

Research Article

Cite this article: Chromý, J., Lacina, R., Brand, J. and Vanek, N. (2025). When the second language attracts but the first does not: A large-scale study of number agreement attraction in Czech learners of English. *Bilingualism: Language and Cognition* 1–12. <https://doi.org/10.1017/S1366728925100126>

Received: 30 March 2024

Revised: 21 March 2025

Accepted: 30 April 2025



Keywords:

agreement attraction; Czech; L2 processing; crosslinguistic differences; unified competition model

Corresponding author:

Jan Chromý;

Email: jan.chromy@ff.cuni.cz

  This research article was awarded Open Data and Open Materials badges for transparent practices. See the Data Availability Statement for details.

© The Author(s), 2025. Published by Cambridge University Press. This is an Open Access article, distributed under the terms of the Creative Commons Attribution licence (<http://creativecommons.org/licenses/by/4.0>), which permits unrestricted re-use, distribution and reproduction, provided the original article is properly cited.



When the second language attracts but the first does not: A large-scale study of number agreement attraction in Czech learners of English

Jan Chromý¹ , Radim Lacina^{2,3}, James Brand^{1,4} and Norbert Vanek^{5,1}

¹Faculty of Arts, Charles University, Prague, Czechia; ²Faculty of Arts, Masaryk University, Brno, Czechia; ³Institute of Cognitive Science, Osnabrück University, Osnabrück, Germany; ⁴Edge Hill University, Omskirk, United Kingdom and ⁵University of Auckland, Auckland, New Zealand

Abstract

We investigate agreement attraction effects in the L2 English of native speakers of Czech, a language that has little-to-no evidence of attraction effects. Our experiments involve two groups of participants. The first group ($N = 415$) participated in an L2 English-only experiment, and the second group ($N = 183$) participated in both L2 English and L1 Czech versions of the experiment (in a randomized order with a two-week interval). Standard attraction effects were observed in L2 English, contrasting with the absence of such effects in L1 Czech. Our results provide unique evidence that an L2 can significantly attract, even when the L1 does not. However, our results also revealed that the attraction effect in L2 English disappeared when the L1 Czech version was completed first. These findings are discussed in relation to the Unified Competition Model and the effects of L2-induced increases in working memory demands.

Highlights

- L1 Czech L2 English speakers show number agreement attraction effects in their L2.
- L1 experiment using translation equivalent stimuli does not yield such an effect.
- Unique evidence that an L2 can significantly attract, even when the L1 does not.

1. Introduction

Agreement attraction has been one of the most studied phenomena in the domain of the processing of speakers' first language (L1). It occurs when a verb inadvertently agrees with an unrelated noun, not the subject, as in sentence (1) below.

- (1) *The blanket on the babies were small. (Bock & Miller, 1991)

Sentence (1) is ungrammatical, since the verb *were* is marked for a plural noun, yet the element with which it ought to agree in number is the singular subject head *blanket*. What makes this sentence special is the presence of another noun phrase (NP) that directly precedes the verb, here *the babies*, which is plural. It has long been noticed by linguists (Bock et al., 2001; Bock & Cutting, 1992) and more traditional grammarians (Fowler, 1937) that English speakers make number agreement errors more often when the preceding noun, known as the *attractor*, is plural as opposed to singular (e.g., *the blanket on the baby*).

Furthermore, and crucially for our present investigation, agreement attraction has also been found during online processing. Wagers et al. (2009) took sentences with constructions similar to those in (2)–(5) and presented them to native speakers of American English using a moving-window self-paced reading paradigm.

- (2) The letter from the investigator allegedly was received in San Francisco in late March.
 (3) The letter from the investigators allegedly was received in San Francisco in late March.
 (4) *The letter from the investigator allegedly were received in San Francisco in late March.
 (5) *The letter from the investigators allegedly were received in San Francisco in late March.

Their results demonstrated that in the region following the verb (*received*), native English speakers were substantially slower when reading the ungrammatical sentences (4) and (5) in comparison to the grammatical sentences in (2) and (3), where the verb *was* is correctly marked for number. This highlights participants' sensitivity to the ungrammatical agreement pattern. However, what they also found was that this slowdown was significantly weaker when the attractor (*investigators*) matched the verb (*were*) in number (5) than when it did not (4). This *facilitatory interference* effect has since been replicated in the processing of English (Cunnings &

Sturt, 2018; Laurinavichyute & von der Malsburg, 2022; Parker & An, 2018; Tanner et al., 2014) as well as many other languages, including Arabic (Tucker et al., 2015, 2021), Armenian (Avetisyan et al., 2020), French (Franck et al., 2015; Franck & Wagers, 2020; Villata & Franck, 2020), German (Lago & Felser, 2018), Greek (Paspali & Marinis, 2020), Korean (Kwon & Sturt, 2016; Sturt & Kwon, 2023), Russian (Slioussar, 2018; Slioussar et al., 2022), Spanish (Lago et al., 2015) and Turkish (Lago et al., 2019; Turk & Logačev, 2021). Additionally, a meta-analysis of 16 studies conducted by Jäger et al. (2017) and a large-scale replication study by Jäger et al. (2020) have shown that number agreement attraction effects in comprehension are robust.

There is, however, a notable exception – Czech. Chromý, Lacina, et al. (2023) ran four web-based self-paced reading experiments to test how native speakers of Czech process structures that had been shown to elicit attraction effects in other languages. Using Bayesian analysis, they found no or negligibly small support for facilitation effects. Furthermore, Chromý, Brand, et al. (2023) conducted a direct experimental comparison between Czech and English. They created translation equivalents of Czech and English sentences of experimental items, where the presence of number agreement attraction was manipulated, as well as fillers. They tested L1 speakers of Czech and English to directly compare the strength of agreement attraction effects between the two languages. The data and their Bayesian models clearly showed support for no effect in Czech and strong support for an effect in English, comparable in size to those found in previous studies (e.g., Lago et al., 2015; Tucker et al., 2015; Wagers et al., 2009).

Two important questions emerge that are critical to our understanding of language processing: What are the crosslinguistic differences, if any, in processing number agreement within speakers of more than one language? And second, to what degree are cognitive resources modulated by the currently used language? One way these questions can be addressed is to investigate what happens in second language (L2) processing. Agreement attraction effects in comprehension have been shown in L2 processing studies, especially in speakers whose L1s exhibit attraction effects too. For example, Jegerski (2016) studied both native and non-native speakers of Spanish and found that highly proficient L1 English speakers of L2 Spanish exhibited native-like patterns when reading attraction sentences. Similar effects were documented by Lago and Felser (2018) in L1 Russian speakers of L2 German. Moreover, Tanner et al. (2012) documented agreement attraction effects in L1 Spanish speakers of L2 English using electroencephalogram (EEG) focusing on the P600 component.

In languages that do not mark number agreement on the verb, such as Chinese or Korean, the evidence for agreement attraction effects in L2 comprehension is mixed. For example, Chen et al. (2007) documented agreement attraction effects in L2 English for L1 Chinese speakers, and comparable effects were found in L2 English for L1 Korean speakers (Lee & Phillips, 2023; Lim & Christianson, 2015). Furthermore, Bian et al. (2021) demonstrated that L1 Chinese speakers of L2 English showed event-related potentials (ERP) patterns qualitatively similar to native English speakers. In contrast, Jiang (2004) failed to find behavioural evidence for agreement attraction effects in L1 Chinese speakers of L2 English. This contrast could be explained by the differences in L2 proficiency between participants across these studies or even low statistical power in the analyses due to a small participant sample ($n < 40$). Overall, it is an important point that even for L1 speakers of languages that do not explicitly mark number agreement, one may expect to see agreement attraction effects in L2 English, at least at high levels of L2 proficiency.

There still remains an open question in this domain. What happens when both the L1 and L2 have number agreement, but attraction effects have only been reliably observed in the second language (such as in Czech as L1 and English as L2)?

Research into this topic has direct implications for two sets of theoretical accounts that concern agreement attraction in L2 processing. The first account focusses on *L2-induced increases in working memory demands* (WMD) (Coughlin & Tremblay, 2013; Cunnings, 2017; Sagarra & Herschensohn, 2013) and the second account is the *Unified Competition Model* (UCM) (MacWhinney, 2001, 2005, 2008, 2018).

Under the WMD account, limitations in working memory give rise to attraction effects through retrieval cue interference (Dillon et al., 2013; Vasishth et al., 2008; Wagers et al., 2009). Cue-based retrieval rests on the idea that sentence comprehension is a memory retrieval task, where syntactic dependencies (e.g., number agreement) are resolved through retrieval cues in working memory (Engelmann et al., 2019; Lewis & Vasishth, 2005; Yadav et al., 2023). The retrieval process is governed by similarity-based interference, meaning that competing items in memory that match retrieval cues can cause delays or errors. In relation to number agreement, the underlying assumption is that the number feature of the verb serves as a cue to search backwards for the subject in working memory. Morphologically realized subject–verb agreement in English acts as an effective retrieval cue to establish an agreement dependency, but cue retrieval gets disrupted in agreement calculations from memory when the verb’s ungrammatical form, such as *were* in (1), is number matched with the attractor, such as *babies* in (1). If grammatical processing in L2 presents additional WMDs (Coughlin & Tremblay, 2013; Sagarra & Herschensohn, 2013) and increases the chances of retrieval interference (Cunnings, 2017), one could expect L2 speakers to exhibit relatively larger attraction effects compared to their processing in L1.

Under the UCM account, the processing of L2 structures is guided by mental maps and cognitive routines already calibrated by L1-based patterns and the corresponding L1 structures. Structural alignment boosts the reliability of cues when L1 and L2 forms match, but cue reliability diminishes when there is a mismatch. To understand how UCM is linked to agreement attraction, we first specify what counts as a cue, what the key crosslinguistic difference in cue weightings is and how cue weightings are assumed to interact with retrieval processes. Cues are linguistic markings. The markings connected to number agreement in Czech and English are singular versus plural distinction marked on the head noun, the attractor and the verb in sentences like (1). In sentence comprehension, Czech speakers in their L1 pay attention to whether the head noun’s number agrees with the verb’s, while placing little reliance on whether the attractor’s number agrees with the verb’s (see Chromý, Brand, et al., 2023; Chromý, Lacina, et al., 2023). In this sense, Czech can be characterized as *attraction-faint* because it shows decreased susceptibility to interference from the plural attractor (*babies* in [1]) due to a greater weight placed on the head noun (*the blanket* in [1]) as a retrieval cue to establish number agreement with the verb (*were* in [1]). English, unlike Czech, can be described as *attraction-prominent* because L1 English speakers place greater weight on the attractor as the retrieval cue when they process a reference to a number. If processing in the L2 follows mental maps strongly committed to L1 patterns (MacWhinney, 2001), one would expect Czech L2 speakers of English to process ungrammatical sentences like (1) free of attraction due to a *transfer of cue strengths* (MacWhinney, 1992: 381). In the current work, we assume that the number of the attractor is used to evaluate

agreement with the verb and, further, that the weighting of the attractor number as a cue in the agreement process can differ across languages or between L1 and L2.

In its original formulation, the UCM considers cues in terms of grammatical properties that readers or listeners can map to specific functions (e.g., agent selection, subject or object identification). Grammatical properties that the original UCM defines as cues can vary in strength across languages. These properties include word order, subject–verb agreement, object–verb agreement, case marking, prepositional case marking, stress, topicalization, animacy, omission and pronominalization (MacWhinney, 2005). As an extension, in this work, we propose that the attractor number is a cue for subject–verb agreement. This extension dovetails with new directions in the Competition Model (Zhao et al., 2025), which widens the scope of what can count as a cue in language comprehension.

In this study, we test whether number agreement features elicit attraction effects in Czech learners of English as a second language. If they do, our aim is to monitor what the intrusion profiles are in various agreement mismatches. We find data from this participant base an informative testbed for two reasons. The first reason is to see whether attraction effects surface in English, a language in which they typically *do*, but when processed by participants with L1 Czech, a language in which they typically *do not*. One hypothesis is that Czech eliminates attraction because richer morphological systems tend to diminish agreement violations, as suggested by results from production (Eberhard et al., 2005; Vigliocco et al., 1995) and also comprehension (Lago et al., 2015). An absence of attraction effects in L1 may persist in L2 processing given that structural overlaps allow L2 processing to build on L1-based patterns (MacWhinney, 2001). Alternatively, attraction could emerge in L2 if the integration of number features is less detailed during online sentence comprehension in L2. This idea aligns with the views that second language processing is often characterized by simpler or shallower representations (Clahsen & Felser, 2006); less detailed linguistic representations constrained by the capacity to what learners can process (Pienemann, 1998); and greater reliance on simpler, less specified linguistic structures (VanPatten, 1996).

The second reason is to examine processing patterns *within* participants in an attraction-prominent versus an attraction-faint linguistic system and reflect on the potential differences within the frame of breakdowns in subject–verb agreement computations due to retrieval cue interference (Dillon et al., 2013; Wagers et al., 2009). To spell out the potential cue weighting differences, a number-matched attractor and verb in English ungrammatical sentences presents a cue with a stronger weight in agreement computations, compared to a number-mismatched attractor and verb. Assuming that L2 learners and native speakers share the cognitive principles underlying language processing, including reliance on working memory and cue weighting (MacWhinney, 2001, 2005, 2008, 2018), Czech L1-based cue weighting may be warped to resemble English L2-based cue weighting when number mismatches are processed in L2 English. Under this scenario, learners, especially those at the higher end of L2 proficiency, may develop new routines that are not eclipsed by the L1 patterns. Alternatively, learners may rely more strongly on L1 processing routines, which would surface as slowdowns in agreement violations, but not attraction effects.

2. Current study

In the current study, we set out to test whether readers with L1 Czech and L2 English exhibit number agreement attraction effects in reaction times (RTs) when reading sentences in their L2. We used a web-based self-paced reading (moving-window) experiment and

examined two groups of participants. The first group participated solely in an L2 English experiment (“L2 English only”). The second group (“both languages group”) participated both in an L2 English and in an L1 Czech experiment using translation equivalents as in Chromý, Brand, et al. (2023). All participants completed a lexical test for advanced learners of English (LexTALE) (Lemhöfer & Broersma, 2012), providing us with a measure of their L2 English proficiency. The first analysis focusses on the group which did not participate in any other attraction experiment. The second analysis then examines the group of participants who participated in both the L1 Czech and the L2 English experiments. This allows us to directly compare the strength of attraction effects in L1 and L2, while keeping the participant groups constant.

Our analyses aim to address four research questions. RQ1 and RQ2 are addressed in the first round of analyses, and RQ3 and RQ4 in the second round.

1. To what extent do Czech learners of English exhibit agreement attraction effects when processing number information in L2 English?
2. How strongly does L2 proficiency among highly proficient learners influence the magnitude of agreement attraction effects?
3. What is the size of the difference in number agreement attraction effects in L1 (Czech) and L2 (English)?
4. Does exposure to agreement attraction in one language influence the processing of translation equivalents in another language?

3. Data availability

The data from this experiment and the scripts used for the analyses are available via the Open Science Framework platform at <https://doi.org/10.17605/OSF.IO/2KJY5> in the folder *L2 English experiment* (“When the second language attracts but the first does not” paper).

4. Method

4.1. Participants

Altogether, we recruited 905 participants from a university-wide student participant pool at Charles University, Czech Republic, who participated for course credit. All participants reported that they were native speakers of Czech, with varying levels of English proficiency. 123 participants were excluded due to their accuracy of comprehension questions being lower than 70%. This was taken to exclude participants who may not have read the stimulus sentences during the experiment carefully enough or participants who may not have understood the presented materials. Moreover, we excluded 184 participants whose LexTALE score was under 65. This exclusion criterion was applied to ensure that the two groups (the L2 English-only group and both-languages group) are as comparable as possible in terms of L2 proficiency and to only sample participants with relatively high L2 proficiency.

The final samples were as follows. The group which participated only in the L2 experiment consisted of 415 participants (335 female, 76 male and 4 participants who did not disclose their gender; mean age = 22.97 years; and the mean LexTALE score = 79.98, *SD* = 9.17). The group that participated both in Czech L1 and English L2 experiments comprised 183 participants (141 female, 38 male and 4 who did not disclose their gender; mean age = 21.67; and the mean LexTALE score = 82.3, *SD* = 9.99). There were 91 participants out of this sample who completed the L1 Czech experiment first, then approximately 2–3 weeks later completed the L2 English

experiment (L1-L2 order). The remaining 92 participants completed the L2 English experiment first, with the L1 Czech experiment completed approximately 2–3 weeks later (L2-L1 order). The Language Experience and Proficiency Questionnaire (LEAP-Q) (Marian et al., 2007) results for both groups are presented in [Supplementary Materials](#) as an RMarkdown output. The L1 experimental data of the L1-L2 order group comes from Chromý, Brand et al. (2023).

4.2. Materials

For the L2 English version of the experiment, we used the items reported in Experiment 2 of Chromý, Brand, et al. (2023), which contained stimuli designed to test number agreement attraction in native English speakers. The experimental items varied in terms of two factors – VERB NUMBER and ATTRACTOR NUMBER – with both factors presented as either singular or plural. The conditions with plural verbs were ungrammatical. There was a total of 24 experimental items, each presented in 4 iterations to different participants. An example item is presented in [Table 1](#).

Each item was followed by a yes–no comprehension question targeting a constituent from the sentence, e.g., *Did the authorities seal the file?* We also used 96 filler items, with comprehension questions. These were again the same as in Chromý, Brand, et al. (2023). The filler items included a variety of syntactic structures and sentence lengths. All filler items were grammatical, resulting in an overall ratio of 0.1 ungrammatical sentences throughout the experiment.

The experimental items were found to elicit standard attraction effects in L1 speakers of English based on the results reported in Chromý, Brand, et al. (2023) and are also comparable to other retroactive interference structures used in previous research, e.g., Wagers et al. (2009). Therefore, we have reason to believe that any lack of effects that could be observed would not be due to confounds related to the construction of the items. We expected attraction effects to appear in the critical plural–plural condition (see example D in [Table 1](#)) in the post-verbal region (i.e., region 8).

The second group of our participants (both-languages group) participated not only in an L2 English experiment but also in an L1 Czech experiment which was identical to Experiment 1 in Chromý, Brand, et al. (2023). Crucially, the items (both experimental and filler) were translation equivalents of the English ones used in the L2 experiment.

4.3. Procedure

The experiment was web based using PCIBex Farm (Zehr & Schwarz, 2018), with all instructions presented in English. Participants first filled out a basic demographic questionnaire (age, gender, native language), a simplified version of the LEAP-Q (Marian et al., 2007) targeting their knowledge of other languages. Then, participants completed a LexTALE test (Lemhöfer & Broersma,

2012), which is an approximately 4-minute-long lexical decision task (60 trials), with participants responding to letter strings based on whether they believe it is an existing word in English or not. The percentage of correct responses is then recorded, and the participants are penalized for each false alarm (a yes–response to a non-word).

Participants then completed the main self-paced reading task, where they were instructed to press the space bar on their keyboard to move from one word to the next within the sentence. Once they had finished reading the sentence, they were presented with a yes–no comprehension question. All participants completed three practice trials, followed by 120 main trials (24 experimental items, 96 fillers). The entire experiment took approximately 25 minutes to complete.

The Czech version of the experiment was presented entirely in Czech and was otherwise identical to the English experiment, but without the English proficiency tasks. Participants in the “both languages group” took part in both the L2 English and L1 Czech experiment, with an interval between the experiments of approximately 14–21 days. They were presented with the same sentences they saw when they participated in the first version (either Czech or English) but presented as translation equivalents in the other language.

4.4. Statistical approach

We approached the analysis based on two different datasets – L2 only and L1 and L2. For all analyses, we used linear mixed-effects models using the linear mixed-effects models using S4 (*lme4*) package in the programming language R (Bates et al., 2014) to investigate the differences in reading times (RTs) in two different sentence regions (i) verb and (ii) verb+1. The analyses only used data from experimental items. The following variables were included in the analysis (the same acronyms are used both here and in the data and R scripts):

LOGRT was the dependent variable. For the purposes of the analysis, raw L2 data were filtered in the following way. First, we excluded all RTs under 100 ms. The remaining RTs were then log-transformed, and any logRTs that were 3 standard deviations above the mean, i.e., 7.375 log(ms) or 1596.37 ms, were also excluded. Altogether, 1.93% of the data were excluded (this was done for both participant groups). The same procedure was used with the L1 data. In this case, the upper cut-off value was 7.23 log(ms), i.e., 1382.29 ms. The trimming excluded 1.54% of the data points.

VERBNUM referred to the verb number and was either singular (*was*) or plural (*were*), with plural verbs always resulting in an ungrammatical sentence. For the mixed-effects modelling, we used sum contrast coding with –1 for singulars and 1 for plurals.

ATRNUM referred to the attractor number and was either singular (e.g., *archiver*) or plural (e.g., *archivers*). Importantly, the

Table 1. Example of an experimental item from the experiment

	1	2	3	4	5	6	7	8	9	10	11	12
A	The	file	for	the	archiver	likely	was	tightly	sealed	by	the	authorities.
B	The	file	for	the	archivers	likely	was	tightly	sealed	by	the	authorities.
C	The	file	for	the	archiver	likely	were	tightly	sealed	by	the	authorities.
D	The	file	for	the	archivers	likely	were	tightly	sealed	by	the	authorities.

Word regions are presented as column names. A: verb = single, attractor = single; B: verb = single, attractor = plural; C: verb = plural, attractor = single; D: verb = plural, attractor = plural.

sentence subject was always singular (e.g., *file*), so it was the plural attractor that caused possible interference. We again used sum contrast coding with -1 for singular and 1 for plural.

SCORE specified the participant's LexTALE score. In the model, it was centred and scaled. This variable was included just in the analysis of the L2 English-only group for two reasons. First, including it in the analysis of the both-languages group would have introduced a four-way interaction that would be too complex to interpret. Second, the L2 English-only group ($n = 415$) provides a sufficiently large sample to reliably examine the effects of SCORE and its potential interactions.

EXPERIMENT was used as a variable only for participants who completed both the L1 Czech and L2 English versions of the experiment, where the results were directly compared. We used treatment contrast coding with 1 for the L1 Czech version and 0 for the L2 English version.

EXPORDER was also used only in the comparison of L1 and L2 versions. It codes the order in which the participants completed the experiment versions, i.e., L1-L2 indicates the Czech version was completed first, then the English version; L2-L1 indicates the English version first, followed by the Czech version. We applied treatment contrast coding with different baselines for the two analyses. For the Czech data, we coded L1-L2 as 0 and L2-L1 as 1 , whereas for the English data, the coding was inverted (L2-L1 as 0 and L1-L2 as 1). In both cases, the baseline corresponds to the order in which the first experiment was conducted in the language targeted by the analysis.

As random effects, participant and item were always used. The random slope structure was set based on the recommendations by Matuschek et al. (2017). First, models with all possible random slope structures that converged and did not contain singularities were calculated. Second, these models were compared using Akaike information criterion (Akaike, 1998) and the model which was evaluated as the best one was then selected.

Since we analyzed two regions in each analysis, we were faced with a risk of inflating the chances of type I error. Therefore, Bonferroni correction was always applied (cf. von der Malsburg & Angele, 2017) and the p -value threshold was set to 0.025 ($0.05/2$).

5. Analysis 1: L2 English only

5.1. Results

Figure 1 shows the raw RTs for all sentence regions.

The linear mixed-effects models included ATRNUM, VERBNUM and SCORE as well as their interactions as fixed effects. The analysis of the verb region (with ATRNUM and SCORE as a random slope for items) yielded three significant effects: (i) SCORE ($\beta = -0.038$, $SE = 0.011$, $t = -3.529$, $p < 0.001$); (ii) VERBNUM ($\beta = 0.013$, $SE = 0.003$, $t = 4.391$, $p < 0.001$); and (iii) an interaction between VERBNUM and ATRNUM ($\beta = -0.007$, $SE = 0.003$, $t = -2.511$, $p = 0.012$). The 3-way interaction between SCORE, VERBNUM and ATRNUM was not significant ($\beta = 0.005$, $SE = 0.003$, $t = 1.729$, $p = 0.083$).

Similar effects were documented also in the verb+1 region (where the best model based on AIC contained only random intercepts, but no random slopes). There was a significant effect of SCORE ($\beta = -0.06$, $SE = 0.015$, $t = -4.083$, $p < 0.001$), VERBNUM ($\beta = 0.014$, $SE = 0.003$, $t = 4.140$, $p < 0.001$) and the interaction between VERBNUM and ATRNUM ($\beta = -0.009$, $SE = 0.003$, $t = -2.605$, $p = 0.009$).

The model estimates (i.e., predicted logRTs) for both regions under analysis are presented in Figure 2. For both regions, the models revealed a significant interaction effect between VERBNUM and ATRNUM. To further explore these interactions, least squares means pairwise comparisons were conducted using the *ls_means()* function from the *lmerTest* package (Kuznetsova et al., 2017). Specifically, we examined the attractor number effects for sentences with plural verbs (i.e., possible facilitatory interference) and attractor number effects for sentences with singular verbs (i.e., possible inhibitory interference).

For the verb region, we found a significant difference for the combination of plural attractor + plural verb versus singular attractor + plural verb ($\beta = -0.02$, $SE = 0.01$, $t = -2.036$, $p = 0.047$), but no effect for singular verbs ($\beta = 0.008$, $SE = 0.01$, $t = 0.84$, $p = 0.405$). For the verb+1 region, the attractor number difference in sentences with plural verbs was not significant ($\beta = -0.015$, $SE = 0.01$, $t = -1.519$, $p = 0.129$), but there was a

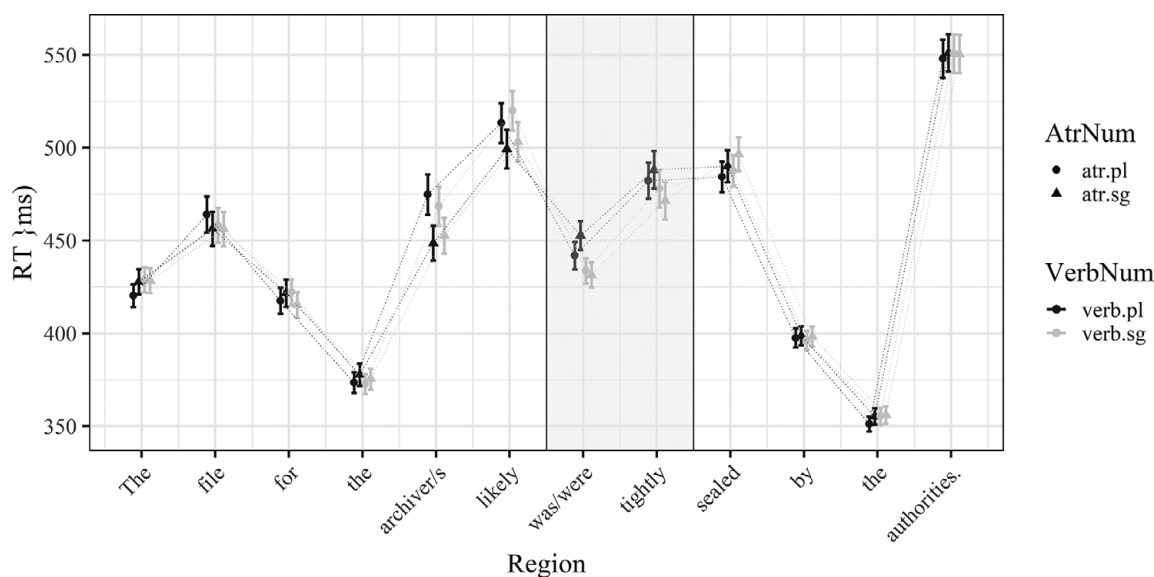


Figure 1. Raw RTs for all sentence regions in the L2 English-only data. Whiskers represent 95% confidence intervals. Colour is used to differentiate verb number (black = plural, grey = singular); shape is used to differentiate attractor number (circles = plural, triangles = singular). The rectangle highlights the regions under analysis (verb and verb+1).

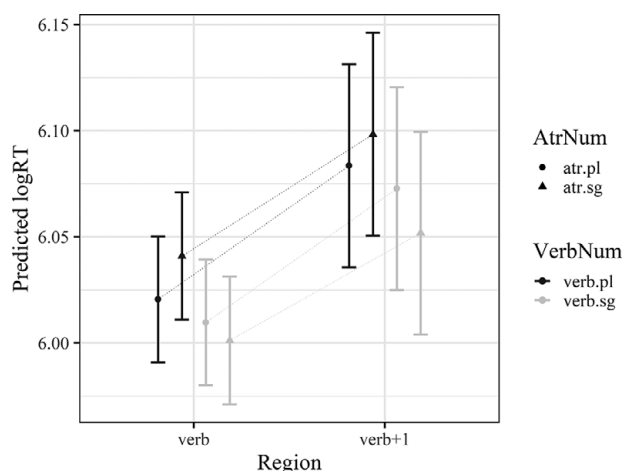


Figure 2. Plotted model estimates (predicted logRTs) for the two regions under analysis (verb, verb+1) for the L2 English-only group.

significant effect of attractor number in sentences with singular verbs ($\beta = 0.02$, $SE = 0.01$, $t = 2.142$, $p = 0.032$).

5.2. Discussion

Our analyses revealed several important findings. First, we found that L2 proficiency (LexTALE score) was related to processing speed. As proficiency increased, RTs were faster in both analyzed regions. Although this is not surprising, it supports the validity of LexTALE scores in capturing differences in L2 syntactic processing ability. Second, we found ungrammaticality (VERBNUM) effects in the verb region and in the region that immediately followed the verb. This may be interpreted as evidence of participants' sensitivity to agreement violations in their L2. However, we should bear in mind that the verb region differs in both length and frequency across conditions (*was* versus *were*) which presents a potential confound. Third, we documented interference effects from the attractor number in the verb and verb+1 regions. In the verb region, the model revealed a typical agreement attraction effect (see Wagers et al., 2009), i.e., a significant difference between the two

ungrammatical (i.e., plural verb) conditions (so-called *facilitatory interference*). When the plural verb was preceded by a plural attractor, reaction times were faster than when it was preceded by a singular attractor. In the verb+1 region, this type of interference did not reach significance. However, the model revealed a significant difference between conditions with singular verbs (grammatical ones). When the singular verb was preceded by a plural attractor, reaction times were slower than when it was preceded by a singular attractor. These results are striking since such inhibitory interference effects were not attested in previous studies on agreement attraction in Czech (Chromý, Brand, et al., 2023; Chromý, Lacina, et al., 2023). Finally, we failed to document any differences in the magnitude of the attraction effect related to L2 proficiency.

Our two main findings – namely the link between L2 proficiency and processing speed, and the lack of evidence for a link between L2 proficiency and attraction/interference effects – enrich insights from previous work into the relationships between L2 proficiency and processing efficiency. This is particularly evident among groups of higher-proficiency L2 speakers, as demonstrated, for example, by L2 English speakers maintaining robust processing abilities under increased cognitive load (McDonald, 2006), L2 Spanish speakers resolving syntactic ambiguities (Dussias & Sagarra, 2007) and L2 English speakers using syntactic cues to resolve temporarily ambiguous syntactic structures (Kaan et al., 2019).

6. Analysis 2: both L1 and L2 languages

In the analysis of the second group of participants, we used linear mixed-effects models with EXPORDER, VERBNUM and ATRNUM as fixed effects as well as their interactions. We first present the analyses of the L1 Czech data, then the L2 English data, and finally compare data from the two experiments.

6.1. Results for L1 Czech

Figure 3 visualizes the raw RTs in the verb region and verb+1 in the L1 Czech experiment. The plotted model estimates (predicted RTs) are presented in Figure 5.

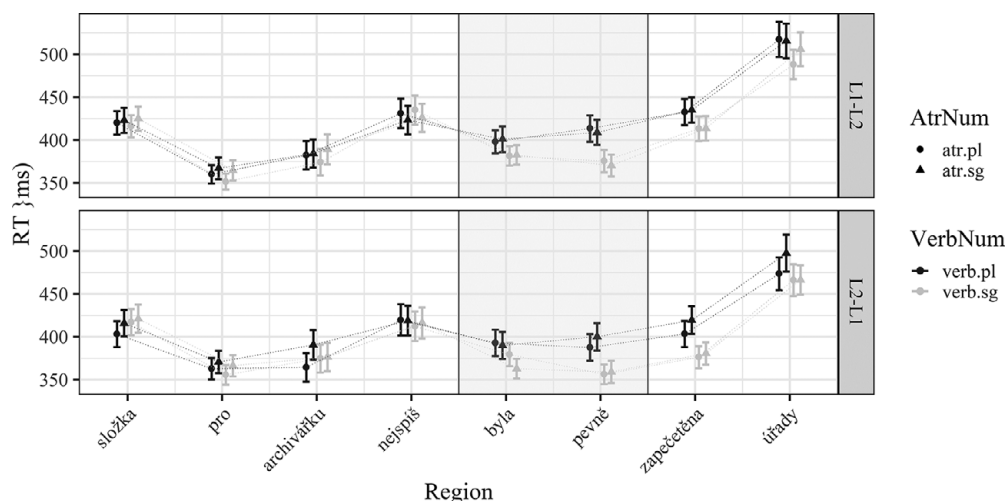


Figure 3. Raw RTs for all sentence regions in the L1 Czech data for both experimental orders (L1-L2 and L2-L1). Whiskers represent 95% confidence intervals. Colour is used to differentiate verb number (black = plural, grey = singular); shape is used to differentiate attractor number (circles = plural, triangles = singular). The rectangle highlights the regions under analysis (verb and verb+1).

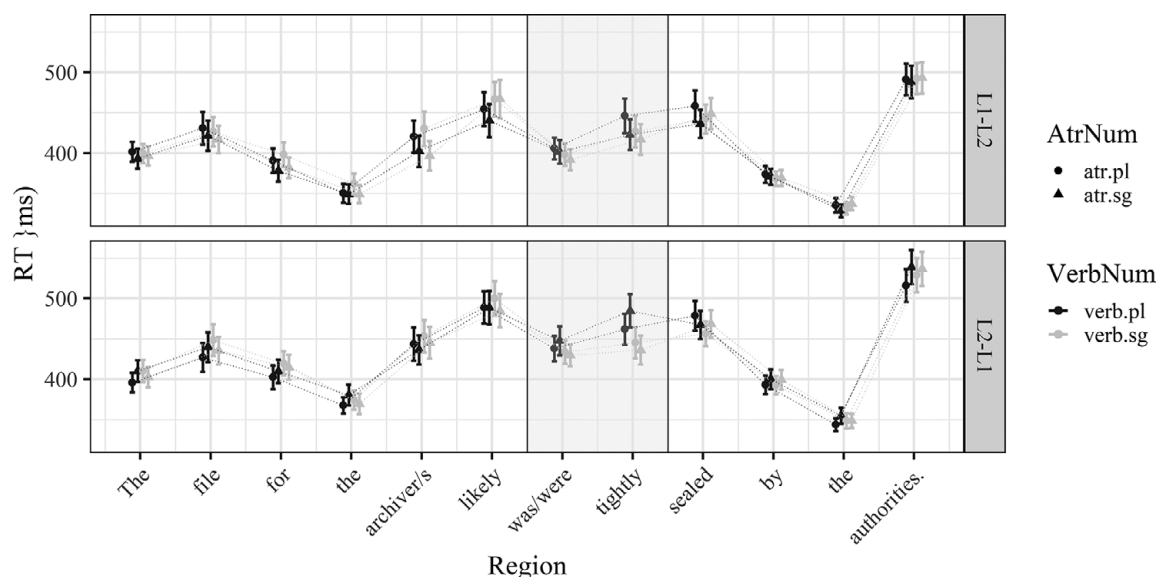


Figure 4. Raw RTs for all sentence regions in the L2 English data for both experimental orders (L1-L2 and L2-L1). Whiskers represent 95% confidence intervals. Colour is used to differentiate verb number (black = plural, grey = singular); shape is used to differentiate attractor number (circles = plural, triangles = singular). The rectangle highlights the regions under analysis (verb and verb+1).

The linear mixed-effects model revealed that only the VERBNUM effect was significant in both regions under investigation. For verb (where the best model contained no random slopes), we documented a significant VERBNUM effect ($\beta = 0.015$, $SE = 0.006$, $t = 2.645$, $p < 0.01$). For the verb+1 region (where VERBNUM served as a random slope for both participant and item), the model yielded only a significant effect of VERBNUM ($\beta = 0.042$, $SE = 0.007$, $t = 6.262$, $p < 0.001$). All the other effects were not significant.

6.2. Results for L2 English

The raw RTs for the two examined regions (verb and verb+1) are presented in Figure 4. The plotted model estimates (predicted RTs) are presented in Figure 5.

In the verb region, the linear mixed-effects model (containing only random intercepts) yielded only a significant effect of EXPORDER

($\beta = -0.09$, $SE = 0.036$, $t = 2.515$, $p = 0.013$), indicating that the L2-L1 group was slower than the L1-L2 group. In the verb+1 region, there was a significant effect of VERBNUM ($\beta = 0.037$, $SE = 0.007$, $t = 5.349$, $p < 0.001$) and a significant interaction between EXPORDER and VERBNUM ($\beta = -0.024$, $SE = 0.01$, $t = -2.467$, $p = 0.014$). The interaction between VERBNUM and ATRNUM was not significant after applying the Bonferroni correction ($\beta = -0.015$, $SE = 0.007$, $t = -2.156$, $p = 0.031$). Based on the visual inspection of the data, it seems that this interaction may be driven by the difference in conditions with plural verbs (i.e., facilitatory interference). We thus ran a simplified model targeting only the sentences with plural verbs. This model yielded a significant interaction between EXPORDER and ATRNUM ($\beta = 0.035$, $SE = 0.014$, $t = 2.463$, $p = 0.014$). Pairwise comparisons then revealed a significant difference between plural and singular attractors, but only in the L2-L1 EXPORDER ($\beta = -0.04$, $SE = 0.02$, $t = -2.03$, $p = 0.043$).

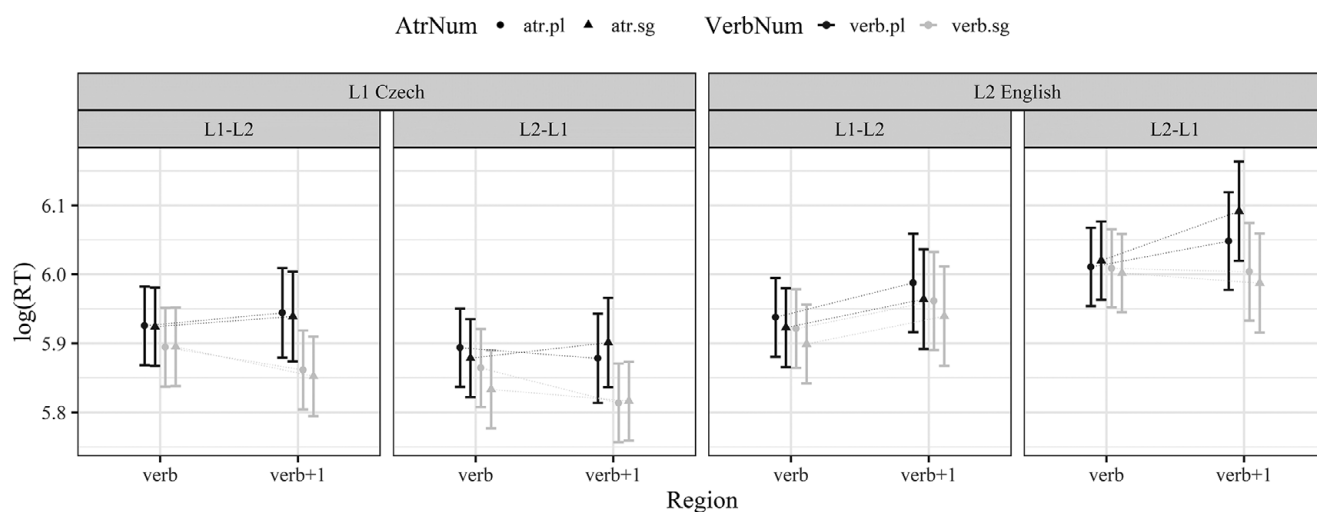


Figure 5. Predicted log RTs in the three analyzed regions together with their 95% confidence intervals in the L1 Czech and L2 English experiments.

6.3. Results for combined analysis

To examine the differences between the L1 Czech and L2 English experiments directly, we conducted an analysis of the combined data from both experiments. To simplify the analysis and to avoid the interference of task adaptation (Chromý & Tomaschek, 2024; Fine *et al.*, 2013; Prasad & Linzen, 2021), we included only data from the first experiment performed by each participant. We thus ran linear mixed-effects models with EXPERIMENT, ATRNUM and VERBNUM as fixed effects in interaction.

In the verb region, the model (including EXPERIMENT as a random slope for items) yielded a significant effect of EXPERIMENT ($\beta = -0.101$, $SE = 0.036$, $t = -2.758$, $p < 0.01$) indicating that the RTs were generally slower in the L2 experiment. No other effects reached statistical significance.

In the verb+1 region (where the model contained EXPERIMENT as a random slope for items), we again observed a significant effect of EXPERIMENT ($\beta = -0.013$, $SE = 0.045$, $t = -2.952$, $p < 0.01$). Additionally, the VERBNUM effect was significant ($\beta = 0.037$, $SE = 0.006$, $t = 6.005$, $p < 0.001$), and we documented a significant interaction between VERBNUM and ATRNUM ($\beta = -0.015$, $SE = 0.006$, $t = -2.397$, $p = 0.016$). No other effects were significant.

Pairwise comparisons revealed a significant difference between conditions with plural and singular attractors for sentences with plural verbs, but only within the L2 experiment ($\beta = -0.043$, $SE = 0.018$, $t = -2.429$, $p = 0.015$). The difference between plural and singular attractors for sentences with singular verbs was not significant in either experiment.

6.4. Discussion

The analysis of the participant groups who took part in both the L1 Czech and L2 English experiments revealed a range of findings that can offer new insights to our understanding of agreement attraction. First, we failed to find any effects in the L1 Czech experiment other than those of ungrammaticality (i.e., VERBNUM), which were significant for both regions (verb, verb+1). This is fully in line with previous studies on L1 Czech that also reported strong ungrammaticality effects, but no (or very weak) evidence supporting the presence of agreement attraction effects (e.g., Chromý, Lacina, *et al.*, 2023; Lacina & Chromý, 2022).

Second, in the L2 English experiment, we also observed ungrammaticality effects, but only in the verb+1 region. This suggests that L1 Czech speakers of L2 English were perceptive to ungrammaticality, although the effect was less prominent or delayed compared to their L1 processing. However, it is important to acknowledge that ungrammatical conditions were one character longer than grammatical ones (*were* versus *was*), which introduces a potential confound. The observed effects might therefore reflect, at least in part, differences in sentence length rather than solely sensitivity to ungrammaticality. We also found evidence for number attraction effects in the verb+1 region. In the full analysis, the interaction between attractor number and verb number approached significance. When we subsequently analyzed only plural verbs, we observed a significant effect of attractor number: Ungrammatical sentences with plural attractors were processed faster than those with singular attractors. While noteworthy, this effect appears to be relatively weak and should not be overinterpreted. One potential limitation is the relatively small sample size ($N = 92$), which may have reduced our statistical power to detect the expected effects. According to Jäger *et al.* (2017), a sample size of at least 120 participants may be necessary for sufficient power. We return to this finding in the General Discussion section in more detail.

Third, in both examined regions, the L2 experiment yielded slower RTs compared to those from the L1 experiment. This means that Czech participants were faster when reading in their L1 than in their L2, which was expected. Moreover, in the L2 experiment, we observed the effects of the order in which the experiments were completed. This suggests that those who participated initially in the L2 experiment were slower in this experiment than those who initially took part in the L1 experiment. This may be interpreted as a sign of task adaptation.

7. General discussion

In the current study, we examined whether L1 Czech speakers exhibit number agreement attraction effects in their comprehension of L2 English sentences and how this processing compares to their L1. We also tested whether exposure to sentences in either L1 or L2 has an effect on the processing of translation equivalents in the other language.

Overall, we find that when exposed to sentences with an agreement attraction configuration in L2 English, native Czech speakers show number agreement attraction effects. The participants speed up their reading of ungrammatical sentences when these include number-matching attractor nouns. Moreover, we also found evidence for a slowdown in their reading of grammatical sentences when the singular verb was preceded by a plural attractor (i.e., inhibitory interference). However, these effects were only observed when their first exposure to the stimuli was in their L2 English. There was no evidence that the effect was present when the Czech speakers had first read the stimuli in their native language and then retaken the experiment in their L2 English, even though the time between completing the L1 and L2 versions of the experiment was relatively long (two to three weeks).

Two main contributions emerge for second language research. First, our findings from a large sample of L1 Czech L2 English participants provide new evidence that the second language can show significant attraction effects even when the first does not. Second, in an experimental setting with a test in the attraction-faint L1 set two weeks apart from a test in the attraction-prominent L2, no attraction effect in the L2 emerged when the attraction-faint L1 was tested first. Evidence for attraction effects in L2 processing has been reported in numerous previous studies (e.g., Jegerski, 2016; Lago & Felser, 2018; Lee & Phillips, 2023; Tanner *et al.*, 2012). Under the WMD account, attraction effects arise due to L2-induced increases in WMDs (Coughlin & Tremblay, 2013; Sagarra & Herschensohn, 2013), which bring about higher chances of cue retrieval interference (Cunnings, 2017; Vasishth *et al.*, 2019). Our WDM-based prediction was that in L2 English, participants would exhibit stronger attraction effects in comparison with their L1 Czech. Results from the group that took the test in the L2-L1 order revealed that participants did indeed show attraction effects in their L2 despite not showing them in their L1. These results are interpreted as support for the WDM account.

But why was the attraction effect asymmetric? How can WMD explain its presence only when there is an L2-L1 test order? We advocate that cue retrieval calculations are more susceptible to disruptions in the weaker L2, but these disruptions get reduced when number agreement processing can benefit from earlier encounters with L1-entrained associations between the head noun and the verb. Associations formed in the stronger L1 can provide relief in memory demands for subsequent L2 processing. The logic behind this idea is that fast commitments to L1-entrained cue

retrieval can free up cognitive resources and lessen the likelihood of overlooking the ungrammaticality of the verb in attraction-prone L2 structures.

The finding that an attraction effect was absent in the L1-L2 testing order aligns with the UCM account (MacWhinney, 2001, 2005, 2008, 2018), under which L2 processing largely follows more entrenched L1-based patterns when the structures in the two languages overlap. However, this account requires another layer of explanation, since attraction seems to have changed as a result of the test order. One UCM-compatible explanation of why attraction emerged when the L2 came first is variation in weighting the attractor as the retrieval cue. In the context of processing number agreement, it is plausible that L2 readers may have placed greater weight on the attractor as it was the more recent retrieval cue, i.e., the one closer to the verb. Greater reliance on the more recently encountered cue could provide a cost benefit of not having to keep the less recent cue, the head noun, ready for retrieval. Instead of keeping the less recent cue readily available for retrieval, more cognitive resources could be used for processing rapid input in the weaker language. Unlike in the L2-L1 order, cue weights in the L1-L2 testing order could follow the L1-based pattern, where arguably more resources could be allocated to upholding the less recent cue in memory. Variation in cue weights in L2 versus L1 is likely under the assumption that second language processing is characterized by less specified linguistic representations (Clahsen & Felser, 2006; Kaan & Grüter, 2021; Pienemann, 1998; VanPatten, 1996), particularly during rapid online sentence comprehension. And variation in cue weights changing with L1-L2 versus L2-L1 task order can be explained as the upregulation of L1-based processing cues in the former and their downregulation in the latter context. In other words, it seems that our high-proficiency learners' "syntactic processing accent" (MacWhinney, 1992: 381) was more firmly L1 driven and transferred to the L2 when the L1 task came first, but no strong L1-based accent was traceable in L2 processing when the task context was initially in the second language. In sum, under the UCM, the observed asymmetry in attraction patterns can be attributed to differences in cue weighting. While the absence of attraction in L1 Czech can be explained by Czech readers paying attention to the head noun cue and placing less reliance on the attractor cue to process number agreement, this type of L1-based cue weighting was found to transfer to L2 English only when upregulated by the task performed in the learners' L1 Czech first.

One might wonder if attraction in L2 English was absent primarily because the attraction-faint Czech was tested first. It is possible that doing the same test twice within two weeks recalibrated participants' expectations and gradually minimized their surprise when reading structures that were initially unexpected. This process of habituation is known as syntactic adaptation (e.g., Chromý & Tomaschek, 2024; Fine et al., 2013; Prasad & Linzen, 2021), which could have decreased the magnitude of the effect in the second session to the point of disappearance. Even though syntactic adaptation in a second language may not be as strong as in the first (Kaan et al., 2019), together with participants adapting to the self-paced reading task itself, it could have affected the results in this study. One way to address this limitation and control for potential adaptation effects could be achieved by comparing the L2 performance of an L1 attraction-prominent group (e.g., L1 Spanish learners of L2 English) with that of an L1 attraction-faint group (e.g., L1 Czech learners of L2 English). A further beneficial addition to control for task adaptation as well would be via comparing L1 Czechs' L2 English performance with the same participants' L2

English performance two weeks later in a structure-insensitive self-paced reading (SPR) task to trace RT changes solely due to experience with the task itself.

Higher L2 proficiency scores often indicate approximations to target-like processing patterns (e.g., Chambers & Cooke, 2009; Dussias et al., 2013). However, our analyses showed that variation in L2 proficiency scores did not predict attraction strength. One plausible reason may be the homogeneity of the group in terms of a relatively narrow range of L2 proficiency levels represented in our sample. This finding further supports the idea that caution is due not to assume a relationship between L2 proficiency and approximation to target-like processing ability as automatic (e.g., Dijkgraaf et al., 2019; Domazetoska & Zhao, 2024; Kim & Grüter, 2021; Vanek et al., 2024). Nevertheless, considering that reading times in our study varied as a function of L2 proficiency, an alternative interpretation could be that research on L2 processing may require sampling from a broader range of linguistic abilities in the target language for an L2 proficiency effect on attraction to surface. Two potential limitations remain in this respect, namely that the proficiency range in our sample was restricted to higher scorers only, and that LexTALE is not the finest proficiency measure as it only captures word recognition ability rather than grammatical knowledge.

Let us now turn to the results of the L1 Czech experiments. Here, only ungrammaticality effects were documented, and no sign of number agreement attraction was detected. This is fully in line with the results of Chromý, Brand, et al. (2023), who used the same stimuli in their Experiment 1. This result lends further credence to the emerging findings that Czech is an anomalous language when it comes to attraction in comprehension, as either no or negligibly small effects have been reported (Chromý, Brand, et al., 2023; Chromý, Lacina, et al., 2023; Lacina et al., 2025; Lacina & Chromý, 2022). What our study contributes to the literature is that the lack of effects is solid and not influenced by participants having gone through the exact same stimuli in terms of content in their L2. Specifically, the same participants showed classic number agreement attraction effects in their L2 English first and then failed to show evidence of them in their L1 when completing the experiment two weeks later. Potentially fruitful extensions in this line of research could involve checking whether attraction effects vary as a function of participants' L2 experience in terms of print exposure and education (Dąbrowska, 2019) or if they change depending on the strength of syntactic engagement as new verbs are learned (García-Castro & Vanek, 2024).

Next, we spell out the implications that our research has for the study of agreement attraction more broadly. One of the most influential theories proposed to explain agreement attraction effects is the cue-based retrieval model (Engelmann et al., 2019; Lewis & Vasisht, 2005; Yadav et al., 2023). The model predicts that the same more general processing mechanisms, such as attending to cues or access and retrieval of stored information from memory, can combine with different cue weightings, which can vary across populations (Vasisht et al., 2019) as well as across different languages (Lago et al., 2015). Our results lend support to these predictions by showing that both English and Czech speakers attend to cues and weigh them, but the values of their weightings differ cross-linguistically. Another contribution of our findings lies in showing that modification to cue weightings in language-specific ways can occur within participants too. We show that for the same people, attraction was present in one of their languages and absent in the other. This suggests that participants' processing was modified in terms of the value of

their cue weightings that varied depending on the currently used language. Caution is needed not to overinterpret within-participant variation as it only surfaced in one (L2-to-L1) direction of testing. On the between-group level across languages, our findings from Czech and English inform cue-based retrieval models by providing evidence consistent with the idea that processing number agreement not only taxes working memory but also modifies processing through cue weighting in language-specific ways.

Finally, our finding that the order of two attraction experiments in L1 and L2 may cause the effect to disappear in one of the languages has implications for the study of the phenomenon overall. It suggests that the specifics of experimental set-up and design as well as participants' prior experience with attraction sentences can influence whether the effect emerges. While the causes of the disappearance of attraction in the L2 English experiment, when presented after the L1 version, may only be speculated on, they do point to factors such as familiarity and attention being relevant. This adds to the emerging literature which suggests that task specifications and other extra-linguistic factors may need to be accounted for when analyzing attraction experiments (Laurinavichyute & von der Malsburg, 2022, 2023).

Supplementary material. The LEAP-Q (Marian et al., 2007) results for both participant groups. Link: <https://osf.io/2kij5/files/osfstorage/676ed8ad3a4f07c0a1962f3b>.

Acknowledgements. This work was supported by the European Regional Development Fund project "Beyond Security: Role of Conflict in Resilience-Building" (reg. no.: CZ.02.01.01/00/22_008/0004595). The second author was supported by the German Research Foundation (DFG) as part of the Emmy-Noether grant awarded to Nicole Gotzner (Nr. GO 3378/1-1) and by the German Academic Exchange Service (DAAD, Scholarship programme Nr. 57694189). The third author was supported by a Primus grant (PRIMUS/21/HUM/015). The last author was supported by the University of Auckland Faculty Research Development Fund (FRDF) grant #3725366.

References

- Akaike, H. (1998). Information theory and an extension of the maximum likelihood principle. In E. Parzen, K. Tanabe, & G. Kitagawa (Eds.), *Selected papers of Hirotugu Akaike* (pp. 199–213). Springer. https://doi.org/10.1007/978-1-4612-1694-0_15
- Avetisyan, S., Lago, S., & Vasisht, S. (2020). Does case marking affect agreement attraction in comprehension? *Journal of Memory and Language*, 112, 104087. <https://doi.org/10.1016/j.jml.2020.104087>
- Bates, D., Mächler, M., Bolker, B., & Walker, S. (2014). Fitting linear mixed-effects models using lme4. *arXiv Preprint arXiv:1406.5823*.
- Bian, J., Zhang, H., & Sun, C. (2021). An ERP study on attraction effects in advanced L2 learners. *Frontiers in Psychology*, 12, 616804. <https://doi.org/10.3389/fpsyg.2021.616804>
- Bock, K., & Cutting, J. C. (1992). Regulating mental energy: Performance units in language production. *Journal of Memory and Language*, 31(1), 99–127. [https://doi.org/10.1016/0749-596X\(92\)90007-K](https://doi.org/10.1016/0749-596X(92)90007-K)
- Bock, K., Eberhard, K. M., Cutting, J. C., Meyer, A. S., & Schriefers, H. (2001). Some attractions of verb agreement. *Cognitive Psychology*, 43(2), 83–128. <https://doi.org/10.1006/cogp.2001.0753>
- Bock, K., & Miller, C. A. (1991). Broken agreement. *Cognitive Psychology*, 23(1), 45–93. [https://doi.org/10.1016/0010-0285\(91\)90003-7](https://doi.org/10.1016/0010-0285(91)90003-7)
- Chambers, C. G., & Cooke, H. (2009). Lexical competition during second-language listening: Sentence context, but not proficiency, constrains interference from the native lexicon. *Journal of Experimental Psychology: Learning, Memory, and Cognition*, 35(4), 1029–1040. <https://doi.org/10.1037/a0015901>
- Chen, L., Shu, H., Liu, Y., Zhao, J., & Li, P. (2007). ERP signatures of subject–verb agreement in L2 learning. *Bilingualism: Language and Cognition*, 10(2), 161–174. <https://doi.org/10.1017/S136672890700291X>
- Chromý, J., Brand, J., Laurinavichyute, A., & Lacina, R. (2023). Number agreement attraction in Czech and English comprehension: A direct experimental comparison. *Glossa Psycholinguistics*, 2. <https://doi.org/10.5070/G6011235>
- Chromý, J., Lacina, R., & Dotlačil, J. (2023). Number agreement attraction in Czech comprehension: Negligible facilitation effects. *Open Mind*, 7, 802–836. https://doi.org/10.1162/opmi_a_00107
- Chromý, J., & Tomaschek, F. (2024). Learning or boredom? Task adaptation effects in sentence processing experiments. *Open Mind*, 8, 1447–1468. https://doi.org/10.1162/opmi_a_00173
- Clahsen, H., & Felser, C. (2006). Grammatical processing in language learners. *Applied Psycholinguistics*, 27(1), 3–42. <https://doi.org/10.1017/S0142716406060024>
- Coughlin, C. E., & Tremblay, A. (2013). Proficiency and working memory based explanations for nonnative speakers' sensitivity to agreement in sentence processing. *Applied Psycholinguistics*, 34(3), 615–646. <https://doi.org/10.1017/S0142716411000890>
- Cummings, I. (2017). Parsing and working memory in bilingual sentence processing. *Bilingualism: Language and Cognition*, 20(4), 659–678. <https://doi.org/10.1017/S1366728916000675>
- Cummings, I., & Sturt, P. (2018). Retrieval interference and semantic interpretation. *Journal of Memory and Language*, 102, 16–27. <https://doi.org/10.1016/j.jml.2018.05.001>
- Dąbrowska, E. (2019). Experience, aptitude, and individual differences in linguistic attainment: A comparison of native and nonnative speakers. *Language Learning*, 69(S1), 72–100. <https://doi.org/10.1111/lang.12323>
- Dijkgraaf, A., Hartsuiker, R. J., & Duyck, W. (2019). Prediction and integration of semantics during L2 and L1 listening. *Language, Cognition and Neuroscience*, 34(7), 881–900. <https://doi.org/10.1080/23273798.2019.1591469>
- Dillon, B., Mishler, A., Sloggett, S., & Phillips, C. (2013). Contrasting intrusion profiles for agreement and anaphora: Experimental and modeling evidence. *Journal of Memory and Language*, 69(2), 85–103. <https://doi.org/10.1016/j.jml.2013.04.003>
- Domazetoska, I., & Zhao, H. (2024). Native and non-native speakers' cue integration in the processing of the English as-predicate construction. *Proceedings of the annual meeting of the Cognitive Science Society*, p. 46. <https://escholarship.org/uc/item/36s8692r>
- Dussias, P. E., Kroff, J. R. V., Tamargo, R. E. G., & Gerfen, C. (2013). When gender and looking go hand in hand: Grammatical gender processing in L2 Spanish. *Studies in Second Language Acquisition*, 35(2), 353–387. <https://doi.org/10.1017/S0272263112000915>
- Dussias, P. E., & Sagarra, N. (2007). The effect of exposure on syntactic parsing in Spanish–English bilinguals. *Bilingualism: Language and Cognition*, 10(1), 101–116. <https://doi.org/10.1017/S13667289060002847>
- Eberhard, K. M., Cutting, J. C., & Bock, K. (2005). Making syntax of sense: Number agreement in sentence production. *Psychological Review*, 112(3), 531–559. <https://doi.org/10.1037/0033-295X.112.3.531>
- Engelmann, F., Jäger, L. A., & Vasisht, S. (2019). The effect of prominence and cue association on retrieval processes: A computational account. *Cognitive Science*, 43(12), e12800. <https://doi.org/10.1111/cogs.12800>
- Fine, A. B., Jaeger, T. F., Farmer, T. A., & Qian, T. (2013). Rapid expectation adaptation during syntactic comprehension. *PLoS One*, 8(10), e77661. <https://doi.org/10.1371/journal.pone.0077661>
- Fowler, H. W. (1937). *A dictionary of modern English usage*. Oxford University Press.
- Franck, J., Colonna, S., & Rizzi, L. (2015). Task-dependency and structure-dependency in number interference effects in sentence comprehension. *Frontiers in Psychology*, 6, 349. <https://doi.org/10.3389/fpsyg.2015.00349>
- Franck, J., & Wagers, M. (2020). Hierarchical structure and memory mechanisms in agreement attraction. *PLoS One*, 15(5), e0232163. <https://doi.org/10.1371/journal.pone.0232163>
- García-Castro, V., & Vanek, N. (2024). Syntactic engagement of new words: The garden-path method applied to track sensitivity to structural ambiguity.

- Research Methods in Applied Linguistics*, 3(3), 100–137. <https://doi.org/10.1016/j.rmal.2024.100137>
- Jäger, L. A., Engelmann, F., & Vasishth, S. (2017). Similarity-based interference in sentence comprehension: Literature review and Bayesian meta-analysis. *Journal of Memory and Language*, 94, 316–339. <https://doi.org/10.1016/j.jml.2017.01.004>
- Jäger, L. A., Mertzen, D., Van Dyke, J. A., & Vasishth, S. (2020). Interference patterns in subject-verb agreement and reflexives revisited: A large-sample study. *Journal of Memory and Language*, 111, 104063. <https://doi.org/10.1016/j.jml.2019.104063>
- Jegerski, J. (2016). Number attraction effects in near-native Spanish sentence comprehension. *Studies in Second Language Acquisition*, 38(1), 5–33. <https://doi.org/10.1017/S027226311400059X>
- Jiang, N. (2004). Morphological insensitivity in second language processing. *Applied Psycholinguistics*, 25(4), 603–634. <https://doi.org/10.1017/S0142716404001298>
- Kaan, E., Futch, C., Fuertes, R. F., Mujcinovic, S., & Fuente, E. Á. D. L. (2019). Adaptation to syntactic structures in native and nonnative sentence comprehension. *Applied Psycholinguistics*, 40(1), 3–27. <https://doi.org/10.1017/S0142716418000437>
- Kaan, E., & Grüter, T. (2021). Prediction in second language processing and learning: Advances and directions. In E. Kaan & T. Grüter (Eds.), *Prediction in second language processing and learning* (pp. 1–24). John Benjamins.
- Kim, H., & Grüter, T. (2021). Predictive processing of implicit causality in a second language: A visual-world eye-tracking study. *Studies in Second Language Acquisition*, 43(1), 133–154. <https://doi.org/10.1017/S0272263120000443>
- Kuznetsova, A., Brockhoff, P. B., Christensen, R. H., & others. (2017). lmerTest package: Tests in linear mixed effects models. *Journal of Statistical Software*, 82(13), 1–26. <https://doi.org/10.18637/jss.v082.i13>
- Kwon, N., & Sturt, P. (2016). Attraction effects in honorific agreement in Korean. *Frontiers in Psychology*, 7, 1302. <https://doi.org/10.3389/fpsyg.2016.01302>
- Lacina, R., & Chromý, J. (2022). No agreement attraction facilitation observed in Czech: Not even syncretism helps. *Proceedings of the Annual Meeting of the Cognitive Science Society*, 44, 2423–2430. <https://escholarship.org/uc/item/2vq4g622>
- Lacina, R., Laurinavichyute, A., & Chromý, J. (2025). Only case-syncretic nouns attract: Czech and Slovak gender agreement. *Journal of Memory and Language*, 143, 104623. <https://doi.org/10.1016/j.jml.2025.104623>
- Lago, S., & Felser, C. (2018). Agreement attraction in native and nonnative speakers of German. *Applied Psycholinguistics*, 39(3), 619–647. <https://doi.org/10.1017/S0142716417000601>
- Lago, S., Gračanin-Yukse, M., Šafak, D. F., Demir, O., Kırkıci, B., & Felser, C. (2019). Straight from the horse's mouth: Agreement attraction effects with Turkish possessors. *Linguistic Approaches to Bilingualism*, 9(3), 398–426. <https://doi.org/10.1075/lab.17019.lag>
- Lago, S., Shalom, D. E., Sigman, M., Lau, E. F., & Phillips, C. (2015). Agreement attraction in Spanish comprehension. *Journal of Memory and Language*, 82, 133–149. <https://doi.org/10.1016/j.jml.2015.02.002>
- Laurinavichyute, A., & von der Malsburg, T. (2022). Semantic attraction in sentence comprehension. *Cognitive Science*, 46(2), e13086. <https://doi.org/10.1111/cogs.13086>
- Laurinavichyute, A., & von der Malsburg, T. (2023). Agreement attraction in grammatical sentences and the role of the task. *PsyArXiv*. <https://doi.org/10.31234/osf.io/n75vc>
- Lee, E.-K. R., & Phillips, C. (2023). Why non-native speakers sometimes outperform native speakers in agreement processing. *Bilingualism: Language and Cognition*, 26(1), 152–164. <https://doi.org/10.1017/S1366728922000414>
- Lemhöfer, K., & Broersma, M. (2012). Introducing LexTALE: A quick and valid lexical test for advanced learners of English. *Behavior Research Methods*, 44, 325–343. <https://doi.org/10.3758/s13428-011-0146-0>
- Lewis, R. L., & Vasishth, S. (2005). An activation-based model of sentence processing as skilled memory retrieval. *Cognitive Science*, 29(3), 375–419. https://doi.org/10.1207/s15516709cog0000_25
- Lim, J. H., & Christianson, K. (2015). Second language sensitivity to agreement errors: Evidence from eye movements during comprehension and translation. *Applied Psycholinguistics*, 36(6), 1283–1315. <https://doi.org/10.1017/S0142716414000290>
- MacWhinney, B. (1992). Transfer and competition in second language learning. In R. J. Harris (Ed.), *Cognitive processing in bilinguals* (Vol. 83, pp. 371–390). North-Holland. [https://doi.org/10.1016/S0166-4115\(08\)61506-X](https://doi.org/10.1016/S0166-4115(08)61506-X)
- MacWhinney, B. (2001). The competition model: The input, the context, and the brain. In P. Robinson (Ed.), *Cognition and Second Language Instruction* (pp. 69–90). Cambridge University Press. <https://doi.org/10.1017/CBO9781139524780.005>
- MacWhinney, B. (2005). A unified model of language acquisition. In J. F. Kroll & A. M. B. De Groot (Eds.), *Handbook of bilingualism: Psycholinguistic approaches* (pp. 49–67). Oxford University Press.
- MacWhinney, B. (2008). A unified model. In P. Robinson & N. C. Ellis (Eds.), *Handbook of cognitive linguistics and second language acquisition* (pp. 341–371). Routledge.
- MacWhinney, B. (2018). A unified model of first and second language learning. In M. Hickmann, H. Jisa, & E. Veneziano (Eds.), *Sources of variation in first language acquisition* (pp. 287–312). John Benjamins.
- Marian, V., Blumenfeld, H. K., & Kaushanskaya, M. (2007). The Language Experience and Proficiency Questionnaire (LEAP-Q): Assessing language profiles in bilinguals and multilinguals. *Journal of Speech, Language, and Hearing Research*, 50, 940–967. [https://doi.org/10.1044/1092-4388\(2007\)067](https://doi.org/10.1044/1092-4388(2007)067)
- Matuschek, H., Kliegl, R., Vasishth, S., Baayen, H., & Bates, D. (2017). Balancing Type I error and power in linear mixed models. *Journal of Memory and Language*, 94, 305–315. <https://doi.org/10.1016/j.jml.2017.01.001>
- McDonald, J. L. (2006). Beyond the critical period: Processing-based explanations for poor grammaticality judgment performance by late second language learners. *Journal of Memory and Language*, 55(3), 381–401. <https://doi.org/10.1016/j.jml.2006.06.006>
- Parker, D., & An, A. (2018). Not all phrases are equally attractive: Experimental evidence for selective agreement attraction effects. *Frontiers in Psychology*, 9, 1566. <https://doi.org/10.3389/fpsyg.2018.01566>
- Paspali, A., & Marinis, T. (2020). Gender agreement attraction in Greek comprehension. *Frontiers in Psychology*, 11, 717. <https://doi.org/10.3389/fpsyg.2020.00717>
- Pienemann, M. (1998). Developmental dynamics in L1 and L2 acquisition: Processability theory and generative entrenchment. *Bilingualism: Language and Cognition*, 1(1), 1–20. <https://doi.org/10.1017/S1366728998000017>
- Prasad, G., & Linzen, T. (2021). Rapid syntactic adaptation in self-paced reading: Detectable, but only with many participants. *Journal of Experimental Psychology: Learning, Memory, and Cognition*, 47(7), 1156–1172. <https://doi.org/10.1037/xlm0001046>
- Sagarra, N., & Herschensohn, J. (2013). Processing of gender and number agreement in late Spanish bilinguals. *International Journal of Bilingualism*, 17(5), 607–627. <https://doi.org/10.1177/1367006912453810>
- Slioussar, N. (2018). Forms and features: The role of syncretism in number agreement attraction. *Journal of Memory and Language*, 101, 51–63. <https://doi.org/10.1016/j.jml.2018.03.006>
- Slioussar, N., Magomedova, V., & Makarova, P. (2022). The role of case syncretism in agreement attraction: A comprehension study. *Frontiers in Psychology*, 13, 829112. <https://doi.org/10.3389/fpsyg.2022.829112>
- Sturt, P., & Kwon, N. (2023). Agreement attraction in comprehension: Do active dependencies and distractor position play a role? *Language, Cognition and Neuroscience*, 39(3), 279–301. <https://doi.org/10.1080/23273798.2023.2269282>
- Tanner, D., Nicol, J., & Brehm, L. (2014). The time-course of feature interference in agreement comprehension: Multiple mechanisms and asymmetrical attraction. *Journal of Memory and Language*, 76, 195–215. <https://doi.org/10.1016/j.jml.2014.07.003>
- Tanner, D., Nicol, J., Herschensohn, J., Osterhout, L., Biller, A., Chung, E., & Kimball, A. (2012). Electrophysiological markers of interference and structural facilitation in native and nonnative agreement processing. *Proceedings of the 36th Boston University Conference on Language Development*, pp. 594–606.
- Tucker, M. A., Idrissi, A., & Almeida, D. (2015). Representing number in the real-time processing of agreement: Self-paced reading evidence from Arabic. *Frontiers in Psychology*, 6, 347. <https://doi.org/10.3389/fpsyg.2015.00347>
- Tucker, M. A., Idrissi, A., & Almeida, D. (2021). Attraction effects for verbal gender and number are similar but not identical: Self-paced reading evidence

- from Modern Standard Arabic. *Frontiers in Psychology*, 3774. <https://doi.org/10.3389/fpsyg.2020.586464>
- Turk, U., & Logačev, P.** (2021). Agreement attraction in Turkish: The case of genitive attractors. *PsyArXiv*. <https://doi.org/10.31234/osf.io/5rmvu>
- Vanek, N., Matic Škorić, A., Košutar, S., Matějka, Š., & Stone, K.** (2024). Looks at what isn't there: Eye movements on a blank screen when processing negation in a first and a second language. *Frontiers in Human Neuroscience*, 18. <https://doi.org/10.3389/fnhum.2024.1457038>
- VanPatten, B.** (1996). *Input processing and grammar instruction in second language acquisition*. Greenwood Publishing Group.
- Vasishth, S., Brüßow, S., Lewis, R. L., & Drenhaus, H.** (2008). Processing polarity: How the ungrammatical intrudes on the grammatical. *Cognitive Science*, 32(4), 685–712. <https://doi.org/10.1080/03640210802066865>
- Vasishth, S., Nicenboim, B., Engelmann, F., & Burchert, F.** (2019). Computational models of retrieval processes in sentence processing. *Trends in Cognitive Sciences*, 23(11), 968–982. <https://doi.org/10.1016/j.tics.2019.09.003>
- Vigliocco, G., Butterworth, B., & Semenza, C.** (1995). Constructing subject-verb agreement in speech: The role of semantic and morphological factors. *Journal of Memory and Language*, 34(2), 186–215. <https://doi.org/10.1006/jmla.1995.1009>
- Villata, S., & Franck, J.** (2020). Similarity-based interference in agreement comprehension and production: Evidence from object agreement. *Journal of Experimental Psychology: Learning, Memory, and Cognition*, 46(1), 170. <https://doi.org/10.1037/xlm0000718>
- von der Malsburg, T., & Angele, B.** (2017). False positives and other statistical errors in standard analyses of eye movements in reading. *Journal of Memory and Language*, 94, 119–133. <https://doi.org/10.1016/j.jml.2016.10.003>
- Wagers, M., Lau, E., & Phillips, C.** (2009). Agreement attraction in comprehension: Representations and processes. *Journal of Memory and Language*, 61(2), 206–237. <https://doi.org/10.1016/j.jml.2009.04.002>
- Yadav, H., Smith, G., Reich, S., & Vasishth, S.** (2023). Number feature distortion modulates cue-based retrieval in reading. *Journal of Memory and Language*, 129, 104400. <https://doi.org/10.1016/j.jml.2022.104400>
- Zehr, J., & Schwarz, F.** (2018). *PennController for internet based experiments (IBEX)*. <https://doi.org/10.17605/OSF.IO/MD832>
- Zhao, H., Vanek, N., & MacWhinney, B.** (2025). *Mental simulation in bilingual and second language processing: New directions in the competition model*. Social Science Research Network. <https://doi.org/10.2139/ssrn.5125974>