

*Re-telling a story in a second language:  
How well do adult learners mine an input text  
for multiword expressions?*

Hien Hoang

Quy Nhon University, Vietnam

[hoangthithuhien@qnu.edu.vn](mailto:hoangthithuhien@qnu.edu.vn)

Frank Boers

Victoria University of Wellington, New Zealand

[frank.boers@vuw.ac.nz](mailto:frank.boers@vuw.ac.nz)

Abstract

Adult second language (L2) learners have often been found to produce discourse that manifests limited and non-native-like use of multiword expressions. One explanation for this is that adult L2 learners are relatively unsuccessful (in the absence of pedagogic intervention) at transferring multiword expressions from input texts to their own output resources. The present article reports an exploratory study where ESL learners were asked to re-tell a short story which they had read and listened to twice. The learners' re-tells were subsequently examined for the extent to which they recycled multiword expressions from the original story. To gauge the influence of the input text on these learners' renderings of the story, a control group was asked to tell the story based exclusively on a series of pictures. The results of the experiment suggest that multiword expressions were recycled from the input text to some extent, but this stayed very marginal in real terms, especially in comparison with the recycling of single words. Moreover, when learners did borrow expressions from the input text, their reproductions were often non-target-like.

*Keywords:* vocabulary; phraseology; formulaic sequences; input; story re-telling

## 1. Introduction

Multiword expressions have attracted a considerable amount of interest in applied linguistics circles in recent years, as evidenced, for example, by the steady stream of monographs and edited volumes devoted to the topic (e.g., Boers & Lindstromberg, 2009; Barfield & Gyllstad, 2009; Lewis, 1997, 2000; Meunier & Granger, 2008; Nattinger & DeCarrico, 1992; Polio, 2012; Schmitt, 2004; Wood, 2010a, 2010b, 2015; Wray, 2002). It is now generally recognized that a large proportion of natural discourse consists of (semi-)fixed word strings and combinations that are predictable by virtue of the frequent co-occurrence of their constituent words (Erman & Warren, 2000; Howarth, 1998; Sinclair, 1991), and that mastery of these sequences is one of the hallmarks of native-speaker competence (Pawley & Syder, 1983).

Multiword expressions come in many shapes and perform a multitude of functions. They include conversational routine formulas (e.g., *How are you doing?*), discourse organisers (e.g., *Having said that*), idioms (e.g., *at the end of the day*), proverbs (e.g., *When the cat's away*), standardized similes (*blind as a bat*) and binomials (*rough and tumble*), phrasal verbs (e.g., *sleep in*), prepositional phrases (e.g., *by car*), compounds (e.g., *baby boom*), and collocations (e.g., *bright sunshine; make an effort*). Some are uninterrupted strings, while others provide sentence frames with slots to be completed (e.g., *Not only . . . , but . . . ; I was wondering if . . .*). Perhaps due to the great diversity of multiword items, both in form and function, many different labels have been used in the literature to refer to multiword lexis in general and to certain categories of expressions in particular (e.g., *lexical phrases, phrasal expressions, multiword units, prefabricated chunks*). At the time of writing her influential book on the topic, Wray (2002) had found over 50 terms to describe instances of multiword lexis. Her own, now much adopted term, was *formulaic sequence*, which she defined as “a sequence, continuous or discontinuous, of words or other elements, which is, or appears to be, prefabricated: that is, stored and retrieved whole from memory at the time of use, rather than being subject to generation or analysis by the language grammar” (p. 9). While this is a comprehensive definition of the kinds of patterns we wish to consider in the present article, the definitional criterion of “holistic” retrieval makes it more suitable when one is dealing with native speakers than for studies such as ours, where L2 learners’ use of multiword lexis is examined. After all, L2 learners’ command of target expressions may not yet be sufficiently proceduralized to warrant the kind of hesitant-free production that suggests the expressions are retrieved from the mental lexicon holistically, as prefabricated units. To avoid this implication, we have therefore opted to use the term *multiword expression* (henceforth MWE) rather than *formulaic sequence* in the present article.

A growing body of experimental research shows that native speakers' knowledge of MWEs facilitates their fluent language processing and that this is a processing advantage many L2 learners do not enjoy (see Siyanova-Chanturia & Martinez, 2015, for a recent review; Sonbul, 2015, for a recent example). There is also evidence that L2 learners who display mastery of MWEs tend to be (perceived as) more proficient or advanced than those whose output exhibits little evidence of such phraseological competence (Bestgen & Granger, 2014; Boers, Eyckmans, Kappel, Stengers, & Demecheleer, 2006; Crossley, Salsbury, & McNamara, 2015; Dai & Ding, 2010; Stengers, Boers, Housen, & Eyckmans, 2011).

Unfortunately, many studies have also shown that even advanced learners' L2 output will typically lack the phraseological richness manifested in native speaker discourse and will usually exhibit phraseological oddities (e.g., Altenberg & Granger, 2001; Laufer & Waldman, 2011; Levitzky-Aviad & Laufer, 2013; Li & Schmitt, 2010; Serrano, Stengers, & Housen, 2015; Siyanova & Schmitt, 2008). And yet, these attested shortcomings often concern relatively common MWEs. Part of the problem, then, seems to lie in learners' failure to pick up the MWEs they encounter and add these to their own repertoire. Several explanations have been proposed that may jointly account for this lack of transfer of MWEs from input to output.

One is that (literate) adult learners are relatively inattentive to phraseological patterns in L2 text, because—unlike children acquiring their mother tongue through aural input—they have become used to treating the word (rather than larger chunks) as the basic unit of meaning (Wray, 2002). Most applied linguists now concur that attention is crucial for intake of information (Schmidt, 2001). It is also well recognized that language users are naturally inclined to process messages first and foremost for their content, not the linguistic packaging of that content (e.g., Sharwood-Smith, 1993; Van Patten, 2002). If, for adult learners, it is single words that provide the key to text content, then it follows that words rather than larger syntagmatic units will be attended to, reducing the likelihood that the latter will leave a durable imprint in memory.

What is more, even when MWEs do receive attention, learners may fail to reproduce them accurately or simply shy away from trying to reproduce them, because MWEs are often longer than single words and thus more challenging to recall (Ellis, 1996; Skrzypek & Singleton, 2013). Often single-word substitutes are available (e.g., *they were furious* vs. *they were up in arms*; *we'll have to accept it* vs. *we'll have to put up with it*; *don't lie* vs. *don't tell lies*) and the learner may consider these easier (and "safer") to use. This may be an especially appealing avoidance strategy in case one is uncertain about the meaning of particular MWEs (e.g., idioms; Laufer, 2000) and in case a given category of L2 MWEs (e.g., phrasal verbs) is absent from the learners' L1 (Dagut & Laufer, 1985; Siyanova &

Schmitt, 2007). But interference from the learners' L1 also occurs when members of a shared category of MWEs (e.g., collocations) are not congruent (Nesselhauf, 2003; Yamashita & Jiang, 2010; Wolter & Gyllstad, 2011).

Given these hindrances to learners' autonomous recycling of MWEs from textual input, one might argue that, instead of relying on incidental uptake, time needs to be invested in the explicit teaching and deliberate study of MWEs (for example by using materials such as Lindstromberg & Boers, 2008; Davis & Kryszewska, 2012; and McCarthy & O'Dell, 2005). On the other hand, there is some evidence that learners *do* spontaneously recycle lexis from instructions and exemplars they are given as prompts for output activities, and that some of that recycling *does* concern lexical items larger than single words (Boston, 2008). In light of that evidence, the above pessimism about learners' spontaneous uptake of MWEs from input texts is perhaps not entirely justified. Perhaps learners do pick up a fair number of MWEs from input texts that they incorporate in their subsequent output, provided the conditions for this to happen are favourable. It is this possibility that is explored in the present article.

## 2. "Mining" input for multiword expressions to fuel an output task

A number of studies have compared L2 learners' rate of acquisition of MWEs under various reading conditions (Pellicer-Sánchez, 2015; Webb, Newton, & Chang, 2013), including reading where target MWEs are typographically enhanced (Boers, Demecheleer, He, Deconinck, Stengers, & Eyckmans, 2016; Sonbul & Schmitt, 2013; Szudarski & Carter, 2014) and/or glossed (Bishop, 2004; Peters, 2009, 2012). Separate to this strand of research into reading-based MWE acquisition, some studies have examined factors that influence the effectiveness of deliberate MWE-focused instruction and learning (e.g., Alali & Schmitt, 2012; Boers, Eyckmans, & Stengers, 2007; Boers, Demecheleer, Coxhead, & Webb, 2014; Boers, Dang, & Strong, 2016; Eyckmans, Boers, & Lindstromberg, 2016; Laufer, 2010; Laufer & Girsai, 2008; Peters, 2016; Szudarski & Conklin, 2014; Webb & Kagimoto, 2011).

What all of these aforementioned studies have in common is that they assess the effects of interventions by means of controlled, discrete-item knowledge tests. To date, much less research has been conducted that addresses the question of how well learners incorporate the MWEs they have been exposed to or have studied in their own meaning-focused communicative output, and what steps can be taken to promote this transfer from input to output. Like Boston (2008, 2010), we borrow the term *mining of input* from Samuda (2001), who used it to refer to learners' use of language elements borrowed from prompts, even though the learners were not explicitly instructed to do so. Focusing specifically on vocabulary, Boston (2008) describes how his Japanese

EFL learners spontaneously recycled items from instructions and examples for output activities. If learners are inclined to mine input materials for language that they feel will help them perform the output task, he argues, then it may be possible to harness this inclination by designing the input materials in ways that promote successful mining. Successful mining is likely to depend on many factors, however, including the kind of language features concerned (e.g., content words are likely to be recycled more than function words and grammar features), the relevance of the input language for the output task (i.e., whether particular elements are task-essential or not), and the way the input language is presented (e.g., written or aural; with or without features that make particular elements perceptually salient).

Several of the studies that address the question of whether adult L2 learners successfully incorporate MWEs from input materials into their own communicative output have been conducted in the context of writing courses. Jones and Haywood (2004) report a study where learners in an EAP course were regularly engaged in activities with a focus on MWEs. While the authors found clear evidence that these students' awareness of MWEs increased, their actual use of MWEs in their end-of-course essays did not differ markedly from that of a comparison group that had not received the MWE-focused treatment. In a partial replication of Jones and Haywood (2004), Peters and Pauwels (2015) also examined the effect of integrating various MWE-focused activities in an EAP course. They found evidence of an effect, but this evidence emerged more clearly in a discrete-item recognition test than in the learners' spontaneous use of the MWEs in their writing assignments. This illustrates that, despite increasing familiarity with MWEs at a receptive level, learners may not adequately deploy this knowledge in communicative output tasks.

The effects of MWE-focused interventions on learners' *oral* output have been investigated by Wood (2009, 2010b), Boers et al. (2006) and Stengers, Boers, Housen and Eyckmans (2010). Wood's investigations focus on the advantages afforded by formulaic sequences specifically for speech fluency. He describes courses where ESL learners took part in intensive fluency workshops involving focused instruction on and sustained practice of formulaic sequences. MWEs were extracted by the instructor from listening input and the learners practised reproducing these sequences in a series of output activities. Detailed analyses of the learners' subsequent oral narratives showed that their speech fluency benefited from using the formulaic sequences they had practised. The intervention tried in Boers et al. (2006) and Stengers et al. (2010) was much less intensive. Students in an EFL course were regularly encouraged to identify MWEs in the authentic reading materials they worked with in class, an awareness-raising activity recommended by Lewis (1997). At the end of the course,

their performance on speaking tasks was compared with that of a comparison group whose attention had not been directed explicitly to the presence of MWEs in the texts. No compelling evidence emerged that the former group made more use *specifically* of the MWEs they had encountered in the course materials. On the positive side, however, if the speaking activity was prompted by a text which they had at their disposal during the speaking activity, these students did appear more inclined than the comparison group to replicate MWEs used in that text.

It is worth noting that these studies present no evidence that MWE learning does *not* occur in the absence of MWE-focused instruction. After all, where comparison groups were included, these were usually also found to make progress, sometimes enough progress to render the differences between MWE-instructed and comparison groups non-significant, thereby casting doubt on the impact that focused MWE instruction has over and above learners' spontaneous mining of input texts. It is also worth noting that, in the aforementioned intervention studies, the learner-participants rarely engaged in communicative output activities making use of the content of (and thus potentially also the wordings used in) the textual input very shortly after having processed that textual input. A study by Lindstromberg, Eyckmans and Connabeer (2016) investigates the potential of text reconstruction activities such as dictogloss for MWE uptake, but the extent to which learners will mine input texts for MWEs spontaneously to help them fulfil a subsequent communicative output task is still insufficiently documented.

In the exploratory study we report below, we gauge the extent to which adult ESL learners spontaneously recycle MWEs from a short story they are asked to retell.

### 3. Method

#### 3.1. Participants

Participants in the study were 34 volunteers who were all international students (aged 19 to 38) in an English proficiency program at a university in New Zealand. They were informed they would be participating in a study on the benefits of story-retelling activities. They came from eight different countries: China (14), Vietnam (6), Japan (4), Thailand (3), Indonesia (2), East Timor (2), Brazil (2), and Germany (1). Their (self-reported) scores for speaking skills on their most recent IELTS test were 5 or above. They all took an in-house placement test at the beginning of their English courses, three weeks prior to the experiment. This placement test consisted of a dictation, C-tests, and a vocabulary test. Taking account of their scores on this placement test, the participants were quasi-randomly assigned

to one of two conditions. In one condition (henceforth *the experimental condition*) the learners read and listened to a short story twice (see Section 3.2 for details) and then re-told the story. In the other condition (henceforth *the control condition*) the participants were presented with the story exclusively in pictorial form before being asked to tell the story. The latter condition served to generate “baseline” data, that is, information about the lexis that same-proficiency learners would resort to for the narrative task without being influenced by the input text. The data of three participants were discarded (for reasons explained further below), which left us with usable data for 17 participants in the experimental condition and for 14 in the control condition. According to their performance on the placement test, the two groups were similar in proficiency, with means of 173.35 ( $SD = 27.75$ ) and 168.07 ( $SD = 32.34$ ), respectively (maximum score = 282). An independent-samples t-test yielded  $t(29) = 0.49$ ;  $p = .63$ .

### 3.2. Materials and procedure

An adaptation of one of Aesop’s fables, *The Donkey and his Masters*, rich in MWEs, served as the input text for the re-tell task (see the appendix). The text consists of 466 words, and the story has six narrative turns: (a) the donkey working for the first master (a herb seller), (b) the donkey asking the god Jupiter for another master, (c) the donkey working for the second master (a brick-maker), (d) the donkey asking Jupiter again for another master, (e) the donkey working for the third master (a tanner), and (f) the donkey recalling Jupiter’s words of caution and learning a lesson.

A PowerPoint presentation was created with 20 slides that told the story in a time span of 4 minutes. For the experimental group the story was presented concurrently in three modes: (a) an audio recording of the story read aloud (relatively slowly, at 117 words per minute) by a female native speaker of English; (b) full captions of the audio recording, i.e., the written text for the participants to read while listening; and (c) pictures (one picture per slide) illustrating the content of the passage being read. For the control group, the same 20 slides with pictures were used, but without the audio recording and without the captions. This was preceded in both conditions by a slide with instructions (see below) and a slide with the title of the story, i.e., *The Donkey and his Masters*.

According to the VocabProfile tool at Tom Cobb’s <http://www.lex Tutor.ca>, only three words (barring the name *Jupiter*) in the story are beyond the 3,000 most frequent word families of English (in the combined Corpus of Contemporary American English [COCA] and British National Corpus [BNC]) and would be considered as mid-frequency rather than high-frequency vocabulary according to the criteria proposed by Schmitt and Schmitt (2015). These are *herbs*, *donkey*,

and *tanner*. The meaning of *donkey* was illustrated pictorially in the slides and the meaning of *tanner* was explained verbally in the text itself. We found no indications in the participants' renderings of the story that the input text posed any challenges for comprehension.

The text was analyzed for the presence of MWEs independently by two experienced English teachers. In cases of disagreement, we resorted to statistical information from COCA to make a decision. For collocations of two content words (e.g., *learn + lesson*), we used mutual information (MI) scores and set the threshold at  $> 4$ . For other word strings (e.g., *no longer*), we used corpus frequency and set the threshold at  $> 100$  occurrences. The combination of these procedures led to the identification of 35 MWEs in the text (see the appendix), henceforth referred to as *target MWEs*. They represent a range of phraseological patterns, including polywords (e.g., *let alone*), complex verbs (e.g., *look forward to . . .*), verb-noun collocations (e.g., *make + request*), and sentence frames with open slots to be completed (e.g., *be too . . . to . . .*).

As already mentioned, the two groups of students were exposed to the same story but in two different conditions: (a) a listening text accompanied by captions and pictures and (b) pictures only. Both groups were given the following instructions:

You are going to watch the story twice. The first time, just enjoy the story. The second time, feel free to take notes. Afterwards, you will be given eight of the pictures to help you remember the story line and three minutes to prepare an oral narrative. Your re-telling of the story should be about three minutes long. You will tell the story twice, for different listeners who have not watched the story themselves.

We believe the task conditions we created for the recall task to be relatively favourable for input mining to occur, for at least the following four reasons. First, reading-while-listening has been found to be more conducive to L2 vocabulary uptake than reading only (Webb & Chang, 2012). Moreover, the availability of the aural input may not only assist learners with the pronunciation of words they might otherwise shy away from using, but the prosodic cues may also assist learners with the chunking of speech into units larger than single words (e.g., Lin, 2012). At the same time, preserving the written mode alongside the aural mode may be helpful for note-taking, and, according to Boston (2008), facilitates mining more than aural input alone. Second, reading and listening to the story twice gives learners the opportunity to first become familiar with its content so that attentional resources can be freed up for taking in the language proper as they process the text a second time. Third, we asked the participants to perform the narrative task twice because that would give them the opportunity to try and modify language forms they may have felt dissatisfied with the first time they performed the task (e.g., Bygate, 2001; Wang, 2014). Evidence of

(accurate) uptake of a language item from the input text might therefore emerge in a second re-tell while it was absent in the learner's first attempt. Finally, presenting the text explicitly as the prompt for a re-tell activity is likely to stimulate engagement with language items in the text that learners anticipate using in the output activity. Moreover, the time given to prepare the narrative allows for a fair amount of (mental) text re-construction prior to the actual oral delivery.

### 3.3. Analysis

The participants' oral narratives were audio-recorded and transcribed. The transcriptions of raw speech were trimmed by removing false starts, repairs, repetitions and filled pauses. These trimmed transcripts were used for the below analyses.

Prior to any between-group comparisons we needed to ascertain whether the experimental and the control group produced narratives of comparable length and structure (i.e., containing six narrative turns, as described above). Recall that participants were instructed to aim at a narrative of about 3 minutes. One (Japanese) participant in the experimental group produced a narrative of only 1.29 minutes. Another (Japanese) participant in the experimental group did deliver narratives of close to 3 minutes but his speech rate was so slow that he produced far fewer words than his peers. This was also the only participant who failed to include six narrative turns in his rendering of the story. One (Brazilian) participant in the control group, by contrast, produced a much longer (and embellished) story than all other participants. We decided to remove these three outliers from the data so as to increase the comparability of the two samples. After removing the data of these three participants, the word counts of the two groups' narratives are very similar: Combining the word counts of the first and the second re-tells yields means of 498.59 ( $SD = 81.98$ ) for the experimental group and 496.79 ( $SD = 80.48$ ) for the control group. In terms of the amount of language produced, then, the two samples are very well matched:  $t(29) = 0.06$ ;  $p = .95$ .

Before turning to an evaluation of the learners' inclination and ability to recycle the phraseological dimension of the input text (operationalized here as the 35 target MWEs), it is worth establishing whether the input text influenced the participants' language at all. More particularly, the extent to which they recycled single words from the input text may help to put the amount of MWE mining into perspective. We resorted to the Text Lex Compare function on Tom Cobb's <http://www.lex tutor.ca> to calculate the proportion of word families shared between the re-tells and the original story, and compare this with the narratives produced by the control group. We complemented this with a count of shared word families from beyond the 2,000 most frequent word families (in the combined COCA and BNC corpus). The proportion of these *beyond-K2* words

is a measure of lexical sophistication, that is, of the extent to which learners have advanced beyond a “basic” vocabulary (Laufer, 1995).

We then proceeded by counting and assessing the target MWEs in the participants’ narratives. We also asked the same two teachers who identified MWEs in the input story to identify any other MWEs in the participants’ narratives. In cases of disagreement, we applied the same corpus-informed criteria as above (MI score > 4; corpus frequency > 100) to determine whether to consider word strings as MWEs.

The analyses reported below concern each participant’s two renderings of the story combined as one document. This was not the original plan, but it follows from the findings that the first and second re-tells were consistently very similar. Only very seldom did participants in the experimental group recycle a word from the input text in one but not the other rendering of the story. As a result, the number of recycled words in the repeated narratives was almost identical (paired-samples  $t(16) = -0.2$ ;  $p = .84$ ). Similarly, only very seldom did participants in the experimental group use MWEs from the input text in the second but not the first of their re-tells. In fact, the number of recycled MWEs tended to be *lower* in participants’ second re-tells, but this trend is not significant ( $t(16) = 0.4$ ;  $p = .69$ ).

## 4. Results and discussion

### 4.1. Recycling of words

The counts of word families shared between the participants’ narratives and the original story indicate that the availability of an input text exerts a strong influence on learners’ choice of words as they deliver their version of the story, as shown in Table 1. Of the word families that make up the narratives of the experimental group, on average 68.77% were shared with the input text. This compares to only 49.47% in the control condition. The effect of having versus not having processed the input text on the participants’ choice of words in their renderings of the story is clearly significant: independent-samples  $t$ -test:  $t(29) = 9.36$ ;  $p < .0001$ ; CI95% exp. [65.59, 71.95]; CI95% ctrl [47.16, 51.78].

Table 1 Descriptive statistics for single words

Group	<i>n</i>	% shared with input text		Number of beyond-K2 word families			
		<i>M</i>	<i>SD</i>	<i>M</i>	<i>SD</i>	<i>M</i>	<i>SD</i>
Experimental group	17	68.77	6.69	5.18	1.98	1.03	0.32
Control group	14	49.47	4.41	5.21	2.89	1.06	0.56

Complementary evidence that learners in the experimental group recycled words from the input text emerges when we focus specifically on the beyond-K2 words from the input text. For example, 13 of the 17 participants in the experimental group reproduced *tanner*, *herbs*, and *request*, three words which none of the participants in the control group used in their renderings of the story. The word *regret* was used by 11 participants in the experimental group but by only three in the control group. Most participants in the control group opted for other means to express the same notions, such as *ask for* instead of *request* and *feel sorry* or *be unhappy about* instead of *regret*.

The fact that the experimental group was strongly influenced in their choice of lexis by the text they had just read and listened to did not necessarily lead them to produce narratives that were lexically more sophisticated than those of their control peers, however. The two groups' narratives actually displayed similar mean numbers of beyond-K2 word families: 5.18 in the experimental condition and 5.21 in the control condition (see Table 1). Given that the total number of words produced by participants varied somewhat, it is perhaps more accurate to compute means per 100 words produced. Also according to this calculation the two groups' narratives contained almost equal proportions of beyond-K2 words: 1.03 and 1.06.

It may seem surprising that the opportunity to mine a text for language did not benefit the overall degree of lexical sophistication of the learners' output. After all, many did use a low-frequency word such as *tanner*, which they would not have used had they not encountered it (and its explanation) in the input text. However, by adhering to the words used in the original story, the learners in the experimental group also recycled many "basic" means of expression (e.g., *tired* and *unhappy*) where their peers in the control group used more "sophisticated" ones, such as *exhausted*, *miserable* and even *disillusioned*. Adherence to the words contained in the input text may thus have led the more advanced participants in the experimental group to produce output exhibiting a lexical profile that masked their real lexical competence. This possibility is lent credibility by the virtual absence of a correlation ( $r = .033$ ) between the experimental participants' scores on their placement test and the number of beyond-K2 words in their narratives. In the control group, however, the correlation between placement test scores and the number of beyond-K2 words used in the narratives is much stronger ( $r = .491$ ), reflecting the expected association between proficiency and lexical richness.

#### 4.2. Recycling of multiword expressions

Having established that mining for language definitely occurred, we can now turn to the data which will help us answer the question we set out to address,

that is, how well L2 learners mine an input text for its *phraseological* dimension, operationalized here as the quantity and quality of MWEs recycled from the text.

As summarized in Table 2, of the 35 target MWEs, the participants in the experimental group reproduced on average only 2.41 (median = 2) (type counts) accurately in their re-tells. The average number of accurate target MWEs per stretch of 100 words in the experimental group was just 0.48 (median = 0.49). The overall likelihood of any given MWE from the input text reappearing in the experimental group's re-tells was only 6.89%. This is obviously very limited use of MWEs from the input text, especially in comparison with the amount of recycling of single words we discussed above. The data nevertheless furnish evidence that learners mine an input text for MWEs to some degree, because the same set of 35 MWEs was virtually absent from the narratives of the control group: Participants in the control group produced on average 0.57 (median = 0.5) of the MWEs that occurred in the input text, or an average of 0.12 (median = 0.09) per 100 words. While this is a significantly smaller proportion than that found in the experimental group ( $z(29) = 2.52$ ;  $p = .01$ , for per-participant counts, and  $z(29) = 2.68$ ;  $p = .007$ , for per-100-word counts), the fact that some of the MWEs also occurred in the control group's narratives does suggest that not all target MWEs were necessarily produced by participants in the experimental condition under the influence of the input text. This makes the mean use of 2.41 out of 35 target MWEs by the experimental group even less impressive: Some of these MWEs may already have been quite familiar and participants might also have used them in their re-tells without exposure to the textual version of the story.

It is perhaps worth mentioning that the amount of successful mining of MWEs from the input text was positively associated with the learners' general proficiency, as gauged by the placement test. This association, however, was rather weak, and statistically non-significant:  $r = .171$  for participants' total number of recycled MWEs, and  $r = .209$  for their mean number of recycled MWEs per 100 words. It therefore looks as though proficiency level does not have a profound impact on learners' inclination to pick up and reproduce MWEs from text.

Table 2 Descriptive statistics for multiword expressions

Group	<i>n</i>	MWEs shared with input text				MWEs overall			
		<i>M</i>	<i>SD</i>	<i>M</i>	<i>SD</i>	<i>M</i>	<i>SD</i>	<i>M</i>	<i>SD</i>
Experimental group	17	2.41	2.03	0.48	0.39	3.88	2.18	0.78	0.44
Control group	14	0.57	0.65	0.12	0.14	6.07	2.76	1.26	0.70

Note. MWE = multiword expressions.

One may of course argue that a number of the MWEs included in the original story (e.g., *it's now or never*) expressed idea units which were not crucial

for the story line. As their use was not task-essential, perhaps participants did not see any reason for recycling these particular bits of text. The same could also be said about some of the single words, though. For example, the information given in the text that the god Jupiter is “the god of the sky and rain” does not appear vital to the story line either, and yet nine of the participants in the experimental group did include this in their re-tells. Another reason why some MWEs, such as collocations, may not have stood a good chance of being recycled is that a whole collocation could be substituted by a single-word constituent, as in the case of *make + request*, which three participants substituted for by the verb *to request*. This does suggest, though, that it is content words and not their phraseological patterning that tend to be successfully mined from input text.

It needs to be clarified that the above MWE counts concern participants' *accurate* production of the target MWEs. In fact, more traces of the target MWEs were present in the re-tells of the experimental group, but many of these were not target-like. Many were incomplete (e.g., *living by* instead of *earning a living by*, *not notice* instead of *take no notice of*, *look forward to the new master* instead of *look forward to meeting the new master*, *make the best situation* instead of *make the best of the situation*, *than* instead of *rather than*) and others manifested erroneous substitutions of constituents (e.g., *earning the living*, *beg a request*, *have a lesson*, *not enough food for eating*, *look forward about*). The participants in the experimental group produced on average 3.41 ( $SD = 2.45$ ) *inaccurate* versions of MWEs they had encountered in the original story. (These means concern type counts; participants tended to produce the same malformed MWEs in both of their re-tells.) This actually exceeds their mean number of MWEs recycled accurately, although the difference falls short of significance ( $t(16) = 1.75$ ;  $p = .099$ ). The ratio of inaccurate versus accurate versions of MWEs produced by the learners appeared not to be a reflection of their proficiency, since we found no correlation between this ratio and the participants' scores on the placement test ( $r = .024$ ).

Besides the target MWEs (i.e., the MWEs that matched those in the input text), the narratives of the experimental group contained a small number of additional MWEs (e.g., *in addition*, *make a mistake*, *feel sorry for . . .*, *adjust to . . .*), such that the total mean number of per-participant MWEs in the experimental group was 3.88. The average per 100 words was 0.78. The fact that the control group used hardly any of the *target* MWEs does not at all mean that their narratives were void of MWEs. Instead, they used *different* MWEs to help them tell the story (e.g., *first of all*, *do business*, *run away from . . .*, *once again*, *at a loss*, *as a result*, *solve the problem*, *make an effort*, *take care of . . .*). The control group's narratives were in fact found to contain a higher number of MWEs overall than the experimental group's: on average 6.07 per participant and 1.26 per

100 words (see Table 2). This is a significant between-group difference:  $t(29) = 2.47$ ;  $p = .02$ ; CI95% exp. [2.84, 4.92]; CI95% ctrl. [4.62, 7.52], for the per-participant count, and  $t(29) = 2.31$ ;  $p = .03$ ; CI95% exp. [0.57, 0.99]; CI95% ctrl. [0.89, 1.63], for the per-100-words count. In sum, despite the availability of an input text that was rich in MWEs, the experimental group's narratives manifested a *lower* degree of native-like phraseology than the control group's narratives, at least as gauged by counting accurate MWEs. Notice, though, that the confidence intervals for the two groups' means show some overlap, suggesting that the between-group difference is far from absolute and should be interpreted with caution.

In both conditions, the more advanced learners (according to the scores on the placement test) tended to display better command of MWEs, and this was especially striking in the control condition. Correlation coefficients between the participants' placement test scores and the number of accurate MWEs they used were  $r = .448$  in the experimental group and  $r = .685$  in the control group. Computing the correlation for all 31 participants together yields  $r = .484$  ( $p = .006$ ). Computing the correlation for the mean number of accurate MWEs used per minute yields parallel results:  $r = .493$  in the experimental group,  $r = .789$  in the control group, and  $r = .577$  ( $p = .0007$ ) for both groups together. These strong correlations demonstrate that a growing command of multiword lexis is an integral part of becoming a proficient language user. The fact that the correlation is somewhat weaker in the case of the experimental condition probably reflects the aforementioned presence of many inaccurate renderings of MWEs from the input text in these learners' narratives.

## 5. Conclusion, implications for pedagogy, and avenues

The above findings suggest that, while an input text will exert an influence on learners' choice of words when they are given the task to reiterate its content, transfer from input to output at the level of phraseology cannot at all be taken for granted. Not only were there relatively few attempts by our participants at recycling MWEs from the input text, but also a considerable number of those attempts resulted in incomplete or otherwise malformed versions of the MWEs. The question of whether this is due to a lack of attention paid to phraseology during input processing or rather to the challenge of recalling encountered MWEs is probably answerable only through collecting online processing data, such as tracking learners' eye-movements during reading, to determine what elements in the text attract learners' attention. Interestingly, we found only a weak and non-significant association between the learners' recycling of MWEs from the input text and their scores on a general proficiency placement test,

which suggests that proficiency is not a major factor when it comes to adult learners' inclination to mine texts for multiword lexis or phraseology.

The control group in our experiment was required to generate speech based on a picture story rather than being given the opportunity to recycle language from an input text. This nevertheless resulted in output displaying a similar lexical sophistication level and at least the same quantity of MWEs as that of the experimental group. Since the MWEs produced by the control group were not copied from the input text, they must have been retrieved from a repertoire these learners had already developed through prior L2 learning experience. The experimental group, by comparison, relied heavily on words borrowed from the input text, but then often failed to string these words together in native-like ways. As a result, the mean number of accurately produced MWEs was lower than in the control condition.

Irrespective of condition, however, the learners' use of MWEs in their narratives and their scores on a general proficiency test were found to be strongly correlated. This is consistent with earlier research that found significant parallels between learners' MWE use and measures of proficiency, and it lends support to earlier assertions about the importance of mastering multiword lexis as an integral part of proficiency development. Given the rather poor incidental uptake of multiword lexis from input, at least as attested here, the data also lend support to arguments that pedagogic interventions are needed to help learners add MWEs to their L2 repertoires.

That MWEs from the story were successfully recycled by the participants in our study to such a limited extent is particularly surprising in light of the fact that, as argued in the method section, the conditions for mining language were quite favourable. If so little successful mining was attested even under these conditions, then this result may indeed be interpreted as support for more explicit, deliberate MWE-focused instruction and productive practice, where teachers (or materials designers) direct learners' attention to particular MWEs and set up communicative or game-like activities where learners are explicitly encouraged to reproduce them. These teacher-led interventions should of course be considered as complements rather than substitutes of approaches that aim to foster autonomous MWE learning (for reviews of work undertaken to facilitate learners' independent study of MWEs, including the development and use of corpus-informed resources, see, e.g., Boers & Lindstromberg, 2012; and Meunier, 2012).

We also need to acknowledge that it would be premature to draw anything but tentative conclusions about the rate of incidental uptake of MWEs during text-based communicative tasks from the outcome of this exploratory study alone. This was a small-scale study, after all. Besides, the conditions created here were perhaps not yet favourable enough for MWE mining to occur, and so

additional steps for stimulating MWE uptake during input-driven communicative tasks should be tried in future studies. It may be helpful, for example, to incorporate multiple instances of the same MWE in the input text, although it may take a fair amount of creativity on the part of the teacher/materials writer to adapt texts in that way. Typographic enhancement (e.g., underlining) of MWEs in the input text is another possibility, and easy to implement.

Apart from manipulating the input so as to make preselected MWEs more salient, one may also raise learners' awareness of the usefulness of multiword phrases more generally (e.g., Lewis, 1997) and explicitly encourage learners to engage in input mining. Some learners may have been told all too often by teachers that they should sum up the content of a text "in their own words," and they may consequently have become reluctant to repeat wording from an input text verbatim, thus missing an opportunity for practicing idiomatic language, expressing content the way a native speaker would.

More specifically in the context of task repetition, such as the repeated narrative task used in our study, it is probably helpful to insert a feedback stage between the first and second task performance (e.g., Hawkes, 2012), where learners are alerted to inaccurate usage of MWEs or where learners are redirected to the input material to compare their output with. It needs to be recognized, though, that with each of these interventions to stimulate MWE mining during a communicative activity (be they textual enhancement, encouragement to imitate phrases, or the provision of corrective feedback), the activity will increasingly be experienced by the learner as language-focused rather than meaning-focused, in effect shifting the nature of the learning experience from incidental towards intentional MWE-focused practice.

## References

- Alali, F., & Schmitt, N. (2012). Teaching formulaic sequences: The same or different from teaching single words? *TESOL Journal*, 3, 153-180. doi: 10.1002/tesj.13
- Altenberg, B., & Granger, S. (2001). The grammatical and lexical patterning of MAKE in native and non-native student writing. *Applied Linguistics*, 22, 173-194. doi: 10.1093/applin/22.2.173
- Barfield, A., & Gyllstad, H. (Eds.). (2009). *Researching collocations in another language: Multiple perspectives*. Basingstoke: Palgrave Macmillan.
- Bestgen, Y., & Granger, S. (2014). Quantifying the development of phraseological competence in L2 English writing: An automated approach. *Journal of Second Language Writing*, 26, 28-41. doi: 10.1016/j.jslw.2014.09.004
- Bishop, H. (2004). The effect of typographic salience on the lookup and comprehension of unknown formulaic sequences. In N. Schmitt (Ed.), *Formulaic sequences: Acquisition, processing and use* (pp. 227-248). Amsterdam: John Benjamins.
- Boers, F., Dang, T., & Strong, B. (2016). Comparing the effectiveness of phrase-focused exercises. A partial replication of Boers, Demecheleer, Coxhead, and Webb (2014). *Language Teaching Research*. Advance online publication. doi: 10.1177/1362168816651464
- Boers, F., Demecheleer, M., Coxhead, A., & Webb, S. (2014). Gauging the effects of exercise on verb-noun collocations. *Language Teaching Research*, 18, 54-74. doi: 10.1177/1362168813505389
- Boers, F., Demecheleer, M., He, L., Deconinck, J., Stengers, H., & Eyckmans, J. (2016). Typographic enhancement of multiword units in second language text. *International Journal of Applied Linguistics*. Advance online publication. doi: 10.1111/ijal.12141
- Boers, F., Eyckmans, J., & Stengers, H. (2007). Presenting figurative idioms with a touch of etymology: More than mere mnemonics? *Language Teaching Research*, 11, 43-62. doi: 10.1177/1362168806072463
- Boers, F., Eyckmans, J., Kappel, J., Stengers, H., & Demecheleer, M. (2006). Formulaic sequences and perceived oral proficiency: Putting a lexical approach to the test. *Language Teaching Research*, 10, 245-261. doi: 10.1191/1362168806lr195oa
- Boers, F., & Lindstromberg, S. (2009). *Optimizing a lexical approach to instructed second language acquisition*. Basingstoke: Palgrave Macmillan.
- Boers, F., & Lindstromberg, S. (2012). Experimental and intervention studies on formulaic sequences in a second language. *Annual Review of Applied Linguistics*, 32, 83-110. doi: 10.1017/S0267190512000050

- Boston, J. S. (2008). Learner mining of pre-task and task input. *ELT Journal*, 62, 66-76. doi: 10.1093/elt/ccm079
- Boston, J. S. (2010). Pre-task syntactic priming and focused task design. *ELT Journal*, 64, 165-174. doi: 10.1093/elt/ccp033
- Bygate, M. (2001). Effects of task repetition on the structure and control of language. In M. Bygate, P. Skehan, & M. Swain (Eds.), *Researching pedagogical tasks: Second language learning, teaching, and assessment* (pp. 23-48). London: Longman.
- Crossley, A. S., Salsbury, T., & McNamara, D. S. (2015). Assessing lexical proficiency using analytic ratings: A case for collocation accuracy. *Applied Linguistics*, 36, 570-590. doi: 10.1093/applin/amt056
- Dai, Z., & Ding, Y. (2010). Effectiveness of text memorization in EFL learning of Chinese students. In D. Wood (Ed.) *Perspectives on formulaic language: Acquisition and communication* (pp. 71-87). New York: Continuum.
- Dagut, M., & Laufer, B. (1985). Avoidance of phrasal verbs—A case for contrastive analysis. *Studies in Second Language Acquisition*, 7, 73-79. doi: 10.1017/S0272263100005167
- Davis, P., & Kryszewska, H. (2012). *The company words keep: Lexical chunks in language teaching*. Peaslake: Delta.
- Ellis, N. C. (1996). Sequencing in SLA: Phonological memory, chunking and points of order. *Studies in Second Language Acquisition*, 18, 91-126. doi: 10.1017/S0272263100014698
- Erman, B., & Warren, B. (2000). The idiom principle and the open choice principle. *Text*, 20, 29-62. doi: 10.1515/text.1.2000.20.1.29
- Eyckmans, J., Boers, F., & Lindstromberg, S. (2016). The impact of imposing processing strategies on L2 learners' deliberate study of lexical phrases. *System*, 56, 127-139. doi: 10.1016/j.system.2015.12.001
- Hawkes, M. L. (2012). Using task repetition to direct learner attention and focus on form. *ELT Journal*, 66, 327-336. doi: 10.1093/elt/ccr059
- Howarth, P. (1998). Phraseology and second language proficiency. *Applied Linguistics*, 19, 24-44. doi: 10.1093/applin/19.1.24
- Jones, M., & Haywood, S. (2004). Facilitating the acquisition of formulaic sequences: An exploratory study. In N. Schmitt (Ed.), *Formulaic sequences* (pp. 269-300). Amsterdam: John Benjamins.
- Laufer, B. (1995). Beyond 2000: A measure of productive lexicon in a second language. In L. Eubank, M. Sharwood-Smith, & L. Selinker (Eds.), *The current state of interlanguage* (pp. 265-272). Amsterdam: John Benjamins.
- Laufer, B. (2000). Avoidance of idioms in a second language: The effect of L1-L2 degree of similarity. *Studia Linguistica*, 54, 186-196. doi: 10.1111/1467-9582.00059

- Laufer, B. (2011). The contribution of dictionary use to the production and retention of collocations in a second language. *International Journal of Lexicography*, 24, 29-49. doi: 10.1093/ijl/ecq039
- Laufer, B., & Girsai, N. (2008). Form-focused instruction in second language vocabulary learning: A case for contrastive analysis and translation. *Applied Linguistics*, 29, 694-716. doi: 10.1093/applin/amn018
- Laufer, B., & Waldman, T. (2011). Verb-noun collocations in second language writing: A corpus analysis of learners' English. *Language Learning*, 61, 647-672. doi: 10.1111/j.1467-9922.2010.00621.x
- Li, J., & Schmitt, N. (2010). The development of collocation use in academic texts by advanced L2 learners: A multiple case study approach. In D. Wood (Ed.), *Perspectives on formulaic language: Acquisition and communication* (pp. 22-46). New York: Continuum.
- Lin, P. (2012). Sound evidence: The missing piece of the jigsaw in formulaic language research. *Applied Linguistics*, 33, 342-347. doi: 10.1093/applin/ams017
- Lindstromberg, S., Eyckmans, J., & Connabeer, R. (2016). A modified dictogloss for helping learners remember L2 academic English formulaic sequences for use in later writing. *English for specific purposes*, 41, 12-21. doi: 10.1016/j.esp.2015.08.002
- Levitzky-Aviad, T., & Laufer, B. (2013). Lexical properties in the writing of foreign language learners over eight years of study: Single words and collocations. In C. Bardel, C. Lindqvist, & B. Laufer (Eds.), *L2 vocabulary acquisition, knowledge and use. New perspectives on assessment and corpus analysis* (pp. 127-148). EUROSLA.
- Lewis, M. (1997). *Implementing the lexical approach*. Hove: Language Teaching.
- Lewis, M. (Ed.). (2000). *Teaching collocation*. Hove: Language Teaching.
- Lindstromberg, S., & Boers, F. (2008). *Teaching chunks of language: From noticing to remembering*. Rum: Helbling Languages.
- McCarthy, M., & O'Dell, F. (2005). *English collocations in use*. Cambridge: Cambridge University Press.
- Meunier, F. (2012). Formulaic language and language teaching. *Annual Review of Applied Linguistics*, 32, 111-129. doi: 10.1017/S0267190512000128
- Meunier, F., & Granger, S. (2008). *Phraseology in foreign language learning and teaching*. Amsterdam: John Benjamins.
- Nattinger, J. R., & DeCarrico, J. S. (1992). *Lexical phrases and language teaching*. Oxford: Oxford University Press.
- Nesselhauf, N. (2003). The use of collocations by advanced learners of English and some implications for teaching. *Applied Linguistics*, 24, 223-242. doi: 10.1093/applin/24.2.223

- Pawley, A., & Syder, F. (1983). Two puzzles for linguistic theory: Nativelike selection and nativelike fluency. In J. Richards & R. Schmidt (Eds.), *Language and communication* (pp. 191-226). London: Longman.
- Pellicer-Sánchez, A. (2015). Learning L2 collocations incidentally from reading. *Language Teaching Research*. Advance online publication. doi: 10.1177/1362168815618428
- Peters, E. (2009). Learning collocations through attention-drawing techniques: A qualitative and quantitative analysis. In A. Barfield & H. Gyllstad (Eds.), *Researching collocations in another language: Multiple perspectives* (pp. 194-207). Basingstoke: Palgrave Macmillan.
- Peters, E. (2012). Learning German formulaic sequences: The effect of two attention-drawing techniques. *Language Learning Journal*, 40, 65-79. doi: 10.1080/09571736.2012.658224
- Peters, E. (2016). The learning burden of collocations: The role of interlexical and intralexical factors. *Language Teaching Research*, 20, 113-138. doi: 10.1177/1362168814568131
- Peters, E., & Pauwels, P. (2015). Learning academic formulaic sequences. *Journal of English for academic purposes*, 20, 28-39. doi: 10.1016/j.jeap.2015.04.002
- Polio, C. (Ed.). (2012). *Topics in formulaic language (Annual Review of Applied Linguistics, 32)*.
- Samuda, V. (2001). Guiding relationships between form and meaning during task performance: the role of the teacher. In M. Bygate, P. Skehan, & M. Swain (Eds.), *Researching pedagogical tasks: Second language learning, teaching, and assessment* (pp. 119-140). London: Longman.
- Schmidt, R. W. (2001). Attention. In P. Robinson (Ed.), *Cognition and second language instruction* (pp. 3-32). Cambridge: Cambridge University Press.
- Schmitt, N. (Ed.). (2004). *Formulaic sequences*. Amsterdam: John Benjamins.
- Schmitt, N., & Schmitt, D. (2014). A reassessment of frequency and vocabulary size in L2 vocabulary teaching. *Language Teaching*, 47, 484-503. doi: 10.1017/S0261444812000018
- Serrano, R., Stengers, H., & Housen, A. (2015). Acquisition of formulaic sequences in intensive and regular EFL programmes. *Language Teaching Research*, 19, 89-106. doi: 10.1177/1362168814541748
- Sharwood-Smith, M. (1993). Input enhancement and instructed SLA: Theoretical bases. *Studies in Second Language Acquisition*, 15, 165-179. doi: 10.1017/S0272263100011943
- Sinclair, J. (1991). *Corpus, concordance and collocation*. Oxford: Oxford University Press.
- Siyanova, A., & Schmitt, N. (2007). Native and nonnative use of multi-word vs. one-word verbs. *International Review of Applied Linguistics*, 45, 109-139. doi: 10.1515/IRAL.2007.005

- Siyanova, A., & Schmitt, N. (2008). L2 learner production and processing of collocation: A multi-study perspective. *Canadian Modern Language Review*, 64, 429-458. doi: 10.1353/cml.2008.0042
- Siyanova-Chanturia, A., & Martinez R. (2015). The idiom principle revisited. *Applied Linguistics*, 36, 549-569. doi: 10.1093/applin/amt054
- Skrzypek, A., & Singleton, D. (2013). Productive knowledge of English collocations in adult Polish learners: The role of short-term memory. *Vigo International Journal of Applied Linguistics*, 10, 105-129.
- Sonbul, S. (2015). Fatal mistake, awful mistake or extreme mistake? Frequency effects on off-line/on-line collocational processing. *Bilingualism: Language and Cognition*, 18, 419-437. doi: 10.1017/S1366728914000674
- Sonbul, S., & Schmitt, N. (2013). Explicit and implicit lexical knowledge: Acquisition of collocations under different input conditions. *Language Learning*, 63, 121-159. doi: 10.1111/j.1467-9922.2012.00730.x
- Stengers, H., Boers, F., Housen, A., & Eyckmans, J. (2010). Does chunking foster chunk uptake? In S. De Knop, F. Boers, & A. De Rycker (Eds.), *Fostering language teaching efficiency through cognitive linguistics* (99-117). Mouton de Gruyter.
- Stengers, H., Boers, F., Housen, A., & Eyckmans, J. (2011). Formulaic sequences and L2 oral proficiency: Does the type of target language influence the association? *International Review of Applied Linguistics*, 49, 321-343. doi: 10.1515/iral.2011.017
- Szudarski, P., & Carter, R. (2014). The role of input flood and input enhancement in EFL learners' acquisition of collocations. *International Journal of Applied Linguistics*. Advance online publication. doi: 10.1111/ijal.12092
- Szudarski, P., & Conklin, K. (2014). Short- and long-term effects of rote rehearsal on ESL learners' processing of L2 collocations. *TESOL Quarterly*, 48, 833-842. doi: 10.1002/tesq.201
- VanPatten, B. (2002). Processing instruction: An update. *Language Learning*, 52, 755-803. doi: 10.1111/1467-9922.00203
- Wang, Z. (2014). On-line time pressure manipulations: L2 speaking performance under five types of planning and repetition conditions. In P. Skehan (Ed.), *Processing perspectives on task performance* (pp. 27-62). Amsterdam: John Benjamins.
- Webb, S. A., & Chang, A, C-S. (2012). Vocabulary learning through assisted and unassisted repeated reading. *Canadian Modern Language Review*, 68, 267-290. doi: 10.1353/cml.2012.0020
- Webb, S. A., & Kagimoto, E. (2011). Learning collocations: Do the number of collocates, position of the node word, and synonymy affect learning? *Applied Linguistics*, 32, 259-276. doi: 10.1093/applin/amq051.

- Webb, S., Newton, J., & Chang, A. C. S. (2013). Incidental learning of collocation. *Language Learning*, 63, 91-120. doi: 10.1111/j.1467-9922.2012.00729.x
- Wolter, B., & Gyllstad, H. (2011) Collocational links in the L2 mental lexicon and the influence of L1 intralexical knowledge. *Applied Linguistics*, 32, 430-449. doi: 10.1093/applin/amr011
- Wood, D. (2009). Effects of focused instruction of formulaic sequences on fluent expression in second language narratives: A case study. *The Canadian Journal of Applied Linguistics*, 12, 39-57.
- Wood, D. (Ed.) (2010a). *Perspectives on formulaic language: Acquisition and communication*. New York: Continuum.
- Wood, D. (2010b). *Formulaic language and second language speech fluency: Background, evidence and classroom applications*. New York: Continuum.
- Wood, D. (2015). *Fundamentals of formulaic language: An introduction*. London: Bloomsbury.
- Wray, A. (2002). *Formulaic language and the lexicon*. Cambridge: Cambridge University Press.
- Yamashita, J., & Jiang, N. (2010). L1 influence on the acquisition of L2 collocations: Japanese ESL users and EFL learners acquiring English collocations. *TESOL Quarterly*, 44, 647-668. doi: 10.5054/tq.2010.235998.

## APPENDIX

### The input text

At the time of writing this article, the original text and the picture story were available at [http://www.e-yep.com/storyEN/060s\\_donkeyandmaster.swf](http://www.e-yep.com/storyEN/060s_donkeyandmaster.swf). The underlining below indicates the target MWEs. The text presented to the participants did not have this underlining.

#### THE DONKEY AND HIS MASTERS

A donkey lived with his master. His master earned his living by selling herbs in the market. Despite the fact that the donkey always did his best and worked all day, his master wanted the donkey to work harder. So he worked on and on.

Every day the donkey wished he had a different master. The master never gave him enough food to eat. Little by little, the donkey lost hope that things would change. He came to a decision that he had had enough. He couldn't help but think that one way or another, he had to leave. He no longer wanted to work for his master. He thought: "It's now or never". There was only one person who could help the donkey. The donkey searched for Jupiter, the god of the sky and rain.

"Could you do me a favor, Jupiter?" begged the donkey, "Please give me a new master." The donkey explained why he wanted a new master. Jupiter did not turn down the donkey's request. "I will give you a new master," Jupiter said finally. "But," Jupiter said firmly, "you will regret making this request". The donkey was too happy to care. He was really looking forward to meeting his new master.

The donkey's new master was a brick-maker by trade. The donkey had to pull heavy carts of brick every day. The donkey's master wanted him to work harder. He did not give the donkey any water, let alone any food. The donkey was also unhappy with his new master. In desperation, he got in touch with Jupiter again. "Please give me a new master," begged the donkey. "I'll give you a new master," Jupiter said finally.

"But," Jupiter said firmly, "this is the last time I will help you. It will be no use coming back again. This will be your last and only master." The donkey was too happy to care. Again he took no notice of Jupiter's words.

The donkey's new master was a tanner. He caught animals and killed them. Then he skinned them and sold their hide. Now, worst of all, the donkey wasn't just tired. He was also frightened of this new master.

He recalled Jupiter's words: "You will regret wishing for a new master."

The donkey's eyes were full of tears. "On second thoughts, it would have been better if I had stayed with the other masters," the donkey cried. "I may have been hungry, or I may have been tired from too much work, but they would not have skinned me like those animals. He will make use of me even after I die."

In the end, the donkey learned a lesson. It is better to make the best of your situation rather than try to change it.