Sergio Conti* and Carmen Lepadat

Situation-bound utterances in Chinese as a foreign language: the effectiveness of the identification task

Abstract: This paper investigates the effectiveness of the identification task on the retention of situation-bound utterances (SBUs) in Chinese as a foreign language (CFL). The participants were Italian CFL learners with different lengths of learning experience, divided into an experimental and a control group. The target SBUs were selected by means of a discourse completion questionnaire previously submitted to Italian and Chinese native speakers and Italian CFL learners. During the treatment, the experimental group was provided with six communicative functions and was asked to identify the corresponding SBUs from eight short dialogues. Subsequently, a posttest was administered to both the experimental and the control groups. Overall, the statistical analysis of the test results suggests a significant effect of the task on the participants' ability to recall the target SBUs, both from an immediate and a longitudinal perspective. However, the data also suggest that, while the participants' length of learning experience and minor linguistic or cultural specificities of the target formulae did not significantly influence the effectiveness of this specific vocabulary learning task, the presence vs. absence of the SBUs in either the L1 or the L2 might be a crucial factor for their successful recall.

Keywords: form recall; formulaicity; identification task; Italian CFL learners situation-bound utterance; vocabulary learning

提要：本文探讨了一个词汇学习任务——识别任务在对外汉语中对于情境专用语(situation-bound utterances, 简称SBU)记忆的有效性。参加实验的为本科二至三年级的意大利汉语学习者,分为实验组和对照组。目标SBU是通过之前提交给意大利语母语者、汉语母语者及意大利汉语学习者的言谈情境填充问卷(discourse completion questionnaire)来选取的。在实验处理的过程中,实验组被提供了6个特定情境,并被要求从8个短对话中找出相应的SBU。随后,我们对实验

*Corresponding author: Sergio Conti, Department of Foreign Languages, Literatures and Cultures, Roma Tre University, via del Valco di S. Paolo 19, 00146, Rome, Italy, E-mail: sergio.conti@uniroma3.it
Carmen Lepadat, Italian Institute of Oriental Studies Department, Sapienza University of Rome, Rome, Italy, E-mail: carmen.lepadat@uniroma1.it

组和对照组都进行了两次后测。测试结果的统计分析表明,无论从即时角度还是从纵向角度来看,该任务对参与者回忆目标SBU的能力都有显著影响。数据也表明,参与者的学习时间长短或目标SBU的语种特殊性 (linguistic specificity)对这一特定词汇学习任务的效力没有显著影响,L1或 L2中某种SBU存在与否,即文化特殊性(cultural specificity),可能才是学习者能够成功地回忆出该SBU的关键因素。

关键词**:** 情景专用语; 识别任务; 语形回忆; 词汇学习; 意大利汉语学习者; 语块性

1 Introduction

According to Kecskés's (2000a, 2010) definition, situation-bound utterances (SBUs) are highly conventionalized, prefabricated pragmatic units whose occurrence is tied to standardized communicative situations. SBUs thus constitute a subtype of formulaic sequences (FSs), which are defined by Wray (2000, 2002) as a sequence, continuous or discontinuous, of words or other elements, which is, or appears to be, prefabricated. The use of SBUs has been observed in many natural languages, including Mandarin Chinese (Kecskés 2016; Zhou 2012). Like FSs, the acquisition of SBUs plays an important role in second (L2) and foreign language (FL), but at the same time it is also very challenging for learners. The issue of FS teaching and learning has been widely investigated in the last decades (see Boers and Lindstromberg 2012; Siyanova-Chanturia and Pellicer-Sánchez 2019). However, the research on the acquisition of SBUs is still at an early stage, especially in the field of Chinese as a foreign language (CFL). The present study is a preliminary attempt to shed more light on this issue. It investigates the effectiveness of a specific vocabulary learning task, an identification task, on the retention of Chinese SBUs with different degrees of cultural specificity. The participants were Italian CFL learners with different lengths of learning experience.

2 Literature review

2.1 FSs and SBUs: background

After Sinclair (1991) proposed his famous language processing model consisting of two principles, the open-choice principle and the idiom principle, the interest of linguists in formulaicity has increased constantly. According to Sinclair's model, the idiom principle consists of the processing of multi-word strings as single unanalysed units, whereas the open-choice principle consists of the word-by-word

processing of language according to syntactic rules. The preferred processing mode is the idiom principle or holistic mode. The open-choice principle or analytic processing mode only intervenes when necessary, as for instance when interpreting lexical choices that are unexpected in a given environment.

The cognitive advantage of the idiom principle is that FSs "facilitate efficient communication, both psycholinguistically and sociolinguistically" (Bell 2012: 193). In other terms, they fulfil two main functions in human language and communication, that is, smoothing social interactions and serving as processing short-cuts (Wray 2002; Wray and Perkins 2000). This also applies to SBUs, a sub-category of FSs that "serve as interactional patterns and rituals that usually mean the same to all speakers of a particular speech community" (Kecskés 2014: 137). However, unlike other kinds of FSs and idioms in particular,[1] the use of SBUs is closely tied to the performance of specific conventionalized pragmatic acts. Therefore, SBUs are obligatory and highly predictable (Kecskés 2010). On a continuum where obligatoriness increases to the right, SBUs will take the rightmost place, to the point that their functional-situational meaning may take over as the most salient meaning of the expression (Kecskés 2000a). In this sense, SBUs are selective and completive, that is, they are preferred over other expressions and evoke the situation in which they typically occur (Kecskés 2010). For instance, the expression *màn màn chī* 慢慢吃,[2] lit. 'eat slowly', corresponding to the English 'enjoy your meal', is selective to the extent that it is the preferred choice in the specific setting in which it is commonly used – starting a meal – and is completive as it can create its own situation without being used in an actual situational context.

Following the most recent trends, linguists' interest in Chinese formulaicity has also increased (see Yin 2013). If research into Chinese idiomatic expressions or *shúyǔ* 熟语 has a long and well-established tradition (e.g., Sun 1989), several studies on the description and classification of Chinese formulaic sequences have also appeared in the last few years (e.g., Bi and Shi 2013; Li 2013). To date, however, the number of studies that investigate Chinese SBUs or *qíngjìng zhuānyòngyǔ* 情境专用语 is still very limited.

The earliest example is probably the study conducted by Zhou (2012). The author adopted the socio-cognitive approach to investigate the unique characteristics of Chinese SBUs and compared them to their English counterparts. In another study,

---

1 Idioms are defined by Liu (2008: 23) as "multiword expressions consisting minimally of two words, including compound words … non-literal or semi-literal in meaning [and] generally rigid in structure". According to Kecskés (2003, 2010, 2014), idioms differ from SBUs in origin, purpose, and use, in that their occurrence is usually unpredictable and mostly depends on the individual speaker.

2 Characters (simplified) and full transliteration (*pinyin*) are only reported at the first occurrence of a Chinese word or expression. In later occurrences, only the *pinyin* without tones is reported.

Kecskés (2016) distinguished Chinese SBUs from three other types of Chinese idioms, i.e. *yànyǔ* 谚语 'proverbs', *guànyòngyǔ* 惯用语 'compounds', and *chéngyǔ* 成语 'four-character idioms', and proposed introducing SBUs as a separate, fourth category of Chinese idiomatic expressions.[3] The scholar classified Chinese SBUs on the basis of their pragmatic extensions or the degree of transparency (see Kecskés 2000a). He distinguished three types of Chinese SBUs: plain, loaded, and charged SBUs. Plain SBUs have a literal meaning which coincides with their functional meaning, so that they can be computed compositionally: An example is *qǐng wèn nín guì xìng* 请问您贵姓, lit. 'please ask you honourable surname', a polite expression used to ask a respected person's name. The pragmatic function of loaded SBUs is more relevant than their literal meaning, which is often difficult to recall. In other terms, loaded SBUs are more function-bound, and therefore less transparent: An example is the above-mentioned *man man chi*. Lastly, charged SBUs are the most ambiguous, because with no context they cannot be distinguished from their freely generated counterparts. In this regard, an interesting phenomenon in Chinese is the great number of multi-charged SBUs: The expression *qǐng biàn* 请便, lit. 'please [do as you find it] convenient', for instance, has three almost equally salient meanings, that is, 'do as you like', 'ask somebody to leave', and 'scorn' (Zhou 2012). This is because in Mandarin the actual situational context plays an even more decisive role in meaning creation than in English (Kecskés 2016). Together with the high degree of cultural specificity, this is one of the factors that make the acquisition task particularly daunting for CFL learners.

2.2 FSs and SBUs in L2/FL

According to Pawley and Syder (1983), reaching a high level of native-like selection is one of the main goals for L2/FL learners. Quoting the authors, learners "must learn not only the grammar [but also] a means for knowing which of the well-formed sentences are nativelike" (p. 194). In other terms, "[p]art of [L2/FL] knowledge entails developing suitable rules to generate all the grammatical utterances […] but another crucial aspect is coming to know which of the feasible grammatical utterances are idiomatic" (Wray 2002: 143). However, the mechanics of how formulaic sequences are acquired are not well understood (Durrant and Schmitt 2010: 163).

---

3 Note that, in addition to *yanyu*, *guanyongyu*, and *chengyu*, Chinese *shuyu* also include *xiēhòuyǔ* 歇后语, a particular idiom form consisting in a sort of riddle and composed of two separated parts (e.g., see Sun 1989). Other idiom forms often included among *shuyu* are *géyán* 格言 'mottoes', and *súyǔ* 俗语 'popular sayings' (see Conti 2019).

The importance of formulaic language in L2/FL acquisition is widely acknowledged by scholars. According to Boers and Lindstromberg (2012), formulaic language plays an important part in enabling the comprehension and expression of messages that might otherwise fail to get across. In addition, there is a growing body of evidence that learners can greatly benefit from mastering FSs, particularly for the development of communicative competence and native-like fluency (see Boers et al. 2006; Conklin and Schmitt 2008; Jiang and Nekrasova 2007; Pang 2020).

SBUs are particularly relevant for learners' pragmatic competence and conceptual fluency. Knowing what expressions are appropriate or inappropriate in different social situations is an important sign of group-inclusiveness and "native-likeness" (Kecskés 2016). However, it was demonstrated that learners of high grammatical proficiency will not necessarily show concomitant pragmatic skills. Kecskés (2000b), for instance, submitted three comprehension and production tasks to 88 non-native speakers of English with different L1 backgrounds and L2 competence levels, and found that even participants who had spent more than two years in the U.S. were non-native-like in their selection of target forms. The author outlined a three-stage developmental scale, comprising strong L1-culture transfer, false generalizations, and individual choice, respectively. Cultural difference played a very significant role, both because some cultures find it important to use an SBU in a particular situation while others do not, and also because the same situational function may be worded differently across languages. Lastly, the study revealed that non-native speakers' use of SBUs was characterized by overuse, oversimplification, verbosity and avoidance, independently from the length of learning experience.

Two studies specifically investigated the development of Chinese SBUs in a study abroad context, both finding that frequency of exposure was not related to SBU production. In the first study, Taguchi et al. (2013) submitted a Discourse Completion Task (DCT) to 31 intermediate-level American learners studying in China. The DCT was administered twice, at the beginning and at the end of the semester. The authors also administered a questionnaire concerning the participants' perceived frequency of encounter with the target situations. The results showed that the reported frequency of encounter did not correlate with the gains in SBU production. In addition, the authors observed that the production of exact target formulae remained limited over time, and that the possible reasons may be lack of lexical and syntactic knowledge, and lack of pragmalinguistic and/or sociopragmatic knowledge. Lastly, the qualitative analysis of the DCT revealed four patterns of change: change toward target formulae, change toward target-like slot-and-frame patterns,[4] change towards non-target formulae, and stabilization of non-target formulae use.

---

4 According to Taguchi et al. (2013: 39), slot-and-frame patterns are variable units with open slots (e.g., locative subject + *zài nǎr* 在哪儿 'in where').

In the second study, Yang (2016) investigated the development of SBU use of 59 CFL learners with different length of classroom instruction experience and different lengths of stay in the target language environment. The participants completed two tests, a production and a recognition test. The statistical analysis of the data showed that length of stay in the target language country significantly correlated with recognition but not with production. In addition, the qualitative analysis revealed that the main challenges arose in three kinds of situations: those requiring different SBUs in the L1 and the L2, those requiring an SBU in the L2 but not in the L1, and those in which the participants overgeneralized a more familiar form.

From the above review, two observations can be made. Firstly, the results of the examined studies confirm that "by the time the learner has achieved a reasonable command of the [L2/FL] lexicon and grammar, the formulaic sequences appear to be lagging behind" (Wray 2002: 182). Secondly, the main difficulty faced by learners when encountering Chinese SBUs is cultural specificity, especially considering the high degree of diatopic variation and the frequent lack of equivalents in learners' L1s (Kecskés 2016). Acquisition, therefore, much depends on how language learners can identify themselves with the function and content SBUs express.

In summary, it appears clear that "simple exposure to the target language environment does not necessarily entail the ability to produce pragmatic routine formulae" (Yang 2016: 52). Formal instruction might be required for the uptake of SBUs to occur.

Indeed, there is enough evidence to conclude that intentional learning activities that explicitly present FSs as objects of language study have statistically significant effects on retention (for a comprehensive review, see Pellicer-Sánchez and Boers 2019). However, the number of studies specifically addressing SBU learning is still limited. To our knowledge, the only study that investigated the teachability of SBUs is that conducted by Yeh (2016). The author examined to what extent explicit versus implicit instruction affected learners' knowledge and ability to use Chinese SBUs. Her findings demonstrated that explicit teaching increased both the students' awareness and their frequency of use of SBUs. Apart from this exception, the research on SBU instruction in a foreign language context is still lagging behind.

In an attempt to fill this gap, this study tested the effectiveness of the identification task on the retention of Chinese SBUs. The identification task is an intentional learning activity consisting of recognising a stimulus or distinguishing it from other stimuli. According to Pellicer-Sánchez and Boers (2019: 159–160), intentional learning activities are those where FSs are explicitly presented as objects of language study. They categorize these activities in three groups: (1) instructions to explore texts for the presence of FSs; (2) decontextualized FS-focused activities; (3) engage learners with particular characteristics of FSs (e.g., form, sound, etc.). The identification task we tested belongs to the first group and consisted of exploring the text to retrieve the target items.

Similar attention-directing activities have already been tested with different types of FSs. An example is 'text-chunking', which requires learners to identify all the FSs in a text. In the study by Boers et al. (2006), the experimental group was explicitly instructed to identify useful phrases in the texts provided. A following interview demonstrated that the text-chunking group used more FSs than the comparison group. However, in a partial replication of this study, pretest-posttest comparisons revealed no evidence of any significant uptake between the control and treatment groups (Stengers et al. 2010). These contrasting results might depend on the fact that "an interview is probably too crude an elicitation instrument to reveal the true extent to which particular items were learned from texts spread over a whole course" (Pellicer-Sánchez and Boers 2019: 161).

Another attention-directing technique which led to sounder results is typographic enhancement. This technique was tested in a study by Peters (2012). The author had the participants read a text which included 24 target items, 12 collocations and 12 single words. Half of these were typographically enhanced, i.e. printed in bold typeface and underlined. The experimental group was explicitly instructed to focus on both vocabulary and FSs and to write down those items that were unfamiliar, whereas the control group was instructed to focus on unfamiliar vocabulary only. An immediate recall test revealed that typographic salience had a significative effect on participants' scores, proving particularly beneficial for FS learning.

Peter's findings show that a task that specifically enforces learners to process new linguistic forms produces significant effects, as students allocate their attentional resources in function of the specificity of the task to perform.

However, each type of FSs has its own characteristics that determine the type of pedagogical intervention required. Unlike other Chinese FSs, SBUs do not show form or sound regularities on which to direct learners' attention.[5] Instead, their close relation with the situational context might provide valuable mnemonic cues to facilitating retention. For these reasons, we hypothesized that the identification task might prove particularly suitable for SBU learning. Based on these premises, the research questions of this study are the following:

1. Is the identification task effective for the retention – active form recall – of Chinese SBUs in CFL?
2. Does length of CFL learning experience have any influence on the retention of the target SBUs?
3. Are the effects – if any – durable from a longitudinal perspective?

---

5 Most *chengyu*, for instance, have a fixed four-character structure and an AABB prosody with a pause after the second character. Their components also display semantic and phonological parallelisms which can be useful for pedagogical purposes (see Conti 2019).

4. Is there any difference in the retention of SBUs with different degrees of lin-
   guistic/cultural specificity?

3 Method

3.1 Design and procedure

Based on Dörnyei (2007) and Phakiti's (2014) terminology, the study adopted a
quasi-experimental, QUAN qual, pretest-posttest-delayed posttest design.

The procedure comprised two main phases, further divided into two sub-
phases. The first phase, the preparatory phase, first consisted of the consultation of
reference studies to identify some of the most recurrent communicative situations
or scenarios which could elicit the use of SBUs. Second, a survey containing 30
scenarios was submitted to three pools of respondents, i.e., Italian native speakers
(INSs), Chinese native speakers (CNSs), and Italian CFL learners (CFLLs). The
survey also served as a pretest, as the same pool of CFLLs were also involved in the
second phase of the study.

The second phase featured the experimental phase proper. First, a treatment
targeting the SBUs selected among the responses to the survey was created and
administered to a group of CFLLs. Lastly, a posttest and a delayed posttest were
submitted to the participants, in order to test the effectiveness of the treatment.

3.2 Preparatory phase: survey design, pretest and selection of target SBUs

In order to select the SBUs for the study, two versions of a discourse-completion
questionnaire were designed, one version in Italian and a second in Chinese.[6] The
questionnaire was based on Taguchi et al. (2013) and Yang (2016) and consisted of
30 scenarios corresponding to various types of communicative functions, such as
greetings and farewells, apologies, requests for information, refusals, etc. The
wording of the scenarios in Chinese were reviewed by a CNS.[7]

---

6 The questionnaire is available on the CFL Research Network (Conti and Lepadat n.d.).

7 In spite of their widespread use in many fields of linguistics, the employment of both written
questionnaires and DCTs has been a matter of debate in the literature due to the "artificial" nature
of these data eliciting methods (e.g. Dohrenwend 1965; Labben 2016). However, as Yang (2016)
pointed out, statistical findings have shown that data collected through discourse-completion
tests are able not only to accurately reflect naturally occurring language (Beebe and Cummings
1996), but also to overcome the unpredictable features of collected natural data (Kwon 2004).

The survey had a twofold objective. The first was that of eliciting INSs and CNSs' preferred formulae for each of the communicative situations involved. The second objective was that of assessing CFLLs' pre-existing knowledge of the considered Chinese SBUs and selecting the unfamiliar ones for the treatment and test phases.

In order to ensure a controlled testing environment, a paper-based questionnaire was submitted to the CFLLs during class. The survey addressed to native speakers was written using the on-line platforms Google Forms and Tencent Wenjuan 腾讯问卷. Following, the two forms were shared on various social media. This allowed us to collect data from a pool of native speakers with diversified provenances and spoken regional varieties.

The final pool of respondents resulted in 48 CFLLs, 65 CNSs, and 85 INSs. The CFLL group was composed of 28 year 2 (58.3%) and 20 year 3 students (41.7%), all enrolled at Roma Tre University, aged 22 on average. Based on the requirements of the Chinese course at Roma Tre University, their language proficiency roughly corresponded to HSK level 2 in year 2 and HSK level 3 in year 3. The CNS group was composed of 23 male (35.4%) and 42 female speakers (64.6%), with an age ranging from 19 to 49 years and residing in different parts of China. They all spoke Mandarin Chinese and the majority of them had completed higher education studies in China. Lastly, the INS sample comprised 16 (18.8%) males and 69 (81.2%) females coming from different regions of Italy. Their ages ranged from 25 to 52 years and the majority of them had obtained advanced degrees in different fields of study. The data collected from this group were used both for the selection of the target SBUs and as a standard of comparison in the analysis of the learners' productions.[8]

After the analysis of the surveys, six Chinese SBUs with different degrees of linguistic and cultural specificity were selected based on two criteria. The first criterion was high frequency of occurrence in native speakers' responses. This ensured that all the target formulae were the preferred choices for NSs. The second criterion was low familiarity to learners. In other terms, only the SBUs for which all the CFLLs obtained a 0 score were selected for the subsequent experiment. The few CFLLs who obtained higher scores with respect to the six items in the survey (pretest) were not included in the final group of 36 CFLLs who took part in the experiment.[9] This ensured the validity of the study, as the results of the posttests only depended on the experimental conditions.

---

8 The comparative analysis between Italian and Chinese SBUs was beyond the scope of this study. Reference was made to the (a)symmetries in the two languages only inasmuch as they were relevant for the CFLLs' retention of specific SBUs. The comparison of Italian, Chinese, and learners' responses to the survey is tackled in a different contribution (Conti and Lepadat in preparation).
9 The scoring criteria are detailed in Section 3.3.2.

The target SBUs and the corresponding Italian SBUs are reported in Table 1. All the selected SBUs present various degrees of specificity, either linguistic of cultural. More specifically, the items were considered as culturally-specific if only one of the two languages presented a specific formula for a communicative situation, whereas the other did not. As is shown in Table 1, only item 2 can be considered as fully culturally-specific when compared to the learners' L1. In fact, this was the only scenario in which CNSs used a situation-specific formula, whereas INSs preferably used a very generic greeting formula which is suitable for different

Table 1: Target SBUs.

Scenario SBU						Cultural spec.	Linguistic spec.
1. Wrong phone call							
CH: (不好意思/对不起)	你/您		打错了。			No	Low (syntax)
Bù hǎo yìsi/duì bu qǐ	*nǐ/nín*		*dǎcuò-le*				
Sorry	2SG/HON		call-wrong-PFV				
IT: (Scus-i) (cred-o	che)	ha		sbagliato	numero		
Sorry-HON think-1SG	that	AUX.HON		mistaken	number		
2. Depart from a friend when outside							High (syntax, semantics, vocabulary)
CH: (路上)	注意		安全!			Yes	
Lùshang	*zhùyì*		*ānquán*				
Road-on	pay.attention		safety				
IT: Ciao							
Hallo/bye							
3. Meet a friend after a long time							Low (syntax)
CH: 好久不见	(了)!					No	
Hǎojiǔbújiàn	*le*						
Very-long.time-not-meet	MOD						
IT: Ciao/ehi, (da	quanto		tempo!				
Hello/hey since	so.much		time				
4. Answer a phone call when with friend							Medium (syntax, vocabulary)
CH: 不好意思/对不起,	(我	先/得)	接			No	
Bù hǎo yìsi/duì bu qǐ	*wǒ*	*xiān/déi*	*jiē*				
Sorry	1SG	first/must	pick.up				
个 电话。							
gè diànhuà							
CLASS phone.call							
IT: Scus-a/ate, devo		(proprio/assolutamente)	rispondere.				
Sorry-2SG/2PL must-1SG		really/absolutely	to.pick.up				
5. Ask if a seat is available							High (syntax, semantics, vocabulary)
CH: (请问/不好意思)	这儿/这里	有	人	吗?		No	
Qǐngwèn/bù hǎo yìsi	*zhèr/zhèlǐ*	*yǒu*	*rén*	*ma*			
Excuse.me/Sorry	here	there.be	people	Q			
IT: Scus-i/ate quest-o/a		posto/sedia	è	liber-o/a?			
Sorry-HON/2PL this-M/F		seat/chair	is	free-M/F			
6. Ask for a ligther							High (syntax, semantics, vocabulary)
CH: (不好意思)	(能/可以)	借	个	火	(吗)?	No	
Bù hǎo yìsi	*néng/kěyǐ*	*jiè*	*gè*	*huǒ*	*ma*		
Sorry	can	to.lend	CLASS	fire	Q		
IT: (Mi scus-i)		ha/avr-ebbe		da	accendere?		
Me forgive-HON		have/have-COND		to	to.light		

scenarios. Other items implied cultural features. For example, both items 1 and 4 contain an apology. However, as these culturally-loaded elements occur in both languages, the two items were considered equivalent.

Linguistic specificity, on the other hand, was established on the basis of three parameters, i.e. syntactic, semantic, and lexical variation between the two languages. Items showing linguistic differences at each of the three levels were considered as having a high linguistic specificity. Those with differences at two of the levels have medium specificity. Low specificity was assigned to those items varying along one single parameter or less.

Note that, in some cases, the situational meaning of the target SBUs does not differ from its propositional meaning (e.g., Item 1). However, we believe that these expressions can be considered as plain SBUs as they satisfy the following assumptions: (1) they are the NSs' preferred choice in that specific situation (i.e., they are selective); (2) they cannot be used in other contexts; (3) they display a high degree of conventionality and predictability, as well as little or no variability; (4) they perform a pragmatic function (e.g., informing the hearer that the number they have called is wrong); (5) they evoke the communicative situation in which they typically occur (i.e., they are completive).

3.3 Experimental phase

3.3.1 Participants

The participants in the posttests were 36 Italian CFLLs who had previously taken part in the survey. They were divided into an experimental group of 17 CFLLs and a control group of 19 CFLLs. The participants in the control group did not receive any instruction. All the participants provided verbal consent to participation in the study.

A short background section was included at the end of each questionnaire in order to collect the participants' biographic data, including age and gender, linguistic background, length of classroom instruction at the university, knowledge of other foreign languages, and length of study in China (if any). These data are reported in Table 2.

Due to natural variation, female participants were represented in a bigger proportion than males. The participants were almost exclusively native speakers of Italian, with the exception of one bilingual speaker. Their ages ranged from 20 to 29, with a mean of 22. Around half of the learners had received 1–2 years of formal classroom instruction in Chinese and the remaining half had received 2–3 years of instruction; only 1/5 of the sample had stayed in China for 1 year or less, whereas the remaining part had never travelled to Chinese-speaking countries.

Table 2: Participants' background information.

Category	Sub-category	*n*	%
Mean age		22	
Gender	Male	5	14%
	Female	31	86%
Language Background	Italian	35	97%
	Filipino	1	3%
Length of classroom instruction (course year)	1–2 years (year 2)	16	44%
	Control	8	50%
	Experimental	8	50%
	2–3 years (year 3)	20	56%
	Control	11	55%
	Experimental	9	45%
Length of stay in China	none	29	81%
	1–6 months	6	16%
	7–12 months	1	3%

The delayed posttest was conducted limitedly to a subgroup of the experimental one. The group was composed of 8 participants, 2 belonging to year two and 6 belonging to year three.

3.3.2 Treatment and posttests

The treatment involved a classroom group activity focused on the communicative functions of the SBUs. The respondents – divided in small groups of 3 or 4 – were presented with 6 dialogues in Mandarin Chinese containing the target expressions plus 2 distractors (dialogues 3 and 6 in the Appendix). They were required to read the dialogues and to identify the target expressions corresponding to the six communicative functions provided in their L1. The dialogues containing the target SBUs were accurately created in order to reflect prototypical situations of occurrence of the expressions and later evaluated by a native speaker of standard Mandarin Chinese. The vocabulary and grammar of the dialogues were calibrated onto HSK level 2, so that they were equally intelligible to participants of both years. A brief discussion among the groups and with the interviewer followed, in order to assess the appropriateness of the responses and provide corrective or confirmation feedback. The control group served as a baseline group, defined as "a control with which treatment effects will be compared" (Riazi 2016: 19). This was in order to ensure that the outcomes of the posttest could only be attributed to the effect of the

treatment. Therefore, the participants in this group were not exposed to the input and only received feedback on the target forms after the posttest.

In order to assess the effectiveness of the activity on the participants' retention of the target SBUs, two posttests were submitted to the participants after two days and after two weeks from the treatment. The posttests were identical, except for the order of the items, and were both written in Italian. The participants were provided with six communicative functions and were required to provide the corresponding Chinese SBUs. The first test was submitted to 36 participants, both in the experimental and the control group, while the second test was carried out on a sub-sample of the experimental group.

CFLLs' responses were scored on a three-point scale ranging from 0 to 2. Target-like expressions were assigned two points. One point was given to expressions which were close to the target ones, but contained little additional linguistic material (e.g. the preposition *zài* 在 in Item 5: *zài zhèr yǒu rén ma?* 在这儿有人吗?; the pronouns *wǒ* 我 and *nǐ* 你 in Item 3: *wǒ hěn jiǔ bú jiàn nǐ* 我很久不见你), small omissions (e.g. the omission of the locative expression *zhèr* 这儿 in Item 5: *yǒu rén ma?* 有人吗?), minor syntactic errors (e.g. the insertion of the structural particle de 得 in the result complement in Item 1: *nǐ dǎ de cuò* 你打得错), or dispreferred lexical choices (e.g. the use of *hěn* 很 instead of *hǎo* 好 in item 3: *hěn jiǔ bú jiàn* 很久不见). Lastly, zero points were given to responses with major syntactic errors (e.g. *nǐ dǎ diànhuà cuò de hào* 你打电话错的号 for item 1), inappropriate lexical choices (e.g. the use of *kàn* 看 instead of *jiàn* 见(面) and *duō* 多 (*shíjiān* 时间) instead of *jiǔ* 久 in Item 3: *hěn duō shíiān wǒmen kàn* 很多时间我们看; the use of *shuōhuà* 说话 instead of *jiē diànhuà* 接电话 in Item 4: *duìbuqǐ, wǒ yào shuōhuà* 对不起,我要说话), or expressions which were clearly non-native-like, even if grammatically correct (e.g. *qǐngwèn, nǐ yǒu méi yǒu huǒjī?* 请问,你有没有火机 for Item 6; *shǒujī hàomǎ shì cuò de* 手机号码是错的 for Item 1). Missing responses were also assigned zero points. Since the treatment focused on the communicative functions of the SBUs rather than the form, only syntactic and lexical well-formedness of the expressions was considered. Minor errors in character writing and the use of *pinyin* were disregarded.

Two independent raters evaluated the tests. Interrater reliability was measured by means of Cronbach's alpha for each of the items, resulting in $α ≥ 0.8$ in all cases. This suggests a high reliability of the assigned scores.

3.3.3 Results

A robust t-test comparison showed that the difference between year 2 and year 3 (including the control group) was not significant (see Section 4 for more details on

the statistic procedures adopted). Year 2: $N = 16$, M = 2.53 (SD = 2.38); year 3: $N = 20$, M = 2.33 (SD = 2.47); t (15.3) = 0.58, p = 0.57 ($p > 0.05$), 95% CI [–1.20, 2.10], $\hat{\xi} = 0.14$.

The same results were found when the comparison was conducted excluding the control group. Year 2: $N = 8$, M = 3.25 (SD = 2.27); year 3: $N = 9$, M = 4.0 (SD = 2.84); t (9.6) = 0.42, p = 0.68 ($p > 0.05$), 95% CI [–2.55, 1.75], $\hat{\xi} = 0.16$. In the subsequent analysis, the two years were thus considered as a single sample numbering 36 participants, further divided into control and experimental groups (see Table 3). The results of the posttest are reported in Table 3 and represented in Figure 1. Scores in the control group that are higher than 0 can be attributed to blind guesses which were accidentally close to the target forms. However, robust statistical methods are not affected by outliers, thus ensuring the validity and reliability of the results.

Table 4 reports the experimental group's results divided by item (see also Figure 2).

The item with the lowest mean score was Item 2 (*zhuyi anquan*), the only culturally-specific one in the sample, whereas the item which obtained the highest mean score was item 5 (*zher you ren ma?*). Other items that contained cultural

Table 3: Posttest results.

Group	n	M	SD
Experimental	17	3.65	2.54
Control	19	1.32	1.66
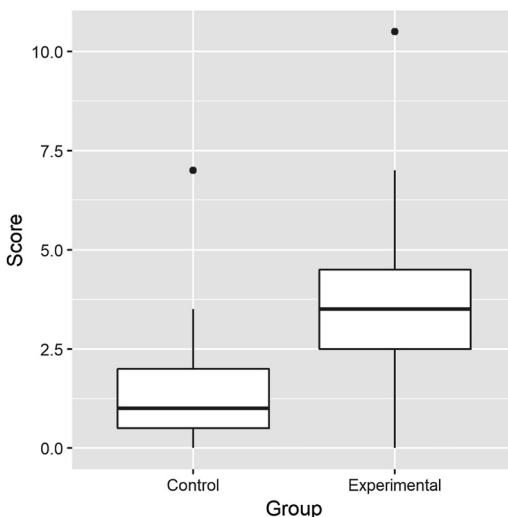
Figure 1: Experimental – control comparison.

Table 4: Posttest item analysis.

Item	M	SD
Item 1	0.76	0.81
Item 2	0.21	0.53
Item 3	0.56	0.77
Item 4	0.59	0.69
Item 5	1.03	0.62
Item 6	0.50	0.64

Note: cf. Table 1 for item details.
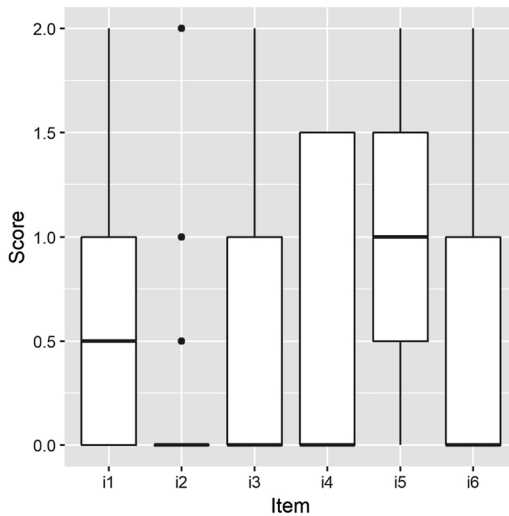
Figure 2: Test score by item.

features did not constitute a challenge, presumably because they had equivalents in the L1 counterparts.

As mentioned in Section 3.3.1, only a sub-sample of 8 students participated in both the posttest and the delayed posttests. Table 5 reports the results of the sub-sample in the two tests (see also Figure 3).

The delayed posttest results divided by item are reported in Table 6 and represented in Figure 4. Again, Item 2 obtained the lowest scores in both tests.

Table 5: Sub-sample immediate and delayed posttest results.

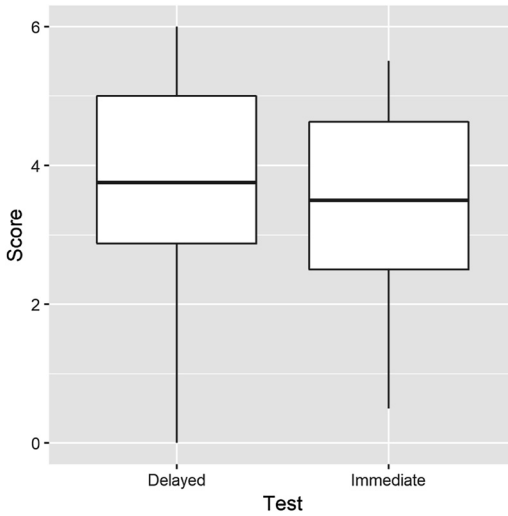| Test	n	M	SD
Immediate posttest	8	3.44	1.64
Delayed posttest	8	3.63	1.87

Figure 3: Immediate – delayed posttest comparison.

Table 6: Immediate – delayed posttest item analysis.

Item	Immediate posttest		Delayed posttest	
	M	**SD**	**M**	**SD**
Item 1	0.50	0.71	0.69	0.84
Item 2	0.13	0.35	0.13	0.35
Item 3	0.56	0.68	0.50	0.76
Item 4	0.38	0.58	0.31	0.46
Item 5	1.25	0.46	1.19	0.84
Item 6	0.63	0.69	0.81	0.53

Note: cf. Table 1 for item details

4 Analysis

All the statistical analyses were carried out by means of the open-source software R, which is particularly useful when dealing with data that are not normally distributed, since it implements robust statistics (Field et al. 2012).

The homogeneity of variance was tested for each item and each group by means of the Levene test, which showed that the variation inside the groups was not always homogeneous. The presence of outliers was also frequent. Therefore, we adopted robust statistics for the analysis of the data (see Mair and Wilcox 2019). Robust 20% mean-trimmed independent t-tests with bootstraps were carried out to measure the differences between year 2 and 3 and between the experimental and the control groups. Robust 20% mean-trimmed RM ANOVA and post-hoc comparisons were conducted to measure between-item differences of both the posttest
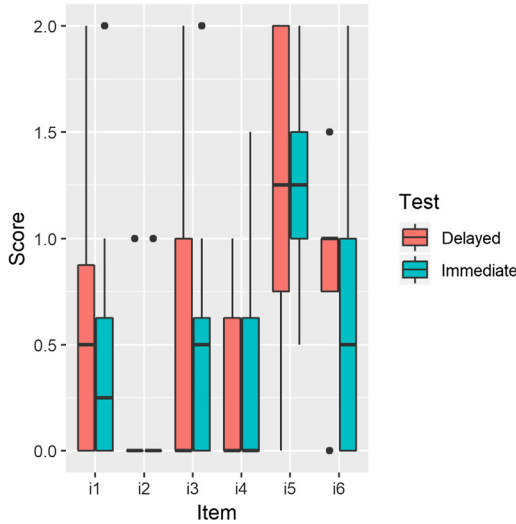
Figure 4: Immediate – delayed posttest comparison by item.

and the delayed posttest. Robust 20% mean-trimmed paired-sample t-test were used to compare the results of the immediate and the delayed posttests. The reported effect size is the explanatory measure of effect size $\hat{\xi}$, proposed by Wilcox and Tian (2011). Values of $\hat{\xi}$ = 0.10, 0.30, and 0.50 correspond to small, medium, and large effect sizes, respectively.

4.1 Posttest

The results of the robust independent-sample t-test are significant for $p < 0.05$; t (15.3) = 4.23, p = 0, 95% CI [1.22, 3.68], with a large effect size ($\hat{\xi}$ = 0.76). These results suggest that the treatment had a pronounced effect on the memorization of the form of the target SBUs.

The scores for each item were compared by means of a robust RM ANOVA. The omnibus test is significant, F (5, 50) = 4.11, p = 0.003 ($p < 0.05$). However, post-hoc comparisons were never significant. Nevertheless, as shown by the $\hat{\Psi}$ values, the biggest differences between items are those involving item 5, particularly in the comparison between items 2 and 5 (Table 7).

4.2 Delayed posttest

The mean scores of the immediate and delayed posttests were compared by means of a robust paired-sample t-test, which resulted as non-significant, t (5) = −0.87,

Table 7: Immediate posttest, between-item post-hoc comparisons.

Comparisons	$\hat{\Psi}$	CI	p	Crit p
i1–i2	0.45	−0.29, 1.20	0.04	0.00
i1–i3	0.14	−0.54, 0.81	0.46	0.01
i1–i4	0.09	−0.27, 0.45	0.36	0.01
i1–i5	−0.27	−1.18, 0.64	0.28	0.01
i1–i6	0.27	−0.44, 0.98	0.17	0.01
i2–i3	−0.23	−0.89, 0.44	0.22	0.01
i2–i4	−0.27	−0.98, 0.44	0.17	0.01
i2–i5	−0.91	−1.92, 0.11	0.01	0.00
i2–i6	−0.14	−0.52, 0.24	0.20	0.01
i3–i4	0.05	−0.53, 0.62	0.77	0.05
i3–i5	−0.59	−1.53, 0.35	0.04	0.00
i3–i6	0.09	−0.51, 0.79	0.58	0.03
i4–i5	−0.45	−1.44, 0.53	0.11	0.00
i4–i6	0.09	−0.51, 0.70	0.58	0.02
i5–i6	0.55	−0.42, 1.51	0.06	0.00

$p = 0.42$ ($p > 0.05$), 95% CI [−0.99, 0.49], $\hat{\xi} = 0.11$. Notice that the CI crosses zero and that the explanatory measure of effective size is negligible. This suggests that after two weeks, the effect of the treatment on the memorization of the target SBUs had not significantly decreased.

Paired comparisons between the mean scores of each item in the two tests were also conducted. The results are reported in Table 8. The t-test was not performed for items 2 and 4, as the results and the corresponding SD did not vary in the two tests.

The between-test item comparisons were never significant, thus confirming the general trend already observed in the analysis of the total scores.

Lastly, a robust RM ANOVA was used to compare the scores obtained in the delayed posttest for each item. As in the immediate posttest, the omnibus test was significant, $F(5, 25) = 2.68$, $p = 0.045$ ($p < 0.05$). Again, however, the post-hoc comparisons did not result as significant. The picture returned by the data reported in Table 9 resembles that described for the immediate posttest (Section 4.1), the highest differences between items involving items 2 and 5.

Table 8: Immediate – delayed posttest comparison.

Item	$t(5)$	p	95% CI	$\hat{\xi}$
Item 1	−1.13	0.31	−0.82, 0.32	0.27
Item 3	0.54	0.61	−0.31, 0.48	0.08
Item 5	0.00	1.00	−1.12, 1.12	0.00
Item 6	−1.83	0.13	−0.80, 0.14	0.33

Table 9: Delayed posttest, between-item post-hoc comparisons.

Comparisons	$\hat{\Psi}$	CI	p	Crit p
i1–i2	0.58	−2.02, 319	0.29	0.01
i1–i3	0.25	−1.60, 2.10	0.51	0.01
i1–i4	0.33	−1.04, 1.70	0.26	0.01
i1–i5	−0.50	−3.48, 2.48	0.42	0.01
i1–i6	−0.17	−2.18, 1.85	0.68	0.03
i2–i3	−0.33	−1.93, 1.27	0.32	0.01
i2–i4	−0.08	−0.80, 0.63	0.57	0.02
i2–i5	−1.08	−4.00, 1.83	0.11	0.00
i2–i6	−0.67	−2.27, 0.93	0.08	0.00
i3–i4	0.08	−1.72, 1.88	0.82	0.05
i3–i5	−0.75	−3.61, 2.11	0.23	0.00
i3–i6	−0.33	−1.93, 1.27	0.32	0.01
i4–i5	−0.83	−3.78, 2.12	0.20	0.00
i4–i6	−0.50	−2.15, 1.15	0.17	0.00
i5–i6	0.33	−1.27, 1.93	0.32	0.01

5 Discussion

The data we collected and analysed seem to suggest that the treatment was effective for the retention of the target forms, also from a longitudinal perspective. On the other hand, the overall scores were generally low, and the responses showed a relatively high degree of inaccuracy.

RQ 1 asked whether intentional identification of the target SBUs from given communicative situations enabled learners to memorize them. Statistical analysis suggests that the treatment was effective. The control group, which had not received any instruction, was not able to provide the required formulae. The experimental group, on the contrary, significantly outperformed the control group, meaning that the identification task did leave some mnemonic traces on the participants.

Even so, the mean scores obtained by the experimental group are still relatively low. This was not surprising and was probably due to the limited exposure to the input. These results are consistent with Durrant and Schmitt's (2010) findings on the retention of English collocations. In their study, the scholars asked the participants to read aloud sentences containing adjective–noun sequences, the control group only receiving a single exposure to the target collocations. In a subsequent posttest, the participants were asked to supply the nouns associated with the adjective provided. The effect of the read-aloud task was very weak in the

control group, as only three collocations out of ten resulted correct. Paraphrasing Boers and Lindstromberg (2012: 89), these findings demonstrate that, as for single words, one or two exposures to a given sequence are hardly enough to leave durable memory traces (see also Pellicer-Sánchez and Boers 2019: 162). Considering the incrementality of vocabulary acquisition (Schmitt 2000), it might be hypothesised that the results would be higher if form recognition was tested instead of recall.

In addition to the low scores, a high rate of inaccuracy also emerged from the qualitative analysis of the participants' responses. An example is provided by the responses to Item 3, which is almost unanimously used by CNSs when meeting someone after a long time. As seen, a frequent response in the posttest was *hen jiu bu jian*, where the emphatic adverb *hǎo* 好 'such a very' of the target formula was replaced by the unmarked degree adverb *hěn* 很 'very'. It is evident that the participants were able to retrieve the formula from memory, and that part of it was adapted to fit their interlanguage by using a more familiar form.

These findings are consistent with Wray's (2000, 2002) observations on the composition of native and non-native speakers' mental lexicon. According to the scholar, native speakers have stored a large amount of non-analyzed or semi-analyzed multi-word sequences, which can be easily retrieved as single units when required. Learners, on the contrary, seem to take an essentially non-formulaic approach to learning. In other terms, it appears that they cannot avoid analytic word-by-word processing, therefore they mostly acquire single-word units. When producing an utterance in FL, they have to compose it out of individual words, and this implies that every option is equally plausible. In our case, we can hypothesize that *hao* was processed as equivalent to *hen*, as both the forms *hǎo jiǔ* 好久 'such a long time' and *hěn jiǔ* 很久 'a very long time' are admitted in Chinese and roughly convey the same meaning. When producing the SBU, the participants opted for the more familiar *hen jiu*, without considering the fixedness of the sequence.

RQ2 addressed the effect of the length of learning experience on the retention of the target forms. As expected, the answer to this question is negative. The difference in the results obtained by the participants in years two and three was not significant, meaning that the effects of the treatment were the same. In fact, all the materials were calibrated onto HSK level 2, which was the assumed competence level of year-two participants. Therefore, year-three participants were not advantaged by their longer learning experience. This suggests that SBUs are teachable even at the earlier stages of CFL acquisition, and that properly designed tasks for SBU learning can also be introduced to lower level learners.

Concerning RQ3 on the longitudinal effects of the considered task, the sample was too small to draw valuable conclusions. Still, the obtained results seem to

suggest that the effects of the identification task were durable. Statistical evidence showed that there was no significant difference between the results of the two posttests. In other terms, the participants retained the target forms even two weeks after exposition to the input. According to Schmitt (2010: 157), in order to be indicative of stable learning, a delayed posttest should be carried out after three weeks from the treatment. In our case, the delayed posttest was administered after two weeks. Even so, it might not be hasty to assume that the identification task did succeed in leaving traces on the participants' long-term memory, especially if we consider the poor exposure to the target forms.

The results reported so far are supported by previous studies on the effects of awareness-raising and attention-directing activities on the retention of formulaic sequences. As mentioned in Section 2.2, research in this area led to contrasting outcomes. Activities such as text-chunking (Jones and Haywood 2004; Stengers et al. 2010) revealed no evidence of differential uptake after pretest-posttest comparisons. Discussing these results, Boers and Lindstromberg (2012: 89–90) noted that the participants had underlined many segments of the texts that none of the native speakers had considered formulaic. According to the authors, it must be hard for learners to autonomously recognize formulaic sequences in the first place.

On the contrary, the results obtained by Durrant and Schmitt (2010) and Peters (2012) are more encouraging. In these studies, the target formulae were either typographically enhanced or presented in decontextualized sentences, and in both cases the participants had a higher success rate in terms of form recall. In line with Boers and Lindstromberg (2012), it appears that a certain degree of focus on form, manipulation of the input, and decontextualization are necessary for learners to commit formulaic sequences to memory. In our study, the participants had to establish a link between a particular communicative function and the formula used to express it. This activity differs from text chunking in at least two ways: first, the participants received a feedback which confirmed or corrected their guesses; second, the aim of the task was that of recognizing specific strings, not to indiscriminately underline any word sequence they assumed to be formulaic. Consequently, the participants' attention was not primarily allocated to the content of the text; rather, it was focused onto the selected linguistic forms, which were further decontextualized through the instructor's feedback. Lastly, the fact that the target formulae were associated to a specific context and a communicative function might have provided extra mnemonic cues that stimulated retention.

The last research question (RQ4) addressed the effect of linguistic and cultural specificity on the memorization of the target SBUs. As reported in Sections 4.1 and 4.2, the comparison of the scores for each item in both posttests only returned a

significant main effect, whereas the results of the post-hoc tests were non-significant. Therefore, we concluded that specificity did not affect the retention of the target SBUs. However, a closer look at the results of the post-hoc tests seem to suggest that the biggest differences were found when Item 2 (*lu shang zhuyi anquan*) was compared to the other items. Recall that Item 2 was the only one which was classified as culturally specific. It is thus safe to hypothesize that minor linguistic or cultural differences did not constitute a big obstacle for retention; rather, the critical factor might be the presence of an SBU in the FL where no specific sequence is expected in the L1 for that particular situation. Indeed, previous studies already demonstrated that cultural specificity constitutes an obstacle to formulaic language learning in FL. A study conducted by Boers and Demecheleer (2001) addressed a different kind of formulaic sequence, that is, imageable idioms. In this study, the scholars observed that idioms reflecting a metaphoric theme that is not salient in the L1 tended to be less easily guessable. Our study offers (tentative) support to the hypothesis that cultural specificity constitutes a big challenge for learners and has a higher impact on the learnability of formulaic sequences compared to linguistic specificity.

6　Limitations, future directions and conclusions

The present study tested the effectiveness of the identification task on the retention of Chinese SBUs. The results, though preliminary, seem to indicate that (i) the task did succeed in leaving some traces on the participant's memories, also in the long-term; (ii) both year 2 and 3 participants benefited from the treatment, without significant differences between the two groups; and (iii) linguistic specificity did not have an influence on the effectiveness of the task, whereas cultural specificity seemed to affect the retainment of the target forms, although no statistical evidence was found. In order to obtain sounder results, it is thus desirable to repeat the experiment involving larger samples of participants with different L1 backgrounds and a larger number of culturally-specific SBUs.

In addition, we also observed that the overall scores were generally low and that the responses were often inaccurate. Although the scarce exposure to the target forms may surely have contributed, our hypothesis is that these poor outcomes should be attributed to the particular composition of FL/L2 learners' mental lexicon and linguistic processing mode. As discussed in Section 5, post-childhood language learning differs from L1 acquisition in that FL/L2 learners tend to process language analytically. This preferred word-by-word processing mode has relevant consequences from the pedagogical perspective. In order to enhance the retention of formulaic sequences, it seems necessary to "look for a way of accommodating

analitycity and formulaicity" (Wray 2000: 483). The identification task adopted in this study prompted the participants to learn the target items as unanalysed strings, that is, holistically. However, paraphrasing Boers and Linstromberg (2012: 100), if it is true that learners find it hard to bypass word-by-word analysis anyway, then it might be more beneficial to focus their attention on the linguistic form of the sequences.

Mnemonic techniques based on the cognitive engagement with the FSs have been proposed in the literature on formulaicity in FL/L2. Such techniques include that proposed by Boers and Lindstromberg (2005), Boers et al. (2012), and Lindstromberg and Boers (2008a, 2008b), based on sound repetition. Another technique proposed by Boers and other scholars is etymological elaboration, which consists in helping learners appreciate the literal sense of figurative idioms by reactivating their literal sense (e.g. Boers 2001; Boers et al. 2004). In CFL, etymological elaboration has been successfully applied to Chinese four-character idiom teaching (see Conti 2020). The problem with SBUs is that they are not figurative (except loaded SBUs) and do not show such features as sound repetition, being often indistinguishable from their freely-generated counterparts. Therefore, hypothetically, techniques such as those just mentioned are not applicable to SBU teaching. A possible solution might be to resort to contrastive analysis and translation, which has already given encouraging results for the retention of English collocations (Laufer and Girsai 2008). Future research will hopefully shed more light on the crucial task of teaching Chinese SBUs effectively.

Appendix

(Translated from Italian) Read the dialogues and find the expressions used in the following situations:

1. Greet a friend that you haven't met for a long time: _____
2. Ask a stranger for a lighter: _____
3. Apologize for answering the phone: _____
4. Ask if a seat is available: _____
5. Greet a friend who is leaving when outside: _____
6. Tell someone that they called the wrong number: _____

一，在开会中 During a meeting
    A: 大家还有什么意见吗？
    B: 我想，下个星期的活动……
    C: 不好意思，我接个电话。
    A: 好，你接吧。
    B: 开会还接电话，太没礼貌了！

> A: 也没什么, 有可能是很重要的电话。
> C: 好消息好消息啊! 我当爷爷了!

二, 在咖啡馆 At the cafeteria
> A: 人太多了, 没位子,我们换一个地方吧。
> B: 哦,你看,那边有空椅子, 你过去问问。
> A: 您好,请问这儿有人吗？
> C: 没有。
> A: 我可以拿走吗？
> C: 可以啊
> A: 谢谢！

三, 给老师打电话 Calling a teacher on the phone:
> A: 喂？
> B: 喂, 您好,请问是王老师吗？
> A: 我就是。
> B: 我是您的学生谢志。
> A: 谢志,你好, 你找我有什么事吗？
> B: 我有点不舒服, 想请一天假,可以吗？
> A: 好的, 那你好好休息。

四, 在路上 In the street
> A: 你是不是小李？
> B: 是啊, 你是小王吧?好久不见了!
> A: 真高兴再见到你!你最近好吗？
> B: 还好, 你呢？
> A: 不错, 我考上博士了!
> B: 太好了, 恭喜你！
> A: 谢谢！

五, 在餐馆外面 Outside the restaurant
> A: 谢谢你请我吃饭。
> B: 不客气！
> A: 下次我请。
> B: 不用了!
> A: 好吧好吧, 天不早了, 我先回家了。
> B: 那行, 你回去吧 !路上注意安全！
> A: 下次见, 拜拜。

六，在机场 At the airport
    A: 小李！
    B: 小王, 你终于到了!怎么样, 累不累?
    A: 不累不累。
    B: 飞了一天还不 累?
    A: 真的不累, 我在飞机上睡得很好。
    B: 哈哈, 好吧!来, 我帮你拿行李吧！
    A: 不用啦！

七，在电话中 On the phone
    A: 喂, 是中意中心吗？
    B: 不好意思, 您打错了。
    A: 哦,是吗 ?我打的是059122865387。
    B: 不是, 我这里是159122865387。
    A: 哦,真抱歉, 打扰了。
    A: 没关系, 再见!
    B: 再见！

八，在外面 Outside
    A: 怎么了?
    B: 我又忘带打火机了!
    A: 你呀!那边有人在抽烟, 你去找人家借一下吧。
    B: 朋友, 能借个火吗？
    C: 当然, 给。
    B: 谢谢！
    C: 不客气。

References

Beebe, Leslie M. & Martha C. Cummings. 1996. Natural speech act data versus written questionnaire data: How data collection method affects speech act performance. In Susan Gass & Joyce Neu (eds.), *Speech acts across cultures*, 65–86. Berlin: Mouton de Gruyter.

Bell, Nancy. 2012. Formulaic language, creativity, and language play in a second language. *Annual Review of Applied Linguistics* 32. 189–205.

Bi, Xiaofang & Chunhong Shi. 2013. Yukuai de xingzhi ji hanyu yukuai xitong de cengji guanxi [The nature of chunks and the gerarchical relationships in the system of Chinese chunks]. *Dangdai Xiucixue* 3. 32–46.

Boers, Frank. 2001. Remembering figurative idioms by hypothesising about their origin. *Prospect* 16(3). 35–43.

Boers, Frank & Murielle Demecheleer. 2001. Measuring the impact of cross-cultural differences on learners' comprehension of imageable idioms. *ELT Journal* 55(3). 255–262.

Boers, Frank & Seth Lindstromberg. 2005. Finding ways to make phrase-learning feasible: The mnemonic effect of alliteration. *System* 33. 225–238.

Boers, Frank & Seth Lindstromberg. 2012. Experimental and intervention studies on formulaic sequences in a second language. *Annual Review of Applied Linguistics* 32. 83–110.

Boers, Frank, Murielle Demecheleer & June Eyckmans. 2004. Etymological elaboration as a strategy for learning idioms. In Bogaards Paul & Batia Laufer (eds.), *Vocabulary in a second language*, 53–78. Amsterdam: John Benjamins.

Boers, Frank, June Eyckmans, Jenny Kappel, Hélène Stengers & Murielle Demecheleer. 2006. Formulaic sequences and perceived oral proficiency: Putting a lexical approach to the test. *Language Teaching Research* 10(3). 245–261.

Boers, Frank, Seth Lindstromberg & June Eyckmans. 2012. Are alliterative word combinations comparatively easy to remember for adult learners? *RELC Journal* 43(1). 127–135.

Conklin, Kathy & Norbert Schmitt. 2008. Formulaic sequences: Are they processed more quickly than nonformulaic language by native and nonnative speakers? *Applied Linguistics* 29(1). 72–89.

Conti, Sergio. 2019. *Chengyu. Caratteristiche e apprenndimento delle espressioni idiomatiche cinesi* [Chengyu. Characteristics and learning of Chinese idioms]. Padua: libreriauniversitaria.it Edizioni.

Conti, Sergio. 2020. Etymological elaboration in *chengyu* 成语 teaching: The role of opacity, type of instruction, and competence level. *Annali di Ca' Foscari. Serie Orientale* 56. 411–438.

Conti, Sergio & Carmen Lepadat. In preparation. Situation-bound utterances in italiano e in cinese: un cofronto tra parlanti nativi e apprendenti italofoni [Situation-bound utterances in Italian and Chinese: comparing native speakers and Italian learners]. In Chiara, Romagnoli & Sergio Conti (eds.), Didattica e acquisizione del cinese in Italia, Rome: Roma TrE-Press.

Conti, Sergio & Carmen Lepadat. n.d. Discourse Completion Test (IT/中/EN). CFL Research Network. https://www.cflnetwork.org/wp-content/uploads/2019/05/test-conti-lepadat.pdf (accessed 15 July 2020).

Dohrenwend, Barbara S. 1965. Some effects of open and closed questions on respondents' awareness. *Human Organization* 24. 175–184.

Dörnyei, Zoltán. 2007. *Research methods in applied linguistics*. Oxford: Oxford University Press.

Durrant, Philip & Norbert Schmitt. 2010. Adult learners' retention of collocations from exposure. *Second Language Research* 26. 163–188.

Field, Andy, Jeremy Miles & Zoë Field. 2012. *Discovering statistics using R*. London: SAGE Publications Ltd.

Jiang, Nan & Tatiana M. Nekrasova. 2007. The processing of formulaic sequences by second language speakers. *The Modern Language Journal* 91(3). 433–455.

Jones, Martha & Sandra Haywood. 2004. Facilitating the acquisition of formulaic sequences: An exploratory study in an EAP context. In Norbert Schmitt (ed.), *Formulaic sequences: Acquisition, processing and use*, 269–300. Philadelphia: John Benjamins.

Kecskés, Istvan. 2000a. A cognitive-pragmatic approach to situation-bound utterances. *Journal of Pragmatics* 32. 605–625.

Kecskés, Istvan. 2000b. Conceptual fluency and the use of situation-bound utterances in L2. *Links & Letters* 7. 145–161.

Kecskés, Istvan. 2003. *Situation-bound utterances in L1 and L2*. Berlin: De Gruyter.

Kecskés, Istvan. 2010. Situation-bound utterances as pragmatic acts. *Journal of Pragmatics* 42. 2889–2897.

Kecskés, Istvan. 2014. The evaluative function of situation-bound utterances in intercultural interaction. In Geoff Thompson & Laura Alba-Juez (eds.), *Evaluation in context*, 137–151. Amsterdam: John Benjamins.

Kecskés, Istvan. 2016. Situation-bound utterances in Chinese. *East Asian Pragmatics* 1(1). 107–126.

Kwon, Jihyun. 2004. Expressing refusals in Korean and in American English. *Multilingua* 23. 339–364.

Labben, Afef. 2016. Reconsidering the development of the discourse completion test in interlanguage pragmatics. *Pragmatics* 26(1). 69–91.

Laufer, Batia & Nany Girsai. 2008. Form-focused instruction in second language vocabulary learning: A case for contrastive analysis and translation. *Applied Linguistics* 29(4). 694–716.

Li, Hui. 2013. Jiyu guanlian leixing de hanyu yukuai fenlei tixi tantao [A study of Chinese language chunks classification system based on correlation type]. *Xihua Shifan Daxue Xuebao (Zhexue Shehui Kexue ban)* 2. 99–103.

Lindstromberg, Seth & Frank Boers. 2008a. The mnemonic effect of noticing alliteration in lexical chunks. *Applied Linguistics* 29(2). 200–222.

Lindstromberg, Seth & Frank Boers. 2008b. Phonemic repetition and the learning of lexical chunks: The power of assonance. *System* 36. 423–436.

Liu, Dilin. 2008. *Idioms: Description, comprehension, acquisition, and pedagogy*. New York & London: Routledge.

Mair, Patrick & Rand Wilcox. 2019. Robust statistical methods in R using the WRS2 package. *Behaviour Research Methods*. https://doi.org/10.3758/s13428-019-01246-w (accessed 15 July 2020).

Pang, Yang. 2020. The cognitive saliency of word associations of verbs of speech in English as a Lingua Franca interactions. *Intercultural Pragmatics* 17(4). 417–443.

Pawley, Andrew & Francis H. Syder. 1983. Two puzzles for linguistic theory: Nativelike selection and nativelike fluency. In Jack C. Richards & Richard W. Schmidt (eds.), *Language and communication*, 191–226. New York: Longman.

Pellicer-Sánchez, Ana & Frank Boers. 2019. Pedagogical approaches to the teaching and learning of formulaic language. In Anna Siyanova-Chanturia & Ana Pellicer-Sánchez (eds.), *Understanding formulaic language: A second language acquisition perspective*, 153–173. New York & Abingdon: Routledge.

Peters, Elke. 2012. Learning German formulaic sequences: The effect of two attention-drawing techniques. *The Language Learning Journal* 40(1). 65–79.

Phakiti, Aek. 2014. *Experimental research methods in language learning*. London & New York: Bloomsbury.

Riazi, A. Mehdi. 2016. *The Routledge encyclopedia of research methods in applied linguistics: Quantitative, qualitative, and mixed-methods research*. Abingdon & New York: Routledge.

Schmitt, Norbert. 2000. *Vocabulary in language teaching*. Cambridge: Cambridge University Press.

Schmitt, Norbert. 2010. *Researching vocabulary: A vocabulary research manual*. New York: Palgrave Macmillan.

Sinclair, John. 1991. *Corpus, concordance, collocation*. Oxford: Oxford University Press.

Siyanova-Chanturia, Anna & Ana Pellicer-Sánchez. 2019. *Understanding formulaic language: A second language acquisition perspective*. New York & Abingdon: Routledge.

Stengers, Hélène, Frank Boers, Alex Housen & June Eyckmans. 2010. Does "chunking" foster chunk-update? In Sabine De Knop, Frank Boers & Antoon De Rycker (eds.), *Fostering language teaching efficiently through cognitive linguistics*, 99–117. Berlin: Mouton de Gruyter.

Sun, Weizhang. 1989. *Hanyu Shuyuxue* [Chinese phraseology]. *Changchun: Jilin Jiaoyu Chubanshe*.

Taguchi, Naoko, Shuai Li & Feng Xiao. 2013. Production of formulaic expressions in L2 Chinese: A developmental investigation in a study abroad context. *Chinese as a Second Language Research* 2(1). 23–58.

Wilcox, Rand R. & Tian S. Tian. 2011. Measuring effect size: A robust heteroscedastic approach for two or more groups. *Journal of Applied Statistics* 38(7). 1359–1368.

Wray, Alison. 2000. Formulaic sequences in second language teaching: Principles and practice. *Applied Linguistics* 21(4). 463–489.

Wray, Alison. 2002. *Formulaic language and the lexicon*. Oxford: Oxford University Press.

Wray, Alison & Michael R. Perkins. 2000. The functions of formulaic language: An integrated model. *Language & Communication* 20. 1–28.

Yang, Jiang. 2016. CFL learners' recognition and production of pragmatic routine formulae. *Chinese as a Second Language* 51(1). 29–61.

Yeh, Shu-Han. 2016. *The teachability of situation-bound utterances in modern Chinese as a foreign language context*. New York: State University of New York at Albany Ph.D. thesis.

Yin, Jiping. 2013. Jin nian lai yukuai yanjiu zai Zhongguo de fazhan, wenti yu quxiang: Jiyu CSSCI kanwu de wenxian fenxi [Development, issues and tendencies of recent researches of language chunks in China – A survey of the CSSCI journal articles]. *Kunming Ligong Daxue Xuebao (Shehui Kexue Ban)* 13(6). 102–108.

Zhou, Honghui. 2012. A study of situation-bound utterances in Modern Chinese. *Chinese as a Second Language Research* 1(1). 55–86.

Bionotes

Sergio Conti
Department of Foreign Languages, Literatures and Cultures, Roma Tre University, Rome, Italy
sergio.conti@uniroma3.it

Sergio Conti is currently an adjunct professor of Chinese language at Roma Tre University (Italy), Siena University (Italy), and the University of Naples "L'Orientale". He obtained his Ph.D. degree (with honors) at Sapienza University of Rome (Italy), discussing an experimental thesis on Chinese idiom teaching. In 2019-2020, he was temporary research fellow at Roma Tre University. His

research interests lie in the area of Chinese acquisition and teaching, with particular regard to formulaicity, vocabulary, and pragmatic competence.

Carmen Lepadat
Italian Institute of Oriental Studies Department, Sapienza University of Rome, Rome, Italy
carmen.lepadat@uniroma1.it

Carmen Lepadat is a Ph.D. candidate at Sapienza University of Rome (Italy), with a research project on right dislocations in spoken Chinese. She obtained her master's degree (with honors) at Roma Tre University (Italy), discussing a thesis on the Chinese modal particle *ma* 嘛. Her research interests include Chinese pragmatics, information structure and modality, and language modeling.