

# Native and nonnative use of multi-word vs. one-word verbs

ANNA SIYANOVA AND NORBERT SCHMITT

## Abstract

*One of the choices available in English is between one-word verbs (train at the gym) and their multi-word counterparts (work out at the gym). Multi-word verbs tend to be colloquial in tone and are a particular feature of informal spoken discourse. Previous research suggests that English learners often have problems with multi-word verbs, and may even avoid their use. This study explores this issue further by comparing the likelihood of using multi-word vs. one-word verbs by both native speakers and advanced nonnatives. A questionnaire of 26 multi-word/one-word verb pairs showed that nonnatives were less likely to use multi-word verbs than native speakers in informal spoken contexts. Moreover, the amount of exposure to native-speaking environments did not have an effect on the likelihood of using the multi-word verbs. However, a corpus analysis of the same verb pairs showed that the one-word verbs are often more frequent in both written and spoken discourse.*

## 1. Introduction

One feature of English which many L2 learners find difficult is multi-word verbs, such as *put off*, *run into*, and *come up with* (Dagut and Laufer 1985; Hulstijn and Marchena 1989; Laufer and Eliasson 1993; Granger 1998; Liao and Fukuya 2004). These multi-word verbs are considered problematic for a number of reasons. At the most basic level, multi-word verbs contain two or more orthographic words working together, making them difficult to *recognize* as a single semantic unit. Unless a learner knows that a string of words is a multi-word verb, they are likely to try to decode the meanings of the individual words. Also, these verbs have to be acquired, stored and retrieved from memory as a holistic unit (Wray and Perkins 2000).

Another problem is that the meanings of multi-word verbs vary on a cline of transparency. Some are quite easy to work out from the individual words (*get back* from a trip = return), while others are virtually impossible (*brush*

*up on* your French = revise). The multi-word verbs with an *idiomatic* meaning can entail difficulty, because of the mismatch of the idiomatic meaning and the meanings of the individual words in the verb. An additional complication is that certain multi-word verbs are *polysemous* (*bring up* the tools from the basement = carry them up; *bring up* children = nurture; *bring up* a suggestion = mention) (Biber et al. 1999). Moon (1998) suggests that around 5 percent of phrasal lexemes in general are polysemous, which leads to multiple meanings being married to a string of multiple words.

There can also be *grammatical* questions concerning the use of multi-word verbs. There may be uncertainty whether the verb in question allows particle movement or not; most phrasal verbs allow particle movement, while most prepositional verbs do not (Carter and McCarthy 2006) (phrasal verb – I will *bring up* this issue tomorrow. I will *bring* this issue *up* tomorrow; prepositional verb – You need to *go up* the stairs and turn left. \*You need to *go* the stairs *up* and turn left).

If the intrinsic difficulty of multi-word verbs were not enough, they may be a totally novel construction for many English learners. Multi-word verbs are a feature of the Germanic language family, to which English belongs (Darwin and Gray 1999; Dagut and Laufer 1985), and so students who are not German or Scandinavian may be particularly unfamiliar with them, and are likely to lack the strategies to deal with them. Laufer (2000), Dagut and Laufer (1985), and Schachter (1974) argue that L2 learners tend to avoid those linguistic categories that do not exist in their L1, and so it is unsurprising that several studies have shown that learners avoid these verb forms, and struggle with them when they do try to use them. For instance, Granger (1998) found that the learners in the ICLE corpus (see below) generally used fewer amplifier collocations (*perfectly natural, closely linked*) than native speakers, but used far more active sentence-builder formulae (*I think that, I would say that*) than natives. De Cock (2000) reported that some formulaic sequences were overused, some underused, and others simply misused by nonnatives when compared to native norms. Dagut and Laufer (1985) found that the majority of intermediate Israeli learners of English exhibited a strong preference for one-word verbs. The authors attributed this finding to L1–L2 structural differences.

In contrast, Liao and Fukuya (2004) suggested that the avoidance or non-avoidance of phrasal verbs could be a manifestation of learners' interlanguage development rather than the L1–L2 differences or similarities. They observed a number of intermediate learners producing phrasal verbs much less frequently than both native speakers and advanced learners. The latter, on the other hand, did not exhibit significant avoidance behavior and demonstrated a near-native use of the target verbs. Similarly, Hulstijn and Marchena (1989) investigated avoidance behavior in advanced and intermediate Dutch learners of English when performing three tasks: multiple-choice test, memorization test and trans-

lation test. The authors concluded that neither advanced nor intermediate learners avoided phrasal verbs as a category.

In some cases, there is no obvious multi-word verb which can replace a one-word verb: to *privatize* a business. Conversely, sometimes only a multi-word verb will do: she *made up* her face, he *put on* his coat. However, the crux of the problem is that there is often a choice between a multi-word verb and a one-word verb which mean roughly the same thing (I will call you when I *get back/return* from my trip). Both members of these verb pairs can be grammatically and semantically correct, with the main difference usually being the register, or *tone* they project. Multi-word verbs tend to be “colloquial in tone” (Biber et al. 1999: 409), and are an important component of spoken, informal English (Freeborn 1995; Altenberg 1998; Biber et al. 1999), especially when compared to written discourse. One-word verbs, on the contrary, tend to have a slightly more formal register, and are thus more at ease in less colloquial contexts, for example, in academic discourse and official written reports.

For learners, the issue is not so much choosing the verb form which carries the correct meaning, but rather choosing the verb which has the appropriate register, and which conforms to the expectations of the speech community. As such, the correct selection between multi-word verbs and their one-word counterparts makes a difference in how *native-like* and communicatively-effective a learner is. For example, by failing to use multi-word verbs in their speech, many learners of English tend to sound unnatural and non-idiomatic. As Wray and Perkins (2000) point out, it is often the failure to use native-like multi-word expressions that give away advanced L2 learners as non-native.

We can conclude from this discussion that there are numerous reasons why multi-word verbs are difficult for learners, but regardless of this, they are an important component of native-like spoken discourse. This study explores multi-word verbs from a number of perspectives, touching on the usage of multi-word verbs in both written and spoken discourse, comparing native and nonnative use of these forms, using both corpus and questionnaire data, and investigating whether longer exposure to L2 environments enhances the use of these verbs. Specifically, the following questions will be addressed:

1. What are the relative frequencies of multi-word vs. one-word verbs in written and spoken English as L1?
2. Are advanced learners of English as likely to use multi-word/one-word verbs as native speakers in spoken and written contexts?
3. How does exposure to an L2 environment affect the likelihood of nonnatives using multi-word vs. one-word verbs?

## 2. Methodology

We will attempt to answer the research questions through two approaches: using corpus data and by surveying natives and nonnatives on their use of multi-word and one-word verbs.

### 2.1. *Multi-word and one-word verbs used in this study*

Of the 26 multi-word verbs in this study, 14 were taken from previous research (Biber et al. 1999; Laufer and Eliasson 1993; Liao and Fukuya 2004) and the remaining arbitrarily taken from a variety of texts and conversations we were exposed to during the time we were planning the study. They were chosen to include a variety of features: they vary in frequency, some are relatively transparent in meaning, some are two-word forms and some are three-word forms, and some are polysemous while others are not. Overall, they are intended to represent the range of multi-word verbs a learner may come across when studying and using English. For each of these multi-word verbs, a one-word verb that is roughly synonymous was chosen (*come up with/suggest*).<sup>1</sup> Thus the corpus-based portion of this study investigates 26 multi- and one-word verb pairs (see Appendix 1).

### 2.2. *Corpora consulted*

In the corpus approach, we will look at native usage of spoken multi-word verbs and their one-word counterparts by referring to the CANCODE corpus,<sup>2</sup> a collection of five million words of mainly informal spoken English. The same verbs will be compared in written discourse by consulting the 90 million word written component of the British National Corpus (BNC). Learner use of the verbs in written discourse will be determined by checking the International Corpus of Learner English (ICLE), a 2.5 million word collection of 3,640 essays written by learners from eleven European languages (Granger, Dagneaux, and Meunier 2002).

### 2.3. *Developing a verb-usage questionnaire*

In order to directly compare native and nonnative usage of the verb pairs, a questionnaire was devised which could be given to both groups. In the pilot version of the questionnaire (Appendix 2), the 26 verb pairs were set in a situation or a short dialogue. Because multi-word verbs are considered a feature of informal spoken discourse, the contexts were written in this style, often by using highly colloquial words and expressions (e.g., *mates, pub, have a chat, yeah*), and/or by using dialogues as context, for example:

- (1) *You are saying to your friend: "It was such a disappointment. We drove all the way to Manchester only to find out that the game was*

- \_\_\_\_\_ *due to the heavy rain!*”  
A. *called off*  
B. *cancelled*

In addition, the instructions to the questionnaire requested participants to use the verb forms they would normally use in colloquial English.

The questionnaire was given to 14 native speakers (undergraduates and postgraduates at British universities/professionals), who were asked to choose which of the two verb forms they were more likely to use in the situation. Because the focus of this study is multi-word verbs, we analyzed the questionnaires to find pairs of verbs in which the majority ( $\geq 60\%$ ) of the natives preferred the multi-word form. Nineteen verb pairs showed such a multi-word preference, and only these were included in the final questionnaire used in the main study. The 19 verb pairs are indicated in Appendix 1. The final questionnaire is identical to that in Appendix 2, except that Items 4, 7, 14, 15, 16, 17, and 24 were deleted.

Previous studies (Dagut and Laufer 1985; Laufer and Eliasson 1993; Liao and Fukuya 2004) required their participants to choose from one of four options: a phrasal verb, a one-word verb, or one of the two distracters. That approach forced participants to choose one of the options even though they may have had equal preference for both multi-word verb and its one-word alternative. In order to obtain a more nuanced measure of the likelihood of use for the verb forms, we opted for a 6-point Likert-type scale ranging from “very unlikely” to “very likely” (Dörnyei 2003). With this scale, participants could indicate if they preferred one verb form, or if they were quite happy to use both verbs, or neither of them. For instance, we found that subjects would often choose “likely” for one verb, and “fairly likely” for the other. This shows that although both verbs might be used by the subject in the given context, we were able to tap into a preference for one verb over the other, even if the preference was rather small. The Cronbach Alpha internal reliability (as taken from the participant ratings in the study) was .864 for the multi-word verbs and .877 for the one-word verbs.

#### 2.4. *Participants*

The questionnaire was given to natives and nonnatives in a number of settings: classrooms, via email, and by approaching subjects on the University of Nottingham campus. Questionnaires filled in incorrectly or incompletely were excluded from data analysis, which left a total of 65 natives and 65 nonnatives. The native speakers were either undergraduate or postgraduate students at the University of Nottingham, or young professionals who graduated in the previous six years. Their age range was 20–35 years old, with 17 males and 48 females.

The 65 nonnatives were chosen on the basis of their advanced English language proficiency, and their L1 being outside the Germanic group of languages, e.g., Arabic, Russian, Italian, and Chinese. They belonged to one of two categories:

- 40 international undergraduate and postgraduate students who were studying a variety of subjects at the University of Nottingham (20–35 years old; 8 M and 32 F).
- 25 young professionals who studied English at the university level at overseas or UK universities and had graduated in the previous six years. Most of them were working as translators or interpreters, teachers of English, or held managerial and administrative posts. They all worked with English on a daily basis, and their proficiency could be considered as advanced, although no direct measures were taken of this. Some of them lived in the UK, while others resided in their home countries (22–35 years old; 5 M and 20 F).

### 3. Results

#### 3.1. *What are the relative frequencies of multi-word vs. one-word verbs in written and spoken English as L1?*

The literature suggests that multi-word verbs are an important feature of English, particularly in informal spoken English (Freeborn 1995; Altenberg 1998; Biber et al. 1999). But few studies have looked at frequencies of multi-word verbs in comparison to frequencies of one-word verbs which could be used in their place. We carried out a frequency analysis on the 26 verb pairs by consulting the CANCODE (native spoken) and BNC (native written) corpora (see Appendix 1). In terms of absolute frequency, half of the multi-word verbs (13) occurred more than 10 times per million running words in the CANCODE, the frequency of words like *passive*, *deer*, and *lounge*. However, 11 of the others occurred rarely, and two (*tell off*, *call off*) did not occur at all. On the other hand, only eight multi-word verbs occurred 10 times or more per million in the BNC, with the remaining (except *tell off*) appearing at low frequencies. These results are consistent with previous findings that multi-word verbs are more frequent in spoken than written discourse. On the other hand, many of the multi-word verbs are relatively infrequent in either mode of discourse.

However, because we are interested in the verb choices learners have, the more interesting analysis is the comparison in frequencies between multi-word verbs and their one-word verb competitors. In most cases, the one-word form was more frequent than the multi-word form in *both* spoken and written discourse (Table 1). When we compared the verb forms in the BNC written corpus, we found that 18 of the 26 verb pairs (69%) were more frequent in their one-word forms in written discourse. This was expected, but when we did the same analysis in the CANCODE spoken corpus, we found that this was also

Table 1. Relative frequencies of multi-word and one-word verb pairs ( $n = 26$ )

	MWV>OWV	MWV=OWV	MWV<OWV
CANCODE	5	4	17
BNC written component	2	6	18
ICLE	0	8	18

MWV multi-word verb; OWV one-word verb; > more frequent than; < less frequent than; = equal within 5 occurrences per million

true for 17 of the verb pairs (65%). This suggests that although multi-word verbs are more frequent in spoken than written discourse, they may *not* be more frequent than their one-word verb alternatives in either mode.

To further explore the choice between multi-word verbs and their one-word verb competitors, we calculated the relative proportion of occurrence of each (Table 2). For example, we added the total occurrences of *mess around* + *misbehave*, and calculated the proportion of each verb. *Mess around* occurred in 71% of the cases in the written BNC corpus, with *misbehave* occurring in 29%. Thus with this verb pair, the multi-word verb is preferred even in written discourse. In the spoken CANCODE corpus, the preference is even stronger: *mess around* = 77%, *misbehave* = 23%. However, when we looked at the proportions in the ICLE corpus, we found that learners had a strong tendency to use *misbehave* (67%) instead of *mess around* (33%).

Looking at the 26 verb pairs in Table 2, we find that some have the multi-word alternative more prominent in spoken discourse, and the one-word alternative more prominent in written discourse. For two pairs (*go on/continue*; *go up/rise*), the preference is very strong. This behavior was expected, but only occurred in three of the 26 pairs. We also find pairs where the preference was for the multi-word alternative in both written and spoken modes, but this occurred in only 5 of 26 cases.

The most noticeable feature of Table 2 is the large number of pairs (18: 69%) in which the one-word alternative was most prominent (i.e., by more than 5 percentage points) in both modes. Moreover, the degree of preference was consistently very large in both written and spoken discourse. The 26 verb pairs in the study were not randomly chosen, but do represent a mixture of the kind of multi-word verbs and their one-word alternatives which are available in English. To the extent that they are representative, the results suggest that one-word verbs generally seem to be preferred to their multi-word verb alternatives, even in spoken discourse.

Table 2. *Percentage of use of multi-word verbs vs. one-word verbs*

	CANCODE		BNC		ICLE	
	MWV %	OWV %	MWV %	OWV %	MWV %	OWV %
turn down : decrease	56	44	35	65	5	95
go on : continue	86	14	40	60	42	58
go up : rise	93	7	31	69	18	82
put off : postpone	100	0	58	42	33	67
work out : train	72	28	67	33	56	44
mess around : misbehave	77	23	71	29	33	67
come up : arise	95	5	60	40	49	51
show off : boast	88	12	56	44	44	56
tell off : reproach	0	100	0	100	0	100
put up with : stand	12	88	9	91	7	93
run into : meet	3	97	5	95	2	98
come around : come	3	97	1	99	0	100
tidy up : organize	17	83	17	83	3	97
pull over : stop	3	97	0	100	0	100
set up : start	12	88	46	54	9	91
get back : return	45	55	25	75	38	62
figure out : understand	2	98	6	94	1	99
walk off : leave	0	100	0	100	0	100
come up with : suggest	35	65	20	80	35	65
hold on : wait	8	92	19	81	16	84
bring up : mention	11	89	14	86	23	77
come along : join	32	68	10	90	2	98
come across as : seem	11	89	0	100	0	100
catch up with : join	3	97	5	95	3	97
call off : cancel	0	100	11	89	0	100
brush up on : revise	5	95	3	97	17	83

Figures are rounded to nearest full percentage.

### 3.2. *Are advanced learners of English as likely to use multi-word/one-word verbs as native speakers in spoken and written contexts?*

3.2.1. *Corpus analysis-written discourse.* In order to judge the relative use of the verb pairs in written discourse between natives and nonnatives, we compared data from the BNC with that in the learner-based ICLE corpus. The analysis shows that the frequency figures for multi-word verbs in the ICLE are remarkably similar to those in the BNC (Appendix 1). Seventeen of the 26 multi-word verbs occur with a frequency within five occurrences per million in the two corpora. Only three multi-word verbs occurred with frequencies markedly different in the two corpora, i.e., with a difference greater than 10

occurrences per million and two of these were more frequent in the ICLE (*go on, bring up*). This means that the learner essays contained the 26 multi-word verbs at a similar rate to the native texts in the BNC. However, the learners used the one-word verbs at a far greater rate than the natives: 15 of the 26 verbs were markedly more frequent in the ICLE. Overall, we find that although the non-natives used the one-word verbs more in their writing than the natives, this did *not* mean that they used multiword verbs less.

In terms of the choice between multi-word verbs and their one-word verb alternatives, natives and learners used similar proportions (within 10 percentage points) with half of the 26 pairs (Table 2). For 10 pairs, the natives used markedly more multi-word verbs than the learners, and conversely, the learners used more of the one-word alternatives for these pairs. The learners used markedly more of the multi-word verb alternative than the natives in only 3 pairs (*get back/return; come up with/suggest; brush up on/revise*). This shows that the learners are making choices between the verb alternatives in similar proportions to the natives in about half of the cases; but in the remainder, they appear much more likely to choose the one-word verb alternative (10/13, 77%). It was relatively uncommon (3/26, 12%) for the learners to use the multi-word alternative more than the natives.

3.2.2. *Corpus analysis-spoken discourse.* We ran a similar analysis to the one above for spoken discourse using the CANCODE data. Whereas most (23/26) of the multi-word verb frequencies were similar (within 10 occurrences per million) between the BNC and the ICLE, only a little over half were in the CANCODE/ICLE comparison (14/26). For 11 of the 26 multi-word verbs, the ICLE learner essays had a markedly lower frequency than the CANCODE. For the one-word verbs, the distribution was more even: 8 verbs had similar frequencies between the corpora, 11 had markedly higher frequencies in the CANCODE, and 7 had markedly higher frequencies in the ICLE. Thus in terms of raw frequency, the learners used fewer multi-word verbs than the natives, but used one-word verbs in a mixed manner compared to the natives: some verbs more frequently, some less frequently, and some with a similar frequency.

Looking at the choice proportions in Table 2, we find very similar results to those of the written analysis. The natives and learners used similar proportions with 13 of the 26 pairs; the natives used markedly more multi-word verbs than the learners with 11 pairs, and for only two pairs (*bring up/mention; brush up on/revise*) did they use fewer. This illustrates the same tendency for learners to choose the one-word verb alternative.

However, we must be careful not to over-interpret the spoken results. Although the written analysis compares like-with-like (learner essays vs. a native written corpus), the spoken analysis compares the written essays in the ICLE with the spontaneous spoken conversations in the CANCODE. A closer com-

Table 3. Differences in the use of verbs for native and nonnative speakers

		Median <sup>a</sup>
one-word verbs	Native speakers <sup>b</sup>	3.36
	Advanced learners <sup>c</sup>	4.21
multi-word verbs	Native speakers <sup>b</sup>	5.15
	Advanced learners <sup>c</sup>	4.47

a 1 = very unlikely to use the type of verb, 6 = very likely to use the type of verb

b  $N = 65$

c  $N = 65$

parison would be to have natives and nonnatives choose from verb pairs embedded in the exact same context (same mode, same passage). This was done in the following questionnaire.

3.2.3. *Questionnaire.* One main goal of this study is to discover whether advanced learners of English use multi-word verbs in a way similar to native speakers. Because multi-word verbs are considered to be a particular feature of informal spoken discourse, we decided to directly ask natives and nonnatives about their use of the 26 verb pairs in colloquial contexts through the use of a questionnaire. Both groups were asked to judge how likely they were to use multi-word verbs and their one-word counterparts in a contextualized situation, ranging from “very unlikely” to “very likely”. The results are shown in Table 3.

The figures show that native speakers were much more inclined to use the multi-word member of the verb pair (5.15) than the one-word alternative (3.36). This result is statistically significant as shown by a Wilcoxon signed ranks test ( $T = 0$ ,  $p < .001$ ). This supports the notion that natives prefer to make use of multi-word verbs, especially in spoken, informal colloquial contexts, such as the questionnaire sought to emulate. However, somewhat surprisingly, we find that the advanced learners also had a slight, but statistically significant, preference for multi-word verbs (4.47) over one-word verbs (4.21) ( $T = 499.5$ ,  $p < .01$ ).

When we compare native usage against nonnative usage, we can see that the advanced learners were *less* likely to use the multi-word verbs (4.47) than the native speakers (5.15), with a Mann–Whitney test confirming that the result is statistically significant ( $U = 991$ ,  $p < .001$ ). We also find that the learners (4.21) were *more* likely to use the one-word verbs than the natives (3.36), a result that is also statistically significant ( $U = 854$ ,  $p < .001$ ).

The median scores show that for the one-word verbs, the native respondents were relatively neutral (3.36 on a 6-point scale), not strongly inclined or dis-

inclined to use these verbs. However, the natives were strongly inclined to use the multi-word verbs (5.15). The learners, on the other hand, reported a slight inclination to use both multi-word (4.47) and one-word verb forms (4.21).

Overall, the results illustrate a strong preference by the native speakers for using multi-word verbs over one-word verbs, while the advanced learners were more likely to use one-word verbs than the natives. However, the learners also showed willingness to use multi-word verbs; it is just that they do not have as strong a preference for them as the natives.

### 3.3. *How does exposure to an L2 environment affect the likelihood of using multi-word vs. one-word verbs?*

The above results show that advanced English learners are less inclined to use multi-word verbs than native speakers, but what can be done to help them become more confident in using them? It has been suggested that phrasal language requires a great deal of exposure to acquire (Adolphs and Durow 2004; Schmitt 2004), so perhaps extended stays in a native English-speaking environment is required. When administering the questionnaire, we collected information on the learners' backgrounds, and are thus able to differentiate between the amount of time they spent in an English-speaking context. This allowed us to investigate if exposure to English in a natural English-speaking environment and its duration had any impact on the learners' likelihood of using multi-word and one-word verbs.

We performed a correlation analysis between the number of months spent in an L2 country and learners' one-word and multi-word verb preferences (subjects with no L2 country experience were excluded from the analysis). A Spearman's test revealed a modest degree of correlation between the number of months spent abroad and one-word scores ( $r_s = -.313$ ,  $p$  (one-tailed)  $< .05$ ). Learners' multi-word scores, on the other hand, revealed no significant correlation with the exposure duration ( $r_s = -.042$ ,  $p$  (one-tailed)  $> .05$ ).

To further explore this issue, we split the learners into groups according to the time spent in an English-speaking country:

- zero months – 17 participants
- up to and including 12 months – 24 participants
- more than 12 months – 24 participants

The questionnaire results for these three groups are given in Table 4.

A Kruskal–Wallis test ( $H(2) = 13.7$ ,  $p < .01$ ) with a Mann–Whitney post hoc test showed that those L2 learners who spent the longest period of time in an English-speaking country, i.e., over 12 months, had lower one-word verb scores than the other two groups. There was no statistically significant difference between having no exposure and having 12 months or less. This indicates that a stay in an English-speaking country for more than 12 months can lead to

Table 4. Differences in the use of verbs for advanced learners ( $N = 65$ ) according to exposure to L2 environment

		Median <sup>a</sup>
one-word verbs	No exposure	4.52
	Up to 12 months	4.31
	More than 12 months	3.71
multi-word verbs	No exposure	4.31
	Up to 12 months	4.63
	More than 12 months	4.60

a 1 = very unlikely to use the type of verb, 6 = very likely to use the type of verb

nonnatives being less likely to use one-word verbs compared to nonnatives who have 12 months or less of this kind of exposure. This suggests the interesting possibility that there is a threshold of input-rich experience which is necessary to begin relying less on one-word verbs. This result is particularly striking in that the subjects in the “no experience” and “less than 12 month” groups are mostly professionals who use English on a daily basis, and who have good exposure to English, but just not in a native English-speaking environment.

Surprisingly, a Kruskal–Wallis test with Mann–Whitney post hoc test showed *no* significant differences in likelihood of multi-word verb usage among the three groups ( $H(2) = 1.4$ ,  $p > .05$ ). Thus while a lengthy stay in an English-speaking environment decreased the likelihood of using one-word verbs, it did *not* increase the likelihood of using the corresponding multi-word verbs.

Overall, it seems that long term exposure to a natural L2 environment does have some effect on the selection of multi-word vs. one-word verbs. There is a modest correlation between longer stays abroad and a lower preference for one-word verbs. Moreover, among the three nonnative groups, long term exposure to a natural L2 environment led to a significant difference between the “over 12 month” group and the other two in their preference for one-word verbs. But even spending more than a year abroad was not enough to bring the nonnatives down to a native-like level; the one-word verb scores of those long-stay learners (3.71) were still higher than the scores of native speakers (3.36) (Mann–Whitney test,  $U = 535.5$ ,  $p < 0.5$ ). However, the effect of the duration of exposure was not straightforward, as it had *no* significant effect on the likelihood of learner multi-word verb production, either in the correlation or time-abroad analyses.

#### **4. Discussion**

Previous corpus analyses have found that phrasal elements in English are common (Biber et al. 1999; Nattinger and DeCarrico 1992), and our investigation confirmed this, as one-third to one-half of the 26 multi-word verbs we studied occurred more than 10 times per million running words, depending on the corpus. However, this does not mean that all or even most multi-word verbs are frequent. Indeed, half of the 26 multi-word verbs we studied occurred rarely or not at all in the CANCODE, while this was the case for 18 out of the 26 in the BNC. The point of the low frequency is even clearer when we compare multi- and one-word verb alternatives. As expected, we found that in almost 70% of our verb pairs, the one-word verbs were more frequent than their multi-word alternatives in written discourse. However, we also found this true for almost the same percentage (65%) in spoken discourse. Furthermore, when we analyzed the total occurrences of the combined verb pairs (i.e., MWV + OWV in Table 2), the one-word verb was used proportionally more than the multi-word verb in both the CANCODE and BNC in 18 out of the 26 pairs (69%). Overall, although many multi-word verbs are reasonably frequent, the BNC and CANCODE data suggests that natives still prefer to use one-word verbs over multi-word verbs in the majority of cases.

This finding must be considered against the questionnaire results. Here the natives generally had a strong preference to use multi-word verbs, at least in the contexts as given in the questionnaire. Previous research shows that phrasal vocabulary is a prominent feature of informal spoken language, and perhaps the questionnaire's strong emphasis on this type of language led to natives' robust preference for multi-word verbs, even though the corpus data indicates an overall preference for one-word verbs. Our conclusion is that in spoken discourse natives do prefer multi-word over one-word verbs in many cases, but we must not assume that this is the trend in general.

The results for learners are more consistent. On a positive note, the frequencies of multi-word verbs are quite similar in the BNC and ICLE, indicating that learners are using these verbs to a similar degree as natives. Also, when comparing the proportions of usage, the nonnatives had similar figures to the natives for half of the verb pairs. However, when there was a discrepancy between the nonnative and native results, the learners had a strong tendency to use the one-word verb. This corpus result was supported by the questionnaire results, where the learners had a much higher median score than the natives (NNS: 4.21 vs. NS: 3.36). Conversely, they had a much lower score for the multi-word verbs (NNS: 4.47 vs. NS: 5.15). Overall, it does seem that learners have a tendency to choose the one-word verb option.

In terms of semantic and grammatical properties, this may not be a problem, as one-word verbs work perfectly well in this regard. But multi-word verbs

tend to have a colloquial tone which makes them particularly appropriate for informal speech, and not using them can make nonnatives sound stilted and unnatural in speech. The result that nonnatives are less likely to use multi-word verbs than natives in the questionnaire contexts suggests that our nonnative respondents will probably not sound as idiomatic in their speech as native speakers.

This is rather surprising as our nonnative respondents were high-level professional users of English. One would not be surprised if beginners were less likely to use multi-word verbs than native speakers; in fact we would expect this because of the difficulties in learning them outlined at the beginning of this article. But we may have expected advanced learners who use English every day in their work or studies to use the target verb pairs with something approaching native-like usage levels. This was not the case, and only reinforces the notion that multi-word verbs (and phrasal language in general) are a tricky element of English, which even highly advanced learners of English may struggle to utilize in a manner congruent to native speakers. This point is further driven home by the result that length of time in a native English-speaking environment had no discernible effect on the likelihood of using multi-word verbs. So, even over a year's worth of native-environment exposure was insufficient to increase the likelihood of using the multi-word verbs in this study.

It is hard to say why longer durations of exposure to native-speaking environments did not result in significant differences in the learners' willingness to use multi-word verbs. We can speculate that the complexity of multi-word verbs discussed in the introduction means that learners require an extremely long period of time to become comfortable with them, and the "more than 12 month" participants were not exposed long enough to show an effect. Interestingly, Liao and Fukuya (2004) found that the amount of contact with the L2 might be a contributing factor to learners' development from avoidance to non-avoidance. However, in their study, nonnative participants were a group of intermediate EFL student with no L2 exposure, and a group of advanced ESL students with at least nine months of natural L2 exposure. Thus the two groups varied considerably in both proficiency level *and* L2 exposure. Participants in our study, on the other hand, varied in terms of their L2 exposure, but all seemed to be relatively proficient in their use of English.

Alternatively, there may also be a cross-linguistic reason. Learners with non-Germanic L1s may take a long time to overcome their discomfort with *alien* multi-word verbs. As one Spanish-speaking participant reported, although he had been living in the UK and USA for over eight years, he still felt more comfortable using one-word verbs "of Latin origin".

Laufer (1997), Moon (1997) and Wray (2000) maintain that both teachers and learners find idiomatic multi-word units more difficult than their non-idiomatic counterparts, which is likely to lead to avoidance behavior. Multi-

word verbs fall in this category, so we might expect avoidance for them as well. Avoidance of certain linguistic structures presupposes that these structures are known to L2 learners, although they may not necessarily use them freely (Kleinmann 1977). In addition, we can only speak of avoidance if there is evidence that native speakers would use the form that we claim is being avoided by L2 learners in the context under consideration (Seliger 1989, cited in Ellis 1994). We did not include a direct measure of the knowledge of the verb pairs in our study, so we cannot address this issue directly. Nevertheless, based on the nonnative participants' advanced proficiency and daily use of English, we can speculate that they knew all/most of the multi-word verbs, if not productively, then receptively. To the extent this is true, we find evidence that could be interpreted as avoidance behavior, as the figures for nonnative likelihood of use were much lower than the native participants', who we assume knew the multi-word verbs in the study.

These results would support those of Dagut and Laufer (1985) who found evidence of avoidance, but differ from the results obtained by Liao and Fukuya (2004). Liao and Fukuya (2004: 210) found that although their advanced Chinese learners showed "a slight tendency" to use fewer phrasal verbs than the native participants, this, nevertheless, was not at the statistically significant level, which meant that no avoidance behavior could be reported. Similar to Liao and Fukuya (2004), Hulstijn and Marchena (1989) did not find categorical avoidance behavior in their Dutch participants. However, the authors report "a tendency to adopt a play-it-safe strategy" (Hulstijn and Marchena 1989: 250), i.e., to prefer one-word verbs with general meaning to phrasal verbs with idiomatic meaning. Overall, on the basis of previous research and the current study, there seems to be converging evidence for the notion that learners tend to avoid multi-word verbs in spoken colloquial, but perhaps not in written, contexts.

However, the corpus data in this study suggests another reason why learners tend to use fewer multi-word verbs. In both the BNC and CANCODE corpora, the one-word verbs were more frequent than their multi-word counterparts. It may be that learners use relatively fewer multi-word verbs simply because this reflects the input they are receiving, rather than because they are actively avoiding them. Undoubtedly, both input and avoidance have a role to play, and this study illustrates the value of using corpus analysis as part of the methodology in studying avoidance.

## **5. Limitations**

The main limitation of the study concerns the corpus data. Many of the verbs in our study, both multi-word and one-word, are polysemous. The corpora we used are not tagged for semantic content, so the concordancer was not able to

give frequency figures for each of the meaning senses. Unfortunately, we did not have the resources to go through the time-consuming process of checking each concordance line to tabulate individual meaning senses. This means that figures reported in Appendix 1 are not a one-to-one comparison of meaning senses. For example, the figures for *turn down* include the meaning senses *turn down the heat* and *turn down an invitation*, while the figures for *decrease* would only include the former sense. There are also some misleading quirks, e.g., one of the topics for the learner essays in the ICLE was joining the European Union, which contributed to a relatively high occurrence of the verb *join*. Thus the corpus figures should be taken as estimates rather than precise figures. Nevertheless, the same procedures were used for all corpora analyses, so the results should still be comparable and informative in showing the general trend of usage of multi- and one-word verbs across native spoken, native written, and nonnative written contexts.

*University of Nottingham*  
<aexas7@nottingham.ac.uk>  
<norbert.schmitt@nottingham.ac.uk>

### Notes

1. As the purpose of the study is to explore the choice made when multi- and one-word verb alternatives are available, the verb pairs needed to be synonymous to the extent that one could replace the other within the contextualized prompts of the questionnaire. However, they are unlikely to be identical semantic alternatives, especially as many of the multi-word verbs are polysemous, e.g., *work out* means “train”, but also has other meanings, such as “have a good result”: *I hope your new job works out*.
2. CANCODE stands for Cambridge and Nottingham Corpus of Discourse in English. It was developed as a joint project between the University of Nottingham and Cambridge University Press, with whom sole copyright resides.

**Appendix 1. Frequencies of multi-word verbs and one-word verbs**

Multi-word verbs	One-word verbs	Used	CANCODE frequency <sup>a</sup>	BNC frequency <sup>a</sup>	ICLE frequency <sup>a</sup>
put off	postpone	Yes	6.0: 0.0	5.9: 4.3	2.6: 5.4
tell off	reproach	Yes	0.0: 0.4	0.0: 0.4	0.0: 1.5
put up with	stand	Yes	22.0: 168.3	7.4: 71.4	10.0: 124.3
turn down	decrease	Yes	3.6: 2.8	2.0: 3.7	1.2: 22.0
run into	meet	Yes	4.4: 169.5	6.2: 113.9	3.5: 146.0
come round	come	Yes	75.8: 2435.2	3.8: 540.5	0.0: 559.1
tidy up	organize	Yes	6.0: 29.6	1.5: 7.2	1.2: 35.0
go on	continue	Yes	311.7: 51.6	68.9: 104.5	95.8: 132.1
pull over	stop	Yes	5.0: 365.4	0.4: 107.0	0.0: 218.0
work out	train	Yes	90.4: 35.2	21.9: 10.6	15.1: 12.0
set up	start	Yes	88.6: 676.5	102.2: 120.3	27.8: 293.6
get back	return	Yes	119.2: 145.8	23.9: 70.1	31.3: 52.0
figure out	understand	Yes	6.4: 381.8	7.8: 129.3	5.8: 382.6
walk off	leave	Yes	1.6: 470.6	0.7: 151.1	0.0: 172.2
mess around	misbehave	Yes	4.0: 1.2	0.5: 0.2	0.4: 0.8
come up	arise	Yes	148.8: 7.2	29.0: 19.6	34.4: 36.3
come up with	suggest	Yes	34.6: 65.6	16.3: 66.6	24.7: 46.7
hold on	wait	Yes	24.2: 269.0	15.3: 67.3	11.2: 58.5
show off	boast	Yes	2.8: 0.4	3.4: 2.7	5.4: 7.0
bring up	mention	No	10.8: 88.5	4.0: 25.6	28.2: 92.3
come along	join	No	31.8: 68.7	6.3: 58.5	1.5: 81.1
come across as	seem	No	41.4: 320.0	0.5: 140.1	0.4: 396.5
go up	rise	No	141.4: 10.3	13.8: 30.7	4.6: 20.6
catch up with	join	No	2.4: 68.7	3.1: 58.5	2.7: 81.1
call off	cancel	No	0.0: 30.6	0.9: 7.6	0.0: 4.3
brush up on	revise	No	0.4: 7.6	0.1: 3.1	0.4: 2.0

a multi-word verb frequency : one word verb frequency (in occurrences per million words).

**Appendix 2. Verb pair questionnaire**

(Note that each question used the 6-point scale. The scale is shown here only once due to space constraints. The 7 verb pairs deleted from the final questionnaire are marked with an asterisk (\*).)

The test is completely anonymous. *Please note that there are NO wrong answers, we are NOT testing your language knowledge.* The aim of the study is to investigate the psychological processes underlying the choice of certain words in colloquial English. Please think for a second and then select those







- Darwin, Clayton M. and Loretta S. Gray (1999). Going after phrasal verbs: An alternative approach to classification. *TESOL Quarterly* 33 (1): 65–83.
- de Cock, Sylvie (2000). Repetitive phrasal chunkiness and advanced EFL speech and writing. In *Corpus Linguistics and Linguistic Theory*, Christian Mair and Marianne Hundt (eds.), 51–68. Amsterdam: Rodopi.
- Dörnyei, Zoltán (2003). *Questionnaires in Second Language Research*. Mahwah, NJ: Lawrence Erlbaum Associates.
- Freeborn, Dennis (1995). *A Course Book in English Grammar*. London: Macmillan.
- Granger, Sylviane (1998). Prefabricated patterns in advanced EFL writing: collocations and formulae. In *Phraseology: Theory, Analysis, and Applications*, Anthony P. Cowie (ed.) 79–100. Oxford: Oxford University Press.
- Granger, Sylviane, Estelle Dagneaux, and Fanny Meunier (eds.) (2002). *International Corpus of Learner English*. Louvain: UCL Presses Universitaires de Louvain.
- Hulstijn, Jan and Elaine Marchena (1989). Avoidance: Grammatical or semantic causes? *Studies in Second Language Acquisition* 11 (3): 241–255.
- Kleinmann, Howard (1977). Avoidance behavior in adult second language acquisition. *Language Learning* 27 (1): 93–107.
- Laufer, Batia (1997). What's in a word that makes it hard or easy: Some intralexical factors that affect the learning of words. In *Vocabulary: Description, Acquisition and Pedagogy*, Norbert Schmitt and Michael McCarthy (eds.), 140–155. Cambridge: Cambridge University Press.
- (2000). Avoidance of idioms in a second language: The effect of L1–L2 degree of similarity. *Studia Linguistica* 54 (2): 186–196.
- Laufer, Batia and Stig Eliasson (1993). What causes avoidance in L2 learning: L1–L2 difference, L1–L2 similarity, or L2 complexity? *Studies in Second Language Acquisition* 15 (1): 35–48.
- Liao, Yan D. and Yoshinori J. Fukuya (2004). Avoidance of phrasal verbs: The case of Chinese learners of English. *Language Learning* 54 (2): 193–226.
- Moon, Rosamund (1997). Vocabulary connections: Multi-word items in English. In *Vocabulary: Description, Acquisition and Pedagogy*, Norbert Schmitt and Michael McCarthy (eds.), 140–155. Cambridge: Cambridge University Press.
- (1998). Frequencies and forms of phrasal lexemes in English. In *Phraseology: Theory, Analysis, and Applications*, Anthony P. Cowie (ed.), 79–100. Oxford: Oxford University Press.
- Nattinger, James R. and Jeanette S. DeCarrico (1992). *Lexical Phrases and Language Teaching*. Oxford: Oxford University Press.
- Schachter, Jacquelyn (1974). An error in error analysis. *Language Learning* 24 (2): 205–214.
- Schmitt, Norbert (ed.) (2004). *Formulaic Sequences*. Amsterdam: John Benjamins.
- Seliger, Herbert (1989). Semantic transfer constraints in foreign language speakers' reactions to acceptability. In *Transfer in Language Production*, Hans Dechert and Manfred Raupach (eds.), Norwood, N.J.: Ablex. Cited in Ellis, Rod (1994). *The Study of Second Language Acquisition*. Oxford: Oxford University Press.
- Wray, Alison (2000). Formulaic sequences in second language teaching: Principle and practice. *Applied Linguistics* 21 (4): 463–489.
- Wray, Alison and Michael R. Perkins (2000). The function of formulaic language: An integrated model. *Language and Communication* 20 (1): 1–28.