Bilingualism: Language and Cognition 14 (3), 2011, 318–331 © Cambridge University Press 2010 doi:10.1017/S1366728910000325

Whose? L2-English speakers' possessive pronoun gender errors*

INÉS ANTÓN-MÉNDEZ
Utrecht University, The Netherlands
& University of New England, Australia

(Received: November 8, 2009; final revision received: June 18, 2010; accepted: July 3, 2010; First published online 15 November 2010)

This article reports the results of an experiment on production of his/her *in English as a second language (L2) by proficient native speakers of Italian, Spanish, and Dutch. In Dutch and English, 3rd person singular possessive pronouns agree in gender with their antecedents, in Italian and Spanish possessives in general agree with the noun they accompany (possessum). However, while in Italian the 3rd person singular possessives overtly agree in gender with the possessums, in Spanish they lack overt morphological gender marking. Dutch speakers were found to make very few possessive gender errors in any condition, Spanish and Italian speakers, on the other hand, behaved like Dutch speakers when the possessum was inanimate, but made more errors when it was animate (e.g.,* his mother*). Thus, even proficient L2 speakers are susceptible to the influence of automatic processes that should apply in their first language alone. The pattern of results has implications for pronoun production and models of bilingual language production.*

Keywords: bilingualism, possessive pronouns, second language speakers, gender errors

Introduction

Speakers of a second language are susceptible to making errors in their L2 for a variety of reasons, and these errors have been interpreted and classified in a variety of ways (e.g., Epstein, Flynn & Martohardjono, 1996; McLaughlin, 1987; Poulisse, 1999). However, independently of which framework one chooses to adopt, L2 errors could be defined as a failure to implement the correct procedure, and could be classified as transient or consistent with respect to the frequency with which they appear, and temporary or persistent with respect to their development over time. Adopting a framework based on cognitive mechanisms (more in line with McLaughlin, 1987) and the automatization of linguistic procedures, the cause of temporary L2 errors is likely to be poor knowledge of the second language – lack of knowledge of the correct forms in the new language or imperfect implementation of the newly acquired knowledge. With increasing proficiency and practice, many of these errors disappear but not all. The errors which appear with certain frequency and persist regardless of proficiency can still be due to either difficulties with the acquisition of particular sorts of knowledge (declarative knowledge) or difficulties with the automatic implementation of certain processes (procedural knowledge) (Ullman, 2004, 2005). At high proficiency levels, however, both types of difficulties seem to be very much dependent on differences between the first and the second languages.

At the risk of oversimplifying, it could be said that persistent errors of proficient second language speakers have one of three possible causes which could be dubbed: unlearnability, insufficient automatization, and excess automatization. The first possible cause, unlearnability, refers to the codification in the second language of a distinction that is not codified in the first language when this distinction is not easily learned later in life. In such cases of acquisition failure, speakers may not even be sure of the appropriate use of the different L2 terms even in off-line tasks that draw on explicit processing and declarative memory, making these errors as likely in edited written production as in spontaneous speech production. An example of this type of error at the lexical level is the confusion in L2 Spanish between the verbs *ser* and *estar* (both "to be" – loosely, *ser* equates and tends to indicate more permanent attributes, *estar* tends to indicate temporary attributes and is also a locative); another example at the syntactic or morphological level is the incorrect use of the indicative and subjunctive verbal modes for English–Spanish L2 speakers, a distinction that is not as salient in their native English. Studying differences in learnability of different linguistic constructs may help find out why some features are sensitive to a

* This research was funded by a grant from the Netherlands Organization for Scientific Research (NWO) to the author (VENI 275–70-019). I would like to thank my participants for their generosity with their time, my Italian informants, Anna Cavallaro, Franko Leoni, Ivano Caponigro, and Giuli Dussias for their valuable help with the preparation of the stimuli, and my research assistants, Jasmijn Claij and Mee Wun Lee, for their great work transcribing the recordings.

Address for correspondence:
Psychology, SBCSS, University of New England, Armidale, NSW 2351, Australia
iantonme@une.edu.au

Table 1. *Examples of third person possessive phrases in the three languages for a female possessor-antecedent.*

	English	Dutch	Italian	Spanish
AN	her-FEM father-MASC	haar-FEM vader-CG	suo-MASC padre-MASC	su-Ø padre-MASC
AN	her-FEM mother-FEM	haar-FEM moeder-CG	sua-FEM madre- FEM	su-Ø madre-FEM
IN	her-FEM dream-Ø	haar-FEM droom-CG	suo-MASC sogno-MASC	su-Ø sueño-MASC
IN	her-FEM house-Ø	haar-FEM huis-NEUT	sua-FEM casa- FEM	su-Ø casa- FEM

AN = animate possessum; IN = inanimate possessum; FEM = feminine; MASC = masculine; CG = common gender; NEUT = neuter; Ø = no gender

critical period of language acquisition while others are easily acquired at any age (Slobin, 1993; but see also Ullman, 2005) possibly due to their being dependent on cognitive skills which remain fully functional in adulthood. The second possible cause of persistent errors is insufficient automatization – i.e., when the second language requires the automatic implementation of a certain syntactic or morphological procedure that is not required in the first, and the automatization is difficult to achieve at all or is difficult to implement in a consistent manner, even when the procedure itself is not difficult to learn on theoretical basis. Such is the case, for example, of gender agreement in L2 when the L1 either lacks a system of gender agreement or possesses a significantly different one (Blom, Polišenská & Weerman, 2008; Sabourin & Stowe, 2008). In this instance, only spontaneous speech production may be affected, while off-line tasks may result in very low error rates. The third possible reason for persistent errors is somewhat different from the previous two in that, while the first two are related to acquisition of L2 features, the third is related to processes that play in L1 – what has traditionally been labeled L1 transfer, and which Truscott and Sharwood Smith (2004) attribute to competition between two automatic procedures. In this case, the first language requires the implementation of an automatic procedure that is difficult to switch off when speaking the L2 (e.g., word order when the two languages differ). Being again a matter of automatization ("excess" automatization in this case), here too the resulting errors are more likely to surface during spoken language production than in samples of edited written production. These errors, being the result of an automatic process inappropriately applied, can inform and constrain models of bilingual production, and can also be a useful source of data to investigate more general processes of language production (Antón-Méndez, 2010).

This article reports an exploration of one particular L2 error which may result from a mix of insufficient automatization and excess automatization at different processing levels: 3rd person singular possessive pronoun gender errors of Italian and Spanish native speakers in L2 English. The correct use of the English possessive pronouns *his* and *her* has been found to be difficult to acquire by French and Spanish speakers learning English (Collins et al., 2009; White, Muñoz & Collins, 2007). What is more, problems with the use of third person singular possessives persist in even proficient L2 English speakers.

As "noun proxies", pronouns usually agree in all features with the nouns they stand in for – person, number, gender and animacy. This is so both in English and in the Romance languages, and it is evident in nominative, accusative and dative pronouns in these languages. Genitive or possessive pronouns are somewhat different because, while they do refer to a noun, in some languages (e.g., English and Spanish) they actually take the function of a determiner, while in other languages (e.g., Italian) they function as adjectives. Possibly as a consequence of this, in a language such as Italian, where there is grammatical gender agreement between nouns and adjectives, and in a language such as Spanish, where there is gender agreement between nouns and determiners, the possessive pronoun (when marked for gender) has to agree in gender with the noun it accompanies instead of with its anaphoric antecedent. This means that an Italian–English bilingual would be faced with a conflict when processing the gender feature of a possessive pronoun – in English it should agree with the possessor antecedent, a form of agreement that can be considered semantic in nature (Corbett, 2006, p. 207), and in Italian it should agree with the accompanying possession noun (henceforth, possessum), a purely syntactic form of agreement (see Table 1 for examples). Spanish patterns with Italian in having nominal grammatical gender as well as syntactic gender agreement between possessives and possessums, although 3rd person possessive pronouns in particular do not present any overt gender markers in this language (*su* is the Spanish equivalent of "his", "her", "its" and "their"). Some Spanish possessive pronouns, however, do have overt gender morphemes agreeing with the possessum in the same noun phrase (NP), namely the 1st and 2nd person plural (*nuestra*/*nuestro* "our" and *vuestra*/*vuestro* "your.PL"). In this sense, Spanish speakers are expected to have to confront the same conflict as Italian speakers

https://doi.org/10.1017/S1366728910000325 Published online by Cambridge University Press

in L2 English if, as would be expected, their L1's lack of overt gender marking for the 3rd person singular possessive is a phonological phenomenon independent of the existence of underlying gender agreement between possessum and possessive. Nevertheless, the possibility exists that the absence of a gender marker for this particular possessive pronoun stems from a genuine lack of gender agreement processing for phrases containing a 3rd person singular possessive pronoun. In that case, Spanish–English speakers should behave differently than Italian–English speakers.

The experiment reported below compares the production rates of possessive pronoun gender errors in L2 English by three different groups of native speakers: Italian and Spanish as the two experimental groups, and Dutch as a control group, since gender agreement for possessive pronouns in Dutch mirrors that of English. The experiment is designed to study the effects of the two conflicting processes in the two experimental groups' production of L2 English to see the degree to which it is possible to automatize the L2 process of semantic gender agreement, and what the influence is of the L1 automatic process of syntactic gender agreement depending on specific linguistic conditions, namely the presence of animate or inanimate possessums. To begin with, the performance of Italian and Spanish speakers is expected to differ from that of Dutch speakers if the different sort of gender agreement in L1 does indeed pose a problem for even advanced L2 speakers of English. Furthermore, because animate nouns and inanimate nouns vary with respect to the nature of the gender features in L1 Spanish, Italian and Dutch, as well as their association with a gender feature in L2 English, the inclusion of the two types of possessums should help us pinpoint the locus in the processing chain where the errors take place. If errors occur only for animate possessums and equally for all speaker groups, the effect is likely to be due to semantic interference between the natural genders of the two nouns involved, the possessor and the possessum (Slevc, Wardlow Lane & Ferreira, 2007). If Italian (and presumably Spanish) speakers produce more errors than Dutch speakers for either animate or inanimate possessums, this difference is likely to be due to the influence of gender features at the grammatical level as a result of transfer of L1 syntactic processes. Finally, if NPs with inanimate possessums are associated with more errors for Italian (and presumably Spanish) than Dutch speakers, it would be of interest to see whether the gender errors reflect the grammatical gender of the L1 noun or not. If they do, it would mean that the erroneous implementation of syntactic gender agreement in L2 English for inanimate nouns depends on accessing the grammatical gender of the Italian (or Spanish) lexical item. This would constitute evidence on the question of the extent of concurrent activation of a bilingual's two lexicons. Alternatively, it may be that possessive gender errors in inanimate possessum NPs are unrelated to the L1 grammatical genders of the inanimate possessums, in which case the L1 syntactic process is applying in L2 production independently of the presence of specific gender features. Finally, the inclusion of a language, Spanish, that does not mark gender overtly for 3rd person singular possessive pronouns but is assumed to require syntactic gender agreement between possessum and possessive pronoun will shed light on whether certain syntactic operations take place even in the absence of morphological evidence. Ultimately, given that these errors are a case of bilingual production, they should be explainable by theories of bilingual language production and should therefore help specify the properties of a valid bilingual production model.

Method

Participants

Sixty-two speakers of English as a second language participated in this experiment: Twenty native speakers of Dutch, 24 native speakers of Italian, and 18 native speakers of Spanish. All had learnt English after the age of 12 (in the sense of starting to use the second language to try to communicate, not just being exposed to it in a classroom setting) and were fluent in English, with a mean proficiency rating of 4.26 ($SD = 0.92$) out of a maximum of 5, equivalent to a lower advanced level, as measured by the Quick Placement Test.[1]

Participants were living in the Netherlands where recruitment took place. The non-Dutch speakers were speaking English on a daily basis. A summary of the groups' characteristics is given in Table 2.

The three speaker groups were equivalent in terms of proficiency, $F(2,59) = 0.61$, $p = .55$, but differed significantly with respect to age, education level, age of English acquisition, years living in an English-speaking environment and frequency of English use. Pairwise comparisons of the most relevant factors carried out at the .05 significance level showed the Dutch group to be different to the two Romance language groups in terms of age of English acquisition (both ps $< .01$), and frequency of English use (both ps $< .01$), while the two Romance

[1] The QPT is a computer-based test commercialized by Oxford University Press, designed to place students of English as a second language in the appropriate level according to their proficiency. It assesses listening, reading, vocabulary and grammar, and provides a score, in accordance with the Association of Language Testers in Europe, of between 0 and 5, which correspond to the following Council of Europe's descriptions: beginner, elementary, lower intermediate, upper intermediate, lower advanced, and upper advanced. More information is available at http://www.oup.com/elt/catalogue/isbn/7162?cc=nl.

Table 2. *Participants' characteristics by speaker group: means (and SDs).*

	Dutch	Italian	Spanish
Age (years)	20.0 (2.1)^z	32.0 (7.7)^y	28.4 (4.3)^y
Proficiency (QPT[a])	4.4 (0.9)^z	4.3 (0.9)^z	4.1 (1.0)^z
Gender (Male %)	15.0%	41.7%	50.0%
Education level[b]	1.1 (0.5)^z	3.1 (0.9)^y	2.8 (0.9)^y
Age English acquired	13.6 (1.1)^z	18.8 (6.0)^y	19.6 (4.4)^y
English immersion[c] (years)	0.1 (0.1)^z	4.8 (5.2)^y	2.2 (1.8)^z
Frequency of English use[d]	3.5 (1.3)^z	4.9 (0.3)^y	4.9 (0.5)^y

[a]Quick Placement Test, see footnote 1 for details.
[b]Scale from 1 to 4, equivalent respectively to High School certificate, Bachelor's degree, Master's degree, and Ph.D. degree.
[c]Years spent in countries where the language of communication was English (including the Netherlands if no Dutch spoken).
[d]Scale from 1 to 5, equivalent respectively to hardly ever, on vacation, every now and then, weekly, and daily.
The superscripts *z* and *y* accompanying the means for each factor and language indicate whether the differences were statistically significant at the .05 level – equal superscripts in each row mean the means are not significantly different; different superscripts means they are.

groups did not differ from each other ($p = 1.00$ for both factors). With respect to years living in an English-speaking environment, the Italian group had spent more years speaking English than the other two, while the Spanish and Dutch groups were similar. In sum, all three language groups were similar in terms of proficiency but different regarding other characteristics. The differences, however, should not constitute a problem in this case given that they would, if anything, reduce the effect of interest – i.e., Italian- and Spanish-speaking participants (who are expected to produce more errors) have an advantage in English due to their speaking it considerably more frequently (Jia & Fuse, 2007); furthermore, Italian speakers had also been immersed in an English-speaking environment for longer periods. And, in any case, the marginal advantage of the younger acquisition age of the Dutch participants (no more than 6 years) is unlikely to have major effects on the results since it does not span the critical age (Johnson & Newport, 1989).

Materials

The experiment consisted of a series of 128 sentences paired with photographs of people. The images depicted a putative speaker (or speakers), and the sentences were meant to be read as an utterance produced by that speaker. There were 64 experimental sentences of the form: Poss-1ST.SG + Noun + Verb-PRESENT + Complement(s) (e.g., *My garden explodes into a million colors*). Sentences were between 7 and 10 words long, with an average of eight words per sentence. Each sentence was paired with two photographs, one of a male speaker and one of a female speaker, to create two conditions – matched (i.e., same gender for possessor and possessum) and mismatched (i.e., different genders for possessor and possessum). Two lists were made so that each sentence appeared only once in each list – paired with a female speaker in list A and with a male speaker in list B or vice versa. The lists were counterbalanced so that each had as many photographs of females as of males, and an equal number of sentences in each condition.

Conditions depended on the gender of the depicted speaker (the possessor antecedent) and on the nature of the noun in the possessive NP (the possessum). These nouns were either animate (e.g., *mother*, *father*) or inanimate (e.g., *dream*, *house*), and either masculine or feminine, with 16 nouns in each of the four conditions. In the case of inanimate nouns, the gender was based on the grammatical gender in the native language for Italian and Spanish, although not for Dutch where the gender system is not based on a masculine–feminine distinction (what were masculine and feminine nouns in old Dutch form in modern Dutch a single common gender class comprising two thirds of the nouns, with the other third belonging to a neuter gender class (see e.g., van Hout, 1996)). In defining the conditions for the Italian speakers, the gender assigned to English inanimate nouns was derived from the translations provided by four native speakers of Italian – only nouns that elicited a gender-consistent translation were included. Most of these nouns had the same gender in Spanish, with only four having to switch conditions in defining the conditions for the Spanish speakers.

All nouns were common English nouns. Inanimate nouns had a frequency of more than 37 occurrences per million according to the CELEX database (Baayen, Piepenbrock & Gulikers, 1995). Because animate nouns that have lexical gender in English are not exactly plentiful, it was necessary to resort to compounds such as *mother-in-law*, *stepfather*, etc., so that, even if the nouns themselves were not very frequent, the gender would still be transparent. Table 3 lists the conditions and their characteristics.

Apart from the experimental sentences, there were 64 fillers. Of these, 18 had the same form as the experimental items except that they contained the 1st person plural possessive pronoun *our*, which was followed by either an inanimate (9 sentences) or an epicene, i.e., a gender-neutral animate noun (9 sentences). Twenty fillers started with either singular or plural 1st person nominative pronouns. The other 26 fillers had other types of NPs as subjects and provided some variety. Finally, there were

https://doi.org/10.1017/S1366728910000325 Published online by Cambridge University Press

Table 3. *Examples of the experimental conditions with the average number of words (SDs) and possessum frequency.*

Possessum	Example	Number of words	Possessum frequency
AN FEM	My mother accompanies the teacher to the school.	7.9 (0.6)	50.1
AN MASC	My father gets a new position at another department.	8.3 (0.6)	52.8
IN FEM	My shirt shrinks after being machine washed.	8.3 (0.7)	200.2
IN MASC	My glass leaves a round mark on the table.	8.6 (0.7)	103.6

AN = animate possessum; IN = inanimate possessum; FEM = feminine possessum; MASC = masculine possessum
Each of these sentences appeared with either a photograph depicting a female or a male speaker giving rise to gender matched and mismatched conditions according to whether the possessor and the possessum had the same or different genders.

10 practice sentences representing the different types of sentences in the main part of the experiment.

Procedure

Participants were tested individually in a sound-proofed booth. Stimuli were presented on a computer screen and responses were digitally recorded. The procedure was as follows. After pressing the spacebar, there appeared on the top part of the screen a photograph of a person (or a group of persons in cases where the sentence had plural reference) and, below that, an English sentence appeared after 500 ms. The sentence disappeared after 3 seconds, and a tone was heard. This was the signal for the participant to start talking. The photograph remained on screen until the participant had finished talking and was ready for the next item. The experiment was self-paced.

The purpose of the experiment was disguised as related to memory in a second language to avoid participants consciously avoiding pronominal gender errors. Participants were told the sentences were statements uttered by the people on the photographs. Their task was to read the sentences silently and, after the sentence had disappeared and the tone was heard, re-tell what the person in the picture had said. During the practice session they were trained to re-tell the sentences bearing in mind they were talking about somebody else (which required them to change the person feature on the pronouns) and that the sentences had been uttered in the past (which required them to change the verb from present to past tense – this served as a further distractor to deflect attention from the pronoun which would otherwise have been the only modified constituent in the recast sentences). In this way, a sentence such as *My father gets a new position at another department* could become something like *His*/*her father got a new position at another department*. The retelling, however, hardly ever resulted in a sentence as similar to the original as the previous example but typically ended up containing the same general idea with different words (*Her*/*his father changed departments*), or being truncated (*His*/*her father got a new position*).

After the experiment was finished, participants completed a language history questionnaire. At this point, they were also probed as to whether they had guessed what the experiment was about. Only two Italian speakers and one Spanish speaker guessed the purpose correctly. This was deemed a small enough number of "guessers" not to jeopardize the presence of the effect (which could otherwise have been too reduced to reach significance) and, therefore, their data were included in the analyses.

Scoring and intrarater reliability

Participants' responses were transcribed twice independently and each transcription was then coded by the author. Responses were coded as correct, pronoun gender error, or other (mainly cases lacking possessive pronouns because the sentence had been recast in a different form altogether, but also sentences with an invalid possessum or where the possessive had not been changed to third person singular). To be counted as correct or pronoun gender error, the sentence had to have a third person singular pronoun and contain the same noun as the stimulus sentence in the case of inanimate nouns, or a noun preserving the natural gender of the stimulus sentence in the case of animate nouns (that is, if participants substituted *mother* for *stepmother*, the response was considered valid). Responses where the pronominal gender agreed with the possessor were coded as correct, and responses where the pronominal gender did not were coded as pronoun gender errors. The two transcriptions were then contrasted and all discrepancies (2.6% of the responses) checked in the original recordings and corrected by the author.

Results and discussion

Three items were found to elicit a very high proportion of errors (more than 2 standard deviations above the mean for the noun type) probably due to semantic anomalies or difficulties assessing the gender of the pictured speaker. One of them had an inanimate possessum (*My school admits only boys into the lower grades*), the other two

Table 4. *Average percentages of each response type per condition and speaker group (and SDs).*

	Animate				Inanimate			
	Feminine		Masculine		Feminine		Masculine	
	Match	Mismatch	Match	Mismatch	Match	Mismatch	Match	Mismatch
DUTCH Correct	92.50	93.75	97.33	94.67	94.67	95.63	96.88	98.00
	(0.13)	(0.08)	(0.06)	(0.08)	(0.08)	(0.06)	(0.07)	(0.05)
Gender errors	3.13	2.50	0.67	3.33	0.67	0.00	0.63	0.67
	(0.08)	(0.07)	(0.03)	(0.06)	(0.03)	(0.00)	(0.03)	(0.03)
Other errors	4.38	3.75	2.00	2.00	4.67	4.38	2.50	1.33
	(0.09)	(0.06)	(0.05)	(0.05)	(0.08)	(0.06)	(0.07)	(0.04)
ITALIAN Correct	85.94	88.02	92.22	81.11	96.67	93.75	95.83	98.33
	(0.18)	(0.15)	(0.12)	(0.20)	(0.07)	(0.09)	(0.06)	(0.05)
Gender errors	4.69	9.38	2.22	15.56	0.56	2.60	1.04	0.00
	(0.07)	(0.14)	(0.08)	(0.17)	(0.03)	(0.07)	(0.04)	(0.00)
Other errors	9.38	2.60	5.56	3.33	2.78	3.65	3.13	1.67
	(0.17)	(0.06)	(0.09)	(0.08)	(0.07)	(0.07)	(0.06)	(0.05)
SPANISH Correct	84.72	85.42	95.56	80.00	94.87	93.21	96.91	92.31
	(0.18)	(0.13)	(0.08)	(0.21)	(0.07)	(0.09)	(0.06)	(0.11)
Gender errors	9.03	9.72	1.75	16.30	0.86	2.47	1.85	2.56
	(0.15)	(0.13)	(0.04)	(0.18)	(0.03)	(0.06)	(0.05)	(0.06)
Other errors	6.25	4.86	2.69	3.70	4.27	4.32	1.23	5.13
	(0.08)	(0.08)	(0.07)	(0.07)	(0.07)	(0.06)	(0.05)	(0.08)

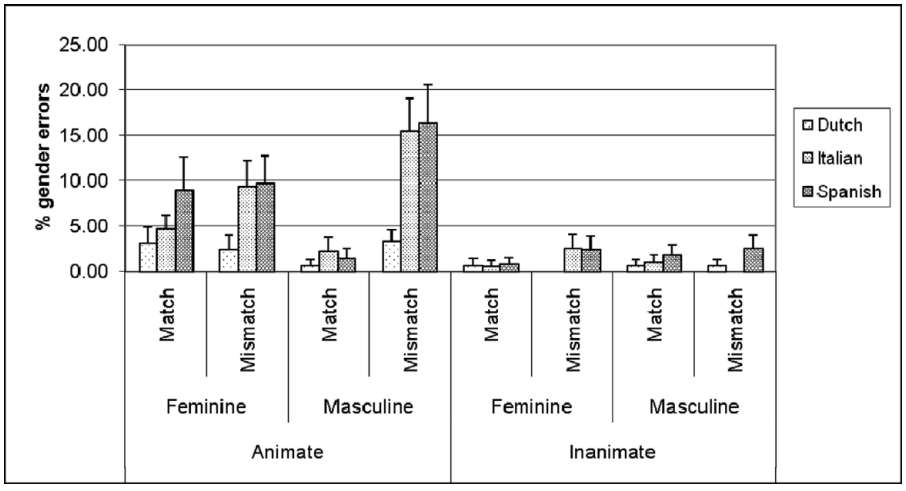

Figure 1. Percentage of possessive pronoun gender errors in the different conditions for the three different speaker populations.

belonged to the animate possessum set (*My sister-in-law feeds the baby before the party*, *My boy-friend brings toys as a peace offering*). These were removed from the analyses. Analyses were carried out on the rates of responses in each category with respect to the total number of responses per condition – i.e., on proportions of correct responses, gender errors, or other error types.

Of a total of 3844 responses, 3556 were correct (92.5%), 147 were gender errors (3.8%), and 144 were classified as other types of error (3.7%). For an overview of the whole data set, see Table 4, and Figure 1 for an overview of the pattern of gender errors.

First, in order to see whether the Italian and Spanish speakers behaved similarly despite their L1s' superficial

https://doi.org/10.1017/S1366728910000325 Published online by Cambridge University Press

differences, the two speaker groups were subjected to a 2 × 2 × 2 × 2 mixed design repeated measures ANOVA with language (Italian and Spanish) as a between subjects factor and animacy (animate vs. inanimate possessum), pronoun gender (feminine vs. masculine possessor), and congruency (same vs. different gender of the possessum) as within subjects factors. Language as a main factor was not significant, $F(1,40) = 0.75$, $p = .39$, and nor were any of the interactions with the other factors – animacy, $p = .85$, pronoun gender, $p = .88$, congruency, $p = .99$; the three-way interactions – language by animacy by pronoun gender, $p = .46$, language by animacy by congruency, $p = .53$, language by pronoun gender by congruency, $p = .35$; or the four-way interaction between language, animacy, pronoun gender and congruency, $p = .62$. This means both Romance languages behave similarly, as expected, despite the apparent differences in the form of the third singular possessive pronouns in the two languages (see introduction). We can therefore conclude that the fact that Spanish 3rd person singular possessives do not carry an overt gender marking must be a morphophonological 'accident' which does not reflect lack of underlying gender agreement within the noun phrase.[2]

Because it can be assumed that Spanish and Italian speakers should behave similarly and the results do not contradict this assumption, gender error data from Italian and Spanish speakers were pooled together and contrasted with Dutch, the control language. Furthermore, given the obvious differences between animate and inanimate nouns, their data were analyzed separately with the aim of maximizing statistical power and avoiding the proliferation of post-hoc tests. It should be stressed here that the two types of nouns were part of completely independent sets of items. Nevertheless, to first establish that indeed the behavior of the different speaker groups followed different patterns in relation to the different types of possessums, a comparison of the effect of animacy on the different languages was performed. A repeated measures ANOVA was conducted with language (Dutch vs. Romance) as a between subjects factor, and animacy (animate vs. inanimate possessum) as a within subjects factor. This resulted in a significant effect of language, $F(1,60) = 15.95$, $p < .01$; a significant effect of animacy, $F(1,60) = 39.59$, $p < .01$; and, crucially, a significant interaction, $F(1,60) = 13.07$, $p < .01$, confirming the already apparent fact that the two types of possessums are behaving quite differently. More specifically, animate and inanimate nouns were found to be associated with significantly different error rates for Romance speakers, $F(1,41) = 56.53$, $p < .01$.

Table 5. *Results of ANOVAs comparing error rates of Romance and Dutch speakers according to the main manipulations.*

	$F(1,60)$	p
Omnibus		
Language	15.99	<.01
Animacy	11.46	<.01
Language × Animacy	12.65	<.01
Animate Nouns		
Language	16.40	<.01
Congruency	14.54	<.01
Gender	0.002	.96
Language × Congruency	8.91	<.01
Language × Gender	0.18	.67
Congruency × Gender	6.50	<.05
Language × Congruency × Gender	1.90	.17
Inanimate Nouns		
Language	2.85	.10
Congruency	0.26	.61
Gender	0.03	.49
Language × Congruency	1.23	.27
Language × Gender	0.49	.49
Congruency × Gender	0.58	.45
Language × Congruency × Gender	2.11	.15

Subsequently, responses to stimuli with animate and inanimate nouns were analyzed separately in a 2 × 2 × 2 mixed design repeated measures ANOVA with language (Dutch vs. Romance) as a between subjects factor, pronoun gender (feminine vs. masculine possessor) and congruency (same vs. different possessum gender) as within subjects factors. Results are summarized in Table 5.

For ANIMATE NOUNS, there were significant main effects of language, $F(1,60) = 16.40$, $p < .01$, and congruency, $F(1,60) = 14.54$, $p < .01$, reflecting the fact that native speakers of one of the two Romance languages made more gender errors than native speakers of Dutch, but all speakers made more errors when the biological genders of the possessor and the possessum differed. On the other hand, there were no significant differences between error rates for masculine and feminine pronouns, $F(1,60) = 0.002$, $p = .96$, which would seem to indicate that there is no default gender for possessive pronouns. Nevertheless, since there is a significant interaction between congruency and pronoun gender, $F(1,60) = 6.50$, $p < .05$, giving statistical validity to the increase in feminine gender errors for the mismatched condition, the no-default gender conclusion should be reconsidered.

[2] Although it must be acknowledged that, given the number of variables and the sample size, it is possible that the lack of significant findings is due to insufficient statistical power, the results were replicated when the animate and inanimate nouns were analyzed separately, lending more credence to these null results.

https://doi.org/10.1017/S1366728910000325 Published online by Cambridge University Press

Looking at the data, it seems Romance speakers have a slight preference for producing masculine pronouns since more errors occur when both genders in the stimulus are feminine than when they are masculine, $t(41) = 2.17, p < .05$. However, for Romance speakers, the production of masculine pronouns seems to be more easily disrupted by the presence of another noun bearing feminine gender since the congruency effect is greater for masculine than for feminine pronouns, $F(1,41) = 9.38$, $p < .01$. This accords with some results in the literature showing masculine to be the default gender and feminine to be the marked gender (Antón-Méndez, Nicol & Garrett, 2002; Harris, 1991), although the effect has not been found consistently (Vigliocco & Franck, 1999).

Congruency also interacts with language, $F(1,60) = 8.91$, $p < .01$, with Dutch showing no consistent congruency effect as opposed to the very clear effect for Italian and Spanish. This is as expected from the fact that in Dutch, like in English, the gender of the possessum does not play any role in the NP. None of the other two-way or three-way interactions was significant (see Table 5).

For INANIMATE NOUNS, there was no significant difference between the two groups of speakers, $F(1,60) = 2.85$, $p = .10$. There were also no significant differences between masculine and feminine pronouns, $F(1,60) = 0.03$, $p = .86$; or in relation to congruency, $F(1,60) = 0.26$, $p = .61$, or any of the other two-way or three-way interactions (see Table 5).

For the sake of completeness, analyses of the other response types were also carried out. For the correct responses to stimuli with ANIMATE POSSESSUMS, there were significant differences between languages, $F(1,60) = 14.99$, $p < .01$, and a significant interaction of gender of the pronoun with congruency, $F(1,60) = 6.28$, $p < .05$, due to a higher effect of congruency on masculine pronouns. No other main effects or interactions were significant (all ps $> .05$). For INANIMATE POSSESSUMS, none of the variables was found to have an effect on the rate of correct responses (all ps $> .05$). These results mirror the results found for error rates. No significant main effects or interactions were found either for stimuli with ANIMATE OR INANIMATE POSSESSUMS on other types of error (all ps $> .05$). It seems thus that these other types of error were unaffected by the experimental manipulations.

General discussion

The main findings can be summarized as follows. First, Italian and Spanish speakers are equally susceptible to making gender errors in English despite superficial differences between these two languages' 3rd person possessive pronouns with respect to gender marking. Second, native speakers of these two Romance languages which, unlike English, have grammatical gender agreement between possessum and possessive pronoun make significantly more errors for animate possessums than native speakers of Dutch, which behaves like English for the purpose of possessive gender agreement, and this effect is modulated by congruency between the genders of possessor and possessum. Third, inanimate possessums did not appear to cause the same level of difficulty for Romance speakers given the apparent lack of a difference with Dutch speakers and the difference between Romance speakers' errors for NPs containing animate and inanimate possessums. Below, I discuss these results and their implications for models of bilingual language production in more detail.

Invisible gender features and the economical language processor

Given that Spanish, like Italian and other Romance languages, has grammatical gender agreement between nouns and determiners in general (*la*/*una*/*esta*/*otra*/*toda*/*vuestra niña* "the.FEM/a.FEM/this.FEM/other.FEM/every.FEM/our.FEM girl.FEM"), those few cases where this agreement is not apparent, as with the third person singular form of the possessive pronoun (see Table 1), are most likely a phonological accident, with the underlying gender agreement still applying between the noun and the determiner even in the absence of overt phonological manifestation. In principle, however, there is an alternative theoretical possibility: A linguistic processor intent on efficiency could have capitalized on the absence of overt marking to avoid enforcing the agreement relation in those cases where it is not needed. Insofar as we can conclude from the similarity between Italian and Spanish speakers shown in this experiment that Spanish speakers are also imposing the possessum's gender on the possessive, it seems that the language production system, instead of economizing on specific cases, may carry out procedures such as syntactic agreement in all instances. On the other hand, it can be argued that, for languages such as Spanish, it may simply be more efficient to automatically apply the regular agreement relations within the NP in every case than to check all cases in order to detect the very few where agreement is not going to be required. In contrast, for languages where the opposite is true – where only few instances of a particular syntactic agreement relation are phonologically realized – it may be more efficient to dispense with the agreement procedure unless it is found to be required. This is in fact what Tokowicz and MacWhinney's (2005) results imply. They found that native English speakers ignore the number agreement relation between noun and determiner when processing L2 Spanish sentences. This form of agreement does exist in English, but it is only overtly marked for the demonstratives *these*/*those*. Thus, apparently, English speakers do not implement the agreement by default in

https://doi.org/10.1017/S1366728910000325 Published online by Cambridge University Press

their native language – a practice they transfer to their second language.

In sum, both Romance languages behaved similarly despite the apparent differences in the production of the third singular possessive pronouns in the two languages. The absence of overt gender marking in Spanish 3rd person singular possessives does not reflect a lack of underlying gender agreement between possessum and possessive within the NP. Rather, it looks like Spanish speakers carry on the same processes of syntactic gender agreement between possessum and possessive pronoun as Italian speakers, for which the 3rd person singular possessive pronoun is morphophonologically marked for gender.

Semantic vs. grammatical gender agreement

The fact that Romance speakers produced significantly more possessive gender errors than Dutch speakers for animate nouns is consistent with the hypothesis that the former, but not the latter, face a sort of processing conflict because their L1s syntactic agreement system differs from their L2 semantic agreement system for specifying the gender of the possessive pronoun. Thus, even for rather proficient speakers, the native language procedures can spill over to the processing of the second language when the two languages differ, proof that some L1 automatic processes may be very difficult to inhibit completely when producing the L2, resulting sometimes in competition between the two alternative procedures in real time (Tokowicz & MacWhinney, 2005; Truscott & Sharwood Smith, 2004) and sometimes in the inappropriate implementation of a certain procedure in a language that does not require it. Alternatively, if the reason for the gender errors had been some kind of interference stemming from the mismatched genders of the nouns' referents at the conceptual level, all speaker groups would have been expected to show the same level of erroneous production. In fact, the error rates for Dutch speakers in English resembles that of native speakers in a study conducted by Slevc et al. (2007), where English speakers were found to make around 5% errors in the gender of possessive pronouns in sentences such as *Victor carried a package to his grandmother*. The authors considered these errors to be due to interference at the conceptual level given that nouns with intrinsic gender (e.g., *sister*) and nouns without intrinsic gender (e.g., *cousin*) resulted in equivalent error rates. Italian and Spanish speakers, however, produced errors over and above that baseline of Dutch speakers' errors which could potentially be attributed to such conceptual interference. Therefore, these errors must be due to interference at the syntactic level and, since this second source of interference is language dependent, it must be related to the implementation of the L1 syntactic agreement procedure.

On the other hand, inanimate nouns did not result in substantial error rates for Romance speakers as attested both by the lack of a significant difference with Dutch speakers in the analyses of inanimate possessums and the significantly lower Romance speakers' error rates for inanimate nouns as compared with their error rates for animate nouns. At first sight, there is something paradoxical about this: If possessive gender errors result from the transfer of an L1 procedure at the syntactic processing level, why do they not surface in the case of inanimate possessums? Differences in animacy itself are unlikely to explain this result since both animate and inanimate nouns are processed in the same manner in the Romance L1s with respect to gender agreement: It is the noun's syntactic gender, not its semantic gender, that has to agree with the possessive pronoun (e.g., a male victim would be *sua vitima* "her/his.FEM victim"). The possessive gender errors with L2 animate nouns could be thought to be the result of the Romance speaker failing to implement semantic gender agreement (i.e., coding the pronoun with respect to the possessor's gender) in accordance with L2 requirements and instead applying the syntactic gender agreement in accordance with L1 requirements. In this case, inanimate nouns would be expected to induce a similar pattern of errors to animate nouns if their L1 syntactic gender features are accessible at the point of implementation of syntactic agreement, or to induce increased number of errors in both matched and mismatched conditions if the L1 syntactic gender is not available. Instead, inanimate nouns appear to cause hardly any trouble to these L2 speakers. It appears, therefore, that the implementation of the L1 syntactic gender agreement process is somehow dependent on the source of the agreement, the possessum, having a gender feature that can be transferred to the pronoun. This is always the case in Italian and Spanish, but it is only the case for certain animate nouns in English (e.g., *actor*/*actress*, *father*/*mother*, etc.; but not *lawyer*, *child*, etc.). In the case of animate possessums, the interference would thus be caused by the lexical gender feature of the English nouns themselves during the preparation of the NP. This account is also compatible with data from another experiment by the author, where Spanish–English speakers had to produce sentences in their L2. In this experiment, some of the filler sentences required participants to produce a third person singular possessive pronoun accompanying an animate possessum which was an epicene, i.e., an invariable noun that can be used to refer to both genders. For these sentences, participants made fewer gender errors (3.6%) than for sentences in the same experiment where the animate possessum bore lexical gender (12.1%), as expected if the interference was mostly triggered by the inherent lexical gender of the possessum.

While inanimate nouns' lack of gender features in English may prevent the implementation of syntactic agreement between possessum and possessive pronoun, it does not completely explain the low error rate for the inanimate conditions. This result means not only that no L1 syntactic agreement was implemented, but also that the L2 semantic agreement unfailingly (more or less) was. If we think of this in terms of competition between two processes, it is not too surprising that preventing one of them from applying would mean the other one can be applied without problem. But the two types of agreement we are considering here are implemented at different processing levels and the conflict they present to the speaker does not have to be due to direct competition between an L1 and an L2 procedure (the only possible competition would be between an L2 procedure and no L1 procedure at the semantic level, and between an L1 procedure and no L2 procedure at the syntactic level). If so, the low error rates in NPs with inanimate possessums are somewhat surprising when we consider that possessives in Spanish or Italian do not appear to require the encoding of the possessor's gender, i.e., semantic gender agreement, at any time. In view of this, one would have expected to find these speakers making gender errors also in the inanimate conditions when mistakenly forgetting to implement semantic agreement and encode the possessor's gender on the possessive pronoun. Such errors, due to insufficient automatization of L2 procedures, should have surfaced in the absence of any further influence from syntactic agreement as more or less random gender errors, as was observed for nominative pronouns in Spanish–English L2 speakers (Antón-Méndez, 2010). One possible way to interpret this lack of errors with inanimate possessums is to assume that the L2 process of semantic gender agreement with the anaphoric antecedent has been fully automatized. This, together with the lack of a lexical gender feature for inanimate English nouns which could have exerted influence further down the line, would have resulted in virtually perfect performance. That is, proficient Italian and Spanish L1 speakers could have mastered the L2 process of semantic possessive gender agreement and only be susceptible to the presence of lexical gender during syntactic processing. A lexical gender feature such as is found in some animate nouns could then act as a catalyst to induce the L1 process of syntactic gender agreement which, in turn, would result in the application of a new gender feature overwriting the gender feature originally applied.

The problem with that account is that it is incompatible with findings of increased nominative pronoun gender errors for proficient Spanish–English speakers as compared with equally proficient French–English speakers (Antón-Méndez, 2010). The former speakers make a significant number of errors with the English nominative pronouns *he* and *she* despite the existence of equivalent pronouns in the L1. The errors appear to be due to native Spanish speakers neglecting to process the gender feature of the pronominal antecedent because it is usually not required in Spanish, which is a *pro*-drop language (e.g., *va a la playa* is literally "goes to the beach"). When Spanish speakers fail to encode the gender of the antecedent during L2 English production, the result is a nominative pronoun gender error (e.g., *he goes to the beach* when talking about a female). If the lack of possessive gender errors in phrases with inanimate possessums is due to speakers having mastered the L2 procedure of reckoning with the gender of the antecedent in order to encode the genitive pronoun, why should they not have achieved the same level of competence for nominative pronoun production? After all, nominative pronouns are even more frequent and afford even more practice with the new L2 pronoun agreement or pronoun coding procedure.

An alternative explanation for the lack of errors observed with possessive pronouns accompanying inanimate nouns is that possessives in the Romance languages are also initially determined, like in English and Dutch, by the semantic/conceptual features of the antecedent, including gender. This is actually how the rest of the overt pronouns in Spanish and Italian must be processed (but apparently not the theoretical construct *pro* that substitutes for a pronoun in *pro*-drop sentences). In this scenario, a basic pronominal form containing the four defining features of person, gender, number and animacy will later be at the receiving end of the procedure in charge of implementing syntactic gender agreement between possessum and possessive pronoun, and will thus receive a grammatical gender feature from the possessum noun. This second gender feature could either come to overwrite the first one or, more likely, be appended to it. If we do not hear *suoa madre* (lit., "his.FEM mother.FEM"), it might simply be due to phonological syncopation. Although this account may at first appear somewhat speculative, it should be noted that this is precisely what happens with number. Possessive pronouns agree in number conceptually with their antecedents and syntactically with the nouns they accompany, with both numbers visible on the final form due to their different morphophonological realization: *mio*/*mios* ("Poss-1ST.SG.MASC.SG"/"Poss-1ST.SG.MASC.PL") vs. *nostro*/*nostros* ("Poss-1ST.PL.MASC.SG"/"Poss-1ST.PL.MASC.PL").

In this account, the lack of errors for stimuli with inanimate possessums is not due to the Romance–English speakers having fully automatized the L2 process of semantic gender agreement, but rather to their being able to make use in their L2 of a procedure they had already automatized for their L1, in combination with the fact that the inanimate English nouns do not provide the speaker

with a grammatical gender feature that can be used for syntactic gender agreement later on.

In sum, a possible interpretation for the differences between Romance speakers' pronoun errors with animate and inanimate possessums is that possessive pronouns are actually processed identically for the two types of nouns at the semantic level, and identically for speakers of Romance and Germanic languages. At this initial processing point, all possessive pronouns would agree with their anaphoric antecedents in gender (as well as the other features). The eventual difference in error rates would be the result of L1 transfer during later syntactic processing due to the inherent gender features of the animate nouns in the L2 triggering syntactic gender agreement between possessums and possessive pronouns. All this also means that syntactic processing must be carried out independently of lexical processes since, in this case, the lexical units are L2 items but the syntactic procedure they are associated with is an L1 procedure.

Extent of parallel activation of a bilingual's two languages

The lack of interference in the case of English inanimate possessums has one other implication: The equivalent L1 lexical items are not active at this point in sentence production. If they were, one would have expected Romance speakers to show increased gender error rates when the translation equivalents in their L1 languages carried conflicting gender features that could trigger syntactic agreement. In fact, effects of the gender of L1 nouns during L2 production has been documented for L2s that also have a gender system (Lemhöfer, Spalek & Schriefers, 2008, and references therein). However, Lemhöfer et al. claim that the effect is not so much due to the activation of the L1 gender as to the instability of the gender representations in L2 – a factor that is not relevant in the case of English as an L2.

Nevertheless, there is sufficient evidence to believe that a bilingual's two languages are activated in parallel during production and comprehension (Colomé, 2001; Dijkstra & van Heuven, 2002). It is therefore rather likely that the L1 equivalents of the English possessums present in these stimuli were active during the production of the L2 sentences even though the experiment was conducted exclusively in English. Even so, it is obvious that the non-target language alternatives must eventually be somehow inhibited so that only the right word is ultimately produced. The critical question is at what point exactly (Abutalebi & Green, 2008; Kroll et al., 2008). Evidence of parallel activation of alternative lexical or phrasal candidates that persists up to the phonological level can be found in monolingual production in the form of blends such as *it doesn't care* (*it doesn't matter* and *I don't care*), or *evoid* (*evade* and *avoid*) (Garrett, 1988), and in bilingual production in the form of cross-linguistic blends, such as *springling* (English *spring* and German *Frühling*) (Green, 1986, cited in Emmorey et al., 2008). These types of error, however, are not very common suggesting that parallel activation of alternatives to such an extent may be an exception. Emmorey et al. (2008), on the other hand, did find a preference for code blending (i.e., simultaneous production of translation equivalents in two languages) over code switches for the only bilinguals who could possibly show full parallel production in two languages: American Sign Language (ASL) and English bilinguals. This means that the two lexical alternatives in ASL and English were active to the point of both being output during sentence production. The authors claim that the absence of blends of this sort in other bilingual production is due to the physical impossibility of producing two spoken words simultaneously. In contrast, the data presented here seem to indicate that the Romance L1 alternatives are already out of the race by the time the English possessive NPs are built. It could be that the sheer physical impossibility of producing two spoken words at the same time results in bilinguals of two spoken languages having to choose between the two alternatives at an earlier point in time.

Models of bilingual language production

The results of this experiment, although not decisive in determining the superiority of any particular model of bilingual production, do provide evidence that can help constrain the properties and characteristics of a valid model of second language acquisition and bilingual production.

In principle, any model that allows the seamless integration of L1 and L2 syntactic procedures within a sentence as well as explaining the decreasing but persisting incorrect implementation of certain L1 procedures alongside the correct implementation of the L2 counterparts even when the L2 is not necessarily computationally more taxing are apt to explain the existence and distribution pattern of these possessive pronoun gender errors. Additionally, only models that consider a separation between syntax and lexicon would be able to account for the fact that L2 lexical items are subject to a syntactic procedure that should only apply in L1 on features of L1 lexical items.

Here, I would like to consider four different bilingual production models in more detail: De Bot's (1992) "Speaking" model adapted for multiple languages, the Declarative/Procedural model (Ullman, 2005), the Acquisition by Processing Theory (Truscott & Sharwood Smith, 2004), and Hartsuiker, Pickering and Veltkamp's proposal (2004). These models were devised within different frameworks, for different reasons, and to account for different kinds of evidence.

De Bot's (1992) model is probably the most abstract and theoretical of the four. In following Levelt's (1989) model of monolingual production, it assumes that there are discrete processing levels devoted to the processing of the different types of information that make up an utterance – conceptual, grammatical, and phonological. Of these, only the conceptualizer level, where utterances are planned, is hypothesized to be shared across languages. The other two processing components, the formulator (in charge of building up the syntactic frame and inserting the lexical items) and the articulator (in charge of finalizing the phonological and prosodic form of the sentence) are thought to be language-specific, and more so the more proficient the bilingual. Although de Bot's model is not described in enough detail to determine with certainty whether it can or cannot account for the pronominal gender errors reported here, it does appear less compatible with these results. Insofar as two language-specific formulators are proposed for sufficiently proficient bilinguals, it is difficult to see how these speakers could have attained a high enough level of proficiency to produce L2 NPs with inanimate possessums without any influence from the L1 but not L2 NPs with animate possessums. What is more, it is difficult to see how a sentence that is supposedly being processed by the L2 formulator is transiently susceptible to a specific process of syntactic agreement that should be restricted to implementation within the L1 formulator.

Truscott and Sharwood Smith's (2004) approach, Acquisition by Processing Theory (APT), is primarily an account of second language acquisition that dispenses with special acquisition mechanisms by postulating that language development is the consequence of processing itself. The model of production on which the theory rests is modular in nature (resembling Levelt's, 1989, model in general characteristics) and assumes that the system is shared between languages. In this model, individual lexical items are responsible for the implementation of specific syntactic procedures which leads the authors to conclude that the apparent L1 influence on L2 is likely the result of parallel activation of the L1 lexicon during L2 processing. The appeal of this model in relation to the present data is its reinterpretation of L1 transfer as the result of competition between alternative procedures for access to the processing chain. In Truscott and Sharwood Smith's (ibid.) view, the intermittent appearance of a syntactic structure from L1 when producing L2 is just an indication of the fact that the L1 structure has at that particular point won the battle by virtue of its higher ease of production, an account that aligns with the perspective adopted here based on procedural automatization. The authors reject the usual interpretation of transfer as a consequence of wrong L2 parameter settings in favor of the existence side by side of the correct forms for each language with the possibility of the inappropriate one being sometimes applied. Transfer is "in effect, chronic involuntary code-switching" (Truscott & Sharwood Smith, 2004, p. 14). Their theory naturally explains why some of these cases of supposed transfer persist despite high proficiency, while it also predicts an effect of proficiency since, the more proficient the speaker, the more the L2 alternative would have been used and the less susceptible it would be to competition from the L1 structure. Furthermore, they also claim that the activation of L1structures is due to the presence of features that are compatible with the use of that particular structure. Most of this is compatible with the results reported here: Even proficient L2 speakers are sometimes liable to implement L1-based syntactic gender agreement, but the rate at which they do so decreases as proficiency increases and, furthermore, all speakers wrongly implement the L1 procedure only when the L2 presents features that justify it, i.e. when the L2 noun in the NP has associated lexical gender. However, the model's reliance on activation of L1 lexical items to explain syntactic transfer fails to account for the different behaviour of animate and inanimate nouns in eliciting possessive gender errors in L2 English, since both types of nouns carry a gender feature in L1 Italian or Spanish. If gender errors in L2 English are the result of L1 nouns being active, both types of nouns should have behaved similarly.

Ullman's Declarative/Procedural model (2004, 2005) represents an attempt to ground accounts of language processing and acquisition on the brain's architecture and physiology. The basic assumption is that knowledge linked to lexical items is stored in long-term declarative memory, while the processes that link these lexical items in grammatically acceptable ways are part of procedural memory. This is in the case of the L1 (Ullman, 2004). L2s are different because, while the declarative system remains fully operational after puberty leaving lexical accretion to be unaffected by age, there is evidence that the procedural system is subject to critical period effects which would affect the procedural learning of grammatical rules. This results in L2 learners relying on declarative memory to carry out functions that, in L1s, depend on procedural memory. With increased proficiency, though, some of these grammatical processes should become proceduralized and be performed in a similar manner as for L1 speakers.

The model is not described at a fine enough level of detail, linguistically speaking, to enable us to hypothesize how possessive pronoun gender errors could have arisen. However, in positing a different processing system for L2 than for L1, it makes the observed integration of the two languages' syntactic procedures difficult to explain. Nevertheless, since Ullman allows for a progressive "proceduralization" of L2 syntactic processing with increasing proficiency, it may be that, given the proficiency level of the participants in this

experiment, these L2 speakers are in essence using the same system and mechanisms to implement both L1-specific and L2-specific procedures which would allow for interaction between the two as observed here.

Finally, Hartsuiker et al.'s (2004) adaptation of Pickering and Branigan's (1998) model proposes the existence of a common repository of syntactic procedures (which they refer to as combinatorial information) for all languages of a speaker which get called in, so to say, when the lexical items they are related to are activated. This explains why they find cross-linguistic syntactic priming since the node to build a certain syntactic structure, e.g., passive, would be linked to both the English and the Spanish verb form. This model, although satisfying the requirement that the syntax of the two languages be shared, fails to account for the difference between inanimate and animate possessum NPs due to the dependence of syntax on the lexicon. In this model, the implementation of syntactic gender agreement within the L2 NP appears to require the activation of the L1 Spanish equivalent of the L2 English noun to, in turn, activate the relevant syntactic procedure or frame which would then be applied when preparing the English sentence. If so, one would have expected inanimate possessums to have been as likely to activate the same syntactic frame via the L1 Spanish equivalents, which was clearly not the case.

It seems thus that none of the models discussed in this section can unambiguously and straightforwardly explain the data on L2 possessive pronouns' gender errors. The two properties identified as necessary for models of second language production given these data can be summarized as follows: (i) shared syntactic processing for different languages, and (ii) independence of syntax and lexicon. De Bot's and Ullman's models satisfy the second requirement but not the first (at least in the case of early L2 learners for Ullman's Declarative/Procedural model), and the APT of Truscott and Sharwood Smith, and Hartsuiker et al.'s model satisfy the first but not the second requirement. However, it may be relatively simple to make both the APT and Hartsuiker et al.'s models compatible with the present data with minor adjustments along the lines of common gender nodes to which nouns with lexical gender are connected (see, e.g., Jescheniak & Levelt, 1994) regardless of language, and which would trigger syntactic gender agreement when activated. Nevertheless, only Ullman's model could account for the data without modification under the very reasonable assumption that the proficiency level of these participants was high enough for their L2 grammatical knowledge to have already achieved proceduralization.

Conclusions

The results of this experiment provide support for the prevalent notion that even highly proficient second language speakers are susceptible to L1 transfer. The reason for the L1 influence on L2 seems to be related to automatic implementation of certain procedures in certain circumstances, and this has implications regarding the properties that an eventual successful model of bilingual production must satisfy.

The results have a number of additional implications. First, it appears that syntactic procedures such as syntactic agreement are implemented regardless of whether the final output is going to overtly display the results of these operations but, apparently, only if the overall balance is favorable in terms of most cases actually presenting phonological proof of the operation. When, on the other hand, it is only a small proportion of the possible instances that display overt phonological marks of the procedure, the language processor seems to opt for limiting its implementation (Tokowicz & MacWhinney, 2005).

Second, the current pattern of errors is most compatible with an account of possessive pronoun production in Spanish and Italian that mirrors that of English and Dutch, including semantic gender agreement with the anaphoric antecedent, despite the lack of phonological evidence in the two Romance languages. In fact, this point relates to the previous one in the sense that, since most pronominal forms (in the nominative and accusative cases, and even the plural forms of the genitive case) carry overt marks of gender agreement with the antecedent, it may be more efficient for all pronouns to be initially also coded from the full set of features even if gender is phonologically unrealized in some cases, such as the third person singular possessives.

References

Abutalebi, J., & Green, D. W. (2008). Control mechanisms in bilingual language production: Neural evidence from language switching studies. *Language and Cognitive Processes*, *23*, 557–582.

Antón-Méndez, I. (2010). Gender bender: Gender errors in L2 pronoun production. *Journal of Psycholinguistic Research*, *39*, 119–139.

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

Baayen, R. H., Piepenbrock, R., & Gulikers, L. (1995). The CELEX lexical database (CD-ROM). Philadelphia, PA: Philadelphia Linguistic Data Consortium, University of Pennsylvania.

Blom, E., Polišenská, P., & Weerman, F. (2008). Articles, adjectives and age of onset: The acquisition of Dutch grammatical gender. *Second Language Research*, *24*, 297–331.

Collins, L., Trofimovich, P., White, J., Cardoso, W., & Horst, M. (2009). Some input on he easy/difficult grammar question: An empirical study. *The Modern Language Journal*, *93*, 336–353.

Colomé, A. (2001). Lexical activation in bilinguals' speech production: Language-specific or language independent? *Journal of Memory and Language*, *45*, 721–736.

Corbett, G. G. (2006). *Agreement*. New York: Cambridge University Press.

De Bot, K. (1992). A bilingual production model: Levelt's 'speaking' model adapted. *Applied Linguistics*, *13*, 1–24.

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

Emmorey, K., Borinstein, H. G., Thompson, R., & Gollan, T. (2008). Bimodal bilingualism. *Bilingualism: Language and Cognition*, *11*, 43–61.

Epstein, S. D., Flynn, S., & Martohardjono, G. (1996). Second language acquisition: Theoretical and experimental issues in contemporary research. *Behavioral and Brain Sciences*, *19*, 677–785.

Garrett, M. F. (1988). Processes in language production. In F. J. Newmeyer (ed.), *Linguistics: The Cambridge survey* (vol. III), pp. 69–96. Cambridge: Cambridge University Press.

Green, D. W. (1986). Control, activation, and resource: A framework and a model for the control of speech in bilinguals. *Brain and Language*, *27*, 210–223.

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

Hartsuiker, R. J., Pickering, M. J., & Veltkamp, E. (2004). Is syntax separate or shared between languages? Cross-linguistic syntactic priming in Spanish–English bilinguals. *Psychological Science*, *15*, 409–414.

Jescheniak, J. D., & Levelt, W. J. M. (1994). Word frequency effects in speech production: Retrieval of syntactic information and of phonological form. *Journal of Experimental Psychology: Learning, Memory, and Cognition*, *20*, 824–843.

Jia, G., & Fuse, A. (2007). Acquisition of English grammatical morphology by native Mandarin-speaking children and adolescents: Age-related differences. *Journal of Speech*, *Language*, *and Hearing Research*, *50*, 1280–1299.

Johnson, J. S., & Newport, E. L. (1989). Critical period effects in second language learning: The influence of maturational state on the acquisition of English as a second language. *Cognitive Psychology*, *21*, 60–99.

Kroll, J. F., Bobb, S. C., Misra, M., & Guo, T. (2008). Language selection in bilingual speech: Evidence for inhibitory processes. *Acta Psychologica*, *128*, 416–430.

Lemhöfer, K., Spalek, K., & Schriefers, H. (2008). Cross-language effects of grammatical gender in bilingual word recognition and production. *Journal of Memory and Language*, *59*, 312–330.

Levelt, W. J. M. (1989). *Speaking: From intention to articulation*. Cambridge, MA: MIT Press.

McLaughlin, B. (1987). *Theories of second-language learning*. London: Edward Arnold.

Pickering, M. J., & Branigan, H. P. (1998). The representation of verbs: Evidence from syntactic priming in language production. *Journal of Memory and Language*, *39*, 633–651.

Poulisse, N. (1999). *Slips of the tongue: Speech errors in first and second language production* (vol. 20). Amsterdam: John Benjamins.

Sabourin, L., & Stowe, L. (2008). Second language processing: When are first and second languages processed similarly? *Second Language Research*, *24* (3), 397–431.

Slevc, L. R., Wardlow Lane, L., & Ferreira, V. S. (2007). Pronoun production: Word or world knowledge? *MIT Working Papers in Linguistics*, *53*, 191–203.

Slobin, D. I. (1993). Adult language acquisition: A review from child language study. In C. Perdue (ed.), *Adult language acquisition: Cross-linguistic perspectives* (vol. II): *The results*, pp. 239–252. Cambridge: Cambridge University Press.

Tokowicz, N., & MacWhinney, B. (2005). Implicit and explicit measures of sensitivity to violations in second language grammar: An event-related potential investigation. *Studies in Second Language Acquisition*, *2*, 173–204.

Truscott, J., & Sharwood Smith, M. (2004). Acquisition by processing: A modular perspective on language development. *Bilingualism: Language and Cognition*, *7*, 1–20.

Ullman, M. T. (2004). Contributions of memory circuits to language: The declarative/procedural model. *Cognition*, *92*, 231–270.

Ullman, M. T. (2005). A cognitive neuroscience perspective on second language acquisition: The Declarative/Procedural Model. In C. Sanz (ed.), *Mind and context in adult second language acquisition*, pp. 141–178. Washington, DC: Georgetown University Press.

van Hout, R. (1996). Waar kwamen hun onze taal binnen? [Where did "them" enter our language?]. In R. van Hout & J. Kruijsen (ed.), *Taalvariatie* [Language variation], pp. 143–156. Dordrecht: Foris.

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

White, J., Muñoz, C., & Collins, L. (2007). The *his*/*her* challenge: Making progress in a 'regular' L2 programme. *Language Awareness*, *16*, 278–299.