Applied Psycholinguistics **36** (2015), 799–830
doi:10.1017/S0142716413000519

Effects of reading speed on second-language sentence processing

EDITH KAAN
University of Florida

JOCELYN C. BALLANTYNE and FRANK WIJNEN
Utrecht University

Received: July 30, 2012 Accepted for publication: May 31, 2013

ADDRESS FOR CORRESPONDENCE
Edith Kaan, Department of Linguistics, University of Florida, Box 115454, Gainesville, FL 32611.
E-mail: kaan@ufl.edu

ABSTRACT
To test the effects of reading speed on second-language (L2) sentence processing and the potential influence of conflicting native language word order, we compared advanced L2 learners of English with native English speakers on a self-paced reading task. L2 learners read faster overall than native English speakers. When differences in reading speed were controlled for, L2 learners were as sensitive to grammaticality manipulations as native English speakers. On-line reading times did not reflect any effect of cross-language conflict in the learners. Results from an end-of-sentence verification task showed a stronger bias toward a subject–object order in the cross-language conflict conditions in speed-matched L2 learners but not in L2 learners reading faster than native speakers. Results are compatible with hypothesized differences in resource allocation between L2 and native language processing.

Several accounts have been proposed regarding the mechanisms underlying differences in sentence comprehension between first-language (L1) native speakers and late second-language (L2) learners. According to the shallow structure hypothesis (Clahsen & Felser, 2006), L2 learners do not build detailed syntactic structures but instead rely on lexical semantics, the context, general heuristics, and local relations. An alternative view is that the main differences between L2 learners and native speakers lie in the allocation of processing resources (e.g., Hopp, 2006; McDonald, 2006). Advanced L2 learners are able to achieve the same level of knowledge and are able to build the same, detailed syntactic representations as native speakers, but they may need to devote more resources to accessing the L2 information and applying parsing routines than do native speakers, rendering L2 processing more effortful, less automatic, and, therefore, slower (Dekydtspotter, Schwartz, & Sprouse, 2006; Segalowitz & Hulstijn, 2005; Sorace, 2011) or

© Cambridge University Press 2014 0142-7164/14 $15.00

https://doi.org/10.1017/S0142716413000519 Published online by Cambridge University Press

less predictive (Havik, Roberts, Van Hout, Schreuder, & Haverkort, 2009; Kaan, Dallas, & Wijnen, 2010; Lew-Williams & Fernald, 2010). As a result, on-line effects observed in L2 learners may be delayed or smaller than those found in native speakers. For instance, the P600 effect in response to syntactic violations may appear later and smaller in L2 speakers (e.g., Hahne, 2001; Sabourin & Stowe, 2008), or gap-filling effects may appear later than in native speakers (e.g., Dallas, 2008; Williams, Möbius, & Kim, 2001). L2 speakers may break down and show processing patterns that are different from those of a native speaker of the language when pressed for time (Hopp, 2010) or when the task requires much processing effort, as when information needs to be combined across multiple domains to form a correct representation (Hopp, 2009; Sorace, 2011; Wilson, 2009).

Part of the discussion on how L1 and L2 processing differ concerns processing speed. Under a processing deficit account, it is expected that L2 learners access and integrate lexical and other information less automatically than native speakers, which may have repercussions at the level of sentence processing. A critical question is whether differences in reading patterns of L2 learners and native speakers can be attributed to reading speed differences. Many studies have observed slower reading times for L2 versus native speakers (for recent studies, see Coughlin & Tremblay, 2012; Roberts & Felser, 2011), although a few studies have reported faster reading times for learners versus native speakers (Felser, Roberts, Gross, & Marinis, 2003). It remains unclear, however, to what extent reading speed is related to qualitative processing differences between native and nonnative speakers.

Clahsen and Felser (2006) claim that the on-line reading pattern of L2 learners was qualitatively different from that of native speakers, despite considerable variation in the reading speed of L2 learners, some of whom read faster than native speakers. This supports the shallow structure hypothesis: differences between L2 and native speakers are not due to processing resources but to the inability of L2 learners to build detailed syntactic structures. In contrast, other studies have reported strong similarities between native speakers and L2 learners, despite differences in reading speed. For instance, Hoover and Dwivedi (1998) explicitly compared fast and slow reading L2 learners of French and observed that both groups were indistinguishable from native French readers in the on-line reading time patterns. The slow reading learners, however, showed a larger end-of-sentence wrap up effect when compared to fast reading learners and native speakers. Hoover and Dwivedi (1998) attributed this pattern to the slow readers being less automatic and needing to apply end-of-sentence compensation strategies. These strategies were not unique to L2 learners because slower native speakers also showed an increase in end-of-sentence wrap up effects in some constructions.

Some evidence that differences between L2 learners and native speakers are at least partially due to differences in reading speed comes from a study by Roberts and Felser (2011). In their study, L2 readers read syntactically ambiguous sentences more slowly than native speakers and were more sensitive to semantic manipulations; however, reading speed affected the pattern in native speakers as well as learners: slower reading native speakers showed on-line effects similar to those of L2 readers, especially faster reading L2 learners. Even in native speakers, then, differences in reading speed may affect the on-line reading pattern.

https://doi.org/10.1017/S0142716413000519 Published online by Cambridge University Press

It is crucial to control for differences in reading speed when comparing L1 and L2 speakers in on-line sentence reading tasks. A smaller and later effect of the experimental manipulation in one group could be ascribed to less automaticity or incomplete knowledge, but it might also be due to more automaticity, enabling faster reading, which could also lead to smaller and later effects. The present study therefore sought to investigate to what extent differences in syntactic processing between L2 and L1 speakers could be attributed to differences in reading speed.

The second aim of the present experiment was to investigate whether L2 learners are sensitive to cross-language conflicts during sentence processing. In a cross-language conflict, a particular sequence of words corresponds to a particular construction in the L2 (e.g., an object relative) but to a different construction in the L1 (e.g., a subject relative) when translated word-by-word. There is some evidence from bilinguals, including late learners, that the processing of one language is affected by the coactivation of the other language, even in situations in which only one language is relevant. Most of the evidence comes from research on lexical processing (e.g., Dijkstra, Grainger, & van Heuven, 1999; Kroll, Bobb, & Wodniecka, 2006; Marian & Spivey, 2003; Schwartz & Kroll, 2006; Van Hell & Dijkstra, 2002; Van Heuven, Dijkstra, & Grainger, 1998). Whether one language can affect the other at the level of sentence processing is more controversial. Some studies have failed to find effects of L1 properties or strategies on L2 sentence processing (e.g., Felser et al., 2003; Jacob, 2009, experiments 5–7; Juffs, 2004; Marinis, Roberts, Felser, & Clahsen, 2005; Papadopoulou & Clahsen, 2003; Williams et al., 2001). Absence of an effect of the first language on the L2 sentence processing is in line with the shallow structure hypothesis (Clahsen & Felser, 2006). According to this view, no detailed syntactic representations are built when processing the L2, so the nature of the L1 should not affect L2 processing. Other studies, however, do find effects of one language on the other in sentence processing (e.g., Dussias & Cramer Scaltz, 2008; Dussias & Sagarra, 2007; Frenck-Mestre, 2005; Frenck-Mestre & Pynte, 1997; Hopp, 2009; Jacob, 2009; Juffs, 2005).

Many studies reporting cross-language effects have been conducted within the framework of the competition model (Kilborn, 1989; Kilborn & Cooreman, 1987; MacWhinney & Bates, 1989; McDonald, 1987; Tokowicz & MacWhinney, 2005). For instance, McDonald (1987) investigated native Dutch speakers, native English speakers, Dutch learners of English, and English learners of Dutch. In a series of experiments, participants were presented with phrases containing at least a verb and two noun phrases in which the word order, animacy, case, or number agreement between the nouns and the verb was varied. When asked to determine who did what to whom, the native Dutch speakers relied more on number agreement and case; the native English relied more on word order. This was also the case when they processed sentences in their L2. However, as they became more proficient in the second language, L2 learners relied more on cues that were more important for native speakers of the L2. Although native language effects on L2 processing are problematic for the shallow structure hypothesis, processing accounts of the difference between L1 and L2 typically do not exclude L1 effects on L2 processing. According to these views, L2 learners need to suppress or ignore their other (typically dominant) language while processing their second language.

https://doi.org/10.1017/S0142716413000519 Published online by Cambridge University Press

This, along with generally slower activation and less automatic processing of L2 information, may lead to the availability of even fewer resources for processing the L2 (Dekydtspotter et al., 2006; Dussias, 2003; Sorace, 2011), or to an increased susceptibility to native language information and parsing routines when processing the L2 (Costa & Caramazza, 1999; MacWhinney & Bates, 1989; Sabourin & Stowe, 2008; Tokowicz & MacWhinney, 2005).

THE CURRENT STUDY

The aim of our study was to investigate whether differences in syntactic processing between L2 learners and native speakers could be attributed to (a) differences in reading speed and (b) the presence of a potential cross-language conflict. We conducted a self-paced reading study in which L1 speakers and L2 learners read sentences phrase-by-phrase. Sentences were followed by a statement verification task to probe (off-line) interpretation. We focused on word order differences in languages that are typologically very close (Dutch and English). In particular, we tested native English speakers and Dutch native speakers who were advanced learners of English on the processing of subject–verb agreement in English object relatives, a construction that allows a potential cross-linguistic ambiguity. In English object relatives, as in Example (1), the object (who) comes before the subject. The finite verb (the auxiliary *have* in (1a)) agrees with the relative clause subject. The sentence in (1b) is ungrammatical, since the verb agrees with the first noun phrase (the object) rather than the second noun phrase (the subject).

(1) a. Mark/ may know/ the instructor/ who/ the students/ **have**/ avoided/ since/ last/ semester.
b. Mark/ may know/ the instructor/ who/ the students/ ***has**/ avoided/ since/ last/ semester.

The sentence in (1b) however corresponds to a grammatical subject relative in Dutch, compare Example (2).

(2) Mark kent misschien de leraar die de studenten heeft vermeden sinds het afgelopen semester.
Mark knows maybe the instructor who the students has avoided since the last semester.
"Mark may know the instructor who has avoided the students since last semester"

In both object and subject relatives in Dutch, the verb follows the two noun phrases. The sequence of a noun phrase followed by a relative pronoun and a second noun phrase is therefore potentially ambiguous between a subject–object and an object–subject order. In Example (2), as in the word-by-word translation of (1a), the finite verb resolves this ambiguity. The verb agrees only with the head of the relative clause in Example (2), and only with the second noun phrase in (1a), indicating that these noun phrases are the subjects. In contrast, in Example (3), where both of the noun phrases and the verb are singular, it is not clear who has avoided whom. The sentence is grammatical, but the relative clause is ambiguous between a subject–object and object–subject order:

https://doi.org/10.1017/S0142716413000519 Published online by Cambridge University Press

(3) Mark kent misschien de leraar die de student heeft vermeden sinds het afgelopen semester.
Mark knows maybe the instructor who the student has avoided since the last semester
"Mark may know the instructor who has avoided the student / who the student has avoided since last semester"

Studies on Dutch and German, which have similar, potentially ambiguous structures, have shown that native readers prefer a subject before object reading, even before the disambiguating information is encountered (e.g., Bader & Meng, 1999; Frazier, 1978; Frazier & Flores d'Arcais, 1989; Havik et al., 2009; Hopp, 2006; Kaan, 2001; Schriefers, Friederici, & Kühn, 1995). Longer processing times are typically seen at the finite verb when it disambiguates for an object–subject interpretation rather than a subject–object one. The Dutch preference for a subject–object interpretation may affect how Dutch L2 learners of English process an English object relative, in which object–subject is the only interpretation. We expected that Dutch L2 learners of English would not notice the ungrammaticality in (1b) if they activate the default Dutch word order, and that they would not slow down at the critical auxiliary, at least, not to the same extent as native English speakers. Furthermore, Dutch L2 learners of English were expected to show longer reading times at the auxiliary in grammatical constructions such as (1a), since the verb disambiguates for the object–subject order, which is less preferred in Dutch. Thus, if L1 word order preference can affect L2 processing, we expected a reduced grammaticality effect when comparing (1b) to (1a) for the Dutch L2 speakers of English, as compared to natives. In contrast, the L2 speakers were expected not to differ from native speakers in the on-line processing of (4b) versus (4a). In (4a), both noun phrases and the verb are singular. The word order and agreement are grammatical both in Dutch and English, although a Dutch sentence would be ambiguous between a subject–object reading and an object–subject one (cf. Example (3)). In (4b) both noun phrases are singular, but the plural verb does not agree with its subject, rendering the sentence ungrammatical in English. In Dutch, too, a finite verb agrees with its subject in number. A word-by-word translation of (4b) in Dutch would therefore also be ungrammatical, under either the subject–object or the object–subject interpretation.

(4) a. Mark /may know/ the instructor/who/the student/ **has**/ avoided/since/last/semester.
b. Mark /may know/ the instructor/who/the student/*__have__/avoided/since/last/semester.

We therefore expected on-line reading times to show a three-way interaction between grammaticality, plurality of the subject, and language group: the L2 learners were expected to show a smaller effect of grammaticality starting at the critical auxiliary for the plural subject conditions, (1b) versus (1a), than for the singular subject conditions, (4b) versus (4a). The native English speakers were expected to show an equally large effect of grammaticality in the singular and plural subject conditions.

In addition, we probed the thematic interpretation of one-third of the relative clauses with an end-of-sentence statement verification task. We expected the Dutch L2 learners of English to assign a subject–object interpretation to the clauses

https://doi.org/10.1017/S0142716413000519 Published online by Cambridge University Press

more often than native English speakers would. This was especially expected for ungrammatical sentences such as (1b), for which a word-by-word translation would yield a grammatical subject–object relative clause in Dutch, hence creating a potential conflict with the English object–subject order. The L2 learners were also expected to assign subject–object interpretations more frequently than the native English group to (4a), for which a word-by-word translation can correspond to a grammatical subject–object reading in Dutch, and to (4b), for which the subject of a word-by-word translation would be unclear (Kilborn & Cooreman, 1987; McDonald, 1987). We also expected the native English speakers to assign the object–subject reading less systematically in the singular subject conditions (4a and 4b) than in the plural subject conditions (1a and 1b): In the singular subject conditions both noun phrases are singular and hence easily confusable (Gordon, Hendrick, & Johnson, 2001).

In order to investigate the effect of reading speed differences between our L2 learners and native speakers, we first conducted an analysis without controlling for reading speed. Then we controlled for reading speed by (a) comparing L2 and native English groups that were matched in reading speed and (b) including reading speed as a continuous measure in the analysis. This analysis allowed us to investigate to what extent differences observed between the L2 and native groups could be attributed to differences in speed.

Methods

Participants. The study included two groups of participants: a native English group, consisting of 39 native speakers of American English, and an L2 group, consisting of 71 native speakers of Dutch who learned English as a second language. All participants were raised monolingually, reported having good (corrected) eyesight, and indicated they did not have dyslexia or other reading or learning related problems. All participants gave written informed consent according to local Institutional Review Board protocol. Participants in the native English group were recruited and run at the University of Florida, US. Participants in this group had taken one or more foreign languages at school after the age of 11, the most popular being Spanish, but did not rate themselves as very fluent in their second language. Six additional participants from the native English group were excluded from analysis because they were early bilinguals or had knowledge of German (which could have affected their processing of our critical items). The L2 group consisted of native Dutch speakers who started learning English in a formal setting around the age of 10, although most had been exposed to English through television and other media before that age. Most used English on a daily basis and were either English majors at Utrecht University, The Netherlands, or were enrolled at University College Utrecht, an English-immersion college in the Netherlands. Only 21% of the L2 participants had lived in an English-speaking country (the average stay in an English-speaking country for the entire L2 group was 1.2 months, range = 0–12 months). Most listed several languages learned in addition to English, the most popular being French and German, but all indicated that English was the foreign language that they started learning first and in which they were most fluent. Two additional participants from the

https://doi.org/10.1017/S0142716413000519 Published online by Cambridge University Press

L2 group were excluded: one because she was an early bilingual, the other because she had excessively long response times in the self-paced reading task (exceeding the mean plus twice the standard deviation of the other L2 speakers tested).

The participants' language background was assessed using the Language Experience and Proficiency Questionnaire (Marian, Blumenfield, & Kaushanskaya, 2007), which was slightly modified to include questions about the parents' education (Hollingshead, 1975) and self-rating of their skills in writing, in addition to reading, listening, and speaking. English proficiency was assessed by the C-test (Grotjahn, 1987; Keijzer, 2007), a word naming task, and the Peabody Picture Vocabulary Task (PPVT; Dunn & Dunn, 2007). The C-test was taken from Keijzer (2007). In this task, participants read five English paragraphs, each including 20 words that were truncated after the first few letters. Participants were given 5 min per paragraph to complete the truncated words. One paragraph was omitted from analysis, since performance on it was much lower than on the other four paragraphs in both the native English and the L2 group. For the naming task, 108 black-and-white drawings were selected (Snodgrass & Vanderwart, 1980) avoiding items with Dutch–English cognates. Two lists of 54 items were constructed such that the two lists were matched in the log and raw word form frequency of spoken use, number of syllables, and number of phonemes of the target words, for English as well as Dutch (Baayen, Piepenbrock, & Gulikers, 1995). The pictures were presented sequentially on a computer screen, and participants were asked to name each picture as quickly, clearly, and accurately as they could in the language under investigation. The order of the two lists was counterbalanced. Native English speakers responded to both lists in English; L2 speakers responded in English to the list presented first and in Dutch to the second list presented later in the study. Accuracy of the responses was determined by consensus of five native speakers of English. We consider below only the proportion of correct responses from the initially presented list (English for all participants). The third proficiency measure, the English PPVT, Version 4a (Dunn & Dunn, 2007), was administered to all participants. This task involves matching an orally presented word to one of four pictures. Participants' scores were converted to standardized scores on this task. To check whether our participants were familiar with the rules of English object relatives, we included a paper-and-pencil repair task at the end of the reading experiment. Participants saw 14 sentences drawn from the experiment (one from each of the four experimental object-relative conditions and 10 distractors; the same items were used for all participants) and were asked to correct any grammatical errors they spotted; they could take as long as they needed.

A comparison of the two participant groups on demographics, self-rating of proficiency, and the English proficiency tests is given in Table 1. The groups did not differ in age or parents' education. The L2 group, however, had learned more foreign languages and had started learning their first foreign language at a younger age than the native English group. Furthermore, the L2 group indicated higher fluency in their L2 English (average over self-rating for reading, writing, listening, and speaking) than the native English participants did for their strongest foreign language. The L2 group also rated themselves as being less fluent in their

https://doi.org/10.1017/S0142716413000519 Published online by Cambridge University Press

Table 1. *Participant demographics and proficiency results*

	Native English	L2
Gender	32 F, 7 M	58 F, 13 M
Age	19.9 (18–28)	19.4 (18–26)
Parents' education	5.6 (1.1)	5.5 (1.2)
Age first L2 learned*	14.2 (11–19)	10.4 (6–12)
No. of foreign lang.*	1.7 (0.8)	4.6 (1.2)
Proficiency in L1, mean self-rating (10 = fluent)*	9.8 (0.5)	9.1 (0.6)
Proficiency in best foreign lang. mean self-rating (10 = fluent)*	4.1 (2.4)	7.8 (0.6)
English proficiency, mean self-rating (10 = fluent)*	9.8 (0.5)	7.8 (0.6)
English naming accuracy*	0.94 (0.04)	0.77 (0.11)
English PPVT standardized score*	111.4 (9.1)	94.7 (10.1)
English C-test (max. 80)*	69.2 (4.1)	65.5 (6.2)
No. of partic. correctly repair. no. in rel. clause		
Plural subject[a]	35/39 (90%)	62/71 (87%)
Singular subject[a]	35/39 (90%)	63/71 (89%)

Note: Range or standard deviation is in parentheses. L2, second language; L1, first language; PPVT, Peabody Picture Vocabulary Test.
[a]Off-line repair task following the self-paced reading task.
*The difference between the two groups is significant, $p < .001$.

native language (Dutch) than the native English rated themselves in their native English. In the analysis below, we determined English proficiency on the basis of the scores obtained in the three English proficiency tasks. L2 learners scored significantly worse than native English speakers on all three tasks. Since there was a strong correlation among all three English proficiency tasks in the L2 learners (rs = .59–.63, ps < .001), we converted the scores for each task to z-scores, and calculated the mean across the tasks. This composite score was calculated separately for the native English and L2 groups and was used in the analysis of the sentence processing data. The last two rows of Table 1 give the numbers of participants who correctly repaired the ungrammatical object relatives with the plural (1b) and singular (4b) relative clause subjects; χ^2 tests yielded no significant differences between the groups (ps > .7). All participants but one, a native English speaker, considered the two grammatical relative clause examples grammatical. The L2 participants performed at the same level as the native English, which suggests that the L2 speakers were aware of the syntactic rules of object-relative formation and subject–verb agreement in English.

Materials. Experimental items were English object relatives with the structure illustrated in Examples (1) and (4). The slashes in those examples indicate the presentation frames. To avoid attraction effects of a plural relative clause head on the verb, as observed by Wagers, Lau, and Phillips (2009), the head of the

https://doi.org/10.1017/S0142716413000519 Published online by Cambridge University Press

relative clause in our study (*instructor*, in the example) was always singular. In the grammatical conditions, (1a) and (4a), the auxiliary in the relative clause agreed with the subject of the relative clause. In the ungrammatical conditions, (1b) and (4b), the auxiliary did not match the subject in number.[1] The critical auxiliary was always in the sixth frame from the start of the sentence and was followed by four to six frames. The critical condition is illustrated in (1b). This condition is grammatical when translated word-by-word into Dutch, and hence it was expected to be susceptible to L1 effects in our Dutch–English participants. The conditions in Example (4) contained two singular noun phrases and served as control conditions. Condition (4b) is ungrammatical in both Dutch and English; Condition (4a) is grammatical in both Dutch and English (although ambiguous in a Dutch word-by-word translation).

We pretested 60 sets of items in two paper-and-pencil questionnaire studies to ensure that the propositions in the relative clauses were thematically reversible, in order to avoid a semantic bias toward interpreting a particular noun phrase as the subject. A group of 36 native speakers of English (18–33 years old; 29 women, 7 men), none of whom participated in the main study, rated sentences for plausibility on a scale from 1 (*highly implausible*) to 7 (*highly plausible*). The first questionnaire study was completed by 18 participants and contained two versions of the experimental items: one with the original object-relative structure, and one in which the word order was changed to a subject-relative structure (e.g., *Mark may know the instructor who has avoided the students since last semester*). In the second questionnaire (completed by another 18 participants), the relative clause was turned into a main clause. Again, two orders were used, one corresponding to the meaning of the object relative (e.g., *The students avoided the instructor*) and one in which the noun phrases were reversed (e.g., *The instructor avoided the students*). Each participant saw half of the experimental items in one version and half in the other, counterbalanced across two lists per study, embedded in plausible and implausible distractor items. The mean plausibility rating was calculated per item. A set of 40 relative clause items was selected for the main study such that (a) the mean rating for each of the four versions seen in the questionnaires was greater than 4.1, (b) the average rating, collapsed over the four versions, was larger than 5.1, and (c) the difference in the average rating between two orders in each study was smaller than 1. Results of *t* tests comparing the ratings between the two orders were not significant, $t\ (39) = 0.53$, $p = .63$. Some sentences contained English–Dutch cognates, which may have heightened L1 activation in our L2 participants. However, since the materials were Latin-squared, potential confounding effects were equally distributed across conditions.

Distractor items in the self-paced reading task consisted of 120 items yielding a total of 160 sentences (60 ungrammatical, 37.5%).[2] Distractor items included an array of constructions, among which were grammatical subject relatives (*The man who fixed my car is a close friend of my father's*), word order violations (**When John in town was last week, we met with him almost every day*), and ungrammatical uses of gerunds or infinitives (e.g., **All of a sudden the car stopped to run in the middle of the road*). Each sentence was followed by a true/false statement. In half of the trials, the correct answer was true, in the other half false. This was counterbalanced by condition. True/false statements mainly probed time, place,

https://doi.org/10.1017/S0142716413000519 Published online by Cambridge University Press

and objects. However, three sentences of each of the experimental conditions (the same items for each list) were followed by questions probing thematic relations, for example, *The instructor avoided the students* was used as a true/false statement following Example (1). Half of these statements were correct, half were incorrect. We kept the number of thematic true/false statements low so as to avoid drawing participants' attention to the experimental manipulations. A list of the experimental items and true/false statements is provided as online Supplementary Materials to this paper (http://journals.cambridge.org/aps).

The experimental items were assigned to four lists according to a Latin square design. Each list contained 10 different sentences of each experimental condition, and each sentence appeared in a different version on another list. Distractors and experimental items were automatically randomized in a different order for each participant, such that no more than two sentences of the same condition could appear in sequence.

Procedure. Sentences were presented in a self-paced noncumulative moving-window reading paradigm, with white letters on black background, and 11 point Courier font. At the beginning of each trial, the contour of the sentence was presented, with the letters of the words replaced with dashes. After the participant pressed a button on a button box, the first word appeared, with the remainder of the sentence still in dashes. At the next button press, the first word was replaced with dashes and the second word appeared, and so on. None of the sentences had any line breaks. After the last word, the next button press displayed the true/false statement, with the letters T and F to the left and right, matching the position of the corresponding response button on the button box. After a response was given, the next trial started. Participants were instructed to read the sentence at a normal pace and to respond to the true/false statement as quickly and accurately as they could. A 30-s break was automatically inserted in the middle of the experiment. Five practice sentences were given before the experiment started (all grammatical, none resembling the experimental items). Automatic feedback in the form of the words "Correct!" and "Incorrect" appeared on the screen for 1500 ms after the participants' response to the true/false statements during the practice block. No feedback was provided during the actual experiment.

GENERAL PROCEDURE. The native English participants were tested by a native speaker of English. The L2 group was tested by a fluent English–Dutch bilingual, who spoke English with the participants during the first five tasks. The first four tasks were the same for both groups: (a) language questionnaire, (b) English C-test, (c) English self-paced reading study followed by the off-line debriefing and repair test, and (d) English picture naming task. The native English group then continued with two cognitive control tasks (Trail Making Task, Reitan & Wolfson, 1995; Stroop Task, Stroop, 1935), the PPVT, and a third cognitive control task (Attention Network Task; Fan, McCandliss, Sommer, Raz, & Posner, 2002) followed by general debriefing. The fifth task for the L2 group was the English PPVT. The remaining tasks for the L2 group were the cognitive control tasks, Dutch proficiency tasks (including the picture naming task),[3] and general debriefing and were administered in Dutch. The study took about 2 hr for the native English participants, and

https://doi.org/10.1017/S0142716413000519 Published online by Cambridge University Press

3 hr for the L2 participants. The cognitive control tasks were included to test previous observations that bilinguals (including late bilinguals) have better cognitive control than monolinguals (Bialystok, 2009; Costa, Hernandez, & Sebastián-Gallés, 2008; Linck, Hoshino, & Kroll, 2008), and that bilinguals with better cognitive control are better at language selection and inhibition (Linck et al., 2008). Since our groups did not differ in cognitive control, and cognitive control did not affect language proficiency, or processing in our study, we will not discuss cognitive control further.

ANALYSIS AND RESULTS

Figure 1 displays the raw reading times for the L2 and native groups starting from the word before the critical word, collapsed over the plurality of the relative clause subject. The effect of grammaticality appears to come later and to be less pronounced in the L2 compared with the native group. However, the L2 group read faster overall than the native group. It can therefore not be excluded that the smaller and later effects observed for the L2 learners are due to their reading faster, regardless of their language background (see e.g., Coughlin & Tremblay, 2012; Roberts & Felser, 2011). To investigate the effect of reading speed, we conducted two types of analysis for both the self-paced reading times and the type of responses given in the end-of-sentence statement verification task. In the first analysis, we did not include any measure of reading speed. In the second, we controlled for differences in reading times between the groups in two ways. We first obtained an independent measure of reading speed by calculating each participant's average reading speed on nonfinal words of grammatical distractor items in the study (about 524 data points per participant). A figure displaying these reading times is given in the online Supplementary Materials to this paper (http://journals.cambridge.org/aps). We then controlled for differences in reading speed thus calculated first by selecting a subset of 37 L2 learners matched with 37 native English speakers on reading speed as defined above and then including reading speed as a continuous measure in the analysis of these speed-matched groups. As was the case with the total sample of L2 learners, the speed-matched L2 group did not differ from the matched native English group in terms of age, parental education, or cognitive control measures and performance in the repair task, but they knew more foreign languages and scored lower on the English proficiency tasks than the native English speakers. To check whether the speed-matched L2 and native English groups were similar in their overall reading speed on the critical conditions and were matched in this respect as well, we conducted an analysis of the raw reading times from the first to the ninth frame of the relative clauses collapsed over position and condition. No main effect of group was found, $t\,(72) = 0.26$, $p = .79$; mean raw reading times for matched native English group was 454 ms ($SD = 218$ ms) and mean for the matched L2 group was 449 ms ($SD = 223$ ms). We first discuss the analyses and results of the self-paced reading data. Then, we present the response type results for the end-of-sentence true/false statement verification task.

https://doi.org/10.1017/S0142716413000519 Published online by Cambridge University Press

https://doi.org/10.1017/S0142716413000519 Published online by Cambridge University Press

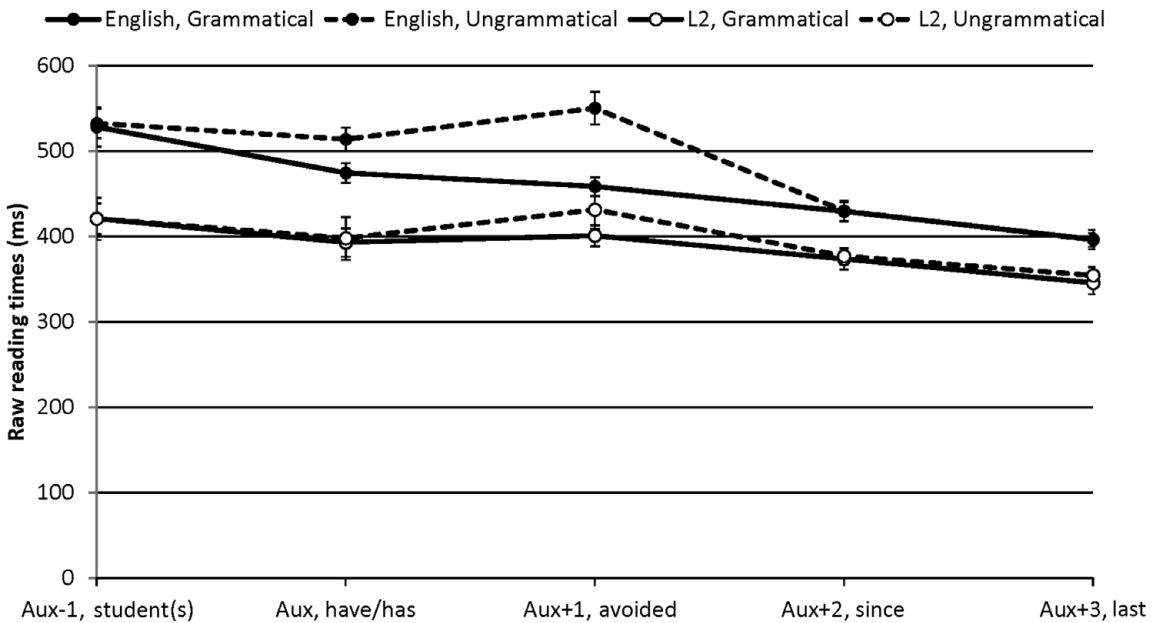


Figure 1. Raw reading times for the native English and second-language groups for the grammatical and ungrammatical conditions, collapsed over plurality of the relative clause subject. Error bars depict the standard error of the mean.

Self-paced reading data

Analysis. In the analysis of the self-paced reading times, response times longer than 4000 ms and shorter than 60 ms were omitted from analysis. For each participant, data exceeding that participant's overall mean + 3 *SD* for that word position (calculated over experimental items as well as distractors) were trimmed to the mean + 3 *SD*. These cutoff procedures affected 3.3% of the data for the word positions to be analyzed in the relative clause conditions (native English group = 2.9%, L2 group = 3.6%). Analysis of the self-paced reading data was conducted on residual reading times to control for differences in word length.[4] We followed the procedure developed by Florian Jaeger (hlplab.wordpress.com/2008/01/23/modeling-self-paced-reading-data-effects-of-word-length-word-position-spill-over-etc/). Residual reading times were calculated on the basis of all trials in the experiment. Reading times were estimated on the basis of a linear model with the following predictors: word length in number of characters, the log of the position (order) of the trial in the experiment, the nature of the item (experimental or filler), and the cubic spline of the position of the word in the sentence. Participant was included as a random effect. All factors contributed significantly to the reading times. Residual reading times were computed by calculating the difference between the reading times observed and the reading times expected on the basis of the regression model above.

We will report the analysis of the self-paced reading times in the experimental conditions for the critical auxiliary and the following word position.[5] For both word positions, a linear mixed-effects model was constructed (lme4 packet; Bates, Maechler, & Bolker, 2011) in R (R Development Core Team, 2007) with the residual reading time as the dependent variable. Fixed effects were the factors grammaticality, the plurality of the relative clause subject, language group, and the two- and three-way interactions between these factors. In the analyses on the speed-matched groups, reading speed (average reading times on grammatical distractor items), and its interactions with other fixed effects factors, were included. In all analyses, the reaction time from three previous regions (spillover) was included as fixed effects. These spillover factors were included to control for differences observed at a preceding word affecting the word position analyzed. Factors were centered around their means to reduce collinearity. A maximum random effects structure was used, with the factors grammaticality, plurality and their interactions nested within participants and items (Barr, Levy, Scheepers, & Tily, 2013). Models were fitted using the restricted maximum likelihood estimation. Since most of the models included random slopes, we could not use Markov chain Monte Carlo sampling to obtain p values. Given that we had more than 2,900 observations for the dependent measure (more than 1,400 when analyzing the language groups separately), we calculated the p value of the effects under the assumption that the t distribution converged with the z distribution (Baayen, 2008). When an interaction involving the factors grammaticality and language group was significant, separate models were estimated for the L2 and native English data. The separate models for the L2 group also included the composite English proficiency score (continuous measure), and its interactions with the factors grammaticality and plurality as fixed effects. The composite English proficiency score did not correlate with reading

https://doi.org/10.1017/S0142716413000519 Published online by Cambridge University Press

speed ($r = –.18$, $p = .13$). We therefore also included English proficiency when separately testing the speed-matched L2 group. The details of the models and the complete results of each analysis are provided as online Supplementary Materials to this paper (http://journals.cambridge.org/aps).

Results

Not controlling for differences in reading speed. Residual reading times for the critical word and surrounding word positions are given in Figure 2 and Figure 3 for the complete native English and L2 group, respectively.

The critical auxiliary and the following word position were read significantly slower in the ungrammatical than in the grammatical condition (critical auxiliary: $\beta = 16.98$, $SE = 6.36$, $t = 2.67$, $p < .01$; following position: $\beta = 49.48$, $SE = 7.37$, $t = 6.71$, $p < .0001$). Language group interacted with grammaticality at both word positions (critical auxiliary: $\beta = 34.33$, $SE = 12.24$, $t = 2.81$, $p < .01$; following position: $\beta = 57.04$, $SE = 15.29$, $t = 3.73$, $p < .001$). Separate analyses for the two language groups at the critical auxiliary and the following position showed an effect of grammaticality at the auxiliary only for the native English participants ($\beta = 38.88$, $SE = 14.36$, $t = 2.71$, $p < .01$; L2: $\beta = 4.82$, $SE = 6.30$, $t = 0.77$, $p = 44$), and at the position following the auxiliary for both groups (native English: $\beta = 85.79$, $SE = 16.57$, $t = 5.17$, $p < .0001$; L2: $\beta = 29.27$, $SE = 8.54$, $t = 3.43$, $p < .001$). We found no interaction of plurality, grammaticality and language at either the critical auxiliary or the following position (ps $> .50$). English proficiency, which was included as a continuous factor in the separate analyses of the L2 data, also had no effect (ps $> .26$).

To further assess the effect of language, we compared linear models with and without language group factors. The explanatory power of the model including language group and its interactions was significantly stronger than that of the model not including language group, critical auxiliary: with language, log likelihood = –28979; without language group, log likelihood = –28984, χ^2 (4) = 9.86; $p < .05$; following word: with language, log likelihood = –29594; without language group, log likelihood = –29604, χ^2 (4) = 20.75; $p < .001$. Summarizing, in the analysis not controlling for reading speed, the L2 group appeared to differ from the native English group in that the grammaticality effect started later and was smaller than in the native English group.

Comparing speed-matched groups. As explained above, we conducted a second analysis in which we controlled for differences in reading speed by taking a subset of the L2 and native English participants, who were matched on reading speed. In addition, reading speed was included as a continuous factor in the analysis. Residual reading times for the speed-matched native English and L2 groups are given in Figure 4 and Figure 5.

Results for the speed-matched groups were different from those of the analysis presented above. First, the effect of grammaticality was largest in participants with a slower average reading speed, leading to an interaction of grammaticality and reading speed (critical auxiliary: $\beta = 0.35$, $SE = 0.10$, $t = 3.36$, $p < .001$; following position: $\beta = 0.36$, $SE = 0.15$, $t = 2.32$, $p < .05$). Second, no interactions

https://doi.org/10.1017/S0142716413000519 Published online by Cambridge University Press

https://doi.org/10.1017/S0142716413000519 Published online by Cambridge University Press

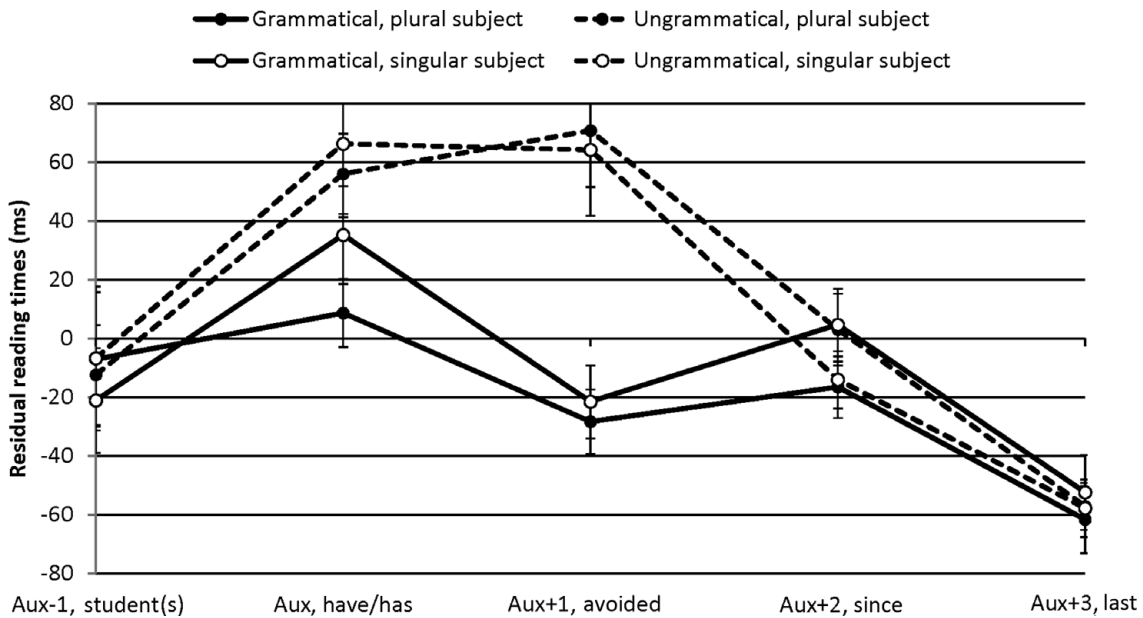


Figure 2. Residual reading times (ms) for the native English group for the four conditions, starting from one word position preceding the critical auxiliary (Aux) until three positions after. Error bars depict the standard error of the mean.

https://doi.org/10.1017/S0142716413000519 Published online by Cambridge University Press

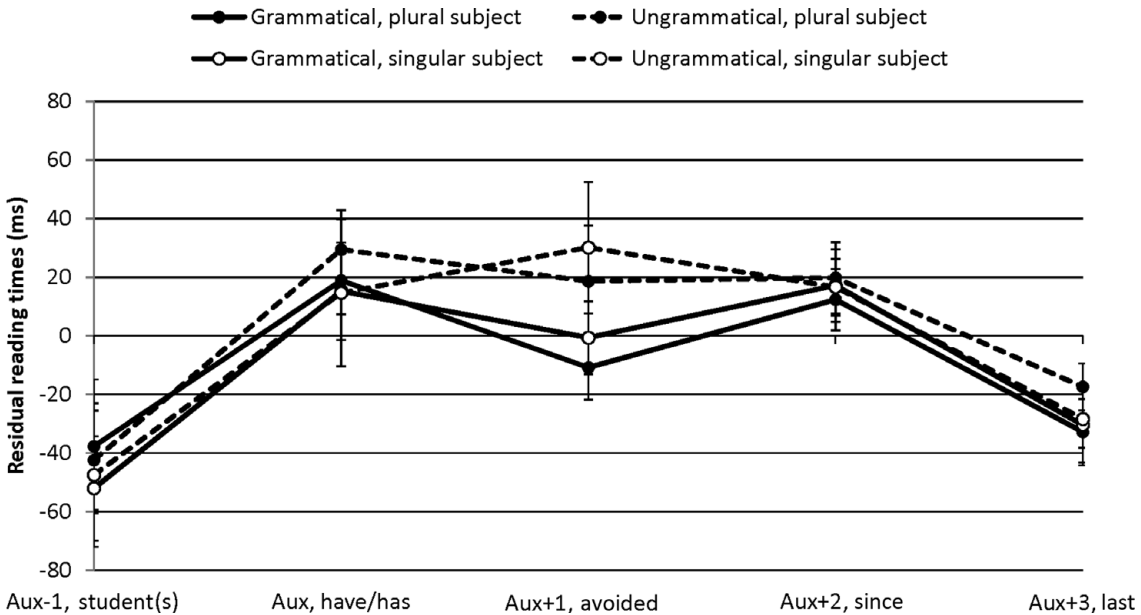


Figure 3. Residual reading times (ms) for the second-language group for the four conditions.

https://doi.org/10.1017/S0142716413000519 Published online by Cambridge University Press

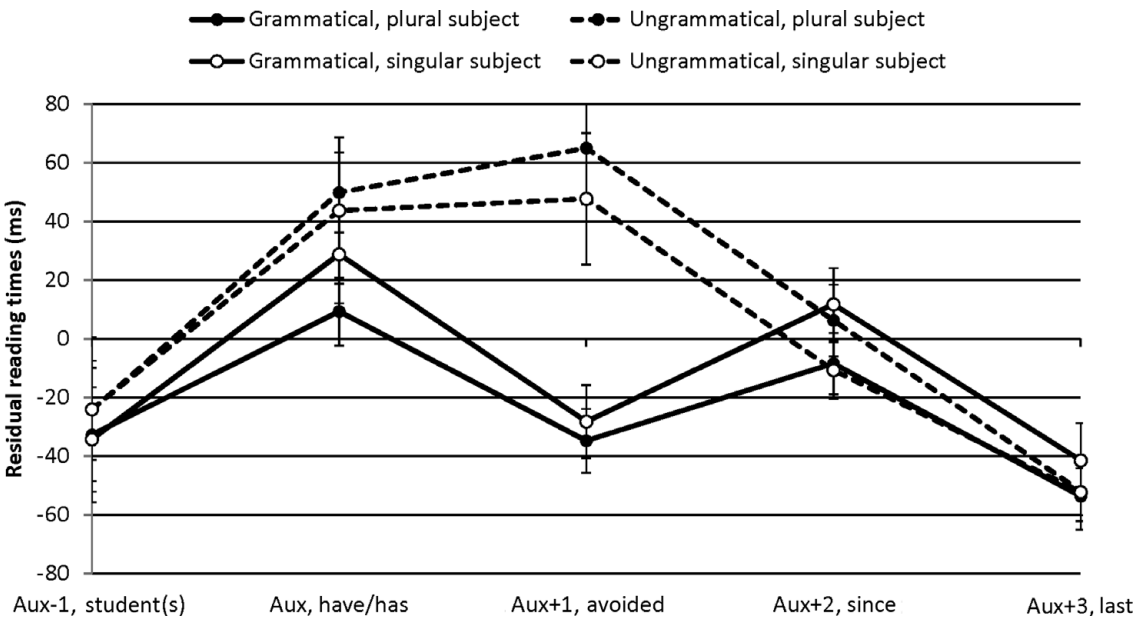


Figure 4. Residual reading times (ms) for the speed-matched native English participants ($N = 37$).

https://doi.org/10.1017/S0142716413000519 Published online by Cambridge University Press

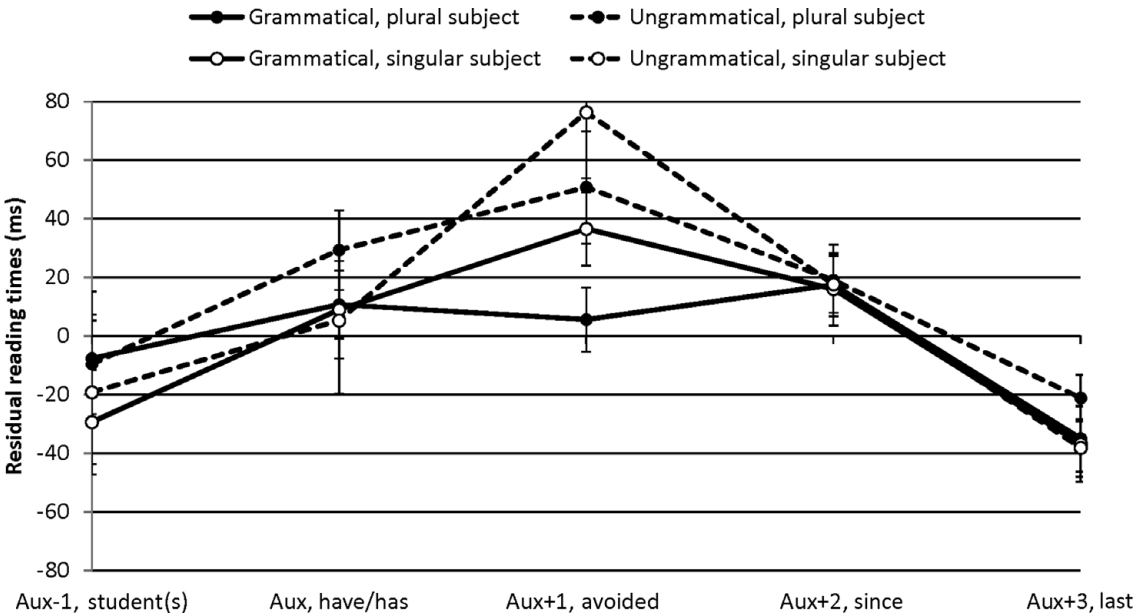


Figure 5. Residual reading times (ms) for the speed-matched second-language participants ($N = 37$).

involving the factors grammaticality and language group were found at the critical word ($ps > .18$). The interaction between grammaticality and language group was significant at the following word ($\beta = 40.58$, $SE = 19.83$, $t = 2.05$, $p < .05$), with separate analyses showing an effect of grammaticality in both native English and L2 groups (native English, $\beta = 83.18$, $SE = 17.93$, $t = 4.64$, $p < .0001$; L2, $\beta = 43.71$, $SE = 15.56$, $t = 2.80$, $p < .01$). As in the first analysis, we found no three-way interaction of plurality, grammaticality, and language group at either position ($ps > .64$), although descriptively, the L2 group showed prolonged reaction times at the critical word to ungrammatical items for the plural subject condition (1b).

When the groups were matched on speed, the L2 and native group thus did not differ in their sensitivity to grammaticality at the critical position. The only difference was that the native group showed a larger difference between the ungrammatical and grammatical conditions at the spillover position than the L2 group. However, a comparison between models for the word following the critical word, with and without language group and its interactions as factors, showed no differences in explanatory power, suggesting that language background did not have any effect above and beyond reading speed, with language: log likelihood = –20114; without language group: log likelihood = –20119, χ^2 (8) = 11.02; $p = .20$.

These results suggest that reading speed can explain a large part of the differences seen in the first analysis between the two language groups: The difference in reading times between the grammatical and ungrammatical conditions was smaller and appeared later for faster readers.[6] In addition, we found no three-way interaction of grammaticality, plurality, and language group. We therefore have no evidence that the L2 group considered a subject–object reading in the language-conflict condition (ungrammatical–plural subject condition, (1b)) during on-line reading.

True/false statements. Each item in the experiment was followed by a true/false statement. On all items in the experiment, including distractors, the native English group performed 91% ($SD = 4.3\%$) correctly; the L2 group performed 88% ($SD = 4.6\%$) correctly. Since we were interested in whether readers assigned an object–subject or subject–object interpretation to the object relative clause, we restricted the analysis of the end-of-sentence true/false judgment to those statements that probed the thematic relations in the relative clause. We analyzed the response type data with a mixed logit model using the lmer function (lme4 packet, Bates et al., 2011) in R (R Development Core Team, 2007), with response type as the dependent variable (1 for an object–subject compatible answer, 0 for a subject–object compatible answer). We followed the same procedures and included the same factors as those described for the analysis of the self-paced reading data. Logit models were fitted by Laplace approximation.

Not controlling for differences in reading speed. An overview of the mean proportion of object-subject compatible responses for the L2 and native English groups is given in Table 2. Overall, the fewest object–subject responses were given to the grammatical condition in which both the object and the subject in the relative clause were singular (4a), leading to an interaction of grammaticality and plurality

https://doi.org/10.1017/S0142716413000519 Published online by Cambridge University Press

Table 2. *Type of response given to true/false statements*

Rel. Clause Subject	Condition	Proportion Object–Subject Responses: Native English ($n = 39$)	Proportion Object–Subject Responses: L2 ($n = 71$)
Plural	Grammatical (1a)	.83 (.03)	.85 (.02)
	Ungrammatical (1b)	.81 (.04)	.65 (.03)
Singular	Grammatical (4a)	.69 (.04)	.57 (.03)
	Ungrammatical (4b)	.80 (.04)	.67 (.03)

Note: L2, second language.

of the relative clause subject ($\beta = 1.45$, $SE = 0.30$, $z = 3.09$, $p < .0001$), and a main effect of plurality ($\beta = –0.71$, $SE = 0.16$, $z = –4.60$, $p < .0001$). The L2 group gave overall fewer object–subject compatible responses than the native English group (language, $\beta = 0.59$, $SE = 0.19$, $z = 3.09$, $p < .01$). In addition, the interaction of grammaticality and language group was significant ($\beta = 0.70$, $SE = 0.31$; $z = 2.30$, $p < .05$). The L2 group gave fewer object–subject responses to the grammatical than ungrammatical conditions when the subject was singular, (4a) versus (4b), but gave fewer object–subject responses to the ungrammatical condition than the grammatical condition when the relative clause subject was plural, (1a) versus (1b). The largest proportion of object–subject responses was observed for the grammatical plural condition (1a). The native English group gave the fewest object–subject responses to the grammatical condition in which both subject and object were singular (4a) but did not show much difference among the other three conditions. The three-way interaction of grammaticality, plurality, and language group failed to reach significance ($p < .10$), however.

Since significant interactions were found involving the factor language group, we conducted separate analyses of the L2 and native English data. The analysis of the L2 group showed that the proportion of object–subject responses was smaller with lower English proficiency ($\beta = 0.28$, $SE = 0.11$; $z = 2.35$; $p < .05$). English proficiency did not interact with grammaticality and/or plurality of the relative clause subject. A main effect of grammaticality was obtained for the English group only: Overall, the native English speakers gave fewer object–subject responses in the grammatical than ungrammatical conditions (Native English, $\beta = 0.81$, $SE = 0.38$, $z = 2.11$; $p < .05$; L2, $\beta = –0.32$, $SE = 0.19$, $z = –1.68$, $p = .09$). This effect may have been driven by the lower proportion of object–subject responses to the grammatical singular condition (4a) in this group. The interaction of grammaticality and plurality was significant only for the L2 group, however (native English: $\beta = 0.40$, $SE = 0.61$, $z = 0.66$, $p = .51$; L2: $\beta = 1.83$, $SE = 0.36$, $z = 5.09$, $p < .0001$).

Follow-up analyses in the L2 group showed significant effects of grammaticality in the plural conditions (more object–subject responses in the grammatical condition, $\beta = –1.25$, $SE = 0.26$; $z = –4.76$; $p < .0001$), as well as in the singular conditions (more object–subject responses in the ungrammatical condition, $\beta = 0.56$,

https://doi.org/10.1017/S0142716413000519 Published online by Cambridge University Press

Table 3. *Mean proportion of object–subject compatible responses for the thematic true/false statements for the speed-matched native English and L2 group, and the fast L2 group*

Rel. Clause Subject	Condition	Matched Native English ($n = 37$)	Matched L2 ($n = 37$)	Faster L2 ($n = 34$)
Plural	Grammatical (1a)	.82 (.22)	.91 (.17)	.78 (.22)
	Ungrammatical (1b)	.80 (.25)	.60 (.27)	.71 (.27)
Singular	Grammatical (4a)	.69 (.24)	.59 (.26)	.54 (.30)
	Ungrammatical (4b)	.79 (.27)	.68 (.25)	.66 (.30)

Note: L2, second language.

$SE = 0.25$; $z = 2.25$; $p < .05$). Although the interaction of grammaticality and plurality was not significant in the native English group, a difference between the grammatical and ungrammatical conditions was seen for the singular conditions (more object–subject responses in the ungrammatical condition, $\beta = 1.04$, $SE = 0.38$; $z = 2.75$; $p < .01$) but not the plural ($\beta = 0.50$, $SE = 0.54$; $z = 0.93$; $p = .36$).[7]

As in the analysis of the on-line reading times, including the factor language group and its interactions greatly improved the explanatory power of the model, with language, log likelihood = –686.19; without language group, log likelihood = –694.44, χ^2 (4) = 9.86; $p < .01$.

The differences described were unique to the thematic true/false statements. Analysis of the performance on the nonthematic true/false statements following the relative clause conditions revealed no differences between the conditions and the language groups. This suggests that the differences found are related to the word order interpretation rather than the processing of the conditions in general.

Comparing speed-matched groups. Table 3 gives an overview of the mean proportion of object–subject compatible answers in the end-of-sentence true/false statements probing thematic relations for the two speed-matched groups. Results for the response type data for the speed-matched groups were largely similar to those of the complete data set, except that the three-way interaction among grammaticality, plurality, and language group was now significant ($\beta = –1.70$, $SE = 0.72$, $z = –.36$, $p < .05$). A separate analysis for the speed-matched L2 group showed a significant effect of grammaticality ($\beta = –1.03$, $SE = 0.32$, $z = –3.23$, $p < .01$), plurality ($\beta = –1.04$, $SE = 0.31$, $z = –3.34$, $p < .001$), and the interaction between these two factors ($\beta = 3.12$, $SE = 0.64$, $z = 4.82$, $p < .0001$). The matched L2 groups gave more object–subject compatible responses following the grammatical than the ungrammatical conditions when the subject was plural, (1a) versus (1b), ($\beta = –2.68$, $SE = 0.48$, $z = –5.48$, $p < .0001$), while in contrast to the analysis on the complete data set, the difference between grammatical and ungrammatical conditions was not significant when the relative clause subject was singular ($\beta = 0.44$, $SE = 0.33$, $z = 1.87$, $p > .18$).[8] Effects for the native English group did not differ from those reported for the complete data set.

https://doi.org/10.1017/S0142716413000519 Published online by Cambridge University Press

The *t* tests on aggregated data between the speed-matched groups showed that the L2 and English groups differed only on the ungrammatical plural (1b) condition, $t\,(72) = 3.25$, $p = .02$; other conditions, $ps > .05$.

In contrast to the on-line reading data, reading speed did not have any effect in the speed-matched groups, but language group did. A comparison of models including and excluding the factor language group and its interactions showed that the inclusion of language group significantly improved the explanatory power of the model: with language, log likelihood = –452.22; without language group, log likelihood = –462.65, $\chi^2\ (8) = 20.86$; $p < .01$. Reading speed did not improve the model of the speed-matched groups, with reading speed, log likelihood = –452.22; without reading speed, log likelihood = –454.48, $\chi^2\ (8) = 4.53$; $p = .97$.

Note, however, that the differences between the groups on the end-of-sentence verification task were more pronounced for the speed-matched groups than in the complete data set. This was partly because the faster reading L2 speakers had been excluded from the speed-matched data. Table 3 also includes the response type data for the (on average) faster reading group of L2 speakers, who were not speed matched with the native English group. The faster reading L2 speakers did not differ from the slower reading L2 group (speed matched with the native English group) in terms of English proficiency or other demographic measures mentioned in Table 1 ($ps > .24$). Even though these faster reading L2 learners gave overall fewer object–subject compatible responses than the native English group ($\beta = 0.65$, $SE = 0.32$, $z = 2.05$, $p < .05$), their response pattern resembled that of the native English speakers (no significant interactions involving the factor group, $ps > .5$).[9]

DISCUSSION

The goal of the present study was to test whether (a) differences between L2 and native processing could be attributed to differences in reading speed and (b) L2 learners are sensitive to cross-language conflict. Results relevant to reading speed showed that our L2 learners read sentences faster than the native English speakers. Differences between native speakers and L2 speakers found in the on-line reading data could be explained largely by differences in reading speed: In the analysis controlling for reading speed, the L2 learners showed on-line reading patterns similar to those of the native speakers. However, even when the groups were matched for reading speed, the L2 and native English groups differed in the end-of-sentence statement verification task. Only the faster reading L2 learners showed nativelike patterns.

Results relevant to the sensitivity to cross-language conflicts included analyses of on-line reading times and an off-line sentence verification task. For the on-line reading time data, we predicted a smaller effect of grammaticality in the L2 learners for the plural subject conditions (1b) versus (1a) than for the singular subject conditions (4b) versus (4a), but we did not expect the grammaticality effect in the native English speakers to be influenced by the number of the relative clause subject. This was not borne out: we found no effects in the on-line reading times that would suggest that the L2 learners were more biased toward a

https://doi.org/10.1017/S0142716413000519 Published online by Cambridge University Press

subject–object interpretation in the language conflict condition than the native English speakers. Our predictions for the end-of-sentence verification statements, however, were largely borne out. We had expected the Dutch L2 learners of English to assign a object–subject interpretation to the clauses less often than native English speakers, especially in the conditions (1b) and (4a), which were compatible with a grammatical subject–object reading in Dutch, and in (4b), in which it is unclear which noun phrase is the subject (Kilborn & Cooreman, 1987; McDonald, 1987). Our analysis showed that the L2 learners, especially those with lower English proficiency, gave fewer object–subject responses than native English speakers. In addition, the speed-matched L2 speakers assigned more object–subject interpretations in the grammatical plural subject conditions (1a) than they did in the other conditions, (1b), (4a), and (4b). A direct comparison of the conditions between the two language groups also showed that the only condition in which the speed-matched L2 learners assigned significantly fewer object–subject interpretations than the native English speakers was the ungrammatical plural-subject condition (1b). This was the condition that is ungrammatical in English but corresponds to a grammatical subject–object ordered relative clause in the L2 learners' native language. We elaborate below on the effect of reading speed and discuss potential accounts of the patterns observed in the end-of-sentence verification task. We will conclude with some theoretical implications.

Reading speed

Some approaches to L2 processing ascribe differences between L2 and native sentence processing to a difference in resources: in L2 learners, access to L2 information and routines is slower and less automatic, which may have repercussions down the line, especially when task demands increase (Hopp, 2009, 2010; Sorace, 2011; Wilson, 2009). One argument in support of this view is that language learners typically read more slowly than native speakers (for recent studies, see Coughlin & Tremblay, 2012; Roberts & Felser, 2011) and that reading speed correlates with proficiency (Roberts & Felser, 2011). The L2 learners in our study, however, read on average faster than the native English speakers (see also Felser et al., 2003; as mentioned in Clahsen & Felser, 2006) and did not show any correlation between reading speed and the proficiency measures used. Differences in processing between the native English speakers and the L2 group can therefore not be ascribed to a general processing slow-down in the L2 learners.

Our results show that reading speed can drastically affect the timing and size of the effects obtained. Faster readers (native speakers as well as L2 learners) showed delayed and smaller differences in reading times between the ungrammatical and grammatical conditions (see also Coughlin & Tremblay, 2012; Roberts & Felser, 2011). Apparent differences between the L2 and native groups in the on-line reading data could be largely reduced to differences in reading speed.

Differences between the L2 group and the native group in the end-of-sentence verification data, though, could not be reduced to reading speed. Even when reading at the same speed as native speakers, L2 speakers appeared to have more difficulty computing and maintaining syntactic representations than the native

https://doi.org/10.1017/S0142716413000519 Published online by Cambridge University Press

English speakers did. The difference between the results from the self-paced reading task and the end-of-sentence statement verification task may be attributed to differences in task demands. The differences between the groups may not have been apparent in the on-line reading task data because the ungrammaticality was relatively easy to detect during reading, especially since the relative clause subject and the critical auxiliary were adjacent (Clahsen & Felser, 2006).[10] The end-of-sentence verification task required more resources: While evaluating the true/false statement at the end of the sentence, the reader needed to maintain in memory a representation of the sentence presented earlier. The meaning expressed in the true/false statement needed to be explicitly compared with the representation in memory. While the thematic relations were probed in the true/false statements, both native and L2 readers may have occasionally not been able to retrieve or maintain a precise representation of the sentence, or may have been confused for reasons to be discussed below. Speed-matched (i.e., relatively slow) L2 learners may have broken down more often than native speakers because fewer resources were available to them, either because access to lexical and other information required more effort and/or because the native language needed to be inhibited. Our findings correspond to a discrepancy between on- and off-line tasks observed by Hopp (2009). In that study, English learners of German were shown to be sensitive to discourse conditions on word order variations in German in an on-line reading task in which they had to verify statements about the content, but they performed much worse than native German speakers in an off-line grammaticality judgment task, which required combining various sources of information and, therefore, may have been more demanding.

The L2 learners in our study gave more object–subject compatible responses to the true/false statements when they scored higher on the English proficiency tasks. This suggests that processing becomes more efficient and that readers are better able to keep track of the thematic roles in the verification task as proficiency increases. Furthermore, the faster L2 readers showed a nativelike pattern in the statement verification task. These faster L2 readers did not differ from the slower (speed matched) L2 group in proficiency, however, suggesting that some L2 learners may become more efficient even though this may not be reflected in proficiency measures used (Hoover & Dwivedi, 1998).

Cross-language conflict?

The second aim of the study was to investigate whether cross-language conflict affects sentence comprehension in L2 learners as compared to native speakers. Differences in response patterns between the L2 learners and native speakers were found only in the end-of-sentence verification task. As mentioned previously, this might be due to the verification task being more demanding than the on-line reading task. When the two groups were compared on each condition separately, the speed-matched L2 learners significantly differed from the native speakers only in the ungrammatical plural subject (1b) condition: The L2 learners assigned fewer object–subject interpretations in this condition than the native speakers. The ungrammatical plural subject condition was the only condition which was ungrammatical in English but corresponded to a grammatical subject–object structure in

https://doi.org/10.1017/S0142716413000519 Published online by Cambridge University Press

the learners native language. This result can therefore be interpreted as evidence supporting a model in which a bilingual's languages are active simultaneously and in which syntactic structures and procedures in one language can influence the processing of sentences in the other language (Costa & Caramazza, 1999; MacWhinney & Bates, 1989; Sorace, 2011; Tokowicz & MacWhinney, 2005). Especially when a word sequence forms an ungrammatical or infrequent structure in the L2 but corresponds to a grammatical and frequent structure in the L1, the L1 structure or parsing procedure may affect the processing of the L2. Our finding, then, corresponds to other observations of cross-language effects in bilingual sentence processing (Dussias & Cramer Scaltz, 2008; Dussias & Sagarra, 2007; Frenck-Mestre, 2005; Frenck-Mestre & Pynte, 1997; Hopp, 2009; Jacob, 2009; Juffs, 2005; Kilborn, 1989; Kilborn & Cooreman, 1987; MacWhinney & Bates, 1989; McDonald, 1987; Tokowicz & MacWhinney, 2005).

However, when the full numeric pattern of results is taken into consideration, several alternative accounts can be provided of our end-of-sentence verification data. We will sketch two below. Accounts differ in the nature of the strategies used by the L2 speakers and in what triggers the use of these strategies. These accounts are not mutually exclusive, and we currently lack data to further support one approach over the other. One account is based on the competition model (MacWhinney & Bates, 1989; Tokowicz & MacWhinney, 2005). Under this approach, L2 learners weigh syntactic cues differently from native speakers. A study by McDonald (1987) showed that native Dutch speakers relied more on subject–verb agreement than word order in their native language, especially when word order and agreement pointed to conflicting interpretations. This reliance on agreement transferred to the processing of the L2 English. In support of the idea that our speed-matched L2 learners relied more on subject–verb agreement than did native English speakers is the observation that the L2 learners showed a high proportion of object–subject compatible responses in the grammatical plural conditions (1a), in which the verb was plural and the subject was the only plural noun. In addition, L2 learners showed a low proportion of object–subject responses in the ungrammatical plural conditions (1b), in which the verb was singular and the relative clause head was the only singular noun. This suggests that the L2 speakers more often interpreted the noun phrase agreeing with the verb as the subject than the native English speakers did. A competition model account also predicts that in the ungrammatical plural condition (singular head, plural subject, singular verb (1b)), in which only the relative clause head agrees with the verb, L2 speakers should assign more subject–object interpretations than in the ungrammatical singular conditions (singular head, singular subject, plural verb (4b)), in which both noun phrases are singular, and agreement is not a cue. We did not find evidence supporting this, however (see Note 7).

There is a third potential account of the differences between the L2 pattern and the native pattern. Given that our L2 speakers noticed the ungrammaticality during on-line reading, the presence of the ungrammaticality may have increased processing difficulty and/or may have led to confusion as to the thematic structure of the relative clause. Under the circumstances, the L2 learners may have defaulted to a subject before object order more often than the native English speakers, accounting for the pattern found in the ungrammatical conditions (1b) and (4b).

https://doi.org/10.1017/S0142716413000519 Published online by Cambridge University Press

This account can also explain the low proportion of object–subject responses in both native English and L2 groups in the grammatical singular subject condition (4a). In this condition, both noun phrases are singular and may be easily confusable (Gordon et al., 2001). The order of the nouns may therefore not have been remembered correctly when the true/false statement was verified, leading to a numerical decrease in the proportion of object–subject interpretations assigned to this condition for both the L2 learners and the native speakers. This account, however, needs to assume that the agreement error in (4b) slightly facilitated the encoding or retrieval of the order of the nouns. This condition also contained two singular noun phrases, but the verb did not agree with either, yet all groups gave more object–subject responses in this condition than in the grammatical singular condition (4a). The agreement error in (4b) may have drawn more attention to the position of the nouns, making the order easier to remember. Under this third interpretation, then, L2 learners build the same structures and may confuse the position of the nouns in the same way as native speakers do. However, the (speed-matched) L2 learners differ from native speakers mainly in how they deal with an ungrammaticality and the resulting processing difficulty: L2 speakers more often default to a subject before object order, whereas native English speakers default to a object before subject order for NP NP V sequences.

The three accounts sketched above differ mainly in what triggers the L2 learners to use the subject before object order (conflict from a grammatical structure in the L1, conflict between cues, general difficulty due to an ungrammaticality). It remains to be determined whether the L2 speaker's default subject before object interpretation is due to transfer from the native language or whether this is a general L2 processing heuristic. In a picture-selection study (Nitschke, Kidd, & Serratrice, 2010), English L2 learners of German more often assigned an object before subject interpretation to an ambiguous NP NP V relative clause in German when compared with native German speakers (see also Bates & MacWhinney, 1981). This suggests that a subject–object pattern is not a general default strategy for L2 learners but may be specific to the language combinations in question. To shed more light on this issue, and on the potential effects of cross-language conflict, studies are needed that compare L2 learners whose native languages differ in word order.

Theoretical implications

According to the shallow structure hypothesis (Clahsen & Felser, 2006), L2 readers do not build detailed syntactic representations but rely instead on lexical and contextual information and local agreement. Under this account, differences between L2 learners and native speakers are not attributed to resource issues, processing speed, or L1 transfer. Our on-line results are compatible with this model: L2 learners are sensitive to local agreement between a noun phrase and an adjacent finite verb, resulting in the nativelike grammaticality effects observed in the self-paced reading data when processing speed is controlled for. However, the shallow structure hypothesis approach cannot readily account for the complexity of our end-of-sentence verification data, in particular the observation that the faster reading L2 speakers showed a nativelike pattern in their end-of-sentence

https://doi.org/10.1017/S0142716413000519 Published online by Cambridge University Press

responses, whereas the slower readers (speed-matched with the native English readers) did not.

Although our results may not be incompatible with the shallow structure hypothesis, our findings can be more easily explained by assuming that L2 learners and native speakers have the same grammar and are able to build the same syntactic structures during on-line processing, but may differ in terms of allocation of processing resources (e.g., Hopp, 2010; McDonald, 2006). Even though some of our L2 learners read at the same speed as the native speakers, they may have had more difficulty computing and maintaining the syntactic representations than native speakers. This may be because their access to the information is slower, information is activated to a lesser extent, or fewer resources are available because they need to inhibit their native language. This might not have affected the on-line reading of the sentences but may well have affected the more resource-demanding end-of-sentence verification task, especially in sentences that contained a cross-linguistic ambiguity or ungrammaticality. The faster readers, by contrast, who were more efficient and had more resources available, showed a nativelike response pattern.

Concluding remarks

Some accounts of the difference between native and L2 sentence processing are based on the assumption that L2 processing is slower and less automatic, leading to a decrease and delay in the effects relative to native speakers (Dekydtspotter et al., 2006; Hoover & Dwivedi, 1998). Contrary to what would be expected on the basis of these account, the L2 speakers in our study read overall *faster* than native speakers. Faster overall reading speed was associated with smaller and later effects of grammaticality. After controlling for reading speed, however, the L2 speakers in our study showed on-line reading patterns that were very similar to those observed for native speakers but showed different effects in the more resource-demanding end-of-sentence interpretation task.

Speed on its own may not be an indication of processing efficiency; participants can read at the same speed, but some people may read more thoroughly than others. As opposed to Roberts and Felser (2011), for example, we also found that differences in speed did not relate to differences in English proficiency (see also Hoover & Dwivedi, 1998). The relation between reading speed, L2 proficiency, and when effects occur during on-line processing is therefore not straightforward. Future research, however, should take differences in reading speed into account when interpreting differences observed between populations.

ACKNOWLEDGMENTS

The authors thank Carlie Overfelt, Rachel Groenhout, Lauren Concepcion, Ania Fiksinski, Kayla Johnson, Renise St. Louis, and Gijsbert Westland for their help with stimulus preparation, data collection, and preliminary data analyses. This research was supported by an Internationalizeringsbeurs from the Netherlands Organization for Scientific Research (Grant 236-70-004) and by the National Science Foundation (Grant 0957178).

https://doi.org/10.1017/S0142716413000519 Published online by Cambridge University Press

NOTES

1. We manipulated agreement on highly frequent, irregular auxiliaries rather than lexical verbs. This was to make the number information more obvious to the reader, and to make comparisons between the conditions at the critical word as clean as possible. In addition, since highly frequent and irregular morphological forms are assumed to be stored in the mental lexicon rather than computed by a procedural system in both L2 and native speakers (Ullman, 2001), differences between the groups at the critical word would not likely be due to differences in the lexical-morphological processing of the critical word form.
2. Because the task was not a grammaticality judgment task, we did not want our participants to focus on the occurrences of ungrammaticalities in the sentences. For that reason we decided to have more grammatical than ungrammatical sentences.
3. The Dutch proficiency tasks were included in the experiment to test the relation between L1 and L2 proficiency and cognitive control. No effect was found of proficiency on cognitive control. Our L2 group showed significant positive correlations between the L1 and L2 proficiency measures used (*r*s ranging from .37, $p < .001$ for the C-test, to .68, $p < .0001$, for the PPVT). This suggests that the English proficiency measures used also reflect general language abilities and experience.
4. Residual reading times also control for differences in overall reading speed among participants. We were, however, interested in the effect of reading speed on the pattern of the differences between the conditions and not so much in the differences in overall speed between the groups. In addition, we wanted to control for differences in word length between the conditions, since the critical auxiliaries differed in length. We therefore considered it more appropriate to conduct the analysis on residual reading times rather than raw reading times and to use an independent measure of reading speed (defined as raw reading times on grammatical distractor items) as one of the predictors.
5. Data were analyzed for the word preceding the critical word until three word positions after the critical auxiliary (positions starting from the fourth word after the auxiliary were confounded by sentence-final words in some items). No effects involving grammaticality were found at positions other than the critical word and the position directly following it.
6. In order to investigate whether faster readers postponed interpreting the sentence until the end of the sentence or until they saw the true/false statement (Roberts & Felser, 2011), we also analyzed response times to the sentence-final words, and to thematic true/false statements. No significant effects were found at the sentence-final words. Responses to the thematic true/false statements (restricted to object–subject compatible responses) were slower when the overall reading speed was slower (for speed-matched groups, $\beta = 3.10$, $SE = 1.43$, $t = 2.17$, $p < .05$). These results therefore provide no supportive evidence for the view that fast readers postpone the interpretation of the sentence.
7. Response times to the true/false statements patterned differently from the response type data: A main effect of grammaticality (slower responses in ungrammatical than grammatical conditions) was seen for the L2 group (L2: $\beta = 578.70$, $SE = 170.55$, $t = 3.39$, $p < .001$) but not for the native English group ($\beta = –26.24$, $SE = 179.02$, $t = –0.15$, $p = .88$). We therefore have no evidence of a speed–accuracy trade-off.

https://doi.org/10.1017/S0142716413000519 Published online by Cambridge University Press

8. In order to check whether the presence of two different noun phrases (one singular, one plural noun phrase (1a,b)) as opposed to two singular noun phrases (4a,b) would affect the processing of the (un)grammaticality in the L2 group, separate analyses were performed on the grammatical and ungrammatical items. The effect of plurality was significant for the grammatical items only (grammatical, $\beta = -2.50$, $SE = 0.49$, $z = -5.15$, $p < .0001$; ungrammatical, $\beta = 0.37$, $SE = 0.31$, $z = 1.20$, $p = 0.23$), suggesting that the presence of an agreement cue biasing toward a subject–object order does not have an additional effect on the processing of an ungrammatical structure, see the discussion.
9. One reviewer suggested that effects of the L1 subject–object order may be more pronounced in responses to true/false statements that were false under an object–subject interpretation, that is, statements in which the order of the noun phrases corresponded to a subject–object interpretation. The mean proportion of object–subject compatible responses (i.e., "false" responses) to false true/false statements probing thematic relations are given in the Supplementary Materials (http://journals.cambridge.org/aps). Although the pattern of the means for only the false statements resembled that of the true and false statements combined, we did not find any interaction of language group with grammaticality and/or number of the relative clause subject when analyses were restricted to responses to false statements only. Since some conditions only contained one false statement in a particular presentation list, we may have lacked power.
10. Alternatively, the self-paced reading technique used may not have been sensitive enough to subtle differences in timing or level or processing between native speakers and L2 learners. A replication of the study using methods with a higher temporal resolution, such as eye-tracking or event-related potentials, may yield different results.

SUPPLEMENTARY MATERIALS

Online Supplementary Materials to this paper can be found at http://journals.cambridge.org/aps.

REFERENCES

Baayen, R. H. (2008). *Analyzing linguistic data. A practical introduction to statistics using R*. Cambridge: Cambridge University Press.

Baayen, R. H., Piepenbrock, R., & Gulikers, L. (1995). *The CELEX Lexical Database* [CD]. Philadelphia, PA: University of Pennsylvania, Linguistic Data Consortium.

Bader, M., & Meng, M. (1999). Subject–object ambiguities in German embedded clauses. An across the board comparison. *Journal of Psycholinguistic Research, 28*, 121–141.

Barr, D. J., Levy, R., Scheepers, C., & Tily, H. J. (2013). Random effects structure for confirmatory hypothesis testing: Keep it maximal. *Journal of Memory and Language, 68*, 255–278.

Bates, D., Maechler, M., & Bolker, B. (2011). lme4: Linear mixed-effects models using S4 classes. R package version 0.9989375-42 [Computer software].

Bates, E., & MacWhinney, B. (1981). Second language acquisition from a functionalist perspective: Pragmatic, semantic and perceptual strategies. In H. Winitz (Ed.), *Annals of the New York Academy of Sciences Conference on Native and Foreign Language Acquisition* (pp. 190–214). New York: New York Academy of Sciences.

Bialystok, E. (2009). Bilingualism: The good, the bad, and the indifferent. *Bilingualism: Language and Cognition, 12*, 3–11.

https://doi.org/10.1017/S0142716413000519 Published online by Cambridge University Press

Clahsen, H., & Felser, C. (2006). Grammatical processing in language learners. *Applied Psycholinguistics, 27*, 3–42.
Costa, A., & Caramazza, A. (1999). Is lexical selection language specific? Further evidence from Spanish–English bilinguals. *Bilingualism: Language and Cognition, 2*, 231–244.
Costa, A., Hernandez, M., & Sebastián-Gallés, N. (2008). Bilingualism aids conflict resolution: Evidence from the ANT task. *Cognition, 106*, 59–86.
Coughlin, C. E., & Tremblay, A. (2012). Proficiency and working memory based explanations for nonnative speakers' sensitivity to agreement in sentence processing. *Applied Psycholinguistics, 34*, 615–646.
Dallas, A. C. (2008). *Influences of verbal properties on second-language filler-gap resolution: A cross-methodological study.* Unpublished doctoral dissertation. Retrieved from http://search.proquest.com/docview/741238885?accountid=10920
Dekydtspotter, L., Schwartz, B. D., & Sprouse, R. A. (2006). *The comparative fallacy in L2 processing research.* Paper presented at the 8th Generative Approaches to Second Language Acquisition Conference (GASLA 2006).
Dijkstra, T., Grainger, J., & van Heuven, W. J. B. (1999). Recognition of cognates and interlingual homographs: The neglected role of phonology. *Journal of Memory and Language, 41*, 496–518.
Dunn, L. M., & Dunn, D. M. (2007). *Peabody Picture Vocabulary Test* (4th ed.). Minneapolis, MN: NCS Pearson.
Dussias, P. E. (2003). Syntactic ambiguity resolution in L2 learners. Some effects of bilinguality on L1 and L2 processing strategies. *Studies in Second Language Acquisition, 25*, 529–557.
Dussias, P. E., & Cramer Scaltz, T. R. (2008). Spanish–English L2 speakers' use of subcategorization bias information in the resolution of temporary ambiguity during second language reading. *Acta Psychologia, 128*, 501–513.
Dussias, P. E., & Sagarra, N. (2007). The effect of exposure on syntactic parsing in Spanish–English bilinguals. *Bilingualism: Language and Cognition, 10*, 101–116.
Fan, J., McCandliss, B. D., Sommer, T., Raz, A., & Posner, M. I. (2002). Testing the efficiency and independence of attentional networks. *Journal of Cognitive Neuroscience, 14*, 340–347.
Felser, C., Roberts, L., Gross, R., & Marinis, T. (2003). The processing of ambiguous sentences by first and second language learners of English. *Applied Psycholinguistics, 24*, 453–489.
Frazier, L. (1978). *On comprehending sentences: Syntactic parsing strategies.* Unpublished doctoral dissertation, University of Connecticut.
Frazier, L., & Flores d'Arcais, G. B. (1989). Filler driven parsing: A study of gap filling in Dutch. *Journal of Memory and Language, 28*, 331–344.
Frenck-Mestre, C. (2005). Ambiguities and anomalies: What can eye movements and event-related potentials reveal about second language sentence processing? In J. F. Kroll & A. M. B. De Groot (Eds.), *Handbook of bilingualism: Psycholinguistic approaches* (pp. 268–281). New York: Oxford University Press.
Frenck-Mestre, C., & Pynte, J. (1997). Syntactic ambiguity resolution while reading second and native languages. *Quarterly Journal of Experimental Psychology A, 50*, 119–148.
Gordon, P. C., Hendrick, R., & Johnson, M. (2001). Memory interference during language processing. *Journal of Experimental Psychology: Learning, Memory, and Cognition, 27*, 1411–1423.
Grotjahn, R. (1987). How to construct and evaluate a C-test: A discussion of some problems and some statistical analyses. *Quantitative Linguistics, 34*, 219–253.
Hahne, A. (2001). What's different in second-language processing? Evidence from event-related brain potentials. *Journal of Psycholinguistics Research 30*, 251–266.
Havik, E., Roberts, L., Van Hout, R., Schreuder, R., & Haverkort, M. (2009). Processing subject–object ambiguities in the L2: A self-paced reading study with German L2 learners of Dutch. *Language Learning, 59*, 73–112.
Hollingshead, A. B. (1975). *Four Factor Index of Social Status*. New Haven, CT: Yale University.
Hoover, M. L., & Dwivedi, V. D. (1998). Syntactic processing by skilled bilinguals. *Language Learning, 48*, 1–29.
Hopp, H. (2006). Syntactic features and reanalysis in near-native processing. *Second Language Research, 22*, 369–397.
Hopp, H. (2009). The syntax–discourse interface in near-native L2 acquisition: Off-line and on-line performance. *Bilingualism: Language and Cognition, 12*, 463–483.

https://doi.org/10.1017/S0142716413000519 Published online by Cambridge University Press

Hopp, H. (2010). Ultimate attainment in L2 inflection: Performance similarities between non-native and native speakers. *Lingua, 120*, 901–931.
Jacob, G. (2009). *The role of native language in second-language syntactic processing.* Unpublished doctoral dissertation, University of Dundee, Scotland.
Juffs, A. (2004). Representation, processing and working memory in a second language. *Transactions of the Philological Society, 102*, 199–225.
Juffs, A. (2005). The influence of first language on the processing of *wh*-movement in English as a second language. *Second Language Research, 21*, 121–151.
Kaan, E. (2001). Effects of NP type on the resolution of word-order ambiguities. *Journal of Psycholinguistic Research, 30*, 529–547.
Kaan, E., Dallas, A. C., & Wijnen, F. (2010). Syntactic predictions in second-language sentence processing. In J.-W. Zwart & M. de Vries (Eds.), *Structure preserved. Festschrift in the honor of Jan Koster* (pp. 207–213). Amsterdam: John Benjamins.
Keijzer, M. (2007). *Last in, first out? An investigation of the regression hypothesis in Dutch emigrants in anglophone Canada.* Doctoral dissertation, Free University of Amsterdam. LOT Dissertation Series 163. Retrieved from http://www.lot.nl
Kilborn, K. (1989). Sentence processing in a second language: The timing of transfer. *Language and Speech, 32*, 1–23.
Kilborn, K., & Cooreman, A. (1987). Sentence interpretation strategies in adult Dutch–English bilinguals. *Applied Psycholinguistics, 8*, 415–431.
Kroll, J. F., Bobb, S. C., & Wodniecka, Z. (2006). Language selectivity is the exception, not the rule: Arguments against a fixed locus of language selection in bilingual speech. *Bilingualism: Language and Cognition, 9*, 119–135.
Lew-Williams, C., & Fernald, A. (2010). Real-time processing of gender-marked articles by native and non-native Spanish speakers. *Journal of Memory and Language, 63*, 447–464.
Linck, J., Hoshino, N., & Kroll, J. F. (2008). Cross-language lexical processes and inhibitory control. *Mental Lexicon, 3*, 349–374.
MacWhinney, B., & Bates, E. (1989). *The crosslinguistic study of sentence processing*. New York: Cambridge University Press.
Marian, V., Blumenfield, H. K., & Kaushanskaya, M. (2007). The language experience and proficiency questionnaire (LEAP-Q): Assessing language profiles in bilinguals and multi-linguals. *Journal of Speech and Hearing Research, 50*, 940–967.
Marian, V., & Spivey, M. (2003). Competing activation in bilingual language processing. *Bilingualism: Language and Cognition, 6*, 97–115.
Marinis, T., Roberts, L., Felser, C., & Clahsen, H. (2005). Gaps in second language sentence processing. *Studies in Second Language Acquisition, 27*, 53–78.
McDonald, J. (1987). Sentence interpretation in bilingual speakers of English and Dutch. *Applied Psycholinguistics, 8*, 379–414.
McDonald, J. (2006). Beyond the critical period: Processing-based explanations for poor grammaticality judgment performance by late second language learners. *Journal of Memory and Language, 55*, 381–401.
Nitschke, S., Kidd, E., & Serratrice, L. (2010). First language transfer and long-term structural priming in comprehension. *Language & Cognitive Processes, 25*, 94–114.
Papadopoulou, D., & Clahsen, H. (2003). Parsing strategies in L1 and L2 sentence processing: A study of relative clause attachment in Greek. *Studies in Second Language Acquisition, 24*, 501–528.
R Development Core Team. (2007). *R: A language and environment for statistical computing*. Vienna: R Foundation for Statistical Computing. Retrieved from http://www.R-project.org
Reitan, R. M., & Wolfson, D. (1995). Category Test and Trail Making Test as measures of frontal lobe functions. *Clinical Neuropsychologist, 9*, 50–56.
Roberts, L., & Felser, C. (2011). Plausibility and recovery from garden paths in second language sentence processing. *Applied Psycholinguistics, 32*, 299–331.
Sabourin, L., & Stowe, L. A. (2008). Second language processing: When are first and second languages processed similarly? *Second Language Research, 24*, 397–430.
Schriefers, H., Friederici, A. D., & Kühn, K. (1995). The processing of locally ambiguous clauses in German. *Journal of Memory and Language, 34*, 499–520.
Schwartz, A. I., & Kroll, J. F. (2006). Bilingual lexical activation in sentence context *Journal of Memory and Language, 55*, 197–212.

https://doi.org/10.1017/S0142716413000519 Published online by Cambridge University Press

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: Psycholinguistics approaches* (pp. 371–388). Oxford: Oxford University Press.

Snodgrass, J. G., & Vanderwart, M. (1980). A standardized set of 260 pictures: Norms for name agreement, image agreement, familiarity and visual complexity. *Journal of Experimental Psychology: Human Learning and Memory, 6*, 174–215.

Sorace, A. (2011). Pinning down the concept of "interface" in bilingualism. *Linguistic Approaches to Bilingualism, 1*, 1–33.

Stroop, J. R. (1935). Studies of interference in serial verbal reactions. *Journal of Experimental Psychology, 18*, 643–662.

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

Ullman, M. T. (2001). The neural basis of lexicon and grammar in first and second language: The declarative/procedural model. *Bilingualism: Language and Cognition, 4*, 105–122.

Van Hell, J., & Dijkstra, T. (2002). Foreign language knowledge can influence native language performance in exclusively native contexts. *Psychonomic Bulletin & Review, 9*, 780–789.

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

Wagers, M. W., Lau, E. F., & Phillips, C. (2009). Agreement attraction in comprehension: Representations and processes. *Journal of Memory and Language, 61*, 206–237.

Williams, J. N., Möbius, P., & Kim, C. (2001). Native and non-native processing of English *wh*-questions: Parsing strategies and plausibility constraints. *Applied Psycholinguistics, 22*, 509–540.

Wilson, F. (2009). *Processing at the syntax–discourse interface in second language acquisition.* Doctoral dissertation, University of Edinburgh.

https://doi.org/10.1017/S0142716413000519 Published online by Cambridge University Press