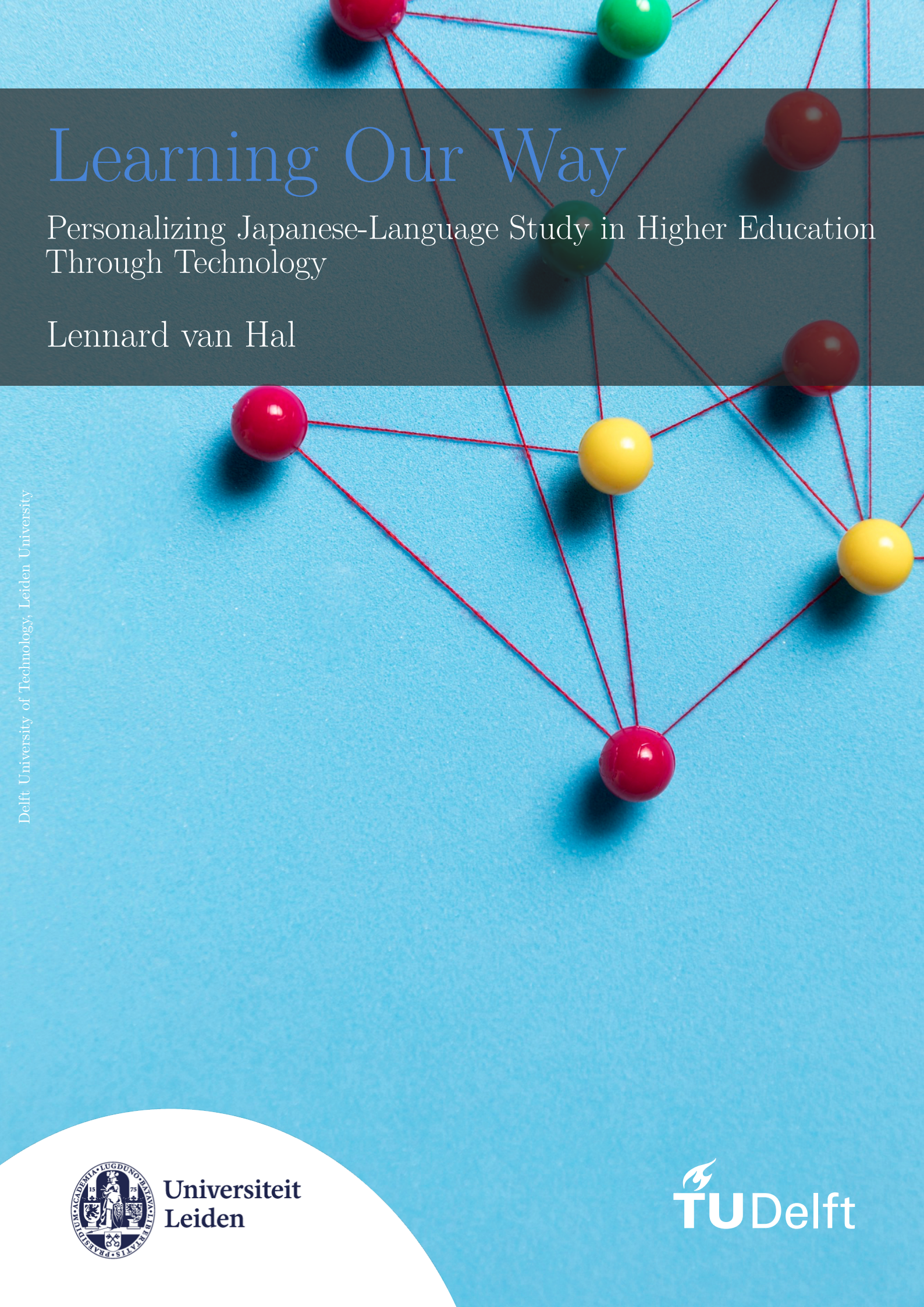

Learning Our Way

Personalizing Japanese-Language Study in Higher Education Through Technology

Lennard van Hal

Delft University of Technology, Leiden University

Universiteit Leiden

TU Delft

Learning Our Way

Personalizing Japanese-Language Study in Higher Education Through Technology

by

Lennard van Hal

to obtain the degree of Master of Science
at the Delft University of Technology,
to be defended publicly on Monday June 29, 2026 at 10:00 AM,
and the degree of Master of Arts
at Leiden University.

Project duration: November 11, 2025–June 29, 2026

TU Delft

Thesis committee:	Dr. Myrthe Tielman,	TU Delft, Interactive Intelligence
	Dr. Keiko Yoshioka,	Leiden University, Institute for Area Studies
	Dr. Gosia Migut,	TU Delft, Web Information Systems

Leiden University

First reader:	Dr. Keiko Yoshioka,	Leiden University, Institute for Area Studies
Second reader:	Dr. Tim Laméris,	Leiden University, Institute for Area Studies

Cover: Blue background and pushpin map, designed by Magnific

An electronic version of this thesis is available at `https://repository.tudelft.nl` and at `https://studenttheses.universiteitleiden.nl`.

Abstract

Software is regarded as a promising means to innovate education, particularly in light of a widely occurring transition from instructor-centered to student-centered instructional approaches. However, with a variously understood notion of student-centeredness and disjointed efforts to study the effects of technology in education, evaluating how technology may help students study a language in their preferred ways to chase their individual interests has remained subject to rigorous inquiry. This thesis takes the study of Japanese as a foreign language (L2) in Dutch higher education as a case in point to investigate through grounded design how technology can facilitate student-personalized learning. To this end, literature reviews were conducted on the topics of human learning, technology in language education, and L2 Japanese-language education. A focus group study was moreover performed with three university instructors of L2 Japanese, as well as an evaluation of a software prototype, *Kamo*, with students of L2 Japanese, to situate the theoretical findings in practice. Strong potentials were identified for software to arrange more individually attuned study materials and help students find peers with similar interests, under the condition that software complements rather than substitutes face-to-face instruction. A key requirement for personalizing technology to effectively integrate into language education is that the curriculum is adjusted to provide sufficient time and technical support to find and use adequate tools. Future research is recommended to take a larger-scale approach so as to prevent the recurring issue of non-generalizable outcomes and to define joint study objectives that can utilize the ever-changing technological landscape in a sustainable way.

Keywords—personalized learning, educational technology, L2 Japanese, higher education

要　旨

ソフトウェアは、講師中心から学生中心への広範な指導法転換を踏まえ、教育を革新する有望な手段と見なされている。しかし、「学生中心性」に対し様々な理解が存在し、教育におけるテクノロジーの効果を研究する取り組みが非協同動作であると考えられる。結果として、テクノロジーが、学生が好む学習方法で言語を学び、個々の関心を追求するのに、どのように役立つかを評価することは、依然として厳密な調査の対象になっている。本論文は、オランダの大学における第二言語（L2）としての日本語学習を事例に取り、テクノロジーが個別的学習にどのように貢献できるかを調査している。よって、人間学習、言語教育におけるテクノロジー、L2 日本語教育に関する文献レビューを実施した。さらに、論理的知見を実践に位置付けるために、三名の L2 日本語教員とフォーカスグループ研究を実施し、L2 日本語の学生を対象に「Kamo」というソフトウェアプロトタイプの評価を行った。ソフトウェアは、個別に合わせた学習教材を用意し、学生が同様の興味を持つ同級生を見つけられるよう支援する大きな可能性があることが特定された。ただし、対面授業に代わるのではなく補完的であることが条件である。尚、適切なツールを見つけて活用できるために、カリキュラムが充分な時間と技術的サポートを提供するように調整されることが、個別化可能なソフトウェアを効果的に言語教育に統合する要件である。今後の研究は、一般化できない結果という繰り返しの問題を防止し、変化しつつある技術環境を支持可能に活用できる共同研究目標を定義するために、より大規模なアプローチを取ることを推奨する。

【キーワード】個別的学習、教育的技術、第二言語としての日本語教育、高等教育

Preface

For a work that starts its introduction with a critique on writing, this thesis has perhaps turned out long. It has been unofficially in the making since early March 2025 when I first spoke Dr. Yoshioka about an idea to combine software with Japanese instruction as a fun study project. Many meetings, brainstorm sessions, and days devoid of sunlight later, it has been done.

Through this work, I have been able to connect many things I like: software development, Japan and Japanese, coffee, a duck (can you find it?), and interesting discussions. It has been an insightful journey.

I express my deepest gratitude to Dr. Yoshioka and Dr. Tielman for their guidance throughout this project, for helping me stay on track content-wise and bureaucracy-wise, and for putting up with increased organizational hassle through this thesis' setup. You have been true teachers in Chapter 2's spirit. Many thanks also to Dr. Migut and Dr. Laméris for spending your valuable time to read and evaluate this work. Thank you to all the instructors and fellow students who participated in my studies to share their valuable insights. Last but not least, thank you to those around me for the continued support, coffee, humor, feedback, and non-thesis-related activities, not necessarily in that order.

Enjoy the read!

Lennard van Hal
Papendrecht, June 15, 2026

Contents

1

Introduction

> Your invention is a potion for jogging the memory, not for remembering. You provide your students with the appearance of intelligence, not real intelligence. Because your students will be widely read, though without any contact with a teacher, they will seem to be men of wide knowledge, when they will usually be ignorant.
>
> *Socrates on the invention of writing* [100]

Learning a foreign language (L2) involves using and polishing a broad range of skills. Depending on how foreign a language is to a learner, one may need to learn new grammatical rules, one or more writing systems with new vocabulary, a phonetic system, a foreign culture, and an understanding of how one's native language may affect one's second-language production [107]. The extent to which each aspect of a language needs to be learned extensively, however, arguably depends on the learner's goals. One who aims to be a salesperson who needs to be able to write business emails in an L2 may be most interested in learning formal language and foreign business etiquette. One who aspires to be a tour guide abroad may want to focus on their spoken communicative skills and the foreign culture. And one who simply wants to understand what the karate forms are actually called in Japanese may be satisfied with learning what some words sound like.

For someone who does not live in an environment where their target L2 is spoken, what are the options to learn the language? A straightforward option would be for them to enroll in an educational program where the language is instructed based on decades of second-language teaching research that has been aimed at developing "the most effective teaching methods" [107, p. 2]. After all, at an early learning stage, the classroom can be seen as more effective for the purpose of language learning than living in a foreign country: instructors can curate study materials that should fit students' proficiency levels, whereas a learner in a foreign country would be confronted with all sorts of difficult-to-interpret language inputs [67]. At the same time, educational institutions are principally constructed to facilitate instruction en masse, not to enable an individual to learn toward their specific interests. Educational systems, in fact, *rely* on massified, standardized instruction to fulfill their purpose, which is to be "the training ground for good state subjects" [12, p. 20], or "to protect corporate profit" [50, p. xvi], or to foster democracy [124], among other purposes. To serve these purposes, standard curricula are employed with prescribed materials and objectives. Where do the language-learning karate practitioner and the aspiring tour guide and salesperson, each with their own individual goals, fit in such systems?

This question appears to be asked more frequently and widely, prompting initiatives that aim to point education in the direction of *student-centered learning* (SCL). On an international, political level, the European Union's Bologna Process, for instance, is supposedly placing increasing importance on respecting individual learners' interests and capabilities [30]. This has in turn led to a 2025 revision of the Dutch Higher Education and Scientific Research Act in an attempt to give students more room

to determine their own study path within higher education [27]. In addition to policy-driven initiatives, software is seen as promising to innovate language education, particularly toward SCL-oriented instruction. It is argued that software brings "innovative pedagogical approaches" [56, p. 251], that it will lead to "better pedagogical design and language education" [75, p. 374], that it will produce "[intelligent virtual agents] that are capable of serving as foreign language instructors" [83, p. 4], and so forth. Importantly, software may help open up possibilities *within* the educational systems to provide students with more personal guidance and attention, for instance by means of a personal virtual language tutor like *Robo-Sensei* [90, 122] or by allowing for personalized study paths that dynamically adapt to students' capabilities [76].

These developments, however, harbor at least three important problems. The first is that in practice, SCL appears to be less about the individual students' personal learning goals and students' agency in achieving those than it purports to be. Instead, it seems to describe a method of *instruction* rather than of *learning*, one that does not escape the standardized nature of education and thus does not necessarily enable instructional personalization for the aspiring salesperson and the to-be tour guide. It follows that SCL-oriented initiatives should not be understood to be synonymous with initiatives that promote personalized study in education. This will be discussed at length in Section 3.2. The second problem pertains to the software-based approaches. UNESCO's 2023 Global Education Monitoring Report on technology in education asserts that "it is debatable whether technology has transformed education as many claim" and that "evidence is mixed on its impact" [120, p. v]. It goes on to state:

> [W]hile there is much general research on education technology, the amount of research into specific applications and contexts is insufficient, making it difficult to prove that a particular technology enhances a particular kind of learning [120, p. 11].

In addition to being difficult to prove, one may wonder whether educational software products are principally aimed at fostering learning in the first place considering their highly financially incentivized nature [17, 33, 75, 95].[1] Third, even when software is applied directly in language education, there appears to be a misalignment between itself and the curriculum, as well as a lack of support in terms of time, technical assistance, and infrastructure. This strongly affects the efficacy of educational software, which will be demonstrated in the present work through a literature review on technology in language education (Section 3.3) and interviews with instructors and students (Chapters 4 and 6). These problems point at research gaps concerning the nature of personalized study and the applicability of software in language education for that purpose.

This thesis investigates how software *can* contribute to personalized language study within the existing educational systems. Specifically, the focus lies on university-level L2 Japanese education in the Netherlands, represented by the following research question:

(RQ) How can software facilitate student-personalized study of Japanese as a foreign language in Dutch university education?

From a language perspective, this presents an interesting case in point because the Dutch and Japanese languages differ significantly in many respects, including their scripts, grammars, and phonetic systems. The Japanese language also has an intricate system of honorifics that affects word choice and morphology depending on the sociocultural situation. Grasping these different aspects demands growth in one's linguistic as well as cultural understanding, but the extent to which a learner wishes to focus on one or the other could highly depend on their learning goals: the aspiring salesperson will need to deeply understand the honorific system, while the to-be tour guide may want to concentrate on non-formal speech. While a consideration of personal interests may be relevant to all levels of education, it may be especially so in the university context wherein students are already expected to make targeted specialization choices and have had more time than younger students to develop an understanding of what they are interested in or not.

To compose the main arguments, this thesis follows the process of grounded design as defined by Stevens et al., which holds that in "IT design, the central quality criterion is its fit with the social

[1] A notorious example is *Duolingo* (www.duolingo.com, accessed January 30, 2026), "[t]he free, fun, and effective way to learn a language!", according to the website—that is, free until one pays a subscription fee to get rid of advertisements and to unlock unlimited practice. Scrolling down, the first two information blocks explain that "research shows that it works." Extensive reviews like [110], however, suggest that this research is inconclusive and pays little attention to language performance.

practices to be supported" and that software designs are to be created upon a theory grounded in the context of such social practices [114, p. 35]—here, the practices of the Dutch university system. This approach is expanded upon through another definition of grounded design by Hannafin et al., oriented specifically toward the design of institutional "learning environments," emphasizing reflection on "psychological, pedagogical, technological, cultural, and pragmatic" foundations that underlie them [43, p. 104]. This thorough, grounded approach ensures the main argument takes into account the abovementioned complexities of education and educational software and their misalignment.

To this end, this thesis is structured as follows. First, Chapter 2 elaborates on the concept of grounded design and provides a background concerning the Dutch university system and the Japanese language. It also explains how seemingly similar terms like 'learner', 'student', 'teacher', and 'instructor' are differentially used throughout this work. Next, Chapter 3 presents three literature reviews to provide insights on the definition of human learning, technologies that have been integrated in language education, and specific instructional approaches that have been taken to instruct Japanese. These findings are situated in practice through a focus group study with instructors of L2 Japanese in Chapter 4. The theoretical and empirical results from Chapters 3 and 4 lead to the design and implementation of a software product in Chapter 5 which is subsequently evaluated in Chapter 6 through an experiment with students of L2 Japanese. Finally, Chapter 7 comprehensively discusses the findings, presents concluding remarks, and offers suggestions for future research.

2

Background

This chapter provides a background that will inform the interpretation of findings in the subsequent chapters. First, grounded design is elaborated on. Next, a description of the Dutch university system is given from the perspective of Dutch laws to sketch the higher-educational context in which **RQ** is situated, followed by a brief introduction of the Japanese language. Finally, terms related to learning and study used throughout this thesis are disambiguated.

2.1. Grounded Design

This thesis follows principles of grounded design introduced in Chapter 1 according to the definitions by Hannafin et al. [43] and Stevens et al. [114]. The definition by Stevens et al. is derived from a viewpoint of computer-scientific system design that "[builds] on situated design knowledge and [deals] with the interplay of the social and the technical" [114, p. 24]. The central idea is that "[t]he design of IT artifacts needs to take the given social practices into account" as well as the social consequences that follow the implementation of said artifacts [114, p. 31]. According to Stevens et al., grounded design is executed in three phases. The first is a context study where relevant social practices and existing software tools are described. In this thesis, this amounts to describing the educational system and the practices in it, practices surrounding learning (Japanese), and existing software products designed for language-educational purposes. This context study is conducted in the present chapter and Chapters 3 and 4. The second phase is a design study (Chapter 5) where the findings from the context study are taken as the basis to design a software product. The final phase is an appropriation study where the produced software is applied in practice and the impact thereof is analyzed to provide directions for future designs. This corresponds to Chapter 6 and subsequent discussions.

Grounded design as advocated by Hannafin et al. relates specifically to the "instructional design" of what they call "learning systems" [43, p. 102]. Grounding the learning system ensures that designers are able to explain and potentially improve upon their rationale for choosing certain foundations over others, which aids future refinements and helps situate designs within applicable contexts. The "roots of any learning environment," according to Hannafin et al., lay in the interrelated "psychological, pedagogical, technological, cultural, and pragmatic" foundations [43, p. 104]. In the theoretical and empirical studies that follow in Chapters 3 and 4, each of these aspects is attended to through thorough reflections on learning, educational systems and their cultures, existing education-technological artifacts, and pragmatic constraints and affordances brought forth by each of the preceding. Finally, "grounded designs are generalizable" and "validated iteratively through successive implementation" [43, p. 103].

A *Scopus* search to find studies that employ grounded design in the educational context yields three studies [19, 22, 64] (excluding [114] itself).[2] Each of the studies refers to grounded design as defined by Hannafin et al. [19] reviews ten studies that each use grounded design to propose hybrid class

[2] www.scopus.com. Publications were filtered in 2026 using the query `PUBYEAR BEF 2026 AND TITLE-ABS-KEY ( "grounded design" ) AND TITLE-ABS-KEY ( instruction OR education )`. The search retrieved 27 publications.

approaches in higher-educational contexts, observing that grounded design helps identify and overcome cultural and pragmatic challenges in the implementation of technology while grounding the choice of technology in established learning theories. [22] uses grounded design to prevent using technology to "[set] the direction of change" in an "ill-structured problem" where "it is not exactly clear what the problem entails, which interventions lead to a solution, and which constraints are imposed on that solution" [22, p. 1057]. Taking the grounded-design approach has allowed the authors to identify and describe their problem, design and implement a software solution grounded in theory and practice, and iteratively improve on the design after practical experimentation [22]. Lastly, [64] (co-authored by Hannafin) uses grounded-design foundations to propose directions to design case-based activities (CBA). The authors do not discuss generalizability and do themselves not take an iterative design approach, but by considering each of the grounded-design foundations, they are able to propose a defensible CBA design framework [64].

[22] demonstrates that a design can be grounded even if the grounding uses limited theory. The authors defend the limited use of theory by arguing that, complementary to theory, empirical observations are especially important to find solution directions for ill-defined problems: it facilitates that "theory does not invent the design but theory and design are interwoven" [22, p. 1059]. Furthermore, while each of the studies in [19] and [22, 64] do ground their designs in learning theories, they appear to presuppose efficacy of some *strand* of learning theories, such as constructivism. Grounding, in these cases, chiefly allows the studies to be defended within some chosen framework, but does not involve a *broad* examination of theory or practice, raising questions as to how the chosen designs relate to designs grounded in other frameworks and how they would deal with issues that other strands of theory and practice potentially bring to light. This thesis attempts to take a more thorough approach by taking a critical look at learning, education, and their theories and practices.

2.2. Legal Aspects of the Dutch University System

This thesis concerns personalization of L2 Japanese education within the Dutch university system. The funding, structure, and functions of official Dutch universities are regulated by legislation, particularly the Higher Education and Scientific Research Act (WHW) and the Implementation Decision WHW 2008 (IDWHW).[3] Since the regulations that shape the university system inevitably affect conduct within it, it is important to understand what regulations are in place. This section briefly introduces key legislation, focusing specifically on academic universities at the *wo*-level (hereafter referred to simply as universities).[4] For reference, law articles are mentioned where relevant and links to legislation are provided in footnotes.

2.2.1. Funding

Most Dutch universities are public and are funded for almost 60% by the state [103]. State funding is determined based on a yearly budget that is divided among all higher-educational institutions across five funding categories.[5] Two of the five categories pertain only to universities, which are respectively related to universities' educational functions and research functions. Funding for the educational functions is directly based on the number of enrolled students and graduates in each academic year. For enrollment-based funding, only students who are enrolled *within* the nominal program duration of their first bachelor's or first master's program are counted.[6] For graduate-based funding, only those are counted who complete their first bachelor's or first master's program, with exceptions for e.g. diplomas in the field of education.

Apart from state funding, universities are also funded by students. The yearly tuition fee is either a lower amount specified in the IDWHW (€2 694 for the academic year of 2026–2027)[7] or an amount set by the university institution, which is typically higher (e.g., €14 300 at Leiden University's Faculty

[3] Wet op het hoger onderwijs en wetenschappelijk onderzoek (wetten.overheid.nl/BWBR0005682) and Uitvoeringsbesluit WHW 2008 (wetten.overheid.nl/BWBR0006152) in Dutch, respectively. Accessed April 14, 2026.

[4] The WHW distinguishes three higher-educational institution categories: academic universities (*wo*-level for *wetenschappelijk onderwijs* 'academic education'), universities of applied sciences (*hbo*-level for *hoger beroepsonderwijs* 'higher vocational education'), and academic medical centers.

[5] IDWHW chapter 4.

[6] Typically, three years for a university-level bachelor's and one year for a master's program as per WHW article 7.5.

[7] duo.nl/particulier/collegegeld.jsp. Stipulated in WHW article 7.45a. Accessed April 15, 2026.

of Humanities for a first bachelor's in 2026–2027).[8] If a student is enrolled in their first bachelor's or master's program, the tuition fee is the lower amount; otherwise, it is the institutional fee.[9]

Finally, university students themselves receive state funding as stipulated in the Student Finance Act 2000, chapter 5.[10] The funding comprises public transport travel expenses, a basic grant, and optional additional grants depending on a student's circumstances. The funding is a loan, but is converted into a gift if the student receives a bachelor's or master's diploma within ten years after the first loan was allocated.[11] The loan duration that is convertible into a gift is equal to the nominal length of the programs a student follows; travel expenses are eligible for one more year of conversion. If the student does not manage to obtain a diploma within ten years, the amount that would otherwise have been converted into a gift must be paid back in full, with compound interest.

Two observations follow from these regulations. First, it is typically (much) cheaper for a student to limit themselves to nominally following one bachelor's and one master's program than two or more of either, unless they manage to uninterruptedly follow multiple programs of the same level at the same time or with overlap. Second, taking more time than the nominal program duration to finish a program yields less funding both for the student and the uiversity.

2.2.2. Enrollment Prerequisites

Students who enroll in a Dutch university's bachelor's program require a high-school diploma at level *vwo*,[12] a bachelor's degree or higher, or some other acknowledged diploma that signifies a student's official readiness, plus potential additional program-stipulated prerequisites.[13] Among additional prerequisites, an important one is a high-school student's *course profile*, which consists of a set of courses that a student chooses to follow from the fourth year of secondary education. To enroll in a master's program, one needs either a bachelor's degree or the skills and knowledge associated with a bachelor's degree, plus potential additional prerequisites.

2.2.3. University Duties and Purposes

While universities determine their own overall strategies, they are also bound by law to minimally "be aimed at facilitating scientific education and conducting scientific research," and to "in part pay attention to the personal development of their students and the fostering of their sense of social responsibility."[14] Universities are to provide their education in the form of educational programs, each consisting of a coherent set of educational units.[15] Each program and unit requires a form of final examination.

2.2.4. Innovation for Improvement of Higher Education

Higher-educational institutions are allowed to deviate from some WHW stipulations, including the articles in WHW chapter 7 referred to above, with the purpose of "[improving] the quality, accessibility or efficacy of higher education."[16] Such deviations are to be formally registered as "experiments" and have a maximum duration of eight years with the possibility of conversion into a new or revised law.

Notably, one such experiment that started in 2016 and ended in 2022 was conducted to introduce more flexibility in part-time and dual (consisting of a practical part) higher education.[17] Participating higher-

[8] www.universiteitleiden.nl/onderwijs/toelating-en-aanmelding/bachelors/collegegeld. Stipulated in WHW article 7.46. Accessed April 15, 2026.

[9] Only if a student follows a second bachelor's or master's program while still enrolled in their first do they, in general, have to pay the lower amount to *one* institution per year until graduation of the second program.

[10] Wet studiefinanciering 2000 (wetten.overheid.nl/BWBR0011453) in Dutch. Accessed April 15, 2026.

[11] Students can also *borrow* in addition to funding, which is always subject to interest and is not converted into a gift.

[12] Dutch high schools are divided into several levels as per the Secondary Education Act 2020 (Wet voortgezet onderwijs 2020 in Dutch, wetten.overheid.nl/BWBR0044212, accessed April 16, 2026). The level determines the duration of the high-school program and a student's possibilities for tertiary education. A program at *vwo*-level (*voorbereidend wetenschappelijk onderwijs* 'preparatory academic education') takes six years to complete and is intended to provide students with the necessary skills and knowledge to be ready for university education.

[13] WHW chapter 7.

[14] WHW article 1.3, members one and five; author's translation.

[15] WHW article 7.3.

[16] WHW article 1.7a; author's translation.

[17] The experiment is described in full at zoek.officielebekendmakingen.nl/stb-2016-145.html (in Dutch). Accessed April 16, 2026.

educational institutions were allowed to provision educational programs based on educational units that consist partly or fully of curriculum-independent *learning-outcome units*. Students were allowed to fill part of their program with study plans consisting of one or more learning-outcome units for a maximum of 30 European Credits per plan, enabling more flexible and modular programs. Participating programs additionally provided students with coaching to assist their planning. The principal goal of the experiment was to attract more students to part-time and dual programs in order to better respond to growing demands of workers with a higher-educational degree.

Positive outcomes resulted in a revision of the WHW, enacted in 2025 to accommodate for learning-outcome units in higher-educational programs, with the modification that also full-time programs are to be eligible if they consist for a significant part of practical elements.[18] This change is in line with changes in educational policy advocated for by the European Union's European Centre for the Development of Vocational Training since 2004,[19] and a joint effort to center education more around the student than the instructor as one of the points that has gained importance in the Bologna Process [30, 31].

2.2.5. Discussion

Dutch universities' duties and principal student admission requirements are stipulated by law. Public universities' funding is moreover largely provisioned by the state and importantly relies on students' nominal enrollment and graduation. In this system, it is financially disadvantageous both for students and universities if students spend more than the nominal duration on their university studies. The introduction of learning-outcome units in the 2025 revision of the WHW can be seen as a partial acknowledgment of the rigidity of this system. The revision introduces some flexibility in students' choice of courses they follow as part of their university-level program (in participating programs only). The principal aim of the revision, however, appears to be to allow more already-employed people to more easily fit higher education in their working lives, such that the rising demands for workers with a higher-educational degree can be met. The objective thus appears motivated economically rather than by some increased sense that programs should be more personally attuned to students. Notwithstanding, WHW article 1.7a provides room for future experiments wherein students are enabled to tailor their studies more closely to their personal interests.

2.3. Japanese Language

This section briefly introduces the Japanese scripts, pronunciation, basic grammar, and honorifics.[20]

Japanese is written in three scripts: ひらがな *hiragana*, カタカナ *katakana*, and 漢字 *kanji*. *Hiragana* and *katakana* are syllabaries with 46 characters each. The characters represent morae which can be understood to correspond to short syllables that are pronounced with a fixed length. For instance, the *hiragana* characters あ and か represent /a/ and /ka/, respectively, both pronounced equally long. In some vowel combinations or when a 長音符 *chōonpu* 'long-vowel mark' is added, the vowel's pronunciation is elongated, like in あー /a:/ or こう /ko:/. *Katakana* is frequently used for loanwords (e.g., コーヒー *kōhii* 'coffee') while *hiragana* is principally used for Japanese words.

The third script, *kanji*, constitutes characters that originate from the Chinese script and that have individual meanings, such as 人 for 'person' and 豆 for 'bean'. In principle, more elaborate *kanji* are built up from simpler *kanji* constituents or *radicals* (e.g., 森 *mori* 'forest' is built up from three times the simpler *kanji* 木 *ki* 'tree') [44]. *Kanji* usually have multiple possible pronunciations. The correct pronunciation depends on the word or phrase it is part of: for instance, 人 is pronounced *hito* on its own, but *jin* in 人工 *jinkō* 'artificial'. To facilitate reading *kanji* that a reader is not expected to know, characters' readings can be added in *hiragana* alongside *kanji*, called *furigana* (as in e.g. 漢字(かんじ)). The Japanese government recognizes 2 136 *kanji* as important for life in Japanese society, although many more exist [45]. All three scripts are used together when writing in Japanese. For non-native Japanese users and learners, Japanese scripts may also be rendered in Latin script or *romaji* (e.g., *kōhii*).

Japanese sentences follow the subject-object-verb order as opposed to subject-verb-object in e.g. En-

[18] See www.eerstekamer.nl/wetsvoorstel/36136_wet_leeruitkomsten_hoger (in Dutch). Accessed April 16, 2026.

[19] See www.cedefop.europa.eu/en/projects/learning-outcomes. Accessed April 17, 2026.

[20] Serving as an introduction, these facets are presented in a simplified manner and do not comprehensively describe the Japanese language.

glish. For various grammatical functions such as tenses and voices, verbs and adjectives are typically conjugated (e.g., -た *-ta* for past tense and -られる *-rareru* for passive voice). Conjugations can be agglutinated to combine different functions (e.g., -られた *-rareta* for past passive). Verbs are categorized in two groups that each have predictable conjugation patterns, plus one group with two irregularly conjugated verbs. Another distinction is made between transitive verbs (which take an object) and intransitive verbs (which do not take an object), where many transitive verbs have an intransitive counterpart (see e.g. [141]). For instance, the transitive 入れる *ireru* 'to put in' has an intransitive counterpart 入る *hairu* 'to enter'. These counterparts can be in different conjugation groups.

Other important grammatical functions are attached to words or word groups by means of trailing particles, for instance to indicate destinations (に *ni*), means (で *de*), the subject (が *ga*), and the object (を *o*). The particle は *wa* is the topic marker and indicates the overarching topic on a phrase-, sentence-, paragraph-, or discourse-level. Once a topic or subject has been indicated or is implied by the context, it need not be repeated or mentioned explicitly. For instance, in a conversation between Alice and Bob, it would be sufficient for Alice to ask 何が飲みたいの *nani ga nomitai no*, literally 'what is it that want to drink' without explicitly mentioning 'you' or 'Bob'; Bob may then reply コーヒーが飲みたい *kōhii ga nomitai*, literally 'want to drink coffee' without mentioning 'I'.

Finally, Japanese has a system of honorifics called 敬語 *keigo*. Honorifics are commonly associated with politeness, although they fulfill many more functions such as delineating social in- and out-groups [116]. They are marked through mechanisms such as simple prefixes like お- *o-* (e.g., in お名前 *onamae* 'your name') up to complex systems of verb alternatives to express humility (e.g., いただく *itadaku* 'humbly receive' versus もらう *morau* 'receive') or respect (e.g., いらっしゃる *irassharu* 'to be (respectful form)' versus いる *iru* 'to be'). It is argued that adequate use of *keigo* is essential to get along as an adult in Japanese society and businesses [29, 116].

2.4. Disambiguation of Terms

This thesis uses learning- and education-related terms that appear similar, but are deliberately distinguished from one another. While the terms' definitions and implications are elaborated on in Chapter 3, this section preliminarily introduces and disambiguates them.

First, in this thesis, a *learner* is considered different from a *student*: anyone who engages in activities that trigger some kind of learning in themselves is a learner, but only those enrolled in formal educational programs are seen as students. Consequently, the *act of learning* is treated differently from the *act of studying*: studying is seen as one's conduct aimed specifically at progressing through some educational program, while the act of learning is understood as engaging in some activity with the purpose of fostering learning in the individual. Here, it is useful to refer to what is called *incidental* and *intentional learning* in literature [53].[21] Incidental learning happens when one learns something without that having been one's explicit purpose; one learns intentionally when one learns as a result of deliberately setting out to learn. There is furthermore a distinction between *explicit* and *implicit learning*. The former refers to learning with the intention to use or point out structure in some information (e.g., learning language with "the aid of grammar rules"), while in the latter, one implicitly infers structure from the information (learning language "without the aid of grammar rules") [53, pp. 131-132]. Incidental and intentional learning on the one hand and implicit and explicit learning on the other are qualitatively different from each other, where the former distinction pertains to one's intention to learn and the latter to where one puts one's focus when learning.

Learning and studying are not mutually exclusive yet do not persistently overlap. When one studies to pass a language exam by memorizing vocabulary lists, it does not immediately mean that one learns how to use that vocabulary in practice or remembers all vocabulary after the exam; it also does not imply that one has learned nothing.

A distinction is furthermore made between *teaching* and *instructing*, and *teachers* and *instructors*. The act of instructing occurs within the educational system as one of its principal tools to inform students on the confines, regulations, and goals of the educational program or to otherwise facilitate students in

[21] The terminologies are variously used and understood in literature [53]. Krashen, for instance, whose particular case is further explored in Section 3.2, seems to refer to incidental learning as *acquisition* and to intentional learning simply as *learning* [67]. In this thesis, we settle with Hulstijn's definitions as proposed in [53].

advancing their studies. This act is performed by instructors who are officially related to the educational program. Teaching is done with the purpose of helping a learner learn, and anyone who teaches is a teacher; but teaching has only occurred if the targeted learner has, in fact, learned as a result. Teaching and instructing can thus, but do not have to, occur simultaneously: one who instructs can be a teacher, but is not if one does not foster learning in a learner; one who teaches can be an instructor if that what one teaches also advances a student's study. In extension, a student does not instruct,[22] but a learner can be a teacher if they somehow aid another's learning.

Students' advancement of study relates to *curricular objectives*, the formal goals imposed by educational programs that students must achieve to finish their programs. These are distinguished from *learning goals* that serve as points of proficiency or understanding toward which a learner sets out to learn. Curricular objectives can, but do not inherently, contain learning goals, and vice versa, but whereas curricular objectives are necessarily stipulated by educational programs, learning goals can be developed through programs as well as by individual learners. An important implication is that achieving advancement in one's studies is not necessarily synonymous with learning or achieving learning goals. Consider, for instance, a course's curricular objective that states that students are to master presenting about academic topics. Passing the course implies meeting the objective; but does it mean that all students have, in fact, learned how to masterfully present about academic topics, even in future courses or outside the educational system?

Education and *pedagogy* are moreover distinguished from each other. Pedagogy is considered the whole of practices concerned with teaching, whereas education is concerned with the practices of instruction. As such, a pedagogist is interested in the science of how a teacher can help a learner learn; an educationalist is interested in maintaining the educational system and facilitating students' passing through the system. Here, too, a pedagogist can be, but does not have to be, an educationalist and vice versa. Importantly, the distinctions between education and pedagogy and between instructing and teaching imply that the quality of instruction and the quality of teaching are not naturally equal, since their objectives and practices are—to smaller or larger degrees, but nonetheless fundamentally—different.

Finally, *personalized study* is considered the provisioning or performance of study attuned to individual students' needs, learning goals, or other personal preferences. This attunement can be of any degree and can be student- or institution-initiated. For instance, if an institution offers students various study materials that cater more to students' interests than a single textbook, or if students are allowed to pick materials autonomously, then this is personalization. Personalization is not limited to study materials but can occur in any facet of educational study, including assessment, curricular objectives, classroom interactions, et cetera. Moreover, personalization is not equal to *individualization* or *differentiated instruction* (the latter being importantly defined by Tomlinson [117]). Individualization implies that distinct instructional modes are devised for each individual student, or that students work individually. Differentiated instruction presupposes a key role for the instructor to guide classes [117]. Personalization, as used in this thesis, can involve either or neither and does not require the instructor or student to fulfill their roles in specific manners.

The distinctions introduced here are categorical—one is either an instructor or not; a teacher or not; a student or not; a learner or not. One can, however, (purport to) act in accordance with multiple categories at once: one can instruct *and* teach. What is important to understand in consideration hereof is that seemingly alike acts or professions may serve significantly different purposes, and that this is not always acknowledged in educational, pedagogical, or other related research.[23] It follows that our understanding of human learning is potentially (adversely) affected, which should be taken into critical consideration when one investigates learning and education, as is done in the rest of this thesis.

[22] Unless, perhaps, the student happens to be an instructor as well.

[23] Hannafin et al. devote a brief paragraph to differences between learning and instruction and lament that they are "seldom differentiated accordingly," but their discussion is situated more strictly within one on instructional versus constructional "learning systems" [43, p. 102] and should be understood to entertain different connotations from those presented here.

3

Literature Review

To build toward an answer to **RQ** posed in Chapter 1, this chapter addresses the following subquestions:

(RQ1) How is human learning defined and how does education facilitate it?

(RQ2) What kinds of technologies have been integrated in language education, for what purposes, and with what consequences?

(RQ3) What instructional approaches have been taken to cater especially to institutional L2 Japanese study contexts?

Each question is approached through its own literature review following the method elucidated in Section 3.1. Findings are subsequently presented in a separate section per topic.

3.1. Method

To answer **RQ1**, a narrative literature review [62] is conducted in Section 3.2, focusing on discussions concerning learning, education, and the roles of students and instructors. **RQ2** and **RQ3** are approached by means of systematic literature reviews [62] to construct a thorough image of existing technologies and pedagogies in the context of (Japanese-)language education. Scopus is used to find and filter relevant literature for each of these topics, and the *Leiden University Libraries Catalogue*[24] is additionally used to find Japanese-language literature to answer **RQ3**.

In the systematic literature reviews, regardless of the topic, literature is filtered (1) on publication year, which should be after 2014 to ensure the literature is up-to-date, yet includes studies from before the introduction of generative AI tools like *ChatGPT* in 2022 [94]; (2) on language, where publications that are not in English, Dutch, or Japanese are disregarded; and (3) on publication type, such that only articles, book chapters, and conference proceedings are retrieved. Furthermore, publications should be peer reviewed.[25] Each topic additionally has its own set of filter queries and inclusion and exclusion criteria, listed in Chapter A.

Scopus and the Leiden University Libraries Catalogue were searched in early 2026 according to **RQ2**'s and **RQ3**'s filter queries. Publications resulting from this search were then, after removal of duplicates, manually evaluated against each topic's inclusion and exclusion criteria by the author, performing initial (dis)qualification based on publication titles and abstracts, followed by full-text evaluation against the criteria. The resulting numbers of studies at each stage are reported in Sections 3.3 and 3.4.

3.2. Human Learning and Learning Methods

Before reflecting on a software design in the realm of language education, this section discusses what we mean when we refer to human learning and education. It also substantiates why the distinctions in

[24] catalogue.leidenuniv.nl.

[25] Some Japanese-language articles in **RQ3** are bulletin papers whose reliability is expected based on university endorsement.

Section 2.4 are warranted. Additionally, insights in the roles of and relationships between instructors and students in the educational context are discussed.

3.2.1. Views on Learning

The question of how learning is defined and perceived does not have a straightforward answer. Learning is attributed several definitions, each in some way related to processes leading to a change in a learner, induced by external stimuli and resulting in the cognitive attainment of something new [68]. The question of what these processes look like has been approached in various fields of study including philosophy, psychology, cognitive science, and neuroscience. Each discipline highlights a different aspect of learning processes and various viewpoints exist within each discipline. Some regularly occurring names in the contexts of development, learning, (language) education, or some combination thereof are John Dewey (1859–1952), Maria Montessori (1870–1952), Jean Piaget (1896–1980), Lev Vygotsky (1896–1934), and Stephen Krashen (b. 1941). Their works are too vast, varied, and complex to discuss them in great depth in this thesis. However, as will be shown, the prominent works that these figures and their disciplines have brought forth have significantly influenced how learning is viewed and how education is shaped. It is suggested that this, unfortunately, has not always been done with proper justification: it appears to be common that bits and pieces of influential works are taken out of their context and applied in pedagogical and educational theories without relevant grounding. Due to the profound impact of the original works and their 'misappropriations' on how learning and education are treated, inescapably affecting the studies discussed in this thesis, it is essential to the central argument that some of these influences and their corresponding views are discussed here.

First, we turn to philosophical and psychological views on learning. Since the late-nineteenth century, Dewey's works have significantly influenced how learning is perceived—his works are still considered "*revolutionary*" today [101, p. 347]. In Dewey's view, people learn from experience, communication, and participation in group events [125]; it is central to his argument that learning is essentially governed by social structures and interactions. All actions happen, namely, in some social context and can only be interpreted relative to that context's norms; by participating in social settings, people *learn* these interpretations and norms [125].[26] Vygotsky similarly observed how children develop through contact with others in social activity [69, 91], as did Montessori who maintained that "children's activity and experience [are] a prerequisite for learning" [2, p. 120]. The (understandings of the) ideas of Vygotsky, though, have perhaps most significantly influenced the studies discussed in this thesis: of the fifty-six studies handled in Section 3.3, eight directly cite Vygotsky (compared to three citing Dewey, two Piaget, and zero Montessori), and at least twenty directly refer to concepts like *scaffolding*, *sociocultural theory*, and the *zone of proximal development* (ZPD), each frequently attributed to Vygotsky [91]. These notions of scaffolding and ZPD are popularly understood to imply that there are things that people cannot learn by themselves, but they become able to if some more knowledgeable other provides 'scaffolds' through which learners can reach previously unattainable knowledge.

It can thus be said that for Dewey, Vygotsky, and Montessori, if anything is key to learning, it is social interaction. Piaget, on the other hand, placed an emphasis on biological development, which, according to him, "*explains* [emphasis added] learning […] contrary to the widely held opinion that development is a sum of discrete learning experiences" [98, p. 58]. Hence, in his view, for any learning to take place, a certain level of biological development needs to have come first. To this end, Piaget divides a child's development into four age-related stages [98]; each subsequent stage builds upon the antecedent stages and a child can only develop into next stages if they successfully attained the previous [68]. The stage a child is at indicates what kind of cognitive structures it is capable of understanding or learning: at the first stage, children develop sensory-motor functions, and only at the fourth stage, they develop the capacity to "reason on hypotheses" [98, p. S9].[27] Against this biological backdrop, learning occurs by means of "active assimilation" or "the integration of any sort of reality into a structure," but only "if you base the more complex structure on simpler structures, that is, when there is […] not simply an

[26] One might say that Wittgenstein's examples of *language games* illustrate this reasoning in terms of learning a language, where a game's rules that make language interpretation possible can be learned "by observing [game participants] play and speak" [14, pp. 21-22].

[27] Notably, Piaget also views lingual communication as "fundamental" to attain new stages of development, but remarks that it "is insufficient because the child can receive valuable information via language or via education directed by an adult only if he is in a [biological] state where he can understand this information" [98, p. S12].

external reinforcement" [98, p. S17].

While the above views pertain to learning and development in general, Krashen particularly influenced how language learning is understood through his *input hypothesis*, with implications for first- and second-language acquisition [66]. The hypothesis posits "that humans acquire language in only one way—by understanding messages, or by receiving 'comprehensible input'," which must happen at times that "[t]he acquirer [is] 'open' to the input" [67, pp. 2-3]. Language input is comprehensible when it is at most at level $i+1$ if one has attained an understanding of level i, implying there are stages to language learning and that they follow some natural, predictable order (the *natural order hypothesis*) [67]. Being 'closed' to comprehensible input means being "unmotivated, lacking in self-confidence, or anxious"; being 'open' to it means not fearing failure and being more involved in the message than in the form [67, pp. 3-4]. Krashen distinguishes between acquisition and learning when he provides evidence that "acquisition can occur without learning" [66, p. 442]. When he writes of learning in this context, he implies studying under formal instruction, suggesting that by acquisition he refers to incidental learning and by learning he refers to study (see Section 2.4).

The above visions seemingly attempt to reveal behavioral or cognitive patterns based on what we see people do from the outside. But what happens *inside* one's brain when one learns, i.e., cognitively or neurologically? There have been profound influences from findings with neurocognitive-scientific origins on how learning is perceived. Perhaps the concept that has gained the most traction in the past century is that of *learning styles*, which points to supposed preferences in individual learners to learn based on visuals, audio, or kinesthetic movement (VAK) [32, 48]. Another is the idea that learners have a "hemispheric dominance" [20, Tab. 1] such that the brain's left or right portion more actively contributes to learning than the other [40]. Such ideas have come into widespread use to guide student learning, owing to a growing urge to acknowledge students' uniqueness and to integrate what we assume to know about the brain in theories on learning [48]. Today, many such ideas are increasingly understood to be unjustified or unrelated to what learning entails neurologically and practically [32, 40, 48]. Some are argued to be blatantly incorrect, such as the idea that people only make use of ten percent of their brain capacity [40]. In the discussion on education in Section 3.2.2, it will become clear why misinformed use of so-called *neuromyths* is potentially detrimental.

It is not strange for there to be an increasing interest to approach learning from a neurological side: increasing technological capabilities have enabled us to gain clearer images of the inner workings of the brain, and the brain surely has *something* to do with learning. Recent neuroscientific insights may provide more concrete evidence to support established or design new learning theories. This is a welcome addition in light of criticism that learning-related theories like those discussed above, or their applications, are essentially based on classroom observations [48] (and see e.g. [91] on Vygotsky) and self-reports [48]. Additionally, researchers have long aspired to construct a "unified account of developmental changes in higher and more basic cognitive processes," which is perhaps one of Piaget's legacies [11, p. 11]—neurological insights may help further this construction.

In fact, findings from theoretical neurobiology have recently led to the suggestion of a state-of-the-art "unified brain theory" referred to as the *free-energy principle* [35, p. 127]. Essentially, in the context of the brain, the free-energy principle revolves around a continuous minimization of *surprise* or *sensory entropy*—conversely, a maximization of *value*; an optimization of energy consumption—where "[a] fish that frequently forsook water would have high entropy" [35, p. 127]. People facilitate this minimization through *active inference*: by taking action or modulating perception in a generatively inferred way based on one's senses and internal belief state [35]. These notions should be understood to pertain to theoretical, mathematical descriptions of the principle, but intuitively, when projected onto the realm of consciousness, active inference manifests itself in "feeling our way in darkness: we anticipate what we might touch next and then try to confirm those expectations" [35, p. 129].

Since the free-energy principle reflects a *unified* brain theory, it suggests that *everything* a human does is essentially motivated or regulated by a natural tendency to biologically minimize surprise. It thus potentially brings profound implications concerning not only how a person physically navigates and perceives the world, but also how and why a person learns and thinks (see e.g. [58] for implications on language acquisition). But how do cognition and consciousness fit in a biological model that pursues surprise minimization? Recent simulations suggest that cognition may serve as a 'meta-controller' of sorts that constantly monitors surprise levels and (dis)engages deliberation accordingly to decide

whether or not action is required [102]. This mechanism is called *cognitive control* [102].

The free-energy principle furthermore hints at a hierarchical functioning of the brain and sensorium. This is suggested theoretically by the cognitive meta-controller's potential role [102] and physically by the "hierarchical arrangement of cortical sensory areas" [35, p. 129]. Conceptually, the idea that *cognition* is layered hierarchically has been proposed independently of the free-energy principle and is intuitively illustrated in a "layer-cake model" in [111, p. 4, Fig. 2]. In this model, cognition operates on (1) a low-level topological layer that solely connects concepts without information on the nature of the connections, (2) a higher-level causal layer that concerns the causal nature of concept relationships from the topological layer, and (3) a complex, high-level metric layer "which elaborates causal structure in a suitable, spatial and temporal setting" [111, p. 4].

Under the assumption that these theoretical and conceptual models are accurate, there are several implications concerning the views on learning and development formulated by the likes of Dewey, Piaget, Vygotsky, and Krashen. First, insights such as those from the free-energy principle are much more neurologically informed than the nineteenth- and twentieth-century psychological views on human development, which importantly base themselves on behavioral observations. There is only so much to be deduced about humans' inner workings through behavioral observations from the outside. Consequently, it is not surprising that the described psychological views are regarded as wildly different at some points. Without concrete knowledge of brain neurology, it is arguably difficult to make accurate and, especially, *consistent* statements concerning human cognition and development. This difficulty is further intensified by the fact that the abovementioned psychological views arose in different social, cultural, and political contexts. These views, in a way, have been filling separate parts of a puzzle whose final form is only now starting to become clearer. To grasp the whole puzzle, it seems counterproductive to view the various psychological accounts of cognition in isolation, let alone as 'rival views'.[28]

A second implication is that various of these observation-based statements on cognition in fact appear consistent with the free-energy principle and layer-cake model. For instance, Piaget maintained that biological development precedes learning [98], which coincides with the idea that biological and physical development of the brain provisions its capacity to perform more complex, higher-layer cognitive functions [111]. It has moreover been suggested that the free-energy principle is already at play *before* birth to steer a fetus' behavior [89] and hence their internal belief state, in a stage where they have not yet been in (cognitive) contact with social constructs. Similarly, Vygotsky's emphasis on the role of social interaction in development [123] can be defended if social interaction is seen as a natural act that serves to minimize expected surprise or cognitive conflict by updating one's internal state [35, 102]. The various aspects of Krashen's input hypothesis [67], too, seem plausible in light of the free-energy principle and cognitive layer cake. The natural order hypothesis can be seen to reflect the layered cognitive hierarchy; the affective filter hypothesis to correspond to one's state of being in need of cognitive conflict resolution or not; the input hypothesis itself to respect that beings minimize expected surprise by sensing the world, where Krashen's comprehensible input serves to update people's internal belief state particularly concerning language. Krashen's argument that "the silent period"—the initial period after e.g. moving to an area where a foreign language is spoken, wherein especially children tend to remain silent—"is not pathological, but normal" [67, p. 9] moreover seems to follow from the free-energy principle: it would simply be a period that children naturally use to update their internal state according to what they sense in their new surroundings in an attempt to minimize surprise. The above consolidations are incomprehensive and subject to more rigorous validity assessment, but the important takeaway is that all these diverse theories and observations can, in the end, be seen as sides of the same coin.

A third, significant implication is that emotional concepts such as *motivation* and *curiosity*, regularly a focal point in pedagogical studies, are all manifestations of a natural tendency to minimize surprise [58]. As will be demonstrated in Section 3.3, enormous efforts have gone into finding ways to consistently motivate people to learn, or students to study. Any such efforts that are successful, even if only at times, are undoubtedly beneficial to learning and study. But at times that students are unmotivated to study and explore *despite* such efforts, it may not only be inefficient to subject students to instruction: it simply would not make sense in light of the free-energy principle if the principal aim is to have

[28] See e.g. [123] for a similar argument concerning demerits of a Vygotsky-Piaget *dichotomy* rather than an assimilation or a unification.

students acquire knowledge or understanding. It would not bring students to a state of cognitive conflict, and so the cognitive meta-controller [102] would not trigger deliberation; there would be no biological motivation for the brain to profoundly update its inner belief state, i.e., to acquire what is instructed. This, too, is presently a preliminary conception, but it is clear that the neurologically-based views on cognition warrant reexamination of pedagogical studies *and* practices that are concerned with attitudes and emotions.

In conclusion, there are several facets to the answer of the question of what learning is. There is a neurobiological side that pertains to the brain's functioning. The recent free-energy principle is a seemingly plausible starting point from which we may understand this. Neurologically, learning can be understood to constitute any change in one's brain that follows from the (subconscious) processing of sensory information or belief states. Then, there is a cognitive side that follows from the brain's neurobiological mechanisms, which implies that one's cognitive state can be understood as a reflection of the brain's physiological state. Cognition is suggested to be built in layers, where the lower layers relate to more primitive and the higher layers to more complex understandings. While the brain's neurobiological mechanisms can be assumed to be fundamentally the same for all, it is implied that cognition develops more distinctly among people, especially at the higher layers. Learning, at the cognitive level, is the restructuring of information in any of these layers. Finally, distinctly developed cognitions contribute to person- and situation-specific behavioral outputs. Learning in terms of behavior can be understood to correspond to making or undergoing changes in behavioral patterns.[29] And since one's behavior or attitude can be seen as a reflection of one's current cognitive state and thus of one's brain's physical state, the free-energy principle invites us to unite and reexamine the diverse behavioral descriptions and their implications brought forth in psychology.

Next, it is discussed how education is understood, how it takes form in practice, and to what extent it takes the abovementioned perspectives on learning into account.

3.2.2. Views on Education

Like learning, education has been approached from various angles. First, it is argued that the philosophical question of 'what is education' has been prompted as early as Plato [12, 50]; that it is a question that warrants critical introspection at all times since education is not static, but rather subject to contemporary discourses that transform through power configurations and "dominant interests" [12, p. 12]. Education is, as such, constituent and driver of discourses in the Foucauldian sense [26]. This has the implications that the state and (capitalist) powers influence educational practices, that knowledge distributed in education is all-encompassing [26, 50], and that it hence shapes "our sense of reality" [70, p. 829]. It is argued, for instance, that in large parts of the world today, it is the hegemonic neoliberalist ideology that determines the form and purpose of education [12, 23]. Any pedagogical side of education is as such aimed at rendering its students ready for capitalist society and working life [12] (i.e., as opposed to *critical* or *wise* [113]) and at protecting corporate interests [50].

In the context of the above observations, education, serving political purposes, is implemented according to formally defined educational systems. In the case of Dutch higher education, this is demonstrated by the laws discussed in Section 2.2, which prescribe what higher education should do, for and by whom, where, and in what ways, thus forming a system; there are similar systems for primary, secondary, and other forms of education. Practically, the higher-educational system is operationalized in institutions like universities, which in turn have their own curricula and policies [23], ways of distributing knowledge, and the state apparatus to back them [12]. Education is hence experienced through (state-funded) institutions with instructors and students and strategies and objectives. In this way, not only do the educational systems operate according to the law: they form and propagate a "grammar of schooling," an institution-level discourse that determines "the very structure of school itself" [109, p. 81].

What is the relation between education and learning? This depends on one's definition of learning. In Dewey's terms, people learn from experience and communication, so intuitively, a student who partakes in education automatically learns by experiencing things in groups. Dewey also had strong views on *democracy* and linked democracy to education and vice versa [50]. The sustenance of a democracy relies on how people live in a society; thus, an important task of education is to raise people such that

[29] It is perhaps at this layer that a distinction between incidental and intentional learning can be made (see Section 2.4) since learning with intention can be seen as a question of behavior.

they understand how to behave in society [124]. As such, "[t]he best way of educating, [Dewey] argues, is by engaging young people directly in activities that are similar or related to adult occupations" [124, p. 6]. Montessori similarly advocated for children's undertaking of "adult tasks" in the school environment and facilitated this by furnishing the school "to resemble a miniature home" wherein students engage in "sweeping the floor, dusting the shelves, tending the flowers," et cetera [2, p. 119]. In these views, education relates to learning by deliberately placing students in an environment wherein they systematically observe or partake in adults' social conducts.

What about learning things that are arguably more cognitively complicated than sweeping the floor, such as a language? Krashen maintains that the classroom is an ideal environment to systematically provide students with comprehensible language input, especially in the earliest stages [67]. Provisioning of comprehensible input is typically done by instructors based on certain materials. But, Krashen laments, students are all too often subjected to vocabulary drills, which is not only less efficient than incidental language learning, but also less fun [66]. Krashen suggests students instead read books or other light materials in the target language and that instructors are to point them toward suitable materials that provide enough comprehensible input [66]. What Krashen's view makes clear in addition to the abovementioned ones is that the educational system brings instructors who, owing to their expertise, can present students with materials from which they may learn.

Section 3.2.3 will discuss the role of instructors in a student's study more elaborately, but here, it is worth pointing out that Vygotsky is regularly cited in this context, as alluded to in Section 3.2.1. A widely circulating thought is, namely, that "Vygotsky's educational theories" [5, p. 1585] have brought us the notions of the ZPD and scaffolding (e.g., [37, 63, 136]) which supposedly suggest that instructors serve to scaffold reachable information for a student to learn. This, however, is suggested to be misplaced because "Vygotsky wrote very little on education" and "Vygotsky himself did not see the zone of proximal development as either important or as an original aspect of his work" [91, p. 11]. In fact, "the ZPD […] is not about learning *per se* but links to his wider notion of development" [91, p. 11]. The problem here is that a vast number of authors in the educational and pedagogical spheres is argued to pick some terms once allegedly mentioned by Vygotsky—*outside* of the educational context—and to use them without critical reflection as if they represent a theory of learning or pedagogy [91].

Something similar is happening with regard to findings from the cognitive sciences concerning human development. These findings with seemingly valid cognitive-scientific origins are argued to have been taken out of their context and to have influenced researchers and instructors in educational circles, resulting in a large-scale acceptance of misconceptions—*neuromyths*—about the brain with regard to learning [20]. For instance, the previously described VAK model was introduced in the context of remedial reading in the early twentieth century, then shown to be gradually taken out of this context until its present inclusion in teaching and learning style theories in an unwarranted and unscientific manner [32, 48]. Yet, despite the lack of evidence that the model can be applied by instructors to accommodate to individual students' preferences so they can learn better [32, 40, 48], the perception of the efficacy of doing so by scholars, educators, and students appears to remain overwhelmingly high (e.g., [20]). Voices of advocates of the model's application are similarly persistently incorporated in educational textbooks [40, 48]. This extends into recent works that forgo critical examination of the model yet vigorously advocate for their integration in instructional models (see e.g. [39, 80]).

This brings us to two reasons why learning and education should be distinguished from one another. The first, straightforwardly touched upon by Krashen in 1989, is "that the teaching profession is controlled by outsiders, by amateurs […] as if hospital administrators dictated to surgeons how to operate" [66, p. 455]. This points at the earlier consideration that educational systems are importantly shaped by politics and industries. If this is the case, then it is perhaps not strange that misconceptions end up in pedagogical models. The unfortunate consequence would be that students, who are promised that they will learn something—be it a language or the knowledge to be ready for democratic society—are subjected to models that do *not* (effectively) foster learning, and that instructors lose valuable time trying to enhance instruction based on ineffective models [40]. The second reason follows from the particular educational objective to deliver graduates to the workforce. This, namely, involves standardized certification (see e.g. [130]) which is in turn enabled by assessments. About this, Krashen again makes a remark: "we do need to get rid of the exams" because "[t]eachers will teach to the exam, and students will study for the exam, and no force is likely to change this" [66, p. 455]. There are many types of assessment and they

have their various purposes [15]. Krashen refers particularly to summative, "discrete-point exams" [66, p. 455]. For language learning, Krashen argues, it makes little sense to subject students to such exams; it only drives them to engage in vocabulary drills while that time could well have been spent reading [66].

Krashen's points were made some four decades ago, but almost ten years later, Hannafin et al. observed that a large number of learning system designs lacked proper grounding [43], and many educational institutions today are still argued to employ out-of-date assessment and instruction techniques [92]. In the meantime, critics have pointed out a perceived disconnect between classroom education, situational practice [16], and the aim to foster flexible competencies in students [47]. Perhaps it does not help that educational discussions on effective learning appear to have a tendency to forgo the practice of learning and skip directly to *instruction* [68]. These issues are so grave precisely because educational systems are set up to instruct many students at once. If a system's mode of instruction is ineffective or counterproductive with respect to learning, then all students are affected. Or, if instruction is only effective to *some* students, the others do not benefit as much. Combined with the earlier observation that education exists principally to connect students to the workforce and protect corporate interests, and not to encourage learning, it seems warranted to see learning separately from education and study.

This is not to say that education serves learning in no way whatsoever, nor that there is no instructional model that facilitates effective learning (see e.g. [57, 68]). As remarked in Section 2.4, learning and study can, and at times likely do, coincide. Whether or not the educational systems ought to be significantly changed is also not the point here. The key point with regard to learning in the educational context is that it is necessarily organized en masse, that concessions must thus be made—to a small or large degree—to sacrifice students' pursuit of personal interests in favor of standardized systems, and that learning must inevitably make way for *study*, because students must strive to obtain their credentials within limited time. Education is, as such, fundamentally not conducive to highly personalized learning at present. The suggestion that education constitutes a discourse that students, instructors, and researchers make part of complicates drawing a clear picture of learning in education because it obscures this distinction between learning and study. And this demonstrates why viewing education critically is essential [71] if one were to build software to personalize study.

It is in this light that studies of education—including the studies discussed in this thesis, and this thesis as a whole—must be viewed in order to understand what parts of them pertain to this educational discourse and what parts to learning. After all, studies that had not taken this into account may have been inadvertently influenced by the discourses of education, yielding outcomes that indeed do not primarily serve learning, but rather the educational (or political, or industrial) systems. It also implies that software that is made for education is not necessarily made for learning. In fact, educational software products, like education itself, are highly financially incentivized, raising the question to what extent a best effort is exerted to have them support learning as opposed to revenue or the educational systems [17, 33, 75, 95].[30] These considerations are crucial, since it is this thesis' aim to design software that personalizes study to facilitate language *learning*, not for the sake of study itself.

3.2.3. Education and the Instructor-Student Relation

Concerning the relation between instructors and students in education, two instructional categories are regularly referred to: *teacher-centered learning* (TCL) and *student-centered learning* (SCL) [7, 118].[31] TCL is seen as an instruction style where instructors act at the center, as in traditional lectures, and SCL as a style where students are more engaged in activities like group assignments or discussions [7, 118]. TCL and especially SCL are 'hot topics' judging from the studies that will be discussed in Section 3.3 and large-scale efforts such as the Bologna Process that have set SCL-based instruction as an important goal [30, 31]. However, while the term SCL suggests that it emphasizes students' personal

[30] This goes both ways: it is argued that "[d]ue to economic and practical pressures in higher education, hybrid language courses […] have been launched as a means to save money" because in this way, "[f]ewer people can be responsible for more students" [95, p. 20]. This demonstrates the necessary, unavoidable trade-off in massified education between deployment of financial resources and attention to an individual's learning. The point to be made here is not that it is 'good' or 'bad'; it is that it should be acknowledged that the trade-off exists so that we can understand (technology-assisted) learning in the context of education better.

[31] In terms of Section 2.4, these concepts would literally translate to 'instructor-centered studying' and 'student-centered studying', but since the terms TCL and SCL are commonly used in literature, they are not altered here. It will moreover follow later that even these translations do not encapsulate the concepts' implied meanings.

learning rather than massified study, it can be argued that TCL and SCL are, in essence, not all that different. To elaborate on this argument, this section briefly explains how TCL and SCL are understood in literature and how this translates to practice.

TCL-based instruction is understood to be characterized by the role of instructors as the highest authority and the principal source and disseminator of information: instructors speak, students absorb, as is traditionally the case in lectures [7]. Some advantages attributed to this instruction style are its structured nature—helpful for students who experience difficulty in guiding their own study—and the efficiency of information dissemination [7]. Disadvantages of TCL are, however, stressed by critics: it is often not tailored to individual students, potentially raising issues for those who require more personal attention, and it does not naturally entice students to think critically [7, 28].

SCL appears to be variedly understood by instructors, students, and scholars, causing it to be interchangeably used with terms like *active* and *collaborative learning* [118]. Nonetheless, it is generally argued that it more broadly represents an instructional style that relies on students as active participants whose individual exploration and autonomy are emphasized, with the instructor fulfilling the role of facilitator [7, 118]. This should encourage students to think critically and supposedly provides students with more opportunities to match their study style with personal needs [7, 28, 96]. At the same time, granting students 'too much freedom' is suggested to bring negative effects for study: for instance, some students find it difficult to autonomously commit to regular class material reviews, causing them to perform relatively worse than others in terms of grades, or to experience cognitive overload in their attempts to organize knowledge by themselves [59].

While the defining criteria for TCL and SCL remain a point of scholarly discussion [118], the idea that the instructor plays a crucial role in the manifestation of both TCL *and* SCL appears to enjoy consensus (e.g., [7, 8, 118, 131]). Consequently and paradoxically, SCL, understood to grant students autonomy and flexibility, appears to be attainable only by grace of some instructor who facilitates it. The instructor, in turn, must rely on the curriculum and institution. In practice, this amounts to situations such as in [4] where instructors struggle to implement somewhat vague, imposed visions of SCL. Or, cases like [78, p. 7] where language instructors have supposedly embraced SCL, yet prohibit students from fast-forwarding instruction videos "to prevent students loafing while they were supposed to be studying." In [8, pp. 17-18], it followed from input from over a hundred language instructors that even though they feel SCL is effective, "they still assign a remarkable amount of importance to the transmission of knowledge from teachers to students," i.e., "spoon-feeding." Perhaps most importantly though, supporters of SCL-like instruction modes do not necessarily take students' preferences and goals into account, as is further substantiated in Section 3.3.2.

These observations raise the question whether the definitions of and distinction between TCL and SCL are as meaningful as they are purported to be and whether SCL, as its name suggests, really involves a centering on students' learning. The terms 'teacher-centered learning' and 'student-centered-learning' appear illusory in the first place, especially with Section 3.2.1 in mind: if anything, the free-energy principle suggests that learning, although informed by perceptions of things outside one's body, happens essentially *within* a person. What does it then mean for learning to be 'centered around' a teacher or student? Instead, TCL and SCL as they are described above pertain more to *instruction* styles; as such, something along the lines of 'instructor-driven instruction' and 'student-driven instruction' may more accurately represent how the terms are practically understood and applied.

In conclusion, this section brings two important implications. The first is that even if some educational instruction mode claims to adhere to SCL, "wherein learners actively participate in their educational journeys" [77, p. 207], it does not automatically entail that the instruction facilitates students' personal learning goals, interests, and preferences with respect to tools and materials. As long as students have no significant say in how they approach their study, it does not seem to be too relevant to personalized learning whether TCL or SCL is employed. This is not to say that in neither TCL nor SCL students have any say in this matter, but that the terms TCL and SCL do not say much about it. Second, if TCL and SCL are seen as instruction styles, there seems to be no reason to limit personalization efforts to either one of them. If TCL offers systematic instruction and SCL offers dynamic activities and both can somehow be effectively used, it makes sense to use both to a degree depending on the instruction goal [8]. These observations are particularly relevant to this thesis because they imply that software that supposedly promotes SCL by no means naturally contributes more to personalized study than

TCL-based instruction, nor that it is naturally 'better' for software to embrace SCL rather than TCL.

3.2.4. Conclusion

This section viewed learning from psychological, cognitive, and neurobiological perspectives. Neurobiologically, learning relates to changes in the brain's belief state; cognitively, to informational changes in several hierarchical cognitive layers; psychologically, to behavioral changes. While the latter are observable from the outside, the former two require deeper insights in the brain, here informed by the free-energy principle and cognitive layer cake. Education helps deliver instruction en masse, arguably focused on 'psychological learning'. Whereas education is said to be in transition from teacher-centered to student-centered learning methods, this section has argued that this does not naturally imply that instruction is becoming more personalized, nor that it shifts focus to 'cognitive learning'. These findings importantly affect interpretation of observations in what follows in this thesis.

3.3. Application of Technology in Language Education

To answer **RQ2** and investigate what kinds of technologies have been implemented in language education, for what purposes, and with what consequences, a systematic literature review was conducted according to the methodology described in Section 3.1. n=56 studies were selected for analysis (see Table 3.1) according to the inclusion and exclusion criteria in Table A.1. The studies were categorized by the author as in Table 3.2. The literature is discussed below per category, followed by a general discussion.

Table 3.1: Number of found publications after filtering for **RQ2**. The number of results after search with the filter queries excludes duplicates. See Chapter A for search queries and inclusion and exclusion criteria.

n **after search with filter query**	n **after manual filter (title and abstract)**	n **after manual filter (full-text)**
146	98	56

Table 3.2: Overview of categorized **RQ2** studies.

(Code) Category	**Category description**	***n***	**Studies**
(T) Technology experiments	Experiments wherein new or existing technologies related to language education are used by participants and evaluated	27	[1, 9, 10, 13, 18, 21, 24, 34, 41, 46, 54, 60, 61, 63, 65, 72, 76, 79, 84, 86, 99, 106, 108, 126, 132, 133, 135]
(P) Studies on perceptions of technology	Studies that evaluate instructors', students', and other relevant actors' perceptions toward the use of technology in education (e.g. through interviews or surveys)	17	[4, 5, 6, 36, 74, 77, 78, 81, 85, 97, 104, 119, 127, 128, 134, 137, 138]
(C) Computer-technical system designs	Studies that design some digital system for educational purposes (e.g. material recommendation systems or learning apps)	8	[25, 37, 38, 52, 82, 115, 139, 140]
(O) Observational studies	Studies that observe participants who use technology for some language-related purpose	2	[42, 55]
(E) Education-technological framework designs	Studies that propose a design of some education-technological framework	2	[87, 121]

3.3.1. (T) Technology Experiments

Of the 56 studies in this review, 27 are studies wherein experiments are conducted with new or existing technologies related to language education. The studies are categorized in Table 3.3 according to the principal aims of each study and the technologies they employ. Since the studies in T largely involve distinct empirical setups, Table A.2 summarizes each study's setup and key results for reference.

The technologies experimented with can be divided in three overarching categories: technologies that facilitate blended study in face-to-face scenarios (T1), those that attempt to personalize language ed-

Table 3.3: Overview of technologies experimented with in T and their purposes.

(Code) Technology purpose	*n*	Technology	Studies
(T1) Facilitating blended (offline and online) study with in-class or otherwise face-to-face interaction	13	Augmented reality (AR) or virtual reality (VR)	[9, 10, 46, 61, 65, 72]
		Blended learning and shared workspaces	[18, 34, 54]
		Gamification (non-AR/VR)	[99, 126, 133]
		Social media	[86]
(T2) Personalization of language education	8	Generative AI	[21, 60, 135]
		Deep-learning models	[76, 132]
		Pedagogical conversational agents	[63]
		Other	[106, 108]
(T3) Facilitating online study or online peer collaboration	6	Online learning management systems or portfolios	[1, 41, 84]
		Language learning platforms	[13, 79]
		E-modules	[24]

ucation (T2), and those that facilitate online study or online peer collaboration (T3). Each group is discussed separately below, followed by general observations.

3.3.1.1. (T1) In-Class Blended Study

A recurring research question is how language education can be made more enticing for students. In T1, the answer is sought in the enhancement of the in-class experience. Technology is presumed effective for this purpose because new technologies are considered innovative—in contrast with traditional educational methods, implied to be boring and inflexible—prompting authors to integrate some piece of technology in a course. [10, 46, 61, 126], for instance, entertain this narrative as a primary motivator to conduct technology experiments in language education. Others elaborate on their choice further by suggesting that specific affordances are only attainable through technology, such as the capability in AR applications to virtually add to physical reality [72] or software's distinct capability to cater "to the unique needs and preferences of students with learning difficulties" [9, p. 2]. Regardless of the motivation, it seems to be generally hypothesized that software will positively innovate language education.

The T1 studies primarily used AR, VR, and blended-learning systems to make the in-class experience more engaging, with gamification regularly employed as a technique regardless of the platform. Of the AR and VR studies, [46] yielded positive results by having children aged five to seven use an AR app to display virtual objects and animations and play sounds on top of a language textbook, increasing their "curiosity and their focus" [46, p. 766]. The study argues that AR can easily enhance interactivity and accessibility in existing materials like textbooks [46]. The remaining AR- and VR-related studies presented more mixed results and cautions with regard to their usage. Studies report that participants experience difficulties and discomfort when interacting with VR environments, partly due to technical issues like lag or inadequate screen resolutions, inhibiting the immersive quality of VR which is precisely seen as its selling point [10, 61, 65]. Furthermore, both VR and AR applications are observed to cause sensory overload or distraction from the learning objective [9, 61, 65], despite the belief that especially VR, through head-mounted displays, should theoretically block distractions that would surface in the physical world [10, 65].

Of the six studies experimenting with AR or VR, [65, 72] explicitly incorporated teamwork, recognizing the idea that language learning importantly involves communication. In [65], leveraging VR avatars to have students work together yielded that the avatars "[boosted] confidence, particularly among those fearful of speaking in foreign languages," because they served as a "mask of anonymity" and could hide physical reactions like blushing when making language mistakes [65, p. 76]. In [72], where teamwork was facilitated by means of an AR game in one participant group and its physical, printed counterpart in another group, no significant difference was found between students' performance when playing the AR game compared to the printed version. In fact, students regarded the physical game as more useful for language learning: they put more energy in understanding English sentences in the printed version than in the AR game since the latter was too immersive [72]. This supports the abovementioned

findings that immersive environments may distract from study objectives. Finally, in a study involving individual consumption of VR video lessons, it was remarked that "many participants [wished] for more opportunities to communicate with peers or native speakers in the virtual environment" [61, p. 532]. These outcomes demonstrate the complexity of building effective, easy-to-use VR or AR applications that truly encourage teamwork and communication in language education scenarios.

The concept of gamification was additionally leveraged through other technologies. [34] yielded positive results by interleaving traditional classical-language instruction with gamified intermezzos that make use of digital collaboration and quizzing tools, among others, resulting in students feeling "reduced academic workload" and reaching higher exam scores [34, p. 954]. The authors emphasize the need for flexible curricula that enable usage of such intermezzos and tools [34]. Two studies, one with kindergarten children [126] and one with university students [133], used a large screen in class on which language-educational games were played in substitution of textbook-based instruction. In both studies, the games were reported to have had positive effects on participants' attitudes. The kindergarten group playing games in [126] demonstrated improvement in areas like initiative-taking and goal awareness relative to the control group studying with a textbook, but not in other areas such as persistence and frustration resistance. The authors moreover warn that animations in games can distract children from the study task at hand [126]. In [133], where university students played motion-based games in groups to study L2 English instead of using a textbook, an improvement in students' self-efficacy beliefs was reported especially regarding speaking and listening skills, thought to originate from the cooperative facets of games. To incorporate game-based learning in class, the authors stress that "[c]areful pedagogic planning [...] is essential" [133, p. 170]. Finally, [99] organized a game involving students walking through a city and interacting with locals in English to complete tasks aided by digital tools, followed by post-game tasks that prompted students to self-learn digital tools necessary to complete them. The game was perceived as enjoyable and was felt to reduce anxiety when communicating in English [99]. Based on these studies employing gamification, it appears that students generally developed positive attitudes toward language study and were more enthusiastic compared to those studying through textbooks.

Other technologies than game-based ones were moreover used to foster in-class peer collaboration or positive group dynamics. [18] had students perform writing tasks in groups, where each student sat together but worked on their own computer in a shared *Google Docs* document. An adjacent television screen displayed each student's screen in real time, such that students could see what their team members were doing at any time [18]. Students did not perceive added value of the television screen since Google Docs already supports live collaboration, but they did value the concept of collaboration in a shared workspace to work on foreign-language writing tasks [18]. [54] attempted to steer the flow of language classes based on students' real-time reports of their feelings submitted through a software tool. This steering took the form of "[implementing] interactive group activities, guided breathing exercises, and reflective journaling strategies" when the tool indicated that students felt tired or stressed [54, p. 265]. While the researchers were able to improve class dynamics to some extent, they observed that some students used the tool superficially and that the curriculum complicated making good use of dynamic class flows [54]. These and aforementioned gamification-related cases demonstrate a potential for complementary cooperative activities to enhance group dynamics and student attitudes, but also that not all cooperative tools yield strictly positive outcomes for all students.

[86] attempted to boost motivation by having students watch educational content on *Instagram* and *TikTok* during the first part of each of their English classes. The authors argue that studying by means of social media content increased students' interest and motivation since it is more dynamic than the "rigid learning content" in textbooks [86, p. 13657], and that students would be interested in watching educational content on social media in their free time on the condition that they would receive instructor guidance on how social media can be used to study English effectively. A large drawback that was mentioned is that social media does not present content as systematically as textbooks and that content quality cannot be controlled [86]. The researchers' motivation for employing social media was that in TikTok, "each video is only fifteen seconds long, which is excellent for young EFL [English as a foreign language] learners with short attention spans" [86, p. 13641]; the possibility that it is apps like TikTok that contribute to people's short attention spans (e.g., [105]), however, is left undiscussed.

In summary, the studies used various means like AR, VR, gamification, and in-class hybrid instruction to improve in-class dynamics. The studies' samples were largely homogeneous and limited in size, limiting

their generalizability. Still, a common denominator across the studies is arguably that participating students valued studying on the basis of something else than a textbook. The studies are, however, inconclusive as to whether there are special benefits in using digital tools as such a basis or not. For instance, [126, 133] do not reflect on the alternative of using non-digital tools to gamify their classes, and [34] mentions it only in passing. This complicates assessment of what makes digital tools especially beneficial to boost motivation and class dynamics. Regardless of the tool employed, multiple studies stress the importance of flexible curricula in order to facilitate integration of tools or concepts [34, 54, 133].

3.3.1.2. (T2) Personalization of Language Education

The T2 studies experimented with technologies that in some way facilitate personalization of language education. [76, 132] set out to personalize curricular study paths by adapting them to individual students by means of deep-learning algorithms. The studies claim to have used recommender systems (although it is left unmentioned *which* recommender systems) to generate personalized study paths based on student behavior measures, leading to higher exam scores in both cases compared to students using traditional study paths [76, 132]. Studying along the recommended paths was furthermore seen to improve task completion rates [76] and reduce cognitive load [132]. These results are highly promising—unfortunately, the studies are unclear concerning the specifics of the systems used and the experiments conducted, making it impossible to replicate or apply them or to assess their reliability.

Another study set up and evaluated a design framework for gamified pedagogical conversational agents (PCA) to boost language students' motivation and provide personal assistance at times that instructors are unavailable [63]. The proposed framework guides PCA designers to act on the aspects of "goal-setting and reflection, novice-expert relationship, performance-related motivation, and learning story narration" [63, p. 1]. The authors implemented a PCA according to their framework, resulting in a chat agent capable of generating short language-related games in accordance with user's self-indicated goals and preferred learning pace [63]. Evaluation of the implemented PCA revealed that it improved students' motivation and well-being, with participants highly valuing the interactivity the PCA provided through chats and games [63].

Three studies investigated whether generative AI can help students study autonomously and in a more personalized way [21, 60, 135]. Each of the studies found that generative AI chatbots made this possible to some extent, attributed to advantages like bots' constant availability [21, 60], their capability to help students autonomously find study materials [21], and their capability to "reduce communication anxiety and [encourage] help-seeking behavior" when facilitating communication practice [135, p. 5]. The studies also point out major concerns raised by instructors and students, such as diminution of students' critical thinking and creativity [21, 135], increased risk of fraud in research and assessment [21, 135], loss of human interaction [21, 60], and the risk of overreliance [21, 60, 135]. Two studies had conflicting findings with regard to speaking practice, where [60] concluded that generative AI is incapable of fostering speaking skills, while [135] observed that generative AI helped "initiating conversation, seeking clarification, and interacting with peers and native speakers" [135, p. 7]. To combat some of the abovementioned issues, the argument is raised that the integration of generative AI in classes needs to be coordinated with due reflection on the curriculum [60, 135]. In summary, the studies suggest that generative AI may have its place in language education when it comes to finding materials that match students' interests or conducting private practice, but that there are significant concerns related to critical thinking, interaction, and fraud, among others, that should be overcome as well.

In [106], researchers experimented with having pre-service language instructors create "digital stories" and observing whether it had an effect on participants' *TPACK* scores [106, p. 444].[32] While it is not clearly explained what digital stories look like in the study's context, it appears that the story-making steps that participants went through facilitated targeted selection of materials and technologies that their future language students would use to study [106]. Adjusting the selection of materials and technologies seems to have been done on the basis of TPACK relations to ensure systematic targeting

[32]TPACK (originally TPCK for Technological Pedagogical Content Knowledge; see [88]) is a framework that is used to assess instructors' (self-belief of their) content knowledge, pedagogical knowledge, technological knowledge, and their intersections in order to gain insights into how instructors' knowledge situates within the context of applying technology in some pedagogical system.

toward the assumed characteristics of future students [106]. The results indicated that creating digital stories in this way significantly improved participants' TPACK self-confidence scores [106], suggesting they felt more confident integrating chosen technologies in a curriculum while taking the students into account.

Finally, [108] investigated how the use of a language corpus to find word contexts and to complete corpus-related tasks influenced language students' "linguistic, subject, communicative, cognitive and intercultural" competences of their language class compared to students who used a traditional textbook [108, p. 75]. The results suggested that students who used the corpus scored significantly higher in all aspects than those who did not [108]. The authors do remark that students had trouble with grammar, differentiating between oral and written styles, and using a wide vocabulary [108], but do not indicate whether this holds for the experiment group, the control group, or both. Despite this and other ambiguities in the study, if it is assumed that the reported results are accurate, then allowing students to, for instance, compose their own vocabulary lists or find example sentences through a corpus could be beneficial for the purpose of language-educational personalization.

Summarizing, technologies like automatically-adjusting study paths, chatbots, and digital corpora potentially contribute to more personalized forms of language education. Particular benefits of the application of such technologies lie in the autonomy they grant students to study at a time convenient to them and to find materials that match their level and interests. However, the studies discussed here all involve interventions *within* existing curricula. Any gains in personalization hence remain constrained within the standardizations inherent to the curricula. This factor remains even if special care is taken to adjust a curriculum to accommodate the introduction of some technology, as prompted by [60, 135], among others. Additionally, like in T1, the generalizability of the T2 studies is limited due to mostly homogeneous samples in each study.

3.3.1.3. (T3) Facilitation of Online Study or Peer Collaboration

The T3 studies attempted to improve or assess the efficacy of language education with complementary use of online learning management systems (LMS) and language learning apps. The common idea behind the choice for such technologies is that they allow students to study outside of class while having full access to instructor-curated materials, or during class as a form of blended learning.

[1, 41, 84] used *Moodle*, *LITE*, and *WebQuest* as LMSs, respectively, each reported to have facilitated online collaboration [41, 84] or foreign-language communication [1, 84] through their collaborative features. In [84], it is suggested that WebQuest is additionally able to foster soft-skill competencies (e.g., presenting and planning) simultaneously with foreign-language skills through its task-based collaboration functions. LITE is observed to be capable of stimulating critical reflection by providing students with means to reach back to their "linguistic histories and trajectories" and offering an explicit reflection module [41, p. 111]. Integration of an LMS in a curriculum was perceived to be particularly challenging in [41] because instructors and students were given little time to interact with or learn how to effectively use the platform. In [1], it is argued that students were able to "progress through the content at different speed and time in accordance with their own needs, pace and abilities" with the help of Moodle, "despite the same curriculum and syllabus for all students" [1, p. 75]. In terms of personalization, it appears that each employed LMS operates largely on instructor-provided materials—notwithstanding LITE's option for students to record their study history—but that they do have the potential to tune courses somewhat to students' needs and to positively impact collaboration.

[13, 79] investigated the use of language learning platforms. [13] used *DynEd*, an English-language learning platform, but found that curricular time constraints and lacking ICT infrastructure heavily inhibited effective use of the platform, thus yielding no positive results. [79] examined the effect of using language learning apps on neural activity and found that, compared to studying with a textbook, studying using such apps "contributed to a more integrated and efficient speech-processing network" in the brain [79, p. 18317]. This study is, however, ambiguous in some regards; for instance, it repeatedly mentions "higher performance" in bilingual students [79, p. 18313] but does not elucidate what this is compared with, making it difficult to interpret the findings.

Lastly, in [24], the authors developed e-modules with the aim to combine foreign-language education with preventive measures "designed to counter anti-immigrant sentiment, xenophobia, and racism" [24, p. 1]. The modules were construed as text documents with group assignments that were executed during three

subsequent online sessions with students from various countries in Europe, hand-picked by the authors [24]. No direct influences of the modules on language learning capabilities are reported, but it was found that the modules prompted reflection, nuanced discussion, and "heightened intercultural critical consciousness" [24, p. 9], which may itself be conducive to appropriate foreign-language communication.

In summary, it appears that online platforms and modules that foster peer collaboration and communication are well-received within the populations that participated in the above studies, but only under the condition that the curriculum provides enough space to use these technologies. Provisioning of class materials outside of class—apart from physical textbooks—furthermore seems to be perceived as a valuable affordance in LMSs. Regarding language learning apps, the studies in T3 were inconclusive.

3.3.1.4. Concluding Observations

Several themes recurred in the discussions in T. First, many handled studies have limited generalizability due to small sample sizes or homogeneous samples, which should be taken into account when interpreting their results and the observations presented here. Second, in each of the categories, multiple studies emphasize the need to re-examine existing curricula: not adjusting the curriculum to a technological intervention is repeatedly seen to inhibit the intervention's efficacy, where especially time constraints are a limiting factor. Third, while some studies report overwhelmingly positive results (e.g., [1, 86]), the majority report more mixed or negative results concerning successful integration of some technology in language education. Although each studied technology was seemingly intended to bring benefit to as broad a student population as possible, their effectivity did not reach all students due to infrastructural or curricular limitations (e.g., [13, 41]) or mixed reception among students (e.g., [61] and expressed discomfort in VR studies in T1). This implies that, among the technologies discussed here, no one-size-fits-all solution has been found yet. Finally, if anything is to be taken from the discussed experiments, it is that regardless of the form of intervention, capabilities that seem to be especially valued by students and instructors are the facilitation of group communication or collaboration and the provisioning of broader ranges of materials and activities other than those offered by traditional textbooks.

3.3.2. (P) Studies on Perceptions of Technology

Seventeen of the filtered studies investigated instructors' and students' perceptions with regard to the use of technology in language education, primarily by means of questionnaires or interviews. It appears that studies on the integration of generative AI-related technologies prompt different considerations in instructors and students than other technologies.[33] The studies are grouped accordingly as in Table 3.4 and handled separately below, followed by a general discussion.

Table 3.4: Overview of technologies handled in P.

(Code) Technology group	*n*	Studies
(P1) Related to AI	6	[6, 81, 104, 128, 137, 138]
(P2) Other	11	[4, 5, 36, 74, 77, 78, 85, 97, 119, 127, 134]

3.3.2.1. (P1) Views on AI in Language Education

In general, the P1 studies reveal that there is cautious but broad optimism with regard to the application of AI, pointing out several advantages for students and instructors. A repeatedly mentioned advantage of AI is that it is capable of providing students with personal support and level-adjusted exercises or goals [6, 128, 137]. In language study contexts, [138] found that students popularly use AI for vocabulary training. [128] found that implementing AI-related features in a language course is "positively associated with learners' emotional health and negatively associated with depressive symptoms" [128, p. 8], a result the researcher attributes to the benefits of personalized attention for students. For instructors, a potential is recognized in using AI as an aid to create lesson plans and complete administrative tasks, saving valuable time [81], and in using it to monitor students for evaluation or for prevention of health issues [6, 128].

[33] For brevity, 'generative AI' is hereafter referred to as 'AI' in this section, since all selected studies that discuss AI refer to generative AI and not to other forms of AI.

The studies also highlight instructors' concerns. One important concern is students' "misuse" of AI [81, p. 21], particularly in the area of plagiarism and cheating [6, 81, 137]. Some also observed an overreliance on AI by students [81, 137]. Consequently, it is emphasized that AI cannot be integrated in language classes without strategic consideration of the (un)ethical sides of AI with respect to the curriculum [6, 81, 128] and careful planning with instructors, students, policymakers, the industry, and educational institutions [128, 137]. Interestingly, it is argued that instructors tend to focus on students' potential misuse of AI rather than misuse among themselves [81]. This seems to contrast with a wide recognition in literature that AI literacy among instructors varies [81, 104, 137] and that instructors should be guided or trained to better understand AI and the implications of integrating it in educational contexts [6, 81, 104, 128, 137]. And if ignorance with respect to the attributes of AI is assumed to cause or otherwise be related to AI-related technophobia, then, according to [104], this significantly hampers AI integration in class. To combat this, [81] demonstrates that (structured) usage of AI, even if by way of experiment, can lead to improved AI literacy.

In conjunction with the above calls for caution, and regardless of whether AI is implemented in some language class, there seems to be a recognition that education is changing or is required to change as a result of students' use of AI. For instance, while [6] argues that students should be prevented from committing AI-facilitated plagiarism in assessment, it also suggests that present forms of written assessment cannot be sustained and should be changed into a system that emphasizes verbal discussion. Similarly, the role of instructors with respect to students is envisioned to change as a result of AI usage and development [81]. In any case, whether or not AI is to be used by students in their language education seems to be regarded in the handled studies as a choice to be made primarily by educators and instructors, and not or to a lesser extent by students.

3.3.2.2. (P2) Views on Other Technologies in Language Education

The remaining studies in P relate to online or blended language education and to general perceptions concerning the implementation of various technologies therein. The sudden COVID-related closure of schools in 2020 is recurringly mentioned as one of the catalysts that brought blended-learning facilities to the foreground [5, 36, 74, 127]. The capacity of blended learning to offer both structured in-class and flexible online instruction [5, 97] or to instigate a more rigorous shift away from rigid traditional education [77] constitutes a second important reason.

Among the affordances of technology in language education, quick access to information and various materials appears highly valued by students [97, 134]. Technology is also observed to help students communicate with instructors without having social hierarchies hold them back [4], which is supported by the finding in [77] that the implementation of technology correlates positively with instructor-student dynamics. Instructors' willingness to incorporate technology in language class moreover seems to positively affect students' performance, while unwillingness appears to impact students' performance negatively [134].

Despite perceived benefits of software technologies in language education, there appear to be high hurdles preventing straightforward implementation. Instructors feel that they require more technical training and support to facilitate technology integration in class [4, 5, 36, 78, 119]. The support would, for instance, pertain to increasing instructors' understanding of how technologies work and how they meet students' and instructors' needs. In addition, instructors are reported to face struggles during the design, implementation, and usage of educational software due to limited available time in curricula [4, 78]. In [4], for instance, instructors lament that they are required to finish prescribed sets of subjects within a semester, causing stress, loss of instructional quality, and a tendency to hold fast to traditional methods. Finally, inadequate ICT infrastructure [4, 78, 119] is reported to hinder effective integration of software technologies. Importantly, in each of these studies, these hurdles are seen to significantly impact the efficacy of educational software in practice.

Multiple studies in P2 allude to student-centered learning (SCL) as an objective to strive for, with software technologies playing an important role in realizing SCL-based instruction [4, 5, 77, 78, 127]. Yet, in the context of perceived effectivity of software for language study, the viewpoint of students is marginally discussed, or predominantly from the perspective of instructors. One study that does consider students' perceptions is [97], which argues that Moodle's user interface and connection stability should be improved to address students' frustrations with the platform. On the other hand, [85] remarks

that student participants reported skepticism toward the effectivity of "computer tests and online tasks" to develop soft skills, but proceeds to conclude that this is merely so because "the students are not acquainted with these types of educational technologies and cannot express their opinion on this issue" [85, p. 11]. Another example is presented in [78, p. 7] where instructors are reported to disable the fast-forward function in instructional videos on an online study platform so as "to prevent students loafing while they were supposed to be studying." The students' motivations for fast-forwarding the videos are not considered. All in all, when it comes to the implementation of technologies within curricular constraints or the expected efficacy of software in language study, various studies in P2 reflect on instructors' experiences and opinions, but none appear to elaborately consider students' side.

3.3.2.3. Concluding Observations

There is a broadly shared vision among instructors that software will take language education to new heights, importantly owing to its potential to provide students with constant access to various materials and exercises. Blending traditional instruction with software-based instructional methods is regarded as a step in the 'right direction', for it facilitates reaching the objective of SCL-based instruction. However, the studies in P reflect a dominant preference for *instructors* to determine whether and how they and students use software; an observation that rhymes with the argument in Section 3.2.3. In the context of AI, students are feared to misuse technology, leading instructors to restrict students' use of AI or to argue for revisions of assessment methods so that misuse can be prevented. Other technologies like LMSs are expressly incorporated in blended-learning classes to reap the 'SCL benefits', despite occurrences of lacking ICT infrastructures or students' reported discomforts operating the technologies. Still, more than half of the studies in P argue that instructors require more training in educational technologies to better understand their uses and students' needs. Instructors' educational institutions are argued to hold a key role to facilitate such training. Whether or not instructors will be given sufficient technical training and support in the future, and time to make use thereof, then, seems to be a problem of institutional or political policy.

3.3.3. (C) Computer-Technical System Designs

Eight studies propose computer-technical designs or implementations to supplement language education in some way, categorized as in Table 3.5 and discussed per category below.

Table 3.5: Overview of computer-scientific design-related studies handled in C.

(Code) Technology purpose	*n*	Studies
(C1) Learning platforms / automated material recommendations based on platform interactions	3	[52, 82, 139]
(C2) Automated assessment	3	[25, 37, 38]
(C3) Automated dialogue generation	2	[115, 140]

3.3.3.1. (C1) Learning Platforms and Automated Material Recommendations

The C1 studies are not completely clear on what data was used and how results were evaluated, making it difficult to assess their reliability. Nonetheless, they each advocate for personalized study paths and propose ideas to implement automation strategies to integrate them in online study platforms. In [52, 139], this is materialized by proposing student-personal course recommendations based on their historically followed courses or their language skills, respectively. [82] proposes to combine the two by recommending study resources based on students' study histories and language proficiencies. It is thus suggested that students' continued use of a learning platform improves personalization options by providing more relevant, activity-based suggestions.

3.3.3.2. (C2) Automated Assessment

[25, 37, 38] propose methods to automate assessment of student inputs. [25] presents a design of a transformer model capable of taking a student's L2 English speech as input and providing corrected speech as output. The model is made to produce outputs that are closer to the speaker's speech characteristics than existing models with the aim to facilitate natural, automated speech assessment and correction that students could use in class or by themselves. For automated guidance in writing practice, [37] proposes a deep-learning model that can be used to detect and correct written grammatical

mistakes and assess essay quality in English. It is suggested that the model could be integrated in online study platforms to provide students with instant and personal feedback on their writing. Finally, to facilitate insight into students' reading skills, [38] proposes a deep-learning model capable of classifying a real-time camera feed of a human face in seven types of emotions, which the authors argue can be used while students are reading some L2 text to assess to what extent they understand it. The studies do not extensively handle ethical considerations, but one can imagine that there are privacy and security implications if students would be placed under continuous supervision of microphones or cameras.

3.3.3.3. (C3) Automated Dialogue Generation

Lastly, two studies aim to automatically provide students with context-relevant practice dialogues in an L2 by means of deep-learning models. The model proposed in [140] is not clearly elucidated, but appears to attempt to automatically respond to a speaker's spoken input with a context-relevant spoken question to facilitate conversation practice. [115] does this in text-based form based on a self-constructed dataset of Chinese-language dialogues in 219 language textbooks. The model specifically generates responses that contain some selected vocabulary word. These models could thus help students gain access to more practice materials.

3.3.3.4. Concluding Observations

The studies in C reflect efforts to personalize language study through personal course recommendations, material suggestions, and automated assessment. The relevance of such efforts is made clear by the suggestions that surfaced in T and P, which indicate that students appreciate flexible use of and access to various study materials. As suggested in C2 and C3, it seems possible to complement existing study platforms with more continuous and personal assessments and context-relevant study materials in conjunction with existing study methods and materials. To ensure context-relevancy, it is suggested that students' historical use of study platforms and developments in language competency could be referenced—ongoing use of study platforms, then, would make the systems more effective. The proposals in C1, however, arguably require more thorough curricular revisions to facilitate more personal study paths. If this is limited to students' more flexible choice of courses, then this seems to be within reach of ongoing higher-educational reforms such as those presented in Section 2.2.4.

3.3.4. (O) Observational Studies

In [42, 55], researchers conducted observations of and interviews with participants as they interacted with language-related technologies in their daily lives [42] or as an instructor at school [55]. The difference with the studies in P is that the participants were not instructed to use specific technologies or complete prescribed tasks.

In [42], three L1 Spanish users who migrated to Canada were observed at their homes as they used technologies to navigate language problems, typically related to job hunting or understanding French-language messages, letters, and websites. The researchers found that participants assertively used various tools but were heavily limited to the Google ecosystem, using *Google Translate* especially to machine-translate texts. Moreover, participants demonstrated varying preferences concerning the use of analog or digital technologies. The observations led the researchers to emphasize some needs in language-educational programs. First, students' various and personal approaches to using software should be recognized so as to be able to better understand students' digital literacies and act thereupon. Second, it should be aimed to expand students' digital literacies and practices beyond single ecosystems like Google's so that students become better able to understand and find other software solutions that might be better suited for their tasks. Finally, the researchers implore that instructors should not frown upon students' use of machine-translation tools, but that they should instead explain how such tools can be used "judiciously rather than indiscriminately" [42, p. 18].

In [55], the researchers conducted a large-scale, longitudinal, quantitative and qualitative study in Mexican secondary schools' English classes across the country. This study was conducted in or around 2017; the situation may thus have changed in the meantime. Nonetheless, the study provides insights into instructors', students', and institutions' attitudes toward technology in language education. Based on their analyses, the researchers drew four conclusions relevant to this thesis. First, the observed instructors were highly motivated to use technology to support instruction, but rarely made use of tools provided by the schools and instead brought their own tools and materials. It is thought that "this be-

haviour [is] related to overwhelming regulations and limited access to facilities, technical support, time investment, and training" [55, p. 38]. Second, it was observed that especially older instructors stuck to simpler technologies like CD players compared to younger instructors, suggesting that any support or training should carefully take instructors' technological literacies into account. Third, if technologies were used in the classroom, they supplemented textbook-based instruction and never replaced it, for which the primary reason is argued to be that the integration of digital technologies was a relatively recent development and could not have replaced all instruction yet. Finally, technologies seem to have been used primarily for motivational purposes, or turned out to be more motivational than instructional. For instance, some instructors experimented with L2 English text messages to communicate with students, but rather than communicating with the aim to use English correctly, students sometimes threw in English phrases in messages that were otherwise mostly in Spanish. The students did perceive this interactional experience positively.

Both studies emphasize that if technology is to be used in a language-educational setting, it is important that educational institutions facilitate technological literacy training. Additionally, people have different digital competencies and preferences, which should be taken into account in literacy training and in language curricula. Lastly, the studies suggest that there exists a potential gap between how instructors desire or envision students to use technology and how students actually use technology in the context of language-related tasks or study. For personalization purposes, it thus seems important not to demand of students that they use some prescribed technology at all costs as part of a curriculum.

3.3.5. (E) Education-Technological Framework Designs

Lastly, [87, 121] propose framework designs aimed at facilitating future implementations of educational systems or technologies. First, [87] refers to the concept of *Assessment for Learning* and to Vygotsky-inspired theory on *teaching praxis* to argue in favor of an educational framework wherein assessment takes a key role in language development of L2 English students. The researchers aimed to promote development of such a framework by preparing a toolkit with assessment materials and facilities that foster continuous evolution of instruction methods through active reflection at a classroom level, institution level, and national or greater level. To achieve this, the researchers implemented an online platform, *TEAL*,[34] in 2015, containing resources on L2 English instruction and assessment, concrete assessment materials, practical instruction examples, and discussion forums for instructors. The platform appears to have been made use of by several dozens of primary and secondary schools in Australia, suggesting a significant interest in the system. Some tools, however, seem to no longer be in use, such as a benchmarking forum which contains outdated resources according to the website, and the instructor discussion forum page, which returns a 404-not-found error. Moreover, in [87] and on the TEAL website, the necessity of assessment in language education appears to be presupposed, without elaborate explanations on the concrete role of assessment in language learning. Nonetheless, the idea to have instructors—not just institutions or policymakers—cooperate to collectively improve on existing instruction methods is unlike any of the previously discussed studies.

[121] proposes a VR-content development framework in response to limited adoption of VR technologies in L2 English instruction scenarios. The framework consists of two "frames": one "immersion frame" that an instructor with little software knowledge can easily edit, and a "base frame" that relates more directly to the VR hardware, software, and interfacing aspects [121, p. 1004]. Concretely, the immersion frame could be some web platform with study materials, and the base frame some device-dependent interface prepared by an educational technologist that displays the web platform. Through this division, VR experiences could theoretically be easily adjusted by an instructor according to specific groups of students or courses at a low cost; meanwhile, the base frame would only need updates if the VR hardware or operating system changes significantly. In practice, however, this system seems to have limited applicability due to the high development rate of VR systems: it was observed that "[e]ven core functions found in a base frame, such as picture-in-picture, may come and go due to [frequent] operating system or program updates" [121, p. 1007]. Still, the proposal to simplify development of educational software by decoupling content and implementation may be of value, perhaps particularly on platforms that can be expected to have a more stable base frame like PCs.

In light of personalization of language education, the systems proposed in [87, 121] seem to hold potential,

[34] *Tools to Enhance Assessment Literacy for Teachers of English as an Additional Language*, teal.global2.vic.edu.au.

both on their own and as a combination. One could, for instance, imagine an online platform where students and instructors continuously reflect on the present educational system from the classroom level to the institutional or (inter)national level. By facilitating active discussions with the possibility of quick inputs from various actors, it may become more feasible to cultivate a united effort aimed at the development, evaluation, and integration of technologies for personalization purposes. Such a platform could be coupled with a system that allows instructors to more easily create digital study experiences without the need for programming knowledge, as in [121]. A platform built by and for a broader range of actors could play an important role in combating the issue of small-size and homogeneous samples that regularly appeared in this chapter.

3.3.6. Discussion and Conclusion

This section discussed 56 studies related to interplays between technology and language education, divided into five categories (Table 3.2): technology experiments (T), studies on perceptions of technology (P), computer-technical system designs (C), observational studies (O), and education-technological framework designs (E). Observations from each category are consolidated in a general discussion below.

3.3.6.1. Perceived Benefits of Educational Technologies

Several benefits of the application of software in language-educational contexts surfaced repeatedly. One important perceived benefit is that software facilitates the provisioning of materials and exercises beyond those offered by the instructor or a textbook (T, P, E). Not only does this enable students to interact with language materials in more varying ways, it also enables them to do so online, outside of class (T3, P2). In this regard, students seemed to especially use or look for materials for vocabulary reference or training. Furthermore, there are developments in areas that provide students with personalized material or course suggestions, which seems increasingly possible as technology becomes more sophisticated (P1, C). Another perceived benefit is that software has the potential to foster communication and collaboration among students and between students and instructors, which appears to be perceived positively (T1, T3, P2). Here, software can provide a layer of anonymity, which is seen to be especially helpful for language students with higher anxiety levels or in contexts where direct student-instructor contact can be socially difficult (T1, P2). These benefits have been combined in multiple studies in blended-learning scenarios, occasionally complemented with gamified study elements, which appeared to have a motivational effect on students (T).

It is, however, difficult to assess to what extent it is the affordances of software technologies that bring the abovementioned benefits, or that the deviation from traditional textbook-based instruction importantly weighs in on this perception. The vast majority of handled studies only compare some technological intervention with the baseline of the ‘traditional’ scenario where students are instructed by way of a textbook. Moreover, there are but few controlled experiments that compare usage of various types of educational technologies or different implementations of some technology with each other.

3.3.6.2. Perceived Disadvantages of Educational Technologies

Perceived disadvantages related to the use of digital technologies also surfaced. For instance, injudicious use of AI tools is perceived to aggravate the risk of overreliance on them and is feared to impair users’ critical-thinking abilities and human connection (T2, P1). Furthermore, allowing students to attain more study materials through software risks confusing or overwhelming them, especially when materials are presented in an unsystematic way (T1). Finally, it has been shown that educational software may work counterproductively with regard to the study tasks students are subjected to by their instructors, for instance in games or highly immersive tools where software potentially forms a distraction (T1).

3.3.6.3. Implementation Difficulties

In addition to these potential disadvantages which may discourage implementation of technologies, instructors are faced with structural hurdles that inhibit the choice and efficacy of technologies in language courses. Studies stress that technological interventions are significantly less effective if they are not taken into formal account in the curriculum (T, P, O). This causes inadequate allotment of resources to train instructors and students to better understand how certain technologies work, what possibilities they bring, and how they answer to instructors’ and students’ needs, if at all. Instructors are also given insufficient time to get to know educational technologies, design courses that can benefit from such technologies, and make good use of them within courses. These issues are not local or specific

to cultural-economic circumstances, but appear among numerous of the discussed studies conducted across the world. Other issues such as lacking ICT infrastructures, uncomfortable software interfaces, and the deployment of tools that do not match students' desires also contribute to mixed perceptions of employed technologies (T1, T3, P2).

3.3.6.4. Technology and Student-Centered Curricula

These structural hurdles seem to contrast with a widespread interest among instructors and institutions to use technologies to transition toward 'student-centered' as opposed to 'teacher-centered' curricula.[35] Despite this interest, the studies indicate that (digital) materials, the pools from which they originate, and language-educational technologies continue to be curated by instructors, not or to a much lesser extent by students. Concerning AI tools, students seem to be met with distrust, driven by suspicions that students misuse AI to 'cheat the system' (T2, P1). The attitude that students are not expected to have as much freedom in the choice of materials and technological tools as their instructors—despite SCL-based instruction—suggests that student-centeredness is importantly delineated by the stipulations of curricula and instructors, rendering it remarkably instructor-centered. The studies thus suggest that student-centeredness is not as much about personalization as it is about the way students are made to engage with usually prescribed materials and tools, supporting the remarks made in Section 3.2.3.

3.3.6.5. Implications for Educational Personalization Through Technology

This section bears important implications for technology-shaped personalized language educations. The first is that present efforts to promote student-centered study are not inherently efforts to support educational personalization, at least not in the sense of allowing students to select materials they would like to use and pursue goals they aim to achieve. Second, the above discussions suggest that not all students appreciate the same kinds of technologies, although students are nonetheless subjected to using the same technologies as their peers once incorporated in a curriculum. It follows that curricular reform may be required to enable technologies to successfully complement language education for personalization purposes, but that it is not sufficient. Third, the studies reflect a persistent desire to perform regular quantitative or summative assessments, both by instructors within a course as well as by academics within research setups. As long as assessment remains a principal means to analyze or advance students' study journeys, the possibility cannot be neglected that students get to desire methods and materials that assist them in passing their assessments rather than those that may assist them in learning.[36] If this is the case, then adjusting curricula to fit in technologies according to students' wishes may contribute to educational personalization, but it risks perpetuating the problem mentioned in Section 3.2 that education facilitates study at the expense of learning (see also Section 2.4).

3.3.6.6. Personalization Possibilities Within Curricula

To conclude, we visit the question of how technology *can* be used to foster personalization in language education, despite curricular and other constraints. The studies discussed here indicate that it seems unavoidable to adjust the curriculum at least to some extent to be able to make effective use of technologies within it—not adjusting the curriculum raises the aforementioned problems of limited time, support, and efficacy, among others. If the aim is to personalize language education to the extent that students have full freedom to choose study materials, curricular goals, technologies, assessments, and so on, then the curriculum arguably requires immense structural changes. If one settles with personalization to a lesser extent, then parts of the curriculum may stay intact.

Digital technologies may prove powerful aids to navigate the implications associated with adjusting curricula based on a chosen level of personalization. For instance, offering materials in a more personalized way can be eased by making use of deep-learning solutions such as those proposed in C to recommend materials fitting for a student. Exercises and assessments can similarly be adjusted to more closely match students' needs and goals by employing generative AI. In tandem, personal guidance could be made more feasible by way of chatbots, or in fact by human instructors who potentially save time if they are no longer concerned with planning and enforcing rigid curricula.

[35] This was explicitly touched upon in P2, but many of the studies in the other categories that attempt to motivate or otherwise assist students seem to do so with student-centeredness in mind, which seems to be treated as an indication of the extent to which students perform group work or use non-textbook materials (e.g., see [1, 34, 41, 46, 133]).

[36] See for instance [48, esp. pp. 16-17] for a similar discussion in the context of learning strategies among students.

Even so, making an informed decision on the extent to which a curriculum is to be reformed should ideally be done before technologies are implemented. This is challenging since it has been shown that institutions and instructors may not be fully aware of the capabilities of digital technologies. Increasing technological literacy among instructors and policymakers thus appears paramount. Additionally, it may be fruitful to investigate the possibility of a network that allows for research on the intersection of education and technology to be more centrally organized and to be better informed by all involved actors, as proposed in E. Such a network may make it more feasible to conduct larger-scale studies with more heterogeneous samples, thus facilitating discussion on curriculum reform or technological personalization interventions. The efforts presented in this section to investigate the role of technology in language education form an important basis, but are also disjoint efforts that show little signs of trans-institutional and transnational coordination, highlighting the potential of the above.

In consideration of Section 3.2, it seems, however, not enough to repeat the same experiments with larger sample sizes. Most of the discussed studies demonstrate no thorough, critical reflection on learning, development, and behavioral theories. Why do VR- and AR-related studies in T1 yield mixed results due to distraction even though they were designed precisely to help students learn? What prompts students to 'misuse' AI (T2, P1)? What, in the first place, does it mean for students to be motivated to study and for students to learn? It is natural that research prompts new questions in response to obtained results, but the above important questions appear to be largely overlooked in the discussed studies. As suggested in Sections 3.2.1 and 3.2.2, examining education and technology therein from angles like the recent free-energy principle may yield important directions to answer such questions.

3.4. Instruction Considerations of L2 Japanese

This section reviews literature on Japanese language-specific considerations in the instruction of L2 Japanese (**RQ3**). Scopus and the Leiden University Libraries Catalogue (LULC) were searched and filtered according to the inclusion and exclusion criteria in Table A.1, resulting in n=4 English-language and n=6 Japanese-language studies, respectively (Table 3.6). Tables A.3 and A.4 present a summary of the studies' focuses and key findings.

Table 3.6: Number of found publications after filtering for **RQ3**. The number of results after search with the filter queries excludes duplicates. See Section A.2 for search queries.

Source	n **after search with filter query**	n **after manual filter (title and abstract)**	n **after manual filter (full-text)**
Scopus	27	6	4
LULC	32	20	6
Total	59	26	10

The studies highlight aspects of the Japanese language that instructors and researchers emphasize in L2 Japanese instruction.[37] These aspects are listed in Table 3.7 and discussed separately below.

Table 3.7: Japanese language aspects highlighted in the selected studies.

Aspect		**Studies**
Scripts	*Kanji* writing	[44, 45]
	Hiragana reading	[144]
Spoken Japanese	Mora-based syllabary	[145]
	Short and long vowels	[144]
	Disfluency in terms of pauses, fillers	[146]
Sociocultural	*Keigo*	[116]
	Various (gender differences, nonverbal communication, etc.)	[29]
Grammar and vocabulary	Transitivity	[141]
	Multi-word units	[142]
	Collocations	[143]

[37] See Section 2.3 for a background on the Japanese language where some of these aspects are preliminarily introduced.

3.4.1. Japanese Scripts

Concerning *kanji*, Hasegawa points out two problems in L2 Japanese-language instruction [44, 45]. First, some fonts render *kanji* differently from what they look like in handwriting [44]. If students study *kanji* through means that use such fonts, there is a risk that students incorrectly learn how they are written. Meanwhile, Hasegawa argues that there is little time in class to focus on writing *kanji* because class hours are limited. As such, "it is important that instructors equip students with strategies which allow them to work effectively in online self-learning environments," with the remark that instructors remain necessary to highlight aspects like stroke directions and radicals [44, p. 171]. Second, since writing *kanji* properly is deemed important to foster reading and writing proficiency, all students are typically expected to physically write *kanji* up to some level of proficiency, even though some students are hindered in doing so due to disabilities [45]. To prevent disadvantaging these students, Hasegawa proposes not to obligatorily assess *kanji* writing and that other skills may be assessed instead depending on students' study purpose and capabilities [45]. Finally, concerning *hiragana*, Motohashi and Ishizawa found that showing a word's reading in *hiragana* helps (especially beginner-level) students pronounce short and long vowels correctly more often compared to when students come up with the reading themselves [144]. They hence suggest that it may be beneficial for instructors to acknowledge the advantageous role of displaying *hiragana* in pronunciation practice [144] (i.e., as opposed to showing only *kanji*, *romaji*, or no reading at all). These kinds of benefits are especially helpful considering the scarce opportunities to spend time on pronunciation guidance in class [144].

3.4.2. Spoken Japanese

As Motohashi and Ishizawa pointed out, correct pronunciation of long and short vowels appears to be challenging for (beginning) L2 Japanese students and is subject to further attention in class and research [144]. At the basis of the long-short distinction in vowels lies the mora-based speech rhythm which, according to Hatasa et al., is relatively unique to Japanese [145]. As such, L2 Japanese students experience difficulties in acquiring a natural rhythm and accent patterns both in terms of production and perception. These difficulties relate partly to a student's native language (L1), where different difficulties are associated with different L1s; however, this does not explain everything because some difficulties arise regardless of L1 [145]. Regarding this, Hatasa et al. point at a research gap to pinpoint pronunciation error sources in L2 Japanese speech [145]. In addition, speech evaluation occurs under the presumption that students aim for native-like pronunciation while students' goals may be different, like to simply be understood without misconceptions or to leave a good impression [145]. It is implied that these goals, the remaining gaps in research, and students' native languages should all be taken into account in evaluation practices. Lastly, Guo argues that although disfluency in speech (e.g., pauses, fillers) has generally been regarded negatively, it in fact fulfills important roles in Japanese conversation such as expressing emotion, navigating politeness, and repairing speech mistakes, both for native and non-native speakers [146]. Disfluency, then, should not be seen as a sign of incapability, but as an important feature of language that should be instructed explicitly so that students can put it to effective use [146]. Explicit instruction is especially important since Guo suggests that different languages have different disfluency-related strategies, and while some strategies may transfer naturally to L2 Japanese, others may not [146].

3.4.3. Sociocultural Aspects of the Japanese Language

Language involves not only speaking, listening, writing, and reading, but also understanding various sociocultural aspects such as those that manifest themselves in expressions of politeness, regional and gender differences, and nonverbal communication. Endo argues that L2 Japanese textbooks pay insufficient attention to these aspects even though they are vital to effective communication [29]. To ensure students do get to know these aspects, Endo suggests that instructors "should stretch beyond the scope of Japanese language and culture presented only in textbooks […] paying more attention to 'authentic' Japanese that the Japanese people speak in the real world," for instance by making use of videos, music, and books [29, pp. 498-499]. Similarly, Takeuchi argues that textbooks and instructors treat *keigo* too simply and ignore its various strategic functions that support individual communication goals and that go beyond acknowledging hierarchical relationships [116]. One reason for a simplistic treatment of *keigo* in L2 Japanese study contexts appears to be a "native-speaker bias" according to which foreign students are not expected to be as fluent in *keigo* as native users because *keigo* is so complex [116, p. 590]. Still, L2 users are regularly not seen as "legitimate speaker[s]" due to insufficient *keigo* proficiency

[116, p. 590]. Takeuchi therefore advocates that instructors "should not apply different standards to L2 speakers or lower expectations for what students can accomplish" concerning *keigo* [116, p. 603].

3.4.4. Japanese Grammar and Vocabulary

One aspect of Japanese grammar that is deemed difficult at all levels is that of transitive and intransitive verbs [141]. Nakaishi points out that these verbs are acquired in two phases—recognition and then application—but that these phases are, in instructional practice, regularly not distinguished from each other, hindering acquisition [141]. By paying specific attention to the two phases, Nakaishi hopes that students can be relieved from the burden of having to simply memorize transitive and intransitive verbs from the start [141]. Liu investigated the processing of multi-word units (MWU; e.g., 'in front of' in English) [142] and incidental acquisition of collocations [143] in L2 Japanese. MWUs in Japanese (e.g., にも関わらず *ni mo kakawarazu* 'in spite of') are often treated simply as grammar points, but Liu argues that it could be more beneficial to treat them as MWUs and to acknowledge that students' L1s and MWUs' frequencies of occurrence in Japanese impact students' processing capabilities [142]. How this would affect classroom instruction is left as a question for future research [142]. Similarly, students' L1s are not taken into account in collocation (e.g., ぐっすり眠る *gussuri nemuru* 'to sleep soundly') acquisition assessment, and collocation acquisition is often researched through "traditional formats such as multiple-choice, fill-in-the-blank, and translation [questions], while frequently lacking contextual information" [143, p. 10]. This is problematic because a large portion of research concerning acquisition of L2 Japanese collocations focuses on L1 English users, while there is a good chance that different L1s and different question types yield different insights [143]. The implication is that students' L1s should be considered when treating MWUs and collocations in research and instruction.

3.4.5. Discussion and Conclusion

The studies discussed in this section handle a wide range of Japanese language features. Two overarching observations stand out. The first is that various studies point at limited adequacy of textbook materials. This surfaces explicitly in the instruction of *kanji* when stylized fonts are used or the existence thereof is not explained [44] and of sociocultural aspects that are treated in an oversimplified manner [29, 116]. Implicitly, it follows from a lack of attention for the practical roles of disfluency [146], acquisition stages of (in)transitive verbs [141], and students' L1s [142, 143, 145, 146]. Second, more attention to the individual student seems warranted for effective L2 Japanese instruction. This follows from a need to acknowledge interplays between students' L1s and peculiarities of the Japanese language, which may affect acquisition of a natural speech rhythm [145], pauses and fillers [146], multi-word units [142], and collocations [143]. Additionally, other factors such as disabilities [45] and personal study or communication goals [116, 145] are argued to require more serious consideration so that students can study to the best of their capabilities and wishes. Important causes of a lack of textbook or instruction quality and personal attention are limited classroom time [44, 144], gaps in research [142, 143, 145] or between research and classroom practice [141], and biases toward L2 students [116]. In conclusion and to answer **RQ3**, the studies propose concrete instructional considerations to instruct the discussed language facets more effectively, summarized in Table 3.8.

This literature review contributes to an understanding of instructional methods concerning specifically the Japanese language, capturing a wide but limited set of features discussed in English and Japanese literature. Nonetheless, the review is limited in several ways. First, the filter queries, inclusion and exclusion criteria, and final selections were made exclusively by the author. Collaboration with additional evaluators could have yielded different results. Second, the selected studies leave a large number of Japanese-language features unhandled. Finally, some of the handled studies are importantly based on instructors' classroom experiences, limiting generalizability. A wider search that moreover includes studies written in different languages could shed a more comprehensive light on L2 Japanese instruction.

Table 3.8: Concrete instructional considerations following from the discussed studies.

	Instruction consideration	**Study**
J1	Ensure students are aware of *kanji* fonts that deviate from handwriting	[44]
J2	Do not force students who are hindered in writing to perform *kanji* writing assessments	[45]
J3	Acknowledge that displaying *hiragana* during pronunciation practice may be beneficial to students	[144]
J4	Take students' L1s and study goals into consideration in the practice and evaluation of rhythm and accent in speech	[145]
J5	Instruct practical uses of disfluency in Japanese and reflect on how these relate to disfluency strategies in students' L1s	[146]
J6	Complementarily employ multimodal study materials to instruct sociocultural aspects	[29]
J7	Do not instruct *keigo* under the assumption that non-native students cannot reach or will not need proficiency	[116]
J8	Instruct (in)transitive verbs in a separate recognition and application phase	[141]
J9	When instructing multi-word units, consider their frequencies of occurrence in Japanese as well as students' L1s	[142]
J10	Consider students' L1s and various assessment methods when instructing collocations	[143]

4

Focus Group Study

The previous chapter provided theoretical and empirical views on learning, language education, and the role of technology therein. To situate these findings in a practical context, this chapter presents a focus group study conducted with instructors of L2 Japanese at Leiden University in the Netherlands. The principal aim of this study is to discover instructors' views on technology-assisted personalization within their program. To this end, a discussion session was held revolving around the following question:

(RQ4) How do university-level instructors of L2 Japanese in the Netherlands perceive the integration of technology for the purpose of personalization of L2 Japanese education?

The methodology used in the focus group study is presented first, followed by the study results and a discussion thereof.

4.1. Method

4.1.1. Participant Recruitment

The author approached staff members instructing Japanese at Leiden University by email. No restrictions were imposed on participation. n=3 participants were recruited. Since Japanese-language courses are tightly adjusted to one another on the basis of regular communication between instructors, one group of three participants was deemed sufficient for this study.[38]

4.1.2. Study Setup

The participants and author took part in a discussion session with **RQ4** as the main topic. The author prepared a discussion guide (Section B.1) to ensure the discussion stayed relevant with respect to **RQ4**. The guide follows the flow proposed in [49], consisting of an introduction, opening questions, introductory questions, key questions, and closing questions, where the key questions serve as the principal source of study-relevant information. This flow facilitates systematic proposition of questions based on the theoretical background from Chapter 3 and is purposefully constructed to encourage participant interaction. The discussion lasted an hour and a half, with one participant leaving after half an hour. The author ensured that the key questions were discussed with all participants.

4.1.3. Study Setting

The discussion group gathered in a private room in Leiden University's main library. Participants sat together around a table and were informed of the study method and purposes before commencing the discussion. Audio was recorded with the participants' unanimous consent through microphones placed on the table. The author served as the moderator during the discussion session.

[38] Participants' views presented in this chapter nonetheless represent their personal views and do not officially represent the university's stance.

4.1.4. Ethical Considerations

At the time of recruitment, potential participants were made aware that their cooperation is voluntary, that any data resulting from their participation is anonymized, and that they may cancel their participation at any time before or during the study. Before commencing the session, these rights were repeated, and participants were asked to sign an informed consent form and whether they permit the author to audio-record the session. After the discussion, audio was transcribed by the author using *Microsoft Word*'s transcription function, followed by manual corrections. Any personally identifiable information such as participants' names or detailed descriptions of personal backgrounds were redacted. The audio recordings were deleted afterwards. The study and tools used during the study were approved by the TU Delft Ethics Committee.

4.1.5. Data Analysis

The data resulting from the discussion session is thematically analyzed in Section 4.2. Thematic analysis lends itself well to this context for its flexible applicability, rigor, and the room that it gives for interpretative depth [3]. Braun and Clarke's widely adopted six-phase framework is followed to systematically perform the analysis [3].

4.2. Results

The transcribed contents of the discussion session were thematically analyzed, resulting in the codes and themes in Table 4.1. Examples of quotes from the discussion transcript that support each code are displayed in Table B.1.

Table 4.1: Codes and themes extracted from the transcribed discussion session data.

Theme	Code	Context
(1) Potential for technology to improve educational processes	(a) Digital technology to save time or achieve greater reach	Digital technology is used instead of some analog alternative to save time or bridge space and time
	(b) Coaching role	Situations wherein a mentor or coach figure could be beneficial to student learning
	(c) Community	Related to the forming of communities among students
	(d) Motivation	Related to cultivating motivation in students (e.g., to get to understand the material rather than to focus on grades). Refers both to intrinsic and extrinsic motivation
(2) Counterproductivity as a result of the application of technology or personalization	(e) Counterproductive use of technology	Some form of technology has been applied in some educational setting for the purpose of aiding students or instructors, but works counterproductively in practice from a pedagogical point of view
	(f) Technology does not fit curriculum	Some form of technology has been applied in some educational setting, but does not fit in the existing curriculum, or does not serve a carefully thought-out purpose
	(g) Counterproductive personalization	Situations wherein personalization is argued to work counterproductively for students
	(h) Purposeful avoidance of technology	The application of some technology is purposefully avoided in the educational setting because its use is deemed counterproductive
(3) Effects of the educational and political system on instructors, students, and learning	(i) Disabling system	The educational system has rules or regulations in place that prevent or disable the implementation of some technology, despite its argued potential to improve education
	(j) Lacking support	Instructors receive inadequate (technical) support to integrate technology in education, or time to discover the technology by themselves
	(k) Student-instructor relation	Related to the relation between students and instructors

Three themes were interpreted in the data. Theme 1 encompasses instructors' beliefs that technology can be used to improve instruction or study in educational settings. Theme 2 points at experienced or expected counterproductivity in instructing or studying resulting from the application of technology or personalization in education. Theme 3 pertains to the control that the educational or political systems are perceived to exert over instructors, instruction, students, and study. Quotes are from the study participants' contributions to the discussion and were translated from Dutch to English by the author.

4.2.1. Theme 1: Potential in Technology

The participants identified several potentials in technology to enhance education in various ways. Some prominently discussed possible improvements relate to saving time for the instructor through digital technologies (a). Particularly, assessment design and evaluation are recurring tasks, but cost instructors tremendous valuable time. Participants suggested that software could help save time by facilitating automated *furigana*[39] generation on exam texts, relaxing some of the rigidity in existing online assessment solutions like *Remindo*,[40] and automatically generating feedback on students' written works. In such cases wherein existing software was deemed too inflexible, integration of generative AI was seen as a possible improvement. Another potential was seen in facilitating access for instructors and students to broader ranges of materials or exercises. For instance, it was suggested that software tools that enable automatic generation of video subtitles and exportation thereof into digital flashcard decks could help broaden the range of materials used in class. Furthermore, the participants referred to resource platform *KITSUNE*[41] which was recognized to make it easier for students to find practice exercises. Finally, applications like chatbots utilizing generative AI were considered to be potentially helpful for students to practice spoken conversation at their convenience.

Apart from these concrete suggestions, the participants frequently considered *roles* that software could play in the educational setting, which were especially referred in relation to the topic of personalization. One frequently discussed role was that of a student coach (b). The consensus was that, if a software tool like a virtual agent were to be implemented in the educational setting, "[i]t should not be a cheerleader, but a sports coach": rather than providing students with on-demand solutions, it should challenge them to do things they find difficult and prompt them to think and reflect, motivating them to go through an actual learning process (d). As such, if students were to engage in dialog with such an agent, the agent should keep track of and move students to reflect about where they stand with respect to their personal learning goals (or curricular objectives) and what their next learning or study step could be. This is deemed especially helpful because "very often, students can't at all articulate what they want and where they want to go, while they do have some kind of gut feeling as to 'I like this, I don't like that, that interests me.'" Apart from coaching a student in this way, it was suggested that this kind of interaction—if combined with frequent contact with an instructor—could pave the way to more flexible assessments: the agent could, for instance, recognize that a student is proficient enough to take an oral examination early on in a course so that they could focus more on reading practice thereafter.

Closely related to this coaching role is software's potential to motivate students (d) and create a sense of community among them (c). Participants agreed that a coaching agent "should not make it more difficult, but more motivating" for students, and that it could help reintroduce the "fun in learning," something that is observed to be missing when students are encouraged to chase passing grades rather than to enjoy and make use of their learning process. Rather than making studying more fun through means such as gamification, the participants emphasized that community formation can be a great driver. Technology, then, should prevent students from engaging in excessive competition among each other and instead encourage students to come together to face (study) challenges. Here, an agent could help by prompting students to discuss some question with their peers, to find conversation partners,

[39] See Section 2.3. Participants suggested that existing tools that automate *furigana* generation are limited in that they tend to cause formatting issues and that they are not easily tuned to students' expected reading proficiency.

[40] *Remindo* (www.remindo.com; accessed February 8, 2026) provides functionalities to conduct online assessments and exams. A problem emphasized by a participant who had experience using the platform is that it is difficult or time-consuming to upload exam question sets with multiple possible answers per question, occasionally causing right answers to be marked as wrong, which the instructor had to correct manually.

[41] KITSUNE (project-kitsune.nl; accessed February 8, 2026) is an online resource platform with study materials, summaries, and exercises related to courses in the Japan Studies program at Leiden University, built by the student association of said program. Although its capability to share resources was recognized, instructors observed shortcomings which are discussed in Section 4.2.2.

or to get in touch with some senior student who might have expertise in some topic, which is argued to be difficult within the present system. In programs like Leiden University's Japan Studies, moving students to reach out to each other in a cooperative manner is deemed especially helpful. On the one hand, language acquisition and cultural understanding play an important role in the program. On the other, instructors also occasionally observed unhealthy competitive tensions in relation to the second-year exchange to Japan, which is only available to a fixed number of students and poses grade requirements that determine the location and length of one's stay in Japan.

In conclusion, the participants identified potentials for software to save time in assessment design, evaluation, and material selection, to encourage students to think about their learning goals and process, and to create a sense of community among students. Technology-wise, generative AI is seen as a promising means to relieve the rigidity in existing software tools that lack the ability to dynamically interpret texts and contexts. A virtual agent-based application that leverages generative AI is thought of as holding potential to facilitate all of the above.

4.2.2. Theme 2: Counterproductivity Due to Technology or Personalization

Alongside the various potentials discussed in Section 4.2.1, the participants cautioned that some forms of technology or personalization could work counterproductively in the educational setting (e). When it came to students' use of technology for course work or in class, the participants had observed that there is a risk of overreliance by students, preventing them from learning to take initiative, be communicative, and take a critical stance. Instructors fear that students will use tools like generative AI for anything and everything "once you allow it for one thing," and "that it is very difficult to teach students [to uphold] a critical stance when working with this kind of [tools]." One participant, for instance, reported having heard of students who exclusively interact with chatbots for study support rather than with peers or instructors, which they believed goes too far and hampers community formation. Similarly, while KITSUNE (see Section 4.2.1) may be a handy resource platform, instructors observe that students tend to be more inclined to reach out to the platform rather than to instructors for help or course content discussions. Contact between students and instructors is furthermore hindered when students use laptops throughout class: one participant observed that "technology often forms a distraction" because students "keep themselves busy with a hundred other things [than the class contents] if they have access to their computer [in class]."

In relation to technology and how it fits in the existing curriculum (f), one participant remarked that "very often it comes down to 'Here we [the university, educational software providers, etc.] have a nice idea,' but how that fits within the curriculum or what one is actually going to do with it that reaches a learning goal, that seems to be lacking a bit"; a view that enjoyed consensus among the participants. For instance, one instructor explained that a recently renovated university building had new "active learning classrooms where there are extra whiteboards and [...] the chairs and tables could be easily moved," but that no one had thought about how to use these rooms in practice for their classes. The same appears to happen with software applications. This results in the existence of large suites of software tools that instructors are more or less aware of but that cannot be readily integrated into their courses as-is, and so they are not. On the other hand, the participants pointed at tools like KITSUNE that *are* actively used, but that are disconnected from the curriculum since they use different terminologies or resources than are used in class and whose quality cannot be controlled. It is thus argued that when purportedly innovative solutions are forced into the curriculum without due consideration, they cause confusion and potentially hinder students from developing the skills that instructors expect students to learn.

An educational development that the participants had observed to be increasingly underway—and to be advocated for by the university—is personalization in curricula and study paths. Personalization, here, is seen as a way to increase the number of places in the curriculum where students have freedom of choice, as well as a general movement to act more on students' individual needs and preferences. In an attempt to follow this development, one participant, had removed their course's mandatory participation requirement for a semester so that students could decide for themselves whether they join classes. As a result, however, students showed up less frequently, and a significantly higher portion of students failed the final exam (g). Another participant had decided to select reading materials for their reading comprehension class that more closely reflect contemporary students' interests, but indicated to make sure at least half of the materials consists of texts outside of their interests as well. This would prevent

students from ending up in a "filter bubble" and ensure they are confronted with a broad range of topics, thereby learning things about Japan they would otherwise not have. As for individual personalization strategies, participants acknowledged benefits of providing personal feedback on writing assignments, but argued against doing so excessively—be it by an instructor or some software tool—because it would be inconducive to creating an environment wherein students take initiative to ask questions about their writing or to join writing support classes.

A solution to counter abovementioned issues that instructors and the university appear to implement widely is the purposeful avoidance of technology in educational settings (h). In first-year language courses, for instance, a participant mentions that students' use of automatic translation tools or generative AI is expressly forbidden so as to prevent students from becoming overly reliant on such tools and to encourage them to maintain skills like critical thinking and handwriting. Similarly, some instructors are said to prohibit use of computers in the classroom entirely under the assumption that this will facilitate students' focus on the class material, the instructor, and peers, and to prevent them from simply pasting questions asked to them into a chatbot and reading the answer aloud without thinking. In the same line of thought, one participant indicated to have reduced the number of presentation slides they use in class in order to promote free discussion.

Summarizing, although there seem to be initiatives to introduce personalization in the curriculum and to increase the use of technology for this purpose, instructors are cautious. The potential deterioration of students' active attitude, community formation, and critical thinking as a result of overreliance on generative AI and other tools is seen as "an enormous problem." It is hence argued that critical reflection is required before integrating new tools or personalization strategies in the curriculum to ensure that students continue to be exposed to materials and instruction strategies that move them toward achieving the established curricular objectives.

4.2.3. Theme 3: Constraints by Educational and Political Systems

In themes 1 and 2, instructors' perceived potentials and pitfalls of software and personalization in Japanese language education surfaced. In connection to this, a recurring theme in the data is that, even in cases where instructors would like to implement some technology or personalization in their classes, their hands are tied due to fundamental limitations and constraints posed by the educational and political systems in which they operate (i). Participants explained that the university relies on funding, and that as far as they knew, the university receives funding for every European Credit students obtain, but only if the Credit counts toward a course within the curriculum and if the student graduates within four years in case of a bachelor's program. A direct consequence of this policy, a participant laments, is that "following extra lectures that are not going to be listed on your diploma, that's good for you [the student], but the university doesn't get anything [for that], so that's why it's discouraged, which is, of course, crazy because [as an instructor] you want to broaden the [interests of the] students." Another consequence is that experiments like one of the participants', where students were no longer obligated to attend class for a semester, cannot be sustained because they potentially cause study delays. This was seen as a loss, since even though the experiment resulted in more students failing the final exam, the participants regarded it as a valuable learning opportunity for the students: by failing the exam after skipping class, students might realize that going to class and interacting with peers and instructors helps them study more effectively. This experiment thus turned counterproductive for the reason that it operates within a system where study delay is unacceptable.

Participants also argued that little time is made available to them to personalize study. For instance, the participants agreed that introducing flexibility in assessment schedules—such that students can choose when during courses they want to take some examination—could be a good way for students to relieve stress and to help them focus on areas they need or want to. However, this would imply that instructors need to be ready all the time to conduct examinations, which is difficult to accomplish because instructors themselves need to operate within tight academic calendars. Initiatives like giving "each of the staff members […] six [students] under their wing at the start of the first year and [having them] guide [the students] for three years" are also implied to be "impossible because half of the [instructors] don't want to go along with that" due to time limitations. This time problem similarly surfaces in the context of the implementation of new technologies in courses (j). A repeatedly mentioned issue is that instructors would need to invest significant time—which they feel they do not have—in figuring out how software

tools work, whether their usage has ethical implications, and how they should fit in the curriculum. While the university hosts workshops on interactive tools and strategies yearly, participants felt that there is little to no adequate support to actually make use of such tools. Consequently, instructors are given little opportunity to act on potentials like discussed in Section 4.2.1 or to consider the benefits and disadvantages of technology and personalization as discussed in Section 4.2.2; innovations that seem interesting to them hence remain unused.

Finally, the relationship between instructors and students frequently came up during the discussion (k). This relationship was primarily viewed from the angle of social and cultural gaps. About the social gap, a participant said that as an instructor, "you have less feeling with the student, because you don't know what Discord is or because you never use ChatGPT, but also because at some point they see 'That dude's hair has already turned gray, there's no need going to him [for questions].'"[42] Similarly, because "students are perhaps slightly less scary than the instructors," instructors observe that students are more inclined to seek help from each other or through online platforms like KITSUNE than from instructors. A similar effect is observed online: students use their own channels to interact informally with each other or discuss course contents, and use official channels like *Brightspace*[43] or *Microsoft Teams* to deal with the necessary, more formal communications with instructors. It was suggested that finding a way to consolidate the two sides by creating a uniform platform or leveraging AI, for instance, might make students more comfortable reaching out to instructors.

The social gap seems to be affected by a structural hierarchical gap as well. Instructors have a certain status in relation to students: they are the ones to decide what materials students are to use, how they should act in class, what tools they are allowed to employ, et cetera; in cases where students pick tools or strategies they want to use themselves, instructors feel a need to exert some kind of control over them. It is not that instructors want to be in control of students per se: it is because there are structural constraints that "we [instructors] are not here to help students develop into their best self, but to deliver some sort of standardized product"; a necessary way of conduct "to survive as a program." As such, the participants felt that the strict funding system and limited time available in the curriculum and academic calendar necessitates the presence of some force that directs students toward their normative graduation—a force necessarily embodied by the instructors who are to take up the authoritative role.

Hence, as discussed in Section 4.2.1, if we were to adjust the program so that it "help[s] students develop into their best self" rather than into a "standardized product," introducing someone who or something that takes up the role of a personal coach—as opposed to a predominantly authoritative figure—is seen as a promising direction. To facilitate this, there must be some sort of technological intervention if the frameworks of the existing system were to remain in place since, after all, instructors have limited time and there are gaps between students and instructors. Nonetheless, participants warned of the danger that once such an intervention becomes normalized, it might end up replacing instruction staff because it is more convenient and cost-efficient to the "managerial layer." This would be problematic because even (or especially) if some form of coaching technology would end up playing a role in students' education, the participants argue that instructors would be required to establish checks and balances in the technology, both to ensure didactic quality and to provide adequate guidance to students.

4.3. Discussion

The participants discussed their views on possible roles of technology in personalized language education, negative sides thereof, and how integration of technology for personalization is affected by system-imposed hurdles. Identified benefits of software in the Japanese-language program were its potential to save time for instructors, fulfill a coaching role for students, foster community, and drive motivation. Software, then, could play a complementary, supportive role, particularly to facilitate things perceived to be difficult within the current system such as connecting with peers and seniors with similar interests, finding more study materials, and helping students discover goals and interests. Still, instructors fear students' overreliance on technology, especially concerning generative AI. Additionally, the participants pointed at a disconnect between university-provided tools and the curriculum and at a lack of technical

[42] *Discord* (discord.com; accessed February 11, 2026) is a platform through which users can create and join online communities and (video) call and chat with each other.

[43] Brightspace (www.d2l.com/brightspace; accessed February 11, 2026) was the learning management platform used at Leiden University at the time of the discussion session.

support and time to investigate or make use of technologies. In some courses, this leads to outright prohibition of students' use of certain tools. All in all, the participants recognized that in the current university system (see also Section 2.2), students are delivered more as "standardized product[s]" than as individuals that "develop[ed] into their best self." The participants indicated that a more personalized system, whether or not supported by technology, would thus be desirable. The degree of personalization would have to be limited though: for instance, providing students with feedback for each writing assignment was imagined to inhibit students' willingness or capacity to ask for targeted feedback after critical self-reflection.

The participants' insights show strong parallels with the findings presented in Section 3.3, demonstrated by their experienced lack of time and technical support, fear of students' overreliance on technology, and envisioned benefits of community formation and the provisioning of wider ranges of materials. Moreover, to make effective use of technology, this and the previously handled studies emphasize that software cannot simply be implemented in curricula without having the curricula take the technology into account. For this, instructors require institutional support.

In deviation from the observations in Sections 3.2 and 3.3, this study's participants did not refer to student-centered learning (SCL) as such, nor did they appear to be principally moved by what is typically implied by SCL. While the participants did seem to feel a responsibility to get students in contact with high-quality pedagogical methods that help them grow into critical and knowledgeable thinkers, they exhibited an understanding that semi-autonomous engagement with the curriculum contributes to the learning process. Hence, rather than controlling exactly what students do how, controlling *quality* is considered particularly important, for instance in study material selection. It followed that the participants did not treat certain tools as one-size-fits-all solutions since their principal aim is not to subject students to standardized methods, but to foster students' ability to think critically and to assertively engage in discussion, among others skills. Lastly, the participants did not discuss Japanese language-specific instruction considerations like in Section 3.4, except perhaps in passing when the importance of handwriting was mentioned (Section 4.2.2).

This study has several limitations. First, the sample was small in size and limited in heterogeneity, inhibiting generalization to other contexts. This setup is coincidentally one of the study's strengths because it highlights the particular views of university-level L2 Japanese instructors. Moreover, the thematic analysis, including the identification of codes, was conducted by the author without establishing some inter-coder agreement. While the author rigorously pursued accuracy and unbiased presentation of the discussion contents by actively reflecting on the author's positionality, the analysis could potentially have benefited from external code and theme reviews. Finally, this study did not directly include students' views, raising the issue of potential bias by the participants or in the results. This issue is, however, addressed in the evaluation study in Chapter 6 which revolves around students' experiences.

5

Application Design and Implementation

Section 2.1 elucidated this thesis' grounded-design approach. So far, Chapters 2 to 4 constructed the necessary background to inform the design. The present chapter concerns the design of a software product and a subsequent prototype implementation. First, the findings concerning technology and personalized language education from the preceding chapters are summarized in Section 5.1, followed by a design informed by these findings. Next, this design serves as the basis to derive an implementation in Section 5.2.

5.1. Application Design

Before constructing the design, the potentials and possible hurdles of software for the purpose of personalization are discussed following from the previous chapters, along with constraints within which the software is to operate. Then, a design is proposed that attempts to maximize benefit of the suggested potentials while avoiding the hurdles and respecting the established constraints.

5.1.1. Potentials and Hurdles

Tables 5.1 and 5.2 summarize the findings from Chapters 3 and 4 in terms of the potentials and hurdles, respectively, of software for personalization of language education. The potentials and hurdles are accompanied by the thesis sections in which they explicitly surfaced.

Table 5.1: Potentials of software in language education for the purpose of personalization.

	Potential	**Source sections**
P1	Present a wider array of study materials	3.3.1 to 3.3.3, 4.2.1
P2	Generate materials (e.g., automate *furigana* or exercise generation)	3.3.2.1, 3.3.3.3, 4.2.1
P3	Provide access to materials outside of the physical class environment	3.3.1.3, 3.3.2.2
P4	Automate feedback on student work to alleviate instructors' workloads and to allow students to practice more often	3.3.3.2, 4.2.1
P5	Track student progress or level to facilitate timely intervention, suggest materials, or propose next steps	3.3.2.1, 3.3.3.1, 4.2.1
P6	Encourage reflection on student progress or study content	3.3.1.3, 4.2.1
P7	Make study more motivating by providing or encouraging content and activities outside of a textbook	3.3.1, 4.2.1
P8	Facilitate student-student and student-instructor communication through a layer of anonymity or indirect contact	3.3.1.1, 3.3.2.2, 4.2.1
P9	Encourage constructive student-student communication or collaboration	3.3.1.1, 3.3.1.3, 4.2.1
P10	Facilitate assessment with flexible timing, content, or modality	3.3.3.2, 4.2.1

Table 5.2: Potential hurdles when applying software in language education for the purpose of personalization.

	Hurdle	**Source sections**
H1	Curricular and infrastructural constraints need to be attuned to technology and vice versa for successful technology integration	3.3.1, 3.3.2, 3.3.4, 4.2.2, 4.2.3
H2	Risk of overreliance on software by students and instructors, especially if software takes away too much work and critical reflection	3.3.1.2, 3.3.2.1, 4.2.2
H3	Generated or other dynamic non-textbook content may lack in quality or systematicity	3.3.1.1, 4.2.2
H4	Instructors have a need to supervise integration and use of software or dynamic content (e.g., for didactic and material quality control) which requires time and expertise	3.3.1, 3.3.2, 4.2.2, 4.2.3
H5	Student needs and capabilities are diverse, inhibiting the application of software as a one-size-fits-all solution	3.3.1.4, 3.3.4, 4.3
H6	Software potentially distracts from study	3.3.1.1, 4.2.2
H7	Software can be difficult to understand or use and may require training to operate	3.3.2, 3.3.4, 4.2.3

As suggested in Chapters 3 and 4, these hurdles seriously obstruct effective integration of technological interventions in educational settings and should thus be avoided. This can arguably be done either from the technological side ('how to design technology so that it avoids the hurdles') or from the curricular side ('how to restrain technology so that it avoids the hurdles'), or both. Table 5.3 presents suggestions that follow straightforwardly from Table 5.2 to alleviate each hurdle.

Table 5.3: Potential solutions to the hurdles in Table 5.2 from the perspectives of technology and curriculum.

Hurdle	**Technology-related solution**	**Curricular solution**
H1	Design technology so that it can be adapted to operate within (changing) curricular constraints	Change curriculum to fully account for the purposes and capabilities of some technological intervention
H2	Ensure technology leaves enough room for students and instructors to do work by themselves	Explicitly limit the (number of) tasks for which technology may or can be used
H3	Enable instructors and students to filter and structure content based on quality assessment and curricular objectives, respectively	Prescribe content (sources) that technology is allowed to use
H4	Design technology to allow instructors to adjust its content or functionalities	See H3; provide instructors with adequate time and technical support to make informed and sufficient decisions
H5	Design technology so that it is dynamically adaptable to individual students' needs	Do not require all students to use some technology; allow usage of different technologies without repercussions
H6	Ensure technology revolves principally around study tasks and content	See H2
H7	Design technology to be simple to understand and use so as to reduce the need for training and technical support	Provide instructors with enough time and resources for training and technical support

5.1.2. Constraints

Section 2.2 and Chapter 4 presented the current situation of the possibilities and constraints in the Dutch university system. There are several requirements that universities need to adhere to by law, such as that they may only grant students a diploma who have obtained a certain number of European Credits organized in a curriculum. It is furthermore financially beneficial for the university to have students graduate nominally. It can be assumed that these constraints are not alleviated in the short term and that any short-term solution should thus operate within them.

However, that does not mean that these constraints will not change later. In fact, in light of rising voices for education tailored more to the individual such as from UNESCO (Chapter 1), the Bologna Process (Section 2.2.4), academics in education (Sections 3.3.6 and 3.4.5), and instructors (Sections 3.3.6, 3.4.5 and 4.2.1), it seems conceivable that reforms will come in the future. Designing technology now that

adheres to the present constraints but does not take future changes into account would only cause H1 to occur repeatedly, each time raising the need for costly new designs and implementations.

In anticipation thereof, it would be useful for the technological design in this thesis to take such future reforms into account. To accommodate this, present and possible future constraints should first be mapped. Table 5.4 sketches three scenarios with constraints ranging from the present situation to a situation with the least possible constraints. Constraints are viewed from the aspects of courses, available time to complete courses and curricula, assessments, materials, types of classes, and employed technologies, reflecting key topics in the discussions in the previous chapters.

Table 5.4: Constraints under several scenarios ranging from the current situation to a situation with the least possible constraints. The constraints listed under the current situation reflect the usual situation at the Japan Studies program at Leiden University, but may deviate slightly depending on the course, program, or university.

Aspect	Current situation	Relaxed (illustrative)	No constraints
Courses	Students need to pass courses and obtain credits according to a curriculum path, with some room for electives	Students can more freely choose existing courses to fill their study path, or propose courses based on interest and available expertise	Courses only exist as suggestions; students can choose how they study what
Time	Students have a fixed amount of time to pass a course and a curriculum	Students have as much time to pass a course as needed to reach a state wherein they understand the course contents; curriculum duration is more flexible	Students are not bound by a specific time frame to complete their study
Assessment	Students are assessed at established intervals according to the curricular objectives; failing implies redoing a course later	Students pick from suitable forms of assessment at a time of choice; students can continue studying for the course and redo assessment if they failed	There are no formal assessments; students can opt for voluntary assessments
Materials	Materials are mostly selected by the instructor to work toward curricular objectives	Students can pick materials to their liking, potentially under instructor supervision, or students get to choose from a wider array of materials	There are no formal requirements to use some set of materials; students can choose with optional supervision
Class	Students and instructors meet physically in class at a fixed frequency for an hour and a half per class; some classes require attendance; content is determined by the instructor in accordance with curricular objectives	Classes are loosely based on courses; voluntary attendance; content is established on the basis of discussion between students and instructors	There are optional classes with no obligation of preestablished content; students and instructors can convene when desired for variable amounts of time
Technology	There is a learning management system (LMS) and some assessment and feedback tools; course materials are shared online; students are required to use the provided tools to complete assignments and assessments	An LMS and tools to facilitate study are provided; course materials and student-chosen materials are shared online; students can choose whether they use provided tools or not	Optional materials and tools are offered for students' use without obligations; if there is an LMS, its contents are flexible and importantly but voluntarily determined by students

Between the constraints in the current scenario and a scenario with none, there are many possibilities to loosen or tighten the constraints, both as a whole and per aspect. The Relaxed column in Table 5.4 may be regarded as some instantiation of a continuum of possible constraints. Similarly, the curricular solutions to the hurdles in Table 5.3 can be seen as specific constraint choices on this continuum. The continuum suggests that a particular implementation need not comply with all constraint situations at once due to the large number of options. Rather, an implementation could allow for simple tweaking of

settings to dynamically respond to constraint changes, or a technology could be conceptually designed so that it is uncomplicated to derive different suitable implementations depending on the constraints.

Now, the extent to which the suggested potentials in Table 5.1 are translated into features in an implementation depends on these constraints, but also on the level of personalization that is sought after. Throughout the design, it is assumed that personalization is maximized under the established constraints. For instance, in the current situation, students enjoy personalization to the extent that they may choose a number of elective courses and have some freedom to choose topics for writing assignments within some prescribed theme. In the situation with the least constraints, personalization can be said to be maximized if students have complete freedom to choose what they do to determine and advance their studies. Anything in between can be seen as the application of some degree of personalization on a continuous scale with respect to each of the aspects in Table 5.4.

5.1.3. Features

For a technological design to wholly overcome the presented hurdles while operating within the established constraints, ideally, all its features should contribute to solutions like proposed in Table 5.3. At the same time, they should make use of the potentials in Table 5.1 and aim for maximum personalization with regard to each aspect and constraint in Table 5.4. Table 5.5 presents a list of features designed on these bases. The possibility of future relaxations is accounted for in each feature with regard to Table 5.4. It is also indicated which aspects in Table 5.4 each feature pertains to ('technology' excluded since that involves the comprehensive design) and what hurdles the features may face themselves. The latter is addressed in the implementation in Section 5.2.2.

Table 5.5: Proposed features, what they do, why they were chosen, which potentials and aspects they relate to and how they could be adjusted for particular constraints (Tables 5.1 and 5.4), and what hurdles they may face (Table 5.2).

Feature		
Knowledge graph	**Potentials (aspects)**	P2, P3, P6 (materials)
	Hurdles to consider	H3, H7
	Basis and function	Allows students to organize their thoughts and materials in the form of nodes and edges in a virtual three-dimensional space. It intends to encourage critical thinking and active involvement of subject matter in language learning. The scarce nature of a three-dimensional graph should prompt students to realize that knowledge is not finite and that concepts can be related in unexpected ways. This, in terms of the free-energy principle, may trigger the student to experience surprise and engage in deliberation
	Constraints	Can be used as an alternative type of notebook regardless of imposed constraints
Chat	**Potentials (aspects)**	P6, P8–P10 (materials, class)
	Hurdles to consider	H6
	Basis and function	Promotes communication among peers and between students and instructors by allowing them to chat with each other. To encourage productive discussion, users can start chats directly related to items in their Knowledge graph
	Constraints	Can be used regardless of imposed constraints
Chatbot	**Potentials (aspects)**	P1–P7, P9 (assessment, materials, class)
	Hurdles to consider	H2–H4, H6, H7
	Basis and function	The Chatbot lives inside the Chat. Its purpose is to guide students to new subjects and materials to explore and to connect students to peers with similar interests. To make informed suggestions, the Chatbot consults the user's Knowledge graph items and their conversation history with the Chatbot
	Current constraints	Bot's material suggestions may be limited to some set of materials allowed by an instructor
	Relaxed constraints	In addition, the Chatbot may be allowed to suggest a wider range of materials it finds on the internet

	No constraints	In addition, the Chatbot may generate materials
PDF viewer	**Potentials (aspects)**	P1, P3, P7 (materials)
	Hurdles to consider	H3
	Basis and function	Enables users to view PDF documents from within the application. Documents can be added as nodes to the Knowledge graph
	Current constraints	Documents may be uploaded only by instructors
	Relaxed constraints	Documents may additionally be uploaded by students, freely or from some curated pool
	No constraints	In addition, the Chatbot may inspect, generate, or suggest documents
Token highlighter	**Potentials (aspects)**	P6, P9 (materials)
	Hurdles to consider	H6
	Basis and function	Service capable of highlighting strings of text that correspond to Knowledge graph node names. This could, for instance, be applied to the Chat or PDF viewer to continuously remind the user of previously handled topics as they chat with other users or peruse materials
	Constraints	Can be used regardless of imposed constraints
Challenges	**Potentials (aspects)**	P5–P10 (courses, assessment, materials)
	Hurdles to consider	H4, H5
	Basis and function	Encourages students to engage in productive collaboration by presenting a set of challenges relevant to the student that they can work on with peers. Students can set and track learning goals per Challenge and see what goals peers have pursued
	Current constraints	Challenges may be strictly instructor-curated and connected to formal assessment
	Relaxed constraints	In addition, students may suggest Challenges and make a request for a form and timing of assessment
	No constraints	In addition, the Chatbot may suggest Challenges and students may voluntarily choose to schedule an assessment
Schedule	**Potentials (aspects)**	P6, P10 (time)
	Hurdles to consider	H2, H4
	Basis and function	Intended to help keep study systematic by providing a personal calendar. May be used to schedule collaborative working sessions based on Challenges, assessments, and other study-related or personal activities
	Current constraints	May be restricted to instructor input, serving as a type of academic timetable
	Relaxed constraints	In addition, students may add their own activities and schedule group items
	No constraints	In addition, the Chatbot may suggest items in line with the user's activity

A key design consideration in the features in Table 5.5 is that they should be easy to use. For users with limited digital literacy, the features are intended to be straightforward to understand, either by replicating real-world phenomena (e.g., the Schedule is simply a digital calendar) or relying on physically informed visuals (the Knowledge graph). Additionally, all features can be used regardless of imposed constraints, although some features' functionalities are affected by constraint choices. Simple interactions and resilience against design changes due to curricular changes should allow instructors and students to utilize the software without having to overcome a steep learning curve (H7).

Further contributing to this, the features are highly integrated. The Chatbot may consult a user's Knowledge graph, Challenge, and Schedule activity to provide relevant suggestions. Chats can be initiated from Challenges or Knowledge graph items. Schedule items can link to Challenges. Lastly, the Token highlighter refers to Knowledge graph items from Chats and the PDF viewer. An intuitive integration of all functionalities is intended to prevent overwhelming users with many different, disconnected tools, as seems to currently be seen as a problem in practice (as evidenced by e.g. H7). In turn, this may help decrease distraction (H6). Figure 5.1 schematically presents how features are interconnected.

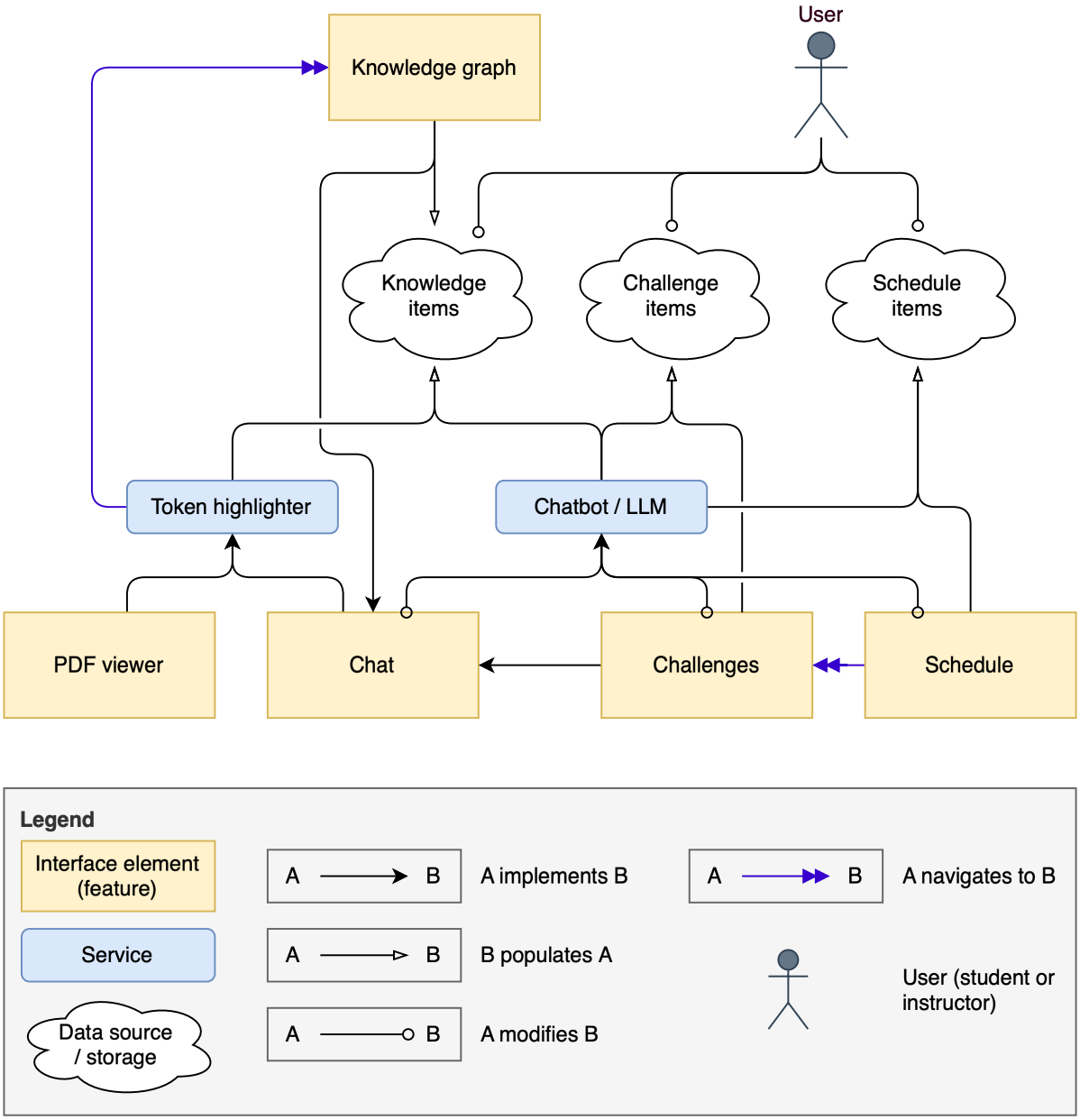

Figure 5.1: Schematic representation of the prototype system.

Finally, the features offer room to tweak the level of personalization. The Knowledge graph, for instance, could be set up so that students can only add items or documents from a curated pool, or so that they can add anything. Challenges could wholly coincide with curricular objectives, or they could be student-initiated. This allows instructors to exercise quality control if required (H3) as well as to keep control over how freely students may study with respect to the curriculum (H4).

It should be noted that each of the identified potentials and hurdles, and thus the features, follow specifically from observations and findings in the context of language education. The features offer room for specific adjustments with respect to a target language. For instance, in this thesis' case of Japanese, the Chatbot could be made to refer to reliable Japanese-language sources. Such implementation choices are explored in the next section.

5.2. Prototype Implementation

5.2.1. Implementation Possibilities for the Japanese-Language Case

How can the designed features cater specifically to an L2 Japanese study context? Table 3.8 and Section 3.4 presented considerations for L2 Japanese instruction. Considering the application itself does not intend to structurally steer instructors or students to (not) using certain types of materials or assessment methods, some of the identified considerations—J1, J3, J5, J6, J7, J8, and J9—have no effect on the implementation. These aspects can be instructed through materials that instructors choose to employ (or students, depending on the level of personalization) and do not seem to require implementation-specific consideration. The remaining items—J2, J4, and J10—relate to assessment. Each of these items can already be addressed through Challenges. For instance, instructors could differentiate Challenge topics and goals according to students' (dis)abilities and L1s (J2, J10) and students could indicate their own goals by creating Challenges themselves (J4). Additionally, none of the features are conceptually bound to inputs in any single language: if instructors and students find it more appropriate to perform activities in the application in Japanese, they could. The identified points may thus not directly influence the implementation, but they may influence how the implementation is used.

What other considerations could be made in relation to Japanese-language study? Two studies handled in Section 3.3, [41, 119], suggest a need to make interfaces available in various languages. However, in these instances, this need arose because users of some technology had a different L1 than offered in the interface, or because it did not offer tasks in a target foreign language. In the case of the present prototype, this does not appear to be an issue since users can add content in any language and eventual participants of the evaluation study can all be expected to be able to read English, making an English-language interface sufficient in that regard. The handled studies moreover do not suggest that having students interact with an interface in a target L2 (Japanese, in this case) particularly assists learning or study.

A final point to consider in light of users' ability to add PDF documents and to chat with the Chatbot is the quality of self-uploaded or suggested materials. Here, it may help to identify reliable sources for users and the Chatbot to refer to. Any number of sources could be employed in practice, but considering the limited duration of the evaluation connected to this prototype and the number of features to be implemented, one source is used in the implementation: the 青空文庫 *Aozora Bunko*.[44] The Aozora Bunko is used by the Chatbot to suggest materials to users who inquire about Japanese-language questions, as is further explained in Section 5.2.2.2.[45]

5.2.2. Prototype Implementation Choices

The prototype is named *Kamo*, which is Japanese for 'perhaps' and refers to the unbounded and indefinite form of knowledge in the Knowledge graph.[46] The purpose of Kamo—the reference point in the evaluation study in Chapter 6—is to serve as a proxy for a more personalized Japanese-language program, or as a tool that offers students a glimpse into such a program. Since a personalized language program may deviate significantly from the present program, a more comprehensively implemented prototype may elicit more relevant insights. To this end, all features in Table 5.5 are functionally implemented. For the features with different options in relation to the level of constraints, implementation leans toward the situation with fewer constraints so as to present evaluation participants with a scenario that differs enough from the present educational setting to form a distinct opinion.[47]

5.2.2.1. Server Hardware and Software

To ensure evaluation participants can easily access the prototype, it is served as a web application. The application front end and back end are hosted on a server running Ubuntu 24.04.4 LTS on an Intel Xeon

[44] www.aozora.gr.jp. Aozora Bunko is an internet archive with 17 333 copyright-free Japanese-language literary works of various levels as of June 2026, mostly written by L1 Japanese authors.

[45] To implement this functionality, the *aozora_corpus* assembled by GitHub user ryancahildebrandt was used (github.com/ryancahildebrandt/aozora_corpus; www.kaggle.com/datasets/ryancahildebrandt/azbcorpus).

[46] It also means 'duck', an animal much beloved by the author.

[47] To facilitate implementation of all features in a relatively short time, LLMs Claude Sonnet 4.6 and GPT-5.2-Codex were used through GitHub Copilot in Visual Studio Code. The LLMs were exclusively used for code implementation based on the author's designs. The author takes full responsibility for the functionality, ethicality, and safety of the code.

Platinum 8358P CPU with 4GB RAM. The front end is served as static HTML, CSS, and JavaScript files to ease implementation. The back end is built in Express.js. User data is stored in a PostgreSQL database on the same server.

5.2.2.2. Implemented Options Per Feature

For each implemented feature in Table 5.5, important implementation choices are discussed below, along with considerations to combat potential hurdles or navigate constraint limitations.

Knowledge graph. To visualize the Knowledge graph, the *3D Force-Directed Graph* JavaScript package[48] was used for four reasons. First, the package features a three-dimensional graph representation in line with the Knowledge graph feature design. Second, since the graph is force-directed, nodes are automatically spatially balanced based on their proximity to other nodes and the presence or absence of edges between them. Groups of related knowledge thus form automatically. This should provide users with insight in which concepts they have worked out elaborately and which not, and it alleviates the need to manually place nodes in space, combating H3 and H7. Third, the visualization is rendered in *WebGL*, facilitating implementation and fast execution within a web page. Fourth, the package comes with a large number of customization options and accompanying examples, making it straightforward to implement interactivity between itself and users and between itself and other features.

To further ease navigation of the Knowledge graph, especially when the number of nodes increases, some helper elements were added. One is the *Explorer* from which users can search, filter, and select nodes. Another is the option to group nodes in *Containers* that can be given different colors. To highlight chains of thought or concepts in the Knowledge graph, users can furthermore create *Pathways* with series of connected nodes that light up when selected. Finally, there are multiple options to create new nodes and edges such as a quick-add field, a more elaborate input card, and a node-specific option menu to directly link new nodes to a selected node (see Figure 5.2). The Knowledge graph window is displayed in Figure 5.3.

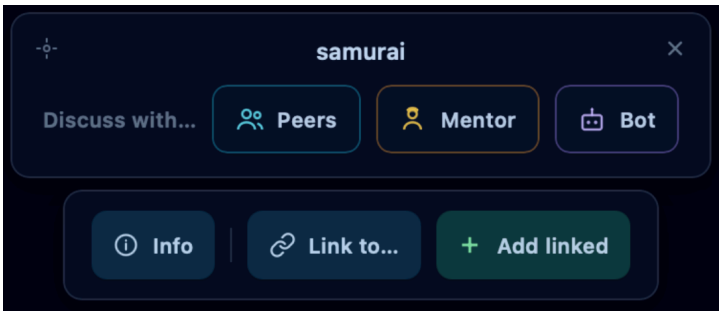

Figure 5.2: Options for a selected node in the Knowledge graph.

Chat. The Chat makes use of PostgeSQL's `LISTEN` and `NOTIFY` triggers to exchange text messages between users asynchronously. To keep the application lightweight, the Chat only supports text messages and not e.g. voice messages or file sharing. In a chat function with complete freedom to exchange messages, distraction from study (H6) may not be completely avoidable. However, study-related communication is encouraged by having users start Chats directly from a Knowledge graph node (see Figure 5.2) or a Challenge. The name of such a node or Challenge is then used as the Chat's title that is always explicitly visible (see Figure 5.4).

To find users to chat with, clicking the 'Peers' and 'Mentor' buttons (Figure 5.2) opens a list of all enrolled classmates and instructors, respectively. Other users with Knowledge graph items with a similar name to the node that the user is selecting are presented as recommended users on the top of these lists, allowing users to find others that potentially have similar interests. For instance, in the case of Figure 5.2, if another user exists that has a Knowledge graph item called 'samurai' or something similar, that user would appear at the top of the Peers or Mentor list.

[48] github.com/vasturiano/3d-force-graph, by GitHub user Vasco Asturiano, MIT license.

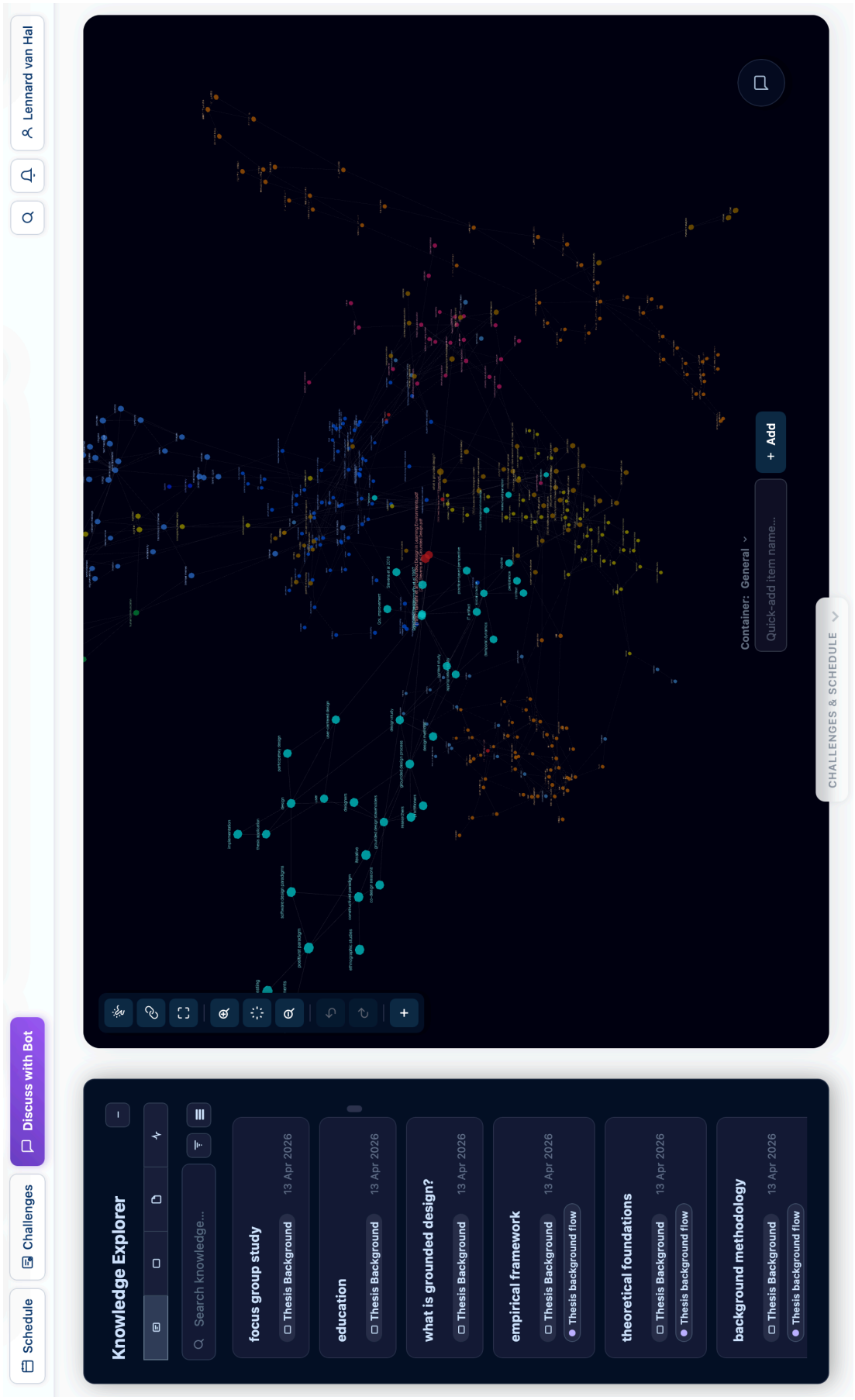

Figure 5.3: Page with the Knowledge graph, presented directly after login.

Chatbot. The Chatbot was implemented using LLMs running locally on a secure server at TU Delft. The LLMs run in stateless inference mode only, i.e., they have static weights, do not learn from inputs, and have no memory. For chat message generation, *gpt-oss-20b* is used; *Qwen3-Coder-30B-A3B-Instruct-FP8* is used for structured generation tasks like generating Challenge or Schedule suggestions.

Users have no API access and can only interact with the LLMs indirectly through the front end. The Chatbot can be chatted with by opening a new Chat from a Knowledge graph node ('Bot' in Figure 5.2) or from the 'Discuss with Bot' button (navigation bar in Figure 5.3). To have the Chatbot reply with relevant information, a portion of its chat history and a user's Knowledge graph items, Challenges, and Schedule items are passed to it as context in every prompt. Users can also use the Chatbot's LLMs through a 'Suggest proposal' button on the Challenges page and 'Propose week schedule' button on the Schedule page, pre-filling a Challenge proposal (Figure 5.6) or week schedule (Figure 5.7), respectively.

Two hurdles that may significantly impact effective use of the Chatbot are H2 (overreliance) and H6 (distraction). To prevent this, the Chatbot is given the instruction to act as a student coach that should encourage students to think independently, opting to ask students questions for reflection rather than giving them answers. In addition, depending on the user's inquiry, the Chatbot may suggest a user to chat with some peer who may have knowledge of the chat topic, inferred from peers' Knowledge graph items. In case the user asks about some Japanese phrase, it may also suggest a Japanese-language literary work from Aozora Bunko (see Section 5.2.1 and Figure 5.5). Works are recommended deterministically that contain the prompted phrase and that match the user's proficiency level, estimated from their previous interactions with the Chatbot and Knowledge graph on the basis of JLPT[49] levels.

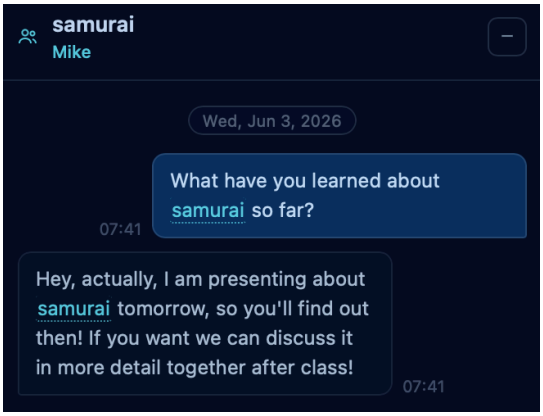

Figure 5.4: Chat between two students (fictional) on the topic of samurai.

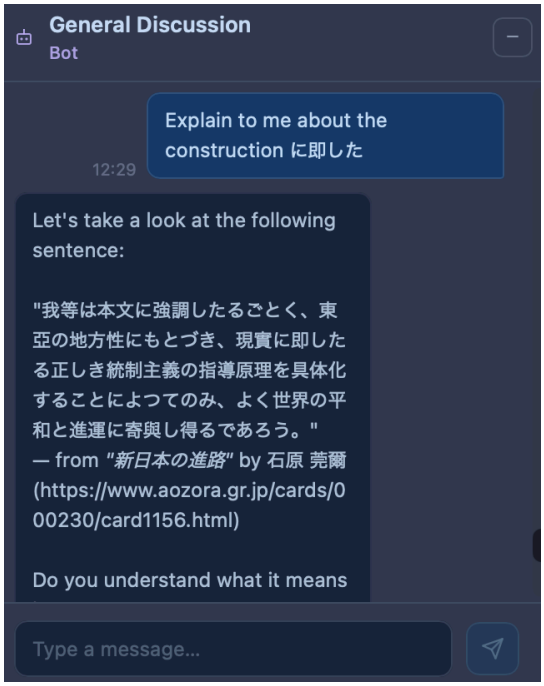

Figure 5.5: Aozora Bunko recommendation from the Chatbot.

Although the Chatbot may still generate answers without referring to peers or literary works, these occasional targeted recommendations should help students collaborate and get them in touch with high-quality works (H3). As for H4 and H7, instructors may want more direct control over sources and users may require some training to understand how to prompt effectively, which is unaccounted for in the present implementation.

[49] Japanese-Language Proficiency Test, www.jlpt.jp. JLPT levels range from N5 ("easy") to N1 ("difficult") and are based on knowledge of certain sets of *kanji*, vocabulary, and grammar rules, among other factors.

PDF viewer. The PDF viewer is facilitated on the front end through the *PDF.js* package[50] which renders PDF documents. On the server back end, user-uploaded documents are stored plainly as files. To ensure all users can upload files equally, a limit is imposed on the maximum size of documents and number of uploadable documents per user which depends on the number of participants in the evaluation study. Documents are limited to the PDF format so that the implementation need not be complicated with additional packages to render or convert to other formats. In the present implementation, users can upload any PDF document within these limits. Depending on whether instructors expect students to use the PDF viewer in some way, H3 may thus be an issue, although this issue is not specific to the prototype since students could already refer to their own selection of documents for study outside of the application.

Token highlighter. The Token highlighter detects tokens in PDF documents and Chats that correspond to names of a user's Knowledge graph items. For texts that contain Japanese characters, the *kuromoji.js* package[51] is used for tokenization; otherwise, tokens default to uninterrupted strings of alphanumeric characters. Highlighted tokens can be clicked to navigate to the corresponding Knowledge graph item in the Explorer and in the graph itself. Figure 5.4 shows an example of token highlighting (observe that the word 'samurai' is underscored).

The Token highlighter serves to help students stay in touch with their studies while interacting with others and materials so as to reduce distraction and foster thought. At the same time, it is imaginable that users are distracted (H6) if they see too many highlights (e.g., because they added a large number of Knowledge graph items). In the context of L2 Japanese study, this may especially be a problem in case students add many individual *kanji* to their Knowledge graph. To prevent excessive highlighting, users are hence given the option to see highlights only for tokens consisting of two or more *kanji*.

Challenges. In a scenario with relaxed constraints, instructors may want to allow students to fulfill (parts of) a course's assessment through completion of Challenges. However, there is a risk of H5 if all students are required to perform assessment in this way, especially if Challenges are curated exclusively by instructors. On the other hand, if students are given full autonomy to create Challenges in this situation, instructors may fear a lack of quality (H4). To meet both ends in the present implementation, students can autonomously propose Challenges, but proposals need to be approved by an instructor before they can be started. Furthermore, once a Challenge with more than two participants ends, an instructor may choose to schedule an interview with the participants for discussion or assessment. Students who perform a Challenge individually can end it immediately; this way, Challenges also serve as a tool to maintain and chase personal goals rather than to facilitate assessment. Figure 5.6 shows a proposed Challenge awaiting approval by an instructor.

Schedule. The Schedule is implemented such that students can freely add items and import calendars in ICS format. Users can also have schedule items dynamically proposed by the LLMs behind the Chatbot based on their existing Schedule items, Knowledge graph items, and active Challenges. Suggestions are offered in two forms: a full-week proposal and individual item proposals. To alleviate H2, generated proposals are limited to weekdays between 9 a.m. and 5 p.m. and respect a midday break (Figure 5.7). Further to this, users can have a maximum of five individual item suggestions per week. This number is lowered if users opt for a lower guidance level as in Figure 5.8.

The Schedule is also used when an instructor schedules an evaluation appointment with students following the completion of a Challenge. Instructors cannot see students' individual plannings, but do get a suggestion for a suitable date and time at which all students in a Challenge group are available. Apart from these options, H2 and H4 were not further addressed in the Schedule implementation.

Lastly, Knowledge graph items, Challenges, and Schedule items are integrated in a single application-wide search function Figure 5.9.

[50] github.com/mozilla/pdf.js, by Mozilla, Apache-2.0 license.
[51] github.com/takuyaa/kuromoji.js, by GitHub user Takuya Asano, Apache-2.0 license.

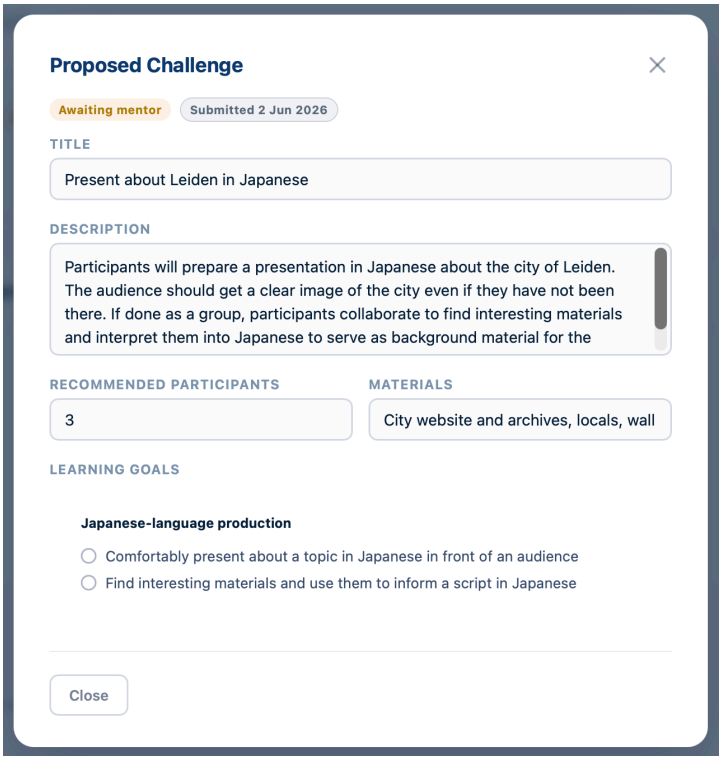

Figure 5.6: Challenge proposal awaiting approval by an instructor.

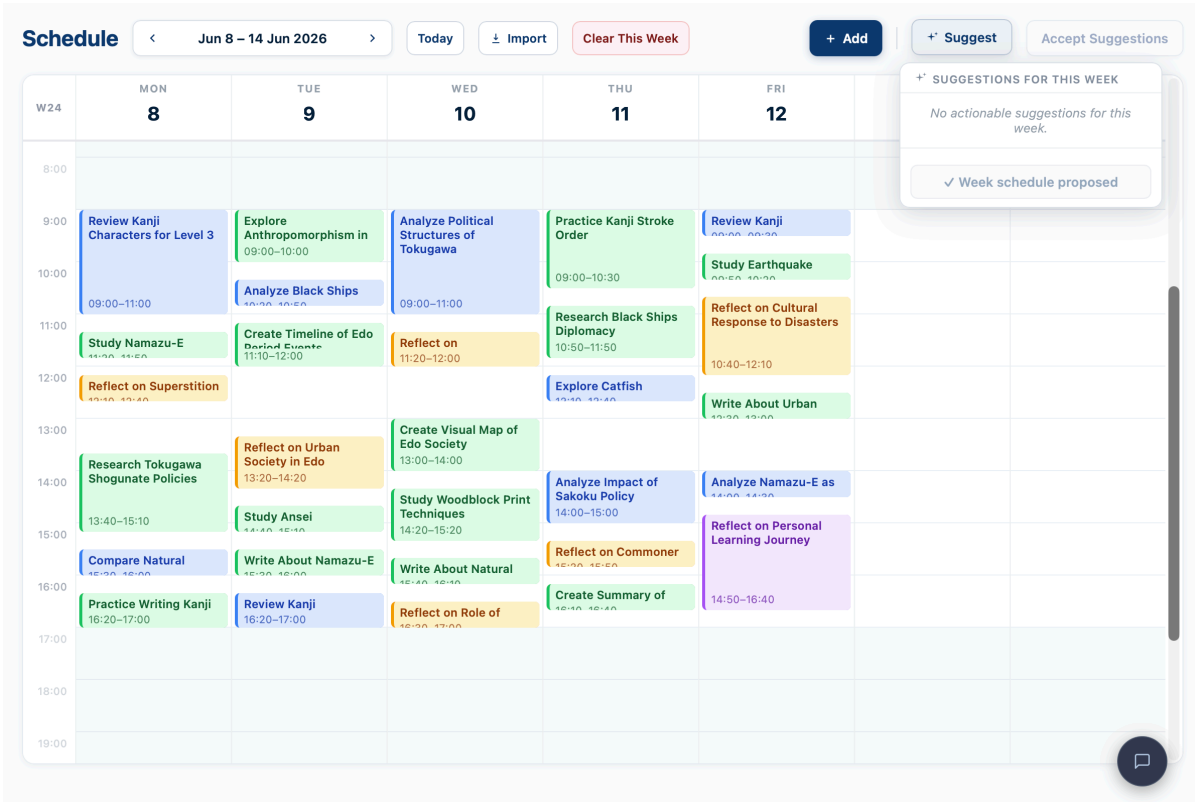

Figure 5.7: A Schedule filled for the week.

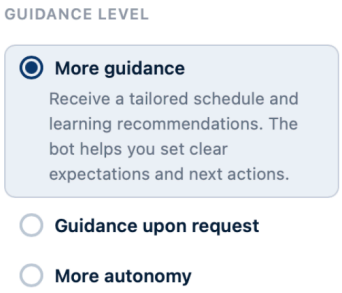

Figure 5.8: Guidance level options. A lower guidance level shows LLM recommendations less visibly and generates less of them. The Chatbot also adjusts its behavior according to the guidance level, where users with a lower guidance level are recommended peers or materials more frequently.

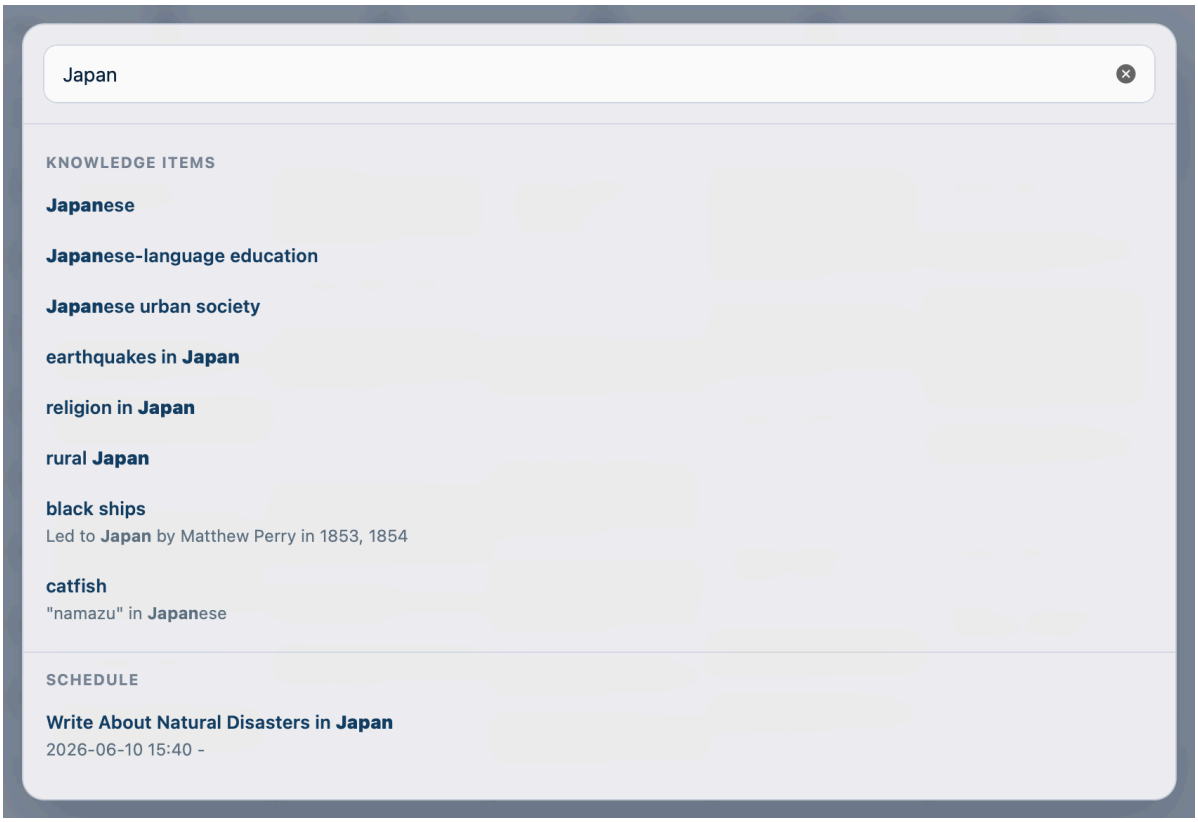

Figure 5.9: Application-wide search function.

6

Evaluation

To evaluate Kamo, the prototype presented in Chapter 5, students were given access to the prototype and participated in a study consisting of closed and open questionnaire items and semi-structured interviews. Kamo's main aim was to have students experience what a more personalized Japanese-language program could look like. In this evaluation, the central questions are whether Kamo succeeds in doing so and how students perceive the prototype and personalized study. This chapter discusses the evaluation process and results.

6.1. Method

6.1.1. Recruitment

Students were recruited from the first-year cohort of the Japan Studies bachelor's program and from the one-year East Asian Studies master's program with Japanese-language specialization at Leiden University. The author approached the students plenarily in person during their class hours in early March 2026 to ask for their voluntary participation in this study starting April 2026. Students were informed about the below study setup. n=47 students indicated an interest to participate by writing their contact details on an interest form or contacting the author directly. At the start of the study in early April, the students were contacted by email to ask for their confirmation that they wish to participate and that the author was permitted to set up an account for them on Kamo. n=12 students responded, of which n=4 were bachelor's students and n=8 were master's students.

6.1.2. Study Setup

Students made use of Kamo starting April 3, 2026 at the earliest. No conditions were attached to usage of Kamo other than that students were able to use it as a tool to support their language studies at their convenience. The author first organized online meetings of one hour each with participants, individually or in pairs, to explain Kamo and the study; all but one participants joined such a meeting. All participants were moreover sent a user manual with short explanations of all features. In addition to the student participants, one instructor of L2 Japanese was enrolled in Kamo so that students' Challenge proposals could be reviewed and so that students could try chatting with an instructor. The participants were encouraged to explore all implemented functionalities. The author made a secured, personal account for each participant. Participants were able to set their own password and choose a username during registration with the possibility to employ an alias or their real name. After usage of Kamo, participants were asked to voluntarily take part in a questionnaire study and a semi-structured interview starting April 27 and May 4, respectively, to inquire about their experiences with and thoughts on the prototype.

6.1.3. Questionnaire

The questionnaire was distributed through Microsoft Forms and consisted of five parts. The first was an informed consent form. Participants could only continue if they gave explicit consent to participate in the study and agreed with the data processing methods. The second part consisted of introductory

questions such as whether the participant joined an online instruction session with the author before using the prototype and how often they used the prototype.

The next part consisted of three sets of closed questions. The first set corresponded to the ten items in the *System Usability Scale* [73, Fig. 1] to measure participants' perceived usability of Kamo. The second consisted of the eleven items in the *Reward Factor* category of the long-form *User Engagement Scale* [93, p. 38] to measure the extent to which users found Kamo rewarding to use. The final set consisted of all items in the *Community of Inquiry Survey* [112, p. 28]. The latter is used to gain insights in the role of online or blended instruction environments in relation to the "interrelating elements" of "cognitive, teaching, and social presence" [112, p. 22]. Each of the employed surveys has been used widely in research, is shown to be reliable, and has benchmarks that can be used in analysis [73, 93, 112].

After the closed questions were six open questions (see Chapter C), of which the first intended to elicit deeper thoughts in correspondence with the Community of Inquiry Survey questions. The other five prompted participants to more elaborately share their experiences with the prototype and how Kamo situates itself within the Japanese-language programs at Leiden University, loosely in line with the questions discussed during the focus group study in Chapter 4. Finally, participants were able to indicate in the fifth part whether they were interested in participating in a subsequent semi-structured interview session with the author and potentially other participants.

6.1.4. Data Analysis

Since there were only seven responses to the questionnaire, statistical analysis on the closed questions was not performed. The open questions were analyzed thematically together with the semi-structured interview notes. This analysis is presented in Section 6.2.

6.1.5. Ethical Considerations

Prior to participation, the participants who filled out the questionnaire or attended an interview session submitted an informed consent form. Participants were made clear that none of their inputs to or usages of the prototype would be analyzed in any way or form and that only their responses in the questionnaire or interviews would be subject to analysis. Students were encouraged not to supply sensitive personal information in any of their answers. Any personally identifiable information in answers was anonymized prior to analysis. Interviews were not recorded; the author made notes during the interviews. The questionnaire answers were securely stored in TU Delft's Microsoft Forms environment. The study setup and used tools were approved by the TU Delft Ethics Committee.

6.2. Results

Of the twelve students who registered their account on Kamo, n=7 responded to the questionnaire. Since this number is not sufficient to draw meaningful conclusions based on a quantitative analysis of the closed questionnaire items, these items were disregarded. Of the seven student respondents, n=6 (hereafter referred to as S1–S6) provided answers to the open questions that are valid for analysis. Two participants, S3 and S5, were in the bachelor's program and the remaining four in the master's program. The participants in the bachelor's program indicated having used Kamo once after registration while S2, S4, and S6 had used Kamo two to four times and S1 had used it ten or more times. All participants had joined the author in an online explanation of Kamo before commencing usage. S1, S2, and S4 additionally partook in a semi-structured interview. Interviews were scheduled according to participants' availabilities, resulting in one session with S1 and one session with S2 and S4.

The answers to the questionnaire's open questions and the notes from the semi-structured interviews were thematically analyzed. Two themes were interpreted in the analyzed data (see Table 6.1). The first relates to participants' experienced and envisioned contributions of software in language study contexts. The second addresses experienced problems or gaps concerning software in language education. The themes are visited separately below, followed by a general discussion. Any citation is taken as-is from the answers to the open questionnaire items, which the participants provided in English.

Table 6.1: Themes, codes, and quotes supporting the codes extracted from the questionnaire's open questions and the semi-structured interview. Quotes come from the questionnaire answers and may contain language errors.

Theme	Code	Supporting quote
(1) (Potential) contributions of software to language study	(a) Connecting students with similar interests	Software could provide a "way for students with similar goals to connect" (S1)
	(b) Knowledge consolidation and language acquisition	"Kamo is great to get theory inside your head" (S3)
	(c) Dynamic materials and exercises	Software that provides "[e]ndless practice tasks of lesson content" is desired (S5)
(2) Current problems and gaps concerning (the absence of) software in language education	(d) No attachment to curriculum and classroom instruction	"[T]here currently is not a lot of software that is actually used by the university for teaching a foreign language in the first place" (S6)
	(e) No guidance in tool selection	"I feel that students are often left on their own to figure what [software] works best for them" (S5)
	(f) Limited possibilities for collaboration and communication	"[T]he programes used in education such as silverpoint, blackboard, brightspace all have their drawbacks and provide little to no social aspects or community tabs" (S2)

6.2.1. Theme 1: Contributions of Software to Language Study

Participants explained how software supports (or future software could support) their language study. Software's potential to encourage communication and collaboration between students with similar interests seemed to resonate (a). Currently, "the fostering of relationships is left to the students themselves," which can be "quite jarring" (S2). To this end, S1 imagined that Kamo's Challenges concept could work well to connect students with similar interests and to present students with more of a challenge complementary to what is offered in class. Similarly, S4 desired a "[p]latform to communicate with classmates for group projects or general questions" and S5 saw "potential in connecting more people together and creating online study groups." Such tools could lower the bar to connect by "quite a bit while still allowing for students to remain independent" (S2). Still, S1 and S6, who had tried out Kamo's Chat together and thought the concept of linking students through shared interests would be interesting, felt messaging someone out of the blue to be somewhat intrusive. S1 especially felt a hurdle to chat with the participating instructor for this reason. To lower such hurdles, S1 and S2 implied that a class-wide chatroom could help.

As for software that would be especially helpful in language study scenarios (b), flashcards and spaced-repetition tools were frequently mentioned. It was remarked that Kamo lacked such features and that it would benefit from integrating them. With Kamo's Knowledge graph alone, S1 was not sure how Japanese characters could be learned effectively; they had tried entering many characters as separate items, but this cluttered the view. "Having different tabs, for instance per country/language," and being able to differ nodes' sizes in the Knowledge graph could help alleviate this problem (S1). At the same time, S1 experienced Kamo's Knowledge graph as innovative and found that it made them "reflect on how different strands of knowledge could be connected." By filling the graph entirely with Japanese terms, S1 thought it helped them acquire the language as well since, according to them, reconceptualizing knowledge in a foreign language is part of the acquisition process. S3 found the feature "great to get theory inside your head," naming grammar as a language-related example, and S6 thought it could "definitely be beneficial in visually showing connections between certain things." S2 and S4 explicitly indicated they would like to see similar mind map features in future software, where S4 imagined it would be fun even outside of class to share "conspiracy board" mind maps with friends. These experiences suggest that a feature like Kamo's Knowledge graph could be beneficial both for language study and to grasp theory by forming connections between concepts.

Lastly, participants indicated that software could help find materials and exercises for language study (c), potentially in the form of a "catalogue of example learning resources" (S1), "a testing platform like quizlet" offered by the university (S3), a recommended material list (S4), "[e]ndless practice tasks of lesson content" (S5), and through ways to enhance media like films and books to turn them into practice material (S6). It was suggested that such features could significantly assist language study, with

S4 emphasizing that "apps and software" with such features "have helped me the most in concretely continuing my language learning" as opposed to university-offered materials and classes. Further, S2 saw an opportunity in a tool that enables students to keep study documents organized "in a way that fits the user better than the standard way provided by the university-delivered alternative [Brightspace]" such that students are "free to explore and customize their own approach [like] in Kamo." S5, too, saw potential for software to "adjust to each individual's need," especially since the "[c]urrent curriculum methods mainly force people into one stream with some side streams alongside it," which "surely does not fit everyone." What becomes clear from the participants' ideas and needs is that they would appreciate the freedom for *them* to select and use relevant materials: they are not simply asking for instructors or software to prescribe more mandatory materials, but for personally relevant recommendations or sources outside of what is currently offered by the program.

In summary, students saw potential in software features that foster connectivity between peers and that support vocabulary study (e.g., flashcard-like tools). A tool like Kamo's Knowledge graph was furthermore seen as especially beneficial to study theory, although some participants saw potential to use it directly for language study as well, for instance by building knowledge in Japanese. Finally, students are clearly looking for additional sources of relevant study materials outside of what is offered through the present curriculum, something the participants thought could be arranged through a software platform.

6.2.2. Theme 2: Problems and Gaps of Software in Language Education

All participants suggested that they—in fact, "most students" according to S4—use or have used software to study language in addition to the curriculum. However, the elaborate ideas they provided concerning useful software tools for language study (theme 1) are indeed mostly ideas and not *reality*. Despite the apparent importance and envisioned benefits of various software tools in students' language study, S6 remarks that "there currently is not a lot of software that is actually used by the university for teaching a foreign language in the first place," pointing at a gap between what is offered by the study program and what is needed (d). To fill part of this gap, "students are often left on their own" (S5) to find commercial language-study tools themselves (e) that they "sometimes pay for" (S4) despite three participants' explicit wishes for free software. While these tools may help vocabulary study (e.g., through flashcards), they do not cater to students' desire for software that promotes collaboration and communication with direct classmates since the tools are not an official part of the program (f). The university's own learning management system itself also provides "little to no social aspects or community tabs. Even when [it] is attempted [to make use of such functions] it is rarely used as this often applies to [students in] the whole course rather than one class" (S2). As such, importantly, in the participants' eyes, a unified platform that effectively combines useful language-study tools with classmate collaboration and that helps students navigate all the technological possibilities currently does not exist. They hinted that this gap could be filled by a single piece of software that "is offered in a package and for free" (S3); as a "platform" (S4) that "can guide students in the right direction [...] to help them find a learning method that suits them best" (S5) and allows students to communicate and collaborate.

As for whether such a platform could also be used during class time, opinions were mixed. With Kamo in mind, S1 and S5 were skeptic:

> [T]here is already too little classroom time in my opinion and using that time on this kind of software would be wasteful. (S1)

> I'm not sure if it useful within the classroom. I feel the current in-class lessons already offer enough explanation. I'm not sure how I imagine more software being integrated in it beneficially. [...] I feel like these software tools should stick mainly to one's own learning time. (S5)

Other participants suggested Kamo could be used during and outside of class, for instance "to really grasps certain theory of for example grammar" (S3), to show "connections between certain things" (S6), or for in-class collaboration if it additionally provided certain tools "such as presentation slides, kahoot kind of class interactive games," et cetera (S4). Finally, S2 suggested that software like Kamo could be a welcome addition inside and outside of class even if it does not suit everyone:

> For those who dislike Brightspace or similar programs they are free to explore and customize their own approach in Kamo, those who cannot be bothered or do not care for this do not HAVE to do this and so it only enriches the experience. (S2)

The above insights from S1, S3, S5, and S6 hint that the Knowledge graph, and not the communicative functions, was experienced as Kamo's most memorable feature, as emphasized by S1 and S4 during their interviews. Consequently, participants seemingly found it difficult to relate activity on Kamo to real-world peer connectivity—S1 was the only participant to explicitly mention a software potential to connect students "face-to-face." S6, for instance, thought that "in-person discussions would likely be more efficient" than online communication, but did not indicate that software could help facilitate in-person discussions. Similarly, S4 and S5 seemed to direct their ideas at online affordances like an online "[p]latform to communicate" and "online study groups." The limited number of users and interactions on Kamo may have contributed to this, as suggested by S4 who wrote "I did not really interact with other through Kamo so my responses [concerning interaction] are based on what I anticipate it to be like." In addition, as for the way Kamo's Chat was implemented such that it would suggest other users with similar Knowledge graph items, S1 and S6 noted that they were not or rarely matched with others due to the inability of the system to match item names written in different languages or as synonyms. This, too, may have inhibited users from connecting in real life based on online interactions.

Regardless of the reason in the context of this study, though, the responses suggest that with the current supply of software, real-world and online study settings are seen as disparate, such that there is little to no social overlap between in-class and online study. They similarly suggest that there is no direct connection between course materials and the online tools students use, other than perhaps that students enter prescribed word lists into a flashcard tool to study—a tool they had to find and potentially pay for themselves, because they feel the university does not assist them therein. The participants thus identified gaps in existing software as well as in the provisioning of software through the curriculum.

6.3. Discussion

In this evaluation study, six participants tried Kamo and shared their experiences. In the discussion of the results above, both potentials and shortcomings of (the provisioning of) software in relation to university-level Japanese-language study surfaced. Software features that were deemed important for language study include flashcard-like practice tools, ways to find relevant materials and store them in one place, and communicative and collaborative tools to find and work with like-minded students. With some interaction improvements, a tool like Kamo's Knowledge graph was also thought to be helpful, particularly for theory-heavy content but also for language acquisition. A platform that could combine all these features in a university-provisioned package was imagined to contribute to solving issues such as the perceived difficulty to find or finance effective tools and to find peers with similar interests. While opinions concerning use of such a platform within class were mixed, all participants found that it could be used well in combination with the existing courses as long as it would not "replace current [instruction] methods" (S5).

What do the results imply for personalized language study? Most prominently, they highlight the importance of taking diverse student needs into account, both in software and in a curriculum. This is reflected in participants' desire to be able to engage with their language study through materials and tools that more closely match their individual interests. It is moreover supported by the fact that not all tools work equally well for all students, which became clear from the diversity in participants' expressed needs with respect to software for language study, but also from students' own recognition that the tools they like are not necessarily suitable for everyone. This underscores the downside of subjecting all students to the same tools and materials.

In connection with the theoretical and empirical backgrounds presented earlier in this thesis, this evaluation yields some noteworthy observations. First, out of the forty-seven students who initially indicated an interest to participate, twelve created an account on Kamo, and six valid responses remained for analysis. While data was not collected on the reasons for students not to participate, participants who did provide a response recurringly indicated that a lack of time to interact with Kamo prevented them from engaging with it as deeply as they would have liked, aptly worded by S3:

> I thought it would be nice to help out and use this [Kamo] as a tool but I just ended up not using it at all cause I was just to busy with university. (S3)

The observation that introducing some new technology without adjusting the curriculum to account for it time-wise is detrimental to its practical efficacy surfaced repeatedly in Section 3.3, which appeared

to be at play in this study as well. Second, it followed from the present evaluation—like many studies in Section 3.3 as well as Section 4.2.3—that the university's current provisioning of technology or technological guidance is deemed inadequate. That is unfortunate: participants highlighted that it is, in fact, the extracurricular technologies they use that importantly keep them motivated to study language. Third, it was implied by the participants that the current program does not provide sufficient means for students to find each other for collaborative study or to share interests, which resonates with the instructors' experiences in Chapter 4. Here, it was argued that technology could play a supportive role, for instance through tools like Kamo's Challenges. Finally, following the earlier observation above that student needs are diverse, it is made clear that no one-size-fits-all software solution exists, supporting the findings in, e.g., Sections 3.3.1.4 and 4.3. Even if all desired software tools are integrated in a single platform in tight collaboration with a curriculum, there may very well be students who prefer offline study methods, which was not examined in this study.

In the interpretation of the above findings, it should be noted that this study is limited in several ways. First, the small sample size and homogeneity of the participants inhibits the generalizability of the study. Second, due to the low number of users on Kamo and limited time available to interact with the prototype, participants expressed a difficulty to experience all of the software's features to their fullest. A larger study with more participants from more varied (academic) backgrounds and during a longer time frame may yield more generalizable results and deeper qualitative insights.

7

Discussion and Conclusion

This thesis has presented theoretical, empirical, and practical views on education and the integration of software in language education to personalize study. Specifically, university-level study of L2 Japanese was taken as a case in point. The main research question this thesis addresses is (from Chapter 1):

(RQ) How can software facilitate student-personalized study of Japanese as a foreign language in Dutch university education?

The findings in this thesis suggest that a degree of software-assisted personalization is possible within the present university system—under the right conditions. Answers to the four subquestions introduced in Chapters 3 and 4 form the basis for this observation:

(RQ1) How is human learning defined and how does education facilitate it?

(RQ2) What kinds of technologies have been integrated in language education, for what purposes, and with what consequences?

(RQ3) What instructional approaches have been taken to cater especially to institutional L2 Japanese study contexts?

(RQ4) How do university-level instructors of L2 Japanese in the Netherlands perceive the integration of technology for the purpose of personalization of L2 Japanese education?

This chapter answers these subquestions to provide a comprehensive image of the possible roles of technology in personalized education, the limits to educational technology and personalization, and the future of personalized language education, leading to the answer to **RQ**. Finally, concluding remarks are presented along with recommendations for future research.

7.1. Answers to the Subquestions

RQ1. Section 3.2 viewed learning from neurobiological, cognitive, and psychological angles. Neurobiologically, learning can be understood as any change in the brain's internal belief state that occurs as a result of the processing of sensory information or of engaging in deliberation. Cognitively, learning is thought to be layered into lower- and higher-complexity processes where the lower layers embody more primitive information and the higher layers highly connected information. Here, learning pertains to modification and creation of pieces of information or their connections in any of the layers. Psychologically, learning is understood to occur through observation, experience, and active replication of behaviors and is said to materialize if this has resulted in a change of behavior or capabilities. With regard to learning, education is seen (Sections 2.2 and 3.2.2) as a means to provide groups of people with a calculated environment wherein information is shared systematically through materials and curricula, curated by institution policies, politics, and industries along a plethora of instructional methods. The time-constricted nature of educational programs that surfaced in Section 2.2 and Chapters 3, 4 and 6 implies that they are generally not conducive to higher-layer cognitive learning, which is understood to require relatively more time and energy than lower-layer learning. They do enable students to 'learn psychologically' through observation and experience of social conduct.

RQ2. As demonstrated in Section 3.3, large varieties of language-educational software have been devised. An important point of study appears to be how technology can be used to motivate students, for instance through gamified classes or cooperative platforms. Another relates to giving students tools to reach more varied study materials from anywhere and at any time so as to encourage engagement and introduce flexibility with respect to students' study needs. Many of the discussed studies reported both positive and negative consequences. Among the positive effects, it stood out that students indeed seemed to appreciate materials and (collaborative) activities outside the textbook; it is, however, not always clear whether this was accomplished by the specific choice of technology or by simply deviating from exclusively textbook-based instruction. As for the negative consequences, many studies reported not being able to achieve effective use of technologies due to curricular constraints and a lack of support and adequate ICT infrastructures. As a result, envisioned benefits of software such as those mentioned above were difficult to materialize. Other important negative effects include overreliance, impaired critical thinking, and loss of human connection, which surfaced particularly in software driven by generative AI or with high immersion.

RQ3. The reviewed studies reveal aspects specific to the Japanese language that are argued to require special attention in L2 Japanese classes, summarized in Table 3.8. The studies suggest that existing textbook materials lack thorough, up-to-date explanations of important issues such as *kanji* writing, *keigo*, pronunciation and (dis)fluency, and how these relate to students' L1s. Increased student-personal attention is desired for each of these aspects, but is hindered by limited classroom time, gaps in research, and biases toward L2 students. Combating this is foremost a question of educational and instructional attitude and content and not necessarily of technological support.

RQ4. Three instructors of L2 Japanese in the Netherlands shared their perspectives which were analyzed in Chapter 4. Importantly, technology was seen to hold potential for personalization through the affordances of chatbots if they could function as a supportive student coach. As a coach, it would have to guide students, but not perform work *for* students, so as to enable them to discover their own interests while fostering critical engagement with study. Another opportunity was seen in using technology as a mediator to connect students to each other and create a positive sense of community, as well as to facilitate finding and using more study materials, for instance derived from media. If technology could fulfill these opportunities to support personalized language education, it was seen to be potentially beneficial. However, the instructors replicated the findings from the literature study in Section 3.3 that there is limited time and support to investigate and implement technologies, while technology can only work effectively if carefully planned with respect to the curriculum and vice versa. It was also argued (both in the literature review and the focus group study) that personalization and technology should not go too far to the extent that students become overly reliant on software, get out of touch with peers and instructors, or lose out on valuable learning opportunities that require active and critical participation. Finally, instructors suggested that implementers should remain vigilant that instructional quality is maintained, which importantly involves not substituting software for instructors.

7.2. Discussion

Returning to **RQ**, what does the above say about software-assisted personalization possibilities of university-level L2 Japanese study? Personalized study implies that a student's study is personally relevant to them (Section 2.4). Depending on the degree of personalization, this could mean anything from shaping study in complete accordance with a student's personal interests to offering a student's favorite coffee at the school café. The answers to **RQ2–4** suggest that some form of study personalization is desirable, whether to provide students with study tools that work best for them, to improve the connection between materials and students' personal situations, or to facilitate students' pursuit of personal goals.

Educational study minimally involves courses and curricular objectives, assessments, some allocation of time to pass them, materials, instructors, and peers, and each of these factors can be personalized, granted the curriculum provides room to do so. The answers to **RQ2** and **RQ4** suggest that software can help here by dynamically providing students with personally relevant materials, getting students in touch with peers, and prompting students to reflect on their progress and to critically engage with materials. However, the integration of software in education also risks achieving the inverse: rendering

students overly reliant on technology, isolating them socially, and impairing critical thinking, among other effects. Achievement of software's positive rather than negative potentials can be contributed to through software designs that are thoroughly grounded in theory and practice, or through the curriculum by limiting the allowed affordances of software. It was the aim in Chapter 5 to combine the two by proposing software features that actively promote exploration, critical engagement, and peer connection while allowing for flexible implementation with respect to established curricular constraints (Table 5.5).

Nonetheless, language-educational practice demonstrates (Section 3.3, Chapters 4 and 6) that these constraints at present strongly inhibit software's efficacy. Importantly, they limit the time and technical guidance instructors and students are given to find and use effective software (e.g., Sections 3.3.2.2 and 6.2.2). Or, they force instructors and students to use certain technologies without relevant connection to course contents (e.g., Sections 3.3.1.3 and 4.2.2). The improved technical guidance and literacy training that instructors and students as a result desire may, as such, be worth extending to all actors who define the curriculum, including those in program management and policymakers.

In relation to L2 Japanese instruction specifically, the answer to **RQ3** suggests that software need not be the sole facilitator of personalization. Instead, it asks for more research into interplays between various L1s and L2 Japanese, updates of outdated study materials, and reflection on instructor attitudes. These issues require attention from researchers and instructors *before* implementing software that purports to solve them.

7.3. Conclusion: A Future of Personalized Language Education

What are we now to make of a future wherein technology supports personalized L2 Japanese education, or the answer to **RQ**? There is clearly not a lack of ideas from instructors and researchers concerning technology in language education. There is, though, a constrained playing field that constitutes the curriculum and the institution. It has been emphasized that there is little space for technology if it has not been accounted for in the curriculum, nor if there is inadequate support for instructors and students in terms of time investment, technical assistance, and technological literacy training. Accounting for these aspects appears to be an absolute necessity for any durable software-supported attempt at personalization to be effective. In light of personalization, it is, of course, also inherently of importance that students' needs and interests are taken into account, which includes not subjecting them to one-size-fits-all solutions.

Still, we live in a reality with these constraints and limitations. It appears that lifting them is a matter of policy and politics. This is typically out of direct reach of the technologist and the instructor. For now, if curricular constraints and personalization are seen to be in a competitive relation, this implies we need to settle for a smaller degree of personalization as long as the curriculum is unchanged. This seems possible and, in fact, a small degree of personalization can be highly impactful. Student perspectives discussed in Section 3.3 and Chapter 6 suggest that students are already helped if they were to be assisted in finding peers, study materials, and tools that match their interests and capabilities. These forms of assistance could be employed regardless of the target language of instruction, although the target language may influence instructors' attitudes and what materials, peers, and tools are recommended, as suggested in the case of L2 Japanese in this thesis.

In conclusion, to answer **RQ**: there are various ways for software to facilitate student-personalized study, but it needs to be done with considerable deliberation concerning students' needs and interests, potential side effects, and curricular flexibility. What makes this problem interesting is that students and their needs and goals are diverse. This is also what makes it complicated to devise some software or instructional mode that works for everyone, an observation this thesis was able to make by taking a grounded-design approach and looking beyond the immediate confines of the classroom, the language, and the software application.

In so doing, this thesis contributes to the fields of second-language instruction and computer science which appear increasingly intertwined with rising attention for technology-supported education. It has clarified what it means for study to be personalized, that present student-centered learning initiatives do not necessarily foster personalization, that interventions that aim to improve study do not naturally improve learning, and that interventions that aim to facilitate personalization do not necessarily do so for all students. Further to this, it has shown that implementing software in a language program is not

as 'simple' as building something and having instructors and students use it: the curriculum first needs to take the software into account, time-wise and content-wise. For second-language instruction, these clarifications provide an important lens through which existing instructional methods and educational software can be reexamined. They demonstrate that one-size-fits-all instructional and software solutions hardly exist and that there is, as such, truly a need to reconsider the individual student's relation to language learning within the university system. Some straightforward questions that follow from this are, for instance, to what extent a language curriculum should urge students to use flashcard tools to prepare for exams and how (if) it helps all students *learn* toward their goals. The clarifications also have important computer-scientific implications since (1) they provide an angle from which existing software designed for (personalized language) education can be critically examined, (2) they point out what factors outside of the immediate educational task affect the efficacy of software interventions, and (3) they thus help streamline directions for future research and software designs in the realm of language education. Section 7.4 suggests some of those future directions.

7.4. Limitations and Future Research

This research is limited in several ways. First, the qualitative studies in Chapters 4 and 6 are based on few participants that moreover have similar backgrounds, replicating the same issue as observed in Section 3.3. Additionally, the evaluation period in Chapter 6 was limited to within a month and participants had little time to thoroughly interact with Kamo alongside their studies, potentially influencing their questionnaire responses. Future studies may benefit from larger sample sizes, longer evaluation periods, and ensuring that a software prototype can be used by students as an official part of a curriculum so that they have enough time to interact with it. Second, the systematic literature reviews in Sections 3.3 and 3.4 discuss studies filtered exclusively by the author. Collaboration with additional evaluators and usage of different inclusion and exclusion criteria may yield different or more insightful results in future studies. Third, this thesis focused on personalization in L2 Japanese education without comparison with other languages or with education in other fields than language. A broader setup involving such elements could provide more elaborate insights into the applicability of personalization strategies through software. An implementation like Kamo's facilitates this already since its features are not tightly coupled with any one field of study. Finally, due to time constraints, the design in Chapter 5 could not be conducted iteratively, although evaluation in Chapter 6 did lead to directions for future improvements.

The issue of limited generalizability that affects this thesis itself and a large portion of the studies in Section 3.3 calls for larger research initiatives. Such initiatives would ideally involve diverse sets of students, instructors, and educational disciplines from different areas and institutions in clear, shared research purposes, combating the issues of small-size and homogeneous samples and of the uncoordinated development of 'software solutions'. Section 3.3.5 provided a preliminary suggestion for such an initiative that directly involves instructors and students in research and design processes. In the Netherlands, article 1.7a of the Higher Education and Scientific Research Act (see Section 2.2.4) may provide room to start such an initiative, for instance by granting instructors more time to design experiments or by granting students additional European Credits if they opt in for regular participation in educational experiments. Future research could investigate and act on such options to further our understanding of personalization and technology in language education.

Statement on the Use of Generative AI

Generative AI was not used by the author in relation to any content produced for this thesis project, except for partial implementation of the software prototype on the basis of the author's original design. The prototype itself features a chatbot with generative AI capabilities that participants of the evaluation study could make use of.

Bibliography

[1] Acar, A., and Kayaoglu, M. N. MOODLE as a Potential Tool for Language Education under the Shadow of COVID-19. *Eurasian Journal of Educational Research 20*, 90 (11 2020), 1–16.

[2] Ahlquist, E.-M. Learning in the Montessori School Environment. In *The Bloomsbury Handbook of Montessori Education*. Bloomsbury Publishing Plc, 2023, ch. 12, pp. 117–128.

[3] Ahmed, S. K., Mohammed, R. A., Nashwan, A. J., Ibrahim, R. H., Abdalla, A. Q., M. Ameen, B. M., and Khdhir, R. M. Using thematic analysis in qualitative research. *Journal of medicine, surgery, and public health 6* (2025), 100198.

[4] Al-Shehri, S. Transforming English Language Education in Saudi Arabia: Why Does Technology Matter? *International Journal of Emerging Technologies in Learning (iJET) 15*, 6 (3 2020), 108–123.

[5] Ali, A. A., AlQarni, K., Migdadi, H. F., Aldosemani, T. I., and Alshraah, S. M. Addressing the Challenge of Hybrid Learning Environment in Foreign Language Education: Training Lecturers for Blended Teaching Approaches. *Theory and Practice in Language Studies 14*, 5 (5 2024), 1582–1594.

[6] Aljabr, F. S., and Al-Ahdal, A. A. M. H. Ethical and pedagogical implications of AI in language education: An empirical study at Ha'il University. *Acta Psychologica 251* (11 2024), 104605.

[7] Altun, M. The Ongoing Debate over Teacher Centered Education and Student Centered Education. *International Journal of Social Sciences & Educational Studies 10*, 1 (1 2023), 106–110.

[8] Badjadi, N. E. I. Learner-Centered English Language Teaching: Premises, Practices, and Prospects. *IAFOR Journal of Education 8*, 1 (2020), 7–27.

[9] Bao, T. AR as an Educational Technology for the Development of Reading Skills in Children With ASD in Korean Language Education. *Journal of Computer Assisted Learning 41*, 4 (8 2025).

[10] Barrett, A., Pack, A., Guo, Y., and Wang, N. J. Technology acceptance model and multiuser virtual reality learning environments for Chinese language education. *Interactive Learning Environments 31*, 3 (4 2023), 1665–1682.

[11] Barrouillet, P. Theories of cognitive development: From Piaget to today. *Developmental Review 38* (2015), 1–12.

[12] Bartlett, A. J., and Clemens, J. Introduction: What is Education? A Polemical Question. In *What is Education?*, A. J. Bartlett and J. Clemens, Eds. Edinburgh University Press, Edinburgh, 2022, pp. 1–32.

[13] Berkant, H. G., and Derer, . K. Teachers'opinions regarding the effectiveness of DynEd language education system in english language teaching. *Turkish Online Journal of Educational Technology 2016*, December Special Issue (2016), 611–620.

[14] Brenner, W. H. *Wittgenstein's Philosophical Investigations*. SUNY series in philosophy. State University of New York Press, Albany, New York, 1999.

[15] Brookhart, S. M. *Classroom Assessment Essentials*, first edition ed. ASCD, Arlington, VA, 2024.

[16] Brown, J. S., Collins, A., and Duguid, P. Situated Cognition and the Culture of Learning. *Educational Researcher 18*, 1 (1989), 32–42.

[17] Broy, M., Selić, B., and Favaro, J. The Effects of Hype in the Software Domain: Causes, Consequences, and Mitigations. *IEEE software 42(2)*, 2 (2025), 98–102.

[18] Buseyne, S., Raes, A., Hurkmans, B., Leemans, M., Rajagopal, K., Eggermont, C., and Montero Perez, M. Scripting synchronous collaborative writing in second language education within a multi-shared visual workspace: a quasi-experimental study. *Pedagogische Studiën 102*, 1 (3 2025), 27–58.

[19] Chang, Y., and Lee, E. Addressing the challenges of online and blended STEM learning with grounded design. *Australasian Journal of Educational Technology 38*, 5 (2022), 163–179.

[20] Dekker, S., Lee, N. C., Howard-Jones, P., and Jolles, J. Neuromyths in Education: Prevalence and Predictors of Misconceptions among Teachers. *Frontiers in Psychology 3* (2012), 429.

[21] Derakhshan, A., and Ghiasvand, F. Is ChatGPT an evil or an angel for second language education and research? A phenomenographic study of research-active EFL teachers'perceptions. *International Journal of Applied Linguistics 34*, 4 (11 2024), 1246–1264.

[22] Diggelen, W. v., and Overdijk, M. Grounded design: Design patterns as the link between theory and practice. *Computers in Human Behavior 25*, 5 (2009), 1056–1066.

[23] Doherty, R. Critically Framing Education Policy: Foucault, Discourse and Governmentality. In *Why Foucault? New Directions in Educational Research*, M. A. Peters and T. A. Besley, Eds., vol. 292. Peter Lang, New York, 2007, ch. 13, pp. 193–204.

[24] Domínguez Romero, E., and Bobkina, J. Stance de-centring in Intercultural Virtual Societal Challenge (IVSC) e-modules for anti-racist foreign language education. *Innovations in Education and Teaching International* (9 2025), 1–16.

[25] Dong, W., Pan, D., and Kim, S. Exploring the integration of IoT and Generative AI in English language education: Smart tools for personalized learning experiences. *Journal of Computational Science 82* (10 2024), 102397.

[26] Downing, L. Life, texts, contexts. In *The Cambridge Introduction to Michel Foucault*, 1st ed., Cambridge introductions to literature. Cambridge University Press, Cambridge, 2008, ch. 1, pp. 1–21.

[27] Eerste Kamer der Staten-Generaal. Wet leeruitkomsten hoger onderwijs. https://www.eerstekamer.nl/wetsvoorstel/36136_wet_leeruitkomsten_hoger, 2024.

[28] Elder, L., and Paul, R. Critical Thinking: Strategies for Improving Student Learning. *Journal of developmental education 32*, 1 (2008), 32–33.

[29] Endo, H. Sociocultural Aspects of Teaching Japanese as a Foreign Language: Transforming Japanese Language Classrooms Into Small Japanese Societies. In *Language Across Disciplinary Boundaries*, M. Mantero, P. C. Miller, and J. L. Watzke, Eds. Information Age Publishing, Charlotte, NC, 2022, ch. 26, pp. 485–500.

[30] European Higher Education Area. Student-centred Learning. https://ehea.info/bologna-policy/student-centred-learning/, 2026.

[31] European Higher Education Area. The European Higher Education Area (EHEA): Brief introduction. https://ehea.info/about-ehea/, 2026.

[32] Fallace, T. The Long Origins of the Visual, Auditory, and Kinesthetic Learning Style Typology, 1921–2001. *History of Psychology 26*, 4 (2023), 334–354.

[33] Fiadeiro, J. L. Software Services: Scientific Challenge or Industrial Hype? In *Lecture Notes in Computer Science* (Berlin, Heidelberg, 2005), K. Araki and Z. Liu, Eds., vol. 3407, Springer, pp. 1–13.

[34] Forero-Álvarez, R., Triana-Perdomo, L. A., and Jiménez-Cuadros, L. K. Integrating Learning and Knowledge Technologies in Hybrid Classical Language Education: A Case Study from the University of La Sabana. *Forum for Linguistic Studies 7*, 1 (1 2025), 947 – 959.

[35] Friston, K. The free-energy principle: a unified brain theory? *Nature Reviews Neuroscience 11*, 2 (2010), 127–138.

[36] Gacs, A., Goertler, S., and Spasova, S. Planned online language education versus crisis-prompted online language teaching: Lessons for the future. *Foreign Language Annals 53*, 2 (6 2020), 380–392.

[37] Gani, S. H. N., Selvaraj, V., Islam, S., Thalapathy, S., Abdulkadhar, M. J., and Kalimuthu, K. Revolutionising English language education:empowering teachers with BERT-LSTM-driven pedagogical tools. *Humanities and Social Sciences Communications 12*, 1 (2025).

[38] Gao, Y., Tao, X., Wang, H., Gang, Z., and Lian, H. Artificial intelligence in language education: introduction of Readizy. *Journal of Ambient Intelligence and Humanized Computing* (3 2021).

[39] Garvey, G. Perspective Chapter: Ungrading, Grading Contracts, Gamification and Game-Based Learning. In *Active Learning: Research and Practice for STEAM and Social Sciences Education*, D. Ortega-Sánchez, Ed. IntechOpen, London, 2022, ch. 11.

[40] Geake, J. Neuromythologies in education. *Educational Research 50*, 2 (2008), 123–133.

[41] Goodier, T., Koslowski, M., Marcel, F., and Karamifar, B. Language Integration Through E-Portfolio (LITE): Building, Mediating and Reflecting on the Application of e-Portfolio Elements to Promote Plurilingual Language Education. In *Activating Linguistic and Cultural Diversity in the Language Classroom*, E. Piccardo, G. Lawrence, Germain-Rutherford Aline, and A. Galante, Eds., vol. 55. 2022, pp. 93–112.

[42] Guichon, N., Bédard, V., and Fortier, V. Digital literacy practices outside language classrooms: insights of adult migrants'language education. *Computer Assisted Language Learning* (3 2025), 1–22.

[43] Hannafin, M. J., Hannafin, K. M., Land, S. M., and Oliver, K. Grounded Practice and the Design of Constructivist Learning Environments. *Educational technology research and development 45*, 3 (1997), 101–117.

[44] Hasegawa, H. The effectiveness of online instruction with regard to acquiring script writing when learning Japanese as a second/foreign language. In *Digital Communication, Linguistic Diversity and Education*, S. Dovchin, Ed. Peter Lang, 2020, ch. 6, pp. 149–174.

[45] Hasegawa, H., and Nonis, K. P. Reflections from a teachers'perspective about the challenges faced by students with disabilities majoring in Japanese as a foreign language. *International Journal of Special Education 33*, 1 (2018), 119–128.

[46] Hashim, N. C., Majid, N. A. A., Arshad, H., Nizam, S. S. M., and Putra, H. M. Mobile augmented reality application for early Arabic language education-: ARabic. In *2017 8th International Conference on Information Technology (ICIT)* (5 2017), IEEE, pp. 761–766.

[47] Hatano, G., and Oura, Y. Commentary: Reconceptualizing School Learning Using Insight From Expertise Research. *Educational Researcher 32*, 8 (2003), 26–29.

[48] Hattie, J., and O'Leary, T. Learning Styles, Preferences, or Strategies? An Explanation for the Resurgence of Styles Across Many Meta-analyses. *Educational Psychology Review 37*, 2 (2025), 31.

[49] Hennink, M. M. *Focus group discussions*, 1st ed. Understanding Qualitative Research. Oxford University Press, New York, New York, 2014.

[50] Hinchey, P. H. Introduction. In *Democracy and Education by John Dewey: With a Critical Introduction by Patricia H. Hinchey*, P. H. Hinchey, Ed., 1st ed., vol. 1. Myers Education Press, United States, 2018, ch. 1, pp. xi–xxiv.

[51] Hsiao, S. Bard becomes Gemini: Try Ultra 1.0 and a new mobile app today. *The Keyword* (2 2024), https://tinyurl.com/2kkmpu9d.

[52] Hu, N. Application of Top-N Rule-based Optimal Recommendation System for Language Education Content based on Parallel Computing. *International Journal of Advanced Computer Science and Applications 14*, 6 (2023), 1027–1037.

[53] Hulstijn, J. H. Theoretical and Empirical Issues in the Study of Implicit and Explicit Second-Language Learning: Introduction. *Studies in Second Language Acquisition 27*, 2 (2005), 129–140.

[54] Ijiri, A., Mori, L., Galante, P. J., and Galante, P. M. Data-Driven Approach to Well Being Tracking in Language Education for Whole Person Learning. In *Lecture Notes in Computer Science*, vol. 15807 LNCS. 2025, pp. 257–270.

[55] Izquierdo, J., de-la Cruz-Villegas, V., Aquino-Zúñiga, S.-P., Sandoval-Caraveo, M.-d.-C., and García-Martínez, V. Teachers'use of ICTs in public language education: Evidence from second language Secondary-School classrooms. *Comunicar*, 50 (1 2017), 33–41.

[56] Jiang, X., Peters, R., Plonsky, L., and Pajak, B. The Effectiveness of Duolingo English Courses in Developing Reading and Listening Proficiency. *CALICO Journal 41*, 3 (2024), 249–272.

[57] Junya, Z., Yan, C., Junting, S., Xueqing, S., Moreira, P., and Cuiping, X. Exploring the Application of Teaching Models to Foster Creativity in Nursing Students: A Systematic Literature Review and Bayesian Network Meta-Analysis. *International Journal of Nursing Practice 31*, 6 (2025), e70078.

[58] Kanazawa, Y. The free energy principle and its implications to language learning and education: 4E cognition, prediction, accuracy-complexity trade-off, intrinsic motivation via epistemic emotions, 4 skills. 大阪大学大学院人文学研究科紀要 *[Memoirs of the Graduate School of Humanities, Osaka University] 1* (2024), 135–157.

[59] Kanuka, H. Independent Learning and the Essential Role of Feedback. In *New Science of Learning: Exploration in Mind, Brain, and Education*, I. M. Saleh and M. S. Khine, Eds. Koninklijke Brill NV, Leiden, The Netherlands, 2023, ch. 4, pp. 40–59.

[60] Karataş, F., Abedi, F. Y., Ozek Gunyel, F., Karadeniz, D., and Kuzgun, Y. Incorporating AI in foreign language education: An investigation into ChatGPT's effect on foreign language learners. *Education and Information Technologies 29*, 15 (10 2024), 19343–19366.

[61] Kereibayeva, A., Sharshova, R., Dyussenova, D., Aliyeva, D., and Toktarova, T. The Effectiveness of VR Glasses as a Tool for Immersive Language Education: insights from questionnaires and interviews. *Procedia Computer Science 272* (2025), 528–533.

[62] Khatri, P., Dabas, V., Thomas, A., Russo, G., and Kallmuenzer, A. Unearthing the Literature Review Dilemma: A Comparative of Typologies of Literature Reviews. *Business Ethics, the Environment & Responsibility 35* (2025), 391–414.

[63] Khosrawi-Rad, B., Keller, P. F., Benner, D., Grogorick, L., Borchers, A., Janson, A., Leimeister, J. M., and Robra-Bissantz, S. Promoting students'motivation in language education with gamified pedagogical conversational agents. *Computers and Education 238* (2025).

[64] Kim, H., and Hannafin, M. J. Grounded design of web-enhanced case-based activity. *Educational Technology Research and Development 56*, 2 (2008), 161–179.

[65] Košatka, D., Košatková, M., Fořtová, N., Šašinková, A., and Černý, M. Implementing Didactic Principles in VR for Language Education. In *Learning and Teaching in Virtual Reality*. Masaryk University Press, Brno, 2023, pp. 57–85.

[66] Krashen, S. We acquire vocabulary and spelling by reading: additional evidence for the input hypothesis. *The Modern Language Journal 73*, 4 (1989), 440–464.

[67] Krashen, S. D. The Input Hypothesis. In *The Input Hypothesis: Issues and Implications*. Longman, London, 1985, ch. 1, pp. 1–32.

[68] Lansdown, M. What is Learning? In *In the Zone: Helping Children Rise to the Challenge of Learning*. University of Buckingham Press, London, United Kingdom, 2019.

[69] Lantolf, J. P., and Xi, J. Let's not get tied into knots: a response to Newman, (2018)—Vygotsky, Wittgenstein and sociocultural theory. *Journal for the Theory of Social Behaviour 49*, 3 (2019), 387–393.

[70] Lazaroiu, G. Besley on Foucault's Discourse of Education. *Educational Philosophy and Theory 45*, 8 (2013), 821–832.

[71] Leask, I. Beyond Subjection: Notes on the later Foucault and education. *Educational Philosophy and Theory 44*, May 2012 (2012), 57–73.

[72] Lee, J. Problem-based gaming via an augmented reality mobile game and a printed game in foreign language education. *Education and Information Technologies 27*, 1 (1 2022), 743–771.

[73] Lewis, J. R. The System Usability Scale: Past, Present, and Future. *International Journal of Human-Computer Interaction 34*, 7 (2018), 577–590.

[74] Li, N., and Lau, K.-L. Factor Structure, Measurement Invariance, and Nomological Network of Teaching Presence in Online Foreign Language Education. *Psychological Reports* (11 2024).

[75] Li, P., and Lan, Y.-J. Digital Language Learning (DLL): Insights from Behavior, Cognition, and the Brain. *Bilingualism (Cambridge, England) 25*, 3 (2022), 361–378.

[76] Liao, Z. Personalized Path to Language Education Supported by Artificial Intelligence. In *Proceedings of the 2nd International Conference on Intelligent Education and Computer Technology* (New York, NY, USA, 6 2025), ACM, pp. 240–246.

[77] Liu, F., and Liu, X. Revolutionizing language education: How advanced technologies in Technology-Assisted Language Learning (TALL) are transforming teacher-student dynamics and elevating learning outcomes. *Porta Linguarum Revista Interuniversitaria de Didáctica de las Lenguas Extranjeras 2025*, XIII (9 2025), 205–224.

[78] Liu, T., and Zhang, Z. Language teachers as pedagogical designers in technology-mediated language education. *System 131* (7 2025), 103662.

[79] Liu, Y. The impact of bi/multilingualism on brain activity in dual language education based on smart technologies: Neurolinguistic aspect. *Education and Information Technologies 29*, 14 (10 2024), 18299–18321.

[80] Lozano-Rodríguez, A., García-Vázquez, F. I., and García-Cué, J. L. Perspective Chapter: The Role of Learning Styles in Active Learning. In *Active Learning: Research and Practice for STEAM and Social Sciences Education*, D. Ortega-Sánchez, Ed., Education and human development. IntechOpen, London, 2022, ch. 8.

[81] Lu, X., Zaitseva, A., and Troyan, F. J. AI literacy: A core practice in world language education. *Foreign Language Annals* (11 2025).

[82] Luo, Z., and Guo, F. Application and Optimization of Semantic-enriched Keyword Prefetching Driven by Intelligent Technology in Language Education Network Platforms. *Journal of Web Engineering 24*, 4 (7 2025), 635 – 654.

[83] Macedonia, M., Groher, I., and Roithmayr, F. Intelligent virtual agents as language trainers facilitate multilingualism. *Frontiers in Psychology 5*, APR (2014).

[84] Medvedeva, O. D. WebQuest as a Means of Students'Soft Skills Development in Engineering Foreign Language Education. In *The World of Games: Technologies for Experimenting, Thinking, Learning*, D. Bylieva and A. Nordmann, Eds., vol. 829 of *Lecture Notes in Networks and Systems*. Springer, Switzerland, 2023, pp. 323–334.

[85] Medvedeva, O. D., and Rubtsova, A. V. Productive Method as the Basis for Soft Skills Development in Engineering Foreign Language Education. *Education Sciences 11*, 6 (6 2021), 276.

[86] Meirbekov, A., Nyshanova, S., Meiirbekov, A., Kazykhankyzy, L., Burayeva, Z., and Abzhekenova, B. Digitisation of English language education: Instagram and TikTok

online educational blogs and courses vs. traditional academic education. How to increase student motivation? *Education and Information Technologies 29*, 11 (8 2024), 13635–13662.

[87] Michell, M., and Davison, C. 'Bringing the Teacher Back In': Toward L2 Assessment Praxis in English as an Additional Language Education. In *Toward a Reconceptualization of Second Language Classroom Assessment*, M. E. Poehner and O. Inbar-Lourie, Eds., vol. 41. Springer Cham, 2020, pp. 23–41.

[88] Mishra, P., and Koehler, M. J. Technological Pedagogical Content Knowledge: A Framework for Teacher Knowledge. *Teachers College Record 108*, 6 (2006), 1017–1054.

[89] Miyagi, Y., Hata, T., and Miyake, T. Fetal brain activity and the free energy principle. *Journal of Perinatal Medicine 51*, 7 (2023), 925–931.

[90] Nagata, N. Robo-Sensei's NLP-Based Error Detection and Feedback Generation. *CALICO Journal 26*, 3 (2009), 562–579.

[91] Newman, S., and Latifi, A. Vygotsky, education, and teacher education. *Journal of education for teaching : JET 47*, 1 (2021), 4–17.

[92] Novak, J. D. A Look to the Future. In *Helping People Learn*. Cambridge University Press, Cambridge, United Kingdom, 2022, ch. 5, pp. 114–156.

[93] O'Brien, H. L., Cairns, P., and Hall, M. A practical approach to measuring user engagement with the refined user engagement scale (UES) and new UES short form. *International Journal of Human-Computer Studies 112* (2018), 28–39.

[94] OpenAI. Introducing ChatGPT. https://openai.com/index/chatgpt/, 11 2022.

[95] Otto, S. E. K. From Past to Present: A Hundred Years of Technology for L2 Learning. In *The Handbook of Technology and Second Language Teaching and Learning*, C. A. Chapelle and S. Sauro, Eds., first ed. John Wiley & Sons, Incorporated, Hoboken, NJ, 2017, ch. 2, pp. 10–25.

[96] Paul, R., and Elder, L. Critical Thinking: Teaching Students How to Study and Learn (Part I). *Journal of developmental education 26*, 1 (2002), 36–37.

[97] Permana, P., Permatawati, I., Khoerudin, E., Muthmainah, N., and Oktapiani, W. Unveiling the user acceptance of Moodle in German language education. *Journal of Education and Learning (EduLearn) 19*, 3 (8 2025), 1687–1697.

[98] Piaget, J. Cognitive Development in Children: Piaget. Development and Learning. *Journal of Research in Science Teaching 40, supplement* (2003), S8–S18.

[99] Pitura, J., and Terlecka-Pacut, E. Action research on the application of technology assisted urban gaming in language education in a Polish upper-secondary school. *Computer Assisted Language Learning 31*, 7 (9 2018), 734–763.

[100] Plato. *Phaedrus*. Oxford world's classics. Oxford University Press, Oxford, 2002.

[101] Pring, R. John Dewey: Philosopher of Education for Our Time. In *John Dewey's Democracy and Education*. Cambridge University Press, 2017, pp. 340–348.

[102] Proietti, R., Parr, T., Tessari, A., Friston, K., and Pezzulo, G. Active inference and cognitive control: Balancing deliberation and habits through precision optimization. *Physics of Life Reviews 54* (2025), 27–51.

[103] Puylaert, R. Rijksbijdrage universiteiten. https://www.universiteitenvannederland.nl/onderwerpen/financien/rijksbijdrage-universiteiten, n.d.

[104] Qadir, S. M., and Omar, R. M. Digital literacy, technophobia, and technophilia: Exploring their roles in artificial intelligence integration into English language education. *E-Learning and Digital Media* (11 2025).

[105] Sada, M. B. Effects of Social Media on Study Efficiency and Academic Performance in Higher Education. In *Knowledge Management in Organisations*, L. Uden and I.-H. Ting, Eds., Com-

munications in Computer and Information Science. Springer Nature Switzerland, Cham, 2025, pp. 284–293.

[106] Sancar-Tokmak, H., and Yanpar-Yelken, T. Effects of creating digital stories on foreign language education pre-service teachers'TPACK self-confidence. *Educational Studies 41*, 4 (8 2015), 444–461.

[107] Savvidou, C. Introductory Chapter: Second Language Acquisition - Pedagogies, Practices and Perspectives. In *Second Language Acquisition - Pedagogies, Practices and Perspectives*, C. Savvidou, Ed. IntechOpen, London, United Kingdom, 7 2020, ch. 1, pp. 1–4.

[108] Senko, A., Khalyapina, L., and Shostak, E. Corpus Technologies for the Creation of Scenarios in Foreign Language Education. In *Lecture Notes in Networks and Systems*, vol. 1204 LNNS. 2024, pp. 71–83.

[109] Shaw, L. F. The Science of Learning Should Include the Study of Habits. In *New Science of Learning: Exploration in Mind, Brain, and Education*, I. M. Saleh and M. S. Khine, Eds. Koninklijke Brill NV, Leiden, The Netherlands, 2023, ch. 6, pp. 81–100.

[110] Shortt, M., Tilak, S., Kuznetcova, I., Martens, B., and Akinkuolie, B. Gamification in mobile-assisted language learning: a systematic review of Duolingo literature from public release of 2012 to early 2020. *Computer Assisted Language Learning 36*, 3 (2023), 517–554.

[111] Signorelli, C. M., Dündar-Coecke, S., Wang, V., and Coecke, B. Cognitive Structures of Space-Time. *Frontiers in Psychology 11* (2020), 527114.

[112] Stenbom, S. A systematic review of the Community of Inquiry survey. *The Internet and Higher Education 39* (2018), 22–32.

[113] Sternberg, R. J. What Is an "Expert Student?". *Educational Researcher 32*, 8 (2003), 5–9.

[114] Stevens, G., Rohde, M., Korn, M., and Wulf, V. Grounded Design. In *Socio-Informatics*, V. Wulf, V. Pipek, D. Randall, M. Rohde, K. Schmidt, and G. Stevens, Eds. Oxford University Press, Oxford, 2018, pp. 23–46.

[115] Sun, J., Song, T., Liu, H., Peng, W., and Song, J. Word-Constrained Response Generation for International Chinese Language Education based on Decoder Backward Attention. In *Proceedings of the 2023 3rd International Conference on Artificial Intelligence, Automation and Algorithms* (New York, NY, USA, 7 2023), ACM, pp. 186–190.

[116] Takeuchi, J. D. Language ideologies among Japanese foreign language teachers: Keigo and L2 speakers. *Foreign Language Annals 54*, 3 (2021), 589–606.

[117] Tomlinson, C. A. What Differentiated Instruction Is—and Isn't. In *How to Differentiate Instruction in Academically Diverse Classrooms*. Association for Supervision & Curriculum Development, United States, 2017.

[118] Trinidad, J. E. Understanding student-centred learning in higher education: students'and teachers'perceptions, challenges, and cognitive gaps. *Journal of Further and Higher Education 44*, 8 (2020).

[119] Tzirides, A. O. Informing Advanced Digital Technologies' Design for Refugee Language Education in Greece. *Proceedings of the 16th International Conference of the Learning Sciences - ICLS 2022* (2022), 1852–1853.

[120] UNESCO. *Global Education Monitoring Report 2023: Technology in education: A tool on whose terms?* UNESCO, Paris, 2023.

[121] Urueta, S. H. Challenges Facing the Adoption of VR for Language Education: Evaluating Dual-Frame System Design as a Possible Solution. *International Journal of Information and Education Technology 13*, 6 (2023), 1001–1008.

[122] Ushida, E. Robo-sensei personal Japanese tutor (version 1.0). *CALICO Journal 24*, 2 (2007).

[123] Vianna, E., and Stetsenko, A. Embracing History through Transforming It: Contrasting Piagetian versus Vygotskian (Activity) Theories of Learning and Development to Expand Constructivism within a Dialectical View of History. *Theory & Psychology 16*, 1 (2006), 81–108.

[124] Waks, L. J. Introduction to Part I. In *John Dewey's Democracy and Education*. Cambridge University Press, 2017, pp. 5–14.

[125] Waks, L. J., and English, A. *John Dewey's Democracy and Education: A Centennial Handbook*. Cambridge University Press, Cambridge, 2017.

[126] Wang, Q., Fan, M., Jin, X., and Wang, X. Research on the design and application of early childhood language education activities supported by information technology. In *2023 International Symposium on Educational Technology (ISET)* (7 2023), IEEE, pp. 148–152.

[127] Wang, X., and Xiong, L. Research on Teaching Strategies of Foreign Language Education Under Multicultural Background. *Eurasian Journal of Educational Research 2022*, 102 (2022), 253–277.

[128] Wang, Z. Redefining Language Education: How Integrating EFL Teacher-Student Relationships With AI-Powered Classrooms Impacts Emotional Health and EFL Learners' Depression. *European Journal of Education 60*, 4 (2025).

[129] Warren, T. Microsoft kills off Kinect, stops manufacturing it. *The Verge* (10 2017), https://tinyurl.com/ympztw2h.

[130] Wheelahan, L., and Moodie, G. Revisiting credentialism - why qualifications matter: a theoretical exploration. *British Journal of Sociology of Education 46*, 6 (2025), 874–892.

[131] Xhomara, N. Critical thinking: Student-centred teaching approach and personalised learning, as well as previous education achievements, contribute to critical thinking skills of students. *International Journal of Learning and Change 14*, 1 (2022).

[132] Yu, M. Application of an Artificial Intelligence-based Adaptive Learning System to Chinese Language Education in Universities. In *Proceedings of the 2024 International Symposium on Artificial Intelligence for Education* (New York, NY, USA, 9 2024), ACM, pp. 416–421.

[133] Yükseltürk, E., Altiok, S., and Başer, Z. Using game-based learning with kinect technology in Foreign Language education course. *Educational Technology and Society 21*, 3 (2018), 159–173.

[134] Yurdagül, C., and Öz, S. Attitude towards Mobile Learning in English Language Education. *Education Sciences 8*, 3 (9 2018), 142.

[135] Zhai, Y., and Nezakatgoo, B. Evaluating AI-Powered Applications for Enhancing Undergraduate Students'Metacognitive Strategies, Self-Determined Motivation, and Social Learning in English Language Education. *Scientific Reports 15*, 1 (2025), 35199.

[136] Zhang, D., Wang, M., and Wu, J. G. Design and Implementation of Augmented Reality for English Language Education. In *Augmented Reality in Education*, V. Geroimenko, Ed. Springer Cham, 2020, pp. 217–234.

[137] Zhang, Y., and Dong, C. Exploring the Digital Transformation of Generative AI-Assisted Foreign Language Education: A Socio-Technical Systems Perspective Based on Mixed-Methods. *Systems 12*, 11 (10 2024), 462.

[138] Zhang, Y., and Li, C. NLP-Driven Insights into the AI Divide: A Mixed-Methods Study of Chinese Language Education in Moroccan Higher Education. In *Proceedings of the 2025 2nd International Conference on Digital Society and Artificial Intelligence* (New York, NY, USA, 5 2025), ACM, pp. 889–893.

[139] Zhou, Y. Research and Application of Artificial Intelligence Assisted Language Education Technology Based on Big Data. In *Proceedings of the 2025 International Conference on Educational Technology and Artificial Intelligence* (New York, NY, USA, 7 2025), ACM, pp. 101–106.

[140] Zhu, K., Jiang, L., Zhang, D., Zhang, X., and Lu, K. Research on the Application of Deep Learning-Based Intelligent Universal Language Listening and Speaking Training Systems

in Minority Language Education. In *Proceedings of the 2024 3rd International Conference on Artificial Intelligence and Education* (New York, NY, USA, 11 2024), ACM, pp. 69–74.

[141] 中石ゆうこ [Yuko Nakaishi]. 第二言語としての日本語の自動詞・他動詞の双方向型教材の開発：形態的区別に着目して [Development of bidirectional instruction materials for transitive and intransitive verbs in Japanese as a second language: Focus on form differentiation]. 県立広島大学総合教育センター紀要 *[Prefectural University of Hiroshima Integrated Education Center Bulletin] 1* (2016), 69–74.

[142] 劉瑞利 [Ruili Liu]. L2 学習者の Multiword units 処理に関する研究の概観: 処理スピード及びその影響要因に着目して [A Review of the Researches on L2 Learners' Multiword Unit Processing: Focusing on the Processing Speed and the Factors Influencing the Speed]. 言語文化と日本語教育 *[Language Culture and Japanese-Language Education] 53* (2018), 11–20.

[143] 劉瑞利 [Ruili Liu]. L2 コロケーションの付随的学習に関する研究の概観 [A Review of the Research on the Incidental Learning of L2 Collocations]. 言語文化と日本語教育 *[Language Culture and Japanese-Language Education] 58* (2023), 1–10.

[144] 本橋美樹 [Miki Motohashi], and 石澤徹 [Toru Ishizawa]. 書字情報が持つ L2 音声習得への影響に関する一考察: 発音の生成に焦点を当てて [Considerations concerning the influence of written information on L2 speech acquisition: Focusing on pronunciation formation]. 関西外国語大学留学生別科日本語教育論集 *[Kansai Gaidai University Exchange Student Department Papers in Teaching Japanese as a Foreign Language] 26* (2016), 13–22.

[145] 畑佐由紀子 [Yukiko Hatasa], 高橋恵利子 [Eriko Takahashi], and 伊東克洋 [Katsuhiro Ito]. 第二言語としての日本語の音声習得と評価研究 [Acquisition and Assessment of L2 Pronunciation Skills in Japanese Learners]. 広島大学大学院教育学研究科紀要. 第二部, 文化教育開発関連領域 *[Bulletin of the Graduate School of Education, Hiroshima University. Part. II, Arts and science education]*, 65 (2016), 177–186.

[146] 郭跻 [Ji Guo]. 第二言語教育と非流暢性: 日本語会話教育への応用を視野に [Second Language Education and Disfluency: Implications for Japanese Conversation Teaching]. 総合文化学論輯 *[Journal of Comprehensive Cultural Studies] 23* (2025), 1–14.

A

Systematic Literature Review Appendix

A.1. Inclusion and Exclusion Criteria

The following inclusion and exclusion criteria were used to filter publications for the systematic literature reviews pertaining to **RQ2** and **RQ3** in Chapter 3.

Table A.1: Inclusion and exclusion criteria for **RQ1** and **RQ2**.

Topic	Inclusion criteria	Exclusion criteria
RQ2	(1) Should be aimed at the application of technology in *language education* for humans; (2) experiments should include interaction with some technology by human instructors or students	(1) Scope is limited to *virtual* technology (e.g., robots fall outside of the scope of this research); (2) publications are excluded that contain only a literature review and not some original design (of e.g. software, instruction methods, etc.), experiment, or original analysis of experiment data
RQ3	Should be about Japanese language-specific considerations concerning the instruction of Japanese as a foreign language	Studies that particularly concern the instruction of L2 Japanese to native users of languages other than Dutch or English are excluded

A.2. Search Queries

The following Scopus advanced search queries were used to filter publications for each of the topics subject to a systematic literature review.

General

Each topic's query is preceded by the following:

```
PUBYEAR > 2014 AND PUBYEAR < 2026
AND LANGUAGE(English OR Dutch OR Japanese)
AND DOCTYPE(ar OR ch OR cp) AND
```

RQ2

```
TITLE-ABS-KEY(experiment* OR trial OR empiric* OR design*)
AND TITLE("language education")
AND TITLE-ABS-KEY(technolog* OR software OR agent)
AND NOT TITLE-ABS-KEY(robot)
```

RQ3

The query used to find English-language literature on Scopus was the following:

```
TITLE(Japanese)
AND TITLE(second* OR foreign* OR L2)
AND TITLE("language")
AND TITLE(teach* OR instruct* OR pedagog*)
```

To find Japanese-language literature in the Leiden University Libraries Catalogue, the following query was used:

```
Title contains 日本語
AND Title contains 第二言語 OR 第2言語 OR 第２言語 OR L2
AND Title contains 教育 OR 教え
NOT Title contains 国語
```

Note that items with the term 国語 *kokugo* in the title were excluded so as to filter out publications concerning the instruction of (typically L1) Japanese within Japanese schools. Furthermore, the search was limited to *articles* with full access since other types of documents were not accessible. The same date range was applied as in the Scopus searches.

A.3. Result RQ2: Technology Experiments

Table A.2 presents a summarizing overview of the studies discussed in Section 3.3, category T.

Table A.2: Overview of studies selected for analysis to answer **RQ2** in category T. Technologies may be analog, digital, or a combination thereof.

Reference	Technology	Purpose (experiment setup)	Results or implications
[1]	Moodle[52] (learning management system)	Complementary tool in a blended learning L2 English high school class (n=42 participants, where half used Moodle over 15 weeks)	- Moodle potentially contributes to higher examination scores; student motivation with respect to foreign language learning was increased; - Students preferred using Moodle online in combination with face-to-face classes.
[9]	AR app *Kimchi Reader*[53] (digital Korean-English dictionary)	Support high school students of L2 Korean with autism spectrum disorder (n=70 students and one parent each; four-month intervention)	- Small but significant increase of academic performance in experiment group; - Students regarded Kimchi Reader positively and were more motivated; - AR applications can cause sensory overload and reduce social interaction.
[10]	VR environment in *Hubs by Mozilla*[54]	Support learning Chinese characters in a university-level L2 Chinese course (n=33 students performed three VR lessons)	- Immersion quality and facilitation of imagination influence perceived usefulness of VR; - Learning environment, hardware limitations, and students' proficiency in VR need to be considered before integrating VR lessons.
[13]	DynEd[55] (English language learning software)	Assess applicability of DynEd in an existing L2 English teaching curriculum in Turkey (n=24 primary, secondary, and high school teachers, each solo-interviewed)	- Schools' ICT infrastructures and technical support were deemed insufficient to incorporate DynEd; - Curriculum is too intensive to allow time for DynEd usage; - Teachers perceived little improvement of students' speaking skills after using DynEd.

[52] moodle.org (accessed February 17, 2026)
[53] kimchi-reader.app (accessed February 17, 2026)
[54] docs.hubsfoundation.org (accessed February 17, 2026; Hubs by Mozilla is obsolete since 2024)
[55] www.dyned.com (accessed February 17, 2026)

Table A.2 (continued): Overview of studies selected for analysis to answer **RQ2** in category T. Technologies may be analog, digital, or a combination thereof.

Reference	Technology	Purpose (experiment setup)	Results or implications
[18]	"Multi-shared visual workspaces" (students sit together with a large screen that shares each student's screen)	Improve group awareness and attention in university-level L2 French collaborative writing (CW) tasks (n=76 students divided in groups of four or five; 60-minute CW task)	- Thematic analysis of participant feedback raised usefulness of the shared screen for coordination as a theme; - Another theme was low perceived benefit of the shared screen because *Google Docs* was also used, which already live-shares users' contributions and positions within a document.
[21]	*ChatGPT*[56]	Identify "potentials and pitfalls of ChatGPT for L2 learning, teaching, assessment, and research" (solo interviews with n=30 L2 English teachers in Iran)	- ChatGPT provides potentials for personalized learning, student autonomy, and saving time when finding or creating study materials or assessments, among others; - Identified pitfalls are the discouragement of critical thought and creativity, loss of human interaction, and fraud in research and assessment, among others.
[24]	Mixed-media "e-module for anti-racist foreign language education"	"[S]caffold foreign language learners' critical interrogation of dominant ideologies and to cultivate intercultural, anti-racist dispositions" (n=17 self-selected university-level students in Europe, participating in "three 90-minute synchronous online sessions")	- The e-module facilitated critical reflection and nuanced discussion; - The "Stance De-Centring Model" upon which the e-module was built is argued to be effective in achieving the above.
[34]	Various (playful) hybrid classroom activities (online activities performed in-class)	Promote "engagement and participation in hybrid classical language courses" (hybrid activities were implemented in university-level Ancient Greek and Latin courses since 2020)	- Students perceived the lessons to be less monotonous and felt "reduced academic workload," while exam scores rose; - Gamification is seen to perform a "critical role […] in optimizing classroom time;" - Authors emphasize "the importance of a curriculum that is flexible, student-centered, and responsive to the specific challenges of hybrid education."
[41]	LITE[57] (digital portfolio)	"[S]upport language learning autonomy [and] language diversity within and, crucially, beyond the classroom" and facilitate community building (implemented in several Canadian schools)	- LITE "[enhances] online collaboration;" - Students aged 13 to 18 experienced difficulty in learning autonomously using LITE due to a "lack of guidance and […] feeling of confusion;" - Teachers and students were enthusiastic, but had limited time to use LITE due to curricular constraints.
[46]	*ARabic* (AR app for Arabic learning and teaching based on textbook)	Enhance interactive learning of L2 Arabic for 5-to-7-year-old children (n=16 children used the app; n=16 parents and teachers completed a questionnaire)	- Participants were positive with respect to use of ARabic and found it easy to use; - The app increased children's "curiosity and focus."

[56] chatgpt.com (accessed February 17, 2026)
[57] *Language Integration Through E-portfolio*; lite.lincdireproject.org (accessed February 17, 2026)

Table A.2 (continued): Overview of studies selected for analysis to answer **RQ2** in category T. Technologies may be analog, digital, or a combination thereof.

Reference	Technology	Purpose (experiment setup)	Results or implications
[54]	*Microsoft Reflect*[58] (emotion reporting and reflection tool)	Adapt classroom dynamics to students' real-time emotions (n=119 university students; Microsoft Reflect was used for three weeks of L2 English classes)	- Students found Microsoft Reflect "effective in increasing emotional literacy;" - Instructors were able to change the flow of a class based on students' reported feelings, although curricular time constraints complicated this; - Some students were seen to use the tool only superficially, and almost half were not sure whether the tool caused positive changes.
[60]	ChatGPT	Assess student perceptions on ChatGPT usage for L2 English education in Turkey (n=13 university students used ChatGPT for four weeks in L2 English class)	- ChatGPT was perceived to facilitate personalized learning, student motivation, and vocabulary learning; - Students expressed concerns about over-reliance and low effectiveness in training speaking and listening skills; - It is suggested that ChatGPT should only be used in class complementarily while "maintaining strong teacher-student relationships."
[61]	360-degree VR videos	Improve L2 English students' motivation (n=17 university students used a VR headset to watch immersive videos in English class for four weeks)	- Participants found VR lessons "to provide more memorable and impactful learning experiences" than "traditional [...] images and videos;" - Participants complained about discomfort wearing VR headsets and wished "for more opportunities to communicate with peers or native speakers in the virtual environment."
[63]	Pedagogical conversational agent (PCA)	Framework proposal for "gamified PCA design" (two-phase application design evaluation with university-level L2 English learners: (1) n=76 survey respondents; (2) n=70 survey respondents and focus groups with n=17 students)	- Provides a framework to design gamified PCAs for language education; - The evaluated gamified PCA appeared to motivate learners and boost their well-being, with "the PCA's role as an experienced classmate" being crucial to achieve this.
[65]	VR lessons based on curriculum	Improve L2 English learners' motivation and proactivity through VR (n=32 university students participated in three VR English lessons)	- Through VR avatars, learners put on a "mask of anonymity," which "[boosted] confidence, particularly among those fearful of speaking in foreign languages;" - Avatars hindered transmittance of non-verbal cues; - Virtual environments augmented traditional teaching methods by facilitating context.

[58] reflect.microsoft.com (accessed February 17, 2026)

Table A.2 (continued): Overview of studies selected for analysis to answer **RQ2** in category T. Technologies may be analog, digital, or a combination thereof.

Reference	Technology	Purpose (experiment setup)	Results or implications
[72]	AR mobile game and printed game	Comparison of effects of AR-based game versus printed game "on student engagement and attitudes toward foreign language learning" (n=42 university students of L2 English; one group played AR game, one group played paper version of AR game)	- No significant differences on effects between the groups, but students perceived paper-based game as more useful; - Authors argue that it is not about the use of new technologies, but innovative teaching strategies.
[76]	Deep learning-driven personalized learning paths	Facilitate language learners' "heterogeneous needs" in a system with "standardized teaching models" (12-week experiment with n=289 students in experimental group using personalized learning paths and n=292 using traditional learning path)	- Experiment group outperformed control group in terms of exam grades, task completion rates, and errors in assignments.
[79]	Language learning apps and MRI	"[E]xamine the influence of bilingualism on neural connections and brain activity in the context of education based on smart technologies" (n=60 L2 Chinese students followed traditional instruction, n=60 used language learning tools; effect on brain activity was measured)	- Experiment group had stronger, broader, more consistent connections in speech-related brain regions than control group.
[84]	WebQuest[59]	"[Develop] students' soft skills while foreign language learning" (n=26 university students following an English course completed a set of WebQuest tasks in groups)	- WebQuest promoted English-language communication between students through a chat functionality; - Students were motivated by points awarded to them for completing tasks; - WebQuest can be used to improve students' foreign language skills and other soft skills (such as teamwork and management) simultaneously.
[86]	Educational content on *Instagram*[60] and *TikTok*[61]	"[I]mprove students' performance and increase their motivation to learn" L2 English (n=187 university students spent the first portion of L2 English classes learning through Instagram and TikTok for four months)	- Students perceived learning English through social networks as positive because it increased interest and motivation compared with textbooks' "rigid learning content;" - Participants indicated interest in using TikTok and Instagram in their own time to learn English, but that they need teachers' guidance to be able to do this effectively.

[59] www.createwebquest.com (accessed June 15, 2026)
[60] about.instagram.com (accessed March 7, 2026)
[61] www.tiktok.com (accessed March 7, 2026)

Table A.2 (continued): Overview of studies selected for analysis to answer **RQ2** in category T. Technologies may be analog, digital, or a combination thereof.

Reference	Technology	Purpose (experiment setup)	Results or implications
[99]	Urban game	Evaluating feasibility of an urban game for language learning in upper-secondary school (n=25 students participated in urban game in Kraków, Poland)	- Student participants reported feeling less anxiety communicating in English; - Participants were able to self-learn usage of digital tools that they needed to complete post-game tasks, using tutorial videos in English; - Participants reported to have enjoyed the game and voluntarily chose to take part in it, despite it having been an extracurricular activity and students' high workloads.
[106]	Digital storytelling	Evaluate "the effects of creating digital stories (DSs) on the self-confidence of foreign language education (FLE) pre-service teachers with regard to [TPACK]" (pre-post TPACK survey with a digital storytelling intervention among n=71 pre-service teachers)	- Results indicate that the digital story creation process improved participants' TPACK self-confidence scores.
[108]	*Sketch Engine*[62] corpus manager	Improve L2 subject competence "using different scenarios created with the help of [Sketch Engine]" (n=20 students in an L2 German course in an experiment group used Sketch Engine to complete assignments; n=20 in a control group used a traditional textbook)	- Experiment group scored better than control group on "linguistic, subject, communicative, cognitive and intercultural" components, measured by evaluation criteria developed by the authors; - Students had trouble with grammar, differentiating between oral and written styles, and using a broad range of vocabulary when composing texts (unclear whether this holds for the control or experiment group or both).
[126]	Interactive virtual games on touch-enabled screen	Make use of kindergarten children's playfulness in language learning (n=20 children in experiment group played a game under teacher supervision to practice spoken language; n=20 in control group used traditional textbook)	- Children in the experiment group significantly outperformed children in the control group in the areas "initiative, goal awareness, imagination and creativity, curiosity and concentration," but not in "persistence, frustration resistance and independence;" - Children in the experiment group displayed more enthusiasm than those in the control group; - Authors warn that animations in games can distract children from the game activity or learning task.

[62] www.sketchengine.eu (accessed March 7, 2026)

Table A.2 (continued): Overview of studies selected for analysis to answer **RQ2** in category T. Technologies may be analog, digital, or a combination thereof.

Reference	Technology	Purpose (experiment setup)	Results or implications
[132]	AI system for personalized language learning	Adapt educational language learning paths to students' individual needs using AI (control-experiment group setup with n=300 university students; experiment group used *unspecified* AI system for one semester)	- "[T]he AI system is able to identify and meet students' individualized learning needs more effectively [than traditional educational methods]," with students in the experiment group achieving higher exam scores, higher recall and precision, and reduced cognitive load; - Performance test (CPU and memory usage, response time, and accuracy) on *unspecified* hardware indicates that the AI system is feasible up to a least a thousand concurrent users.
[133]	Motion-based *Kinect*[63] games	Develop Kinect-based learning activities and evaluate their effects on "learner's self-efficacy beliefs in and attitude toward learning English as a foreign language" (n=62 university students in English class, where n=32 in an experimental group played Kinect-based language learning games on several occasions during a semester)	- Experiment group had enhanced beliefs and attitudes toward learning English compared to control group; - Self-efficacy improvement in listening and speaking skills outperformed that in reading and writing skills; - Authors argue that "[c]areful pedagogic planning of game-based learning activities is essential" and can be used to "meet [students'] needs and interests."
[135]	ChatGPT, *Google Bard*[64] and *Poe*[65]	Explore AI applications' "potential to support autonomous learning" and effects on "metacognitive strategies, social strategies, and self-determined motivation" (n=310 university students, where n=139 "received AI-integrated instruction" to perform "language learning activities")	- AI tools were found to stimulate "*self-directed reading habits*" (emphasis in original) and motivation; - Practicing L2 English with AI tools helped "reduce communication anxiety and [encouraged] help-seeking behavior;" - Due to observations that students merely used AI as a shortcut, among others, the authors argue that curricula that make use of AI tools should be designed such that "AI tools are part of a cyclical process, encouraging students to reflect on their progress with each iteration."

A.4. Result RQ3

Tables A.3 and A.4 present an overview of the studies selected from Scopus and LULC for **RQ3** (Section 3.4). For each study, the table indicates what the study discusses and what its key findings are.

[63] Motion-based interactive system by Microsoft, discontinued in 2017 [129].

[64] Google Bard is *Gemini* (gemini.google/about; accessed March 8, 2026) since February 2024 [51].

[65] Poe (poe.com/about; accessed March 8, 2026) is an online platform that grants users online access to AI models from various providers.

Table A.3: Overview of studies filtered from Scopus search results for analysis to answer **RQ3**.

Reference	Focus	Key findings
[29]	Sociocultural aspects of Japanese (instructor's suggestions to improve L2 Japanese instruction)	Textbooks tend to pay insufficient attention to sociocultural aspects while their understanding is vital to one's ability to communicate in Japanese. Important aspects include politeness in speech, gender differences, nonverbal communication, and regional dialects. L2 Japanese instructors should present students with authentic materials (videos, music, books) to explicitly and implicitly demonstrate these sociocultural aspects.
[44]	Effectiveness of online study tools to study writing *kanji*	Due to constraints in available instruction hours, there is limited time for instructors to discuss the topic of writing *kanji*. This prompts students to use online tools to practice in their own time. One frequently recurring problem is that students who study in this way practice the wrong handwriting because they copy computer fonts that display *kanji* in a non-handwriting style. The author argues that online study should not replace face-to-face instruction and that such instruction should guide students according to their abilities to prevent such issues.
[45]	How classroom practice may be adjusted to accommodate students of L2 Japanese with a disability	Importance of the *kanji* script in fostering both receptive and productive skills is emphasized. To ensure students with disabilities hindering them from writing *kanji* can successfully participate in class, instructors are suggested to continuously discuss with students how this may affect their study of Japanese and their prospects after the study. The authors also call for consideration of curricular (e.g., assessment) changes to ensure students with a (writing) disability can obtain a degree and to provide them with alternative materials that strengthen other skills than writing.
[116]	Instructors' beliefs concerning *keigo*, usage and comprehension of *keigo* by L2 Japanese users, and how this affects classroom practice	Textbooks and instructors regard *keigo* principally as a means to express or enact hierarchical relationships, while *keigo* has been shown to have more dynamic uses. Instructors' belief that L2 users are not expected to master *keigo* as much as L1 users seems to inhibit instructors' willingness to overcome difficulties in instructing *keigo*. The author argues that L2 users should be seen as "legitimate speakers" and that they should be instructed as if they have the potential to master *keigo*.

Table A.4: Overview of studies filtered from LULC search results for analysis to answer **RQ3**.

Reference	Focus	Key findings
[141]	Study of transitive and intransitive verbs as an L2 Japanese student	Argues that transitive and intransitive verbs should be instructed in two stages: (1) identifying whether some word is transitive or intransitive; (2) judging whether in some scenario, a transitive verb or an intransitive verb is appropriate. Oftentimes, this distinction is not made, thus instruction of these verbs is mixed, hindering a solid grasp of the differences.
[142]	Multi-word unit (MWU) processing speed by L1 English students of L2 Japanese	(1) Students' L1 and (2) the frequency of MWU appearances impact processing, which could be taken into account by an instructor. For instance, the instructor could increase inputs with MWUs that occur infrequently to stimulate their acquisition.
[143]	Incidental learning of collocations in L2 Japanese	Surveyed studies all used outdated collocation assessment methods (e.g., fill in the blank, choose the right option, etc.) that regularly omit context. The author argues that other types of evaluations (e.g., with audio or moving pictures) are necessary, especially since English and Japanese have different orthographies.
[144]	Influence on L1 English users' L2 Japanese pronunciation of long- and short-vowel words based on the absence or presence of written *hiragana*	Study participants were better able to correctly pronounce long-vowel words (e.g., りょこう *ryokō*) if they were shown the written form in *hiragana* compared to students who relied on their own translation into *hiragana* from English. Especially beginning students' pronunciation may improve if both students and instructors are aware of the advantages of displaying written hiragana in pronunciation practice.
[145]	How various L1s may influence L2 Japanese pronunciation	Mainly, disfluency with respect to Japanese's mora-based rhythm stands out when L2 speakers pronounce Japanese. Since the mora rhythm appears relatively unique to Japanese, it should be taken into account relative to a student's L1 in language instruction.
[146]	Treating 'disfluency' (e.g., in terms of fillers or pauses during speech) as a part of language acquisition that should be explicitly instructed	Disfluency is not simply a sign of inability or low proficiency, but performs important functions even for L1 speakers: (1) transmitting feeling, (2) managing politeness, and (3) repairing speech. For L2 speakers, effectively managing 'practical disfluency' needs to be instructed to sound more natural. This especially holds for repair strategies and types of disfluency that are natural in Japanese, but not in one's L1.

B

Focus Group Study Appendix

This appendix contains the discussion guide that was used to lead the focus group discussion handled in Chapter 4 and a list of quotes from the discussion that support the codes in the resulting thematic analysis.

B.1. Discussion Guide

The below discussion guide was used in the focus group discussion during the pre-design stage of the application, handled in Chapter 4. The guide is written in Dutch, since that is the language in which the discussion was held.

Introductie Ik werk aan een scriptie over de integratie van software in Japans taalonderwijs. Het doel hiervan is om te ontdekken hoe we, zonder de druk op docenten verder op te voeren, het onderwijs persoonlijker kunnen maken: meer afgestemd op eigen doelen van de student, op individuele benodigdheden (bijvoorbeeld voortkomend uit leermoeilijkheden), en op voorkeuren ten opzichte van leren. Software introduceren in een programma met een bestaand curriculum, op een dusdanige manier dat het daadwerkelijk nut heeft voor de student en de docent, is niet eenvoudig: er moet onder andere rekening gehouden worden met de middelen en taken die studenten al hebben en de mate waarin studenten en docenten overweg kunnen met technologie. In deze discussiesessie wil ik graag bespreken hoe u zich voorstelt dat software in deze context past en hoe het kan bijdragen aan het verpersoonlijken van taalonderwijs.

Voordat we beginnen, wil ik enkele zaken benadrukken. Ten eerste ben ik uitsluitend als student betrokken aan de universiteit. Ik vervul geen adviserende rol en heb geen invloed op de toekomstige vormgeving van uw vakken. Ten tweede is uw deelname geheel vrijwillig, en kunt u zich op elk gewenst moment terugtrekken uit deze discussie en het onderzoek. Ten derde maakt deze discussiesessie onderdeel uit van een onderzoek, en zodoende zal wat er ter sprake komt later door mij geanalyseerd en verwerkt worden binnen de context van mijn scriptie. Alles wat hier besproken wordt, wordt echter geanonimiseerd vóórdat het gebruikt wordt in de analyse. Dat wil zeggen dat fragmenten die u persoonlijk zouden kunnen identificeren, zoals namen, woonplaatsen, en andere persoonlijke achtergronden, worden verwijderd voordat ze op enige wijze worden geanalyseerd of verwerkt in de scriptie.

Om ten behoeve van de analyse later terug te kunnen verwijzen naar onze discussie zou ik graag een audio-opname van deze sessie maken. Deze opname wordt uitsluitend lokaal opgeslagen op mijn computer, en door mij gebruikt om achteraf een tekstuele transcriptie te kunnen maken, waarna de opname wordt verwijderd. Gaat u ermee akkoord dat ik deze sessie opneem?

Tot slot: dit is een discussiesessie, dus iedereen kan vrijelijk op elkaar inhaken, elkaar aanvullen, vragen stellen, en meningen uiten, en ieders inbreng is even waardevol. Ik fungeer gedurende de sessie als moderator en leid het gesprek aan de hand van enkele vragen. Om ervoor te zorgen dat we de belangrijkste onderwerpen binnen de tijd kunnen behandelen, kan het voorkomen dat ik een gesprek tussentijds moet onderbreken om door te gaan naar een volgend onderwerp of om het

gesprek terug te leiden naar het te bespreken onderwerp, mocht het daarvan afgeweken zijn.

Openingsvragen (5 min.)

1. We beginnen met een korte zelfintroductie.
   (naam, vakgebied, hoofdinteresse binnen Japanstudies)

Introductievragen (10 min.) Over ervaringen met leren en doceren gebruikmakend van technologie.

2. **Gebruikt u zelf weleens bepaalde technologieën** ter verdere ontwikkeling van uw Japanse taalvaardigheid, en waarom?
   (hoe past u personalisatie toe; leer-apps op telefoon, games, media, onderzoekstools, online woordenboeken, etc.; welk effect hebben deze)

Hoofdvragen (35 min.) Over ideeën aangaande nieuwe technologieën voor het doceren en door studenten laten leren van Japans op de universiteit.

3. **Wat ontbreekt er** wat u betreft in het bestaande aanbod van technologie ten behoeve van het leren van de Japanse taal?
   (voor studenten/docenten; belang/nut/(on)handigheden van bestaande middelen, aansluiting op leermoeilijkheden/-sterktes, ondersteuning/persoonlijke begeleiding voor studenten; in welke Japanse taalverwervingsgebieden is technologie meest behulpzaam)
4. Als de universiteit aan iedere student Japanstudies een app zou aanbieden, wat zouden dan **behulpzame functionaliteiten** van zo'n app kunnen zijn?
   (bijv. in relatie tot lesstof, lesvoorbereiding, huiswerk, contact, persoonlijke doelen of moeilijkheden/sterktes, en de vraag hiervoor)
5. Uitgaande van een app die over voldoende kennis beschikt: Op wat voor manier ziet u in dat zo'n app **binnen het klaslokaal** ingezet zou kunnen worden?
   (gelimiteerd tot 'teksten', 'grammatica' etc.) (rol in relatie tot docent/student, leren in groepen/alleen; los onderdeel in les of continu gebruik tijdens les)

Ter afsluiting (10 min.) Samenvattende vragen.

6. Denkt u dat de **rol van docenten en studenten gaat veranderen** in een toekomst met meer technologieën zoals die we vandaag besproken hebben?
7. Zijn er overige zaken die u nog wilt bespreken?

B.1.1. Scenarios

The following scenarios were prepared to support questions 4 and 5 during the discussion in case instructors required clarification. Bob's scenario (Q4.2) was actually used. Names and circumstances depicted in the scenarios are fictional.

Q4.1 Anna (BA1) moet over twee weken een korte presentatie opnemen in het Japans voor conversatieles. Ze is angstig dat haar Japanse accent niet goed genoeg is. In de les krijgt ze weleens feedback over haar uitspraak, maar ze weet niet zo goed wat ze moet doen om haar uitspraak te verbeteren. De docent denkt graag mee en geeft regelmatig tips, maar heeft geen tijd om persoonlijke oefeningen of begeleiding te bieden, en Anna heeft er moeite mee om zelfstandig de tips van de docent in te passen in haar agenda en leerschema.

Q4.2 Bob (BA2) is erg geïnteresseerd in de Japanse politiek, maar heeft moeite met de taal. Vorige week kreeg hij zijn tussentoetsresultaten terug van Teksten 2a. Die vielen hem tegen. Bij de toetsinzage vandaag zijn de vragen besproken die minder goed gingen. Na afloop was Bob echter overweldigd. Hij weet niet goed hoe hij de volgende toets beter kan maken en is een hoop motivatie verloren.

Q5.1 Clara (BA2) is met drie medestudenten in een groepje ingedeeld om samen in de klas een Japans krantenartikel te vertalen. Er is geen formele rolverdeling afgesproken. Ze merkt echter dat er twee groepsgenoten zijn die angstig lijken te zijn om fouten te maken, waardoor deze niet goed weten wat bij te dragen. Ze doet daarom vooral het werk samen met de vierde groepsgenoot.

Q5.2 Dirk (BA2) doet in de conversatieles mee aan een rollenspel. Tussen de karakters in het rollenspel zitten (bewust) sociale statusverschillen, zodat de studenten kunnen oefenen wanneer ze welke beleefdheidsvormen dienen te gebruiken. Hij heeft echter nooit eerder gesprekken in het Japans gevoerd met moedertaalsprekers en is ook niet eerder in Japan geweest, en weet niet goed welke vorm wanneer gewenst is. Daarom houdt hij het tijdens de oefening op de standaard beleefdheidsvorm (*desu*, *masu*).

B.2. Quotes Supporting Codes and Themes

Table B.1 displays quotes from the focus group discussion session that support each code identified in the thematic analysis (see Table 4.1).

Table B.1: Quotes associated with each code. The quotes are from the discussion transcript and are translated from Dutch to English by the author.

Code	Supporting quote
(a) Digital technology to save time or achieve greater reach	"That's a tool [FeedbackFruits] that is offered by the university, that allows you to generate standard feedback yourself […]. I haven't tried it myself, but I've heard that, indeed, you can just, react to things in a standardized way […] to save time in that way."
(b) Coaching role	"[…] and then you return at the coaching function again, so that you on the one hand have technology that coaches and that tracks 'Where do you stand now, where, what is your next step' and on the other hand instructors with whom you [the student] are constantly in conversation in the abstract sense about where you stand right now."
(c) Community	"[…] you can use technology for that, but you don't have to, […] but […] connectedness is very important there, I think, for motivation when it comes to learning."
(d) Motivation	"[…] some degree of standardization seems to be taking place, that causes students to want more exercises, because 'I want to pass the exam,' and [they] are indeed not concerned with 'I want to know how it works, because I want to know that what I came here for, that's what I'm interested in.'"
(e) Counterproductive use of technology	"They [students] don't necessarily need to sit [in the classroom] to learn something from me, as long as they are there to learn something from each other. And that won't happen if they're all having a screen on their tables."
(f) Technology does not fit curriculum	"I think that for many of this kind of applications […] very often it comes down to 'Here we have a nice idea,' but how that fits within the curriculum or what one is actually going to do with it that reaches a learning goal, that seems to be lacking a bit."
(g) Counterproductive personalization	"[…] of course it's very nice if you can provide each student with feedback for everything they write, but they also need to learn to ask questions about 'Why wasn't this a pass.' Or learn that they should go to the Writing Lab."
(h) Purposeful avoidance of technology	"In some lectures this year, I've further reduced the PowerPoint presentations so as to try to have more of a discussion with a, with the students […]."
(i) Disabling system	"But, we don't allow them [students] [to decide for themselves if they attend lectures or not], because then they will have a year of study delay and we don't get any money."
(j) Lacking support	"If you would want to integrate such materials, that there are also web extensions, […] but […] how should one integrate such things in lectures […], there is a lot lacking there, let's say in guidance."
(k) Student-instructor relation	"I think, maybe, the advantage of a digital variant [of a mentor] is that it is somewhat more approachable, to the extent that it is not an eminent lecturer who you consequently need to address formally."

C

Evaluation Appendix

The below open questions were used in the evaluation questionnaire in Chapter 6. The first question was asked in the context of the closed Community of Inquiry Survey items.

1. Do you feel that your use of Kamo affected your responses to the above statements, or that the use of Kamo or similar software would contribute positively or negatively to any of your responses to the statements? How and why, or why not?
   Think about how (software like) Kamo could affect communication and discussion with your instructors and peers within the course context, and how it could affect how you handle or think about course contents.
2. Suppose the university grants each student in your study program access to some app, website, or other software that is supposed to help you study Japanese in a way that allows you to pursue your own interests and to use your own learning methods or materials, if you so wish.
   What would be **helpful functionalities** that you would like the software to have for that purpose?
   You may refer to functionalities that already exist in Kamo, but may also come up with your own ideas or extensions.
3. Do you think such software would **fill a gap in the university's current provisioning of software technologies** in relation to learning Japanese as a foreign language, or not? Why?
   For example, think about how it may or may not help you in learning Japanese in addition to existing technologies provided by the university, and whether some areas in language acquisition (e.g. practicing vocabulary, grammar, spoken or written dialog, etc.) could potentially benefit from the introduction of some new software technology (or not).
4. Do you think such software could be used well **in combination with existing Japanese language courses and the current curriculum** (i.e., like how you used Kamo in this experiment *alongside* your courses)? Why or why not?
5. Do you feel that such software might be beneficially used **within the classroom** alongside existing instructional methods? How and why, or why not?
6. Do you have any other comments, thoughts, ideas, or suggestions about Kamo, personalized language education, or this study?