

Philipp Stöckle, Wolfgang Koppensteiner,
Katharina Korecky-Kröll, Maria Luisa Krapp, Tanja Wissik (eds.)

LEXICOGRAPHY IN THE AGE OF AI

PROCEEDINGS OF THE XXII EURALEX INTERNATIONAL CONGRESS 2026

29 September - 3 October 2026, Vienna | Austria



Editors: Philipp Stöckle, Wolfgang Koppensteiner, Katharina Korecky-Kröll, Maria Luisa Krapp, Tanja Wissik

Cover & title page design: Sandra Lehecka

Some rights reserved.

Copyright © Austrian Academy of Sciences, Vienna 2026

Typesetting: Sunrise Setting, Brixham, UK

ISBN: 978-3-7001-5137-1

<https://epub.oeaw.ac.at/5137-1>

<https://verlag.oeaw.ac.at>

Zongyuan Wang, Valeria Caruso & Anna De Meo

COLLABORATIVE LEXICOGRAPHY FOR HERITAGE LANGUAGE LEARNING

Abstract This study explores the potential of collaborative dictionary compilation to promote vocabulary retention and contributes to research in Chinese heritage language education. A pilot intervention was conducted with 22 second-generation Chinese students attending a complementary school in Italy. The participants studied literary vocabulary from Lu Xun's *Village Opera* and collaboratively compiled dictionary entries using Lexonomy. To assess the broader effectiveness of the protocol and compensate for the absence of a control group of heritage learners in Italy, the framework was replicated with two age-matched groups in China: an experimental group of 50 students engaged in dictionary-compilation activities and a control group of 53 students receiving traditional teacher-led instruction. Vocabulary knowledge was assessed through a pre-test and a delayed post-test administered three months after the intervention. A mixed-effects logistic regression analysis showed that the experimental group was approximately twice as likely as the control group to improve its retention of the literary vocabulary studied. The findings provide preliminary evidence that collaborative lexicography can support medium-term literary vocabulary acquisition and culturally oriented language learning, although larger-scale studies involving students from complementary schools are needed.

Keywords Heritage language learning; collaborative lexicography; Chinese as a heritage language; vocabulary acquisition; dictionary compilation; lexonomy; literacy development

1 Preliminary Remarks on Dictionaries and Societies

This paper explores the role that lexicography can play in societies and among specific groups of people whose lives and social engagement may be improved through the use of dictionaries. If lexicography is to become a discipline more closely concerned with people and real-world contexts, it may have to sacrifice some degree of precision, run the risk of relying on unbalanced data, and even produce partial or negative evidence. However provisional or even disappointing a real-world context may appear because of its imbalanced components, it should not be disregarded but rather used as a starting point for further improvement, accumulating evidence, replication studies, and clearer demonstrations of the role dictionaries play in the lives of their users.

The current study addresses a complex global context, populated by migrant communities with strong cultural identity that are rapidly evolving in relation both to their host societies and to their cultures of origin. Despite the common features that newly established Chinese communities abroad may share, and the Chinese state's "resilient capacity" to facilitate "diasporic participation in China's socio-economic and political processes" (Liu, 2011, p. 813), the Chinese diaspora has been described as "one of the most hyperdiverse groups, with substantial subgroup differences that distinguish it from most other diasporas" (Guo, 2022, p. 847).

Against this social background, Chinese complementary schools¹, which are found across the world, play a particularly important role. Their social mission varies from country to country,

¹ A complementary school is a community-based organisation, often established by parents, that provides language and cultural education outside regular school hours to maintain a community's linguistic and cultural heritage. Such schools are a widespread and distinctive feature of the Chinese diaspora (Curdt-Christiansen & Hancock, 2014). These schools are also referred to as *heritage language schools*, *community language schools*, and *supplementary schools*. It is worth noting that a UK government report commissioned by the Department for Children, Schools and Families uses the term *supplementary schools* (Maylor et al., 2010).

depending largely on the characteristics of the host society and of the migrant community itself. In Italy, these schools began to appear in the second half of the 1990s and, because they are self-governed and are often established through the initiatives of private members of the immigrant community – such as entrepreneurs, they have developed without coherent planning, display highly diverse characteristics, and still lack a comprehensive census (Pellin, 2025), let alone a clear overview of the teaching practices adopted in each institution.

With respect to the Italian context, Brigadoi Cologna (2023) notes the tendency to promote a standardised, supranational culture centred on Pǔtōnghuà (普通话), with the support of the PRC government, and the difficulties that this may create within migrant communities whose members predominantly speak Wu varieties of Sinitic, which belong to a branch distinct from Mandarin.

Setting aside the complex issue of which language should be taught in complementary schools, it should be emphasised that conducting coherent and structured research within these institutions is as complex as it is desirable. Their geographical dispersion and the absence of coordinating bodies² make it difficult to establish collaborations and carry out research involving their students. Such research is nevertheless highly desirable, since support for the development of young learners' L1 is recognised in the relevant literature as a key factor in the educational success of students from immigrant backgrounds (see, at least, Cummins, 1979; Thomas & Collier, 2002).

However, given the current organisation of Chinese complementary schools in Italy and the objective difficulty of recruiting large numbers of learners, the present study should be regarded as a starting point for investigating the benefits of integrating innovative dictionary-based programmes into this highly specific educational context. It provides preliminary evidence of both the effectiveness of the learning protocol and its feasibility for second-generation Chinese students attending these schools across Italy. Nevertheless, larger-scale studies involving more students and a greater number of complementary schools are still needed.

2 Introduction

In the teaching of Chinese as a heritage language, particular importance is attached to the systematic integration of cultural and community dimensions into language instruction (He, 2010; Carreira & Kagan, 2011). Building on this premise, the present study proposes a pedagogical approach that combines digital lexicographic technologies with the study of Chinese literature. Its aim is to strengthen heritage language learners' connection with the culture of their country of origin while improving their command of different language registers, given that their linguistic repertoires are often unevenly distributed across domains and varieties (Carreira, 2004; Valdés, 2001). To implement this dictionary-compilation framework, which transforms the dictionary from a source of information into a repository of structured lexical data compiled by the learners themselves, the study addresses two interrelated research questions: whether collaborative dictionary compilation is an effective strategy for learning new vocabulary, and whether this approach can also be successfully implemented in more complex social settings, such as heritage language education in a diasporic context.

² Two recent initiatives are seeking to provide a clearer picture of Chinese complementary education in Italy, thereby addressing the limited attention paid by Italian institutions to migrant-community multilingualism and to the educational success of second-generation students. These are the PRIN 2022 PNRR project ECCO-Italy – Education of Children of Chinese Origin in Italy (Pellin, 2025) and the Observatory on Italian Heritage Schools currently being developed by the Centre for Research on Minorities (CERM) at the University of Insubria (Brigadoi Cologna, 2023).

A pilot study was conducted with a group of second-generation students of Chinese origin (Heritage Group, HG) attending a Chinese complementary school in Italy. The experiment aimed to assess the overall design of the teaching protocol and the types of activities to be carried out alongside dictionary compilation. Regrettably, the limited number of participants made it impossible to establish a control group, while the lack of computers for compiling the learners' dictionary prevented the study from meeting all the conditions required for a fully robust experimental design.

The participants attended a Chinese complementary school in Italy that offers extracurricular classes alongside mainstream public education, also preparing learners for the HSK language certification (Hànyǔ Shuǐpíng Kǎoshì, 汉语水平考试). The participants, who were enrolled in Grades 7 and 8 of the Chinese school, had only recently begun studying literary texts, thereby minimising the risk of bias arising from prior knowledge of the vocabulary selected for testing.

Within the Chinese education system, literary education—including the reading and appreciation of classical and modern poetry, prose, fiction, and other culturally significant texts—is regarded as part of the broader development of linguistic competence, cultural awareness, and students' core competencies. This conception is also reflected in the terminology used within the nine-year compulsory education system. Literature is not taught as a separate subject but is integrated into Yuwen (语文), the national curriculum area conventionally rendered in English as “Chinese Language” or “Chinese Language and Literature”³.

Within this educational setting, a module on Chinese language and culture based on literary reading was piloted, adopting a lexical approach and a lexicographic methodology that frames dictionary compilation as an experiential and cooperative learning practice rather than as pure consultative activity (Campoy Cubillo, 2002; Caruso, 2024; Caruso & De Meo, 2025; Chaves-Yuste & Márquez-Cruz, 2022, 2024).

To alleviate the practical difficulty of establishing a control group for the Chinese heritage learners in Italy, and to assess whether the Heritage Group's vocabulary knowledge was comparable to that of age-matched students in China, the same framework was replicated with two groups of learners in China. One group followed the same activities implemented in Italy (Experimental Group, EG), while the other received more teacher-centred instruction based on textbook activities (Control Group, CG).

Results from a delayed post-test administered three months after the classes point to the efficacy of the protocol among Chinese learners, with the experimental group showing significantly greater gains over time. These findings encourage further investigation into the efficacy of the protocol also among heritage learners in Italy.

3 The learning framework

3.1 Collaborative lexicography for learning

Building on and adapting the protocol proposed in Caruso (2024), an experimental group of 22 Chinese heritage language learners (HG, mean age 12.7) read and analysed the short story *Village Opera* (社戏; *Shèxì*) by Lu Xun (1922), a canonical text in school curricula for the

³ The relevant national framework is the *Compulsory Education Curriculum Plan and Curriculum Standards (2022 Edition)*, issued by the Ministry of Education of the People's Republic of China and implemented from the autumn semester of 2022. See http://en.moe.gov.cn/news/press_releases/202205/t20220507_625532.html.

transmission of modern Chinese literary language and cultural identity (He, 2006; Curdt-Christiansen & Hancock, 2014). The instructional intervention pursued a dual objective: on the one hand, to improve learners' linguistic and cultural competences through the acquisition of literary vocabulary; on the other hand, to foster the development of effective learning strategies for vocabulary acquisition and Chinese character literacy (He, 2020).

The experimental intervention focused on the learning of 15 target words over a ten-hour instructional period. The course combined teacher-led activities devoted to guided reading and textual analysis with laboratory-based work designed to familiarize learners with the use of Lexonomy (Mechura, 2017; Jakubiček et al., 2018, 2022). Through lexicographic training and collaborative dictionary compilation, students were expected to develop greater awareness of structured lexical description as a means of facilitating vocabulary retention, while also benefiting from peer-supported use of the digital tool (Elia, 2009; Caruso & De Meo, 2025). Target words were selected based on their frequency in contemporary Chinese, as determined through the zhTenTen21 corpus (Kilgariff et al., 2014). Only the highest-frequency words were included among the target items for assessment, although learners were free to choose any lemma when compiling entries for their dictionary.

In line with Papert's constructivist pedagogy (Papert, 1980; Papert & Harel, 1991), widely adopted in educational technology, students were encouraged to explore Lu Xun's short story, *Village Opera* (社戏; *Shèxì*), autonomously and to select, working in small groups, the lemmas to be included in their dictionary. Papert's learning model, grounded in Piagetian learning theory, emphasizes the active involvement of learners in constructing knowledge through experiential activities and task-based learning. Within this framework, technology is effective when it supports learners' active participation and helps them create meaningful products. The collaborative compilation of digital dictionary entries was therefore expected to promote vocabulary growth and encourage learners to develop new ways of practicing unfamiliar words, such as searching for example sentences online and providing Italian translations.

3.2 Participants and Group Characteristics

Due to infrastructural constraints, the 22 heritage students could only use tablets to carry out the lexicographic activities, and some insisted on using their own smartphones instead of the devices provided by the Chinese school in Italy. This represented a significant limitation for the learning framework, since Lexonomy has been designed for larger PC screens, particularly with respect to dictionary compilation tasks. Only recently has the tool become more suitable for smartphone-based consultation and not for its compilation. A relevant finding of the pilot study involving the HG concerned the duration of the instructional intervention: the initially estimated ten hours proved to be excessive.

To address the difficulty of establishing a control group of Chinese heritage learners in Italy, and to assess the knowledge of the literary vocabulary targeted in the study among Chinese students of the same age and school grade, the same instructional framework was replicated with two groups of learners in China. Their results showed that students in this age group were also largely unfamiliar with the language of Lu Xun's short story.

To evaluate the effectiveness of the dictionary compilation protocol, a control group of 53 Chinese students (Control Group, CG) received traditional instruction based on teacher-led explanations and textbook activities. A second group of 50 students from the same school (Experimental Group, EG) followed the same protocol as the heritage learners in Italy but used computers to compile dictionary entries with Lexonomy. The instructional module

implemented with the Chinese students, in both the experimental and control groups, lasted half as long as the one conducted in Italy.

The control group did not engage in any lexicographic activities. Instead, students completed textbook exercises and received more detailed teacher explanations of the text. The textbook activities also had a lexical focus: students were asked to summarize episodes from the story using only four characters, and to reflect on the different narrative techniques employed by the author, including descriptions, the expressions of feelings, and his reflections.

To evaluate the efficacy of the experimental framework, all groups completed an initial cloze test on the 15 target words, which required them to insert the words into sentences drawn from Lu Xun's literary text (literary context) and from other usage examples extracted from the zhTenTen21 corpus (non-literary context). All tests were administered online via Google Forms. Except for the HG, the number of students completing the cloze post-test was lower than the original group size: 37 out of 50 in the experimental group and 50 out of 53 in the control group. Students in the HG and EG also completed an acceptance questionnaire designed to assess their satisfaction with the learning framework, the group work, and the Lexonomy dictionary editor.

TABLE 1 Participant group composition and activities undertaken

Student Type	School Year	Participants in the pre-test	Participants in the delayed post-test	Gender Distribution	Mean Age	Activities
Heritage students	Seventh and eighth school year	22	22	15 females, 7 males	12.7	10 meetings (500 minutes); collaborative dictionary compilation
Experimental Group	Seventh school year	50	37	24 females, 26 males	13.3	5 meetings (235 minutes); collaborative dictionary compilation
Control Group	Seventh school year	53	50	27 females, 26 males	13	5 meetings (235 minutes); predominantly teacher-led activities

3.3 *Learning objectives and Assessment of Students' Vocabulary Knowledge*

The teaching protocol implemented in the experimental and heritage group was designed as an interdisciplinary learning experience combining cultural and literary knowledge within a lexical approach supported by state-of-the-art lexicographic technologies. The project aimed to enhance students' lexical competence through guided exposure to Chinese literary vocabulary and the compilation of lexicographic entries including definitions, usage examples, collocations, and translations. Within this framework, vocabulary was conceived not merely as a set of linguistic units to be learned, but also as a means of accessing cultural and metalinguistic knowledge through reflection on the types of lexical information included in dictionary entries.

By way of example, the entries compiled by the **Heritage** and **Chinese** students (EG) for the same lemma—朦胧 (*ménglóng*, ‘hazy’, ‘dim’, or ‘blurred’)—can be compared, while the total number of entries compiled by each group is summarised in Table 2.

The number of students who contributed to the dictionary compilation differs from the number of students who participated in the tests. In the Heritage Group, 27 students took part in the instructional intervention and contributed to the collaborative dictionary, but only 22 completed both the pre-test and the delayed post-test. In the Experimental Group, all 50 students participated in the instructional intervention; however, only 47 entered their names when compiling the dictionary entries, and only 37 completed both the pre-test and the delayed post-test. Consequently, only the performance of these 37 students was included in the statistical analysis. The number of compiled dictionary entries is therefore reported on the basis of the students who actually contributed to the dictionary compilation.

TABLE 2 Composition of the groups and number of dictionary entries compiled

Group	Number of groups compiling entries	Average group size	Number of entries compiled	Average number of entries per group
Heritage	10	2.7	33	3.3
Experimental	14	3.6	29	2.1

From a cultural perspective, the protocol sought to introduce students to the culture of *Shèxì* (or ‘village opera’, 社戏)⁴ and, more broadly, to the heritage of traditional Chinese opera, with particular attention to Peking Opera (*Jīngjù* 京剧)⁵. Lu Xun’s short story *Village Opera* (社戏; *Shèxì*) describes the young protagonist’s experience of attending a traditional opera performance in rural southern China, where stages were set up along the riverbanks and spectators watched the performances from boats on the water. During classroom activities, students were introduced to this theatrical tradition through audiovisual materials that enabled them to explore its main character types, the symbolic significance of masks, and the relationship between theatre, collective ritual practices, and cultural memory.

In this context, the literary text and its lexicon served as valuable entry points into cultural traditions from which even younger generations in China are gradually becoming detached. The experimental framework, which combined vocabulary learning with literary study, therefore proved relevant not only for heritage learners but also for students in China. While the control group received the same culture-oriented explanations and was exposed to the same audiovisual materials, only the experimental group participated in the collaborative dictionary-compilation activities, following the same protocol as the heritage group in Italy.

Moreover, the integration of literature, culture, and vocabulary learning was supported by collaborative group work and the critical use of digital tools. These activities were designed to promote students’ active participation in the creation of a cultural product, such as a dictionary. Role rotation and peer review further encouraged cooperative, reflective, and meta-linguistic skills, helping learners take an active role in the co-construction of linguistic and cultural knowledge.

⁴ Shèxì (社戏, literally “community performance” or “community opera,” from shè ‘community’ and xì ‘theatrical performance’) refers to traditional operatic performances held during village festivals and communal celebrations in rural China.

⁵ Peking Opera is included in UNESCO’s Representative List of the Intangible Cultural Heritage of Humanity: <https://ich.unesco.org/en/RL/peking-opera-00418>.

Results from the pre-test questionnaire showed the unfamiliarity with the target literary vocabulary in all groups (19% correct responses for EG; 18% for CG; 13% for HG), not only for heritage learners. Familiarity scores were higher when the same words were provided in a non-literary context (or *general*, see Figure 1). Baseline differences among groups on the pretest were assessed using a mixed-effects logistic regression model (Table 3). The initial model included register, group, and their interaction, with random intercepts for question and word:

$$Pre \sim 1 + Register + Group + Register \times Group + (1|Question) + (1|Word)$$

A random intercept for participant ID was deliberately excluded, as the focus was on inter-group differences. Model selection identified the following reduced model as the best-fitting model:

$$Pre \sim 1 + Group + (1|Question) + (1|Word)$$

TABLE 3 Predicted probabilities of success estimated using a mixed-effects logistic regression model. Although Group significantly predicted pretest performance, the 95% confidence intervals around the predicted success probabilities overlapped across all groups, revealing no clear pairwise differences between specific groups

Group	Predicted Probability of success	95% CI
Control	0.11	[0.06, 0.21]
Experimental	0.11	[0.06, 0.21]
Heritage	0.07	[0.03, 0.14]

Group was retained as a predictor in the final model. However, confidence intervals for the predicted scores all overlap with predicted success rates, suggesting no differences between particular groups. The predicted probabilities of success were 0.11 for both the Control and Experimental groups and 0.07 for the Heritage group.

To assess medium-term lexical retention, groups were administered the same sentence-completion test used in the pre-test three months after the instructional intervention. On the delayed post-test, students obtained better scores, suggesting improvement relative to baseline performance (Figure 1).

4 Results

To assess the validity of the proposed learning framework, statistical analyses were conducted using mixed-effects logistic regression models. These models are used to analyse binary outcomes (e.g., success vs. failure) while simultaneously accounting for both fixed effects (experimental predictors of theoretical interest) and random effects (variability associated with participants, items, or tasks). This approach is particularly appropriate for repeated-measures linguistic and psycholinguistic data, where observations are not fully independent. Compared to classical inferential statistics (e.g., ANOVA or standard logistic regression), mixed-effects models offer the advantage of simultaneously modeling experimental effects and the variability associated with participants and linguistic items. This is particularly important in psycholinguistic and lexical research, where observations are not fully independent, since the same participants respond to multiple items and the same items are seen by multiple participants. By accounting for these sources of random variation, mixed-effects models provide more reliable and generalizable estimates of the effects of interest. To analyze the data, we began with a maximal model including four fixed predictors — Group, Register of context, Compiled, and Frequency — and all their interactions, alongside random intercepts for

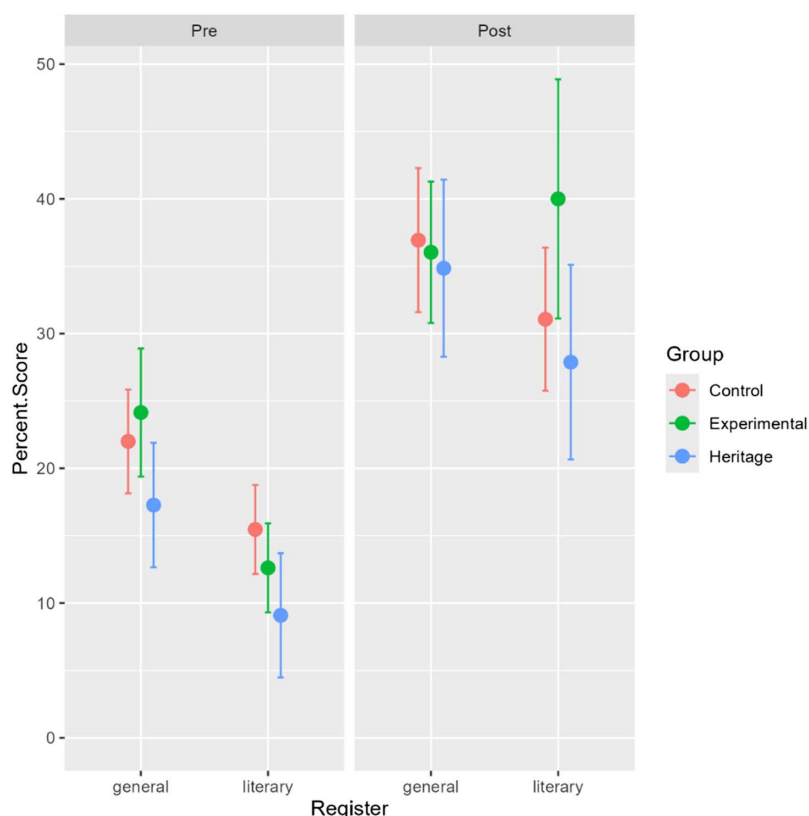


FIGURE 1 Percentage of correct item completion in the pre-test and delayed post-test in the heritage, experimental and control groups for both contexts: literary and non-literary

participants, questions, and words. The predictor *Compiled* indicated whether a given target word had been selected by students for dictionary compilation. The predictor *Frequency* captured differences in word frequency in the zhTenTen21 corpus. The 15 target words were classified as high-, medium-, or low-frequency based on their corpus frequency values. We then applied a principled model reduction procedure which progressively removes terms that do not contribute to model fit, arriving at the most parsimonious model supported by the data. The final model retained Group, Register, and their interaction as fixed effects, with random intercepts for participants ($N = 109$), questions ($N = 30$), and words ($N = 15$).

Importantly, most of the variability in our data came from these random effects — that is, from individual differences between participants, differences in how questions were phrased, and differences between the target words themselves — rather than from the experimental manipulation. This is a normal and expected finding in lexical research, and our model (as *mixed* or *hierarchical* models in general) accounts for all of this background noise before estimating the effects of interest.

Looking at the fixed effects, neither register nor group membership alone made a difference. What did matter was their combination: participants in the Experimental Group were roughly twice as likely to show improvement when the context was literary ($OR = 2.08$, $p < .001$), compared to the Control group in the same condition (see Figure 2). No such effect was found for the Heritage group, whose sensitivity to context register did not differ reliably from Controls. Predicted improvement was broadly comparable across groups in the general register (approximately 13–17%), whereas in the literary register the model predicted the highest improvement (about 27%) for the Experimental group, compared with the Control (17%) and Heritage (19%) groups (see Figure 3).

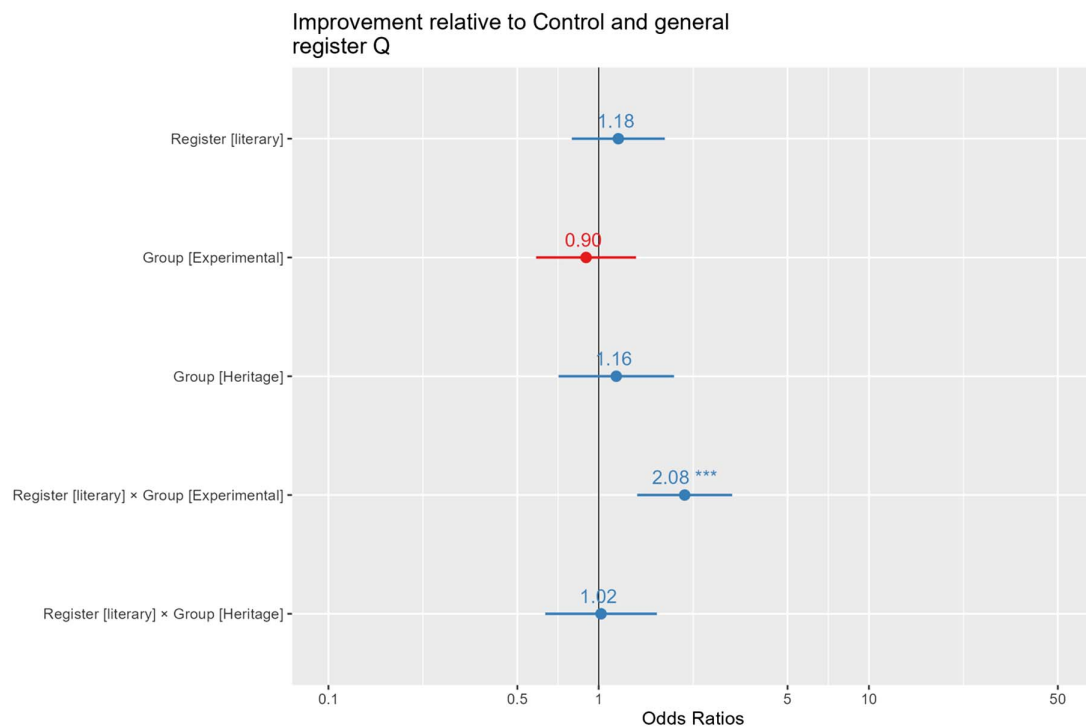


FIGURE 2 Forest plot of odds ratios from the mixed-effects logistic regression model predicting improvement relative to the Control group and the general register condition. Only the interaction between literary register and the Experimental group was statistically significant ($OR = 2.08$, $p < .001$), suggesting that participants in the Experimental condition were approximately twice as likely to improve when working with literary-register items

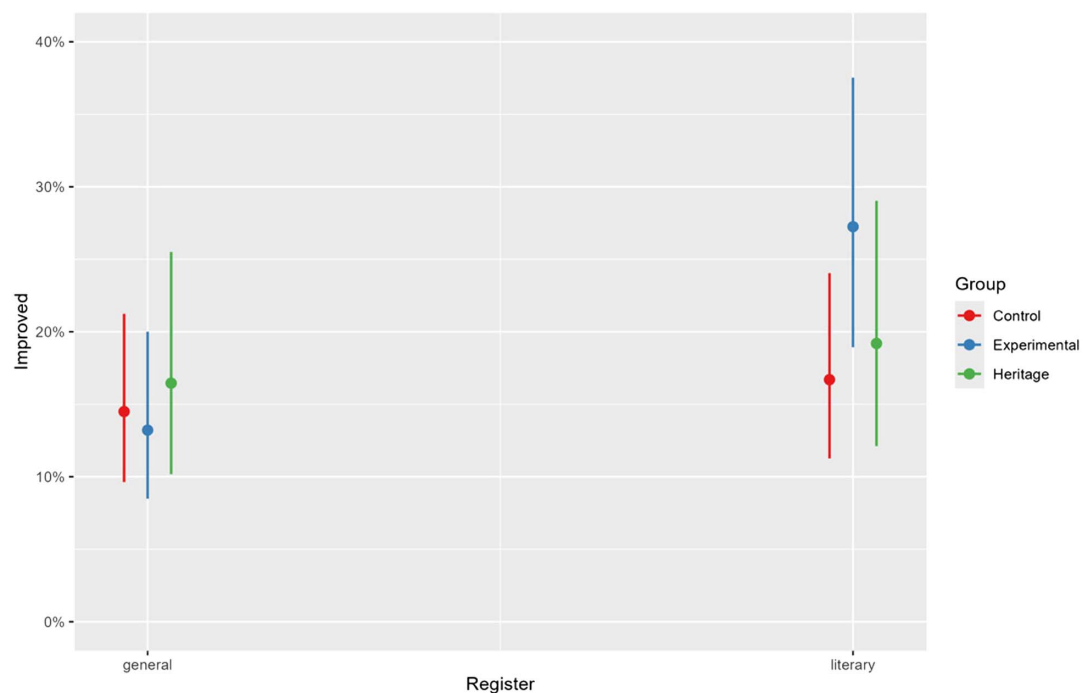


FIGURE 3 Improvement from pre-test to post-test across registers (general vs. literary) and participant groups (Control, Experimental, and Heritage). Error bars indicate 95% confidence intervals

5 Concluding remarks

Given the numerous limitations involved in assessing the effectiveness of the learning framework with heritage students in Italy, the initial research question concerning its potential to enhance L1 vocabulary development and strengthen cultural identity of Heritage Learners could not be conclusively answered on the basis of the present data. Two major factors have limited the effectiveness of the intervention.

The first concerns the lack of access to computers, which likely reduced students' engagement with the dictionary-compilation task. Participants could only use tablets, and some even preferred completing the activity on their smartphones, despite Lexonomy having been designed primarily for larger screens. The second limitation was the absence of a control group of Chinese heritage learners in Italy.

More conclusive evidence regarding the broader question of whether dictionary compilation is an effective strategy for learning new vocabulary was obtained by replicating the framework in China. The study compared the performance of an experimental group following the dictionary-compilation framework with that of a control group receiving traditional, culturally oriented instruction. Notably, students who participated in the dictionary-compilation activities (EG) showed substantial gains in vocabulary acquisition. Over the three months following the instructional intervention, their scores nearly doubled relative to the initial test, and they significantly outperformed the students who received traditional teacher-led instruction.

Another relevant finding relates to the vocabulary selected for testing, which proved to be largely unfamiliar to both groups in China (EG and CG) as well as to the heritage learners (HG), thereby supporting the appropriateness of the lexical selection.

References

- Brigadoi Cologna, D. (2023). Sviluppo, prospettive e problemi delle scuole ereditarie cinesi in Italia. *OrizzonteCina*, 14(1), 71–75. <https://doi.org/10.13135/2280-8035/8528>
- Campoy-Cubillo, M. C. (2002). Dictionary use and dictionary needs of ESP students: An experiential approach. *International Journal of Lexicography*, 15, 206–228.
- Carreira, M. (2004). Seeking explanatory adequacy: A dual approach to understanding the term “heritage language learner.” *Heritage Language Journal*, 2(1), 1–25.
- Carreira, M. & Kagan, O. (2011). The results of the National Heritage Language Survey: Implications for teaching, curriculum design, and professional development. *Foreign Language Annals*, 44(1), 40–64.
- Caruso, V. (2024). Between collocation and colligation: An experiment in collaborative lexicography. *International Journal of Lexicography*, 37(1), 75–94.
- Caruso, V. & De Meo, A. (2025). Tecnologie e apprendimento esperienziale-cooperativo: Dizionari fraseologici digitali per le lingue seconde. *Rassegna Italiana di Linguistica Applicata*, 57(1–2–3), 145–168.
- Ceccagno, A. (2003). New Chinese migrants in Italy. *International Migration*, 41(3), 187–213. <https://doi.org/10.1111/1468-2435.00246>
- Chaves-Yuste, B. & Márquez Cruz, M. (2022). Lexicografía digital y adquisición de léxico en inglés: Una experiencia b-learning en los tiempos de Covid-19. In B. Puebla Martínez, R. Vinader Segura & N. Navarro Sierra (Eds.), *Covid-19, multiverso de disciplinas. Una mirada desde la salud, la educación y la comunicación* (pp. 350–367). Dykinson.
- Chaves-Yuste, B. & Márquez Cruz, M. (2024). A cooperative digital dictionary to improve EGP reading comprehension in secondary education. In J. Bobkina & E. Domínguez Romero (Eds.), *Transformative trends in language education: Pioneering autonomous learning technologies* (pp. 171–188). Peter Lang.

- Confucius Institute (孔子学院). (2014). *HSK zhen ti ji = Official examination papers of HSK*. Gao.
- Cummins, J. (1979). Linguistic interdependence and the educational development of bilingual children. *Review of Educational Research*, 49(2), 222–251.
- Curdt-Christiansen, X. L. & Hancock, A. (Eds.). (2014). *Learning Chinese in diasporic communities: Many pathways to being Chinese*. John Benjamins.
- Elia, A. (2009). Can a collaborative wiki Weblish dictionary project help academic writing of ICT language learners. In I. González-Pueyo., C. Foz Gil., M. Jaime Siso & M. J. Luzón Marco (Eds.), *Teaching academic and professional English online* (pp. 153–180). Peter Lang.
- Guo, S. (2022). Reimagining Chinese diasporas in a transnational world: Toward a new research agenda. *Journal of Ethnic and Migration Studies*, 48(4), 847–872.
- Harel, I. & Papert, S. (Eds.). (1991). *Constructionism: Research reports and essays, 1985–1990*. Ablex Publishing Corporation.
- He, A. W. (2006). Toward an identity theory of the development of Chinese as a heritage language. *Heritage Language Journal*, 4(1), 1–28.
- He, A. W. (2010). The heart of heritage: Sociocultural dimensions of heritage language learning. *Annual Review of Applied Linguistics*, 30, 66–82.
- He, A. W. (2020). *Heritage language education: A research-based approach*. Routledge.
- Jakubiček, M., Měchura M., Kovář, V. & Rychly P. (2018). Practical post-editing lexicography with Lexonomy and Sketch Engine. In *Proceedings of the XVIII EURALEX International Congress* (pp. 431–440). Ljubljana University Press.
- Jakubiček, M., Kovář, V., Měchura, M. & Rambousek, A. (2022). Using NVH as a Backbone Format in the Lexonomy Dictionary Editor, in A. Horák, P. Rychlý & A. Rambousek (Eds.), *Proceedings of Recent Advances in Slavonic Natural Language Processing, RASLAN 2022* (pp. 55–61). Tribun EU.
- Kilgariff, A., Baisa, V., Bušta, J., Jakubiček, M., Kovář, V., Michelfeit, J., Rychlý, P. & Suchomel, V. (2014). The Sketch Engine: Ten years on. *Lexicography*, 1, 7–36.
- Leeman, J. (2015). Heritage language education and identity in the United States. *Annual Review of Applied Linguistics*, 35, 100–119.
- Liu, H. (2011). An emerging China and diasporic Chinese: Historicity, state, and international relations. *Journal of Contemporary China*, 20(72), 813–832.
- Maylor, U., Glass, K., Issa, T., Kuyok, K. A., Minty, S., Rose, A., Ross, A., Tanner, E., Finch, S., Low, N., Taylor, E., Tipping, S. & Purdon, S. (2010). *Impact of supplementary schools on pupils' attainment: An investigation into what factors contribute to educational improvements* (Research Report No. DCSF-RR210). Department for Children, Schools and Families. <https://dera.ioe.ac.uk/id/eprint/810/>
- Mechura, M. (2017). Introducing Lexonomy: An open-source dictionary writing and publishing system. Online. In I. Kosem, C. Tiberius, M. Jakubiček, J. Kallas, S. Krek & V. Baisa (Eds.), *Electronic lexicography in the 21st century: Proceedings of eLex 2017 conference* (pp. 662–679). Lexical Computing
- Papert, S. (1980). *Mindstorms: Children computers and powerful ideas*. Basic Books.
- Pellin, T. (2025). Le scuole complementari cinesi in Italia dopo la pandemia, fra grandi città, provincia e l'Europa. *OrizzonteCina*, 16(1), 4–15. <https://doi.org/10.82002/117661>
- Valdés, G. (2001). Heritage language students: Profiles and possibilities. In J. K. Peyton, D. A. Ranard & S. McGinnis (Eds.), *Heritage languages in America: Preserving a national resource* (pp. 37–77). Delta Systems.
- Thomas, W. P. & Collier, V. P. (2002). *A national study of school effectiveness for language minority students' long-term academic achievement*. Center for Research on Education, Diversity & Excellence, University of California, Santa Cruz.

Contact information

Zongyuan Wang
Università di Napoli L'Orientale
z.wang1@unior.it

Valeria Caruso
Università di Napoli L'Orientale
vcaruso@unior.it

Anna De Meo
Università di Napoli L'Orientale
ademeo@unior.it

Acknowledgements

This paper is the result of ongoing collaboration among the authors, who jointly endorse its methodology, findings, and claims. Responsibility for drafting the individual sections is distributed as follows: Sections 1, 3.2, 3.3, and 4 were written by Valeria Caruso; Sections 2 and 3.1 by Zongyuan Wang; and Section 5 by Anna De Meo.