Abstract

Personal User Models for Life-long, Life-wide Learning (PUML, Kay & Kummerfeld, 2019) are gaining increasing traction as individuals and government legislation seek greater control over their learning paths and personal data.

This thesis builds on these ideas by proposing a PUML for language learning that can serve as the personalised user/learner[1] model for a plurality of CALL/TELL tools for use in both formal and informal learning contexts, such as spaced-repetition systems for vocabulary, input-enhanced extensive reading for vocabulary and grammar development, text selection, learning games, auto-generated grammar tasks, and many more. By decoupling and sharing the learner model between many different CALL/TELL tools, tools can leverage the significant amounts of data generated over the lifetime of the learner, from early K12 into the workforce, and provide highly personalised learning over all the necessary skills as they develop, all the while leaving ownership and control over the data in the hands of the learners (or their legal guardians). By providing a means to take advantage of data generated in both in-class and out-of-class environments, learners would be empowered to take significantly greater control over their learning, all the while allowing much greater precision and personalisation for both IT and human teaching resources to utilise.

Previous systems (Chen & Meurers, 2019; SUNG, DYSON, CHEN, LIN, & CHANG, 2015; Yeung & Lee, 2018) for selecting appropriately levelled authentic texts for use in language learning (extensive reading, analysis in classrooms, etc.) have relied on expert levelled-datasets and test or teacher/class-based levelling for the learners (i.e, the texts in a year 10 English textbook). In order to validate the utility of the fully-personalised approach afforded by a PUML, a novel system for selecting texts based on the current lexico-semantic and grammatical knowledge of each individual learner, their content preferences, and the lexico-semantic and grammatical properties of the texts is proposed and tested with real CSL learners.

Hanyu Shuiping Kaoshi (Hanban, 2014) resources are used wherever possible in the study. The proposed system includes components for collecting and populating the PUML with data from the learners for meaning-level lexical entries and grammatical rules using a binary known/unknown scale for simplicity. The grammatical rules are taken from the set of rules defined for the HSK. Study participants' HSK levels are tested and assigned and are interviewed on their topics of interest and expertise, and documents are retrieved that correspond to their particular knowledge. These texts are then analysed to determine complexity according to a non-personalised system and those corresponding to an equal or higher level on the standard scale are kept. Standard, levelled texts are used as controls. The learners then read and are

1 The terms “learner model” and “user model” are used interchangeably in the text. Both terms refer to the same thing in the literature when applied to learning applications and contexts.

tested on their comprehension of both sets of texts and their perceived complexity. We compare results for the control and the personalised texts and see whether, as predicted, the learner finds the curated texts easier and achieves higher comprehension scores than the control texts of the same level. The results for specially curated texts at a higher level are also predicted to be considered easier than/as easy as texts at the test-assigned level.

A pilot study will be conducted to validate the core components with a small group of learners.

While the desired difficulty level for each individual learner can be set directly by the learners or teachers when selecting the texts, it is expected that real use of the system would focus on targeting the ZPD for the individual, as contrasted with previous systems working only with personalisation on one dimension (e.g., knows/doesn't know vocabulary items), or using expert-levelled systems that can't fully adapt to individuals' differing existing competencies.

Table of Contents

Introduction

Technology in SLA

Technology use in SLA has a long and productive history over many decades (Chun, 2016; Reinders & Stockwell, 2017), with technology use being strong both inside and outside the classroom (Lai, 2017). Education companies (McGraw Hill[2], etc.) and hundreds/thousands of desktop (Innovative Language[3], etc.), web (Living Language[4], Open Culture[5], etc.), and

2 http://www.mhlanguagelab.com/

3 https://www.innovativelanguage.com/products/pc

mobile/multi-platform (DuoLingo[6], Memrise[7], Rosetta Stone[8], etc.) applications exist specifically for second language learning, including both free/open-source (Openwords[9], etc) and proprietary technology and SAAS offerings by for-profit companies (DuoLingo, etc). Technology use has also been the focus of a great deal of research in the Applied Linguistics and Education fields, with regular articles in the most important journals, book series (Advances in Digital Language Learning and Teaching[10]) and even entire journals dedicated to the subject (Language Learning and Technology[11], Technology in Language Teaching and Learning[12], Journal of Foreign Language Education and Technology[13], etc.).

It is fair to say that technology is available for developing all the major language skills (reading/writing/listening/speaking, cultural knowledge, etc.), and considerable research has been done on the use of technology to develop these individual skills, as evidenced in the voluminous available literature. However, the market and literature are notably lacking in technology or research that integrates the learning of all of these aspects into a single, coherent system for use in both formal/structured (i.e, "schooling") and informal/unstructured environments (free extensive reading, watching video/film content, etc. for pleasure + learning). Evidence exists to suggest that learning each of the skills influences learning the other skills (Bozorgian, 2012; Krashen, 2004), and a system that allows learners to benefit from mechanisms that can track their progress in all the relevant domains could provide great benefit and scientific interest. With a great deal of the world's formal language learning going on in primary and secondary classrooms, and lasting for a decade or more, insights into the long-term role and capabilities of technology is critical. With the rapidly changing technological landscape, changing fashions, and apps and platforms coming and going, there is little research on observed (as opposed to reported) usage-level details of the impact of the use of these

---

4 https://www.livinglanguage.com/

5 http://www.openculture.com/freelanguagelessons

6 https://www.duolingo.com/

7 https://www.memrise.com/

8 https://www.rosettastone.com/

9 https://www.openwords.com/

10 https://www.bloomsbury.com/us/series/advances-in-digital-language-learning-and-teaching/

11 https://www.lltjournal.org/

12 https://journals.castledown-publishers.com/index.php/tltl/

13 http://jflet.com/jflet/

technologies over the timescales we currently observe for instruction and learning of foreign languages.

PUML for SLA

In order to provide the necessary continuity over time and learning contexts, as well as empowering learners to take a more active role in their learning, it is necessary to allow learners to share the data generated between all of the different digital contexts they interact with the L2 in. Any and every interaction with the L2 provides an opportunity for technology-based systems to enrich an evolving learner model, so long as these systems have an opportunity to share this information. In addition to standard protocols and APIs, a standard, extensible mechanism for representing competencies in and between the various skills and knowledge elements is necessary for applications to act upon.

A PUML (Kay & Kummerfeld, 2019) for language learning (see, for example, Al-Jadaa, Abu-Issa, Ghanem, & Hussein, 2017; Rudzewitz, Ziai, Nuxoll, De Kuthy, & Meurers, 2019), is a good candidate for fulfilling the role described above (Kay, personal communication) and provides the theoretical foundation for this work. One of the key challenges in a rapidly changing technological landscape is to enable new applications and technologies to integrate seamlessly, without requiring a learner to repopulate a new model with each new iteration and for each language skill. Evidence of knowledge of a given lexical or grammatical item gathered by an extensive reading application or Learning Management System (LMS) can clearly usefully inform applications dedicated to listening comprehension, or even a grammar-checker that could provide help in production settings.

By separating the learner model into a distinct, network-connected unit, many existing applications could be extended to benefit from it, and new applications can fully utilise the capabilities exposed, including those in emerging technologies such as AR, VR and ubiquitous computing (mobile and IOT/connected objects). Because the creation of a complete technology platform is a very significant engineering and research undertaking, the focus of this work is on the core PUML and certain technologies for extensive reading.

Reading in L2

Extensive reading, particularly when the learner has an element of freedom in the choice of subject matter, is widely recognised as having significant learning value for L2 learners (Grabe, 2008; Krashen, 1982, 2004; I. S. P. Nation, 2001; P. Nation, 2015). Reading, however, is also a complex process, and successful reading can be influenced by a number of factors. Extensively citing the literature, Jung (2009) provides an excellent overview of these, and includes orthography, vocabulary, grammar, background knowledge and metacognitive strategies. As mentioned there and elsewhere (Lee, 2016, etc.), finding research techniques for accurately determining the various contributing roles is an ongoing research question.

Individualised extensive reading technologies and research

The use of technology and network-accessible and searchable texts has significantly increased the world of content available to today's learners, meaning intelligent systems can now be developed to pair a learner's current level with appropriate material. Liaw and English (2017) cite several studies looking into the automated levelling of digital texts, with a focus on lexical complexity. Research into analysing the complexity of texts selected for individualised extensive reading has also been reported, though again with a focus on lexical complexity (Azab et al., 2013; Walmsley, 2015; Yeung & Lee, 2018).

This research has typically focussed on simple measures of lexical knowledge, i.e., "knowing" or "not knowing" the meaning of a word (and potentially providing glosses for unknown words), which may also pose problems in itself by over-simplifying a complex and multi-layered situation with relation to lexical depth (Gyllstad, 2013; Pignot-Shahov, 2012).

Grammatical knowledge and reading comprehension

The study of the role of grammatical knowledge in L2 reading comprehension has received some renewed attention over the last decade or so (Jung, 2009; Lee, 2016; Shahi, 2016; Shiotsu, 2010; Shiotsu & Weir, 2007; Zhang, 2012) after a relatively long period where grammatical knowledge was considered to only play a minor role (Jung, 2009). Though some of this research continues to suggest grammar only plays a minor role (e.g., Zhang, 2012), there is also strong evidence that at least for some types of reading, grammatical knowledge is indeed an important factor for comprehension.

A possible confounding factor in previous research has been pointed out by Lee (2016), in that many of the studies use different testing techniques, these techniques appear to isolate different aspects of the reading process and the effects can differ according to mastery level and task complexity (Jung, 2012; Lee, 2016).

Most of this research deals with L2 English, and the literature appears to have relatively little on the influence of grammatical knowledge for non-English L2 reading, including Chinese. This work hopes to further our knowledge of the influence of grammatical knowledge on L2 reading comprehension for Chinese and show that using PUMLs for such research could greatly increase our ability to tease apart the multiple contributing influences on the complex process of reading.

Mandarin Chinese

The use of Mandarin Chinese as an international language is increasing and the Chinese government is actively promoting the teaching, learning and use of the language for business, travel and cultural purposes, particularly via the network of Confucius Institutes worldwide

(Goh, 2017). China's economic, cultural and technological importance is also growing rapidly, yet many factors combine to thwart its widespread adoption as an international *lingua franca*.

The difficulty of learning a character-based writing system, tones, and significant cultural differences, among others, are some of the reasons why the US Foreign Service Institute puts Chinese into their categorisation system at level 5 (Doughty, 2019) – the level reserved for the most difficult languages for American native speakers of English to learn. The number and economic clout of the combined countries of speakers of Western European languages is vast, and many of the difficulties native speakers of English learning Chinese face are the same as native speakers of French, Spanish and Portuguese. As such, learning Chinese to a level where they can competently operate in a wide range of business, educational and cultural contexts is a major challenge.

As the current dominant international *lingua franca*, English has received the lion's share of interest from researchers and technologists alike. With the growing interest in Chinese as an international language (Goh, 2017) and the particular issues facing non-Sinic language learners, more research to create innovative technologies for learning Chinese is clearly needed. While the the initial focus in this work is on learning Chinese for speakers of English, the technologies investigated are all expected to apply similarly for most of the world's widely spoken/written languages.

HSK

One of the issues raised by scholars surrounding teasing out the role of grammatical knowledge in L2 reading is the definition and scope of what is considered "grammatical knowledge" (Shiotsu & Weir, 2007). In order to add practical utility and significance, it was decided to focus on an existing examination framework for this question – the *Hanyu Shuiping Kaoshi,* or HSK (Hanban, 2014). By using an existing standard framework, the knowledge elements tested have practical utility for students of Mandarin Chinese, and pre-validated testing resources are available which can be drawn from.

The HSK is now modelled on the CEFR system and is the recognised standard for testing and levelling in Mainland China[14]. The Confucius Institute Headquarters or Hanban, as it is also known, is "a public institution affiliated with the Chinese Ministry of Education"[15]. As the major international exam of Mandarin Chinese, and a requirement for proving competence for such things as entrance to Mainland Chinese universities, using the rules as directly defined in the standard has clear practical and theoretical justification.

14 http://english.hanban.org/node_8002.htm, retrieved 12/11/2019.

15 http://english.hanban.org/node_7719.htm, retrieved 12/11/2019.

International curriculum for Chinese language education (Hanban, 2014) describes the pedagogical content (vocabulary and grammar) expected for each of the 6 levels of the exam, which are claimed by Hanban to correspond to the six well-known Common European Framework of Reference for Languages (CEFR) levels A1-C2[16]. Appendix 4 (pp115-146) contains descriptions of the common grammar points at each level, and these descriptions are used as the basis in this work.

The Zone of Proximal Development

The Soviet scholar Leo Vygotsky proposed the idea of the Zone of Proximal Development early in the 20th century. Much to the dismay of educators, philosophers and psychologists the world over who have had to struggle with his many undeveloped theories, Vygotsky suffered an untimely death at an early age. One of his ideas that still has significant traction with many education scholars is the Zone of Proximal Development. He suggested that children learn best in an environment where things are too hard to do without help but with the help of another (usually older) person, be it a sibling or teacher, but not so hard that it is still the learner performing the task (Schaffer, 2006). Vygotsky and many subsequent proponents were relatively clear that this ZPD was enabled and maintained by other human beings – human teachers in the broad sense – but the notion has been extended by others to encompass a broader notion of both human and technological aids (McLoughlin, 1999, among others). This extra help, or “scaffolding”, that the learner receives is just help from somewhere that turns something impossible for the learner to achieve alone into something that can be achieved with some effort.

ZPD and i+1

Krashen’s Input Hypothesis (Krashen, 1982) introduced a notion with strong parallels to the idea of the ZPD – i+1. While it is true that scholars such as Kinginger (2001) argue there are fundamental philosophical differences between the two concepts, for the purposes of this work they can be considered “equivalent”. The basic notion of a “zone” or level “+1” just above the current level of the learner where learning is said to occur optimally is common to both, and sufficiently similar for our purposes.

Vygotsky’s work has also been applied to L2 research, in works such as Lantolf (2000), Lantolf and Thorne (2006) and Lantolf and Poehner (2014), among others. The term ZPD is used in this work, as it encompasses a more general notion of the development and learning which is enabled by a PUML.

16 http://english.hanban.org/node_8002.htm, retrieved 12/11/2019.

PUML and the ZPD

The learning situations of L2 learners are as many and varied as the learners are themselves. In any given classroom or lecture theatre, the social, educational and material resources and experiences of the learners vary greatly. As such, targeting, and hitting, the ZPD over a group of learners over the timescales involved in language learning using traditional instructional methods is an almost intractable task. Individuals' learning preferences in a given group can be markedly different, and can also change over time. As such, even attempting to optimise the learning over an entire group can be significantly sub-optimal for a large proportion of the individuals that make it up.

Technology, and a PUML in particular, could play a very important role in personalising both learning and instruction, and provide many tools for optimising the learning experience across several axes. An accurate and evolving PUML, accompanying the learner across the many learning contexts they encounter over their language learning journey, could allow tools to adapt in real time to a learner's learning needs.

Like a personal human teacher, tools that get constant feedback from the learner can adapt in real time, providing extra help, slowing down or speeding up as needed by the learner at that precise moment. A key aspect of a PUML compared with most existing tools is its potential independence from a particular learning institution, set of technologies or company. By providing a standard API, and freedom for the learner to control many aspects such as who and what client technologies can access and analyse their data, learners are empowered in their learning journey (Al-Jadaa et al., 2017). Research has shown that this freedom can have a significantly positive effect on learning outcomes for many learners (Benson, 2011; Krashen, 1982; Lai, 2017; see also Rudzewitz et al., 2019).

With a PUML and sufficient quantities of data, it becomes possible to use Educational Data Mining/Learning Analytics techniques to research and optimise many aspects of the L2 learning process (Rudzewitz et al., 2019), independent of particular client technologies. With a sufficient variety of client learning applications, rich veins of research into the interaction of the different language skills open up, allowing educators a much deeper understanding of how to design and implement wholistic, technology-mediated educational experiences.

Pilot Study

The goal of the pilot study is to validate the mechanisms to be used for gathering evidence for knowledge of a particular grammatical structure and evaluating the complexity of a text depending on this information.

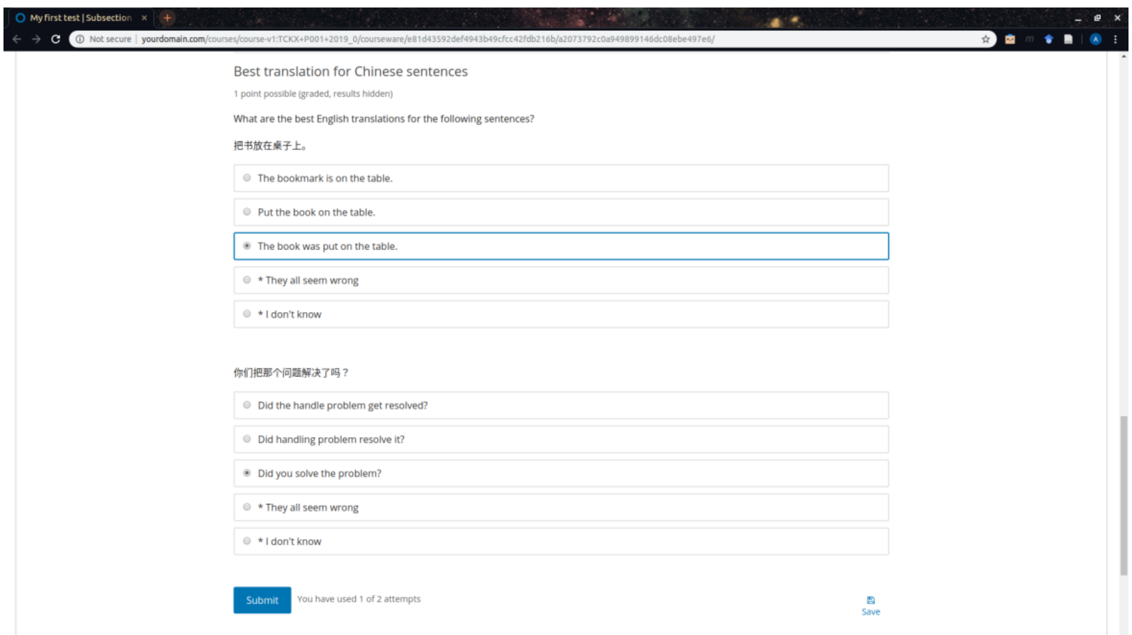
Software components

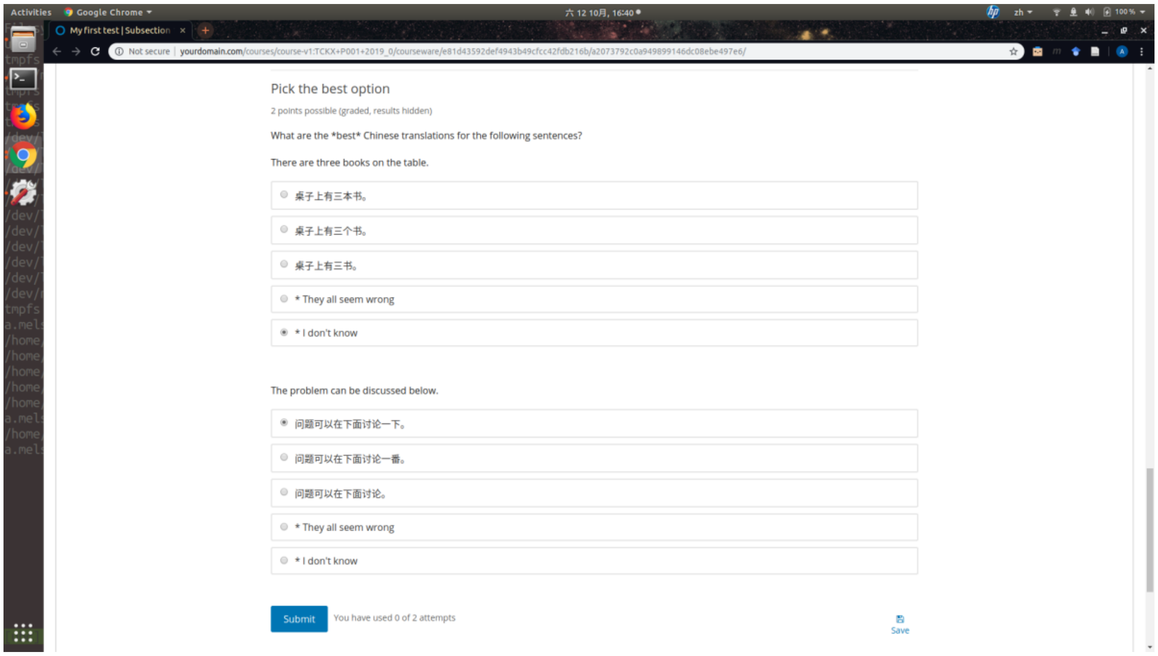
Four main software components are needed for the pilot study:

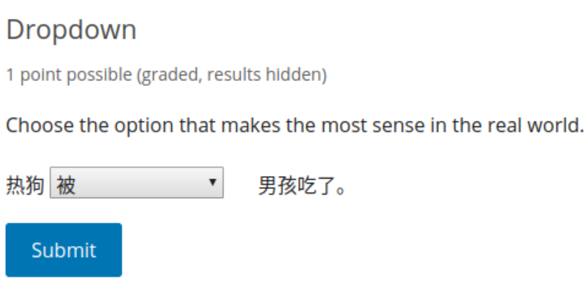
1. Evidence gathering/testing interface (Learning Management System, LMS)
A bespoke interface for testing knowledge of the selected grammatical features will be developed. The following minimal set of features is required:

- User management (administrators and learners)
- Question entry
- Test generation and administration (including user complexity judgement feedback)
- Test evaluation/grading
- User-model update

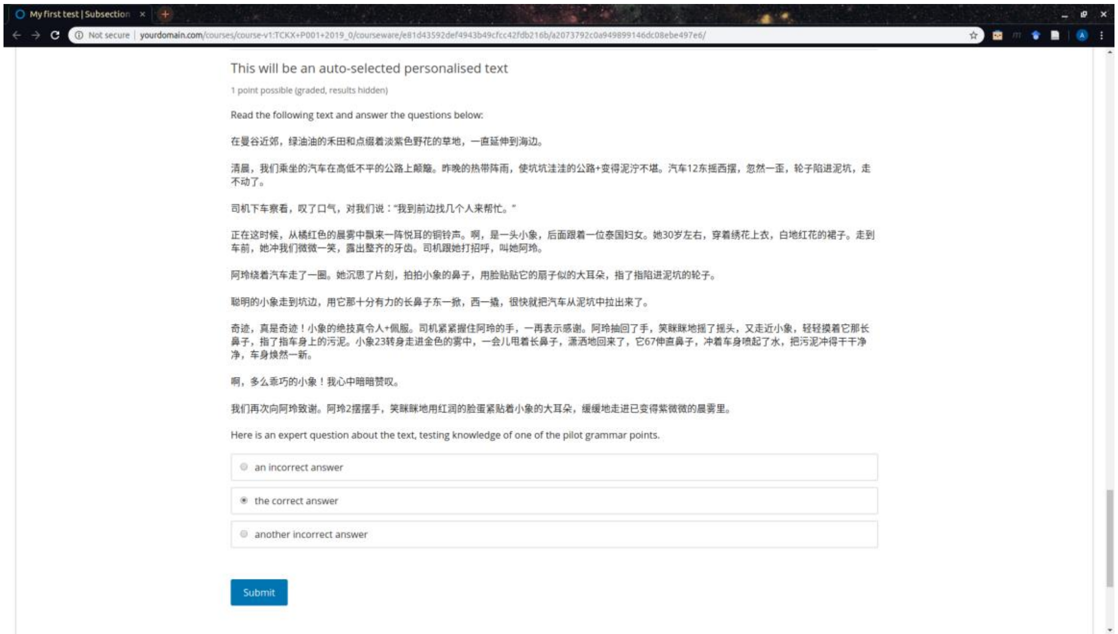
As the requirements for this component are fairly standard, it is expected that relying on an existing LMS will reduce development time, increase quality and stability and provide for easy export of data. Below are some example screenshots from an LMS with questions for gathering evidence of knowledge of grammar points:







In order to simplify the initial phase, this interface will also allow the learners to perform the reading comprehension verification tasks (reading).



2. PUML

The Pilot Study will only implement the very basic rudiments of a PUML, including mainly the user model ontology and basic interfaces available in the LMS. Analytics and direct user update feature will be added after the pilot.

The following minimal set of features is required:

- User management (connection to the LMS)
- User model documentation
- Datastore
- Network-accessible API

3. Text analysis

In order to identify the study's target grammatical constructions in texts, these texts must be analysed. CoreNLP (Manning et al., 2014) is widely used for the analysis of Chinese (for example, Hu et al., 2019; Wang, Li, & Xue, 2018; Yeung & Lee, 2018) and can identify the constructions in the target group (via POS tagging and dependency parsing). Personalisation will be performed by other components, meaning CoreNLP can be used off-the-shelf.

4. Complexity calculator.

This system will determine the personalised complexity measure of a text, given the information from the LMS and the text analysis. This will be developed in Python 3.6+. The following minimal set of features is required:

- User management (connection to the LMS)
- Complexity Calculation Engine (CCE)

- Network-accessible API

Datasets

1. A set of considered personalisable grammatical constructions. Proposed features from the Penn Chinese Treebank (3.0) POS tagging guide (Xia, 2000), and HSK (Hanban, 2014):

- Classifiers (PCT 2.6; HSK 1.7, 2.4, 3.3, 4.1.1, 4.2.1, 4.3, 5.1, 5.3, 6.2)
- "ba" sentences (PCT 2.11.5; HSK 3.20, 4.10, 5.10, 6.7)
- "bei" sentences (PCT 2.11.3-4; HSK 3.21)

2. A set of non-personalisable grammatical constructions. Proposed features:
- Ratio of conjunctions (fixed list, see PCT 4.1, 4.4, 4.6)
- Average sentence length

3. Mappings of instruction sourced constructions to analysed structures. Via CoreNLP:
- Classifiers = POS "M", weight CW (TBD)
- "ba" = POS "BA", weight BA (TBD)
- "bei" = POS "SB", "LB", weight BEI (TBD)
- Conjunctions = POS "CS"/"CC"/"AD"(see list in relevant PCT sections)/Nb of words, weight CONJ (TBD)
- Average sentence length = Total number of words / Total number of sentences, weight ASL (TBD)

4. A set of questions/answers for gathering evidence. These will be validated with an experienced CSL teacher to ensure the material focusses only on the targeted structures and doesn't introduce additional complexities. TBD.

5. A set of graded texts extracted from previous HSK exams. This provides easy access to texts that have been validated for a given level by the testing authority itself, guaranteeing accuracy – HSK exam texts of a given level are by definition at that level.

Complexity Calculation Engine (CCE)

Personalised measures:
- Each sentence has a default complexity value of zero.
- For each unknown pattern, each sentence that contains the pattern is marked as 1.
- For each unknown pattern, the fraction of all sentences containing the pattern is calculated, and is multiplied by the Weight.

Non-personalisable features:
- The total number of conjunctions is identified divided by the total number of words and multiplied by the weight.

- If the average sentence length is less than or equal to 6, the value is zero (??? motivate this choice). If larger than 6, the difference is multiplied by the Weight.

The weighted pattern scores are added together to form a total Complexity Score.

Pilot Study Validation Procedure

A cohort of 5 active learners of Mandarin Chinese will be asked to test the system. Participants will then be interviewed by an experienced CSL teacher to validate the accuracy of the system in identifying knowledge of the Pilot grammar points. The following aspects will thus be directly or indirectly tested:

- Ability to identify knowledge of a particular grammar point (accuracy of the PUML)
- Ability to rank/filter texts from the set of Pilot texts corresponding to the grammar points known by the learner and those present in the text (accuracy of the CCE)
- Suitability/Capabilities of the LMS interface for collecting, storing and making available evidence of knowledge (user input entry and testing)
- Communication between the components (technical validation)

References

Al-Jadaa, A. A., Abu-Issa, A. S., Ghanem, W. T., & Hussein, M. S. (2017). Enhancing the intelligence of web tutoring systems using a multi-entry based open learner model. *Proceedings of the Second International Conference on Internet of Things, Data and Cloud Computing - ICC ’17*, 1–5. https://doi.org/10/ggczjv

Azab, M., Salama, A., Oflazer, K., Shima, H., Araki, J., & Mitamura, T. (2013). *An NLP-based reading tool for aiding non-native English readers*.

Benson, P. (2011). *Teaching and researching autonomy* (2nd ed.). Harlow ; New York.: Harlow ; New York. : Longman/Pearson.

Bozorgian, H. (2012). The Relationship between Listening and Other Language Skills in International English Language Testing System. *Theory and Practice in Language Studies*, *2*(4), 657–663. https://doi.org/10/ggczh9

Chen, X., & Meurers, D. (2019). Linking text readability and learner proficiency using linguistic complexity feature vector distance. *Computer Assisted Language Learning*, *32*(4), 418–447. https://doi.org/10.1080/09588221.2018.1527358

Chun, D. M. (2016). The Role of Technology in SLA Research. *Language Learning & Technology*, *20*(2), 98.

Doughty, C. J. (2019). Cognitive Language Aptitude: Cognitive Language Aptitude. *Language Learning*, *69*, 101–126. https://doi.org/10/ggcp55

Goh, Y.-S. (2017). *Teaching Chinese as an International Language: A Singapore Perspective*. https://doi.org/10.1017/9781107280472

Grabe, W. (2008). *Reading in a Second Language: Moving from Theory to Practice*. https://doi.org/10.1017/CBO9781139150484

Gyllstad, H. (2013). Looking at L2 vocabulary knowledge dimensions from an assessment perspective—Challenges and potential solutions. *Bardel, C., Lindqvist, C., & Laufer, B.(Eds.) L*, *2*, 11–28.

Hanban. (2014). *International curriculum for chinese language education.* Place of publication not identified: Beijing Language & Cultur.

Hu, H., Li, W., Zhou, H., Tian, Z., Zhang, Y., & Zou, L. (2019, June). *Ensemble Methods to Distinguish Mainland and Taiwan Chinese*. 165–171. Retrieved from https://doi.org/10.18653/v1/W19-1417

Jung, J. (2009). Second language reading and the role of grammar. *Studies in Applied Linguistics and TESOL*, *9*(2).

Jung, J. (2012). Relative Roles of Grammar and Vocabulary in Different L2 Reading Tasks. *English Teaching*, *67*, 57–77. https://doi.org/10/ggcz2f

Kay, J., & Kummerfeld, B. (2019). From data to personal user models for life-long, life-wide learners. *British Journal of Educational Technology*, bjet.12878. https://doi.org/10/gf8547

Kinginger, C. (2001). I+ l# ZPD. *Foreign Language Annals*, *34*(5), 417–425. https://doi.org/10/fv4h4b

Krashen, S. D. (1982). *Principles and practice in second language acquisition* (1st ed). Oxford ; New York: Pergamon.

Krashen, S. D. (2004). *The power of reading: Insights from the research* (2nd ed.). Westport, Conn. : Portsmouth, NH: Westport, Conn. : Libraries Unlimited.

Lai, C. (2017). *Autonomous language learning with technology: Beyond the classroom*. London: London : Bloomsbury Academic.

Lantolf, J. P. (2000). *Sociocultural theory and second language learning* (Vol. 78). Oxford university press.

Lantolf, J. P., & Poehner, M. E. (2014). *Sociocultural theory and the pedagogical imperative in L2 education: Vygotskian praxis and the research/practice divide*. New York ; London: Routledge.

Lantolf, J. P., & Thorne, S. L. (2006). *Sociocultural theory and the genesis of second language development*. Oxford: Oxford : Oxford University Press,.

Lee, J.-W. (2016). The Role of Vocabulary and Grammar in Different L2 Reading Comprehension Measures. *English Teaching*, *71*(3).

Liaw, M., & English, K. (2017). Technologies for Teaching and Learning L2 Reading. In C. A. Chapelle & S. Sauro (Eds.), *The Handbook of Technology and Second Language Teaching and Learning* (1st ed., pp. 62–76). https://doi.org/10.1002/9781118914069.ch5

Manning, C. D., Surdeanu, M., Bauer, J., Finkel, J., Inc, P., Bethard, S. J., & Mcclosky, D. (2014). The Stanford CoreNLP natural language processing toolkit. *In ACL, System Demonstrations*. https://doi.org/10/gf3xhp

McLoughlin, C. (1999). *Scaffolding: Applications to learning in technology supported environments*. 1826–1831. Retrieved from https://www.learntechlib.org/primary/p/7213/

Nation, I. S. P. (2001). *Learning Vocabulary in Another Language*. Retrieved from https://academic.microsoft.com/paper/2114282361

Nation, P. (2015). Principles Guiding Vocabulary Learning through Extensive Reading. *Reading in a Foreign Language*, *27*(1), 136–145.

Pignot-Shahov, V. (2012). Measuring L2 receptive and productive vocabulary knowledge. *Language Studies Working Papers*, *4*(1), 37–45.

Reinders, H., & Stockwell, G. (2017). Computer-assisted SLA. *The Routledge Handbook of Instructed Second Language Acquisition*, 361–375. https://doi.org/10/ggczps

Rudzewitz, B., Ziai, R., Nuxoll, F., De Kuthy, K., & Meurers, D. (2019). *Enhancing a Web-based Language Tutoring System with Learning Analytics*. Proceedings of the 2nd workshop on Learning Analytics: Building bridges between the Education and the Computing communities (LABBEC) held at Educational Data Mining (EDM) 2019.

Schaffer, R. H. (2006). ZONE OF PROXIMAL DEVELOPMENT and: CULTURAL TOOLS SCAFFOLDING GUIDED PARTICIPATION. In *Key concepts in developmental psychology.* Retrieved from https://search.credoreference.com/content/entry/sageukdp/zone_of_proximal_development_and_cultural_tools_scaffolding_guided_participation/0

Shahi, M. A. J. (2016). Improving Reading Comprehension of ESP Learners: Grammar practice. *International Journal of Language Learning and Applied Linguistics World, 11 (2), 135*, *145*.

Shiotsu, T. (2010). *Components of L2 reading: Linguistic and processing factors in the reading test performances of Japanese EFL learners* (Vol. 32). Ernst Klett Sprachen.

Shiotsu, T., & Weir, C. J. (2007). The relative significance of syntactic knowledge and vocabulary breadth in the prediction of reading comprehension test performance. *Language Testing*, *24*(1), 99–128. https://doi.org/10/bvkjvk

SUNG, Y.-T., DYSON, S. B., CHEN, Y.-C., LIN, W.-C., & CHANG, K.-E. (2015). Leveling L2 Texts Through Readability: Combining Multilevel Linguistic Features with the CEFR. *The Modern Language Journal*, *99*(2), 371–391. https://doi.org/10/f7mwhd

Walmsley, M. (2015). *Learner Modelling for Individualised Reading in a Second Language* (Thesis, University of Waikato). Retrieved from https://researchcommons.waikato.ac.nz/handle/10289/10559

Wang, C., Li, B., & Xue, N. (2018, June). *Transition-Based Chinese AMR Parsing*. 247–252. Retrieved from https://doi.org/10.18653/v1/N18-2040

Xia, F. (2000). The part-of-speech tagging guidelines for the Penn Chinese Treebank (3.0). *IRCS Technical Reports Series*, 38.

Yeung, C. Y., & Lee, J. (2018). Personalized text retrieval for learners of chinese as a foreign language. *Proceedings of the 27th International Conference on Computational Linguistics*, 3448–3455.

Zhang, D. (2012). Vocabulary and Grammar Knowledge in Second Language Reading Comprehension: A Structural Equation Modeling Study. *Modern Language Journal*, *96*(4), 558–575. https://doi.org/10/f4p2nm