Hawkins & Chan, 1997), that L1 and L2 are fundamentally different. The implication is that only L1 learners have access to Universal Grammar (UG), and thus only they are capable of fully acquiring grammatical features, such as gender. Adult L2 learners, who are past the critical period and therefore no longer have access to UG, cannot acquire grammatical features that are not present in their L1. From this perspective, variability, which involves "an alternation between correct and incorrect morphology" (McCarthy, 2008, p. 462), i.e., the failure to achieve native-like proficiency is explained in terms of a deficit in grammatical competence or representation, which reflects either a developmental problem (e.g., Vainikka & Young-Scholten, 1994, 1996) or some type of permanent impairment (e.g., Clahsen, 1988; Meisel, 1991). Since age of acquisition is a factor under these representational deficits accounts, heritage Spanish speakers, who have acquired Spanish gender in early childhood, should be able to perform with native-like competence, and thus be more accurate than L2 learners in both comprehension and production tasks.

Alternatively, claims such as the Full Transfer /Full Access Hypothesis (Schwartz & Sprouse, 1994, 1996) and the Missing Surface Inflection Hypothesis (Prévost & White, 2000) hypothesize similarities between L1 and L2 acquisitions, including equal access to UG for both L1 and L2 learners. Within this view, adult L2 learners are capable of fully acquiring grammatical features absent in their L1, including gender. In terms of mental representation, gender is considered acquired despite occasional errors in the surface realization of its abstract features. This position implies that such errors among advanced L2 learners are better explained by a breakdown in the morphosyntactic computation of gender rather than a deficit in its underlying representation (Hawkins, 2000). They are errors of performance rather than competence, and stem from difficulties in accessing the abstract features of gender from the lexicon and in mapping those features onto their surface forms (see Lardiere, 2007, for a full discussion). Evidence supporting the Missing Surface Inflection Hypothesis comes from comprehension and production data in which both intermediate and advanced L2 learners, with L1s both with (French) and without (English) grammatical gender, behaved similarly to native Spanish speakers (e.g., White, Valenzuela, Kozlowska-Macgregor & Leung, 2004). Such results indicate no effects for L1. Since these findings suggest that age of acquisition is irrelevant for acquiring gender features not instantiated in a learner's native language, under full access accounts there should be no difference between heritage Spanish speakers and L2 learners with respect to their implicit knowledge of gender as reflected by comprehension tasks.

In the only published study thoroughly comparing these two groups, Montrul et al. (2008) found that incompleteness of gender acquisition was not unique to adult L2 learners, but also characterized heritage speakers. Their results also indicated selective advantages for each group depending on task and modality: L2 learners were better in written comprehension, while heritage speakers performed better in oral production. Although matched in overall language proficiency, Montrul et al. compared L2 learners of various proficiency levels with exclusively lower-proficiency heritage speakers (i.e., exposed to English before age 5). This leaves open the question of whether heritage speakers and L2 learners matched at a higher-proficiency level would also display incomplete acquisition of gender, and, if so, whether they would display similar patterns of errors. More important, if advanced proficiency heritage speakers, who were exposed to English after age 5, display completely native-like gender behavior both in comprehension and production, while advanced L2 learners do not, one could argue that maturational constraints play a role in the underlying representation of gender. But if L2 learners display greater gender proficiency in comprehension than in production, then their production errors would be more plausibly attributable to computational difficulties than to a representational deficit. In this case, age of acquisition would play a role only in the syntactic mapping (computation) of gender morphology.

The present study addresses these issues by replicating two of Montrul et al.'s (2008) tasks. The goal is to verify whether their asymmetrical findings hold for more advanced proficiency-matched heritage speakers, with early and naturalistic exposure to Spanish, and L2 learners, whose first exposure was in the post-critical period and in a formal classroom setting. As predicted by the Missing Surface Inflection Hypothesis, results reveal that both heritage and L2 learners have acquired gender in their underlying grammars, since there is no significant difference between the two groups in written gender comprehension. Nonetheless, results also indicate that even advanced L2 learners have difficulties with the

syntactic manifestations of gender in oral production. Contrary to Montrul et al.'s results, the present study found no advantage for L2 learners in the written comprehension task, since both groups performed at ceiling, but agreed with Montrul et al. that heritage speakers had an advantage over L2 learners in an oral production task. These findings support the Missing Surface Inflection Hypothesis as a full access account of L2 morphosyntax acquisition, in which abstract knowledge of gender exists despite impaired performance. The present results also provide evidence that age of acquisition is fundamental in predicting gender computation difficulties because the advanced L2 learners, but not the advanced heritage speakers, displayed errors in spontaneous oral production. Furthermore, the difference between the heritage speakers in Montrul et al.'s and in the present study, who were successful even on the written task, highlights the link not only between age of first exposure to the heritage language, but also the connection between age of L2 acquisition (English) and (in)complete L1(Spanish) acquisition.

Spanish grammatical gender

Spanish has a binary gender system (see Gómez, 1998; Lloret & Viaplana, 1998; and Alarcos, 1999, for detailed discussions of Spanish gender), in which one of two genders, masculine or feminine, is assigned to all nouns. Animate nouns are those in which gender is assigned in accordance with biological sex, and thus is semantically motivated, and inanimate nouns are those whose gender is assigned in a purely grammatical, semantically arbitrary fashion (see Corbett, 1991, for a full discussion). As opposed to animate nouns, gender assignment to inanimate nouns is a puzzle, since "there is no correlation with either meaning or [the] phonological shape of the stem" (Harris, 1991, p. 36). Gender acquisition with inanimate nouns occurs later than with animate nouns, and both L1 and L2 learners tend to be less accurate with them (e.g., Andersen, 1984; Fernández-García, 1999; Hernández-Pina, 1984). Therefore, the current study focuses exclusively on gender agreement with inanimate nouns.

In Spanish, noun gender and morphological marking appear to co-occur in a systematic and predictable way (Green, 1988): the noun endings *-o* and *-a* are strongly correlated with masculine and feminine gender, respectively. According to the *Dicccionario de la Lengua Española* de la Real Academia, 99.8% of nouns ending in *-o* are masculine and 96.3% of nouns ending in *-a* are feminine (Teschner & Russell, 1984). Such nouns thus have morphologically marked gender, and are called overt gender nouns. All other nouns are referred to as non-overt gender nouns, for which gender is displayed by the agreeing elements in the noun phrase, as in, for example, *la pared blanca* "the(FEM) white(FEM) wall(FEM)".

The acquisition of grammatical gender, in Spanish or any other language with gender, involves acquiring gender both at the lexical level, by learning the meaning of the noun with its inherent gender feature (gender assignment), and at the syntactic level, by being able to establish correct agreements between the noun and the other elements in the phrase or sentence (gender agreement). As Carroll (1989) contended, gender assignment is "a permanent characteristic of nouns independent of the context of occurrence, whereas gender agreement involves a variable characteristic of modifiers, and their gender specification depends entirely on their occurring with a nominal element elsewhere in a sentence" (p. 546). Learners of Spanish thus need to acquire the nominal feature of gender in their implicit knowledge systems before being able to make valid form–function mappings in the language. Therefore, if gender knowledge is part of the learners' underlying representation, then agreement errors in production must involve difficulties in mapping the abstract gender features to the appropriate surface manifestations. These mapping errors are syntactic, not lexical, in nature, since they involve variable forms, such as determiners and modifiers, which take their gender from the nominal element. Nouns, the genders of which are considered inherent and invariable, determine the gender of the accompanying elements in the noun phrase or sentence (Comrie, 1999; Corbett, 1991).

Investigating both lexical and syntactic gender acquisition requires a practical method for distinguishing the two (see Montrul et al., 2008, for such a method), since gender concord in the noun phrase or sentence might reflect both lexical and syntactic gender. For example, the phrase *la torre inclinada* "the(FEM) leaning(FEM) tower(FEM)" displays both correct feminine assignment and agreement, as indicated by the corresponding correct feminine article and adjective. There is no way to distinguish assignment from agreement here. Now consider the phrases **una lápiz amarillo* "a(FEM) yellow(MASC) pencil(MASC)" and **un lápiz amarilla* "a(MASC) yellow(FEM) pencil(MASC)". These errors of gender concord between the noun and either the article or the adjective, respectively, strongly indicate a syntactic breakdown, an agreement rather than an assignment error, because in both cases at least one of the items correctly classified the noun as masculine. But if both the article and the adjective were incorrect, as in **una lápiz amarilla* "a(FEM) yellow(FEM) pencil(MASC)", then a lexical assignment error is indicated, since the noun has been incorrectly classified as feminine. Learning Spanish as either an L1 or L2 entails acquiring both the lexical and the syntactic properties of gender as specified in the lexicon.

L1 and L2 acquisitions of Spanish gender

Studies in L1 acquisition reveal that gender is acquired by age 3–4 (among numerous examples, see Hernández-Pina, 1984; López-Ornat, 1997; Mariscal, 2009; Pérez-Pereira, 1991, for Spanish; Carroll, 1989, for French; Maratsos, 1988, for Polish, German, and Russian; and Comrie, 1999, for Isangu, a Bantu language). However, research in L2 Spanish acquisition indicates that gender is problematic and is acquired relatively late in the learning process. Even advanced L2 learners display persistent errors with gender agreement, primarily in their spontaneous oral production (e.g., Montrul et al., 2008). This is also true for L2 gender acquisition in other languages, including French (Dewaele & Veronique, 2001), Italian (Oliphant, 1998), German (Rogers, 1987), and Russian (Taraban & Kempe, 1999). Some L2 research, though, suggests that gender does not involve a learnability problem (see Gass & Selinker, 2008, for a detailed discussion), but rather a mapping problem (Lardiere, 2007), since even beginning learners of Spanish have acquired the concept of gender in their developing systems, as shown by their high accuracy rates, but display difficulties with its syntactic manifestations (e.g., Alarcón, 2006; Bruhn de Garavito & White, 2002). Other studies suggest a critical role for the learner's L1 in the L2 acquisition of gender, with native-like attainment influenced by the status of gender features in the L1 (e.g., Franceschina, 2005; Sabourin, Stowe & de Haan, 2006).

Regarding the linguistic variables in the present study, previous findings on the L2 acquisition of Spanish grammatical gender show that learners operate with a default gender value, which can be either masculine or feminine, and over-generalize the masculine forms of determiners and modifiers, but do show evidence for incremental development as exposure/level increases (e.g., Shlig, 2003; White et al., 2004). Agreement between articles and nouns occurs earlier than agreement between adjectives and nouns, and some studies have found that L2 learners are significantly less accurate in gender agreement with adjectives than with determiners (e.g., Bruhn de Garavito & White, 2002). Furthermore, animate nouns, such as *tía* "aunt", are acquired earlier, and learners show higher rates of accuracy with animate than with inanimate nouns, such as *cuchara* "spoon" (e.g., Fernández-García, 1999; Finnemann, 1992). Concerning morphology, overtly marked nouns, such as *libro* "book", are acquired earlier, and learners display higher accuracy with them than with non-overtly marked nouns, such as *flor* "flower" (e.g., Franceschina, 2001). Most of this research, though, falls within a traditional framework, employing data collected through oral interviews of learners at only one proficiency level, and relying on error analysis techniques for assessing acquisition.

Current psycholinguistic research on L2 Spanish gender indicates that learners are sensitive to a variety of linguistic variables when assessing gender agreement correctness in online comprehension tasks (e.g., Alarcón, 2006, 2009; Keating, 2009; Sagarra, 2007; Tokowicz & MacWhinney, 2005). Most of these studies either measure real time processing, including eye tracking, moving window, and event related potentials, or use highly constrained instruments to measure accuracy and reaction times. This research focuses exclusively on the linguistic behavior of L2 learners. To date, however, there are relatively few studies exploring gender agreement behavior between heritage speakers and L2 learners, and the little research that does exist focuses on lower-proficiency heritage learners (e.g. Montrul et al., 2008; Montrul & Potowski, 2007).

The present study compares the performances of advanced proficiency-matched heritage language learners and L2 learners of Spanish, with native Spanish speakers as a baseline, on comprehension and production tasks. Linguistic similarities and differences between advanced heritage and L2 learners have not been thoroughly explored, and the connections between heritage and L2 research have not been consolidated. This study tries to establish some of those connections by investigating one grammatical feature, gender, which is a persistent source of errors for adult L2 learners. Not only is the explanation for these errors debated in L2 acquisition research (see above), emerging studies of heritage learners suggest that gender is also a difficult grammatical feature for lower-proficiency heritage speakers (Montrul et al., 2008). Therefore, a primary orientation of the present study is to investigate whether proficiency-matched advanced heritage and L2 learners acquire the Spanish gender system in similar ways. Among the many relevant questions to address are: What kinds of errors are associated with each type of learner? How similar or different are L2 and heritage learners' grammatical gender behaviors in written comprehension and oral production? Does variability of gender agreement extend to both comprehension and production? Is the behavior of advanced heritage learners more similar to that of native speakers or to advanced L2 learners'? Is age of acquisition of fundamental importance in gender acquisition for both heritage and L2 learners? These questions, and others, are important for research in both grammatical gender and L2 acquisition. The present study attempts to address most of them by focusing on gender agreement within the noun phrase, and by examining linguistic behavior assessed by different tasks (comprehension and production) and different modalities (oral and written).

Method

Research questions

The specific research questions guiding the present study are:

https://doi.org/10.1017/S1366728910000222 Published online by Cambridge University Press

1. Regarding both written comprehension and oral production, are advanced heritage learners more accurate than advanced L2 learners with gender agreement in the noun phrase?
2. What gender patterns do heritage and L2 learners display with respect to agreement domain (article or adjective), noun morphology (overt or non-overt ending), and noun gender (masculine or feminine)?

Before addressing the predictions for the research questions, consider the difference in the population in Montrul et al.'s (2008) and in the present study. In terms of proficiency test results, both the L2 and the heritage learner groups in Montrul et al.'s study displayed similar means, *SD*s, and ranges: 70.1, 9.24 and 32–100; and 73.7, 8.17 and 30–96, respectively. The corresponding figures for the native speakers were 97, 1 and 90–100. Although these experimental groups were proficiency-matched, as indicated by their similar means, both groups contained subjects of widely varying proficiency levels. This is not true of the present study, for which the ranges of the L2 and heritage speaker groups on a proficiency test were 75–92 and 80–97, respectively, and the means were closer to the native speaker baseline (86.4 and 89.9, respectively, compared with 97.7). Moreover, Montrul et al.'s heritage speakers rated their own Spanish abilities with a mean of 78% ($M = 3.9$ out of 5) whereas the present study heritage speakers self-rated their Spanish skills with a mean of 90% ($M = 3.6$ out of 4). Furthermore, only 48% of Montrul et al.'s heritage speakers felt they knew Spanish as a native language, as opposed to 97% of the heritage speakers in the present study.

Based on Montrul et al.'s (2008) findings, which focused on L2 and heritage learners proficiency-matched at a lower level, on average, with respect to their native speaker baseline than those of the present study, the predictions for the first research question suggest an asymmetrical pattern of acquisition: advantages in comprehension for the L2 learners and in production for the heritage speakers. This result would suggest that gender is problematic even at the advanced proficiency level, and that both groups display incomplete gender acquisition. Regarding the second research question, based on previous findings of studies investigating the effects of these variables on accuracy, including L1 (e.g., Hernández-Pina, 1984; López-Ornat, 1997), L2 (e.g., Fernández-García, 1999; Franceschina, 2001), and heritage learners (Montrul et al., 2008), both groups are expected to be more accurate on gender agreement with articles than with adjectives, with overt than with non-overt ending nouns, and with masculine than with feminine gender. If these patterns were consistent across comprehension and production tasks for both groups, this would offer strong evidence that gender is represented in the implicit systems of both heritage and L2 learners because the distribution of gender patterns, including gender errors, would be systematic rather than random.

Participants

All participants (n = 53) completed an extensive language background questionnaire and took lexical and grammar tests (described below) in addition to completing the two required tasks. They were compensated monetarily for their participation in the study.

The L2 learners (7 males, 11 females) were all raised as English monolinguals, implying that they spoke exclusively English at home, learned how to read first in English, and had their primary and secondary education entirely in English. Most of them (78%) started to learn Spanish formally around puberty (middle/high school), and had been studying Spanish for over seven years. All but one were Spanish majors and had studied Spanish abroad. At the time of the data collection, all but one were enrolled in advanced Spanish courses at their university. Their average age was 21.2 (range: 20–24 years). In general, these L2 learners reported feeling comfortable with their Spanish language skills: 94% with speaking and reading in Spanish, 89% with listening to Spanish, and 83% with writing. Their self-ratings were slightly lower: 89% self-assessed their reading skills as advanced, 83% did the same with listening, and 78% with their speaking and writing abilities.

The heritage language learners (8 males, 10 females; mean age: 20.3, range: 18–27 years) participating in the study were all students at the same university that the L2 learners attended. This common educational background reduced the heterogeneity characteristic of heritage learners (see Valdés, 1997, and elsewhere) by creating a proficiency-matched group of heritage individuals. For most of them (78%), both parents were native Spanish speakers who represented a wide variety of national origins, including Chile, Colombia, Cuba, Ecuador, Guatemala, Mexico, Nicaragua, Panama, Peru, and the Dominican Republic. 61% of the heritage learners were born in the U.S., and the rest were pre-pubescent when they first came to live in the U.S. (range of age-of-arrival: 3–10 years), but are still considered bilinguals (see Valdés, 2001, for a discussion of heritage speakers' profiles). All spoke Spanish at home, particularly with parents, and were exposed exclusively to Spanish until age 5. They started learning English after their pre-school years, but before puberty, and their formal schooling was in English. When speaking with their siblings, they reported using either Spanish (11%) or English (50%), or both (39%); and the language they most frequently used on a daily basis was English (89%). Since they had been exposed to Spanish throughout their lives, 100% of them indicated feeling comfortable when listening and reading in Spanish, and 94% of them indicated the same when speaking and writing. Furthermore, all of them

rated their Spanish listening and reading abilities as either native or advanced, 94% gave that rating to their Spanish speaking skills, and 83% evaluated their writing as native or advanced. Although their overall Spanish skills could be considered native-like, and 83% identified themselves as Latino/Hispanic, only 44% of the heritage learners were either majoring or minoring in Spanish. Of the rest, half had taken at least one university course in Spanish. Consequently, 72% of the heritage speakers had been exposed to formal Spanish at the college level.

The native Spanish speakers (5 males, 12 females) were included in the study to serve as a baseline for comparisons. They came from a variety of national backgrounds (Argentina, Chile, Colombia, Spain, Honduras, and Mexico, and were recruited from the same university as the other participants: eight were language instructors and the rest were from the university community (instructors' friends and/or relatives). Although they currently live in the U.S., and had a mean length of residence of 13.1 years (range: 2–27 years), they all arrived as adults, and had received their K-12 schooling entirely in Spanish. Their average age was 42.5 (range: 18–63 years). Due to the diverse backgrounds of the native speakers, the tasks avoided using vocabulary restricted to a specific Spanish variety. The native speakers completed the same instruments administered to the learners, and were tested under the same conditions.

Procedure

To investigate the research questions and predictions, the study consisted of two tasks: a Written Gender Recognition Task, and an Oral Picture Description Task, both modeled after Montrul et al. (2008).[1] Each task was administered in a separate testing session. Overall, the study involved a total of three sessions: in the first, a language background questionnaire, and a lexical test were administered; the second included a grammar proficiency test and Task 1; and Task 2 was given in the third session.

Lexical test

To verify knowledge of the meanings of the words included in the experiment, a lexical test was administered. It contained a list of the 112 Spanish nouns that were included in either of the two tasks, and participants were asked to provide the English equivalent or an English explanation of the meaning of each Spanish word. Although the native speakers also took the test, their results are not relevant for this study. The heritage speakers had a mean of 99.0% ($SD = 1.26$, range: 95–100), and the L2 learners obtained a mean of 97.1% ($SD = 2.02$, range: 93–100). These means were compared using an independent samples t-test, which showed that the two groups were only marginally different from each other ($t(34) = 3.459$, $p = .051$). So both experimental groups were highly familiar with the task words, which themselves are among the most frequently used words in Spanish (Davies, 2006).

Grammar proficiency

According to Montrul et al. (2008, pp. 519–520), "if we want to compare L2 learners and heritage speakers, we need to have a basic measure to equalize the groups at the outset and see how that measure correlates or not with other aspects of the groups' linguistic performance". Thus, in order to control the influence of grammatical proficiency on the accuracy of gender agreement, all participants, including native Spanish speakers, took a grammar test. The same test had been used in a previous study with L2 learners (Alarcón, 2006), but was refined and expanded for the present population. It consisted of 60 multiple-choice questions presented in brief, meaningful, and familiar contexts, including the copula *ser*/*estar* "to be", demonstratives, object pronouns, and prepositions. The native speakers had a mean of 97.7% ($SD = 2.02$, range: 93–100), heritage learners obtained a mean of 89.9% ($SD = 6.18$, range: 80–97), and L2 learners 86.4% ($SD = 4.49$, range: 75–92). Heritage and L2 learners' scores were compared with an independent samples t-test, which indicated that the two groups did not differ significantly from each other ($t(34) = 1.944$, $p = .060$). Thus these two groups could be regarded as advanced proficiency-matched.

Task 1: Written Gender Recognition

The goal of this task was to assess correct identification of the masculine or feminine form of articles and adjectives within the noun phrase based on the ending of the noun. The first research question (first part) is addressed: Are heritage learners more accurate than L2 learners on written comprehension of gender agreement? Also considered is the following question stemming from the second research question: Are participants more accurate with masculine than with feminine nouns; with overt than with non-overt nouns; and with articles than with adjectives? The task followed closely that of Montrul et al. (2008), but with some significant modifications. Montrul et al. mixed together attributive and predicative adjectives, which could have affected their results, so the present study included only attributive adjectives. In

[1] Ideally, as one reviewer advised, both tasks would have been presented in the same modality, preferably oral, to avoid methodological issues. The comprehension-written and production-oral combinations were maintained in the present study solely for replication purposes.

https://doi.org/10.1017/S1366728910000222 Published online by Cambridge University Press

addition, the present task had a larger number of tokens per experimental condition.

Task and materials

The task was a Written Gender Recognition Task consisting of a short story (*Último Acto*, by Pablo de la Torriente Brau) adapted for the task by the researcher. Adapting the story involved (i) replacing non-familiar with highly frequent nouns and (ii) shortening long sentences and paragraphs. The story displayed 48 gaps in the article and adjective positions. Example (1) below shows an excerpt from the text. (The entire text is in the Appendix.) The nouns that were manipulated in the task were the most frequent nouns from a frequency dictionary of Spanish (Davies, 2006). Only inanimate nouns and only adjectives in attributive position ending in *-o*/*-a* were used. To test the agreement domain (article or adjective), there were 24 article gaps and 24 adjective gaps. To assess the effect of gender, half the nouns were masculine and half feminine. Finally, to test the effect of noun ending on gender recognition, half the nouns in each gender group were overtly marked and half non-overtly marked for gender.[2]

(1)

Último Acto

En (1) un/una [ ] sector del jardín, en (2) el/la [ ] lugar donde se elevaba (3) el/la [ ] árbol de la palma, el hombre esperaba. La noche (4) profundo/profunda [ ] lo envolvía todo. Sólo se escuchaba el ruido (5) monótono/monótona [ ] del silencio (6) lejano/lejana [ ].

"In a section of the yard, where the palm tree stood, the man waited. The black night wrapped everything. Only the monotonous sound of the distant silence could be heard."

Procedure

The task was presented on a computer screen using the program *Blackboard*. Participants were asked to select the correct gender form of the article or adjective from two given options (one masculine and one feminine), and were instructed to enter their response by typing the corresponding article and adjective. The program was not sensitive to accent marks or capital letters. The task was graded automatically by the software, but was checked manually by the researcher, and by an independent rater. The task was untimed, but took 10–15 minutes to complete.

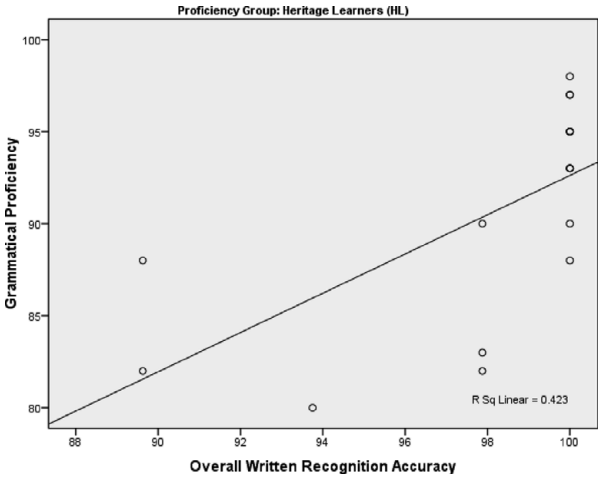

Figure 1. Correlations between accuracy on grammar proficiency and the Written Gender Recognition Task: Heritage language learners.

Results

For the 48 nouns on this task, correct responses received a 1, and incorrect responses a 0. The average percentage accuracy score for the native speakers was 99.6% (SD = 0.83, range: 97.9–100), which confirmed that gender agreement is not a problem for native speakers of Spanish (cf. Montrul et al., 2008). The heritage learners obtained an overall score of 98.0% (SD = 3.55, range: 89.6–100), and the L2 learners 96.9% (SD = 2.62, range: 91.6–100). Independent samples t-tests showed that the heritage and L2 groups did not differ significantly from each other in written gender recognition behavior ($t(33) = 1.032$, $p = .310$). The scatter plots in Figures 1 and 2 show the distribution of the individual participants in these two groups, and the correlation between the grammar proficiency scores and the overall scores on the task, respectively. One-tailed Pearson correlations between the grammar and task scores were significant for the heritage group ($r = .650$, $p = .002$), but not for the L2 learners ($r = .291$, $p = .121$).

To investigate accuracy rates and variability in performance, the task results were analyzed in a mixed ANOVA with group (L2 learners, heritage language learners, and native speakers) as the between-subjects factor, and the linguistic variables as the within-subjects factors (gender (masculine vs. feminine), agreement domain (article vs. adjective), and morphology (overt vs. non-overt)). Results showed a main effect of noun ending on accuracy of gender agreement ($F(1,33) = 16.366$, $p = .000$), indicating that overall agreement accuracy was significantly higher with overt than with

[2] One reviewer pointed out that some contexts in this task provide a gender cue through a determiner (article, possessive, or demonstrative) or an overtly marked adjective, which might have affected the results. This is a valid criticism: neither Montrul et al. nor the present study (in which 16/48 of the written contexts contained a gender cue: 2/24 for determiners and 14/24 for adjectives) controlled for this effect. Further research manipulating those contexts is required.

https://doi.org/10.1017/S1366728910000222 Published online by Cambridge University Press

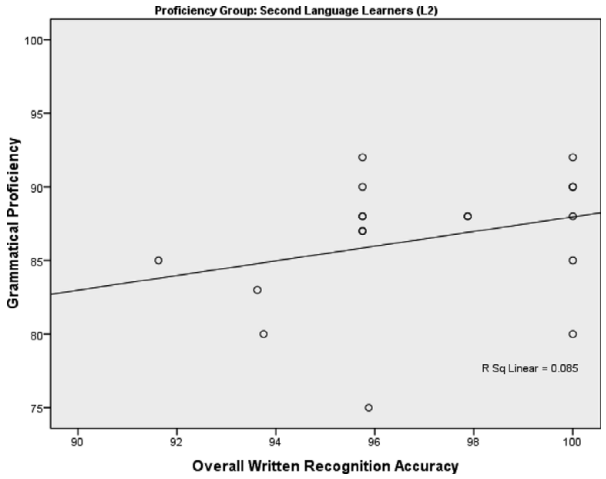

Figure 2. Correlations between accuracy on grammar proficiency and the Written Gender Recognition Task: L2 learners.

non-overt morphology. No other main effects were found, indicating that the gender of a noun, and whether the agreement is with an article or an adjective, had no significant effect on overall group performance on written gender comprehension. The average accuracy scores were high, suggesting that the experimental task was insufficiently demanding to produce a main effect of advanced proficiency group differences on accuracy rates (compare accuracy rates in Montrul et al., 2008, in which L2 learners performed better than lower-proficiency heritage speakers on a similar task). Consequently, even though both the heritage and the L2 groups were more proficient with overt than with non-overt nouns, with masculine rather than with feminine nouns, and with articles rather than with adjectives (although for the L2 group the domain distinction was extremely small), most of these differences were statistically insignificant. The native speakers performed at ceiling with all the binary linguistic variables, so their results are not reported, in order to focus instead on the experimental groups.

There was a significant two-way interaction between noun ending and group ($F(1,33) = 7.544$, $p = .010$). For heritage speakers, accuracy scores with overt and non-overt nouns were 98.6% and 97.6%, respectively, while for L2 learners the analogous accuracy scores were 99.5% and 94.4%. Moreover, a significant three-way interaction was found between gender, agreement domain, and noun ending ($F(1,33) = 5.427$, $p = .026$). Tables 1 and 2 show detailed results for the masculine and feminine conditions, respectively.

Summarizing the results of Task 1, although heritage learners displayed slightly higher accuracy scores than the L2 learners on the Written Gender Recognition Task, the difference was not significant. The two groups displayed largely similar patterns of gender agreement:

Table 1. *Written Gender Recognition Task: Masculine conditions (means and standard deviations).*

	Article				Adjective			
	Overt		Non-overt		Overt		Non-overt	
Group	M	SD	M	SD	M	SD	M	SD
HLL	100	–	100	–	98.1	8.0	97.1	12.1
L2L	100	–	97.2	6.5	100	–	93.4	10.2

HLL = heritage language learners; L2L = L2 learners

Table 2. *Written Gender Recognition Task: Feminine conditions (means and standard deviations).*

	Article				Adjective			
	Overt		Non-overt		Overt		Non-overt	
Group	M	SD	M	SD	M	SD	M	SD
HLL	100	–	95.1	9.8	96.1	9.4	98.1	8.0
L2L	98.1	5.5	92.6	11.7	100	–	94.3	8.2

HLL = heritage language learners; L2L = L2 learners

slightly more accurate with masculine than with feminine, and with overtly than with non-overtly marked nouns. In agreement domain, though, heritage learners were slightly more accurate with articles than with adjectives, while L2 learners were equally accurate with both. Nonetheless, the difference between the accuracy scores of the two groups on the Written Gender Recognition Task was statistically insignificant. Despite the differences in maturational constraints and acquisition contexts, both groups displayed similar underlying knowledge of Spanish gender.

Task 2: Oral Picture Description

The goal of this task was to test the oral production of gender agreement in noun phrases. This addresses the other half of the first research question: Are heritage speakers more accurate than L2 learners in oral production? Also addressed is the following question stemming from the second research question: Are speakers more accurate with masculine than with feminine gender, with overt than with non-overt ending nouns, and with articles rather than with adjectives? To investigate these questions, an Oral Picture Description Task, modeled on Montrul et al. (2008), was used. The experimental conditions were modified substantially, however. For example, the present study does not include nouns with exceptional endings, such as *mano* "hand(FEM)", or animate nouns, such as *oveja*

https://doi.org/10.1017/S1366728910000222 Published online by Cambridge University Press

"sheep(FEM)". Nouns showing dialectal variation, such as *radio* "radio", which can take either the masculine or the feminine article depending on the variety of Spanish, were also excluded. Instead, for a more precise comparison, the present task focused on the same experimental conditions as Task 1 (the nouns were not identical, but belonged to the same noun categories). In addition, more tokens were analyzed than in the Montrul et al. study.

Task and materials

For the Oral Picture Description Task, participants were asked to describe, using the structure [*Veo un*/*una* "I see a" + noun + adjective], what they saw in a series of pictures presented one by one on a computer screen.[3] The stimuli consisted of 80 photographs of objects and places: the first 16 were practice to become familiar with the procedure, and the remaining 64 were target items. Of these target nouns, half were masculine and half feminine. Within each gender, half the nouns had overt and half non-overt endings. Within each of the four resulting subcategories, half the nouns were examined for article agreement and half for adjective agreement. As with Task 1, all the nouns included in this task were taken from Davies' (2006) frequency dictionary of Spanish. Each picture was accompanied by the target noun written on top of the photograph to avoid dialectal differences between Spanish varieties with which participants might have been familiar (such as *habitación* vs. *pieza* for "bedroom"), as well as to ensure that participants produced the specific target noun and not another (such as *cárcel* instead of *prisión* for "prison").

Procedure

Participants were instructed to describe what they saw in each slide by using the sentence *Veo un*/*una* "I see a" + noun + adjective. For example, *Veo un volcán nevado*/*majestuoso*/*hermoso* "I see a(MASC) snow-covered/majestic/beautiful(MASC) volcano(MASC)". Participants were indirectly encouraged to produce exclusively adjectives with overt endings. Whenever they produced an unwanted response (including an adjective with a non-overt ending or a prepositional phrase), they were prompted with either: *¿Qué más me puedes decir?* "What else can you tell me?" or *Dame otra característica del objeto que ves* "Give me another characteristic of the object you see". Participants were tested individually, and their responses were audio-taped, but not timed, which replicated the conditions of Montrul et al. (2008). Including the practice trials, the task took most participants 15–20 minutes. All responses, which were fast, brief, and spontaneous, were evaluated in real time by the researcher using a prepared written answer key, and later checked by an independent rater.

[3] A sample picture is available online, on the Journal's website, as Supplementary Materials accompanying the present article (see the online version of the present article at journals.cambridge.org/bil).

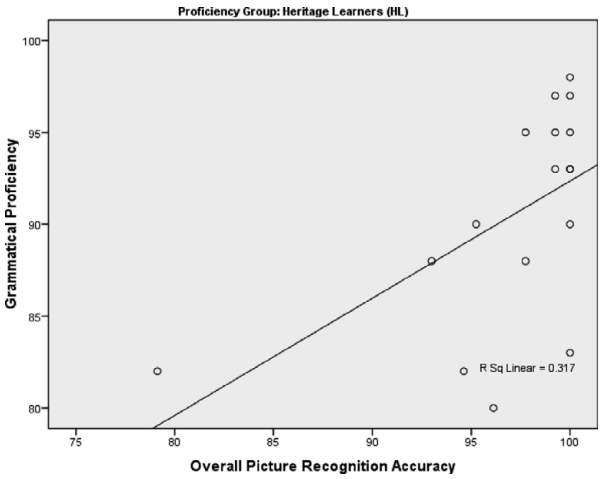

Figure 3. Correlation between grammar proficiency and accuracy on the Oral Picture Description Task: Heritage language learners.

Results

Correct agreement responses received 1 point, incorrect answers were assigned a 0. All native speakers scored 100%, indicating clearly that this task was extremely easy for them. The heritage learners scored 97.1% ($SD = 5.1$, range: 79.1–100), and the L2 learners 85.6% ($SD = 5.2$, range: 72.0–94.5). An independent samples t-test indicated that the difference in means between the two groups was statistically significant ($t(33) = 6.621$, $p = .000$).

The scatter plots in Figures 3 and 4 illustrate the individual distribution of heritage and L2 learners by grammar proficiency and accuracy in the Oral Picture Description Task, respectively. Positive correlations between grammar scores and overall task accuracy were significant for both the heritage ($r = .563$, $p = .009$) and the L2 ($r = .526$, $p = .013$) learners. The scatter plots show also that substantially more heritage than L2 learners achieved 90% accuracy or higher.

As with Task 1, results of the Picture Description Task were analyzed with a mixed ANOVA design, with group (L2 learners, heritage language learners, and native speakers) as the between-subjects factor, and the linguistic variables (noun gender, noun morphology, and agreement domain) as within-subjects factors. Tables 3 and 4 show the accuracy scores on the masculine and feminine conditions, respectively. Since the native speakers had perfect scores, they are excluded from the present report in order to focus on the two groups of learners.

https://doi.org/10.1017/S1366728910000222 Published online by Cambridge University Press

Table 3. *Oral Picture Description Task: Masculine conditions (means and standard deviations).*

	Article				Adjective			
	Overt		Non-overt		Overt		Non-overt	
Group	M	SD	M	SD	M	SD	M	SD
HLL	100	–	99.3	1.9	98.2	5.3	96.4	7.4
L2L	99.7	1.4	93.5	10.1	99.3	1.9	86.6	11.8

HLL = heritage language learners; L2L = L2 learners

Table 4. *Oral Picture Description Task: Feminine conditions (means and standard deviations).*

	Article				Adjective			
	Overt		Non-overt		Overt		Non-overt	
Group	M	SD	M	SD	M	SD	M	SD
HLL	99.6	1.5	91.6	16.1	99.3	1.9	92.8	13.8
L2L	95.2	19.1	54.3	21.1	97.3	3.7	58.8	17.1

HLL = heritage language learners; L2L = L2 learners

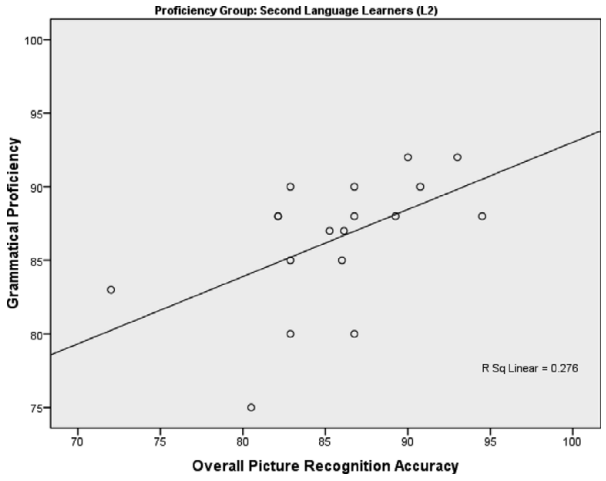

Figure 4. Correlation between grammar proficiency and accuracy on the Oral Picture Description Task: L2 learners.

Findings revealed main effects for group ($F(1,33) = 43.832$, $p = .000$), gender ($F(1,33) = 27.332$, $p = .000$), and noun morphology ($F(1,33) = 138.961$, $p = .000$). Overall accuracy on oral production, therefore, was significantly higher for heritage than for L2 learners, for masculine rather than feminine nouns, and with overt rather than non-overt ending nouns.

Considering secondary effects, gender interacted significantly with group ($F(1,33) = 15.356$, $p = .000$), with morphology ($F(1,33) = 31.534$, $p = .000$), and with agreement domain ($F(1,33) = 6.556$, $p = .015$). Concerning the gender vs. group interaction, the respective accuracy scores with masculine and feminine nouns were 98.5% and 95.8% for the heritage speakers, but for L2 learners the gap was significantly wider, 94.8% to 76.5%. Regarding the interaction between gender and noun morphology, the difference in accuracy between overt and non-overt nouns was greater for feminine than for masculine nouns. In the gender vs. domain interaction, performance with masculine nouns was significantly higher with articles than with adjectives, but this was not the case with feminine nouns. There was also a significant interaction between group and noun ending ($F(1,33) = 68.767$, $p = .000$). Heritage speakers were more accurate with overt (99.3%) than non-overt (95.0%) nouns, but, again, as with the group vs. gender interaction, the analogous gap for the L2 group was wider (from 97.9% to 73.3%). Finally, gender and morphology were also components in two significant three-way interactions, with domain ($F(1,33) = 5.271$, $p = .028$), and with group ($F(1,33) = 14.065$, $p = .001$).

Error analysis: Gender assignment and agreement

Because there were significant differences between heritage and L2 learners in their oral production, Task 2 errors were examined in more detail. With the understanding that gender acquisition involves both a lexical level (assignment) and a syntactic level (agreement), errors were categorized as pertaining to gender assignment if both the article and the adjective were incorrect, or as a gender agreement error if only one of the two were incorrect. Examples of assignment errors produced by the heritage and L2 learners in the study are shown in (2)–(3), and examples of agreement errors in (4)–(5).

(2) *Veo una puente peligrosa

"I see a(FEM) dangerous(FEM) bridge(MASC)."

(3) *Veo un catedral antiguo

"I see an(MASC) antique(MASC) cathedral(FEM)."

(4) *Veo una pie blanco

"I see a(FEM) white(MASC) foot(MASC)."

(5) *Veo un sillón cómoda

"I see a(MASC) comfortable(FEM) armchair(MASC)."

The ratio of the total number of errors to the total number of opportunities (the number of noun phrases produced by all the participants in Task 2) was 271/2,304. Native speakers made no errors in their oral production, so their results are not included in this analysis. The heritage learners produced 62 (22.9%) of the errors, and the L2 learners produced the remaining 209 (77.1%). (Montrul et al., 2008, found similar proportions.) 51.6% (32/62) of the errors by the heritage learners were in assignment, compared with 59.8% (125/209) for the L2 learners.

https://doi.org/10.1017/S1366728910000222 Published online by Cambridge University Press

Table 5. *Oral production: Percentage of error types by noun gender and morphology.*

		Heritage language learners				L2 learners			
		Assignment		Agreement		Assignment		Agreement	
Gender	Morphology	%	Count	%	Count	%	Count	%	Count
MASC	Overt	0	0/32	16.7	5/30	0	0/125	3.6	3/84
	Non-overt	37.5	12/32	33.3	10/30	12.8	16/125	28.5	24/84
FEM	Overt	0	0/32	13.3	4/30	0	0/125	28.5	24/84
	Non-overt	62.5	20/32	36.7	11/30	87.2	109/125	39.3	33/84

Both groups made more errors with feminine than with masculine nouns, and of those errors with feminine nouns, more were with assignment than with agreement. Both of those results are consistent with the previous findings, in both L1 and L2 Spanish grammatical gender research, of a masculine default setting. In contrast, errors with masculine nouns were mostly at the syntactic level, i.e., in gender agreement with either the article or the adjective, but not both. These patterns of gender errors were observed in both the heritage and the L2 groups, and are consistent with Montrul et al.'s (2008) findings.

The patterns of error distribution according to noun morphology (overt or non-overt) were also similar for both heritage and L2 learners: more errors were made with non-overt than with overtly marked nouns. This is also consistent with previous acquisition literature. In the present study the percentage of errors associated with non-overt nouns for the heritage and L2 groups were very close (85.5% and 87.1%, respectively). In both groups, 100% of the assignment errors occurred with non-overt nouns. When errors with overt nouns did occur, they were exclusively agreement errors.

Table 5 offers a multivariate depiction of the patterns of oral production errors by the heritage and L2 learners. This perspective, which includes all the variables simultaneously, confirms the above results: both groups were more accurate with masculine than with feminine nouns, and with overtly rather than with non-overtly marked nouns. Also, for both groups all assignment errors were with non-overt nouns, and more of those errors were with feminine rather than masculine nouns. Moreover, agreement errors were also more frequent with feminine non-overt nouns, suggesting that a masculine default, which we have already seen at the lexical level, also occurs at the syntactic level. Furthermore, when errors with masculine nouns occur, they are more likely agreement rather than assignment errors, and exclusively with non-overt nouns. These findings accord with Montrul et al.'s (2008) data of lower-proficiency heritage learners. Contrary to their results, however, both advanced heritage and L2 learners produced more assignment than agreement errors.

Results also showed patterns of errors specific to each group. For example, with masculine non-overt nouns, heritage learners made more assignment errors, whereas L2 learners made more agreement errors. In agreement, heritage learners displayed about the same level of errors with both genders, whereas L2 learners made more errors with feminine than with masculine nouns. With non-overt nouns, heritage learners showed similar gender patterns in both assignment and agreement: they are better in agreement with both genders; but L2 learners exhibited an asymmetric gender pattern: better with masculine nouns in assignment and with feminine nouns in agreement.

Summarizing the findings of the oral production task, there was a clear distinction between the two groups: heritage learners had significantly higher scores than the L2 learners. Regarding agreement domain, heritage learners were more accurate with articles than with adjectives, but L2 learners were equally accurate with both, just as they were on the gender recognition task. Both groups were more accurate with masculine and overt ending nouns than with feminine and non-overt ending nouns. As for the detailed error analyses, including the error types (assignment or agreement) and the patterns of distribution of these errors according to gender and morphology, both groups displayed similar linguistic behaviors overall. Both produced more errors with assignment than with agreement, with feminine than with masculine nouns, and with non-overt rather than with overt ending nouns. Within both groups, assignment errors were found only with non-overt nouns, and primarily with feminine nouns. Although agreement errors were distributed more widely, they were also more frequent with feminine non-overt nouns. Consequently, the results strongly support a masculine default that operates at both the lexical and the syntactic levels, as well as a facilitatory role for overt morphology in the accuracy of gender assignment and agreement.

Comparing the results of the two tasks: Comprehension and production

The results of the two tasks confirm that adult native Spanish speakers have virtually no problems with gender,

https://doi.org/10.1017/S1366728910000222 Published online by Cambridge University Press

Table 6. *Overall accuracy on the two tasks (means, standard deviations and ranges).*

	Written Recognition Task			Picture Description Task		
Group	M	SD	Range	M	SD	Range
NS	99.6	.8	97.9–100	100	.0	100–100
HLL	98.0	3.6	89.6–100	97.1	5.1	79.1–100
L2L	96.9	2.6	91.6–100	85.6	5.2	72.0–94.5

NS = native speakers; HLL = heritage language learners; L2L = L2 learners

since they performed at ceiling level on both tasks. In general, the heritage and L2 groups made few gender errors, in accordance with their advanced proficiency-matched levels. Nonetheless, results also showed that for L2 learners, even at this level of proficiency, spontaneous oral production of gender was still problematic. Although heritage speakers displayed similar performance on both tasks (above 95% accuracy on each), the L2 learners were better in comprehension than in production (error rates of about 5% and 15%, respectively). An error rate of even 10% by L2 learners is significant and demands an explanation (White et al., 2004). Thus, the type of task made a significant difference both between the two groups and within the L2 group. These results are displayed in Table 6, which includes native speakers for comparison.

A mixed ANOVA was conducted on the average scores of the two tasks. Results revealed main effects for task ($F(1, 49) = 63.395$, $p = .000$) and group ($F(2, 49) = 37.940$, $p = .000$), as well as a task by group interaction ($F(2, 49) = 56.918$, $p = .000$). For each group, the difference in accuracy on the two tasks was also examined. For the native speakers, the less than 1% difference between scores on the written recognition and oral production tasks was not significant ($t(16) = -1.852$, $p = .083$). Likewise for the heritage learners, for whom the 1% difference between the two tasks was also statistically insignificant ($t(17) = 1.455$, $p = .164$). But for the L2 learners, the difference between the scores on the two tasks, slightly above 10%, was highly significant ($t(17) = 9.150$, $p = .000$). Furthermore, pairwise comparisons among the proficiency groups' mean differences on the two tasks indicated that the three groups were all statistically distinguishable. Native speakers ($M = 99.8\%$) and heritage learners ($M = 97.6\%$) had a mean difference of 2.22, $p = .037$, and heritage speakers and L2 learners ($M = 91.3\%$) showed a mean difference of 6.321, which was also significant, $p = .000$.

To investigate how close individual heritage and L2 participants came to native-like results on the two tasks, the number of individuals in each of the learner groups who scored within the range of variation of native speakers was considered (cf. Montrul et al., 2008). The minimum score among the native speakers was 97.9% on the written recognition task and 100% on the oral picture description. Although the means for the heritage and L2 groups on the written recognition task were not significantly different, 14 of the 18 heritage speakers (77.8%) scored above the native speaker minimum, but only 8 of the 18 (44.4%) L2 learners managed to do so. On the oral production task, 7 of the 18 heritage learners (38.9%), but none of the L2 participants, scored 100%, which was the native speaker minimum. So even at the advanced proficiency level, the accuracy of gender agreement in oral production by L2 learners was substantially below the level of native speakers. This was not the case with the advanced heritage learners, whose results were significantly closer to native-like performance.

In summary, the findings of the study show that there was no significant difference between advanced-proficiency heritage and L2 learners on the gender recognition task (the lexical dimension of gender acquisition, more deeply implicated in language competence than performance). The picture description task, however, revealed that spontaneous oral production remains problematic for L2 learners. Results also show that, in both comprehension and production, heritage and L2 learners were more accurate with masculine and with overt nouns than with feminine and non-overt nouns. The results for agreement domain were also consistent across tasks, though they differed in each group: heritage speakers were more accurate with agreement on the article than on the adjective, while L2 learners were equally accurate with both articles and adjectives. In relation to types of errors in production, both groups made more errors in assignment (lexical gender) than in agreement (syntactic gender), particularly with feminine non-overt ending nouns.

Discussion

The original research questions guiding this study can now be addressed. As predicted by the Missing Surface Inflection Hypothesis, and contrary to representational deficit accounts such as the Failed Functional Features Hypothesis, both heritage and L2 learners were comparably accurate in the gender comprehension task (research question 1, first part). This finding supports the full access accounts of L2 acquisition (cf. White et al., 2003, 2004), since the age of first exposure to Spanish did not affect the L2 acquisition of gender, a grammatical feature not present in their L1. Nonetheless, a look at the individual results per group indicated that more heritage than L2 learners scored within the native-speaker range. Therefore, even though both groups displayed similar abstract knowledge of gender in comprehension, the performance of heritage speakers was more native-like

https://doi.org/10.1017/S1366728910000222 Published online by Cambridge University Press

than that of the L2 learners. This might be explained by several factors, including the context in which learning Spanish took place (natural vs. formal classroom setting) and the amount and frequency of input (daily language exposure and interaction at home vs. limited classroom input and opportunities for using the language). Also consistent with the Missing Surface Inflection Hypothesis, which considers differences between L2 comprehension and production as mapping problems, results revealed a significant statistical difference in the oral production of the two groups: heritage speakers were more accurate than L2 learners on the production task (research question 1, second part). While the heritage learners were equally accurate in gender comprehension and production, the L2 learners were significantly more accurate with gender comprehension than production. Although some of the heritage learners scored within the native speakers' range on the production task, none of the L2 learners achieved this. These results suggest that, for the L2 learners, there is a deficit in the computation of gender, a divergence between their acquisition of abstract features, as indicated by their high accuracy scores on the comprehension task, and their continuing difficulties with surface manifestations, as revealed by their significantly lower scores on the production task. As the Missing Surface Inflection Hypothesis predicts, the L2 learners displayed performance errors. And, in contradiction of a prediction of the Failed Functional Features Hypothesis, the current findings provide evidence against a deficit in the underlying representation of the morphosyntactic features of gender, since both heritage and L2 learners were statistically indistinguishable in their comprehension of gender. With high accuracy scores in written comprehension, both groups appear to have gender in their underlying grammars.

Nonetheless, one might contend that the written comprehension task measured more explicit knowledge of gender, since participants were given the options from which to choose, while the oral production task measured more implicit knowledge because it required fast, spontaneous, and automatic processing. Then, since the L2 learners performed better on the written than the oral task, it could be claimed that L2 learners had not yet internalized gender rules as part of their implicit knowledge system (e.g., Ellis, 2005). Accordingly, the present results could also be explained by the explicit/implicit knowledge distinction, as one reviewer suggested, because they show that advanced heritage speakers, but not advanced L2 learners, display both explicit and implicit knowledge of gender, and hence complete acquisition. This alternative view offers a plausible account of the differences in performance between high-proficiency heritage and L2 learners on written and oral tasks. One could also argue, however, that at advanced proficiency levels it is difficult to determine whether a linguistically non-demanding task, such as the written task in this study, taps explicit or implicit knowledge of gender, particularly when carried out smoothly and quickly by all participants, and with a focus on meaning (since the noun phrases were contextualized). As Ellis (2008) suggests, the distinction between implicit and explicit is itself controversial. One contentious issue, for example, is whether the distinction is a continuum or a dichotomy. The current findings are explained nicely by the Missing Surface Inflection Hypothesis, which claims that advanced L2 learners do have implicit knowledge of the target language, even though this knowledge might not be accessible when performing a specific task due to computation problems. If gender is part of their linguistic competence, then that knowledge must comprise an implicit gender representation. This implies that both L2 and heritage speakers might possess gender competence, but still display performance errors, particularly in oral tasks.

The present results differ from Montrul et al.'s (2008) finding that L2 learners were more accurate than lower-proficiency heritage speakers in gender comprehension on a similar written gender recognition task. This could be attributable to the difficulty of the written stimuli: Montrul et al.'s task included adjectives in both attributive and predicative positions. In contrast, the present study focused only on attributive adjectives, with which gender agreement is acquired earlier than with predicative adjectives (e.g., Bruhn de Garavito & White, 2002). Another possible set of factors relates to the differences between lower-proficiency and advanced heritage speakers in terms of literacy and exposure to formal instruction (e.g., Valdés, 1997). Alternatively, the varying results of the two studies could also be explained by the differing profiles of the two L2 populations. Montrul et al. included graduate students of Spanish and Spanish instructors, who had a relatively strong command of the language compared with the L2 group in the present study, which was composed exclusively of undergraduate language learners. Since differences in knowledge, input exposure and experience do affect gender acquisition (Ellis & Schmidt, 1997; Taraban & Kempe, 1999), Montrul et al.'s L2 group was linguistically more advanced than the L2 learners in the present study.

In accordance with Montrul et al.'s (2008) findings, however, the heritage speakers in the present study were superior to the L2 learners and closer to the native speakers in their oral production. This finding strongly suggests that heritage speakers, whether of lower (Montrul et al.) or more advanced (the present study) proficiency, have an advantage over advanced L2 learners when producing gender agreement in oral production. L2 grammatical gender research has demonstrated that L2 learners consistently display errors in the oral production (e.g., Fernández-García, 1999; Finnemann, 1992). These errors, though, are systematic rather than random, since the same

patterns of errors observed in production also occur in comprehension. How can we account for these persistent errors in oral production in advanced L2 learners, who have already acquired gender in their underlying grammars? According to Lardiere (2000, p. 121), "the most coherent explanation for the L2 data is . . . figuring out how (and whether) to spell out morphologically the categories they already represent syntactically, i.e., the 'mapping problem'". This explanation does not imply that the problems L2 learners have with gender production will be solved. Some learners never overcome these errors, despite further input, meaningful language interactions, and continued feedback, and eventually fossilize. Fossilization, though, "is attributable not to a breakdown in the grammar as such but, rather, to some kind of unreliability in the interface between the syntax and other areas of the grammar" (White, 2003, p. 201). There is still no widely accepted explanation of the fact that some gender agreement production becomes fossilized in some L2 learners and not in others; this remains a controversial issue in L2 acquisition research.

With respect to the second research question, the gender-related pattern distributions observed in the data were very similar in both the heritage and the L2 groups. This was unsurprising, since previous L1 and L2 Spanish gender findings provide ample evidence for significant effects of noun gender, noun morphology, and agreement domain on the accuracy of gender agreement. For example, the existence of a masculine default setting, in both the lexical and the syntactic levels of gender, has substantial support (e.g., Cain, Weber-Olsen & Smith, 1987; Schlig, 2003; White et al., 2001). In the present study, however, the effects of the masculine gender on accuracy were significant only in the oral production task. L2 learners, in particular, were more likely to incorrectly assign masculine gender to feminine nouns than feminine gender to masculine nouns during oral elicitation. This implies that the masculine gender emerges as the default in Spanish (see Harris, 1991) when learners are unable to link the abstract gender feature to its appropriate forms in spontaneous oral production of gender agreement. The use of a default in oral production reflects a mapping problem, a performance issue, rather than a problem in the underlying representation of gender in the speaker's grammar. Evidence supporting this claim comes from longitudinal (e.g., Hernández-Pina, 1984; López-Ornat, 1997) and cross-sectional L1 Spanish studies (e.g., Brisk, 1976; Cain et al., 1987) showing that monolingual children tend to overuse and to be more accurate with the masculine forms of determiners and modifiers. A significant masculine default effect has also been found in psycholinguistic studies of adult Spanish monolinguals. For example, lexical-decision experiments indicate that differences in reaction times are more significant with masculine than with feminine words (Domínguez, Cuetos & Segui, 1999). In L2 acquisition, this masculine default is pervasive and persistent in the learners' interlanguage, and in some cases, for unknown reasons, becomes a long-term or even permanent problem. Moreover, the emerging body of research on Spanish heritage language (particularly Montrul et al., 2008), including the present study, shows that the masculine default is also found in heritage speakers' oral production, indicating that it is part of both complete and incomplete grammars. Therefore, the masculine default is a linguistic strategy used by monolinguals and early and late bilinguals that significantly affects oral production data as a result of a variety of performance factors, including processing demands and communication pressure.

In relation to noun morphology, the present results show that heritage and L2 learners were significantly more accurate with overtly than with non-overtly marked nouns, both in comprehension and in production. An examination of previous L1 and adult L2 production data suggests that the gender of overt ending nouns is acquired before that of non-overt ending nouns, and that accuracy is higher with overt than with non-overt nouns (e.g., Fernández-García, 1999; Finnemann, 1992; Franceschina, 2001; González, 1978; Hernández-Pina, 1984). The present findings confirm those results. Noun morphology aids in gender agreement, and is a reliable cue for establishing correct agreements in the phrase and sentence. According to the Competition Model (Bates & MacWhinney, 1982, 1987), an input processing model that tries to explain the ways in which speakers of different languages are affected by a variety of linguistic cues during syntactic processing, when encoding the notion of agency, English speakers rely primarily on word order, while Japanese speakers prefer animacy as a cue, and German speakers use case marking. The strength of competing linguistic cues varies with the language. In Spanish, a morphologically rich language, native speakers use overt morphology as a strong linguistic cue for gender agreement in both comprehension and production. Psycholinguistic research measuring both accuracy and reaction times of gender agreement demonstrate forcefully that morphology is a valid cue for native speakers of Spanish: they are more accurate and faster with overt than with non-overt ending nouns (e.g., Alarcón, 2006, 2009). The present results show that overt morphology is also a valid cue for advanced heritage and L2 learners when establishing gender agreement in the comprehension and production of noun phrases. That both groups responded to the same linguistic gender cue (noun ending) that native Spanish speakers use for establishing correct gender agreement suggests that advanced heritage and L2 learners process gender agreement in the same way as native Spanish speakers.

Previous L1 and L2 studies examining gender agreement within the noun phrase indicate that gender

https://doi.org/10.1017/S1366728910000222 Published online by Cambridge University Press

agreement is acquired earlier with articles than with adjectives, and that accuracy with articles is higher than with adjectives (e.g., Bruhn de Garavito & White, 2002; Fernández-García, 1999; Hernández-Pina, 1984). The present results for heritage speakers support these findings: heritage speakers were more accurate with agreement with articles than with adjectives. But the present findings for L2 learners are inconsistent with those earlier results, since the L2 learners were equally accurate with both adjectives and articles. The salient question is why the agreement domain did not affect the L2 learners in the present study as it affected the comprehension and production gender behavior of L2 learners in the earlier findings? One possible explanation is that much of the earlier research exploring noun + adjective agreement focused exclusively on either predicative (e.g., Alarcón, 2006, 2009) or attributive (e.g., Sagarra & Herschensohn, 2008) adjective agreement. Results might have been different if both types of adjectives had been investigated simultaneously. For example, Bruhn de Garavito and White (2002) found that low-proficiency French-speaking learners of L2 Spanish produced more gender agreement errors with predicative adjectives (34.56%) than with attributive adjectives (26.95%). Moreover, Keating (2009) investigated the effects of syntactic distance on the processing of gender agreement by comparing agreement with adjectives that were adjacent to the noun (attributive position) and adjectives in another syntactic phrase and clause (predicative position). He measured the eye movements of native speakers and of L2 learners at three proficiency levels as they read sentences on a computer screen. Results showed that only the advanced learners, and only with attributive adjectives, performed like native speakers. Since attributive and predicative adjectives differ in their distance from the target noun, the difference in the processing of gender agreement involving these two types of adjectives can be explained by the syntactic distance between the target noun and the agreeing adjective. The Structural Distance Hypothesis and the Linear Distance Hypothesis, which have been used to explain the difficulty in processing of grammatical structures such as relative clauses in L2 Korean by English-speaking learners (O'Grady, Lee & Cho, 2003), also predict that agreement rates with predicative adjectives will be lower than with attributive adjectives. Since the present study was limited to noun phrase agreement, only attributive adjectives were included. But in terms of syntactic distance, there is no difference between agreement with articles and with attributive adjectives. Therefore, the present results, in which advanced L2 learners, on both comprehension and production tasks, performed equally well with articles and attributive adjectives, are consistent with the prediction derived from the distance hypotheses.

Conclusion

This study supports the view that Spanish gender agreement is acquirable irrespective of the age of acquisition and the status of the gender features in the learner's L1. The advanced L2 learners in the study displayed the same level of implicit grammatical gender knowledge on the comprehension task as the advanced heritage speakers. Not only were the accuracy scores of the two groups high, their patterns of gender agreement were consistent in both comprehension and production. Therefore, as predicted by full access accounts, learners who are not exposed to grammatical gender until after the critical period might still acquire abstract gender representation in their underlying grammar.

Nonetheless, the L2 learners alone showed deficits in spontaneous oral production. Gender agreement variability at advanced levels of L2 proficiency stems from performance issues (i.e., production-based factors) rather than competence, and is predicted by full access accounts such as the Missing Surface Inflection Hypothesis. But why is there a divergence between these two advanced proficiency-matched groups in oral production? The fundamental difference between the two groups involves the age of first exposure to Spanish. Therefore, I argue that post-critical period learners, due to their late acquisition of gender, are more susceptible to computation deficits in gender agreement than those who acquire the language at birth, and who have mastered the gender system by the time they are exposed to English. The present findings suggest that maturational constraints affect gender agreement performance, but not its representation in L2 acquisition. Consequently, the L2 learners showed incomplete acquisition of Spanish gender, which is, indeed, expected in L2 acquisition (Bley-Vroman, 1989; VanPatten, 2003).

The present study also provides evidence that heritage speakers do achieve complete acquisition of gender agreement, as shown by their consistently high accuracy rates on both written comprehension and oral production tasks, which approached the levels of native speakers. Why, then, did Montrul et al.'s (2008) heritage speakers exhibit problems on their gender comprehension task, and thus incomplete acquisition? By definition, language acquisition occurs early in childhood in all heritage speakers, so there must be factors other than age of exposure that affect the acquisition of gender. The heritage learners in Montrul et al. (2008) were exposed to English before the age of 5 (p. 518), so it is possible that their Spanish gender systems were not fully developed by that time, and subsequently either fossilized or atrophied. Another possibility is that the gender system was already in place, but was in some way attenuated by the early Spanish–English contact situation (see Silva-Corvalán, 1994). In both cases, those heritage learners would have

experienced a reduction in meaningful Spanish input that blocked their ultimate attainment of gender. In contrast, the heritage speakers in the present study were exposed exclusively to Spanish during their pre-school years, and were introduced to English after age 5. By that time, their gender agreement system had been fully developed and consolidated, and was impervious to English influence. Consequently, their acquisition of Spanish gender was complete.

Several interesting pedagogical questions for further research emerge from these conclusions, the answers to which would also substantially enhance our understanding of early bilingualism in the L2 context. What are the effects, if any, of formal instruction on gender agreement in the implicit knowledge systems of heritage speakers? For example, do lower-proficiency heritage speakers receiving formal training overcome their difficulties with grammatical gender? Can advanced-proficiency learners become even more native-like if exposed to formal Spanish instruction in gender? The biggest challenge for heritage language teaching programs is to design curricula that effectively help heritage learners to (re)acquire, develop, and maintain their heritage languages according to their specific sociolinguistic profiles and concrete linguistic needs. To date, as Valdés, Fishman, Chávez and Pérez (2008, p. 21) have acknowledged, "very little empirical research is available about the outcomes of different kinds of instruction given particular goals". There is a pressing need to investigate which types of instruction best help heritage speakers reacquire gender and/or maintain their existing gender systems.

Appendix. *Último Acto, by Pablo de la Torriente Brau*

Adapted by the researcher from: http://www.lainsignia.org/2003/mayo/cul_040.htm.

En (1) un/una sector del jardín, en (2) el/la lugar donde se elevaba (3) el/la árbol de la palma, el hombre esperaba. La noche (4) profundo/profunda lo envolvía todo. Sólo se escuchaba el ruido (5) monótono/monótona del silencio (6) lejano/lejana. Su traje (7) oscuro/oscura, lo convertía en sombra (8) intenso/intensa. Sus brazos (9) poderosos/poderosas, manchados por (10) el/la aceite de las máquinas, apenas se distinguían. Estaba inmóvil. Esperaba.

Aquélla era su casa, pero en (11) el/la medianoche llena de frío él esperaba. Dentro del (12) amplio/amplia bolsillo de su pantalón, su mano ruda de hombre de las máquinas contenía (13) el/la papel, hallado casualmente en la oficina, hacía apenas (14) un/una hora, cuando fue a hacer una consulta (15) serio/seria al Ingeniero Jefe. Había visto un sobre (16) dirigido/dirigida a su mujer, abandonado en la mesa, lo había cogido y ahora estaba detrás de la palma, al momento de la cita (17) trágico/trágica. La carta decía: "Esta noche tu marido está de guardia en la casa de máquinas y a las doce iré de todas maneras . . . " "de todas maneras" estaba subrayado. Era el "Administrador del Ingenio" quien la firmaba. El hombre sólo había tenido tiempo para correr a su casa y esconderse en (18) el/la patio. Todavía su cerebro (19) confundido/confundida por la sorpresa, por (20) el/la rabia y por la humillación (21) completo/completa no reaccionaba.

Y poco antes de las doce apareció "el otro" que, después de asomar su cabeza por (22) el/la muro, miró cuidadosamente el entorno. Luego, salió con mucha precaución. Venía con (23) un/una camisa de kaki. "El otro" se detuvo un momento para escuchar (24) los/las sonidos de la noche, el latido de su corazón (25) asustado/asustada . . . (Desde detrás de la palma (26) los/las ojos de acero que lo espiaban llegaron a esta conclusión (27) descriptivo/descriptiva: "¡Si es un cobarde! . . . ") "El otro" fue avanzando con cuidado y llegó hasta la misma palma . . . Es extraño, pero "el otro" no percibió la presencia del enemigo . . .

Fue todo rápido, eléctrico. La mano de acero del hombre de las máquinas apresó su garganta y ahogó (28) el/la grito terrible. Y el golpe (29) violento/violenta lo dejó en el acto sin sentido. El hombre de las máquinas, furioso, no tuvo (30) el/la paciencia que se había propuesto y ahora estaba de pie, al lado de "el otro", contemplando su puño lleno de sangre (31) rojo/roja y con la mente (32) vacío/vacía de impresiones. Así estuvo un rato, inmóvil, como un tronco, cuando pensó: "Si no he podido hablar con él, hablaré con ella". Le pegó una patada brutal al muerto y se dirigió a la casa . . . Iba con (33) el/la invisible actitud y (34) silencioso/silenciosa velocidad de un gato negro.

Cerca ya de la puerta del fondo, se detuvo. Un miedo (35) raro/rara lo había paralizado. Por un momento le asaltó (36) el/la emoción perturbadora de que él era en realidad el amante, que era a él a quien ella esperaba. Y su corazón se le agitó con (37) perverso/perversa esperanza y tuvo (38) el/la temor del burlador. Cuando llegó a la puerta se puso a escuchar y no oía nada. Sin embargo, sintió que cerca estaba de ella. Hizo (39) un/una suave presión sobre la puerta para que se abriera . . . Pensó: "¡Ella lo esperaba!" Y la ira le hizo empujar la puerta con fuerza . . . Pero antes de llegar a dar dos pasos, sintió el balazo en (40) el/la cuerpo y (41) el/la voz de ella que decía: "¡Canalla, te lo dije! . . . "

Luego, se escuchó un "¡Ay!" de dolor (42) inmenso/inmensa y de sorpresa (43) espantoso/espantosa. Ella llenó un espacio de silencio y asombro. Luego, cuando encendió corriendo (44) el/la luz, él vio (45) el/la cara de su esposa reflejando una pena (46) infinito/infinita. Arrodillada estaba a su lado y, con (47) un/una angustia insoportable, le decía "¿Por qué, por qué . . . ?" Sin comprender el error (48) extremo/extrema y fatal . . . Pero ya el rostro del hombre comenzaba a ponerse alegre, alegre, como un niño que mejora . . .

https://doi.org/10.1017/S1366728910000222 Published online by Cambridge University Press

References

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

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

Alarcos, E. (1999). *Gramática de la lengua española*. Madrid: Espasa Calpe.

Andersen, R. (1984). What's gender good for, anyway? In R. Andersen (ed.), *Second languages: A cross-linguistic perspective*, pp. 77–99. Rowley, MA: Newbury House.

Bates, E., & MacWhinney, B. (1982). Functionalist approaches to grammar. In E. Wanner & L. R. Gleitman (eds.), *Language acquisition: The state of the art*, pp. 173–218. Cambridge: Cambridge University Press.

Bates, E., & MacWhinney, B. (1987). Competition, variation, and language learning. In B. MacWhinney (ed.), *Mechanisms of language acquisition*, pp. 157–193. Hillsdale, NJ: Lawrence Erlbaum.

Bley-Vroman, R. (1989). What is the logical problem of foreign language learning? In S. M. Gass & J. Schachter (eds.), *Linguistic perspectives on second language acquisition*, pp. 41–68. Cambridge: Cambridge University Press.

Brisk, M. E. (1976). The acquisition of Spanish gender by first grade Spanish speaking children. In G. D. Keller, R. V. Teschner & S. Viera (eds.), *Bilingualism in the bicentennial and beyond*, pp. 143–160. New York: Bilingual Press.

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

Cain, J., Weber-Olsen, M., & Smith, R. (1987). Acquisition strategies in a first and second language: Are they the same? *Journal of Child Language*, *14* (2), 333–352.

Carroll, S. (1989). Second language acquisition and the computational paradigm. *Language Learning*, *39*, 535–594.

Clahsen, H. (1988). Parameterized grammatical theory and language acquisition: A study of the acquisition of verb placement and inflection by children and adults. In S. Flynn & W. O'Neil (eds.), *Linguistic theory in second language acquisition*, pp. 47–75. Dordrecht: Kluwer.

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

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

Davies, M. (2006). *A frequency dictionary of Spanish: Core vocabulary for learners*. New York: Routledge.

Dewaele, J.-M., & Veronique, D. (2001). Gender assignment and gender agreement in advanced French interlanguage: A cross-sectional study. *Bilingualism: Language and Cognition*, *4* (3), 275–297.

Domínguez, A., Cuetos, F., & Segui, J. (1999). The processing of grammatical gender and number in Spanish. *Journal of Psycholinguistic Research*, *28* (5), 485–498.

Ellis, N. C., & Schmidt, R. (1997). Morphology and longer distance dependencies. *Studies in Second Language Acquisition*, *19*, 145–171.

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

Ellis, R. (2008). Investigating grammatical difficulty in second language learning: Implications for second language acquisition research and language testing. *International Journal of Applied Linguistics*, *18* (1), 4–22.

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. *International Review of Applied Linguistics in Teaching*, *30* (2), 121–136.

Franceschina, F. (2001). Against an L2 morphological deficit as an explanation for the differences between native and non-native grammars. *EUROSLA Yearbook*, *1*, 143–158.

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

Gass, S. M., & Selinker, L. (2008). *Second language acquisition: An introductory course*. New York: Routledge.

Gómez, L. (1998). *Gramática didáctica del español*. Madrid: Ediciones SM.

González, G. (1978). *The acquisition of Spanish grammar by native Spanish speaking children*. Rosslyn, VA: InterAmericana Research Associates.

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

Harris, J. (1991). The exponence of gender in Spanish. *Linguistic Inquiry*, *22* (1), 27–62.

Hawkins, R. (2000). Persistent selective fossilization in second language acquisition and the optimal design of the language faculty. *Essex Research Reports in Linguistics*, *34*, 75–90.

Hawkins, R., & Chan, Y.-H. C. (1997). The partial availability of Universal Grammar in second language acquisition: The 'failed functional features hypothesis'. *Second Language Research*, *13*, 187–226.

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

Hoekstra, T., & Schwartz B. D. (eds.) (1994). *Language acquisition studies in generative grammar*. Amsterdam: John Benjamins.

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

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: Lawrence Erlbaum.

Lloret, M. R., & Viaplana, J. (1998). El binarismo del género gramatical en castellano y catalán. *VERBA*, *25*, 71–91.

López-Ornat, S. (1997). What lies in between a pre-grammatical and a grammatical representation? Evidence on nominal and verbal form–function mappings in Spanish from 1;7 to 2;1. In A. T. Pérez-Leroux & W. R. Glass (eds.), *Contemporary perspectives on the acquisition of Spanish* (vol. 1): *Developing grammars*, pp. 3–20. Somerville, MA: Cascadilla Press.

Maratsos, M. (1988). The acquisition of formal word classes. In Y. Levy, I. Schlesinger & M. Braine (eds.), *Categories and processes in language acquisition*, pp. 31–44. Hillsdale, NJ: Lawrence Erlbaum.

Mariscal, S. (2009). Early acquisition of gender agreement in the Spanish noun phrase: Starting small. *Journal of Child Language*, *36*, 143–171.

McCarthy, C. (2008). Morphological variability in the comprehension of agreement: An argument for representation over computation. *Second Language Research*, *24* (4), 459–486.

Meisel, J. M. (1991). Principles of Universal Grammar and strategies of language learning: Some similarities and differences between first and second language acquisition. In L. Eubank (ed.), *Point counterpoint: Universal Grammar in second language*, pp. 324–376. Amsterdam: John Benjamins.

Montrul, S., Foote, R., & Perpiñán, S. (2008). Gender agreement in adult second language learners and Spanish heritage speakers: The effects of age and context of acquisition. *Language Learning*, *58* (3), 503–553.

Montrul, S., & Potowski, K. (2007). Command of gender agreement in school-age Spanish–English bilingual children. *International Journal of Bilingualism*, *11* (3), 301–328.

O'Grady, W., Lee, M., & Cho, M. (2003). A subject–object asymmetry in the acquisition of relative clauses in Korean as a second language. *Studies in Second Language Acquisition*, *25*, 433–448.

Oliphant, K. (1998). Acquisition of grammatical gender in Italian as a foreign language. *The Canadian Modern Language Review*, *54* (2), 239–262.

Ortega, L. (2005). For what and for whom is our research? The ethical as transformative lens in instructed SLA. *The Modern Language Journal*, *89* (3), 427–443.

Pérez-Pereira, M. (1991). The acquisition of gender: What Spanish children tell us. *Journal of Child Language*, *18* (3), 571–590.

Polinsky, M. (2008). Gender under incomplete acquisition: Heritage speakers' knowledge of noun categorization. *Heritage Language Journal*, *6* (1), 40–71.

Prévost, P., & White, L. (2000). Missing surface inflection or impairment in second language acquisition? Evidence from tense and agreement. *Second Language Research*, *16* (2), 103–133.

Rogers, M. (1987). Learners' difficulties with grammatical gender in German as a foreign language. *Applied Linguistics*, *8* (1), 48–74.

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

Sagarra, N. (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.), *Romance Linguistics 2006: Selected papers from the 36th Linguistic Symposium on Romance Languages (LSRL), New Brunswick, March–April 2006*, pp. 241–254. Amsterdam and Philadelphia: John Benjamins.

Sagarra, N., & Herschensohn, J. (2008). Processing gender in L2 Spanish. In H. Chan, H. Jacob & E. Kapia (eds.), *Proceedings of the 32nd Annual Boston University Conference on Language Development*, pp. 427–437. Somerville, MA: Cascadilla Press.

Schlig, C. (2003). Analysis of agreement errors made by third-year students. *Hispania*, *86* (2), 312–319.

Schriefers, H., & Jescheniak, J. D. (1999). Representation and processing of grammatical gender in language production: A review. *Journal of Psycholinguistic Research*, *28* (6), 575–600.

Schwartz, B. D., & Sprouse, R. (1994). Word order and nonnative case in nonnative language acquisition: A longitudinal study of (L1 Turkish) German interlanguage. In Hoekstra & Schwartz (eds.), pp. 317–368.

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

Silva-Corvalán, C. (1994). *Language contact and change: Spanish in Los Angeles*. Oxford: Oxford University Press.

Taraban, R., & Kempe, V. (1999). Gender processing in native and nonnative Russian speakers. *Applied Psycholinguistics*, *20*, 119–148.

Teschner, R. V., & Russell, W. M. (1984). The gender patterns of Spanish nouns: An inverse dictionary-based analysis. *Hispanic Linguistics*, *1*, 115–132.

Tokowicz, N., & MacWhinney, B. (2005). Implicit and explicit measures of sensitivity to violations in second language grammar. *Studies in Second Language Acquisition*, *27*, 173–204.

Tucker, G. R., Lambert, W. E., & Rigault, A. A. (1977). *The French speaker's skill with grammatical gender: An example of rule-governed behavior*. The Hague: Mouton.

Vainikka, A., & Young-Scholten, M. (1994). Direct access to X′ theory: Evidence from Korean and Turkish adults learning German. In Hoekstra & Schwartz (eds.), pp. 265–316.

Vainikka, A., & Young-Scholten, M. (1996). Gradual development of L2 phrase structure. *Second Language Research*, *12*, 7–39.

Valdés, G. (1997). The teaching of Spanish to bilingual Spanish-speaking students: Outstanding issues and unanswered questions. In M. C. Colombi & F. X. Alarcón (eds.), La *enseñanza del español a hispanohablantes: Praxis y teoría*, pp. 8–44. Boston, MA: Houghton Mifflin.

Valdés, G. (2000). *Spanish for native speakers* (vol. 1) (AATSP Professional Development Series Handbook for Teachers K-16). New York: Harcourt College Publishers.

Valdés, G. (2001). Heritage language students: Profiles and possibilities. In J. K. Peyton, D. A. Ranard & S. McGinnis (eds.), *Heritage languages in America: Preserving a national resource*, pp. 37–77. Washington, DC: Center for Applied Linguistics/Delta Systems.

Valdés, G. (2005). Bilingualism, heritage language learners, and SLA research: Opportunities lost or seized? *The Modern Language Journal*, *89* (3), 410–426.

Valdés, G. (2006). Making connections: Second language acquisition research and heritage language teaching. In R. Salaberry & B. Lafford (eds.), *The art of teaching Spanish: Second language acquisition from research to praxis*, pp. 193–212. Washington, DC: Georgetown University Press.

Valdés, G., Fishman, J. A., Chávez, R., & Pérez, W. (2008). Maintaining Spanish in the United Status: Steps toward the effective practice of heritage language re-acquisition/development. *Hispania*, *91* (1), 4–24.

VanPatten, B. (2003). *From input to output: A teacher's guide to second language acquisition*. Boston, MA: McGraw-Hill.

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

White, L., Valenzuela, E., Kozlowska-Macgregor, M., & Leung, I. (2003). Gender agreement in L2 Spanish: Evidence against failed features. *McGill Working Papers in Linguistics*, *17* (2), 153–182.

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

White, L., Valenzuela, E., Kozlowska-Macgregor, M., Leung, I., & Ayed, H. B. (2001). The status of abstract features in interlanguage: Gender and number in L2 Spanish. In A. H.-J. Do (ed.), *Proceedings of the 25th Annual Boston University Conference on Language Development*, pp. 792–802. Somerville, MA: Cascadilla Press.