Learning Low-Resource Languages Through NLP-Driven Flashcards: A Case Study of Hokkien in Language Learning Applications

Tai Zhang[1], Lucie Yang[1], Erin Chen[1], Karen Riani[1], Jessica Zipf[3], Mariana Shimabukuro[2], En-Shiun Annie Lee[1, 2],

[1]University of Toronto, [2]Ontario Tech University, [3]University of Konstanz

Abstract

LANGLEARN is an open-source framework designed to facilitate autonomous learning of low-resource languages (LRL). By combining a language-agnostic approach with AI-enhanced flashcards, LangLearn empowers users to generate custom flashcards for their vocabulary, while offering structured learning through both pre-curated and self-curated decks. The framework integrates six key components: the word definition, corresponding Hanji characters, romanization with numeric tones, audio pronunciation, a sample sentence, as well as a contextual AI-generated image. LangLearn currently supports English and Taiwanese Hokkien (a variety of Southern Min), with plans to extend support for other dialects. Our preliminary study demonstrates that LangLearn positively empowers users to engage with LRLs using their vocabulary preferences, with a comprehensive user study currently underway. LangLearn's modular structure enables future expansion, including ASR-based pronunciation practice. The code is available on GitHub.

1 Introduction

Language learning programs have becoming increasingly popular, with apps such as Duolingo reaching over 200 million users 2023 (Curry, 2024). However, a key limitation in many of these apps is their focus on popular languages, leaving low-resource languages (LRLs) such as Taiwanese Hokkien underrepresented. Taiwanese Hokkien, a variety of Southern Min, is facing a dramatic decline in speakers — less than 10% of people aged 18–24 can speak Hokkien compared to 66% of those 65 and older (Directorate General of Budget, Accounting and Statistics, 2021).

While recent advances in generative Artificial Intelligence (genAI) models have enabled more personalized language learning experiences with shifts to learner-centered experiences (Joshi et al., 2020; Ranathunga et al., 2023), LRLs remain poorly supported by popular large language models due to the lack of training data (Zhong et al., 2024; Chang et al., 2024). Dialect differences and the lack of written material make it particularly difficult to develop language models or learning tools for these languages, creating a critical gap in technology support. Notably, while personalization was not shown to increase vocabulary learning, it was shown to improve learner motivation (Leong et al., 2024).

To address these challenges, we present LangLearn, a modular framework designed to facilitate the autonomous learning of LRLs. LangLearn integrates multiple AI models into a flashcard generation pipeline, producing personalized AI-enhanced flashcards that support vocabulary acquisition, pronunciation practice as well as contextualized learning. The framework is built to be extensible, allowing for easy adaption for other LRLs through language-specific translation/speech synthesis models (Lee et al., 2022; Su et al., 2024; Khiu et al., 2024; Nayak et al., 2023). As an autonomous self-study app, LangLearn utilizes AI models to to meet the user at their interests.

This paper describes the design and implementation of LangLearn, focusing particularly on how it supports the self-study of Hokkien as a second language (Lu et al., 2024). Results of preliminary user studies showed that 82% of users agreed that the app was easy to use, with numerous users enjoying the novelty of the contextualized definitions. Reliability issues around AI-generated content such as inconsistent translations and image relevance are being continually addressed with a community rating system and user feedback, with future updates planned to improve the accuracy of AI models, enhance pronunciation support, and expand LangLearn's language offerings.

Proceedings of the 2025 Conference of the Nations of the Americas Chapter of the Association for Computational Linguistics: Human Language Technologies (System Demonstrations), pages 303–312

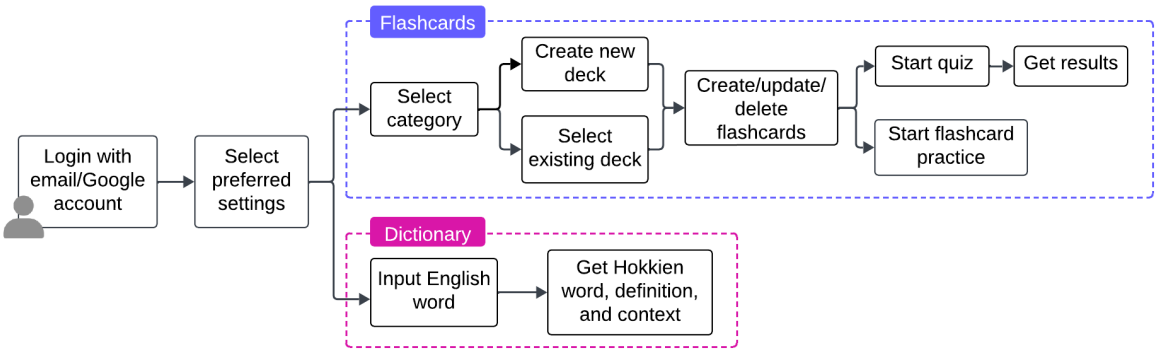

Figure 1: Overview of the user flow through the LangLearn application, illustrating user interactions within the two major components: a dictionary, and flashcards.

2 LangLearn — Learner-centered Design

LangLearn is designed to address the unique challenges of learning LRLs like Hokkien, which are predominantly oral with limited written resources (Chen et al., 2023). The application has two major components: (1) a dictionary and (2) flashcard features. Figure 1 illustrates the primary user flow through the application, from signing up to practicing with flashcards and testing knowledge with quizzes.

Learner account. Users can create an account with either their email or Google account. Once the user is authenticated, they can access LangLearn's features. The user account enables personalization of the learner's experience and safe storage of their progress.

2.1 Dictionary and Dynamic Content Generation

The dictionary enables users to input an English word and retrieve its corresponding vocabulary in Hokkien. As shown in Figure 2, the dictionary offers different output options; learners may customize the output options in the application's settings. The application has six output options, including (1) word definition, (2) the corresponding Hokkien word in characters (Hanji), (3) romanization in numeric tones, (4) audio of the pronunciation, (5) an example sentence in Hokkien (also with characters, romanization, and audio), and (6) an AI-generated image representing the word. The last two output options, in particular, were designed to leverage contextual learning and help learners observe their target vocabulary in usage.

Hokkien Models These functionalities are supported by integrating the APIs developed for the Hokkien language by Lu et al. (2024). By providing this collection of output options, LangLearn aims to support autonomous language learning while at the same time emphasizing oral language skills — especially as LRLs like Hokkien lack a widely adopted writing system (Chen et al., 2023; Sheu et al., 2024). We first use the *Taiwanese Hokkien Translation Model* to translate the English input word to Hokkien characters (Hanji). Then, we use the *Hokkien Transliteration Model* to convert the characters to the romanization. Finally, we use the *Text2Speech Model* to generate realistic audio for the word based on the romanization. For images, we use the *Text2Image Model*.

GenAI considerations for education. While integrating AI-based APIs enables personalized and contextual learning through dynamic content generation, the accuracy of the output is not guaranteed due to the inherent challenges of the genAI models. To mitigate this issue, all images and contextual sentences for starter vocabulary in the dictionary have been reviewed for appropriateness. We also have a 'feedback' option for learners to report inaccuracy or content issues.

2.2 Flashcards & Quizzing

The core language learning and practicing functionality of LangLearn are enabled by the Flashcard feature. From the flashcard feature, learners can select from 11 pre-defined flashcard categories, such as 'Daily Conversations' and 'Shopping.', or simply create a new deck.

Creating and organizing decks. Users are also able to customize their learning by editing flashcard

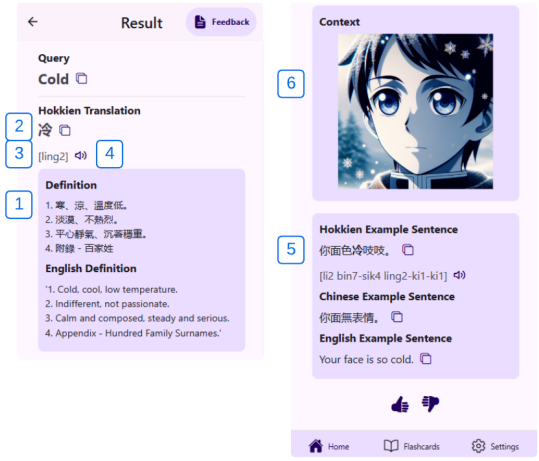

Figure 2: Example of linear Dictionary output, with all six components, for the vocabulary word of *'cold'*.

(a) (b)

Figure 3: Example of Flashcards and Quiz interface for the vocabulary word of *'sorry'*: **a)** the flashcard front with Hokkien characters, romanization and audio and **b)** the quiz for the Daily Conversations Starter deck, with English answers option.

decks and individual flashcards. In deck management, users can update or delete existing decks or create new ones by selecting a list of initial words from the category. They can also choose the visibility of their deck: public decks are visible to all users, whereas private decks are only available to the owner. Once in a selected deck, users can perform various flashcard operations such as creating, updating, or deleting flashcards. When creating flashcards, users have the option to "auto-fill" the card, where the Hokkien is automatically generated by the Hokkien APIs, and the multiple choice options are generated with the OpenAI API using GPT-4o-mini.

Practicing vocabulary. Users can start flashcard practice sessions where the app presents words in randomized order from their chosen deck. By default, the front of the flashcard shows the Hokkien word in characters with the corresponding romanization and audio, as shown in Figure 3a, and the English translation on the back.

Testing vocabulary. Users can then test their knowledge in the Quiz section, where the app generates multiple choice questions based on the deck. The user can choose to answer with English, as shown in Figure 3b, Hokkien characters, or Hokkien romanization, depending on their learning needs. Quizzes are complemented by a progress tracker and results summary, providing immediate feedback and promoting consistent improvement.

2.3 Technology Stack

LangLearn is a mobile application with web compatibility, written with JavaScript. The database is implemented with *Firebase* for managing real-time data, including users, flashcards, and audio. The front-end is built with React and *Expo* frameworks, supporting iOS and Android from a single code base, as well as *Native Base* to provide consistent and reusable interface components. APIs handle translation, sentence generation, and audio synthesis. Deployment is managed through *Netlify*, which automates builds and hosting. The system uses a modular architecture that connects the front-end, back-end, APIs, and deployment services, as shown in Figure 4.

3 User Study

To help assess the usability of LangLearn, we completed a preliminary study over three weeks using an early version of the application to learn Hokkien, we recruited 39 participants. The participants were undergraduate students from two Canadian universities without prior knowledge of Hokkien. They were asked to use the application for three weeks and provide feedback for the application through usability surveys at the end of each week.

3.1 Methodology

The study was entirely online and participation was entirely voluntary.

Upon entry of the study, participants were provided with login credentials and general instructions on accessing the app. In addition, they were asked to complete an **entry survey** which inquires their previous experience experience with language

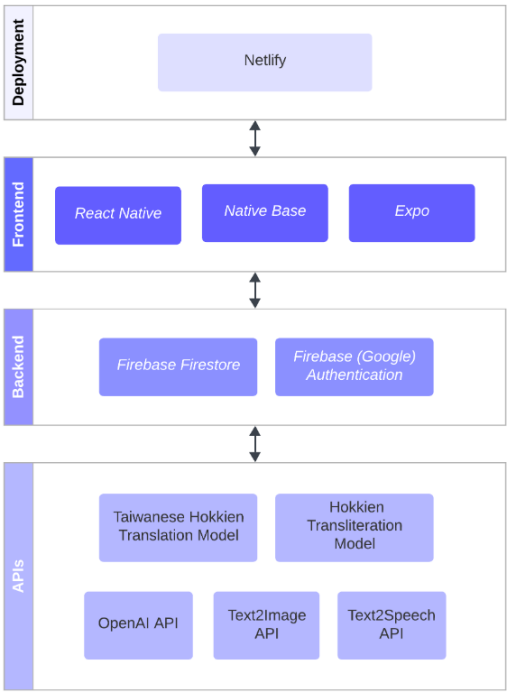

Figure 4: System architecture showing the interactions between components.

learning apps and language backgrounds. We assessed the usability of the app through an usability survey for which the participants were reminded through email to complete at the end of each week of the three weeks study period. The usability survey focuses on user perception on two main areas: whether the app is intuitive and easy to use,and the effectiveness of the app in facilitating Hokkien-learning. Participants were asked to answer 5-Point Likert scale agreement statements ("strongly agree", "agree", "neutral", "disagree", or "strongly disagree") related to the two categories. Participants were also asked to provide feedback on any obstacles/bugs they encountered while using the app and general comments. To conclude the study, participants were asked to complete an **exit survey**, which inquired participants' overall experience with the app, whether they were satisfied with their progress in learning Hokkien, and if they would continue using the app after the study. Participants were also asked to comment on how to improve features of the application, and if they would recommend the app to other Hokkien-learners.

3.2 Results

Since participation in the study was entirely voluntary, we were unable to control the number of survey responses and amount of usage; which is a common issue for longitudinal studies. As shown in Table 1, the number of survey responses decreased over weeks of the study. Since much of the usability survey focused on the experience with using the app, this paper focuses on the results from participants that used the application at least once during the study week.

Number of Times Used During Week				
	None	Once	More Than Once	Total Responses
Week 1	10	7	12	29
Week 2	11	6	3	20
Week 3	8	2	2	12

Table 1: Number of usability survey responses with the number of participants that used the app once, more than once, or did not use the app each study week.

Intuitive Use Statements that focused on intuitive use of the app included:

1. It was easy to interact with the app's features.
2. The layout of the app was intuitive.
3. Overall, the platform is easy to use.
4. The flashcard system was intuitive and easy to use.

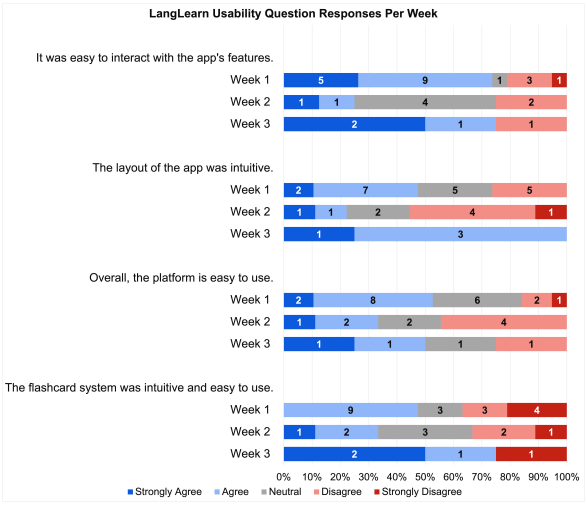

Figure 5: Responses from users who used the app at least once during the study week on the questions related to intuitive use of the app, with the count of users that selected each option.

As shown in Figure 5, the majority of responses were "Strongly Agree", "Agree", or "Neutral" on all of the questions. An exception was the week 2 response of question 2, where the majority of responses were "disagree" and "strongly disagree".

Hokkien Language Learning Questions that focused on whether the app facilitated Hokkien-

learning included:

1. The platform has been helpful for learning Hokkien.
2. The text-to-speech on the home page helps with understanding the translated words/sentences.
3. The image generation on the home page helps with understanding the translated words/sentences.
4. The generated images for the translations on the home page are relevant.
5. The feedback provided while using flashcard system was helpful.

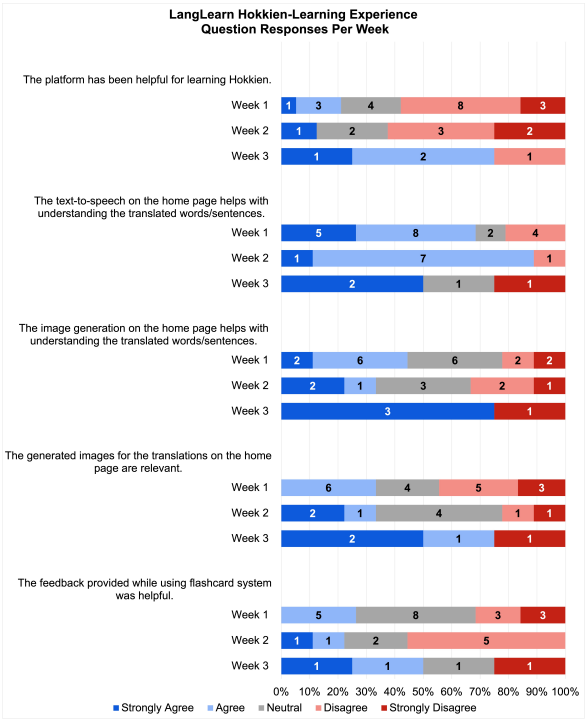

Figure 6: Responses from users who used the app at least once during the study week on the questions that asked about Hokkien-learning features, with the count of users that selected each option.

As shown in Figure 6, the majority of responses for question 2–5 were "Strongly Agree", "Agree", or "Neutral", except Week 2 of question 5. The only question that did not follow this pattern was question 1, where majority of responses in Week 1 and 2 were "disagree" and "strongly disagree".

3.3 Discussion

Intuitive Use Overall user feedback indicated that the app was intuitive to use and easy to interact with. Out of the 19 users who used the app at least once in Week 1 of the trial, 84.2% (n = 16) answered "agree", "strongly agree", or "neutral" to the question "Overall, the platform is easy to use.", with 52.6% (n = 10) of participants that used the app at least once that week answering "agree" or "strongly agree". 73.7% (n = 14) of the users that used the app at least once in Week 1 answered "agree" or "strongly agree" to the statement "It was easy to interact with the app's features". Though improvements could be made as some participants suggested that more guidance in the app would be helpful in supporting Hokkien beginners, overall results suggest our app is easy to interact.

Hokkien Language Learning The usability survey asked for user feedback on whether certain features facilitated Hokkien language learning, with focus on the flashcard system, and the text-to-speech and contextualization components of the dictionary function. Overall user feedback indicated that our app at the stage of the study lacked certain features that would support Hokkien learning. Insight was provided through user feedback in the open-ended general feedback question, where much of the feedback was focused on whether the application was beginner friendly, since all participants had no prior knowledge of Hokkien before using the app. Our initial version of the app did not include romanization or audio, which was challenging for beginners, since the pronunciations of the characters cannot be easily inferred. Some comments from participants who answered "disagree" to the question "The platform has been helpful for learning Hokkien" included:

> Participant P2: "... *I wish the flashcard system had an audio component since I cannot read Chinese characters...*"
> Participant P5: "*I feel like the way it is now could be useful for someone who already knows a bit of Hokkien, but as someone who doesn't have any prior Hokkien knowledge at all, I had to use additional external resources to understand how the language worked even at the basic level (e.g., understanding the tones of the language).*"

To address this, we have updated the application to include speech and pronunciation features as a part of the flashcard system. In addition, an user study is being conducted, using the updated application, which could potentially provide insight to the effect of speech and pronunciation in facilitating Hokkien language learning.

4 Related Work

4.1 Translation based language tools and multi-modal engagement

Translation tools are important in vocabulary acquisition, whether through machine translation, parallel corpora, or crowd-sourcing (Jiménez-Crespo, 2017). Studies have shown that first language (L1) translation is particularly effective for vocabulary expansion compared to second language (L2) definition approaches (Joyce, 2015; Jee-Young Shin and Choi). Additionally, digital flashcards have been shown to be an effective tool for learning vocabulary (Byrd and Lansing, 2016; H. Gülru Yüksel and Yılmaz, 2022).

LangLearn utilizes these findings by creating L1 translation flashcards with one side containing the L1 definition, and the other side containing the L2 definition. The Object Detection Translation (ODT) app demonstrated that combining visual and verbal presentation enhanced vocabulary acquisition for English learners, achieving higher success rates compared to traditional methods (Pei-Lin Liu and Chiu-Jung Chen, 2023). Based on ODT's findings on the importance of multi-modal learning, LangLearn integrates visual and verbal elements into its translation system. However, while ODT relies on user-uploaded images, LangLearn generates contextual images automatically, making it more scalable for vocabulary acquisition.

4.2 Support for LRL Acquisition

Current popular online Hokkien dictionaries such as moedict.tw[1] and itaigi.tw[2] are in Mandarin Chinese, which creates a significant prerequisite barrier for English-speaking learners. Existing AI-powered language learning solutions such as Glossika [3] attempt to address these challenges through human-recorded audio on a standardized base of words and sentences across numerous languages including Hokkien. This approach faces inherent scalability limitations. Reliance on native speakers for audio recording limits vocabulary coverage and makes expanding content expensive. These challenges are particularly visible for low-resource languages such as Hokkien, which do not have a wide existing base of annotated audio. By utilizing existing pre-trained text-to-speech models, LangLearn can instantly generate pronunciation audio to user-requested vocabulary words (Liao et al., 2022), creating rich contextual content on the fly.

4.3 Contextualized language learning engagement

In terms of promoting greater language acquisition efficacy, recent work has explored various approaches to engaging learners by contextualizing perspectives from the users themselves. Vocabulometer creates word lists based on user knowledge levels (Yamaguchi et al., 2020) while MicroMandarin generates contextualized flashcards based on user location and categories (Edge et al., 2011). Additionally, QuickLearn implements microlearning through notification-based flashcard delivery (Dingler et al., 2017).

Following MicroMandarin's category-based approach, LangLearn organizes content into daily life categories, allowing for specific learning that can be extended. Additionally, LangLearn empowers learners to rate flashcard quality, creating a community-driven quality control system. Finally, by being a mobile application with web compatibility, further work into generating notifications and contextualized learning suggestions could promote user learning on-the-go.

5 Conclusion

As an open-source framework, LangLearn is designed to facilitate autonomous low-resource language (LRL) learning. LangLearn empowers learners by giving them the power to generate contextualized flashcards complete with definitions, sample sentences, as well as audio pronunciation in the target language. These flashcards can be put into decks, or added to the existing categories to support learners in their daily life, with support for quizzes and sharing flashcards out-of-the-box. LangLearn currently supports English and Taiwanese Hokkien, with plans to expand towards ASR-based pronunciation practice to emphasize the spoken nature of many low-resource languages.

6 Future Work and Limitations

The current version of LangLearn has only been tested with a translation model between English and Hokkien. Future work would involve adding additionally language support by bringing in more translation models in the back-end. To support a new language, at least four pre-trained models

[1] https://moedict.tw

[2] https://itaigi.tw

[3] https://ai.glossika.com/language/learn-taiwanese-hokkien

would be needed: a translation model, context generation model, transliteration model, as well as a text to speech model (Lu et al., 2024; Liao et al., 2022).

The current decks of flashcards found in the LangLearn categories focus on vocabulary and basic phrases as opposed to grammar or other language components. Future work could be done similar to (Shimabukuro et al., 2023) where the flashcards can be used to improve grammar or verb conjugations.

Finally, a future two part study built off of the initial findings of the preliminary study is being conducted, focused on highlighting the difference in effectiveness between the simple flashcards containing just the translated word and pronunciation vs the enhanced flashcards with additional genAI components (sample sentences and image).

Ethics Statement

All evaluations conducted and reported in this paper were reviewed and conducted according to the Research Ethics Board of Canada. All participants were volunteers and gave us informed consent for data collection and de-identified reporting of results including quotes, survey, and software logs.

Acknowledgments

We thank Vincent Shuai Zhu for his initial leadership and contributions to the project. We also thank Fatimeh Hassan, Ricky Zhang, Mohammad Abrar, Emma Zhuang, and Yi Yang for their contributions to the application development and preliminary study. We are also thankful to Winnie Chu, Bo-Han Lu, and Prof. Richard Tzong-Han Tsai for development and maintenance of the Hokkien APIs.

We acknowledge the support of the Natural Sciences and Engineering Research Council of Canada (NSERC) Discovery Grant and University of Toronto's Research Opportunities Program and CSC494/495 Program.

References

David R. Byrd and Branden Lansing. 2016. Electronic flashcards inside the classroom: Practical and effective. *Journal of Language Teaching and Learning*, 6(2):1–13.

Tyler A. Chang, Catherine Arnett, Zhuowen Tu, and Ben Bergen. 2024. When Is Multilinguality a Curse? Language Modeling for 250 High- and Low-Resource Languages. In *Proceedings of the 2024 Conference on Empirical Methods in Natural Language Processing*, pages 4074–4096, Miami, Florida, USA. Association for Computational Linguistics.

Peng-Jen Chen, Kevin Tran, Yilin Yang, Jingfei Du, Justine Kao, Yu-An Chung, Paden Tomasello, Paul-Ambroise Duquenne, Holger Schwenk, Hongyu Gong, Hirofumi Inaguma, Sravya Popuri, Changhan Wang, Juan Pino, Wei-Ning Hsu, and Ann Lee. 2023. Speech-to-speech translation for a real-world unwritten language. In *Findings of the Association for Computational Linguistics: ACL 2023*, pages 4969–4983, Toronto, Canada. Association for Computational Linguistics.

David Curry. 2024. Language learning app revenue and usage statistics (2024). *Business of Apps*. Accessed: December 12th, 2024.

Tilman Dingler, Dominik Weber, Martin Pielot, Jennifer Cooper, Chung-Cheng Chang, and Niels Henze. 2017. Language learning on-the-go: opportune moments and design of mobile microlearning sessions. In *Proceedings of the 19th International Conference on Human-Computer Interaction with Mobile Devices and Services*, MobileHCI '17, pages 1–12. Association for Computing Machinery.

Directorate General of Budget, Accounting and Statistics. 2021. 2020 population and residence census: Summary and analysis of initial statistical findings [109]. Technical report, Executive Yuan, R.O.C. (Taiwan).

Darren Edge, Elly Searle, Kevin Chiu, Jing Zhao, and James A. Landay. 2011. MicroMandarin: mobile language learning in context. In *Proceedings of the SIGCHI Conference on Human Factors in Computing Systems*, pages 3169–3178, Vancouver BC Canada. ACM.

H. Güldem Mercanoğlu H. Gülru Yüksel and M. Betül Yılmaz. 2022. Digital flashcards vs. wordlists for learning technical vocabulary. *Computer Assisted Language Learning*, 35(8):2001–2017.

L. Quentin Dixon Jee-Young Shin and Yunkyeong Choi. An updated review on use of l1 in foreign language classrooms. *Journal of Multilingual and Multicultural Development*, 41(5):406–419.

Miguel A. Jiménez-Crespo. 2017. The role of translation technologies in spanish language learning. *Journal of Spanish Language Teaching*, 4(2):181–193.

Pratik Joshi, Sebastin Santy, Amar Budhiraja, Kalika Bali, and Monojit Choudhury. 2020. The State and Fate of Linguistic Diversity and Inclusion in the NLP World. In *Proceedings of the 58th Annual Meeting of the Association for Computational Linguistics*, pages 6282–6293, Online. Association for Computational Linguistics.

Paul Joyce. 2015. L2 vocabulary learning and testing: the use of l1 translation versus l2 definition. In *The*

Language Learning Journal, 46(3), page 217–227. Association for Language Learning.

Eric Khiu, Hasti Toossi, David Anugraha, Jinyu Liu, Jiaxu Li, Juan Flores, Leandro Roman, A. Seza Doğruöz, and En-Shiun Lee. 2024. Predicting machine translation performance on low-resource languages: The role of domain similarity. In *Findings of the Association for Computational Linguistics: EACL 2024*, pages 1474–1486, St. Julian's, Malta. Association for Computational Linguistics.

En-Shiun Lee, Sarubi Thillainathan, Shravan Nayak, Surangika Ranathunga, David Adelani, Ruisi Su, and Arya McCarthy. 2022. Pre-trained multilingual sequence-to-sequence models: A hope for low-resource language translation? In *Findings of the Association for Computational Linguistics: ACL 2022*, pages 58–67, Dublin, Ireland. Association for Computational Linguistics.

Joanne Leong, Pat Pataranutaporn, Valdemar Danry, Florian Perteneder, Yaoli Mao, and Pattie Maes. 2024. Putting things into context: Generative AI-enabled context personalization for vocabulary learning improves learning motivation. In *Proceedings of the CHI Conference on Human Factors in Computing Systems*, CHI '24, pages 32:1–32:15. ACM.

Yuan-Fu Liao, Jane S. Tsay, Peter Kang, Hui-Lu Khoo, Le-Kun Tan, Li-Chen Chang, Un-Gian Iunn, Huang-Lan Su, Tsun-Guan Thiann, Hak-Khiam Tiun, and Su-Lian Liao. 2022. Taiwanese across taiwan corpus and its applications. In *2022 25th Conference of the Oriental COCOSDA International Committee for the Co-ordination and Standardisation of Speech Databases and Assessment Techniques (O-COCOSDA)*, pages 1–5.

Bo-Han Lu, Yi-Hsuan Lin, Annie Lee, and Richard Tzong-Han Tsai. 2024. Enhancing Taiwanese hokkien dual translation by exploring and standardizing of four writing systems. In *Proceedings of the 2024 Joint International Conference on Computational Linguistics, Language Resources and Evaluation (LREC-COLING 2024)*, pages 6077–6090, Torino, Italia. ELRA and ICCL.

Shravan Nayak, Surangika Ranathunga, Sarubi Thillainathan, Rikki Hung, Anthony Rinaldi, Yining Wang, Jonah Mackey, Andrew Ho, and En-Shiun Annie Lee. 2023. Leveraging auxiliary domain parallel data in intermediate task fine-tuning for low-resource translation. *Preprint*, arXiv:2306.01382.

Pei-Lin Liu and Chiu-Jung Chen. 2023. Using an AI-Based Object Detection Translation Application for English Vocabulary Learning. *Educational Technology & Society*, 26(3).

Surangika Ranathunga, En-Shiun Annie Lee, Marjana Prifti Skenduli, Ravi Shekhar, Mehreen Alam, and Rishemjit Kaur. 2023. Neural machine translation for low-resource languages: A survey. *ACM Comput. Surv.*, 55(11).

Jeng-Shin Sheu, Aftab Ahmad, and Jian-Min Li. 2024. Bilingual speech recognition for taiwanese hokkien and mandarin: Computational intelligence in complex acoustic environments. In *2024 International Conference on Consumer Electronics - Taiwan (ICCE-Taiwan)*, pages 215–216.

Mariana Shimabukuro, Jessica Zipf, Shawn Yama, and Christopher Collins. 2023. Evaluating classroom potential for Card-it: Digital flashcards for studying and learning Italian morphology. In *Proceedings of the 18th Workshop on Innovative Use of NLP for Building Educational Applications (BEA 2023)*, pages 130–136, Toronto, Canada. Association for Computational Linguistics.

Tong Su, Xin Peng, Sarubi Thillainathan, David Guzmán, Surangika Ranathunga, and En-Shiun Lee. 2024. Unlocking parameter-efficient fine-tuning for low-resource language translation. In *Findings of the Association for Computational Linguistics: NAACL 2024*, pages 4217–4225, Mexico City, Mexico. Association for Computational Linguistics.

Kohei Yamaguchi, Motoi Iwata, Andrew Vargo, and Koichi Kise. 2020. Mobile vocabulometer: a context-based learning mobile application to enhance English vocabulary acquisition. In *Adjunct Proceedings of the 2020 ACM International Joint Conference on Pervasive and Ubiquitous Computing and Proceedings of the 2020 ACM International Symposium on Wearable Computers*, pages 156–159, Virtual Event Mexico. ACM.

Tianyang Zhong, Zhenyuan Yang, Zhengliang Liu, Ruidong Zhang, Yiheng Liu, Haiyang Sun, Yi Pan, Yiwei Li, Yifan Zhou, Hanqi Jiang, Junhao Chen, and Tianming Liu. 2024. Opportunities and Challenges of Large Language Models for Low-Resource Languages in Humanities Research. *arXiv preprint*. ArXiv:2412.04497 [cs].

A Appendix

A.1 Database Design

The utilization of a database in our application serves a critical role in generating the application's flashcard user flow and enables the user experience to be seamless.

Our database aims to be user centric to ensure the best learning experience and outcome for users to learn their desired languages. The database is also designed to be optimal in retrieval and storage and scalable to provide future feature developments. The final database schemas contains four primary collections, `flashcard`, `flashcardList`, `flashcardQuiz`, and `category` with details shown in Table 2.

Collection	Field	Type	Description
Flashcard	`origin`	String	Hokkien word in Hanji.
	`destination`	String	English translation of the word.
	`otherOptions`	List[String]	Three incorrect English choices.
	`type`	String	Type of card (word, sentence, image, speech).
	`audioUrl`	String	URL of pre-generated audio of the `origin`.
	`romanization`	String	Pre-generated romanization of `origin` in numeric tones.
	`categoryId`	String	Reference to ID in Categories collection.
	`createdBy`	String	Reference to a user (email).
	`createdAt`	String	UTC Timestamp (ISO 8601 format).
FlashcardList	`name`	String	Flashcard list/deck name.
	`cardList`	List[String]	List of Flashcard IDs from Flashcard collection.
	`categoryId`	String	Reference to ID in Categories collection.
	`shared`	Boolean	True if shared, False if user-specific.
	`createdBy`	String	Reference to a user (email).
	`createdAt`	String	UTC Timestamp (ISO 8601 format).
FlashcardQuiz	`flashcardListId`	String	Reference to ID in FlashcardList collection.
	`scores`	Dict[String, List[Dict]]	Maps users to quiz scores and timestamps.
	`flashcardScores`	Dict[String, Dict]	Maps user IDs to flashcard correctness.
Categories	`name`	String	Category name.
	`description`	Optional[String]	Category description.
	`createdAt`	String	UTC Timestamp (ISO 8601 format).
	`flashcardList`	List[String]	List of FlashcardList IDs.

Table 2: Database Schema for the Flashcard feature with collection type, data structure type, and each name of the fields with descriptions.

A.2 Preliminary Study Material and Participant Data

Our three week long preliminary study included an entry survey that participants filled out when they entered the study, an usability survey that participants filled out at the end of each study week, and an exit survey that participants filled out at the conclusion of the study (end of Week 3).

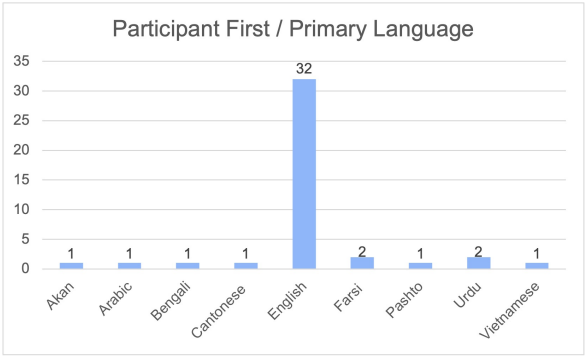

Figure 7: First languages spoken by participants in the user study, collected through the entry survey.

Figure 8: Other languages that participants in the user study have knowledge of, collected through the entry survey.

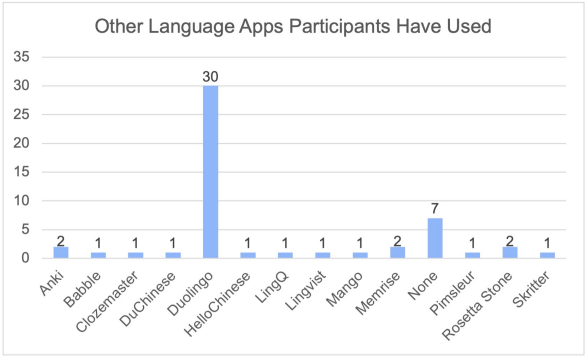

Figure 9: Other language learning apps that participants in the user study have used in past, collected through the entry survey.

Entry Survey	
Question	**Response type**
Please list your native/primary language.	Short Answer
Please list any other languages you speak and your proficiency from 1 to 5 (1 being basic sentences and understanding and 5 being complete fluency).	Short Answer
Have you used any language-learning apps before? If yes, which ones?	Short Answer

Table 3: Entry Survey completed by participants when beginning the study.

Usage Frequency	
Question	**Response type**
How many times did you use the app in the past week?	Short Answer
If you did not use it (0 times) in the past week, why did you not use it? Choose the reason that best applies to you.	Multiple Choice Options: a. I was not able to due to time or personal reasons b. I did not remember to use it c. I had technical issues or computer-related problems d. The app was too difficult to use e. I did not want to use it f. I did not feel the need to use it g. Other (Short Answer): _____

Flashcard System	
Question	**Response type**
The flashcard system was intuitive and easy to use.	5-Point Likert Scale (“strongly agree”, “agree”, “neutral”, “disagree”, or “strongly disagree”)
The feedback provided while using flashcard system was helpful.	5-Point Likert Scale (“strongly agree”, “agree”, “neutral”, “disagree”, or “strongly disagree”)

Dictionary System	
Question	**Response type**
The generated images for the translations on the home page are relevant.	5-Point Likert Scale (“strongly agree”, “agree”, “neutral”, “disagree”, or “strongly disagree”)
The image generation on the home page helps with understanding the translated words/sentences.	5-Point Likert Scale (“strongly agree”, “agree”, “neutral”, “disagree”, or “strongly disagree”)
The text-to-speech on the home page helps with understanding the translated words/sentences.	5-Point Likert Scale (“strongly agree”, “agree”, “neutral”, “disagree”, or “strongly disagree”)

Platform User Experience	
Question	**Response type**
Overall, the platform is easy to use.	5-Point Likert Scale (“strongly agree”, “agree”, “neutral”, “disagree”, or “strongly disagree”)
The platform has been helpful for learning Hokkien.	5-Point Likert Scale (“strongly agree”, “agree”, “neutral”, “disagree”, or “strongly disagree”)
The layout of the app was intuitive.	5-Point Likert Scale (“strongly agree”, “agree”, “neutral”, “disagree”, or “strongly disagree”)
It was easy to interact with the app’s features.	5-Point Likert Scale (“strongly agree”, “agree”, “neutral”, “disagree”, or “strongly disagree”)
Did you encounter any obstacles when trying to perform tasks (e.g., taking quizzes, reviewing flashcards)?	Short Answer
Do you have any general comments or questions about your experience using the app?	Short Answer

Table 4: Usability Survey completed by participants at the end of each study week.

Exit Survey	
Question	**Response type**
How would you rate your overall experience with the app?	5-Point Quality Scale (“very poor”, “poor”, “average”, “good”, or “excellent”)
How satisfied are you with the progress you made in learning Hokkien??	5-Point Satisfaction Scale (“very satisfied”, “satisfied”, “neutral”, “unsatisfied”, or “very unsatisfied”)
Would you recommend this app to others interested in learning Hokkien? Why or why not?	Short Answer
What aspects of the app did you like the most?	Short Answer
What aspects of the app could be improved or changed?	Short Answer
Do you plan to continue using the app after the study?	Multiple Choice ("Yes", "No", "Maybe", "Other" (short answer))
Do you have any final comments or suggestions for the research team?	Short Answer

Table 5: Exit Survey completed by participants at the end of the study (end of Week 3).