



HAL
open science

A systematic literature review of Technology-Assisted Vocabulary Learning

Enzo Simonnet, Mathieu Loiseau, Elise Lavoué

► **To cite this version:**

Enzo Simonnet, Mathieu Loiseau, Elise Lavoué. A systematic literature review of Technology-Assisted Vocabulary Learning. *Journal of Computer Assisted Learning*, In press, 10.1111/jcal.13096 . hal-04795580

HAL Id: hal-04795580

<https://hal.science/hal-04795580v1>

Submitted on 28 Nov 2024

HAL is a multi-disciplinary open access archive for the deposit and dissemination of scientific research documents, whether they are published or not. The documents may come from teaching and research institutions in France or abroad, or from public or private research centers.

L'archive ouverte pluridisciplinaire **HAL**, est destinée au dépôt et à la diffusion de documents scientifiques de niveau recherche, publiés ou non, émanant des établissements d'enseignement et de recherche français ou étrangers, des laboratoires publics ou privés.



Distributed under a Creative Commons Attribution - NonCommercial - NoDerivatives 4.0 International License

A systematic literature review of Technology-Assisted Vocabulary Learning

Enzo Simonnet^{1,2}[0009-0000-9740-5212], Mathieu Loiseau¹[0000-0002-9908-0770], and Élise Lavoué^{1,2}[0000-0002-2659-6231]

¹ Université Jean Moulin Lyon 3, iaelyon school of Management, Lyon, France

² Univ Lyon, INSA Lyon, CNRS, UCBL, LIRIS, UMR5205, F-69621, Villeurbanne, France
enzo.simonnet@liris.cnrs.fr, mathieu.loiseau@insa-lyon.fr, elise.lavoue@univ-lyon3.fr

Abstract. Background: Vocabulary learning is an essential dimension of foreign language learning. The learners have a huge responsibility in this task. Among other factors, time constraints and an overall lack of maintained motivation on the part of students makes it especially challenging. Consequently, improving vocabulary learning requires strategies to motivate learners and encourage autonomous learning. Technology-Assisted Vocabulary Learning (TAVL) aims to provide solutions to this set of problems.

Objectives: We aim to provide an overview of the different types of TAVL tools developed and studied since 2015 with regard to (1) tool features, (2) observed variables and research methods used in the studies, and (3) impact on learners' vocabulary acquisition, motivation, and engagement.

Method: We conducted a systematic review of the literature between 2015 and 2023 on TAVL. After delineating our field of study and research query, we followed a rigorous methodology to extract relevant papers to answer our research questions. A total of 44 papers of different languages and academic levels were selected for this review.

Results and conclusions: We report a wide variety of features provided by TAVL tools in various contexts. We observe an overall positive influence of TAVL tools on learners' vocabulary learning, motivation, flow and engagement. We also discuss potential biases in the studies, which do not allow generalization of the results. We conclude with recommendations for developers and researchers, including the use of more standardized experimental protocols with comparable measures, the integration of collaborative and social features, and a more central role for teachers and academic learning contexts.

Keywords: Vocabulary learning · Motivation · Engagement · Flow · Literature review

Lay summary

– *What is currently known about this topic?*

- Many reviews on Technology-Assisted Vocabulary Learning (TAVL) highlighted the effectiveness of technology-assisted approaches in enhancing vocabulary learning for second language learners.
- Various types of digital tools are used for vocabulary learning (AI-based tools, mobile apps, games, VR and AR tools among others) with different impacts on learners' motivation, attitudes and perceptions.
- Several key factors influence the effectiveness of TAVL, such as the type of device, the use of game features, and the learning settings.

- TAVL are more efficient when they include incidental instruction, when the target language is similar to learners' first language, and when they use mobile devices.
- *What does this paper add?*
 - We provide a detailed description of 42 TAVL tools: target languages, economic model, interaction devices, self-sufficiency, intended uses, linguistic content.
 - We identify the features integrated into the 42 tools for vocabulary retention, learners' assessment, collaboration and gamification.
 - We report a detailed description on 44 papers published between 2015 and 2023 on the impact of TAVL tools on learners' vocabulary acquisition, motivation, flow, and engagement.
 - In line with previous reviews, we highlight the positive impact of TAVL on these dimensions and discuss potential biases and external factors that could impact these outcomes.
- *Implications for practice/or policy*
 - We recommend to include more social and collaborative features into TAVL tools.
 - We suggest to provide more features dedicated to teachers in order to guide and monitor learners, and to integrate the use of TAVL in academic curriculum.
 - We recommend the use of more standardized questionnaires and protocols to obtain generalizable results.
 - We recommend to explore the potential of the analysis of interaction traces to provide insights into the evolution of learners' motivation and behaviours with TAVL over time.

1 Introduction

It is commonly recognised that “learning vocabulary is an essential part of mastering a second language” (Schmitt, 2008, p. 329). This crucial importance has been acknowledged for decades: “no language acquisition [or learning], whether first, second, or foreign; child, or adult, can take place without the acquisition of lexis” (Laufer, 1986, p. 69). Still, owing to the principles of the communicative approach (Hilton, 2008, p. 162), the position of vocabulary learning has not always been in line with this status, neither in France (Hilton, 2002, § 39) nor more globally (Meara, 2002, p. 393). The resurgence of vocabulary learning in academic research (Meara, 2002, p. 393) is in synchronicity with the emergence of various digital tools targeting specifically vocabulary learning.

1.1 Vocabulary learning and “word consciousness”

Vocabulary learning is a complex task that surpasses plain memorisation of the association of a word form in L2 (second language, *i.e.* learned language) to an L1 (native language) counterpart. The lexical competence as characterised by Tremblay and Anctil (2020) involves knowledge, skills and attitude. The elements of vocabulary knowledge are in line with how Nation describes what is involved in learning a word (Nation, 2013, pp. 48–50) and covers various aspects of word knowledge that can be grouped into three categories: form, meaning and use. Form designates word forms and what Tremblay and Anctil (2020) call “implicit knowledge of the lexical system”: organisation of the lexicon, derivation, inflexion. Meaning deals with either associating a concept to a word form (and *vice versa*) but also addresses polysemy and associations to the concept/word form. Finally, “use” — or “combining” for Tremblay and Anctil (2020) — covers grammatical functions, collocations and constraints on use (register, frequency, style, connotations, etc.).

To each of the aforementioned knowledge element, Tremblay and Anctil (2020) associate skills that cover reception and production (again formalising Nation’s description (Nation, 2013)) that interpret knowledge elements in terms of oral and written procedural skills, both considering message reception and utterance production. These skills put in action the knowledge described earlier.

Finally, lexical competence for Tremblay and Anctil (2020) encompasses the dimension of attitudes towards vocabulary learning. They proposed examples of such manifestations and refer to it as “*sensibilité lexicale*”³. It corresponds to the concept of “word consciousness” viewed as an attitude. Word consciousness is defined as “interest and awareness in words” (Scott and Nagy, 2009, p. 127), thus encompasses also a dimension of knowledge. Among the attitudinal manifestations of word consciousness suggested by (Tremblay, 2021, § 20), we can highlight the following attitudes:

- be enthusiastic about learning new words and expressions;
- be motivated to learn new words;
- enjoy sharing their lexical discoveries with others;
- show an interest in learning the meaning of a new word encountered in a text and understanding its subtleties.

These examples underline the role of attitude in vocabulary learning, which is also inherent in integrating that attitude to the overall lexical competence. These examples also show how motivation intertwines with this vocabulary learning, not only in terms of engagement with the task but also in terms of opportunity for practice (*e.g.* sharing with others). In a way these reported attitudes reify the role(s) of motivation in vocabulary learning.

The importance of motivation in vocabulary learning is also supported by the model of motivated vocabulary learning proposed by Tseng and Schmitt (2008). In this model, motivation is described as a dynamic component fluctuating during the course of the learning process rather than a static feature intrinsic to the learner. Therefore, “motivation is not just an ‘initial state’ factor; it is an integral part of the whole system that drives the vocabulary learning cycle along” (Tseng and Schmitt, 2008, p. 383). Indeed, mastering new words often unfolds as a monotonous and repetitive task, which can possibly result in learners losing their motivation (Tseng and Schmitt, 2008, pp. 359–360). Thus, sustaining learners’ motivation and engagement over extended periods can be an interesting leverage point for vocabulary learning progress. Maintaining learner motivation then emerges as a key component for vocabulary acquisition.

In the broader context of second language acquisition, technological tools offer ways and solutions for facing these challenges: grasping the complexity of vocabulary learning, staying motivated, but also fostering student autonomy (Hao et al., 2021).

1.2 TAVL

In this paper, we present a systematic review of the literature between 2015 and 2023 that provides an overview of what we refer to as Technology-Assisted Vocabulary Learning (TAVL)⁴. We followed a rigorous methodology to extract 44 articles relevant to the scope of the study. In section 4, we begin

³ Literally, “lexical sensitivity”.

⁴ Although the acronym TAVL is not widely used, we believe it makes a logical addition to the acronyms Mobile-Assisted Vocabulary Learning (MAVL) (Ma, 2017; Ye et al., 2023) and Computer-Assisted Vocabulary Learning (CAVL) (A. Al-Jasir, 2019), which are too restrictive for our meta-review. It is noteworthy that the expression “Technology-Assisted Vocabulary Learning” has already been used in (Hao et al., 2021).

by developing a detailed description of TAVL tools by classifying them according to their target languages, business model, interaction devices on which they rely, self-sufficiency, intended uses and linguistic content. We also pay particular attention to the different features offered by the tools, distinguishing between features directly dedicated to the acquisition of lexical skills (separated into assessment and reception skills); features specially designed for teachers and other specific additional features (game elements, feedback, use of social media, use of AI). We thus report a wide variety of features provided in different contexts, indicating a very diverse and rich field.

After this broad overview of the field, section 5 takes a closer look at the empirical studies conducted and the tools experimented with. We identify the moderating variables that may influence the effectiveness of TAVL and their measurement using various methods. We then present the detailed results of various types of intervention and observe an overall positive impact of TAVL on vocabulary acquisition, motivation, engagement, flow, attitudes and learners' perceptions of these interventions and tools.

We finally discuss our findings and make some recommendations in the field. Beyond the almost unanimous positive impacts observed in the experiments, we critically examine the potential biases, the external systemic influences that could affect these results and the diversity of the protocols, which make it impossible to generalise the results. We recommend that developers and researchers include the use of more standardised experimental protocols with comparable measurements, the integration of more gamification features into vocabulary learning tools and/or the design of more vocabulary games, the integration of collaborative and social features, and a more central role for teachers and academic learning contexts with the integration of the use of TAVL into the academic curriculum. This information could be used to optimise the design of TAVL tools and contribute to more effective and engaging vocabulary learning.

2 Literature reviews on TAVL

In order to support learners in vocabulary learning tasks, many TAVL tools have been developed and studied by researchers and companies (Wang et al., 2021), as shown by seven literature reviews and meta-analyses published in recent years (Lin and Lin, 2019; Klimova, 2021; Wang et al., 2021; Hao et al., 2021; Zou et al., 2021; Alhuwaydi, 2022; Yu and Trainin, 2021). These include four meta-analyses that have been conducted to shed light on the impact of MAVL in the context of English as a Foreign Language (EFL) (Lin and Lin, 2019; Wang et al., 2021; Klimova, 2021; Alhuwaydi, 2022).

Alhuwaydi (2022) made a review of previous studies on the impact of mobile applications designed for MAVL in the EFL context from 2010 to 2022. The paper examines the research features and impact of these applications and emphasises the role that vocabulary learning tailored applications have played in impacting learners' EFL vocabulary learning process. The analysis showed that specifically designed mobile applications play an important role in boosting learners' EFL vocabulary knowledge, motivation, attitudes towards vocabulary learning, and perceptions. In a meta-analysis conducted between 2005 and 2018, Lin and Lin (2019) examined the effectiveness of MAVL for L2 vocabulary acquisition in 33 studies. They found that MAVL has a positive and large effect size on L2 word retention and they identified four factors that may influence the effect of mobile learning for L2 vocabulary acquisition, including research settings, treatment duration, task-afforded autonomy, and outcome measurements. Wang et al. (2021) reviewed more specifically commercial EFL vocabulary learning apps, which include cognitive approaches, multimedia input, learning materials, and game elements. They conclude that most EFL vocabulary learning apps use

word association to convey word knowledge to their users, primarily because of its learning speed advantages, and suggest that these components are effective in enhancing vocabulary learning. Finally, Klimova (2021) reviewed five research studies on the effectiveness of mobile applications on EFL university learners' vocabulary learning. The results indicate that mobile applications are effective, useful and appropriate learning tools for retaining and practising new vocabulary, when their implementation is guided and controlled. The reviewed mobile apps were mainly used as an additional support to the traditional instruction. These four studies offer significant synthetic and annotated bibliographies related to MAVL, and have examined this field from different perspectives. However, the context of mobile learning presents specifics that do not enable us to cover the whole field of technologies for vocabulary learning.

Another aspect of TAVL has been studied by Zou et al. (2021), who present a systematic review of studies on digital game-based vocabulary learning from five perspectives: a general overview of published studies, digital games for vocabulary learning, theoretical frameworks, research issues and findings, and implications. The authors found that digital games generally demonstrate positive effects in promoting short-term and long-term vocabulary learning, facilitating reading and listening comprehension, increasing motivation and engagement, decreasing anxiety and fostering interactions among learners. This literature review is of high interest because games are becoming increasingly widespread in the field of language learning. We take this aspect into account in our literature review, but extend it to other approaches and functionalities.

Finally, two studies have taken a more global approach to the field of TAVL. Hao et al. (2021) made a meta-analysis of 45 studies conducted between 2012 and 2018 on TAVL for EFL learners. The overall effect was found to be large compared to traditional instructional methods, indicating that technology-assisted learning was more beneficial than non-technology-assisted instruction and could enhance learners' long-term vocabulary retention. The effectiveness was found to be affected by several variables, including device type, game condition, setting, test format, and reported reliability. On a similar scope, Yu and Trainin (2021) conducted a meta-analysis to examine the effectiveness of second language TAVL on L2 vocabulary learning and identified factors that may play a role in their effectiveness. The study found 34 studies with 2,511 participants yielding 49 separate effect sizes, the overall effect size being moderate. The results of the moderator analysis revealed that learners benefited more from TAVL with incidental teaching than with intentional teaching; assessment types were not significant moderators of the effect on TAVL; TAVL is more effective when the target language is close to the learner's first language; higher education students benefited more from TAVL than K-12 students; and finally, MAVL was more effective than CAVL.

In summary, the previous meta-reviews provided valuable insights into the field of TAVL and MAVL. They highlighted the effectiveness of technology-assisted approaches in enhancing vocabulary learning for second language learners, especially in the context of EFL, and emphasised their impact on vocabulary retention, motivation, attitudes, and perceptions. They identified several key variables influencing the effectiveness of TAVL and MAVL, such as the type of device used, the game conditions, and the learning settings. Additionally, they showed that TAVL is more effective when it includes incidental instruction, when the target language is similar to the learner's first language, and when mobile devices are used. However, five of the reviews and meta-analysis focus on specific technologies and approaches, e.g. mobile technologies and game-based learning, which makes it difficult to generalise the results. Furthermore, the two more general literature reviews focused specifically on the impact of the use of technologies on specific aspects of vocabulary learning, such as efficiency and retention, but were not interested in the characteristics and functionalities of these

technologies. In addition, these reviews primarily covered studies conducted up to 2018, whereas other technologies have emerged in recent years, such as augmented and virtual reality.

3 Method

Relying on Chong and Plonsky (2024)’s typology of literature reviews, we opted for the ”Research-focused systematic literature review” methodology as described in their article. A systematic literature review needs to follow a method that analyses the literature available of a specific phenomenon of interest, focusing on providing a background for new studies, identifying gaps for further investigation, and summarising evidence concerning a technology (Kitchenham and Charters, 2007). The methodology used to retrieve the studies fitting our scope, selecting and excluding respectively relevant and irrelevant papers, is based on the PRISMA guidelines (Page et al., 2021). The next steps of the systematic review process, namely extracting and analysing the data from the papers, were based on the guidelines and processes described in (Kitchenham and Charters, 2007).

First, we defined our review scope by specifying a search query that we estimated to be the most able to retrieve all the literature corresponding to our research questions (we explain this in more details in section 3.2). Then, we ran our search query and filtered the papers that did not fit our review scope (see section 3.3). Following the guidelines previously described by Kitchenham and Charters (2007), we answer our research questions through a defined, reproducible protocol describing: (1) research questions, (2) keywords, (3) search engines, (4) inclusion criteria, and (5) extracted data.

3.1 Research questions

To address the gaps identified in the previous literature reviews, our systematic review aims to provide a comprehensive overview of recent studies on TAVL from 2015 to 2023. We provide a broader view of the field, including the characteristics of the tools, the variables that have been studied, the methods used, and the impact of the tools on all dimensions of the user experience (motivation, engagement, and flow), in addition to vocabulary learning. We more specifically address the following research questions:

- RQ1: What types of TAVL tools have been developed and studied to date?
We describe and classify the tools according to their characteristics and features.
- RQ2: Which variables have been evaluated and through which methods?
We provide an overview of the empirical studies conducted and the tools experimented with, in order to then identify moderating variables that may influence the effectiveness of TAVL.
- RQ3: What are the effects of using TAVL?
We explore the overall effectiveness of various types of interventions for vocabulary learning, motivation, engagement, flow, learning attitudes and learners’ perceptions towards those interventions and tools.

3.2 Definition of our review scope

After defining our three research questions, we then clearly identified our search terms. More particularly, we used the following search query:

(vocabulary OR lexic) AND (learning OR teaching OR training) AND (application OR Web OR environment OR technolog*)*

where * is a wildcard. This query was strategically designed to encompass the range of concepts and themes vital to our research scope. The first segment of the query, "(vocabulary OR lexic*)," was chosen to capture literature pertaining to vocabulary and lexicon-related studies, ensuring inclusiveness in this domain. The subsequent segment, "(learning OR teaching OR training)," was chosen to ensure that the papers we retrieved were relevant to educational contexts. The keywords "learning" and "teaching" are particularly important as they help us capture a wide array of instructional methods and educational approaches. Additionally, the term "training" serves a similar purpose, broadening our scope to encompass the continuing professional development in language-related skills. Lastly, the combination of "application, web, environment, technolog*" in the final segment of the query was chosen to specifically target literature related to technological applications and environments. This set of keywords effectively encompasses the technological tools, platforms, and spaces that are integral to modern educational technology.

3.3 Search engines

The search engines were defined based on other systematic reviews in the field (Mahdi, 2018; Al-huwaydi, 2022) and recommendations from Gusenbauer and Haddaway (2020) on the most suitable academic search systems for systematic reviews methodologies. We selected ACM, IEEE, ScienceDirect and ERIC. Recognized journal in the field of Computer-Assisted Language Learning (CALL) such as ReCALL, CALL, ILLT and CALICO, are indexed in these libraries, ensuring access to most of the specialist literature on the subject. All queries were searched on titles and keywords. We decided to restrict the publication date from 2015 to 2023 in order to prioritise the most recent tools and studies about vocabulary learning.

3.4 Inclusion criteria

After executing the query in the search engines, all 170 papers returned were screened according to the inclusion criteria (IC) to make the results more accurate (see Fig. 1).

- IC1: Non-duplicate papers (i.e., with the same Digital Object Identifier - DOI)
- IC2: Papers written in English
- IC3: Full papers (with six or more pages);
- IC4: Primary studies (i.e., not surveys, meta-analysis, systematic; mappings or reviews);
- IC5: Papers which present one or more technological tool used for vocabulary learning;
- IC6: Papers available for download.

3.5 Data to be extracted

To extract the relevant data from the papers, a lecture grid was created with all the authors, then the main author proceeded to parse all the papers to report the corresponding information. The grid included information such as the name, characteristics and features of the studied tools (RQ1), the experimental protocols, studied variables and their corresponding measuring instruments (RQ2), the results of said experiments (RQ3).

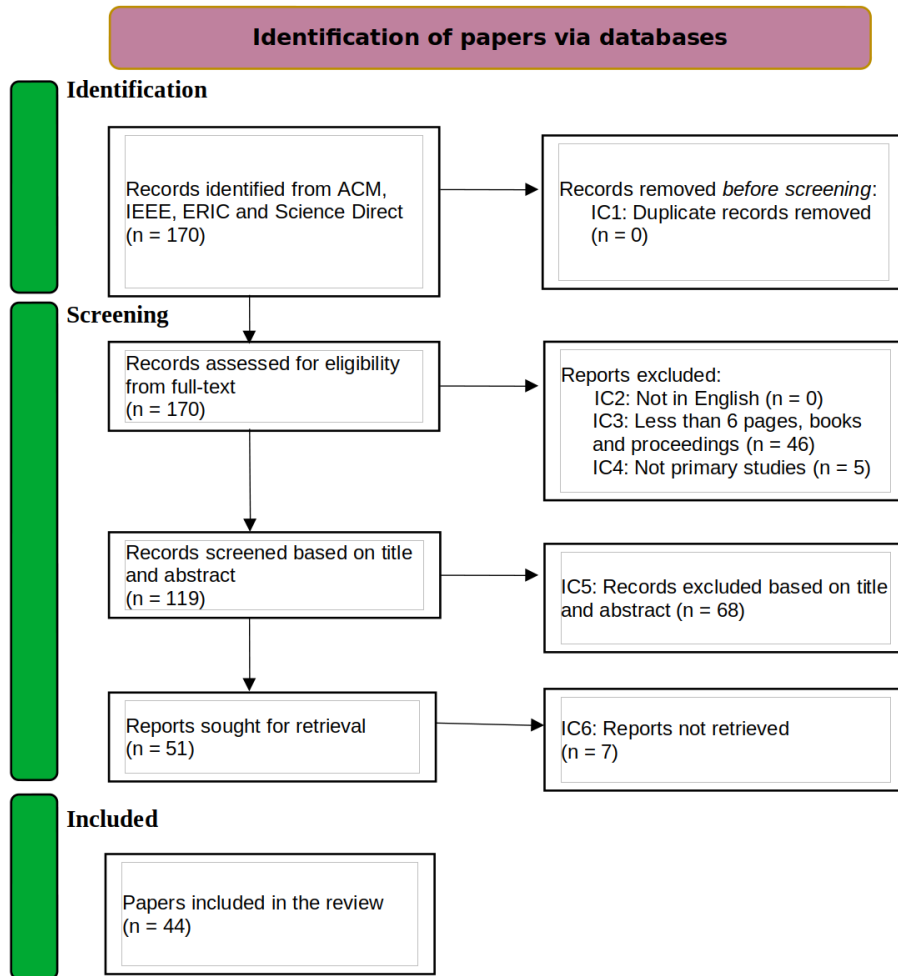


Fig. 1: PRISMA flow diagram (Page et al., 2021)

4 Description of the TAVL tools

Among the 44 papers retrieved, most of them (36) present a unique tool. 2 other papers report studies on the same tool: a commercial messaging application WhatsApp (Basal et al., 2016; Hasan et al., 2022), and a tool named Angličtina Today (Klimova and Polakova, 2020; Pol’akov’a, 2022). The 4 remaining papers present a set of tools used in combination for an activity:

- Alhajaji et al. (2020) uses a combination of a tangible game, a mind-mapping activity and Twitter Hashtags;
- Eren (2015) relies on Facebook, Blogs, Word Exercise and Wiki services;
- Kazu and Kuvvetli (2023) uses ”Games to Learn English” and Google Speech Recognition;

- Wolsey et al. (2015) includes a platform for creating e-dictionary entries, a platform for the e-dictionary itself, and support technology for finding, creating, and embedding images, audio, or video.

We thus examined the characteristics of a total of 42 different tools presented in the selected papers (the features of the 4 set of tools used in combination are examined as if they represent a single tool). In this section, we report descriptive information to give an overview of the different TAVL tools:

- **Economic model:** We made a distinction between tools made available with a commercial purpose, *Commercial Off The Shelf (COTS)*, and the prototypes solely used in research studies (to this point). In table 1, the 23 lines that do not have a “✓” are research prototypes.
- **Interaction devices:** The device for which the tool is intended strongly determines the differences in characteristics and intended use, in the sense described below.
- **Self-sufficiency:** This criteria allows differentiating tools that can be used on their own and constitute in themselves an activity, from tools that require external material in order to be used. Not self-sufficient tools include for instance extensions collecting content on visited web pages.
- **Languages:** We looked at the available languages in order to identify the most represented ones.
- **Intended uses:** We identify tools that are specifically designed for use in the classroom.
- **Linguistic content:** The tools offer different types of media content displayed to convey vocabulary learning.
- **Features:** As a key component of the studied tools, we detail the different features of each tool to highlight the most commonly used, and missing ones.
- **Social interactions:** The degree of collaboration and social interaction supported by the different tools has caught our attention because it can be both a motivating factor (Deci and Ryan, 1985). We isolated three types of “social” setting for the applications: **I**ndividual, **C**ollaborative and **S**ocial.

The information is summed up in Table 1. We precise that we have not tested the tools ourselves, mainly because many of them cannot be found on the Internet, but also to avoid introducing a bias in favour of tools that we know or that we could test.

Table 1: Characteristics of the technology-assisted vocabulary learning (TAVL) tools

Tool name	Presentation							Device				Linguistic content			
	Languages	Language learning	COTS	Gameful	Social ^a			Computer	Mobile phone	AR	VR	Text	Audio	Video	Image
					I	C	S								
Counts	EFL: 29	34	18	17	29	7	6	20	29	3	2	41	28	15	31
First word II Abdul Wahdi and Dzulkifli (2018)	EFL	✓	✓	✓	✓			✓				✓			✓
ThinkEnglish! Ahmad et al. (2015)	EFL	✓				✓			✓			✓	✓		✓
House of Languages Alfadil (2020)	5 languages ^b	✓	✓	✓	✓						✓	✓			✓
GMT (Twitter Hashtags) Alhajaji et al. (2020)	Any		✓	✓		✓		✓	✓			✓			
ICONS simulation platform Angelini and García-Carbonell (2019)	EFL	✓		✓		✓		✓				✓		✓	
QR codes and tablet PCs Arikan and Ozen (2015)	EFL	✓			✓			✓	✓			✓	✓		
WhatsApp Basal et al. (2016); Hasan et al. (2022)	Any		✓			✓			✓			✓	✓		✓
ARLang Caetano et al. (2023)	Portuguese	✓		✓	✓				✓	✓		✓	✓		✓
MOOC 'Empieza con el inglés' Chacón-Beltrán (2018)	EFL		✓			✓		✓				✓	✓		✓
VALRS-VLM Chen et al. (2022)	EFL	✓			✓			✓				✓	✓		✓
My-Pet-Shop Chen and Lee (2018)	EFL	✓		✓				✓				✓	✓		✓
Smart Notebook™ Dennis and Whalon (2021)	EFL	✓	✓		✓			✓				✓	✓		✓
Vocabulary Flashcards 2016 Ebadi and Bashiri (2018)	EFL	✓	✓		✓				✓			✓	✓		
VocabGame Elaish et al. (2019)	EFL	✓		✓	✓				✓			✓			✓
HiroTan Enokida et al. (2018)	EFL	✓		✓		✓		✓				✓	✓		
Web 2.0 tools (Facebook, Blogs, Word Exercise, Wiki services) Eren (2015)	Any		✓			✓		✓	✓			✓			
Memrise Gay (2022)	8 languages ^c	✓	✓	✓	✓			✓	✓			✓	✓		✓
3D Virtual Environment Vocabulary Framework Hameed (2020)	EFL	✓		✓	✓						✓	✓	✓		✓
Detective ABC Hao et al. (2019)	EFL	✓		✓	✓				✓			✓	✓		✓
Wordhyve Hasnine and Wu (2021)	3 languages ^d	✓			✓			✓	✓			✓			✓
Exploring Verbs; Exploring Nouns Huang (2015)	EFL	✓	✓		✓			✓				✓			✓
Nearpod application Huang (2021)	Any		✓					✓				✓			
MAVR Jalaluddin et al. (2021)	EFL	✓		✓	✓				✓	✓		✓			✓
Games to Learn English + Google Speech Recognition Kazu and Kuvvetli (2023)	EFL	✓	✓	✓	✓			✓				✓	✓		✓
WCFLASMA Lian et al. (2017)	EFL				✓			✓				✓	✓		✓
AI-based object detection translation (AI-based ODT) application (app) Liu and Chen (2023)	English, Chinese	✓			✓				✓			✓	✓		✓
Endless Alphabet Loftus-Rattan and Furey (2021)	EFL	✓	✓	✓	✓				✓			✓	✓		✓
electronic books (e-books) Love et al. (2017)	EFL	✓	✓		✓			✓	✓			✓	✓		✓
VLASTWA Mirzaei et al. (2021)	EFL	✓			✓			✓	✓			✓			✓
Telegram Motlagh et al. (2020)	Any		✓			✓			✓			✓	✓		✓
WeChat Pamintuan et al. (2018)	Any		✓			✓			✓			✓	✓		✓
Angliština Today Pol'akov'a (2022); Klimova and Polakova (2020)	EFL	✓			✓				✓			✓	✓		✓
Animated schematic image as visual gloss Sato (2016)	EFL	✓			✓			✓				✓			✓
Photographe Tangworakithaworn et al. (2019)	EFL	✓		✓	✓				✓			✓	✓		✓
Mobile Vocabulary Learning System Uz Bilgin and Tokel (2019)	EFL	✓			✓				✓			✓			✓
Interactive syllable-based English vocabulary learning board (I SEE) Wang and Huang (2017)	EFL	✓				✓			✓			✓	✓		✓
Vocabulary Self-Collection Strategy Plus (VSS+) (many tools) Wolsey et al. (2015)	Any		✓			✓		✓				✓	✓		✓
MyCLOUD Wong et al. (2016)	Chinese	✓				✓			✓			✓	✓		✓
AWL builder Xodabande and Atai (2022)	EFL	✓	✓		✓				✓			✓	✓		✓
FinDo Yamamoto et al. (2019)	EFL	✓			✓				✓			✓			✓
AR-supported educational toys Yilmaz et al. (2022)	EFL	✓		✓	✓				✓	✓		✓	✓		✓
Digital game for EFL vocabulary learning Zhang et al. (2023)	EFL	✓		✓	✓			✓				✓	✓		✓

^a Social interactions: Individual, Collaborative, Social

^b English, French, German, Russian, Spanish

^c English, Korean, Italian, Spanish, Japanese, Chinese, German, Russian, Slovene

^d English, Japanese, Spanish

4.1 Languages

The 42 tools show a limited diversity among taught languages. 31 tools target only one language, which is EFL in 94% of cases (29/31). The other single language tools target Portuguese (Caetano et al., 2023), and Chinese (Wong et al., 2016). This does not come as a surprise since English is currently the most taught and learned language over the world.

4 tools among those designed specifically for vocabulary learning are multilingual (Alfadil, 2020; Hasnine and Wu, 2021; Gay, 2022; Liu and Chen, 2023). The AI-based ODT application (Liu and Chen, 2023), House of Languages (Alfadil, 2020), Memrise (Gay, 2022) and WordHyve (Hasnine and Wu, 2021) offer a wide variety of languages (respectively 5, 8, 3 and 2). All four applications include English, which is completed with Chinese for the AI-based ODT application, Japanese and Spanish for Wordhyve, French, German, Russian and Spanish for House of Languages and Korean, Italian, Spanish, Japanese, Chinese, German, Russian, Slovene in the case of Memrise (see table 1). 7 generic tools, not designed for vocabulary learning (as described in next section), can be used for any language (Wolsey et al., 2015; Eren, 2015; Basal et al., 2016; Pamintuan et al., 2018; Motlagh et al., 2020; Alhajaji et al., 2020; Huang, 2021; Hasan et al., 2022). They can theoretically support any language teaching as they all depend on the content input by the users. Yet, among the 40 tools with EFL as an available language, all of them have been tested for EFL learning in the papers considered in our review.

4.2 Intended Uses

As illustrated in Table 1, we differentiated between language learning tools that can be used on their own and constitute in themselves an activity, and tools that require external lexical material or organisation in order to be used. Tools specifically designed for vocabulary learning are largely predominant as they represent 81% (34) of the studied tools. We can for instance cite First Word II (Abdul Wahdi and Dzulkifli, 2018), VLASTWA (Mirzaei et al., 2021) and VocabGame (Elaish et al., 2019) among others.

Among the remaining tools, 2 of them are intended for learning (Chacón-Beltrán, 2018; Huang, 2021), but not specifically vocabulary, and 6 tools were originally designed for other purposes than learning, such as messaging (Wolsey et al., 2015; Basal et al., 2016; Pamintuan et al., 2018; Motlagh et al., 2020; Alhajaji et al., 2020; Hasan et al., 2022), creating wikis and shared online resources (such as wikis) (Wolsey et al., 2015; Eren, 2015). Those tools therefore require external lexical information in order to be used as vocabulary learning tools. Indeed, WhatsApp (Basal et al., 2016; Hasan et al., 2022), Telegram (Motlagh et al., 2020), Twitter (Alhajaji et al., 2020), and WeChat (Pamintuan et al., 2018) can be used in vocabulary learning activities but an external input in lexical information or activity organisation from the teacher is mandatory. Similarly, the "VSS+strategy" (Wolsey et al., 2015) relies on many tools and serves a vocabulary learning purpose only when their use is organised as part of a vocabulary learning activity. Same goes for the Nearpod application (Huang, 2021), a "Student Response System" that can be used merely as a support tool for a vocabulary learning activity.

It is also interesting to note that many tools are not designed with the intent of being used in a classroom environment. For instance, VR devices (Alfadil, 2020; Hameed, 2020), given the current state of the technology, require an infrastructure that would be complicated to deploy for dozens of students in the same space. Software such as FinDo (Yamamoto et al., 2019) and ARLang (Caetano et al., 2023) are even thought to be exclusively played outside, and require students to stroll or

go from one place to another to discover more vocabulary, which are incompatible with an in-class activity and can only be used as a complement or occasional activity.

4.3 Economic model

We regrouped under the denomination "commercial" or COTS (Commercial Off The Shelf) all the tools that are made available to the large public and have an economic intent. Hence, a tool that we labelled as "commercial" can still be free or partially free to use. This category of commercial tools represents 43% (18) of the studied tools. The remaining 57% (24) are prototypes developed for research purpose. None of the research-based tools can be found on the web, as it is most common that the tools remain at the prototype stage or are not made available to the large public.

4.4 Interaction devices

The vast majority of the vocabulary learning tools are software. Among the 42 tools, 20 are available on computer and 29 on mobile phone. It is also noteworthy that 9 software are mobile and PC compatible (Huang, 2015; Eren, 2015; Arikan and Ozen, 2015; Love et al., 2017; Enokida et al., 2018; Elaish et al., 2019; Alhajaji et al., 2020; Mirzaei et al., 2021; Kazu and Kuvvetli, 2023). To classify a tool as mobile or computer compatible, we resorted to the information provided by the authors, in order to treat all tools identically. This creates a bias for some web platforms and tools (cf. Telegram, WhatsApp), which are classified as "Mobile phone" even though a computer version is available for both tools. This exemplifies the recent importance granted to MAVL.

A new trend emerges with AR that adds digital content on top of the already existing environment and incorporates text, images, videos or 3D objects into a real-world scenario in real time. In our review, 3 studies are based on AR (Jalaluddin et al., 2021; Yilmaz et al., 2022; Caetano et al., 2023). As there are many definitions of AR, we relied on the authors of the papers describing their tools as such. The 3 tools add virtual content on the scanned environment. Sometimes the content is related to the environment, like in ARLang (Caetano et al., 2023) that adds virtual labels to physical objects, or unrelated to the environment, such as the AR-supported educational toys (Yilmaz et al., 2022) that project flashcards and animations on a table. VR technology is also receiving special attention with two studies dedicated to their impact on vocabulary learning (Alfadil, 2020; Hameed, 2020). The tools employ pose tracking and 3D near-eye displays to give the user a feeling of immersion in a simulated virtual world

4.5 Linguistic content

Classroom integration does not only depend on the functionalities of the tools, but also on the content itself. In 62% (26) of the tools, the content is already provided by the software and is integrated into a progression. They have the advantage of guiding learners and teachers according to a predefined and consistent learning path. But this may lack adaptability for experienced teachers as the content might not fit the curriculum. This can impede classroom integration, but does not prevent the tools from being considered as complementary vocabulary learning tools.

Another variable in terms of content is its form. As mentioned in the introduction, the lexical competence integrates many dimensions of each word: written and oral forms, meaning and use. In order to convey these types of information, various media can be mobilised: textual, audio and pictorial (image or video). Diversifying the representation is known to reinforce memorisation (Mayer,

2002). The most natural is text, as the written form of a word is an unavoidable characteristic to learn said word's written form. Every analysed tool presents written vocabulary content, whether it is through lists of words, song lyrics, in game speeches, or plain text. The pronunciation of the words and other vocabulary audio content (speeches, stories, musics (with lyrics)) are made available in 67% (28) of the tools, whereas video is used by 36% (15) of them. Images are also used as a complementary linguistic content for vocabulary learning in 74% (31) of the tools. A large part of the tools uses multi-modal cues: 90% (38) combine at least 2 different media contents, 57% (24) combine 3 and finally 29% (12) use all 4.

4.6 Features

Now that we have presented in general terms the applications available, we delve more precisely into the actual features offered by the tools. We distinguish between the features directly dedicated to acquire lexical skills, separated into assessment and reception skills, and additional specific features (game elements, feedback, use of social media, use of AI and features designed for teachers). The features listed in Table 2 for each tool are not exhaustive, as some of them are tool-specific. For example, the talking head virtual assistant is only found in First Word II (Abdul Wahdi and Dzulkiffi, 2018), the formulating of compound words from images is specific to Photograffe (Tangworakitthaworn et al., 2019), while collaborative multimedia annotation is proper to WCRASMA (Lian et al., 2017).

Table 2: Table of features

Tool name	Counts	Assessment (including production)					Reception		Other features				
		Pronunciation activities	Written activities ^a	Quizzes and tests ^a	Flashcards	Word lists	RP scenario and exploration	Detailed lexical information	Game elements	Feedback	Social media	Use of AI	Teacher features
First word II Abdul Wahdi and Dzul kifli (2018)		3	12	24	7	10	4	27	10	20	8	2	10
ThinkEnglish! Ahmad et al. (2015)	✓		✓					✓	✓				
House of Languages Alfadil (2020)				✓			✓	✓	✓	✓			
GMT (Twitter Hashtags) Alhajaji et al. (2020)			✓								✓		
ICONS simulation platform Angelini and García-Carbonell (2019)							✓			✓			
QR codes and tablet PCs Arikan and Ozen (2015)			✓	✓						✓			
WhatsApp Basal et al. (2016); Hasan et al. (2022)											✓		
ARLang Caetano et al. (2023)				✓				✓	✓				
MOOC 'Empieza con el inglés' Chacón-Beltrán (2018)				✓		✓		✓					✓
VALRS-VLM Chen et al. (2022)						✓		✓					
My-Pet-Shop Chen and Lee (2018)				✓		✓		✓	✓	✓			
Smart Notebook™ Dennis and Whalon (2021)								✓					
Vocabulary Flashcards 2016 Ebadi and Bashiri (2018)				✓	✓	✓		✓	✓				
VocabGame Elaish et al. (2019)			✓	✓				✓	✓				✓
HiroTan Enokida et al. (2018)				✓		✓		✓	✓				✓
Web 2.0 tools (Facebook, Blogs, Word Exercise, Wiki services) Eren (2015)								✓	✓		✓		
Memrise Gay (2022)			✓	✓	✓	✓		✓	✓				✓
3D Virtual Environment Vocabulary Framework Hameed (2020)				✓	✓		✓	✓					
Detective ABC Hao et al. (2019)				✓			✓	✓	✓				
Wordhyve Hasnine and Wu (2021)				✓				✓	✓				
Exploring Verbs; Exploring Nouns Huang (2015)				✓				✓					✓
Nearpod application Huang (2021)									✓				✓
MAVR Jalaluddin et al. (2021)													
Games to Learn English + Google Speech Recognition Kazu and Kuvvetli (2023)	✓			✓		✓		✓	✓			✓	✓
WCRASMA Lian et al. (2017)				✓				✓					
AI-based object detection translation (AI-based ODT) application (app) Liu and Chen (2023)				✓								✓	
Endless Alphabet Loftus-Rattan and Furey (2021)								✓					
electronic books (e-books) Love et al. (2017)								✓					
VLASTWA Mirzaei et al. (2021)				✓				✓					
Telegram Motlagh et al. (2020)									✓				
WeChat Pamintuan et al. (2018)										✓			
Angliĕtina Today Pol'akov'a (2022); Klimova and Polakova (2020)				✓					✓				✓
Animated schematic image as visual gloss Sato (2016)													
Photografie Tangworakitthaworn et al. (2019)				✓				✓	✓				
Mobile Vocabulary Learning System Uz Bilgin and Tokel (2019)				✓				✓					
Interactive syllable-based English vocabulary learning board (I SEE) Wang and Huang (2017)				✓				✓	✓				✓
Vocabulary Self-Collection Strategy Plus (VSS+) (many tools) Wolsey et al. (2015)				✓				✓		✓			
MyCLOUD Wong et al. (2016)				✓				✓		✓			
AWL builder Kodabande and Atai (2022)	✓			✓	✓	✓		✓	✓				✓
FinDo Yamamoto et al. (2019)				✓		✓		✓	✓				
AR-supported educational toys Yilmaz et al. (2022)					✓				✓				
Digital game for EFL vocabulary learning Zhang et al. (2023)				✓	✓		✓	✓	✓				

^a The main difference between the two modalities is the recipient of the assessment (used for auto-assessment in the case of "text activities" vs grading for "quizzes and tests").

79% (33) of the tools have at least one *feature dedicated to assessment* (see Table 2). In this category, we regroup very diverse activities. These activities provide the learner with a score that allow them to measure the difference between their skills and the expected level. This evaluation is not necessarily the core aspect of the activities, often meant to provide practice opportunities and knowledge expansion. The most popular feature included in 57% (24) of the tools, named *Quizzes and Tests*, regroups a wide variety of different evaluative vocabulary exercises. They could also fall under the denomination of *Written activities*, but the pedagogical objective of tests and evaluations makes it important to be distinguished. In this case, the evaluation is external and the assessment function is more central in the perception of the users.

The category *Written activities*, (*i.e.* fill-in the gaps, find the correct word, pair the word and its translation, put the word in its correct form), allows to work on form and meaning, sometimes integrating a use dimension, and covers 29% (12) of the tools.

By adding *Pronunciation activities* to the previous two categories, we cover most, if not all⁵, the productive vocabulary skills trained by the different tools.

Flashcards and *wordlists* are two reflexive tools. The latter is more explicitly associated with creating a set of objectives for the learning task.

While the categorisation in terms of reception *vs* production is difficult to evaluate in the aforementioned “assessment activities”, certain features explicitly focus on *receptive skills*, which we categorize as follows:

- *Role playing scenario and exploration*: scenarios in which two or more people act out in specific roles. In the studied tools, this occurs through dialogues between the player and non-playing characters. In this case, the production part consists in choosing a predefined utterance, thus not qualifying as a true production task. Exploration relates to the nature of the game, which gives the players the possibility to navigate through a virtual world.
- *Detailed lexical information* about the words: information that can typically be found in a dictionary such as definitions, examples, and part of speech.

61% (23) of the tools include at least one of these solely receptive skill features.

Tseng and Schmitt (2008) insist on the crucial role of motivation in language learning; it is therefore important to examine the *features that can motivate learners* during the vocabulary learning tasks. In proportion, 45% (17) of the tools are categorised as “gameful”, meaning that the software are games or use game elements. It illustrates that games and playfulness are rather popular vocabulary learning strategies currently studied by many researchers in the domain. The *Role playing scenarios and exploration* category regroups 5 games (Hao et al., 2019; Angelini and García-Carbonell, 2019; Alfadil, 2020; Hameed, 2020; Zhang et al., 2023), in which the players make decisions within a pre-defined narrative or explore the virtual environment of the game. *Virtual rewards* are used in 10 (24%) different software. They can take various forms (various gamified rewards (Abdul Wahdi and Dzulkifli, 2018; Elaish et al., 2019; Hao et al., 2019; Alfadil, 2020), points and scores (Enokida et al., 2018; Hao et al., 2019; Tangworakitthaworn et al., 2019; Uz Bilgin and Tokel, 2019; Gay, 2022; Zhang et al., 2023; Kazu and Kuvvetli, 2023; Caetano et al., 2023), badges (Elaish et al., 2019; Kazu and Kuvvetli, 2023; Zhang et al., 2023), timer (Wang and Huang, 2017), in-game

⁵ When a learner reads the signification side of a flashcard, evaluating whether they no the signifier, the task leans on the productive side. Likewise, a wordlist, when used as a constraint for a textwriting activity, targets production.

leveling (Alfadil, 2020)) and are used to reward the learners' actions that benefit their learning.

Broadly defined, *feedback* is “information given to students about their performance that guides future behavior” (Ambrose et al., 2010). Feedback can help to direct learners' attention to areas of improvement, and to connect them with future learning opportunities. It is provided to learners in 48% (20) of the tools and is exclusively a backward-looking and corrective type of feedback. In practice, it consists in saying if the learner's answer is right or wrong with sometimes explanations. While most *written activities*, *tests* and *quizzes* provide a score, they do not necessarily provide a question per question feedback.

The *Social media* category regroups the interactive technologies that facilitate the creation, sharing and aggregation of content, ideas, interests, and other forms of expression through virtual communities and networks, as defined by Kietzmann et al. (2011). It includes well known social media such as WhatsApp (Basal et al., 2016; Hasan et al., 2022), Telegram (Motlagh et al., 2020), Twitter (Alhajaji et al., 2020), WeChat (Pamintuan et al., 2018), but also MyCLOUDNet, described by the authors as “a social network for students to share social media, and respond to others' social media” (Wong et al., 2016).

Only two studies investigate *the use of AI tools* (Liu and Chen, 2023; Kazu and Kuvvetli, 2023) for vocabulary learning. It is noticeable that both papers are published in 2023, as the advances in the field of AI is very recent. In the first paper, the “Google Speech Recognition” tool is used to monitor and correct the students' pronunciation (Kazu and Kuvvetli, 2023). In the second paper, (Liu and Chen, 2023) use a AI-based object detection system to identify the objects filmed by a mobile phone and to display the corresponding word.

Finally, we count 10 (26%) different tools that have *features explicitly designed for teachers*. Those features can be again separated into two distinct categories. The first one is the possibility to monitor students, track their progress, either directly or via analytics. 6 tools present a monitoring option for teachers, with various levels of details and information (Huang, 2015; Wang and Huang, 2017; Enokida et al., 2018; Pamintuan et al., 2018; Elaish et al., 2019; Xodabande and Atai, 2022). The second type of feature thought for teachers is the ability for them to share their own content and to upload their personal learning material. It differs from the content sharing features of the “Social media” category in that it allows to upload either a course directly, as in the case of MOOC platforms (Chacón-Beltrán, 2018), or specifically vocabulary content in “Games to Learn English” (Kazu and Kuvvetli, 2023) or Memrise (Gay, 2022). The Student response system allows the teacher to interact with their class in an engaging manner (Huang, 2021). Finally, we can highlight the effort of Pol'akov'a (2022) who developed on their application a web interface for the teachers. Each teacher can manage several lessons, register their students, distribute news or alerts through notifications, respond to their comments and visualise the results of all students with a dedicated dashboard.

4.7 Social interactions

We classified the tools into 3 distinct categories according to the level and type of interactions available:

- *Collaborative*: 17% of the tools have features that make collaborative work possible between learners and with the teacher through diverse features. Wikis allow collaboration not only between learners but also with the teacher through shared contents and editing features (Eren, 2015; Wolsey et al., 2015). Multiplayer games and activities allow learners to exercise or play

together, generally without the teacher. For instance, the ICONS platform offers a multiplayer simulation (Angelini and García-Carbonell, 2019). In Photographe, each learner has to select one word that they want to use as a base word, then the system analyse all selected words and use a combination algorithms to return a new word and its image as the result (Tangworakitthaworn et al., 2019). Another example is WCRASMA, which offers features for learners to collaboratively generate different combinations of multimedia annotations on a given content (Lian et al., 2017).

- *Social*: these tools offers various functionalities to become aware of the presence of other learners and to interact with them. The most obvious examples are the social media tools. We could argue that they enable collaborative learning, but the functionalities offered, while technically enabling learners to work or play together, are limited, do not reflect their primary use and may not be appropriate for learning. We categorised 6 (14%) tools as social, even though some tools classified as collaborative also present social features (Wolsey et al., 2015; Eren, 2015).
- *Individual*: these tools do not offer any feature supporting interaction between learners. It represents 69% (29) of the tools.

The categories in question are not mutually exclusive, as collaborative tools often incorporate individual and social features to varying degrees. For the purpose of this review, we categorised the tools based on their highest feature on the social scale, with individual features considered the least prominent and collaborative features considered the most prominent.

As a summary, we can underline that the TAVL presented across the 44 papers are very diverse in form and content. The overwhelming majority of activity addresses the written form of the words, but half of the tools integrate an audio component. Pronunciation is even present in 3 tools. Most systems define a progression but few (7%) provide authoring tools for the teachers or learners to take control over the content. The social dimension of learning is addressed in just above 31% of the systems. Some systems integrate various gaming elements (24%).

5 Experiments

Among the 44 papers, 4 do not report any experiment (Hasnine and Wu, 2021; Love et al., 2017; Tangworakitthaworn et al., 2019; Wolsey et al., 2015). We will therefore detail 40 experiments.

5.1 Duration

The experiments were conducted during different time spans, the shorter being a single session in one day, the longer being 13 months (converted to 58 weeks), with an average of 12,1 weeks (slightly less than a semester) and a standard deviation of 15,8 weeks. We categorised them in four groups (see Table 3):

- express, when the experiment is conducted in a single session;
- short, when the experiment lasts less than a month (4 weeks);
- medium, when the experiment lasts between one month and one semester;
- long, when the experiment lasts more than a semester.

During the experiments, the exposure time to the system, i.e. the time learners spent using the system is very variable. In 52,5% (21) of the experiments, this duration is not declared. In a

Table 3: Duration distribution

Express	Short	Medium	Long
5	15	15	5

few papers, it is a simple omission, but in most papers, the precise duration is unknown to the researchers when they let the study subjects use the tools freely. In the remaining experiments, the mean exposure time is 6,2 hours with a standard deviation of 7,2 hours.

5.2 Participants

We classified the participants according to 3 variables: their age (see table 4), their nationality (see table 5) and their estimated level of proficiency in the studied language using the Common European Framework of Reference for Languages (CEFR) (see table 6).

Table 4: Age distribution grouped by scholar level

Primary school (4-10 y.o)	Middle school (11-14 y.o)	High school (15-18 y.o)	University	Non university adults (18+ y.o)
12	3	3	18	4

Table 5: Nationality distribution of participants

Taiwan	Iran	Turkey	USA	Saudi Arabia	Japan	Various Asian and Oceanian countries	Various European countries	Not specified
7	5	5	4	3	3	8	3	2

One paper used the HSK classification for the level of students learning Chinese (Pamintuan et al., 2018). The vast majority of the experiments did not specify the estimated level of the participants according to a referential. Therefore, the distribution above is statistically not representative of the actual distribution of the levels.

5.3 Variables and measures

Across the experiments, 4 categories of observed variables were identified (see Table 7). First, 92,5% (37) of the experiments measure one aspect or another of the participants' vocabulary learning. Then, the participants' opinions about the tools are collected by surveys, interviews and observations in 25% (10) of the experiments. Students' motivation, flow or engagement, combined in a same category, are measured in 8 (20%) experiments. Learning attitudes, approached as a combination of diverse variables such as self-regulation, concern 8 experiments.

Finally, a wide variety of other variables were measured in the experiments such as cognitive load, learning behavior, learning anxiety, and difference in gender performance. Since these variables are each studied in one or two experiments, we could not generalise the eclectic and specific findings

Table 6: Level distribution (CEFR)

Unspecified	A1 and below	B1	B2+	Varied
34	1	2	1	1

of the results. The variables are grouped in the category "Others", which concern 45% (18) of the experiments.

5.3.1 Vocabulary learning 36 out of the 37 experiments testing vocabulary skills test at least one written skill. On the opposite, only 9 experiments test oral abilities (speech, pronunciation or conversation comprehension) (Lian et al., 2017; Chacón-Beltrán, 2018; Chen and Lee, 2018; Pamintuan et al., 2018; Angelini and García-Carbonell, 2019; Hao et al., 2019; Loftus-Rattan and Furey, 2021; Yilmaz et al., 2022; Liu and Chen, 2023) and one solely focuses on pronunciation (Angelini and García-Carbonell, 2019), while the others measure both oral and written skills.

Testing someone's vocabulary is a complex task (Webb and Sasao, 2013): a complete lexicon contains tens of thousands of different words depending on the language, it is impossible to be exhaustive and test even a significant part of the lexicon. Therefore, a big uncertainty arises as it is possible that the words used in the tests overlap completely, partially or not at all the portion of lexicon known by the subject. To address the difficulties regarding the measure of the different vocabulary skills, three solutions are used in the experiments:

1. Crafting a test specifically for the experiment. This allows to focus solely on the selected vocabulary skills and to target the words studied during the experiment.
2. Using a standardised test as it is, or adapting it to fit the experiment more closely.
3. Using the logs left by the users on the tools. This is a recent approach that does not require test and takes advantage of the technology environment. It can be used as a primal source of information (Enokida et al., 2018) or to complement data obtained through tests (Chen and Lee, 2018).

The experiments reviewed in this study largely favored the "hand-crafted" *ad hoc* test method for measuring vocabulary learning, with 65% (24/37) of them using their own vocabulary test. The 35% (13) remaining studies used various existing tests.

- Alfadil (2020) adapted a test based on the Vocabulary Knowledge Scale developed by Paribakht and Wesche (1993);
- Arikan and Ozen (2015) used the Vocabulary Check List (VCL) developed by Ünal (2006);
- Chacón-Beltrán (2018) used the X-Lex from Meara and Milton (2003);
- Ebadi and Bashiri (2018) relied on the Dialang online diagnostic test ⁶;
- Gay (2022) adapted sections from the published Cambridge B2 First and C1 Advanced practice tests;
- Huang (2015) used the Expressive Vocabulary Test-2 (EVT-2) (Williams, 2007);
- Huang (2021) used a sample of the TOEIC Bridge (no reference provided by the authors);
- Jalaluddin et al. (2021) used the British Picture Vocabulary Scale III, a vocabulary test produced by GL Education Group UK (no reference provided by the authors);

⁶ <https://dialangweb.lancaster.ac.uk/>

Table 7: Variables tested in vocabulary learning studies^a

Study	Vocabulary learning	Motiv./ Engage./ Flow	Perceptions about the tools	Learning attitudes	Others
Counts	37	8	10	6	19
Abdul Wahdi and Dzulkifli (2018)	+				
Ahmad et al. (2015)	+		+		✓
Alfadil (2020)	+				
Alhajaji et al. (2020)	+	+			✓
Angelini and García-Carbonell (2019)	+				✓
Arikan and Ozen (2015)	+		+		
Basal et al. (2016)	+				
Caetano et al. (2023)	=		+		
Chacón-Beltrán (2018)	+			+	✓
Chen et al. (2022)	+				
Chen and Lee (2018)	=	+		+	✓
Dennis and Whalon (2021)	+				✓
Ebadi and Bashiri (2018)			+		✓
Elaish et al. (2019)	+	+			
Enokida et al. (2018)	=			+	✓
Eren (2015)	+		+		✓
Gay (2022)	+				
Hameed (2020)	+				
Hao et al. (2019)	+				✓
Hasan et al. (2022)	+		+	+	✓
Huang (2015)	+	+		+	
Huang (2021)	+		+		✓
Jalaluddin et al. (2021)	+				
Kazu and Kuvvetli (2023)	+				✓
Klimova and Polakova (2020)			+		
Lian et al. (2017)	+				✓
Liu and Chen (2023)	+				
Loftus-Rattan and Furey (2021)	+	+		+	
Mirzaei et al. (2021)	+				
Motlagh et al. (2020)	+				✓
Pamintuan et al. (2018)	+				
Pol'akov'a (2022)	+	+	+		✓
Sato (2016)	=				
Uz Bilgin and Tokel (2019)					✓
Wang and Huang (2017)	+	+	+		
Wong et al. (2016)	+				✓
Xodabande and Atai (2022)	+				
Yamamoto et al. (2019)	+				✓
Yilmaz et al. (2022)	+		+		
Zhang et al. (2023)	+	+			

^a A colored cell with + indicates a positive effect was found for that variable. An empty cell indicates the variable was not studied or reported in the paper. A ✓ in the "Others" column indicates additional variables were studied.

- Loftus-Rattan and Furey (2021) used the Peabody Picture Vocabulary Test-Fourth Edition (Dunn and Dunn, 2007)
- Liu and Chen (2023) used the first Step Level of Anglia ESOL Exams (England, 2014);
- Motlagh et al. (2020) took a test from the Essential World for TOEFL Practice Test Book (Matthiesen, 2017);
- Xodabande and Atai (2022) used the New vocabulary levels test (NVLTL) developed by Mclean et al. (2015);
- Pol’akov’a (2022) finally, claims to have used a standardised vocabulary test but without naming it.

5.3.2 Motivation, flow and engagement The impact of TAVL environments on learners’ motivation, flow and/or engagement is measured in 20% (8) of the experiments (Huang, 2015; Wang and Huang, 2017; Chen and Lee, 2018; Elaish et al., 2019; Alhajaji et al., 2020; Loftus-Rattan and Furey, 2021; Pol’akov’a, 2022; Zhang et al., 2023). Researchers have used various tests, in most cases without defining the concepts being measured nor the theoretical background:

- Alhajaji et al. (2020) used a qualitative approach by asking the participants how motivated they were through a single question on a survey.
- Chen and Lee (2018) measured the participants’ perception of flow using the EGameFlow flow scale Fu et al. (2009) that they adapted to their needs.
- Elaish et al. (2019) and Wang and Huang (2017) both used the Motivated Strategies for Learning Questionnaire (MSLQ) (Pintrich et al., 1993). Elaish et al. (2019) explicitly claims to rely on Self-Determination theory and develops the theoretical background of motivation in language learning, while Wang and Huang (2017) does not develop its definition of the concept.
- Huang (2015) conducted interviews at the end of the experiment and weekly field notes allowed them to observe the participants’ engagement.
- Loftus-Rattan and Furey (2021) used field observations for monitoring the participants’ engagement without defining nor developing the concept.
- Pol’akov’a (2022) coded and created categories to analyse the answers to questionnaires, minute papers and focus groups in order to analyse learners’ motivation.
- Zhang et al. (2023) used a questionnaire adapted from Gu and Johnson (1996); Mizumoto (2013); Wang and Bai (2017) to measure cognitive engagement in Digital Game Based Vocabulary Learning (DGBVL) and a questionnaire on learners’ emotional engagement adapted from Wu (2018). They rely on a definition of engagement from Ge and Ifenthaler (2018), and describe their theoretical background about engagement in DGBVL.

5.3.3 Perceptions about the tools The perceptions about the tools are collected either via interviews or handmade satisfaction questionnaires. Generally, in the reviewed papers a more qualitative approach is preferred in order to collect open and more nuanced answers. The researchers rely on open-ended questions during interviews or questionnaires using questions such as “How would you rate the enjoyment/efficiency/usability of the tool you used” or “How likely are you to keep using the tool in the future?”. Some of them are directly taken or inspired from the SUS questionnaire, but only one used the full questionnaire (Mirzaei et al., 2021).

For instance, the studies reported in (Huang, 2015; Wolsey et al., 2015; Eren, 2015; Ahmad et al., 2015; Arikan and Ozen, 2015; Ebadi and Bashiri, 2018; Hao et al., 2019; Yilmaz et al., 2022) used interviews with the participants. Yilmaz et al. (2022) also supplemented the interviews with

diaries and logs, and Ahmad et al. (2015) with field observations and audio recordings. Hasan et al. (2022) used an attitude questionnaire, which involved questions about the attitude of the students towards the tool.

5.3.4 Learning attitudes and behaviors In this section, we have grouped together the studies that paid close attention to the learners' behaviors, learning strategies, and attitudes towards the tool (Eren, 2015; Huang, 2015; Lian et al., 2017; Chacón-Beltrán, 2018; Chen and Lee, 2018; Enokida et al., 2018; Loftus-Rattan and Furey, 2021; Hasan et al., 2022). Among others, Chacón-Beltrán (2018) presented various learning strategies to the students before the experiment, then asked them which one they used and with which frequency; Chen and Lee (2018) investigated the influence of self-regulation, by adopting a Chinese version of a self-regulation scale (not cited in the paper), and also monitored the learners' behaviors on the app. Enokida et al. (2018) tracked the total learning duration, the outcome, and learning efficiency using the platform logs. Eren (2015) conducted semi-structured interviews regarding learners' attitudes towards the use of web-based platforms for improving their vocabulary knowledge. Hasan et al. (2022) used an attitude questionnaire from a previous study (Heidari Tabrizi and Onvani, 2018) to focus on vocabulary learning attitudes from a cognitive and metacognitive perspectives. Huang (2015) recorded the vocabulary teaching instruction and learning attitudes of the students toward vocabulary games in observational notes. The purpose of observations in the study conducted by Loftus-Rattan and Furey (2021) was to monitor students' behaviors and to capture qualitative observations regarding their interest, attention, and independence in completing the technology-based intervention sessions. Finally, Lian et al. (2017) analysed the frequency of some specific actions on the platform using the logs, to observe differences in behaviors among different groups.

5.4 Results

We report in this section the results obtained in the 40 studies regarding the variables identified in previous section.

5.4.1 Effects on learners' lexical skills Of the 37 studies on vocabulary learning, all but four (Sato, 2016; Enokida et al., 2018; Chen and Lee, 2018; Caetano et al., 2023) reported a significant positive impact of the tool. The results hold different significance as they were measured in different ways.

In the first category of experiments, including 13 studies (35%), the experimental design compare a test group that used a TAVL tool and a control group that relied on conventional learning. This group performed regular classroom activities but not necessarily similar to the activities performed by the test group with the tool. The TAVL tools showed greater results on vocabulary learning than a traditional approach, using various devices: VR (Alfadil, 2020), AR (Hameed, 2020), AI-based (Kazu and Kuvvetli, 2023), mobile (Basal et al., 2016; Pamintuan et al., 2018; Elaish et al., 2019; Alhajaji et al., 2020; Motlagh et al., 2020; Pol'akov'a, 2022; Xodabande and Atai, 2022) or computer-based (Huang, 2015; Eren, 2015; Abdul Wahdi and Dzulkifli, 2018; Angelini and García-Carbonell, 2019).

In 12 (32%) experiments, the evolution of variables was measured using pre and post tests, with no control group, and all showed an increase in vocabulary knowledge. In experiments where the software is considered by researchers to be the main vocabulary learning tool for students, it attests that the tool enables vocabulary acquisition, but not that it is more effective than more traditional

means (Arikan and Ozen, 2015; Chacón-Beltrán, 2018; Ebadi and Bashiri, 2018; Hao et al., 2019; Uz Bilgin and Tokel, 2019; Klimova and Polakova, 2020; Jalaluddin et al., 2021; Yilmaz et al., 2022; Zhang et al., 2023). In experiments where the tool is used as a complementary means of vocabulary learning, the studies focus on the overall setting (tool+courses+extra-curricular activities) and not on whether the tool is more effective than more traditional means (Ahmad et al., 2015; Wang and Huang, 2017; Wong et al., 2016).

In 4 (11%) experiments, the test group used a tool and the control group used a different one. These studies thus allowed researchers to compare the effect of the tool they propose with another "classical" vocabulary learning tool, whose efficacy is already proven. In their experiment, Caetano et al. (2023) compared the effect of ARlang, the experimental tool, with Anki, and found no significant difference between the two interventions in terms of short term vocabulary recall. Two experiments (Wang and Huang, 2017; Chen and Lee, 2018) compared their experimental tool, respectively My-Pet-Shop and "Interactive syllable-based English vocabulary learning board (I SEE)" with a quiz activity. Wang and Huang (2017) found increased performance in the group following the experimental condition while Chen and Lee (2018) found that their their intervention involving the application did not contribute to learning performance. Yamamoto et al. (2019) compared FindDo, a system based on Location-Context, with another application called VocList that does not contain location-based or context related content, and found a significantly better vocabulary score in the experimental group in post test.

Four other experiments compared different tool features: they provide two versions of a same tool and evaluate the effect in terms of vocabulary learning outcomes. The comparison between the use of animated visual glosses and pictorial images showed no positive impact of the animated glosses (Sato, 2016). In the case of the VLASTWA application, the control group used the app without access to the "KeyWord Method" strategy training environment, while the other participants had the "KeyWord Method" instruction. This latter experimental group yielded better results in terms of vocabulary learning. The AI-Based Object Detection Translation application was opposed to a simple translation software and the experimental group displayed better performances (Liu and Chen, 2023). In their experiment, Enokida et al. (2018) compared the use of HiroTan, a web application for vocabulary learning as a cross-platform application (experimental group) and an older, PC-based version (control group). No significant differences in outcome or learning efficiency were observed between the groups. Finally, Chen et al. (2022) investigated the effect of a new "Vocabulary Learning Mechanism" (VLM), composed of an online dictionary, word lists, searching and annotation features, added to an other existing system, VALRS. They compared the experimental group using VALRS+VLM to the control group using only VALRS and reported that students who used VALRS-VLM significantly performed better than the control group.

Finally, several studies followed specific protocols. A within-subjects design with pretest-posttest measures compared vocabulary knowledge gains between interventionist-led extended instruction and technology-based instruction (Loftus-Rattan and Furey, 2021). Students learned more words in the extended instruction condition than in the technology-based condition, but meaningful gains were observed in both conditions. Another study conducted by Huang (2021) used 3 intact classes to make 3 groups: non- SRS (Student Response System) group, high-tech SRS group (Nearpod) and low-tech SRS group (mini-whiteboards). Significantly higher comprehension scores were found for the two SRS groups but no significant difference in vocabulary tests. In a third study, participants were divided into 4 groups: control group, matched VE (Virtual Environment) group, independent VE group, teacher-centered VE group. All the VE groups performed better on vocabulary learning compared to the control group. Lian et al. (2017) tested the web-based collaborative reading anno-

tation system with multimedia annotations (WCRASMA) in 3 different conditions: voice and text annotations (VT), voice and picture (VP), and voice, text, and picture (VTP). The VTP group outperformed VP and VT groups in vocabulary learning.

The effect sizes are generally not provided in the studies, with 29 (72.5%) studies not explicitly reporting a comparable effect size. Among the others, one used Hedges' g , 3 relied on Cohen's d and 7 studies reported the effect size using partial eta squared η_p^2 (see Table 8).

Table 8: Reported effect sizes of interventions on vocabulary learning

Loftus-Rattan and Furey (2021)	$g = 0.66$
Motlagh et al. (2020)	$d = 0.74$
Wang and Huang (2017)	$d = 1.78$
Hasan et al. (2022)	$d = 7.08$
Hasan et al. (2022)	$d = 7.08$
Abdul Wahdi and Dzulkifli (2018)	$\eta_p^2 = 0.607$
Alfadil (2020)	$\eta_p^2 = 0.09$
Elaish et al. (2019)	$\eta_p^2 = 0.148$
Jalaluddin et al. (2021)	$\eta_p^2 = 0.75$
Kazu and Kuvvetli (2023)	$\eta_p^2 = 0.19$
Uz Bilgin and Tokel (2019)	$\eta_p^2 = 0.82$
Hasan et al. (2022)	$\eta_p^2 = 0.356$
29 remaining studies	Effect size not reported

5.4.2 Motivation, flow and engagement 8 studies measured a dimension of motivation, flow or engagement. 2 of them analysed motivation using qualitative approaches. In the study conducted by Alhajaji et al. (2020), students answered to an open-ended question and declared that they liked using the tool and felt motivated using it. Pol'akov'a (2022) relied on the analysis of a questionnaire, minute papers and focus groups to identify that students' motivation was increased thanks to better learning outcomes, and that the mobile learning application Angličtina Today enhanced students' motivation to learn.

Using the MSLQ motivation scale, Elaish et al. (2019) showed that the group that used the VocabGame mobile game did not report significantly higher motivation than the group that followed traditional training in the post-test. Time had a statistically significant effect, which demonstrates changes in motivation during the 3 weeks learning process. Wang and Huang (2017) used the same questionnaire to show that the tool used in the experimental group promoted greater motivation in comparison to the control group with a large effect size (Cohen's d greater than 0.8). They also noted that the students who were primarily most interested in the course content were motivated by using the tool to learn vocabulary.

Flow was measured using the EduFlow scale by Chen and Lee (2018) in their experiment. The findings demonstrated that the game My-Pet-Shop provided students with enhanced goals, feedback, autonomy, and immersion over the quiz game. The authors posited that goals and feedback were the most helpful to facilitating the flow and immersive experience, and also contributed to control and autonomy; all of this enhancing the students' engagement experience.

Through field observations, Huang (2015) observed that the use of vocabulary learning software contributed to students' interest and motivation to learn. Students were more engaged in the vocabulary software programs as they played with animated characters and interacted with words. Similarly, Loftus-Rattan and Furey (2021) observed during their experiment that students were engaged and attentive in the technology-based condition. Finally, a last study (Zhang et al., 2023) revealed the overall active learner engagement in DGBVL with their tool (no control group). The researchers identified three main phases of learner engagement in the DGBVL: exploring the game, developing vocabulary knowledge, and struggling to end the game. Statistically significant, positive effects of learner engagement in DGBVL on vocabulary development were identified, while engagement from different dimensions (cognitive engagement, emotional engagement) demonstrated different influences.

5.4.3 Perceptions about the tools The second most measured variable is the perception of the learners about the tools used, measured in 25% (10) of the experiments.

The results are unanimously positive among the 10 studies. All the same, within each experiment, perceptions are not unanimous. Certain quotes and drawbacks might give us more insight in this apparent positive consensus. A teacher involved in the study described in (Arikan and Ozen, 2015) declared that "It causes children to loose concentration, and it is not possible for them to regain it during the activities". Ebadi and Bashiri (2018) reported the negative aspects of the tool "Vocabulary Flashcards 2016", mentioned by the participants in a dedicated section, mostly directed towards the quiz and the levels. In the study of Klimova (2021), 21% of the students disagreed with the statement "I enjoyed using a mobile app to learn.", 33% agreed and 46% were neutral. Some other papers do not report enough details about the learners' perceptions and use vague sentences such as "quite positive" or "most of the students enjoyed...".

Therefore, it appears that in many cases, if not all, a minority of the subjects expressed some concerns or negative perceptions about the tools. This deserves to get more attention from the researchers, since that minority can represent up to 20% of the subjects (Klimova, 2021). In the terms of Pol'akov'a (2022), in their experiment "the number of negative findings were outnumbered by the positive ones". Regarding students' perceptions, a number of negative perceptions have to be considered, which means either that the tools can be improved, or that a certain percentage of the participants are systematically reluctant towards the tools for various reasons. Nevertheless, the clear majority of positive responses underline the massive approval by the general public of computer and mobile tools for learning vocabulary.

5.4.4 Learning attitudes and behaviors Chacón-Beltrán (2018) highlighted the importance of vocabulary learning strategies in the vocabulary learning process. Three strategies were reported by students before joining the course, namely: (1) memorising lists of words, (2) checking meanings of words in the dictionary and (3) using new words in sentences. After the course, the authors identified that eleven strategies were new to participants and were learned during the course. They conclude that those results are in line with previous research demonstrating high levels of achievement in relation to strategy use. Regarding self-regulation strategies, Chen and Lee (2018) revealed that their tool contributed to the forethought and performance phases. The analysis of behavior patterns showed that learning regulation is significantly related to both the application scenario and subject learning. Hasan et al. (2022) reported that the students who utilized WhatsApp in order to improve their vocabulary knowledge showed a more favourable viewpoint than the control group relying on

traditional classroom teaching toward learning new words in EFL. Enokida et al. (2018) relied on the platform logs to monitor students behavior, including the frequency, periodicity and learning efficiency of the participants. They observed a higher frequency and periodicity of logins for the cross platform application compared with the PC-based web application. Finally, Huang (2015) highlighted that integrating technology to a course changed some students' attitudes about topics they had not liked or areas where they did not feel competent. Moreover, they had more positive attitudes toward working with peers and reading activities while working with different digital word games.

6 Discussion

We have examined 44 papers that discuss TAVL tools and systems. The findings from these studies indicate general positive outcomes.

When looking at the results of the experiments regarding vocabulary learning, a clear tendency stands out. Indeed, most studies report a positive effect of the use of the digital vocabulary learning tools on the acquisition of vocabulary knowledge, only four studies (Sato, 2016; Enokida et al., 2018; Caetano et al., 2023) showing no effect.

It could be expected that some external factors independent from the tools themselves partially caused the observed effects. Firstly, the attention of the participants being suddenly drawn to vocabulary learning is very likely to trigger by itself an improvement in their performances. Vocabulary learning is often left outside the classroom as a responsibility of the learner, creating methodological gaps for students (Hilton, 2002, § 40). Vocabulary is learned in an implicit way (Hilton, 2002, § 39), which is known to be insufficient (Khezrlou et al., 2017). Therefore, the simple shift from a peripheral and implicit learning to explicit learning could be for a significant part in the observed results. Secondly, the novelty effect (Pisapia et al., 1993) is a phenomenon that arouses the curiosity of an individual confronted with a new experience for the first time. It has been shown in education that it can lead to an increase in attention span and, finally, an increase in effort or perseverance, which in turn leads to better performances (Pisapia et al., 1993).

A broader look at the methodology used in all the studies shows that they are based on different measurement methods and experimental designs. A large number of experiments (32%) rely on quasi-experimental designs, without a control group. In that case, these studies do not allow identifying clear cause-and-effect relationships between the variables. In fact, without a control group it becomes challenging to attribute changes in the outcome variable solely to the intervention or treatment being tested.

Another element that make the results obtained in the studies difficult to generalise is the predominant use of *ad hoc* questionnaires to measure the variables, which does not allow for comparisons of the results between experiments. The use of standardised tests in language learning experiments, especially within the field of computer-assisted vocabulary learning, presents a great challenge. Standardised tests are meticulously designed to comprehensively evaluate a multitude of language aspects, spanning grammar, vocabulary, syntax, and more. This thoroughness is intended to ensure that the tests can be broadly applied across various learning contexts and populations, providing a standardised measure of language proficiency. Moreover, they offer a solution to the generalisation problem, making it possible to compare the results of experiments using the same test. We thus recommend that standardised tests be utilised as a measurement instrument in ex-

periments when it is possible, in order to improve the generalisability and comparability of research results.

In addition to standardised and *ad hoc* tests, a third means of measurement was used in the experiments: the analysis of learners' interactions with the environment, in order to track their progress and behaviors over time with regard to the use of the environment. As it is non-intrusive, it can be used in conjunction with other methods. The drawbacks are that it is difficult to compare results between studies. In addition, this method can be difficult for researchers to implement, especially when using systems that they did not design (*e.g.* COTS). Finally, the combination of vocabulary tests with this non-intrusive method of traces analysis, can provide a comprehensive understanding of learners' behaviors and progress over time in technology-enhanced learning contexts.

The review reveals that 90% of the tools combine at least two different media contents for vocabulary presentation, with 29% using all four (text, audio, video, and images). This multimodal approach aligns with cognitive theories of multimedia learning (Mayer, 2002) and could potentially enhance vocabulary retention. However, more research is needed to determine the optimal combination of media types for different learning contexts and learner profiles. Future studies could explore how different multimodal presentations affect vocabulary acquisition, retention, and transfer to productive language use.

Regarding the features offered by the TAVL tools, our review reveals a wide variety. While this diversity provides learners and educators with many options, it also presents challenges in terms of comparing effectiveness across studies. The predominance of tools focusing on EFL highlights a potential gap in resources for other languages. Future research could explore the development and efficacy of TAVL tools for a broader range of languages, particularly those with different writing systems or linguistic structures from English.

We also observed that an important portion of the TAVL tools (40%) are games and gameful tools, which illustrates the interest in the digital game-based vocabulary learning domain and gamification for vocabulary learning. One of the common advantages attributed to games is to maintain a high level of motivation among learners (Garris et al., 2002, p. 443–444; Hess and Gunter, 2013, p. 382; Reinhardt and Thorne, 2019, p. 203; Schmoll, 2016, p. 18). Previous studies on serious game-aided vocabulary learning demonstrated the positive impact of serious games on vocabulary learning (Miller and Hegelheimer, 2006; Smith et al., 2013; Hassinger-Das et al., 2016). This is strongly supported by the mostly positive results of the experiments carried on games and gameful tools studied in this review. On the 16 experiments carried on games and gameful tools, 14 found a positive impact of the intervention. Moreover, 6 experiments that compared the effects of a game or gamelike tool with the effects of conventional classroom learning settings unanimously reported significant impact of the games or gameful tools on vocabulary learning (Abdul Wahdi and Dzulkifli, 2018; Elaish et al., 2019; Angelini and García-Carbonell, 2019; Alfadil, 2020; Alhajaji et al., 2020; Gay, 2022). In agreement with Wang et al. (2021), these results advocate for the integration of more gamification in vocabulary learning tools and/or the design of more vocabulary games, as the results above make them likely to improve students' learning experiences and learning efficiency.

It comes as a surprise that 7 papers focus on the use of a social media for vocabulary learning. Even though such systems do not necessarily present specific features for vocabulary teaching and learning, the results of the studies highlight that social media is a great way to practice writing skills in context. Indeed, regular written exchanges between social media users are an interesting

way of using vocabulary in a social way, which can certainly be seen as more fun and engaging than learning words individually from word lists, for instance. This interest from researchers and teachers for social media might therefore stem from a similar concern that leads them to look into games and gameful tools: as developed previously, it is crucial to catch the learners' interest and motivate them over the long term in order to positively influence their vocabulary learning. Therefore, as social media are very popular and commonly used, particularly by teenagers and young adults, they could be a way of engaging learners and making this new use of a tool they already use or are familiar with part of their daily routine.

Our analysis also reveals a lack of collaborative features in many TAVL tools, with only 17% designed to enhance collaboration between learners. This might be partially explained by the fact that vocabulary learning is mainly an individual autonomous task. This represents a significant opportunity for improvement, as collaborative learning has been shown to have numerous benefits in language acquisition. Collaborative learning is an educational approach to teaching and learning, which involves groups of learners working together to solve a problem, complete a task, or create a product (Laal and Laal, 2012) — which is in line with the principles underlying task-based learning. Several studies highlight the usefulness and benefits of "cooperative teams". For example, these teams achieve higher levels of thinking and retain information longer (Johnson and Johnson, 1987), which is underlined in CEFRL for vocabulary learning (Trim, 2002, § 5.4). In these collaborative activities, the coordination of individual actions within each group is important. Attitudes such as enjoying sharing lexical discoveries with others, noticing new words used by others and inquiring about their meanings, are entailed into "word consciousness". According to Tremblay (2021), it is an essential component of lexical competence, encompassing attitudes that contribute to the simultaneous development of knowledge and skills associated with word mastery. Thus, we make the hypothesis that implementing more collaborative features in TAVL tools, such as peer review systems, group vocabulary challenges, or collaborative storytelling exercises using target vocabulary, might be beneficial for learners' vocabulary learning. Future research into the effectiveness of such collaborative TAVL approaches compared to individual learning could yield valuable insights.

Finally, even though some tools analysed in this study take into consideration teachers and their role in the vocabulary learning process, the large majority still do not. Moreover, with the exception of Angličtina Today, which has a web interface specially developed for teachers, most of the tools do not go beyond a few analytics of student progress or, in even fewer cases, an integrated feature for sharing content such as lessons, vocabulary lists and more. They do not provide opportunities to articulate their use with a pedagogical scenario orchestrated by a teacher or a curriculum designer. Indeed, the tools seem to be thought as independent from the classroom and university settings, which may be explained by the aim of reaching a larger number of users. The support of teacher input in TAVL as a learning complement and not as a standalone activity can be interesting to address all the different lexical skills, recommended by Tremblay and Anctil (2020), and to address the methodological gaps found in student practices (Hilton, 2002). The role of the teacher should be considered when developing a TAVL tool in order to improve its integration into academic curriculum and more generally, learners' vocabulary learning⁷. This recommendation is in line with Alhuwaydi (2022) who suggests that teachers should guide their students to use mobile applications effectively to achieve the maximum benefit of vocabulary learning. Klimova (2021) also observed that only guided and controlled implementation of vocabulary learning apps can lead to an effective learning process.

⁷ See (Nation, 2013, chap. 14) for insight on the role of teachers and course designer in vocabulary learning.

7 Limitations

Our systematic literature review relies on specific methodological and keyword choices to ensure reproducibility. Alternative types of literature reviews, guided by different criteria, such as the relevance of papers within the domain, or alternative keywords, may have produced different results. To mitigate this specific threat, and with regard to the validity of the conclusions, we carefully discussed the inclusion criteria together. Nevertheless, we believe that this threat will not strongly affect the main results, as we have considered a rather large number of studies in our literature review. We decided to settle for reproducibility over exhaustiveness and opted for the PRISMA methodology that is widely adopted in the literature.

In addition, there was a time delay between the data collection process and the writing of the paper - some recent studies published after 2023 were not considered in our review process. This could explain why the field of AI, including LLMs, chatbots and machine translation, is barely represented despite its recent popularity. In fact, our review identified only two studies investigating the use of AI tools for vocabulary learning, both published in 2023. Moreover, some widely renowned applications, such as Duolingo, Quizlet and Rosetta Stone, do not appear in the retrieved papers. This may result from the fact that those applications are not regarded by researchers as vocabulary learning tools, therefore ending outside the scope of our search query applied on titles and keywords.

This being said, our literature review demonstrates the diversity and productivity of the field of CALL and TAVL in recent years, which makes it impossible to conduct an exhaustive systematic review of the field that would lead to the analysis of hundreds of papers. Therefore, our approach aims to retrieve a sample of vocabulary learning tools with as little bias as possible, in a reasonable quantity, to conduct an in-depth study and analysis of the retrieved papers, which provides a broad overview of the domain.

8 Conclusion

To conclude, this systematic review of literature demonstrates the overall positive influence of digital tools on vocabulary learning, encompassing outcomes, motivation, flow, and engagement. We have also identified various potential biases and external factors that could impact these outcomes. Those biases underline the importance of explicit focus on vocabulary, considering difficulties of students, at least in France (Hilton, 2002). Based on the research findings, we propose recommendations for designers and researchers. Firstly, we suggest the development of experiments that specifically neutralise the factors we have highlighted as potentially contributing to the observed effects on participants' vocabulary learning, namely the novelty effect and attention given to vocabulary learning. Furthermore, the use of more standardised experimental protocols with comparable measures, such as standardised tests, would greatly benefit the entire field of TAVL by facilitating generalisation of results and drawing conclusions from experiments conducted with different tools.

Based on these findings, we suggest to give a more central place to teachers and academic learning settings. Indeed, a previous existing review insisted on the fact that technologies show their greatest impact on students vocabulary learning when the tools are introduced and explained by the teachers and the usage is accompanied by them (Klimova, 2021). Therefore, it is of great importance to think this student-teacher relationship as a primordial aspect of the tools. We also suggest to include more social and collaborative features in order to stimulate the learners, multiply the points of view on the lexicon and improve their motivation to learn vocabulary. Finally, we propose that analysing usage traces could be highly valuable in observing the evolution of variables

over time and providing a more detailed and higher quality interpretation of results obtained from pretests and post-tests.

9 Supplementary materials

The corpus of papers studied in this systematic analyses is available as a resource (where and how to be determined with the editors)

10 Data availability statement

Data sharing not applicable to this article as no datasets were generated or analysed during the current study.

11 Ethical statement and competing interests

The research conducted adheres to ethical standards regarding the validation of experimental design. No participants are involved in this study. The author(s) declare(s) no competing interests.

Bibliography

- Masha'el A. Al-Jasir. Computer Assisted Vocabulary Learning: A Case Study on EFL Students at Al-Imam Muhammad Ibn Saud Islamic University. *Arab World English Journal*, (249):1–63, November 2019. ISSN 22299327. <https://doi.org/10.24093/awej/th.249>.
- Elok Vradizza Feni Abdul Wahdi and Mariam Adawiah Dzulki'fi. The Effectiveness of Computer-Assisted Instruction in Vocabulary Learning. In *2018 International Conference on Information and Communication Technology for the Muslim World (ICT4M)*, pages 219–226, July 2018. <https://doi.org/10.1109/ICT4M.2018.00048>.
- Kham Sila Ahmad, Fay Sudweeks, and Jocelyn Armarego. Learning English Vocabulary in a Mobile Assisted Language Learning (MALL) Environment: A Sociocultural Study of Migrant Women. *Interdisciplinary Journal of e-Skills and Lifelong Learning*, 11:25–45, January 2015. ISSN 23752084.
- Mohammed Alfadil. Effectiveness of virtual reality game in foreign language vocabulary acquisition. *Computers & Education*, 153:103893, 2020.
- Banan Hassan Alhajaji, Jalila Saleh Algmedi, and Amal Abdelsattar Metwally. Exploring the Success of GMT Technique: Games, Mind-Mapping, and Twitter Hashtags in Teaching Vocabulary in EFL Higher Education Environment. *International Journal of Higher Education*, 9(3):290–299, January 2020. ISSN 19276044.
- Ahmed Alhuwaydi. A Review on Vocabulary Learning-Designed MALL Applications in the EFL Context. *Theory and Practice in Language Studies*, 12(10):2191–2200, 2022. ISSN 17992591. <https://doi.org/10.17507/tpls.1210.27>.
- Susan A. Ambrose, Michael W. Bridges, Michele DiPietro, Marsha C. Lovett, and Marie K. Norman. *How Learning Works: Seven Research-Based Principles for Smart Teaching*. May 2010. ISBN 978-0-470-48410-4.
- M. Laura Angelini and Amparo García-Carbonell. Developing English Speaking Skills through Simulation-Based Instruction. *Teaching English with Technology*, 19(2):3–20, January 2019. ISSN 16421027.
- Yuksel Deniz Arikan and Sevil Orhan Ozen. A Learning Environment for English Vocabulary Using Quick Response Codes. *Educational Sciences: Theory and Practice*, 15(2):539–551, April 2015. ISSN 13030485.
- Ahmet Basal, Selahattin Yilmaz, and Asli Tanriverdi. Effectiveness of Mobile Applications in Vocabulary Teaching. *Contemporary Educational Technology*, 7(1):47–59, January 2016. ISSN 1309517X.
- Arthur Caetano, Alyssa Lawson, Yimeng Liu, and Misha Sra. ARLang: An Outdoor Augmented Reality Application for Portuguese Vocabulary Learning. In *Proceedings of the 2023 ACM Designing Interactive Systems Conference*, DIS '23, pages 1224–1235, New York, NY, USA, 2023. Association for Computing Machinery. ISBN 978-1-4503-9893-0. <https://doi.org/10.1145/3563657.3596090>.
- Rubén Chacón-Beltrán. Vocabulary Learning Strategies outside the Classroom Context: What Adults Learn in a Technology-Based Learner-Centred Environment. *Language Learning Journal*, 46(5):583–593, January 2018. ISSN 09571736. <https://doi.org/10.1080/09571736.2018.1503135>.
- Chih-Ming Chen, Ming-Chaun Li, and Mei-Feng Lin. The Effects of Video-Annotated Learning and Reviewing System with Vocabulary Learning Mechanism on English Listening Comprehension and Technology Acceptance. *Computer Assisted Language Learning*, 35(7):1557–1593, January 2022. ISSN 17443210. <https://doi.org/10.1080/09588221.2020.1825093>.
- Zhi-Hong Chen and Shu-Yu Lee. Application-Driven Educational Game to Assist Young Children in Learning English Vocabulary. *Educational Technology & Society*, 21(1):70–81, January 2018. ISSN 14364522.
- Sin Wang Chong and Luke Plonsky. A typology of secondary research in Applied Linguistics. *Applied Linguistics Review*, 15(4):1569–1594, July 2024. ISSN 1868-6311. <https://doi.org/10.1515/applirev-2022-0189>. URL <https://www.degruyter.com/document/doi/10.1515/applirev-2022-0189/html>.
- Edward L. Deci and Richard M. Ryan. *Intrinsic Motivation and Self-Determination in Human Behavior*. Boston, MA, 1985. ISBN 978-1-4613-4448-3. <https://doi.org/10.1007/978-1-4899-2271-7>.

- Lindsay R. Dennis and Kelly J. Whalon. Effects of Teacher- versus Application-Delivered Instruction on the Expressive Vocabulary of At-Risk Preschool Children. *Remedial and Special Education*, 42(4):195–206, August 2021. ISSN 07419325. <https://doi.org/10.1177/0741932519900991>.
- Lloyd M. Dunn and Leota M. Dunn. *Peabody Picture Vocabulary Test*. American Guidance Service, 4th edition, 2007.
- Saman Ebadi and Saba Bashiri. Investigating EFL Learners’ Perspectives on Vocabulary Learning Experiences through Smartphone Applications. *Teaching English with Technology*, 18(3):126–151, July 2018. ISSN 16421027.
- Monther M. Elaish, Norjihani Abdul Ghani, Liyana Shuib, and Ahmed Al-Haiqi. Development of a Mobile Game Application to Boost Students’ Motivation in Learning English Vocabulary. *IEEE Access*, 7: 13326–13337, 2019. ISSN 2169-3536. <https://doi.org/10.1109/ACCESS.2019.2891504>.
- Anglia Examinations England. *Handbook for teachers: Full examination syllabus and specifications for the revised exams 2014/2015*. Anglia Examinations Syndicate Limited, 2014.
- Kazumichi Enokida, Kunihiro Kusanagi, Shusaku Kida, Mitsuhiko Morita, and Tatsuya Sakaue. Tracking Online Learning Behaviour in a Cross-Platform Web Application for Vocabulary Learning Courses. December 2018.
- Omer Eren. Vocabulary Learning on Learner-Created Content by Using Web 2.0 Tools. *Contemporary Educational Technology*, 6(4):281–300, January 2015. ISSN 1309517X.
- Fong-Ling Fu, Rong-Chang Su, and Sheng-Chin Yu. EGameFlow: A scale to measure learners’ enjoyment of e-learning games. *Computers & Education*, 52(1):101–112, January 2009. ISSN 0360-1315. <https://doi.org/10.1016/j.compedu.2008.07.004>.
- Rosemary Garris, Robert Ahlers, and James E. Driskell. Games, motivation, and learning: A research and practice model. *Simulation & Gaming*, 33(4):441–467, 2002. ISSN 1046-8781. <https://doi.org/10.1177/1046878102238607>.
- Freddie Gay. Investigating the effects of technology-enhanced vocabulary learning strategy instruction on supporting mixed-ability EMI learners studying a journalism and communication major: An action research project at a university in China. *Journal of English for Academic Purposes*, 55:101068, 2022.
- Xun Ge and Dirk Ifenthaler. Designing Engaging Educational Games and Assessing Engagement in Game-Based Learning. In *Ophthalmology: Breakthroughs in Research and Practice*, pages 1–18. January 2018. ISBN 978-1-5225-5199-7. <https://doi.org/10.4018/978-1-5225-5198-0.ch001>.
- Yongqi Gu and Robert Keith Johnson. Vocabulary Learning Strategies and Language Learning Outcomes. *Language Learning*, 46(4):643–679, 1996. ISSN 1467-9922. <https://doi.org/10.1111/j.1467-1770.1996.tb01355.x>.
- Michael Gusenbauer and Neal R. Haddaway. Which academic search systems are suitable for systematic reviews or meta-analyses? Evaluating retrieval qualities of Google Scholar, PubMed, and 26 other resources. *Research Synthesis Methods*, 11(2):181–217, 2020. ISSN 1759-2887. <https://doi.org/10.1002/jrsm.1378>.
- Paiker Fatima Mazhar Hameed. Learner-Centered Language Learning Strategy in Digital Environment: Its Effect to Students’ Vocabulary, Collaboration and Independence. *TESOL International Journal*, 15(4): 102–123, January 2020. ISSN 20943938.
- Tao Hao, Zhe Wang, and Yuliya Ardasheva. Technology-Assisted Vocabulary Learning for EFL Learners: A Meta-Analysis. *JOURNAL OF RESEARCH ON EDUCATIONAL EFFECTIVENESS*, 14(3):645–667, July 2021. ISSN 1934-5747. <https://doi.org/10.1080/19345747.2021.1917028>. Place: 2-4 PARK SQUARE, MILTON PARK, ABINGDON OX14 4RN, OXON, ENGLAND Publisher: ROUTLEDGE JOURNALS, TAYLOR & FRANCIS LTD Type: Article.
- Yungwei Hao, Kathryn S Lee, Szu-Ting Chen, and Sin Chie Sim. An evaluative study of a mobile application for middle school students struggling with English vocabulary learning. *Computers in Human Behavior*, 95:208–216, 2019.
- Md Kamrul Hasan, Abdul-Hafeed Fakih, Prodhan Mahbub Ibna Seraj, and Hasmirati. The effect of technology-assisted language programme on vocabulary learning among EFL students at the tertiary level. *Heliyon*, 8(8):e10313, 2022.

- Mohammad Nehal Hasnine and Junji Wu. Wordhyve: A context-aware language learning app for vocabulary enhancement through images and learning contexts. *Procedia Computer Science*, 192:3432–3439, 2021.
- Brenna Hassinger-Das, Katherine Ridge, Amira Parker, Roberta Michnick Golinkoff, Kathy Hirsh-Pasek, and David K. Dickinson. Building Vocabulary Knowledge in Preschoolers Through Shared Book Reading and Gameplay: Vocabulary Game. *Mind, Brain, and Education*, 10(2):71–80, June 2016. ISSN 17512271. <https://doi.org/10.1111/mbe.12103>.
- Hossein Heidari Tabrizi and Nareges Onvani. The Impact of Employing Telegram App on Iranian EFL Beginners' Vocabulary Teaching and Learning. *Applied Research on English Language*, 7(1):1–18, January 2018. ISSN 2252-0198. <https://doi.org/10.22108/are.2017.103310.1087>.
- Taryn Hess and Glenda Gunter. Serious game-based and nongame-based online courses: Learning experiences and outcomes. *British Journal of Educational Technology*, 44(3):372–385, 2013. ISSN 1467-8535. <https://doi.org/10.1111/bjet.12024>.
- Heather Hilton. Modèles de l'acquisition lexicale en L2 : où en sommes-nous ? *ASp*, (35-36):201–217, 2002. ISSN 1246-8185. <https://doi.org/10.4000/asp.1668>.
- Heather Hilton. The link between vocabulary knowledge and spoken l2 fluency. *The Language Learning Journal*, 36(2):153–166, 2008. <https://doi.org/10.1080/09571730802389983>.
- Joan Wan-Ting Huang. Is SRS (Student Response System) Icing on the Cake? Comparing Efficacy of Different Modalities of SRS Engagement Incorporated into Collaborative Reading in an EFL Classroom. *Reading Matrix: An International Online Journal*, 21(1):38–50, April 2021. ISSN 1533242X.
- SuHua Huang. Mixed-Method Research on Learning Vocabulary through Technology Reveals Vocabulary Growth in Second-Grade Students. *Reading Psychology*, 36(1):1–30, January 2015. ISSN 02702711. <https://doi.org/10.1080/02702711.2013.808723>.
- Ilyana Jalaluddin, Ramiza Darmi, and Lilliati Ismail. Application of Mobile Augmented Visual Reality (MAVR) for Vocabulary Learning in the ESL Classroom. *Asian Journal of University Education*, 17(3):162–173, July 2021. ISSN 18237797.
- David W. Johnson and Roger T. Johnson. *Learning together and alone: Cooperative, competitive, and individualistic learning*, 2nd ed. Learning together and alone: Cooperative, competitive, and individualistic learning, 2nd ed. Prentice-Hall, Inc, Englewood Cliffs, NJ, US, 1987. ISBN 978-0-13-527871-0.
- Ibrahim Yasar Kazu and Murat Kuvvetli. The Influence of Pronunciation Education via Artificial Intelligence Technology on Vocabulary Acquisition in Learning English. *International Journal of Psychology and Educational Studies*, 10(2):480–493, January 2023. ISSN 21489378.
- Sima Khezroul, Rod Ellis, and Karim Sadeghi. Effects of computer-assisted glosses on EFL learners' vocabulary acquisition and reading comprehension in three learning conditions. *SYSTEM*, 65:104–116, April 2017. ISSN 0346-251X. <https://doi.org/10.1016/j.system.2017.01.009>.
- Jan H. Kietzmann, Kristopher Hermkens, Ian P. McCarthy, and Bruno S. Silvestre. Social media? Get serious! Understanding the functional building blocks of social media. *Business Horizons*, 54(3):241–251, May 2011. ISSN 0007-6813. <https://doi.org/10.1016/j.bushor.2011.01.005>.
- Barbara Kitchenham and Stuart Charters. Guidelines for performing systematic literature reviews in software engineering. (EBSE-2007-01), 2007.
- Blanka Klimova. Evaluating Impact of Mobile Applications on EFL University Learners' Vocabulary Learning – A Review Study. *Procedia Computer Science*, 184:859–864, January 2021. ISSN 1877-0509. <https://doi.org/10.1016/j.procs.2021.03.108>.
- Blanka Klimova and Petra Polakova. Students' Perceptions of an EFL Vocabulary Learning Mobile Application. *Education Sciences*, 10, January 2020. ISSN 22277102.
- Marjan Laal and Mozghan Laal. Collaborative learning: What is it? *Procedia - Social and Behavioral Sciences*, 31:491–495, 2012. ISSN 18770428. <https://doi.org/10.1016/j.sbspro.2011.12.092>.
- Batia Laufer. Possible changes in attitude towards vocabulary acquisition research. 24(1):69–75, 1986. ISSN 0019-042X. <https://doi.org/10.1515/iral.1986.24.1-4.35>.
- Ying-You Lian, Chih-Ming Chen, and Yi-Rong Li. Effects of Collaborative Multimedia Annotations on Elementary School Students' Vocabulary Learning Performance. In *2017 6th IIAI International Congress on Advanced Applied Informatics (IIAI-AAI)*, pages 649–654, July 2017. <https://doi.org/10.1109/IIAI-AAI.2017.52>.

- Jen-Jiun Lin and Huifen Lin. Mobile-assisted ESL/EFL vocabulary learning: a systematic review and meta-analysis. *Computer Assisted Language Learning*, 32(8):878–919, November 2019. ISSN 0958-8221. <https://doi.org/sc>.
- Pei-Lin Liu and Chiu-Jung Chen. Using an AI-Based Object Detection Translation Application for English Vocabulary Learning. *Educational Technology & Society*, 26(3):5–20, July 2023. ISSN 14364522.
- Susan M. Loftus-Rattan and Jenlyn Furey. Word Learning in First Grade: Evaluating the Effectiveness of a Technology-Based and an Extended Vocabulary Intervention. *Reading & Writing Quarterly*, 37(4):317–335, January 2021. ISSN 10573569. <https://doi.org/10.1080/10573569.2020.1800540>.
- Matthew L. Love, Tracy G. Spies, and Joseph J. Morgan. Using E-Books to Acquire Foundational Academic Vocabulary. *Intervention in School and Clinic*, 53(2):88–93, November 2017. ISSN 10534512. <https://doi.org/10.1177/1053451217693368>.
- Qing Ma. *Technologies for Teaching and Learning L2 Vocabulary*, chapter 4, page 45–61. 2017. ISBN 9781118914069. <https://doi.org/10.1002/9781118914069.ch4>.
- Hassan Saleh Mahdi. Effectiveness of Mobile Devices on Vocabulary Learning: A Meta-Analysis. *Journal of Educational Computing Research*, 56(1):134–154, March 2018. ISSN 0735-6331.
- Steven J. Matthiesen. *Essential Words for the TOEFL*. Simon and Schuster, New York, 2017.
- Richard E. Mayer. Multimedia learning. In *Psychology of Learning and Motivation*, volume 41, pages 85–139. Elsevier, 2002. ISBN 978-0-12-543341-9. [https://doi.org/10.1016/S0079-7421\(02\)80005-6](https://doi.org/10.1016/S0079-7421(02)80005-6).
- Stuart Mclean, Brandon Kramer, M. Kanzaki, and J. D. Brown. The Creation of a New Vocabulary Levels Test. 2015.
- Paul Meara. The rediscovery of vocabulary. 18(4):393–407, 2002. <https://doi.org/10.1191/0267658302sr211xx>.
- Paul Meara and J Milton. *The Swansea Levels Tests X-LEX*. Express Publishing, 2003. ISBN 978-1-4715-4246-6.
- Megan Miller and Volker Hegelheimer. The SIMs meet ESL Incorporating authentic computer simulation games into the language classroom. *Interactive Technology and Smart Education*, 3(4):311–328, January 2006. ISSN 1741-5659. <https://doi.org/10.1108/17415650680000070>.
- Siamak Mirzaei, Trent Lewis, Mirella Wyra, and Brett Wilkinson. Usability Testing of VLASTWA: A Vocabulary and Strategy Teaching Web App. In *Proceedings of the 32nd Australian Conference on Human-Computer Interaction, OzCHI '20*, pages 614–621, New York, NY, USA, 2021. Association for Computing Machinery. ISBN 978-1-4503-8975-4. <https://doi.org/10.1145/3441000.3441056>.
- Atsushi Mizumoto. Effects of self-regulated vocabulary learning process on self-efficacy. *Innovation in Language Learning and Teaching*, 7(3):253–265, November 2013. ISSN 1750-1229. <https://doi.org/10.1080/17501229.2013.836206>.
- Hossein Saadabadi Motlagh, Halimeh Khafaie, Ali Asghar Arastoo, Maria Cheraghi, and Morteza Abdullatif Khafaie. Application of Social Network in Traditional Sciences Education on the Vocabulary Acquisition of Secondary English Learner Students. *Education and Information Technologies*, 25(4):3071–3085, July 2020. ISSN 13602357. <https://doi.org/10.1007/s10639-020-10108-4>.
- I. S. P. Nation. *Learning Vocabulary in Another Language*. Cambridge Applied Linguistics. Cambridge University Press, 2 edition, 2013. <https://doi.org/10.1017/CBO9781139858656>.
- Matthew J. Page, Joanne E. McKenzie, Patrick M. Bossuyt, Isabelle Boutron, Tammy C. Hoffmann, Cynthia D. Mulrow, Larissa Shamseer, Jennifer M. Tetzlaff, Elie A. Akl, Sue E. Brennan, Roger Chou, Julie Glanville, Jeremy M. Grimshaw, Asbjørn Hróbjartsson, Manoj M. Lalu, Tianjing Li, Elizabeth W. Loder, Evan Mayo-Wilson, Steve McDonald, Luke A. McGuinness, Lesley A. Stewart, James Thomas, Andrea C. Tricco, Vivian A. Welch, Penny Whiting, and David Moher. The PRISMA 2020 statement: an updated guideline for reporting systematic reviews. *BMJ*, 372, March 2021. ISSN 1756-1833. <https://doi.org/10.1136/bmj.n71>.
- Cavin F. Pamintuan, Donnabelle G. Mallari, Nicole T. Garcia, Jeniezen P. Galang, and Reggie Mark B. Buduan. The Use of WeChat Application on CFL Learners' Vocabulary Acquisition. *TESOL International Journal*, 13(4):26–38, January 2018. ISSN 20943938.

- T. Sima Paribakht and Marjorie Bingham Wesche. Reading Comprehension and Second Language Development in a Comprehension-Based ESL Program. *TESL Canada Journal*, 11(1):09, October 1993. ISSN 1925-8917, 0826-435X. <https://doi.org/10.18806/tesl.v11i1.623>.
- Paul R. Pintrich, David A. Smith, Teresa Garcia, and Wilbert J. McKeachie. Reliability and predictive validity of the Motivated Strategies for Learning Questionnaire (MSLQ). *Educational and Psychological Measurement*, 53(3):801–813, 1993. ISSN 1552-3888. <https://doi.org/10.1177/0013164493053003024>.
- John Pisapia, Jeanne Schlesinger, and Amanda Parks. Learning Technologies in the Classroom: Review of the Literature. Technical report, February 1993.
- Petra Pol'akov'a. Use of a mobile learning application in the process of foreign vocabulary learning. *Procedia Computer Science*, 207:64–70, 2022.
- Jonathon Reinhardt and Steven L. Thorne. *Digital Games as Language-Learning Environments*, chapter 17, page 409–435. MIT Press, Cambridge, MASS, 2019. ISBN 9780262043380.
- Takeshi Sato. Applicability of technology-enhanced visual glosses for explicit L2 vocabulary learning: The enhancement of metaphoric competence through the learning of English polysemous words. *Ampersand*, 3:43–51, 2016.
- Norbert Schmitt. Review article: Instructed second language vocabulary learning. *Language Teaching Research*, 12(3):329–363, 2008. <https://doi.org/10.1177/1362168808089921>.
- Laurence Schmoll. L'emploi des jeux dans l'enseignement des langues étrangères : Du traditionnel au numérique. *Sciences du jeu*, (5), February 2016. ISSN 2269-2657. <https://doi.org/10.4000/sdj.628>.
- Judith A. Scott and William E. Nagy. *Developing Word Consciousness*, page 106–117. International Reading Association, 2009. ISBN 978-0872078062.
- Glenn Gordon Smith, Mimi Li, Jack Drobisz, Ho-Ryong Park, Deoksoon Kim, and Stanley Dana Smith. Play games or study? Computer games in eBooks to learn English vocabulary. *Computers & Education*, 69:274–286, November 2013. ISSN 0360-1315. <https://doi.org/10.1016/j.compedu.2013.07.015>.
- Preecha Tangworakitthaworn, Preeyapol Owatsuwan, Nutsima Nongyai, and Nongnapas Arayapong. An Image-Based Vocabulary Learning System Based on Multi-Agent System. In *2019 16th International Joint Conference on Computer Science and Software Engineering (JCSSE)*, pages 324–329, July 2019. <https://doi.org/10.1109/JCSSE.2019.8864170>.
- Ophélie Tremblay. « Sensibilité lexicale » : l'émergence d'un concept en didactique du lexique. *Pratiques. Linguistique, littérature, didactique*, (189-190), July 2021. ISSN 2425-2042. <https://doi.org/10.4000/pratiques.9999>.
- Ophélie Tremblay and Dominic Anctil. Introduction. *Recherches actuelles en didactique du lexique : avancées, réflexions, méthodes*, (62), November 2020. ISSN 1146-6480.
- John Trim, editor. *Cadre européen commun de référence pour les langues: apprendre, enseigner, évaluer — guide pour les utilisateurs*. Division des Politiques Linguistiques, Strasbourg, Avril 2002.
- Wen-Ta Tseng and Norbert Schmitt. Toward a Model of Motivated Vocabulary Learning: A Structural Equation Modeling Approach. *Language Learning* 58:2, June 2008. ISSN 0023-8333.
- Cigdem Uz Bilgin and S. Tugba Tokel. Facilitating Contextual Vocabulary Learning in a Mobile-Supported Situated Learning Environment. *Journal of Educational Computing Research*, 57(4):930–953, July 2019. ISSN 07356331. <https://doi.org/10.1177/0735633118779397>.
- Chuang Wang and Barry Bai. Validating the Instruments to Measure ESL/EFL Learners' Self-Efficacy Beliefs and Self-Regulated Learning Strategies. *TESOL Quarterly*, 51(4):931–947, 2017. ISSN 0039-8322.
- Fu Lee Wang, Ruofei Zhang, Di Zou, Oliver Tat Sheung Au, Haoran Xie, and Leung Pun Wong. A Review of Vocabulary Learning Applications: From the Aspects of Cognitive Approaches, Multimedia Input, Learning Materials, and Game Elements. *Knowledge Management & E-Learning*, 13(3):250–272, September 2021. ISSN 20737904. URL <https://search.ebscohost.com/login.aspx?direct=true&db=eric&AN=EJ1326130&lang=fr&site=ehost-live>.
- Wen-Yen Wang and Yueh-Min Huang. Interactive Syllable-Based English Vocabulary Learning in a Context-Aware Environment. *Journal of Educational Computing Research*, 55(2):219–239, April 2017. ISSN 07356331. <https://doi.org/10.1177/0735633116663212>.

- Stuart A. Webb and Yosuke Sasao. New Directions In Vocabulary Testing. *RELC Journal*, 44(3):263–277, December 2013. ISSN 0033-6882, 1745-526X. <https://doi.org/10.1177/0033688213500582>.
- Kathleen T. Williams. *EVT-2: Expressive Vocabulary Test*. Pearson Assessments, 2007.
- Thomas DeVere Wolsey, Linda Smetana, and Dana L. Grisham. Vocabulary plus Technology: An After-Reading Approach to Develop Deep Word Learning. *Reading Teacher*, 68(6):449–458, March 2015. ISSN 00340561. <https://doi.org/10.1002/trtr.1331>.
- Lung-Hsiang Wong, Ronnel B. King, and Ching Sing Chai. 'Seamlessly' Learning Chinese: Contextual Meaning Making and Vocabulary Growth in a Seamless Chinese as a Second Language Learning Environment. *Instructional Science: An International Journal of the Learning Sciences*, 44(5):399–422, October 2016. ISSN 00204277. <https://doi.org/10.1007/s11251-016-9383-z>.
- Ting-Ting Wu. Improving the effectiveness of English vocabulary review by integrating ARCS with mobile game-based learning. *Journal of Computer Assisted Learning*, 34(3):315–323, 2018. ISSN 1365-2729. <https://doi.org/10.1111/jcal.12244>.
- Ismail Xodabande and Mahmood Reza Atai. Using Mobile Applications for Self-Directed Learning of Academic Vocabulary among University Students. *Open Learning*, 37(4):330–347, January 2022. ISSN 14699958. <https://doi.org/10.1080/02680513.2020.1847061>.
- Keiko Yamamoto, Jesus Rodriguez, and Yoshihiro Tsujino. FinDo: A Foreign Language Vocabulary Learning System Based on Location-Context. In *2019 20th IEEE/ACIS International Conference on Software Engineering, Artificial Intelligence, Networking and Parallel/Distributed Computing (SNPD)*, pages 302–307, July 2019. <https://doi.org/10.1109/SNPD.2019.8935794>.
- Shelly Xueting Ye, Jia Shi, and Linyu Liao. An evaluative review of mobile-assisted L2 vocabulary learning approaches based on the situated learning theory. *Journal of Curriculum and Teaching*, 2023.
- Rabia Meryem Yilmaz, Fatma Burcu Topu, and Aysegül Takkaç Tuglar. An Examination of Vocabulary Learning and Retention Levels of Pre-School Children Using Augmented Reality Technology in English Language Learning. *Education and Information Technologies*, 27(5):6989–7017, June 2022. ISSN 13602357. <https://doi.org/10.1007/s10639-022-10916-w>.
- Aiqing Yu and Guy Trainin. A Meta-Analysis Examining Technology-Assisted L2 Vocabulary Learning. *ReCALL*, 34(2):235–252, May 2021. ISSN 09583440. <https://doi.org/10.1017/S0958344021000239>. URL <https://search.ebscohost.com/login.aspx?direct=true&db=eric&AN=EJ1333441&lang=fr&site=ehost-live>.
- Ruofei Zhang, Di Zou, and Gary Cheng. Learner engagement in digital game-based vocabulary learning and its effects on EFL vocabulary development. *System*, 119:103173, 2023.
- Di Zou, Yan Huang, and Haoran Xie. Digital game-based vocabulary learning: where are we and where are we going? *Computer Assisted Language Learning*, 34(5-6):751–777, July 2021. ISSN 0958-8221. <https://doi.org/10.1080/09588221.2019.1640745>.
- Evsen Ünal. An investigation into the role of interactive vocabulary instruction on learners' vocabulary knowledge. Master's thesis, Çukurova University, Institute of Social Sciences, Adana, Turkey, 2006.