

German for Specific Purposes: Developing lexical knowledge in commerce and marketing classes at university level

Tim Hammrich

Universidad de Oviedo (Spain)
hammrichtim@uniovi.es

Abstract

The article analyses the effectiveness of two different learning approaches on the development of lexical knowledge in LSP classes at the university. In a GFL lecture in the Bachelor's degree programme in Marketing and Commerce, different tasks were used within an intentional and an incidental learning setting to develop lexical knowledge. The experimental group used an application to memorise the specialised vocabulary (intentional learning), while the control group worked on writing tasks (incidental learning). All other relevant learning factors were identical for both groups. After completing learning units 1 and 2, the experimental and control groups were swapped for learning units 3 and 4 in order to avoid bias. Learning progress of each group was measured throughout the entire semester and took place at the end of each learning unit using a total of 4 tests. Each test consisted of 3 subtests: vocabulary test, writing test and C-test. Tests 1 & 2 were repeated 8 weeks later as delayed tests. The results show that the task format used in the learning phase to acquire vocabulary has an influence on the different competences measured. Furthermore, it was shown that intentional learning using an application led to better learning gains in the areas of general language proficiency and vocabulary knowledge. Delayed tests also showed that learning gains are maintained over time and remain consistent even after swapping the experimental and control groups. Didactic implications can be derived from the results obtained in this study.

Keywords: Language for Specific Purposes, German as a Foreign Language, lexical knowledge, intentional learning, ICT.

Resumen

Alemán para fines específicos: El desarrollo de la competencia léxica en una asignatura de Marketing y Comercio

El artículo analiza la eficacia de dos enfoques de aprendizaje distintos en el desarrollo de la competencia léxica en clases de lenguas para fines específicos en el ámbito universitario. Para ello, en una asignatura de alemán como lengua extranjera del Grado en Marketing y Comercio, se aplicaron diferentes tareas en un contexto de aprendizaje intencional frente a un contexto de aprendizaje incidental. El grupo experimental utilizó una aplicación para memorizar el vocabulario especializado (aprendizaje intencional), mientras que el grupo de control realizó tareas de redacción (aprendizaje incidental). Todos los demás factores relevantes de aprendizaje se mantuvieron idénticos en ambos grupos. Tras completar las unidades de aprendizaje 1 y 2, los grupos experimental y de control se intercambiaron en las unidades 3 y 4 con el fin de evitar sesgos. El progreso de aprendizaje de cada grupo se midió a lo largo de todo el semestre y al finalizar cada unidad didáctica mediante un total de cuatro pruebas. Cada prueba constaba de tres subpruebas: un test de vocabulario, un test de redacción y un C-test. Asimismo, las pruebas 1 y 2 se repitieron ocho semanas más tarde como pruebas diferidas. Los resultados muestran que el formato de las tareas empleadas en la fase de aprendizaje para la adquisición de vocabulario influye en las diferentes competencias evaluadas. Por otra parte, se evidenció que el aprendizaje intencional mediante una aplicación condujo a mayores resultados en las áreas de competencia lingüística general y conocimiento del vocabulario. Las pruebas diferidas confirmaron, además, que los progresos de aprendizaje se mantienen a lo largo del tiempo y se mantienen consistentes incluso después del intercambio entre el grupo experimental y el de control. De los resultados obtenidos en este estudio se derivan implicaciones didácticas relevantes.

Palabras clave: Lengua para fines específicos, alemán como lengua extranjera, competencia léxica, aprendizaje intencional, TIC.

1. Rationale

In literature, it is often claimed that vocabulary knowledge is crucial in language learning (Barcroft, 2016; Boers, 2021; Laufer & Nation, 2012). This seems to apply in particular to the area of Languages for Specific Purposes (LSP), as their specific characteristics are particularly prominent at the lexical level (Coxhead, 2018; Hoffmann, 1987). Furthermore, specialised vocabulary presents a significant portion in specialized texts (Buhlmann & Fearn, 2000; Chung & Nation, 2003) and becomes essential for the

understanding of professional communication (Knoch, 2014) and the development of subject knowledge (Liu & Lei, 2020). As a consequence, learners need to master a considerable amount of vocabulary in their subject area (Coxhead, 2018). Against this background, didactics of specialised vocabulary becomes increasingly important and terminology a key area of language learning (Charles & Pecorari, 2016). Nevertheless, studies have evidenced that learners fail to acquire a satisfactory number of specialised vocabulary (Coxhead et al., 2016; Evans & Morrison, 2011) and that vocabulary learning represents a considerable burden for them (Coxhead, 2000; Webb & Nation, 2017). At the same time, teachers often complain about a lack of sufficient guidance provided by textbooks and curricula (Schmitt, 2008) as well as insufficient teaching time to deal with vocabulary issues during classes (Chacón-Beltrán et al., 2010). Research results even suggest that there is a considerable gap between the actual lexical knowledge of L2 learners and the lexical requirements of specific subject areas (Nekrasova & Beker et al., 2019).

The existing didactic models for LSP usually follow a communicative approach, with comparatively little attention paid to vocabulary issues. Altogether, there is a strong preference for teaching specialized vocabulary in context, while the usefulness of learning isolated word structures remains controversial (Boers, 2021; Laufer, 2005b; Reynolds et al., 2021). At the same time, a number of desiderata is indicated (Breindl, 2021; Woodrow, 2022), in particular the lack of studies investigating the effectiveness of methods that place a greater focus on learning and teaching processes in the classroom, including virtual environments (Coxhead, 2016; Flowerdew, 2015). In addition, it has been objected, that much of the research focuses on English as a Foreign Language, while results for other languages are lacking. Against this background, the present study examines the benefits of isolated memorization of specialised vocabulary in the context of a business German course at university level.

2. Lexical knowledge

There are different approaches when it comes to the development of lexical knowledge. In the following, Laufer's Focus on Form approach and the concepts of incidental and intentional vocabulary learning are examined, insofar as they relate to the present work.

2.1. Focus on form

In the area of teaching vocabulary, Laufer distinguishes between two major types of form-focused instruction (FFI): Focus on Form (FonF) and Focus on Forms (FonFs).

In a FonF approach, it is assumed that focusing on meaning alone is not sufficient to ensure adequate lexical development. Therefore, the learner's attention must be drawn to certain lexical structures. This is achieved through form-focused tasks that are embedded in a communicative context (task embedded Focus on Form, FonF). These may include activities such as filling in previously deleted vocabulary in a text or answering comprehension questions about the text that require the use of specific lexical structures. All such activities, however, always take place within a communicative context and their main purpose is to contribute to the fulfilment of the communicative activity: "In terms of vocabulary instruction, FonF refers to attending to lexical items [...] within a communicative task environment when these lexical items are necessary for the completion of a communicative, or an authentic language task" (Laufer, 2010, p. 17).

The term FonFs, on the other hand, refers to "teaching and practicing discrete lexical items in non-communicative, non-authentic language tasks" (Laufer, 2010, p. 17). In this approach, therefore, work with lexical structures is isolated from the text and the focus is not primarily on communication, but on words as the object of learning. Laufer further subdivides FonFs into form-focused tasks that refer to (but are not embedded in) a communicative context (Task-related Focus on Forms, related FonFs) on the one hand, and purely form-focused tasks (pure Focus on Forms, pure FonFs) on the other. The difference lies in how the vocabulary work is linked to a communicative activity: "In task-related FonFs activities, words are the objects of learning, but they are, nevertheless, related to, though not embedded in, a meaning-based task which is central in a lesson" (Laufer, 2005a, p. 238). An example for this type of task would be an activity in which the learner is asked to relate the new vocabulary in the text to its corresponding synonyms after reading. Pure FonFs, on the other hand, involve a complete decontextualization of the target vocabulary, i.e. these exercises "require learners to work with isolated words that are not related to any meaning-based task" (Laufer, 2005a, p. 238). Examples of tasks located within pure FonFs are crossword puzzles or learning words with lists or flashcards. To further exemplify the distinction between Focus on Form (FonF) and Focus

on Forms (FonFs) in vocabulary teaching, we will consider Laufer's example of the introduction of 10 new lexical items to a group of learners. These words appear in a text and must be understood before the learners can participate in a group discussion.

Therefore, learners decide to look them up in a dictionary. Looking them up constitutes FonF, since these words are attended to in order to complete a communicative/authentic language task. However, the same 10 words may be presented with their L1 translations in a decontextualized list and supplemented by two vocabulary exercises: (1) match each word in column A to its definition in column B and (2) fill in the words in 10 sentences, one word in each sentence. In these two tasks, attending to the words constitutes FonFs as the tasks are not connected with a communicative or authentic language activity, but the words are the objects of study (Laufer 2010, p. 17-18).

2.2. Incidental and intentional learning of vocabulary

There are different definitions of the terms incidental learning and intentional learning (Webb, 2020). In this paper, we will understand these terms according to Barcroft and Boers. Barcroft refers to incidental learning as the acquisition of new words from context without the learner clearly intending to do so, i.e. "when one 'picks up' new words and lexical phrases from context without intending to do so" (Barcroft, 2016, p. 12). It should be mentioned, however, that the distinction between intentional and incidental learning can sometimes be difficult (Boers, 2021) and, to some extent, problematic (Webb, 2020), as "learners do shift their attention occasionally to items or features of the target language even if they are primarily engaged with communicative content" (Boers, 2021, 17). That is, when reading a text, the learner may look up the meaning of a word with the intention of understanding the text. Consequently, the terms incidental and intentional do not represent two exactly opposite terms with a strict dividing line. We therefore prefer to speak of "incidentally oriented" and "intentionally oriented" (Barcroft, 2015, pp. 44-46), assuming that the two concepts are not dichotomously opposed, but are to be understood as a continuum. Learning activities are thus neither purely intentional nor incidental and show gradual differences in terms of attention directed towards vocabulary learning versus other possible aspects (Barcroft, 2015, p. 51). In this sense, we understand that a task focused on learning words tends to be more towards the intentional learning end of the spectrum than a

reading and discussion task. Tasks within the pure FonFs approach, as defined by Laufer, would therefore be situated “at the most intentional end of the incidental-intentional continuum” (Lindstromberg, 2020, p. 242).

3. Objectives

The present study investigates the learning impact for two different learning approaches, incidental versus intentional (always conceived according to the conception of both approaches as a continuum described above), in a Specialised German as a Foreign Language class in the Bachelor’s degree of Marketing and Commerce. The aim is to investigate whether the respective learning approaches lead to different results in the different test types and whether the learning outcomes are sustainable, i.e. remain relevant over time. Research questions (RQ) are as follows:

- RQ1: Are there differences in the learning gains of the experimental and the control group, each learning with a different learning approach during the various learning phases (intentional versus incidental)?
- RQ2: In which way do the different learning approaches impact the results of the different test types (subtest 1: vocabulary test, subtest 2: writing test and subtest 3: C-test¹)?
- RQ3: Is there a difference in the sustainability of the learning gains across the different subtests?

4. Context of research study

The study was carried out within the Bachelor’s degree in Marketing and Commerce at a university in northern Spain in 2022. Participants are between 19 and 22 years old and in the third year of their bachelor’s degree, and therefore have basic knowledge of the various areas of economics taught in the degree program. To attend German lectures, no previous knowledge of the German language is required and lessons start at language level zero. In 2022, 26 students took part in the course. Nevertheless, to avoid bias, students who already have some German language skills before

starting the course, as well as those who submit less than 75% of the tasks to be completed during the learning phases, are excluded from the study. Altogether 20 students (N=20) took part in the present study, 11 in Group 1 (G1) and 9 in Group 2 (G2). The course is divided into a total of 24 sessions of 90 minutes each, 2 sessions per week during the 14 weeks of the semester. Lessons follow a communicative approach, in which professional and subject-relevant language skills are developed at a beginner's level. With a particular focus on future professional activities, students are to be introduced to business-related communication. The vocabulary of the units covered can primarily be described as general and technical vocabulary (see Coxhead, 2013; Liu & Lei, 2020).

5. Didactic intervention

For the didactic intervention, the class is divided into an experimental and a control group. Assignment to each group takes place based on the alphabetical order of the students' last names and is carried out by the university administration. All relevant learning factors are the same for both groups (lecturer, learning content, objectives, textbook, assessment, etc.). The only difference lies in the approach to learning the vocabulary during the 14 weeks of the course, i.e. students engage in different exercises that are carried out as autonomous learning activities outside class time and are evaluated by the lecturer after submission.

On the one hand, the experimental group uses a learning app (www.leo.org) to learn the vocabulary intentionally: During classes, new vocabulary is looked up by each student individually in the online dictionary and then saved in a personal vocabulary list. The words in this list can then be practised and learned through the various exercises offered by the app (e.g. memory, flash cards, dictate, etc.). Doing this, the app automatically creates learning statistics in which various parameters (time spent on exercise, exercise type, words, etc.) are documented. At the end of each learning unit, the student hands in her/his learning statistics for evaluation and monitoring of the learning process. In contrast to this, the control group learns the vocabulary incidentally by working on writing tasks: During each learning unit, the students must write texts. These contain the vocabulary relevant to the unit and are submitted for evaluation at the end of each learning unit.

To avoid bias, the experimental group (intentional learning with app) and the control group (incidental learning through writing tasks) were swapped after 2 learning units, i.e., in units 3 and 4, the experimental group became the control group and vice versa.

Learning phase	Unit of the textbook	Didactic intervention
Phase 1	Learning Unit 1	Vocabulary acquisition: Group 1: Experimental Group: Intentional Learning with app Group 2: Control Group: Incidental Learning with Writing Tasks Submission of tasks at the end of Learning Unit 1: Group 1: Experimental group: Submission of vocabulary learning statistics Group 2: Control group: Submission of Writing Tasks
	Test 1*	Vocabulary Test 1 + Writing Test 1 + C-Test 1
Phase 2	Learning Unit 2	Vocabulary acquisition: Group 1: Experimental Group: Intentional Learning with app Group 2: Control Group: Incidental Learning with Writing Tasks Submission of tasks at the end of Learning Unit 2: Group 1: Experimental group: Submission of vocabulary learning statistics Group 2: Control group: Submission of Writing Tasks
	Test 2*	Vocabulary Test 2 + Writing Test 2 + C-Test 2
Group Change		Experimental Group becomes Control Group Control Group becomes Experimental Group
Phase 3	Learning Unit 3	Vocabulary acquisition: Group 1: Control Group: Incidental Learning with Writing Tasks Group 2: Experimental Group: Intentional Learning with app Submission of tasks at the end of Learning Unit 3: Group 1: Control group: Submission of Writing Tasks Group 2: Experimental group: Submission of vocabulary learning statistics
	Test 3*	Vocabulary Test 3 + Writing Test 3 + C-Test 3
Phase 4	Learning Unit 4	Vocabulary acquisition: Group 1: Control Group: Incidental Learning with Writing Tasks Group 2: Experimental Group: Intentional Learning with app Submission of tasks at the end of Learning Unit 4: Group 1: Control group: Submission of Writing Tasks Group 2: Experimental group: Submission of vocabulary learning statistics
	Test 4*	Vocabulary Test 4 + Writing Test 4 + C-Test 4
Group Change		Original group assignment as in Learning Phase 1 and 2
Delayed Phase 1	Delayed Test 1**	Delayed Vocabulary Test 1 + Delayed Writing Test 1 + Delayed C-Test 1
Delayed Phase 2	Delayed Test 2**	Delayed Vocabulary Test 2 + Delayed Writing Test 2 + Delayed C-Test 2

* Same test for both groups

** Delayed Test 1 = Test 1, Delayed Test 2 = Test 2, 8 weeks after.

Table 1. Scheme of didactical intervention.

During class time, vocabulary is addressed solely by means of incidental learning. No special attention is placed on intentional memorizing of vocabulary during the session.

6. Assessment of learning gain

At the end of each learning unit, a test to measure learning gain is applied. Each of these tests consists of three different subtests to be able to detect

learning gains across different areas of knowledge. Subtest 1 is a conventional vocabulary test (Vocabulary Test, VT), in which students are asked to translate 10 words. 5 words must be translated from L1 (Spanish) into L2 (German), i.e. students must provide form from meaning (active recall) and 5 words from L2 to L1, i.e. students must provide meaning from form (passive recall) (Laufer et al., 2004; Laufer & Goldstein, 2004). Hence, knowledge of word meaning and word form is measured. Subtest 2 is a writing test (Writing Test, WT), which requires students to write phrases or short texts with given information. A writing test measures students' writing ability, which relates to a series of subcompetencies (knowledge of syntax, vocabulary, textual knowledge, and pragmatic knowledge (Bachmann & Palmer, 1996). Thus, students are required to activate that knowledge efficiently in written communication. Subtest 3 is a C-Test. A C-test is a cloze test, i.e. a text in which from the second sentence onwards, every second word is incomplete (the second half of the word is eliminated) and needs to be completed by the student. A C-Test is an instrument to measure general language proficiency (Cantera Adrados, 2010; Eckes & Grotjahn, 2006; Grotjahn et al., 2002).

7. Analysis method

All statistical calculations were conducted with SPSS. As part of the descriptive statistics, mean value, standard deviation (SD), and mean standard error (SEM) are calculated for the test results of both groups. To compare the results of learning gain of the experimental and the control group, we used the t-test and the Mann-Whitney U test. Significance level is set to $\alpha < 0.05$ (significant) and $\alpha < 0.01$ (strong significance), respectively. Before applying the independent samples t-test, we verified that the necessary assumptions were met—specifically, normality (Shapiro–Wilk test) and homogeneity of variances (Levene's test). Although the sample sizes were relatively small ($N = 11$ and $N = 9$), the large differences between group means, combined with low within-group variability, resulted in statistically significant p-values as reported in the results section. Prior research supports the validity of using the t-test with small samples under these conditions (de Winter, 2013; Rhiel & Wilkie, 1996). In cases where assumptions were not met, we used the Mann–Whitney U test, whose suitability for small-sample contexts has also been documented (Zhu, 2021).

To determine effect size, Cohen’s *d* was calculated for *t*-test results and Pearson’s *r* for Mann-Whitney *U* test results. With regards to the interpretation of the *d* values, we assume that effect size is small with *d* > 0.2, medium with *d* > 0.5 and large with *d* > 0.8 . For the interpretation of *r* values, we follow Gignac and Szodorai considering *r* > 0.1 a small, *r* > 0.2 a medium and *r* > 0.3 a large effect size.

8. Results

8.1. Vocabulary Test

As can be seen from the descriptive statistics in Table 2, the experimental group (G1) achieved better results in the Vocabulary Tests 1 and 2 (VT1 & VT2) and in the Delayed Vocabulary Tests (DVT1 & DVT2) than the control group (G2). In contrast, after switching the experimental and control groups (see Table 1), G2 achieved higher test results in VT3 and in VT4. In other words, the respective group that used the app to memorize the vocabulary through intentional learning during the learning phase also achieved higher learning gains in the Vocabulary Test.

Descriptive statistics						Inferential statistics	
Test	Group	N	Mean	Standard deviation	Mean standard error	t-test	effect size
VT 1	experimental (G1)	11	7,18	2,523	,761	<i>t</i> (18) =2.667, <i>p</i> = .016*	<i>d</i> = 1.199, large
	control (G2)	9	4,44	1,944	,648		
VT 2	experimental (G1)	11	6,82	2,272	,685	<i>t</i> (18) = 3.408, <i>p</i> = .003**	<i>d</i> = 1.532, large
	control (G2)	9	3,22	2,438	,813		
VT 3	control (G1)	11	5,55	2,382	,718	<i>t</i> (18) = -.124, <i>p</i> = .902	<i>d</i> = -.056, medium
	experimental (G2)	9	5,67	1,871	,624		
VT 4	control (G1)	11	4,73	1,679	,506	<i>U</i> =30,000, <i>Z</i> =-1,518, <i>p</i> = .152	<i>r</i> = .34, large
	experimental (G2)	9	5,78	1,394	,465		
DVT1	experimental (G1)	11	6,27	3,003	,905	<i>U</i> = 44,000, <i>Z</i> = -.427, <i>p</i> = .710	<i>r</i> = .1, small
	control (G2)	9	5,89	2,667	,889		
DVT2	experimental (G1)	11	5,73	2,494	,752	<i>t</i> (18) = 1.943, <i>p</i> = .068	<i>d</i> = .873, large
	control (G2)	9	3,78	1,856	,619		

Note: * significant; ** strongly significant, *r* = Pearson correlation coefficient, *d* = Cohen’s *d*.

Table 2. Results of vocabulary tests and delayed vocabulary tests.

The results of the interference statistics (Table 2) are as follows: Students from G1 compared to students from G2 demonstrated significantly better scores in VT1 ($t(18) = 2.667$, $p = .016^*$) and in VT2 ($t(18) = 3.408$, $p = .003^{**}$) with a large effect size in both, being $d = 1.199$ in VT1 and $d = 1.532$ in VT2. In contrast, the picture is reversed in VT3 and VT4 with G1 attaining lower scores compared to G2, although tests did not show a statistic significance, with VT3 being $t(18) = -.124$, $p = .902$ with a medium effect size of $d = -.056$ and VT4 being $U = 30,000$, $Z = -1,518$, $p = .152$ with a large effect size of $r = .34$. In both delayed tests G1 attained higher scores than G2, although tests didn't show a significance with DVT1 being $U = 44,000$, $Z = -.427$, $p = .710$ with a small effect size $r = .1$ and DVT2 being $t(18) = 1.943$, $p = .068$ with a large effect size of $d = .873$.

8.2. Writing test

As shown in Table 3, the control group (G2) achieved better results in the Writing Tests 1 and 2 (WT1 & WT2) and in the Delayed Writing Tests (DWT1 & DWT2) than the experimental group (G1). In contrast, after the group change, G1 achieved higher test results in WT3 and in WT4. This indicates that the group that worked with the writing tasks during the learning phase respectively also achieved higher scores in the associated subtest.

Descriptive statistics						Inferential statistics	
Test	Group	N	Mean	Standard deviation	Mean standard error	t-test	effect size
WT1	experimental (G1)	11	5,636	1,5181	,4577	$t(18) = -1,868$, $p = .078$	$d = -.840$, large
	control (G2)	9	6,722	,9391	,3130		
WT2	experimental (G1)	11	5,545	1,4740	,4444	$t(18) = -1,594$, $p = .128$	$d = -.716$, large
	control (G2)	9	6,778	1,9861	,6620		
WT3	control (G1)	11	6,818	1,9010	,5732	$t(18) = 3,833$, $p = .001^{**}$	$d = 1,723$, large
	experimental (G2)	9	3,889	1,4093	,4698		
WT4	control (G1)	11	7,864	1,4848	,4477	$U = 17,500$, $Z = -2,448$, $p = .012^*$	$r = 0,55$, large
	experimental (G2)	9	5,389	2,4972	,8324		
DWT1	experimental (G1)	11	5,955	1,8091	,5455	$t(18) = -.683$, $p = .503$	$d = -.307$, small
	control (G2)	9	6,556	2,1279	,7093		
DWT2	experimental (G1)	11	6,818	1,4190	,4279	$t(18) = -.026$, $p = .980$	$d = -.012$
	control (G2)	9	6,833	1,1726	,3909		

Note: * significant; ** strongly significant, r = Pearson correlation coefficient, d = Cohen's d .

Table 3. Descriptive statistics for results of writing tests and delayed writing tests.

As for interference statistics (Table 3) the results show that students from G2 compared to students from G1 demonstrated higher scores in writing tests WT1 ($t(18) = -1,868$, $p = .078$) and WT2 ($t(18) = -1.594$, $p = .128$) and the two delayed writing tests DWT1 ($t(18) = -.683$, $p = .503$) and DWT2 ($t(18) = -.026$, $p = .980$), whereas G1 outperformed G2 with statistically significantly better scores in WT3 ($t(18) = 3.833$, $p = .001^{**}$) and WT4 ($U = 17.500$, $Z = -2,448$, $p = .012^{*}$). Effect sizes are large in all four writing tests, being $d = -.840$ in WT1, $d = -.716$ in WT2, $d = 1.723$ in WT3 and $r = 0.55$ in WT4 and there is a small effect size in the delayed writing test DWT1 ($d = -.307$).

8.3. C-Test

Table 4 shows that the experimental group (G1) achieves higher scores than the control group (G2) in C-Test 1 and 2 (CT1 & CT2) and the two delayed C-Tests (DCT1 & DCT2). Whereas G2 obtains higher scores in C-Test 3 and 4 (CT3 and CT4), meaning that the group that worked with the app (intentional learning) during the learning phases, respectively, also achieved higher scores in the C-Test.

Test	Group	N	Descriptive statistics			Inferential statistics	
			Mean	Standard deviation	Mean standard error	t-test	effect size
CT1	experimental (G1)	11	26,09	7,993	2,410	$t(18) = 1.655$, $p = .115$	$d = .744$, large
	control (G2)	9	20,39	7,236	2,412		
CT2	experimental (G1)	11	27,09	5,009	1,510	$t(18) = 2.210$, $p = .040^{*}$	$d = .993$, large
	control (G2)	9	21,67	5,979	1,993		
CT3	control (G1)	11	23,27	7,862	2,371	$t(18) = -.356$, $p = .726$	$d = -.160$
	experimental (G2)	9	24,44	6,598	2,199		
CT4	control (G1)	11	19,00	4,099	1,236	$t(18) = -1.556$, $p = .137$	$d = -.700$, large
	experimental (G2)	9	21,56	3,005	1,002		
DCT1	experimental (G1)	11	27,27	6,084	1,834	$t(18) = 2.106$, $p = .049^{*}$	$d = .947$, large
	control (G2)	9	21,56	5,981	1,994		
DCT2	experimental (G1)	11	26,09	5,338	1,609	$t(18) = 2.533$, $p = .021^{*}$	$d = 1.138$, large
	control (G2)	9	20,89	3,371	1,124		

Note: * significant; ** strongly significant, $d = \text{Cohen's } d$.

Table 4: Descriptive statistics for results of C-Tests and delayed C-Tests.

Interference statistics (Table 4) show that higher scores of G1 are statistically significant in CT2 ($t(18) = 2.210$, $p = .040^*$) and both delayed tests DCT1 ($t(18) = 2.106$, $p = .049^*$) and DCT2 ($t(18) = 2.533$, $p = .021^*$), but not so in CT1 ($t(18) = 1.655$, $p = .115$), neither are they in CT3 and CT4, where G2 obtained better results than G1 with CT3 ($t(18) = -.356$, $p = .726$) and CT4 ($t(18) = -1.556$, $p = .137$). Effect sizes are large in all test results being in CT1 $d = .744$, CT2 $d = .993$, CT4 $d = -.700$, DCT1 $d = .947$ and DCT2 $d = 1.138$, except in CT3 ($d = -.160$).

9. Discussion

9.1. Learning impact of tasks in the Vocabulary Tests and Writing Tests

Considering the results of the Vocabulary Tests (Table 2), it can be noted that the group that studied the vocabulary with the app, memorising isolated vocabulary (intentional learning), achieved better results in the Vocabulary Test. This applies to both groups, G1 and G2. Whereas regarding the Writing Tests (Table 3), results show that the group that worked with the writing tasks during the learning phase achieved a better test result in the Writing Tests. Again, this applies to G1 and G2. In my opinion, three fundamental conclusions can be drawn from these results. First, the didactic intervention carried out in the present study proves to have a measurable impact on learning. This is illustrated by the better test results that each group achieved in the subtest that corresponds to the specific task type that the group dealt with during the learning phase. It can therefore be deduced that the tasks implemented during learning phases represent a feasible strategy for improving the respective competences in LSP classes. Secondly, results confirm that practising a particular skill using a particular task type also tends to lead to better learning outcomes when tested by using the same or a similar task type in the test. This is consistent with research findings (Webb, 2009). LSP courses usually have homogeneous learner groups, which permits to clearly delimit specific competencies as learning objectives (Anthony, 2018) (reading technical instructions, on-flight communication with passengers, etc.). Knowing that the relation between tasks during the learning phase and test tasks is straightforward provides a sound basis for designing specific tasks for the course syllabus in LSP courses (Boers, 2021). Finally, the results show that neither of the two groups is more effective at

acquiring knowledge than the other, as the effect persists even after the experimental and control groups have been switched. A distortion of the test results in the sense that the better students happen to be in the same group can therefore be ruled out.

9.2. Learning impact of tasks in the C-Test

As results of the C-Tests show (Table 4), the experimental group (be it G1 or G2) always outperformed the control group, i.e. the group that used the app for memorisation of discrete vocabulary outperformed the group that was working with the writing tasks. Some basic didactic considerations can be derived from this:

9.2.1. Efficiency of intentional learning

It has been questioned whether memorising isolated vocabulary can lead to a better general language proficiency (Anthony, 2018; Kühn, 2010; Laufer & Nation, 2012; Liu & Lei, 2020). Especially in the light of a communicative approach, memorisation, list learning and vocabulary cards are dismissed as outdated or obsolete (Laufer, 2009; Nation & Meara, 2020; Read, 2004). In the present study, general language proficiency is measured using the C-test, a sufficiently verified and tested instrument for measuring general language proficiency (see above). Results of the C-test reflect higher learning gains of the group that memorised discrete vocabulary during the learning phase, indicating that memorising isolated words does indeed have a measurable impact on general language proficiency. However, this should not lead to the conclusion that memorisation is the best or only approach. Rather, the results must be interpreted in the context of the study. For one thing, the respective tasks (app or writing tasks) are carried out outside of class time and are intended to support the contents taught during class time, where no attention is given to the memorisation of isolated words. Foremost, therefore, the results indicate that in this particular context, intentional learning tasks (pure Focus on Forms) prove to be comparatively more effective than the incidental learning tasks (Task embedded Focus on Form) in terms of general language proficiency. Similar results are reported in comparable empirical studies (Hamrich, 2022, 2025), and they are also consistent with prominent approaches of teaching vocabulary, which point out that vocabulary learning should consist of different activities, including intentional learning (Laufer & Nation, 2012; Nation, 2013, 2020).

9.2.2. The influence of language level

One of the aspects that, in our opinion, explains the comparatively higher effectiveness of intentional learning is the actual language level of the students in the present study. The course is aimed at beginners with no previous language skills. At this level, learning words may prove particularly difficult due to the high degree of novelty in the various aspects that constitute a word (e.g. phonetics, orthography, etc.) (see table of Nation, . At the same time, it is assumed that the capacity of learners to simultaneously process L2 meaning and form is limited (Laufer, 2009, 2010; VanPatten, 1990). Isolated learning, in contrast, can offer the novice learner the opportunity to concentrate only on certain aspects and thus facilitate learning (Barcroft, 2002, 2020; Schmitt, 2008). Studies that suggest that isolated learning is particularly beneficial for beginners seem to underline this (Coady, 1993; Hulstijn, 2001; Schuetze, 2017). Against this background and in terms of transferability to other contexts, it would be interesting to conduct a similar study in language classes with students of a higher language level to confirm or refute the results obtained here.

9.2.3. Vocabulary learning burden in LSP courses

Taking into account the observation that incidental learning can only make a small contribution to the learning of words in purely quantitative terms (Laufer, 2005a, 2020), intentional vocabulary learning by means of memorising discrete words as an out of class activity as presented here in the study can contribute to the effective learning of larger amounts of vocabulary, with the described impacts in terms of general language proficiency and sustainability (see below). For LSP courses, which usually involve large amounts of vocabulary, it therefore represents a viable option for supporting learning processes.

9.3. Learning impact of tasks in the Delayed Tests

The delayed tests were carried out 8 weeks after the corresponding original tests (see Table 1). A comparison of the mean values of the two groups in the Delayed Tests (DT1 & DT2) shows that the learning effect is maintained over time across all 3 test forms (see Tables 2, 3 & 4). While G1 achieves higher scores in both Delayed Vocabulary Tests (DVT1 & DVT2) and both Delayed C-Tests (DCT1 & DCT2), G2 does so in both Delayed Writing Tests (DWT1 & DWT2). The results of the Delayed Tests, therefore, reflect

the results from both previous Tests 1 and 2 (in which G1 was the experimental group and G2 the control group), i.e. a higher score was achieved in the test part that was previously trained by the corresponding tasks in each respective case. Hence, these results show that engaging with the tasks has a sustainable learning gain, which is still visible after the above-mentioned period.

Nevertheless, the picture is somewhat more nuanced regarding statistical significance and effect size and therefore requires a more careful analysis. Regarding the VT and the WT, it can be observed that although there are higher means of G1 (see Table 2), differences are no longer statistically significant (see VT1 vs DVT1 and VT2 vs DVT2) and the effect size is also partially reduced (see VT1 vs DVT1, WT1 vs DWT1 and WT2 vs DWT2). Overall, it seems that the differences in scores of the two groups become smaller in the two delayed subtests (vocabulary and writing test). If one takes a closer look at the results in the DVTs, it can be deduced from the means (Table 2) that there is a decrease in G1, while G2 was able to improve the result. In our view, this could be due to the fact that G1 was able to memorise a large number of words during the first 2 learning phases and was also able to recall them (passive & active recall) in VT1 & VT2. In the following 8 weeks before the respective DVTs, however, these words were not repeated regularly using the app; instead, focus changed to the writing tasks to be resolved by G1 in learning phase 3 & 4. Vocabulary from learning phase 1 & 2 was thus partially forgotten. This would be consistent with various learning theories that stress the importance of repetition (Barcroft, 2016; Boers & Lindstromberg, 2008; Nation, 2013). In contrast, the control group was able to expand their vocabulary somewhat, as classes continued, and certain vocabulary appeared repeatedly in class activities and could therefore be picked up by G2. With regard to the DWTs, a very similar picture can be derived on the basis of the means achieved (Table 3). Competences acquired by each group also appear to converge during the 8 weeks between WT 1 & 2 and the two DWTs. While G1 improved its results in the DWTs, G2 achieved slightly lower scores (exception: WT2 vs DWT2). Here too, we attribute this effect to the circumstance that, on the one hand, G1 was able to acquire certain relevant knowledge through general in-class activities and was therefore able to improve its results in the two DWTs, whereas, on the other hand, specific knowledge acquired by G2 during learning phase 1 & 2, which had led to the better results in WT1 & WT2, was not repeatedly practised and therefore forgotten.

Strikingly, the picture is different with regard to the C-Tests. Here, the differences between G1 and G2 measured in CT1 and CT2 also remain in the delayed tests (DCT1 & DCT2), both in terms of statistical significance and effect size; the difference even increases in the case of DCT1 (see Table 4). The competences measured in this test format (general language proficiency), therefore, did not converge between the two groups, as they did in the VTs and WTs, but instead G1 continues to perform (statistically significantly) better than G2. One possible explanation might lie in the format of the two test formats, VT and CT. The C-Test measures general language proficiency, as a series of competences are necessary for its processing that go beyond pure vocabulary knowledge (Read 2000), which is supported by a number of empirical studies that have shown a high correlation between the C-test and different test formats (e.g. writing, listening, etc.) (Eckerth & Tschirner, 2010). Nevertheless, there are studies that indicate that the C-test has particularly high correlations with tests that primarily measure vocabulary knowledge (Daller & Xue, 2007; Daller & Xue, 2009). From the test results presented here, we assume that on the one hand, G1 had acquired solid vocabulary knowledge using the app during the learning phases 1 & 2, as shown in VT 1 & VT2. Nevertheless, some of this knowledge got lost due to attrition (Schmid, 2011), as can be observed in DVT1 and DVT2. In the vocabulary tests, words had to be reproduced through translation without the support of context or any other kind of clue, contrary to what is the case with the C-Test, where context (in the form of phrases and text) and clues (50% of the word) are provided. We assume that it was the additional support provided in the C-Test that made it easier for the students to recall vocabulary and successfully complete the C-Test 8 weeks after memorising took place during the learning phase.

10. Conclusion

The present work is a longitudinal study, which compares the impact of two different learning settings (intentional versus incidental, RQ1) on learning specialised vocabulary in a German Business course at university level. It has been shown that memorisation of discrete vocabulary during didactic units has a positive effect both on vocabulary knowledge and on general language skills (RQ2). Learning gains are also present after switching the experimental and control group, and they are still visible after a period of eight weeks, as a delayed test showed (RQ3). We believe that the findings of this study will

contribute to a better understanding of the field of vocabulary acquisition at a beginner's level in LSP courses and, therefore, allow didactic conclusions to be drawn with a view to a sound and validated design of specific tasks.

Nevertheless, the number of individuals that make up the study (N=20) makes it necessary to repeat and further investigate vocabulary acquisition in similar settings. One specific issue that arises from the results is whether they would be similar in a setting with advanced instead of beginner learners. Bearing in mind the aforementioned limitations, in our opinion, these results are transferable to similar settings to the one described in the present study.

Acknowledgements

This publication is part of the R&D&I project PID2023-151333NB-I00 “CLAVDAP-MULTLEX: Verb classes in (semi-)academic and professional discourse: multidisciplinary perspectives on German and Spanish lexicon”, funded by MICIU/AEI/10.13039/501100011033/ and “FEDER/EU”.

Article history:

Received 16 December 2024

Received in revised form 24 June 2025

Accepted 05 September 2025

References

- Anthony, L. (2018). *Introducing English for Specific Purposes*. Routledge.
- Bachmann, L. F., & Palmer, S. (1996). *Language testing in practice*. Oxford University Press.
- Barcroft, J. (2002). Semantic and structural elaboration in L2 lexical acquisition. *Language Learning*, 52(2), 323-363. <https://doi.org/10.1111/0023-8333.00186>
- Barcroft, J. (2015). *Lexical input processing and vocabulary learning*. John Benjamins.
- Barcroft, J. (2016). *Vocabulary in language teaching*. Routledge.
- Barcroft, J. (2020). Key issues in teaching single words. In S. A. Webb (Ed.), *The Routledge handbook of vocabulary studies* (pp. 479-492). Routledge.
- Boers, F. (2021). *Evaluating second language vocabulary and grammar instruction: A synthesis of the research on teaching words, phrases, and patterns*. Routledge.
- Boers, F., & Lindstromberg, S. (2008). How cognitive linguistics can foster effective vocabulary teaching. In F. Boers & S. Lindstromberg (Eds.), *Cognitive linguistic approaches to teaching vocabulary and phraseology* (pp. 1-64). Mouton de Gruyter.
- Breindl, E. (2021). Forschungsansätze der Linguistik des Deutschen als Fremd- und Zweitsprache. In C. Altmayer, K. Biebighäuser, S. Haberzettl, & A. Heine (Eds.), *Handbuch Deutsch als Fremd- und Zweitsprache: Kontexte – Themen – Methoden* (pp. 105-147). J.B. Metzler.
- Buhlmann, R., & Fearn, A. (2000). *Handbuch des Fachsprachenunterrichts: Unter besonderer Berücksichtigung naturwissenschaftlich-technischer Fachsprachen*. Gunter Narr Verlag.
- Cantera Adrados, S. (2010). *Modelos del C-Test y su correlación con los tests de la comprensión lectora de la E.O.I.* PhD dissertation. Universidad

Complutense de Madrid.

Chacón-Beltrán, R., Abello-Contesse, C., & Torreblanca-López, M. del M. (2010). Vocabulary teaching and learning: Introduction and overview. In R. Chacón-Beltrán, C. Abello-Contesse & M. del M. Torreblanca-López (Eds.), *Second language acquisition. Insights into non-native vocabulary teaching and learning* (pp. 1-14). Multilingual Matters.

Charles, M., & Pecorari, D. (2016). *Introducing English for Academic Purposes*. Routledge.

Chung, T. M., & Nation, I. S. P. (2003). Technical vocabulary in specialised texts. *Reading in a Foreign Language*, 15(2), 103-116. <https://doi.org/10.64152/10125/66770>

Coady, J. (1993). Research on ESL/EFL vocabulary acquisition: Putting it in context. In T. Huckin, M. Haynes, & J. Coady (Eds.), *Second language reading and vocabulary learning* (pp. 3-23). Ablex.

Cohen, J. (1988). *Statistical power analysis for the behavioural sciences*. Erlbaum.

Coxhead, A. (2000). A new academic word list. *TESOL Quarterly*, 34(2), 213. <https://doi.org/10.2307/3587951>

Coxhead, A. (2013). Vocabulary and ESP. In B. Paltridge, & S. Starfield (Eds.), *The Handbook of English for Specific Purposes* (pp. 115-132). Wiley.

Coxhead, A. (2016). Acquiring academic and disciplinary vocabulary. In K. Hyland, & P. Shaw (Eds.), *The Routledge Handbook of English for Academic Purposes* (pp. 177-190). Routledge.

Coxhead, A. (2018). *Vocabulary and English for Specific Purposes Research: Quantitative and qualitative perspectives*. Routledge.

Coxhead, A., Demecheleer, M., & McLaughlin, E. (2016). The technical vocabulary of Carpentry: Loads, lists and bearings. *TESOLANZ Journal*, 24, 38-71.

Daller, H., & Xue, H. (2007). Lexical richness and the oral proficiency of Chinese EFL students. A comparison of different measures. In H. Daller, J. Milton, & J. Treffers-Daller (Eds.), *Modelling and assessing vocabulary knowledge* (pp. 150-164). Cambridge University Press.

Daller, M., & Xue, H. (2009). Vocabulary knowledge and academic success: A study of Chinese students in UK higher education. In B. J. Richards (Ed.), *Vocabulary studies in first and second language acquisition: The interface between theory and application* (pp. 179-193). Palgrave Macmillan.

de Winter, J. (2013). Using the Student's t-test with extremely small sample sizes. *Practical Assessment. Research & Evaluation*, 18(10), 1-12.

Eckerth, J., & Tschirner, E. (2010). Review of recent research (2002–2009) on applied linguistics and language teaching with specific reference to L2 German (part 2). *Language Teaching*, 43(1), 38-65. <https://doi.org/10.1017/S0261444809990255>

Eckes, T., & Grotjahn, R. (2006). A closer look at the construct validity of C-tests. *Language Testing*, 23(3), 290-325. <https://doi.org/10.1191/0265532206lt3300a>

Evans, S., & Morrison, B. (2011). The first term at university: Implications for EAP. *ELT Journal*, 65(4), 387-397. <https://doi.org/10.1093/elt/ccq072>

Flowerdew, L. (2015). Corpus-based research and pedagogy in EAP: From lexis to genre. *Language Teaching*, 48(1), 99-116. <https://doi.org/10.1017/S0261444813000037>

Gignac, G. E., & Szodorai, E. T. (2016). Effect size guidelines for individual differences researchers. *Personality and Individual Differences*, 102(7), 74-78. <https://doi.org/10.1016/j.paid.2016.06.069>

Grotjahn, R., Klein-Braley, C., & Raatz, U. (2002). C-Test: an overview. In J. A. Coleman, R. Grotjahn, & U. Raatz (Eds.), *University language testing and the C-test* (pp. 91-114). AKS-Verlag.

Hacking, J. F., Rubio, F., & Tschirner, E. (2018). Vocabulary size, reading proficiency and curricular design: The case of college Chinese, Russian and Spanish. In P. Winke, & S. M. Gass (Eds.), *Foreign language proficiency in higher education* (pp. 25-44). Springer. https://doi.org/10.1007/978-3-030-01006-5_3

Hammrich, T. (2022). El desarrollo de la competencia léxica: la eficacia de un aprendizaje intencional versus un aprendizaje incidental. *Revista Internacional de Lenguas Extranjeras / International Journal of Foreign Languages*, 18, 36-54. <https://doi.org/10.17345/rile18.3501>

Hammrich, T. (2025). La memorización de vocabulario y el desarrollo de la competencia lingüística general: Rendimiento, preferencias de aprendizaje y percepción de los estudiantes. *Revista Española de Lingüística Aplicada/Spanish Journal of Applied Linguistics*, 38(1), 252-282. <https://doi.org/10.1075/resla.22053.ham>

Hoffmann, L. (1987). Research on Languages for Special/Specific Purposes. In U. Ammon, N. Dittmar, & K. J. Mattheier (Eds.), *Sociolinguistics* (pp. 653-660). De Gruyter.

- Hulstijn, J. (2001). Intentional and incidental second language vocabulary learning: A reappraisal of elaboration, rehearsal and automaticity. In P. Robinson (Ed.), *Cambridge applied linguistics. Cognition and second language instruction* (pp. 258-286). Cambridge University Press. <https://doi.org/10.1017/CBO9781139524780.011>
- Knoch, U. (2014). Using subject specialists to validate an ESP rating scale: The case of the International Civil Aviation Organisation (ICAO) rating scale. *English for Specific Purposes*, 33, 77-89. <https://doi.org/10.1016/j.esp.2013.08.002>
- Kühn, P. (2010). Wörterbücher/ Lernerwörterbücher. In H.-J. Krumm, C. Fandrych, B. Hufeisen, & C. Riemer (Eds.), *Deutsch als Fremd- und Zweitsprache* (pp. 304-315). De Gruyter Mouton.
- Laufer, B. (2005a). Focus on form in second language vocabulary learning. In S. H. Foster-Cohen, M. del P. García Mayo, & J. Cenoz (Eds.), *EUROSLA yearbook* (pp. 223-250). John Benjamins.
- Laufer, B. (2005b). Instructed second language vocabulary learning: the fault in the 'default hypothesis'. In A. Housen, & M. Pierrard (Eds.), *Investigations in instructed second language acquisition* (pp. 311-332). Mouton de Gruyter.
- Laufer, B. (2009). Second language vocabulary acquisition from language input and from form-focused activities. *Language Teaching*, 42(3), 341-354. <https://doi.org/10.1017/S0261444809005771>
- Laufer, B. (2010). Form focused instruction in second language vocabulary learning. In R. Chacón-Beltrán, C. Abello-Contesse, & M. del M. Torreblanca-López (Eds.), *Second language acquisition. Insights into non-native vocabulary teaching and learning* (pp. 15-27). Multilingual Matters.
- Laufer, B. (2020). Evaluating exercises for learning vocabulary. In S. Webb (Ed.), *The Routledge handbook of vocabulary studies* (pp. 351-369). Routledge.
- Laufer, B., & Goldstein, Z. (2004). Testing vocabulary knowledge: Size, strength, and computer adaptiveness. *Language Learning*, 54(3), 399-436. <https://doi.org/10.1111/j.0023-8333.2004.00260.x>
- Laufer, B., & Nation, P. (2012). Vocabulary. In S. M. Gass, & A. Mackey (Eds.), *The Routledge handbook of second language acquisition* (pp. 163-176). Routledge.
- Laufer, B., Elder, C., Hill, K., & Congdon, P. (2004). Size and strength: do we need both to measure vocabulary knowledge? *Language Testing*, 21(2), 202-226. <https://doi.org/10.1191/0265532204lt277oa>
- Lindstromberg, S. (2020). Intentional L2 vocabulary learning. In S. Webb (Ed.), *The Routledge handbook of vocabulary studies* (pp. 240-254). Routledge.
- Liu, D., & Lei, L. (2020). Technical vocabulary. In S. Webb (Ed.), *The Routledge Handbook of Vocabulary Studies* (pp. 111-125). Routledge.
- Nation, I.S.P. (2013). *Learning vocabulary in another language* (2nd ed.). Cambridge University Press.
- Nation, I.S.P. (2020). The different aspects of vocabulary knowledge. In S. Webb (Ed.), *The Routledge handbook of vocabulary studies* (pp. 15-29). Routledge.
- Nation, I.S.P., & Meara, P. (2020). Vocabulary. In N. Schmitt & M. P. H. Rodgers (Eds.), *An introduction to applied linguistics* (3rd ed., pp. 35-54). Routledge.
- Nekrasova, Beker, T., Becker, A., & Sharpe, A. (2019). Identifying and teaching target vocabulary in an ESP course. *TESOL Journal*, 10(1), e00365. <https://doi.org/10.1002/tesj.365>
- Read, J. (2000). *Assessing vocabulary*. Cambridge University Press.
- Read, J. (2004). Research in teaching vocabulary. *Annual Review of Applied Linguistics*, 24, 146-161. <https://doi.org/10.1017/S0267190504000078>
- Reynolds, B. L., Zhang, X., & Ding, C. (2021). A mixed-methods study of English vocabulary for medical purposes: medical students' needs, difficulties, and strategies. *Applied Linguistics Review*, 14(3), 643-678. <https://doi.org/10.1515/applirev-2020-0119>
- Rhiel, G., & Wilkie, W. (1996). An investigation of the large-sample/small-sample approach to the one-sample test for a mean (sigma unknown). *Journal of Statistics Education*, 4(3), 1-7. <https://doi.org/10.1080/10691898.1996.11910515>
- Roelcke, T. (2015). Besondere Wörter II: Fachwörter, Termini. In U. Haß, & P. Storjohann (Eds.), *Handbücher Sprachwissen: Bd. 3. Handbuch Wort und Wortschatz* (pp. 371-393). De Gruyter.
- Schmid, M. (2011). *Language attrition*. Cambridge University Press.
- Schmitt, N. (2008). Review article: Instructed second language vocabulary learning. *Language Teaching Research*, 12(3), 329-363. <https://doi.org/10.1177/1362168808089921>

- Schuetze, U. (2017). Efficiency in second language vocabulary learning. *Die Unterrichtspraxis/Teaching German*, 50(1), 22-31. <https://doi.org/10.1111/tger.12019>
- VanPatten, B. (1990). Attending to form and content in the input. *Studies in Second Language Acquisition*, 12(3), 287-301.
- Webb, S. (2009). The effects of pre-learning vocabulary on reading comprehension and writing. *Canadian Modern Language Review*, 65(3), 441-470. <https://doi.org/10.3138/cmlr.65.3.441>
- Webb, S. (2020). Incidental vocabulary learning. In S. A. Webb (Ed.), *The Routledge handbook of vocabulary studies* (pp. 225-239). Routledge.
- Webb, S., & Nation, I.S.P. (2017). *How vocabulary is learned. Oxford handbooks for language teachers*. Oxford University Press.
- Woodrow, L. (2022). *Introducing researching English for Specific Purposes*. Routledge.
- Zhu, X. (2021). Sample size calculation for Mann-Whitney U test with five methods. *International Journal of Clinical Trials*, 8(3), 184-195. <https://doi.org/10.18203/2349-3259.ijct20212840>

Tim Hammrich is an Associate Professor in the Department of English, French and German Philology at the University of Oviedo. His main research interests include foreign language learning, computer-assisted language learning, and learning analytics. In these fields, he has delivered numerous presentations at international conferences and published articles in peer-reviewed journals.

NOTES

¹ A detailed description of the C-Test is provided in Section 6.

