pedocs
Open Access Erziehungswissenschaften

DIPF

Lentner-Matron, Daniela

How words become knowledge

2026, 13 S.

Quellenangabe/ Reference:
Lentner-Matron, Daniela: How words become knowledge. 2026, 13 S. - URN: urn:nbn:de:0111-pedocs-366800 - DOI: 10.25656/01:36680

https://nbn-resolving.org/urn:nbn:de:0111-pedocs-366800
https://doi.org/10.25656/01:36680

Nutzungsbedingungen

Gewährt wird ein nicht exklusives, nicht übertragbares, persönliches und beschränktes Recht auf Nutzung dieses Dokuments. Dieses Dokument ist ausschließlich für den persönlichen, nicht-kommerziellen Gebrauch bestimmt. Die Nutzung stellt keine Übertragung des Eigentumsrechts an diesem Dokument dar und gilt vorbehaltlich der folgenden Einschränkungen: Auf sämtlichen Kopien dieses Dokuments müssen alle Urheberrechtshinweise und sonstigen Hinweise auf gesetzlichen Schutz beibehalten werden. Sie dürfen dieses Dokument nicht in irgendeiner Weise abändern, noch dürfen Sie dieses Dokument für öffentliche oder kommerzielle Zwecke vervielfältigen, öffentlich ausstellen, aufführen, vertreiben oder anderweitig nutzen.
Mit der Verwendung dieses Dokuments erkennen Sie die Nutzungsbedingungen an.

Terms of use

We grant a non-exclusive, non-transferable, individual and limited right to using this document.
This document is solely intended for your personal, non-commercial use. Use of this document does not include any transfer of property rights and it is conditional to the following limitations: All of the copies of this documents must retain all copyright information and other information regarding legal protection. You are not allowed to alter this document in any way, to copy it for public or commercial purposes, to exhibit the document in public, to perform, distribute or otherwise use the document in public.

By using this particular document, you accept the above-stated conditions of use.

Kontakt / Contact:

peDOCS
DIPF | Leibniz-Institut für Bildungsforschung und Bildungsinformation
Informationszentrum (IZ) Bildung
E-Mail: pedocs@dipf.de
Internet: www.pedocs.de

Mitglied der Leibniz-Gemeinschaft

How Words Become Knowledge

Daniela Lentner-Matron

November 2026

Abstract

The Vocabulary Loop is a didactic model that describes how an initial encounter with a word can develop into robust vocabulary knowledge. It comprises seven phases that are not strictly linear: initial encounter, attention and perception, building meaning, multiple connections, retrieval and consolidation, application and automation, and renewed encounter with feedback and adaptation. The model is a synthesis for practical use that draws on research on multidimensional word knowledge, retrieval practice, spaced practice, and feedback; it is neither an experimentally validated stage model nor a description of a neurological sequence. From it, the article derives a four-step learning cycle and five practical principles. It then examines what meta-analyses and experiments show about digital and AI-supported vocabulary learning. Most of this evidence comes from adolescents and adults learning English as a foreign language; evidence for preschool and primary-school children is sparse. The article closes with the limitations of the model and age-sensitive suggestions for teachers and parents, including data-protection considerations in the German context.

Keywords: vocabulary learning; vocabulary acquisition; didactic model; retrieval practice; spaced practice; feedback; digital media; artificial intelligence; foreign language learning

Zusammenfassung

Der Vocabulary Loop ist ein didaktisches Modell, das beschreibt, wie aus einer ersten Begegnung mit einem Wort robustes Wortschatzwissen entstehen kann. Er umfasst sieben nicht strikt lineare Phasen: erste Begegnung, Aufmerksamkeit und Wahrnehmung, Bedeutungsaufbau, vielfältige Verknüpfungen, Abruf und Festigung, Anwendung und Automatisierung sowie erneute Begegnung mit Rückmeldung und Anpassung. Das Modell ist eine Synthese für die praktische Anwendung, die sich auf Forschung zu mehrdimensionalem Wortwissen, Abrufübungen, verteiltem Üben und Feedback stützt; es ist weder ein experimentell validiertes Stufenmodell noch die Beschreibung einer neurologischen Abfolge. Daraus werden ein Lernzyklus mit vier Schritten und fünf praktische Prinzipien abgeleitet. Anschließend wird geprüft, was Metaanalysen und Experimente zum digitalen und KI-gestützten Wortschatzlernen zeigen. Die Befunde stammen überwiegend von Jugendlichen und Erwachsenen, die Englisch als Fremdsprache lernen; für Kinder im Vorschul- und Grundschulalter ist die Evidenz dünn. Abschließend werden die Grenzen des Modells sowie altersdifferenzierte Anregungen für Lehrkräfte und Eltern dargestellt, einschließlich datenschutzrechtlicher Hinweise für den deutschen Kontext.

Schlagwörter: Wortschatzlernen; Wortschatzerwerb; didaktisches Modell; Abrufübung; verteiltes Üben; Feedback; digitale Medien; Künstliche Intelligenz; Fremdsprachenlernen

Introduction

The Vocabulary Loop can be understood as a didactic model that illustrates how an initial encounter with a word can lead to increasingly robust vocabulary knowledge. The following phases serve as a practical guide and should not be interpreted as a fixed neurological sequence. The model is a synthesis for practical use: it draws on research on multidimensional word knowledge (Nation, 2022; Nation, 2024), retrieval practice and spaced practice (Dunlosky et al., 2013; Cepeda et al., 2006), and feedback (Hattie & Timperley, 2007), but it is not an experimentally validated stage model. Learners are able to develop different aspects simultaneously, return to earlier phases, or understand a word in a specific context without yet being able to use it independently. What does it mean to be familiar with a word? Knowledge of words is multidimensional (Nation, 2022). A person may be able to identify a word they read or hear, but may not yet be able to speak or write it on their own. They may be familiar with a translation without having a firm grasp of the typical word combinations, grammatical structures, pronunciation, or nuances of meaning. Receptive knowledge (recognition and comprehension) and productive knowledge (autonomous, appropriate use) can develop at different rates (Nation, 2022).

An Overview of the Phases of the Vocabulary Loop

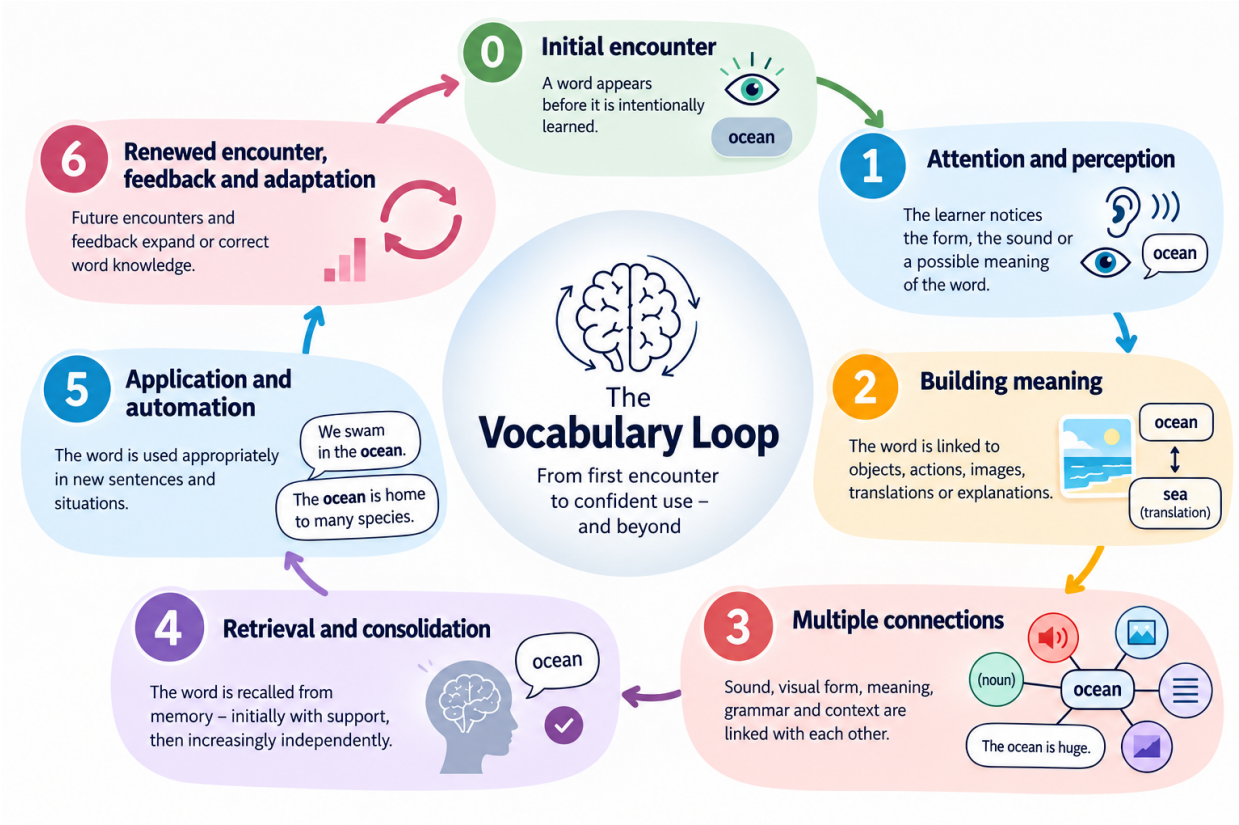


Figure 1: The Vocabulary Loop as a didactic model, from the first encounter to confident use (phases 0–6). This illustration is intended as a guide and does not represent a neurological sequence. The figure was created using ChatGPT (AI-generated)

- **0 – Initial Encounter:** A word appears before it is intentionally learned.
- **1 – Attention and Perception**: The learner perceives the word's form, sound, or possible meaning.
- **2 – Building Meaning:** The word is linked to objects, actions, images, translations, or explanations.
- **3 – Multiple Connections:** Sound, visual representation, meaning, grammar, and context are linked together.
- **4 – Retrieval and Consolidation:** The word is retrieved from memory—initially with assistance, then increasingly independently.
- **5 – Application and Automation:** The word is used appropriately in new sentences and situations.
- **6 – Renewed Encounter, Feedback, and Adaptation:** Future encounters and feedback expand or correct the learner's knowledge of the word.

This list (see also Figure 1) is provided for clarity. In practice, these processes are not strictly linear. A child may first pronounce a word, later learn how to spell it, and only much later understand a figurative meaning.

Phase 0 – Initial Encounter

Definition: The learner encounters the word incidentally, for example in a song, a story, a game, on a poster, or in a conversation.

Key concept: An earlier encounter can influence later attention and recognition; however, it does not in itself indicate understanding or lasting learning.

What happens? The learner may perceive the sound, spelling, or context without intentionally learning the word.

Why is this important? Repeated encounters can create opportunities to notice a word again. However, attention, processing of meaning, and further learning opportunities are also important for building robust vocabulary knowledge (Nation, 2024).

Duration: It can vary from a brief moment to repeated encounters over the course of days or weeks; there is no fixed duration.

Risk: The word remains background noise. If attention and meaning are lacking, casual exposure should not be considered a guaranteed learning success.

Example: A child hears “bridge” in a song. An adult caregiver later brings up the word again during a walk past a bridge.

Note: A song can facilitate a casual encounter with a word and lead learners to recognize its sound or spelling later on. However, exposure to a word alone does not indicate an understanding of its meaning or lasting learning. Conscious learning of meaning, subsequent active recall, and further encounters remain important.

Phase 1 – Attention and Perception

Definition: The learner consciously perceives the word and focuses on at least one aspect of its form or meaning.

Key mechanisms: Attention and targeted processing. Vocabulary learning requires learners to direct their attention to what they are learning; however, attention alone does not guarantee long-term retention (Nation, 2024).

What happens? The learner perceives the sound, sees the letters, observes a gesture, or recognizes the situation in which the word is used. The pronunciation and spelling do not need to be fully grasped right away.

Why is this important? An encounter without sufficient attention provides fewer opportunities for targeted learning.

Duration: Seconds to minutes per interaction; several brief interactions can be more helpful than an overloaded explanation.

Risk: Too many new words, distractions, time pressure, or unclear explanations can make processing more difficult.

Example: The teacher says “enormous,” shows a very large picture, and pronounces the word clearly.

Phase 2 – Building Meaning

Definition: The form of the word is linked to a comprehensible meaning. Various approaches can be created using pictures, actions, translations, examples, and child-friendly explanations.

What happens? The learner forms an initial idea of what the word refers to. For abstract or ambiguous words, this idea may be incomplete at first.

Why is this important? Reciting a word by heart without knowing its meaning is not enough for flexible use. Vocabulary learning involves targeted processing of meaning, repetition, and linking form and meaning (Nation, 2024).

Duration: This can range from a brief explanation to several learning and everyday situations, depending on the complexity and prior knowledge.

Risk: A single translation may be too narrow; images alone can be ambiguous and should be supported by verbal explanations when necessary.

Example: “Borrow” is not only translated as “leihen,” but is also used in the sentence “May I borrow your pencil?” and then explained.

Phase 3 – Multiple Connections

Definition: The learner connects various aspects of a word: pronunciation, spelling, meaning, part of speech, typical word combinations, and context of use. Depending on the word and learning objective, listening, speaking, reading, writing, movement, and context can be combined.

What happens? The word is not merely learned as an isolated label. The learner develops a more nuanced understanding of how it sounds and is spelled, which words frequently appear alongside it, and in which sentences it fits.

Why is this important? Vocabulary knowledge encompasses several aspects, including form, meaning, and usage; learning activities should take these aspects into account in a way that aligns with the learning objective (Nation, 2022).

Duration: Several short practice sessions over an extended period can provide opportunities for repetition and consolidation; the optimal interval depends on the learning objective and the circumstances (Cepeda et al., 2006).

Risk: A single activity may target only one dimension. For example, simply copying words does not necessarily promote pronunciation or spontaneous recall.

Example: For “decide,” learners listen to the pronunciation and read the word, practice “decide to do something,” and form their own sentence.

Phase 4 – Retrieval and Consolidation

Definition: The learner attempts to retrieve the word from memory without being immediately given a cue and receives feedback when necessary.

Key mechanisms: Retrieval practice and learning over an extended period. A review article evaluates practice tests and spaced practice as effective learning methods (Dunlosky et al., 2013). Retrieval can improve subsequent retention and serves not only to measure current knowledge (Dunlosky et al., 2013).

What happens? The learner identifies what is shown in a picture, recalls a translation, fills in a blank, or uses the word independently in a sentence. Effort is typically required; if mistakes are made, feedback after the recall attempt is helpful. The best time to provide feedback depends on the task and the learner (Hattie & Timperley, 2007; Metcalfe, 2017).

Why is this important? Recognition is often easier than free recall. Actively producing a word trains the process that will later be needed when speaking or writing (Dunlosky et al., 2013).

Duration: Short recall sessions, spread out over several days and repeated later (Cepeda et al., 2006). The appropriate interval depends, among other things, on how long the word is to be retained (Cepeda et al., 2006).

Risk: Simply looking at the answer can create a sense of familiarity without actually testing retrieval ability (Dunlosky et al., 2013). Hints and feedback can be helpful.

Example: First, the picture or meaning side of a flashcard is shown; the child responds; then the answer is checked together, and the word is tested again later.

Phase 5 – Application and Automation

Definition: The word is used with increasing confidence and appropriateness in meaningful sentences and various situations.

Key concept: Practice, active retrieval, and use in varying contexts can strengthen different aspects of vocabulary knowledge; the didactic model does not prescribe a fixed number of repetitions or a specific point in time at which automation automatically occurs.

What happens? The learner uses the word in dialogue, in a narrative, during a writing exercise, or when comprehending a new text. Availability may vary depending on the context.

Why is this important? The goal is not merely to work through a list of vocabulary words, but to apply words in comprehension and communication. Receptive knowledge may develop differently than productive knowledge (Nation, 2022).

Duration: Repeated use over weeks, months, and across new topics; a fixed number of repetitions cannot be derived from this.

Risk: Even if a word is successfully used in a practiced sentence, this does not mean it is already reliably available in new situations.

Example: A child has learned "borrow" in the sentence "May I borrow your pencil?" and later, during a role-play activity in the classroom, asks, "Can I borrow your eraser?"

Phase 6 – Renewed Encounter, Feedback, and Adaptation

Definition: The learner encounters the word again in later situations, uses it once more, and receives feedback that allows them to expand or correct their knowledge of the word.

Key learning activities: spaced repetition (Cepeda et al., 2006), exercises that support long-term retention and transfer (Soderstrom & Bjork, 2015), and feedback, the effectiveness of which depends on its type and design (Hattie & Timperley, 2007).

What happens? The word appears in new texts, conversations, and contexts. In the process, word combinations, nuances of meaning, and figurative uses may be added, and previous misunderstandings can be corrected.

Why is this important? Vocabulary knowledge develops gradually. Spaced exposure and helpful feedback can help ensure that a word remains accessible for longer and is used in new contexts (Cepeda et al., 2006; Hattie & Timperley, 2007).

Duration: This can span weeks, months, or longer; there is no fixed duration or fixed interval between repetitions.

Risk: A correct response immediately after learning provides only limited insight into what will be retained later (Soderstrom & Bjork, 2015). Feedback has varying effects depending on its type and design (Hattie & Timperley, 2007).

Example: "Run" is first learned as a verb denoting movement and later rediscovered in phrases like "run a business" or "the river runs." The new uses are explicitly explained.

What Neuroscience Can Contribute—and What It Cannot

The processing of word meanings relies on a network of several brain regions. A meta-analysis of 120 imaging studies identified a distributed network involved in semantic processing (Binder et al., 2009). This research describes the neural correlates of semantic processing, not the Vocabulary Loop or a fixed sequence of learning. No neurological sequence of the didactic phases can be derived from this.

A Learning Cycle Suitable for Everyday Use

The learning cycle is a practical aid, not an additional neurological phase. It translates selected learning principles into four steps:

- **Clarify meaning:** Link a word to an understandable example, image, object, or translation. Vocabulary learning involves attention, linking form to meaning, repetition, and processing usage (Nation, 2024).
- **Link form and usage:** Hear, read, and—if appropriate—pronounce the word; add typical word combinations or an example sentence (Nation, 2024).
- **Actively retrieve:** First recall the word independently or form a sentence, then check the answer and use helpful feedback (Dunlosky et al., 2013; Hattie & Timperley, 2007).
- **Encounter it again later:** Repeat the word at intervals and, if possible, in a new context. Research supports spaced practice; the optimal interval depends, among other things, on the retention interval, not solely on the assumed level of learning (Cepeda et al., 2006).

Five Practical Principles Derived from the Vocabulary Loop

The five principles summarize practical implications drawn from the learning processes described. They are not additional brain phases and should not be understood as a comprehensive program that has been tested on its own.

1. Consistency Instead of Cramming

Short, recurring learning opportunities can help keep words accessible over time. Research on spaced repetition supports this principle; however, it does not specify an optimal number of minutes per day for all learners (Cepeda et al., 2006).

2. Start Early and Practice Active Recall

Learners should not wait until their grammar is perfect before using a word. They can start with simple sentences and take advantage of opportunities to recall words on their own. Recall exercises can support long-term retention; they are not proof that every early speaking activity automatically leads to fluency (Dunlosky et al., 2013).

3. Learning Words in Context and Through Common Expressions

Associate words with their meanings, typical word combinations, and example sentences, rather than relying solely on an isolated translation. Vocabulary learning benefits from focused attention, repeated learning opportunities, and processing methods that align with the learning objective (Nation, 2024).

4. Use a Variety of Effective Approaches

Listening, reading, speaking, pictures, or movement can—depending on the learning objective—offer different ways to approach a word. Which combination is helpful depends on the learning objective and on which aspects of vocabulary knowledge are to be practiced (Nation, 2022).

5. Use Mistakes as Feedback

Mistakes can reveal which words a learner cannot yet recall with confidence. It is helpful to take the time to look up the word independently, check the answer, and try again. Mistakes alone do not guarantee learning success; what helps is combining them with clear feedback and another opportunity to practice (Hattie & Timperley, 2007; Metcalfe, 2017).

In summary: encounter, perceive, construct meaning, make connections, retrieve, apply, and encounter again later. These terms describe learning activities that are helpful from a pedagogical perspective; they do not represent a fixed neurological sequence.

Implications for Parents and Teachers

Parents and teachers can support the learning processes described without turning them into a rigid sequence:

- Give learners enough time to think and do not provide the solution immediately. If additional support is needed, hints can be offered step by step. This is presented here as a recommendation for teaching practice.
- Provide feedback that is as objective and specific as possible, and allow students to try again. What kind of feedback is helpful depends, among other things, on its nature and its relevance to the task (Hattie & Timperley, 2007).
- Do not assess only immediate recognition: check later whether the word can be recalled without a cue and used in a different context. The benefits of delayed testing can be derived from research on recall exercises and the difference between learning and immediate performance (Dunlosky et al., 2013; Soderstrom & Bjork, 2015).

Limitations of the Model and Unanswered Questions

There is no universal timeline by which every word is permanently learned. Retention depends, among other things, on the word itself, the learner, the practice method, and subsequent use. A correct response immediately after learning provides only limited insight into what can be recalled days or weeks later. Many learning techniques have been studied; their effectiveness depends on content, age, language proficiency, feedback, and learning objectives. The results of individual experiments cannot be readily generalized to every child and every classroom situation. The Vocabulary Loop is intended as a planning and discussion aid, not as a test, diagnostic tool, or strict sequence. The seven didactic stages are a synthesis for practical use and not an experimentally validated step-by-step model.

Research on Digital and AI-Supported Vocabulary Learning

Research on digital learning resources varies and must be evaluated based on the specific activity. An openly accessible meta-analysis on digital reading summarized 21 empirical studies on vocabulary learning in a second or foreign language and examined, among other things, additional digital resources such as word definitions, quiz questions, and audio (Zhu et al., 2024). These findings pertain to digital reading and do not demonstrate a general advantage of all digital tools or a specific advantage of AI. Nor do they confirm the phases of the Vocabulary Loop as a neurological stage model.

Who Was Studied? Age Groups and Validity

When evaluating digital tools, the most important factor is the age group to which the findings apply. Studies conducted, commissioned, or funded by app providers themselves are not included. The table presents a selection of the remaining studies; the selection is illustrative rather than exhaustive and does not claim to be a systematic review. Almost all findings on digital tools and AI come from adolescents and adults in foreign language classes; for preschool- and elementary-school-aged children, the data is significantly more limited.

Table 1: Groups Studied, Findings, and Limitations (Selected)

Study	Study Sample	Findings	Limitations
Yu & Trainin, 2022 (Meta-analysis)	34 studies (49 effect sizes), 2,511 participants, studies from 2006–2017; second-language and foreign-language learners: university students (20 studies) and K–12 students (13 studies, age range 6–18 years); one further study did not report the age range	Average overall effect (d = 0.64). University students d = 0.84; schoolchildren: d = 0.30. Mobile learning methods: d = 0.85; computer-assisted learning: d = 0.46.	Liberal inclusion criteria; asymmetric funnel plot (possible publication bias); primary and secondary levels not analyzed separately; “mobile” includes learning via text message.
Zhu et al., 2024 (Meta-analysis on digital reading)	21 studies, 77 effect sizes; foreign language learners; age groups are not reported separately; most participants had an intermediate language proficiency level; there were no advanced learners	Large effects, most pronounced for digital word definitions. Beginners benefited less than intermediate-level learners.	The authors describe their conclusions as preliminary; there is only one study on personalized reading programs.
Şimşek et al., 2026 (Meta-analysis on AI)	18 studies, 1,823 participants: preschool (3 comparisons), elementary school (2), secondary school (4), college (14); 21 of 23 comparisons involved English as a foreign language	Overall effect g = 1.309. Preschool: no significant effect (g = 0.029); elementary school g = 1.927; secondary school g = 1.358; college g = 1.548. The effect is stronger in the classroom than outside of school (g = 1.606 versus 0.758).	Very high heterogeneity (I^2 = 93.868%); small subgroups, particularly in elementary school; most comparisons involve university students; the authors urge caution.
Roseberry et al., 2014 (Experiment)	36 toddlers aged 24 to 30 months; new verbs	The children learned the verbs only in interactive situations (in person or via video chat), not from recorded videos.	Small sample size; a single learning step, no apps.
Bastani et al., 2025 (Field experiment)	Just under 1,000 high school students (three grade levels); subject: mathematics	When practicing with AI, performance was higher (+48% with unprotected access, +127% with protected access). In the subsequent test without AI, the scores of the group with unprotected access were 17% lower than those of the control group; with the AI tutor featuring learning protection, this decline was largely avoided.	Not language learning; there is no evidence that this applies to vocabulary.

The Meta-Analysis by Yu & Trainin in Detail

The meta-analysis by Yu & Trainin (2022) is the most comprehensive of the studies mentioned above. It breaks down the overall effect by age group, device, and learning format. All comparisons contrast digital offerings with instruction without technology.

- **Age:** University students benefited more than schoolchildren (d = 0.84 versus d = 0.30). The authors cite motivation and self-regulation, as well as a minimum level of language proficiency required to benefit from digital resources, as possible reasons. The studies involving schoolchildren covered the age range of 6 to 18 years; the meta-analysis provides no data for younger children (Yu & Trainin, 2022).
- **Device:** Mobile learning methods (apps and text messages) had a greater impact than computer-based methods (d = 0.85 versus d = 0.46). The authors suggest that immediacy, flexibility, and portability are the reasons for this (Yu & Trainin, 2022).
- **Learning approach:** Incidental learning in context—such as through games or digital communication—performed better than deliberate memorization, such as using word definitions or digital dictionaries (d = 1.04 versus d = 0.57). The value for incidental learning is based on only eight studies; the majority of the studies examined deliberate learning (Yu & Trainin, 2022).
- **Linguistic relatedness:** Learners benefited more when both their native language and the target language were Indo-European (d = 0.85) than when there was a great linguistic distance between them (d = 0.48) (Yu & Trainin, 2022).
- **Receptive and productive:** The effect on word recognition (d = 0.69) and on active use (d = 0.47) was numerically higher than the effect on active use, but the difference was not statistically significant (Yu & Trainin, 2022).

Individual Apps: Why There Are No Brand Comparisons Here

For well-known language-learning and flashcard apps, there are hardly any independent studies available that are large enough and methodologically sound. Several of the widely cited effectiveness studies were commissioned or funded by the providers, or authored by their employees; others have very small sample sizes or are purely correlational. Such studies are not cited as evidence here.

The following example shows how easily a correlation can be misinterpreted as a causal effect: In a non-experimental study involving 134 adults (mid-20s, college graduates, business English program), the group that used only a flashcard app initially performed best over a four-month period. In the subsequent model, however, only the time spent in the app was associated with test scores ($p = .005$), while total study time was not ($p = .24$); the advantage of the strategy itself disappeared (Daly, 2022). The study shows a correlation, not a causal relationship.

Hirsh-Pasek et al. (2015) had already pointed out that more than 80,000 apps were described as educational apps without any scientific standards in place; more than 72% of the top apps for sale in Apple's App Store came from the Toddler/Preschool category. Their framework identifies four "pillars" of effective educational apps: active involvement, engagement (staying on task without distraction), meaningfulness (relevance to everyday life), and social interaction, within the context of a learning goal. It is a conceptual framework and does not provide evidence of effectiveness for individual apps.

Statements regarding effectiveness therefore apply to categories—namely, digital learning offerings in general, mobile learning, digital reading with word definitions, and AI-supported methods—rather than to individual products. When choosing an app, the design should be more important than the brand: The decisive factor is whether the app encourages learners to retrieve words independently, provides spaced repetition and helpful feedback, and incorporates opportunities for conversation with other people (Cepeda et al., 2006; Hirsh-Pasek et al., 2015).

Possible Classification by Phase: A Didactic Framework

Phase 0 – Initial Encounter: Digital stories, songs, or videos can introduce words before they are specifically taught. However, early exposure alone does not demonstrate an understanding of meaning or lasting learning. Even with digital resources, a distinction should be made between exposure to a word and verifiable knowledge of the word (Nation, 2024).

Phase 1 – Attention and Perception: Digital activities can draw attention to sound, spelling, or word form. Focusing attention on what is being learned is a central principle of vocabulary acquisition, but it does not, on its own, guarantee long-term retention (Nation, 2024). Whether learners actually follow a prompt depends on the task and the situation.

Phase 2 – Building Meaning: Images, audio, dictionary functions, translations, and example sentences can provide learners with opportunities to connect word forms with meanings. The meta-analysis by Zhu et al. (2024) examines, among other things, digital resources for reading; it provides no evidence that a specific function automatically triggers meaning processing. The assignment to this phase is therefore a pedagogical classification and not a directly tested effect of a single feature.

Phase 3 – Multiple Connections: Some applications combine text, sound, images, and example sentences. From a pedagogical perspective, such combinations can be used to reinforce various aspects of word knowledge. Vocabulary knowledge encompasses various aspects, such as form, meaning, and usage; which combination is most effective depends on the learning objective (Nation, 2022). The meta-analysis by Zhu et al. (2024) describes digital resources in the context of reading, but does not examine the Vocabulary Loop as a model.

Phase 4 – Retrieval and Consolidation: Flashcards, fill-in-the-blank exercises, or quiz questions can facilitate active retrieval if learners first generate an answer on their own and only then see the solution. Multiple-choice questions often test recognition rather than free retrieval. General studies on recall support the learning principle, but not necessarily every specific digital implementation (Dunlosky et al., 2013). Spaced-repetition systems can support both recall and re-exposure and therefore overlap with Phase 6.

Phase 5 – Application and Automation: Dialogues and sentence-building exercises can encourage the use of words. A correctly completed dialogue alone does not demonstrate flexible application in new situations or full automation. A meta-analysis of AI-supported vocabulary instruction reports positive average effects on measured vocabulary performance ($g = 1.309$) with very high heterogeneity ($I^2 = 93.868\%$; Şimşek et al., 2026). The results do not confirm a fixed sequence of phases.

Phase 6 – Renewed Encounter, Feedback, and Adaptation: Digital systems can present words again, correct answers, or schedule review sessions. These functions are appropriate for this phase when they actually facilitate another encounter with the word or provide meaningful feedback. Spaced practice is supported by research on the intervals between learning opportunities; the effectiveness of feedback depends on its type and design. The studies cited do not prove that every digital system operates adaptively or that there is a universally optimal schedule (Cepeda et al., 2006; Hattie & Timperley, 2007). Spaced-repetition systems may overlap with Phase 4.

Limitations of the Findings

The classifications above represent a didactic interpretation of the learning activities described, not a direct experimental test of the Vocabulary Loop phases. The meta-analysis by Zhu et al. (2024) examined 21 empirical studies on digital reading and vocabulary learning. It supports a cautious conclusion that certain digital reading features—such as word definitions—can supplement learning opportunities. It does not prove that digital or AI-supported tools are generally superior or that the individual phases of the Vocabulary Loop have been scientifically validated.

A meta-analysis of Şimşek et al. (2026) reports an overall positive effect for AI-supported vocabulary instruction, but also a very high degree of heterogeneity ($I^2 = 93.868\%$). In the preschool setting, the effect was not significant ($k = 3$; $g = 0.029$). The small subgroup does not allow for far-reaching conclusions regarding children or individual digital activities. Overall, these results do not demonstrate an equally large benefit in every learning environment, nor do they point to a specific neurological phase.

Conclusion: Digital and AI-supported tools do not constitute a distinct phase of brain development. Depending on how they are designed, they can support individual learning activities. What remains decisive is what learners actually do and what learning outcomes were measured in each study.

Opportunities and Risks of Digital Tools and AI

The assessment depends heavily on the age of the learners.

Opportunities for adolescents and adults: Meta-analyses find, on average, moderate to large benefits of digital learning resources for foreign language vocabulary (Yu & Trainin, 2022; Şimşek et al., 2026). Digital tools can make retrieval, spaced repetition, and feedback more readily available (Cepeda et al., 2006; Dunlosky et al., 2013). The effects were greater in classroom settings than in out-of-school learning (Şimşek et al., 2026).

Opportunities for children: The evidence base is significantly weaker. For elementary school students, the AI meta-analysis reports large effects, but is based on only two studies; no significant effect was found for preschool children (Şimşek et al., 2026). In a small study of toddlers aged 24–30 months, children learned new words only through interactive engagement (Roseberry et al., 2014).

Generalizability: Almost all findings pertain to adolescents and adults in English language instruction; there is little evidence regarding the first language or younger children (Şimşek et al., 2026).

Outsourcing of thinking: In a field experiment with nearly 1,000 high school students, access to AI improved performance during practice (+48% with a standard chatbot, +127% with a tutor featuring learning protection). When access was withdrawn, the grades of the group with unprotected access were 17% lower than those of the control group; this group had used the AI as a "crutch" during practice. With the tutor featuring learning protection, this decline was largely absent (Bastani et al., 2025). The study focuses on mathematics. It remains unclear whether these findings can be generalized to vocabulary.

Short-term evidence: A large proportion of the studies cover only short time periods. In the AI meta-analysis, only 8 of the 22 comparisons that could be classified by duration (23 in total) lasted longer than one month (Şimşek et al., 2026). Little is therefore known about long-term retention.

Quality of content: Şimşek et al. (2026) point out that, to date, there have been very few studies comparing AI-generated learning materials with those created by humans, and that the quality of AI-generated content must be evaluated.

Age and data protection: UNESCO recommends a minimum age of 13 for independent interactions with generative AI and calls for data protection (Miao & Holmes, 2023). The KMK calls for the protection of students' personal rights and for the guarantee of their right to informational self-determination (KMK, 2024, p. 9). SWK points out that with commercial providers, it is often unclear where data is stored and whether German data protection regulations are being followed (SWK, 2024, p. 16).

State of research: Studies on learning effectiveness are available for intelligent tutoring systems and adaptive learning environments; according to the KMK, the state of research on generative AI applications is still insufficient (KMK, 2024, p. 4). The SWK also notes that there are still no reliable empirical findings on the impact of language models on teaching and learning processes (SWK, 2024, p. 8).

Recommendations for schools in Germany: The SWK recommends largely avoiding the use of language models in elementary school and at the beginning of lower secondary school, and instead focusing on the systematic development of basic reading and writing skills; if language models are introduced too early in text production, this hinders the development of these skills (SWK, 2024, pp. 11, 20). Starting in 8th grade, regular use of writing assistance should be considered, while students should continue to produce texts without a language model (SWK, 2024, p. 20). The recommendation pertains to text-generating language models, not specifically vocabulary apps. The KMK considers the SWK's recommendation to largely refrain from using such models in elementary school and during the early years of lower secondary school to be a matter for discussion, but emphasizes the importance of evidence-based and age-appropriate use (KMK, 2024, p. 4). When it comes to AI-supported learning, the KMK initially focuses on basic skills in German, mathematics, and foreign languages (KMK, 2024, p. 5).

Responsibility and performance assessment: The SWK requires that control and responsibility for assessment remain in human hands, and recommends distinguishing between sections of exams that do not allow the use of aids and those that do allow the use of AI (SWK, 2024, pp. 16–17, 20). According to the KMK, performance assessment is an educational and sovereign task that can be carried out exclusively by teachers; AI systems may only serve as decision-support tools. The KMK refers in this context to Article 22(1) of the GDPR, which concerns decisions based solely on automated processing

that produce legal effects concerning a person or similarly significantly affect that person. The KMK identifies data protection, a lack of transparency, and the risk of discrimination due to bias as key challenges (KMK, 2024, p. 6).

Reducing teachers' workload and setting requirements: The KMK sees opportunities to reduce teachers' workload in lesson preparation, delivery, and follow-up, and to strengthen formative assessment; the teacher's central role in shaping the learning environment remains unchanged (KMK, 2024, pp. 3, 6). At the same time, it calls for integrating the use of AI as an integral part of all three phases of teacher education (KMK, 2024, p. 8). The SWK emphasizes that the dynamic development of these tools poses particular challenges for teachers and recommends practicing and assessing students' proficiency in using AI as a learning objective (SWK, 2024, pp. 15–16).

Equity of opportunity and access: The KMK distinguishes three levels of the digital divide: material access to technologies (first level), the nature of media use and motivation to use them (second level), and the quality of effective use—that is, the ability to purposefully utilize digital resources for one's own learning (third level) (KMK, 2024, p. 11). It aims to provide data-protection-compliant and free access to language models in the school setting (KMK, 2024, p. 11). The SWK recommends integrating language models into suitable learning platforms or developing open-source alternatives whose functional scope and biases are known and can be freely researched; it notes that whether higher- or lower-performing learners benefit remains an open empirical question (SWK, 2024, pp. 17, 21).

Practical Examples for Teachers and Parents

The following examples are suggestions for instructional use. They are derived from the findings mentioned above but have not been tested as a complete package. The target audience indicated shows the age group for which the underlying research is most appropriate.

- **Understand first, then practice digitally:** The word "borrow" is introduced using an object, a gesture, and a sentence ("May I borrow your pencil?"). Afterward, learners practice recall using a flashcard app in which they first enter the answer before the solution appears (Dunlosky et al., 2013). Target audience: Elementary through secondary school; research on elementary school is limited.
- **Spaced repetition:** The same ten words are tested again in a short quiz after one day, after about a week, and later. There are no fixed ideal intervals (Cepeda et al., 2006). Target audience: elementary school age and up; studies on spaced practice primarily involve adults.
- **Digital texts with word definitions:** Clickable definitions are available while reading on screen. In studies, learners with intermediate language proficiency benefited more than beginners (Zhu et al., 2024); beginners need shorter, more heavily supported texts. Target group: Learners with intermediate language proficiency. The study does not specify the age groups.
- **AI with learning safeguards:** Learners first write their own sentence; the AI provides only follow-up questions or hints, not the solution. This corresponds to the tutor with safeguards described in Bastani et al. (2025). It must be clarified in advance whether the application complies with data protection regulations and whether the minimum age requirement is met (Miao & Holmes, 2023; KMK, 2024, p. 9). Target group: Adolescents aged 13 and older.
- **Review AI content:** Sample sentences, images, and explanations from AI tools are reviewed before use (Şimşek et al., 2026). Target audience: all age groups.
- **Plan for phases without tools:** In addition to tasks involving AI, there are tasks without tools in which learners use words from memory (Bastani et al., 2025; SWK, 2024, pp. 16, 20). Target audience: secondary school; for elementary school, the SWK recommends largely avoiding language models anyway.
- **Check for transfer:** Whether a word has been retained becomes apparent in conversation or role-play ("Can I borrow your eraser?"), not through app points (Soderstrom & Bjork, 2015). Target audience: all age groups.

Examples for Parents

- **Conversation instead of a video by itself (toddlers):** A video call with grandparents during which they talk about objects is more effective for vocabulary learning than a pre-recorded video that does

not respond to the child (Roseberry et al., 2014). Target audience: Children aged about two to two and a half years; small study.

- **Evaluate apps using four questions:** Does the app encourage active thinking rather than just tapping and swiping? Does the child stay focused on the learning objective without being distracted by sounds, movements, or side activities? Does what is learned relate to everyday life? Does the app stimulate conversations, for example, with parents or teachers? (Hirsh-Pasek et al., 2015). Target audience: Children; the framework is not limited to vocabulary and predates AI chatbots.
- **Ask questions together:** After app time, ask, "What new word did you learn? Can you make a sentence with it?" This helps the child recall and apply the word (Dunlosky et al., 2013). Target group: Children of elementary school age and older.
- **Practice briefly and frequently:** Several short practice sessions throughout the week are more effective than one long session all at once (Cepeda et al., 2006). Target audience: The studies primarily involve adults; the applicability to children has not been established.
- **Interpreting points and streaks:** Points and daily streaks indicate activity, not whether the word can be recalled later. Test recall without the app from time to time (Soderstrom & Bjork, 2015). Target audience: all age groups.
- **AI chatbots as guides:** According to UNESCO, independent conversations with generative AI are only recommended for ages 13 and up (Miao & Holmes, 2023). Younger children should only use them together with adults, should not enter personal data, and should review the responses together. Target audience: adolescents aged 13 and older; younger children only when accompanied by an adult.

Conclusion

Repeated exposure alone does not guarantee robust vocabulary knowledge. Comprehensible meanings, attention, connections, active retrieval, feedback, and reuse can create learning opportunities that support the development of vocabulary knowledge. As a practical guide, this cycle can be summarized as follows: encounter, understand, remember, apply, and encounter again later. Well-supported repetition can help make words more readily available later on; how well this works depends, among other things, on the word, the learner, the exercise, and the time interval. Digital tools and AI can support some of these learning activities. The findings on this topic come primarily from adolescents and adults; for younger children, personal interaction remains important (Roseberry et al., 2014; Yu & Trainin, 2022).

References and Further Reading

Bastani, H., Bastani, O., Sungu, A., Ge, H., Kabakcı, Ö., & Mariman, R. (2025). Generative AI without guardrails can harm learning: Evidence from high school mathematics. *Proceedings of the National Academy of Sciences, 122*(26), Article e2422633122. https://doi.org/10.1073/pnas.2422633122

Binder, J. R., Desai, R. H., Graves, W. W., & Conant, L. L. (2009). Where is the semantic system? A critical review and meta-analysis of 120 functional neuroimaging studies. *Cerebral Cortex, 19*(12), 2767–2796. https://doi.org/10.1093/cercor/bhp055

Cepeda, N. J., Pashler, H., Vul, E., Wixted, J. T., & Rohrer, D. (2006). Distributed practice in verbal recall tasks: A review and quantitative synthesis. *Psychological Bulletin, 132*(3), 354–380. https://doi.org/10.1037/0033-2909.132.3.354

Daly, N. P. (2022). Investigating learner autonomy and vocabulary learning efficiency with MALL. *Language Learning & Technology, 26*(1), 1–30. https://doi.org/10.64152/10125/73469

Dunlosky, J., Rawson, K. A., Marsh, E. J., Nathan, M. J., & Willingham, D. T. (2013). Improving students' learning with effective learning techniques: Promising directions from cognitive and educational psychology. *Psychological Science in the Public Interest, 14*(1), 4–58. https://doi.org/10.1177/1529100612453266

Hattie, J., & Timperley, H. (2007). The power of feedback. *Review of Educational Research, 77*(1), 81–112. https://doi.org/10.3102/003465430298487

Hirsh-Pasek, K., Zosh, J. M., Golinkoff, R. M., Gray, J. H., Robb, M. B., & Kaufman, J. (2015). Putting education in “educational” apps: Lessons from the science of learning. *Psychological Science in the Public Interest, 16*(1), 3–34. https://doi.org/10.1177/1529100615569721

Kultusministerkonferenz [KMK]. (2024). *Handlungsempfehlung für die Bildungsverwaltung zum Umgang mit Künstlicher Intelligenz in schulischen Bildungsprozessen* [Beschluss der Bildungsministerkonferenz vom 10. Oktober 2024, redaktionell überarbeitete Fassung März 2025]. https://www.kmk.org/fileadmin/Dateien/veroeffentlichungen_beschluesse/2024/2024_10_10-Handlungsempfehlung-KI.pdf

Metcalfe, J. (2017). Learning from errors. *Annual Review of Psychology, 68*, 465–489. https://doi.org/10.1146/annurev-psych-010416-044022

Miao, F., & Holmes, W. (2023). *Guidance for generative AI in education and research*. UNESCO. https://doi.org/10.54675/EWZM9535

Nation, I. S. P. (2022). *Learning vocabulary in another language* (3rd ed.). Cambridge University Press. https://doi.org/10.1017/9781009093873

Nation, P. (2024). Re-thinking the principles of (vocabulary) learning and their applications. *Languages, 9*(5), Article 160. https://doi.org/10.3390/languages9050160

Roseberry, S., Hirsh-Pasek, K., & Golinkoff, R. M. (2014). Skype me! Socially contingent interactions help toddlers learn language. *Child Development, 85*(3), 956–970. https://pmc.ncbi.nlm.nih.gov/articles/PMC3962808

Şimşek, B., Direkci, B., Koparan, B., Akbulut, S., Şimşek, E. E., Gülmez, M., & Ökmen, D. N. (2026). Artificial intelligence-assisted vocabulary teaching: A meta-analysis study. *Frontiers in Artificial Intelligence, 9*, Article 1806700. https://doi.org/10.3389/frai.2026.1806700

Soderstrom, N. C., & Bjork, R. A. (2015). Learning versus performance: An integrative review. *Perspectives on Psychological Science, 10*(2), 176–199. https://doi.org/10.1177/1745691615569000

Ständige Wissenschaftliche Kommission der Kultusministerkonferenz [SWK]. (2024). *Large Language Models und ihre Potenziale im Bildungssystem: Impulspapier*. https://doi.org/10.25656/01:28303

Yu, A., & Trainin, G. (2022). A meta-analysis examining technology-assisted L2 vocabulary learning. *ReCALL, 34*(2), 235–252. https://doi.org/10.1017/S0958344021000239

Zhu, T., Zhang, Y., & Irwin, D. (2024). Second and foreign language vocabulary learning through digital reading: A meta-analysis. *Education and Information Technologies, 29*(4), 4531–4563. https://doi.org/10.1007/s10639-023-11969-1

Author Information

Daniela Lentner: https://www.linkedin.com/in/daniela-lentner-800920445/