RESLA
Revista Española de Lingüística Aplicada

https://matrix.aesla.org.es/RESLA
ISSN 0213-2028 | e-ISSN 2254-6774

Integrating Vocabulary Goals in University Level, Spanish L2 Content-based Courses: Flashcards for Learning Frequent Words

Nausica Marcos Miguel [1,2]* & Melissa Huerta [2]

[1]Ghent University, Ghent, Belgium
[2]Denison University, Granville, USA

Abstract: Integrating language goals in upper-level content courses in language majors at U.S. universities has been encouraged in the literature. Nevertheless, few studies have examined students' perceptions and gains when language learning goals are integrated. This action research study explores students' perceptions and their actual gains when using flashcards in two content courses to learn the 4K and 5K bands of most frequent Spanish words. These results are compared with the same and other students learning the 2K and 3K bands of most frequent Spanish words while enrolled in a previous language-focused course in the major. Both groups reacted similarly to the flashcard activities, even though the students in the language course found them better integrated with their class goals. Learning of frequent words was higher in the language course than in the content courses, as measured by the *ITT-Receptive Vocabulary Test*. Pedagogical recommendations are given about how to better integrate vocabulary learning goals in content-courses for language majors.

Keywords: *Spanish L2, literature courses, higher education, world language majors, vocabulary development.*

Un análisis de la integración de objetivos léxicos en cursos de contenido en español L2 a través de tarjetas digitales para aprender vocabulario frecuente

Resumen: Aunque se recomienda integrar la enseñanza de vocabulario en los objetivos curriculares de las clases de contenido impartidas en español L2 en universidades estadounidenses, pocos estudios han examinado el aprendizaje y las percepciones de los estudiantes cuando se sigue esta recomendación. Este estudio de investigación-acción explora las percepciones y el aprendizaje en dos cursos de contenido donde los estudiantes tenían que aprender la banda de vocabulario de las cuatro mil palabras más frecuentes y la banda de las cinco mil con tarjetas de vocabulario. Los resultados se compararon con el aprendizaje de las palabras de las bandas de dos mil y tres mil en un curso de lengua previo. En los dos tipos de cursos, los estudiantes percibieron de forma similar la actividad, aunque los estudiantes del curso de lengua la consideraron mejor integrada con los objetivos curriculares. El *ITT-Receptive Vocabulary Test* mostró mejores resultados en el curso de lengua que en los cursos de contenido. Se dan recomendaciones pedagógicas sobre cómo integrar objetivos léxicos en este contexto.

Palabras clave: *ELE, cursos de literatura, educación universitaria, programas de lengua, desarrollo de vocabulario.*

How to cite: Marcos Miguel, N., & Huerta, M. (2026). Integrating Vocabulary Goals in University Level, Spanish L2 Content-based Courses: Flashcards for Learning Frequent Words. *Revista Española de Lingüística Aplicada, 39*(1), 184-211. https://doi.org/10.58859/resla.868

*Corresponding author: nausica.marcosmiguel@ugent.be

AESLA
Asociación Española de Lingüística Aplicada

1. Introduction

Content-based language instruction is a popular approach to work on language development (see Ruiz de Zarobe & Cenoz, 2015). In the case of US universities, content-based instruction is the backbone of language majors, as students in the second and third-year are exposed to language learning through content in courses focusing on history, literature, linguistics, and culture studies. In this context, content learning is the main curricular goal, and language development is considered a by-product of content learning. Therefore, grammatical or lexical goals are not typically set in these courses (see Donato & Brooks, 2004; Marcos Miguel, 2022; Polio & Zyzik, 2009; Thoms, 2014).

However, in the last decades, there has been a call for better integrating the teaching of content and language to strengthen both areas (e.g., Maxim, 2014; Paesani, 2017). For example, to comprehend and produce texts, L2 users need to have both a large vocabulary breadth and depth (see Qian & Lin, 2020). Without knowledge of the most frequent words (1k to 5k), readers will have a hard time understanding texts. Although each text studied in a class will have specific vocabulary related to its genre and content, the most frequent vocabulary will be constant across texts. Therefore, since L2 learners need more than five years to master the 3,000 most frequent words (e.g., Coxhead & Boutorwick, 2018; Webb & Chang, 2012; Zhang & Lu, 2014), a certain degree of explicit instruction and planning is necessary. Whereas instructors can not explicitly teach all these words because of time constraints, they can point out resources for learners to acquire them.

This study proposes and analyses an explicit vocabulary learning intervention in two content courses, an introduction to literary analysis and an introductory culture course, taught at a Liberal Arts College (LAC)[1] in the United States. We explore how students perceived an explicit vocabulary learning goal (i.e., learning the 4k most frequent words in the literature course and the 5k most frequent words in the culture course) in these courses. We compare their perceptions and learning with the perceptions and learning of the same and other students in a previously taken language course focusing on the 2k and 3k most frequent words. The vocabulary learning goal serves two functions: 1) increasing the vocabulary breadth of the students, and 2) making them aware of the importance of this goal.

This study employs an action-based approach (see Burns, 2005) by documenting the integration of the language-learning goal through several iterations of the courses. Following this approach, the teaching practice always has priority over research methodology (Banegas & Villacañas de Castro, 2019). Because of the prioritization of the educational practice, this article can provide insights for researchers and practitioners. Especially, the study seeks to motivate discussion of how to integrate language goals for practitioners working in similar contexts.

[1] LACs are four-year, BAs granting institution that are characterized by small class sizes and faculty with strong interest in teaching.

In the following section, the instructional context will be described in detail, comparing it to other contexts with different degrees of language and content integration. Next, the importance of systematic vocabulary instruction for language development will be discussed.

1.1 World language major programs in the United States

World language major programs are two-tier programs: the first two years are devoted to language learning, while the last two years focus on content learning (Geisler et al., 2007; Lomicka & Lord, 2018). At approximately 50% of the universities in the United States, all students need to complete two or three semesters of language study unless this requirement is satisfied with a placement test showing prior knowledge of a language other than English (Lusin, 2012). Most students who continue with the language program beyond the required two or three semesters become majors or minors. Majors typically finish the program with a proficiency level of intermediate-mid or high in speaking, advanced low in reading, and intermediate high in listening on the ACTFL proficiency scale (Winke et al., 2020). As minors typically have less exposure to the language than majors, their proficiency levels are likely to be lower.

To facilitate comparison with research conducted in other instructional contexts combining L2 development and content learning, Macaro's (2018) classification system is helpful. Macaro identifies the following contexts: immersion, Content- and Language Integrated Learning (CLIL) (in Europe), Content-based Instruction (CBI/CBLT), and English Medium of Instruction (EMI). Content courses in a typical Spanish major in the United States are close to the EMI label, except for the fact that Spanish is the language medium of instruction instead of English. Features shared between EMI and the content courses from which data for this study were drawn are as follows:

1. "the L2 is not spoken locally" (i.e., in the LAC where the study took place),
2. "the majority of students are nationals of, or reside in the country,"
3. "teachers are bilinguals,"
4. "teacher is always a content specialist,"
5. "the L2 is [Spanish] only,"
6. There can be variability by institution and country (e.g., Spanish majors in the UK, Germany, etc.),
7. "few officially produced policies,"
8. "almost entirely content-driven," and
9. "program objectives in terms of L2 proficiency [are] rarely stated" (Macaro, 2018, pp. 39–40).

There are some differences when comparing EMI with content-based courses in Spanish programs. For Spanish majors, program objectives include cultural understanding as well as social objectives, which does not seem to be the case in EMI as Macaro (2018) describes it. Contrary to Macaro's description of EMI, teachers in content-based courses for a Spanish major received training in L2 and content teaching, and teaching materials can consist of both authentic materials, such as short stories written for native speakers, and pedagogical materials developed for L2 learners. Interestingly, in the case of Spanish in the United States, depending on the area where the university is placed, Spanish can be locally spoken. Moreover, Spanish heritage speakers, i.e., students who speak the language at home but have received little or no formal education in Spanish, are a frequent group among the Spanish majors.

Previous studies have shown that Spanish majors are not exposed to advanced language patterns in their content courses. This limited input suggests that learners cannot develop advanced proficiency (Darhower, 2014; Donato & Brooks, 2004). Simultaneously, content learning can be reduced if the learner has low proficiency (Roussel et al., 2017; Yang & Farley, 2019). Although little empirical research has been conducted on language development in content courses, some studies have explored vocabulary teaching and learning. Regardless of the instructional focus, students believe there is an increase in their vocabulary knowledge (see Polio & Zyzik, 2009). Students also frequently ask for vocabulary clarifications for both high-frequency and specialized words related to the subject matter (see Marcos Miguel, 2022).

In these courses, learners are exposed to content through lectures and through authentic readings and audiovisual materials, which can contribute to incidental vocabulary learning (see Webb, 2020b). By providing additional support (e.g., including glossaries, captions, or subtitles) and explicit instruction, educators can help learners overcome the challenges of reading (see Vander Beken et al., 2020) and listening to audiovisual materials (see Montero Pérez, 2022). In brief, whereas these courses provide plenty of input for incidental vocabulary learning, explicit learning through a more thoughtful and goal-oriented approach guided by the instructor should also take place. The next section will delve into current recommendations for vocabulary instruction.

1.2 Recommendations for vocabulary instruction

Vocabulary researchers agree that explicit instruction is necessary for successful vocabulary acquisition in L2 teaching contexts (see Newton, 2019). Needs analysis and systematic instruction are fundamental for developing vocabulary knowledge in any language program (Barcroft, 2012; Nation, 2001). Furthermore, teachers and learners also need to find the instruction appealing (Lindstromberg, 2020).

The acquisition of high-frequency vocabulary allows for understanding more words in a text, as frequent words tend to offer greater lexical coverage than lower-frequency ones. For example, in Spanish, the 3,000 most frequent words can cover 90% of the words in any given text (Davies, 2005a). In English, knowing the 5,000 most frequent words allows

readers to understand 98% of the words in a text (see Barclay & Schmitt, 2019). That is, the more words a student knows, the better their reading comprehension can be. For content-based courses where reading is a fundamental part, larger vocabulary sizes will facilitate students' comprehension.

However, students are not explicitly taught frequent vocabulary in the language courses. Lexical analyses of Spanish textbooks have revealed that beginner and intermediate textbooks do not select their target vocabulary based on general frequencies (see Davies & Face, 2006; López Bastidas & Sánchez-Gutiérrez, 2020). Even though instructors might be aware of the importance of teaching frequent vocabulary, they do not tend to prioritize it (see Sánchez-Gutiérrez et al., 2022).

In an analysis of the Spanish vocabulary development of students at a large U.S. university who did not receive explicit instruction on high-frequency words, learners in the fourth year showed knowledge of around 2,000 of the 3,000 most frequent words. Learners in this year acquired vocabulary slower than their counterparts in the first two years of the major when instruction focused on language development rather than content (Robles-García, 2022). This speed reduction can be attributed both to the lack of explicit vocabulary instruction in content courses and to the increasingly specialized nature of the vocabulary in such courses.

Clearly, the nature of word frequencies influences learning (see Marcos Miguel et al., 2023; Robles-García, 2022). As the words decrease in their frequency, learners will be less exposed to them, even if they are still relatively frequent. Since the recommended learning target is the 5,000 most frequent words, instructors need to be aware that learners will be more exposed to the first frequency bands than to the following bands. For example, in a previous study with third-semester students (Marcos Miguel et al, 2023), learners in the same institution recognized over 80% of the 1,000 most frequent words, but slightly over 50% of the 2,001–3,000 most frequent words.

Because of limited classroom time, using flashcards (both paper-based and digital) is a good strategy for learning individual words outside of the classroom (see Dizon & Tang, 2017; Yüksel et al., 2020). Flashcards facilitate the creation of links between form and meaning while increasing vocabulary size (see Barclay & Schmitt, 2019; Nation & Webb, 2011; Lei & Reynolds, 2022). However, flashcard work may be less effective if students do not use them for formal or informal self-testing (see Nation & Webb, 2011).

Given that sustainable teaching practices must be easy for instructors to implement (see Kennedy & Kennedy, 1996), digital flashcards, which are easier to create and distribute than paper-based flashcards, can be a helpful tool for instructors who teach content-based courses. In this study, the sets were created by one of the instructors who then shared them with their colleagues.

2. Research Questions

Spanish L2 majors in the United States do not master the 5,000 most frequent words in Spanish (Blake, 2020; Robles-García, 2022). Therefore, this study compares students learning vocabulary from two bands of Spanish high-frequency vocabulary (2k and 3k) in a writing-intensive language course, which was the last course of the language-development sequence, with students learning additional bands of Spanish high-frequency vocabulary in the following two content courses. These two courses were consecutively taken: an introduction to literary analysis (learning the 4K band) and an introduction to culture (learning the 5k band) (see section 3.2 for a more detailed description of the courses and the motivation for the teaching approach).

To examine whether this intervention was successful, the following questions are explored:

1. What are students' perceptions regarding the use of flashcards for explicit vocabulary study of high-frequency words in content courses?
    a. And in the last language course?
    b. Are there differences and similarities between both groups (content versus language)?
2. Using longitudinal data, to what extent do 1) students in content courses and 2) students in a language course make gains in high-frequency vocabulary by using flashcards?
3. What is the relationship between students' perceptions of high-frequency vocabulary learning with flashcards and their actual vocabulary knowledge in content courses?
    c. And in the last language course?
    d. Are there differences and similarities between both groups?

3. Methodology

3.1 Procedure following the action-based research paradigm

This pedagogical practice was integrated into a Spanish major at a LAC. Two instructors-researchers (A and C) collaborated in the data collection and analysis efforts, whereas four additional instructors shared data from their classes and provided feedback (see Tables 1 and 2). Professors A and C are the authors of this study. While professor A researches L2 acquisition and language instruction, professor C's research focuses on cultural studies. At the time of the study, each had over twelve years of experience teaching language courses. However, only professor C taught the introductory literary analysis class and the culture class. One of the strengths of this action-based research study is the collaboration between a language researcher and a cultural studies researcher.

Following the action-based research paradigm (see Burns, 2005; Burns & Khalifa, 2017), each study should include four phases—initiation, planning, implementation, and evaluation—and have several iterations (see Figure 1, Table 1 & Table 2). Although qualitative data is the gold standard for action-based research (see Banegas & Villacañas de Castro, 2019), quantitative data was also collected, as quantitative methods are the favored approach in vocabulary research (e.g., Webb, 2020a). The qualitative and quantitative data-collection instruments—a questionnaire and the *ITT-Receptive Vocabulary Test* (ITT-Test) for Spanish (see Tschirner, 2021)—and their administration will be described in section 3.4.

Figure 1 shows the semesters in which the questionnaire and the *ITT- Test* were given to the students. The first time the vocabulary test was given is marked in blue. The red text indicates the courses in which the data was collected, i.e., language, literature, or culture.

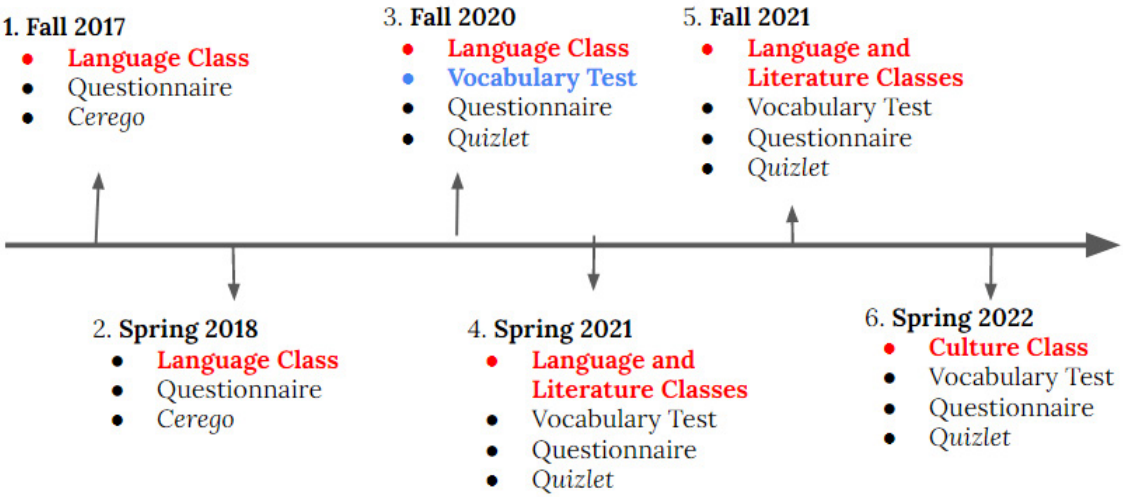

Figure 1. Timeline of analyzed data

During the Fall 2017 and Fall 2018 semesters, only the language courses participated in the study by completing the end-of-the-semester questionnaire. No additional data was collected from these students that semester or in the following semesters. This gap in data collection comes from the fact that professor A, who was the originator of the instructional plan, was not assigned to teach the writing language course during the two following academic years (i.e., 2018–2019; 2019–2020).[2] Professor A and C co-coordinated a fourth-semester conversation course in the major during Spring 2020 which started their common reflection on vocabulary development and contributed to their planning for the intervention during the 2020–2021 academic year.

Starting in Fall 2020, the *ITT-Test* was also given to the students in the language course at the beginning and the end of the semester. Since Spring 2021, data was collected in the content courses. Students who took the language course in Fall 2020 and after could

[2] Professor A was on leave for the first of those two years.

participate in the study through more than one course since the language course was a pre-requisite for the content courses. Therefore, it was possible to collect data longitudinally through the three courses.

3.2. Description of the classroom curriculum and the flashcard activities (pedagogical intervention)

Before describing the flashcard activities in detail, the content, and goals of each of the three courses will be presented. Although each instructor could revise the syllabus of their section as they saw fit (see Table 1 and Table 2 for a list of instructors), the materials and activities were identical or similar across sections of the same course. In all courses, students were exposed to a variety of written and audio-visual materials that could facilitate incidental vocabulary learning.

First, the language course focused on developing students' writing abilities by working on several essays during the semester. As this class fulfilled a writing requirement set by the university, 50% or more of the course grade was determined by these writing activities. Process writing was the favored teaching approach, i.e., including planning, drafting, editing, peer-review, and instructor's feedback at several points in the process (see Seow, 2002). In addition, the class focused on developing students' grammar competence and included explicit grammar instruction. The textbook used was *Gramática para la composición* by Stanley and González (2015). The course also included listening, reading, and speaking activities, but to a lesser extent.

Second, the literature course was an introduction to literary analysis. Students read several short stories and other texts to illustrate different periods and genres of literature written in Spanish. For the first time in the language major, students were asked to read a full-length novel in Spanish, such as *Como agua para chocolate* by Laura Esquivel (1995). This course also fulfilled the university writing requirement. In addition to completing writing assignments, students presented and participated in class discussions in Spanish.

Third, the culture course was an introduction to the culture of the Spanish-speaking world. This course was structured much like the literature course. For example, students read several literary and academic texts to understand cultural and historical moments. Students also presented and discussed texts. Like the other two courses, this course fulfilled the university-wide writing requirement.

The need to initiate the vocabulary intervention was determined by an analysis of the vocabulary gains of lower-level students who took an instructor-designed test to show their knowledge of the 3,000 most frequent Spanish words (Marcos Miguel et al, 2023). Like the findings of Robles-García (2022), the results indicated that learners receptively acquired around 60% of those words after three semesters of study. However, while students could easily recognize the first 1,000 most frequent words (with an average of around 90%), the 2,000 most frequent words were slightly less recognized (around 80%), and the 3,000 most frequent words were less known (slightly over 50%). Given these findings, the training intervention started on the 2k band in the language course.

The format of the vocabulary activities was kept as similar as possible across iterations. Instructors A, B, C, D, and F (see Table 1 and Table 2 for a timeline of the courses they taught) included the flashcard activity in their syllabus and/or calendar. Apart from studying the flashcards on their own, the students completed in-class vocabulary tests (before the Fall of 2020) and *Quizlet* tests (from the Fall of 2020 on). These tests had a practice and accountability function for learning vocabulary and are not part of the data reported in this study. The frequency bands were taken from Davies (2005b).

During the first two iterations (see Figure 1), the online flashcard program *Cerego* was used. In these two iterations, the bands were divided into 18 sets, i.e., from the 1,001 word up to the 2,800th most frequent word. Each set included a selection of 11 to 22 words taken from each group of 100 consecutive words from each frequency band. The words were selected based on the instructor's perceptions of what their students did not know.[3] Students were assigned one or two sets to learn on their own every week and took a vocabulary test in class every two weeks. This vocabulary test was graded by the instructor and was part of the students' grades.

Since the COVID-19 pandemic influenced the design of the program, the testing format and the sets varied in the following iterations. First, in Fall 2020, the program was changed to *Quizlet* given that it was free for the students. Second, as some students were attending class remotely and the institution called for a reduction in paper-based assignments, students used the "test" function in *Quizlet* to complete one weekly test and posted their scores in the course Learning Management System (LMS). This retrieval practice was fundamental to student's learning (Karpicke & Roediger, 2008). Students could repeat the tests as many times as they wished, and their scores usually ranged between 90% and 100%.[4]

Given that the testing using *Quizlet* was less demanding, each vocabulary set contained a larger number of words (100) compared to those in earlier iterations of the course. That is, all words from the 2k and 3k frequency bands were targeted. Every week, students were assigned two sets with a total of 200 words. The sets were progressively studied, i.e., starting from the most frequent to the less frequent words. In addition, students curated one set on their own and shared it with their classmates. This new activity aimed at giving students ownership of their learning by allowing them to choose the vocabulary they thought relevant.

Next, as the use of flashcards seemed to be well received by students in the language course, we implemented the vocabulary activities in the content courses (see Figure 1). In the literature course, we introduced 10 sets comprised of 100 words from the 4k-word

[3] Although instructors do not always predict with full accuracy the vocabulary their students know and do not know, Spanish L2, university instructors can show strong correlations with students' knowledge (Sánchez-Gutiérrez et al., 2023). If the instructors know the course, they might be able to predict students' knowledge better than frequency ranges (low-, mid- and high-frequency ranges) (Robles-García et al., 2023).

[4] Students in the Fall 2021 course completed the vocabulary tests in class with the sets of 200 words. This format was professor B's choice.

frequency band. Similarly, in the culture course, we introduced 10 sets of 100 words from the 5k-frequency band. The students worked on their sets independently and submitted their test scores weekly.

Figure 2 is a screenshot of the sets in *Quizlet* to illustrate how students accessed them. Each set was assigned during one week of the semester for a total of ten weeks. As all courses lasted 15 weeks, the students started working with the flashcards in the second or third week of the semester.

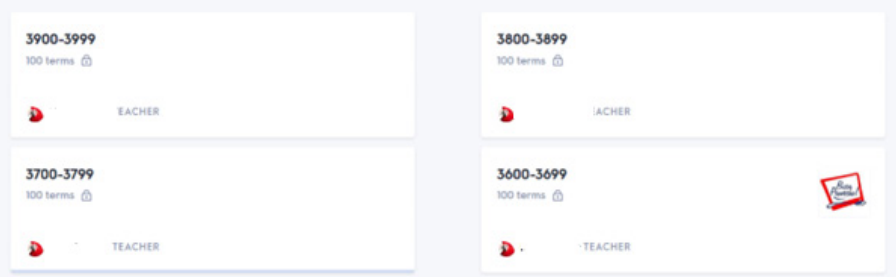

Figure 2. Screenshot of the vocabulary sets for the literature course.

Figure 3 shows a screenshot of a set with a flashcard and the different practice activities, including the "test" option. Once a student clicked on the flashcard, they could read the translation of the word into English, i.e., *pañuelo – handkerchief*.

Figure 3. Screenshot of a flashcard for the literature course.

3.3 Participants in the research study

Table 1 and Table 2 offer a summary of the participating classes. Each class enrolled a maximum of 16 students. As not all students completed each data-collection instrument or decided not to share their results with the researchers, the third column indicates the

number of participants who completed each instrument in each class. Because of logistics issues, students in the Spring 2021 language course did not complete the questionnaire, while students in the Fall 2021 language course and the culture course did not complete the *ITT-Test*.

Table 1. Data reported in this study from the participating language sections.

Course	Semester	Instruments and Number of Participants	Instructor
Language Course 1,000-2,999 (Selection, in class tests)	Fall 2017 (2 sections)	Questionnaire* (n=30)	Professor A (researcher)
	Spring 2018	Questionnaire (n=12)	Professor A (researcher)
Language Course 1,000-2,999 (All words, online tests)	Fall 2020 (2 sections)	Questionnaire (n=9) Vocabulary Test (n=13)	Professor A (researcher)
	Spring 2021	Vocabulary Test (n=9)	Professor A (researcher)
	Fall 2021	Questionnaire (n=9)	Professor B

*The questionnaire data analyzed in the language courses consisted of one open-ended question.

Table 2. Data reported in this study from the participating content sections .

Course	Semester	Instruments and Number of Participants	Instructor
Literature Course (3,000-3,999)	Spring 2021	Questionnaire* (n=16) Vocabulary Test (n=15)	Professor C (researcher)
	Fall 2021 (2 sections)	Questionnaire (n=18) Vocabulary Test (n=9)	Professor D
Culture Course (4,000-4,999)	Fall 2021	Questionnaire (n=3) **	Professor E
	Spring 2022 (2 sections)	Questionnaire (n=3) Vocabulary Test (n=14)	Professor F

*The questionnaire data analyzed in the content courses consisted of one open-ended question and several closed questions (see Figure 5).
**Two of the three students also took the Literature Course at the same time.

The language course was the last course in the language sequence and the first course that fulfilled credits for the major. Students in this course had already completed the general language requirement of the university. However, for some students, this class was still a requirement, as certain university majors required students to complete two additional semesters of language beyond the university-wide language requirement. Students in the content courses were interested in pursuing a major or minor in Spanish and/or wanted to fulfill the writing requirement. These two content courses were the first courses in the major to focus on content rather than language.

Students in the language course completed a proficiency assessment at the end of the semester. This test, STAMP[5], measures the four skills (reading, speaking, writing, and listening), and provides a composite score of the learners' proficiency according to the ACTFL proficiency scale. The average proficiency level of the 22 students who completed the *ITT- Test* in Fall 2020 and Spring 2021 was 6.6 (SD=.88), which means that most students at the time of the study had achieved the intermediate-high level in the ACTFL proficiency scale.

To complement this proficiency information, Table 3 reports 1) the number of years participants had studied Spanish—both at the university and in other settings—before enrolling in the language course, and 2) the percentage of students who claimed that they spoke Spanish outside of class.

Table 3. Previous language experiences.

	Language Courses (N=51)
Years of Spanish before the University (% of students)	1 to 2 years = 8% 2 to 3 years = 4% 3 to 4 years = 8% 4 to 5 years = 31% Over 5 years = 49%
Spanish semester at the LAC (% of students)	1st semester = 14% 2nd semester = 16% 3rd semester = 43% 4th semester = 22% 5th semester = 6%
Students who spoke Spanish outside of class	49%

3.4 Instruments and analysis

3.4.1 Questionnaire

Participants in each course completed an anonymous[6] questionnaire at the end of the semester to report their perceptions of vocabulary learning. In addition to the demographic information in Table 3, the questionnaire included closed questions measured with a five-point Likert-Scale (see Figure 5) for the content courses as well as an open-ended question for all courses;[7] namely *What is your opinion about the implementation of (the online flashcard system)? (e.g., what has worked well, what needs to improve?).*

For the closed questions, the data was averaged to show tendencies. For the open-ended question, the two instructor-researchers read participants' responses and searched for recurring topics. After this initial common reading, the instructor-researchers

[5] https://avantassessment.com/stamp

[6] The anonymity is important as the students were in contact with the researchers because they were either their professors or colleagues of their professors.

[7] The questionnaire in the language course included some additional questions about other class activities, which are beyond the scope of this paper.

individually revised their codes while they performed an initial coding of the literature course. Next, they met again to discuss and review the codes. Individually, the instructor-researchers coded all data for the literature sections and met again to discuss their coding. For cases in which the codes did not exhibit a 90% agreement, the instructor-researchers discussed participants' responses until the coding reached an agreement of 90% or higher. The same process to reach agreement was repeated for the language-course data and the culture-course data.

3.4.2 ITT-Receptive Vocabulary Test

The vocabulary test used in this study is the *ITT-Test* for Spanish. According to Tschirner (2021), the test format follows the English Vocabulary Level Test (Nation, 1990) and uses the Spanish frequency list created by Davies and Davies (2017). The *ITT-Test* measures the five bands of the 5,000 most frequent words. This test has been measured for reliability and internal validity and has shown a Cronbach's Alpha over .9 (Tschirner, 2021).

This test is available online[8] and must be completed within 30 minutes. Figure 4 shows a screenshot of the section testing the first 1k words. Testers should choose the word that matches best the definition. Definitions with the same number (i.e., 1a, 1b, 1c) have the same six words to choose from. For every band of thousand words, the maximum score is 30, i.e., one point for each correct answer.

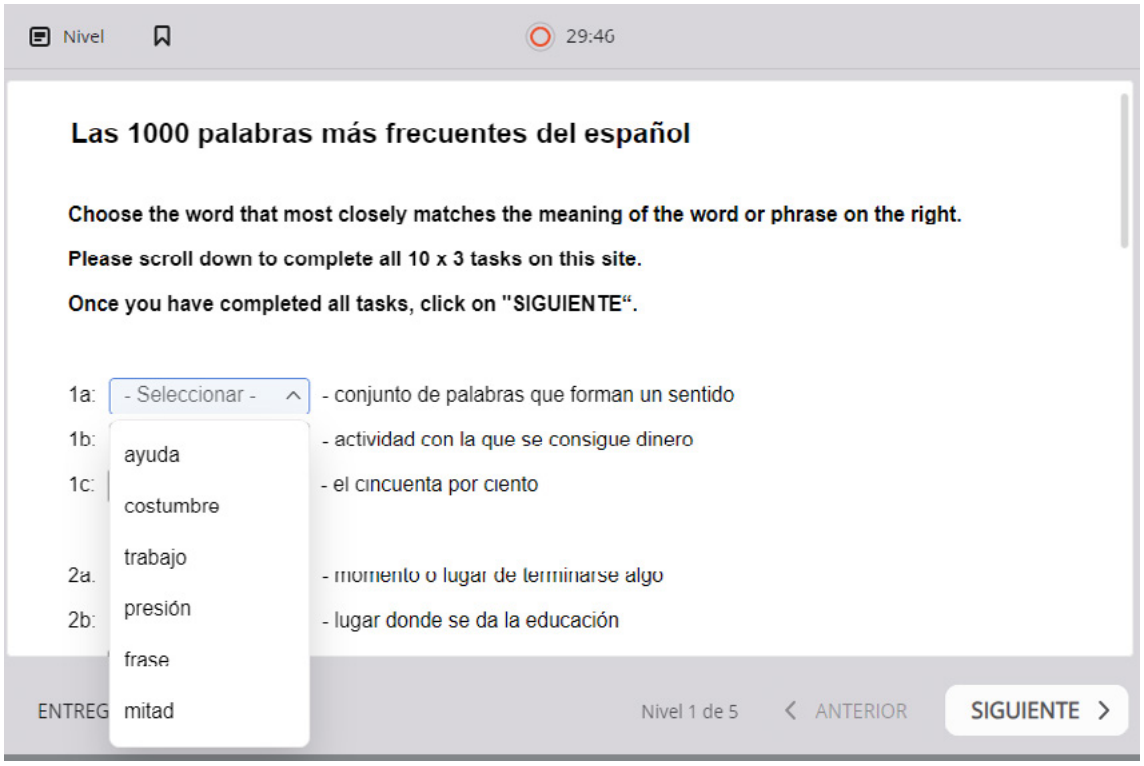

Figure 4. Screenshot of test items in the ITT-Receptive Vocabulary Test.

[8] https://itt-leipzig.de/about-the-vocabulary-tests-2-2/?lang=en

Once the test is completed, participants see a report of their scores by frequency band as well as a total score. For this study, students were asked to submit a PDF version or screenshot of the report through their course LMS. The cumulative score of the test for all 5,000 words was used as a proxy for vocabulary learning.

In the language course (see Table 1), participants completed the test both at the beginning and at the end of the semester, while participants enrolled in the content courses completed the test only at the end of the semester. That is, data from some of these students was collected longitudinally. Thus, three paired t-tests were performed to compare the longitudinal scores: (1) learning in the language course, as measured by comparing pre- and post-test scores, (2) learning between the two content courses, as measured by comparing the test scores at the end of each content course, and (3) learning between the language course and a content course, as measured by comparing the post-test in the language course and the last test taken in either one of the content courses.

4. Results

4.1 Students' perceptions: content courses

Figure 5 shows the average scores (n=40) in response to the closed questions, which were only included in the questionnaire for the content courses. Responses to the first three questions are consistent with one another. Students showed agreement with the fact that they learned more vocabulary than in other semesters and disagreement with the fact that they learned less than in other semesters. Similarly, students seemed to agree that they learned the target words, but they were slightly more neutral with the statement that they saw the target words in their readings.

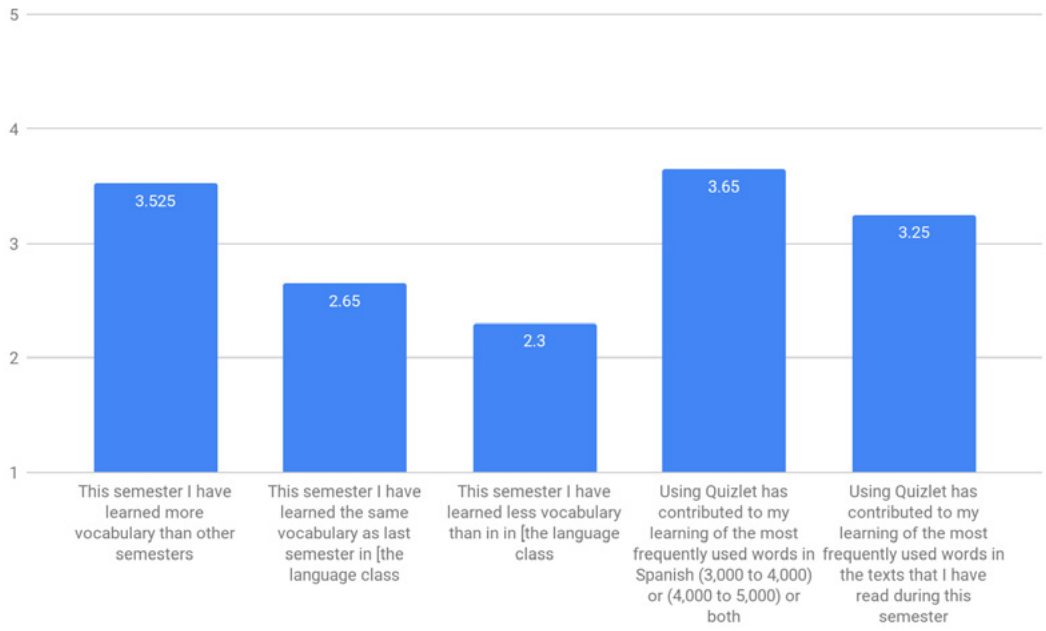

Figure 5. Closed-Questions in the Questionnaire, Likert Scale from 1 (strongly disagree) to 5 (strongly agree).

4.2 Students' perceptions: Comparing content and language courses

In Table 4, the codes used to categorize responses to the open-ended question are shown. The number of responses differs from the number of responses in Figure 5, as not all students

Table 4. Students' Perceptions in Content and Language Courses.

		Content Courses (N=33)	Language Course (N=60)	Comment from the Content Course *(Examples)*
Positive codes	Help in learning (e.g., explicit comment saying that)	55% (18)	82% (49)	Quizlet has worked well because it provides different ways for students to practice and memorize the words (Comment 1, Student 4).
	They like that there is a variety of activities with the flashcards	9% (3)	6% (4)	Quizlet is a great tool for practicing vocabulary because there are so many different activities. This allows people to find what type of learning is best for their retention (Comment 2, Student 9).
Negative codes	Not helpful for learning	18% (6)	2% (12)	[...] I wasn't incentivized to actually learn these words by heart. I think that there could have been more incorporation of the quizlet frecuencias during class to make it actually somewhat important that we learn the words (Comment 3, Student 11).
	Issues with the platform and the format of the activity (e.g., pictures instead of words, glitches)	24% (8)	48% (29)	The opportunity to learn more vocabulary is excellent, I think that having assignments to submit the Quizlets in [the LMS] would be helpful to stay on top of them (Comment 4, Student 18).
	Not appropriate number of words (e.g., too many words)	9% (3)	7% (4)	I think that Quizlet is a good way to test knowledge but the number of words in each set can be overwhelming and hard to carve out time to study them fully (Comment 5, Student 26).
	Difficulties with memorization	21% (7)	18% (11)	I think Quizlet is a good way to learn the vocabulary, especially because it introduces us to a lot of new vocabulary. However, I would only remember the words for a short period of time. I think that the Quizlet along with another practice tool would be a better approach. Rather than just learning the new words of the week, we could work to remember and use the new words (Comment 6, Student 17).
Neutral Codes	More integration with other class activities	24% (8)	2% (1)	(see Comment 3)
	More vocabulary tests in class	9% (3)	0	(see Comment 3)

who completed the questionnaire provided an answer to the open-ended question. However, because any given comment could be tagged with more than one code, the number of total coded responses is higher than the number of responses collected. Percentages are given with raw numbers between brackets. The right column shows some illustrative quotations from students' responses. The left column indicates whether these codes and comments reflect a positive, negative, or neutral view of the vocabulary activities.

Consistent with the results illustrated in Figure 5, most students considered the flashcards to be a helpful component of the learning process. This process, however, was not perfectly smooth, as students encountered issues with the format and/or the platform used to access the flashcards. This challenge seems to have arisen more frequently among students in the first two iterations of the language course, which is probably because these students were less familiar with the first software they used (*Cerego*) as well as with the format of the vocabulary intervention. Some students also preferred other formats for learning (e.g., on paper) and liked the idea of in-class tests to ensure accountability. Moreover, some students, mostly from the content courses, also mentioned the importance of integrating the flashcard activities with other topics in the class.

4.3 Quantitative results: ITT-Receptive Vocabulary Size Test

Table 5 shows the descriptive statistics of the *ITT-Test* each time the test was administered. The *ITT-Test* was given both at the beginning and the end of the semester in the language course (two sections in Fall 2020 and one section in Spring 2021), but only at the end of the semester in the content courses (one section of the literature course each in Spring 2021 and Fall 2021, and one section of the culture course in Spring 2022) (see Tables 1 and 2). By sequencing testing this way, the longitudinal data collection follows the basic tenet of the traditional pre- and post-study design. The *ITT-Test* cumulative score for all 5,000 words of each test administration was used as a proxy for learning.

Table 5. Descriptive statistics of the ITT-Test (all participating students).

Course and Test	Participants	Mean (100%)	SD
Pre_language course	22	59.77	17.20
Post_language course	22	68.86	15.56
Literature	33	68.75	14.88
Culture	15	74.80	17.19
Content course (literature + culture)	35	71.26	15.60

To explore whether the increase in vocabulary learning was statistically significant, paired t-tests were performed. These analyses were limited to the scores of those students who completed the vocabulary test more than once. Table 6 depicts the descriptive statistics for the analyzed students in each t-test, as the sample of students differed by t-test. For example, the *post-language course* score is different in the second row and the fifth row. In the second row, the score belongs to the sample of students (n=22) who took the test at the

Table 6. Descriptive statistics of the ITT-Test (paired t-tests).

Paired T-test	Course and Comparison	Mean (100%)	SD
Language course (n=22)	Pre_language course	59.77	17.21
	Post_language course	68.86	15.57
Content courses (n=11)	Literature	65.64	15.59
	Culture	70.18	17.74
Comparison: language versus content courses (n=18)	Post_language course	70.39	14.19
	Content course (literature + culture)	72.72	14.72

beginning and the end of the language course (*pre-* and *post-language course*). In the fifth row, the score refers to the students (n=18) who took it at the end of the language course and the end of one of the content courses. By pairing the tests this way, the analysis follows a pre- and post-test design.

First, to gauge learning in the language course, the pre- and post-test scores for students enrolled in that course were compared. The mean scores between both tests were significantly different (see Figure 6), $t(21)=-6.002$, $p<.001$, suggesting that there was an increase in receptive vocabulary knowledge. There was a large effect size ($d=-1.28$).

Second, the scores in the content courses were compared to each other (see Figure 7).

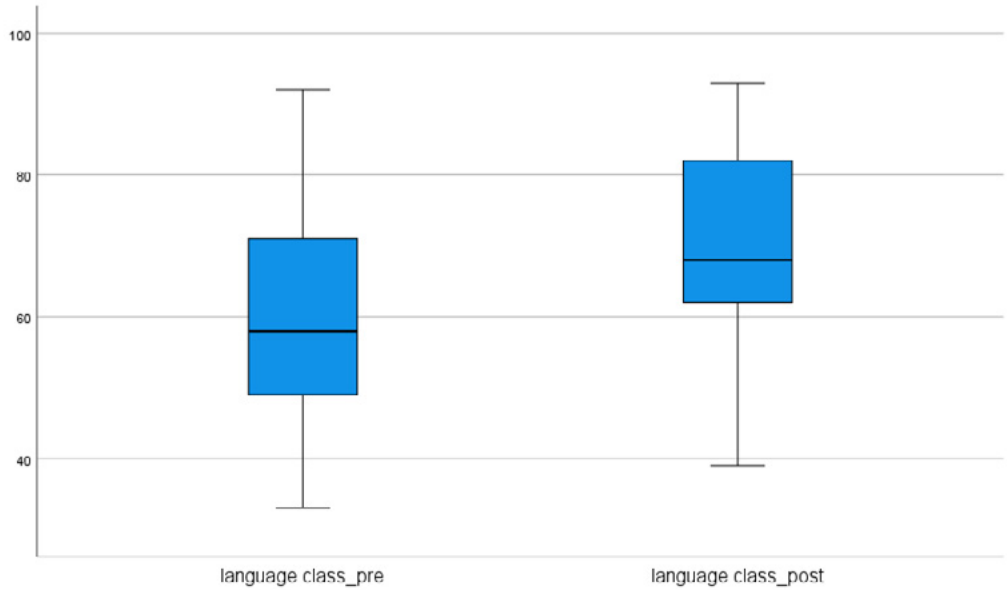

Figure 6. Descriptive statistics of the pre- and post-test in the language course (n=22).

The results were not statistically significant, $t(10)=-2.26$, $p=.047$, which suggests that there was no statistically significant increase in knowledge from one content course to the other. The p-value is established at <.005 in this study (see Benjamin et al., 2018). There was a medium effect size ($d=-.68$).

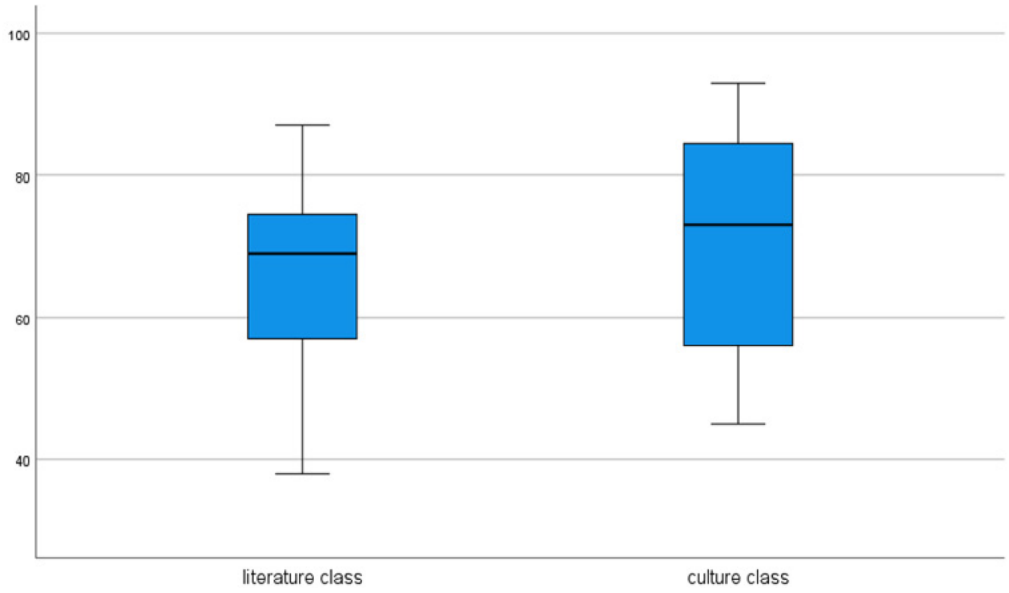

Figure 7. Descriptive statistics comparing the two content courses (n=11).

As the differences in the scores between the two content courses were not statistically significant, the scores were considered comparable and were interchangeably used as a proxy for vocabulary learning in content courses (see Figure 9). This strategy resulted in a slightly larger sample size than the one that would have been obtained by comparing the language course score with each content course score individually.

Figure 8 displays the post-test scores in the language course and the content course scores. The paired t-test result was not statistically significant, $t(17)=-1.209$, $p=.243$, which

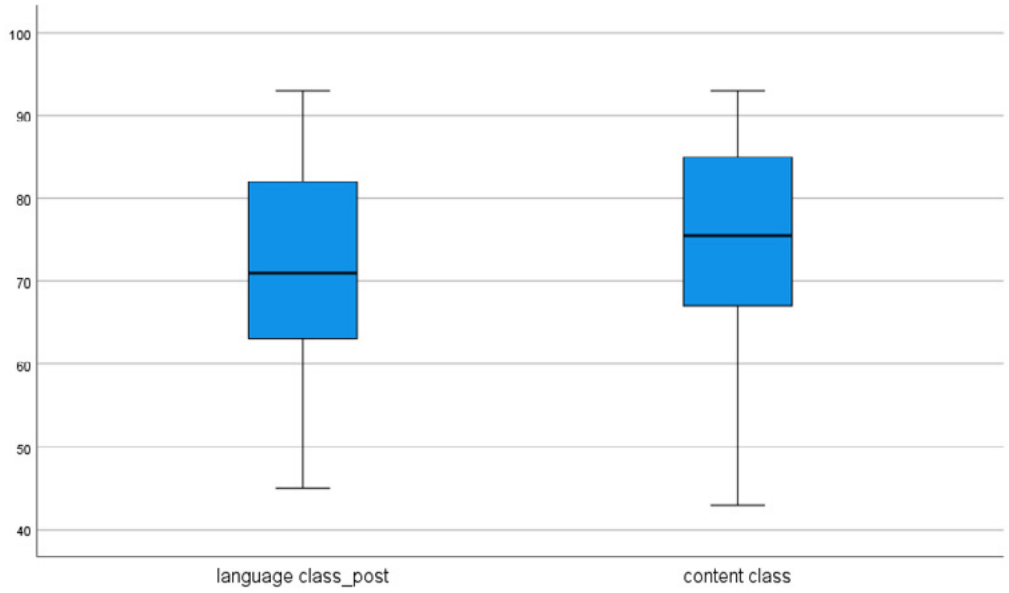

Figure 8. Descriptive statistics comparing the score from the last content course taken and the post-test score in the language course (n=18).

suggests no statistically significant increase in knowledge in the content course compared to the language course. There was a small effect size (*d*=-.28).

As Tables 5 and 6 have shown, there were large differences in SDs. To better illustrate the differences in students' developmental patterns, Figure 9 displays the growth, attrition, and/or retention of the 18 students who took the language course and either one or two of the content courses. The order of the tests is shown in the sequence they were taken. While all students scored higher on the *ITT-Test* taken at the end of the language course than on the one taken at the beginning, only 12 students showed higher scores at the end of the last content course than at the end of the language course.

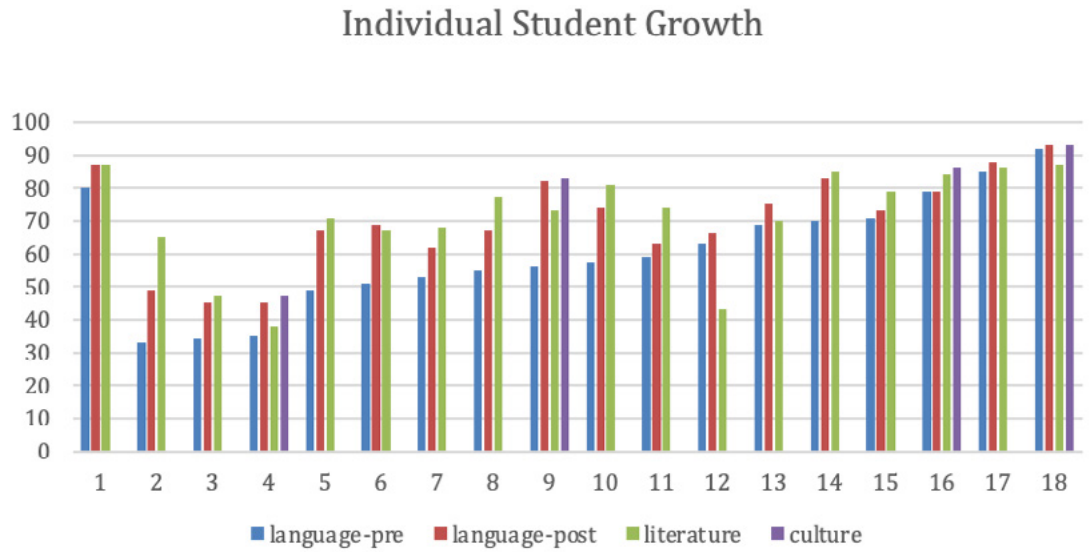

Figure 9. Individual Student Growth (Language and Content Courses, n=18).

5. Discussion

The main goal of the intervention was to integrate intentional learning of the most frequently used Spanish words into two content courses following a flashcard activity that started in the last language course the students took. The success of the intervention from the point of view of the students (their perceptions) and their actual gains will be discussed in detail in this section. The first research question and its subsections will be revised next: *What are students' perceptions regarding the use of flashcards for explicit vocabulary study of high-frequency words in content courses? And in the last language course? Are there differences and similarities between both groups (content versus language)?*

Table 4 suggests that both groups perceived they learned vocabulary using the flashcards. However, the positive perceptions were slightly lower in the content courses than in the language course. Figure 5, nevertheless, suggests that students in the content courses tended to agree with the idea that they learned more vocabulary during the

semester in which this study was conducted than in other semesters. We take these perceptions of learning as a positive effect of the flashcards, and of students' awareness of the vocabulary goals.

These results are consistent with students' perceptions in three similar content courses described in Polio and Zyzik (2009). In these three courses, vocabulary instruction came primarily from preemptive lexical focus on form, but not from explicitly planned activities in or outside the classroom. Moreover, the students in the course with the most focused instructor on vocabulary instruction perceived higher vocabulary learning than students in the two other courses with instructors less focused on vocabulary development.

In this current study, a crucial difference between the perceptions of students in the language course and students in the content courses is that the latter would like to see more integration of the flashcard activity in their courses. We assume the absence of similar comments in the language course comes from its clearer goals of language learning. In the language course, there was, however, one student who asked for more integration of the activity where students created their set in *Quizlet*. From our perspective, asking students to create their own set should provide them with the opportunity to better integrate their learning with the course activities. This particular learning activity deserves further exploration in the future.

Students in the content courses also highlighted the need to include more in-class tests, although this idea was not raised by students in the language course. We interpret this perception as a call for more accountability. Although allowing students to take the test on *Quizlet* as many times as they want is a productive idea for those students who are focused on their learning, it might be a way to pass the test with minimal effort for those who have other priorities at the time of taking the test (e.g., they can consult additional materials). At the end of the day, learners' time investment depends on their individual choices. In-class tests, on the other hand, require all students to be on top of their learning if they want to learn and/or obtain a high grade.

Something to consider, as well, is the number of words to be included in the tests. Smaller sets might be more appropriate for in-class tests than for online-based tests. In the first iterations (before Spring 2020), the sets were smaller, with around 20 words, to ensure time for learning in a week. After the move to online teaching due to COVID-19, some students found the 100-word sets too large, especially if they were tested in class. Future classroom-based studies could assess what combinations of set sizes, testing modes, and set creators (student vs. instructor) can better help learners.

Another interesting difference in perceptions relates to the ease of the platform. The percentage of students who found issues with the platform doubled in the language course compared to the content courses (24% vs. 48%). These differences may be attributed to students' familiarity with the platform: the students in the content courses were

accustomed to the activity and the platform since they had used it in previous courses. Moreover, students reported in informal conversations that they used *Quizlet* in high school; thus, this online tool may have already been familiar to most of the participants.

To address the second question, *Using longitudinal data, to what extent do 1) students in content courses and 2) students in a language course make gains in high-frequency vocabulary by using flashcards?,* the scores in the *ITT-Test* were compared using paired t-tests in three ways; namely (1) scores at the beginning of the language course with scores at the end of the language course (see Figure 6), (2) scores at the end of the two content courses (see Figure 7), and (3) scores at the end of the language course and the end of the last content course taken (see Figure 8). While the difference between pre-test and post-test scores for students in the language course was statistically significant, with post-test scores being higher than pre-test scores, the difference between post-test scores in the language course and scores in the last content course was not statistically significant. There was also no statistically significant difference in vocabulary learning between both content courses.

These findings suggest that students did not grow in their learning of the 5,000 most frequent words after the language course. The lack of a statistically significant increase in learning in the content courses does not mean that students learned no vocabulary. For example, they may have learned lower-frequency words specific to the course topics or they may have increased their depth of vocabulary knowledge. Overall, these results demonstrated that students retained the vocabulary knowledge they gained in their language courses.

With the answers to both questions, we can now answer the third research question and its sub-questions, *What is the relationship between students' perceptions of high-frequency vocabulary learning with flashcards and their actual vocabulary in content courses? And in the last language course? Are there differences and similarities between both groups?*

Learners in the content course perceived that they had learned the most frequent words and other words (see Figure 5), yet their scores in the *ITT-Test* were not consistent with those perceptions. Learners in the language course perceived they learned with the flashcards, and their vocabulary scores also showed learning.

Additional factors affecting these results need to be considered. First, the sample size comparing the content courses with the language course is smaller than the sample size of the language course (*pre-* and *post-*). Second, students' scores vary widely (see Figure 9). In brief, although we cannot conclude that students' actual learning in the content courses is higher than in the language courses, knowledge did not decrease from one course to the other, which suggests retention. The individual patterns reflected in Figure 9 also show a general increasing trend. As mentioned earlier, students' depth of vocabulary could have increased, and their learning may have included recognition and/or production of more specialized, lower-frequency words.

Based on these findings and additional anecdotal data, we offer pedagogical reflections and recommendations as well as ideas for further research. In informal conversations, some students pointed out that their classmates might not be aware of the plethora of activities available in *Quizlet*.[9] Increased training on how to use the website might be a way to integrate the vocabulary goals more clearly. *Quizlet* also includes two games that allow for informal, in-class assessment of students' learning. While some instructors might feel these games are more appropriate for younger students, they can easily be used with university students, since they require little class time and contribute to vocabulary learning. As the students suggested, implementing formal and informal in-class tests can also improve their learning by increasing their accountability.

Furthermore, since students were not completely convinced that the learned words appeared frequently in their readings, an additional activity we recommend is introducing applications that provide information on texts' word frequencies, such as "my texts" in *Corpus del Español* (Web/Dialects) (Davies, 2016) and *MultilingProfiler* (Finlayson et al., 2023). Instructors could ask students to use these applications when reading texts so that they can select words that might be new to them from the different frequency bands. This kind of activity can be integrated with social reading tools, which allow for commenting collectively on the same text while reading through an online platform.[10] Working with these two resources—the multiple options of *Quizlet* and the applications to display word frequencies—can be a way to foster the lifelong autonomous learning that is a primary goal in any language class.

In addition to these tools, vocabulary notebooks can be a helpful strategy for incorporating both general and academic vocabulary into content classes (see Dubiner, 2017; Walters & Bozkurt, 2009). By explicitly telling students to notice whether they are targeting academic vocabulary or general vocabulary, we can make learners more aware of frequent vocabulary, as learners might not target high-frequency vocabulary in their notebooks without explicit training (see McCrostie, 2007).

In our informal conversations with participating instructors, we observed different degrees of enthusiasm for the vocabulary activity. One instructor objected that students must learn specialized vocabulary rather than general vocabulary to do well in the content courses. While specialized vocabulary is certainly fundamental in content courses, the importance of high-frequency vocabulary should not be underestimated. Therefore, the abovementioned applications and students' notebooks can be especially helpful in informing both students and instructors about the importance of frequent vocabulary.

The participating instructors' varied levels of enthusiasm might serve as an important topic for future research as there is a lack of research on instructors' perspectives on teaching and learning vocabulary (Newton, 2021). The number of studies on Spanish L2 in both language-focused (e.g., Sánchez-Gutiérrez et al., 2022) and content courses (e.g., Polio

[9] There are other online flashcards that can be used instead of *Quizlet* such as *Anki*, *Lingco*, *Brainscape*, *Wordwall*, and *Memrise*.

[10] See https://libguides.denison.edu/c.php?g=312050&p=6475125/ for a list of social reading tools.

& Zyzik, 2009) is limited. Understanding instructors' perspectives is necessary to design and include sustainable language practices in content courses. For future research, knowing when and how instructors could be interested in combining opportunities for incidental learning with their own designed materials and other materials for explicit learning would provide insights into what these instructors need to further foster vocabulary development.

Because of the action-based nature of the educational practice, the data presented has some limitations. First, although including a control group is a methodological recommendation in vocabulary research, none was included here. Nevertheless, working with flashcards is better than not working with them as previous research has already demonstrated (see Nation & Webb, 2011; Webb et al., 2020). Second, the analysis carried out here focuses only on the cumulative score of all frequency bands. While the *ITT-Test* measures each frequency band individually, students' submitted PDFs and/or screenshots of their scores did not always show disaggregated data for each band. Further classroom interventions should emphasize the importance of each band, not only to raise students' awareness of the distinctions between bands and the need to learn them but also to ensure that the data is disaggregated for research purposes. Third, we did not collect any information on how the students worked on their own with *Quizlet*. Although *Quizlet* does not track the time spent on a set, it shows whether the students practice the sets on a given week. Collecting these data points (i.e., use of the flashcards and the other *Quizlet* activities for each set by the student, see Figure 3) could be helpful to better understand how they learned and what learners' behaviors can contribute to greater success.

Finally, research in courses that are not part of the required language sequence typically results in smaller sample sizes than research conducted in lower-level courses. Although we were able to increase the sample size by including more iterations, this solution presented another challenge; namely, participating teachers used a variety of teaching practices. While the resultant data reflect the reality of our program, they may not be generalizable to other contexts. Nevertheless, the use of both quantitative and qualitative methods allowed for triangulation of data, thereby increasing the likelihood that similar patterns may be seen in other programs.

6. Conclusion

This action-based study has explored the feasibility of including an explicit language learning goal in content courses, i.e., learning high-frequency words with flashcards. This study also compared perceptions of learning with actual learning within and between content- and language-focused courses. While perceptions of learning seem higher than actual learning, we take this finding as a positive result, given the contingency of vocabulary learning. Overall, using flashcards seems an appropriate tool for systematic vocabulary instruction in content courses.

Acknowledgments

We would like to thank our students and colleagues who agreed to participate in this teaching intervention and allowed us to use their classroom experiences for this article.

CRediT Author contribution / *Contribución de los autores*

Conceptualization / *Conceptualización:* Nausica Marcos Miguel & Melissa Huerta.
Formal Analysis / *Análisis formal:* Nausica Marcos Miguel & Melissa Huerta.
Methodology / *Metodología:* Nausica Marcos Miguel & Melissa Huerta.
Writing / *Redacción:* Nausica Marcos Miguel & Melissa Huerta.
Research dataset / *Datos de investigación:* Nausica Marcos Miguel & Melissa Huerta.

Funding, data availability, and copyright / *Financiación, disponibilidad de datos y derechos de autoría*

Funding / *Financiación:* No Funding / *No se ha recibido financiación*.
Image use consent / *Consentimiento para el uso de imágenes:* Permission from Quizlet and ITT-Leipzig / *Se ha consultado a Quizlet y a ITT-Leipzig para confirmar que se podían incluir las capturas de pantalla en este artículo.*
Conflict of interest / *Conflicto de intereses:* No conflict of interest / *Sin conflicto de interés*.
Data availability statement/ *Declaración de disponibilidad de datos:* Anonymized data available under request / *Se pueden compartir los datos anonimizados con investigadores cualificados que los soliciten a las autoras.*

Editorial history / *Fechas del proceso editorial*

Received / *Recibido:* 04/12/2023
Accepted / *Aceptado:* 15/09/2024
Published / *Publicado:* 01/04/2026

References

Banegas, D.L., & Villacañas de Castro, L.S. (2019). Action research. In S. Mann & S. Walsh (Eds.), *The Routledge handbook of English language teacher education* (pp. 570–582). Routledge. https://doi.org/10.4324/9781315659824-44

Barcroft, J. (2012). *Input-based incremental vocabulary instruction*. TESOL International Association.

Barclay, S., & Schmitt, N. (2019). Current perspectives on vocabulary teaching and learning. In X. Gao (Ed.), *Second handbook of English language teaching* (pp. 799–819). Springer. https://doi.org/10.1007/978-3-319-58542-0_42-1

Benjamin, D.J., Berger, J.O., Johannesson, M., Nosek, B.A., Wagenmakers, E.J., Berk, R., … Johnson, V.E. (2018). Redefine statistical significance. *Nature Human Behaviour, 2*(1), 6–10. https://doi.org/10.1038/s41562-017-0189-z

Blake, R.J. (2020). Vocabulary and the upper-division language curriculum: The case of non-native and heritage Spanish majors. *L2 Journal, 12*(3). https://doi.org/10.5070/L20048986

Burns, A. (2005). Action research: An evolving paradigm? *Language Teaching, 38*(2), 57–74. https://doi.org/10.1017/S0261444805002661

Burns, A., & Khalifa, H. (2017). Engaging action research to explore use of assessment for improving language ability. In A. Burns & H. Khalifa (Eds.), *Second language assessment and action research* (pp. 1–30). Cambridge University Press.

Coxhead, A., & Boutorwick, T.J. (2018). Longitudinal vocabulary development in an EMI international school context: Learners and texts in EAL, maths, and science. *TESOL Quarterly, 52*(3), 588–610. https://doi.org/10.1002/tesq.450

Darhower, M. (2014). Literary discussions and advanced-superior speaking functions in the undergraduate language program. *Hispania, 97*(3), 396–412. https://doi.org/10.1353/hpn.2014.0081

Davies, M. (2005a). Vocabulary range and text coverage: Insights from the forthcoming Routledge frequency dictionary of Spanish. In D. Eddington (Ed.), *Selected proceedings of the 7th Hispanic Linguistics Symposium* (pp. 106–115). Cascadilla Proceedings Project.

Davies, M. (2005b). *A frequency dictionary of Spanish: Core vocabulary for learners*. Routledge.

Davies, M. (2016). *Corpus del Español: Web/Dialects*. Retrieved 29 October 2025, from https://www.corpusdelespanol.org/web-dial/

Davies, M., & Davies, K.H. (2017). *A frequency dictionary of Spanish: Core vocabulary for learners* (2nd ed.). Routledge. https://doi.org/10.4324/9781315542638

Davies, M., & Face, T.L. (2006). Vocabulary coverage in Spanish textbooks: How representative is it? In N. Sagarra & A.J. Toribio (Eds.), *Selected proceedings of the 9th Hispanic Linguistics Symposium* (pp. 132–143). Cascadilla Proceedings Project.

Dizon, G., & Tang, D. (2017). Comparing the efficacy of digital flashcards versus paper flashcards to improve receptive and productive L2 vocabulary. *The EUROCALL Review, 25*(1), 3–15. https://doi.org/10.4995/eurocall.2017.6964

Donato, R., & Brooks, F.B. (2004). Literary discussions and advanced speaking functions: Researching the (dis)connection. *Foreign Language Annals, 37*(2), 183–199. https://doi.org/10.1111/j.1944-9720.2004.tb02192.x

Dubiner, D. (2017). Using vocabulary notebooks for vocabulary acquisition and teaching. *ELT Journal, 71*(4), 456–466. https://doi.org/10.1093/elt/ccx008

Esquivel, L. (1995). *Like water for chocolate: A novel in monthly installments with recipes, romances, and home remedies* (C. Christensen & T. Christensen, Trans.). Anchor Books. (Original work published 1989)

Finlayson, N., Marsden, E., & Anthony, L. (2023). Introducing MultilingProfiler: An adaptable tool for analysing the vocabulary in French, German, and Spanish texts. *System, 118*, 103122. https://doi.org/10.1016/j.system.2023.103122

Geisler, M., Kramsch, C., McGinnis, S., Patrikis, P., Pratt, M.L., Ryding, K., & Saussy, H. (2007). Foreign languages and higher education: New structures for a changed world. *Profession*, 234–245. https://www.jstor.org/stable/25595871

ITT e. V. (2023). Receptive Vocabulary Levels Test Spanish 1.2. Leipzig: Institute for Test Research and Test Development (ITT e. V.). Retrieved November 7th, 2025, from https://itt-leipzig.de

Karpicke, J.D., & Roediger III, H.L. (2008). The critical importance of retrieval for learning. *Science, 319*(5865), 966–968. https://doi.org/10.1126/science.1152408

Kennedy, C., & Kennedy, J. (1996). Teacher attitudes and change implementation. *System, 24*(3), 351–360. https://doi.org/10.1016/0346-251X(96)00027-9

Lei, Y., & Reynolds, B.L. (2022). Learning English vocabulary from word cards: A research synthesis. *Frontiers in Psychology, 13*, 984211. https://doi.org/10.3389/fpsyg.2022.984211

Lindstromberg, S. (2020). Intentional L2 vocabulary learning. In S. Webb (Ed.), *The Routledge handbook of vocabulary studies* (pp. 240–254). Routledge. https://doi.org/10.4324/9780429291586-16

Lomicka, L., & Lord, G. (2018). Ten years after the MLA report: What has changed in foreign language departments. *ADFL Bulletin*, *44*(2), 116–120. https://doi.org/10.1632/adfl.44.2.116

López Bastidas, L., & Sánchez-Gutiérrez, C.H. (2020). Vocabulary selection and word repetitions in beginner L2 Spanish textbooks. *Revista Electrónica de Lingüística Aplicada*, *19*(2), 48–63.

Lusin, N. (2012, March). *The MLA survey of postsecondary entrance and degree requirements for languages other than English, 2009–10*. Modern Language Association of America. https://www.mla.org/Resources/Guidelines-and-Data/Reports-and-Professional-Guidelines/Teaching-Enrollments-and-Programs/The-MLA-Survey-of-Postsecondary-Entrance-and-Degree-Requirements-for-Languages-Other-Than-English-2009-10

Macaro, E. (2018). *English-medium instruction: Language and content in policy and practice*. Oxford University Press.

Marcos Miguel, N. (2022). Vocabulary in Spanish L2 content-based courses: How do classroom elaborations provide a space for learning general and specialised vocabulary? *The Language Learning Journal*, *50*(6), 684-697. https://doi.org/10.1080/09571736.2021.1889647

Marcos Miguel, N., Edge, J.R., Beaton, M., & Sánchez-Gutiérrez, C.H. (2023). 3,000 Words in Spanish L2 Basic Language Courses: A Reachable Goal? *Second Language Research & Practice, 4*(1), 17–47.

Maxim, H.H. (2014). Curricular integration and faculty development: Teaching language-based content across the foreign language curriculum. In J. Swaffar & P. Urlaub (Eds.), *Transforming postsecondary foreign language teaching in the United States* (pp. 79–101). Springer. https://doi.org/10.1007/978-94-017-9159-5_4

McCrostie, J. (2007). Examining learner vocabulary notebooks. *ELT Journal*, *61*(3), 246–255. https://doi.org/10.1093/elt/ccm032

Montero Pérez, M. (2022). Second or foreign language learning through watching audio-visual input and the role of on-screen text. *Language Teaching*, *55*(2), 163–192. https://doi.org/10.1017/S0261444821000501

Nation, I.S.P. (1990). *Teaching and learning vocabulary*. Newbury House.

Nation, I.S.P. (2001). *Learning vocabulary in another language*. Cambridge University Press. https://doi.org/10.1017/CBO9781139524759

Nation, I.S.P., & Webb, S.A. (2011). *Researching and analyzing vocabulary*. Heinle Cengage Learning.

Newton, J. (2019). Approaches to learning vocabulary inside the classroom. In S. Webb (Ed.), *The Routledge handbook of vocabulary studies* (pp. 255–270). Routledge. https://doi.org/10.4324/9780429291586-17

Newton, J. (2021). Teacher and learner perspectives on vocabulary learning and teaching. In H. Mohebbi & C. Coombe (Eds.), *Research questions in language education and applied linguistics* (pp. 137–141). Springer. https://doi.org/10.1007/978-3-030-79143-8_26

Paesani, K. (2017). Redesigning an introductory language curriculum: A backward design approach. *L2 Journal*, *9*(1), 1–20. https://doi.org/10.5070/L29130408/

Polio, C., & Zyzik, E. (2009). Don Quixote meets *ser* and *estar*: Multiple perspectives on language learning in Spanish literature classes. *The Modern Language Journal*, *93*(4), 550–569. https://doi.org/10.1111/j.1540-4781.2009.00930.x

Qian, D.D., & Lin, L.H. (2020). The relationship between vocabulary knowledge and language proficiency. In S. Webb (Ed.), *The Routledge handbook of vocabulary studies* (pp. 66-80). Routledge. https://doi.org/10.4324/9780429291586-5

Robles-García, P. (2022). Receptive vocabulary knowledge in L2 learners of Spanish: The role of high-frequency words. *Foreign Language Annals*. Advance online publication. https://doi.org/10.1111/flan.12630

Robles-García, P., Stewart, J., Nicklin, C., Vitta, J.P., McLean, S., & Kramer, B. (2023). The wisdom of crowds: When teacher judgments outperform word frequency as a predictor of students' vocabulary knowledge. *Language Teaching Research*. Advance online publication. https://doi.org/10.1177/13621688231176067

Roussel, S., Joulia, D., Tricot, A., & Sweller, J. (2017). Learning subject content through a foreign language should not ignore human cognitive architecture: A cognitive load theory approach. *Learning and Instruction*, *52*, 69–79. https://doi.org/10.1016/j.learninstruc.2017.04.007

Ruiz de Zarobe, Y., & Cenoz, J. (2015). Way forward in the twenty-first century in content-based instruction: Moving towards integration. *Language, Culture and Curriculum*, *28*(1), 90–96. https://doi.org/10.1080/07908318.2014.1000927

Sánchez-Gutiérrez, C.H., Robles-García, P., & Álvarez, C.H. (2023). Vocabulary in the L2 Spanish classroom: What students know and what their instructors believe they know. In I. Checa-García & L. Marqués-Pascual (Eds.), *Current perspectives in Spanish lexical development* (pp. 245–268). De Gruyter Mouton. https://doi.org/10.1515/9783110730418-010

Sánchez-Gutiérrez, C.H., Robles-García, P., & Pérez Serrano, M. (2022). L2 Spanish vocabulary teaching in US universities: Instructors' beliefs and reported practices. *Language Teaching Research*, *29*(2), 677-699. https://doi.org/10.1177/13621688221074443

Seow, A. (2002). The writing process and process writing. In J.C. Richards & W.A. Renandya (Eds.), *Methodology in language teaching: An anthology of current practice* (pp. 315–320). Cambridge University Press. https://doi.org/10.1017/CBO9780511667190.044

Stanley, W.M., & González, L. (2015). *Gramática para la composición* (3rd ed.). Georgetown University Press.

Thoms, J.J. (2014). An ecological view of whole-class discussions in a second language literature classroom: Teacher reformulations as affordances for learning. *The Modern Language Journal*, *98*(3), 724–741. https://doi.org/10.1111/j.1540-4781.2014.12119.x

Tschirner, E. (2021). Examining the validity and reliability of the IIT vocabulary size tests. *Research Papers in Assessment*. Institut für Testforschung und Testentwicklung e.V. Leipzig. https://nbn-resolving.org/urn:nbn:de:bsz:15-qucosa2-763177

Vander Beken, H., De Bruyne, E., & Brysbaert, M. (2020). Studying texts in a non-native language: A further investigation of factors involved in the L2 recall cost. *Quarterly Journal of Experimental Psychology*, *73*(6), 891–907. https://doi.org/10.1177/1747021820910694

Walters, J., & Bozkurt, N. (2009). The effect of keeping vocabulary notebooks on vocabulary acquisition. *Language Teaching Research*, *13*(4), 403–423. https://doi.org/10.1177/1362168809341509

Winke, P., Zhang, X., Rubio, F., Gass, S., Soneson, D., & Hacking, J. (2020). The proficiency profiles of language students: Implications for programs. *Second Language Research & Practice*, *1*(1), 25–64. http://hdl.handle.net/10125/69840

Webb, S. (2020a). Incidental vocabulary learning. In S. Webb (Ed.), *The Routledge handbook of vocabulary studies* (pp. 225–239). Routledge. https://doi.org/10.4324/9780429291586-15

Webb, S. (Ed.). (2020b). *The Routledge handbook of vocabulary studies*. Routledge. https://doi.org/10.4324/9780429291586

Webb, S., & Chang, A.C. (2012). Second language vocabulary growth. *RELC Journal*, *43*(1), 113–126. https://doi.org/10.1177/0033688212439367

Webb, S., Yanagisawa, A., & Uchihara, T. (2020). How effective are intentional vocabulary-learning activities? A meta-analysis. *The Modern Language Journal*, *104*(4), 715–738. https://doi.org/10.1111/modl.12671

Yang, H.H., & Farley, A. (2019). Quantifying the impact of language on the performance of international accounting students: A cognitive load theory perspective. *English for Specific Purposes*, *55*, 12–24. https://doi.org/10.1016/j.esp.2019.03.003

Yüksel, H., Mercanoğlu, H., & Yılmaz, M. (2020). Digital flashcards vs. wordlists for learning technical vocabulary. *Computer Assisted Language Learning*, *35*(8), 2001–2017. https://doi.org/10.1080/09588221.2020.1854312

Zhang, X., & Lu, X. (2014). A longitudinal study of receptive vocabulary breadth knowledge growth and vocabulary fluency development. *Applied Linguistics*, *35*(3), 283–304. https://doi.org/10.1093/applin/amt014