Porta Linguarum 38, June 2022 159-172

Acquisition of L3 Spanish combinations: Development in multilingual contexts

Hui-Chuan Lu
National Cheng Kung University, Taiwan
An Chung Cheng
University of Toledo, U.S.A.
Meng-Hsin Yeh
National Cheng Kung University, Taiwan

Received: 28 September 2021 / Accepted: 19 June 2022
DOI: 10.30827/portalin.vi38.22338
ISSN paper edition: 1697-7467, ISSN digital edition: 2695-8244

ABSTRACT: There has been little research on the acquisition of Spanish collocation, particularly concerning collocation combinations in a multilingual context. Therefore, this corpus-based study attempted to provide new insights for learning Spanish collocation. The research questions are: (1) What are the usage and error tendencies in acquiring Spanish combinations by learners of L3 Spanish? (2) How do the contrasts between L3 Spanish and the learners' prior languages (L1 Chinese and L2 English) play a role in acquiring L3 Spanish collocation? The results showed a developmental sequence moving from N-Adj, V-N, to Adj-N combinations. Furthermore, the results of the qualitative analysis suggested that most errors were associated with the learners' L1 (Chinese) and L2 (English). On the other hand, lexical errors might be attributed to the formal similarity between L2 English and L3 Spanish and the meaning transfer between Chinese (L1) and English (L2). Finally, the analysis suggested the similarity among the three languages in form, and the connections among the linguistic features played an essential role in learning L3 Spanish combinations.
Key words: collocation, L3, learner corpus, parallel corpus, Spanish

La adquisición de las combinaciones en español L3: El desarrollo en contextos multilingües

RESUMEN: Este estudio basado en corpus tiene como objetivo explorar nuevas perspectivas para la adquisición de las colocaciones en el español. Para ello, se establecen dos preguntas de investigación: (1) ¿Cuáles son los usos incorrectos y las tendencias de error en la adquisición de las combinaciones en español por aprendices cuya L3 es español? (2) ¿Qué papel juegan las similitudes y las diferencias entre el español y la L1 (el chino) para la adquisición de la lengua meta?, y ¿qué papel juegan las similitudes y las diferencias entre el español y la L2 (el inglés)? Los resultados señalan el desarrollo del lenguaje partiendo desde las combinaciones de N-Adj, V-N a Adj-N. Los mismos demuestran que la mayoría de los errores observados se asocian tanto a la influencia de la L1 como de la L2 de los aprendices. Por otra parte, los errores léxicos podrían atribuirse a la conexión morfológica entre la L2 y la L3 y la transferencia de significado entre la L1 y la L2. Por último, la similitud estructural entre los tres idiomas, y la conexión de sus características lingüísticas juegan un papel importante en el aprendizaje de las combinaciones de la L3.
Palabras clave: colocación, L3, corpus de aprendices, corpus paralelo, español

1. Introduction

Despite this growing interest in collocation over the years, there has been little research on the acquisition of Spanish collocation, particularly with different collocation combinations in a multilingual context (Ferraro et al., 2014). Previous studies have focused more on second language learners, especially English learners in general. Therefore, this study explores new insights into learning collocations of a foreign language other than English among multilingual learners in Asia. The present study aimed to investigate the acquisition of Spanish collocations focusing on more than one combination by adult learners. This study examined the learners whose first language (L1) was Chinese, the second language (L2) was English, and the third language (L3) was Spanish.

Firth (1935) studied language from the semantic perspective and coined the term "collocation" in the 1950s. In corpus linguistics, "collocation" is defined as a group of words that co-occur more frequently than would be expected by chance (McKeown & Radev, 2000). For example, in the expression of "*take a walk*" in English ("*dar un paseo*" in Spanish), "*take*" not "*give*" or "*do*" collocates with "*a walk*". According to Aisenstadt (1979), Cowie (1986), Mel'čuk (1998), and Nesselhauf (2003, 2005), collocations are habitually occurring lexical combinations that are characterised by the restricted co-occurrence of elements. For example, "*tea*" collocates with "*strong*" but not with "*powerful*" in Halliday´s example of "*strong* vs. *powerful tea*" (1966: 150) and have relative transparency in terms of meaning. Furthermore, "*face*" in "*face a problem*" is not used with its original meaning but is an idiom. Collocation, then, can be denoted as a type of word combination in a phraseological sense (e.g., Cowie, 1994). Collocation can also be defined as the co-occurrence of words in a frequency-based sense (e.g., Sinclair, 1991). A collocation has two types of constituents: (1) open collocations, which are often called "lexical collocations," such as Verb-Noun or Adjective-Noun combinations, and (2) grammatical collocations, which are formed by an open element and a closed element, such as Verb-Preposition (Aisenstadt, 1979; Cowie, 1986; Fontenelle, 1994). Different researchers have used diverse definitions to delimit collocation from other types of combinations. In this study, based on the chi-square statistic results as well as the Google and corpora frequency searches, we analysed collocations including "*leer libros*" ("*read books*") and semantically restricted collocations "*prestar atención*" ("*pay attention*"), but excluding free combinations such as "*comprar cortinas*" ("*buy curtains*").

The term "combination" is used in this study to include different types of structures: Adjective-Noun (Adj-N), Noun-Adjective (N-Adj), and Verb-Noun (V-N). The purpose of this study is to investigate the learning of Spanish combinations by examining Taiwanese L3 learners' uses and errors about Spanish collocations. To further examine possible cross-language influences, this cross-linguistic study also examines Spanish combinations with contrastive data in English and Chinese.

The research design addressed the two research questions: (1) What are the usage and error tendencies in acquiring Spanish combinations by learners of L3 Spanish? (2) How do the similarities and differences between learners' target language Spanish (L3) and first language Chinese (L1) and those between Spanish and English (L2) play a role in the acquisition of L3 Spanish collocation based on contrastive analysis in corpus linguistics?

2. Previous research

Previous studies related to L2 English (Sinclair, 1991; Nattinger & DeCarrico, 1992; Wray, 2000; among others) showed that collocation plays an important role in foreign language acquisition. However, the process of lexical acquisition causes frustration for both beginners and advanced learners, and it is characterised by various degrees of difficulty. In Laufer and Waldman (2011), learners at three different proficiency levels used less collocation than native speakers in terms of usage tendency. Correct use of collocations remains an obstacle even for advanced learners (Granger, 1998; Lorenz, 1999; Nesselhauf, 2003, 2005).

Among different types of collocations, previous research (Alfahadi et al., 2014; Siyanova & Schmitt, 2008) indicated that the Adj-N construction causes L2 English learners to make mistakes easily. On the other hand, Nesselhauf (2003) pointed out that the V-N construction is more problematic for learners. Furthermore, around 25% of learner-written productions demonstrate incorrect uses of N-V collocations in English. In a corpus-based study on Spanish V-N collocations, Laufer and Waldman (2011) found learning collocations to be a slow process. The speed only became faster when learners reached a more advanced proficiency level. Alonso Ramos et al. (2010) went one step further in pointing out that almost two-thirds of incorrect uses are classified as lexical errors. More than half of these lexical errors are associated with verbs and adjectives other than nouns. The majority of previous studies seemed to focus on one type of combination. This study, however, investigated three kinds of collocations to include a broader base of comparisons to advance our understanding of the Spanish collocations acquisition.

As suggested in previous research, factors that might affect learners' usage of collocations include learner proficiency level, cross-linguistic transfer from a learner's first language to a target language, and word-for-word correspondent translation between first language and the target language. Keshavarz and Salimi (2007) and Thewissen (2008) concluded that collocation knowledge is related to the proficiency level of the learners' target language. Also, Bahns and Eldaws (1993), Nesselhauf (2003, 2005), Fan (2009), Laufer and Waldman (2011), and Phoocharoensil (2013) suggested that collocation errors were related to the native language of learners, which might interfere with L2 collocation acquisition. Laufer and Waldman (2011) pointed out that 50% of incorrect collocational uses are affected by the native language.

Research has shown that both the native (L1) and nonnative (L2) languages can be sources of influence when acquiring a new language (L3) (Cenoz, 2001; Hammarberg, 2001; Möhle, 1989; Ringbom, 1987, 2001). Ringbom (2001) indicates two types of transfers: formal transfer and semantic transfer. Formal transfer refers to morphological errors, including the use of false cognates, lexical borrowing, or lexical inventions. Semantic transfer refers to using a target language word with a meaning that reflects an influence from the semantics of a corresponding word in another language.

In the literature on cross-linguistic influence, it has been shown that factors that influence the acquisition of a third language or another language may involve typological similarity, the frequency of language use, and the level of proficiency of a second language. Linguistic distance plays a vital role in L3 acquisition (Odlin & Jarvis, 2004; Ringbom, 1987 & 2001). These studies suggested that typologically similar languages provide much more information for the learner than linguistically distant languages. Ringbom (2001) found that the linguistic transfer to L3 English was mainly from Swedish as L1 or L2 (another L1 or

L2 was Finnish). Cenoz (2001) found that learners with L1 Basque more often transferred words from their L2 Spanish into their L3 English than the L1 Spanish, L2 Basque, L3 English group. Thus, Cenoz suggested that language distance played a more significant role in cross-linguistic transfer than L2 proficiency. The source of meaning transfer can also be a nonnative language, but only when a high (near-native) level of proficiency was reached. A high proficiency level in L2 may not be enough for L2 to become automatised, and L2 exposure in the input may be essential (De Angelis & Selinker, 2001; Ringbom, 1987). Also, Hammarberg (2001) observed that while L1 influence persisted over a more extended period, L2 influence tended to fade away twice as rapidly.

Thus, the present corpus-based study compares different types of collocations, considering influencing factors such as learner proficiency levels and structural differences between previous and target languages, to better understand the acquisition of Spanish combinations as a third language of Chinese-speaking learners.

3. Research Method

3.1 Tools

The data sources of this study are from a learner corpus, the "Taiwanese Learners' Written Corpus of Spanish" / "Corpus Escrito de Aprendices Taiwaneses de Español (CEATE)" in Spanish and a trilingual corpus, the "Parallel Corpus of Spanish, English and Chinese" / "Corpus Paralelo de Español, Inglés y Chino (CPEIC)" in Spanish. The research team from National Cheng Kung University in Taiwan developed the CEATE (available at http://corpora.flld.ncku.edu.tw/). This learner corpus compiled texts (446,694 words) written by Taiwanese university students of L3 Spanish and the corresponding corrected versions annotated by native speakers to identify learners' correct usage and errors by contrasting learners' and native speakers' texts. The CEATE includes two major features: POS (parts of speech)-tagging and annotations of error correction, allowing users to search Spanish collocations through multi-function options for research and teaching/learning purposes.

To address the first research question, we used the Spanish Collocation Tool (Herramienta de Colocaciones Españolas (HCE) in Spanish)[1] to extract data from the learner corpus (CEATE) through the chi-square statistic. The extracted data were manually checked for error analysis.

To address the second research question, the trilingual parallel corpus (CPEIC) was used to check the frequency of combinations as we compared the data across the three languages in this study. The CPEIC was created in 2007 by the same research team from National Cheng Kung University, which has continued collecting data from diverse sources. The CPEIC is a POS-tagged parallel translation corpus that contains three of the most spoken languages in the world, Spanish, English, and Chinese. The CPEIC compiles written and oral texts from three different sources, the Bible, United Nations documents, and fairy tales.

[1] The Spanish Collocation Tool /HCE was developed by the same research team from National Cheng Kung University in Taiwan to facilitate the analysis of Spanish collocations.

The parallel corpus contains 961,000 Spanish words, 980,000 English words, and 1,250,000 Chinese characters.

3.2. Research procedure

3.2.1. Using the learner corpus CEATE for error analysis

In this study, we filtered the texts from the learner corpus, CEATE, using the following criteria for consistent data characteristics: length of 100-200 words, with a textual type of description, and themes related to leisure and routine life. Then, we extracted three types of Spanish combinations (Adj-N, N-Adj, and V-N) from (1) texts written by learners at three proficiency levels and from (2) the revised corresponding learner texts corrected by native Hispanic speakers. The proficiency levels in Spanish were determined according to hours of instruction in Spanish at the participants' universities. Although the proficiency level of English is not the variable of data analysis, the average scores on the University Entrance Exam showed that all the participants had at least upper-elementary level. Furthermore, the participants of the elementary Spanish group reached a higher English proficiency level than that of the intermediate Spanish group, and the intermediate Spanish group reached a higher English level than that of the upper-intermediate Spanish group. A total of approximately 105,314 words were analysed in this study. Table 1 shows the data extracted from the CEATE for the three different groups under the study.

Table 1. Analysed data extracted from CEATE

Proficiency levels	Learning hours	Background	Original texts (words)	Revised texts (words)
Elementary	128	Department of Foreign Languages or Department of English	18,291	18,535
Intermediate	576-1,088	Department of Spanish (2nd year)	16,607	16,792
Upper -Intermediate	1,500-2,000	Department of Spanish (3rd and 4th year)	17,437	17,652

The data extracted from the learner corpus were initially imported into the HCE collocation tool to obtain collocation output through data processing, POS tagging, collocation extraction, and calculation. In this study, three types of combinations (Adj-N, N-Adj, and V-N) from the learner corpus and revised texts were analysed to find the similarities and differences between the usage tendencies of L3 learners and native speakers of Spanish. Next, learner errors were compared manually by contrasting the collocations uses between the learners' original and revised texts to facilitate the error-type classification. Finally, the learner errors in the three types of combinations were categorised into grammatical or lexical errors.

3.2.2. Using the parallel corpus (CPEIC) for contrastive analysis

The second corpus used as a data source was the trilingual parallel corpus (CPEIC). Since Chinese is the first language of the participants of this study, and English is the first foreign language learnt at schools, the process of learning Spanish might be affected by their prior knowledge of Chinese and English. The second part of the study examined the linguistic similarities and differences between Chinese and Spanish as well as between English and Spanish by comparing and contrasting correspondent Adj-N, N-Adj, and V-N combinations extracted from the trilingual parallel corpus.

The data in this part of the study contained 9,870 Spanish words, 9,932 English words, and 12,313 Chinese characters, extracted from the sub-corpus of fairy tales for the posterior analysis. Subsequently, the Spanish combinations and their parallel translations in English and Chinese were manually listed for contrastive analysis. Finally, the phrase structures expressing the same meaning in the three different languages were compared and contrasted.

4. Results

4.1. Learning development: Usages and errors

To investigate the development of the acquisition of Spanish combinations, we compared the correct and incorrect uses of different types of combinations by learners at three different proficiency levels. The error type and linguistic similarity were used as two variables to examine the learners' development of Spanish combinations.

4.1.1. Results of usage analysis

The frequency of learner usage of the three combinations (Adj-N, N-Adj and V-N) appearing in the learner corpus, CEATE, was tallied under different proficiency levels: elementary, intermediate, and upper-intermediate. The usage of three combinations by native Spanish speakers was employed as a baseline for determining the correct and incorrect learner production. The frequency data for the correct and incorrect uses were tabulated under learner proficiency levels and types of combinations, as shown in Table 2.

Table 2. Accuracy counts of learners' Spanish combinations

Level	Elementary		Intermediate		upper-intermediate	
Usage	*Correct*	*Incorrect*	*Correct*	*Incorrect*	*Correct*	*Incorrect*
Adj-N	49	18	84	8	56	5
	(73%)	(27%)	(91%)	(9%)	(92%)	(8%)
N-Adj	113	18	160	38	177	31
	(86%)	(14%)	(81%)	(19%)	(85%)	(15%)
V-N	99	23	100	22	32	8
	(81%)	(19%)	(82%)	(18%)	(80%)	(20%)

According to Table 2, in the comparisons between the correct and incorrect uses of the Adj-N combination, the accuracy rates of the elementary (73%), intermediate (91%), and upper-intermediate (92%) level learners increased gradually. For example, for the usage of Adj-N combination, "*próxima vez*" (literal translation in English "*next time*") was found in the text at the elementary level; "*mala nota*" ("*bad grade*") at the intermediate level, and "*densa niebla*" ("*dense fog*") at the upper-intermediate level.

However, different tendencies were observed in the other two combinations. For the N-Adj combination, the accuracy rate at the elementary level (86%) was higher than that at the intermediate level (81%) and upper-intermediate level (85%). In comparison, the accuracy rate at the intermediate level (82%) was slightly higher than the elementary level (81%) and the upper-intermediate level (80%) for the V-N combination. In the N-Adj combination, learners from the elementary, intermediate, and upper-intermediate levels used examples such as "*mensaje electrónico*" ("*message electronic*"), "*fuegos artificiales*" ("*fire artificial*"), and "*seguridad social*" ("*security social*"), respectively. As with the V-N combination, "*sacar fotos*" ("*take pictures*"), "*componer versos*" ("*compose verses*"), and "*tener lugar*" ("*take place*") are the instances in the texts produced by learners from the elementary, intermediate, and upper-intermediate levels, respectively.

For the elementary level across all types of combinations, the results seemed to suggest that the development of learning N-Adj (86%) and V-N (81%) occurred earlier than Adj-N (73%) according to the accuracy rate of the three types of combinations. Furthermore, through the development of each combination, we observed that learners developed the Spanish Adj-N combinations moving from the adjective with a post-nominal position at the elementary level to the prenominal position at posterior stages. That is, they developed the knowledge for the adjective with noun combinations that allow for the Adj-N construction at the intermediate (91%) and upper-intermediate (92%) levels.

To address research question one regarding the usage of Taiwanese learners of L3 Spanish at different proficiency levels, we derived the following developmental sequence ranked from N-Adj, V-N, to Adj-N moving from high to low accuracy of uses. The results are shown in Table 2. No prior study has examined different combinations in learner production; these results suggested a learning sequence of combinations in the Spanish collocation acquisition.

4.1.2. Results of error analysis on the CEATE

The first variable used to examine the acquisition of Spanish combinations in this study was error type. Lexical, grammatical, and orthographic errors in the uses of Spanish Adj-N, N-Adj, and V-N combinations were examined in the analysis of learner errors, in contrast to the usages revised by native speakers. The grammatical errors included number (*"*diferente universidades*" instead of "*diferentes universidades*") and gender (*"*buen compañía*" instead of "*buena compañía*") agreement, misplacement (*"*entero cuerpo*" instead of "*cuerpo entero*") and ungrammatical omission of obligatory elements (*"*coger barco*" instead of "*coger un barco*"), etc.. Lexical errors refer to word choice errors (*"*hacer vino*" instead of "*hacer viento*" or *"*ver cine*" instead of "*ver películas*"). Orthographic category refers to spelling (*"*aprender langules*" instead of "*aprender lenguas*") and capitalization (uppercase or lowercase) errors (*"*hablar Español*" instead of "*hablar español*"). Finally, if a collocation involved more than one error type from the previously mentioned categories (lexical, grammatical, or orthographic), it was assigned to the "mixed" category. Table 3 shows the distribution of error types, including lexical, grammatical, orthographic, and mixed categories.

Table 3. Distribution of learner errors

Level	elementary				Intermediate				upper-intermediate			
Error type	*Lexical*	*Gram-matical*	*Ortho-graphic*	*Mixed*	*Lexical*	*Gram-matical*	*Ortho-graphic*	*Mixed*	*Lexical*	*Gram-matical*	*Ortho-graphic*	*Mixed*
Adj-N	0	12 (67%)	2 (11%)	4 (22%)	1 (12.5%)	6 (75%)	1 (12.5%)	0	4 (80%)	0	0	1 (20%)
N-Adj	0	7 (30%)	14 (61%)	2 (9%)	8 (21%)	18 (47%)	5 (13%)	7 (19%)	12 (39%)	7 (22.5%)	7 (22.5%)	5 (16%)
V-N	12 (52%)	7 (31%)	3 (13%)	1 (4%)	2 (9%)	16 (73%)	0	4 (18%)	4 (50%)	4 (50%)	0	0

As shown in Table 3, within each proficiency level, the intermediate learners made more grammatical than lexical errors across the three combinations (75% vs. 12.5% for Adj-N, 47% vs. 21% for N-Adj, and 73% vs. 9% for V-N combinations, respectively). However, upper-intermediate learners made more lexical than grammatical errors in the Adj-N (80% vs. 0) and N-Adj (39% vs. 22.5%) combinations.

Furthermore, learners at the elementary and intermediate levels made more grammatical errors (67% and 75%) than other errors in the Adj-N combination, whereas this was not the case for learners at the upper-intermediate (0%) level. In the N-Adj combination, the major errors made by learners at different proficiency levels were in distinct categories: The elementary learners had more orthographic errors (61%), the intermediate learners had more grammatical errors (47%), and the upper-intermediate learners had more lexical errors (39%). In the V-N combination, learners at the elementary level made more lexical (52%) errors, while learners at the intermediate level made more grammatical (73%) errors.

In further examining the data where there were more grammatical errors than lexical ones, the following discussion provides examples of the grammatical errors found in the results. Learners at the elementary level misplaced the adjective in the prenominal position and ignored number agreement such as *"*mayor hermanos*" ("*elder brothers*") instead of "*hermanos mayores* ("*brothers elder[s]*"). Learners at the intermediate level wrote the Adj-N combination *"*grandes ojos*" ("*big[s] eyes*") instead of the N-Adj combination "*ojos grandes*" ("*eyes big[s]*"). In contrast, learners at the same level used the N-Adj combination *"*relaciones buenas*" ("*relations good[s]*") instead of the Adj-N combination "*buenas relaciones*" ("*good[s] relations*"). Both errors involved the misplacement of adjectives and nouns but for different reasons, which will be discussed in the following sections. Furthermore, learners at the same level also made grammatical errors in the V-N combination *"*celebrar cumpleaños*" ("*celebrate birthday*") due to the unacceptable omission of an article "*un*" ("*a*") or a possessive pronoun "*mi*" ("*my*") between the verb and the noun in Spanish.

On the other hand, there were more lexical than grammatical errors in the different combinations. Examples of lexical errors involving the selection of appropriate words in three combinations are presented as follows. At the upper-intermediate level, "*buen cerebro*" ("*good brain*") was used instead of "*buena memoria*" ("*good memory*") in the Adj-N combination,

and "*día caliente*" ("*day hot*") was used instead of "*día caluroso*" ("*day hot*") used in the N-Adj combination. The lexical error "*he visto cine*" ("*have seen theatre*") instead of "*he visto la película*" ("*have seen the movie*") was also observed in the V-N combination for learners at the elementary level. Finally, in the N-Adj combination, learners at the elementary level wrote *"*languas extranjeras*" ("*languages foreign[s]*" with spelling errors and *"*pais antiguo*" ("*country old*") without the accent on "país" ("*country*").

In summary, although the L3 Spanish learners at three proficiency levels seemed to behave differently in producing Adj-N, N-Adj, and V-N combinations, overall, these results suggested that learner errors changed from grammatical to lexical type errors in the Adj-N and N-Adj combinations for the upper-intermediate learners.

4.2. Results of the contrastive analysis on CPEIC

To address the research question two regarding possible influences from Chinese or English, we mapped learners' errors in Spanish (L3) to characteristics of Chinese (L1) and English (L2) as discussed in Section 4.1. This section discusses the comparisons across the three languages, Spanish, English, and Chinese, in the three Spanish combinations. Some Spanish combinations correspond to the same structure in English and/or Chinese, thus sharing similar semantic and syntactic features with English and Chinese combinations. However, some combinations do not have one-to-one matching across the three languages. Table 4 shows the results of the contrastive analysis of the Adj-N, N-Adj, and V-N combinations among the three languages based on the data analysis of fairy tales from the trilingual parallel corpus CPEIC.

Table 4. Frequency of Adj-N, N-Adj, and V-N combinations in the contrasts between Spanish and English and between Spanish and Chinese

ENGLISH		SPANISH		CHINESE	
Adj-N	73 (95%)	Adj-N	77	Adj-N	31 (40%)
				Adj/N*de*-N_{head}	38 (49%)
Adj-N	125 (86%)	N-Adj	145	Adj-N	74 (51%)
N	9 (7%)			Adj/N*de*-N_{head}	45 (31%)
V-N	160 (69%)	V-N	232	V-N	162 (70%)
V-Prep (N)	33 (14%)			*Ba*-N-V	22 (9%)
V	16 (7%)			V∅	13 (6%)
V-Adj	10 (4%)			V-Prep-(N)	10 (4%)
				V-Adj	4 (2%)

As shown in Table 4, in the comparisons between Spanish and English, both Spanish Adj-N and N-Adj combinations could be translated into Adj-N structures, 95% and 86% in English, respectively. On the other hand, it should be noted that the particle "*de*" followed

a noun can form an adjective like "N *de*" in "*lianghao de*" ("*goodness de*") "*guanxi*" ("*relationship*") in Chinese and is functioned as an adjective that modified the N. Therefore, the Adj-N and N*de*-N combinations could be considered the most similar and major types in the distribution, 89% (40% and 49%) and 82% (51% and 31%). According to the distribution of the correspondent structures, the Adj-N combination is more similar than the N-Adj combination in the three parallel languages because the latter has the reversed structure of L1 and L2.

The following examples extracted from the CPEIC illustrate parallel comparisons of Adj-N and N-Adj combinations across the three languages. In the Adj-N combination, the phrases "*buen amigo* (Spanish)*/ good friend* (English)*/ hao pengyou* (Chinese)" and "*nueva idea / new idea / xin de zhuyi*" shared the same syntactic structure and meaning with the only difference being the marker "*de*" inserted between the modifying and modified elements in Chinese.

In the N-Adj combination, the phrase "*animales salvajes*" ("*animals wild[s]*") does not share the same syntactic structure with its English equivalent, "*wild animals*" (Adj-N), nor with its Chinese equivalent, "*yesheng dongwu*" (Adj-N, "*wild animal*"). Similar contrasts can also be found in "*fuego brillante*" ("*fire brillant*"), except that in the Chinese equivalent, "*xongxong de huo*" the particle "*de*" is inserted to form the adjective "*xongxong de*" ("*flaming*") to modify the head-noun, "*huo*" ("*fire*").

The head of Adj-N and N-Adj combinations in Spanish is variable, whereas the head position is fixed for English and Chinese. The instances of the Adj-N combination in English are overwhelmingly more than other combinations, and the noun is a fixed head as a second element in the Chinese Adj-N and N/Adj*de*-N combinations. Furthermore, most Spanish N-Adj combinations have the reversed structure of the Adj-N in English and Adj-N and N/Adj*de*-N in Chinese. One-to-one matching between form and meaning is expected to facilitate prior language transfer, particularly before elementary learners. However, considering the similarity transfer, the results seemed contrary to the conclusion in Section 4.1.1., which indicates that the acquisition of the N-Adj combination occurred earlier than the Adj-N combination. The L1 or L2 influence in the acquisition of N-Adj combination is not evident among the participants of this study. The participants of this study had learned Spanish for 128 hours or about two or four semesters (four or two hours per week respectively). The learner corpus did not include those first-semester learners who are most likely to transfer the L1 (Chinese) Adj-N structure to L3 (Spanish). Furthermore, the elementary group performed better than the intermediate and upper-intermediate groups in the N-Adj combination. It is plausible that the elementary group had internalized the N-Adj construct and applied it to all adjective and noun constructs. That is, the N-Adj construct was overgeneralized to all N-Adj and Adj-N combinations. In addition, by observing the frequency of Spanish Adj-N and N-Adj combinations, we would argue that the earlier acquisition of N-Adj than Adj-N combination might result from interaction with the input in Spanish. The frequency of Spanish N-Adj (145) was almost twice that of Spanish Adj-N (77) as shown in the data analysis of the trilingual corpus in Table 4. This argument is further supported by the search result using the Corpus del Español (Davies, 2002) as a reference corpus, i.e., the Adj-N combination is used less frequently than the N-Adj combination in natural Spanish. The incorrect use of the Adj-N combination (e.g., *"*relaciones buenas*" ("*relations good[s]*") instead of the Adj-N combination "*buenas relaciones*" ("*good[s] relations*") could result

from the overgeneralization of the N-Adj combinations because N-Adj was taught first and stressed more in the classroom for its difference with learners' first language. Besides, the frequency of occurrences is overwhelmingly high in comparison with that of the Adj-N combination. Learning the Spanish Adj-N and N-Adj combinations is associated with the distinction between various adjective types and subtle differences in meaning. Therefore, the frequency of input and classroom instruction overpower the first language influence on the acquisition of Adj-N combinations. It will take more time for learners to differentiate the meanings of different adjective types in Adj-N and N-Adj combinations.

As with the V-N combination, the majority (over 69%) of Spanish V-N combinations corresponded to the V-N structure in English and Chinese. For example, there are word-for-word correspondent structures in the three languages such as "*sacaba excusas* / *made excuses* / *bianzao jiekou*" and "*construir una casita* / *build a house* / *gai yidong fangzi*". In addition, there are also cases with correspondent structures in only two of the three languages "*hacer preguntas* / *ask questions* / *xunwen*" and "*encendió el fuego* / *made-up a fire* / *shenghuo*".

Furthermore, some V-Ns involving articles (Art) or possessives (Poss), such as "*sacudía la cabeza*" ("*shake the head*") / "*shake his head*" / "*yao tou*" ("*shake head*") are also classified as V-N. However, more similarities were observed between Spanish and English. In contrast, more differences were observed between Spanish and Chinese. The cross-linguistic differences in the V-N combination would require special attention in a contrastive analysis and might play a role in determining the dominant role of learners' previous languages. Taking a closer look, we further analysed 121 Spanish examples with parallel V-N structures both in English and Chinese at the same time. The results showed that 49 (40%) of them have a similar structure (V-Art-N or V-Poss-N) between Spanish and English, but only 2 (2%) have a similar structure between Spanish and Chinese. There were examples such as "*hacer* (V) *los* (Art) *quehaceres* (N) / *do* (V) *the* (Art) *housework* (N) / *zuo* (V) *jiashi* (N)" and "*perder* (V) *el* (Art) *apetito* (N) / *lose* (V) *his* (Poss) *appetite* (N) / *shiqu* (V) *shiyu* (N)."

Finally, around 30% of the Spanish V-N combinations corresponded to other structures, such as V and V-Adj in English and Chinese. For example, in "*hacer daño*" ("*do hurt*") and "*dar pena*" ("*give pain*") Spanish collocation belongs to the V-N structure, but their equivalents in English "*hurt*" and "*feel sorry*" are V and V-Adj structures; the same meaning is expressed in "*shanghai*" ("*hurt*") and "*gandao baoqian*" ("*feel sorry*") with a simple V and a V-Adj structures in Chinese. In comparison with Adj-N and N-Adj combinations, we observed that Spanish V-N combinations varied greatly in English and Chinese translations. The detailed results of the contrastive analysis provided further explanation for the grammatical difficulty that intermediate learners demonstrated.

5. CONCLUSION

This study examined learner corpus in L3 Spanish to investigate learners' acquisition sequence of L3 Spanish combinations of N-Adj, V-N, and Adj-N. It also compared the similarities and differences in three combinations among the three languages, which provided a window through which we observed possible influences of cross-language transfer. The employment of error analysis and a contrastive analysis on a learner corpus and a trilingual parallel corpus presented a venue for research on the acquisition of the Spanish combinations.

The mixed research methods provided an innovative approach to a better understanding of learner production in the three types of Spanish combinations cross-linguistically.

To answer the research question addressing the usage and errors in the Spanish collocations of multilingual learners, the results of the usage analysis in the learner corpus (Section 4.1.1.) showed an acquisition sequence ranked from high to low usage accuracy: N-Adj, and then V-N, followed by Adj-N combinations. This order of the Spanish combinations acquisition might associate with natural language input where the instances of the N-Adj combination are about twice as many as those of Adj-N combinations in Spanish in the natural corpus and in learning input as well.

Furthermore, the results of error types shown in the section 4.1.2. suggest that the upper-intermediate learners tended to make more lexical errors than grammatical ones in general. In contrast, intermediate learners showed more grammatical than lexical errors in the opposite tendency. Grammatical errors related to the post-nominal adjectives in the N-Adj combination could be attributed to the overgeneralization of Adj-N in L3 Spanish. In contrast, grammatical errors related to the prenominal adjectives in the Adj-N combination and omission of elements in the V-N combination were associated with learners' L1 (Chinese) and L2 (English), both of which share the same linguistic features as their L3 (Spanish).

Spanish adjective and noun constructions differ from English and Chinese in that many Spanish adjectives allow for prenominal and post-nominal positions, but English and Chinese adjectives typically precede the nouns they modify. The overwhelming amount of the N-Adj combination in the learners' input in Spanish, the binary options of prenominal and post-nominal positions for the adjective placement in Spanish, and learners' habitual usage of an adjective in the prenominal position in Chinese and English may all contributed to the errors in the Spanish collocations acquisition. Learners at the intermediate and upper-intermediate levels had a broader vocabulary than elementary learners. Thus, intermediate learners and started to see the subtle differences in the adjective's meanings in prenominal and post-nominal positions, which explained that more proficient learners had more lexical errors than grammatical errors. By the same token, upper-intermediate learners tended to make more lexical errors in word choices than grammatical errors. In contrast, elementary learners with usually used formulaic expressions or only the limited vocabulary they knew, and thus they produced more grammatical errors than lexical errors in Adj-N and N-Adj combinations.

Furthermore, the second research question explored how the similarities between L3 Spanish and Taiwanese learners' L1 (Chinese) and those between Spanish and learners' L2 (English) affect the acquisition of Spanish collocation. The contrastive analysis was conducted to address the learners' prior linguistic knowledge in the cross-language transfer. The results indicated that grammatical errors were associated with learners' L1 and L2 when neither L1 nor L2 shared the same linguistic features with L3, i.e., the case for learners at the intermediate level in the N-Adj combination. In addition, L1 played a more dominant role in the grammatical errors of the V-N combination for the intermediate learners because their L1 differed more from L3 than L2 did.

Built upon our previous outcomes of various corpora, the CEATE, CPEIC, and HCE, this study explored the learning of Spanish collocation as a third language in multilingual acquisition. Needless to say, there are limitations in this study. Future research could in-

vestigate learner corpus regarding the uses of these three combinations written by learners at a lower proficiency level to capture learner production, particularly in the Adj-N and N-Adj combinations to see the L1 (or even L2) influence on the N-Adj combination errors. Future study could also exam further the written production of advanced learners to clarify lexical vs. grammatical errors in these three combinations. It will also be necessary to explore the collocations in learner writings in various themes and different text types such as informational (expository), narrative, argument, descriptive writings to elicit more varieties of learner productions for a better understanding of the acquisition of Spanish collocations.

6. ACKNOWLEDGMENTS

We wish to extend our sincere gratitude for the generous support of a project grant from the Ministry of Science and Technology of Taiwan (MOST103-2410-H-006-059-MY2) and the technical support from the Computer Science and Information Engineering team at National Cheng Kung University in Taiwan.

7. REFERENCES

Aisenstadt, E. (1979). Collocability restrictions in dictionaries. *ITL, 45-46*, 71-74.

Alfahadi, AM., Zohairy, SA., Momani, MM., & Wahby, MH. (2014). Promoting awareness of teaching collocations techniques to beginners (Adjective-Noun collocations). *European Scientific Journal*, *10*(10), 389-396.

Alonso Ramos, M., Wanner, L., Vincze, O., Casamayo del Bosque, G., Vázquez Veiga, N., Mosqueira Suárez, E., & Prieto González, S. (2010). Towards a motivated annotation schema of collocation errors in learner corpora. In N. Calzolari, K. Choukri, B. Maegaard, J. Mariani, J. Odijk, S. Piperidis, M. Rosner and D. Tapias (Eds.), *Proceedings of the Seventh conference on international language resources and evaluation* (pp. 3209–3214). European Language Resources Association (ELRA).

Bahns, J., & Eldaw, M. (1993). Should we teach EFL students collocations? *System, 21*(1), 101-114.

Cenoz, J. (2001). The effect of linguistic distance, L2 status and age on cross-linguistic influence in third language acquisition. In J. Cenoz, B. Hufeisen & U. Jessner (Eds.), *Cross-linguistic influence in third language acquisition: Psycholinguistic perspectives* (pp.8-20). Multilingual Matters.

Cowie, A. P. (1986). Collocational dictionaries—A comparative view. In Murphy (Ed.), *Fourth joint Anglo-Soviet seminar* (pp. 61-69). British Council.

Cowie, A. P. (1994). Phraseology. In R. E. Asher (Ed.), *Encyclopedia of language and linguistics* (pp. 3168-3171). Pergamon.

Davies, M. (2002). *Corpus del Español: 100 million words, 1200s-1900s*. Available online at http://www.corpusdelespanol.org.

De Angelis G., & Selinker L. (2001). Interlanguage transfer and competing linguistic systems in the multilingual mind. In J. Cenoz, B. Hufeisen & U. Jessner (Eds.), *Cross-linguistic influence in third language acquisition: Psycholinguistic perspectives* (pp.42-58). Multilingual Matters.

Fan, M. (2009). An exploratory study of collocational use by ESL students—A task based approach. *System, 37*, 110-123.

Ferraro, G., Nazar, R., Alonso Ramos, M. & Wanner, L. (2014). Towards advanced collocation error correction in Spanish learner corpora. *Language resources and evaluation, 48*, 45-64.

Firth, J. R. (1935). The technique of semantics. *Transactions of the philological society, 34,* 36-72.

Fontenelle, T. (1994). What on earth are collocations? An assessment of the ways in which certain words co-occur and others do not. *English Today*, *10*(04), 42-48.

Granger, S. (1998). Prefabricated patterns in advanced EFL writing: Collocations and formulae. In A. Cowie (Ed.), *Phraseology. Theory, analysis and applications* (pp. 145-160). Clarendon Press.

Halliday, M.A.K. (1966). Lexis as a linguistic level. In J. R Firth, C. E Bazell, J. C. Catford, M. A. K. Halliday & R. H. Robins (Eds.), *In memory of J.R. Firth* (pp. 150-161). Longman.

Hammarberg, B. (2001). Roles of L1 and L2 in L3 production and acquisition. In J. Cenoz, B. Hufeisen & U. Jessner (Eds.), *Cross-linguistic influence in third language acquisition: Psycholinguistic perspectives* (pp. 21-41). Multilingual Matters.

Keshavarz, M. H., & Salimi. H. (2007). Collocational competence and cloze test performance: A study of Iranian EFL learners. *International Journal of Applied Linguistics*, *17*(1), 81-92.

Laufer, B., & Waldman, T. (2011). Verb-noun collocations in second language writing: A corpus analysis of learners' English. *Language Learning, 61*(2), 647-672.

Lorenz, G. R. (1999). *Adjective intensification - Learner's versus native speakers: A corpus study of argumentative writing.* Editions Rodopi B.V.

McKeown, K. R., & Radev, D. R. (2000). Collocations. In R. Dale, H. Moisl, and H. Somers (Eds.), *A Handbook of Natural Language Processing* (pp. 507-523). Marcel Dekker.

Mel'čuk, I. (1998). Collocations and lexical functions. In A. P. Cowie (Ed.), *Phraseology: Theory, Analysis and Applications* (pp. 24-53). Clarendon Press.

Möhle, D. (1989). Multilingual interaction in foreign language production. In H.W. Dechert & M. Raupach (Eds.), *Interlingual processes* (pp. 179-194). Gunter Narr Verlag.

Nattinger, J., & DeCarrico, J. (1992). *Lexical phrases and language teaching*. Oxford University Press.

Nesselhauf, N. (2003). The use of collocations by advanced learners of English and some implications for teaching. *Applied Linguistics, 24*(2), 223-242.

Nesselhauf, N. (2005). *Collocations in a learner corpus* (Vol. 14). John Benjamins Publishing.

Odlin, T., & Jarvis, S. (2004). Same source, different outcomes: A study of Swedish influence on the acquisition of English in Finland. *International Journal of Multilingualism, 1*, 123-140.

Phoocharoensil, S. (2013). Cross-linguistic influence: Its impact on L2 English collocation production. *English Language Teaching*, *6*(1), 1-10.

Ringbom, H. (1987). *The role of the first language in foreign language learning*. Multilingual Matters.

Ringbom, H. (2001). Lexical transfer in L3 production. In J. Cenoz, B. Hufeisen & U. Jessner (Eds.), *Cross-linguistic influence in third language acquisition: Psycholinguistic perspectives* (pp. 59-68). Multilingual Matters.

Sinclair, J. (Ed.) (1991). *Corpus, concordance, collocation*. Oxford University Press.

Siyanova, A., & Schmitt, N. (2008). L2 learner production and processing of collocation: A multi-study perspective. *The Canadian Modern Language Review, 64*(3), 429–458.

Thewissen, J. (2008). The phraseological errors of French-, German- and Spanish-speaking EFL learners: evidence from an error-tagged learner corpus. In Associação de Estudos e de Investigação Científica do ISLA-Lisboa (Ed.), *Proceedings from the 8th Teaching and Language Corpora Conference (TaLC 8)* (pp. 300-306).

Wray, A. (2000). Formulaic sequences in second language teaching: Principle and practice. *Applied Linguistics, 21*(4), 463-489.