The development of verb constructions in spoken learner English

Tracing effects of usage and proficiency

Ute Römer and James R. Garner
Georgia State University | University of Florida

Based on datasets of L1 Italian and Spanish learner language culled from the Trinity Lancaster Corpus Sample, this paper investigates how verb-argument constructions (VACs) develop in the spoken English of L2 learners across proficiency levels. In addition to proficiency and L1 effects, we focus on the potential influence of native English usage on learner VAC production. Insights into learners' productive knowledge of five target VACs and the verbs used in those VACs are gained through (1) comparisons of normalized entropy scores for verbs in VACs; (2) correlation analyses comparing for each VAC the verbs produced by groups of learners and by native English speakers; and (3) regression analyses comparing learner verb-VAC associations against indices of VAC usage, including verb-VAC frequency, VAC-verb association strength and contingency. Results indicate that, across L1 backgrounds, more proficient learners are more productive in their VAC use and closer to patterns in L1 English usage than less proficient learners.

Keywords: second language development, Construction Grammar, usage-based SLA, L1 Italian and Spanish learners, frequency effects

1. Introduction and background

Recent research in first language (L1) and second language (L2) acquisition has demonstrated that we learn language by learning constructions, described as the building blocks of language (e.g. Ambridge & Lieven 2015; Ellis, Römer & O'Donnell 2016; Li et al. 2014; Tomasello 2003). Within the framework of Construction Grammar, constructions are defined as form-meaning pairings that are conventionalized in the speech community and entrenched as language knowledge in the speaker's mind (Bybee 2010; Goldberg 1995, 2003; Trousdale & Hoffmann 2013). Constructions range from morphemes (e.g., *un-* in *unhappy*), over words (e.g.,

https://doi.org/10.1075/ijlcr.17015.rom
International Journal of Learner Corpus Research 5:2 (2019), pp. 207–230. issn 2215-1478 | e-issn 2215-1486

apple) and phrases (e.g., *the apple of my eye*), to more complex and abstract syntactic frames (e.g., the ditransitive construction).

Studies in L2 English acquisition have begun to examine construction development in learner production data (Ellis & Ferreira-Junior 2009; Eskildsen 2009; Eskildsen & Cadierno 2007; Li et al. 2014). Eskildsen (2009) used a longitudinal corpus of oral classroom language produced by one L1 Spanish learner of English to study the use of the modal *can* and the emergence of its constructions. Eskildsen and Cadierno (2007) used the same corpus to study the development of negation constructions. The same longitudinal dataset was also used by Li et al. (2014) who examined developing realizations of motion constructions around verbs such as COME and GO. In all three studies the authors found that, over time, the learner's inventory of constructions grew from a small set of initially dominant fixed patterns such as *I don't know* (see also Eskildsen 2012) to a larger set of increasingly productive and abstract ones. Also working with longitudinal learner data, Ellis and Ferreira-Junior (2009) studied the development of three types of verb constructions (verb locative, verb object locative, and verb object object) in a subset of the European Science Foundation corpus. They discuss how seven learners of English (L1s Italian and Punjabi) first used predominantly semantically general verbs in each construction (e.g., GO, PUT) before expanding their repertoire to more specific verbs. They also observed a strong correlation between the frequencies of verbs in VACs produced by learners and the frequencies of verbs in VACs in the input they received, indicating a strong usage effect on learner verb-VAC associations. While these studies have provided us with interesting and important insights into L2 learner VAC development, they have been based on small sets of data produced by small numbers of learners.

Other studies on constructions in L2 learner production are based on larger corpora and data collected in psycholinguistic experiments (Ellis, O'Donnell & Römer 2014; Römer, O'Donnell & Ellis 2014; Römer, Roberson, O'Donnell & Ellis 2014; Römer, Skalicky & Ellis 2018), but do not focus on language development. Ellis, O'Donnell and Römer (2014) and Römer, O'Donnell and Ellis (2014) collected responses given by advanced L2 English learners (L1s Czech, German, and Spanish) and native English speakers in free association tasks that asked them to generate the first verb that came to mind to fill the blank when presented with VAC frame prompts such as 'she _____ over the…' or 'it _____ about the…'. In correlation analyses, the learners' verb responses to the selected VAC frames were then compared with verb frequency information for the same VACs derived from 100 million words of native English usage as captured in the British National Corpus (BNC; Burnard 2007). Using the same VAC prompts, Römer, Skalicky and Ellis (2018) collected additional production data from advanced English language learners (L1s German and Spanish) and native English speakers in verbal fluency

tasks which asked participants to generate as many verbs as possible for a given VAC frame over a span of 60 seconds. In this study, verb responses for each VAC were compared across participant groups (L1 German, L1 Spanish, native speakers), again utilizing correlation analysis. The same study, as well as Römer, Roberson, O'Donnell and Ellis (2014), also used data from learner corpora (subsets of ICLE and LINDSEI) to gain further insights into L2 learners' verb-VAC associations. All of these studies indicate that L2 English learners at advanced proficiency levels have strong constructional knowledge that is influenced by usage (as captured in the BNC), and that overlaps to a considerable extent with that of L1 English speakers, but differs from it in ways that can be explained on the basis of cross-linguistic transfer and language typology effects. A limitation of these studies is that they exclusively focus on language produced by advanced L2 learners and do not include a developmental or cross-level dimension.

The current study addresses limitations of previous work on L2 verb-argument constructions by using data produced by hundreds of learners and by including data from learners at different proficiency levels (low intermediate to high advanced). Based on subsets of the cross-sectional Trinity Lancaster Corpus Sample (TLCS, further described in the following section), the study investigates how VACs (e.g. the 'V *about* n' construction, as illustrated by 'she talked about her favorite music') develop in the spoken English of L2 learners across CEFR levels B1 to C2 (Common European Framework of Reference for Languages; Council of Europe 2001). It aims to answer the following three research questions:

- RQ 1: How does L2 learners' productive verb-construction knowledge develop with increasing proficiency?
- RQ 2: Can we observe effects of the learners' L1 on verb-VAC associations?
- RQ 3: Is L2 learners' productive verb-construction knowledge affected by L1 usage and, if so, in what ways?

In addressing these questions, we hope to gain insights into how learner level and L1 background as well as variables concerning L1 construction usage affect the development of the target VACs in spoken L2 learner English. In what follows, we will give an overview of the data and methods used, present selected results, discuss those results with reference to our research questions and previous related studies, and lay out future directions for research on L2 construction development.

2. Data and methods

The data for our study come from a subset of the Trinity Lancaster Corpus Sample (TLCS), a cross-sectional corpus of spoken EFL test-taker production at intermediate and advanced proficiency levels, described in detail in Gablasova, Brezina, McEnery and Boyd (2015) and Gablasova, Brezina and McEnery (this issue). Since our previous work on VACs in advanced learner production across different L1s has shown that Spanish learners produce less target-like verbs (especially in directed motion constructions) than L1 Czech and L1 German learners at the same proficiency level (Römer, O'Donnell & Ellis 2014), we were particularly interested in studying the development of constructions in L2 English learners from Romance language backgrounds. Because of this specific interest, we created TLCS subcorpora that capture language produced by L1 Italian and L1 Spanish learners, and based our VAC analyses on those subcorpora. We included data from all three task types that the learners completed during their test session (discussion and conversation tasks at B1, B2 and C1/C2; interactive task at B2 and C1/C2; presentation task only at C1/C2). We decided to include all available tasks to ensure sufficient sample sizes for our quantitative analyses. We were also interested in seeing what learners at a particular proficiency level are capable of in terms of VAC production, independent of whether the task they completed was more monologic or more dialogic in nature. Although we did not treat TLCS tasks separately and systematically investigate the data with respect to task, we paid attention to possible task-specific patterns of verb usage in the qualitative analysis of VAC concordances and did not notice any dominant patterns.

The nine TLCS subcorpora used in this study are listed in Table 1, together with information on the learners' first language, country of residence, proficiency level (operationalized as the CEFR band at which performances were rated), and subcorpus size in number of test takers and words. The name of each subcorpus reflects the learners' country of residence and proficiency level, which are the two main variables used in our study. Subcorpus Ita-B1, for instance, contains language produced by learners from Italy (L1 Italian) at CEFR level B1 ("intermediate"); Mex-C1/C2 consists of language produced by learners from Mexico (L1 Spanish) at levels C1 ("advanced") and C2 ("proficiency"). The TLCS groups the two C levels together, so it was not possible for us to carry out separate C1 and C2 searches. Our subcorpora range in size from around 81,000 to 144,000 words, with learners at higher proficiency levels producing more words per test session than their lower proficiency peers. Overall we worked with datasets adding up to just over one million words of spoken L2 learner language produced by 750 test takers.

Table 1. Overview of TLCS subsets

Subcorpus	Learner L1	Country	Proficiency level	No. of speakers	Word count
Ita-B1	Italian	Italy	B1	100	81,367
Ita-B2	Italian	Italy	B2	100	135,408
Ita-C1/C2	Italian	Italy	C1 and C2	60	144,099
Mex-B1	Spanish	Mexico	B1	100	83,474
Mex-B2	Spanish	Mexico	B2	70	96,388
Mex-C1/C2	Spanish	Mexico	C1 and C2	60	140,945
Spa-B1	Spanish	Spain	B1	100	75,432
Spa-B2	Spanish	Spain	B2	100	122,396
Spa-C1/C2	Spanish	Spain	C1 and C2	60	127,493
Total				**750**	**1,007,002**

We also use data collected in previous studies on the distribution of VACs in the BNC. In the present study, the BNC serves as a proxy for L1 usage, not as a direct comparison corpus for TLCS. The BNC VAC data used here include verb frequency lists for selected VACs, as well as verb-VAC contingency, directional association, and faithfulness statistics further described below.

From each of the nine TLCS subcorpora listed in Table 1, we exhaustively retrieved instances of five VACs that previous research on a larger set of constructions (Römer, Roberson, O'Donnell & Ellis 2014; Römer, Skalicky & Ellis 2018) found to be particularly frequent in L2 learner speech. The VACs we included in our analysis are: 'V *about* n' (e.g., "I don't care about fashion"; Ita-B1), 'V *for* n' (e.g., "well I ask for a taxi driver"; Mex-B2), 'V *in* n' (e.g., "I don't believe in ghosts"; Spa-B2), 'V *like* n' (e.g., "they look like real people"; Spa-C1/C2), and 'V *with* n' (e.g., "you must talk with your cousin"; Ita-C1/C2). Concordance searches for these five VACs were carried out in SketchEngine (Kilgarriff et al. 2014). We used CQL (Corpus Query Language) syntax searches in SketchEngine to extract any word tagged as a verb, followed by zero or one word (to capture instances with hesitation markers such as *er* or *erm* between the verb and preposition), followed by either *about, for, in, like*, or *with*. The resulting 45 concordances were saved as text files, opened in Excel, and filtered manually for true hits of each construction (a verb form, followed by one of the five prepositions, followed by a noun or noun phrase). Both authors checked all concordances and worked with clear definitions of what constitutes an instance of each target VAC. The manual filtering reduced the number of concordance lines from 19,406 initially retrieved results to 7,067 across the five target VACs. The token distribution across VACs post filtering was as follows: 1,843 for 'V *about* n'; 704 for 'V *for* n'; 2,399 for 'V *in* n'; 274 for 'V *like* n'; and 1,847 'V *with* n'.

Based on the filtered results files, we created frequency-sorted lemmatized verb lists and verb type-token overviews for each VAC and subcorpus (e.g., 'V *for* n' in Ita-B1). These VAC- and learner-group-specific type-token and verb frequency lists served as input for various analyses and comparisons, including entropy analysis, correlation analyses, and regression analyses. A first step towards addressing RQ 1 was a comparison of the verb type and token distributions in each of the five VACs across learner proficiency levels and L1s. We did this by comparing lead verbs (i.e. the most frequent verbs) in each VAC across datasets and by calculating normalized entropy scores for each VAC and L1-proficiency combination. Normalized entropy measures the uncertainty of a probability distribution, in our case the distribution of verb types within a construction, and ranges from 0 to 1 (Kumar, Kumar & Kapur 1986). Values closer to 1 indicate more even distributions in which all verbs are equally likely to occur, while values closer to 0 indicate a more uneven and predictable distribution of verb types (potentially with one or two very frequent verb types). We chose to use normalized entropy instead of type-token ratio because it is more sensitive to Zipfian frequency distributions (Eeg-Olofsson & Altenberg 1994; Gries & Ellis 2015). Ellis and O'Donnell (2014) found that VACs in the BNC have significantly lower entropy scores than type-token frequency matched control constructions, with the most frequent verb types in each VAC accounting for the lion's share of tokens. The authors' results also showed that the most frequent verb types in each VAC were prototypical of the construction and general in their semantics. In addition to calculating normalized entropy scores for each L1-proficiency subset, we also calculated scores for each VAC in the BNC, serving as our L1 usage benchmark. This allowed us to compare the distribution of verbs in VACs in each learner subset against their distributions in a large corpus that captures native English usage, the BNC.

In a second step, and in order to further address RQ 1 as well as RQ 2, we carried out correlation analyses on verb-in-VAC distributions that allowed us to systematically compare, for each VAC, the verbs that were produced by groups of learners at different proficiency levels and of different L1 backgrounds. It also allowed us to compare the preferred verb-VAC associations of learner groups with those of native English speakers, using data retrieved from the 100-million word BNC as a proxy for L1 usage (described in Ellis, Römer & O'Donnell 2016). Correlation analyses were selected because previous work has shown them to be useful in measuring how strongly correlated L1 and L2 speaker production data are to L1 usage (Ellis, O'Donnell & Römer 2014; Ellis, Römer & O'Donnell 2016). We carried out three types of comparisons of verb usage in a VAC: (1) learner groups vs. native usage (e.g. Ita-B2 vs. BNC), (2) learner groups across L1s, with proficiency kept constant (e.g. Ita-B2 vs. Mex-B2), and (3) learner groups across proficiency levels, with L1 kept constant (e.g. Ita-B2 vs. Ita-C1/C2). For each of the 27 com-

parisons, we calculated Pearson correlation coefficients (r) in R (R Development Core Team 2012) and generated scatterplots to visualize distributions of sets of verbs. All calculations were based on the log10 transformations of the verb token frequencies. To avoid missing responses as a result of logging zero, all values were incremented by 0.01.

In order to investigate how learner VAC production is predicted by VAC usage in a large reference corpus of native English language production, we conducted regression analyses comparing learner VAC production against indices of VAC usage in the BNC. We carried out separate analyses for each VAC for each of the three proficiency levels represented in our dataset (B1, B2, C1/C2). For the current study, the focal VAC usage indices were based on usage information from the BNC and included (1) verb frequency in the VAC, (2) two verb-VAC directional association strength indices, (3) two verb-VAC contingency indices, and (4) verb-VAC faithfulness. Verb-VAC directional association strength was assessed using directional mutual information (MI word to construction (MIwc) and MI construction to word (MIcw)). Verb-VAC contingency was measured using ΔP word to construction (ΔPwc) and ΔP construction to word (ΔPcw). ΔP is a contingency measure that calculates the probability of an outcome given a cue minus the probability of the outcome occurring in the absence of the cue. For instance, ΔPwc measures how strongly the verb predicts the construction by calculating the probability of the construction given the verb minus the probability of the construction without the verb. Verb-VAC faithfulness measures the proportion of a verb's total tokens that occurs in a given VAC. These indices have been used in previous studies of VAC production in L1 and L2 research. For more information concerning these indices, see Ellis, Römer and O'Donnell (2016).

The first step in our regression analysis involved log10 transformations of the learner usage and BNC usage indices. Some of the verbs produced by the learners either did not occur in the BNC or had negative MI scores, which indicates repulsion between verbs and VACs, in the BNC. In order to avoid missing responses resulting from logging these negative values, all values were incremented by 5 before being log transformed. The number 5 was chosen because the lowest negative MI score was −4.73. Next, Pearson correlations between the verb-VAC frequencies in the learner datasets and the six selected BNC usage indices were conducted. Indices that were significantly positively correlated with learner verb-VAC frequency ($p < .05$) and above $r = .10$ (i.e., a small effect size) were then retained as predictors in subsequent regression analyses. Before being included in the analysis, correlations were conducted between the significantly correlated indices in order to check for multicollinearity. If two measures were correlated above $r = .70$, only the measure with the highest correlation with learner verb-VAC frequency was retained. These indices were then included as independent

variables in a backwards stepwise regression analysis with learner verb-VAC frequency as the dependent variable. In the event that only one of the BNC usage indices significantly correlated with verb-VAC frequency in a learner dataset, that index was included as an independent variable in a simple regression with learner verb-VAC frequency as the dependent variable.

3. Results

3.1 Distribution of verbs in VACs across TLCS subsets

As a first step in addressing RQ 1, we created type-token verb lists for each VAC in each subset of the TLCS dataset and the BNC. Table 2 presents the most frequent (or lead) verb in each VAC across the 10 datasets and their percentage of tokens in the VAC. For instance, the lead verb in 'V *about* n' in the Italian B1 dataset is TALK and it accounts for 34.3% of all tokens of this VAC. For two of the VACs ('V *about* n' and 'V *like* n'), the same verb was the most frequent across all 9 learner datasets and in the BNC. For 'V *about* n', the most frequent verb across datasets was TALK, while for 'V *like* n' the most frequent verb was BE. The most frequent verb in 'V *for* n' alternated between LOOK in four datasets and BE in six. For 'V *in* n', BE was the most frequent verb in seven of the ten datasets, while LIVE was the most frequent in two datasets and GO was the most frequent in one. 'V *with* n' showed the most variation in terms of lead verbs across datasets. AGREE was the most frequent verb in four learner datasets, TALK was the most frequent in three learner datasets, GO was the most frequent in two learner datasets, and BE was the most frequent in the BNC. We observed a lot of variation in lead verb percentages across L1-level groups, but fairly consistent choices in lead verbs. The overall strong overlap in lead verbs across datasets suggests that learners at all proficiency levels have an awareness of appropriate verb-VAC combinations and do not randomly slot verbs into these constructions. However, to gain a fuller understanding of learners' VAC knowledge at different levels of proficiency, we need to look beyond the most frequent verb in each dataset and study dominant verb-VAC associations in their production more systematically. We do this in the following sections of our article.

Table 3 gives an overview of the normalized entropy scores for each VAC in each L1-proficiency subset as well as in the BNC. A lower entropy score indicates a more predictable distribution of verbs in the VAC, with the most frequent and prototypical verbs for the VAC taking the largest share of the distribution.

The entropy scores in Table 3 show that intermediate and advanced Italian, Mexican, and Spanish learners are less predictable in their production of four of

Table 2. Most frequent verbs in VACs across datasets

	V *about* n		V *for* n		V *in* n		V *like* n		V *with* n	
	Verb	**%**	**Verb**	**%**	**Verb**	**%**	**Verb**	**%**	**Verb**	**%**
Ita B1	TALK	34.33	LOOK	18.18	GO	15.87	BE	50.00	GO	15.05
Ita B2	TALK	29.55	BE	24.53	LIVE	20.70	BE	73.08	AGREE	17.87
Ita C1/C2	TALK	35.37	BE	22.50	BE	29.82	BE	72.34	AGREE	37.12
Mex B1	TALK	27.54	BE	31.91	BE	37.34	BE	76.47	TALK	12.95
Mex B2	TALK	35.35	LOOK	28.39	LIVE	28.92	BE	51.51	TALK	16.87
Mex C1/C2	TALK	34.17	LOOK	24.39	BE	30.90	BE	66.04	AGREE	13.20
Spa B1	TALK	32.50	BE	36.36	BE	24.54	BE	83.33	GO	14.29
Spa B2	TALK	36.23	BE	23.71	BE	24.71	BE	52.63	TALK	11.34
Spa C1/C2	TALK	45.29	LOOK	21.28	BE	30.89	BE	54.76	AGREE	31.01
BNC	TALK	15.81	BE	14.33	BE	31.22	BE	26.11	BE	8.16

Table 3. Normalized entropy scores for VACs across datasets

	V *about* n	V *for* n	V *in* n	V *like* n	V *with* n
Ita B1	0.76	0.87	0.79	0.86	0.89
Ita B2	0.68	0.88	0.73	0.56	0.78
Ita C1/C2	0.61	0.84	0.75	0.53	0.73
Mex B1	0.74	0.82	0.71	0.56	0.87
Mex B2	0.69	0.81	0.68	0.70	0.84
Mex C1/C2	0.64	0.81	0.72	0.57	0.84
Spa B1	0.73	0.80	0.76	0.51	0.85
Spa B2	0.67	0.80	0.71	0.79	0.82
Spa C1/C2	0.58	0.87	0.73	0.67	0.76
BNC	**0.55**	**0.63**	**0.67**	**0.57**	**0.68**

the five target VACs than the speakers represented in the BNC. For those VACs, normalized entropy scores in the BNC were lower than those in all TLCS datasets. 'V *like* n' does not follow this pattern, most likely because of low verb type and token frequencies in the learner production data (learners only use between 3 and 11 different verb types and between 8 and 53 tokens in this VAC). Additionally, Table 3 shows that, for all target VACs except 'V *like* n', there was a general trend of the higher proficiency learners exhibiting greater predictability in their verb-VAC production. For two of the VACs ('V *about* n' and 'V *with* n'), C1/C2 entropy scores were the lowest while the B1 entropy scores were the highest across all three L1 groups. For 'V *for* n' and 'V *in* n', lower entropy scores at higher proficiency levels were found for two of the L1 groups. For 'V *for* n', the Italian and Mexican C1/

C2 learners had lower entropy scores than their B1 counterparts. However, in the Italian subset, the B2 learners had the highest entropy scores. In the Mexican subset, the B2 dataset had the same entropy score as the C1/C2 dataset. In the case of 'V *in* n', entropy scores were lower for the Italian and Spanish C1/C2 learners compared to scores for their B1 counterparts. However, the learners with the lowest entropy scores in both L1 groups were the B2 learners. In response to RQ 1, these results indicate that higher proficiency learners are more in line with native English usage (as captured by the BNC) in their VAC production in that their repertoire of verbs for each VAC is more predictable, with the most frequently used verb types accounting for a larger percentage of tokens in the VAC.

3.2 Learner verb constructions across proficiency levels and L1s

To gain further insights into whether and how verb-VAC associations differ in learners across proficiency levels (RQ 1) as well as across first language backgrounds (RQ 2), we correlated verb distributions in a VAC in two datasets at a time. Table 4 displays Pearson correlations for verb-VAC associations between datasets. The possible range of values for r is 0 to 1, with 0 indicating no correlation and 1 indicating a perfect correlation. This means that the closer the value is to 1, the stronger the correlation between two datasets is. For each of the five target VACs, Table 4 lists the r-values for 33 comparisons grouped into three types. The three comparison types allow us to examine potential effects of usage, L1, and proficiency on learner VAC usage. All correlations are significant at $p < .001$.

The first nine correlations for each VAC are based on comparisons of TLCS subsets with the BNC. Overall r-values for these comparisons are fairly low, as is to be expected, given that the BNC-derived verb lists are much longer than those retrieved from the TLCS. This means that we included a large number of verbs (data points) in our comparisons for which there are no occurrences in the TLCS. This is especially true for 'V *like* n' for which we found very few occurrences in most of the TLCS subcorpora. We notice however that, except for 'V *like* n' in the Mexican learner data, correlations are always higher at C1/C2 than at B1 level, and usually somewhere in between at level B2. This indicates that, across learner L1s and VACs, the production of verbs in VACs is closer to L1 English usage patterns for high proficiency learners than for lower proficiency learners, confirming findings from our entropy analysis and addressing RQ 1.

The next type of comparison (learner groups across L1s) aimed at addressing RQ 2 resulted in strong correlations (high r-values) overall, which implies that learners from different Romance language backgrounds (Italian, Mexican Spanish, Iberian Spanish), but at the same proficiency level, use the target VACs with very similar verbs, both in terms of types and tokens. This observation led us to

combine data from the three L1 groups together in additional analyses, with the added benefit of this providing us with larger, more robust datasets for each proficiency level (B1, B2, and C1/C2). If we now go back to the BNC comparisons and look at correlations for Italian, Mexican, and Spanish learner data grouped together (correlations 10 through 12 in Table 4, shaded grey), we notice higher *r*-values throughout than for the individual comparisons. These correlations are non-trivial at the advanced proficiency levels, especially for 'V *about* n' ($r = .56$) and 'V *for* n' ($r = .51$), and indicate effects of usage frequencies on advanced L2 learners (confirming findings reported in Ellis, O'Donnell & Römer (2014), which were based on different types of learner production data collected in experimental settings). All five VACs show higher *r*-values at C1/C2 than at B1 levels.

The third and final type of comparison focused on learners across proficiency levels with L1 background controlled for. If there is a development in learner verb-VAC associations from lower to higher proficiency levels (B1 to B2 to C1/C2), we should find higher correlation values for two adjacent levels (B1 vs. B2, and B2 vs. C1/C2) than for two levels that are further apart (B1 vs. C1/C2). If we look at the last 12 rows of Table 4, we find this to be true for all combined correlations (the last three rows, shaded grey) and for most of the other L1-separated comparisons (except for 'V *about* n' in the Italian and Spanish datasets, and 'V *like* n' in the Italian datasets). Differences between correlations (for adjacent vs. non-adjacent levels) are particularly noticeable for 'V *for* n', 'V *in* n', and 'V *with* n' and more marginal for 'V *about* n', and 'V *like* n'. Learners at proficiency levels that are further apart (B1 and C1/C2) produce sets of verbs in VACs that are less similar (overlap less in terms of type and token frequencies) than learners at adjacent proficiency levels.

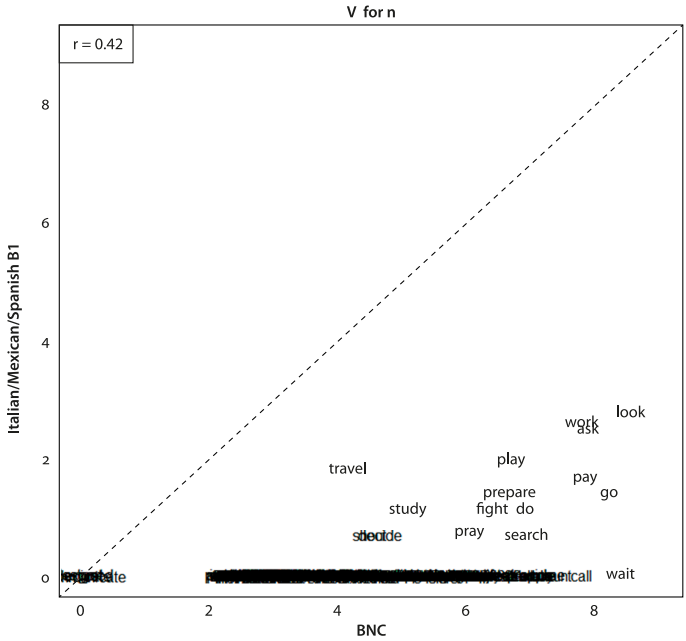
We also generated plots to visualize the correlations for all comparisons listed in Table 4. Figure 1 shows the scatterplots comparing verbs used in the 'V *for* n' VAC in the TLCS datasets at all three proficiency levels with the BNC data. The x-axis displays the logarithmic frequency of each verb type in native English usage (BNC); the y-axis shows the logarithmic frequency of each verb type in learner production (TLCS subsets). If there were perfect overlap in verbs across two datasets (i.e. an *r*-value of 1), all verb labels would be placed neatly along the diagonal that runs through the middle of the plot. Verbs that are plotted below the diagonal are markedly less frequent in the TLCS than in the BNC data; verbs that appear above the diagonal are markedly more frequent in the TLCS than in the BNC data.

In all three plots in Figure 1, the majority of verbs appear below the diagonal, which indicates that, except for the ones plotted on top of each other at $y = 0$ (resulting in a black bar), they are used by the learners but not as frequently as by native speakers. We notice that the higher the proficiency level, the more verbs move out

Table 4. Correlations (*r*-values) for verb usage in focus VACs across datasets

Type of comparison	V *about* n	V *for* n	V *in* n	V *like* n	V *with* n
Learner groups vs.NS usage					
Ita-B1 vs. BNC	0.40	0.30	0.35	0.25	0.29
Ita-B2 vs. BNC	0.49	0.39	0.39	0.30	0.33
Ita-C1/C2 vs. BNC	0.49	0.40	0.40	0.37	0.32
Mex-B1 vs. BNC	0.52	0.32	0.33	0.36	0.29
Mex-B2 vs. BNC	0.53	0.41	0.33	0.34	0.32
Mex-C1/C2 vs. BNC	0.57	0.47	0.39	0.34	0.36
Spa-B1 vs. BNC	0.45	0.33	0.34	0.25	0.30
Spa-B2 vs. BNC	0.55	0.37	0.37	0.38	0.33
Spa-C1/C2 vs. BNC	0.49	0.43	0.38	0.40	0.32
Ita/Mex/Spa-B1 vs. BNC	0.52	0.42	0.38	0.36	0.34
Ita/Mex/Spa-B2 vs. BNC	0.57	0.45	0.43	0.44	0.39
Ita/Mex/Spa-C1/C2 vs. BNC	0.56	0.51	0.46	0.46	0.41
Learner groups across L1s					
Ita-B1 vs. Mex-B1	0.74	0.54	0.73	0.88	0.72
Ita-B2 vs. Mex-B2	0.81	0.80	0.71	0.97	0.79
Ita-C1/C2 vs. Mex-C1/C2	0.85	0.73	0.77	0.79	0.75
Ita-B1 vs. Spa-B1	0.91	0.70	0.79	1.00	0.82
Ita-B2 vs. Spa-B2	0.88	0.87	0.84	0.87	0.76
Ita-C1/C1 vs. Spa-C1/C2	0.93	0.72	0.78	0.87	0.81
Mex-B1 vs. Spa-B1	0.84	0.63	0.78	0.88	0.80
Mex-B2 vs. Spa-B2	0.86	0.85	0.83	0.83	0.81
Mex-C1/C2 vs. Spa-C1/C2	0.90	0.82	0.74	0.82	0.81
Learner groups across levels					
Ita-B1 vs. Ita-B2	0.80	0.71	0.83	0.90	0.70
Ita-B1 vs. Ita-C1/C2	0.85	0.60	0.75	0.86	0.68
Ita-B2 vs. Ita-C1/C2	0.90	0.78	0.82	0.78	0.81
Mex-B1 vs. Mex-B2	0.92	0.71	0.67	0.88	0.76
Mex-B1 vs. Mex-C1/C2	0.88	0.65	0.66	0.85	0.68
Mex-B2 vs. Mex-C1/C2	0.91	0.83	0.82	0.88	0.81
Spa-B1 vs. Spa-B2	0.87	0.71	0.77	0.84	0.78
Spa-B1 vs. Spa-C1/C2	0.90	0.66	0.67	0.77	0.67
Spa-B2 vs. Spa-C1/C2	0.90	0.83	0.72	0.93	0.80
Ita/Mex/Spa-B1 vs. Ita/Mex/Spa-B2	0.89	0.84	0.84	0.94	0.82
Ita/Mex/Spa-B1 vs. Ita/Mex/Spa-C1/C2	0.88	0.79	0.76	0.87	0.74
Ita/Mex/Spa-B2 vs. Ita/Mex/Spa-C1/C2	0.91	0.89	0.80	0.89	0.88

of the black bar of verbs that appear only in the BNC to populate the plot, showing that they now also appear in the learner production data. This observation of plots becoming more populated is indicative of an increase in VAC productivity from lower to higher learner proficiency levels. We also notice that, with increasing proficiency, verbs move up along the y-axis and closer to the diagonal, indicating that learner verb use in this VAC moves closer to native speaker usage. The only verb that occurs relatively more often in the TLCS than the BNC data is EXPORT at level B2. Learners at this level repeatedly use this verb in the 'V *for* n' construction (e.g., "we export for other regions") instead of in the more idiomatic 'V n *for* n' ("we export it for other regions") or 'V *to* n' ("we export to other regions") constructions. During the manual filtering of concordance results we noticed additional unidiomatic choices made by learners across all L1 groups. For 'V *for* n' they include "depend for (something)" instead of "depend on (something)", "travel for new countries" instead of "travel to new countries", "die for cancer" instead of "die of/from cancer", and "go for myself" instead of "go by myself". These unidiomatic choices which are more common at B1 and B2 than at C1/C2 level are not necessarily highlighted in the scatterplots because they either only occur once or they actually are possible choices in the VAC but inappropriate in the larger lexicogrammatical context (e.g. "go for myself" vs. "go for it"; "die for cancer" vs. "die for you").



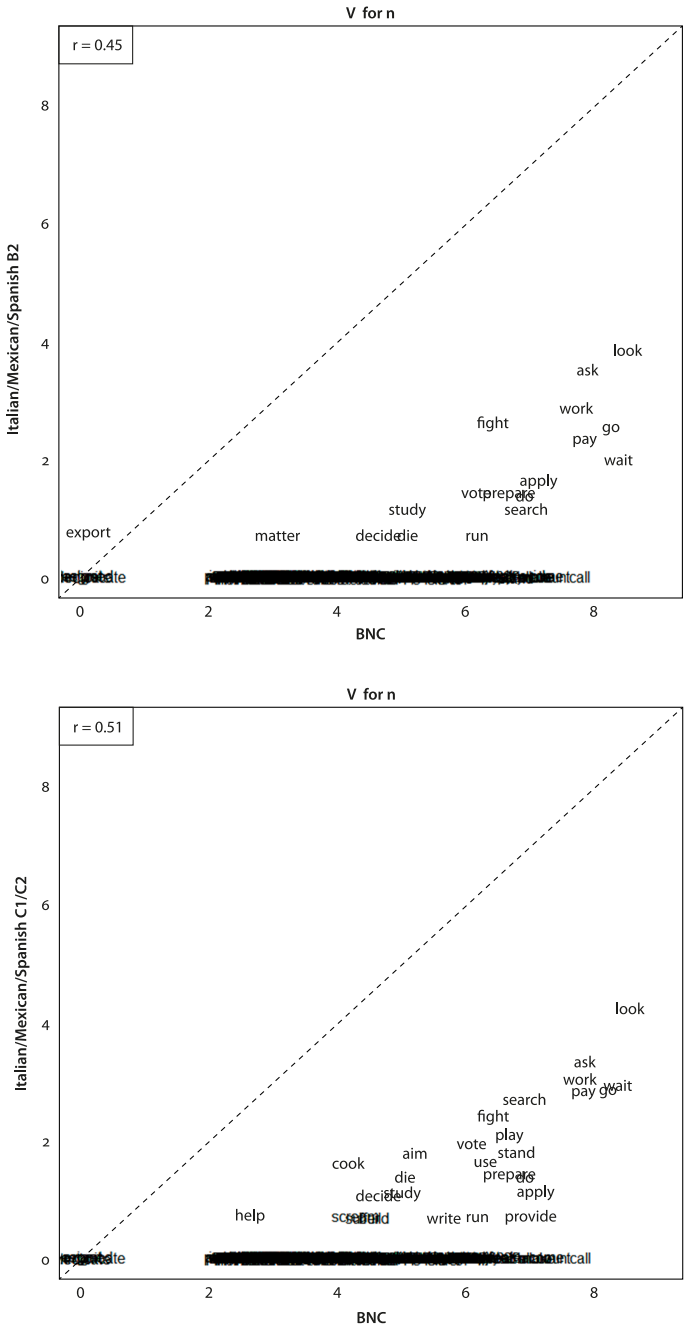
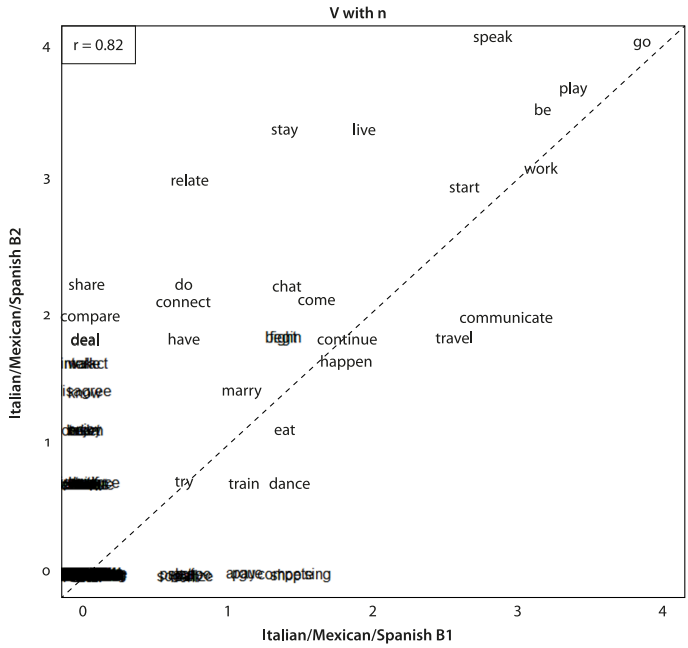


Figure 1. Correlations of verbs in learner language at three levels (B1, top plot; B2, middle plot; C1/C2, bottom plot) and native English usage (BNC) for 'V *for* n'

To further illustrate the differences in verb-VAC associations across proficiency levels, Figure 2 shows three scatterplots comparing verbs used in the 'V *with* n' constructions by learners at levels B1 to C1/C2 (all L1s grouped together). As discussed earlier, the correlations for the three comparisons differ in that levels that are closer together (B1 vs. B2, B2 vs. C1/C2) have a higher *r*-value than levels that are further apart (B1 vs. C1/C2). Additionally, the plots highlight which verbs contribute to these *r*-values and which ones are used comparatively more frequently by lower or higher level learners (or are used equally frequently by both groups). For 'V *with* n' for example, STAY, SPEAK, and SHARE are among the verbs that are preferred by B2 level learners (compared to their higher and lower level peers). These verbs appear above the diagonal in the top plot in Figure 2 and below the diagonal in the bottom plot. In the first two plots, we also notice a higher number of verbs plotted above than below the diagonal, which points to a higher productivity of the construction in the B2 and C1/C2 than the B1 learner data. In the third plot, the distribution of plotted verbs is more even and more verbs appear closer to the diagonal, pointing to similarities in 'V *with* n' usage between B2 and C1/C2 level learners.



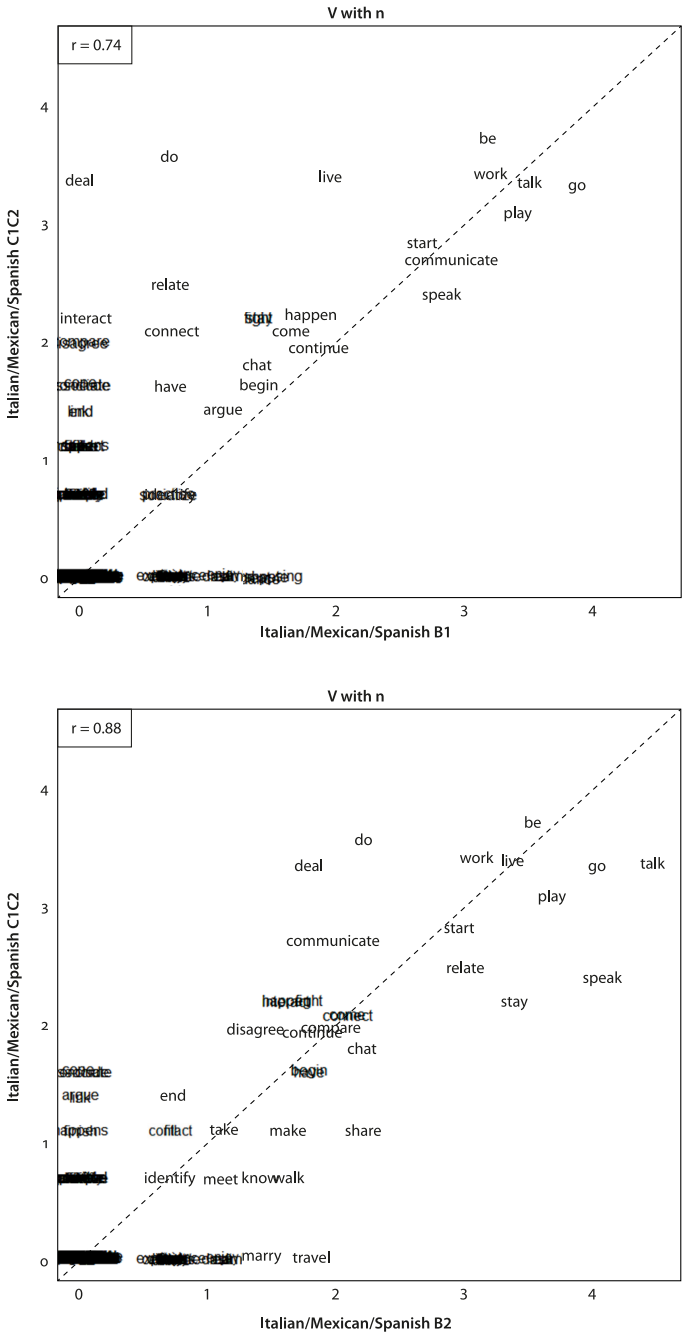


Figure 2. Correlations of verbs in learner language across proficiency levels (B1 vs. B2, top plot; B1 vs. C1/C2, middle plot; B2 vs. C1/C2, bottom plot) for 'V *with* n'

3.3 Effects of L1 usage on learner verb construction development

3.3.1 *Correlations*

With reference to RQ 3, correlations were conducted between the six selected BNC usage indices and verb-VAC frequencies at each learner proficiency level for 'V *about* n', 'V *for* n', 'V *in* n', and 'V *with* n'. 'V *like* n' was excluded due to its small sample size in the TLCS datasets. The six selected BNC usage indices include verb-VAC frequency, two measures of verb-VAC directional association (MIwc and MIcw), two measures of verb-VAC contingency (ΔPwc and ΔPcw), and one measure of verb-VAC faithfulness (i.e. the proportion of total verb usage accounted for by its use in the VAC). *R*-values and their significance for each correlation are shown in Table 5. Verb-VAC frequency in the BNC was the only index to demonstrate significant positive correlations for all target VACs at all three proficiency levels. Results for the other BNC indices varied across VACs. 'V *about* n' showed the greatest number of positive correlations, with MIwc, MIcw, ΔPwc, ΔPcw, and faithfulness significantly correlated with verb-VAC frequency in at least one proficiency level. None of these indices were found to positively correlate in 'V *for* n'. For 'V *in* n', ΔPcw demonstrated a significant positive correlation with verb-VAC frequency at each proficiency level. Results for 'V *with* n' showed a significant positive correlation between MIwc and verb-VAC frequency for the B1 learners.

Table 5. Correlations for verb production and BNC usage information

	MIwc		MIcw		ΔPwc		ΔPcw		Faith		Freq	
VAC	*r*	*p*	*r*	*p*	*r*	*p*	*r*	*p*	*r*	*p*	*r*	*p*
V about n												
B1	0.39	*	0.21	*ns*	0.40	*	0.46	*	0.40	*	0.70	**
B2	0.45	**	0.23	*ns*	0.54	**	0.59	**	0.54	**	0.79	**
C1/C2	0.36	*	0.27	*	0.46	**	0.43	**	0.13	*ns*	0.54	**
V for n												
B1	0.09	*ns*	−0.41	*	−0.10	*ns*	−0.22	*ns*	−0.10	*ns*	0.66	**
B2	0.17	*ns*	−0.36	*ns*	0.09	*ns*	0.02	*ns*	0.09	*ns*	0.72	**
C1/C2	0.26	*ns*	−0.21	*ns*	0.10	*ns*	0.16	*ns*	0.10	*ns*	0.70	**
V in n												
B1	0.07	*ns*	−0.30	*	−0.01	*ns*	0.56	**	−0.01	*ns*	0.66	**
B2	0.19	*ns*	−0.23	*ns*	0.08	*ns*	0.63	**	0.08	*ns*	0.60	**
C1/C2	0.15	*ns*	−0.32	*	0.06	*ns*	0.71	**	0.06	*ns*	0.61	**
V with n												
B1	0.22	*	−0.02	*ns*	0.04	*ns*	−0.07	*ns*	0.04	*ns*	0.52	**
B2	0.10	*ns*	−0.15	*ns*	0.08	*ns*	−0.13	*ns*	0.08	*ns*	0.47	**
C1/C2	0.15	*ns*	−0.12	*ns*	0.13	*ns*	−0.03	*ns*	0.13	*ns*	0.56	**

Signif. code:
* < 0.05. ** < 0.01.

3.3.2 *Regression analysis*

To further address RQ 3, we conducted separate backwards stepwise regression analyses for each target VAC, excluding 'V *like* n'. Results of these analyses, including coefficients for each of the included predictors, are presented in Tables 6 to 13. For three of the VACs ('V *about* n', 'V *for* n', 'V *with* n'), the regression analyses yielded models in which BNC verb-VAC frequency was the only significant predictor of verb-VAC frequency in the learner data. As can be seen by their R^2s, these models demonstrated that BNC verb-VAC frequency explained between 27% ('V *with* n' in the B1 dataset) and 62% ('V *about* n' in the B2 dataset) of the variance in the frequency of verbs produced in the VACs by the Italian, Mexican, and Spanish intermediate and advanced learners in the TLCS. In the case of 'V *in* n', the regression analyses produced models that included BNC verb-VAC frequency and BNC ΔPcw as significant predictors of verb-VAC frequency in the TLCS data. The R^2s for these models demonstrate that these two indices explain between 37% (B1 dataset) and 55% (C1/C2 dataset) of the variance in the frequency of verbs produced in the VACs by the Italian, Mexican, and Spanish intermediate and advanced learners in the TLCS. Overall, these results indicate that, across VACs, verbs that occur more frequently in VACs in L1 usage are more likely to be produced by L2 speakers at both intermediate and advanced proficiency levels. They also indicate that, for some VACs, verbs that are more strongly contingent on the VAC in L1 usage are more likely to be produced by intermediate and advanced L2 speakers.

Table 6. Regression analyses for V *about* n across proficiency levels

Model	*R*	R^2	Adjusted R^2	Std. Error	*df1*	*df2*	*F*	*p*
B1	0.70	0.49	0.47	0.23	1	29	27.96	<0.01
B2	0.79	0.62	0.61	0.22	1	51	83.00	<0.01
C1/C2	0.54	0.29	0.28	0.31	1	62	25.87	<0.01

Table 7. Coefficients for regression analyses for V *about* n across proficiency levels

Model	Coefficient	B	Std. Error	*t*	*p*
B1	(Intercept)	0.51	0.10	5.16	<0.01
	BNC VAC Freq	0.24	0.05	5.29	<0.01
B2	(Intercept)	0.45	0.07	6.74	<0.01
	BNC VAC Freq	0.32	0.04	9.11	<0.01
C1/C2	(Intercept)	0.60	0.08	7.39	<0.01
	BNC VAC Freq	0.23	0.05	5.09	<0.01

Table 8. Regression analyses for V *for* n across proficiency levels

Model	*R*	R^2	Adjusted R^2	Std. Error	*df1*	*df2*	*F*	*p*
B1	0.66	0.44	0.42	0.16	1	26	20.12	<0.01
B2	0.72	0.52	0.50	0.21	1	25	27.12	<0.01
C1/C2	0.70	0.49	0.48	0.20	1	42	40.61	<0.01

Table 9. Coefficients for regression analyses for V *for* n across proficiency levels

Model	Coefficient	B	Std. Error	*t*	*p*
B1	(Intercept)	0.48	0.10	4.64	<0.01
	BNC VAC Freq	0.17	0.04	4.49	<0.01
B2	(Intercept)	0.24	0.15	1.58	0.13
	BNC VAC Freq	0.28	0.05	5.21	<0.01
C1/C2	(Intercept)	0.30	0.11	2.74	0.01
	BNC VAC Freq	0.27	0.04	6.37	<0.01

Table 10. Regression analyses for V *in* n across proficiency levels

Model	*R*	R^2	Adjusted R^2	Std. Error	*df1*	*df2*	*F*	*p*
B1	0.61	0.37	0.35	0.25	2	70	20.61	<0.01
B2	0.69	0.48	0.46	0.26	2	64	29.05	<0.01
C1/C2	0.74	0.55	0.54	0.19	2	84	50.79	<0.01

Table 11. Coefficients for regression analyses for V *in* n across proficiency levels

Model	Coefficient	B	Std. Error	*t*	*p*
B1	(Intercept)	−102.09	31.46	−3.25	<0.01
	ΔPcw	146.94	45.12	3.26	<0.01
	BNC VAC Freq	0.13	0.05	2.46	<0.01
B2	(Intercept)	−124.39	32.45	−3.83	<0.01
	ΔPcw	178.73	46.52	3.84	<0.01
	BNC VAC Freq	0.18	0.06	3.18	<0.01
C1/C2	(Intercept)	−150.35	26.06	−5.77	<0.01
	ΔPcw	215.94	37.36	5.78	<0.01
	BNC VAC Freq	0.13	0.04	3.19	<0.01

Table 12. Regression analyses for V *with* n across proficiency levels

Model	R	R^2	Adjusted R^2	Std. Error	*df1*	*df2*	F	p
B1	0.52	0.27	0.26	0.18	1	88	32.81	<0.01
B2	0.47	0.22	0.21	0.21	1	88	25.20	<0.01
C1/C2	0.56	0.31	0.31	0.21	1	120	54.00	<0.01

Table 13. Coefficients for regression analyses for V *with* n across proficiency levels

Model	Coefficient	Estimate	Std. Error	t	Sig.
B1	(Intercept)	0.62	0.05	11.74	<0.01
	BNC VAC Freq	0.14	0.04	5.73	<0.01
B2	(Intercept)	0.68	0.05	13.10	<0.01
	BNC VAC Freq	0.12	0.02	5.02	<0.01
C1/C2	(Intercept)	0.55	0.05	10.34	<0.01
	BNC VAC Freq	0.18	0.02	7.35	<0.01

4. Conclusion and outlook

Inspired by recent research in the usage-based SLA tradition (UB-SLA) and by learner corpus research (LCR), the goal of our corpus-based study was to provide new insights into the development of constructions in L2 learner production, hence contributing to both UB-SLA and LCR and attempting to bring the two closer together. While previous studies either focused exclusively on advanced learners (CEFR level C1) and their constructional knowledge or only collected data from individual or very small groups of learners, we based our analyses on data produced by altogether 750 learners at various proficiency levels ranging from CEFR levels B1 through C2. We examined five types of verb-argument construction (VAC) in the spoken English of low-intermediate, high-intermediate, and advanced L1 Italian and L1 Spanish learners. The research questions we addressed were:

- RQ 1: How does L2 learners' productive verb-construction knowledge develop with increasing proficiency?
- RQ 2: Can we observe effects of the learners' L1 on verb-VAC associations?
- RQ 3: Is L2 learners' productive verb-construction knowledge affected by L1 usage and, if so, in what ways?

To address the first research question, we first determined and compared the frequencies of verb types and tokens in each VAC across proficiency levels, including lead verbs and their percentages. A normalized entropy analysis revealed that, for four of the five VACs, verb type and token distributions for the advanced

learners (C1/C2) were usually more predictable and closer to L1 English usage (as captured by the BNC) than for the low-intermediate (B1) learners. The strongest results were found for 'V *about* n' and 'V *with* n', with the most advanced learners (level C1/C2) having the lowest entropy scores across all three L1 groups. The results for the other three VACs, however, were more mixed. For 'V *for* n', only the L1 Italian and L1 Mexican Spanish advanced learners had lower entropy scores than their low-intermediate counterparts. For 'V *in* n', only the L1 Italian and L1 Iberian Spanish advanced learners had lower entropy scores than their low-intermediate counterparts. In correlation analyses based on verb-frequency lists for each VAC, we also compared the preferred verb-VAC associations of groups of learners with those of native English speakers. Our findings revealed that, across all L1 groups, students at C1/C2 level were closer to native usage in their VAC production than students at B1 level. The advanced learners from all three L1 groups produced more verbs in the VACs that are also used in the BNC, and in more similar distributions to those found in the BNC. Taken together, these results indicate that higher proficiency L2 speakers produce VACs more in line with native English usage than lower proficiency L2 speakers. The results also confirm findings from previous studies (Ellis, O'Donnell & Römer 2014; Römer, O'Donnell & Ellis 2014; Römer, Roberson, O'Donnell & Ellis 2014; Römer, Skalicky & Ellis 2018) that found strong similarities between advanced L2 learners' and L1 English speakers' constructional knowledge. The findings from these analyses expand on previous research by not just showing how advanced but also how low- and high-intermediate L2 learners differ from L1 English speakers.

To further study proficiency effects, we created scatterplots that revealed that, as their proficiency increased, the L2 speakers in our datasets became more productive in their VAC use. B2 and C1/C2 learners produced a wider range of verbs in the target VACs than B1 learners and were closer to L1 English usage in their verb choices. These findings support those from smaller-scale longitudinal studies of constructional knowledge development (Ellis & Ferreira-Junior 2009; Eskildsen 2009; Eskildsen & Cadierno 2007; Li et al. 2014). In each of these studies, the L2 learners' constructional inventory developed over time from a small set of fixed patterns to a larger set of more productive constructions.

The correlation analysis also helped us to address our second research question. Concerning L1 influence, we found strong correlations between Italian, Mexican Spanish, and Iberian Spanish learner groups when proficiency level was controlled for. This suggests that, at least when it comes to our five focal constructions, learners from these Romance language backgrounds are very similar in their VAC knowledge and in their verb-in-VAC production.

Addressing the third and final research question, we conducted stepwise regression analyses that compared learner VAC production against indices of VAC

usage in the BNC. Overall results indicate that verb-VAC frequency in native English usage is the only index that is consistently predictive of verb-VAC frequencies across L2 proficiency levels. For three of the five VACs analyzed in this study, BNC verb-VAC frequency accounted for between a quarter and over half of the variance in verb-VAC frequency in the learner datasets. For one VAC ('V *in* n'), BNC verb-VAC frequency and ΔPcw, a measure of contingency between the construction and the verbs that occur in it, accounted for between a third and half of the variance in verb-VAC frequency. These findings confirm those of Ellis and Ferreira-Junior (2009) and Ellis, O'Donnell and Römer (2014), who also found strong effects of native English usage on L2 construction production. Our work, however, goes beyond these studies in that it is based on a large dataset of L2 spoken English production, whereas previous findings have come from either small-scale datasets or from data collected in experimental settings.

In interpreting these results, it is important to consider potential effects of the speaking tasks learners were required to complete at each proficiency level. B1 learners completed only two speaking tasks as part of the exam. Both of these were dialogic tasks in which learners discuss a range of different topics with the examiner. B2 learners completed these two dialogic tasks and an additional interactive task that is dialogic as well but candidate-led rather than led jointly by the candidate and the examiner. C1/C2 learners completed all of these tasks and were also required to give a short presentation about a topic of their choice. The nature of these different tasks (more or less dialogic) could have influenced the learners' use of the target VACs. We chose not to control for task effects in this study because this would have left us with datasets that were insufficiently large for robust VAC analyses. Future VAC research with larger datasets of L2 speech would benefit from controlling for task effects in order to provide even stronger evidence for VAC development across proficiency levels and over time.

In addition to controlling for task effects, future research on L2 productive constructional knowledge would benefit from expanding focus to learners from additional L1 backgrounds and at even lower proficiency levels (CEFR levels A1 and A2). By including spoken data produced by learners from other, perhaps typologically different L1 backgrounds, researchers could provide a stronger analysis of possible first language effects on L2 VAC development. Another valuable step to include in future research on constructions in learner production would be to look more closely at unidiomatic VAC choices. Follow-up studies should also include additional types of VACs to see if they follow the same (or similar) developmental patterns as the target VACs selected here. Despite these limitations, we think that our study has provided relevant new insights into second language learners' developing constructicon as they progress from low intermediate to advanced levels of proficiency. These insights have helped us gain

a better understanding of some of the processes that underlie second language acquisition and highlighted the strong impact of usage on L2 development.

References

Ambridge, B. & Lieven, E. 2015. "A constructivist account of child language acquisition". In B. MacWhinney & O'Grady, W. (Eds.), *The Handbook of Language Emergence*. Malden MA: Wiley-Blackwell, 478–510.

Burnard, L. 2007. *Reference Guide for the British National Corpus* (XML edition). Retrieved from: <http://www.natcorp.ox.ac.uk/docs/URG>

Bybee, J. 2010. *Language, Usage and Cognition*. Cambridge: Cambridge University Press. https://doi.org/10.1017/CBO9780511750526

Council of Europe. 2001. *Common European Framework of Reference for Languages: Learning, teaching, assessment*. Cambridge: Cambridge University Press.

Eeg-Olofsson, M. & Altenberg, B. 1994. "Discontinuous recurrent word combinations in the London-Lund Corpus". In U. Fries, G. Tottie, & P. Schneider (Eds.), *Creating and Using English Language Corpora: Papers from the Fourteenth International Conference on English Language Research on Computerized Corpora*. Amsterdam: Rodopi, 63–77.

Ellis, N. C. & Ferreira-Junior, F. 2009. "Constructions and their acquisition: Islands and the distinctiveness of their occupancy". *Annual Review of Cognitive Linguistics 7*, 188–221. https://doi.org/10.1075/arcl.7.08ell

Ellis, N. C. & O'Donnell, M. B. 2014. "Construction learning as category learning: A cognitive analysis". In T. Herbst, H. J. Schmid & S. Faulhaber (Eds.), *Constructions Collocations Patterns*. Berlin: De Gruyter Mouton, 71–97.

Ellis, N. C., O'Donnell, M. B. & Römer, U. 2014. "Second language processing of verb-argument constructions is sensitive to form, function, frequency, contingency, and prototypicality". *Linguistic Approaches to Bilingualism* 4(4), 405–431. https://doi.org/10.1075/lab.4.4.01ell

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*. Malden MA: Wiley Blackwell.

Eskildsen, S. W. 2009. "Constructing another language: Usage-based linguistics in second language acquisition". *Applied Linguistics* 30(3), 335–357. https://doi.org/10.1093/applin/amn037

Eskildsen, S. W. 2012. "L2 negation constructions at work". *Language Learning* 62(2), 335–372. https://doi.org/10.1111/j.1467-9922.2012.00698.x

Eskildsen, S. W. & Cadierno, T. 2007. "Are recurring multi-word expressions really syntactic freezes? Second language acquisition from the perspective of usage-based linguistics". In M. Nenonen & S. Niemi (Eds.), *Collocations and Idioms 1. Papers from the First Nordic Conference on Syntactic Freezes*. Joensuu: Joensuu University Press, 86–99.

Gablasova, D., Brezina, V., McEnery, T. & Boyd, E. 2015. "Epistemic stance in spoken L2 English: The effect of task and speaker style". *Applied Linguistics Advanced Access*. https://doi.org/10.1093/applin/amv055

Goldberg, A. E. 1995. *Constructions. A Construction Grammar Approach to Argument Structure*. Chicago IL: University of Chicago Press.

Goldberg, A. E. 2003. "Constructions: A new theoretical approach to language". *Trends in Cognitive Science 7*, 219–224. https://doi.org/10.1016/S1364-6613(03)00080-9

Gries, S. T. & Ellis, N. C. 2015. "Statistical measures for usage-based linguistics". *Language Learning* 65(S1), 228–255. https://doi.org/10.1111/lang.12119

Kilgarriff, A., Baisa, V., Bušta, J., Jakubíček, M., Kovář, V., Michelfeit, J. & Suchomel, V. 2014. "*The Sketch Engine: Ten years on*". *Lexicography* 1(1), 7–36.

Kumar, U., Kumar, V. & Kapur, J. N. 1986. "Normalized measures of entropy". *International Journal of General Systems* 12(1), 55–69. https://doi.org/10.1080/03081078608934927

Li, P., Eskildsen, S. W. & Cadierno, T. 2014. "Tracing an L2 learner's motion constructions over time: A usage-based classroom investigation". *The Modern Language Journal* 98, 612–628. https://doi.org/10.1111/modl.12091

Römer, U., O'Donnell, M. B. & Ellis, N. C. 2014. "Second language learner knowledge of verb-argument constructions: Effects of language transfer and typology". *The Modern Language Journal* 98(4), 952–975. https://doi.org/10.1111/modl.12149

Römer, U., Roberson, A., O'Donnell, M. B. & Ellis, N. C. 2014. "Linking learner corpus and experimental data in studying second language learners' knowledge of verb-argument constructions". *ICAME Journal* 38, 59–79. https://doi.org/10.2478/icame-2014-0006

Römer, U., Skalicky, S. & Ellis, N. C. 2018. "Verb-argument constructions in advanced L2 English learner production: Insights from corpora and verbal fluency tasks". *Corpus Linguistics and Linguistic Theory*. Advanced online access. https://doi.org/10.1515/cllt-2016-0055

Tomasello, M. 2003. *Constructing a Language. A Usage-based Theory of Language Acquisition*. Cambridge MA: Harvard University Press.

Trousdale, G. & Hoffmann, T. (Eds.). 2013. *Oxford Handbook of Construction Grammar*. Oxford: Oxford University Press.

Address for correspondence

Ute Römer
Department of Applied Linguistics and ESL
Georgia State University
25 Park Place NE, Suite 1500
Atlanta, GA 30303
USA

uroemer@gsu.edu

Co-author information

James R. Garner
Department of Linguistics
University of Florida

james.garner@ufl.edu