

Universidad de Puerto Rico
Recinto de Rio Piedras
Facultad de Humanidades
Programa Graduado de Lingüística

L2 Acquisition of Splitable Compounds of Mandarin Chinese: The “Split” vs “Unsplit” Use

Leisha Marie Guzmán Torres
801-10-3096

Monografía final de grado
Fecha de aprobación: 1 de diciembre de 2020

Comité de monografía final:

Dra. Meili Deng (directora)

Dr. Nicholas Faraclas (lector)

Dr. Patrick-André Mather (lector)

Index

1.	Introduction.....	1
2.	Literature Review.....	1
	2.1 The Mandarin Chinese lexicon.....	1
	2.2 What is considered a compound word in Chinese?.....	2
	2.2.1 How is the concept of “word” defined?.....	2
	2.2.2 What are “compound words” in Chinese?.....	4
	2.3 The debate on collocations.....	4
	2.4 What is a Splittable Compound (líhé cí)?.....	5
	2.5 Previous studies examining acquisition of SCs.....	7
	2.6 Objectives of the current study	8
3.	Methodology.....	9
	3.1 Research Question	10
	3.2 Participants.....	10
	3.3 Design.....	11
	3.3.1 Selection of Verbs.....	11
	3.3.2 Designing the survey.....	13
	3.4 Procedure	15
	3.5 Data Analysis.....	15
4.	Results.....	16
5.	Discussion, Conclusion and Recommendations for Further Research.....	22
	Bibliography.....	25
	Appendices.....	30
	Appendix 1: The survey.....	30
	Appendix 2: Selected Verbs.....	40
	Appendix 3: Coding Sheet.....	42
	Appendix 4: Two variable t-test results.....	48

Tables and Graphs

Table 3.2.1 Division of Participant Groups.....	10
Table 3.2.2 Participants in the Control Group.....	11
Table 4.1 Accuracy in Participant’s use of Split entries.....	16
Table 4.2 Accuracy in Participant’s use of Unsplit entries.....	16
Table 4.3 Selection of ungrammatical entries.....	17
Table 4.4 The Control Group’s use of Split vs Unsplit entries.....	17
Table 4.4.1 The Control Group’s use of Ungrammatical entries.....	18
Graph 4.5 Acquisition Pattern of Split and Unsplit Uses by Group.....	18
Table 4.6 Two sample t-test results for correlations between the Basic and Advanced Groups in their accuracy using SCs.....	20
Table 4.7 Two sample t-test results for correlations between the control group and Advanced group in their accuracy using SCs.....	21

List of Abbreviations

SCs- Splitable Compounds

L2- Second Language

N- Noun

V- Verb

O- Object

Adj.-Adjective

Pron.- Pronoun

ASP- Aspect marker

MOD- Modifier

Q- Question marker

QUANT- Quantifier

About the researcher

Leisha Guzmán studied at University of Puerto Rico, Rio Piedras Campus, where she graduated with a bachelor's degree of Art in Creative Writing. She has been an English as a Second Language teacher for adults since 2014. As a recipient of the Critical Language Scholarship Program to study Chinese in Suzhou, China in 2017, she experienced firsthand the challenges her students face when acquiring a new language. She has had experience in Second Language Acquisition from both a pedagogical and learner perspective.

Dedication

I would like to thank Kathy Nevarez and Veronica Duque for all their help and encouragement. Your friendship is a beacon of light in my life. Thank you.

Another special thanks goes to my project director, Prof. Meili Deng. Thank you for pushing through with me until the end.

Finally, I would like to thank Prof. Patrick Mather and Prof. Nicholas Faraclas for reading my monograph and giving me helpful insights.

Abstract

Splitable Compounds (SC) of Chinese are VO construction that can be used as both a compound word (in its Unsplit form) and a phrase (in its Split form). It is known that learning complex structures poses a challenge for L2 learners. Due to this duality, SCs could be especially difficult to acquire. This study aimed to answer: “Does the language proficiency of the learner affect the acquisition of the Split vs Unsplit uses of SCs?”.

For this, three groups of participants at different proficiency levels of Mandarin Chinese answered an online survey which consisted of multiple-choice elicitation tasks. Significant improvement was seen from Group 1 to Group 3 in their accuracy using the Unsplit form of SCs, but not when using the Split (phrase/collocation) form. Thus, SCs Split and Unsplit forms are not acquired evenly. One (the Unsplit) is much easier to acquire than the other (the Split).

Keywords: second language acquisition, Mandarin Chinese second language acquisition, acquisition of splitable compounds, acquisition of complex structures, acquisition of *like ci*

Resumen

Los Compuestos Separables (CS) del chino mandarín son construcciones de Verbo+Objeto que pueden ser utilizados como una palabra compuesta (cuando ambas partes aparecen unidos) y, y a la vez, una frase (cuando aparecen ambas partes separadas). Es reconocido en la literatura que aprender estructuras complejas posa un gran reto para aprendices de una L2. Debido a su dualidad los CSs pudiesen ser particularmente difíciles de adquirir. El objetivo de este estudio fue contestar la siguiente pregunta: “¿el nivel de fluidez de los aprendices afecta la adquisición de las formas Unidas y Separadas de los CS?”

Para este propósito, tres grupos de participantes a diferentes niveles de fluidez de chino mandarín contestaron una encuesta en línea que consistía en varias secciones con tareas de elicitación. Hubo mejoría significativa del Grupo Básico al Grupo Avanzado en su uso de la forma Unida de los CS, pero no significativa en su uso de la forma Separada. Por lo tanto, la forma Unida y Separada de los CS no se adquieren equitativamente. La forma Unida es considerablemente más sencilla de adquirir que la forma Separada.

Palabras clave: adquisición de segundas lenguas, adquisición del chino mandarín como segunda lengua, adquisición de compuestos separables, adquisición de estructuras complejas, adquisición de *lihe ci*

L2 Acquisition of Splittable Compounds of Mandarin Chinese: The “Split” vs “Unsplit” Use

1. Introduction

There has been a long-standing interest in the Splittable Compounds (SCs) of Mandarin Chinese. Decades of research have focused on the question of whether SCs are compound words, phrases, or collocations (Arcodia, 2007; Cai, 2017; Siewierska et al., 2010; Feng, 1998). The multiple classification of these compounds might cause learners of Chinese as an L2 to struggle more when acquiring them, given that that learning collocations in general tends to be difficult, and they are important for achieving fluency (Cai, 2017; Palmer, 1981; Nesselhauf, 2003; etc.).

Understanding the acquisition of SCs (known as 离合词 (*líhé cí*) in Chinese) in Second language learners is one of the primary objectives of this study. For this purpose, a survey was designed to study the accuracy of learners of Mandarin Chinese when using both Split and Unsplit forms of SCs. In the process of defining SCs, it is necessary to discuss the different perspectives on the distinctions made between compound word, collocation, and phrase.

2. Literature Review

2.1 The Mandarin Chinese lexicon

Several studies have shown that there have been many debates about what constitutes a word, a compound, a collocation, and a phrase. This is especially true in Mandarin Chinese as compounding seems to be such a common phenomenon. The Modern Chinese Dictionary (1996 edition) has 2960

entries of compound words, which comes to about 51% of all verb-object words (Siewierska et al., 2010; Shi, 1999; Zhu, 2006).

However, according to Cai (2017) “The updated 2002 version of the dictionary had 3,326 splittable compounds” (p.9). Verb-object words, in turn, constitute 97% of all compounds in Chinese (Zhu, 2006; as cited in Siewierska et al., 2010). Definitions of compound words range from words created by the union of two or more bound stems to those created by two or more existing, independent words. This definition makes classifying some terms as compounds challenging. Moreover, Bauer (2005: 106-107, cited in Arcodia, 2007) makes the following observation: “[G]iven the difficulty that there has been for many years in defining a word, it is not surprising that there should be difficulty with the borderline of compounding. Items which fit poorly into the category of word should also fit poorly in the category of possible compound element” (p.80).

So, what is an SC? Three lines of research are relevant to this question: 1) What is a compound word in Chinese?, 2) What is a phrase?, and 3) What is a collocation?

2.2 What is considered a compound word in Chinese?

2.2.1 How is the concept of “word” defined?

Before we define a compound word, we must discuss a smaller unit: “word”. There has not been much consensus among linguists as to what constitutes a “word”. This situation is made even more difficult by the fact that there are no inflectional markers in Chinese to help us identify word boundaries (Arcodia, 2007).

The issue with SCs stems from the fact that in Western linguistics, words cannot be discontinuous. However, due to the disyllabic nature of most of the lexicon of Modern Chinese, words are often composed of two lexical morphemes bound to each other.

“In fact, the vast majority of Chinese morphemes have a lexical nature, and the great part of them are bound, which are termed by Packard (2000: 77-78) “bound roots”; these may be compared to the so-called “neoclassical constituents” of Standard Average European languages (henceforth, SAE), such as *philo-*, *-logy* or *-phobia*, having lexical (rather than grammatical) meaning and always bound to some other constituent. ” (Arcodia, 2007, p. 81)

That is, most words are composed of one lexical root always found bound to another constituent (Arcodia, 2007), such as (1) below. This is contrasted with a monosyllabic word, like in (2).

(1) 食

Shí

‘to eat, food’

食欲

Shíyù

eat+desire

‘appetite’

(2) 吃

Chī

‘to eat’

吃鸡。

Chī jī

Eat V chicken N

‘to eat chicken’

As seen in example (1) above, although “shi”(食) means ‘to eat’, it is found as part of a disyllabic word (‘appetite’) as a bound root. It cannot be separated. In SCs, on the other hand, their lexical morphemes can be split, thus breaking with the traditional Western perspective of words not being discontinuous in nature.

2.2.2 What are “compound words” in Chinese?

The term compound “is used as a cover term for a collection of related, but not necessarily identical, phenomena in the literature,” (Dai, 1998, p.125, as quoted in Arcodia, 2007, p.80). This also means there is no distinction between related constructions such as words composed of two or more bound stems and those comprised of two or more words (Arcodia, 2007).

Packard (2000) (quoted in Zhang et al., 2012, p.755) proposed that a good way to analyze compounds in Chinese is based on the form-class of the morpheme. As such, compound nouns can be NN, NV, VN, VV; while compound verbs can be VN or VV. He also used traditional Chinese grammar to define the compounds. Meaning that SCs do not fall under this category in traditional Chinese grammar since they are VO compound verbs.

2.3 The debate on collocations

Since the coinage of the term “collocations” in 1957 by John Rupert Firth, there have been many linguists debating its definition. Researchers have discussed collocations in numerous ways, from Palmer (1993) “restrict[ing] the definition of collocations to idioms whose meaning is not obvious from its

components, such as blow the gaff” (Cai, 2007, p.4), to opposing views that categorize collocations as distinct from idioms and phrasal verbs altogether (Hill, 2000) (quoted in Cai, 2007, p.6).

Two of the main approaches used to define collocations were summarized by Nesselhauf (2005) (quoted in Cai, 2007, p.6) as being the frequency-based approach and the phraseological approach. The frequency-based approach defines collocations based on the frequency of co-occurrence. On the other hand, the phraseological approach considers the syntactic relationship between the elements of the collocations. From a syntactic view, “[Hausmann (1989)] went as far as to conclude that only six collocational combinations appear in a pre-defined set of syntactic relations; these are: adjective+noun, noun+verb, noun+noun, adverb+adjective, verb+adverb, verb+noun.” (Cai, 2007, p.6).

However, other researchers take a completely different approach and perceive collocations as lexical in nature, not grammatical (Shui, 2007; Arcodia, 2007; and Packard, 2000). Thus, the definition of collocations is still widely debated and encompasses a myriad of different phenomena. This, in turn, means, that SCs could arguably be seen as collocations from some perspectives. However, it is important to note that SCs and collocations have different names in Chinese, “Collocations” translates to 搭配 (dāpèi), while “Splittable Compound” translates to 离合词 (líhé cí). So, they will not be considered as the same phenomenon in this study.

2.4 What is a Splittable Compound (*líhé cí*)?

SCs are very common in Chinese. According to Zhang et al. these separable compounds are a special type of compound verb (2012). Packard (2003, as cited in this Siewierska et al., 2010, p.467) states that, when it comes to grammar in Mandarin Chinese, the verb-object paradox cannot be ignored.

Many attempts have been made to establish criteria for classifying and defining these VO constructions (Siewierska et al., 2010; Lu, 1957; Packard, 2003; Lu, 1979; Zhu, 1982).

In order to understand the debate, we must understand how SCs work. Like compounds in other languages such as English and German, Chinese SCs have a “head” and a “tail” that can be separated into two independent lexemes. However, in SCs when the *head* and the *tail* appear together, they are considered a word, such as in (3) and (5a); yet, they exhibit qualities of a phrase (without affecting the semantics of the word) whenever the *head* and the *tail* are interposed by other elements, as shown in (4) and (5b).

(3) 睡觉 *Shuìjiào*, sleep V sleep N ‘to sleep’

(4) 睡了四个小时的觉 *Shuìle sì gè xiǎoshí de jiào*, sleep V ASP Marker four Quantifier hour N sleep N, ‘I slept four hours’

In other words, an SC’s verb may function as transitive because SCs can take objects (direct objects, subordinate clauses, etc), for example (5) despite their intransitive use being more common (Li & Thompson, 1981, quoted in Siewierska et al., 2010, p.466).

(5) a. 吃饭

Chīfàn

Eat V rice N

‘to eat’

(5) b. 吃了他做的饭

chīle tā zuò de fàn

Eat V ASP he doV MOD rice N

‘ate his cooking’

Due to its duality, Lu (1957) (quoted in Siewierska et al., 2010, p.466) coined the term *líhé cí* to define this type of construction that is something in between a word and a phrase.

Despite much debate, though, the overall view remains that which we have discussed: SCs are a word when the *head* and *tail* are together and a phrase when they are apart (as seen in examples (3), (4), and (5) above) (Lu, 1957; Zhu, 1982, among others). As can be expected, this duality makes SCs challenging to acquire for learners.

2.5 Previous studies examining acquisition of Splittable Compounds

Most studies on acquisition of *líhé cí* have been conducted in Chinese. However, some scholars have started conducting research in English, making some of the available literature’s main take-aways accessible to non-Mandarin speakers. One such scholar is Cai (2017) who summarized that most of the research done on Splittable Compounds has focused on the acquisition of SCs different split forms.

“For instance, Ma (2008) discovered that learners' acquisition of different discrete forms of splittable words did not develop at equivalent levels, and also did not align with the improvement in their language ability. Moreover, the acquisition of split forms was slower than the acquisition

of unsplit forms at all three proficiency levels. Zhou and Li (2015) carried out an empirical study on the effect of synchronous teaching and sequential teaching on the acquisition of splittable words by Chinese language learners, and concluded that learners exposed to sequential teaching performed better.” (Cai, 2007, p.14)

2.6 Objectives of the current study

The goal of the present study is to contribute to the knowledge of whether proficiency level influences correct use of both Split and Unsplit forms of high frequency SCs (离合词 *líhé cí*). Our research question is “does the language proficiency of learners affect the correct use of Split vs Unsplit *líhé cí*?”

“A major aim of existing research on L2 collocations has been to identify the problems in learners' production and discover the causes of these problems. The findings of the previous studies indicate that three factors are related to the use of collocations: first, the influence of L1; second, intralingual factors, such as word frequency; third, learners' language proficiency.” (Cai, 2007, p.15)

In this study, the influence of learner's L2 proficiency level and of word frequency on accuracy of both Split and Unsplit SCs was researched. Since “...Zhang (1993) and Al-Zahrani (1998) argued that the use of collocations is related to the L2 learner's proficiency level, while Bahns and Eldaw (1993), and Howarth (1998) argued that there is no relationship between the use of collocations and proficiency” (Cai, 2007, p.17) the following hypothesis was formulated: learner's proficiency level is related to their accuracy in the use of Split and Unsplit *líhé cí*.

In order to test the hypothesis that learner's L2 proficiency level is related to the production of L2 SCs (*líhé cí*), a survey was designed to test the accuracy of learners' production of both forms.

In this literature review, we looked at the different perspectives of linguists concerning the prototypical word in Chinese, the compound nature of the Chinese lexicon, the category of collocations, and previous studies looking at the L2 acquisition of *líhé cí*.

In the next section, we will discuss the design of the survey and the selection of the participants in a study on the acquisition of SC by second/foreign language learners of Mandarin. This will be followed by a discussion of the collected data. From there, conclusions about our hypothesis will be drawn, and recommendations for future studies will be made.

3. Methodology

In the previous section, a review of the existing literature on “word, compound word, and collocation” was conducted to provide a framework for defining **Splittable Compounds in Chinese**¹ (SCs). This study focuses on Splittable VO Compounds since their specialized use leads us to believe that learners experience difficulty in acquiring them, given the debated status of SCs and their classification as being somewhere in between a compound word and a phrase.

The objective of this study is to further explore the acquisition of SCs by learners of Chinese at various L2 proficiency levels.

¹ As for the definition of “Splittable Compound” used in this study, we adopt the view put forth by various Chinese scholars which describes SCs as somewhere in between a compound and a phrase.

3.1 Research Question

This study tested high frequency SCs found in the textbook *New Practical Chinese Reader Textbook 1 & 2* to explore the relationship between verb-object SCs and L2 proficiency levels.² The aim of this study is to answer the research question previously mentioned in this paper: does the language level of the learners affect their correct use of Split vs. Unsplit SCs?

Answering this research question will provide more evidence for the correlation, or lack thereof, between L2 proficiency level and accurate use of complex structures such as SCs.

3.2 Participants

Table 3.2.1 Division of Participant Groups

Group	Length of Study	Participants	Age	Sex
Basic	1 semester	4	20-32	1 female, 3 male
Intermediate	3 semesters	4	17-24	3 female, 1 male
Advanced	5 semesters	4	22-28	4 female

The participants recruited for this research project were students who self-identified as advanced or native English speakers who were studying or had studied Chinese at the university level in any university in the world. The participants were divided into three groups. As with Cai's study (2017), membership in each of the three groups was determined by how many semesters they had studied. The

² Using the SCs that appear in one of the popular textbooks for teaching Chinese ensures that the selected compound verbs are taught within the first few semesters of study. In other words, learners have a higher chance of encountering them within the first two years of language study.

Basic group had participants who had studied Mandarin Chinese in a formal classroom environment for 1 semester; the Intermediate group had studied for 3 semesters, and the Advanced group for 5 or more semesters. Each group had 4 participants in it.

Participants ranged between 17-32 years of age and majored in widely different fields. Additionally, 6 of the participants had experience studying in Mainland China, and 1 in Taiwan.³

Table 3.2.2 Control Group

Control Group	Participants	Age	Sex
Native speakers	4	18-38 years-old	2 female, 2 male

The control group⁴ consisted of 4 native speakers of Mandarin Chinese who were around the same age as the participants in the experimental groups.

3.3 Design

3.3.1 Selection of Verbs

The 17 SCs used in this study were selected from the glossary of the *New Practical Chinese Reader Textbook 1 & 2*, as shown in examples (1), (2), (3), (4), and (5) below. This was done to ensure

³ In the Basic group, none of the participants had gone to China. In the Intermediate group, all 4 had gone. While, in the Advanced group, 3 out of 4 participants had gone to the target language country.

⁴ The native speaker participant pool was very diverse, including speakers from Taiwan, Hong Kong, and Mainland China. However, the native speakers selected did not mention where they were from.

that there was a high probability that participants would have encountered the verbs selected during their studies at university.

1. 下雨

啊，下雨了！

Xià yǔ

A, xià yǔle!

V Fall + N rain

V rain ASP marker

‘to rain’

‘Oh, it’s raining’

2. 上课

我3点上课

Shàngkè

Wǒ 3 diǎn shàngkè

V Go up + N class

I 3 o’clock take class

‘to take class’

‘I have class at 3 o’clock’

3. 帮忙

请帮我忙

Bāngmáng

qǐng bāng wǒ máng

V Help + Adj busy

Please V help Pron me Adj busy

‘to help’

‘Please help me out’

4. 吃饭

我饿了，要吃饭了

Chīfàn

wǒ èle, yào chīfànle

V Eat + N rice

I hungry ASP marker, want V eat ASP

‘to eat’

marker

‘I’m hungry, I’m going to eat’

5. 看书

你看了这本书吗？

Kànshū

nǐ kànle zhè běn shū ma?

V look + N book

You V see ASP marker this Quantifier N

‘to read’

book Q marker

‘Have you read this book?’

3.3.2 *Designing the survey*

In order to test the participants’ knowledge of SCs, a survey was designed on Google Forms⁵ to elicit their use of Splittable Compounds. The survey consisted of three parts: 1) a participant background questionnaire, 2) an exercise asking participants to select the correct answers after reading a short scenario, such as in (1), and 3) an exercise asking participants to select the correct answers based on the actions shown in a Graphics Interchange Format (GIF), such as in (2). The sections of the language test consisted of 10 exercises each. All the entries had at least one correct use of both the Split and Unsplit form of one of the 17 SCs used in this study and a minimum of one incorrect use of the Unsplit form, as shown in Appendix 1.⁶

⁵ The Republic of China blocks all Google services from being accessed on the Mainland. For this reason, having created the survey on Google Forms posed a challenge when trying to get participants residing in Mainland China to successfully access the survey. Additionally, time constraints did not allow for the survey to be reformatted and re-approved by the Institutional Research Board (IRB) to overcome this hurdle.

As a work-around to this constraint, participants were mainly from Taiwan and Hong Kong. This is due to the fact that they can easily access the survey on Google Forms. However, there is a noticeable dialectal difference between speakers of the Taiwanese, Hong Kongese, and Standard dialects of the language. Usually the Standard Language is the one which is taught to learners. This means that, while some Unsplit uses of SCs were considered to be incorrect for this test, some natives found them to be grammatically acceptable in their dialect, while some of the ones that were deemed “correct” for this study were considered ungrammatical in the dialects of these native speakers. Despite such dialectal differences, though, the natives performed better than the learners.

⁶ An English translation option was also offered. First, this option was offered because there were no translation tasks due to the space constraints of the present monograph. Whether they needed the English translation or not can aid us in measuring learners’ understanding of the vocabulary. Thus, it can work as a work-around for not having a full translation exercise to test participants’ understanding of the vocabulary. Second, the translation option was included to aid participants in answering accurately when they did not know the vocabulary of the prompt.

(1) 请问，力波在这儿吗？

Qǐngwèn, lì bō zài zhè'er ma?

Excuse me, Libo is here ma Q

Marker?

‘Excuse me, is Libo here?’

a. 不在，他在上汉语课。

Bùzài, tā zài shàng hànyǔ kè.

Is not, he is taking English

class.

‘No, he’s taking an English

class’

b. *不在，他在上课汉语。

Bùzài, tā zài shàngkè hànyǔ.

Is not, he is taking class

English.

c. *不在，他在汉语上课。

Bùzài, tā zài hànyǔ shàngkè.

Is not, he is at English taking

class.

d. 不在，他在上课。

Bùzài, tā zài shàngkè.

Is not, he is taking class.

(2)



你想做什么？

Nǐ xiǎng zuò shénme?

You want do what?

‘What do you want to do?’

a. *我想游泳一会儿。

Wǒ xiǎng yóuyǒng yīhuǐ'er.

*I want swim in a moment.

b. 我想游一会儿泳。

Wǒ xiǎng yóu yīhuǐ'er yǒng.

I want swimV in a moment

swimmingO.

c. 我想游泳。

Wǒ xiǎng yóuyǒng.

I want swim.

‘I want to swim’

‘No, he’s taking a class’

3.4 Procedure

All of the survey data was collected online. The survey was made available on Google Forms. The link for the survey was distributed online using social media and Prolific. The survey was carried out in the Spring semester of 2020. There was no time limit for completing the survey, but completion time ranged from 7 to 25 minutes.

3.5 Data Analysis

The data obtained from the survey was coded and then processed using SPSS. As previously mentioned, the dependent variables were the accuracy of SCs in their Unsplit and Split forms⁷, while the independent variable was time studied (proficiency level).

The accuracy per exercise for each participant was coded by the researcher herself following the coding sheet, which can be found in Appendix 3. Then, the data was analyzed to look at the average accuracy rate for each group. Finally, a t-test was done on the data in SPSS to look at the statistical significance of the correlations under study.

⁷ Unsplit: 唱歌 (Chànggē)- to sing

Split: 唱了一首歌 (Chàngle yī shǒu gē)- Sing a song

4. Results

The data was submitted to several different analyses in order to answer our research question. First, the total accuracy was summed up for every participant, and then it was summed up per group. Subsequently, the average of each group's accuracy was calculated. The results were as follows:

Table 4.1 Accuracy in the use of Split items

Group	Total Accuracy	Average of Total Correct	Average of Total Missed
Basic	43/80	54%	46%
Intermediate	51/80	64%	36%
Advanced	55/80	69%	31%

As can be seen in Table 4.1, accuracy for the 20 exercises ranged between 54-69%, while the average of total correct answers that participants missed per group ranged between 31-46%. An improvement of 15% can be seen from the Basic group to the Advanced group in terms of the accuracy of responses.

Table 4.2 Accuracy in the use of Unsplit items

Group	Total Accuracy	Average of Total Correct Uses	Average of Total Missed
Basic	39/80	49%	51%
Intermediate	49/80	61%	39%
Advanced	65/80	81%	19%

Table 4.2 demonstrates that for the Unsplit uses, a 32% improvement was observed from the Basic to the Advanced group. This improvement in accuracy for Unsplit items is slightly more than double that for the Split uses.

Table 4.3 Selection of ungrammatical items

Group	Total Incorrect Selected	Average of Total Incorrect Selected
Basic	53/128	41%
Intermediate	27/128	21%
Advanced	29/128	23%

In general agreement with the data in table 4.2, the data in Table 4.3 shows that the Intermediate and Advanced groups showed an average 19% decrease in selection of ungrammatically Unsplit items, when compared to the Basic group. However, the Intermediate group showed a minimally better performance than the Advanced group by 2%.

Table 4.4 The Control Group

Type of SC	Total Accuracy	Average of Total Correct	Average of Total Missed
Split Uses	73/80	91%	9%
Unsplit Uses	66/80	82%	18%

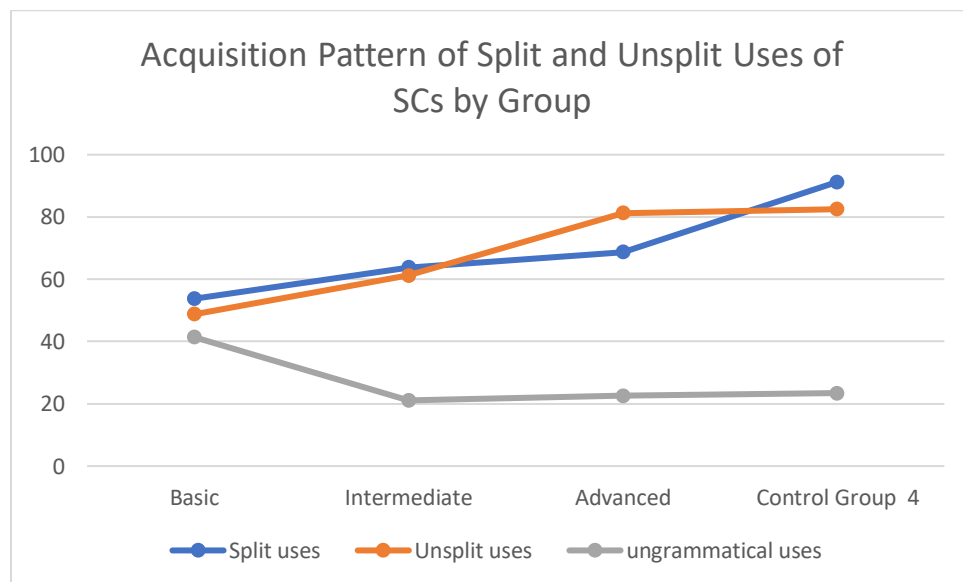
As shown in Table 4.4, the control group displayed a generally high degree of accuracy which exceeded that of all of the L2 groups in terms of Split SCs. However, the control group performed 9% better in their use of Split items than in their use of Unsplit items, performing similarly to the Advanced L2 group in this respect.

Table 4.5 The Control Group's selection of ungrammatical items

Ungrammatical Unsplit uses	Total ungrammatical uses selected	Average of ungrammatical uses selected
Incorrect Unsplit	30/128	23%

Finally, Table 4.5 the control group selected ungrammatical entries 23% of the time, at the same rate as the L2 Advanced group.

Graph 1. Acquisition Pattern of Split and Unsplit Uses by Group



As can be seen in Graph 1, the Basic group's results were minimally more accurate for the correct uses of the Split (54% accuracy) SCs than for the correct uses of the Unsplit (49% accuracy) SCs. They also selected ungrammatical entries more often than all other groups (41% selected an incorrect use of Unsplit SCs). In the Intermediate group, an improvement can be seen for both the correctly Split (64% accuracy) and Unsplit (61% accuracy) SCs. In hand with this improvement, a decrease in ungrammatical uses can also be seen in Intermediate group (20% less than the Basic group). Meanwhile, the Advanced group managed to surpass the performance of the Basic group in their use of correctly Unsplit SCs by 32%, but only by 15% in their use of Split entries. Not only is their improvement in for the Split SCs minimal, but their accuracy using the Split SCs is also 12% worse than for the Unsplit forms. The Advanced group also displayed a low number of ungrammatical uses of Unsplit SCs (only 23% selected incorrect uses).

Finally, the use of both correctly Split and Unsplit SCs was most accurate for the control group overall (91% accuracy for Split SCs; 82% accuracy for Unsplit SCs). Both the control group and the Advanced group were close in their use of Unsplit SCs (only 1% difference in accuracy). However, the control group did noticeably better than the Advanced group in their use of Unsplit entries, with a 22% difference. Their responses for ungrammatical uses of Unsplit SCs were similar to those of the Intermediate group.

The data was submitted to a two sample t-test on SPSS to determine if there were statistically significant differences between the Basic and Advanced groups in their use of Unsplit SCs and Split SCs, as well as in their responses for ungrammatical uses of Unsplit SCs. Two sample t-test was also used to determine if there was statistically significant difference between the Advanced group and the control group.

The results were as follows:

Table 4.6 Two sample t-test results for correlations between the Basic and Advanced Groups in their accuracy using SCs.⁸

	Accuracy for Basic:Advanced	Is there statistically significant difference?	Is there correlation between language level and accuracy?
Split Uses	43:55	No (.48)	No
Unsplit Uses	39:65	Yes (.04)	Yes
Incorrect Uses	53:29	No (.195)	No

Table 4.6 shows that the results from these analyses indicated that the improvement from the Basic group to the Advanced group in the use of Unsplit forms was significant. Thus, despite there not being a significant correlation between language level and accuracy for the Split uses, there was significant improvement from Basic group to Advanced group in the use of Unsplit SCs.

As for the differences seen in Table 4.6 between the mastery of Split vs Unsplit uses from the Basic to the Advanced groups, a t-test was performed to test if it was significant. However, the results turned out to be not significant with a result of $P=0.52$.

⁸ The details of the results of the two sample t-tests are in the Appendices.

Table 4.7 Two sample t-test results for correlations between the control group and Advanced group in their accuracy using SCs

	Accuracy for Advanced:Control	Is there statistically significant difference?	Is there correlation between language level and accuracy?
Split Uses	55:73	No (.21)	No
Unsplit Uses	65:66	No (.93)	No
Incorrect Uses	29:30	No (.93)	No

The difference between the control group and Advanced group was also not significant for all categories (Split, Unsplit, and ungrammatical uses) meaning that they both performed at similar levels in their use of SCs. These results could indicate that learners are acquiring the language of the young generations of Chinese, despite their textbooks teaching the standard. This could be explored in future research.

5. Discussion, Conclusion and Recommendations for Further Research

In this study, we investigated whether the number of semesters spent studying Mandarin Chinese as an L2 affects accuracy when using Split vs. Unsplit forms of Splittable Compounds in the language. To this end, we collected data on the use of SCs in three groups at different L2 proficiency levels in an elicitation task. As seen in the previous section, there is significant improvement from the Basic group (1 semester) to the Advanced group (5 or more semesters) for Unsplit SCs. However, results showed no corresponding significant improvement in learners' uses of Split SCs.

There are two possible explanations for these results. In the case of the Split SCs, our findings converge with previous findings, such as Bahns and Eldaw (1993), and Howarth (1998), which found that there is no significant correlation between L2 language proficiency and accurate use of collocations. Another possible reason is that the level of mastery of Mandarin Chinese among the L2 learners at all three levels in our sample was very low. At the time that we administered our instrument, the level of the Advanced group was more or less at HSK 3, which may have been insufficient for satisfactory performance in relation to Split SCs, which are generally understood to be difficult to acquire. If we had included an L2 group at HSK 5 or HSK 6, we might have seen a significant correlation between language proficiency and accurate use of SCs in their Split form.

On the other hand, the patterns of use of Unsplit SCs that we observed generally confirmed our hypothesis, with statistically significant improvement registered from the Basic group to the Advanced group. Thus, we can argue that there is a relationship between the number of semesters of study of Mandarin Chinese as L2 and mastery of Unsplit SCs, which corroborates previous studies which have found a relationship between L2 proficiency and formal L2 learning (Cai, 2017; Zhang, 1993; Al-Zahrani, 1998, among others). Contrary to the findings of Cai (2010) and Kellerman (1979), however, our data do not indicate a “u-shaped” pattern of acquisition of Unsplit SCs. Instead, a linear progression in accuracy from the Basic group to the Advanced group and the control group was observed.

The Basic group showed almost no difference in their levels of accuracy in relation to Split versus Unsplit SCs. However, the Advanced group performed noticeably worse in their use of Split SCs when compared to their performance with Unsplit SCs. This divergence between levels of L2 learners’ performance regarding Split vs. Unsplit SCs has also been attested in the literature. For example, Ma (2008) found that the Split form was acquired more slowly than the Unsplit form, concluding that “...learners' acquisition of different discrete forms of splittable words did not develop at equivalent

levels, and also did not align with the improvement in their language ability.” (Cai, 2017, p. 14).

Although a two sample t-test showed the difference between the Advanced group’s performance with Split and Unsplit SCs not be significant, the results highlight the difficulty in acquiring Split SCs.

Our study also had several limitations. The most important limitation was that the sample size was small, and thus not necessarily representative, making it difficult to demonstrate statistically significant correlations. Another limitation was the proficiency level of the L2 students of Mandarin Chinese included in the study. We mechanically equated number of semesters of Mandarin study with proficiency, instead of administering proficiency tests to all of the participants and grouping them according to the results of those tests. Even within the same group, the study backgrounds of the participants were vastly different and involved different materials and methods, with some having studied in China and others not, some exposed to more hours per week of instruction than others, etc. In general, however, the proficiency levels achieved by the members of all three groups seemed to be relatively low, suggesting that if we had included participants with higher levels, our results might have been appreciably different. Thus, future studies should include a larger number of participants who are grouped according to proficiency tests with membership in the Advanced group being restricted to learners at HSK 5 or HSK 6.

In terms of the control group, the limitations were largely due to our use of Google Forms to collect our data, which resulted in a lack of participants from Mainland China as well as a lack of older speakers. Because most control group participants were from Taiwan and Hong Kong, it is possible that they spoke Cantonese as their L1, rather than Mandarin. In terms of age, it is important to note that Standard Mandarin is spoken mainly by the older generations of Chinese people, not the young ones. According to Lening Liu (2020), younger Chinese generations tend to use SCs as transitive verbs, meaning that they add another object to SCs (VO +O), which is traditionally considered ungrammatical.

Thus, a possible explanation for the control group's performance could be the age of the participants. Future studies should therefore avoid using Google Forms and other such tools in order to reach people in Mainland China in general, and a more generationally diverse sample in particular.

Finally, this study demonstrated a gap in the accuracy of learners' performance with Unsplit versus Split SCs as students' general proficiency levels increased, indicating higher and more persistent rates of difficulty in learners' acquisition of Split SCs. This might mean that as learners progress in their studies of Mandarin, they master the use of Unsplit SCs better and faster because Unsplit SCs function to some degree as compound words, while they master the use of Split SCs less well and more slowly because Split SCs function to some degree as collocations, which have generally been shown to be very challenging to learn. Could Unsplit SCs and Split SCs therefore belong to two distinct categories in learners' minds? A study that tests the treatment by grammarians of Unsplit SCs and Split SCs as two manifestations of the same phenomenon would therefore be highly recommended.

Bibliography

- Al-Zahrani, Mohammed Said (1998). Knowledge of English Lexical Collocations Among Male Saudi College Students Majoring in English at a Saudi University. (PhD). Ann Arbor, MI:UMI. In Cai, Y. (2017). *Second Language Acquisition of Chinese Verb-Noun Collocations*. <https://doaj.org/article/f820bd6e28cf44988e96d72e946a06ff>
- Arcodia, G. G. F. (2007). Chinese: a language of compound words? *Selected Proceedings of the 5 Th Décembrettes: Morphology in Toulouse*, 79–90.
<http://www.lingref.com/cpp/decemb/5/paper1617.pdf>
- Bahns, Jens & Eldaw, Moira (1993). “Should we teach EFL students collocations?” *System*, 21(1), 101-114. In Cai, Y. (2017). *Second Language Acquisition of Chinese Verb-Noun Collocations*. <https://doaj.org/article/f820bd6e28cf44988e96d72e946a06ff>
- Bauer, Laurie 2006. Compound In K. Brown, ed.. *Encyclopedia of Language and Linguistics*. Oxford: Elsevier. 719-726 (II vol.). In Arcodia, G. G. F. (2007). Chinese: a language of compound words? *Selected Proceedings of the 5 Th Décembrettes: Morphology in Toulouse*, 79–90.
<http://www.lingref.com/cpp/decemb/5/paper1617.pdf>
- Cai, Y. (2017). *Second Language Acquisition of Chinese Verb-Noun Collocations*.
<https://doaj.org/article/f820bd6e28cf44988e96d72e946a06ff>
- Dai, John X. 1998. Syntactical, phonological and morphological words in Chinese. In J.L. Packard, ed. *New Approaches to Chinese Word Formation*. Berlin – New York: Mouton de Gruyter. 103-133. In Arcodia, G. G. F. (2007). Chinese: a language of compound words? *Selected Proceedings of the 5 Th Décembrettes: Morphology in Toulouse*, 79–90.
<http://www.lingref.com/cpp/decemb/5/paper1617.pdf>

- Feng, Shengli (冯胜利) 1998. Prosodic structure and compound words in Classical Chinese, In J.L. Packard, ed. In Arcodia, G. G. F. (2007). Chinese: a language of compound words? *Selected Proceedings of the 5th Décembrettes: Morphology in Toulouse*, 79–90.
<http://www.lingref.com/cpp/decemb/5/paper1617.pdf>
- Hausmann, Franz J. (1989). “Le dictionnaire de collocation.” In Franz J. Hausmann, Herbert E. Wiegand, & Ladislav Zgusta (Eds), *Wörterbücher, Dictionaries, Dictionnaires. Ein internationales Handbuch zur Lexikographie* (pp. 1010-1019). Berlin: de Gruyter. In Cai, Y. (2017). *Second Language Acquisition of Chinese Verb-Noun Collocations*. <https://doaj.org/article/f820bd6e28cf44988e96d72e946a06ff>
- Hill, J. (2000). Revising priorities: from grammatical failure to collocational success. In Lewis, M. (ed.), *Teaching Collocation*, 47-69. Hove: Language Teaching Publications. In Cai, Y. (2017). *Second Language Acquisition of Chinese Verb-Noun Collocations*.
<https://doaj.org/article/f820bd6e28cf44988e96d72e946a06ff>
- Howarth, P. (1998). Phraseology and second language proficiency. *Applied Linguistics*, 19, 24-44. In Cai, Y. (2017). *Second Language Acquisition of Chinese Verb-Noun Collocations*.
<https://doaj.org/article/f820bd6e28cf44988e96d72e946a06ff>
- Lewis, M. (Ed.). (2000). *Teaching collocations. Further developments in the lexical approach*. Hove, England: Language Teaching Publications. In Cai, Y. (2017). *Second Language Acquisition of Chinese Verb-Noun Collocations*. <https://doaj.org/article/f820bd6e28cf44988e96d72e946a06ff>
- Li, Charles N. and Sandra A. Thompson 1981. *Mandarin Chinese. A Functional Reference Grammar*. Berkeley: University of California Press. In Siewierska, A., Xu, J., & Xiao, R. (2010). Bang-le yi ge da mang (offered a big helping hand): a corpus study of the splittable compounds in spoken and written Chinese. *Language Sciences*, 32, 464–487. <https://doi.org/10.1016/j.langsci.2009.08.002>

Liu, L. (16, 09, 2020). *What is the "Future" of Chinese Language?* [video-conference via Zoom].

Confucius Institute's Academic Seminar University of Macau China.

Lu, Zhiwei, 1957. *The Word-formation of Chinese Language*. The Scientific Press, Beijing. In Siewierska, A.,

Xu, J., & Xiao, R. (2010). Bang-le yi ge da mang (offered a big helping hand): a corpus study of the splittable compounds in spoken and written Chinese. *Language Sciences*, 32, 464–487.

<https://doi.org/10.1016/j.langsci.2009.08.002>

Lu¹, Shuxiang, 1979. *Hanyu yufa fenxi wenti* (Problems in grammatical analysis in Chinese). Reprinted in

Shuxiang, Lu¹. (Ed.), 1984. *Hanyu Yufa Lunwenji* (Anthology of Chinese Grammar). The Commercial Press, Beijing. In Siewierska, A., Xu, J., & Xiao, R. (2010). Bang-le yi ge da mang (offered a big helping hand): a corpus study of the splittable compounds in spoken and written Chinese.

Language Sciences, 32, 464–487. <https://doi.org/10.1016/j.langsci.2009.08.002>

Ma Ping 马萍 (2008)留学生动宾式离合词习得研究——以统计学为视角, 《汉语学习》第5期。In Cai,

Y. (2017). *Second Language Acquisition of Chinese Verb-Noun Collocations*.

<https://doaj.org/article/f820bd6e28cf44988e96d72e946a06ff>

Nesselhauf, N. (2003). The use of collocations by advanced learners of English and some implications

for teaching. *Applied Linguistics*, 24, 223-242. In Cai, Y. (2017). *Second Language Acquisition of Chinese Verb-Noun Collocations*. <https://doaj.org/article/f820bd6e28cf44988e96d72e946a06ff>

Packard, J. L. (2000). *The morphology of Chinese: A linguistic and cognitive approach*. Cambridge:

Cambridge University Press. In Zhang, J., Anderson, R. C., Wang, Q., Packard, J., Xinchun, W.,

Shan, T., & Xiaoling, K. (2012). Insight into the structure of compound words among speakers of Chinese and English. *Applied Psycholinguistics*, 33(4), 753–779.

<https://doi.org/10.1017/S0142716411000555>

- Packard, Jerome, 2003. *The Morphology of Chinese: A Linguistic and Cognitive Approach*. Cambridge University Press, Cambridge. In Siewierska, A., Xu, J., & Xiao, R. (2010). Bang-le yi ge da mang (offered a big helping hand): a corpus study of the splittable compounds in spoken and written Chinese. *Language Sciences*, 32, 464–487. <https://doi.org/10.1016/j.langsci.2009.08.002>
- Palmer, H. E. (1993) *Second Interim Report on English Collocations*. Tokyo: Kaitakusha. In Cai, Y. (2017). *Second Language Acquisition of Chinese Verb-Noun Collocations*. <https://doaj.org/article/f820bd6e28cf44988e96d72e946a06ff>
- Shui, Lian. 税莲. (2007) 现代汉语词语搭配原理与动宾搭配研究[D]. 四川大学, 2007:1-2. In Cai, Y. (2017). *Second Language Acquisition of Chinese Verb-Noun Collocations*. <https://doaj.org/article/f820bd6e28cf44988e96d72e946a06ff>
- Siewierska, A., Xu, J., & Xiao, R. (2010). Bang-le yi ge da mang (offered a big helping hand): a corpus study of the splittable compounds in spoken and written Chinese. *Language Sciences*, 32, 464–487. <https://doi.org/10.1016/j.langsci.2009.08.002>
- Zhang, Xiaolin (1993). *English Collocations and Their Effect on the Writing of Native and Non-native College Freshmen*. Ann Arbor, MI: UMI. In Cai, Y. (2017). *Second Language Acquisition of Chinese Verb-Noun Collocations*. <https://doaj.org/article/f820bd6e28cf44988e96d72e946a06ff>
- Zhang, J., Anderson, R. C., Wang, Q., Packard, J., Xinchun, W., Shan, T., & Xiaoling, K. (2012). Insight into the structure of compound words among speakers of Chinese and English. *Applied Psycholinguistics*, 33(4), 753–779. <https://doi.org/10.1017/S0142716411000555>
- Zhu, Dexi, 1982. *Yufa Jiangyi (Lecture Notes on Grammar)*. The Commercial Press, Beijing. In Siewierska, A.,

Xu, J., & Xiao, R. (2010). Bang-le yi ge da mang (offered a big helping hand): a corpus study of the splittable compounds in spoken and written Chinese. *Language Sciences*, 32, 464–487.

<https://doi.org/10.1016/j.langsci.2009.08.002>

Zhu, Kunlin, 2006. Xiandai Hanyu Cidian zhong de Lihe Ci Yanjiu (A study of Splittable Compounds in Modern Chinese Dictionary). Jilin Sheng Jiaoyu Xueyuan Xuebao (Journal of Educational Institute of Jilin Province) 22 (2), 29–30. In Siewierska, A., Xu, J., & Xiao, R. (2010). Bang-le yi ge da mang (offered a big helping hand): a corpus study of the splittable compounds in spoken and written Chinese. *Language Sciences*, 32, 464–487. <https://doi.org/10.1016/j.langsci.2009.08.002>

Appendices

Appendix 1. The Survey

Part 1: Demographics

How old are you?

What is your gender?

Mark only one oval.

Female

Male

Non-binary

I prefer not to answer

Other: _____

What is your major?

Do you have a minor? If yes, what is it?

How many Mandarin Chinese courses have you taken?

Mark only one oval.

1 course or less

- 2 courses
- 3 courses
- 4 courses
- 5 courses or more
- I prefer not to answer
- Other: _____

How many hours a week were your courses?

Have you ever been to China for language study? If yes, for how long?

When was the last time you took a formal Mandarin Chinese course?

Do you speak any other languages? If yes, which one and what is your proficiency level?

What do you consider is your English proficiency?

Mark only one oval.

- Native
- Advanced
- Intermediate
- Basic
- I prefer not to answer

Part 2: Choose the correct answers.

Please read each exercise and select the correct answers. More than one option is possible in each exercise. At the end of each exercise you will be offered a translation. only use it if you need it to answer that exercise. YOU MUST MARK THE TRANSLATION IF YOU USE IT TO AID YOU.

1. 我去上课，可是我没有雨伞。我问妈妈，有雨伞吗？妈妈问，为什么要雨伞。我回答：

Check all that apply.

因为在下大雨。

因为在下雨大。

因为在天下雨。

因为在下雨。

I need a translation for this one: I need to go to class, but I don't have an umbrella. I ask my mom, do we have an umbrella? Mom asked, why do you want an umbrella? I answered:

2. 请问，力波在这儿吗？

Check all that apply.

不在。他在上课。

不在。他在上汉语课。

不在。他在上课汉语。

不在。他在汉语上课。

I need a translation for this one: Excuse me, is Libo here?

3. 我很忙。我让我弟弟帮我买饭。我问弟弟，

Check all that apply.

你可以帮我忙吗？

你可以帮忙我吗？

你可以帮忙吗？

你可以我帮忙吗？

I need a translation for this one: I'm very busy. I ask my younger brother to help me by buying food. I ask my brother:

4. 我的同事做事情做得很快。我问他，你在这儿上班上了多长时间了？他回答

Check all that apply.

我上了两年的班了。

我上班两年了。

我上班上了两年了。

我两年上班了。

I need a translation for this one: My coworker does things very quickly. I ask him, How long have you worked here? He answered:

5. 我吃饭的时候，我的男朋友来我家。我问他：

Check all that apply.

你在家吃饭了吗？

你在家吃了饭了吗？

你吃饭在家了吗？

你吃饭了在家吗？

I need a translation for this one: When I was eating, my boyfriend came to my house. I ask him:

6. 我的朋友很喜欢唱歌。我问他，

Check all that apply.

你最喜欢唱什么歌？

你最喜欢唱歌什么？

你喜欢什么唱歌？

I need a translation for this one: My friend really likes to sing. I ask him:

7. 昨天我跟朋友去跳舞。很晚回家。今天妈妈问我：

昨天你跳了几个小时的舞？

昨天你跳舞跳了几个小时？

昨天你跳舞了几个小时。

昨天你几个小时跳舞了。

I need a translation for this one: Yesterday I went dancing with a friend. We returned home very late. Today mom asked me:

8. 我生病了。头疼，不想学习。我朋友给我打电话，说

Check all that apply.

你想聊一下天吗？

你想聊天吗？

你想聊天一下吗？

你想一下聊天吗？

I need a translation for this one: I got sick. My head hurts, I don't want to study. My friend called me and said:

9. 我有很多事情问我的男朋友。我给他发信息，问他：

Check all that apply.

什么时候方便给你打一个电话？

什么时候方便给你打电话。

什么时候方便给你打电话一个？

什么时候方便给你一个打电话？

I need a translation for this one: I have many things I need to ask my boyfriend. I sent him a text asking:

10. A: 这个周末你有没有安排？ B: 有，睡觉。我累死了。 A: 你昨晚睡得不好吗？

没睡好。昨晚只睡了四个小时的觉。

没睡好。昨晚只睡觉了四个小时。

没睡好。昨晚只四个小时睡觉了。

我昨晚没睡觉。

I need a translation for this one: A: Do you have any plans for this weekend? B: Yeah, sleep. A: Did you sleep badly last night?

Part 3: What do you see?

Please use the gifs to answer the prompts. More than one answer is possible. Remember, if you need to use the translation to assist you, mark it. 加油!

1. 你想做什么?



Check all that apply.

- 我想游泳。
- 我想游一会儿泳。
- 我想游泳一会儿。
- I need a translation for this one: What do you want to do?

2. 我忘了小虎的生日。所以现在:



Check all that apply.

- 他生气了。
- 他生我的气了。
- 他生气我了。
- I need a translation for this one: I forgot Xiaohu's birthday, so now:

3. 你想做什么？



Check all that apply.

- 我想打电话。
- 我想打一个电话。
- 我想打电话一个。
- I need a translation for this one: What do you want to do?

4. 你必须去医院，因为：



Check all that apply.

- 你发烧了。
- 你发高烧了。
- 你发烧高了。
- 你高烧了。
- I need a translation for this one: You have to go to the hospital because:

5. 他在做什么？



Check all that apply.

- 他在开门。
- 他在开大门。
- 他在大开门。
- 他在开门大。
- I need a translation for this one: What is he doing?

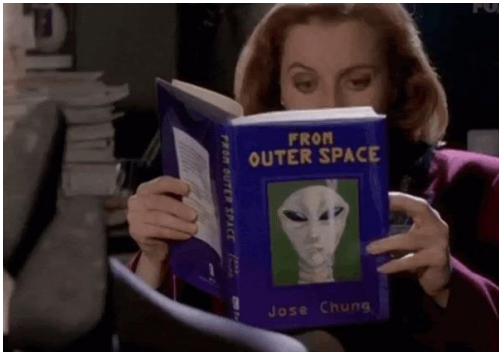
6. 我想去游泳,



Check all that apply.

- 可是下雨了。
- 可是下这么大的雨。
- 可是下雨这么大。
- 可是这么大下雨。
- I need a translation for this one: I want to go swimming,

7.她在做什么？



Check all that apply.

- 她在看书。
- 她在看有意思的书。
- I need a translation for this one: What is she doing?

8.周末他们喜欢做什么？



Check all that apply.

- 他们喜欢散步。
- 他们喜欢散一下步。
- 他们喜欢散步一下。
- 他们喜欢一下散步。
- I need a translation for this one: What do they like to do on weekends?

9. 他们在做什么？



Check all that apply.

- 他们在上课。
- 他们在上英语课。
- 他们在上课英语。
- I need a translation for this one: What are they doing?

10. 我很热，很不舒服。



Check all that apply.

- 我想洗澡。
- 我想洗一会儿澡。
- 我想洗澡一会儿。
- I need a translation for this one: I'm very hot and uncomfortable.

Appendix 2. Selected Verbs

Verbs Used in Study

下雨 Xià yǔ V Fall + N rain 'to rain'	生气 Shēngqì V give birth + N air 'to be angry'
上课 Shàngkè V Go up + N class 'to take class'	发烧 Fāshāo V emit + N fever 'to have a fever'
帮忙 Bāngmáng V Help + Adj busy 'to help'	开门 Kāimén V open + N door 'to open the door'
上班 Shàngbān V Go up + N work shift 'to work'	看书 Kànshū V look + N book 'to read'
吃饭 Chīfàn V Eat + N rice 'to eat'	散步 Sànbù V scatter + V walk 'to take a stroll'

<p>唱歌</p> <p>Chànggē</p> <p>V Sing + N song</p> <p>‘to sing’</p>	<p>洗澡</p> <p>Xǐzǎo</p> <p>V wash + N bath</p> <p>‘to take a bath’</p>
<p>跳舞</p> <p>Tiàowǔ</p> <p>V jump + N dance</p> <p>‘to dance’</p>	<p>游泳</p> <p>Yóuyǒng</p> <p>V swim + N swimming</p> <p>‘to swim’</p>
<p>聊天</p> <p>Liáotiān</p> <p>V chat + N day</p> <p>‘to chat with’</p>	<p>打电话</p> <p>Dǎ diànhuà</p> <p>V make + N phone</p> <p>‘to call’</p>
<p>睡觉</p> <p>Shuìjiào</p> <p>V sleep + N sleep</p> <p>‘to sleep’</p>	

Appendix 3. Coding Sheet

The following information shows the results per participant per group.

Legend:

CRS= Group

1=Basic

2=Intermediate

3= Advanced

6=Control

A= Exercises from part 2 of the survey

B= Exercises from part 3 of the survey

Correctly Split (CS)

Participant	CRS	A1CS	A2CS	A3CS	A4CS	A5CS	A6CS
1	1	0	0	1	0	0	0
2	1	0	0	0	0	0	1
3	1	1	1	1	1	0	1
4	1	1	0	1	0	1	1
5	3	1	1	0	1	0	1
6	3	0	1	1	1	0	1
7	3	1	1	1	0	1	1
8	3	1	0	0	0	0	1
9	5	1	1	1	1	0	0
10	5	1	1	1	0	0	1
11	5	1	1	1	1	1	1
12	5	0	1	0	0	0	1
13	6	1	1	1	1	1	1
14	6	1	1	0	1	1	1
15	6	1	1	1	0	0	1
16	6	1	1	1	1	1	1

A7CS	A8CS	A9CS	A10CS	B1CS	B2CS	B3CS	B4CS
1	0	1	0	0	0	1	0
0	0	0	0	0	1	1	0
1	1	1	0	0	0	1	1
1	0	1	0	0	1	1	1

1	0	1	1	1	0	1	0
1	0	1	1	0	1	1	1
0	1	1	1	1	1	1	1
0	0	0	0	0	0	0	0
1	1	1	1	1	1	1	1
0	0	1	1	0	1	1	1
1	1	1	1	1	1	1	1
0	0	0	0	0	0	1	0
1	1	1	1	1	1	1	1
1	1	1	1	1	1	1	1
1	1	1	1	1	1	1	1
1	1	1	1	1	1	1	1

B5CS	B6CS	B7CS	B8CS	B9CS	B10CS	Total Correct
0	1	1	0	1	0	7
1	1	0	0	0	1	6
1	0	1	1	1	1	15
1	1	1	1	1	1	15
1	1	1	0	1	1	14
1	1	1	0	1	0	14
1	1	1	1	1	1	18
1	0	1	0	1	0	5
1	1	0	0	1	1	16
1	1	1	1	1	0	14
1	1	1	1	1	1	20
0	1	0	0	1	0	5
0	1	0	1	1	1	18
1	1	1	1	0	0	17
1	1	1	1	1	1	18
1	1	1	1	1	1	20

Correctly Unsplit (CU)

Participant	CRS	A1CU	A2CU	A3CU	A4CU	A5CU	A6CU	
1		1	1	0	0	1	1	0
2		1	0	1	0	0	0	0
3		1	1	1	0	0	1	1
4		1	1	1	0	0	1	0

5	3	1	1	1	1	1	0
6	3	1	0	0	0	1	0
7	3	0	0	0	1	0	0
8	3	0	1	0	1	1	0
9	5	1	1	1	0	1	0
10	5	0	1	1	1	1	0
11	5	1	1	1	1	1	0
12	5	1	0	0	0	1	0
13	6	1	1	1	1	1	0
14	6	0	0	1	1	1	0
15	6	0	0	0	0	1	0
16	6	1	1	1	1	1	0

A7CU	A8CU	A9CU	A10CU	B1CU	B2CU	B3CU	B4CU
0	0	0	0	1	1	1	0
0	0	1	0	1	0	0	1
1	0	0	1	1	1	0	1
0	0	0	1	1	1	0	0
1	1	1	1	1	1	1	1
0	0	0	0	1	1	1	1
1	0	0	0	0	0	0	0
0	1	1	1	1	1	1	1
1	1	1	1	1	1	0	1
1	1	1	1	1	1	1	1
1	1	1	1	1	1	1	1
0	0	0	0	1	1	0	1
1	1	1	1	1	1	1	1
1	1	1	1	1	1	1	1
0	1	1	0	0	1	1	1
1	1	1	1	1	1	1	1

B5CU	B6CU	B7CU	B8CU	B9CU	B10CU	Total Correct
1	1	1	1	0	1	11
0	0	1	0	1	0	6
1	1	1	0	1	1	14
1	0	1	0	0	0	8
1	1	1	1	1	1	19
1	1	1	1	1	1	12

Appendix 4. Two variable t-test results

Basic Group (53.75%) and Advanced Group (68.75%)

Group Statistics

	groups	N	Mean	Std. Deviation	Std. Error Mean
values	group1	4	53.75	24.622	12.311
	group3	4	68.75	31.721	15.861

Independent Samples Test

		Levene's Test for Equality of Variances		t-test for Equality of Means						
		F	Sig.	t	df	Sig. (2-tailed)	Mean Difference	Std. Error Difference	95% Confidence Interval of the Difference	
								Lower		Upper
values	Equal variances assumed	.004	.950	-.747	6	.483	-15.000	20.078	-64.129	34.129
	Equal variances not assumed			-.747	5.652	.485	-15.000	20.078	-64.871	34.871

From the t-test analysis, there was no statistically significant mean difference between the basic group (M=53.75, SD = 24.62, N =4) and the advanced group (M=68.75, SD=31.72, N =4), $t(6) = -0.75$, $P = 0.48$.

Basic Group (48.75%) and Advanced Group (81.25%)

Group Statistics

	groups	N	Mean	Std. Deviation	Std. Error Mean
values	group1	4	48.75	17.500	8.750
	group3	4	81.25	17.970	8.985

Independent Samples Test

		Levene's Test for Equality of Variances		t-test for Equality of Means						
		F	Sig.	t	df	Sig. (2-tailed)	Mean Difference	Std. Error Difference	95% Confidence Interval of the Difference	
								Lower		Upper
values	Equal variances assumed	.011	.921	-2.591	6	.041	-32.500	12.542	-63.188	-1.812
	Equal variances not assumed			-2.591	5.996	.041	-32.500	12.542	-63.193	-1.807

From the analysis, the Basic group (M=48.75, SD = 17.50, N =4) had a statistically significant mean difference from the Advanced group (M =81.25% SD = 17.97, N=4), $t(6) = -2.59$, $P = 0.04$.

Advanced Group 's Correct Split (68.75%) and Un-split (81.25%)

Group Statistics

groups	N	Mean	Std. Deviation	Std. Error Mean
values split	4	68.75	31.721	15.861
un_split	4	81.25	17.970	8.985

Independent Samples Test

		Levene's Test for Equality of Variances		t-test for Equality of Means						
		F	Sig.	t	df	Sig. (2-tailed)	Mean Difference	Std. Error Difference	95% Confidence Interval of the Difference	
								Lower	Upper	
values	Equal variances assumed	.664	.446	-.686	6	.518	-12.500	18.229	-57.104	32.104
	Equal variances not assumed			-.686	4.746	.525	-12.500	18.229	-60.124	35.124

From the analysis, there was no statistically significant mean difference between correct split (M =68.75, SD =31.72, N=4) and un-split (M =81.25, SD = 17.97, N =4), $t(6) = -0.69$, $P = 0.52$.

Basic Group (41.4%) and Advanced Group (22.6%)

Group Statistics

groups	N	Mean	Std. Deviation	Std. Error Mean
values group1	4	41.4063	19.49476	9.74738
group3	4	22.6563	16.80444	8.40222

Independent Samples Test

		Levene's Test for Equality of Variances		t-test for Equality of Means						
		F	Sig.	t	df	Sig. (2-tailed)	Mean Difference	Std. Error Difference	95% Confidence Interval of the Difference	
								Lower	Upper	
values	Equal variances assumed	.730	.426	1.457	6	.195	18.75000	12.86891	-12.73908	50.23908
	Equal variances not assumed			1.457	5.872	.196	18.75000	12.86891	-12.90571	50.40571

From the analysis, there was no statistically significant mean difference between the Basic group (M =41.41, SD = 19.50, N =4) and the Advanced group (M =22.66, SD = 16.80, N = 4), $t(6) = 1.46$, $P = 0.195$.

Advanced Group's correct split uses (68.75%) and control group correct split uses (91.25%)

Group Statistics

groups	N	Mean	Std. Deviation	Std. Error Mean
values group3_correct_splits	4	68.75	31.721	15.861
control_group	4	91.25	6.292	3.146

Independent Samples Test

		Levene's Test for Equality of Variances		t-test for Equality of Means						
		F	Sig.	t	df	Sig. (2-tailed)	Mean Difference	Std. Error Difference	95% Confidence Interval of the Difference	
									Lower	Upper
values	Equal variances assumed	3.204	.124	-1.391	6	.213	-22.500	16.170	-62.066	17.066
	Equal variances not assumed			-1.391	3.236	.252	-22.500	16.170	-71.902	26.902

From the analysis, there was no significant mean difference between the Advanced group correct split use (M =68.75, SD = 31.72, N = 4) and the control group's correct split uses (M=91.25, SD= 6.29, N = 4), $t(6) = -1.39$, $P = 0.21$.

Advanced Group's un-split uses (81.25%) and control group un-split uses (82.5%)

Group Statistics

groups	N	Mean	Std. Deviation	Std. Error Mean
values group3_unsplit_uses	4	81.25	17.970	8.985
control_group	4	82.50	18.930	9.465

Independent Samples Test

		Levene's Test for Equality of Variances		t-test for Equality of Means						
		F	Sig.	t	df	Sig. (2-tailed)	Mean Difference	Std. Error Difference	95% Confidence Interval of the Difference	
									Lower	Upper
values	Equal variances assumed	.008	.932	-.096	6	.927	-1.250	13.050	-33.183	30.683
	Equal variances not assumed			-.096	5.984	.927	-1.250	13.050	-33.204	30.704

From the analysis, there was no statistically significant mean difference between the Advanced group's Unsplit uses (M = 81.25, SD = 17.97, N= 4) and control group Unsplit uses (M =82.50, SD = 18.93, N=4), $t(6) = -0.096$, $P= 0.93$.

Advanced Group's ungrammatical Unsplit use (22.6%) and control group ungrammatical Unsplit uses (23.4%)

Group Statistics

groups	N	Mean	Std. Deviation	Std. Error Mean
values group3_incorrect_unsplit_use	4	22.6563	16.80444	8.40222
incorrect_unsplit_uses	4	23.4375	1.80422	.90211

Independent Samples Test

		Levene's Test for Equality of Variances		t-test for Equality of Means							
		F	Sig.	t	df	Sig. (2-tailed)	Mean Difference	Std. Error Difference	95% Confidence Interval of the Difference		
										Lower	Upper
values	Equal variances assumed	4.156	.088	-.092	6	.929	-.78125	8.45051	-21.45891	19.89641	
	Equal variances not assumed			-.092	3.069	.932	-.78125	8.45051	-27.33497	25.77247	

There was no statistically significant mean difference between the Advanced group's ungrammatical Unsplit use (M =22.66, SD =16.80, N =4) and the control group ungrammatical Unsplit uses (M =23.44, SD= 1.80, N =4), $t(6) = -0.92$, $P = 0.93$.