Nikolic, J. (2026). Lexical transfer in l3 Chinese: a qualitative study of collocational errors among Serbian-speaking learners, *SCIENCE International journal, 5*(1), 161-166.
doi: 10.35120/sciencej0501161n UDK: 811.581'243:37.091.3

LEXICAL TRANSFER IN L3 CHINESE: A QUALITATIVE STUDY OF COLLOCATIONAL ERRORS AMONG SERBIAN-SPEAKING LEARNERS

Jelena Nikolic[1*]

[1]University of East Sarajevo, Bosnia and Herzegovina,
e-mail: jelena.nikolic@ffuis.edu.ba

Abstract: This paper investigates lexical transfer in L3 Chinese through a qualitative analysis of collocational errors produced by Serbian-speaking learners. The study focuses on how cross-linguistic influence shapes learners' use of Chinese collocations and how conceptual patterns from the first and second language, as well as dictionary-mediated lexical choices, lead to semantically deviant combinations. A corpus of learner collocations, gathered from written production across several HSK proficiency levels, was compiled and manually annotated. Each erroneous collocation was analysed in terms of presumed source expression in Serbian or English, and type of transfer involved. The analysis identifies two major groups of errors: (1) conceptual transfer from Serbian or English and (2) errors driven by "false synonyms", where near-synonymous items treated as equivalent in dictionaries and English-based materials are not interchangeable in Chinese. The results indicate that collocational competence in Chinese as a third language is strongly constrained by L1 and L2 conceptualisation, the mediating role of English, bilingual lexicographic practices and structural properties of Chinese, rather than by overall proficiency alone. The study concludes that teaching materials and dictionaries for Serbian speaking learners of Chinese should be designed to highlight collocations rather than isolated words, systematically contrast Serbian, English and Chinese ways of expressing high-frequency concepts, and make subtle semantic and collocational differences between near-synonymous items explicit. As additional data, the study offers a qualitatively annotated set of learner collocations, their reconstructed Serbian and English source expressions and their target equivalents in Chinese, which can be used as a resource for further research and for the development of classroom activities focusing on collocational awareness.

Keywords: *lexical transfer; L3 Chinese; collocational errors; cross-linguistic influence; conceptual transfer; false synonyms; qualitative error analysis*

Field: Humanities

1. INTRODUCTION

Collocations, which are recurrent, conventional combinations of content words, are a central component of vocabulary knowledge and play a key role in fluent, idiomatic language use. For learners of Chinese as a third language, collocational competence poses particular challenges. Chinese is a typologically analytic language without inflectional morphology, and it relies heavily on function words and fixed multi-word patterns to express grammatical and semantic relations. This means that the learner cannot rely on inflectional cues to signal argument structure or word class, and small changes in lexical choice often have large effects on acceptability. At the same time, high-quality, easily accessible collocation resources for Chinese have historically been scarce, which has prompted the recent developments of tools based on a large-scale collocational knowledge base, intended to support both teaching and self-directed learning (Hu, et al., 2016). In terms of research, collocational errors in Chinese have been studied far less extensively than in better-researched target languages. In comparison with the extensive body of work on collocations and collocational errors in L2 English, empirical research on phraseology and collocation use in L2 Chinese remains relatively limited, a situation reflected in the small number of available L2 Chinese learner corpora (Wu et al., 2025; Iurato, 2022).

Research on crosslinguistic influence shows that prior languages affect learners' lexical choices and structural preferences in both facilitating and interfering ways, leading to conventional target-like uses as well as systematic deviations from target norms (Jarvis & Pavlenko, 2007). Research on L3 acquisition indicates that cross-linguistic influence from previously acquired languages is conditioned by factors such as typological similarity, learners' proficiency in those languages, their recency of use, and L2 status (Hammarberg, 2001). Recent large-scale research on Chinese as an additional language has shown that the linguistic distance between learners' L1 and Chinese affects the acquisition of different components of language knowledge, including characters, vocabulary, and grammar (Chai & Bao, 2023). While there is growing evidence that L1 also shapes learners' collocational knowledge in Chinese (e.g. the studies reviewed in Chai & Bao, 2023), systematic work on collocational transfer, especially for learners from

*Corresponding author: jelena.nikolic@ffuis.edu.ba

South Slavic backgrounds, remains limited. Experimental work on dictionary use for collocation learning shows that electronic dictionaries can significantly improve learners' productive collocation knowledge, but many learners struggle to exploit dictionary information effectively; they often fail to use hyperlinks, have difficulty distinguishing between subsenses, and tend to rely on the first sense in overcrowded entries (Chen, 2017). In such situations, it is plausible that learners will treat near-synonymous items as interchangeable and overlook subtle collocational and register constraints, especially when these are not made salient in the reference tools they consult.

The present study addresses these gaps by examining lexical transfer in the collocational choices of Serbian-speaking learners of Chinese, for whom Chinese functions as an L3 learned after at least one foreign language. Building on a corpus of written production compiled across several HSK-based proficiency levels, the analysis focuses on collocational errors and explores how cross-linguistic influence shapes learners' use of Chinese collocations. In particular, the study investigates two interrelated sources of transfer: (a) conceptual transfer from Serbian, where learners map L1 conceptualisations and phraseological patterns directly onto Chinese; (b) transfer mediated by "false synonyms", where dictionary or English-based materials suggest equivalence between near-synonymous items that are not interchangeable in Chinese, also sometimes linked to overgeneralisation of Chinese word-class flexibility in interaction with Serbian morphosyntax. By providing a qualitative account of these error types in the specific context of Serbian learners of L3 Chinese, the paper aims to deepen our understanding of how L1 conceptualisation, lexicographic practices and structural properties of Chinese jointly shape the development of collocational competence, and to draw pedagogical implications for the design of materials and classroom activities that address collocational constraints and cross-linguistic influence more explicitly.

2. METHODS AND MATERIALS

The corpus for this study consists of written texts produced by learners of Chinese as L3 over the course of one academic year (2024/2025) at the University of East Sarajevo (Bosnia and Herzegovina). At the beginning of the academic year, the students completed a language repertoire questionnaire (L1, L2, L3) to establish their linguistic backgrounds and eliminate the influence of other languages. Of the 47 students in total (third- and fourth-year students studying Chinese as a major), 42 informants were included in the study, as they reported Serbian as their L1, English as their L2, and Chinese as their L3. The remaining five students, who reported German or Russian as their L2, were excluded from the analysis. They were excluded based on possible cross-linguistic influence (CLI) from different L2s and in order to keep the sample as homogeneous as possible.

HSK (Hànyǔ Shuǐpíng Kǎoshì) is the standardised proficiency test for learners of Chinese as a foreign language (Chen, 2022). Broadly speaking, HSK levels 3 and 4 correspond to intermediate levels of proficiency, while HSK 5 indicates a lower-advanced level. Among the 42 students included in the study, 17 (40.5%) held an HSK 3 certificate, 9 (21.4%) held HSK 4, 8 (19.0%) held HSK 5, and the remaining 8 students (19.0%) had not yet taken the HSK exam. Reporting these levels is important because they provide an external, independent indication of the learners' overall command of Chinese, which helps to contextualise their collocational performance and to distinguish errors that may be related to general developmental stage from those more clearly attributable to cross-linguistic influence.

In total, more than 400 homework essays were collected; from these, 112 texts that contained at least one non-target-like collocation were selected for analysis. Only texts in which a collocational error could be clearly identified were included in the corpus examined in this paper. The texts were primarily homework essays, each no longer than one A4 page, written on specific topics closely related to the lessons covered in class.

Collocations are habitual combinations of words that recur frequently in texts and form an important part of speakers' mental lexicon (Yu et al., 2025). They are understood here as frequent, conventional combinations of at least two content words that stand in a syntactic or semantic relationship (Xu, 2007) and, in corpus-based work, as co-occurrences whose frequency is statistically higher than would be expected by chance (Lehecka, 2015; Gablasova et al., 2017). These definitions underlie the selection of target and non-target combinations in the learner data.

The first step in the analysis was to note all multi-word combinations in the learner corpus that appeared to be non-idiomatic or structurally unusual in Chinese. Only combinations fulfilling the above definition of collocation were retained for further analysis. Following the researcher's initial identification and analysis of collocations, their correctness was further verified through native-speaker judgement by two Chinese lecturers. All candidate collocations were manually examined by these lecturers, who

assessed whether each combination was acceptable in contemporary standard Chinese. In cases where a collocation was judged non-target-like, they proposed the conventional expression used in actual language use, and this was adopted as the target form for the subsequent analysis.

3. RESULTS

The qualitative analysis of the learner corpus revealed a set of recurrent non-target-like collocations in the Chinese written production of L1 Serbian-speaking students. These collocations could be grouped into two main types according to their primary source: (a) combinations shaped by the "logic" of Serbian (L1) or English (L2), and (b) false synonyms and dictionary-mediated transfer.

a) A number of the non-target-like collocations were traced to conceptual transfer from L1 Serbian or L2 English. In these cases, learners appeared to translate L1 or L2 expressions into Chinese in a largely word-for-word fashion, preserving Serbian- or English-based conceptual structures, "logic" and lexical choices even where Chinese uses different patterns. For instance, the verb–object collocation 吃汤 chī tāng 'eat soup' was used instead of the target-like 喝汤 hē tāng 'drink soup': while Serbian or English conceptualises jesti supu / eat soup, Chinese encodes the consumption of soup with the verb 'drink'. A similar pattern emerged in expressions of language proficiency, where learners produced 知道汉语 zhīdào hàny 'know Chinese' modelled on the Serbian expression 'znati iezik' or the English 'to know a language', whereas in Chinese the relevant ability is expressed by 会(说)汉语 huì (shuō) hàny 'be able (to speak) Chinese'. Another similar example is 喝药 hē yào 'drink medicine', where the default Chinese collocation is 吃药 chī yào 'take (lit. eat) medicine'. This mirrors the Serbian expression piti lijek 'drink medicine', which is the usual way to describe taking medicine regardless of its physical form. In Chinese, however, 药 yào 'medicine' most often collocates with 吃 chī 'eat', and 喝药 hē yào 'drink medicine' is reserved for cases where the medicine is specifically in liquid form. There was another representative example: Serbian uses the verb nositi both for wearing clothes (nositi jaknu "wear a jacket") and for carrying bags (nositi torbu "carry a bag"), whereas Chinese systematically distinguishes between 穿 chuān "wear (clothes, shoes, etc.)" and 背 / 提 bēi / tí "carry (on the back / in the hand)". In the corpus we had the example of the collocation 穿包 chuān bāo "wear a bag" by analogy with 穿衣服 "wear clothes", instead of the idiomatic 背包 bēi bāo "carry a bag / backpack". The collocation 生活生活 shēnghuó shēnghuó, intended as a direct transfer of the Serbian pattern živjeti život (verb 'live' + noun 'life') or the English to live life, also stood out. Although 生活 shēnghuó can function as both a verb and a noun in Chinese, these two uses do not normally appear side by side, and the idiomatic expression would be (过)生活 (guò) shēnghuó '(to) lead/live one's life'. Such examples indicate that students overapplied the possibility of word-class conversion in Chinese, combining it with Serbian phraseological patterns. Most of the examples from the corpus include learners combining the adjective 漂亮 (piàoliang, 'beautiful') with nouns such as 茶 (chá, 'tea'), 咖啡 (kāfēi, 'coffee') and 天气 (tiānqì, 'weather'), producing collocations like 茶漂亮 (chá piàoliang, lit. 'tea [is] beautiful'), 咖啡漂亮 (kāfēi piàoliang, lit. 'coffee [is] beautiful') and 天气漂亮 (tiānqì piàoliang, lit. 'weather [is] beautiful'). These expressions are formally grammatical, but they are not idiomatic in contemporary standard Chinese. The problem lies in the mismatch between the semantic profile of 漂亮 piàoliang and the preferred evaluative patterns for these nouns. In Chinese, 漂亮 (piàoliang) is primarily used for visual attractiveness (e.g. 漂亮的衣服 piàoliang de yīfu, 'beautiful clothes'; 漂亮的风景 piàoliang de fēngjing, 'beautiful scenery'), and is not normally used to evaluate the taste of drinks or the quality of the weather.

These cases, among others, showed that learners sometimes relied on L1 or L2 phraseological templates when constructing collocations in L3 Chinese, even when the resulting expressions were pragmatically unnatural or conceptually unnatural in the target language.

b) The second group of errors involved false synonyms, where learners selected near-synonymous Chinese items as if they were fully interchangeable. For example, the collocation 住生活 zhù shēnghuó 'live life' frequently appeared instead of the idiomatic 过生活 guò shēnghuó 'lead/live (one's) life'. In the learners' principal reference tools, the Pleco Chinese Dictionary (version 3.2.80) and the textbook New Practical Chinese Reader (2003, p. 209), 住 zhù is glossed in English as "live, reside, stay" or "to live, to stay", respectively. Such broad translations may have blurred its core meaning of 'reside in a place' and thereby encouraged learners to treat it as synonymous with more general verbs meaning 'to live'. There were many other examples of errors that showed transfer mediated by English, reflecting the role of English both as the learners' L2 and as the metalanguage of many dictionaries and textbooks. One clear example involved the verb 回 huí. In the above-mentioned reference tools used by the students, the Chinese verb 回 was translated into English simply as "return, go back". When this English gloss was subsequently translated into Serbian, the reflexive nuance present in Chinese - 'to return oneself (to a

place)' - was in some cases further obscured. As a result, learners produced collocations such as 回钱 huí qián 'return (self) + money' and 回书 huí shū 'return (self) + book' instead of the target 还钱 huán qián and 还书 huán shū, where 还 huán is the appropriate verb for 'giving something back'. In another case, learners produced the collocation 国家表面 (guójiā biaomiàn, lit. 'country surface') instead of the idiomatic 国家面积 (guójiā miànjī, lit. 'country area'). This error can be traced to the English (and Serbian) noun area, which can refer both to a geographical expanse (the area of a country, površina države) and to a more general surface (surface area, površina stola). In Chinese, however, these two meanings are lexicalised by different nouns: 面积 (miànjī) is used for the size/area of land, rooms, countries, etc., whereas 表面 (biaomiàn) typically refers to the surface of an object (e.g. the surface of a table, the surface of water) and does not collocate with 国家 'country' in this sense. Other examples of this type include 开自行车 kāi zìxíngchē (lit. 'drive a bicycle') instead of the idiomatic 骑自行车 qí zìxíngchē 'ride a bicycle'; 打扫街道 dasao jiēdào 'clean/sweep the street' instead of 打扫房子 qīngsao jiēdào 'sweep/clean the street'; 打扫家 dasao jiā 'clean the home/family' instead of 打扫房子 dasao fángzi 'clean the house/flat'; and 打开饭馆 dakāi fànguan (lit. 'open [like a door] a restaurant') instead of 开饭馆 kāi fànguan 'start/run a restaurant'. For learners with Serbian as L1 and English as L2, these errors are readily interpretable as transfer from overlapping patterns in the previously acquired languages: Serbian uses voziti both for vehicles and bicycles (voziti auto, voziti bicikl), which encourages the choice of 开 'to drive, to operate' rather than 骑 'to ride'; the verb 'čistiti' is used broadly for cleaning both houses and streets, which supports extending 打扫 to contexts where Chinese prefers 清扫 or a more specific object such as 房子; and otvoriti 'to open' in Serbian, like 'to open' in English, is used both for physical opening and for starting a business (otvoriti restoran / open a restaurant), which makes 打开饭馆 appear plausible even though Chinese distinguishes 打开 for concrete, physical opening and 开 for starting or running an enterprise.

This pattern highlighted the limitations of using English as a universal metalanguage in lexicographic and pedagogical materials and underscored the need for language-specific resources that are tailored to Serbian learners and explicitly address the transfer risks arising from their L1 and L2. The absence of a clear distinction in the translation sometimes encouraged learners to treat similar Chinese verbs as interchangeable. In such cases, students chose the item that matched the translation they associated with the Serbian source expression, even when its collocational and semantic restrictions in Chinese were narrower or different.

4. DISCUSSIONS

The findings of this study indicate that collocational errors in L3 Chinese among L1 Serbian, L2 English learners were not accidental slips, but systematic reflections of how prior languages and learning tools shaped their emerging collocational knowledge. At the core of many erroneous combinations was the projection of entrenched L1/L2 phraseological frames onto Chinese, such as 'eat soup', 'know a language', 'live life' or 'drink medicine', with only partial adjustment to target norms. Learners activated familiar multilingual templates and filled them with Chinese lexical material, producing combinations that were grammatically possible but pragmatically non-idiomatic. This pattern suggests that, for these learners, collocational transfer operated primarily at the level of conceptual and phraseological units, not isolated lexemes: once an L1 or L2 frame was selected, it strongly guided the choice of Chinese verbs and nouns, even when Chinese encoded the same situation through different, often more specialised, collocational patterns.

The study also showed that lexicographic mediation was not a neutral support, but an additional source of shaping the collocations. English glosses and Serbian translations often compressed several Chinese items into a single broad equivalent ('to live', 'to return', 'area'), which made near-synonymous verbs and nouns appear mutually interchangeable. In such cases, learners tended to select the Chinese item that best matched the translation they had internalised, rather than the one that best matched actual Chinese collocational norms. This mapping - Serbian ⇄ English ⇄ Chinese - was especially problematic when Serbian itself did not lexicalise a contrast that is overtly encoded in Chinese (e.g. "reside" vs "lead one's life", "return oneself" vs "give something back", "surface" vs "area of land"). The errors involving 住生活, 回钱 / 回书 or 国家表面 are therefore best understood not simply as vocabulary gaps, but as consequences of how dictionary entries and textbook glosses package semantic space: where the tools present a single, undifferentiated label, learners are encouraged to treat distinct Chinese items as false synonyms in collocation.

A further contribution of the study lies in its treatment of word-class flexibility in Chinese. The data showed that learners were aware that many Chinese forms (e.g. 生活) can function as both verbs and

nouns, but they sometimes overgeneralised this flexibility, combining two possible uses in ways that are structurally conceivable yet not idiomatic, as in 生活生活. This finding nuances the common view of Chinese as "simpler" because of its lack of inflection: the absence of overt morphological marking does not reduce complexity, but rather shifts it into the domain of lexical patterning and collocation. For L3 learners coming from a morphologically rich L1 like Serbian, where case and inflection help signal argument structure and word class, the burden of interpretation in Chinese is redistributed onto fixed multi-word patterns and conventional pairings. The present results suggest that, without sustained exposure to authentic collocations and explicit attention to their distribution, learners may use word-class flexibility as a licence for creative recombination that exceeds what is acceptable in actual usage.

Taken together, these observations highlight that the development of collocational competence in L3 Chinese is constrained by a network of interacting factors: L1 or L2 conceptualisation, L2 English as a mediating system, lexicographic practices, and language-internal properties of Chinese. They also point to several implications for research and pedagogy. For research, the study underscores the value of qualitative, learner-specific analyses that trace collocational errors back to concrete L1/L2 source expressions and dictionary entries, particularly for under-represented language pairings such as South Slavic L1s and L3 Chinese. For pedagogy, the results argue for materials and tasks that (a) foreground collocations rather than single words, (b) systematically contrast Serbian, English and Chinese ways of expressing high-frequency meanings, and (c) make semantic, register and collocational differences between near-synonymous Chinese items explicit.

Finally, the study's limitations, such as informants from a single institution, written homework data, and a focus on intermediate to lower-advanced learners, suggest clear directions for future work, including analyses of spoken collocations, longitudinal tracking of collocational development, and experimental studies of how different dictionary designs affect learners' collocational choices in L3 Chinese.

5. CONCLUSIONS

This study examined lexical transfer in L3 Chinese through a qualitative analysis of collocational errors in the written production of Serbian-speaking learners with English as their L2. The analysis showed that non-target-like collocations were not sporadic mistakes, but systematic outcomes of how learners mobilised their existing linguistic resources when writing Chinese. Two main sources of deviation were identified. First, many erroneous collocations arose from the direct projection of entrenched L1 and L2 phraseological frames onto Chinese: expressions such as 'eat soup', 'know a language', 'live life' or 'drink medicine' were transparently mapped onto Chinese verbs and nouns, yielding combinations that were structurally possible but not idiomatic (e.g. 吃汤, 知道汉语, 生活生活, 喝药). Second, a substantial group of errors involved false-synonym choices and dictionary-mediated transfer: broad English glosses and Serbian translations encouraged learners to treat distinct Chinese items (e.g. 住 vs. 过生活, 回 vs. 还, 表面 vs. 面积) as interchangeable, leading to collocations such as 住生活, 回钱, or 国家表面 that are semantically interpretable yet infelicitous. A smaller set of examples also pointed to overextension of Chinese word-class flexibility, where learners combined verb and noun uses of the same form in ways that go beyond accepted usage.

These findings have several implications for research on L3 Chinese and for the teaching of Chinese to learners with South Slavic backgrounds. They show that collocational competence in L3 Chinese is jointly shaped by L1 and L2 conceptualisation, the mediating role of English, the way bilingual reference tools package semantic space, and language-internal properties of Chinese, rather than by overall proficiency or vocabulary size alone. For the field, the study underlines the importance of examining collocation and phraseology in under-researched language constellations and of tracing collocational errors back to concrete L1/L2 source patterns and dictionary entries.

For pedagogy, the results point to the need for language-pair-sensitive materials and tasks that foreground collocations instead of isolated words, systematically contrast Serbian, English and Chinese ways of expressing high-frequency concepts, and make subtle semantic and collocational differences between near-synonymous Chinese items explicit. Such an approach can support the gradual restructuring of learners' multilingual mental lexicon around Chinese-specific patterns of lexical combination and, ultimately, more successful development of idiomatic collocational use in L3 Chinese.

ACKNOWLEDGEMENTS

I would like to express my sincere gratitude to prof. Xiao Jingfen and prof. Su Xiaolei, who generously contributed their time and expertise as native-speaker consultants.

REFERENCES

Chai, X., & Bao, J. (2023). Linguistic distances between native languages and Chinese influence acquisition of Chinese character, vocabulary, and grammar. Frontiers in Psychology, 13, 1083574. https://doi.org/10.3389/fpsyg.2022.1083574

Chen, J. Y.-W. (2022). Implementation of the Hanyu Shuiping Kaoshi and its home edition during the COVID-19 pandemic: A survey of European test centres. International Journal of Chinese Education, 11(2), 1–21. https://doi.org/10.1177/2212585X221100877

Chen, Y. (2017). Dictionary use for collocation production and retention: A CALL-based study. International Journal of Lexicography, 30(2), 225–251. https://doi.org/10.1093/ijl/ecw005

Gablasova, D., Brezina, V., & McEnery, T. (2017). Collocations in corpus-based language learning research: Identifying, comparing, and interpreting the evidence. Language Learning, 67(S1), 155–179. https://doi.org/10.1111/lang.12225

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.

Hu, R., Chen, J., & Chen, K.-h. (2016). The construction of a Chinese collocational knowledge resource and its application for second language acquisition. In Proceedings of COLING 2016, the 26th International Conference on Computational Linguistics: Technical Papers (pp. 3254–3263). The COLING 2016 Organizing Committee.

Iurato, A. (2022). Learner corpus research meets Chinese as a second language acquisition. Annali di Ca' Foscari. Serie orientale, 58, 709–742.

Jarvis, S., & Pavlenko, A. (2007). Crosslinguistic influence in language and cognition. Routledge. https://doi.org/10.4324/9780203935927

Lehecka, T. (2015). Collocation and colligation. In P. Hanks & G.-M. de Schryver (Eds.), International handbook of modern lexis and lexicography (pp. 1–17). Springer.

Wu, J., Lu, X., Hu, R., Lin, Y., & Liu, F. (2025). Syntactic and phraseological complexity in Chinese as a second language adapted teaching materials. The Modern Language Journal, 109(2), 326–343. https://doi.org/10.1111/modl.12994

Xu, R., Lu, Q., Wong, K.-F., & Li, W. (2007). Annotating Chinese collocations with multi information. In B. Boguraev, N. Ide, A. Meyers, S. Nariyama, M. Stede, J. Wiebe, & G. Wilcock (Eds.), Proceedings of the Linguistic Annotation Workshop (pp. 61–68). Association for Computational Linguistics.

Yu, M., Xu, S., Yang, L., & Chen, S. (2025). The Influence of Input Frequency and L2 Proficiency on the Representation of Collocations for Chinese EFL Learners. Behavioral Sciences, 15(1), 46. https://doi.org/10.3390/bs15010046