



The Development of EFL Learners' Syntactic Knowledge and its Relationship to Writing Quality

Brahim Ait Hammou

*English Department, School of Arts and Humanities,
Moulay Ismail University, Meknes, Morocco
hammou76@gmail.com*

Mohammed Larouz

*English Department, School of Arts and Humanities,
Moulay Ismail University, Meknes, Morocco
m.larouz@flsh.umi.ac.ma*

Mustapha Fagroud

*Department of Agronomy, National School of Agriculture,
Moulay Ismail University, Meknes, Morocco
mfagroud@hotmail.com*

Fouad Akki

*English Department, School of Arts and Humanities,
Moulay Ismail University, Meknes, Morocco
fouad.akki@hotmail.fr*

Abstract

Knowledge of verb-argument constructions (VACs) has been examined in a variety of contexts where English is mostly taught from the early stages of their schooling. No study has so far examined EFL learners' knowledge of VACs in a context where English as a foreign language is taught only late at the secondary school level. The current study fills this gap by exploring construction knowledge in a context where teaching EFL is postponed until other languages have been introduced to learners. Using written essays of 180 students of 3 different proficiency groups, three indices of syntactic sophistication (i.e. the frequency of the Verb-argument constructions, the frequency of the verb-construction combinations and the verb-construction strength of association) were examined for their development across the three grade levels. As a second objective, this study examined the ability of these indices to predict students' writing quality, through their grades. The results suggest that while the higher proficiency group showed a significant improvement in the rareness of their VACs, the three groups showed slow progress in the use of less frequent and less strongly associated verb-construction combinations. Similarly, only the index of construction sophistication (i.e. rareness) correlated significantly with and could predict variance in the writing score. This suggests that EFL learners follow a 'necessary first' principle in their learning of constructions. Similarly, these results indicate that the

use of rare constructions enhances raters' judgements of students' writing quality.

Keywords: Verb-Argument Constructions (VACs); Frequency; Strength of Association; Assessment; Writing

Received: 2022-09-24

Accepted: 2022-12-17

Available Online: 2022-12-17

DOI: 10.22034/efl.2022.363483.1199

* Corresponding Author

1. Introduction

The present study is situated within the usage-based tradition to language learning, which is centered around the principal claim that language learning emerges from language use (e.g., [Bybee, 2010, 1995](#); [Ellis, 2002, 2016](#); [Ellis et al., 2015](#); [Langacker, 1988, 2000](#); [Tomasello, 1998a, 1998b](#)). Unlike the generativist tradition which assumes that language develops following a domain-specific faculty, the usage-based linguists assume that knowledge of language is shaped by the type of language which learners frequently encounter in their environment. Within the usage-based model, the development of language is “done with general cognitive processes, and universals of linguistic structure derive from the fact that all humans are endowed with similar general cognitive processes” ([Bavin, 2009](#), p. 85). A basic assumption of the usage-based approaches is that learning a language is the result of conducting online computational operations which are based on categorizing language elements following their frequency distributions in the input (e.g., [Bybee, 1985](#); [Ellis, 2002](#); [Langacker, 1988](#)). That is to say, learners gradually learn language based on the frequency of meaningful linguistic elements (i.e. constructions) in the linguistic environment. Higher frequency items are learnt first, then only come other lower frequency ones. [Ellis \(2006, p.8\)](#) maintains that “[l]earners FIGURE language out: their task is, in essence, to learn the probability distribution [...], the probability of an interpretation given a formal cue in a particular context (...)” Language is, therefore, seen as a process of continuous online computations and mappings of forms to meanings based on the type of input to which language learners are exposed.

The current study is concerned with the development of the sophistication of EFL university undergraduates' knowledge of syntactic constructions, known as verb-argument constructions (VACs) in the context of delayed EFL. The usage-based linguists define a construction as a form-meaning pairing (e.g., [Boyd & Goldberg, 2009](#); [Ellis, 2002, 2005](#); [Goldberg, 1995, 2003, 2006](#); [Langacker, 1987](#)). That is, any form which can be associated with meaning in language is considered a construction. A VAC refers to a syntactic construction which consists of a verb and all the arguments it takes ([Goldberg, 1995, 2006](#); [Kyle, 2016](#); [Kyle & Crossley, 2017](#)). Any simple sentence is a VAC as it contains a



verb and its arguments which might be a subject, a direct object, an indirect object, a locative, etc. An SVO construction, such as ‘*John kicks the ball*’, contains an agent argument (*John*), which causes an object (*the ball*) to transfer (*Kicks*) to a recipient argument.

2. Review of the Literature

Various studies reported that VACs are not merely linguistic forms with semantics of their own, but they also reflect a speaker’s knowledge of language in the sense that they have a psycholinguistic reality (e.g., [Boyd & Goldberg, 2009](#); [Ellis et al., 2015](#); [Ellis & Cadierno, 2009](#); [Goldberg, 1995](#)), and that language learners can make sense of syntactic constructions even in cases where they don’t know the meaning of the main verb. This shows that VACs have a meaning of their own independently of the meanings of the lexical items which occupy them (i.e. the verb). In this sense, as [Goldberg \(1995\)](#) maintains, syntactic frames are directly associated with semantics, independently of the verbs which may occur in them. [Goldberg \(1995\)](#) conducted an experiment in which she asked 10 participants about the meaning of the nonce word ‘*topamased*’ in the following sentence:

‘She topamased something to him.’

Six of ten subjects said that ‘*topamased*’ means ‘give’. Goldberg concluded that although there are other verbs which are far more frequent in the corpus than ‘give’ such as ‘tell’, ‘take’ and ‘get’, most subjects preferred ‘give’ because it is the most frequent in the ditransitive construction. This result was confirmed by [Ninio \(1999\)](#) who concluded that language learners learn verb-construction combinations first as fixed constructions which represent frequent and strongly associated verb-construction combinations in the input before they begin to learn the schematicity of the verb slot and start using novel verbs.

[Gries and Wulff \(2005\)](#) used a sentence completion task in which advanced German learners were required to complete a sentence by producing either a prepositional dative or a ditransitive construction based on a particular subject-verb combination. The results of the study indicated that German learners of English were likely to sort constructions based on their overall meaning rather than the meaning of the particular verbs. This is supported by findings from other studies which reported that even when nonce verbs are inserted in the construction, learners can still assign a semantic meaning to the construction (e.g., [Ellis & Ferreira-Junior, 2009a, 2009b](#); [Goldberg, 1995, 2003, 2006](#); [Römer, 2019](#); [Römer & Berger, 2019](#)).

Investigating how three constructions (verb locative (VL), verb object locative (VOL) and the ditransitive (VOO)) are acquired, [Ellis and Ferreira-](#)

Junior (2009a, 2009b) compared the English oral performance of non-native children of Italian and Punjabi L1 (using data from the European Science Foundation (ESF) corpus; Perdue, 1993) to their interviewer native counterpart productions. The study concluded that for each construction there is one exemplar that accounts for a large number of the verb tokens in the total productions of that construction, and that after each leading exemplar, the frequency of each subsequent verb declines, a finding which confirms the Zipfian (Zipf, 1935) distribution where a prototypical exemplar verb dominates the appearances of verb lemmas in each construction and the appearance of subsequent verbs declines by almost half of the appearances of the previous verb in the list. The study outlined very similar findings about the verbs which dominated the target constructions both in native and L2 performance. For native speaker conversation, 380 instances of 'go' constituted 42% of the total tokens of VL construction, 106 instances of 'put' constituted 35% of the total tokens of VOL and 75 instances of 'give' constituted 53% of the total tokens of the VOO construction. For the L2 English production, 380 instances of 'go' constituted 53% of the total tokens of the VL construction, 153 instances of 'put' constituted 68% of the total tokens of VOL, and 22 instances of 'give' constituted 64% of the total tokens of VOO. These results suggest that the input has a crucial role in affecting the type of verb-construction combinations that are produced by learners. Similarly, the study shows that in addition to frequency, the strength of association between the verb and the construction affects how that construction is learnt and produced. Learners usually tend to associate the construction with the verb with which it is strongly associated to the extent that the verb becomes prototypical to the construction.

In connection with learning prototypical verb-construction combinations, Eskildsen (2009) investigated the development of 'can' construction in the oral production of one Spanish native learner of English. Although Eskildsen concluded that this learner could use a variety of 'can' constructions over time, there was no evidence that he fully moved to abstract constructions. Using data from the same learner, Eskildsen and Cadierno (2007) also investigated the development of English negation construction. The study concluded that the negation construction moved from a fixed construction such as 'I don't know' to more abstract patterns. Based on this, the researchers noted that the development of constructions in L2 generally proceeds from a single potentially fixed construction to a more schematic one. Similarly, Li et al. (2014) investigated the development of motion constructions in the oral production of the same subject studied in Eskildsen and Cadierno (2007) and Eskildsen (2009) and they concluded that the development of constructions evolves first around a small set of high-frequency motion verbs such as 'go' and 'come' before learners expand their knowledge to more varied verbs. In a similar study, Eskildsen et al. (2015) examined the development of motion constructions in two Spanish speakers of L2 English and, for comparison, reviewed the study conducted by Li (2014) on



two Chinese (Mandarin and Cantonese) speaking learners of English. The researchers concluded that the linguistic means used to express motion were restricted at the early stage of learning the construction and they became more varied over time with the inclusion of more verbs. This study highlighted that while the use of ‘go’ patterns in the motion construction developed towards a general increase in their productivity, the use of ‘come’ patterns was chaotic in development. The results of the study support the idea of an exemplar-based approach to learning: ‘come’ and ‘go’ with recurrent path lexicalization (mainly ‘to’) dominated the use of verbs in the construction at the early stage of learning to use the motion construction.

Römer and Berger (2019) investigated how native German and Spanish learners of English develop their knowledge of 19 English verb-argument constructions of the type of “V with n”, “V in n” and “V of n”. The researchers observed that at the lowest proficiency level A1, for the “V in n” construction, the most frequent verb is ‘live’, followed by ‘work’ and ‘be’. At this level, these three verbs are the primary occupants of the construction. At the proficiency levels A2 and B1, additional verbs emerged in this construction (including ‘stay’, ‘sleep’, ‘stand’), while it was noticed that ‘live’, ‘work’, and ‘be’ are still the most frequent ones. The researchers noted that the three verbs continue to increase consistently across proficiency levels; at the same time, the use of the lower frequency verbs in this VAC rapidly increased. This finding supports previous results (e.g., Eskildsen & Cadierno, 2007; Gries & Wulff, 2005; Römer et al., 2014) about the early emergence of high-frequency verbs before the construction is open for less general and less frequent verbs. Römer and Berger (2019) concluded that comparisons across proficiency levels and also with native speaker norms (using BNC) showed that learners’ verb-construction associations do not only become more varied at higher proficiency levels, but they also move closer to a native usage norm.

Römer (2019) compared the verb-construction combinations in two proficiency groups of L1 German English learners (levels corresponding to CEFR A1 and B2 levels) and concluded that the most frequent combinations between the two levels indicated that A1 learners rely mostly on VAC realizations with high-frequency verbs whereas at B2 the top-10 VAC realizations also include lower frequency verbs. Again this supports the schematicity of learning VACs. Learners start from using the most frequently associated verb with the construction before using less frequent ones. This conclusion is also outlined by Römer and Yilmaz (2019) and Kyle (2016) Kyle and Crossley (2017) and also Römer and Garner (2019) who concluded that as learners advance in their L2 proficiency, they produce more VACs that are in line with native speaker usage. Römer and Garner (2019) note that “L2 learners’ constructional inventory developed over time from a small set of fixed patterns to a larger set of more productive constructions” (p. 227). Ellis et al. (2015) also noted that constructions

with more faithful association between the verb and the construction are more readily acquired by L2 learners.

3. Rationale and aim of the study

Although different studies have investigated the development of learners' syntactic knowledge from a usage-based perspective worldwide, most of these studies were conducted in contexts where English is used as a first or second language, hence using quite advanced language as their data. The present study contributes to further understanding how EFL learners develop their syntactic knowledge from a usage-based perspective by using data from learners with limited language proficiency, which is more likely to allow us notice how syntactic constructions develop at early stages of L2 (i.e. any language after L1) development. At a second level, very few studies (e.g., Kyle, 2016; Kyle & Crossley, 2017) have examined the role of knowledge of verb-argument constructions in predicting the quality of a student's writing. The present study shall contribute to dismantling how construction knowledge contributes to the quality of writing by examining the ability of three indices of construction knowledge to predict variance in the holistic scores assigned to students' writings. The current study is guided by the following research questions:

1. Do learners of different undergraduate EFL proficiency levels differ concerning the frequency of their produced VACs?
2. Do learners of different undergraduate EFL proficiency levels demonstrate similar or different knowledge in terms of the frequency and strength of association of their verb-construction combinations and the verbs which occupy them?
3. Which of these three aspects of syntactic knowledge (VAC frequency, verb-construction combination frequency and verb-construction combination strength of association) affect(s) raters' judgements of quality writing?

4. Material and Methods

4.1. Sampling and data collection procedures

To answer the three research questions, this study adopted a cross-sectional design. Written essays were collected from 180 EFL students who belong to three different undergraduate groups of EFL university learners. The target groups were semester 1 (S1; N= 60), semester 3 (S3; N= 60) and semester 5 (S5; N= 60) students majoring in English at a Moroccan university. S1 learners are university freshmen who just joined university from secondary school. Hence, their



proficiency is supposed to reflect the knowledge they built during the few hours they were exposed to English over the last four years of the secondary school (an average of two hours a week). S3 learners just started their second year at university at the time of data collection and S5 learners just started their third year.

To collect data for this study, an expository writing test was administered at the beginning of the academic year. The participants were asked to write an essay of approximately 300 words about *'the reasons which push young people to go to school besides the dream of having a job in the future'*. The expository genre was adopted for data collection because it is the genre which is dealt with most often by students, especially in secondary schools. Similarly, the subjects were required to write their essays in response to a unified topic to avoid the effect of topic complexity or lack of background knowledge on any of the groups. Participants were all given a maximum of one hour to complete their essays.

4.2. Data Analysis

After collecting the data from the informants, the hand-written texts were word-processed and then saved under a *.txt* format, as required by the corpus analysis tool. During the word-processing stage, spelling, punctuation and capitalization mistakes were corrected. Syntactic mistakes which affect the structure of the constructions were not corrected. In the very few cases where wrong words (not wrong spelling) were used, no attempt was made to correct these words. The informants were told that issues of layout such as the number of written paragraphs were not of importance to the study.

For data analysis, [Kyle's \(2016\)](#) Tool for the Automatic Analysis of Syntactic Sophistication and Complexity (TAASSC) was used. As reported in [Kyle \(2016\)](#), TAASSC uses The Stanford Parser ([Klein & Manning, 2003](#)) and Tregex ([Levy & Andrew, 2006](#)) to annotate the sentences in a text. Kyle reported that the Stanford Parser creates a constituency representation of each sentence in a text, and Tregex is used to find particular patterns in that representation. Similarly, the tool uses the Stanford Neural Network Dependency parser to provide dependency representation such as the subject and the direct object of each sentence.

The first research question examines learners' use of rare or sophisticated VACs. The second research question examines the extent to which learners use sophisticated verb-construction combinations through the use of rare and less frequent and also less strongly associated verb-construction combinations as they develop their language proficiency. The three indices are computed based on the written section of the American Corpus of Contemporary American English (COCA) ([Davies, 2008](#)). The verb-construction frequency index is related to computing the frequency of the verb-construction combinations. The index of verb-construction strength of association measures how strong is the probability

that a verb and a construction co-occur from a non-directional perspective. To measure the strength of association between the verbs and the constructions in which they occur, TAASSC uses collostructional collexeme strength (the log to the base 10 of the p-value of the Fisher-Yates exact test), (Stefanowitsch & Gries, 2003). An analysis of variance was used to test the development of the three target indices among the three proficiency groups (S1, S3 and S5 students). The frequencies which are provided in this study are normed for a million for each student's essay. Both the constructions and the verb-construction combinations are examined as types and the verbs are considered as lemmas.

The participants' essays were scored by four independent raters who were given guidelines and were directed not to pay attention to minor writing issues such as punctuation, capitalization and spelling mistakes if there are any; these mistakes were corrected beforehand. The scores of one of the raters were ultimately removed from computing the final mean writing score of each learner as he showed the lowest inter-rater reliability with the other three. The three raters whose scores were retained for the final analyses reached inter-rater reliability of $\alpha = .65$. To answer the third research question, correlations were conducted between the three target construction sophistication indices and the holistic writing scores. Because only one variable (the frequency of the VACs) correlated significantly with the holistic writing score, a simple linear regression was conducted, after checking the assumptions, to determine the amount of contribution of this variable to the students' writing quality.

5. Results

5.1. *The frequency of learners' VACs*

This section examines the sophistication of the produced VACs in the essays of the three proficiency groups. Table 1 presents the mean frequencies of learners' VACs. The frequency of the constructions which S1 learners deployed in their writing is 516.529.14. For S3, the mean VACs frequency is 456.088.67, which is a bit lower than that of S1. For S5, the mean VAC frequency decreased to 447.999.51. Note that these mean frequencies are computed for all the produced constructions based on COCA corpus. The decreasing construction mean frequency shows that learners tend to produce less frequent syntactic constructions as they enhance their overall language proficiency. Hence, the difference in frequency scores gets larger as learners improve their proficiency in language.

Table 1.

The Frequencies of Learners' Produced VACs (per a million)



Proficiency		N	Mean	Std. Dev.
Level				
S1	VAC frequency	60	516.529.15	181.874.07
S3	VAC frequency	60	456.088.67	159.264.46
S5	VAC frequency	60	447.999.51	149.149.85

Using an analysis of variance, the difference between the three proficiency groups is statistically significant ($F_{(2, 177)} = 3.129, p. < .05$). The decreasing mean frequencies as learners move to higher proficiency levels suggests that they are enhancing their syntactic knowledge and they are making use of more sophisticated constructions. The post hoc Tukey LSD test showed that only the difference between S1 and S5 is significant ($p. < .05$). The differences between both S1 and S3 and between S3 and S5 are not significant. The results suggest that the production of more sophisticated VACs seems to take at least two years before it is reflected in learners' productions. The comparisons suggest that one year of instruction does not significantly improve learners' use of rare constructions.

5.2. Verb-construction combination frequency and strength of association

This section examines the frequency of the verb-construction combinations which are produced by the students of the three groups. In this section, we check the extent to which EFL learners' syntactic knowledge becomes more sophisticated through the production of less frequent and less strongly associated verb-construction combinations. Table 2 provides the descriptive statistics of the verb-construction combinations.

Table 2.

The Verb-Construction Combination Frequency and Strength of Association

Proficiency	Index	N	Mean	Std. Dev.
Level				
S1	Verb-construction frequency	60	82.771.58	45.001.47

	Verb- construction strength of association	60	26.642.59	49.192.77
S3	Verb- construction frequency	60	82.926.57	50.788.07
	Verb- construction strength of association	60	18.051.25	47.096.38
S5	Verb- construction frequency	60	70.947.75	44.574.82
	Verb- construction strength of association	60	17.805.71	36.285.43

Table 2 shows that there's a decreasing tendency in the means of students' verb-construction combinations both at the level of their frequencies in COCA and their strength of association. At the level of frequency, S5 (mean freq. = 70.947.76) produced less frequent combinations compared to the two lower proficiency groups (S1 and S3), while S3 produced combinations which are a bit more frequent than those of S1. Using an analysis of variance, we checked the significance of the differences between the three proficiency groups. The results suggest that the observed differences were not significant ($F_{(2, 177)} = 1.28, p. >.05$) which indicates that after two years of studying English, learners still rely on the use of frequent verb-construction combinations.

Looking at the verb-construction strength of association, we could see that there is also a tendency to produce less strongly associated combinations as learners move towards their second year at university (S5). S1 students produced verb-construction combinations which are more strongly associated (mean strength of association = 26242.59) compared to S3 (mean = 18051.25) and S5 (mean = 17805.71). The means show that S5 students produced less strongly associated verb-construction combinations. The analysis of variance shows that the differences, though large, are not significant ($F_{(2, 177)} = 0.76, p. >.05$). Similar to the frequencies of the verb-construction combinations, this result suggests that



after two years of instruction, EFL university learners produced constructions mainly with their most frequent and strongly associated verbs. Hence, it can be concluded that even after two years of studying English at university, S5 learners still rely on the use of highly frequent and strongly associated verb-construction combinations.

The non-significant difference between the three groups in the verb-construction frequency and strength of association indices was further explored by the analysis of a set of the most frequent constructions in the productions of the three groups (Table 3). The analysis shows that irrespective of their proficiency level, learners use very similar verbs in their constructions and they continue to rely on the use of the most frequent and the most strongly associated verbs within the constructions.

Table 3.

A Sample of the Most Frequent Verb-Construction Combinations As Produced By the Three Groups

S1		S3		S5	
Verb-construction combination	Verb-construction Frequency	Verb-construction combination	Verb-construction Frequency	Verb-construction combination	Verb-construction Frequency
have_v-dobj	88	have_v-dobj	110	have_v-dobj	106
be_nsubj-vcop-	181	be_nsubj-vcop-	165	have_nsubj-v-dobj	55
have_nsubj-v-dobj	46	have_nsubj-v-dobj	43	be_nsubj-vcop-ncomp	155
be_nsubj-vcop-	45	be_nsubj-vcop-	135	be_nsubj-vcop-acomp	123
want_nsubj-v-	39	want_nsubj-v-	21	want_nsubj-v-xcomp	37
get_v-dobj-conj_and	9	go_v-prep_to	73	go_v-prep_to	88
make_nsubj-v-	23	make_nsubj-v-	24	be_vcop-acomp	81
read_v	8	have_v-dobj-	8	educate_v	14
know_v-ccomp	8	be_vcop-acomp	72	have_v-dobj-conj_and	11
go_v-prep_to	50	push_nsubj-v-dobj-	17	make_nsubj-v-ccomp	17
want_mark-nsubj-v-	47	be_vcop-ncomp	53	go_advmod-nsubj-v-	62
be_vcop-acomp	47	educate_v	6	make_v-xcomp	11

be_expl-vcop-nsubj	45	know_v-ccomp	8	be_vcop-ncomp	55
grow_advmod-	44	have_v-dobj-dobj	11	push_nsubj-v-dobj-	13
have_mark-nsubj-v-	15	want_mark-nsubj-v-	28	have_v-dobj-dobj	9

5.3. The relationship between learners' VACs sophistication and the quality of their writing

This section reports results related to the 3rd research question which aims at checking if there's any relationship between the three indices of syntactic sophistication as discussed above and the students' holistic writing scores. We first examined the mean writing score of each proficiency group. S1 learners had a score of 11.96 (out of a scale of 20) (sd. = 1.35), S3 learners had a mean score of 13.16 (sd. = 1.54) and finally S5 learners had a mean score of 13.93 (sd. = 1.19). The differences between the three groups are statistically significant ($F_{(23, 156)} = 3.86, p. < .05$) which suggests that the perceived quality of learners' writing by the raters matches the differences in proficiency level.

The correlations (Table 4) indicate that among the three indices (the VACs frequency, the verb-construction frequency and the verb-construction strength of association) only the index of construction frequency is significantly correlated with the holistic writing score ($r. = -.204, p. < .01$) which is consistent with the previous finding (section 2.1) which showed that this index reflected significant between-groups differences, unlike the indices of the verb-construction combination frequency and verb-construction strength of association. The results of the correlations indicate that obtaining better scores in one's writing is related to producing less frequent VACs, which suggests that the use of sophisticated, less common, syntactic constructions can be considered as an indicator of better language proficiency.

Table 4

Correlations between learners' knowledge of VACs and the holistic writing score

		Average holistic writing score	Mean construction frequency	Mean verb-construction combination frequency	Mean construction-combination strength of association
Average holistic writing score	Pearson correlation	1			
					Sig. (2-tailed)



	N	180			
Mean construction frequency	Pearson correlation	-.204**	1		
	Sig. (2-tailed)	.006			
	N	180	180		
Mean verb-construction combination frequency	Pearson correlation	-.009	.262**	1	
	Sig. (2-tailed)	.905	.000		
	N	180	180	180	
Mean construction-combination strength of association	Pearson correlation	-.083	.090	-.078	1
	Sig. (2-tailed)	.268	.227	.301	
	N	180	180	180	180

To determine the amount of variance which this variable could explain in the holistic writing score, a simple linear regression was used. Before running the analysis, we checked a few necessary assumptions. The holistic writing score, which is the dependent variable, is normally distributed. Similarly, the VIF values were checked. Because only the variable of the frequency of the VACs. The analysis turned a significant model ($F_{(1, 178)} = 7.75, p. < .01, r. = .20, R^2 = .042$). The sophistication of the produced VACs in one’s writing could, therefore, significantly predict the writing score by approximately 4%. Although the predictive ability of this variable is not very large, it could be concluded that the use of rare constructions in learners’ writing is an indicator of learners’ proficiency and the quality of their writing.

6. Discussion

Situated within the usage-based model, this study examined the development of EFL university learners’ syntactic sophistication using three syntactic indices, namely (1) the frequency of the constructions, (2) the verb-construction combination frequency and (3) the verb-construction combination strength of

association. Similarly, the relationships between these three indices and the quality of learners' writing was also examined. For construction frequency, the results showed that only the difference between university freshmen (S1) and semester 5 learners (S5) is significant while the difference between both S1 and S3 and also between S3 and S5 students did not turn any significant difference. For the indices of verb-construction combination frequency and verb-construction combination strength of association, we noticed a tendency to use less frequent and less strongly associated combinations as we compare lower to higher proficiency groups. The differences, however, are not significant.

These results are also supported by the correlations between the three indices and the holistic writing score. Only the index of construction frequency significantly correlated with learners' scores in writing. The results suggest that at early stages of learning English it might be the rareness of the produced constructions which best reflects learners' proficiency as manifested by differences both in their grade level and in their writing quality, as supported by their scores.

The results of the current study suggest that learning more sophisticated constructions might be a slow process. It takes at least the first two years at university before learners demonstrate a significant difference in the use of rare constructions. While previous research showed that learners produce more sophisticated constructions as they get exposed to larger amounts of input, our results show that this process might be slow especially in contexts where studying EFL starts later in the secondary school. It seems that the development of construction knowledge is a gradual process through which learners focus first on the most important syntactic aspects for their communicative needs before they step up to other elements of the target linguistic aspect. Similar to this conclusion, previous studies (e.g., [Ellis & Ferreira-Junior, 2009a, 2009b](#); [Kyle, 2016](#); [Römer & Burger, 2019](#)) noted that second language learners produce frequent constructions before they manage to produce less frequent ones as an effect of the input they receive.

Because we noticed that the three proficiency groups are very similar concerning their use of frequent and strongly associated verbs within the constructions, we believe that the first step of developing syntactic knowledge focuses first on learning and using more communicatively-desired constructions before learners redirect their attention, with larger amounts of input, towards learning and using the same constructions with less frequent and less prototypical verbs ([Ellis & Ferreira-Junior, 2009a, 2009b](#); [Goldberg et al., 2004](#); [Römer et al., 2018](#)).

Although more proficient learners in this study showed important improvements in their use of less frequent constructions, our analysis of the verb-



construction combinations suggests that not all aspects of constructional knowledge develop at the same pace. The high verb-construction combination frequency and strength of association scores for the three groups suggest that after two years of instruction in EFL, learners have not yet expanded their knowledge of the constructions towards the use of less frequent and less strongly associated verb-construction combinations. This is consistent with previous research (e.g., [Ellis & Ferreira-Junior, 2009a, 2009b](#); [Ellis et al., 2014a, 2014b](#); [Kyle, 2016](#); [Kyle & Crossley, 2017, 2018](#); [McDonough & Kim, 2009](#)) which maintains that more frequent and more strongly associated verb-construction combinations are learnt earlier than less frequent and less strongly associated ones. This suggests that at the level of the lexical items which are used within the verb argument slot in the constructions, learners start their production following the ‘easy first’ principle (e.g. [Beaty et al., 2021](#); [Ellis, 2002](#)). Hence, they tend to rely first on the most frequent verbs before they start producing less frequent and less strongly associated verbs with the construction. [Ellis et al. \(2014b\)](#) noted that in their study only advanced L2 learners are similar to native speakers in their use of less frequent and less strongly associated verb-construction combinations. This suggests that, in the current study, the two-year period which learners spent studying EFL is not sufficient to enhance all aspects of construction knowledge.

The analysis of a set of the most frequent constructions in learners’ productions (Table 3) shows that the three proficiency groups do not only have very similar constructions at the top of their list of produced constructions, but they also use very similar verbs within each construction. By comparing the list of constructions on the table below to the list of constructions which are most frequently used by native speakers, as generated from COCA corpus in [Kyle \(2016\)](#) (see appendix), we see that there is an overlap between the constructions which are most frequently produced by EFL students in this study and those which are produced by native speakers. This suggests that the frequency of constructions in the input affects how learners learn and produce them and that learners follow the same frequency-biased procedures as reported in previous studies (e.g., [Ellis, 2002](#); [Ellis & Ferreira-Junior, 2009a, 2009b](#); [Goldberg, 1995, 2003, 2006](#); [Goldberg et al., 2004](#); [Tomasello, 1992, 2000](#)) irrespective of the context of L2 learning. [Ellis \(2005, p.16\)](#) maintains that “language use shapes grammar through frequent repetitions of usage.”

Because we could observe very strong similarities in the list of the first most frequently used constructions for the three groups of learners, and because there are also significant differences in the construction mean frequency between university freshmen (S1) and the upper proficiency group (S5), we can conclude that there is an effect of proficiency on EFL learners’ syntactic knowledge especially at the level of the rareness of the produced constructions. In other words, the most frequently used constructions are produced by learners in very

similar ways and as we move towards less frequent constructions, learners demonstrate significant differences.

This conclusion is supported by our answer to the third research question which aimed at examining the relationship between learners' syntactic sophistication and the holistic writing score. The results show that only the index of VAC frequency correlated significantly, though with small effect size, with the writing score. This index could also explain around 4% of variance in the holistic writing score, which suggests that the use of less frequent constructions in writing does not only reflect learners' language proficiency but it is also an indicator of writing quality. The two other syntactic indices (the verb-construction combination frequency and strength of association) which are examined in this study didn't turn significant correlations with the writing score and they couldn't also significantly discriminate between the three proficiency levels. Linked to what we said above, it seems that the sophistication of the verbs which are connected with the constructions becomes an indicator both of language proficiency and better writing only at more advanced levels of learning EFL. At early stages of language development learners use data which is readily available to them, and this includes having access to the most frequently used verbs with the constructions. These results are less coherent with previous research (Kyle, 2016; Kyle & Crossley, 2017, 2018; Kyle et al., 2021) which showed that it's not only the rareness of the constructions which could predict the quality of learners' writing but also the frequency and strength of association of verb-construction combinations. This is so, probably because L2 learners in these studies are more advanced and are studying the language in contexts where English is mostly used as L1 or is studied at early stages of schooling.

It might be concluded that the development of constructional knowledge starts from focusing first on the most communicatively required linguistic aspects. Learners first use very frequent constructions with frequent lexical items and then they seem to take some time to broaden their repertoire of constructions with more sophisticated ones. The use of frequent lexical items with the constructions takes even longer before learners widen their vocabulary and start using less frequent and less strongly associated verbs with the constructions. This step seems to be of less emergency in FL learning contexts where the focus is first put on communicating one's ideas using the most readily available items, following the 'easy-first' principle.

7. Conclusion and Recommendations

The results of this study suggest that since construction frequency affects how syntactic knowledge develops, materials writers and language teachers are invited



to take this aspect into consideration in their planning of the syllabi and the teaching materials. This can be done by making use of the lists of constructions which are produced by L2 learners based on their frequency appearances in native speaker corpus. Because grammar lessons today deal with structures in traditional, non-usage-based ways, it would also be more helpful to structure grammar lessons in a usage-based fashion following the lists of constructions as they are observed to occur through various corpora. In this way, construction learning will be accelerated at a speed which is beyond that in which learners would develop the constructions by relying on a non-usage-based syllabus. Because our results showed that learners are slow in using less frequent and less strongly associated verb-construction combinations even after two years at university, it is also necessary not to heavily rely on the natural input which is presented through learners' formal language classes and to focus on including a kind of constructional knowledge in grammar lessons where learners study constructions in an organized way with a progressive focus on adding in a variety of less frequent verbs with each target construction, while the early focus should always be on the most prototypical verb with the construction. This is likely to facilitate construction learning.

The results of the current study also showed that the use of sophisticated constructions in one's writing enhances the rater's judgment of its quality. Therefore, it is important to encourage EFL learners to edit their writing productions and to include a variety of syntactic structures with a more focus on the use of less frequent ones. This shall enhance the quality of learners' writing and make it more native-like.

This study adopted a cross-sectional design in its examination of university EFL learners' constructional knowledge. Further research is needed in similar contexts where English is taught late with a more longitudinal focus. This is likely to provide a more comprehensive picture of how constructional knowledge develops in EFL. Similarly, in its assessment of the quality of learners' writing, this study used only indices of syntactic sophistication. It would also be important to complement the results of the present study with a comparison of the joint effect of sophisticated and structural or complexity-based indices (such as clausal complexity, mean length of T-unit, etc.) on the quality of the produced writing. While this study overlooked the possible effect of L1 on learners' construction knowledge, future research can take this aspect into consideration.

References

- Bavin, L., (ed.) (2009). *The Cambridge handbook of child language*. Cambridge University Press.
- Beaty, R. E., Frieler, K., Norgaard, M., Merseal, H. M., MacDonald, M. C., &

-
- Weiss, D. J. (2021). Expert musical improvisations contain sequencing biases seen in language production. *Journal of Experimental Psychology: General*.
- Boyd, J. K., & Goldberg, A. E. (2009). Input Effects Within a Constructionist Framework. *The Modern Language Journal*, 93, 418-429. <https://doi.org/10.1111/j.1540-4781.2009.00899.x>
- Bybee, J. L. (1985). *Morphology: A study of the relation between meaning and form*. John Benjamins.
- Bybee, J. L. (1995). Regular morphology and the lexicon. *Language and Cognitive Processes*, 10(5), 425-455
- Bybee, J. L. (2010). *Language, usage, and cognition*. Cambridge University Press.
- Davies, M. (2008). The Corpus of Contemporary American English (COCA): 1.0 billion words, 1990-present. <https://www.english-corpora.org/coca/>. Accessed 18th January 2022.
- Ellis, N. C. (2002). Frequency effects in language processing: A review with implications for theories of implicit and explicit language acquisition. *Studies in Second Language Acquisition*, 24(2), 143-188.
- Ellis, N. C. (2005). At the interface: Dynamic interactions of explicit and implicit language knowledge. *Studies in Second Language Acquisition*, 27, 305-352.
- Ellis, N. C. (2006). Language acquisition as rational contingency learning. *Applied Linguistics*, 27(1), 1-24.
- Ellis, N. C. (2016). Online processing of verb-argument constructions: Lexical decision and meaningfulness. *Language and Cognition*, 8(3), 391-420. doi:10.1017/langcog.2016.18
- Ellis, N. C., & Ferreira-Junior, F. (2009a). Construction learning as a function of frequency, frequency distribution, and function. *The Modern Language Journal*, 93, 370-385
- Ellis, N. C., & Ferreira-Junior, F. (2009b). Constructions and their acquisition: Islands and the distinctiveness of their occupancy. *Annual Review of Cognitive Linguistics*, 7(1), 187-221. <https://doi.org/10.1075/arcl.7.08ell>



- Ellis, N. C., O'Donnell, M. B., & Römer, U. (2014a). The processing of verb-argument constructions is sensitive to form, function, frequency, contingency, and prototypicality. *Cognitive Linguistics*, 25(1), 55-98.
- Ellis, N. C., O'Donnell, M. B., & Römer, U. (2014b). Second language verb-argument constructions are sensitive to form, function, frequency, contingency, and prototypicality. *Linguistic Approaches to Bilingualism*, 4(4), 405-431.
- Ellis, N. C., O'Donnell, M. B., & Römer, U. (2015). Usage-based language learning. In Brian MacWhinney and William O'Grady (Eds.), *The handbook of language emergence* (pp. 163-180). John Wiley & Sons, Inc.
- Ellis, N.C., & Cadierno, T. (2009). Constructing a Second Language: Introduction to the Special Section. *Review of Cognitive Linguistics. Published under the auspices of the Spanish Cognitive Linguistics Association*, 7, 111-139.
- Eskildsen, S. W. (2009). Constructing another language: Usage-based linguistics in second language acquisition. *Applied Linguistics*, 30, 335–357
- 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* (Vol. 41) (pp. 86–99). Joensuu University Press.
- Eskildsen, S. W., Cadierno, T., & Li, P. (2015). On the development of motion constructions in four learners of L2 English. In T. Cadierno, & S. W. Eskildsen (Eds.), *Usage-based perspectives on second language learning* (pp. 207–232). Walter de Gruyter.
- Goldberg, A. E. (1995). *Constructions: A construction grammar approach to argument structure*. 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](https://doi.org/10.1016/S1364-6613(03)00080-9)
- Goldberg, A. E. (2006). *Constructions at work: The nature of generalization in language*. Oxford University Press.
- Goldberg, A. E., Casenhiser, D. M., & Sethuraman, N. (2004). Learning argument structure generalizations. *Cognitive Linguistics*, 15(3), 289–316.

<https://doi.org/10.1515/cogl.2004.011>

- Gries, S. T., & Wulff, S. (2005). Do foreign language learners also have constructions? Evidence from priming, sorting, and corpora. *Annual Review of Cognitive Linguistics*, 3, 182-200
- Klein, D., & Manning, C. D. (2003). Accurate unlexicalized parsing. *Proceedings of the 41st Annual Meeting on Association for Computational Linguistics – ACL*, 03(1), 423–430. <http://doi.org/10.3115/1075096.1075150>
- Kyle, K. (2016). *Measuring syntactic development in L2 writing: Fine-grained indices of syntactic complexity and usage-based indices of syntactic sophistication* [Unpublished doctoral dissertation]. Georgia State University. Retrieved from http://scholarworks.gsu.edu/alesl_diss/35
- Kyle, K., & Crossley, S. (2017). Assessing syntactic sophistication in L2 writing: A usage-based approach. *Language Testing*, 34(4), 513-535. <https://doi.org/10.1177/0265532217712554>
- Kyle, K., & Crossley, S. (2018). Measuring syntactic complexity in L2 writing: Using fine-grained clausal and phrasal indices. *The Modern Language Journal*, 102(2), 333-349. doi: 10.1111/modl.124680026-7902/18
- Kyle, K., Crossley, S., & Verspoor, M. (2021). Measuring longitudinal writing development using indices of syntactic complexity and sophistication. *Studies in Second Language Acquisition*, 43(4), 781-812. <https://doi.org/10.1017/S0272263120000546>
- Langacker, R. (1987). *Foundations of cognitive grammar, Vol. 1. Theoretical prerequisites*. Stanford University Press.
- Langacker, R. (1988). A usage-based model. In B. Rudzka-Ostyn (Ed.), *Topics in cognitive linguistics* (pp.127–161). Benjamins.
- Langacker, R. (2000). A dynamic usage-based model. In M. Barlow & S. Kemmer (Eds.), *Usage-based models of language* (pp. 1-63). CSLI.
- Levy, R. & Andrew, G. (2006). Tregex and Tsurgeon: Tools for querying and manipulating tree data structures. *Proceedings of the Fifth International Conference on Language Resources and Evaluation (LREC'06)*. European Language Resources Association (ELRA). Retrieved from: http://www.lrec-conf.org/proceedings/lrec2006/pdf/513_pdf.pdf. Last accessed: October 1st, 2021



- Li, P. (2014). *On the development of second language learners' English motion constructions – a longitudinal usage-based classroom investigation* [Unpublished doctoral dissertation]. University of Southern Denmark.
- 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.
- McDonough, K., & Kim, Y. (2009). Syntactic priming, type frequency, and EFL learners' production of wh-questions. *The Modern Language Journal*, 93(3), 386-398.
- Ninio, A. (1999). Pathbreaking verbs in syntactic development and the question of prototypical transitivity. *Journal of Child Language*, 26(3), 619-653. <https://doi.org/10.1017/S0305000999003931>
- Perdue, C. (Ed.). (1993). *Adult language acquisition: Crosslinguistic perspectives*. Cambridge University Press.
- Römer, U. (2019). A corpus perspective on the development of verb constructions in second language learners. *International Journal of Corpus Linguistics*, 24, 268-290. doi: 10.1075/ijcl.00013.roe
- 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(5), 1089-1110. doi:10.1017/S0272263119000202
- 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 10.1075/ijlcr.17015.rom
- Römer, U., & Yilmaz, S. (2019). Effects of L2 usage and L1 transfer on Turkish learners' production of English verb-argument constructions. *Vigo International Journal of Applied Linguistics*, 16, 107-133. 10.35869/vial.v0i16.95
- Römer, U., O'Donnell, M., & Ellis, N. C. (2014). Using COBUILD grammar patterns for a large-scale analysis of verb-argument constructions: Exploring corpus data and speaker knowledge. In M. Charles, N. Groom, and L. Suganthi (Eds.), *Corpora, Grammar, Text and Discourse: In Honour of Susan Hunston* (pp. 43-71). John Benjamins.
- Römer, U., Skalicky, S. C., & 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*, 0(0) doi:10.1515/cllt-2016-0055

Stefanowitsch, A., & Gries, S. T. (2003). Collostructions: Investigating the interaction between words and constructions. *International Journal of Corpus Linguistics*, 8(2), 209–243 <https://doi.org/10.1075/ijcl.8.2.03ste>

Tomasello, M. (1992). *First Verbs: A case study of early grammatical development of cognition and action*. MIT Press. <https://doi.org/10.1017/CBO9780511527678>

Tomasello, M. (1998b). The return of constructions. *Journal of Child Language*, 25(2), 431–442. doi:10.1017/s0305000998003493

Tomasello, M. (2000). First steps toward a usage-based theory of language acquisition. *Cognitive Linguistics*, 11, 61–82. 10.1515/cogl.2001.012

Tomasello, M. (Ed.). (1998a). *The new psychology of language: Cognitive and functional approaches to language structure*. Erlbaum.

Zipf, G. K. (1935). *The psycho-biology of language: An introduction to dynamic philology*. Cambridge, MA: The M.I.T. Press.



Appendix

A list of the most frequent constructions in COCA corpus (Kyle, 2016)

Rank	Frequency per million	Verb-argument construction
1	64733.43	Verb - direct object
2	48780.10	Subject - verb - direct object
3	34540.26	Subject - verb - nominal complement
4	33315.86	Subject - verb - adjective complement
5	21321.88	Subject - verb
6	20297.22	Subject - verb - clausal complement
7	15960.63	Subject - verb - external complement
8	11788.37	Verb - clausal complement
9	11117.08	Verb
10	9879.52	Subordinator - subject - verb - direct object

This Page is Intentionally Left Blank.