Cross-linguistic differences in processing double-embedded relative clauses: Working-memory constraints or language statistics?

Stefan L. Frank
s.frank@let.ru.nl
Centre for Language Studies, Radboud University Nijmegen
P.O. Box 9103, 6500 HD Nijmegen, The Netherlands

Thijs Trompenaars
Centre for Language Studies, Radboud University Nijmegen

Shravan Vasishth
Department of Linguistics, University of Potsdam
and
School of Mathematics and Statistics, University of Sheffield

Abstract

An English double-embedded relative clause from which the middle verb is omitted can often be processed more easily than its grammatical counterpart, a phenomenon known as the grammaticality illusion. This effect has been found to be reversed in German, suggesting that the illusion is language specific rather than a consequence of universal working memory constraints. We present results from three self-paced reading experiments which show that Dutch native speakers also do not show the grammaticality illusion in Dutch, whereas both German and Dutch native speakers do show the illusion when reading English sentences. These findings provide evidence against working memory constraints as an explanation for the observed effect in English. We propose an alternative account based on the statistical patterns of the languages involved. In support of this alternative, a single recurrent neural network model that is trained on both Dutch and English sentences indeed predicts the cross-linguistic difference in grammaticality effect.

Keywords: bilingualism, cross-linguistic differences, sentence comprehension, relative clauses, centre embedding, grammaticality illusion, self-paced reading, recurrent neural network model

1 Introduction

Because of long-distance relations between the words of a sentence, readers or listeners often need to keep material in mind while waiting for future input that is to be connected to earlier parts of the sentence. The occurrence of a noun phrase, for example, is often predictive of a to-be-encountered verb phrase. Such a prediction can be costly to maintain in working memory, in particular when a subordinate clause is embedded in between the noun phrase and its verb. Embedding a second subordinate clause within the first yields a double centre-embedded sentence structure that requires the retention of three verb-phrase predictions, making such structures notoriously difficult to process (Hamilton & Deese, 1971). An example (from Frazier, 1985) is sentence (1a), which constitutes a grammatical sentence in English yet is regarded by native speakers as very difficult to understand and is often judged unacceptable. In fact, native speakers rate the ungrammatical sentence in (1b), derived from (1a) by omitting the second verb phrase, as equally acceptable (Gibson & Thomas, 1999) or even as more acceptable (Christiansen & MacDonald, 2009) compared to its grammatical counterpart. This effect has also been observed in French (Gimenes, Rigalleau, & Gaonac'h, 2009) and in artificial grammar learning (De Vries, Geukes, Zwitserlood, Petersson, & Christiansen, 2012).

(1a) The apartment that the maid who the service had sent over was cleaning every week was well decorated.

(1b) *The apartment that the maid who the service had sent over was well decorated.

Gibson and Thomas (1999) argue that this phenomenon, which Vasishth, Suckow, Lewis, and Kern (2010) call the *grammaticality illusion*, arises from a process of structural forgetting caused by working memory capacity limitations. During the processing of sentences such as (1a), the working memory system is required to keep the predictions made by the three noun phrases active in order to integrate the three verb phrases that will follow. Gibson and Thomas (1999) conclude that the equal acceptability of sentences (1a) and (1b) is a result of one of the verb-phrase predictions being lost from working memory. With only two predictions left, the omission of the second verb phrase is not noticed; rather, the occurrence of three verb phrases where only two are expected may lead to parsing difficulty with the grammatical sentence. Crucially for the illusion to occur, it is the second verb that has to be omitted as the second noun phrase is most likely to be forgotten due to the relatively high memory load it causes (Gibson & Thomas, 1999).

Measuring reading times (as opposed to acceptability ratings), Vasishth et al. (2010) find additional evidence for the grammaticality illusion in English. In four self-paced reading and one eye tracking study, native English speakers who read sentences such as (1a) and (1b) show longer reading times on the grammatical compared to the ungrammatical versions. In other words, encountering an ungrammaticality makes them speed up rather than slow down. Interestingly, native German speakers reading German sentences analogous to the English materials show the opposite pattern: They are slower in the ungrammatical than in the grammatical condition, in both the self-paced reading and eye tracking studies. This difference between

English and German suggests that the grammaticality illusion is not a language independent consequence of universal working memory constraints.

Vasishth et al. (2010) argue that the language specific effect of grammaticality is related to the difference in word order between German and English. German has SOV order in subordinate clauses, which leads to a much higher frequency of verb-final constructions than in English, forcing the working memory system to keep verb-phrase predictions active more often and for a longer duration. This, in turn, makes working memory more robust against structural forgetting, hence the reversed effect of grammaticality compared to English.[1]

The first objective of the present study is to replicate the German results in another SOV language: Dutch. If Dutch speakers, like Germans, show slower reading for the ungrammatical sentences, this would support the idea that the difference between German and English is caused by differences in word order. Second, we investigate the working memory robustness account by testing whether German and Dutch native speakers show the grammaticality illusion when reading English. If the absence of the illusion in SOV languages is indeed due to a more robust working memory for SOV structures, we may expect this to carry over to a second language. In that case, German and Dutch native speakers should also show slower reading for ungrammatical sentences in English. Conversely, if German and Dutch speakers reading English behave like English natives, then this would mean that they process English using statistical information specific to that language. Hence, this would cast doubt on the working memory robustness hypothesis.

In a series of three self-paced reading experiments, we first investigate whether the grammaticality effect in Dutch mirrors that in German (Experiment 1). Next, we test German (Experiment 2) and Dutch (Experiment 3) native speakers in English. Dutch speakers tested in Dutch indeed behave like German speakers in their L1, but both groups' reading-time effects in English resemble those of native English speakers. This strongly suggests that the grammaticality illusion observed in reading times on English sentences is not caused by cross-linguistically invariant limits on verbal working memory capacity. In the Discussion, we propose an alternative explanation based on readers' sensitivity to differences between the statistical language patterns of English and Dutch/German. This explanation predicts that readers who have been more exposed to English as an L2 will display a stronger grammaticality illusion, which is indeed what we find. Moreover, we show that the grammaticality effects on reading times in Dutch and English are predicted by a recurrent neural network that has learned the statistical patterns of these two languages from large text corpora.

[1] The language's headedness may not be all there is to it: Gimenes et al. (2009) presented some evidence that French (which is head first, like English) behaves like German insofar as the reading times on double embedded sentences are concerned: The grammatical sentences were read more quickly than the ungrammatical versions. In contrast, the latter were rated as more understandable, which suggests that the subjective illusion of grammaticality may be dissociated from the grammaticality effect on reading times.

2 Experiment 1: Dutch speakers tested in Dutch

2.1 Method

2.1.1 Materials

Sixteen target sentences were constructed, based on Dutch translations of Vasishth et al.'s (2010) English items. Table 1 displays an example of a grammatical Dutch target item (all 16 target sentences are listed in the Appendix). As in Vasishth et al. (2010), ungrammatical versions were constructed by removing the second verb. If encountering the ungrammaticality causes processing difficulty, we would expect this to occur at the first post-verbal determiner (labeled 'Det1' in Table 1) because this is where the sentence becomes ungrammatical. Conversely, the grammaticality illusion will result in processing difficulty at the final verb ('V3') of *grammatical* sentences.

Vasishth et al.'s (2010) English stimuli included sentences in which the second noun phrase (i.e, 'the daughter' in the example of Table 1) was inanimate. We left these out because they can lead to semantic anomalies in the ungrammatical condition: When the second verb phrase is absent, the last verb will be attached to the second (i.e., inanimate) noun phrase, possibly resulting in anomalous subject-verb combinations. This may cause a slowdown in reading which could be mistaken for an effect of ungrammaticality, while it is in fact a semantic effect.

Dutch relative clauses are ambiguous between subject-relative and object-relative readings. Without clear contextual, semantic, or prosodic cues that support an object-relative reading, the subject-relative reading is strongly preferred (Mak, Vonk, & Schriefers, 2002, 2006). This means that word-for-word translations of Vasishth et al.'s (2010) English object-relative target sentences will mostly be understood as subject relatives in Dutch (we will get back to this issue in the General Discussion). To prevent the subject/object-relative clause ambiguity from affecting the comprehension process, an adverb was inserted directly after each relative pronoun. This disambiguates towards the subject-relative reading (see Mak et al., 2002, p. 67).

In addition, a three-word prepositional phrase was attached to the end of the sentence to capture any grammaticality effect that is not yet visible at the post-verbal noun phrase because of the delay caused by the spillover effect. Self-paced reading generally results in stronger spillover than does eye tracking (Frank, Monsalve, Thompson, & Vigliocco, 2013) and this may be true in particular when words are presented at a fixed location (Just, Carpenter, & Woolley, 1982), as they were in the current experiment (see Section 2.1.3).

The 16 target sentences were interspersed among 56 fillers. Each target and filler sentence was paired with a yes/no-comprehension question, intended to ensure that participants try to read for comprehension. For target sentences, different comprehension questions inquired about different described events so that participants could not learn to strategically direct their attention to just one part of these sentences. Correct answers to comprehension questions were evenly divided between Yes and No. Questions never referred to the second verb, and leaving this verb out did not change the correct answer, under the reasonable assumption

Table 1: Example of a grammatical Dutch target sentence and its English translation. The labels below the words in the second part of the sentence are used in the Results section to refer to word locations.

De	moeder	die	vrijdag	de	dochter	die	toen	de	zus
The	*mother*	*who*	*Friday*	*the*	*daughter*	*who*	*then*	*the*	*sister*
		vond	beangstigde	begroette	de	oma	op	de	driewieler.
		found	*frightened*	*greeted*	*the*	*granny*	*on*	*the*	*tricycle.*
		V1	V2	V3	Det1	N1	Det2	Prep	N2

'The mother who on Friday frightened the daughter who then found the sister greeted the granny on the tricycle.'

that the first noun (being the sentence's topic) has a much stronger representation than the second noun and will therefore always be interpreted as the subject of the final verb. The comprehension questions for the 16 targets are listed in the Appendix.

2.1.2 Participants

Twenty-five native Dutch-speaking students at Radboud University Nijmegen took part in Experiment 1. One participant who reported having reading difficulties was excluded from the analysis, leaving 24 participants (18 female; mean age 22.3 years).

2.1.3 Procedure

Reading times were collected using word-by-word self-paced reading, where each word was presented at a fixed, central screen location. This has the advantage over the more common moving-window technique (as well as eye tracking) that the participant does not receive information about the remaining sentence length, which confounds with grammaticality.

The experiment was prepared using the Linger software (http://tedlab.mit.edu/~dr/Linger/), which controlled stimuli presentation and response time recording.[2] After on-screen instructions and three practice sentences, the 72 sentences were presented in pseudo-random order such that no two consecutive sentences were target items. The comprehension question was displayed immediately after the button press on the sentence-final word.

[2]For one participant, only part of the data (including five target sentences) were recorded due to a technical malfunction.

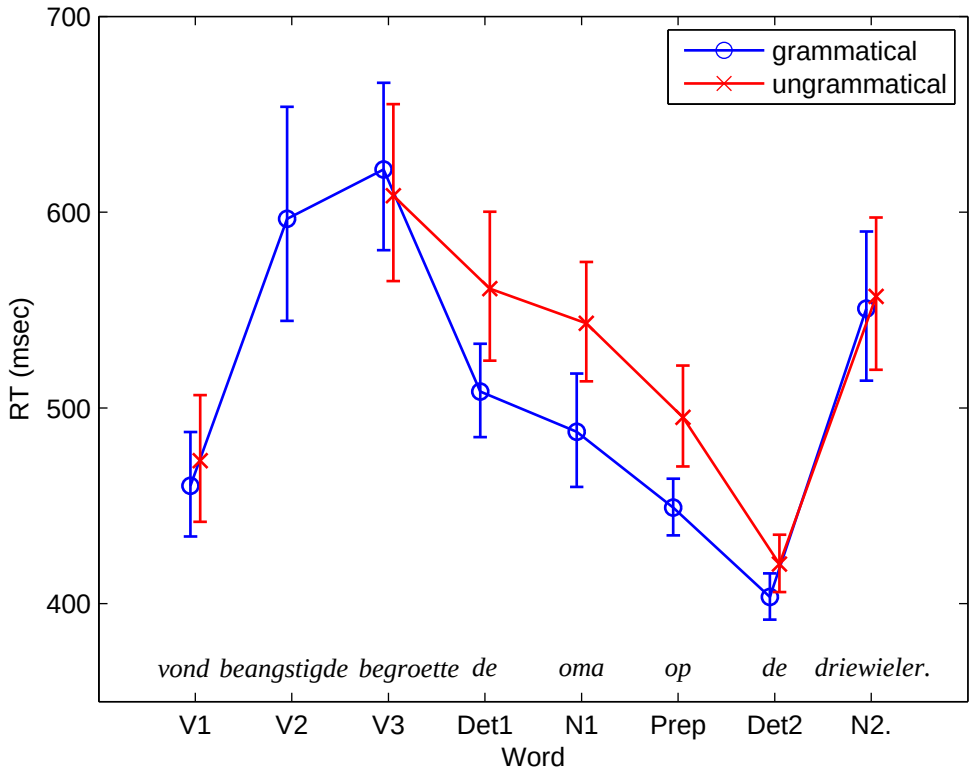

Fig. 1: Mean reading times at each point of target sentences (from the first verb onwards) for Dutch speakers reading Dutch (Experiment 1). V1, V2, and V3 denote the three verbs; Det1 and N1 are, respectively, the determiner and noun of the post-verbal noun phrase; Prep, Det2, and N2 are, respectively, the preposition, determiner, and noun of the sentence-final prepositional phrase. Error bars indicate 95% confidence intervals. Means and CIs were computed over log-transformed RTs.

2.2 Results

2.2.1 Accuracy

Per-participant percentages of correct answers to comprehension question varied between 72% and 99% (mean: 84.3%). The difference between grammatical and ungrammatical target items (79% and 69%, respectively) was marginally significant ($z = 1.95; p < 0.06$ in a logistic mixed-effects model).

2.2.2 Reading times

One item (0.3% of the data) was removed from the analysis because of an extremely long reading time of over 12 seconds.

Mean reading times over the course of the target sentences are displayed in Figure 1, which shows a slowdown on ungrammatical compared to grammatical sentences. In order to evaluate the pattern statistically,

hierarchical linear models were fitted to log-transformed reading times at each of the five word positions from V3 (where the two conditions begin to differ) until Det2 (the sentence's pre-final word). Linear mixed models were fit using `lme4` (version lme4_1.1-7; Bates, Mächler, Bolker, & Walker, in press), with Grammaticality as a fixed effect: The grammatical condition was coded as $+1$ and ungrammatical as -1, so negative coefficients indicate slower reading in the ungrammatical condition. These models always had varying intercepts and varying slopes for subject but no correlation was estimated, as including these often led to degenerate variance matrices. In several models, `lme4` was unable to estimate variance components for varying slopes by item, so these were excluded for all models. When varying intercepts by items had non-zero variance, these were included.

In addition to the frequentist analyses, we also fit Bayesian hierarchical models using Stan (Stan Development Team, 2014). This allows us to compute, given the data, the posterior probability (in other words, our degree of belief) of the coefficient being positive or negative. In these models, we fit a full variance-covariance matrix for subjects and for items (Barr, Levy, Scheepers, & Tily, 2013), including correlations. We used uninformative priors on the correlation matrices (so-called LKJ priors) in order to obtain priors for the variance-covariance matrices.[3]

For each word position of interest, Table 2 lists the estimated coefficient of the Grammaticality factor using the Bayesian models (the frequentist estimates were similar). From the Baysian models we also computed the 2.5th and 97.5th percentiles of the coefficient's posterior distribution, and the posterior probability that the coefficient is negative. The percentiles and $P(b < 0)$ were estimated by MCMC sampling.[4] The effect of grammaticality in the frequentist models was assessed using likelihood ratio tests (Pinheiro & Bates, 2000). Accordingly, we report the χ^2_1- and p-values.

An effect of grammaticality is apparent on the post-verbal noun: This word is read more slowly in the ungrammatical than in the grammatical condition. This effect may also be present on the previous and next words, although it is less reliable at those positions.

2.3 Discussion

Dutch speakers do not show the grammaticality illusion when reading Dutch: The ungrammatical sentences were read more slowly than their grammatical counterparts, from the point at which the sentence becomes ungrammatical (although the effect is strongest at N1). The German participants in the Vasishth et al. (2010) study also slowed down on the grammatical compared to the ungrammatical sentences, but showed this effect already at the matrix verb (V3), irrespective of whether reading times were measured by self-paced reading or eye tracking. Vasishth et al. (2010) explained this early effect by the fact that German (unlike Dutch) requires commas before the relative pronoun and after the verb that closes the relative clause. This

[3] The tutorial by Sorensen and Vasishth (2014) on fitting such models provides more details. Complete code and data for the present paper are available as online supplementary materials.

[4] All the Stan models had four chains, a warm-up (burn-in) of 500, and 2000 iterations. Convergence was checked visually and by using the Gelman-Rubin convergence diagnostic (Gelman et al., 2014).

Table 2: Regression analysis results for Experiment 1. A negative coefficient of the posterior mean b (in log milliseconds) indicates slower reading in the ungrammatical condition. The percentiles show 95% credible intervals computed from a Bayesian hierarchical model.

			Percentiles				
Region	b	SD	2.5th	97.5th	$P(b < 0)$	χ^2_1	p-value
V3	0.01	0.04	−0.06	0.08	0.37	0.14	0.71
Det1	−0.05	0.03	−0.11	0.01	0.94	3.14	0.08
N1	−0.06	0.03	−0.12	0.01	0.96	3.87	0.05
Prep	−0.05	0.03	−0.11	0.01	0.95	3.04	0.08
Det2	−0.02	0.02	−0.07	0.02	0.85	1.36	0.24

means that a comma is expected at the second verb. However, when V2 is missing, the second verb to appear is V3, which does not have a comma. It is because of the absence of this expected comma that reading slows down on V3 in German.

3 Experiments 2 and 3: German and Dutch speakers tested in English

If German and Dutch speakers process double-embedded structures in English as they do in their native language, we expect reading time patterns similar to those of Experiment 1: Reading difficulty occurs in the ungrammatical items at Det1, where the sentence becomes ungrammatical, but may be measured at a later word (i.e., N1) because of the spillover effect in the self-paced reading paradigm. Conversely, if German and Dutch speakers reading the English target sentences behave more like the English native speakers of the Vasishth et al. (2010) study, reading difficulty should occur at the V3 position in the *grammatical* condition, although it may only be measurable from Det1. Indeed, Vasishth et al. (2010) found that the effect of grammaticality appeared in the post-verbal region of English sentences in self-paced reading. When reading times were measured with eye tracking, however, the effect did appear on V3.

3.1 Method

3.1.1 Materials

Target sentences were the same as the English items of Vasishth et al. (2010) (using only animate nouns and the relative pronoun *who*) except that three-word sentence-final propositional phrases were attached, as in the Dutch items. One example is shown in Table 3, and all 16 target sentences are listed in the Appendix.

3.1.2 Participants

Forty-one native German speakers (34 female; mean age 24 years) took part in Experiment 2. All were students at the University of Potsdam.

Twenty-nine students of English Language and Culture at Radboud University Nijmegen took part in Experiment 3. Five were excluded from the analysis, either because their native language was not (only) Dutch or because they reported having reading difficulties. This left 24 participants (16 female; mean age 20.4 years).

3.1.3 Procedure

The procedure was identical to that of Experiment 1, except that the instructions were presented in English. Also, all participants completed an English proficiency test, comprising 30 four-choice 'fill in the blank' questions. The test questions are available from the authors.

3.2 Experiment 2 Results

3.2.1 Proficiency test

Due to a technical error, for the English proficiency test, only 15 of the 30 questions were presented to the German participants. These participants were asked two months later to fill out the full 30-questionnaire test. In this re-test, proficiency data were not gathered from four of the 41 participants because they did not respond to the request for a re-test. The correlation between the 15-question and 30-question proficiency scores was 0.80; that is, the scores were very stable across the two tests. The 37 German participants' accuracy on the 30-question English proficiency test ranged from 10 to 27 (mean: 20; SD: 4.6).

3.2.2 Accuracy

Two participants with exceptionally low comprehension question accuracy scores (62% and 69%) were excluded from further analysis. For the remaining 39 participants (33 female; mean age 24 years) accuracies varied between 72% and 95% (mean: 83.1%). The difference between error rates on grammatical and ungrammatical target items (66% and 68%, respectively) was not statistically significant ($z = 0.39; p > .6$ in a logistic mixed-effects model).

Table 3: Example of a grammatical English target sentence. The labels below the words in the second part of the sentence are used in the Results sections to refer to word locations.

The mother who the daughter who the sister								
	found	frightened	greeted	the	grandmother	on	the	tricycle.
	V1	V2	V3	Det1	N1	Det2	Prep	N2

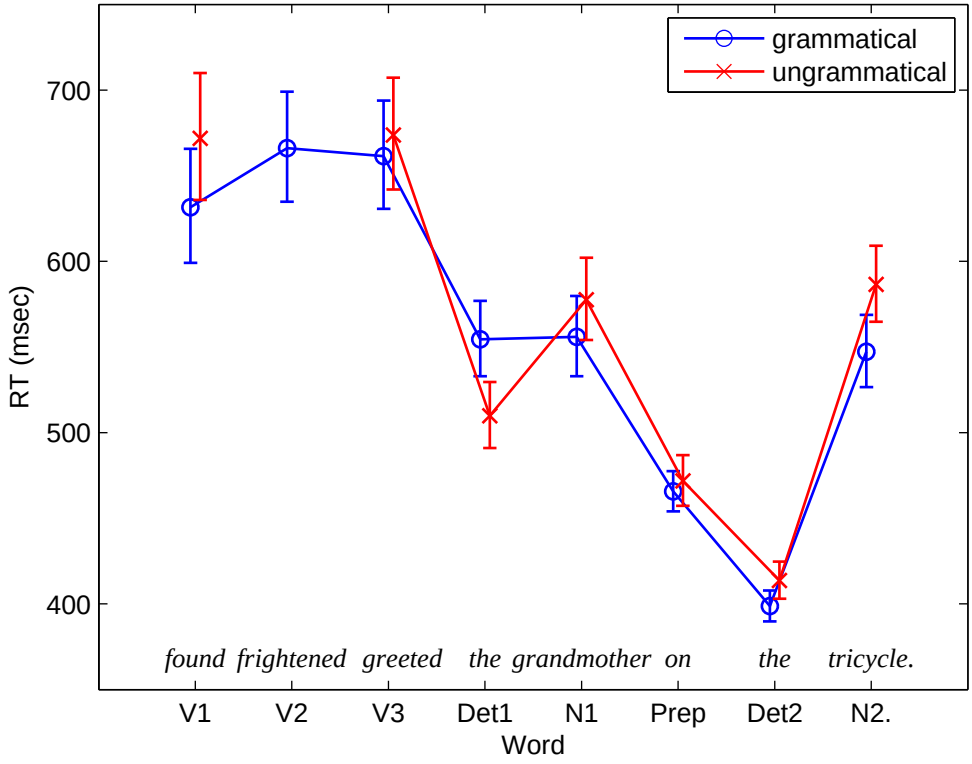

Fig. 2: Mean reading times at each point of target sentences (from the first verb onwards) for German speakers reading English (Experiment 2). V1, V2, and V3 denote the three verbs; Det1 and N1 are, respectively, the determiner and noun of the post-verbal noun phrase; Prep, Det2, and N2 are, respectively, the preposition, determiner, and noun of the sentence-final prepositional phrase. Error bars indicate 95% confidence intervals. Means and CIs were computed over log-transformed RTs.

3.2.3 Reading times

Ten items (1.6% of the data) were removed from the analysis because of extremely long (over 10 seconds) or short (under 40 msec) reading times.

Figure 2 shows the mean reading times for grammatical and ungrammatical target sentences, from the first verb onwards. In contrast to Dutch and German speakers reading in their L1, German speakers tested in English display the grammaticality illusion like English native speakers. This is confirmed by the regression analysis results in Table 4. As before, the regression models included Grammaticality as a fixed effect, as well as by-subject varying intercepts and slopes, and by-item varying intercepts (included when these were non-zero).

At Det1, reading was reliably slower in the grammatical compared to the ungrammatical condition. However, near the end of the sentence (at Det2) we find weak evidence for the reversed effect (slower reading

Table 4: Regression analysis results for Experiment 2 (German speakers reading English). A positive coefficient b indicates slower reading in the grammatical condition. The percentiles show 95% credible intervals computed from a Bayesian hierarchical model.

			Percentiles				
Region	b	SD	2.5th	97.5th	$P(b > 0)$	χ^2_1	p-value
V3	−0.01	0.02	−0.06	0.04	0.34	0.22	0.64
Det1	0.04	0.02	−0.00	0.09	0.97	7.06	0.01
N1	−0.02	0.02	−0.06	0.02	0.17	1.35	0.24
Prep	−0.01	0.02	−0.03	0.02	0.35	0.27	0.60
Det2	−0.02	0.02	−0.07	0.02	0.15	3.79	0.05

in the ungrammatical condition). In addition, a regression analysis at Det1 that included the main effect of Proficiency and the Grammaticality × Proficiency interaction revealed only very weak evidence that the grammaticality illusion is stronger for participants with higher English proficiency (Bayesian analysis: $b = 0.005$, 95% credible interval $[-0.004, 0.014]$, $P(b > 0) = 0.84$; likelihood ratio test: $\chi^2_1 = 1.30, p = 0.30$).

3.3 Experiment 3 Results

3.3.1 Proficiency test

Dutch participants' accuracy on the 30-question English proficiency test ranged from 25 to 30 (mean: 28; SD: 1.2). A Welch two-sample t test showed that the Dutch speakers had a significantly higher English proficiency score than the German speakers ($t(44) = 10.3; p < 0.0001$; 95% confidence interval for the difference in means: [6.5, 9.7]).

3.3.2 Accuracy

Two participants with exceptionally low comprehension question accuracy scores (68% and 47%) were excluded from further analysis. For the remaining 22 participants (15 female; mean age 20.5 years) accuracies varied between 71% and 97% (mean: 84.7%). The difference between error rates on grammatical and ungrammatical target items (69% and 68%, respectively) was not statistically significant ($z = 0.15; p > 0.8$ in a logistic mixed-effects model).

3.3.3 Reading times

One item (0.3% of the data) was removed from the analysis because of an extremely long reading time of over 80 seconds.

Figure 3 shows the mean reading times across the grammatical and ungrammatical target sentences. As was the case for the German participants in Experiment 2, there appears to be a strong effect of grammatical-

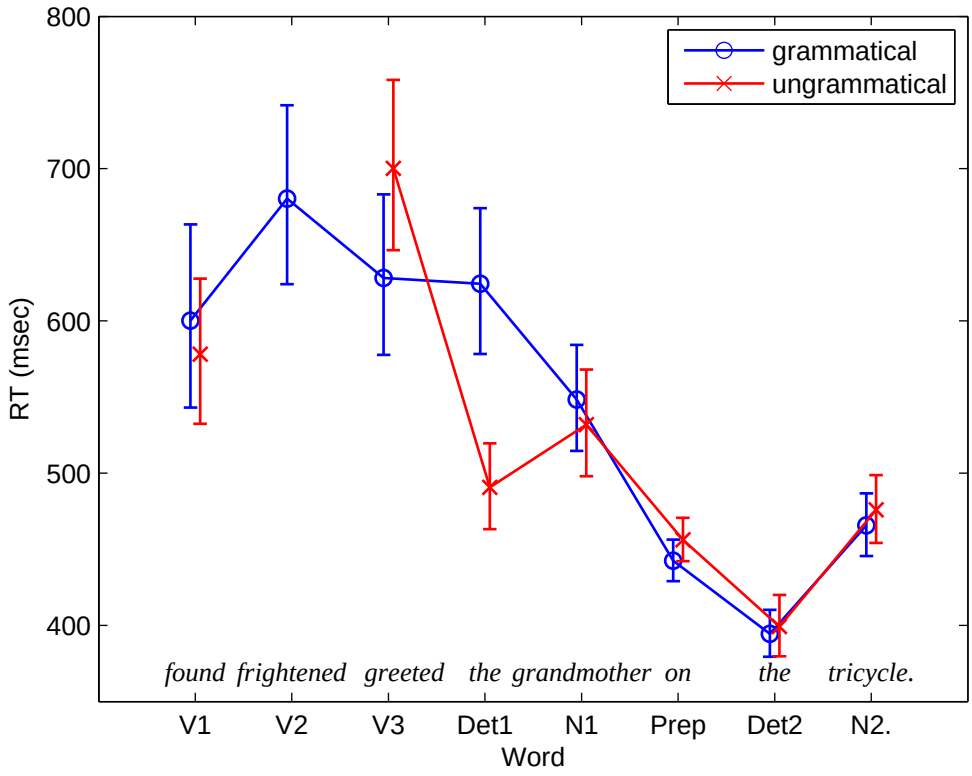

Fig. 3: Mean reading times at each point of target sentences (from the first verb onwards) for Dutch speakers reading English (Experiment 3). V1, V2, and V3 denote the three verbs; Det1 and N1 are, respectively, the determiner and noun of the post-verbal noun phrase; Prep, Det2, and N2 are, respectively, the preposition, determiner, and noun of the sentence-final prepositional phrase. Error bars indicate 95% confidence intervals. Means and CIs were computed over log-transformed RTs.

ity at the Det1 position, which is confirmed by the regression analyses (see Table 5). Reading times on Det1 are slower in the grammatical compared to the ungrammatical condition. Grammaticality and Proficiency did not interact at Det1 (Bayesian analysis: $b = -0.006$, 95% credible interval $[-0.09, 0.08]$, $P(b > 0) = 0.43$; likelihood ratio test: $\chi^2_1 = 0.13, p = 0.72$) but proficiency scores were near ceiling so no effect of Proficiency was to be expected.

3.4 Discussion

When tested in English, native German and Dutch speakers slow down on the grammatical relative to the ungrammatical sentences. In this respect, they behave like the native English speakers in the study by Vasishth et al. (2010) and unlike German and Dutch speakers reading in their L1. In further agreement with results from Vasishth et al.'s (2010) English self-paced reading experiments, the slowdown occurred on Det1

Table 5: Regression analysis results for Experiment 3 (Dutch speakers reading English). A positive coefficient b indicates slower reading in the grammatical condition. The percentiles show 95% credible intervals computed from a Bayesian hierarchical model.

			Percentiles				
Region	b	SD	2.5th	97.5th	$P(b > 0)$	χ^2_1	p-value
V3	−0.05	0.04	−0.13	0.03	0.10	2.75	0.10
Det1	0.12	0.04	0.05	0.20	1.00	10.66	< 0.002
N1	0.02	0.03	−0.04	0.08	0.70	0.43	0.51
Prep	−0.01	0.02	−0.05	0.02	0.17	1.00	0.32
Det2	−0.01	0.02	−0.05	0.04	0.41	0.10	0.75

rather than V3, most likely because of spillover.

4 General discussion

4.1 Language specificity of the grammaticality illusion

Vasishth et al. (2010) showed that the grammaticality illusion is apparent in the reading times of native speakers of English who read English double-embedded sentences: When the sentence was rendered ungrammatical due to a missing verb, reading was faster than when all three required verbs were present. In contrast, native speakers of German displayed the reversed effect: Reading was faster when the double-embedded sentences were grammatical than when a verb was missing. Our Experiment 1 replicated this study in Dutch, which, like German, is verb-final in subordinate clauses. The results for Dutch speakers tested in their L1 mirrored those for Germans: Reading times were lower on grammatical than on ungrammatical sentences.

In German orthography, commas are obligatory around relative clauses, possibly allowing participants to rely on a 'comma counting'-strategy for the German target sentences. Indeed, simulations with a connectionist model (Engelmann & Vasishth, 2009) suggest that commas may be crucial in German (which is difficult to test with human participants because of the obligatory presence of commas). Consequently, the difference between the German and English results could be due to these commas rather than to German being verb final.[5] However, such an explanation raises the question why the grammaticality effect would be the same in Dutch as it was in German. In Dutch, as in English, commas are not required around relative clauses and

[5] Vasishth et al. (2010) also tested English speakers on English sentences with 'German style' commas included and found that it made no difference to the results: The grammaticality illusion still appeared in English. The imperviousness of English to the presence of commas in such structures has also been shown to hold in connectionist simulations on an artificial language with English-like relative-clause structure (Engelmann & Vasishth, 2009). Nevertheless, as Vasishth et al. (2010) acknowledged, the possibility remains that English speakers do not rely on comma counting because these commas are not as reliable a cue in English as they are in German.

they were not present in the experimental items. Yet, the reading times showed the same grammaticality effect in Dutch as Vasishth et al. (2010) found in German. Hence, if the German results are due to the presence of commas, an alternative explanation is required for the findings in Dutch. One such alternative may be found in the semantic difference between our Dutch and English stimuli.

Dutch relative clauses are ambiguous between subject- and object-relative readings, but become unambiguously subject-relative when an adverb is introduced after the relative pronoun. Consequently, our Dutch target sentences necessarily contained only subject-relative clauses (SRCs). The English items, in contrast, were necessarily object relatives (ORCs). It is well known that SRCs are easier to process than ORCs, even without a difference in word order (Mak et al., 2002, 2006), possibly because they are much more frequent. The relative ease of SRCs may result in more successful memory retention of the three noun phrases (and, consequently, the absence of a grammaticality illusion) in the Dutch SRCs compared to the English ORCs. In that case, the apparent cross-linguistic difference would be an artifact of the semantic difference between the Dutch and English sentences. However, memory demands for processing the Dutch sentences are at least as high as for their English counterparts. This becomes clear when each verb's subject and object argument are indicated, as in sentences (2) and (3).

(2) Dutch: De moeder$_1$ die de dochter$_2$ die de zus$_3$ vond$_{2,3}$ beangstigde$_{1,2}$ begroette$_{1,4}$ de oma$_4$

(3) English: The mother$_1$ who the daughter$_2$ who the sister$_3$ found$_{3,2}$ frightened$_{2,1}$ greeted$_{1,4}$ the grandmother$_4$

The verbs' dependents are the same in the Dutch and English sentences (although subject and object are swapped for the first two verbs) so at each point in the sentence, the number of predicted heads and to-be-integrated dependents is identical for sentences (2) and (3). Consequently, memory load is the same across the two languages, both under Gibson and Thomas's (1999) explanation of the grammaticality illusion and under Dependency Locality Theory (Gibson, 2000). In fact, the Dutch target stimuli could result in higher memory load than the English sentences because of the adverbs following the relative pronouns (see Table 1). Under Lewis and Vasishth's (2005) memory-retrieval based sentence comprehension model, these intervening words result in working memory decay for the nouns and memory-retrieval interference at the verbs, making it more difficult for a verb to be integrated with its dependents. Self-paced reading and eye-tracking studies have indeed shown that increasing the distance between a verb and its dependent results in larger processing difficulty (Bartek, Lewis, Vasishth, & Smith, 2011; Grodner & Gibson, 2005).

The most parsimonious account of the reversed grammaticality effect in Dutch/German relative to English is one in which the same cause underlies both the Dutch and German effects. One obvious candidate is the languages' verb-final word order in relative clauses. Vasishth et al. (2010) argue that speakers of verb-final languages are more used to retaining a noun phrase in verbal working memory until the verb appears, which may make their working memory more robust to the structural forgetting hypothesized by Gibson and Thomas (1999). As a result, the sentence processing system detects the missing verb, causing a

slowdown in the ungrammatical condition. However, as we discuss below, our results suggest an alternative explanation that does not rely on working memory robustness but is based on the statistical word-order patterns of English on the one hand, and Dutch and German on the other.

4.2 The grammaticality illusion and language transfer

Our main finding is that when German and Dutch L1 speakers who are proficient in English are tested in English, they do display the grammaticality illusion like native English speakers. If the reversed grammaticality effect in German and Dutch is due to an increased ability to retain the three noun phrases in working memory (and, consequently, to expect three verb phrases), this ability apparently does not carry over from L1 to L2 sentence comprehension.

In second language learners, language transfer from L1 to L2 is a well known phenomenon in which aspects that are typical to L1 affect L2 comprehension or production. For example, White (1991) found that French speakers learning English regularly accept and produce English sentences with an verb-adverb order that is ungrammatical in English but correct in French. Other studies have shown that some aspects of syntax may not be susceptible to L1 transfer in highly proficient L2 speakers (Hoshino, Dussias, & Kroll, 2010; Ojima, Nakata, & Kakigi, 2005; Rossi, Gugler, Friederici, & Hahne, 2006). The case of the grammaticality illusion is particularly interesting in this respect because transfer from German or Dutch L1 to English L2 would be beneficial rather than detrimental. After all, having a more robust verbal working memory will not hurt, but probably help, English sentence comprehension.

It seems very unlikely that our German and Dutch participants have learned *not* to use a more robust verbal working memory for head-final structures when reading English, in order to achieve native-like (i.e., suboptimal) performance on unnaturally complex sentences. Three more viable explanations for the lack of L1 transfer remain. First, the ability to retain linguistic material in memory may be language specific, that is, when working memory is trained to retain German or Dutch noun phrases while waiting for their verbs, this ability does not generalize to another language (e.g., English). However, this explanation seems unlikely in light of evidence that a bilingual's working memory capacity is not language specific. For example, Lanfranchi and Swanson (2005) investigated the relation between working memory and vocabulary size of Spanish/English bilingual children. A number of working memory capacity measures were collected using tests in both Spanish and English, after which a factor analysis revealed a single working memory factor to underlie scores in both languages. Likewise, Keijzer (2013) presents evidence that the working memory capacity of adult L1 Dutch speakers who are fluent in L2 English is language independent. Osaka and Osaka (1992; Osaka, Osaka, & Groner, 1993) found the same for Japanese/English and German/French bilinguals.

Second, L2 processing may be more taxing on working memory than L1 processing, which leads to structural forgetting in English (as an L2) but not in German or Dutch (as L1). This is not the most parsimonious explanation, however, because it requires an alternative account for the grammaticality illusion in English native speakers. Moreover, the overall accuracy scores were nearly identical between Experiments

1 and 3, indicating that the Dutch participants had no particular difficulty with the English materials. Likewise, accuracy scores were numerically higher for our German and Dutch participants in L2 English than for Vasishth et al.'s (2010) native English speakers, both overall and for the target sentences.[6] Also, the grammaticality effects in Experiments 2 and 3 were not modulated by proficiency scores, suggesting that the effects are not caused by non-nativeness. If anything, the grammaticality illusion seems to be stronger for the more proficient English speakers: A post-hoc comparison between the German and Dutch groups (Experiments 2 and 3) revealed an interaction between Group and Grammaticality, such that the effect was stronger for the Dutch participants ($b = 0.04; \chi^2_1 = 5.86; p = .016$). Note that the Dutch participants were all students of English and, consequently, their average proficiency scores were significantly higher than those of the German group, who were students in areas such as Psychology and Linguistics. Moreover, our Dutch participants read a large amount of English literature as part of their study program, and much of the German students' English input comes from university textbooks and academic papers. Consequently, our participants' exposure to English is likely to be biased towards longer and more complex sentences. To the extent that their L2 reading is not native like, this bias will make the target sentences easier to process rather than more difficult. Hence, there is no reason to believe that their non-nativeness causes a grammaticality illusion in English.

Third, the cross-linguistic effect may be caused by differences in the languages' word-order patterns rather than differences in working memory requirements for the different languages. Native speakers and proficient L2 speakers could be similarly sensitive to the particular statistics of a language, which depends on its word order in subordinate clauses. For example, one is much less likely to encounter a three-verb sequence in English than in German or Dutch. Consequently, after having read two consecutive verbs, the occurrence of a third verb would be less predictable in English than in German/Dutch, which leads to slower reading. This may even be the case if that third verb is grammatically required (as in our target sentences) because local information can trump global, hierarchical structure when generating expectations about the upcoming word's syntactic category (Frank & Bod, 2011).

4.3 The grammaticality illusion and language statistics

According to the language-statistics account of the grammaticality illusion, our non-native participants are sensitive enough to English word-order patterns for the illusion to emerge. Our account predicts that increased exposure to the statistics of English would lead to a more pronounced grammaticality illusion. Indeed, we did find that the illusion was stronger for the Dutch participants (who were students of English) than for the German group.

The effect of language exposure on subsequent sentence processing can be illustrated by computational models that extract statistics from training sentences and use these to process novel input. MacDonald and

[6] However, the accuracies cannot directly be compared between the two studies because some of the questions for our target sentences referred to the sentence-final prepositional phrase, which was not present in Vasishth et al.'s (2010) materials.

Christiansen (2002; Wells, Christiansen, Race, Acheson, & MacDonald, 2009) used a simple recurrent neural network (SRN) to show that individual differences in relative clause processing can be caused by differences in language exposure rather than differences in working memory capacity. Likewise, a connectionist modeling study into the grammaticality illusion (Engelmann & Vasishth, 2009) confirmed that a language's word-order patterns, in combination with the learning and processing biases of an SRN, can affect how double-embedded relative clauses are processed: The network became more sensitive to the grammaticality illusion when it was trained on a head-first compared to a head-final artificial miniature language.

More recently, Frank (2014) trained a single SRN on over 10 million Dutch and English sentences from natural text corpora, resulting in a bilingual model that was able to process the actual stimuli from our experiments and compute a surprisal value (Hale, 2001; Levy, 2008) for each word. These values correlated significantly with the word reading times on filler sentences of Experiments 1 and 3, showing that the network successfully captured relevant aspects of human sentence processing, both in Dutch and English.

We tested Frank's (2014) model on the Dutch and English target sentences, extracting surprisal values on the critical words V3 and Det1. As can be seen in Table 6, the model correctly predicts the general pattern of results: For Dutch, surprisal (i.e., predicted reading time) is lower at Det1 for grammatical compared to ungrammatical sentences, while there is no reliable difference on V3. This matches the reading times from Experiment 1. For English, the reversed pattern is observed: Surprisal is higher for grammatical than ungrammatical sentences. This effect is most reliable at Det1, which may appear to correspond to the results of Experiment 3. However, recall that we interpreted the reading time effect at Det1 as spillover from comprehension difficulty occurring at V3. As such, the model does not fully predict our findings. Nevertheless, it does demonstrate that a difference in language statistics can yield a reversal of the grammaticality effect between English and Dutch.

If, as we suggest, the differences in language statistics between English on the one hand and German and Dutch on the other is responsible for the difference in grammaticality effects, it may be possible to reduce (or even reverse) the grammaticality illusion in English relatively easily by repeatedly exposing participants to double-embedded structures. Language users are highly sensitive to the particular statistics of the language

Table 6: Mean surprisal values on critical words of Dutch and English target sentences in the grammatical and ungrammatical conditions, with results of likelihood ratio tests comparing surprisal across items between the two conditions.

		Mean surprisal			
Language	Word	grammatical	ungrammatical	χ^2_1	p-value
Dutch	V3	15.62	15.88	0.35	0.55
	Det1	3.49	4.40	16	< 0.001
English	V3	14.18	13.59	3.1	0.08
	Det1	2.46	2.15	5.01	0.03

environment and if these statistics change, readers' syntactic expectations can adapt surprisingly quickly (Fine & Jaeger, 2013). Fine, Jaeger, Farmer, and Qian (2013) demonstrated this by exposing participants to sentences with a local main-verb/relative-clause ambiguity (e.g, 'The soldiers warned about the dangers conducted the midnight raid'). Readers generally prefer the (much more frequent) main verb initial reading, as is apparent from a garden path effect in reading times when the sentence turns out to have relative clause. However, after reading just 18 locally ambiguous sentences with a 50/50 division between the two global readings, the participants' initial preference for the main verb reading had disappeared, as was apparent from the absence of a garden path effect.

If our participants, too, adapt to the statistics of the experimental items, we would expect that the exposure to double centre-embeddings over the course of the experiment makes such structures more acceptable, weakening the grammaticality effect in English.[7] In order to ascertain whether such weakening of the effect occurs in our data, we reanalyzed the reading times on Det1 in Experiments 2 and 3, this time including in the regression model an interaction term between Grammaticality and the number of grammatical target items presented to the subject so far. A negative coefficient for this term means that the grammaticality effect weakens over the course of the experiment, which would be indicative of syntactic adaptation to the experiment's language statistics. However, there was no such interaction in either the German ($b = -0.005$; $\chi^2_1 = 0.63; p = 0.43$) or Dutch group ($b = 0.003$; $\chi^2_1 = 0.07; p = 0.8$). Possibly, more exposure to double centre-embeddings in English is required to noticeably weaken the grammaticality illusion because the near absence of such structures may make the effect of exposure harder to detect than was the case for the structural ambiguities of the Fine et al. (2013) study. This may be true in particular for the Dutch participants, whose extensive experience with English will have resulted in prior beliefs about English sentence structure that are stronger and, consequently, harder to overcome than those of the German group.

5 Conclusion

In a series of self-paced reading experiments, we studied the relation between language structure and the grammaticality illusion; the phenomenon that English sentences with double-embedded relative clauses are read more quickly when they are made ungrammatical by removing one of the required verbs. In Experiment 1, we found that Dutch native speakers do not show the illusion when presented with sentences in Dutch, consistent with the results for German reported by Vasishth et al. (2010).

According to current theories of the grammaticality illusion in English, the underlying cause of the failure to notice the ungrammaticality is structural forgetting (Gibson & Thomas, 1999). Under this account, the reversed effects in German and Dutch reveal that speakers of these languages suffer less from structural forgetting than do speakers of English. This difference would be due to the head-final nature of German and

[7] Whether the same weakening effect should also occur in German or Dutch is unclear, because repeated exposure to the ungrammatical sentences does not make them more grammatical.

Dutch relative clauses, which forces German/Dutch speakers to keep verb predictions in mind more often and for longer duration. This, in turn, makes working memory more robust against structural forgetting. However, our second and third experiment showed that German and Dutch participants who were presented with English double-embedded sentences behaved as English speakers do: Reading times of the grammatical sentences exceeded those of their ungrammatical counterparts. This finding suggests that it is the difference between languages (i.e., the frequency of verb-final structures in English versus German/Dutch), rather than the difference between participants, that underlies the reversal of the grammaticality effect between English on the one hand and German and Dutch on the other. The lack of transfer from L1 to L2 sentence processing therefore sheds doubt on working memory capacity limitations as an explanation of the illusion in English. This does not imply that working memory suffices for accurate processing of double-embedded structures: The extent to which comprehension of these sentences taxes working memory cannot be determined from our current results.

We have argued that the grammaticality effects are caused by speakers' sensitivity to a language's word-order patterns. For example, one is much less likely to encounter a three-verb sequence in subordinate clauses in English than in German or Dutch and, given this distribution, the occurrence of a third verb phrase can be more surprising in English, even when the third verb phrase is required to form a grammatical sentence. Readers' sensitivity to the language specific statistical patterns can therefore result in the differences in reading times observed. This was confirmed by simulations using connectionist models that are exposed to languages with different word-order patterns (Engelmann & Vasishth, 2009; Frank, 2014).

The language-statistics account is highly parsimonious in that it assumes a single underlying cause for the effects in L1 German, L1 Dutch, L1 English, and L2 English. In contrast, alternative explanations for these findings rely on different causes for the effects in German versus Dutch, or in L1 versus L2 comprehension. In further empirical support of our account, we found a stronger grammaticality effect in English for the participant group that was more exposed to English. However, a longitudinal study is needed to investigate if the grammaticality illusion is indeed caused by exposure to English word-order patterns.

Acknowledgments

The research presented here was funded by the University of Potsdam, and by the European Union Seventh Framework Programme under grant number 334028 awarded to the first author. We would like to thank Manuel Augustin, Eline Bekkers, Merel Burghouwt, Rick Godschalk, Tessa van Ingen, Noud Maas, and Sytse van der Veen for their contribution to preparing and running Experiments 1 and 3. For Experiment 2, thanks go to Johanna Thieke.

References

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.

Bartek, B., Lewis, R. L., Vasishth, S., & Smith, M. (2011). In search of on-line locality effects in sentence comprehension. *Journal of Experimental Psychology: Learning, Memory and Cognition*, *37*, 1178–1198.

Bates, D., Mächler, M., Bolker, B., & Walker, S. G. (in press). Fitting linear mixed-effects models using lme4. *Journal of Statistical Software*. Retrieved from `http://arxiv.org/abs/1406.5823`

Christiansen, M. H., & MacDonald, M. C. (2009). A usage-based approach to recursion in sentence processing. *Language Learning*, *59*, 126–161.

De Vries, M. H., Geukes, S., Zwitserlood, P., Petersson, K. M., & Christiansen, M. H. (2012). Processing multiple non-adjacent dependencies: Evidence from sequence learning. *Philosophical Transactions of the Royal Society B: Biological Sciences*, *367*, 2065–2076.

Engelmann, F., & Vasishth, S. (2009). Processing grammatical and ungrammatical center embeddings in English and German: A computational model. In A. Howes, D. Peebles, & R. Cooper (Eds.), *Proceedings of 9th International Conference on Cognitive Modeling* (pp. 240–245).

Fine, A. B., & Jaeger, T. F. (2013). Evidence for error-based implicit learning in adult language processing. *Cognitive Science*, *37*, 578–591.

Fine, A. B., Jaeger, T. F., Farmer, T. A., & Qian, T. (2013). Rapid expectation adaptation during syntactic comprehension. *PLoS ONE*, *8*, e77661.

Frank, S. L. (2014). Modelling reading times in bilingual sentence comprehension. In P. Bello, M. Guarini, M. McShane, & B. Scassellati (Eds.), *Proceedings of the 36th Annual Conference of the Cognitive Science Society* (pp. 1860–1861). Austin, TX: Cognitive Science Society.

Frank, S. L., & Bod, R. (2011). Insensitivity of the human sentence-processing system to hierarchical structure. *Psychological Science*, *22*, 829–834.

Frank, S. L., Monsalve, I. F., Thompson, R. L., & Vigliocco, G. (2013). Reading time data for evaluating broad-coverage models of English sentence processing. *Behavior Research Methods*, *45*, 1182–1190.

Frazier, L. (1985). Syntactic complexity. In D. R. Dowty, L. Karttunnen, & A. M. Zwicky (Eds.), *Natural language parsing: Psychological, computational, and theoretical perspectives* (pp. 129–189). New York: Cambridge University Press.

Gelman, A., Carlin, J. B., Stern, H. S., Dunson, D. B., Vehtari, A., & Rubin, D. B. (2014). *Bayesian data analysis* (Third ed.). Chapman and Hall/CRC.

Gibson, E. (2000). The Dependency Locality Theory: A distance-based theory of linguistic complexity. In Y. Miyashita, A. Marantz, & W. O'Neil (Eds.), *Image, Language, Brain* (pp. 95–126). Cambridge, MA: MIT Press.

Gibson, E., & Thomas, J. (1999). Memory limitations and structural forgetting: The perception of complex ungrammatical sentences as grammatical. *Language and Cognitive Processes*, *14*, 225–248.

Gimenes, M., Rigalleau, F., & Gaonac'h, D. (2009). When a missing verb makes a French sentence more acceptable. *Language and Cognitive Processes*, *24*, 440–449.

Grodner, D., & Gibson, E. (2005). Consequences of the serial nature of linguistic input. *Cognitive Science*, *29*, 261–291.

Hale, J. T. (2001). A probabilistic Earley parser as a psycholinguistic model. In *Proceedings of the 2nd Conference of the North American Chapter of the Association for Computational Linguistics* (Vol. 2, pp. 159–166). Pittsburgh, PA: Association for Computational Linguistics.

Hamilton, H. W., & Deese, J. (1971). Comprehensibility and subject-verb relations in complex sentences. *Journal of Verbal Learning and Verbal Behavior*, *10*, 163–170.

Hoshino, N., Dussias, P. E., & Kroll, J. F. (2010). Processing subject-verb agreement in a second language depends on proficiency. *Bilingualism: Language and Cognition*, *13*, 87–98.

Just, M. A., Carpenter, P. A., & Woolley, J. D. (1982). Paradigms and processes in reading comprehension. *Journal of Experimental Psychology: General*, *111*, 228–238.

Keijzer, M. (2013). Working memory capacity, inhibitory control and the role of L2 proficiency in aging L1 Dutch Speakers of near-native L2 English. *Brain Sciences*, *3*, 1261–1281.

Lanfranchi, S., & Swanson, H. L. (2005). Short-term memory and working memory in children as a function of language-specific knowledge in English and Spanish. *Learning and Individual Differences*, *15*, 299–319.

Levy, R. (2008). Expectation-based syntactic comprehension. *Cognition*, *106*, 1126–1177.

Lewis, R. L., & Vasishth, S. (2005). An activation-based model of sentence processing as skilled memory retrieval. *Cognitive Science*, *29*, 375–419.

MacDonald, M. C., & Christiansen, M. H. (2002). Reassessing working memory: A comment on Just & Carpenter (1992) and Waters & Caplan (1996). *Psychological Review*, *109*, 35–54.

Mak, W. M., Vonk, W., & Schriefers, H. (2002). The influence of animacy on relative clause processing. *Journal of Memory and Language*, *47*, 50–68.

Mak, W. M., Vonk, W., & Schriefers, H. (2006). Animacy in processing relative clauses: The hikers that rocks crush. *Journal of Memory and Language*, *54*, 466–490.

Ojima, S., Nakata, H., & Kakigi, R. (2005). An ERP study of second language learning after childhood: Effects of proficiency. *Journal of Cognitive Neuroscience*, *17*, 1212–1228.

Osaka, M., & Osaka, N. (1992). Language-independent working memory as measured by Japanese and English reading span tests. *Bulletin of the Psychonomic Society*, *30*, 287–289.

Osaka, M., Osaka, N., & Groner, R. (1993). Language-independent working memory: Evidence from German and French reading span tests. *Bulletin of the Psychonomic Society*, *31*, 117–118.

Pinheiro, J. C., & Bates, D. M. (2000). *Mixed-effects models in S and S-PLUS*. New York: Springer-Verlag.

Rossi, S., Gugler, M. F., Friederici, A. D., & Hahne, A. (2006). The impact of proficiency on syntactic second-language processing of German and Italian: Evidence from event-related potentials. *Journal of Cognitive Neuroscience*, *18*, 2030–2048.

Sorensen, T., & Vasishth, S. (2014). *A tutorial on fitting Bayesian linear mixed models using Stan.* (Manuscript submitted for publication)

Stan Development Team. (2014). Stan modeling language users guide and reference manual, version 2.5.0 [Computer software manual]. Retrieved from `http://mc-stan.org/`

Vasishth, S., Suckow, K., Lewis, R. L., & Kern, S. (2010). Short-term forgetting in sentence comprehension: Crosslinguistic evidence from verb-final structures. *Language and Cognitive Processes*, *25*, 533–567.

Wells, J., Christiansen, M. H., Race, D. S., Acheson, D., & MacDonald, M. C. (2009). Experience and sentence processing: Statistical learning and relative clause comprehension. *Cognitive Psychology*, *58*, 250–271.

White, L. (1991). The verb-movement parameter in second language acquisition. *Language Acquisition*, *1*, 337–360.

A Target sentences

Dutch and English target sentences and associated comprehension questions, including English glosses and translations of Dutch items, are listed below. Brackets indicate the verb that was removed to create the ungrammatical version. Two words that are connected by an underscore were presented simultaneously in the self-paced reading task. Brackets and underscores were not visible to the participants.

Dutch

1. De timmerman die eergisteren de vakman die zaterdag de boer droeg [bezeerde] begeleidde de leerling in de tuin.
   The carpenter who two_days_ago the craftsman who Saturday the peasant carried [hurt] supervised the apprentice in the garden.
   'The carpenter who two days ago hurt the craftsman who carried the peasant on Saturday supervised the apprentice in the garden.'
   Q: Droeg de boer de vakman in de tuin? 'Did the peasant carry the craftsman in the garden?' N
2. De moeder die vrijdag de dochter die toen de zus vond [beangstigde] begroette de oma op de driewieler.
   The mother who Friday the daughter who then the sister found [frightened] greeted the granny on the tricycle.
   'The mother who on Friday frightened the daughter who then found the sister greeted the granny on the tricycle.'
   Q: Vond de zus de dochter? 'Did the sister find the daughter?' N
3. De arbeider die onlangs de huurder die toen de voorman zocht [verwondde] ondervroeg de herder in het kantoor.
   The worker who recently the tenant who then the foreman looked_for [injured] questioned the shepherd in the office.
   'The worker who recently injured the tenant who then looked for the foreman questioned the shepherd in the office.'
   Q: Werd de herder in het kantoor ondervraagd? 'Was the shepherd questioned in the office?' Y
4. De handelaar die net de zakenman die destijds de professor inhuurde [verwarde] irriteerde de investeerder in de ochtend.
   The trader who just_now the businessman who back_then the professor hired [confused] annoyed the investor in the morning.
   'The trader who just now confused the businessman who back then hired the professor annoyed the investor in the morning.'
   Q: Werd de investeerder door de handelaar geïrriteerd? 'Was the investor annoyed by the trader?' Y
5. De schilder die laatst de muzikant die al zo lang de vader miste [beschutte] kookte_voor de kunstenaar in de keuken.
   The painter who recently the musician who already so long the father missed [sheltered] cooked_for the artist in the kitchen.
   'The painter who recently sheltered the musician who missed the father for a long time already cooked for the

artist in the kitchen.'

Q: Kookte de schilder voor de kunstenaar? 'Did the painter cook for the artist?' Y

6. De saxofonist die zondag de trompettist die altijd de dirigent meebracht [afleidde] bedankte de violist in een toespraak.

   The saxophonist who Sunday the trumpeter who always the conductor brought_along [distracted] thanked the violinist in a speech.

   'The saxophonist who on Sunday distracted the trumpeter who always brought_along the conductor thanked the violinist in a speech.'

   Q: Werd de violist bedankt in een toespraak? 'Was the violinist thanked in a speech?' Y

7. De apotheker die woensdag de opticien die gisteren de vreemdeling zag [verontrustte] ondervroeg de klant aan de balie.

   The pharmacist who Wednesday the optician who yesterday the strager saw [troubled] questioned the customer at the counter.

   'The pharmacist who on Wednesday troubled the optician who saw the stranger yesterday questioned the customer at the counter.'

   Q: Werd de klant ondervraagd aan de balie? 'Was the costumer questioned at the counter?' Y

8. De schoonmaker die vanochtend de conciërge die snel de dokter herkende [verwondde] verraste de patiënt in de gang.

   The cleaner who this_morning the janitor who quickly the doctor recognized [hurt] surprised the patient in the hallway.

   'The cleaner who this morning hurt the janitor who quickly recognized the doctor surprised the patient in the hallway.'

   Q: Verraste de patiënt de schoonmaker? 'Did the patient surprise the janitor?' N

9. De danser die gisteren de zanger die laatst de toeschouwer bewonderde [bezeerde] beloonde de portier met een tientje.

   The dancer who yesterday the singer who recently the bystander admired [hurt] tipped the doorman with a tenner.

   'The dancer who yesterday hurt the singer who recently admired the bystander tipped the doorman with a tenner.'

   Q: Bewonderde de zanger de toeschouwer? 'Did the singer admire the bystander?' Y

10. De artiest die vanmorgen de sportman die soms de bewaker riep [irriteerde] instrueerde de nieuwslezer in de studio.

    The artist who this_morning the sportsman who sometimes the guard called [annoyed] instructed the newscaster in the studio.

    'The artist who this morning annoyed the sportsman who sometimes called the guard instructed the newscaster in the studio.'

    Q: Werd de nieuwslezer geïnstrueerd op de wc? 'Was the newscaster instructed in the bathroom?' N

11. De klerk die gisteren de ambtenaar die soms de bezoeker vergat [hielp] irriteerde de buurman op het stadhuis.

    the clerk who yesterday the civil_servant who sometimes the visitor forgot_about [helped] annoyed the neighbor

at the town_hall.

'The clerk who yesterday helped the civil servant who sometimes forgot about the visitor annoyed the neighbor at the town hall.'

Q: Werd de buurman geïrriteerd op het kantoor? 'Was the neighbor annoyed at the office?' N

12. De zoon die laatst de vader die vaak de leraar zag [stoorde] bezocht de grootvader in het bejaardenhuis.

    The son who recently the father who often the teacher saw [disturbed] visited the grandfather in the nursing_home.

    'The son who recently disturbed the father who often saw the teacher visited the grandfather in the nursing home.'

    Q: Bezocht de vader de grootvader? 'Did the father visit the grandfather?' N

13. De conducteur die vanmorgen de dirigent die altijd de arbeider negeerde [sloeg] bekritiseerde de musicus op het festival.

    The conductor who this_morning the choirmaster who always the worker ignored [hit] criticized the musician at the festival.

    'The conductor who this morning hit the choirmaster who always ignored the worker criticized the musician at the festival.'

    Q: Heeft de dirigent de arbeider genegeerd? 'Has the conductor ignored the worker?' Y

14. De verdediging die vanmiddag de aanklager die eventjes de spion aankeek [verraste] overtuigde de rechter in de rechtbank.

    The defence who this_afternoon the prosecutor who briefly the spy looked_at [surprised] convinced the judge in the courtroom.

    'The defence who this afternoon surprised the prosecutor who briefly looked at the spy convinced the judge in the courtroom.'

    Q: Was de verdediging in de rechtbank? 'Was the defence in the courtroom?' Y

15. De neef die eergisteren de zus die ooit de boer in detail beschreef [plezierde] haatte de oom na de familieruzie.

    The cousin who two_days_ago the sister who once the peasant in detail described [pleased] hated the uncle after the family_fight.

    'The cousin who two days ago pleased the sister who once described the peasant in detail hated the uncle after the family fight.'

    Q: Beschreef de boer de zus? 'Did the peasant describe the sister?' N

16. De schilder die toentertijd de muzikant die ooit een vriend sloeg [stoorde] bewonderde de dichter op het poëziefestival.

    The painter who back_then the musician who once a friend hit [disturbed] admired the poet at the poetry_festival.

    'The painter who back then disturbed the musician who once hit a friend admired the poet at the poetry_festival.'

    Q: Werd de dichter bewonderd in het zwembad? 'Was the poet admired in the pool?' N

English

1. The carpenter who the craftsman who the peasant carried [hurt] supervised the apprentice in the garden.
   Q: Did the carpenter supervise the apprentice in the garden? Y
2. The mother who the daughter who the sister found [frightened] greeted the grandmother on the tricycle.
   Q: Did the sister find the daughter? Y
3. The worker who the tenant who the foreman looked_for [injured] questioned the shepherd in the office.
   Q: Was the shepherd questioned in the office? Y
4. The trader who the businessman who the professor hired [confused] annoyed the investor in the morning.
   Q: Did the professor hire the businessman? Y
5. The painter who the musician who the father missed [sheltered] cooked_for the artist in the kitchen.
   Q: Did the father miss the musician? Y
6. The saxophonist who the trumpeter who the conductor brought_along [distracted] thanked the violinist in his speech.
   Q: Was the violinist thanked in a speech? Y
7. The pharmacist who the optician who the stranger saw [troubled] questioned the customer at the counter.
   Q: Was the customer questioned at the counter? Y
8. The cleaner who the janitor who the doctor recognized [hurt] surprised the patient in the hallway.
   Q: Did the doctor recognize the janitor? Y
9. The dancer who the singer who the bystander admired [hurt] tipped the doorman at the door.
   Q: Did the singer admire the bystander? N
10. The artist who the sportsman who the guard shouted_at [annoyed] instructed the newscaster in the studio.
    Q: Did the guard shout at the sportsman on the football field? N
11. The clerk who the bureaucrat who the visitor forgot_about [helped] annoyed the neighbor at the town_hall.
    Q: Did the bureaucrat annoy the neighbor at the office? N
12. The son who the father who the teacher saw [disturbed] visited the grandfather in the nursing_home.
    Q: Did the father see the teacher? N
13. The conductor who the choirmaster who the worker ignored [hit] berated the musician at the festival.
    Q: Was the musician berated at the concert hall? N
14. The defence who the prosecutor who the spy looked_at [surprised] convinced the judge in the courtroom.
    Q: Did the prosecutor look at the spy? N
15. The cousin who the brother who the peasant described [pleased] hated the uncle from the farm.
    Q: Did the brother describe the peasant? N
16. The painter who the musician who the friend liked [disturbed] admired the poet in the pyjamas.
    Q: Did the poet wear a suit? N