

An Empirical Analysis Based on a Corpus: Lexical Collocation Errors in the Writings of Chinese Learners of Italian L3 from a Trilingual Acquisition Perspective

Yisu Wang

Department of Classical Philology and Italian Studies, University of Bologna, Bologna 40126, Italy

Abstract

Lexical collocation is an essential aspect of learning Italian, but it remains a challenging area for Chinese learners. This study, grounded in the theories of trilingual acquisition and collocation, employs a combined methodology of error analysis and corpus analysis. Using a self-compiled corpus of Chinese learners of Italian L3, the writings of 296 university students majoring in Italian from seven universities in China were analyzed both quantitatively and qualitatively. The research explores specific causes of lexical collocation errors and proposes teaching strategies to address these issues.

Keywords

Chinese learners, Italian acquisition, lexical collocation, error analysis.

1. Introduction

Italian writing, as one of the fundamental skills that learners of Italian should master, reflects a comprehensive integration of students' linguistic and cognitive abilities. It is an extremely important yet particularly challenging aspect of learning Italian. Various factors affect writing quality, among which collocation errors are significant and cannot be overlooked. Incorrect or non-native-like Italian collocations can impact the overall quality of compositions. Research on collocation errors can shed light on the difficulties learners face in acquiring Italian vocabulary. Such findings can be promptly incorporated into teaching practices to enhance teaching effectiveness and alleviate learners' challenges. At present, studies on trilingual acquisition in Italian predominantly focus on learners whose native languages are European (De Angelis & Selinker 2001; De Angelis 2005; Bardel & Lindqvist 2007; Izzo et al. 2017; Badan et al. 2022; Zambianchi 2021). Research on errors specific to Chinese learners of Italian L3 in a trilingual context remains fragmented, and targeted studies on lexical collocation errors in Italian writing are even scarcer.

2. Theoretical Framework and Research Methodology

This study is based on the theories of trilingual acquisition and collocation. Combining error analysis with corpus analysis, it quantitatively and qualitatively examines the compositions of Chinese students majoring in Italian to identify and categorize collocation errors, explore their underlying causes, and propose corresponding solutions.

2.1. Theoretical framework

Research on trilingual acquisition started relatively late. Initially, scholars regarded trilingual acquisition as merely an extension of second language acquisition (SLA), paying little attention to distinguishing between the two concepts. However, as research advanced, it became evident that trilingual acquisition is more complex, exhibiting characteristics such as nonlinearity, complexity, and diversity (Zeng & Li, 2010).

For Chinese university students majoring in Italian, their language acquisition sequence is Mandarin (L1) > English (L2) > Italian (L3). Since English is a compulsory subject in Chinese middle and high schools, students typically have at least six years of English learning experience before starting Italian. During this period, they develop their own strategies and experiences in foreign language acquisition. These experiences and strategies from their first language (L1) and second language (L2) can significantly influence their trilingual acquisition process.

Corder (1967), one of the earliest proponents of error analysis, argued that errors are deviations from the norms of the target language caused by learners' incomplete knowledge of its rules. These deviations are often systematic rather than random. According to Corder, error analysis involves five steps: 1. Collection of a sample of learner language; 2. Identification of errors; 3. Description of errors; 4. Explanation of errors; 5. Evaluation of errors. This study adheres to these five steps in its analysis.

In 1957, Firth first introduced the concept of “collocation,” defining it as “the co-occurrence relationship between words.” He argued that there exists a natural network of associations between words and their meanings, though this network is often implicit. While it is an integral part of a native speaker's lexical competence, it remains unfamiliar to second-language learners. Numerous studies have shown that collocation errors constitute a significant proportion of errors among second-language learners (Howarth, 1998; Nesselhauf, 2003). Hausmann (1989) proposed six categories of lexical collocations: noun + noun, noun + verb, verb + noun, adjective + noun, verb + adverb, and adverb + adjective. These classifications have been widely recognized by researchers. Consequently, this study adopts Hausmann's classification of lexical collocations for annotating the corpus.

2.2. Research methodology

With the development of computer technology, corpus linguistics has emerged as a valuable field. In recent years, corpora have been widely applied in research on second language writing instruction. However, research combining corpus analysis with the study of lexical collocation errors in Italian writing remains insufficient.

For this study, the author collected written materials from 296 students majoring in Italian at seven universities in China to compile a small-scale corpus of Chinese learners of Italian, consisting of 56,108 words. Using Hausmann's classification of lexical collocations, the corpus was annotated to identify types of lexical collocation errors in writing. The annotated corpus was then imported into the AntConc (AntConc is a free software created and developed by Laurence Anthony of Waseda University in 2002. For further information, please visit the website <https://www.laurenceanthony.net/index.html>.) concordance tool for retrieval and statistical analysis.

This study aims to address the following questions through an analysis of Italian writings by Chinese students:

1. What is the distribution of the six types of lexical collocation errors in student writing?
2. What are the causes of these collocation errors?
3. What measures can teachers and students take to reduce lexical collocation errors in Italian writing?

3. Statistical Results and Analysis

3.1. Statistical results

The statistical results show that there are a total of 1,087 collocation errors in the corpus, distributed as follows:

Noun + Noun (EC1): 47 occurrences

Noun + Verb (EC2): 90 occurrences

Verb + Noun (EC3): 506 occurrences

Adjective + Noun (EC4): 380 occurrences

Verb + Adverb (EC5): 44 occurrences

Adverb + Adjective (EC6): 20 occurrences

As shown in Fig.1, the errors in Verb + Noun (EC3) and Adjective + Noun (EC4) account for a larger proportion, while the other types of errors are less frequent.

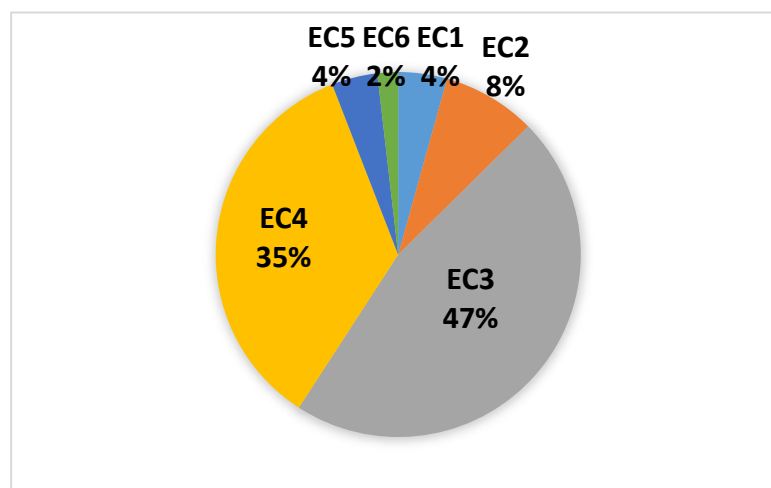


Figure 1. Proportion chart of six types of collocation errors

As shown in Fig. 2, the errors mainly originate from Chinese (L1), followed by the overuse of commonly used Italian words (L3), and finally, the influence of fixed expressions in English (L2).

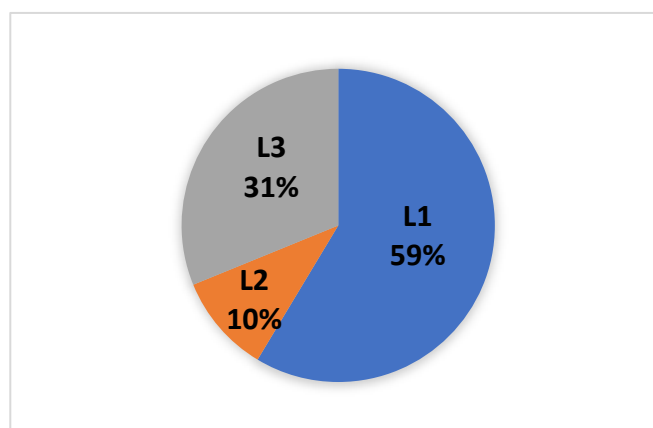


Figure 2. Proportion of language causing errors

3.2. Analysis and discussion

Next, this section will comprehensively consider the sources and types of collocation errors. A selection of prominent and representative errors will be analyzed.

1. Noun + Noun

These errors are mainly influenced by the word order in Chinese and fixed expressions in English. Typical errors include:

1. O spesso leggo romanzi, guardo la TV serie.
2. I miei genitori mi ha permesso di andare dal parrucchiere a fare un nuovo stile dei capelli.

In sentence 1, the learner intends to express “serie TV (TV series)”, but the correct expression in Italian should be “serie TV” rather than “TV serie”. This error is influenced by the word order in Chinese, where the sequence “电视” (TV) and “剧集” (series) directly corresponds to the structure of “TV serie”. In sentence 2, the learner wants to express “hairstyle”. The correct expression in Italian should be “taglio dei capelli”, but “stile dei capelli” appears in the sentence, which is clearly influenced by the English term “hairstyle”, where “stile” corresponds to “style” and “capelli” corresponds to “hair”.

2. Noun + Verb

There are three main reasons for these errors: First, the existence of a one-to-many semantic correspondence between Chinese vocabulary and Italian vocabulary often leads learners to base their word choices on Chinese semantics, resulting in inappropriate collocations in Italian. Second, there is an overuse of common Italian words; learners tend to rely on simple, frequently used words and may not have a proper grasp of more natural expressions. Finally, learners are influenced by negative transfer from English, often transferring English word usage to Italian expressions. Specific examples include:

3. Anche se il cane continua chiamare, non si aiuto.

4. Il tempo fa subito.

5. Come va? io sono bene.

In sentence 3, the learner starts with the Chinese verb “叫” (call) and selects the common verb “chiamare (call)” in Italian. In Chinese, the verb “叫” can refer to the sounds made by animals, such as “狗叫” (the dog barks) or “猫叫” (the cat meows). However, in Italian, the verb for the sound a dog makes is “abbaiare (bark)”, not “chiamare (call)”. In sentence 4, the learner overuses the Italian delexical verb “fare (do)”, extending its meaning beyond its proper context. The correct collocative verb should be “passare (pass)”. In sentence 5, the learner directly transfers the English expression “I am fine” into Italian as “io sono bene”, but the correct verb should be “stare”, as in “io sto bene”.

3. Verb + Noun

In the corpus search, it was found that errors of this type make up the highest proportion of collocation errors. Typical errors include:

6. Quando ero un bambino, circa a terzo anno nella scuola elementare, ho già chiesto una domanda a mia mamma nel letto, una sera.

7. Voglio spendere una giornata divertente con te.

8. Inoltre, ho studiato a prendere la macchina che penso non molto difficile.

These types of errors are mainly influenced by Chinese thinking patterns, English expression styles, and common fixed collocations in Italian. In sentence 6, “chiedere una domanda” is clearly a direct translation from the Chinese phrase “问一个问题 (ask a question)”, where “问” means “chiedere (ask)”, “一个” means “una (a)”, and “问题” means “domanda (question)”. However, in Italian, the correct collocation should be “chiedere una cosa (ask something)” or “fare una domanda (ask a question)”. In sentence 7, “spendere una giornata” is influenced by the English collocation “spend + time”, but in Italian, “spendere” is generally used with money and cannot be used with time. In sentence 8, the learner uses the common Italian collocation “prendere + a means of transport” to mean “take a mode of transportation,” using “prendere la macchina” to express “driving a car”. The correct collocation should be “guidare la macchina”.

4. Adjective + Noun

In this type of collocation, the most typical errors involve the confusion of “buono (good)”, “bello (beautiful, nice, fine)”, “bravo (good at something, capable)”, and “bene (well)”, as well as the overuse of common Italian adjectives. Specific examples include:

9. Io ho una idea brava.

10. Confucius aveva molto studenti bellissimi.

11. Il freddo sconosciuto divenne caldo.

12. Il film racconta le storie più calde delle ragazzine che hanno perso i genitori e che sono state mandate a vivere con il nonno.

In sentence 9, the learner intends to express “buona idea (good idea)”, and the correct adjective should be “buona” or “bella”, not “brava”. Similarly, in sentence 10, the learner wants to express “studenti bravissimi (good students)”, and the correct adjective should be “bravissimi”, not “bellissimi”. The main reason for Chinese learners confusing “buono”, “bello”, “bravo”, and “bene” lies in the fact that the Chinese word “好” corresponds to several different Italian words, creating a one-to-many mapping. In Chinese, “好” can be translated into “buono”, “bello”, “bravo”, and “bene” in Italian, and its use is very broad, functioning both as an adjective and an adverb. However, these four words in Italian have different grammatical categories and specific contexts for use. Learners often overlook these distinctions, leading to collocation errors.

The second typical error involves the semantic extension of common Italian adjectives, such as in sentences 11 and 12 with “caldo (hot, passionate, warm)”. In sentence 11, the learner uses “freddo (cold)” in the figurative sense of “indifferent” (“freddo sconosciuto” for “indifferent stranger”), and then incorrectly chooses “caldo” to convey the idea of “close” or “familiar”, which “caldo” does not mean in Italian. A similar error occurs in sentence 12, where the learner attempts to use “caldo” to mean “commovente (touching, moving)”.

5. Verb + Adverb

Errors involving verb and adverb collocations are relatively less frequent. A typical error is the confusion between “no (no)” and “non (not)”, such as in the following examples:

13. Perché no andare in palestra nel sabato pomeriggio?

The confusion between “no” and “non” typically appears in the corpus of beginner learners, and the main reason for this is the influence of Chinese. In Chinese, both words can function as negators for “没有” (meaning “not” or “none”). In Chinese, it is common to say “没有书 (no book)” or “没有吃 (did not eat)”, where both the noun and the verb can be negated. However, in both English and Italian, there is a distinction: “no” negates nouns (like in English “no books”), while “non” (like in English “not”) negates other sentence components.

6. Adverb + Adjective

Collocation errors involving adverbs and adjectives are the least frequent, typically manifesting as the misuse of the absolute superlative, such as:

14. Ho mangiato il gelato, molto buonissimo.

In sentence 14, “buonissimo” is the absolute superlative, and in Italian, the suffix “-issimo” is used to emphasize the degree of a quality. Here, “buonissimo” already conveys the meaning of “molto buono (very good)”, so it is unnecessary to further modify “buonissimo” with “molto (very)”.

3.3. Causes of collocation errors

By analyzing the typical examples of collocation errors above, we can identify several factors that contribute to these errors.

Influence of native language transfer. In the process of using a foreign language, when learners encounter a word, they do not know, their first instinct is to look for the corresponding word in their native language. They then try to find a corresponding word in the target language based on the meaning of the native language word. The problem arises when the meaning of the target language word does not perfectly match. Chinese learners of Italian tend to translate word-for-word from Chinese into Italian when they cannot recall an appropriate Italian

collocation, and due to the influence of their native language, the translation often becomes incomprehensible or even unacceptable.

Influence of English negative transfer. From a linguistic proximity perspective, while Chinese and Italian belong to different language families, English and Italian are both Indo-European languages, and there are many Latin-based words that are similar or identical in form between the two languages. Over time, however, words with the same origin and form may have different meanings and uses in each language. On one hand, the similarity between languages can help learners acquire a new language faster; on the other hand, words that appear similar but have different meanings in the two languages can cause errors. For instance, learners may translate “二年级 (second grade)” as “secondo grado”, which is clearly influenced by the English “second grade”. The correct expression in Italian should be “secondo anno”.

Limited input and output of Italian collocations. By analyzing collocation errors, we can see that learners tend to overuse simple, common words, leading to improper collocations. On one hand, students do not have enough input of collocations. The curriculum of Italian language programs in Chinese universities generally does not include courses focused on collocations. Furthermore, learners do not always pay attention to distinguishing synonyms that have similar meanings but different collocational patterns. As a result, when choosing words and collocations, learners often make inaccurate choices. On the other hand, learners still tend to choose simple or self-created collocations. A lack of sufficient input of authentic Italian collocations leads to a shortage of collocational knowledge. Infrequent output of natural Italian collocations causes learners to be unfamiliar with the collocations they have learned. Limited input and output of Italian collocations result in frequent collocation errors.

4. Conclusion

In response to the issues mentioned above, the author proposes three strategies to help learners better master Italian vocabulary collocations and reduce collocation errors in writing. **Cultivating learners' Italian thinking ability.** For Chinese learners of Italian L3, understanding and mastering the differences between Chinese and Italian, and gradually forming an Italian way of thinking, can effectively reduce collocation errors. First, schools should offer courses related to intercultural communication. Language is a carrier of culture, and it is also an embodiment of culture. During teaching, teachers should pay attention to explaining the differences between Italian and Chinese languages, as well as the cultural differences between China and Italy. Second, Italian language teachers should aim to teach entirely in Italian to provide students with an Italian learning environment, thus reducing the negative transfer from Chinese to Italian. Finally, students should read Italian literature, newspapers, and other materials. This will broaden their international perspective, help them understand foreign cultures, and gradually cultivate their ability to think in Italian.

Making full use of learners' previous language learning experience. A learner's native language and previously acquired foreign languages, regardless of whether they are similar or not to the language the learner is currently acquiring, should not be seen as "enemies" of the new language. Once learners understand the differences between these languages, cross-linguistic comparison can help them achieve positive transfer while being aware of the potential for negative transfer. This awareness can help learners avoid vocabulary errors caused by interlingual interference during the language acquisition process. Additionally, learners' foreign language learning awareness should be expanded. Since they have already learned a foreign language, this experience should be applied in the classroom and further developed. Teachers can help learners analyze the similarities and differences between their native language, their first foreign language, and their third language. Understanding these structural differences will help learners better distinguish between languages, reduce confusion between

similar languages, and utilize the structural similarities to facilitate their language acquisition, improving their third language proficiency.

Increasing input and output of Italian collocations. First, schools should offer courses related to vocabulary, explaining authentic Italian collocations and the distinction between near synonyms to increase students' input of Italian collocations. Second, corpora can be introduced into the classroom. By learning vocabulary in specific contexts through corpora, learners can explore the semantic tendencies of words, that is, the frequent co-occurrence patterns of words with surrounding words. This approach allows learners to visually grasp the common collocations of a word, enabling them to better master and use these collocations. Additionally, teachers should encourage students to create sentences, write essays, and craft dialogues using the collocations they have learned in class to reinforce their knowledge of collocations and promote their output of Italian collocations.

References

- [1] Angelis G D. Interlanguage transfer of function words[J]. *Language learning*, 2005, 55(3): 379-414.
- [2] Badan L, Irene C, Giuliano I. Fenomeni di interferenza nell'italiano L3 di nederlandofoni: conseguenze per la didattica[J]. *INCONTRI*, 2022, 36(2): 1-16.
- [3] Bardel C, Lindqvist C. The role of proficiency and psychotypology in lexical cross-linguistic influence. A study of a multilingual learner of Italian L3[C]//Atti del VI Congresso di Studi dell'Associazione Italiana di Linguistica Applicata. Perugia: Guerra Editore, 2007: 123-145.
- [4] Corder S.P. The Significance of Learners'Errors [J]. *International Review of Applied Linguistics*, 1967 (04).
- [5] De Angelis G, Selinker L. Interlanguage transfer and competing linguistic systems in the multilingual mind[J]. *Bilingual Education and Bilingualism*, 2001: 42-58.
- [6] Ellis R. *The Study of Second Language Acquisition*[M]. Shanghai: Shanghai Foreign Education Press 1999.
- [7] Firth J. R. *Modes of Meaning* [A]. *Essays and Studies* [C]. Oxford: Oxford University Press 1951.
- [8] Hausmann F J. Le dictionnaire de collocations[J]. *Wörterbücher, Dictionaries, Dictionnaires*, 1989, 1: 1010-1019.
- [9] Howarth P. The phraseology of learners' academic writing[J]. *Phraseology*, 1998: 161-187.
- [10] Izzo G, Cenni I, De Smet J. Third language acquisition and its consequences for foreign language didactics: the case of Italian in Flanders[C]//Van Schooten tot Scriptie III-Een colloquium over universitair taalvaardigheidsonderwijs. 2017: 61-72.
- [11] Nesselhauf N. The use of collocations by advanced learners of English and some implications for teaching[J]. *Applied linguistics*, 2003, 24(2): 223-242.
- [12] Zambianchi L. «RIPOSO LA MIA VALIGIA, VOSTRO ONORE». INTERFERENZE NELL'ITALIANO L3 DI APPRENDENTI ANGLOFONI[J]. *LinguaInAzione*, 33.
- [13] Zeng Li, Li Li. Reflections on "Trilingual Acquisition" as an Independent Research Field [J]. *Foreign Languages and Their Teaching*, 2010 (2): 6-9.