Studies in Second Language Acquisition (2024), 1–19
doi:10.1017/S027226312400024X

REPLICATION STUDY

Development of verb argument constructions in L2 English learners: *A close replication of research question 3 in Römer and Berger (2019)*

Yingying Liu[1] and Xiaofei Lu[2]

[1]Department of English, College of Foreign Languages, Ocean University of China, Qingdao, Shandong, China; [2]Department of Applied Linguistics, The Pennsylvania State University, University Park, PA, USA
Corresponding author: Xiaofei Lu; Email: xxl13@psu.edu

(Received 30 June 2023; Revised 24 February 2024; Accepted 05 March 2024)

Abstract

This study closely replicates the analyses of the third research question in Römer and Berger (2019), which reported that the associations between verbs and verb argument constructions (VACs) used by German and Spanish learners of English move closer to a native usage norm as the learners' proficiency increases. This study conducted the same correlation analyses from the original study but with a substantially expanded version of the learner corpus used therein. Additionally, we conducted zero-inflated negative binomial analyses to estimate the relationship between the frequencies of verb-VAC combinations in the British National Corpus (BNC) and in the learner subcorpora representing different proficiency levels. Our findings were consistent with the original study in showing significant positive correlations between frequencies of the verb-VAC combinations in the BNC and in the learner subcorpora but further revealed the potential effect of topic on the learners' VAC usage. Implications for future studies are discussed.

Introduction

With the growing availability of large-scale learner corpora, the synergy between corpus linguistics and second language acquisition (SLA) has seen tremendous growth in recent years. Learner corpora have been increasingly used to investigate diverse aspects of second language (L2) development (Hunston, 2022; Lu, 2023; McEnery et al., 2019), including not only lexical and syntactic development but also the development of form-meaning pairings, known as constructions (Goldberg, 1995, 2006). This latter line of scholarship has paid much attention to the developmental trajectory of verb argument constructions (VACs), such as transitive constructions (e.g., "V n"), ditransitive constructions (e.g., "V n"), locative constructions (e.g., "V n *onto* n"), and constructions of the "VERB PREPOSITION NOUN" pattern (e.g., "V *about* n"), to name but a few (see Römer, 2019, for a comprehensive review). Usage-based construction grammar sees language as a structured, hierarchical inventory of constructions interconnected with one another through various semantic and syntactic links; consequently, the learning of L2 constructions is a central component of SLA. A sizable body of L2 construction development research has provided evidence that L2 development involves a dynamic and multifaceted process of accumulating an expanding repertoire of constructions (Ellis & Ferreira-Junior, 2009; Eskildsen, 2009; Hwang & Kim, 2023; Liu & McManus, 2020; Römer, 2019; Römer & Berger, 2019).

A recent large-scale corpus-based study in this line is Römer and Berger (2019; henceforth R&B). Using the first version of EF-Cambridge Open Language Database (EFCAMDAT1; Alexopoulou et al., 2015), R&B examined the emergence and development of 19 VACs in Spanish and German learners of English. Among other analyses of the use of VACs by L2 learners across different L2 proficiency levels and first language (L1) backgrounds (research questions 1 [RQ1], 2 [RQ2], and 4 [RQ4]), they correlated the frequencies of verb-VAC combinations in learner subcorpora representing different proficiency levels and in the British National Corpus (BNC) representing L1 English usage (research question 3 [RQ3]). The results showed that the verb-VAC associations for individual constructions, particularly "V *about* n", became more varied and closer to the L1 usage norm at higher proficiency L2 levels. Their findings lend important empirical support to usage-based approaches to SLA by confirming the emergence of more productive verb constructions and L2 learners' move toward an L1 usage norm in terms of verb-VAC associations. In light of the contribution of learner corpora to SLA research and the importance of understanding L2 construction development, it is necessary to verify the generalizability of their results with more L2 learners and triangulate their findings with additional statistical analyses. The second release of EF-Cambridge Open Language Database (EFCAMDAT2; Alexopoulou et al., 2017) provides such an opportunity for replicating R&B with a highly comparable but substantially larger learner corpus. With modifications in the use of an expanded dataset and the statistical method, this study is a close replication (Porte & McManus, 2019) of R&B.

Motivation for replication

The proposed replication is broadly motivated by the growing interest in the use of learner corpora to understand L2 development (McEnery et al., 2019; Römer & Garner, 2022). Lu (2023: 3) succinctly summarizes four major issues in corpus-based SLA research: (a) variation in L2 use, (b) factors influencing L2 processing and production, (c) group-level developmental trajectories, and (d) variability and variation in L2 development. The original study of R&B and this replication study contribute to the understanding of the third major issue in the field. A number of studies concerning this question have revealed a move toward the L1 usage norm (Römer, 2019; Römer & Berger, 2019), more productive usage of individual constructions (Liu & McManus, 2020; Römer, 2019; Römer & Berger, 2019; Xu, 2016), fewer erroneous cases (Römer & Berger, 2019; Xu, 2016), and more lower-frequency constructions (Ellis & Ferreira-Junior, 2009; Kyle, et al., 2021) with increasing proficiency. For example, focusing on transitive and modal VACs, Römer (2019) found the correlations for verb-VAC combinations between German learners of English and L1 English users became stronger from A1 (via A2 and B1) to B2 level. Römer and Garner's (2019) analysis of L2 spoken data showed a similar developmental pattern, that is, more proficient L2 speakers were more productive in verb-VAC combinations and closer to L1 English usage than less proficient speakers. Through a contrastive corpus-based analysis of *give* ditransitive constructions, Xu (2016) showed Chinese English as a foreign language (EFL) learners' movement from less to more semantically varied *give*-ditransitives as well as a decline in unidiomatic usage. As L2 learners' repertoire of constructions expands, the token frequency of already familiar constructions may decrease with L2 proficiency, reflecting a departure from the repetitive use of such familiar constructions (Biber et al., 2016). Hwang and Kim (2023) reported that while the frequency proportion of caused-motion, ditransitive, and passive constructions in 11 constructions increased with proficiency in L2 writing, that of attributive constructions (i.e., copular *be* + adjective) decreased with proficiency.

In addition to advancing our understanding of group-level L2 development trajectories, this replication study is specifically motivated by the need to grapple with data sparsity in corpus-based studies (McEnery et al., 2019). The small scale of most existing learner corpora has frequently led to data sparsity issues; even with recent, larger learner corpora, the problem persists when researchers consider multiple factors in selecting target subcorpora for analysis. The issue of data sparsity also manifested in R&B's correlation analyses of EFCAMDAT1 and the BNC, with a very large group of verb-VAC combinations in the BNC absent in their learner data. Such a data distribution can compromise the accuracy of Pearson correlation analysis, which assumes homoscedasticity. While R&B addressed this issue by log-transforming those counts, there exist several limitations for log-transformation, including (a) heteroscedasticity is oftentimes not addressed even after log-transformation, (b) the primary quantity of interest (i.e., the counts) is not modelled directly, and (c) log-transformation runs into issues for count data with many zeros (O'Hara & Kotze, 2010; Winter & Bürkner, 2021), necessitating additional data processing procedures to avoid logging zeros. Commonly employed techniques include adding a small constant to all values before log-transforming them (e.g., Liu, 2021) or replacing all zeros directly with a small constant without log-transformation (e.g., Ellis et al., 2016; Römer & Berger, 2019). However, the variation in data processing techniques and the choice of the specific constant value (e.g., –0.1, 0.1,0.01) can affect statistical inferences (e.g., Ekwaru & Veugelers, 2018). It will therefore be desirable to build a more robust statistical model that can well handle count data with many zeros to triangulate the results of correlation analyses in R&B. Considering the ubiquity of zero-abundant count data in corpus-based SLA research, this replication may hold methodological implications for future corpus-based research using count data with many zeros.

The replication

This study replicates the analyses performed for RQ3 in R&B, which constituted the bulk of the analysis in the original study. We only replicated RQ3 because this was the only question that pertained to the data sparsity problem, while the remaining three research questions did not face the same challenge, as they were based solely on the learner corpus dataset without comparison to a larger L1 reference corpus. We followed the same procedures as reported in the initial study unless otherwise noted. The main changes between the initial study and this replication study are summarized in Table 1. The first main part of our replication examines how R&B's findings regarding the L2 development trajectory with EFCAMDAT1 extend to EFCAMDAT2. Specifically, the type-token frequency distribution of a set of VACs following the pattern of "VERB PREPOSITION NOUN" and their associated verbs were derived from relevant EFCAMDAT2 subcorpora and the BNC. The results were then compared against

Table 1. Comparison between the initial study and the replication study

	R&B	This replication
L1 background	German, Spanish	German, Spanish
Proficiency level	CEFR A1–C1	CEFR A1–C1
Corpus	EFCAMDAT1	EFCAMDAT2
Retrieved VACs	19 VACs	19 VACs
Frequency profile	19 VACs	9 VACs
Focal VAC	V *about* n	V *about* n
Analyses	Correlation analysis	Correlation analysis, ZINB regression analysis

R&B's results to determine whether their results hold with the larger corpus. This part of analysis is useful both for verifying the generalizability of R&B's findings and for understanding the effect of learner corpus scale on corpus-based SLA research in general.

Second, zero-inflated negative binomial (ZINB) regression analyses were conducted to examine the role of L2 input and L2 proficiency in learners' use of verb-VAC combinations. The ZINB regression model was chosen because of its suitability for modeling count data with overdispersion and excess zeros (O'Hara & Kotze, 2010; Winter et al., 2018). The datasets we used for comparing L1 speakers' and L2 learners' usage contained a large number of zeros, as many verb-VAC combinations in the BNC were not used by L2 learners. The results of ZINB models can help reveal whether using a more robust model for handling count data with overdispersion and excess zeros modifies the results.

Research questions

This replication study explores the same research question as RQ3 in the original study of R&B (p. 1092): "Do learners' verb-VAC associations move closer to a native usage norm as language proficiency increases?" The investigation of the present RQ1 follows the same approach adopted in the original study, while that of the present RQ2 takes a different approach to further validate the results of RQ1.

RQ1: Based on correlation analyses, do learners' verb-VAC associations move closer to an L1 usage norm as their language proficiency increases?

RQ2: Based on ZINB models, is learners' usage of verb-VAC associations affected by the L1 usage norm and L2 proficiency? If so, in what ways?

Method

Data

Following R&B, we used the BNC XML version[1] as the reference corpus to approximate the language input received by learners. This version of the BNC consists of 4,049 texts of 96,986,707 words. For the purpose of extracting the target verb constructions, we prepared a syntactically parsed version of the BNC using the Stanford Parser via the Stanford CoreNLP pipeline (Manning et al., 2014).

The learner corpus for the present study was EFCAMDAT2 (Alexopoulou et al., 2017), which is available upon request[2]. EFCAMDAT2 consists of essays submitted to

[1]http://www.natcorp.ox.ac.uk/

[2]https://ef-lab.mmll.cam.ac.uk/EFCAMDAT.html

https://doi.org/10.1017/S027226312400024X Published online by Cambridge University Press

Table 2. Overview of the EFCAMDAT subcorpora used in R&B and in this replication study

Writer group	Number of learners		Number of texts		Number of words	
	R&B	This study	R&B	This study	R&B	This study
Mexican A1	4,043	8,073	24,275	51,571	1,533,012	2,305,788
Mexican A2	1,596	3,090	10,572	21,050	1,012,049	1,616,176
Mexican B1	808	1,743	3,903	10,400	471,543	1,120,267
Mexican B2	273	633	1,158	3,476	178,907	524,724
Mexican C1	34	102	186	667	37,225	124,157
Mexican all levels	*6,754*	*11,537*	*40,094*	*87,164*	*3,232,736*	*5,691,112*
German A1	2,072	3,655	10,721	19,520	728,275	1,006,399
German A2	1,580	2,718	8,507	14,850	811,842	1,167,962
German B1	1,240	2,333	5,222	10,982	631,338	1,210,386
German B2	926	1,736	3,092	6,776	488,431	1,027,181
German C1	202	430	930	2,163	186,176	413,722
German all levels	*6,020*	*8,992*	*28,472*	*54,291*	*2,846,062*	*4,825,650*
Total	*12,774*	*20,529*	*68,566*	*141,455*	*6,078,798*	*10,516,762*

Note: A1 = levels 1–3 in EFCAMDAT; A2 = levels 4–6; B1 = levels 7–9; B2 = levels 10–12; C1 = levels 13–15. The dataset for each L1 group at each CEFR level includes texts produced for 24 prompts (3 EFCAMDAT levels × 8 prompts).

the online school of EF Education First. All essays were written by L2 English learners from different L1 backgrounds and proficiency levels. By a placement test or successful progression through coursework, writers of the essays were allocated to 16 proficiency levels aligned with the six levels of the Common European Framework of Reference for Languages (CEFR). Each of the 16 proficiency levels has eight different writing tasks. Table 2 provides an overview of the subcorpora of EFCAMDAT1 used in R&B and those of EFCAMDAT2 used in this replication study. As can be seen, EFCAMDAT2 is much larger than EFCAMDAT1 in terms of the number of learners, texts, and words. We wrote a Python script to retrieve the textual content from the XML files of EFCAMDAT2 and syntactically parsed the retrieved texts using the Stanford Parser (Klein & Manning, 2003), thus obtaining the same type of syntactic parse trees as those from the BNC.

Identification of VACs

To retrieve instances of VACs from the corpora, we used Stanford Tregex (Levy & Andrew, 2006) to traverse the syntactic parse trees and identify instances of VACs as defined by the Tregex search patterns. The Tregex patterns used in the present study targeted verb patterns consisting of a verb that immediately dominates a specific preposition. Through an iterative trial-and-error process, the original search patterns were further modified to avoid expected noises. For example, the following pattern was used for the retrieval of "V *as* n" constructions, in which "!<# VBN" excluded passive constructions (e.g., "the old logo was regarded as an old fashion"), "!<< (NP << (JJ|RB < much|little|few|many))" avoided cases containing *"as much/little/few/many…as"* (e.g., "to eat as much peanut butter as they would"), and "&!< (NP << (NN < result)))" removed cases of "*as a result*" (e.g., "the impressionist movement began as a result to [of] a very formal and rigid style of paintings").

$$(\mathrm{VP!}<\#\mathrm{VBN})<\begin{matrix}2(\mathrm{PP}<1(\mathrm{IN}<\mathrm{as})!<<(\mathrm{NP}<<(\mathrm{JJ}|\mathrm{RB}<\mathrm{much}|\mathrm{little}|\mathrm{few}|\mathrm{many}))\&!\\<(\mathrm{NP}<<(\mathrm{NN}<\mathrm{result})))\end{matrix}$$

https://doi.org/10.1017/S027226312400024X Published online by Cambridge University Press

Table 3. Identification accuracy for VACs in each dataset

	BNC	L1 German	L1 Spanish
V *about* n	**100%**	**100%**	**96%**
V *across* n	**96%**	**100%**	**94%**
V *after* n	58%	48%	34%
V *against* n	**92%**	**100%**	**100%**
V *among* n	90%	75%	80%
V *around* n	64%	90%	94%
V *as* n	70%	92%	86%
V *between* n	**82%**	**100%**	**98%**
V *for* n	76%	34%	76%
V *in* n	64%	66%	82%
V *into* n	**98%**	**80%**	**88%**
V *like* n	96%	66%	62%
V *of* n	74%	68%	72%
V *off* n	**94%**	**100%**	**100%**
V *over* n	78%	80%	58%
V *through* n	**92%**	**100%**	**100%**
V *toward* n	**92%**	**96%**	**95%**
V *under* n	58%	88%	70%
V *with* n	**92%**	**86%**	**84%**

Note: Bolded rows represent VACs with an accuracy rate of 80% or more in all three datasets.

All Tregex patterns used in this study are presented in Appendix A in the supplementary material. For 18 of the 19 VACs, a single Tregex pattern was used to identify their instances, while the "V *off* n" construction involved two patterns, as *off* received two different tags, RP (participle) and IN (preposition), both of which were included in this study.

To check the accuracy of the automatic parsing and retrieval procedure in this study, we manually examined 100 randomly selected sentences for each VAC in the BNC and 50 for each in the L1 German and L1 Spanish subcorpora, respectively (i.e., 200 sentences per VAC total). Table 3 shows the accuracy for each VAC in each of the three datasets. For VACs occurring fewer than 100 or 50 times in the dataset, we reported the overall accuracy of all retrieved cases. We found comparatively low accuracy rates for certain VACs, but the accuracy rate reached 80% or above in all three datasets for the following nine VACs: "V *about* n," "V *across* n," "V *against* n," "V *between* n," "V *into* n," "V *off* n," "V *through* n," "V *toward* n," and "V *with* n". We retrieved the head verbs and created the frequency-sorted verb lemma list for each VAC in each L1 group and L2 proficiency level. TagAnt (Anthony, 2015) was used for the lemmatization of the verb list. These frequency lists of verb lemmas served as the datasets for the statistical analyses in this study.

While we used the same procedure to automatically extract VACs from syntactically parsed versions of the BNC and the learner subcorpora, R&B used different procedures to extract VACs from them. For the BNC, they used a different parser, namely, the Robust Accurate Statistical Parsing System (Andersen et al., 2008), to parse the BNC files, and then relied on the same part-of-speech (POS) categories (i.e., verbs, prepositions, and nouns) and syntactic relationships (i.e., between a target preposition and a noun/verb) to extract VACs. Our BNC VAC files can be expected to be highly comparable to those obtained in R&B, with the caveat that some minor variation may have resulted from the disparities in the tagging and parsing accuracies between the two parsers. For the learner subcorpora, R&B first POS-tagged the files using

https://doi.org/10.1017/S027226312400024X Published online by Cambridge University Press

Table 4. Overview of VACs across L1 German learner datasets

	German A1				German A2				German B1				German B2				German C1			
	Tps	Tks	CTTR	Tks 100K	Tps	Tks	CTTR	Tks 100K	Tps	Tks	CTTR	Tks 100K	Tps	Tks	CTTR	Tks 100K	Tps	Tks	CTTR	Tks 100K
V *about* n	22	302	.90	43.5	48	437	1.62	51.7	70	1029	1.54	112.2	41	1238	.82	159.4	30	266	1.30	85.2
V *across* n	2	20	.32	2.9	5	27	.68	3.2	8	10	1.79	1.1	6	10	1.34	1.3	4	4	1.41	1.3
V *against* n	4	13	.78	1.9	7	80	.55	9.5	17	67	1.47	7.3	24	81	1.89	10.4	10	18	1.67	5.8
V *between* n	9	97	.65	14.0	12	52	1.18	6.1	15	45	1.58	4.9	22	34	2.67	4.4	11	19	1.78	6.1
V *into* n	18	67	1.55	9.7	33	271	1.42	32.0	37	146	2.17	15.9	46	155	2.61	20.0	36	118	2.34	37.8
V *off* n	2	2	1.00	.3	11	23	1.62	2.7	11	49	1.11	5.3	22	131	1.36	16.9	8	11	1.71	3.5
V *through* n	9	25	1.27	3.6	23	88	1.73	10.4	21	63	1.87	6.9	32	66	2.79	8.5	27	58	2.51	18.6
V *toward* n	—	—	—	—	1	1	.71	.1	2	2	1.00	.2	11	25	1.56	3.2	9	24	1.30	7.7
V *with* n	112	754	2.88	108.7	158	1592	2.80	188.2	216	1796	3.60	195.8	200	1575	3.56	202.8	125	489	4.00	156.6

Note: Tps = types; Tks = tokens; CTTR = corrected type-token ratio; Tks 100K = tokens per 100K words.

https://doi.org/10.1017/S027226312400024X Published online by Cambridge University Press

Table 5. Overview of VACs across L1 Spanish learner datasets

	Spanish A1				Spanish A2				Spanish B1				Spanish B2				Spanish C1			
	Tps	Tks	CTTR	Tks 100K	Tps	Tks	CTTR	Tks 100K	Tps	Tks	CTTR	Tks 100K	Tps	Tks	CTTR	Tks 100K	Tps	Tks	CTTR	Tks 100K
V *about* n	31	211	1.51	13.0	72	695	1.93	59.3	104	1220	2.11	141.8	61	669	1.67	162.2	28	119	1.81	122.7
V *across* n	8	79	.64	4.9	5	54	.48	4.6	—	—	—	—	2	2	1.00	.5	1	1	.71	1.0
V *against* n	1	2	.50	.1	11	61	1.00	5.2	9	103	.63	12.0	11	22	1.66	5.3	4	12	.82	12.4
V *between* n	7	305	.28	18.9	12	31	1.52	2.6	12	19	1.95	2.2	7	10	1.57	2.4	6	8	1.50	8.2
V *into* n	16	38	1.84	2.3	38	299	1.55	25.5	35	93	2.57	10.8	29	75	2.37	18.2	19	35	2.27	36.1
V *off* n	5	12	1.02	.7	9	39	1.02	3.3	10	31	1.27	3.6	11	55	1.05	13.3	2	5	.63	5.2
V *through* n	4	8	1.00	.5	10	55	.95	4.7	19	23	2.80	2.7	16	27	2.18	6.5	9	10	2.01	10.3
V *toward* n	—	—	—	—	6	6	1.73	.6	2	2	1.00	.2	2	4	.71	1.0	3	8	.75	8.2
V *with* n	168	1258	3.35	77.8	219	2290	3.24	195.4	249	1979	3.96	230.0	243	1972	3.87	478.1	66	176	3.52	181.5

Note: Tps = types; Tks = tokens; CTTR = corrected type-token ratio; Tks 100K = tokens per 100K words.

https://doi.org/10.1017/S027226312400024X Published online by Cambridge University Press

TagAnt, then exhaustively retrieved all VAC candidates (i.e., sequences consisting of a verb followed by a target preposition) via concordance searches on the POS-tagged files, and finally manually filtered the candidates to identify true hits of each VAC. In this case, our automated procedure proved more efficient, but it came with limitations in terms of precision and recall that could affect the full comparability of our results and those of R&B.

Statistical analyses

In terms of statistical analyses, we first correlated the verb-VAC combination frequency in learner writing and the BNC to see whether R&B's observation about learners' move toward the L1 usage norm still holds in the expanded new version of the corpus. Following R&B, we used the log10 transformed frequencies of verb-VAC combinations in this analysis, with all non-zero values increased by . 01 before log-transformation and all zero values directly replaced by . 01 without log-transformation[3].

We then conducted additional ZINB regression analyses, in which L2 proficiency level and the verb-VAC combination frequency in the BNC were entered as explanatory variables, and the verb-VAC combination frequency in each learner subset served as the dependent variable. Given the varying sizes of the BNC and the learner subcorpora, the frequencies in each learner subcorpus and the BNC were normalized to 1,000,000 words. As the ZINB model only deals with integers (number of counts), the normalized frequencies were rounded to integers.

The *pscl* package in R was used for fitting ZINB models and the *sjPlot* package was used to obtain the incidence rate ratios (IRRs) for the ZINB models. The *ggplot2* package was used for plotting, and *cor.test*, *lm* and *r.squared* functions were used to compute the Pearson r values, p values and R^2 values of correlation analyses. In addition to using normalized frequencies to account for the size differences among the corpora, we also fit ZINB models using raw frequencies with the log10 of the corpus size (i.e., number of word tokens in the corpus) as the offset variable. As detailed in Appendix B of the supplementary material, these models yielded comparable results as those obtained from the models with normalized frequencies.

Overall distribution of VACs

Following the original study, we computed the type and token frequencies of each VAC across proficiency levels. To better reveal the changes in construction diversity across proficiency levels, we additionally calculated the corrected type-token ratio for each VAC construction (VAC CTTR; Park, 2017), a transformed measure of type-token ratio ($VAC\ CTTR = \frac{types}{\sqrt{2 * tokens}}$) that approximately addresses the impact of text length. Tables 4 and 5 present an overview of the frequencies of VACs in L1 German learners and L1 Spanish learners, respectively. A striking similarity in terms of construction

[3]As R&B (2019: 1095) reported, "To avoid missing responses as a result of logging zero, all values were incremented by . 01". The absence of negative values in Figures 5–10 in R&B suggested that the zero values were replaced by . 01 directly without further log-transformation. Our personal communication with the authors confirmed that this was the case. This procedure differed from the one in which zero values were increased by . 01 and then log-transformed (i.e., log10(0 + .01) = –2.0), as was the case in Liu (2021). Similar to R&B, Ellis et al. (2016: 130) replaced zero values directly with – .1. The practices in these studies further illustrate the variation in data processing techniques related to log-transformation.

frequency profile can be found between this replication and the original study. R&B reported particularly low type and token frequencies for "V *across* n," "V *off* n" and "V *toward* n", while much higher frequencies for "V *about* n" and "V *with* n". Our analysis confirmed the same frequency profile of the above-mentioned VACs in the EFCAMDAT2 subsets, although more tokens of VACs were found in this expanded version of corpus. "V *about* n" and "V *with* n" were used most frequently by learners and "V *with* n" was the most productive one with the highest CTTRs among the nine VACs. Except for the three VACs with a small number of tokens (i.e., "V *across* n," "V *off* n," and "V *toward* n"), the other six VACs (i.e., "V *about* n," "V *against* n," "V *between* n," "V *into* n," "V *through* n," and "V *with* n") showed an increase with fluctuations in type frequency and CTTR from A1 to B2 proficiency levels, indicating learners' more varied and productive usage and expanding knowledge repertoire of these VACs. A particularly interesting case is "V *between* n". Despite a decline in its token frequency with proficiency, its type frequency and CTTR increased steadily with L2 proficiency in L1 German learners from A1 to B2 and in Spanish learners from A1 to B1. The observed decrease in the token frequency of "V *between* n" confirms previous findings that construction token frequency is not always positively correlated with L2 proficiency (Biber et al., 2016; Hwang & Kim, 2023). On the other hand, the steady increase in type frequency and CTTR suggests that learners at higher proficiency possess a larger repertoire of verbs for this VAC despite the lower frequency in use.

Results and discussion for RQ1

To address RQ1 (whether learners' verb-VAC associations move closer to an L1 usage norm as their language proficiency increases), we will focus on the "V *about* n" VAC, the results of which were discussed in detail in R&B. The automatic identification procedure for "V *about* n" reached a high accuracy rate (see Table 3), and we further manually checked the list of lemmas against the retrieved sentences to ensure the reliability of the analysis. Table 6 compares the Pearson correlation *r* values reported by R&B and by the present study. Both studies showed a stronger correlation between learners' and L1 English verb-VAC pairings at A2 than A1. While the correlation values remained at the same level from A2 to B2 in R&B's dataset, this study demonstrates an increase from A2 to B2 in both L1 groups of learners. The lower correlations at the C1 level in R&B (particularly for L1 Spanish learners) and in this study can be due to the smaller size of the C1 subcorpus in both studies (as previously shown in Table 2). Overall, the correlations in Table 6, as argued by R&B, are non-trivial and provide empirical support for the important role of L2 input frequency in VAC acquisition.

To further examine the variations of verb-VAC combinations across different learner groups and proficiency levels, we created the same type of correlation plots as

Table 6. Correlations for verbs in "V *about* n" between L2 learners' and L1 English usage in R&B and this replication study

	R&B		This Study	
Comparison	L1 German	L1 Spanish	L1 German	L1 Spanish
A1 vs. BNC	.46	.49	.47	.55
A2 vs. BNC	.55	.53	.56	.58
B1 vs. BNC	.55	.51	.61	.58
B2 vs. BNC	.54	.51	.66	.61
C1 vs. BNC	.56	.32	.61	.55

https://doi.org/10.1017/S027226312400024X Published online by Cambridge University Press

in R&B. Figure 1(a), for example, shows the correlations between the logarithmic frequencies of verb types in the "V *about* n" VAC used by A1 level L1 German learners (y-axis) and L1 English users (x-axis). The full statistics for all correlation analyses are reported in Tables 7 and 8. The black bar at y = .01, as all values were incremented by . 01, indicates a large number of verb types in the BNC that were not used by the learners. R&B reported two major findings about learners' use of verbs in "V *about* n" across proficiency levels. First, learners produced a wider variety of verbs in "V *about* n" as their proficiency increased. Specifically, verbs became more populated in the plots for higher proficiency levels, and A2-level learners started to use legitimate verbs that were absent in A1-level writings (e.g., *hear*, *like*, *report*, *laugh*, and *complain*). Second, learners at B1 and B2 levels produced verbs that seldom co-occur with the "V *about* n" in L1 English usage. These verbs, nonetheless, are semantically related to the core verbs of the VAC, including the *talk* group, the *think* group, and the *learn* group as categorized by Francis et al. (1996),[4] each named after the most frequent member in the respective meaning group. Such unidiomatic examples in learner writing, as R&B found, included *consider* in the *think* group and *discuss* and *request* in the *talk* group. The use of unidiomatic verb-VAC combinations, however, notably decreased at the C1 level.

The first finding of R&B still holds in our dataset. As shown in Figures 1 and 2, the plots for higher proficiency levels are populated with a larger number of verb types. In line with R&B, we also noted that verbs such as *laugh*, *report*, and *complain* were used by learners at the A2 level and above but not by those at the A1 level. That is, learners seem to accumulate more knowledge of verb types legitimate for the target VAC with more L2 experience. In line with the second finding of R&B, the current analysis revealed fewer learner-specific, unidiomatic instances at the C1 level. Both studies found that (a) "*discuss about* n" was used by L1 German learners at the A2, B1, and B2 levels (e.g., "we will discuss about the topic in several teams") and by L1 Spanish learners at the B2 level (e.g., "The purpose of this meeting is to discuss about company problems") but were absent from the C1 learner dataset; and (b) "*consider about* n"

(e.g., "we start to consider about a job change") and "*request about* n" (e.g., "we request about a loan because we want to buy a house") were used by B2-level L1 German learners but not by other groups of learners.

Two observations not discussed in the original study are worth noting. The first one is that L1 German and Spanish learners seemed to be able to use all three meaning groups of verbs (i.e., the *talk* group, the *think* group, and the *learn* group) for "V *about* n" as summarized by Francis et al. (1996). Realizations such as "*talk about* n" and "*think about* n" remained as the top two in token frequency in the BNC and in 9 of the 10 learner datasets, with the exception of Spanish A2 where *be* ranked the second. While the use of "*learn about* n" was comparatively less frequent than the top two, both L1 German and Spanish learners started to use this expression at the beginner level (A1). L1 Spanish learners, in particular, used "*learn about* n" fairly frequently from A1 to B2 levels (see Figure 2). Second, plots for the B2 level in our study and in R&B both exhibited a high frequency of "*enquire about* n" at the B2 level, which ranked the fourth in token frequency after *think*, *talk*, and *be* in L1 German learners (Figure 1(d)) and the third after *think* and *talk* in L1 Spanish learners (Figure 2(d)). The frequent use of *enquire* may be attributed to the influence of writing topic. Among the writing tasks at the B2 level, there was a prompt asking learners to write a formal request letter for a

[4]https://grammar.collinsdictionary.com/us/grammar-pattern/v-about-n_4

https://doi.org/10.1017/S027226312400024X Published online by Cambridge University Press

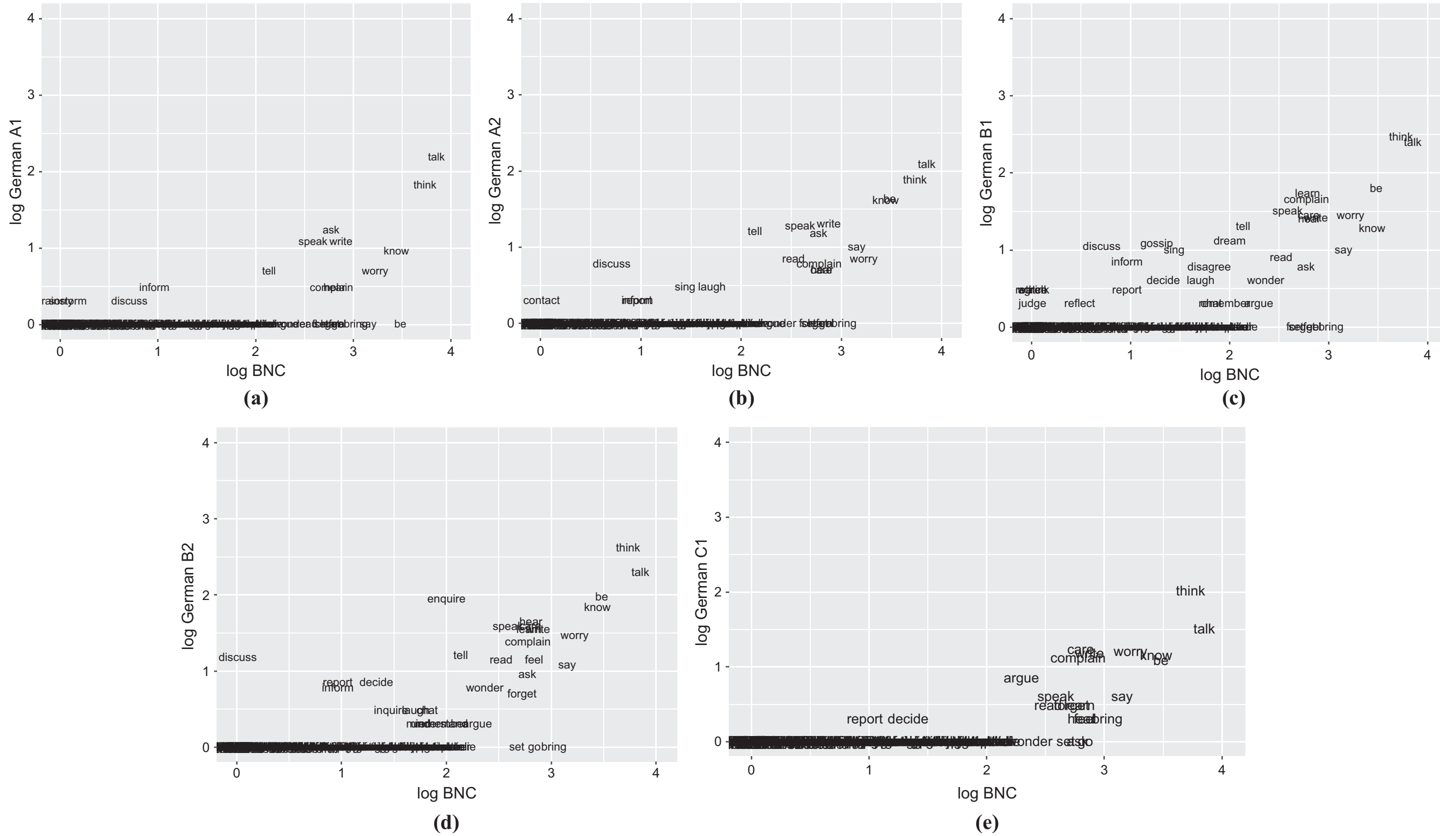

Figure 1. Correlations of verbs in L1 German learners' writing and the BNC for "V *about* n".

Table 7. Statistics for the correlation analyses between the BNC and L1 German learners

	Pearson's *r*	*df*	95% CI	*p*	R^2
German A1	.467	311	[.375, .550]	< .001	21.8%
German A2	.564	331	[.486, .632]	< .001	31.8%
German B1	.608	337	[.536, .671]	< .001	36.9%
German B2	.657	311	[.589, .716]	< .001	43.2%
German C1	.607	312	[.532, .672]	< .001	36.8%

Table 8. Statistics for the correlation analyses between the BNC and L1 Spanish learners

	Pearson's *r*	*df*	95% CI	*p*	R^2
Spanish A1	.548	322	[.467, .620]	< .001	30.1%
Spanish A2	.583	346	[.509, .648]	< .001	33.9%
Spanish B1	.576	367	[.503, .640]	< .001	33.1%
Spanish B2	.613	329	[.541, .677]	< .001	37.6%
Spanish C1	.551	312	[.468, .623]	< .001	30.3%

bank loan, and the model answer to this task contained an instance of "*enquire about* n", thus possibly promoting learners' use of this expression. Topic effect, therefore, seems to be visible in the results of this study and R&B. Although few corpus-based studies to date have examined the effects of writing topic on L2 constructions, this is a critical issue to understand for making L2 written and/or spoken corpora more fruitful for L2 acquisition and writing research (Alexopoulou et al., 2017), as different topics/tasks may create unequal opportunities of construction use (Caines & Buttery, 2017). R&B and Römer and Garner (2019) also acknowledge the potential task effect on L2 VACs and call for further investigation into the impact of writing topic on L2 construction use.

Results and discussion for RQ2

Tables 9 and 10 present the results of ZINB models for learners' verb-VAC associations for "V *about* n" as a function of L1 usage norm and L2 proficiency level. The ZINB model consists of two parts: the count model that determines the count values (i.e., verb-VAC frequencies) and the zero-inflated model that predicts excess zeroes (i.e., verb-VAC combinations in the BNC that were not in the learner subset).

The count models for L1 German and L1 Spanish learners consistently indicated the BNC verb-VAC frequency as a significant predictor of the verb-VAC frequency in L2 learners' written production (estimate = .117, IRR = 1.124, $z = 6.601$, $p < .001$ for L1 German learners; estimate = .136, IRR = 1.145, $z = 7.522$, $p < .001$ for L1 Spanish learners). Nonetheless, the IRRs of this predictor for the two L1 groups of learners suggested an increase by an estimated factor of 1.124 and of 1.145 in learners' verb-VAC frequency for each unit increase in the BNC. This small estimated increase in learners' use of VAC instances with the BNC frequency can be partly attributed to the scope and the nature of writing topics in the learner corpus, which may limit the use of specific VAC instances despite their high frequencies in L1 usage. For example, *set*, *grumble*, *shout*, and *mutter* were used in "V *about* n" in the BNC, but such cases were not found in the learner corpus. "*Set about* n" means "to undertake a task" while the latter three verbs (i.e., *grumble*, *shout*, and *mutter*) entail specific ways of speaking. With

https://doi.org/10.1017/S027226312400024X Published online by Cambridge University Press

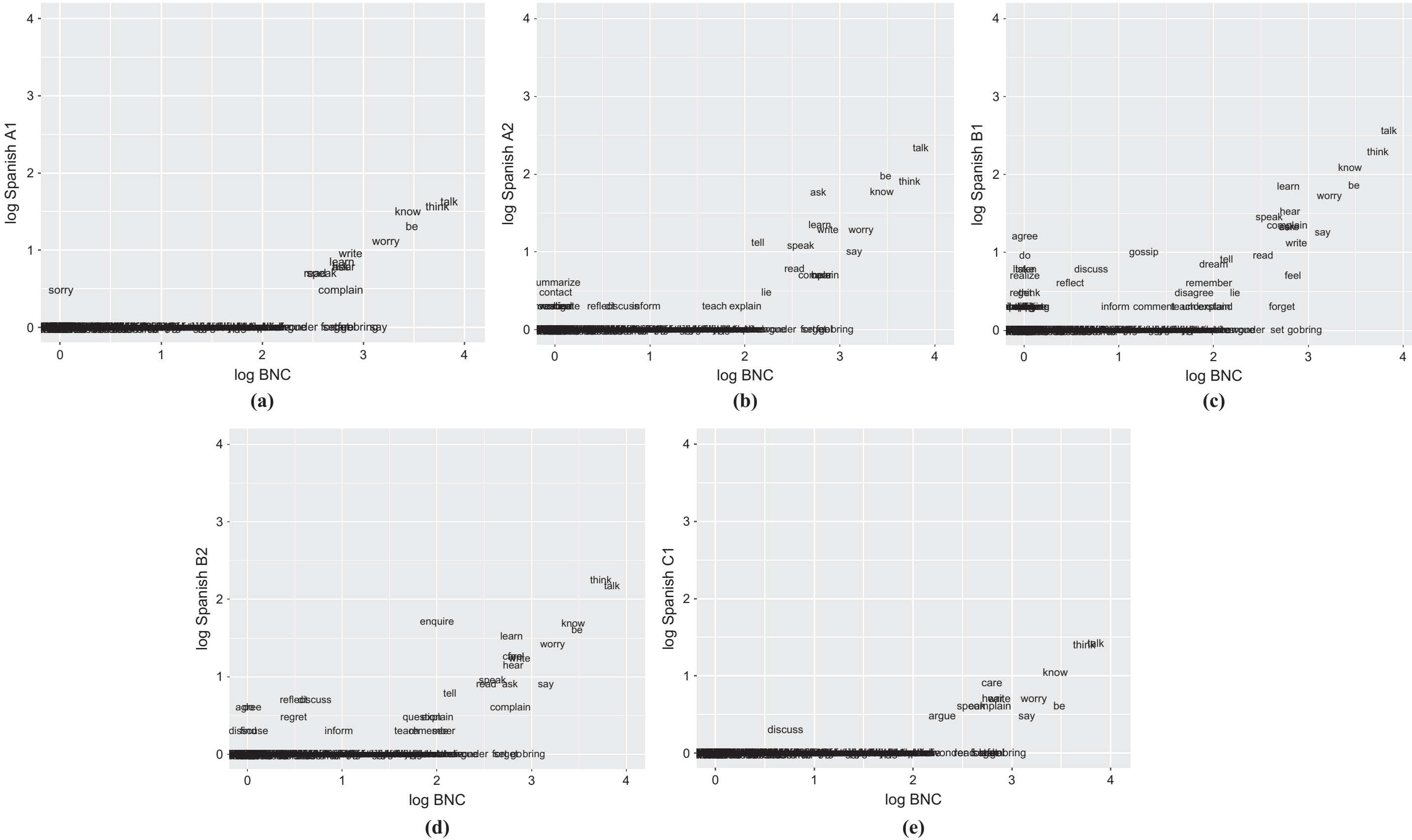

Figure 2. Correlations of verbs in L1 Spanish learners' writing and the BNC for "V *about* n".

Table 9. Results of the ZINB model for Verb-VAC combinations in L1 German learners

Predictors	Estimates	IRR	95% CI for IRR	*z*	*p*
Count model					
(Intercept)	–.073	.929	[.486, 1.776]	–.222	.824
BNC Freq	.117	1.124	[1.086, 1.164]	6.601	< .001***
Group [GA2]	–.100	.905	[.431, 1.903]	–.263	.793
Group [GB1]	.645	1.906	[.944, 3.848]	1.800	.072
Group [GB2]	1.011	2.749	[1.381, 5.473]	2.878	.004**
Group [GC1]	1.171	3.225	[1.425, 7.298]	2.810	.005**
Zero–inflated model					
(Intercept)	2.653	14.195	[6.337, 31.798]	6.447	< .001***
BNC Freq	–1.696	.183	[.091, .372]	–4.706	< .001***
Group [GA2]	–1.595	.203	[.077, .533]	–3.240	.001**
Group [GB1]	–1.761	.172	[.071, .416]	–3.907	< .001***
Group [GB2]	–1.917	.147	[.062, .350]	–4.334	< .001***
Group [GC1]	.229	1.257	[.459, 3.440]	.445	.656
Theta	.369				
AIC	2406.296				

Note: Freq = frequency; AIC = Akaike Information Criterion; IRR = incidence rate ratio; * $p < .05$; ** $p < .01$; *** $p < .001$.

Table 10. Results Of The ZINB Model For Verb-VAC Combinations In L1 Spanish Learners

Predictors	Estimates	IRR	95% CI for IRR	*z*	*p*
Count Model					
(Intercept)	–.943	.389	[.171, .886]	–2.249	.025*
BNC Freq	.136	1.145	[1.105, 1.186]	7.522	< .001***
Group [SA2]	.406	1.501	[.639, 3.523]	.933	.351
Group [SB1]	1.228	3.414	[1.493, 7.808]	2.910	.004**
Group [SB2]	2.295	9.922	[4.201, 23.436]	5.233	< .001***
Group [SC1]	2.764	15.859	[6.207, 40.523]	5.774	< .001***
Zero–inflated Model					
(Intercept)	4.729	113.128	[15.358, 833.286]	4.641	< .001***
BNC Freq	–1.921	.146	[.070, .307]	–5.089	< .001***
Group [SA2]	–4.599	.010	[.001, .084]	–4.250	< .001***
Group [SB1]	–4.674	.009	[.001, .073]	–4.453	< .001***
Group [SB2]	–3.141	.043	[.006, .322]	–3.067	.002**
Group [SC1]	–1.315	.269	[.035, 2.048]	–1.268	.205
Theta	.438				
AIC	2498.349				

Note: Freq = frequency; AIC = Akaike Information Criterion; IRR = incidence rate ratio; * $p < .05$; ** $p < .01$; *** $p < .001$.

very specific meanings, these VACs are arguably less likely to occur in argumentative writings than in narrative essays and novels. In addition to the lack of opportunities to use specific verbs arising from the limited range of topics, the absence of the above-mentioned verbs could also be partially attributed to the potential discrepancies between the L1 English corpus and the L2 input received by the learners (Kyle, et al., 2021; Lu, 2023: 85), which is particularly relevant for learners in EFL contexts, such as the majority of learners represented by EFCAMDAT.

In terms of the effects of L2 proficiency, Group [GB2] (where “G” denotes “L1 German” and “B2” denotes the proficiency level) and Group [GC1] of L1 German learners tended to have a significantly higher mean frequency of verb-VAC combinations than the reference A1 group (estimate = 1.011, IRR = 2.749, $z = 2.878$, $p = .004$ for

https://doi.org/10.1017/S027226312400024X Published online by Cambridge University Press

Group [GB2]; estimate = 1.171, IRR = 3.225, $z = 2.810$, $p = .005$ for Group [GC1]). In a similar vein, for L1 Spanish learners, Group [SB1] (B1-level; estimate = 1.228, IRR = 3.414, $z = 2.910$, $p = .004$), Group [SB2] (estimate = 2.295, IRR = 9.922, $z = 5.233$, $p < .001$) and Group [SC1] (estimate = 2.764, IRR = 15.859, $z = 5.774$, $p < .001$) were positively related to a higher mean count of verb-VAC combinations.

The zero-inflated models for both learner groups showed that BNC frequency was negatively associated with excess zeros, with IRRs much lower than 1 (estimate = –1.696, IRR = .183, $z = -4.706$, $p < .001$ for L1 German learners; estimate = –1.921, IRR = .146, $z = -5.089$, $p < .001$ for L1 Spanish learners). That means the probability of having an excess zero decreased with every unit increase in the predictor BNC frequency, regardless of L1. In addition, as shown by the negative estimates and IRRs less than 1, higher proficiency levels, such as Group [GA2], Group [GB1], Group [GB2], Group [SA2], Group [SB1], and Group [SB2], were significantly negatively related to the odds of having an excess zero in comparison to the reference level of A1. In short, the results of the ZINB models corroborated our and R&B's findings from the correlation analyses that BNC verb-VAC frequency was significantly positively associated with learners' verb-VAC frequency as well as the findings regarding the increased overlap between the verb-VAC combinations used by L1 users and learners with higher proficiency. The estimated increase in verb-VAC associations in learners with BNC frequency, however, was relatively small, suggesting the existence of other important explanatory factors for the frequency of verb-VAC combinations in the learner corpus.

Future replication research

This replication study holds several implications for future research. First, the findings of this replication suggest the potential influence of writing topic on the use of verb-VAC associations in L2 writing. While corpus-based investigations into lexical and syntactic features of L2 writing (e.g., Alexopoulou et al., 2017; Gablasova et al., 2017; Hinkel, 2009; Michel et al., 2019; Yang et al., 2015; Yoon, 2021) have consistently supported the topic effect on L2 use, less attention has been paid to the topic effect on L2 form-meaning pairings at the lexis-grammar interface (e.g., VACs). Future replication studies are recommended to examine L2 constructions across proficiency levels with better control for writing/speech topics. Another possible approach to alleviate the potential topic effect is to incorporate data from psycholinguistic experiments, such as the free association tasks in Ellis et al. (2016), that provide frame prompts with minimum context information (e.g., *she* _____ *over the* …). Integrating corpus investigations with free association tasks can also help address the data sparsity problem in learner corpus analysis, which persists even in large-scale learner corpora, such as EFCAMDAT1 and EFCAMDAT2, as demonstrated in R&B and in this replication, respectively.

Second, as a pioneering L1 English corpus compiled in the 1990s, the BNC may fall short in capturing recent language usage. Future research may explore the potential effect of using the BNC to derive the native usage norm in R&B and our study by considering a more up-to-date reference corpus as a proxy learner input corpus. An example of such a corpus may be the Corpus of Contemporary American English (Davies, 2010), which encompasses language data from 1990 to 2019.

Third, constructing more representative L2 input corpora for learners in specific EFL contexts could constitute a valuable addition to usage-based SLA research. For instance, English textbooks, classroom interactions, and extracurricular readings recommended to or selected by learners may serve as important sources of English input for many EFL learners. A collection of these and other materials learners are actually exposed to may better represent their L2 input than general-purpose L1 English corpora such as the BNC. Compiling such a corpus of learner input, however, is no easy undertaking, as it necessitates recording and transcribing classroom interactions as well as tracking and collecting out-of-class materials that learners encounter over time. Additionally, the input corpus should best be compiled in tandem with a corpus of language output by the same cohort of L2 learners, as the language input to different cohorts of learners would likely vary. For this reason, it was not feasible for R&B or our study to construct such an input corpus for the learners represented in EFCAMDAT post hoc. Future endeavors that simultaneously build learner input and output corpora could test the value of such L2 input corpora for usage-based SLA research.

A final recommendation for future replication is the inclusion of VACs which are expected to decrease with L2 proficiency but with sufficient tokens for subsequent analysis. While the token frequency of certain construction (i.e., "V *about* n," "V *with* n") increased with L2 proficiency in the present learner corpus, some other constructions (e.g., "V *that*," "*be* ADJ," etc.) are expected to decrease in use with L2 proficiency (Biber et al. 2016; Hwang & Kim, 2023). An examination of these types of constructions will enable a more thorough understanding of the L2 construction developmental trajectory.

Conclusion

As a close replication of the third research question in Römer and Berger (2019), this study investigated L1 German and L1 Spanish English learners' use of VACs across five proficiency levels (i.e., CEFR A1–C1) in comparison to L1 English usage. Overall, the frequency profile of VACs observed in the present study was consistent with that reported in the original study, although more tokens of VACs were identified in the current study. Furthermore, the results of correlation analyses enabled us to examine whether learners' VAC usage moves closer to the L1 usage norm as their proficiency increases, and the results of ZINB regression analyses estimated the influence of L1 usage norm and L2 proficiency on L2 VAC use. Addressing our first research question, correlation analyses of "V *about* n" constructions in this study confirmed the findings from R&B, including (a) learners with higher proficiency tended to use more varied types of verb-VAC combinations and (b) learners moved closer to an L1 usage norm as they became more proficient. While the correlations between the verb-VAC frequency in the BNC and the learner subcorpora remained comparatively steady across various proficiency levels in R&B, our results showed stronger correlations at higher proficiency levels, lending further support to the expanding overlap between learners' and L1 verb-VAC combinations with L2 development. We additionally found that (a) L1 German and L1 Spanish learners were able to use all three core meaning groups of verbs for "V *about* n," namely, the *talk* group, the *think* group, and the *learn* group (Francis et al., 1996) from the beginner level, and (b) the writing topic seemed to have an effect on verb-VAC combinations (e.g., the "requesting a bank loan" topic at the B2 level leads to remarkably more use of "*enquire about* n" in learner writing). Addressing RQ2, results of ZINB models on "V *about* n" indicated that BNC verb-VAC frequency and L2 proficiency were significant positive predictors of the count of verb-VAC combinations and negative predictors of the excess zeros in L2 use (i.e., verb-VAC pairings in the BNC that were not used by learners). In other words, for this focal construction, high-frequency instances in the BNC have a greater chance to be used and tend to be used more frequently by learners; and more proficient learners generally use a wider range of verb-VAC combinations attested in the BNC.

This replication study has corroborated the findings of R&B pertaining to L2 development of VACs on a larger-scale learner corpus and contributed new corpus-derived evidence in support of usage-based approaches to SLA. It also contributes to the ongoing empirical and methodological efforts toward linking learner corpus research and SLA (McEnery et al., 2019; Römer & Garner, 2022) in the following two ways. First, this study showcases the potential advantages of ZINB models in analyzing corpus frequency data with overdispersion and excess zeros, which helps enrich the statistical toolkit for corpus-based SLA research. Second, the potential influence of topic effect in the learner corpus and the caveats of using the BNC as a proxy of L2 input are highlighted.

Supplementary material. The supplementary material for this article can be found at http://doi.org/10.1017/S027226312400024X.

Competing interest. The authors declare none.

References

Alexopoulou, T., Geertzen, J., Korhonen, A., & Meurers, D. (2015). Exploring big educational learner corpora for SLA research: Perspectives on relative clauses. *International Journal of Learner Corpus Research*, *1*, 96–129.

Alexopoulou, T., Michel, M., Murakami, A., & Meurers, D. (2017). Task effects on linguistic complexity and accuracy: A large-scale learner corpus analysis employing natural language processing techniques. *Language Learning*, *67*, 180–208.

Andersen, Ø., Nioche, J., Briscoe, E. J., & Carroll, J. (2008). The BNC parsed with RASP4UIMA. In *Proceedings of the Sixth International Conference on Language Resources and Evaluation* (pp. 865–869). European Language Resources Association.

Anthony, L. (2015). *TagAnt* (Version 1.2. 0) [Computer software]. Waseda University.

Biber, D., Gray, B., & Staples, S. (2016). Predicting patterns of grammatical complexity across language exam task types and proficiency levels. *Applied Linguistics*, *37*, 639–668.

Caines, A., & Buttery, P. (2017). The effect of task and topic on opportunity of use in learner corpora. In V. Brezina & L. Flowerdew (Eds.), *Learner corpus research: New perspectives and applications* (pp. 5–27). Bloomsbury.

Davies, M. (2010). The Corpus of Contemporary American English as the first reliable monitor corpus of English. *Literary and Linguistic Computing*, *25*, 447–464.

Ekwaru, J. P., & Veugelers, P. J. (2018). The overlooked importance of constants added in log transformation of independent variables with zero values: A proposed approach for determining an optimal constant. *Statistics in Biopharmaceutical Research*, *10*, 26–29.

Ellis, N. C., & Ferreira-Junior, F. (2009). Construction learning as a function of frequency, frequency distribution, and function. *The Modern Language Journal*, *93*, 370–385.

Ellis, N. C., Römer, U., & O'Donnell, M. B. (2016). *Usage-based approaches to language acquisition and processing: Cognitive and corpus investigations of construction grammar*. Wiley-Blackwell.

Eskildsen, S. W. (2009). Constructing another language-usage-based linguistics in second language acquisition. *Applied Linguistics*, *30*, 335–357.

Francis, G., Hunston, S. & Manning, E. (1996). *Collins cobuild grammar patterns 1: Verbs*. HarperCollins.

Gablasova, D., Brezina, V., & McEnery, T. (2017). Collocations in corpus-based language learning research: Identifying, comparing, and interpreting the evidence. *Language Learning*, *67*, 155–179.

Goldberg, A. E. (1995). *Constructions: A construction grammar approach to argument structure*. University of Chicago Press.

Goldberg, A. E. (2006). *Constructions at work: The nature of generalization in language*. Oxford University Press.

Hinkel, E. (2009). The effects of essay topics on modal verb uses in L1 and L2 academic writing. *Journal of Pragmatics*, *41*, 667–683.

Hunston, Susan. (2022). *Corpora in applied linguistics*. Cambridge University Press.
Hwang, H., & Kim, H. (2023). Automatic analysis of constructional diversity as a predictor of EFL students' writing proficiency. *Applied Linguistics*, *44*, 127–147.
Klein, D., & Manning, C. D. (2003). Accurate unlexicalized parsing. In *Proceedings of the 41st annual meeting of the association for computational linguistics* (pp. 423–430). Association for Computational Linguistics.
Kyle, K., Crossley, S., & Verspoor, M. (2021). Measuring longitudinal writing development using indices of syntactic complexity and sophistication. *Studies in Second Language Acquisition*, *43*, 781–812.
Levy, R., & Andrew, G. (2006). Tregex and Tsurgeon: Tools for querying and manipulating tree data structures. In N. Calzolari, K. Choukri, A. Gangemi, B. Maegaard, J. Mariani, J. Odijk, & D. Tapias (Eds.), *Proceedings of The Fifth International Conference on Language Resources and Evaluation* (pp. 2231–2234). European Language Resources Association.
Liu, Y. (2021). *Multiword constructions in published and student mechanical engineering research articles* [Unpublished doctoral dissertation]. The Pennsylvania State University.
Liu, Y., & McManus, K. (2020). Investigating the use of article-adjective-noun constructions in EFL writing. *TESL-EJ*, *24*. http://www.teslej.org/wordpress/issues/volume24/ej94/ej94a8/
Lu, X. (2023). *Corpus linguistics and second language acquisition: Perspectives, issues, and findings*. Routledge.
Manning, C. D., Surdeanu, M., Bauer, J., Finkel, J. R., Bethard, S., & McClosky, D. (2014). The Stanford CoreNLP natural language processing toolkit. In *Proceedings of 52nd annual meeting of the association for computational linguistics: System demonstrations* (pp. 55–60). Association for Computational Linguistics.
McEnery, T., Brezina, V., Gablasova, D., & Banerjee, J. (2019). Corpus linguistics, learner corpora, and SLA: Employing technology to analyze language use. *Annual Review of Applied Linguistics*, *39*, 74–92.
Michel, M., Murakami, A., Alexopoulou, T., & Meurers, D. (2019). Effects of task type on morphosyntactic complexity across proficiency: Evidence from a large learner corpus of A1 to C2 writings. *Instructed Second Language Acquisition*, *3*, 124–152.
O'Hara, R. B., & Kotze, D. J. (2010). Do not log-transform count data. *Methods in Ecology and Evolution*, *1*, 118–122.
Park, J. H. (2017). *Syntactic complexity as a predictor of second language writing proficiency and writing quality* [Unpublished doctoral dissertation]. Michigan State University.
Porte, G., & McManus, K. (2019). *Doing replication research in applied linguistics*. New York: Routledge.
Römer, U. (2019). A corpus perspective on the development of verb constructions in second language learners. *International Journal of Corpus Linguistics*, *24*, 270–292.
Römer, U., & Berger, C. (2019). Observing the emergence of constructional knowledge: Verb patterns in German and Spanish learners of English at different proficiency levels. *Studies in Second Language Acquisition*, *41*, 1089–1110.
Römer, U., & Garner, J. (2019). The development of verb constructions in spoken learner English: Tracing effects of usage and proficiency. *International Journal of Learner Corpus Research*, *5*, 207–230.
Römer, U., & Garner, J. (2022). What can corpus linguistics tell us about second language acquisition? In A. O'Keeffe & M. McCarthy (Eds.), *The Routledge handbook of corpus linguistics* (2nd ed.) (pp. 328–340). Routledge.
Winter, B., & Bürkner, P. (2021). Poisson regression for linguists: A tutorial introduction to modelling count data with brms. *Language and Linguistics Compass*, *15*, e12439. https://doi.org/10.1111/lnc3.12439
Winter, B., Perlman, M., & Majid, A. (2018). Vision dominates in perceptual language: English sensory vocabulary is optimized for usage. *Cognition*, *179*, 213–220.
Xu, Q. (2016). Item-based foreign language learning of give ditransitive constructions: Evidence from corpus research. *System*, *63*, 65–76.
Yang, W., Lu, X., & Weigle, S. C. (2015). Different topics, different discourse: Relationships among writing topic, measures of syntactic complexity, and judgments of writing quality. *Journal of Second Language Writing*, *28*, 53–67.
Yoon, H. J. (2021). Interactions in EFL argumentative writing: Effects of topic, L1 background, and L2 proficiency on interactional metadiscourse. *Reading and Writing*, *34*, 1–21.

Cite this article: Liu, Y., & Lu, X. (2024). Development of verb argument constructions in L2 English learners: *A close replication of research question 3 in Römer and Berger (2019). Studies in Second Language Acquisition,* 1–19. https://doi.org/10.1017/S027226312400024X