

Gender Bender: Gender Errors in L2 Pronoun Production

Inés Antón-Méndez

Published online: 29 September 2009

© The Author(s) 2009. This article is published with open access at Springerlink.com

Abstract To address questions about information processing at the message level, pronoun errors of second language (L2) speakers of English were studied. Some L2 pronoun errors—*he/she* confusions by Spanish speakers of L2 English—could be due to differences in the informational requirements of the speakers' two languages, providing a window into the composition of the preverbal message that guides grammatical encoding during language production. To study this, Spanish and French speakers of L2 English were made to answer questions designed to elicit pronouns. Spanish speakers produced significantly more gender errors than any other type of pronoun error, and significantly more gender errors than French speakers. The results are as expected given that Spanish is a pro-drop language where many sentences would not have required gender to be encoded had they been uttered in the L1 instead of the L2. The implications for theories about the preverbal message are discussed.

Keywords Language production · Pre-verbal message · Second language · Pronouns

Introduction

In order to formulate a thought in linguistic terms, one has to first choose the information that is going to be included in the linguistic output. Apart from having to conform to general linguistic requirements, the information to include is obviously going to depend heavily on one's communicative intention, and, importantly, it will also have to be tailored to the target language's specific requirements. This early processing stage, the preparation of the preverbal message (Levelt 1989), has not been as well researched as other processes related to language production such as lexical retrieval, or building of syntactic structures, and it is only

I. Antón-Méndez

UiL OTS, Utrecht University, Utrecht, The Netherlands
e-mail: ines.anton-mendez@let.uu.nl

I. Antón-Méndez (✉)

BCSS, University of New England, Armidale, NSW, 2351, Australia
e-mail: iantonme@une.edu.au

lately that it is beginning to receive more attention (Brown-Schmidt and Tanenhaus 2006). The problem with doing research into “conceptualization” has been being able to determine what speakers intend to say. Until not long ago, the most researchers could do was see how conceptual properties of the input such as perceptual saliency would affect the choice of syntactic structure, inferring from that what conceptual features played a role in planning utterances and some of the temporal aspects of conceptual processing (Bock et al. 1992; Bock and Warren 1985; Ferreira 1994; Levelt and Maassen 1981; Sridhar 1988). More recently, eye tracking and the visual world paradigm have enabled researchers a measure of access to speakers’ communicative intentions (Arnold and Griffin 2007; Griffin and Bock 2000; Haskell 2005). Research employing this methodology has contributed to our understanding of the timing of sentence planning by showing that it is incremental (e.g., Bock et al. 2003), and that it gets updated in real time as required (Brown-Schmidt and Tanenhaus 2006). What are still missing, however, are details regarding the structure and composition of the preverbal message (henceforth PVM) itself—how the information is organized, and exactly which information is included. For example, it is not known whether the PVM consists of discrete units which more or less correspond to those at the grammatical level, or if the PVM packs information in excess of what would be needed for the final utterance (e.g., always including tense in the case of languages where tense is normally required irrespective of whether the particular utterance being planned is going to be inflected for tense or not), or only includes the specific information that would be actually needed to build the target sentence.

In this paper, I introduce a possible tool to address precisely these sorts of questions about the PVM: Analysis of L2 speech errors. The tool itself is not new, but it has never been used before for this particular purpose.

L2 speech errors have been used to study both questions related to L2 processing, such as L2 acquisition and bilingual speech production (e.g., Poullisse 1999), and more general psycholinguistic questions, such as the role of universal grammar after the critical period (Toribio et al. 1993). Also, analyses of spontaneous as well as experimentally elicited L1 speech errors have been a very valuable source of information about the stages in, and the mechanisms responsible for, building sentences (Bock and Levelt 1994), such as grammatical encoding (e.g., Bock and Miller 1991; Garrett 1988), phonological processing (e.g., Dell 1990; Levelt et al. 1999), and lexical access (e.g., Brown 1991). However, although there are also some categories of errors which can be unambiguously ascribed to the message level (e.g., blends, cognitive intrusions and environmental contaminants; Garrett 1988), they do not offer the kind of detailed evidence necessary for addressing questions about the fine-grained composition of the PVM.

Certain L2 errors, on the other hand, could provide the more specific information needed for this. The rationale is as follows: In cases where the linguistic expression of a certain message would require inclusion of a particular feature in one language but not in another, speakers of both languages may sometimes make errors due to this cross-linguistic discrepancy. Observed errors affecting the feature in question could be considered to have their origin at the conceptual processing level if alternative explanations can be theoretically or experimentally rejected, and could therefore help us understand processing at the conceptual level. The reasoning rests on the, as yet experimentally untested but theoretically well grounded, assumption that the PVM is language specific (De Bot 2000; Levelt 1989), implying that the communicative intentions of bilinguals can be given expression by preparing a PVM for language A when speaking in language A, or a PVM for language B when speaking in language B. Naturally, this will offer a chance for bilingual speakers to make an error when they mistakenly prepare a PVM for language A while intending it to serve as the basis for an utterance in language B. Evidence of interference of L1 in L2 at various levels has often been

reported (e.g., Poulisse and Bongaerts 1994; Rodriguez-Fornells et al. 2005), and even very proficient L2 speakers will be expected to occasionally make errors due to their preparing a PVM according to L1 requirements when the target language for the particular intended utterance is L2, in the same way that they sometimes produce unintentional code-switches for example (Broersma and De Bot 2006; Poulisse and Bongaerts 1994). In those cases, the discrepancy between what is needed and what is available at the point of encoding the sentence will surface as a semantic anomaly but, crucially, not as a grammatical error.

The question is how to ascertain that an anomalous utterance is anomalous because of mismatches between the information required to grammatically encode the sentence and the information provided by the PVM and not, for example, because of lemma misselection, or something else taking place during grammatical encoding itself. While it may be difficult to conclusively tease these alternative explanations apart, a good case can be mounted in favour of specific errors being due to misconstruction of the PVM if the following conditions are met:

1. the error is expected given a particular discrepancy in information requirements between two languages, therefore pointing to a likely origin at the message level;
2. the error is common to all proficient speakers of the same two languages, and not to other L2 speakers whose native language does not share the critical discrepancy, so that it cannot be attributed to inherent difficulties of the second language;
3. the error does not result in ungrammatical sentences, so that it cannot be attributed to inappropriate L1 grammatical transfer—i.e., the error must be an error in the sense of being a semantic anomaly.

One potential candidate put forward in this paper is an error involving the 3rd person singular nominative pronouns *he* and *she* which seems to be characteristic of, at least, Spanish speakers of English as a second language. As some readers will probably already have noticed, even very proficient Spanish-English L2 speakers seem to confound the two pronouns, calling husbands *she* and mothers *he*—often without any awareness of having said anything wrong. It is a particularly odd phenomenon because Spanish has two equivalent pronouns (*él* and *ella*), and it is, furthermore, a language in which natural gender figures prominently (most nouns referring to animate entities in Spanish are composed of a root and a suffix denoting the biological gender of the referent, as in *panadero/panadera* [baker-masc/baker-fem]). Therefore, we should look for the etiology of the problem neither in inherent difficulties of the English pronoun system for L1 Spanish speakers, nor in poor knowledge of the L2, nor in lack of familiarity with the distinction—all of them, for the rest, likely reasons for L2 errors in general. Neither can the error be strictly construed as a syntactic error, since the resulting sentence is grammatically correct, albeit conveying the wrong meaning.

A pertinent difference between the two languages, on the other hand, is the pro-drop status of Spanish but not English. In Spanish, nominative personal pronouns are omitted unless they are stressed or required for contrastive purposes (Luján 1997). This means that many of the sentences containing these pronouns in English would have contained no pronoun at all if uttered in Spanish. What is also important to bear in mind is that, in terms of specific information, what would be missing in the Spanish pro-drop equivalent of a sentence containing a nominative pronoun in English is the gender of the referent (assuming there are no other sentence components, such as adjectives, which refer to the subject and are gender marked in Spanish), while the other two defining features, person and number, are needed in Spanish in order to correctly inflect the sentence's finite verb. Compare:

1. Sacó la espada [Verb-past.3rd.sg. + direct object]
2. He drew the sword [Pron-nom.3rd.sg.masc. + verb-past + direct object]

This means that, if the L2 speaker of English had prepared a PVM for the wrong language, the information available would allow him to retrieve at least two pronouns in English: *she* and *he* (possibly *it* as well, although the neuter pronoun differs from the other two also with respect to animacy—a feature without an explicit role in the Spanish pronoun system), but the specific information about natural gender necessary to choose between the two would be absent. There is another feature to take into account when producing a pronoun—case. Case, however, is not necessarily a conceptual feature defining the pronoun at a message level, but rather a syntactic feature—the result of the pronoun being assigned a thematic role during grammatical encoding (Bock and Levelt 1994; Bock et al. 1992). The fact that pronoun errors do not affect the case feature could be related to the fact that pronoun errors do not seem to manifest themselves as omissions in L2 either: If the error is occurring, as hypothesized, at the message level, grammatical processing would proceed for the rest unimpeded and should result in a grammatically correct output. Since English is not a pro-drop language and requires an overt pronoun to occupy the subject position, one would be inserted, even if not the correct one, and it should take the case which corresponds to its syntactic function.

Having established an a priori theoretical reason why we could suspect the pronoun gender errors of Spanish-English speakers to originate at the message level (condition i in p. 6), it is important to also make sure that they are not the result of difficulties inherent to the second language which are shared by some or most L2 speakers, or to L1 grammatical transfer (conditions ii and iii above). For this, a set of what could be called semi-spontaneous English utterances was experimentally collected from two groups of L2 speakers: French and Spanish native speakers. Spoken utterances containing nominative pronouns of all types were elicited by means of leading questions (presented auditorily) about images (presented visually), although speakers were for the rest free to answer as they pleased. The data were then analysed in order to, first of all, confirm the anecdotal observation of the occurrence of he-she errors for Spanish speakers of English as L2; and also to contrast Spanish speakers' pronoun errors with those of another group of L2 speakers comparable in most aspects to the experimental group of Spanish speakers, but whose native language does not possess the critical pro-drop parameter, as well as to confirm that pronoun errors are not associated with ungrammatical pronoun omission.

Experiment

Method

Participants

Above 20 native Spanish speakers and 20 native French speakers participated in this study in exchange for a payment of 7.50 euros. The groups' characteristics are summarized in Table 1. They were all competent in English, with a mean proficiency level of, respectively, 3.5 and 3.6 out of 5, as measured by a commercially available English test: the Quick Placement Test.¹ Most of them used English on a daily basis in their studies or their jobs. Most participants

¹ This 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 corresponds to the following Council of Europe's descriptions: beginner, elementary, lower intermediate, upper intermediate, lower advanced and upper advanced. More information available at <http://www.oup.com/elt/catalogue/isbn/7162?cc=nl>

Table 1 General characteristics of the two groups of participants

	French	Spanish
Mean age	25.2 (SD 5.6)	28.9 (SD 5.1)
Education	20% Attending university 80% Bachelor degree or higher	20% Attending university 80% Bachelor degree or higher
Gender	50% Males 50% Females	50% Males 50% Females
Age English acquisition ^a	17.9 (SD 3.2)	18.5 (SD 3.1)
Age 1st non-prodrop L2 acquisition ^b	17.4 (SD 3.1)	17.9 (SD 3.5)
Age 1st L2 acquisition ^c	16.9 (SD 3.0)	17.8 (SD 3.5)
Frequency of English usage	90% Daily 10% Weekly	85% Daily 5% Weekly 10% Occasionally
English proficiency ^d	3.6 (SD 1.1)	3.5 (SD 0.8)

^a Age at which participant felt competent in English, even if s/he had started learning the language earlier in a classroom setting

^b Age at which the first non-native non-pro-drop language (English, German or French in the present sample) had been acquired

^c Age at which the first non-native language, irrespective of pro-drop parameter, had been acquired

^d As measured by the Quick Placement Test[®], on a scale of 1–5

had been monolingual at least until the age of 12, with the exception of two French–Dutch, four Spanish–Catalan, and one Spanish–Galician bilinguals. In all these cases, the bilinguals' two languages were either both pro-drop, or both non-pro-drop, therefore not affecting the contrast of interest. For these early bilinguals, it was also the case that the first non-native language had been acquired not earlier than age 12. The groups were very similar in their overall composition. Crucially, there were no significant differences in English proficiency scores as measured by the Quick Placement Test, $F(1, 19) = 0.74$, $p > .05$, or in mean ages of English acquisition, $F(1, 19) = 0.55$, $p > .05$, first non-pro-drop non-native language acquisition, $F(1, 19) = 0.60$, $p > .05$, or first non-native language acquisition, $F(1, 19) = 0.39$, $p > .05$.

Additionally, a smaller group of ten native speakers of English participated in the experiment. Nine were monolingual speakers, one had learnt Thai, Indonesian, Tetun, and Tatuá at age 24. Mean age was 33. Their results will only be brought to bear when necessary. The analyses and discussion are focused on the L2 speakers only.

Materials

Materials consisted of 43 visually presented illustrations by Norman Rockwell which were accompanied by a series of questions related to the depicted scene (from 2 to 12 per illustration, for a total of 241 questions). The questions were presented auditorily. See Table 2 for an example (the full list of questions is available from the author upon request).

Questions were spoken by two English native speakers, a female and a male, and recorded digitally with a Sennheiser microphone in a sound-proof booth. In the final experimental set up, half the illustrations' questions were spoken by a female, and the other half by a male. The voice-illustration pairing was sometimes random, and sometimes based on appropriateness of gender according to the questions asked and the illustration they referred to.

Table 2 Illustration and accompanying set of questions

Visual presentation	Auditory presentation
Illustration “Marbles Champion” can be found at http://nostalgia.tistory.com/search/marbles%20champion	What are the children doing? Who is winning? Is the girl indifferent? What is that white bag beside the girl? Is there another bag? Are the two boys happy? Did you and your friends like to play marbles when you were kids? Did you and your friends like any other competitive games? Was there something you were particularly good at when you were a kid? Did you like to show off your talents?

Experimental questions were designed to prime the inclusion of particular pronouns in the answer. Each pronoun was targeted on between 20 and 28 questions (except the expletive *it* which was targeted in only nine of the questions), but participants were actually free to respond in any way they pleased.

Apart from the experimental questions designed to elicit pronouns, there were also 43 filler questions for which no pronoun was targeted.

Procedure

Participants first received instructions emphasizing that their answers should consist of full sentences, that is, their answers should contain a verb. Nothing was mentioned about the use of pronouns. They first answered a series of practice questions while the experimenter was still present in order to correct any misunderstandings about the task and to make sure that participants got used to answering in full sentences. For the remainder of the experiment, participants were left alone. Pictures were presented in a randomized order for each participant, while the order of questions accompanying each picture was fixed. Participants had as much time as they needed to answer the questions, having to press the space bar in order to hear the next question whenever they were ready for it. Their answers were digitally recorded on digital audiotapes (DAT) with the same sound system used for recording questions.

Having done the experiment, L2 speakers of English had to complete a short English proficiency test—the Quick Placement Test (see Footnote 1).

Scoring

Spanish and French speaking participants' responses were transcribed twice independently by two different transcribers. Each transcription was subsequently and independently coded by one of two coders (the author being one of them).

Native speakers' responses were transcribed only once by one of the transcribers and coded by the author.

The coding proceeded in the following manner. For each answer, the instances of each of the nominative pronouns uttered was counted, and any error noted down. Uttered pronouns were coded as errors when they differed from what the coder considered to be the speaker's intended pronoun (e.g., in answering the question "Is the girl indifferent?", if the participant had answered "No, we are not", the coder would have counted the pronoun *we* as an error). Whether a pronoun was an error or not was mostly evident from the question and the picture to which the answers referred. There were, however, a number of cases in which participants seemed to take the wrong perspective when replying—answering as if they were part of the scene when not asked to put themselves in such a position, or answering as if the questioner was one of the characters in the illustration when this had not been indicated. This might seem strange, but it must be remembered that, in order to elicit the full range of nominative pronouns, participants were indeed sometimes asked to imagine themselves as one of the characters, and that questioners did sometimes frame the questions as if they themselves were one of the characters in the illustration (e.g., "Imagine you are the tall boy and I am the short one. What is happening?"). This confusion gave rise to switches affecting the pronoun person (e.g., *I* instead of *you* or *he*). For coding purposes, these pronouns were considered correct when: (a) another pronoun in the same sentence indicated consistency in the alternative perspective (e.g., "They are laughing at us because we are arguing" for the question "What are the other children doing and why?" when the questioner had earlier indicated that the girl in the picture was herself and the answer should thus have been "They are laughing at you because you are arguing"), or (b) the same "erroneous" pronoun was persistently employed across a whole series of questions for a particular illustration (e.g., "I don't pay attention at all to the young man", "I'm very concentrated about what the guy is saying, and I look at him as if he was like kind of Christ", "I'm going to dance with the older guy", as replies for a series of questions accompanying an illustration where the questioner had indicated that she was the depicted young girl). To make sure these person errors were indeed due to the way the experiment was set up, and not to problems with the L2, the performance of the native speakers of English was contrasted with that of the non-native speakers. Native speakers had made 10 of these detectable perspective switches (0.45% of all 2234 pronouns produced); L2 speakers made 78 detectable perspective switches (0.68% of all 11559 pronouns produced). Native speakers also made an additional 7 person errors (0.31% of all pronouns produced). Given that native speakers are unlikely to make so many errors due to problems with the person feature, this seems to imply that most of the remaining person errors of L2 speakers were probably also perspective switches, and not real pronoun errors. Nevertheless, they were analysed as person errors.

The different error types are: person errors (e.g., *I* instead of *you*), number errors (e.g., *I* instead of *we*), gender errors (e.g., *he* instead of *she*), animacy errors (e.g., *he* instead of *it*), omission errors (pronoun absent in a sentence with a finite verb), insertion errors (e.g., *the boy he played* instead of *the boy played*), and other errors (e.g., *it* instead of *there* in *there is*). Pronouns in the answer which had been present in the question were excluded from the final counts to avoid the possible bias introduced in those cases by the recently heard pronoun. This led to the exclusion of 503 data points (4.2% of the initial 12,062 pronouns produced)—264 for the Spanish group (out of a total of 6,317 pronouns produced initially), and 239 for the French group (out of a total of 5,745 pronouns produced initially); they affected the pronouns *I* (289), *you* (153), and *we* (61), which were the only pronouns appearing in questions. A chi-squared test comparing the deletion pattern in the two groups of speakers showed no significant differences, $X^2 = 5.23$, $df = 2$, $p > .05$. Most cases occurred for the 1st person singular pronoun (the 2nd person pronoun appeared in about twice as many questions since

the singular and plural cases had to be pooled together) due to the tendency to start answers with “I think”, “I believe”, etc.

The two coded transcripts of the L2 speakers were then contrasted, and the discrepancies classified as either transcription discrepancies or coding discrepancies.

Transcription discrepancies consisted of: general misses, which were differences in the number of pronouns (except when affecting *it's* contractions) written down by the two transcribers; misses re. *it's*, when only one of the two transcribers had detected the presence of the phoneme *t* in an *it's* contraction, while the other one had considered it missing; general mishearing, referring to cases in which the two transcribers had heard different things. Of the mishearing cases, those affecting the critical pronouns *he* and *she* were noted separately as mishearing re. *he/she*. All transcription discrepancies were checked directly at the source and resolved by the author. Those for which the author could not discern between the two interpretations and those affecting the critical pronouns *he* and *she* were further checked at the source by a fourth person.

Coding discrepancies were of the following types: disagreement about whether *you* was meant as singular or plural by the participant; disagreement about whether a particular instance of *it* was an expletive or not; disagreement about whether the use of the third person singular pronoun *he* was appropriate for dogs and babies; or about whether the third person pronoun *it* in certain types of copular clauses (as in “who is the man?”, “it’s the soldier’s father”) was appropriate when talking about persons; critical disagreement about a gender error; and other miscellaneous differences of interpretation. Due to the difficulty in elucidating when the 2nd person pronoun *you* was meant as singular or plural by the participant, and to the recurrent discrepancy between the two coders with respect to this feature as well as to whether instances of *it* were expletives or not, the two *you's* and the two *it's* were eventually collapsed. With respect to the other discrepancies, it was deemed that the use of *he* was correct when referring to dogs and babies, and that the use of *it* for animate referents was correct in copular clauses referred to as specificational or specifically-identifying (Declerk 1983; Mikkelsen 2005). The remaining coding discrepancies were reconsidered by the two coders together on a one by one basis. They were mainly due to oversights by one or the other coder and posed no great difficulty in reaching consensus.

Results and Discussion

Overview of Pronoun Production

French and Spanish speakers produced very similar proportions of each pronoun with no significant differences between the mean pronoun production for the two groups of speakers, $X^2 = 11.93$, $df = 6$, $p > .05$.

Of the total of 11,559 pronouns included, there were 11,134 correct pronouns (96.32%), 37 person errors (0.32%), 17 number errors (0.15%), 78 gender errors (0.67%), 15 animacy errors (0.13%), 256 omission errors (2.21%), 14 insertion errors (0.12%), and 7 errors of other types (0.07%). There were only 6 cases of pronouns showing a combination of two errors—2 animacy-insertion errors, 1 person-omission error, 2 person-number errors, and 1 person-animacy error. These combination errors were counted as two errors and included in each of the relevant categories. Tables 3 and 4 offer an overview of the distribution of the different pronoun errors produced by speakers of French and Spanish respectively.

For contrast purposes, the performance of native speakers of English is given in Table 5.

Table 3 Pronoun distribution (% of correct and different types of pronoun errors) for native speakers of French

Pron.	Correct	Error						
		Person	Number	Gender	Animacy	Omission	Insertion	Other
I	99.02	0.15	0.08			0.76	0.00	0.00
You	99.39	0.41				0.20	0.00	0.00
He	98.04	0.00	0.14	0.42	0.28	0.98	0.14	0.00
She	98.55	0.26	0.13	0.92	0.00	0.00	0.13	0.00
It	95.24	0.00	0.55		0.55	3.54	0.11	0.00
We	99.10	0.30	0.00			0.60	0.00	0.00
They	96.71	2.06	0.62			0.62	0.00	0.00
Total	98.07	0.35	0.20	0.18	0.13	1.02	0.05	0.00

Table 4 Pronoun distribution (% of correct and different types of pronoun errors) for native speakers of Spanish

Pron.	Correct	Error						
		Person	Number	Gender	Animacy	Omission	Insertion	Other
I	99.35	0.33	0.00			0.33	0.00	0.00
You	99.31	0.59				0.10	0.00	0.00
He	93.02	0.00	0.00	5.68	0.13	1.16	0.00	0.00
She	96.03	0.37	0.12	2.98	0.12	0.37	0.00	0.00
It	78.33	0.00	0.21		0.64	19.10	1.18	0.54
We	99.19	0.54	0.27			0.00	0.00	0.00
They	98.21	0.33	0.33			0.65	0.00	0.49
Total	94.73	0.30	0.10	1.12	0.13	3.30	0.18	0.13

Table 5 Pronoun distribution (% of correct and different types of pronoun errors) for native speakers of English

Pron.	Correct	Error						
		Person	Number	Gender	Animacy	Omission	Insertion	Other
I	97.54	0.00	0.00			2.46	0.00	0.00
You	98.77	0.99				0.25	0.00	0.00
He	99.64	0.00	0.00	0.36	0.00	0.00	0.00	0.00
She	99.64	0.36	0.00	0.00	0.00	0.00	0.00	0.00
It	94.06	0.00	0.00		0.00	5.94	0.00	0.00
We	100.00	0.00	0.00			0.00	0.00	0.00
They	99.07	0.46	0.00			0.46	0.00	0.00
Total	97.94	0.27	0.00	0.04	0.00	1.75	0.00	0.00

Pinpointing the Locus of Errors for the Spanish Speakers

A look at the Spanish speakers' total number of errors for each pronoun (Table 4) shows that most errors occurred in the 3rd person singular pronouns. A repeated measures ANOVA on error proportions with the different pronouns as independent variables confirms that the differences are significant, $F(6, 14) = 4.85, p < .01$. It is obvious that the 3rd person singular neuter pronoun *it* on its own can be responsible for this outcome, but the results for that particular pronoun are not only unexpected, but are also somewhat unreliable and uninformative with respect to the aims of this paper. Of the 930 cases of *it*, 200 were errors—a striking 21.5% error rate for a group of proficient L2 speakers. The overwhelming majority of the errors were omissions: 19.1% (178 cases). Although the most straightforward interpretation for the increase in *it* omission errors would be that they are attributable to the fact that Spanish speakers usually omit nominative personal pronouns in Spanish and could, therefore, be transferring this parameter to their L2, the general distribution of omission errors belies this conclusion. Omission errors are not homogeneously distributed (see Tables 3 and 4). On the contrary, they cluster on the 3rd person neuter pronoun *it*—89% of the omissions made by the Spanish speakers involved the pronoun *it*. Since all personal pronouns tend to be omitted in Spanish, a parameter transfer from the first to the second language should result in widespread omission of pronouns, which is clearly not the case. This, in combination with the fact that the two transcribers had already demonstrated to have difficulties in deciding whether the Spanish speakers had uttered *it's* or *is* in many instances, makes an alternative explanation for the high rate of *it* omission errors much more likely—phonological difficulties with the pronunciation of the *it's* contraction by Spanish speakers together with possible detection difficulties during transcription of this same contraction. In fact, of the 178 *it* omissions, 146 were indeed in sentences where *it's* was expected instead of *is*—that is, 82% of the omissions were for 3rd person singular neuter pronouns which should have accompanied *is*. In contrast, of the correct instances of *it*, only approximately 66% preceded *is*. It is also worth mentioning that more than a quarter of the *it* omission errors (49 out of 178, 38 of which were in cases of *it's*) were made by just one participant with a marked foreign accent in his English who, furthermore, also spoke particularly fast. This phenomenon of syllable simplification by consonant deletion has already been reported in the literature on L2 production for Spanish speakers of German (Tropf 1987), and for Portuguese (a language with similar phonotactic constraints in this sense as Spanish) speakers of English (Tarone 1987). In sum, the cause of the increase in omission errors for the pronoun *it* is more likely phonological than syntactic or semantic. As phonological errors are not the focus of this paper and are obscuring the overall results, the 3rd person singular neuter pronoun was excluded from further analyses.

The remaining amount of omission errors for Spanish speakers (0.44% of all pronouns produced except *it*) is similar to that of French speakers (0.52% of all pronouns, once the pronoun *it* is removed), and less than that of native English speakers (0.77% of all pronouns except *it*). Some omissions are indeed expected since they are acceptable in casual speech in English (Gerken 1991; Harvie 1998; Valian 1990).

An analysis of variance of the other six pronouns indicated that the differences were significant, $F(5, 15) = 3.40, p < .05$. To keep the number of tests at a minimum, three planned comparisons were performed. An ANOVA of error rates for all pronouns except the two 3rd person singular animate pronouns, showed them to be not significantly different, $F(3, 17) = 0.92, p > .05$. Neither were the two 3rd person singular animate pronouns different from each other, $F(1, 19) = 1.43, p > .05$, which, incidentally, suggests that there is no default pronoun—if there were, the non-default pronoun would have been overrepresented in, at least, the gender error category. Finally, comparing error rates for the two 3rd

Table 6 Error distribution according to error type for French and Spanish speakers

Pron.	Person	Number	Gender	Animacy	Omission	Insertion	Other
French	19 (0.41%)	6 (0.15%)	10 (0.68%)	2 (0.14%)	24 (0.52%)	2 (0.04%)	0 (0.000%)
Spanish	18 (0.43%)	4 (0.09%)	68 (4.30%)	2 (0.13%)	22 (0.43%)	0 (0.000%)	3 (0.06%)
Total	37 (0.384%)	10 (0.13%)	78 (2.55%)	4 (0.13%)	46 (0.47%)	2 (0.02%)	3 (0.03%)

Error frequency and percentage in brackets; percentages calculated with respect to the total amount of pronouns produced which are susceptible to the particular error

person animate pronouns versus the rest of the pronouns yielded a significant difference, $F(1, 19) = 10.66$, $p < .01$, due to an increase in errors for the two 3rd person singular animate pronouns and proving that they are particularly susceptible to error.

To find out whether it is indeed the gender errors that are responsible for the increase in error rates for the 3rd person singular animate pronouns, a repeated measures ANOVA was carried out on error rate with error type as independent variable. There was a marginally significant difference between the error types, $F(4, 16) = 3.01$, $p = .05$. Pairwise analyses using Bonferroni's correction showed the difference to be due to gender errors being different from all other types, while none of the other types of errors were different from each other.

In conclusion, Spanish speakers make more errors for *he* and *she* (respectively 7 and 4% of the total instances of each) than for the other pronouns (0.7% for *I*, 0.7% for *you*, 0.8% for *we*, and 1.8% for *they*), these errors being predominantly gender errors (see Table 6), confirming anecdotal observations and pointing to specific problems with the gender feature. They also make a considerable amount of errors for *it* due to incorrect omissions. It is important to note here that, whatever the cause of the errors involving the pronoun *it* (and they may need further analysis and a paper all of their own), errors made by Spanish speakers for the two 3rd person singular animate pronouns are not omission errors. This means that, at least for sentences containing either *he* or *she*, the majority of errors could not be the result of straight grammatical transfer, and they result in grammatically correct sentences which are anomalous only conceptually.

Comparing French and Spanish Speakers

To investigate whether Spanish speakers' error pattern is common to other L2 speakers of English, their performance was compared to that of a similar group of L2 speakers: French native speakers. Like Spanish, French is a Romance language, and, as a consequence of their common origin, the two languages share many traits, such as the pronoun system, the existence of grammatical gender, and the gender agreement system (La *grammaire pour tous* 1990). However, the two languages differ in that French, unlike Spanish, is not a pro-drop language.

Table 6 shows the frequencies (and percentages with respect to the number of pronouns susceptible to the error in question) of the different types of errors produced by the two L2 speaker groups collapsed across pronouns (excluding the 3rd person singular neuter pronoun).

A one-way multivariate analysis of variance was conducted to compare the different error rates for the two groups of speakers. Error rates were calculated over the total number of pronouns produced by the speaker which could have been subject to the error in question, since not all errors affect all the pronouns. The difference between the two languages was marginally significant, $F(7, 32) = 2.27$, $p = .05$. Follow-up tests using the Bonferroni

method, at the 0.01 confidence level, showed gender error rates to be significantly different for the two groups, $F(1, 38) = 9.93$, $p < .01$, while none of the other error types differed with language group, all p 's $> .05$.

In sum, Spanish speakers are significantly more likely than French speakers to make gender errors, but not other kinds of pronoun errors. Therefore Spanish speakers are making gender errors over and above those expected from putative general difficulties posed by the characteristics of the L2 English, implying that the increase in gender errors for Spanish speakers is most likely due to processing idiosyncrasies of this group.

Correlation with Proficiency

To check whether the rate of pronoun gender errors diminishes as speakers' proficiency increases as would be expected if speakers are able to adjust the automatic selection of information for linguistic expression, the correlation between error production and proficiency level was calculated. The Pearson correlation coefficient between proficiency and gender errors is not significant either for the population of Spanish speakers, $r = -0.250$, $p > .05$, or the population of French speakers, $r = -0.132$, $p > .05$. However, given the fact that the L2 speakers included were all fairly fluent and relatively homogeneous in their proficiency scores, it is possible that a negative correlation would be found for L2 learners in different stages of acquisition, but that once speakers have reached a certain level of proficiency the error rate plateaus.

General Discussion

To summarize the results: Spanish speakers of English as a second language make significantly more errors for the 3rd person singular animate pronouns *he* and *she* than for other pronouns (with the exception of *it* which is associated with a high rate of phonologically induced omission errors); the increase in errors for these two pronouns is due to confusion involving the gender feature. Furthermore, the pattern of pronoun errors for the Spanish L2 speakers differs from that of French native speakers of English as L2—a group comparable to the Spanish speakers in many aspects, but not with respect to the pro-drop status of their native language.

In relation to the three conditions stated in the introduction which would have to be met in order to be able to consider a production error as resulting from lapses affecting the information content of the PVM, we can conclude that:

1. pronoun gender errors are indeed particularly common for Spanish speakers of L2-English; furthermore, these errors are to be expected on the basis of a discrepancy in information requirements between Spanish and English, since Spanish equivalents of many English sentences would lack a pronoun and, therefore, would very often not require gender information about the covert subject of the sentence, while person and number information, on the other hand, would be required for verb inflection;
2. pronoun gender errors are not due to inherent difficulties of L2 usage, since they are characteristic of Spanish speakers of L2-English, but not of French speakers of L2-English, who would have been expected to show a similar pattern of pronoun errors if the problem lay in the second language rather than the native language;
3. errors affecting most pronouns are not due to inappropriate transfer of L1 grammar, since the Spanish L1 grammar allows the omission of all pronouns in subject position, but most

pronouns were not involved in errors of omission, and errors affecting most pronouns did not result in ungrammaticalities (as would have been the case for inappropriate L1 grammatical transfer).

In short, we know that pronoun gender errors are not due to lack of knowledge or to application of the wrong grammatical rules; and we also know that the pro-drop parameter of Spanish should allow speakers to dispense with the gender feature in many cases when preparing sentences. We can therefore hypothesize that a likely cause of pronoun gender errors is faulty processing at the conceptual level which affects the informational content of the PVM. Such interpretation has implications related to the way pronouns are produced, the properties and composition of the PVM, and the availability of feedback between levels of processing.

Before proceeding to discuss the implications in more detail, there are a couple of alternative explanations in principle compatible with the data which should be contemplated. One possibility would be that the error arises not during preparation of the PVM, but rather as a consequence of Spanish-English speakers fully processing the sentence in Spanish and then translating the end result into English at the last moment. The speaker would realize that the English sentence requires a pronoun but, if basing the translation on the final Spanish sentence, there would be no gender information to choose the right one, in which case, pronoun gender errors would be expected. Although this explanation is theoretically possible, it is rather unparsimonious and seems less likely than a defective PVM hypothesis for several reasons. Firstly, all the available evidence suggests that proficient second language speakers do not translate from their native languages (Kroll and Tokowicz 2005). Secondly, even if these speakers are only translating in a small proportion of the produced sentences, and those would be the ones subject to pronoun gender errors, the translation mechanism is unclear—the sentence should not go through a second conceptual processing because the speaker would then be able to contribute the missing feature since the speaker knows who the pronoun's referent is; but the translation could not be performed strictly word by word, since then the animacy, person and number features for the pronoun would also be missing and should be just as likely to be part of a pronoun error. Furthermore, if translation is being carried out, one would expect to find other errors such as code-switches given that the Spanish word forms would have been retrieved. However, the only case of code-switching found in sentences with a pronoun gender error concerns the insertion not of a Spanish word, but a Dutch word.

A second possibility is that the source of the pronoun gender feature is not conceptual but lexical—part of the lemma. This is certainly very likely to be the case for grammatical gender (Meyer and Bock 1999; Schmitt et al. 1999), and, in principle, it could also be so even if the pronominal gender feature is based on the natural gender of the referent. Consequently, if the gender feature is being derived from the lemma, pronoun gender errors would not be a reflection of processing at the conceptual level, but rather of grammatical encoding and lexical access. It could be that native Spanish speakers access the lemma of the antecedent but, when processing it for pronoun selection and being used to not requiring a gender feature in their native language, they would sometimes fail to include it when looking for the pronoun in English. This possibility is discussed in the next section.

Preparing Pronouns for Production

Pronoun production has received a fair amount of attention from psycholinguists. Pronouns have been studied with respect to agreement (e.g., Bock et al. 1999), to which conditions

influence choosing a pronoun over other form of referring expressions (Arnold and Griffin 2007; Sridhar 1988), and, crucially, to whether pronouns are retrieved with respect to their linguistic (discourse) antecedent or to their conceptual referent (Meyer and Bock 1999; Schmitt et al. 1999). Both Meyer and Bock (1999), and Schmitt et al. (1999) conclude on the basis of their results that pronoun production requires the activation of the antecedent at a syntactic level. This conclusion is at odds with the claim made here that pronoun gender errors are probably the result of speakers failing to include gender information in the PVM and then finding themselves one feature short for correct third person singular pronoun selection. If pronouns are produced from the antecedent's features, Spanish speakers would also have had to include gender when speaking their native language since such information would be needed for accessing the antecedent. Therefore, pronoun gender errors would have to be interpreted as resulting from a failure to process the antecedent's features properly in English, forgetting to take gender into account because in Spanish this feature is required less often.

One problem with this approach is that the putative mechanism for selecting pronouns' features on the basis of a lexical entry would only work for third person pronouns, but not for first or second person pronouns, for which there would be no lexical item from which to extract the pronoun's features. And even for third person pronouns, Corbett (1991, p. 246) has argued on the basis of cross-linguistic evidence on what he calls hybrid nouns (nouns that seem to take different genders for different agreement targets; e.g., *team* is grammatically singular but can take a plural personal pronoun, as in "the team is very good, but they always lose") that they are the targets most likely to take semantic agreement. Furthermore, there are other cases when pronouns' features are clearly conceptually derived: when the referent is not lexically gendered (e.g., "there was only one other *person*—*she* was wearing blue"), when the pronoun is used deictically, without antecedent; and even when the pronoun's meaning is arbitrary and no specific referent can be identified (e.g., "the fact that you're paranoid doesn't mean *they* are not after you"). This would mean that first and second person pronoun features are conceptually determined, and third person pronouns are sometimes conceptually and sometimes lexically determined. It not only seems unnecessarily complicated, but also begs the question of how the language processor would know a priori whether to base the coming pronoun's features on conceptual information or wait until the lexical entry has been accessed, making it much more plausible that conceptual features are used for nominative pronoun production whenever such features are available, i.e. for animate referents. Indeed, Corbett (1991, p. 244) argues that personal pronouns are likely to refer to what their antecedent refers to, i.e. to the concept denoted by the antecedent, rather than to the antecedent itself.

On the other hand, pronouns referring to inanimate objects could hardly be expected to depend on the conceptual features of the referent, since grammatical gender does not necessarily have a conceptual counterpart (but see Konishi 1993, for evidence that even grammatical gender might have semantic correlates of femininity and masculinity), and even number is unreliable—e.g., mass nouns and collectives such as *rice* and *fleet* are singular despite being composed of a multitude of individual objects, *pluralia tantum* such as *scissors* are plural despite being only one object. In fact, in accordance with the proposal that animates and inanimates are processed differently with respect to pronoun production, only collectives denoting a collection of humans (also called 'corporate nouns'), can be referred to by a plural personal pronoun in English, while collectives denoting a group of objects cannot (Corbett 2000, footnote 10, p.188).

Postulating a different mechanism for pronouns referring to animate or inanimate entities would also explain why Meyer and Bock (1999) and Schmitt et al. (1999) found evidence of pronouns being processed from their lexical antecedents since they studied pronoun

production with respect to inanimate nouns in Dutch and German. Given that, in both German and Dutch, grammatical gender is not reliably predicted by conceptual properties of the referent (for example, plate is neuter in Dutch but masculine in German), it is to be expected that its retrieval would depend on the activation of the grammatical properties of the referent which may only be possible by accessing the antecedent's lemma. In contrast, the critical pronouns in the present study, *he* and *she*, always have an animate referent for which all the features can be found at a pre-linguistic conceptual level.

Intriguingly, a different mechanism for producing animate and inanimate pronouns would naturally account for the reduced number of animacy errors in cases of pronoun gender errors in the current data set (see “A remaining question” below). If only pronouns referring to animate entities are processed from conceptual features, the distinction between *he* and *she*, on the one hand, and *it*, on the other, is already established from the beginning. If all third person pronouns are processed from lexical features, this result remains puzzling—in those cases in which the second language speaker has neglected to include the gender feature associated with the lemma of the antecedent, the information available to select a pronoun (third person singular, no gender) would be most compatible with the pronoun *it*, rather than with either one of the two gendered pronouns *he* and *she*.

It could maybe be argued that the present results differ from those of other researchers with respect to pronoun production because what we are dealing with here are not just pronouns in the usual sense of overt pronouns, but either features meant for the purposes of verb inflection rather than to encode a pronoun or, alternatively, features meant for the purposes of encoding the covert element *pro* (see “Correspondence between PVM and grammatical constituents” below). If the former, these results may not contribute anything to the research on pronoun production. If the latter, however, the production of *pro*, is likely to share production mechanisms with the overt pronouns since *pro* shares all the characteristics of overt pronouns except phonological realization (Haegeman 1991).

However unlikely, a possibility remains that Spanish speakers' pronoun gender errors are the result not of conceptual processing but of grammatical processing. In this case, the present results would indicate that, with respect to pronoun production, the antecedent is accessed but its features are processed differently according to the target language and, more interestingly, according to the specific pronominal form and its function in the final sentence—that is, a full 3rd person pronoun would require the gender specification, but *pro* would not need gender to satisfy its function as sentential subject when it just has to agree with the verb, while it might if the sentence contains a predicate adjective which needs to agree in gender with its antecedent. Furthermore, the existence of these errors could then be interpreted as proof of a step during pronoun production where the necessary features have been extracted from the antecedent but where the antecedent itself may not be further available.

While a lexical origin for the pronoun gender errors cannot be definitively ruled out at this point, a conceptual origin for pronoun gender errors stands as the most parsimonious and most likely explanation for the present data. It is quite probable that the reason for these gender errors of Spanish-English speakers is a PVM where the natural gender of the referent is not specified, which in turn would mean that Spanish speakers do not routinely include gender information when preparing for a sentence without an overt pronoun. As already mentioned, this has several implications. First and foremost, the finding would mean that the information included in the PVM is not only language specific as expected, but also sentence specific at a very fine grained level. The finding could also be revealing as to the correspondence between information in the PVM and specific syntactic constituents. Finally, the existence of this pronoun gender error seems to point to the unavailability of feedback mechanisms which, if present, would prevent the error from occurring by allowing the speaker, at the point during

grammatical encoding where the two pronouns are found to be equally compatible with the specifications, to return to the conceptual level and check on the conceptual referent in order to retrieve the disambiguating feature.

Each of these implications will be discussed in more detail in the following sections.

PVM is Sentence Specific and Language Specific

At first sight it may not seem very surprising that the PVM is sentence specific—after all it is postulated to reflect in propositional form the specific communicative intention of the speaker at a particular point (Levelt 1989, p.70), an intention whose output would eventually be a sentence (or some other kind of expression). Therefore, one PVM would be representing one sentence and, in that sense, would need to be sentence specific. Furthermore, the PVM is sentence specific because it is the blueprint from which the sentence will be constructed—thus, the resulting sentence will be moulded according to the information provided by the PVM underlying it. Nevertheless, the PVM could plausibly contain more information than that strictly necessary for the grammatical encoding of the intended sentence. For example, in English most present tense verbs need to be inflected for third person singular, but modal verbs do not. It could be that the PVM of a sentence for which eventually a modal verb is to be selected does nevertheless contain information about person—information that would later not be realized grammatically. This could be simply because this information is normally necessary for the construction of most finite present tense sentences, and it is more economical for the speaker to include it in every case than to determine whether it will not be necessary in a particular case. Once the verb is selected during grammatical encoding, it will be found whether it does or does not require third person inflection and, if it does not, the person information provided would not find realization in the final sentence. If this were so, PVMs would be sentence specific in broad terms, but they could contain, in cases like the above, information in excess of that required for the final formulation. Alternatively, it could be that, during conceptualization, this particular subset of verbs is treated differently (provided they share some distinctive conceptual feature) allowing the PVM to dispense with person information when one of them is going to be the main verb. Whether one or the other possibility is actually the case may depend on the type of information, on whether a conceptual distinction can be established at the level of conceptualization that would allow for different sentence projections at this level, on whether there is feed-back for grammatical constraints to be able to determine the contents of the PVM, and on whether it is more efficient for the language production system to generalize first and waste later, or to detail first and economize later.

With respect to referential expressions in Spanish, the present data are compatible with a PVM which includes exclusively the features that would be required in the final utterance. Thus, if a full noun phrase was going to be employed, the speaker would have supposedly included enough information in the PVM to retrieve the appropriate lexical item from the lexicon, including natural gender; if an overt pronoun would have been needed, the speaker would have included gender information in their PVM, not only when speaking in their L2 English, but also had they been speaking L1 Spanish and planning a sentence with a stressed overt pronoun—*él* or *ella* (*he* or *she*). But, on the other hand, for sentences without an overt pronoun, where gender would not be required in the final sentence, it seems that Spanish speakers do not necessarily include gender information in their PVMs.

The second claim is more straightforward—the PVM must indeed be language specific since Spanish speakers of English as L2 are significantly more susceptible to producing this sort of error than French speakers of English. In fact, the claim should not be particularly

controversial and, as already mentioned, Levelt (1989, p. 103) had already proposed this on theoretical grounds for monolinguals, as had De Bot (2000) for bilinguals. Slobin (1996) has also convincingly argued for language specific procedures to encode communicative intentions which, furthermore, he claims are difficult to acquire for adult learners (Slobin 1993). That PVMs were language specific had been assumed as a premise justifying the comparison between French and Spanish speakers. In this sense, the present data constitute only the first empirical corroboration of this claim.

Correspondence Between PVM and Grammatical Constituents

Although the discussion has so far focused on the missing feature (i.e., gender), there is the twin question of the provenance of the other two features: person and number.

Up to this point, we have mostly treated the missing pronoun in pro-drop contexts in Spanish as if it was completely absent from the sentence at a grammatical level. In principle, this is a plausible psycholinguistic hypothesis with some empirical support from Japanese (Yamashita et al. 2005). If the pronoun is indeed completely absent in the Spanish equivalent of the English sentences where the pronoun gender errors appear, the verbal inflections would have to be determined directly from information at the PVM level. This would mean that the subject-verb agreement in this case is not a syntactic procedure, but rather conceptually guided. Such a hypothesis has been put forward before (Vigliocco et al. 1996) for languages with relatively rich verbal morphology to explain what appeared to be cross-linguistic differences with respect to the role played by conceptual features on subject-verb number agreement. The conceptual influence on agreement (the so-called distributivity effect) was eventually found in other languages with poorer verbal inflectional morphology (Eberhard 1999). Nevertheless, although the possibility of having the verb directly inherit its agreement features from the conceptual representation exists (Vigliocco and Franck 1999), a fair amount of the evidence points to subject-verb agreement being a syntactic phenomenon (Bock et al. 2001; Franck et al. 2008).

If the verb inherited its agreement features from the conceptual representation, it would mean that features in the PVM are not specifically tagged for any one syntactic constituent. This is because in cases of pronoun gender errors, when the PVM is supposed to have been prepared with a Spanish sentence in mind, the person and number features would have been destined for the verb's inflection, but in the final utterance in English, these features would have been used instead to retrieve the pronoun.

However, in line with the evidence for agreement as a purely syntactic phenomenon, the likely source for the person and number features on the verb is not the conceptual features encoded in the PVM but a syntactic constituent in subject position: *pro*. *Pro* has a long history in the linguistic literature as a substitute for missing arguments which are otherwise deemed obligatory to fulfil syntactic requirements such as realization of the verb theta roles (Haegeman 1991). On theoretical grounds, *pro* has been amply justified (e.g., Jaeggli and Safir 1989). Although there seems to be no direct experimental evidence as to the psychological reality of *pro*, there is evidence for the existence of other silent elements both behavioural (e.g., Swinney et al. 1988), and physiological (Demestre et al. 1999; Featherston 2000), which makes it plausible that this one too is endowed with a covert existence in Spanish speakers' minds. Furthermore, Oshita (2004) has convincingly argued that production of certain L2 English constructions by L1 pro-drop speakers is proof of the psychological reality of, at least, the null expletive.

If so, the present data suggest that *pro* is not just the muted version of an overt pronoun, but a different constituent comprising only the features needed to carry out its syntactic

functions—in this case, subject-verb agreement. It will be interesting to see whether the composition of *pro* varies according to syntactic need—that is, whether in sentences where *pro* has to act as agreement source for a predicate adjective, for example, it will also carry a gender feature, since in Spanish adjectives agree in gender and number with their antecedents.

Unavailability of Feedback

The existence of the type of errors we claim pronoun gender errors to be—errors derived from a discrepancy between the information included in the conceptual blueprint of the sentence, the PVM, and the information ultimately required at a grammatical level—hints at the unavailability of conceptual information during later grammatical encoding. If the information of previous levels were available because of regular feedback loops during sentence production, at the point where two pronouns seem to be compatible with the initial specifications, the speaker would be able to check the gender of the referent at a conceptual level and select the correct one. However, this does not seem to be happening. Of course, we cannot be sure from these data that it does not happen part of the time but, if it only happens part of the time, it should be viewed more as an emergency repair mechanism than a regular fixture of the normal production process. In fact, the processing of the PVM has been found to be incremental (e.g., Bock et al. 2003; Griffin and Bock 2000), and hesitations have been linked to repairs (Brown-Schmidt and Tanenhaus 2006). It is quite probable that something similar happens with L2 speakers, and hesitations could very well be linked to reduced gender error rates. This, in any event, would not be the same as claiming that conceptual information is playing a direct role during grammatical encoding via feedback mechanisms.

A Remaining Question

A question remains as to why the 3rd person singular pronoun errors do not also affect the animacy feature—i.e., why the neuter pronoun, *it*, does not seem to participate in the 3rd person singular pronoun mishaps. In principle, if the only information required for the Spanish equivalent of the English sentences where pronoun gender errors appear is person and number, it could have been expected that the resulting errors involve the three 3rd person singular pronouns, not only *he* and *she*. But the scarcity of animacy errors (both as instantiated by inanimate pronouns taking the place of animate ones, and of animate pronouns taking the place of inanimate ones; see Table 4), suggests that this is not the case. One possibility is that the verbs themselves or some other sentence constituents are responsible for specifying a requirement of animacy on the pronoun, therefore precluding such animacy errors. A look at the verbs and predicates accompanying pronoun gender errors provides some support for such a possibility: Of the 45 different verbs appearing with gender errors, 39 could be said to require animate subjects (accept, arrange, wear, have a shower, etc.), and only 6 are generally used with both animate and inanimate subjects (be there, get, go, have, let, and take).

A second possibility is that in Spanish inanimate nouns are never referred to with a nominative pronoun (Luján 1997)—that is, nominative pronouns are only used to refer to animate entities, while inanimate objects can be referred to, when necessary, with demonstratives or alternative referential expressions. Still, while this difference could underlie the successful acquisition and use of the animacy distinction in English between the pronouns *he* and *she*, on the one hand, and the pronoun *it*, on the other, and explain the low rate of animacy errors in general, it is not clear how it could have an effect in cases of pronoun gender errors, when the Spanish equivalent of the sentence would have been pro-drop and would have been just as valid for animate as for inanimate subjects.

There is a third possibility already discussed in the section “Preparing pronouns for production”. Suppose all pronouns referring to animate entities (*he* and *she*, but also first and second person pronouns) are retrieved from conceptual features, but pronouns referring to inanimate objects (*it* and sometimes *they*), lacking the conceptual features necessary for encoding the pronoun directly, are processed only after having accessed the antecedent’s lemma and retrieved the grammatical features. This would explain why Spanish speakers do not also make animacy errors when they make pronoun gender errors in English, since pronoun gender errors involve cases in which the pronoun is being processed from conceptual features which would limit the choice to those pronouns referring to animate entities.

Conclusion

The data presented here provide initial evidence of the usefulness of studying L2 errors of proficient speakers to investigate the mechanisms of language production.

The specific case studied here—pronoun gender errors of Spanish speakers of English as L2, apart from likely shedding light on the composition of the PVM, impinges on several other aspects of language production, such as pronoun production, the psychological reality of covert *pro*, and the availability of feedback. Hopefully, contributions from the study of other pro-drop and non pro-drop language combinations will help support and refine these conclusions.

More generally, the proposed approach for detecting errors at the PVM level—focusing on discrepancies between languages as to informational requirements and analysing the expected L2 errors, could also be applied to other language pairs. Other types of L2 errors should provide us with more data from which to gain further insights into the structure and composition of the PVM.

Acknowledgments This research was supported by a grant from the Netherlands Organization for Scientific Research (NWO) to the author. Thanks to Wandert Logcher, Megan Gilliver, and Geert Geesink for their help with transcription and coding, and to Bruce Stevenson, Oren Sadeh-Leicht, Will Coventry and Don Hine for their advice on statistics. Many thanks also to Frank Wijnen and Tamar Gollan for their useful comments and suggestions.

Open Access This article is distributed under the terms of the Creative Commons Attribution Noncommercial License which permits any noncommercial use, distribution, and reproduction in any medium, provided the original author(s) and source are credited.

References

- Arnold, J., & Griffin, Z. M. (2007). The effect of additional characters on choice of referring expression: Everyone counts. *Journal of Memory and Language*, 56(4), 521–536.
- Bock, K., Eberhard, K. M., Cutting, J. C., Meyer, A. S., & Schriefers, H. (2001). Some attractions of verb agreement. *Cognitive Psychology*, 43, 83–128.
- Bock, K., Irwin, D. E., Davidson, D. J., & Levelt, W. J. M. (2003). Minding the clock. *Journal of Memory and Language*, 48, 653–685.
- Bock, K., & Levelt, W. J. M. (1994). Language production. Grammatical encoding. In M. Gernsbacher (Ed.), *Handbook of psycholinguistics* (pp. 945–984). New York: Academic Press.
- Bock, K., Loebell, H., & Morey, R. (1992). From conceptual roles to structural relations: Bridging the syntactic cleft. *Psychological Review*, 99(1), 150–171.
- Bock, K., & Miller, C. A. (1991). Broken agreement. *Cognitive Psychology*, 23, 45–93.
- Bock, K., Nicol, J., & Cutting, J. C. (1999). The ties that bind: Creating number agreement in speech. *Journal of Memory and Language*, 40, 330–346.

- Bock, K., & Warren, R. K. (1985). Conceptual accessibility and syntactic structure in sentence formation. *Cognition*, 21, 47–67.
- Broersma, M., & De Bot, K. (2006). Triggered codeswitching: A corpus-based evaluation of the original triggering hypothesis and a new alternative. *Bilingualism: Language and Cognition*, 9(1), 1–13.
- Brown-Schmidt, S., & Tanenhaus, M. K. (2006). Watching the eyes when talking about size: An investigation of message formulation and utterance planning. *Journal of Memory and Language*, 54, 592–609.
- Brown, A. S. (1991). A review of the tip-of-the-tongue experience. *Psychological Bulletin*, 109, 204–223.
- Corbett, G. G. (1991). *Gender*. Cambridge: Cambridge University Press.
- Corbett, G. G. (2000). *Number*. Cambridge: Cambridge University Press.
- De Bot, K. (2000). A bilingual production model: Levelt's 'speaking' model adapted. In L. Wei (Ed.), *The bilingualism reader* (pp. 420–442). London: Routledge.
- Declerck, R. (1983). 'It is Mr. Y' or 'He is Mr. Y'? *Lingua*, 59, 209–246.
- Dell, G. S. (1990). Effects of frequency and vocabulary type on phonological speech errors. *Language and Cognitive Processes*, 5, 313–349.
- Demestre, J., Meltzer, S., García-Albea, J. E., & Vigil, A. (1999). Identifying the null subject: Evidence from event-related potentials. *Journal of Psycholinguistic Research*, 28(3), 293–312.
- Eberhard, K. M. (1999). The accessibility of conceptual number to the processes of subject-verb agreement in English. *Journal of Memory and Language*, 41(4), 560–578.
- Featherston, S., Gross, M., Münte, T. F., & Clahsen, H. (2000). Brain potentials in the processing of complex sentences: An ERP study of control and raising constructions. *Journal of Psycholinguistic Research*, 29(2), 141–154.
- Ferreira, F. (1994). Choice of passive voice is affected by verb type and animacy. *Journal of Memory and Language*, 33, 715–736.
- Franck, J., Vigliocco, G., Antón-Méndez, I., Collina, S., & Frauenfelder, U. H. (2008). The interplay of syntax and form in sentence production: A cross-linguistic study of form effects on agreement. *Language and Cognitive Processes*, 23(3), 329–374.
- Garrett, M. F. (1988). Processes in language production. In J. F. Newmeyer (Ed.), *Linguistics: The Cambridge Survey* (Vol. 3, pp. 69–96). Cambridge: Cambridge University Press.
- Gerken, L. A. (1991). The metric basis for children's subjectless sentences. *Journal of Memory and Language*, 30, 431–451.
- Griffin, Z. M., & Bock, K. (2000). What the eyes say about speaking. *Psychological Science*, 11(4), 274–279.
- Haegeman, L. (1991). An inventory of empty categories. In: *Introduction to government and binding theory* (pp. 393–438). Cambridge, MA: Blackwell.
- Harvie, D. (1998). Null subject in English: Wonder if it exists? *Cahiers Linguistiques d'Ottawa*, 26, 15–25.
- Haskell, T. R. (2005). Speaker's eye movements during the production of adjectives, *Eighteen annual CUNY sentence processing conference*. Tucson, AZ.
- Jaeggli, O., & Safir, K. (1989). *The null subject parameter*. Dordrecht: Kluwer.
- Konishi, T. (1993). The semantics of grammatical gender: A cross-cultural study. *Journal of Psycholinguistic Research*, 22(5), 519–534.
- Kroll, J. F., & Tokowicz, N. (2005). Models of bilingual representation and processing: Looking back and to the future. In J. F. Kroll & A. M. B. deGroot (Eds.), *Handbook of bilingualism: Psycholinguistic approaches* (pp. 531–553). New York: Oxford University Press.
- La grammaire pour tous. (1990). (Vol.3). Paris: Librairie Hatier.
- Levelt, W. J. M. (1989). *Speaking: From intention to articulation*. Cambridge, MA: MIT Press.
- Levelt, W. J. M., & Maassen, B. (1981). Lexical search and order of mention in sentence production. In W. Klein & W. J. M. Levelt (Eds.), *Crossing the boundaries in linguistics* (pp. 221–252). Dordrecht: Reidel.
- Levelt, W. J. M., Roelofs, A., & Meyer, A. S. (1999). A theory of lexical access in speech production. *Behavioral and Brain Sciences*, 22(1), 1–75.
- Luján, M. (1997). La expresión y omisión del pronombre personal. In V. Demonte & I. Bosque (Eds.), *Nueva gramática descriptiva de la lengua española*. Madrid, Spain: Instituto Universitario Ortega y Gasset.
- Meyer, A. S., & Bock, K. (1999). Representations and processes in the production of pronouns: Some perspectives from Dutch. *Journal of Memory and Language*, 41, 281–301.
- Mikkelsen, L. (2005). *Copular clauses: Specification, predication and equation*. Amsterdam: John Benjamins.
- Oshita, H. (2004). Is there anything there when there is not there? *Second Language Research*, 20(2), 95–130.
- Poullisse, N. (1999). *Slips of the tongue. Speech errors in first and second language production* (Vol. 20). Amsterdam: John Benjamins.

- Poullisse, N., & Bongaerts, T. (1994). First language use in second language production. *Applied Linguistics*, 15, 36–57.
- Rodriguez-Fornells, A., van der Lugt, A., Rotte, M., Britti, B., Heinze, H.-J., & Münte, T. F. (2005). Second language interferes with word production in fluent bilinguals: Brain potential and functional imaging evidence. *Journal of Cognitive Neuroscience*, 17(3), 422–433.
- Schmitt, B. M., Meyer, A. S., & Levett, W. J. M. (1999). Lexical access in the production of pronouns. *Cognition*, 69, 313–335.
- 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, MA: Cambridge University Press.
- Slobin, D. I. (1996). From “thought and language” to “thinking for speaking”. In J. J. Gumperz & S. C. Levinson (Eds.), *Rethinking linguistic relativity* (pp. 71–96). Cambridge, MA: Cambridge University Press.
- Sridhar, S. N. (1988). *Cognition and sentence production: A cross-linguistic study*. New York: Springer.
- Swinney, D. A., Ford, M., Frauenfelder, U., & Bresnan, J. (1988). Coreference assignment during sentence processing. In B. Grosz, R. Kaplan, M. Macken & I. Sag (Eds.), *Language structure and processing*. Stanford: CSLI.
- Tarone, E. E. (1987). Some influences on the syllable structure of interlanguage phonology. In G. Ioup & S. H. Weinberger (Eds.), *Interlanguage phonology* (pp. 232–247). Cambridge, MA: Newbury House.
- Toribio, A. J., Roebuck, R., & Lantolf, J. P. (1993). On the role of UG in adult second language acquisition: Evidence from pro-drop. In *Linguistic Society of America*. Los Angeles.
- Tropf, H. (1987). Sonority as a variability factor. In A. James & J. Leather (Eds.), *Sound patterns in second language acquisition*. Dordrecht: Foris.
- Valian, V. (1990). Null subjects: A problem for parameter-setting models. *Cognition*, 35, 105–122.
- Vigliocco, G., Butterworth, B., & Garrett, M. F. (1996). Subject-verb agreement in Spanish and English: Differences in the role of conceptual constraints. *Cognition*, 61, 261–298.
- 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.
- Yamashita, H., Chang, F., & Hirose, Y. (2005). Producers build structures only for overt arguments. In *18th Annual CUNY sentence processing conference*. Tucson, AZ.