ISSN 1799-2591
Theory and Practice in Language Studies, Vol. 9, No. 6, pp. 610-619, June 2019
DOI: http://dx.doi.org/10.17507/tpls.0906.02

The Application of Conceptual Metaphors to Teaching English Idioms to English-majored Students in Viet Nam

Pham Thai Bao Ngoc
University of Social Sciences and Humanities, Ho Chi Minh City, Viet Nam

To Minh Thanh
Hoa Sen University, Ho Chi Minh City, Viet Nam

***Abstract*—The discovery that several figurative idioms are semantically motivated by a common conceptual metaphor (CM) has opened up a path to more systematic and insightful learning. However, it was still unclear to what extent the elaboration of conceptual metaphors (CMs) could facilitate learners' reception and production of idioms over time. To address this issue, a quasi-experiment was conducted, with the pre-test – post-test design, on a sample of 69 Vietnamese undergraduates. Results revealed that the explanation of CM was especially beneficial for the students' idiom reception over time, and to a lesser extent for their use of idioms. Though not outstanding in the short term compared with the traditional method, this cognitive approach showed its relatively long-lasting value in terms of both idiom reception and production.**

***Index Terms*— idioms, conceptual metaphors, CM-inspired instruction, reception, production**

I. Introduction

An English native speaker may use approximately 20 million idioms throughout his or her lifetime of 60 years (Cooper, 1998). Due to their pervasiveness, the lack of idiomatic knowledge can be a great hindrance to EFL learners' communication with native speakers and their progress in achieving native-like fluency.

The problem is that English idioms are "notoriously difficult" to learn due to their "rather rigid structure, quite unpredictable meaning and fairly extensive use" (Liu, 2003, p. 671). Besides, idioms are not only cross-linguistic, but also cross-cultural phenomena (Kövecses, 2002). Even a learner with profound knowledge of English grammar and vocabulary still has difficulties in comprehending and using idioms if he or she is not aware of the cultural diversity underlying these expressions.

Despite the importance of learning English idioms and the learners' difficulties when facing them, this area of language is often ignored in EFL classrooms and textbooks. Many teachers in Vietnam even have a tendency to avoid using or teaching idioms in classroom since they believe that idioms are too difficult for learners, which leads to Vietnamese students' poor idiomatic competence (Tran, 2013).

Due to the alleged arbitrary nature of idioms and their fixed structures, rote memorization appeared to be the only way for learners to acquire these expressions, which were believed to be "isolated" and "independent of any conceptual system" (Kövecses, 2002, p. 200). However, since 1980, cognitive linguistics has proved that idioms are not arbitrary; it is, in fact, rooted in human thought and perception (Lakoff & Johnson, 1980). According to the Conceptual Metaphor Theory (CMT), idioms are not dead metaphors and the occurrence of particular words in idiomatic expressions is, to some extent, semantically motivated by conceptual metaphor (CM) – a central concept in Cognitive Linguistics, opening up a new path to more systematic and insightful learning. This finding is an inspiration for the current research, the aims of which can be addressed through the following research questions:

(1) To what extent does the application of conceptual metaphors facilitate English-majored students' reception of the target idioms over time?

(2) To what extent does the application of conceptual metaphors facilitate English-majored students' production of the target idioms over time?

In this study, the application of CMs is proposed as a promising method to present idioms to EFL students, shortly referred to as the CM-inspired instruction as compared to the traditional instruction that encourages rote learning. Due to the limitation of the study, the term "reception" is used to refer to the recognition and comprehension of idiomatic expressions in reading context (rather than listening) and "production" is involved in the recall and use of idioms in writing (rather than speaking).

II. Literature Review

A. Conceptual Metaphor

The primary function of metaphor is to aid the understanding of one concept in terms of another, typically a more abstract and unfamiliar concept (i.e. target domain) in terms of a more concrete, physical and familiar concept (i.e. source domain) (Kövecses, 2002). Owing to its important function, metaphor is found to be "pervasive in everyday life, not just in language but in thought and action" (Lakoff & Johnson, 1980, p. 3). It is, therefore, necessary for EFL learners to build a large repertoire of figurative expressions for comprehension and use. This inevitably leads to the quest for effective methods of teaching and learning idiomatic language.

A conceptual metaphor is defined as a cross-domain mapping, i.e. "a fixed set of ontological correspondences between entities in a source domain and entities in a target domain" and is expressed as TARGET-DOMAIN IS SOURCE-DOMAIN or TARGET-DOMAIN AS SOURCE-DOMAIN, in which capital letters is used as mnemonics to name mappings (Lakoff, 1993, p. 245). These cognitive mappings of metaphors are tightly structured and asymmetric. The following table shows an example of the mapping of ANGER AS HEATED FLUID IN A CONTAINER.

TABLE I.
ONTOLOGICAL CORRESPONDENCES OF ANGER AS HEATED FLUID IN A CONTAINER

ANGER IS HEATED FLUID IN A CONTAINER		Metaphorical expressions	
Source domain: HEATED FLUID IN A CONTAINER	→ Target domain: ANGER		
The container	The body	• He was *filled with* anger.	• She was *brimming with* rage.
The heated fluid	The anger	• His pent-up anger *welled up* inside him.	• Her rudeness *made my blood boil.*
Heat scale	Anger scale	• Let him *stew*.	• Racial tension *reached boiling point.*
Pressure in container	Internal pressure in the body	• He was *bursting with* anger.	• She'll *flip her lid* when she finds out.
Explosion	Loss of control	• I'm sorry I *blew up* at you.	
Coolness in the fluid	Lack of anger	• *Simmer* down!	

(adapted from Kövecses, 1986, pp. 17-18)

It can be seen that a considerable number of idioms can be traced back to a limited number of CMs, forming a coherent system of metaphorical concepts. For instance, all the expressions *flip one's lid*, *reach boiling point* and *simmer down* relate to one single conceptual metaphor ANGER IS HEATED FLUID IN A CONTAINER. Specifically, the container, e.g. a pressure cooker or an enclosed kettle, is our body. The increase in temperature and the rising of the fluid inside the container are associated with the growth in anger intensity. Intense heat creates pressure on the container, corresponding to internal pressure in the body, and an attempt being made to keep the pressure back is similar to how a person tries to control his anger. When the pressure in the container becomes extremely high, the container explodes. Likewise, when anger is too intense, the person loses control and explodes. These reasonable ontological correspondences and elaborations thanks to the discovery of CMs have provided a guarantee for a more systematic and insightful learning of idiomatic expressions.

B. The Application of Conceptual Metaphors to Teaching Idioms

In recent years, there has been growing interest in contrastive analysis of CMs in English and Vietnamese metaphorical and idiomatic expressions (Nguyen, 2015; Nguyen, 2016; Pham, 2016; Dinh, 2017; Ha, 2018). The findings from their research generally support and further develop the CMT, building a sound foundation for the application of CMs to EFL teaching in Vietnamese context. However, so far, there seems to be little empirical research on the pedagogic advantage of the application of CMs to idiomatic acquisition.

Beyond the boundary of Vietnam, several empirical researches have been conducted in order to explore how to employ the cognitive approach effectively to teach English vocabulary and idiomatic language in particular.

The major theme running through several studies is that metaphor awareness can enhance idiomatic competence (Kömür & Çimen, 2009; Vasiljevic, 2011; Doiz & Elizari, 2013; Khoshniyat & Dowlatabadi, 2014; Kartal & Uner, 2017; Pérez, 2018; Chen, 2019). Though the use of CMs is proved to be beneficial to idiom acquisition, their effects over time are still a matter for debate.

In Vasiljevic (2011), the students' idiomatic knowledge was measured in terms of both recognition and production. There were two groups: the experimental group (EG), which received the CM-grounded instruction, and the control group (CG), which was exposed to the traditional method, i.e. memorizing idioms in a list. As regards the receptive knowledge test administered immediately after the instruction and four weeks later, the EG only performed better than the CG when imperfect answers, i.e. "correctly selected idioms" without entirely correct forms, together with perfect answers, were also considered as correct. The result of the productive knowledge test was rather different: the EG significantly outperformed the CG regardless of whether the imperfect answers were counted or not in both immediate and delayed post-tests, except the case of strict marking in the delayed post-test. However, a study by Doiz and Elizari (2013) provided different results from the findings of Vasiljevic (2011). In this study, the EG outperformed the CG in terms of comprehension and short-term retention of idiomatic meaning and form. However, concerning the longer-term effect, the EG did not maintain its superiority. In Kömür & Çimen's (2009), Kartal & Uner (2017), and Chen (2019) the EG outperformed the CG in the acquisition and retention of figurative idiomatic expressions, including phrasal verbs,

but the question about the learners' long-term retention of the taught expressions remained unanswered. In sum, the conflicting findings among previous studies seemed to add to the general confusion about this issue rather than solve it adequately.

Such inconsistent conclusions from prior studies may have resulted from problems in research methodology, including the lack of a delayed post-test (as in Kömür & Çimen, 2009; Kartal & Uner, 2017; and Chen, 2019), no control group (as in Kömür & Çimen, 2009 and Pérez, 2018), inadequate exposure to the CM instruction, i.e. only one or two lessons (as in Doiz & Elizari, 2013 and Chen, 2019). Besides, the treatment in Khoshniyat & Dowlatabadi (2014) included both the explicit teaching of CMs and Disney movies as the manifestation of the CMs, which led to the uncertainty about whether the positive results from the tests were actually due to the combination of CM instruction and Disney movies or simply the fascination of the movies themselves. Moreover, not all CMs can be illustrated with movie scenes, which made the method inapplicable to several cases.

Given the unresolved problems in the research methodology and the inconsistent findings of previous studies, the researcher felt there was a need to improve the research methodology for more valid conclusions about the effect of the cognitive semantic approach over time. Moreover, due to lack of research and practice regarding this issue in the context of Vietnam, a study conducted in Vietnam was of importance to promote the pedagogic application of CMs to teaching idiomatic expressions.

C. Theoretical and Empirical Guidelines for the Study

Based on the findings and suggestions from previous research, this study proposed the following sequence of steps when applying CMs to teaching idioms:

(1) introduce the notion of CM
(2) activate the source domain vocabulary
(3) explain the ontological mappings of CMs
(4) guide learners to apply CMs to interpret idiom meaning
(5) refine or rectify their interpretations

III. Research Methodology

A. Participants

This study was conducted at a public university named Thu Dau Mot University, Binh Duong Province, Viet Nam. 99 English-majored students from two intact classes volunteered to take extra hours to participate in the research. One class was randomly selected as the control group who was instructed in the traditional way, i.e. rote learning, whereas the other was the experimental group who learned the target idioms under the CM-inspired instruction.

In order to guarantee the comparability of the two groups in terms of English proficiency and idiomatic knowledge, the Quick Placement Test (QPT) (UCLES, 2001) and the Idiom Knowledge pre-test were administered to the classes. Among 99 sophomores, only 71 students who reached the intermediate or upper-intermediate level (equivalent to B1 and B2 CEFR) were selected as participants of the study because idioms are multi-word and, in most cases, non-literal fixed expressions, which require learners to have a good command of English to fully understand their figurative meanings (Liu, 2003; Boers, 2013). Throughout the experiment, two students who missed the post-tests were excluded from the study. As a result, the total number of the participants was 69.

The description of the participants is summarized in Table II, which shows that the two groups shared a large number of similarities regarding their age range, gender and English proficiency.

Table II.
Demographic Characteristics Of The Participants

		Control group (CG)		Experimental group (EG)	
		Raw count	Percentage	Raw count	Percentage
Number of students		34	100%	35	100%
Gender	Male	5	15.6%	5	14.7%
	Female	29	84.4%	30	85.3%
Age	20	30	87.5%	32	91.2%
	21	4	12.5%	3	8.8%
English Proficiency	B1	25	75%	28	79.4%
	B2	9	25%	7	20.6%

The two groups were also very similar in terms of idiomatic knowledge before the treatment. In the Idiom Receptive Knowledge pre-test, the Independent T-test showed that there was no significant difference (t = .362, df = 67, two-tailed p = .719> 0.05) between the mean scores of the CG and the EG (Table VII). Likewise, in the Idiom Productive Knowledge pre-test, no statistically significant difference was found, via the Mann-Whitney U Test, in the two groups' scores irrespective of marking methods (U = 577.000, $N1$ = 34, $N2$ = 35, two-tailed p = .541> 0.05) (Table III) as the students could hardly deal with any question in the test before the teaching instruction.

TABLE III.
RESULTS OF THE MANN-WHITNEY U TEST IN THE PRODUCTIVE IDIOM KNOWLEDGE PRE-TEST

	Mann-Whitney U	Wilcoxon W	Z	Asymp. Sig. (2-tailed)
Strict marking	577.000	1207.000	-.612	.541
Less strict marking	577.000	1207.000	-.612	.541

B. Teaching Materials

Ten CMs, i.e. HAPPY IS UP and SAD IS DOWN, ANGER IS HEATED FLUID IN A CONTAINER and ANGER IS FIRE, LIFE IS A JOURNEY and LIFE IS A GAMBLING GAME, RELATIONSHIP IS A JOURNEY and RELATIONSHIP IS A LIVING ORGANISM, BUSINESS IS WAR and AN ORGANIZATION IS A GARDEN, were selected for the study, together with 60 idiomatic expressions, each of which was motivated by one of the CMs above. In each lesson, the EG was introduced to 2 CMs related to one semantic topic, e.g. happiness/sadness, anger, life, etc.

Rather than teaching all the target idioms within one or two intensive lessons as previous research, the experiment of this study was spread out over five lessons so that the students were not overloaded with such heavy memory work. In each lesson, the two groups were introduced to 11-13 idioms related to one semantic topic. The handouts of each lesson were distributed to the participants at each meeting. There were three sections in their handouts, including (1) a short story or extracts from newspapers, books, etc. that provided contexts and illustrated the use of most of the target idioms in the lesson, (2) the explanation of the idioms that appeared in the story or extracts as well as other expressions of the same topic, and (3) controlled and freer practice of the target idioms. The first and the third sections in the handouts were the same for two groups; the only difference was found in the second section, concerning the inclusion of CM-inspired instruction and the organization of the target idioms.

Unlike the EG which received the CM-inspired instruction, following a sequence of 5 steps, the CG inferred the meaning of the new idioms by using the contexts given in the reading texts, and then got correction from the instructor. In addition, while the CG had the target idioms organized in alphabetical order in the list of new phrases and was given time to memorize them, the EG learned the idioms in two subsets, categorized by CM and was encouraged to apply CM to interpret and elaborate the meanings of the target expressions.

C. Research Instrument

The Idiom Knowledge Test was designed to measure the subjects' knowledge, including receptive knowledge and productive knowledge, of the idioms taught in the experiment.

TABLE IV
SUMMARY OF ITEM TYPES IN THE IDIOM KNOWLEDGE TEST

Idiom Knowledge Test		**Item type (in each sub-test)**	**Section**	**Item sub-type (in each section)**
Sub-test 1	RECEPTIVE KNOWLEDGE TEST	30 selected-response items	Section A	20 multiple-choice items
			Section B	10 gap-filling items (with selection from banks)
Sub-test 2	PRODUCTIVE KNOWLEDGE TEST	30 constructed-response items	Section C	15 gap-filling items (with no bank)
			Section D	15 sentence transformation items

The validity of the test was confirmed by the supervisor and advisor of the researcher. To ensure its reliability, the test was piloted on 25 students of the same population, using the test-retest method. The correlation coefficient of .805 ascertained the high reliability of the test. The internal consistency among the test items was also evaluated with the alpha coefficient of reliability. The Cronbach's alpha (α) for each subtest and the average alpha for the whole test was all above .800, which were above the borderline of acceptability. Because of its stability over time and high internal consistency, this test was used as pre-test and post-tests in the main study.

The total score of the test was 60 for 60 items, including 30 for receptive knowledge test and 30 for productive knowledge test. In order to ensure the consistency in scoring, strict grading was a requisite. However, an initial examination of the data in the productive knowledge post-tests showed that a large number of the students' responses were mostly, but not completely, correct. They managed to provide the right targeted keywords or content words but had a tendency to make a mistake in using articles and possessive adjectives, and confusing singular with plural noun form. These mistakes were taken into special consideration in this study as such mistakes did not seem to be serious as they were grammatical rather than lexical mistakes. Besides, learning is a process and learners are inclined to make several mistakes before they can become a competent language user. Thus, this study proposed two ways of marking for the productive knowledge test: (1) considering only completely correct responses and (2) including idioms which did not have an entirely correct form due to unserious grammatical mistakes.

D. Procedure

The following steps were taken to collect the necessary data. First, the students in both groups took the QPT and the pre-test to ensure their homogeneity in terms of English proficiency and idiomatic competence. Then each group went through five lessons during five weeks in a row. There was one meeting each week, and the lesson about idiomatic language last about 60-75 minutes. After five weeks of learning idioms, the students did the immediate post-test, including two subtests, for measuring their reception and production of the taught idioms, within 60 minutes. Then, on

the fifth week after the immediate post-test, without being informed in advance, the subjects were required to take the second post-test.

Regarding the data analysis procedure, in order to find out whether the CM-inspired instruction had a significant effect on the students' receptive and productive knowledge of idioms over time, as stated in Research Question 1 and 2, a Repeated-measures ANOVA was expectedly conducted on the mean scores of the three test administrations, i.e. pre-test, post-test 1 and post-test 2, for within-group comparison. The evaluation of the effects of the application of CMs would have been incomprehensive without between-group comparison. Specifically, the Independent-Samples T Test was expected to be performed to examine the mean scores of the EG and the CG, to find out which type of instruction was more effective in facilitating the students' idiomatic knowledge.

An initial analysis of the data revealed that not all the data in this study were normally distributed. Specifically, the data in the Productive Idiom Knowledge Pre-test were found to deviate from the normal distribution. Thus, the Friedman Test and the Mann-Whitney U Test, which do not rely on the data with a normal distribution, were performed (instead of the Repeated-measures ANOVA and the Independent-Samples T Tests) whenever the statistical tests had the scores from the Productive Idiom Knowledge Pre-test as part of their data.

IV. Results

A. Results of the Receptive Idiom Knowledge Test over Time

To examine whether the CM-inspired instruction facilitated the student subjects' reception of the target idioms over time, within-group and between-group comparison were made via Repeated-measures ANOVAs and Independent Samples T-Tests.

Descriptive statistics of all three test administrations and the results of the within-group comparison were presented in Tables V, VI and VII.

TABLE V.
DESCRIPTIVE STATISTICS OF THE RECEPTIVE IDIOM KNOWLEDGE TEST IN 3 TEST ADMINISTRATIONS

Group	N	Pre-test		Post-test 1		Post-test 2	
		Mean	Std. Deviation	Mean	Std. Deviation	Mean	Std. Deviation
CG	34	5.15	3.295	24.38	3.660	18.00	3.861
EG	35	4.86	3.362	25.20	3.376	23.91	3.364

The total score of the Receptive Idiom Knowledge Test is 30.

TABLE VI.
RESULTS OF THE REPEATED-MEASURES ANOVA FOR THE RECEPTIVE IDIOM KNOWLEDGE TEST IN 3 TEST ADMINISTRATIONS

	Source		Type III Sum of Squares	df	Mean Square	F	Sig.	Partial Eta Squared
CG	time	Sphericity Assumed	6527.196	2	3263.598	4474.652	.000	.993
	Error(time)		48.137	66	.729			
EG	time	Sphericity Assumed	9084.362	2	4542.181	12047.241	.000	.997
	Error(time)		25.638	68	.377			

Table VI revealed that the scores at three different testing stages were statistically different in the case of the CG ($F(2,66) = 4474.652$, $p = .000 <. 05$) and the EG ($F(2,68) = 12047.241$, $p = .000 <. 05$). Overall, the scores of both groups followed a similar pattern. As compared to the pre-test results, the participants in the CG and EG showed considerable improvement in their reception of the target idioms in both post-tests. However, five weeks of inexposure to the instruction caused a marked decline in both groups' retention of idiom meaning as compared to the immediate post-test, though the mean scores of the delayed post-test were still significantly higher than those of the pre-test ($p = .000 <. 05$), as shown in the Bonferroni post-hoc test result (Table VII).

TABLE VII.
PAIRWISE COMPARISONS FOR THREE ADMINISTRATIONS OF THE RECEPTIVE IDIOM KNOWLEDGE TEST

(I) Time	(J) Time	Mean Difference (I-J)		Std. Error		Sig.	
		CG	EG	CG	EG	CG	EG
1	2	-19.235	-20.343	.184	.147	.000	.000
	3	-12.853	-19.057	.232	.164	.000	.000
2	1	19.235	20.343	.184	.147	.000	.000
	3	6.382	1.286	.203	.127	.000	.000
3	1	12.853	19.057	.232	.164	.000	.000
	2	-6.382	-1.286	.203	.127	.000	.000

Notes: The mean difference is significant at the .05 level.
1 represents the first test administration (pre-test), 2 the second test administration (post-test 1), and 3 the third test administration (post-test 2).

To find out whether the result gained by the CG was statistically different from that of the EG in each test administration, the Independent Samples T-Test was run for between-group comparison at each time point, as summarized in Table VIII.

TABLE VIII
RESULTS OF THE INDEPENDENT SAMPLES T-TEST FOR THE RECEPTIVE IDIOM KNOWLEDGE TEST IN 3 TEST ADMINISTRATIONS

Receptive Knowledge Test			Equal variances assumed	Equal variances not assumed
Pre-test	Levene's Test for Equality of Variances	F	.036	
		Sig.	.849	
	t-test for Equality of Means	t	.362	.362
		df	67	66.994
		Sig. (2-tailed)	.719	.719
Post-test 1	Levene's Test for Equality of Variances	F	.059	
		Sig.	.809	
	t-test for Equality of Means	t	-.965	-.964
		df	67	66.202
		Sig. (2-tailed)	.338	.339
Post-test 2	Levene's Test for Equality of Variances	F	1.045	
		Sig.	.310	
	t-test for Equality of Means	t	-6.790	-6.776
		df	67	65.200
		Sig. (2-tailed)	.000	.000

Results indicate that the mean score of the CG (M = 24.38, SD = 3.660) was not significantly higher (t = -.965, df = 67, two-tailed p = .338 > .05) than that of the EG (M = 25.20, SD = 3.376). In other words, the two types of instruction seemed to exert similar effect on the students' acquisition of idiom meanings immediately after the teaching stage. However, 5 weeks after the instruction, the EG (M = 23.91, SD = 3.364) gained a significantly higher score (t = -6.790, df = 67, two-tailed p = .000< .05) than the CG (M = 18.00, SD = 3.861). Moreover, the smaller standard deviation of the EG (3.364) shown in Table V suggests that the scores of this group were more homogeneous and more consistent than that of the CG (3.861). In short, these results indicate that the CM-instruction offered more long-term benefit and yielded more consistent results than the traditional instruction in terms of idiom reception.

B. Results of the Productive Idiom Knowledge Test over Time

To investigate whether the CM-inspired instruction facilitated the students' production of the target idioms over time, within-group and between-group comparison were made via Friedman Test and Independent Samples T-Tests.

Descriptive statistics of all three test administrations and the results of the within-group comparison were presented in Tables IX, X and XI.

TABLE IX
DESCRIPTIVE STATISTICS OF THE PRODUCTIVE IDIOM KNOWLEDGE TEST IN 3 TEST ADMINISTRATIONS

			Idiom Productive Knowledge Test					
			Pre-test		Post-test 1		Post-test 2	
			Strict marking	Less strict marking	Strict marking	Less strict marking	Strict marking	Less strict marking
CG	Mean		.06	.06	21.91	23.03	13.85	14.53
	Std. Deviation		.239	.239	3.980	3.896	3.669	4.407
	Percentiles	25th	.00	.00	20.00	21.00	11.75	12.00
		50th (Median)	.00	.00	22.00	24.00	15.00	15.00
		75th	.00	.00	25.00	25.50	16.25	18.25
EG	Mean		.03	.03	21.29	23.37	15.86	18.29
	Std. Deviation		.169	.169	3.494	3.255	2.982	3.885
	Percentiles	25th	.00	.00	19.00	21.00	14.00	16.00
		50th (Median)	.00	.00	21.00	24.00	16.00	18.00
		75th	.00	.00	24.00	26.00	18.00	21.00

The total score of the Productive Idiom Knowledge Test is 30.

TABLE X
RESULTS OF THE FRIEDMAN TEST ON THE RESULTS OF THE PRODUCTIVE IDIOM KNOWLEDGE TEST IN 3 TEST ADMINISTRATIONS

		Strict marking	Less strict marking
CG	N	34	34
	Chi-Square	68.000	68.000
	df	2	2
	Asymp. Sig.	.000	.000
EG	N	35	35
	Chi-Square	70.000	70.000
	df	2	2
	Asymp. Sig.	.000	.000

Results indicate that there was an overall statistically significant difference among the mean ranks of the CG ($\chi^2(2)$ = 68.000, p = 0.000) and among the mean ranks of the EG ($\chi^2(2)$ = 70.000, p = .000) in the three testing administrations, irrespective of the marking methods. As the Friedman test does not pinpoint exactly where those differences lie, post hoc tests were run to find out which specific pairs of means differed, as follows.

TABLE XI
RESULTS OF THE WILCOXON SIGNED-RANK TESTS ON THE PRODUCTIVE IDIOM KNOWLEDGE TEST OVER TIME

		Strict marking			Less strict marking		
		Post-test 1 – Pre-test	Post-test 2 – Pre-test	Post-test 2 – Post-test 1	Post-test 1 – Pre-test	Post-test 2 – Pre-test	Post-test 2 – Post-test 1
CG	Z	-5.092	-5.096	-5.113	-5.096	-5.094	-5.126
	Asymp. Sig. (2-tailed)	.000	.000	.000	.000	.000	.000
EG	Z	-5.165	-5.167	-5.374	-5.167	-5.164	-5.201
	Asymp. Sig. (2-tailed)	.000	.000	.000	.000	.000	.000

As can be seen from Table XI, the *p*-value of each pair in the case of each group was below .05, irrespective of marking procedure. Thus, it can be concluded that, as regards the Productive Knowledge Test, the scores of three test administrations for the EG were significantly different from one another. A similar conclusion can be drawn for the CG.

To clarify which type of instruction had a statistically greater effect on the learners' performance, the Independent Samples T-Test was run for between-group comparison at each time point, as summarized in Table XII.

TABLE XII
RESULTS OF THE INDEPENDENT SAMPLES T-TEST FOR THE PRODUCTIVE IDIOM KNOWLEDGE POST-TESTS

Productive Idiom Knowledge Test			Strict marking		Less strict marking	
			Equal variances assumed	Equal variances not assumed	Equal variances assumed	Equal variances not assumed
Post-test 1	Levene's Test for Equality of Variances	F	.223		.413	
		Sig.	.638		.523	
	t-test for Equality of Means	t	.695	.694	-.396	-.396
		df	67	65.356	67	64.249
		Sig. (2-tailed)	.490	.490	.693	.694
Post-test 2	Levene's Test for Equality of Variances	F	1.483		.243	
		Sig.	.228		.624	
	t-test for Equality of Means	t	-2.493	-2.486	-3.806	-3.800
		df	67	63.541	67	65.848
		Sig. (2-tailed)	.015	.016	.000	.000

Results reveal that, in the case of strict marking, there was no statistical difference (t = .695, df = 67, two-tailed p = .490> .05) between the mean score of the CG (M = 21.91, SD = 3.980) and that of the EG (M = 21.29, SD = 3.494). The same conclusion was drawn for less strict marking as the p-value (.693) was also higher than .05. All things considered, the two types of idiom instruction seemed to have a similar impact on the students' productive knowledge of idioms right after a 5-week treatment.

The single most striking observation to emerge from the data comparison is that a more lenient way of marking could bring more points for the EG and even reverse the situation in the immediate post-test. When only complete idioms were coded as right answers, the CG mean score (M = 21.91, SD = 3.980) was slightly higher than that of the EG (M = 21.29, SD = 3.494). However, when a less strict way of marking was applied and those that did not have an entirely correct form were also considered as correct, the EG was at the advantage with its mean score (M = 23.37, SD = 3.255) even higher than the CG score (M = 23.03, SD = 3.896). A closer look at the answers of the students to the Productive Idiom Knowledge Test discloses that, in some cases, the EG students, though unable to recall the whole idiom, managed to provide the right keywords of the target idiom and gained the score for that item. Meanwhile, some in the CG remembered the wrong keywords or just left the space blank and got no point for that even when a more lenient scoring system was adopted. Further analysis is given to the standard deviation. Though the EG mean was higher, its standard deviation (3.255) was smaller than that of the CG (3.896). This finding reveals that the scores of the EG were more tightly clustered around the mean and, therefore, more homogeneous. For these reasons, the arguments in support of the CM-inspired instruction were more valid and strong.

Unlike the results from the immediate post-test in which the CM's effect was not particularly outstanding, there was a statistically significant difference (strict marking: t = -2.493, df = 67, two-tailed p = .015<.05; less strict marking: t = -3.806, df = 67, two-tailed p = .000<.05) between the mean score of the CG (strict marking: M = 13.85, SD = 3.669; less strict marking: M = 14.53, SD = 4.407) and that of the EG (strict marking: M = 15.86, SD = 2.982; less strict marking: M = 18.29; SD = 3.885) in the delayed post-test. In two cases of scoring, the higher means justified the superiority of the EG over the CG regarding long-term benefit of the application of CMs to teaching idioms.

Briefly, as regards the second research question, it can be concluded that the CM-inspired instruction had a considerable effect on the students' production of the target idioms over time. Though it seemed not to be superior to the traditional instruction immediately after the treatment, the CM-inspired instruction turned out to be outstanding in the long term.

V. DISCUSSION

A. Idiom Reception over Time

Overall, these findings are in agreement with a great deal of previous work in this field, which proved that the awareness of these underlying metaphors can facilitate the learner's comprehension and retention of idiomatic meaning. However, a closer examination reveals that there are some important differences in the level of effectiveness of the CM-inspired instruction and its benefits over time among the studies.

The similarity of the two groups' results in the immediate post-testing stage is in contradiction with Vasiljevic (2011). Specifically, in the immediate post-test, her analysis indicated a significant difference between the two conditions, whereas no statistically significant difference was found between the two groups' results in this study. The contrast in the findings of the two studies might stem from the difference in the presentation of the target idioms for the CG. In Vasiljevic (2011)'s research, the EG learned idioms organized along three metaphoric themes, i.e. IDEAS ARE FOOD, LOVE IS A JOURNEY, and LIFE IS A GAMBLING GAME; one theme was taught each week. Unlike the CM group, the CG were exposed to these idioms in a mixed order, which means each week they had to learn idioms of the three topics (IDEAS, LOVE and LIFE) at the same time and all of them were not categorized even into topics. It was apparent that the CG was at a great disadvantage due to such disorganization. Thus, any superior learning effects under the experimental condition could be merely the result of the idioms being grouped for the EG and the mere ineffectiveness of the comparison treatment for the CG.

Different from Vasiljevic (2011)'s research, the CG in the current study learned idioms organized under topic headings, and, like the EG, they only learned about one topic in each lesson. This ensured the input given to the CG was, to some extent, systematized. Such lexis grouping allowed a meaningful comparison between the effects of the traditional treatment with the experimental treatment. As a result, both types of idiom instruction were found to be quite effective in facilitating the students' idiom reception right after the instruction, though it should be emphasized that the CM-inspired instruction showed superior strength in the longer term.

These findings were also in contrast with Doiz and Elizari (2013). Specifically, they found that the EG obtained a significantly higher score mean than the CG in the immediate post-test, while there was no significant difference between two groups' performance in the delayed post-test. Such a contradiction between the two studies may be due to the difference in the experiment duration. Whereas the teaching stage in Doiz and Elizari (2013) last within only one lesson, the instruction in the current study ran over 5 lessons during 5 weeks. The experiment duration in Doiz and Elizari (2013) seems to be relatively short for the researchers to draw a conclusion about the effect of CMs. As the concept of CM was entirely new to the participants, it might be quite demanding for them to comprehend and employ it to learn the target idioms within only one lesson. In other words, a greater time length for the experiment seems to be a necessary condition for the students' familiarization with the treatment so that they could get the full benefit from it.

There are four main reasons for the success of this approach. Firstly, thanks to the teaching of CMs, the students had the chance to learn idioms in a more systematic way thanks to the metaphorical sub-themes in each topic. Secondly, the CM-inspired instruction raised the students' awareness of the semantic motivation behind the target expressions; therefore, they viewed these expressions as meaningful parts of certain structured networks rather than rigid and isolated pieces of language. Thirdly, the cognitive approach in teaching idioms assisted the subjects to create mental images, allowing dual coding of information. Since CMs are grounded in bodily experience and in cultural and social practices (Kövecses, 2002), the explicit instruction of these metaphors could possibly stimulate the learners' visualization of the input and improve their comprehension and memory. Lastly, the students' mental processing of the information at a deeper level was facilitated thanks to the employment of CMs. During the instruction, the students were encouraged to activate their prior knowledge about a familiar, concrete or physical concept to understand an unfamiliar and abstract concept, and simultaneously associate the verbal information with a mental image. This process provided a stimulus for the input to be converted into intake which was then stored in the learners' memory.

B. Idiom Production over Time

In comparison with the Receptive Idiom Knowledge Test, both groups' performance in the Productive Idiom Knowledge Test seemed to be less satisfactory. It seems that the success obtained through CM-inspired instruction was less pronounced in the case of idiom production. In fact, this result is not surprising as the production test was more demanding than the reception test. Since idioms are fixed in their structures, the students in both groups faced considerable difficulty in recalling the precise composition of those multi-word expressions, though several students could identify the right idioms for the given situations, which means the students could understand more idioms than they could really use.

The most striking result to emerge from the data of this study is the relatively long-lasting effect of the CM-inspired instruction five weeks after the treatment. Though the number of idioms that they were able to recollect inevitably decreased over time, the EG was still at a significant advantage over the CG in both scoring procedures. A closer look at the students' answers to the Productive Idiom Knowledge Test discloses that a large number of the EG's errors were involved in the wrong use of articles, possessive adjectives, and prepositions as well as the wrong addition or omission of the noun plural suffix. This may be due to the semantic orientation of the CM-inspired instruction as it tends to draw the students' attention to the motivation behind the meaning of the idiomatic expressions rather than their precise lexical make-up. Nevertheless, thanks to insightful learning, the EG correctly recalled more content words than the CG, as

evidenced by the higher mean score of the EG in both the immediate post-test and delayed post-test when a more lenient scoring procedure was adopted.

The universality of CMs and the role of the participants' first language (L1) were also of importance in the in-depth analysis of the students' responses. As several CMs taught in this experiment exist in the participants' L1, the activation of these CMs might facilitate L1 transfer. This influence may have usefully accelerated the learning process, as in the case of HAPPY IS UP / SAD IS DOWN (*with a heavy heart = với cõi lòng nặng trĩu, raise your spirits = lên tinh thần*). Nevertheless, L1 transfer is also a common source of errors. Some idioms have very close equivalents which share the same CMs, but have a slight difference in the surface structure; it is this minor difference in wording that may have led to the wrong substitution of keywords. Take the expression *to add fuel to the fire* as an example. Its translation term in Vietnamese is *thêm dầu vào lửa (to add oil to the fire)*, which is also based on ANGER IS FIRE as its English CM equivalent. Perhaps due to L1 interference, 18 EG students and 11 CG students substituted *fuel* with *oil* in the delayed post-test. More EG students made this transfer error probably because the CM-inspired instruction tended to draw the students' attention to the motivation behind each idiomatic expression rather than rote learning the precise idiomatic form. Since the EG were encouraged to activate their prior knowledge to analyze the new target idioms, several of them had recourse to their L1 perception and tended to translate directly from L1 to the target language when they needed to recollect the form of the idioms for production. However, this does not mean that the traditional instruction was more effective over time in this case. Actually, more students in the CG were unable to recall this idiom; they tended to give no answer or replace it with a non-idiomatic phrase. Overall, despite L1 interference in some cases, the CM-inspired instruction showed its relative effectiveness over time in facilitating the students' recollection of the form of the idioms, though partial in certain cases, thanks to the use of CMs as a mnemonic device.

To foster the retention of the exact form of the target idioms, more activities aimed at structural elaboration must be integrated with the CM-inspired instruction. According to Boers and Lindstromberg (2008), the lexical composition of several (semi-)fixed phrases is explicable in terms of alliteration and assonance. In fact, some common mistakes in word selection made by the students in this study can be rectified via this technique. For example, it is hard to explain, via semantic elaboration, why the idiom is *add fuel to the fire* rather than *add oil to the fire* as in the Vietnamese equivalent, or *flip your lid* rather than *flip your top* or *flip your roof* though "lid," "top" and "roof" are nearly synonymous. Phonological motivation may be helpful in these cases as in word combinations "euphonious word strings are preferred over same-meaning but non-euphonious word strings" (ibid., p. 330). Thus, the teacher can explain that "fuel" was selected rather than "oil" in the idiom *add **f**uel to the **f**ire* because it alliterates with the word "fire" at the end of the phrase; likewise, in *fl**i**p your l**i**d*, "flip" rhymes with "lid" and, therefore, sound pleasant when standing together. Similar cases can be found in several idioms taught in this study: ***f**an the **f**lames, **b**ite the **b**ullet, **g**ain **g**round on, **s**et your **s**ights on*, etc. (alliteration); *s**e**v**e**nth h**ea**v**e**n, a d**ea**d **e**nd, d**o**wn in the m**ou**th, a fl**a**sh in the p**a**n*, etc. (assonance). Hopefully, this technique can help the students better understand the motivation, i.e. both semantic and phonological, behind the target idioms and later effectively recall the precise components in these idioms for production.

VI. Conclusion

It can be concluded that the application of CM to teaching idioms could have significant effects on the students' reception and production of idiomatic language over time. Though its effectiveness might not be particularly outstanding relative to the traditional instruction in the short term, the CM-inspired instruction is likely to provide a relatively substantial long-term benefit for the students. However, every method has its own limitations, and this type of instruction is not an exception. To overcome its shortcomings, structural elaboration is proposed to be used concurrently with the CM-inspired instruction.

The fact that CMs can be used as a strategy to learn idiomatic language may have radical implications for EFL learners, teachers as well as textbook designers. By explaining the motivation behind several idiomatic expressions, the CM-inspired instruction indeed brings many problems concerning idioms to light and helps relieve students' burden of rote learning. With the knowledge of CMs, the learner no longer has to depend upon only one single channel of acquiring idioms, i.e. encountering idioms in context and blindly memorizing them. In fact, a variety of learning activities can be designed with CMs as an organizer and motivator of the target idioms for systematic and insightful learning.

It is also noteworthy that the application of CMs should be considered as part of a learning program and be integrated with other approaches to teach vocabulary in general and idioms in particular. Rather than adopting a single method in teaching idioms, EFL teachers can use a variety of techniques in order to enhance their students' idiomatic knowledge and inspire them with innovative activities, and one of the best options is the employment of CMs. EFL syllabus designers as well as textbook writers can include CMs in their language learning program as a viable method to develop the learner's metaphorical competence in general and idiomatic competence in particular.

References

[1] Boers, F. (2013). Cognitive semantic ways of teaching figurative phrases: An assessment. In F. Gonzalvez-Garcia, M. S. Pena Cervel, & L. Perez Hernandez (eds.), *Metaphor and Metonymy revisited beyond the Contemporary Theory of Metaphor*. Amsterdam: John Benjamins, 229–263.

[2] Boers, F., & Lindstromberg, S. (2008). Structural elaboration by the sound (and feel) of it. In F. Boers & S. Lindstromberg (eds.), *Applications of cognitive linguistics: cognitive linguistic approaches to teaching vocabulary and phraseology*. Berlin: Mouton de Gruyter, 329–354.
[3] Cooper, C.T. (1998). Teaching idioms. *Foreign Language Annals* 31.2, 255–266.
[4] Chen, Y. (2019). Teaching figurative language to EFL learners: an evaluation of metaphoric mapping instruction. *The Language Learning Journal*, 47.1, 49–63.
[5] Dinh, T. M. A. (2017). An investigation of conceptual metaphors denoting "life" in American and Vietnamese short stories. *Journal of Development Research* 1.1, 29–35.
[6] Doiz, A. & Elizari, C. (2013). Metaphoric competence and the acquisition of figurative vocabulary in foreign language learning. *Estudios ingleses de lingüística aplicada (ELIA) [English Studies of Applied Linguistics]* 13, 47-82.
[7] Ellis, R. (1985). Understanding second language acquisition. Oxford: Oxford University Press.
[8] Ha, L. H. H (2018). A contrastive study of conceptual metaphor of love in English and Vietnamese songs in the 20th century. M.A. dissertation, The University of Da Nang.
[9] Kartal, G. & Uner, S. (2017). The effects of conceptual metaphors on the acquisition of phrasal verbs by Turkish EFL learners. *European Journal of Foreign Language Teaching* 2.2, 34–51.
[10] Khoshniyat, A. S. & Dowlatabadi, H. R. (2014). Using Conceptual Metaphors Manifested in Disney Movies to Teach English Idiomatic Expressions to Young Iranian EFL Learners. *Procedia-Social and Behavioral Sciences* 98.1, 999-1008.
[11] Kömür, S., & Çimen, S. S. (2009). Using conceptual metaphors in teaching idioms in a foreign language context. *Sosyal Bilimler Enstitüsü Dergisi (İLKE) [Journal of the Institute of Social Sciences (Policy)]* 23, 205-222.
[12] Kövecses, Z. (1986). Metaphors of anger, pride, and love: a lexical approach to the structure of concepts. Amsterdam: John Benjamins.
[13] Kövecses, Z. (2002). Metaphor: A practical introduction. Oxford: Oxford University Press.
[14] Liu, D. (2003). The most frequently used spoken American English idioms: a corpus analysis and its implications. *TESOL Quarterly* 37.4, 671–700.
[15] Lakoff, G. (1993). The contemporary theory of metaphor. In A. Ortony (ed.), Metaphor and thought (2nd edn.). Cambridge: Cambridge University Press, 202–251.
[16] Lakoff, G. & Johnson, M. (1980). Metaphor we live by. Chicago: University of Chicago Press.
[17] Nguyen, T. N. N. (2015). Ẩn dụ trong văn bản diễn thuyết chính trị Mỹ và dịch ẩn dụ từ tiếng Anh sang tiếng Việt [Metaphors in American political discourse and translating metaphors from English to Vietnamese]. Ph.D. dissertation, University of Social Sciences and Humanities_Ho Chi Minh City.
[18] Nguyen, V. T. (2016). The cross-cultural analysis of the metaphorical conceptualization of happiness in English and Vietnamese. *Review of Cognitive Linguistics* 14.2, 275–302.
[19] Pham, H. P. Q. (2016). A contrastive analysis of the conceptual metaphor "IDEAS ARE FOOD" in English and Vietnamese. *Journal of Science_Ho Chi Minh City University of Education* 80.2, 12–21.
[20] Pérez, R. G. (2018). The development of a metaphoric competence: A didactic proposal of educational innovation. *Innovation in Language Learning and Teaching*, 1–27.
[21] Tran, H. Q. (2013). Figurative idiomatic competence: An analysis of EFL learners in Vietnam. *Language Education in Asia* 4.1, 23–38.
[22] UCLES (University of Cambridge Local Examinations Syndicate Quick Placement Test) (2001). Oxford: Oxford University Press.
[23] Vasiljevic, Z. (2011). Using conceptual metaphors and L1 definitions in teaching idioms to non-native speakers. *The Journal of Asia TEFL* 8.3, 135-160.

Pham Thai Bao Ngoc received a Master's degree in TESOL from the University of Social Sciences and Humanities, Ho Chi Minh City, Viet Nam in 2016 and is currently a lecturer at the Faculty of English Linguistics and Literature at the same university. Her research interests include English language teaching and English linguistics.

To Minh Thanh received a Master's degree in Bilingual Education from the University of Massachusetts, the USA, in 1998 and a PhD's degree in Contrastive Linguistics from the University of Social Sciences and Humanities, Ho Chi Minh City, Viet Nam in 2006. She was promoted to Associate Professor in Linguistics in 2012 and was awarded with the title of Meritorious Teacher by the President of Viet Nam in 2014 for her outstanding scientific research achievements and significant contribution to the educational development of the country. She is currently teaching at Hoa Sen University, and her research interests include Applied Linguistics, Contrastive Linguistics, TESOL, and Measurement and Assessment in Higher Education.