A Study on Instructed MWEs With Reference to the Movie *After*

Chung, Yunjeong¹

Abstract

The purpose of this paper is to observe how EFL learners use multiword expressions (MWEs) or multiword combinations (MWCs) in their writings. Three subjects participated in this writing activity. After, an American romantic drama film, was chosen as material. Three scenes, which were translated into Korean, were provided to the subjects. They were then assigned to write in English with the Korean compositions. The results indicated various information about the strategies of the subjects' use of MWEs. Subject A depended on one-word, two-word, or three-word utterances due to his lack of English grammatical knowledge. He simply showed sequences based on meaning relations. Subject B used phrase-long MWEs (P-MWEs) relatively more often and combined a P-MWE with another P-MWE to generate sentences. His grammatical knowledge motivated such combinations, but the sentences hardly looked natural. Subject C primarily used sentence-long MWEs. They seemed to make his composition effortless. Through these writing activities, it was found that MWCs can be developed to MWEs in the case of subject A and that grammar learning before MWCs or MWEs learning may be harmful to learn L2 naturalness in B's case, and that sentence-long MWEs should be an ultimate goal in C's case.

Keywords: multiword expressions, multiword combinations, sentence-long MWEs, phrase-long MWEs

Applicable level: college

---

¹ Corresponding author, Lecturer, Department of English Education, Daegu National University of Education, 219, Jungang-daero, Nam-gu, Daegu, Republic of Korea (E-mail: ychung80@gmail.com)

Received: April 15, 2021

Revised: May 15, 2021

Accepted: May 23, 2021

Copyright: © 2021 The Society for Teaching English through Media (STEM)

This is an open access article distributed under the terms of the Creative Commons Attribution Non-Commercial License (<http://creativecommons.org/licenses/by-nc/4.0/>), which permits unrestricted non-commercial use, distribution, and reproduction in any medium, provided the original work is properly cited.

I. INTRODUCTION

Multiword expressions (MWEs)¹ are assumed to play an important role in EFL learners' language use (Alali & Schmitt, 2012; Rafieyan, 2018a). They behave as single units such as idioms, binominals, collocations, lexical bundles, proverbs, and probably more (Conklin & Schmitt, 2012). Therefore, their mastery is supposed to be deeply related to a key component of language proficiency (Rafieyan, 2018b). Since Pawley and Syder (1983), MWEs research has been increasingly accumulated, but it is a good time to see MWEs from different perspectives.

A new perspective comes from EFL learners themselves. Suppose a low-intermediate learner who lacks grammatical knowledge and vocabulary of English. If they are instructed to write a composition in English, what kind of knowledge do they depend on? How do they construct a sentence without grammatical knowledge and a limited vocabulary? Most likely, they will have to count on MWEs or multiword expressions (MWCs)² (Berk & Lillo-Martin, 2012), which they often find from maybe Google Translate and Naver Papago. Such language forms must be different from MWEs mentioned in current literature. The language forms are simply sentence-long MWCs, not MWEs right now. However, nobody denies that their language forms can become MWEs soon or later (Ellis, 1996). Language is considered dynamic. The rules of human languages are not static but change over time. "New words come into the language all the time and others become obsolete" (Steels & Szathmáry, 2018, p. 128). Like this, L2 learners' MWCs come into MWEs. Over time, change will occur again. Their MWEs can disappear and they reinvent new MWEs.

In this paper, several aspects of MWEs will be described: definition of MWE, subtypes of MWE, merits of MWE, etc. After this, one important observation will be revealed. Three EFL college-level students will participate in English writing session. One student is low-intermediate, another is high-intermediate, and the last one is very high-advanced. It will be observed how each student uses MWEs in their writing, which is the purpose in this paper. The material for this paper will be a movie, which was selected due to the rich context it provides to help improve student writing.

II. LITERATURE REVIEW

1. Definition of MWEs

Sag, Baldwin, Bond, Copestake, and Flickinger (2002) define MWEs as "idiosyncratic interpretation that cross word boundaries (or spaces)" (p. 2). *Idiosyncratic* means strange, which refers to deviation from normal. For example, "ad hoc" is lexically idiosyncratic (Baldwin & Kim, 2010). The meaning of "ad hoc" cannot be predicted from separate meanings of its component words, "ad" and "hoc." This is idiosyncrasy which means "idiomaticity." There are lexical idiomaticity, syntactic idiomaticity, semantic idiomaticity, pragmatic idiomaticity, and statistical idiomaticity. Yet, this paper does not review such concepts because such a review is not the concern of the paper. "Cross word boundaries" means at least two words. Put together, MWEs are at least two-word expressions which are lexically, syntactically, semantically, pragmatically, and statistically idiosyncratic in nature. One definition which has a reputation in this field is Wray's (2002):

a sequence, continuous or discontinuous, of words or other elements, which is, or *appears to be*, [my italic] prefabricated: that is, stored and retrieved whole from memory *at the time of use*, [my italic] rather than being subject to generation or analysis by the language grammar. (p. 9)

Wray's definition corresponds to the purpose of this paper in two ways. First, "appears to be" means "seems like." In other words, it is true or maybe not. Thus, a sequence can be an MWE or not. In this paper, subjects use MWEs or MWCs. It does not matter whether they use either one. For the moment, both are accepted as MWEs. Second, "at the

¹ There are so many terms for MWE. This paper uses MWE as an umbrella term for such many terms.

² MWCs simply refer to a sequence of words which are not collocated each other. MWEs reflect "any kind of phraseological unit[s]" (Sailer & Markantonatou, 2018, p. 4).

time of use” refers to the moment a language user utters a sequence. Turning back to the paper, if subjects use sequences like MWEs at the time of use, the sequences will be accepted as MWEs. Therefore, Wray’s definition seems close to individual inclinations. These are the principles this paper follows.

2. Classification of MWEs

MWEs are crucial for Natural Language Processing³ (NLP) because they occur in any natural language (Kumar, Behera, & Jha, 2017). Furthermore, MWEs are heterogeneous, so a unified classification of MWEs is difficult. Laporte (2018) tries to design a satisfactory classification of MWEs, saying “current practice routinely uses fuzzy features, or features defined in an imprecise way” (p. 180). It is hard to identify MWEs without appropriate classification. The current state of MWE research is split up into diversified categories from different perspectives (Wang, 2020). Specifically, Wang describes three different categories: pedagogically-oriented, linguistically-oriented, and NLP-oriented. In this paper, the pedagogically-oriented category will be selected for the EFL classroom. He also shows part of categories in English as below:

[...] (1) polywords that function like individual lexical items, such as *by the way*; (2) phrasal constraint, such as *a ___ ago*, *dear ___*, *the ___er*; (3) sentence builders, containing slots for parameters or arguments, like *I think that X*, *That reminds me of X*, *Have you heard about X?*; (4) collocations (*noun + adjective*, *verb + noun*, *verb + adverb*, etc.); (5) institutionalized expressions that have pragmatic functions. They stand as separate utterances in distinct social situations, such as *How do you do?*; (6) discourse devices, such as logical connectors — *as a result of*, *in spite of*; temporal connectors — *the next is Y*; spatial connectors — *at the corner*; fluency devices — *you know*; exemplifiers — *in other words*; summarizers — *to sum up*, *and so on*. These categories are very wide, from morphemes to sentences, not committing to any linguistic status of this multiword unit phenomenon. (p. 43)

It seems that polywords, phrasal constraint, sentence builders, collocations, institutionalized expressions, discourse devices, logical connectors, temporal connectors, spatial connectors, fluency devices, exemplifiers, and summarizers are unnecessary for EFL learners to remember. Therefore, this paper will only focus on the pedagogical approach.

Siyanova-Chanturia (2017) reviewed the current pedagogy-oriented contribution papers of 2014, *Annual Review of Applied Linguistics*, 32 and 2014, *The Mental Lexicon*, 9(3). Macis and Schmitt (2017a, 2017b) did research on collocations. Macis and Schmitt (2017a) investigated whether Chilean Spanish-speaking college students of English have knowledge of the figurative meanings of collocations that can be originally both literal and figurative. The result indicates that they have limited collocation knowledge. If they spend time reading English books or stay in English-speaking countries, their knowledge of the figurative meanings of collocations would be enhanced. Macis and Schmitt (2017b) did research on how to teach different meaning senses of collocations. The authors asserted that collocations can be split up into 3 categories from the meaning perspective: literal collocations, figurative collocations, and duplex collocations. They analyzed 54 collocations and found that the majority of the collocations are literal, but a significant number of collocations have literal and figurative meanings. Also, they discovered relatively few collocations are solely figurative. Thus, if instructors teach collocations, they must consider the meaning. Eyckmans and Lindstromberg (2017) experimented the effect of sound in L2 idiom learning, which is example one of MWEs. In this study, 26 advanced-level EFL learners of Dutch show superiority over phonologically dissimilar control idioms when they learn significantly more phonologically similar idioms. Alliteration (*miss the mark*) and assonance (*get this show on the road*) show phonological similarity.

³ Natural Language Processing is “a theoretically motivated range of computational techniques for analyzing and representing naturally occurring texts at one or more levels of linguistic analysis for the purpose of achieving human-like language processing for a range of tasks or applications” (Liddy, 2001).

3. Types of MWE

In Table 1, Baldwin and Kim's (2010) description of types of MWE is tabulated below. However, it is a simplification of the original version of Baldwin and Kim (pp. 274-279).

TABLE 1
Types of MWE

Types	Subtypes	Examples
Nominal MWEs		golf club, computer science department stress avoidance connecting flight open secret
Verbal MWEs	Verb-particle constructions	play around, take off cut short, band together let go, let fly
	Prepositional verbs	refer to, look for come across, grow on
	Light-verb constructions	do a report give a sigh have a drink make a mistake take a bath
	Verb-noun idiomatic combinations	kick the bucket, shoot the breeze
Prepositional MWEs	Determinerless-prepositional phrases	on top vs *on bottom by car/foot/bus at high expense, on summer vacation *at level vs at eye level *at expense vs at company expense on top of
	Complex prepositions	on top of, in addition to

Though many types of MWEs are introduced here, teachers cannot teach this kind of information in the classroom. If this is instructed to students, it will be a big burden to them. Additionally, the types of MWEs of this paper will be introduced in Chapters III and IV.

III. RESEARCH DESIGN

The goal aims at observation of how three subjects use MWEs in completing their composition. Specifically speaking, the goal is to find the role of MWEs for each subject. It will also be observed how they combine phrase-long MWEs with sentence-long ones and whether their grammatical knowledge plays a certain role in the combination.

1. Subjects

The three subjects are college-level male students. They are divided into three levels based on their TOEIC scores. Subject A is low-intermediate with a score around 350, Subject B is high-intermediate with a score of around 750, and Subject C is high-advanced with a score around 960. It is presumed that low-intermediate subject A only will use sentence-long MWEs or MWCs because he lacks grammatical knowledge and vocabulary. Moreover, he will depend on translation machines to get MWEs. Presumably the high-intermediate and high-advanced subjects will use phrase-long MWEs. The high-advanced subject C has much knowledge of grammar and vocabulary, so it is presumed that he will create grammatical constructs. In his composition, it will be observed how the ratio of MWEs to non-MWEs is expressed.

2. Material

In this paper, the 2019 American romantic drama film, *After* (Gage, 2019), will be used as material. However, the subjects do not watch the movie, nor does the teacher tell them that their writing compositions are based on a movie script. Though English scenes are shown in Chapter IV, the scenes are for this research, not for the subjects. All that subjects receive for writing is pieces of paper which include dialogue and the relevant context written in Korean. The subjects will not be able to guess that some pieces of paper in Korean are from the movie, *After*.

There is no particular reason to choose this movie. It is simply chosen to provide rich context with the subjects. This is because romantic movies usually have rich context about feelings among characters. Three scenes are selected for writing. However, the scenes in English will not be provided to the subjects. Instead, they are given the relevant context in Korea. The three subjects are supposed to read both context and utterances in Korean. Then they will write a composition in English.

3. Evaluation

Two features to identify MWEs in this experiment are *single units* and *length of MWEs*. Single units refer to MWEs, and the length of MWEs refers to sentence-long MWEs (S-MWEs) or phrase-long MWEs (P-MWEs).

There are some excuses to admit in this paper. Let us examine single units first. The idea that a single unit is an MWE refers to taking out the multiword string as a whole from the mental lexicon (Hüning & Schlücker, 2015; Wray, 2008). In this paper, if a subject utters multiword strings from his memory, it is considered as a single unit. Moreover, even though a multiword string is combined unconventionally (e.g., *rich coffee* rather than *strong coffee*), it can be both an MWE and a single unit for this paper. Many EFL learners, including subjects in this paper, are not familiar with MWEs. That is why some errors for combination will be ignored.

Another feature *creative construct* that is not related to MWEs in this paper refers to computation of individual words. Though their computation turns out ungrammatical, such an error will also be excused.

It is believed that each subject's composition consists of MWEs and creative constructs. The number of MWEs and creative constructs will be counted separately. Each subject writes 3 compositions and each composition has about 5 sentences⁴. In total, each subject writes 15 sentences in English. The ratio of MWEs to creative constructs will be interpreted and discussed in Chapter IV.

IV. RESULT AND ANALYSIS

In this chapter, three scenes from the movie *After* are used for analysis. As mentioned in the previous chapter, each scene or dialogue is provided in Korean and then subjects are supposed to rewrite in English.

Tessa: Um, uh, excuse me? Uh, I th... I think that you're in the wrong room.

Hardin: I'm in the right room.

Tessa: How did you even get in here?

(Hardin shows keys of the room.)

Tessa: Okay. Can you please go out into the hall so I can get dressed?

Hardin: Don't flatter yourself. I'm not looking.⁵

(Conversation 1, *After*)

As introduced at evaluation section of III Research Design, conversations of Tables are going to be analyzed according to features, such as phrase-long and sentence-long MWEs, unit words, and constructs.

⁴ It should be admitted that more data leads to better results. However, space limitations should be considered, too.

⁵ In the context of this Conversation 1, Tessa is a freshman, and Hardin is an enrolled student. They are the main characters and fall in love later. After taking a shower in the communal bathroom, Tessa returns to her room. She casually opens the closet door and is startled to see, reflected in the closet mirror, that a man is sitting down in the room.

TABLE 2
Conversation 1 Analysis

Sub	P-MWEs	S-MWEs	P-construct	S-construct	S-single unit words	P-single unit words	One-word grammars	Total
A	0	7	0	0	1	1	1	10
B	1	3	0	4	0	1	0	9
C	0	5	0	3	0	1	0	9

Note. Short form for phrase-long MWEs = P-MWEs, sentence-long MWEs, phrase-long creative constructs = S-MWEs, sentence-long creative constructs = S-construct, sentence-long single unit words = S-Single unit words, phrase-long single unit words = P-single unit words

Subject A used 7 MWEs, but they were all from Google Translate and Naver Papago. This means they cannot be accepted as MWEs. However, he used 3 single words as functional single units. One was “anyway.” It is an adverb, and it is used as a discourse marker. It belongs to phrase-long single unit word functionally. It is called a phrase-long single unit word here. Another was “please.” It belongs to sentence-long single unit word functionally in this context because it is used as a social formula. The final was ‘for’. He used this word as grammatical conjunction even though he was wrong. The next is how A uses “for” in his sentence.

Please you go out to the hall for I change my clothes → S + (grammatical conjunction, “for”) + S

Subject A found two sentences in the translation machines. The first sentence was “you go out to the hall” and the second one, “I change my clothes.” He found that there is cause-effect relation between the two sentences. Therefore, he needed to use the conjunction “because,” but he ended up choosing “for” instead of “because” due to his lack of grammatical knowledge.

Basically, he depends on the translation machines for multiword expressions because he has no MWEs in his memory. If he starts English learning, he needs sentence-long MWEs in the first place. As shown in Table 2, he used 7 S-MWEs, but he did not use P-MWEs at all. In a sense, it is very reasonable because people want to express whole meanings, not part meanings. When a teacher wants to teach MWEs to low-intermediate students, they had better teach S-MWEs. Since MWEs come out from learners’ mental lexicon, that is, memory, memorization activities are recommended.

Next, let us consider subject B. As mentioned before, his TOEIC score was about 750. Interestingly though, he used 4 sentence-long constructs the most. Using S-constructs means B makes sentences with his grammatical knowledge. He used 3 sentence-long MWEs the second most. The three S-MWEs were “Excuse me,” “I’ve got it right,” and “How did you get in here?.” B has the first MWE in his memory, but the other two MWEs seem to come from different sources. Also, the first P-MWE, “wrong room” is a compound word. A also used this compound word, but his use is different from B’s use. Subject A does not have the compound word in his memory. He borrowed “I think you came to the wrong room” from the translation machines. For A, the sentence was a single unit. On the other hand, in B, the compound word, “wrong room” was a single unit. Another example of a single unit was when B used “anyway.”

In sum, B could make sentences with grammatical knowledge. When he faced some expressions in Korean which seem to be difficult in translating English, B used the translation machines. When a teacher teaches MWEs to learners like B who are intermediate, they can instruct their learners with phrase-long and sentence-long MWEs.

Lastly, let us investigate what Subject C, the student who scored about 960 on the TOEIC test, wrote. C used 5 sentence-long MWEs the most. It is believed that the five MWEs are in his memory. Compared to A’s 7 MWEs, C’s S-MWEs are easily split into P-MWEs. Since A’s MWEs are based on translation machines, their flexibility is low. C also uses three sentence-long constructs, compared to B’s 4. Yet, C’s composition is more natural because the number of MWEs is greater than that of constructs. Whereas in the case of B, the number of constructs is more than that of MWEs.

To sum up Table 2, the three subjects mainly used S-MWEs and S-constructs. Since the use of S-MWEs was the main concern in this paper, only they were summarized. Subject A used many MWEs, but they were not in A’s memory. Such MWEs or MWCs were admittedly single units, but they dropped flexibility. B’s MWEs were relatively small,

but his composition did not sound natural. As expected, C was the best. He used 5 MWEs, which were more than those of S-constructs. However, C's MWEs are qualitatively suspicious, so all subjects need to learn MWEs first for one reason or another.

(dormitory room)

Tessa: What do you think?

Steph: What? It's... It's pretty. Maybe it's just... a little too formal?

Tessa: You said be myself.

Steph: You know what? I love it. I love it.⁶

(Conversation 2, *After*)

TABLE 3
Conversation 2 Analysis

Sub	P-MWEs	S-MWEs	P-constructs	S-constructs	S-Single unit words	P-Single unit words	One-word grammars	Total
A	0	7	0	0	1	0	0	8
B	0	6	0	0	3	0	0	9
C	0	7	0	0	2	0	0	9

Note. Short form for phrase-long MWEs = P-MWEs, sentence-long MWEs, phrase-long creative constructs = S-MWEs, sentence-long creative constructs = S-construct, sentence-long single unit words = S-Single unit words, phrase-long single unit words = P-single unit words

Table 3 seems unequally distributed. All subjects mainly used S-MWEs. Subject A depended on the translation machines 100%. B also counted on the translation machines and extra material. Yet, C was different. Considering his language skills, we are quite sure C would use S-MWEs and S-single unit words without the help of the translation machines.

Then why did the subjects heavily depend on S-MWEs? There might be at least two reasons. First, the subjects are using Korean writings to create English compositions. Therefore, it is unsurprising that all three subjects find the Korean writings to be very natural in the Korean sense. This is because all the writings are written by a Korean native speaker, making them very conventional and natural. On the other hand, when they compose the writings in English, it is not easy for the subjects to do it because of their unfamiliarity with English. To the subjects, every utterance in the Korean writings looks like single units. That is why they must depend on S-MWEs.

Subject A had no S-MWEs in his memory, so he had no choice to compose on his own. Instead, he must use the translation machines. Subject B was half and half. For example, one of his writings, "You told me to live on my way," was proceeded in the following way.

You told me/to live on my way.

Subject B can write "you told me" on his own. However, the second part "to live on my way" requires the use of translation devices.

Second, when people write or speak, they think of what to write or say in terms of meaning, not of syntax or grammar. Grammar cannot catch up with the relevant meaning, so language users need linguistic devices to absorb huge meanings at once. Sentence-long MWEs can be an essential linguistic device (Church, 2013; Shudo, Kurahone, & Tanabe, 2011; Tanabe, Takahashi, & Shudo, 2014), while phrase-long MWEs are not enough to fully convey meanings which language users express.

In Table 3, subjects prefer sentence-long MWEs regardless of their language abilities. It is called *single big words* (Ellis, 1996). Therefore, when a teacher teaches MWEs in the classroom, they had better focus on S-MWEs. Yet, it could be asked, What about P-MWEs? In those cases, teachers should utilize sentence-long MWE. Then learners will

⁶ In the context of this Conversation 2, Steph and Tess are roommates. Steph is an enrolled student like Hardin. She wants to introduce Tessa to her friends, and also Steph wants Tessa to participate in school activities. Now they prepare themselves to go to the party.

naturally find phrase-long MWEs from the analysis of S-MWEs (Ellis, 1996).

The next is a phone conversation between Tessa and Noah. This is a long conversation for analysis, so only five underlined utterances are handed out to the subjects.

Tessa-Noah conversation on the phone

(Tessa calls Noah outside)

Tessa: Hey.

Noah: Hey. Thought we were gonna FaceTime.

Tessa: Yeah, sorry.

Noah: Where are you right now? It's really loud.

Tessa: Um... I'm with Steph and her friends, but they're all just, like... I don't know.

Noah: So, uh, you're at a party? Have you been drinking?

Tessa: I just had one drink.

Noah: Okay, so you go to college and now you drink. That's... That's really great, Tessa.

Tessa: Noah, can you not be so, like...

Noah: So, like, what? I'm not the one who's out partying right now.

Tessa: Just forget it.

Noah: Tessa, I wanna... ⁷

(Conversation 3, *After*)

TABLE 4
Conversation 3 Analysis

Sub	P-MWEs	S-MWEs	P-constructs	S-constructs	S-Single unit words	P-Single unit words	One-word grammars	Total
A	0	3	0	2	1	0	0	6
B	3	2	0	1	1	1	0	8
C	0	4	0	1	0	0	0	5

Note. Short form for phrase-long MWEs = P-MWEs, sentence-long MWEs, phrase-long creative constructs = S-MWEs, sentence-long creative constructs = S-construct, sentence-long single unit words = S-Single unit words, phrase-long single unit words = P-single unit words

In Table 4, let us look at A's composition. He had little knowledge of English grammar, so he used three S-MWEs and made two sentence-long constructs. The three S-MWEs are "very loud," "You drunk?," and "only one drink." His MWEs are like the two-word or three-word utterances of children. He cannot write "It's very loud," "Are you drunk?," and "I had only one drink." Instead, he uses shortened forms of the three MWEs, but they are all functionally utterances.

One thing which is amazing about A is the fact that he has some MWEs in his memory. When he is asked how he knows the three MWEs, he answers, he remembers the first one from song lyrics and the second and the third ones from children's books and an American animated sitcom. However, he only remembers informal shortened forms, not full sentence forms. What does this mean? How can the phenomenon be interpreted? Answers might be found from the two-word utterance stage. This stage occurs within the age range of 19-26 months (Berk & Lillo-Martin, 2012). Like this, subject A experiences the same phenomenon. He produces one-word utterances (e.g., *drink*) or two-word utterances. These occurrences promise multiword expressions later (Clark, 2009). Subject A has linguistic limitations right now, which cause him to produce one-word or two-word utterances, but sometime later, his production will become linguistically stable MWEs.

Now let us consider A's creative constructs. The followings are his constructs.

⁷ In the context of this Conversation 3, Noah is Tessa's boyfriend, who her mother permits them to be together. Tessa is not in a good mood at the party, so she goes outside to give Noah a call. Noah acts like her mother. He worries about Tessa, and he is afraid that she is going out with ungroomed college friends to drink instead of studying.

You place in party now. → You + place + in party + now
 You entered college. → You + entered + college

The first sentence is completely ungrammatical. According to subject A, he does not know which verb needs to be inserted in the position of the verb. This leads him to put a noun (*place*) there instead. The second sentence is perfect. Since he knows the right verb, he can complete a sentence. From the two example sentences, we can guess that he can make a transitive construction as long as he knows a verb. In fact, a transitive construction is a basic construction in the world's languages (Ibbotson, Theakston, Lieven, & Tomasello, 2012). If lexical items are supported, subject A can make grammatical constructions of transitive constructions.

In sum, subject A has some MWCs in his memory. Theoretically, one cannot say that they are all MWEs. However, it is likely that sometime later, his MWCs will turn into MWEs.

Now let us turn our attention to B. He used three phrase-long MWEs in his composition. No phrase-long MWEs exist in subject A and subject C in Table 4. What does it mean? The followings are his utterances.

You've got drunken? → You've got + drunken
 You've got in the college. → You've got + in the college

"Have got" and "have" mean the same. The former is simply informal. Subject B believes that "have got" emphasizes the meaning of process, and that instead "have drunken" represents a state. This shows that he prefers "have got" in both utterances. Though he used "you've got" inappropriately, "you've got" is phrase-long MWEs to him. In the second sentence, "in the college" is another P-MWE. Here he connected "you've got" and "in the college" mistakenly. His grammatical knowledge failed to monitor grammaticality.

Subject B used two sentence-long MWEs.

It's really noisy.
 Only one shot

In the first sentence, subject B produced a full-length utterance. In the second sentence, B also used a short form with a three-word utterance. He did this because he couldn't think of a full sentence, so the short form was substituted. However, both examples are regarded as sentence-long MWEs from the functional point of view.

In addition to this, he shows an interesting functional contrast of one lexeme (drunken vs drink).

You've got drunken. → You've got + drunken
 You've got in the college and drink. → You've got + in the college + and + (you) + drink

In the first sentence, "drunken" is an adjective and part of a sentence. On the other hand, "drink" in the second sentence is a verb and it leads to another sentence, "you drink." In these cases, "drunken" is a P-single unit word and "drink" is a one-word utterance which is an S-single unit word.

In sum, one characteristic of subject B is that he uses three phrase-long MWEs and one phrase-long single unit word in making sentences. They are relatively done often. Also, they are supposed to combine with each other, such as "P-MWE + P-single unit word" and "P-MWE + P-MWE." The problem is that the resultant sentences do not look natural. His grammatical knowledge fails to achieve naturalness. He needs to learn MWEs first and then better to learn grammar.

Last, let us observe subject C. The components of his composition are very simple: four sentence-long MWEs and one sentence-long construct. Since he has sentence-long MWEs in his memory, it is not difficult for him to make a sentence. Also, he has good grammatical knowledge allowing him to make a long, grammatically correct sentence.

To summarize Conversation 3, subject A used three sentence-long MWEs, and subject C used four. However, MWEs of subject A were all either two-word or three-word utterances. Technically speaking, his MWEs are MWCs right now. His linguistic level is like a 2-year-old child, but these utterances are in his memory. This is a positive point. Someday his MWCs would turn into MWEs if he keeps language use. On the other hand, C used four S-MWEs.

They are good linguistic devices to generate sentences. While B has phrase-long MWEs, he often fails to make larger size MWEs. He had better start English learning with sentence-long MWEs.

V. CONCLUSION AND PEDAGOGICAL IMPLICATIONS

Considering the important role of MWEs in language use (Constant, et al., 2017; Schmitt & Carter, 2004), we cannot avoid identifying, classifying, and labeling the MWEs (Kumar, Behera, & Jha, 2017). Moreover, MWEs are heterogeneous, and then such workings are essential (Laporte, 2018; Wang, 2020).

However, in the EFL classroom, some teachers are dubious about the fact that identification, classification, and labeling of MWEs are necessary for L2 learners. This paper deals with information on MWEs, which learners need to know, the definition of MWEs, the types of MWEs, and the learners' strategies for using MWEs. In this paper, multiword combinations are called MWEs, with regards to how learners use them as single units. Meaning that multiword combinations and multiword expressions are compatible in this paper. Types of MWEs are sentence-long and phrase-long multiword combinations, single words as functionally sentence-long, and phrase-long multiword combinations. Regarding learners' strategies, this paper observes when learners use sentence-long, phrase-long multiword combinations, and one-word utterances and why they use such combinations or utterances. Three subjects (beginner A, high-intermediate B, and high advanced C) were invited in this paper, and the material was the movie, *After*.

The result showed much information about each subject. Regarding subject A, he mainly used sentence-long MWEs because he had to use translation machines that would give him sentence-long expressions. When he faced inappropriate situations not to use the machines, he used one-word or two-word utterances⁸. It turned out that subject A had pieces of language like them in his memory. Subject B showed combinatorial strategies between phrase-long MWEs. However, his expressions were not natural. He needs to learn sentence-long MWEs before phrase-long MWEs to overcome his unnaturalness. Subject C mainly used sentence-long MWEs like subject A. The difference between C and A was that C's MWEs were in his memory and A's MWEs were not. Thanks to C's memory, he wrote sentences without much effort. However, he did not use phrase-long MWEs much. Yet, this observation cannot tell why.

There is one last important thing about MWEs or multiword combinations. Subject A showed characteristics of 2-year-old children who produce two-word utterances: a lack of grammatical markers and a preponderance of meaning relations (Berk & Martin, 2012). Moreover, meaning relations may bring forth sequence learning which serves as the database for grammar learning. Ellis (1996) said that "language learning is the learning and analysis of sequences" (p. 92).

Returning to the EFL classroom. A teacher hopefully encourages their students to express their intention or meaning through meaning-based sequence learning. Awareness of MWEs might be harmful to students. Thus, teachers should notice one thing. MWEs should be learned in the speech community, not in the classroom.

REFERENCES

- Alali, F. A., & Schmitt, N. (2012). Teaching formulaic sequences: The same as or different from teaching single words? *TESOL Journal*, 3(2), 153-180. <https://doi.org/10.1002/tesj.13>
- Baldwin, T., & Kim, S. N. (2010). Multiword expressions. In N. Indurkha & F. J. Damerau (Eds.), *Handbook of natural language processing* (2nd ed., pp. 267-292). Boca Raton, FL: CRC Press.
- Berk, S., & Lillo-Martin, D. (2012). The two-word stage: Motivated by linguistic or cognitive constraints? *Cognitive Psychology*, 65(1), 118-140. <https://doi.org/10.1016/j.cogpsych.2012.02.002>
- Church, K. (2013). How many multiword expressions do people know? *ACM Transactions on Speech and Language Processing*, 10(2), Article 4. <https://doi.org/10.1145/2483691.2483693>

⁸ Two-word utterances are functionally sentence-long or phrase-long MWEs, depending on the context.

- Clark, E. V. (2009). *First language acquisition*. Cambridge, UK: Cambridge University Press.
- Conklin, K., & Schmitt, N. (2012). The processing of formulaic language. *Annual Review of Applied Linguistics*, 32, 45-61. <https://doi.org/10.1017/S0267190512000074>
- Constant, M., Eryigit, G., Monti, J., van der Plas, L., Ramisch, C., Rosner, M., & Todirascu, A. (2017). Multiword expression processing: A survey. *Computational Linguistics*, 43(4), 837-892. https://doi.org/10.1162/COLI_a_00302
- Ellis, N. C. (1996). Sequencing in SLA: Phonological memory, chunking, and points of order. *Studies in Second Language Acquisition*, 18(1), 91-126. <https://doi.org/10.1017/S0272263100014698>
- Eyckmans, J., & Lindstromberg, S. (2017). The power of sound in L2 idiom learning. *Language Teaching Research*, 21(3), 341-361. <https://doi.org/10.1177/1362168816655831>
- Gage, J. (Director). (2019). *After* [Motion picture]. Unite States: Voltage Pictures.
- Hüning, M., & Schlücker, B. (2015). Multi-word expressions. In P. O. Müller, I. Ohnheiser, S. Olsen, & F. Rainer (Eds.), *Word-formation. An international handbook of the languages of Europe* (Vol. 1, pp. 450-467). Berlin, Germany: de Gruyter Mouton. <https://doi.org/10.1515/9783110246254-026>
- Ibbotson, P., Theakston, A. L., Lieven, E. V. M., & Tomasello, M. (2012). Semantics of the transitive constructions: Prototype effects and developmental comparisons. *Cognitive Science*, 36(7), 1268-1288. <https://doi.org/10.1111/j.1551-6709.2012.01249.x>
- Kumar, S., Behera, P., & Jha, G. N. (2017). A classification-based approach to the identification of multiword expressions (MWEs) in Magahi applying SVM. *Procedia Computer Science*, 112, 594-603. <https://doi.org/10.1016/j.procs.2017.08.059>
- Laporte, E. (2018). Choosing features for classifying multiword expressions. In M. Sailer & S. Markantonatou (Eds.), *Multiword expressions: Insights from a multi-lingual perspective* (pp. 143-186). Berlin, Germany: Language Science. <https://doi.org/10.5281/zenodo.1182597>
- Liddy, E. D. (2001). Natural language processing. In M. Drake (Ed.), *Encyclopedia of library and information science* (2nd ed., pp. 51-89), New York, NY: Marcel Decker. Retrieved from <https://surface.syr.edu/istpub/63/>
- Macis, M., & Schmitt, N. (2017a). Not just ‘small potatoes’: Knowledge of the idiomatic meanings of collocations. *Language Teaching Research*, 21(3), 321-340. <https://doi.org/10.1177/1362168816645957>
- Macis, M., & Schmitt, N. (2017b). The figurative and polysemous nature of collocations and their place in ELT. *ELT Journal*, 71(1), 50-59. <https://doi.org/10.1093/elt/ccw044>
- Pawley, A., & Syder, F. H. (1983). Two puzzles for linguistic theory: Nativelike selection and nativelike fluency. In J. C. Richards & R. W. Schmidt (Eds.), *Language and communication* (pp. 191-226). New York, NY: Longman.
- Rafieyan, V. (2018a). Knowledge of formulaic sequences as a predictor of language proficiency. *International Journal of Applied Linguistics & English Literature*, 7(2), 64-69. <https://doi.org/10.7575/aiac.ijalel.v.7n.2p.64>
- Rafieyan, V. (2018b). Role of knowledge of formulaic sequences in language proficiency: Significance and ideal method of instruction. *Asian-Pacific Journal of Second and Foreign Language Education*, 3(9). <https://doi.org/10.1186/s40862-018-0050-6>
- Sag, I. A., Baldwin, T., Bond, F., Copestake, A., & Flickinger, D. (2002). Multiword expressions: A pain in the neck for NLP. In A. Gelbukh (Ed.), *Computational linguistics and intelligent text processing. CICLing 2002. Lecture notes in computer science* (Vol. 2276, pp. 1-15). Berlin, Germany: Springer. https://doi.org/10.1007/3-540-45715-1_1
- Sailer, M., & Markantonatou, S. (2018). Multiword expressions: Insights from a multi-lingual perspective. In M. Sailer & S. Markantonatou (Eds.), *Multiword expressions: Insights from a multi-lingual perspective* (pp. 3-31). Berlin, Germany: Language Science. <https://doi.org/10.5281/zenodo.1186597>
- Schmitt, N., & Carter, R. (2004). Formulaic sequences in action: A introduction. In N. Schmitt (Ed.), *Formulaic sequences: Acquisition, processing and use* (pp. 1-22). Amsterdam, The Netherlands: John Benjamins.
- Shudo, K., Kurahone, A., & Tanabe, T. (2011). A comprehensive dictionary of multiword expressions. *Proceedings of the 49th Annual Meeting of the Association for Computational Linguistics: Human Language Technologies*, 1, 161-170. Retrieved from <https://dl.acm.org/doi/abs/10.5555/2002472.2002494>

- Siyanova-Chanturia, A. (2017). Researching the teaching and learning of multi-word expressions. *Language Teaching Research*, 21(3), 289-297. <https://doi.org/10.1177/1362168817706842>
- Steels, L., & Szathmáry, E. (2018). The evolutionary dynamics of language. *BioSystems*, 164, 128-137. <https://doi.org/10.1016/j.biosystems.2017.11.003>
- Tanabe, T., Takahashi, M., & Shudo, K. (2014). A lexicon of multiword expressions for linguistically precise, wide-coverage natural language processing. *Computer Speech and Language*, 28(6), 1317-1339. <https://doi.org/10.1016/j.csl.2013.09.001>
- Wang, S. (2020). *Chinese multiword expressions: Theoretical and practical perspectives*. Gateway East, Singapore: Springer. <https://doi.org/10.1007/978-981-13-8510-0>
- Wray, A. (2002). *Formulaic language and the lexicon*. Cambridge, UK: Cambridge University Press.
- Wray, A. (2008). *Formulaic language: Pushing the boundaries*. Oxford, UK: Oxford University Press.

APPENDIX

Subject A's Compositions

Composition 1

Tessa: Hello?/ I think you came to the wrong room. → S-single + S
 Hardin: I've come to the right place. → S
 Tessa: How did you get in here? → S
 Tessa: Anyway,/ please you go out to the hall/ for I change my clothes.
 → P-single + S + conjunction + S
 Hardin: You think your body is awesome./ I don't watch it. → S + S

Composition 2

Tessa: How about this? → S
 Steph: What?/ It's pretty/. But doesn't it look like a suit? → S-single + S + S
 Tessa: You told me like I am. → S
 Steph: You know/, I like it./ I like it. → S + S + S

Composition 3

Noah: Very loud → S
 Noah: You place in party now/. You drunk? → SC + S
 Tessa: Only one drink → S
 Noah: You entered college, drink. → SC + S-single

Subject B's Compositions

Composition 1

Tessa: Excuse me?/ You've found/ wrong room. → S + SC + P
 Hardin: I've got it right. → S
 Tessa: How did you get in here? → S
 Tessa: Anyway,/ would you go out of here for me to put on my clothes? → P-single + SC
 Hardin: You have that confidence in your body. /I don't look at it. → SC + SC

Composition 2

Tessa: How does it look like? → S
 Steph: What?/ Pretty/. But isn't it like/ kind of a suit? → S-single + S-single + S
 Tessa: You told me/ to live on my way. → S + S
 Steph: Hey/ I like it./ I like it. → S-single + S + S

Composition 3

Noah: It's really noisy. → S
 Noah: You are at the party now./ You've got drunken? → SC + P + P-single
 Tessa: Only one shot. → S
 Noah: You've got/ in the college/ and drink. → P + P + S-single

Subject C's Compositions

Composition 1

Tessa: Excuse me?/ I think you've got the wrong room. → S + S

Hardin: I've got the right room. → S

Tessa: How could you enter here? → SC

Tessa: Anyway/, could you please leave to the hall/ because I have to change my clothes?
→ P-single + S + SC

Hardin: You think your body fits great./ I won't see. → S + SC

Composition 2

Tessa: What do you think? → S

Steph: What?/ Cool./ But seems like too formal,/ isn't it? → S-single + S-single + S

Tessa: You said/ I should be myself. → S + S

Steph: You know what?/ I /kind of/ like it. /I love it. → S + S + S

Composition 3

Noah: It's so loud. → S

Noah: now in party./Are you wasted? → S + S

Tessa: Just one drink. → S

Noah: You can drink/ cause you are in college now. → SC