Bilingualism: Language and Cognition 13 (2), 2010, 87–98 © Cambridge University Press 2009 doi:10.1017/S1366728909990034

Processing subject–verb agreement in a second language depends on proficiency*

NORIKO HOSHINO
PAOLA E. DUSSIAS
JUDITH F. KROLL
Pennsylvania State University

(Received: March 9, 2007; Revised: April 8, 2009; Accepted: April 27, 2009; First published online 30 November 2009)

Subject–verb agreement is a computation that is often difficult to execute perfectly in the first language (L1) and even more difficult to produce skillfully in a second language (L2). In this study, we examine the way in which bilingual speakers complete sentence fragments in a manner that reflects access to both grammatical and conceptual number. In two experiments, we show that bilingual speakers are sensitive to both grammatical and conceptual number in the L1 and grammatical number agreement in the L2. However, only highly proficient bilinguals are also sensitive to conceptual number in the L2. The results suggest that the extent to which speakers are able to exploit conceptual information during speech planning depends on the level of language proficiency.

In languages such as English and Spanish, if the subject is singular, the verb must be singular; if the subject is plural, the verb must be plural. Although agreement between the subject and verb of a sentence might seem as if it should follow a very simple rule, even highly proficient speakers producing sentences in their native language make agreement errors more often than might be expected. One factor that complicates the production of subject–verb agreement is that grammatical number and conceptual number do not always match. Past research has used a sentence completion task to examine the consequence of grammatical and conceptual mismatches in the production of subject–verb agreement (e.g., Bock and Miller, 1991; Vigliocco, Butterworth and Garrett, 1996; Nicol, Teller and Greth, 2001). In this task, speakers are given a sentence fragment and are asked to produce a possible completion. For example, when shown the fragment, "The drawing on the posters", the participant might complete the sentence by adding "was colorful". Of interest is that the subject of the sentence can refer to either a single referent or multiple referents. Consider the examples below:

(1) The author of the novels (Single-Referent Mismatch)

(2) The drawing on the posters (Distributive-Referent Mismatch)

Here the number of the head noun phrase (NP) is mismatched with that of the local NP in both cases. The verb to be produced must agree with the head NP in number. Although the grammatical number is the same in both examples (i.e., singular), the conceptual number differs. In the first case, our conceptual interpretation is that there can be only a single author who wrote several novels. In the second case, we understand that each of the posters has a drawing and that there are, therefore, multiple drawings, one corresponding to each poster. In the sentence completion task, the most critical dependent measure is the rate of subject–verb agreement errors (e.g., "The author of the novels *are* clever" and "The drawing on the posters *are* colorful") relative to responses to control preambles in which the number of the nouns in the complex NP matches, such as "The author of the novel is clever" and "The drawing on the poster is colorful", respectively. If speakers are sensitive to the conceptual number in representing the subject phrase in the production of subject–verb agreement, the rate of subject–verb agreement errors should be higher for the distributive-referent than for the single-referent because the plural notional number of the distributive-referent presents a conflict with the grammatical number of the phrase. In contrast, if speakers are not sensitive to the conceptual number in representing the subject when computing subject–verb agreement and rely only on grammatical number, the rate of agreement errors should be no different for these two fragment types.

In early research on this topic, there was a suggestion that the production of subject–verb agreement in English was unlike other languages in that speakers were sensitive to grammatical number only (e.g., Bock and Miller, 1991; Vigliocco, Butterworth and Semenza, 1995; Vigliocco, Butterworth and Garrett, 1996; Vigliocco, Hartsuiker,

* This study was completed by the first author as a master's thesis under the direction of the second and third authors. The research reported in this paper was supported in part by NSF Grant BCS-0418071 and NIH Grant R01-HD053146 to Judith F. Kroll, by NIH Grant HD50629 and NSF Grant BCS-0821924 to Paola E. Dussias and by NSF Dissertation Grant BCS-0518814 to Noriko Hoshino and Judith F. Kroll. We thank James Burns, Natalie De Rosa, Mark Minnick, Judith Pirela and Raul Rios for research assistance. Preliminary results were presented at the 43rd Annual Meeting of the Psychonomic Society, the 4th International Symposium on Bilingualism, and the 4th International Conference on the Mental Lexicon.

Address for correspondence:
Noriko Hoshino, ESRC Centre for Research on Bilingualism in Theory and Practice, Bangor University, Bangor Gwynedd LL57 2DG, UK
n.hoshino@bangor.ac.uk

Jarema and Kolk, 1996). However, subsequent research has shown that native English speakers, like speakers of Spanish, Italian, French and Dutch, appear to be sensitive to both grammatical and conceptual number (e.g., Eberhard, 1999; Bock, Eberhard and Cutting, 2004).

If an individual is bilingual in two languages that compute grammatical and conceptual number similarly, then bilingualism itself might not be expected to affect the ability to produce agreement correctly in each language. Only a small number of studies have examined the production of subject–verb agreement in bilinguals. In each case, these studies show that bilinguals appear to exploit grammatical and conceptual information in each of their languages (e.g., Nicol et al., 2001; Nicol and Greth, 2003; Van Hell and Mensies, 2004). An aspect of the bilingual results that is notable is that the sensitivity to grammatical and conceptual number has been observed for late bilinguals (Nicol and Greth, 2003), for early bilinguals (Nicol et al., 2001) and for both speaking and writing (Van Hell and Mensies, 2004). The only exception to this general pattern has been reported for bilinguals who speak native languages such as Chinese, in which there is no comparable computation of subject–verb agreement in number. A recent study of the comprehension of subject–verb agreement in English for Chinese–English bilinguals also provides evidence for sensitivity to agreement in offline but not in online measures (Jiang, 2004) and in event-related potentials (ERPs) but not in behavioral measures such as response time (RT) (Chen, Shu, Liu, Zhao and Li, 2007). Nicol and Greth (2003) argued that bilinguals tend to transfer agreement strategies from the first language (L1) to the L2. If the L1 does not easily enable transfer to occur, then the acquisition of agreement in the L2 will presumably be more difficult.

In addition to constraints that are imposed by distinct cross-language preferences, there is another respect in which the performance of bilinguals may differ for their two languages. Recent studies suggest that even for relatively proficient bilinguals, the L2 is likely to be processed less automatically (e.g., Segalowitz and Hulstijn, 2005) and to make additional demands on cognitive resources compared to the L1 (e.g., Miyake and Friedman, 1998; Hasegawa, Carpenter and Just, 2002; Michael and Gollan, 2005). A number of past studies of individuals speaking their native language have shown that skill in producing subject–verb agreement is influenced by available working memory resources. Hartsuiker, Kolk and Huinck (1999) found that Dutch aphasics, hypothesized to have severely reduced cognitive resources, were sensitive to grammatical but not conceptual number in producing subject–verb agreement (but see Vigliocco and Zilli, 1999). A control group of unimpaired Dutch speakers showed sensitivity to both grammatical and conceptual number. In a study of Dutch university students, Hartsuiker and Barkhuysen (2006) reported further evidence for a role of working memory resources in computing subject–verb agreement. In the presence of an increased memory load, individuals with low speaking span committed more agreement errors than high-span individuals. However, unlike aphasics, the overall performance on the subject–verb agreement task was affected by cognitive resources, but not specifically the sensitivity to conceptual number; agreement errors were simply higher overall under conditions of reduced resources. Hartsuiker and Barkhuysen (2006) proposed that cognitive resources do not affect the mapping of conceptual information to lemmas and/or the maintenance of the mapping but rather a secondary control process that checks for the presence of conflict in agreement. In the presence of an additional memory load or relatively reduced working memory span, participants will effectively skip the control process and therefore make an increased number of agreement errors. They speculate that only when cognitive resources are pathologically limited (e.g., for aphasics) will the mapping of conceptual number or the maintenance of that mapping itself be impaired.

Unlike aphasics, bilinguals are certainly not pathologically impaired. To the contrary, there is recent evidence suggesting that life experience as a bilingual confers a set of cognitive benefits, specifically to executive control functions (e.g., Bialystok, Craik, Klein and Viswanathan, 2004). However, in addition to the potentially increased memory demands on processing in the L2, bilinguals may also experience some constraints in their ability to fully compute the conceptual nuances that are normally available quite automatically in the L1 (e.g., Finkbeiner, Forster, Nicol and Nakamura, 2004; Kotz and Elston-Güttler, 2004; Silverberg and Samuel, 2004). For example, using a semantic priming paradigm with both RT and ERP measures, Kotz and Elston-Güttler (2004) showed that even advanced L2 learners were sensitive only to associative relations but unable to exploit categorical information, whereas early bilinguals and native speakers were sensitive to both associative and categorical information. If the computation of subject–verb agreement is indeed a two-stage process, as Hartsuiker and Barkhuysen (2006) propose, the consequences of bilingualism for processing in the L2 should be evident at each of these stages. At the first stage, there may be restrictions on the range of conceptual representations that can be accessed and maintained. At the second stage, the ability to verify agreement in a post-hoc check may be limited by available cognitive resources in L2. When bilinguals attain a sufficiently high level of proficiency in the L2 (Nicol et al., 2001) and/or are living in a bilingual environment in which there is a high degree of support for the L2 (Nicol and Greth, 2003; Van Hell and Mensies, 2004), either or both of the two hypothesized processing components may be completed adequately. The question we address in the present paper is whether

bilinguals who differ in their relative proficiency in the L2 will also differ in their sensitivity to grammatical and conceptual number in producing subject–verb agreement.

If bilinguals transfer agreement strategies from the L1 to the L2, as Nicol and Greth (2003) propose, then the ability to compute subject–verb agreement in the L2 in an online task should be a function of whether the bilingual has skilled access to that computation in the L1. Hoshino, Dussias and Kroll (in preparation) tested the transfer hypothesis by examining the performance of Japanese–English bilinguals on a sentence completion task in English, their L2. The results showed that although Japanese speakers do not have a similar construction in their L1, they were able to successfully compute grammatical number. These findings suggest that a simple version of cross-language transfer from the L1 to the L2 does not account for L2 performance. Likewise, the two-stage model of subject–verb agreement proposed by Hartsuiker and Barkhuysen (2006) was intended to account for the performance of aphasic patients and non-aphasic adults who differ in available working memory resources. Although the results in that study were complex, they are consistent with a model in which cognitive resources must be recruited to enable sensitivity to conceptual number. However, as noted above, the two stages, one involved in the mapping and maintenance of grammatical and conceptual information and the other engaging a later checking mechanism, do not themselves make differential predictions about the role of cognitive resources constrained by language proficiency. Either or both stages might suffer in the absence of sufficient resources.

In the present paper, we test the hypothesis that access to conceptual number is more easily disrupted when cognitive resources are stressed by virtue of the cognitive load imposed by processing in an L2. This hypothesis is based on the observation that virtually no past subject–verb agreement studies fail to observe effects of grammatical number whereas effects of conceptual number are apparently vulnerable to brain damage (Hartsuiker et al., 1999; but see Vigliocco and Zilli, 1999), and relative imageability of the head noun (Eberhard, 1999). As noted earlier, recent studies of subject–verb agreement in English suggest that English functions much like other languages with respect to sensitivity to conceptual number, although the initial studies on this topic suggested otherwise (e.g., Bock and Miller, 1991). However, critical to the present hypothesis, the result in those early studies was always the same: English speakers were sensitive to grammatical but not conceptual number. Taken together, the empirical evidence suggests that the computation of conceptual number will be more affected by the availability of cognitive resources than the computation of grammatical number. Note that this hypothesis is agnostic with respect to debates concerning the independence or interaction between syntactic and semantic information during the planning of spoken sentences (e.g., Vigliocco and Hartsuiker, 2002). Conceptual information may be more vulnerable to disruption due to the nature of the required computation rather than to the sequencing with which that information is available.

In the present study, we tested this hypothesis by examining the consequences of proficiency in the L2 for bilinguals performing a sentence completion task in both English and in Spanish. In two experiments, we compared the production of subject–verb agreement for a group of less proficient English–Spanish bilinguals (Experiment 1) and a group of less proficient Spanish–English bilinguals (Experiment 2) with a group of more proficient Spanish–English bilinguals (Experiment 2). If the degree to which bilinguals require additional cognitive resources to process the L2 is a function of proficiency, then we predicted that the highly proficient bilinguals would be more likely to show sensitivity to both grammatical and conceptual number agreement in their L2, whereas bilinguals who are only relatively proficient in the L2 would show sensitivity to grammatical number but not to conceptual number. We also expected bilinguals to demonstrate sensitivity to grammatical and conceptual number in the L1 regardless of their L2 proficiency. In L1, they should produce more subject–verb agreement errors for the distributive-referent mismatch preambles (e.g., "the drawing on the posters") than for the single-referent mismatch ones (e.g., "the author of the novels").

Experiment 1: English–Spanish bilinguals

Method

Participants

Forty-two English–Spanish bilinguals participated in Experiment 1 for payment. The participants were all native speakers of English who spoke Spanish as an L2 and were living in an environment in which English was predominant at the time of testing. Their self-assessed abilities in English and Spanish are provided in Table 1. Four participants were excluded from all analyses due to poor performance on the sentence completion task

Table 1. *Characteristics of English–Spanish bilinguals in Experiment 1.*

	English–Spanish bilinguals ($n = 38$)
Age (years)	22.2 (4.8)
L1 self-ratings (10-point scale)	9.2 (1.0)
L2 self-ratings (10-point scale)	6.5 (1.4)
Length of L2 immersion (months)	6.1 (2.2)

Note. Standard deviations are in parentheses.

(less than 30% overall accuracy in either or both of their languages). The remaining 38 participants were included in the data analyses.

Materials

Sentence completion task. Sixty-four pairs of sentence preambles consisting of a singular head NP followed by a modifying PP were developed in English and in Spanish. In one of each pair, the grammatical number of the head NP was mismatched with that of the adjunct NP within the PP modifier (e.g., *The author of the novels*), whereas in the other one the grammatical number of the head NP was matched with that of the adjunct NP (e.g., *The author of the novel*). Thus, there were four types of preambles in this experiment, as shown in (3) and (4). The Spanish materials were translations of the English sentences that were checked by a native Spanish speaker. In Spanish, the gender of the head NP and the adjunct NP as well as gender agreement between the head NP and the adjunct NP were counterbalanced. The complete list of the experimental preambles is provided in the Appendix.

(3) a. The author of the novels
a′. La autora de las novelas
(Single-Referent, Mismatch)
b. The author of the novel
b′. La autora de la novela
(Single-Referent, Match Control)

(4) a. The drawing on the posters
a′. El dibujo de los carteles
(Distributive-Referent, Mismatch)
b. The drawing on the poster
b′. El dibujo del cartel
(Distributive-Referent, Match Control)

In addition to these experimental preambles, 64 pairs of fillers consisting of a plural head NP followed by a modifying PP were developed. In one of each pair, the grammatical number of the head NP was matched with that of the adjunct NP (e.g., *The rooms in the apartments*), whereas in the other of each pair the grammatical number of the head NP was mismatched with that of the adjunct NP (e.g., *The rooms in the apartment*). None of the fillers was the same as the experimental preambles.

Four 64-item lists were constructed and two lists were paired so that each participant received the critical conditions in the context of a different item in each of their languages, with lists counterbalanced across participants. In each list, there were 32 experimental items (eight items for each condition) and 32 filler items (16 items for each condition). All preambles were paired with an adjective that could modify the preamble NP plausibly (e.g., "fat" for "the author of the novels"). None of the adjectives was repeated. The order of the presentation of items was randomized for each participant.

The one difference between the English and Spanish materials was that in the Spanish version, adjectives were presented in their stems so that they could not carry information about number and gender (e.g., *estresad_* instead of *estresado/estresada*). In other words, these Spanish adjectives were modified to play the same role in subject–verb agreement as English adjectives.

Language history questionnaire. A questionnaire was designed to obtain information about participants' language learning experiences. The questionnaire asked participants about their native and home languages, the amount and type of their language learning experience, the length of living in target language cultures, and self-rated proficiency for each of the four language skills of reading, writing, speaking and listening in their L1 and L2. Proficiency was rated on a 10-point scale with 1 being not proficient and 10 being very proficient.

Procedure

Participants were tested individually in a quiet room. They were seated in front of a computer monitor, a button box, a microphone connected to the computer and an audiocassette tape recorder. Participants were given the sentence completion task in English first and then in Spanish. The English-first design was adopted to give the participants practice on the task before having to produce sentences in their less proficient L2. At the end of the experiment, participants were asked to complete the language history questionnaire.

In the sentence completion task, participants first received written instructions on the computer monitor. They were informed that each trial would begin with the presentation of an adjective which would then be replaced with the sentence preamble. Following the offset of the preamble, a beep signaled to the participant that they should initiate their response. Their task was to repeat the preamble and complete the sentence using the presented adjective as quickly as possible after the beep. One example was given in the instructions. However, no other instruction was given about the form of the responses. Eight practice trials preceded the actual experimental session to familiarize participants with the experimental procedure.

On each trial, as illustrated in Figure 1, a fixation sign (+) was presented at the center of the computer screen. At the self-paced press of a button, the fixation sign was replaced with an adjective, which was present for 600 ms. A sentence preamble followed the adjective and remained on the computer screen for 1800 ms. A beep was inserted at the end of the presentation of the sentence preamble to signal the initiation of the response. Responses were recorded on tape. Because the task was self-paced, participants were required to press a button in order to proceed to the next trial.

Table 2. *Distribution of response (in percent) by scoring category in Experiment 1 (N = 38).*

	English (L1)				Spanish (L2)			
	Single		Distributive		Single		Distributive	
Category	Mismatch	Match	Mismatch	Match	Mismatch	Match	Mismatch	Match
Correct	73.1	76.4	72.1	84.4	55.0	65.2	57.6	68.1
Agreement error	3.3	1.2	8.0	0.0	6.2	0.8	4.9	1.2
Other error	23.6	22.4	19.9	15.6	38.8	34.0	37.6	30.6

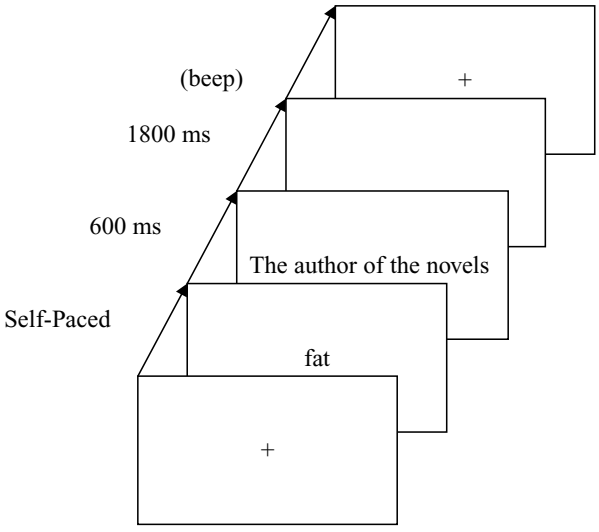

Figure 1. An illustration of a trial in the sentence completion task.

Scoring

Tape-recorded responses were transcribed for the critical sentence trials. Responses that the microphone did not detect were eliminated as technical errors. The rest of the responses were scored based on the following three categories for the purpose of the present paper: (1) responses were scored as CORRECT if participants repeated the presented preamble correctly and completed a sentence in a sensible manner with a verb marked as singular. Even if they used a different adjective from the one provided, it was scored as correct if the completed sentence was semantically congruent; (2) responses were scored as AGREEMENT ERRORS if participants repeated the presented preamble correctly but used a verb form that did not agree in number with the head NP of the preamble (i.e., marked a verb as plural); (3) responses were scored as OTHER ERRORS if participants made responses that did not meet any criteria that are described above. In the present paper, we focus on agreement errors because that has been the main dependent measure in past studies, but data for the other scoring categories are also reported.

Analyses

A 2 × 2 analysis of variance was conducted by participants (F_1) and by items (F_2) on agreement errors, with number match (mismatch vs. match control) and distributivity (single-referent vs. distributive-referent) as independent variables. In the participant analyses, number match and distributivity were within-participants variables. In the item analyses, number match was a within-items variable, whereas distributivity was a between-items variable.

Results and discussion

English (L1)

After trimming 256 trials for six excluded pairs of items based on the norming data (21.1% of the data),[1] 960 responses were scored according to the categories described above and mean percentages for each scoring category were computed. The distribution of responses in the different scoring categories for each experimental condition is given in Table 2. Again, the focus in the analyses to be reported is on agreement errors.

The main effect of number match emerged by participants and by items [$F_1(1,37) = 14.49$, $MSE = 67.38$, $p < .01$; $F_2(1,50) = 25.84$, $MSE = 8.36$, $p < .001$]. The main effect of distributivity was marginally significant by participants [$F_1(1,37) = 3.48$, $MSE = 33.36$, $p = .07$; $F_2(1,50) = 2.29$, $MSE = 30.67$, $p > .10$]. More critically, there was a significant interaction between number match and distributivity [$F_1(1,37) = 7.51$, $MSE = 43.72$, $p < .01$; $F_2(1,50) = 8.36$, $MSE = 25.40$, $p < .01$]. A test of simple main effects showed that more agreement errors were produced for distributive-referent mismatched items than for single-referent mismatched items [$F_1(1,37) = 6.09$, $MSE = 68.54$, $p < .05$; $F_2(1,50) = 5.62$, $MSE = 50.09$, $p < .05$]. However, the difference between

[1] A group of 10 English monolinguals judged all the single-referent and distributive-referent mismatched items according to whether they refer to one thing or to more than one thing, with "1" being "one thing" and "2" being "more than one thing". The mean for each preamble was computed. Six distributive-referent mismatched items whose mean rating was smaller than 1.4 and six single-referent mismatched items whose mean rating was 1.5 or greater than 1.5 and their matched controls were excluded from all the analyses.

the two matched control conditions was only marginally significant [$F_1(1,37) = 3.15$, $MSE = 8.54$, $p = .08$; $F_2(1,50) = 3.19$, $MSE = 6.01$, $p = .08$].

The important result in this analysis is that a significant interaction was obtained between number match and distributivity, with the performance of English–Spanish bilinguals impaired for the distributive-referent mismatched condition but not for the single-referent mismatched condition. Like the results reported by Nicol and Greth (2003), the present findings suggest that English–Spanish bilinguals are sensitive to the conceptual as well as grammatical number of the subject phrase in their L1 English.

Spanish (L2)

After trimming 244 trials for the six excluded pairs of items (20.1% of the data), 972 responses were scored according to the categories described above and the distribution of responses in the different scoring categories for each experimental condition is given in Table 2. Only the analyses on agreement errors are reported below.

The main effect of number match was reliable both by participants and by items [$F_1(1,52) = 11.25$, $MSE = 68.92$, $p < .01$; $F_2(1,50) = 17.88$, $MSE = 26.25$, $p < .001$] with more agreement errors for number mismatched items than for number matched items. As can be seen in Table 2, the main effect of distributivity and the interaction between number match and distributivity were not significant [$F_2(1,50) = 1.19$, $MSE = 26.25$, $p > .10$ for the interaction between number match and distributivity; all the other Fs < 1].

These results suggest that, contrary to Nicol and Greth (2003), the English–Spanish bilinguals were not sensitive to the distributive number of the subject phrase, but only to the grammatical number in Spanish, their L2. The most plausible explanation for the lack of an interaction between number match and distributivity in Spanish is that the English–Spanish bilinguals in the present study may not have been as highly proficient in Spanish as those in Nicol and Greth's study. The participants in this experiment had taken 4.6 semesters of Spanish at the college level on average and 26 out of 38 participants had taken it for less than six semesters. This contrasts with the English–Spanish bilinguals in Nicol and Greth's study, who had taken at least six semesters of Spanish. Furthermore, the English–Spanish bilinguals in the present study lived in the environment where Spanish was rarely used in everyday life, whereas those in Nicol and Greth's study lived in the environment where both English and Spanish were widely spoken.

In Experiment 2, we examined the performance of Spanish–English bilinguals who, like the English–Spanish bilinguals, were late bilinguals but who varied in their level of proficiency in English as the L2. Using an independent measure of L2 skill, the Spanish–English bilinguals were grouped by their relative proficiency in English. If the absence of a conceptually distributive number effect in Spanish in Experiment 1 was due to the properties of the translation materials in Spanish, then neither group should show the interaction between number match and distributivity in Spanish although it is their L1. Alternatively, if the lack of the effect in Spanish in Experiment 1 was due to the participants' level of proficiency, then Spanish–English bilinguals should show the effect in Spanish, their L1, but the effect in English, their L2, should depend on their proficiency.

Experiment 2: Spanish–English bilinguals

Method

Materials

Sentence completion task and language history questionnaire. The materials were identical to Experiment 1 except for the format of the language history questionnaire, which was modified for non-native speakers of English.

Lexical decision task in English. A lexical decision task in English was included in Experiment 2 as an independent online proficiency measure. Forty English words and 40 English pseudo-homophones were taken from Azuma and Van Orden (1997). The order of the presentation of items was randomized for each participant.

Procedure

Participants performed the sentence completion task followed by the lexical decision task in English. They completed the language history questionnaire at the end of the experiment.

Sentence completion task. The procedure was identical to Experiment 1 except that the order of the two language blocks was counterbalanced. Because the participants in this experiment were living in the L2 environment at the time of testing and were more proficient in the L2, we followed the convention of counterbalancing the order of the language blocks in the sentence completion task.

Lexical decision task in English. Participants first received written instructions on the computer monitor. They were informed that strings of letters would be presented one at a time on the computer screen. Their task was to decide whether each string was a real English word or not. At the beginning of each trial, a fixation sign (+) was presented for 500 ms at the center of the computer screen. When the fixation sign was replaced with the string of letters, participants were required to judge whether the string was a real English word, and to indicate their decision by pressing the "yes" button on the left if it was a real word, or the "no" button on the right if it was not a real word. After they responded, a fixation sign appeared for 500 ms and the next string of letters was

Table 3. *Characteristics of more and less proficient Spanish–English bilinguals in Experiment 2.*

	More proficient Spanish–English ($n = 20$)	Less proficient Spanish–English ($n = 15$)	p value t-test
Age (years)	26.8 (4.7)	26.9 (5.1)	n.s.
L1 self-ratings (10-point scale)	9.8 (0.4)	9.5 (0.6)	n.s.
L2 self-ratings (10-point scale)	8.7 (1.1)	7.6 (1.2)	<.01
Length of L2 immersion (months)	63.9 (34.3)	37.8 (29.4)	<.05
Lexical decision in L2 English			
Accuracy (%) for nonwords	89.5 (3.6)	74.3 (7.3)	.001
Accuracy (%) for words	93.5 (3.7)	92.0 (4.2)	n.s.
Response latencies (ms) for nonwords	899 (239)	1198 (331)	<.01
Response latencies (ms) for words	744 (147)	841 (150)	<.07

Note. Standard deviations are in parentheses.

presented. Ten practice trials preceded the experimental session.

Scoring

Sentence completion task. The scoring procedure was identical to Experiment 1.

Lexical decision task in English. After trimming RTs for correct responses that were less than 300 ms or greater than 3000 ms, RTs that were 2.5 standard deviations above or below the mean were identified as outliers and excluded from the analyses. Finally, we calculated accuracy for correct responses again.

Participants

Fifty-two Spanish–English bilinguals participated in Experiment 2 for payment. However, data from 17 participants were excluded from the analyses because they did not perform the lexical decision task in the L2 English which was used to divide Spanish–English bilinguals into two proficiency groups. The remaining 35 participants were all native speakers of Spanish who learned English as an L2 and had been living in the United States at the time of testing. These 35 Spanish–English bilinguals were divided into two proficiency groups based on their performance in the lexical decision task in the L2 English. Because the median accuracy for nonwords was 85.0%, we grouped those whose nonword accuracy was 85% or higher as more proficient and those whose nonword accuracy was less than 85% as less proficient. Table 3 provides a direct comparison between more proficient and less proficient Spanish–English bilinguals. The comparison makes it clear that the more proficient group performed more accurately in the English lexical decision task and had lived longer in an English-speaking environment than the less proficient group. Furthermore, the mean ratings for L1 were significantly higher than those for L2 regardless of the proficiency group [$t(17) = 4.01$, $p < .01$ for more proficient group and $t(12) = 7.10$, $p < .001$ for less proficient group],[2] reflecting the fact that Spanish was still their dominant language even for more proficient Spanish–English bilinguals, even though they were living and working in an English-speaking environment.

Analyses

A $2 \times 2 \times 2$ analysis of variance was conducted by participants (F_1) and by items (F_2) on agreement errors, with number match (mismatch vs. match control), distributivity (single-referent vs. distributive-referent) and proficiency group (more proficient vs. less proficient) as independent variables. In the participant analyses, number match and distributivity were within-participants variables, whereas proficiency group was a between-participants variable. In the item analyses, number match and proficiency group were within-items variables, whereas distributivity was a between-items variable. In the following section, we consider the performance in the L1 Spanish followed by the performance in the L2 English.

Results and discussion

Spanish (L1)

After trimming 212 trials for six excluded pairs of items (18.9% of the data), 908 responses were scored with the same criteria as Experiment 1 and mean percentages for each scoring category were computed. The distribution of responses in the different scoring categories for each experimental condition is provided in Table 4. In

[2] Two more-proficient Spanish–English bilinguals and two less-proficient Spanish–English bilinguals did not answer the questions on the self-assessed ratings of proficiency.

Table 4. *Distribution of response (in percent) by scoring category in Experiment 2 (N = 35).*

	More proficient Spanish–English bilinguals ($n = 20$)							
	Spanish (L1)				English (L2)			
	Single		Distributive		Single		Distributive	
Category	Mismatch	Match	Mismatch	Match	Mismatch	Match	Mismatch	Match
Correct	85.4	90.7	81.3	89.9	68.8	81.8	60.3	87.5
Agreement error	1.7	0.0	8.2	0.7	4.8	0.0	13.9	0.0
Other error	13.0	9.3	10.5	9.4	26.4	18.2	25.7	12.5
	Less proficient Spanish–English bilinguals ($n = 15$)							
	Spanish (L1)				English (L2)			
	Single		Distributive		Single		Distributive	
Category	Mismatch	Match	Mismatch	Match	Mismatch	Match	Mismatch	Match
Correct	82.1	90.6	67.2	86.7	38.1	68.3	50.9	70.1
Agreement error	3.2	0.0	10.3	0.0	13.8	1.0	8.0	1.7
Other error	14.8	9.4	22.5	13.3	48.1	30.8	41.1	28.3

the analyses reported below, we focus on agreement errors.

The main effect of number match was reliable by participants and by items [$F_1(1,33) = 12.13$, $MSE = 89.97$, $p < .01$; $F_2(1,50) = 24.43$, $MSE = 70.07$, $p < .001$]. The main effect of distributivity was marginally significant by participants and by items [$F_1(1,33) = 7.69$, $MSE = 57.07$, $p < .01$; $F_2(1,50) = 9.32$, $MSE = 69.37$, $p < .01$]. Critically, these main effects were qualified by a significant interaction between number match and distributivity [$F_1(1,33) = 10.49$, $MSE = 33.91$, $p < .01$; $F_2(1,50) = 8.25$, $MSE = 70.07$, $p < .01$]. A test of simple main effects showed that more agreement errors were produced for distributive-referent mismatched items than for single-referent mismatched items [$F_1(1,34) = 9.35$, $MSE = 85.49$, $p < .01$; $F_2(1,50) = 10.17$, $MSE = 58.82$, $p < .001$]. However, the difference between the two matched control conditions was only marginally significant [$F_1(1,34) = 1.00$, $MSE = 2.92$, $p > .10$; $F_2(1,50) = 1.00$, $MSE = 2.37$, $p > .10$]. Neither main effect of proficiency group nor other interactions was significant [$F_2(1,50) = 1.00$, $MSE = 61.76$, $p > .10$ for proficiency; $F_2(1,50) = 1.61$, $MSE = 70.35$, $p > .10$ for the interaction of number match and proficiency; all the other Fs < 1].

An important result in Experiment 2 is that, like the performance of the English–Spanish bilinguals in Experiment 1, all bilinguals showed the interaction between number match and distributivity in their L1. The results are consistent with the performance of Spanish monolinguals in the past studies (Vigliocco, Butterworth and Garrett, 1996). Furthermore, these data demonstrate that the failure to obtain the effect of distributivity in Experiment 1 was not due to the Spanish translations of the materials. These results suggest that if speakers are proficient in the target language, they will be sensitive to the distributive as well as grammatical number of the subject phrase during the process of subject–verb agreement. However, it is also possible that speakers are sensitive to both grammatical and conceptual number only in their L1. To determine whether proficiency modulates the interaction between number match and distributivity in their L2, we now consider the performance of these more and less proficient Spanish–English bilinguals in their L2 English.

English (L2)

After trimming 217 trials for six excluded pairs of items (19.4% of the data), 903 responses were scored with the same criteria as Experiment 1 and mean percentages for each scoring category were computed. The distribution of responses in the different scoring categories for each experimental condition is provided in Table 4. Again, only the analyses on agreement errors are reported below.

The main effect of number match was significant by participants and by items [$F_1(1,33) = 33.29$, $MSE = 92.25, p < .001$; $F_2(1,50) = 28.04, MSE = 168.60$, $p < .001$]. The main effect of distributivity or proficiency group or the interaction between number match and distributivity or between number match and proficiency group did not emerge [$F_2(1,50) = 1.01$, $MSE = 124.77$, $p > .10$ for proficiency; all the other

Fs < 1]. However, the interaction between distributivity and proficiency group was significant by participants [$F_1(1,33) = 5.12$, $MSE = 85.87$, $p < .05$] and marginally significant by items [$F_2(1,50) = 3.99$, $MSE = 124.77$, $p = .05$]. More importantly, there was a significant three-way interaction of number match, distributivity and proficiency group [$F_1(1,33) = 8.41$, $MSE = 63.22$, $p < .01$; $F_2(1,50) = 5.05$, $MSE = 124.49$, $p < .05$]. To follow up the significant three-way interaction, separate 2 (number match) × 2 (distributivity) ANOVAs were run for each proficiency group. For the more proficient group, the interaction between number match and distributivity was significant [$F_1(1,19) = 5.63$, $MSE = 74.97$, $p < .05$; $F_2(1,50) = 5.00$, $MSE = 97.46$, $p < .05$], such that more agreement errors were produced for distributive-referent mismatched items than for single-referent mismatched items [$F_1(1,19) = 5.63$, $MSE = 149.95$, $p < .05$; $F_2(1,50) = 5.00$, $MSE = 194.92$, $p < .05$], but there was no difference between the two matched control conditions [Fs < 1]. For the less proficient group, the interaction between number match and distributivity was marginally significant by participants but not by items [$F_1(1,14) = 3.42$, $MSE = 47.28$, $p = .09$; $F_2 < 1$]. If anything, however, this interaction was opposite to the one for the more proficient group, such that the less proficient bilinguals tended to produce more agreement errors for single-referent mismatched items than for distributive-referent mismatched items [$F_1(1,14) = 1.98$, $MSE = 129.75$, $p > .10$; $F_2 < 1$].

An important result in this analysis is that a significant interaction between number match and distributivity was obtained in the L2 English for more proficient Spanish–English bilinguals, but not for less proficient Spanish–English bilinguals. This interaction indicates that the performance of more proficient Spanish–English bilinguals was impaired more for distributive-referent preambles than for single-referent preambles. Contrary to the finding in Experiment 1 that the English–Spanish bilinguals did not show sensitivity to conceptual number in their L2 Spanish, the more proficient Spanish–English bilinguals were sensitive to both grammatical and conceptual number of the subject phrase during the process of subject–verb agreement in their L2 English. Similarly to the English–Spanish bilinguals in Experiment 1, on the other hand, the less proficient Spanish–English bilinguals showed sensitivity only to the grammatical number. These results suggest that speakers are able to access the conceptually distributive number of the preamble in the L2 if they are highly proficient in the target language.

General discussion

In two experiments, we elicited subject–verb number agreement errors to test whether bilinguals' access to conceptual information during the computation of subject–verb agreement is modulated by the level of language proficiency. The signature finding taken as evidence for the sensitivity to the conceptual number of the subject phrase is a higher rate of subject–verb agreement errors for distributive-referent items than for single-referent items when the head noun and the local noun are mismatched in number. In both experiments, all bilingual speakers showed sensitivity to both grammatical and conceptual number in their L1. However, bilinguals who were relatively, but not highly, proficient in the L2, failed to demonstrate sensitivity to conceptual number. That finding held regardless of whether Spanish or English was the L2. In contrast, the more proficient Spanish–English bilinguals in Experiment 2 produced a pattern of results in L2 that was identical to their performance in L1, with sensitivity to both grammatical and conceptual number. Taken together, these results suggest that adequate cognitive resources are required to maintain the conceptual representation of the subject during the computation of subject–verb number agreement.

As described earlier, Hartsuiker and Barkhuysen (2006) proposed a two-stage model for the production of subject–verb agreement, such that cognitive resources do not influence the first stage of mapping the conceptual information to lemmas but the second stage of checking the presence of conflict in agreement. They further argue that the processing at the first stage is impaired only when cognitive resources are pathologically limited. How can we determine whether the failure of the English–Spanish bilinguals in Experiment 1 and the less proficient Spanish–English bilinguals in Experiment 2 to process the conceptual number in their L2 is due to a problem with the mapping and/or maintenance of conceptual number at the first stage of processing or in checking for conflict at the second stage of processing?

Studies of both comprehension and production in bilinguals have shown that the slower time course associated with the L2 reflects greater vulnerability to interference from competing information (e.g., Van Heuven, Dijkstra and Grainger, 1998; Costa and Caramazza, 1999). To establish the locus of the observed effects for the English–Spanish bilinguals and less proficient Spanish–English bilinguals in the present study, we needed to first determine whether these bilinguals were able to represent the conceptual number accurately. It is possible that the English–Spanish bilinguals and less proficient Spanish–English bilinguals did not appreciate the distributive-referent readings in the L2.

To examine the possibility of the incomplete representation of the conceptually distributive number in bilinguals whose L2 proficiency is not high enough, an independent group of 14 English–Spanish bilinguals who

were sampled from the same population as Experiment 1 performed an offline questionnaire task. The mean self-rated L2 proficiency for this group of English–Spanish bilinguals was 6.4 out of 10, which did not differ from the 6.5 for the English–Spanish bilinguals in Experiment 1 [$t < 1$]. The offline questionnaire included 26 single-referent mismatch (e.g., *la autora de las novelas* "the author of the novels") and 26 distributive-referent mismatch preambles (e.g., *el dibujo de los carteles* "the drawing on the posters") from the sentence completion task. In the offline questionnaire, the participants were asked to judge the preambles according to whether they refer to one thing or more than one thing with "1" being "one thing" and "2" being "more than one thing."[3] The mean for each preamble was computed. The mean for distributive-referent preambles was significantly higher than for single-referent preambles [$t(14) = 7.81$, $p < .001$, 1.62 vs. 1.21 respectively], showing that although the English–Spanish bilinguals were not highly proficient in L2 Spanish, they can indeed distinguish between single-referent and distributive-referent readings in the off-line task.[4] This result suggests that less proficient bilinguals can represent the conceptually distributive number accurately in their L2 when the task demands are low (i.e., offline) and that the mapping and maintenance of the conceptual number can be interfered when the task demand is high (i.e., online).

If the locus of the effect of the conceptual agreement is at the second checking stage, higher proficiency should lead to a greater distributivity effect in the context of a smaller grammatical number effect because more errors can be detected and corrected. In contrast, if the conceptual number is lost at the first stage, the size of the distributive number effect can be independent of the size of the grammatical number effect. We calculated the effect sizes of the sensitivity to the conceptually distributive number (i.e., the interaction between number match and distributivity) and to the grammatical number (i.e., the main effect of number match) in the L1 Spanish and in the L2 English for more proficient Spanish–English bilinguals. It is important to note that more proficient Spanish–English bilinguals were the only group in the present study who showed the sensitivity to grammatical number as well as conceptual number in the online sentence completion task. The effect size of the conceptual number effect was similar in both languages ($\partial\eta^2 = .22$ for L1 Spanish and $\partial\eta^2 = .23$ for L2 English), whereas the effect size of the grammatical number effect was smaller in L1 Spanish ($\partial\eta^2 = .27$) than in L2 English ($\partial\eta^2 = .54$). These results suggest that the size of the distributive number effect is independent of the size of the grammatical number effect and the locus of the effect of the conceptual agreement appears to be at the first stage rather than the second stage.

The critical finding in the present study is that the dissociation between the online and offline tasks resembled that of Hartsuiker et al.'s (1999) aphasic patients who demonstrated sensitivity to the conceptual number during offline processing but not during online processing. As described in the introduction, Hartsuiker and Barkhuysen (2006) assume that an additional memory load or relatively reduced working memory span affects the second control stage of checking the presence of conflict in agreement, and the first stage of mapping and maintaining the conceptual number can be constrained only when cognitive resources are pathologically limited. If greater demands on the computation of subject–verb agreement in less proficient L2 are more analogous to pathologically limited cognitive resources in aphasics, it is possible to assume that the mapping and maintenance of conceptual number can be impaired in online processing. In other words, the limited cognitive resources appear to affect not only the second stage but also the first stage, such that when the bilingual's L2 proficiency is not high enough, the mapping and/or maintenance of conceptual number is constrained and the checking mechanism is impaired. That is, when the task is less demanding, even less proficient bilinguals can represent and maintain conceptual number for processing. However, once the task becomes more demanding, conceptual number is lost at an earlier stage in the process.

In conclusion, the two experiments reported here indicate that the ease of access to the conceptual number of a subject phrase during the process of subject–verb agreement is constrained by level of language proficiency. We have argued that sensitivity to conceptual number depends upon the extent to which cognitive resources are available to the individual speaker and how easily the linguistic information facilitates conceptual access. The pattern of data suggests that there is an intricate interplay between cognitive resources and language processing. A goal in future research will be to examine the precise consequences of L2 proficiency in more detail and to better understand its relation to cognitive resources. It will also be critical to identify factors that may indirectly determine the cognitive resources that are available to perform a given task in a given language. Resolving these issues holds implications for models of language processing, for models of how the bilingual's two languages interact with one another, and for models of how cognitive resources are recruited in planning well-formed utterances.

[3] Participants were also told that multiple items may all be identical so that they understood a token reading.

[4] The ratings in English by English monolinguals and those in Spanish by English–Spanish monolinguals were correlated ($r = .632$, $p < .001$), and this correlation also suggests that the lack of sensitivity to the distributive reading in the online task in a less proficient language is unlikely to result from "misjudged" items.

Appendix

Single-Referent Mismatch Items	
English	Spanish
The author of the novels	La autora de las novelas
The mother of the girls	La madre de las niñas
*The recording of the singers	*La canción de los cantantes
The lawsuit against the owners	La demanda contra los propietarios
*The bread for the sandwiches	*El pan para los bocadillos
The scent of the almond trees	El olor de los almendros
The boy with the crutches	El niño de las muletas
The fisherman with the nets	El pescador con las redes
The offense to the women	La ofensa a las mujeres
*The girl of the photographs	*La chica de las fotografías
The office of the employees	La oficina de los empleados
The suggestion to the directors	La sugerencia a los directores
The witness for the lawyers	El testigo de los abogados
The gift for the babies	El regalo para los bebés
The inventor of the motorcycles	El inventor de las motocicletas
The director of the films	El director de las películas
The house on the hills	La casa de las colinas
The complaint from the students	La queja de las estudiantes
The affliction in the men	La enfermedad de los hombres
The copying machine of the professors	La fotocopiadora de los profesores
The assault on the ministers	El atendado contra los ministros
*The advice of the experts	*El consejo de los expertos
The debate over the drugs	El debate sobre las drogas
The grandfather of the girls	El abuelo de las niñas
*The scholarship for the students	*La beca a las estudiantes
*The lamp on the tables	*La lámpara sobre las mesas
The trap for the mice	La trampa para los ratones
The position of the lawyers	La posición de los licenciados
The van for the tourists	El minibus de los turistas
The stroll by the lakes	El paseo por los lagos
The physician for the ill women	El médico de las enfermas
The teacher of the girls	El maestro de las chicas

Distributive-Referent Mismatch Items	
The skirt in the store windows	La falda en las vitrinas
The label on the bottles	La etiqueta de las botellas
The experience of the presidents	La experiencia de los presidentes
The grade for the students	La nota de los estudiantes
The insurance for the cars	El seguro de los carros
The drawing on the posters	El dibujo de los carteles
*The villain in the films	*El malo de las películas
The postmark on the envelopes	El sello de las cartas
*The box for the bracelets	*La caja de las pulseras
*The advertisement in the magazines	*La propaganda en las revistas
The pollution of the rivers	La contaminación de los ríos
The crack on the planes	La avería de los aviones
*The driver of the buses	*El conductor de los autobuses
The cap on the men	El gorro de los hombres
The problem in the schools	El problema de las escuelas
The color of the flowers	El color de las flores
The door on the houses	La puerta de las casas
The music of the composers	La música de las compositoras
The tie for the clowns	La corbata de los payasos
The illustration in the manuals	La ilustración en los manuales
The uniform for the soldiers	El uniforme de los soldados
The coat on the men	El abrigo de los hombres
The costume for the women	El disfraz de las mujeres
The number on the cards	El número de las tarjetas
The agenda of the professors	La agenda de las profesoras

Distributive-Referent Mismatch Items	
English	Spanish
The computer for the offices	La computadora de las oficinas
The medal for the athletes	La medalla de los atletas
The breed of the dogs	La raza de los perros
The bell at the entrances	El timbre de los portales
The name of the boys	El nombre de los niños
*The mayor of the cities	*El alcalde de las ciudades
*The purse for the girls	*El bolso de las chicas

Note. The items excluded from analyses are indicated with an asterisk.

References

Azuma, T. & Van Orden, G. C. (1997). Why safe is better than fast: The relatedness of a word's meanings affects lexical decision times. *Journal of Memory and Language*, 36, 484–504.

Bialystok, E., Craik, F. I. M., Klein, R. & Viswanathan, M. (2004). Bilingualism, aging and cognitive control: Evidence from the Simon task. *Psychology and Aging*, 19, 290–303.

Bock, K., Eberhard, K. M. & Cutting, J. C. (2004). Producing number agreement: How pronouns equal verbs. *Journal of Memory and Language*, 51, 251–278.

Bock, K. & Miller, C. A. (1991). Broken agreement. *Cognitive Psychology*, 23, 45–93.

Chen, L., Shu, H., Liu, Y., Zhao, J. & Li, P. (2007). ERP signatures of subject–verb agreement in L2 learning. *Bilingualism: Language and Cognition*, 10, 161–174.

Costa, A. & Caramazza, A. (1999). Is lexical selection language specific? Further evidence from Spanish–English bilinguals. *Bilingualism: Language and Cognition*, 2, 231–244.

Eberhard, K. M. (1999). The accessibility of conceptual number to the processes of subject–verb agreement in English. *Journal of Memory and Language*, 41, 560–578.

Finkbeiner, M., Forster, K., Nicol, J. & Nakamura, K. (2004). The role of polysemy in masked semantic and translation priming. *Journal of Memory and Language*, 51, 1–22.

Hartsuiker, R. J. & Barkhuysen, P. N. (2006). Language production and working memory: The case of subject–verb agreement. *Language and Cognitive Processes*, 21, 181–204.

Hartsuiker, R. J., Kolk, H. H. J. & Huinck, W. J. (1999). Agrammatic production of subject–verb agreement: The effect of conceptual number. *Brain and Language*, 69, 119–160.

Hasegawa, M., Carpenter, P. A. & Just, M. A. (2002). An fMRI study of bilingual sentence comprehension and workload. *NeuroImage*, 15, 647–660.

Hoshino, N., Dussias, P. E. & Kroll, J. F. (in preparation). Producing subject–verb agreement: Does L1 syntax influence L2 performance? Unpublished manuscript, Pennsylvania State University, University Park.

Jiang, N. (2004). Morphological insensitivity in second language processing. *Applied Psycholinguistics*, 25, 603–634.

Kotz, S. A. & Elston-Güttler, K. (2004). The role of proficiency on processing categorical and associative information in the L2 as revealed by reaction times and event-related potentials. *Journal of Neurolinguistics*, 17, 215–235.

Michael, E. B. & Gollan, T. H. (2005). Being and becoming bilingual: Individual differences and consequences for language production. In J. F. Kroll & A. M. B. De Groot (eds.), *Handbook of bilingualism: Psycholinguistic approaches,* pp. 389–410. Oxford: Oxford University Press.

Miyake, A. & Friedman, N. F. (1998). Individual differences in second language proficiency: Working memory as language aptitude. In A. F. Healy & L. E. Bourne (eds.), *Foreign language learning: Psycholinguistic studies on training and retention*, pp. 339–364. Mahwah, NJ: Erlbaum.

Nicol, J. & Greth, D. (2003). Production of subject–verb agreement in Spanish as a second language. *Experimental Psychology*, 50, 196–203.

Nicol, J., Teller, M. & Greth, D. (2001). Production of verb agreement in monolingual, bilingual and second-language speakers. In J. Nicol (ed.), *One mind, two languages: Bilingual language processing*, pp. 117–133. Malden, MA: Blackwell.

Segalowitz, N. & Hulstijn, J. (2005). Automaticity in bilingualism and second language learning. In J. F. Kroll & A. M. B. De Groot (eds.), *Handbook of bilingualism: Psycholinguistic approaches*, pp. 371–388. Oxford: Oxford University Press.

Silverberg, S. & Samuel, A. (2004). The effect of age of second language acquisition on the representation and processing of second language words. *Journal of Memory and Language*, 51, 381–398.

Van Hell, J. & Mensies, M. (2004). Subject–verb agreement in beginning L2 learners and fluent bilinguals. Paper presented at the 45th Annual Meeting of the Psychonomic Society of America, Minneapolis, November 18–21.

Van Heuven, W. J. B., Dijkstra, A. & Grainger, J. (1998). Orthographic neighborhood effects in bilingual word recognition. *Journal of Memory and Language*, 39, 458–483.

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., Butterworth, B. & Semenza, C. (1995). Constructing subject–verb agreement in speech: The role of semantic and morphological factors. *Journal of Memory and Language*, 34, 186–215.

Vigliocco, G. & Hartsuiker, R. J. (2002). The interplay of meaning, sound, and syntax in sentence production. *Psychological Bulletin*, 128, 442–472.

Vigliocco, G., Hartsuiker, R. J., Jarema, G. & Kolk, H. H. J. (1996). One or more labels on the bottles? Notional concord in Dutch and French. *Language and Cognitive Processes*, 11, 407–442.

Vigliocco, G. & Zilli, T. (1999). Syntactic accuracy in sentence production: The case of gender disagreement in Italian language-impaired and unimpaired speakers. *Journal of Psycholinguistic Research*, 28, 623–648