Pronunciation Scaffolder 3.0: A User Experience and Usability Study

Nicholas Carr*†, John Blake*
*The University of Aizu, Aizu-Wakamatsu, Japan
†Email: carrnick@u-aizu.ac.jp

***Abstract*—This paper presents the results of a user experience and usability study on version 3.0 of the *Pronunciation Scaffolder*, a web application designed to help Japanese computer science majors read presentation scripts aloud. The web app annotates scripts using symbols and colours, indicating pauses, intonation, rhythm, word stress and so forth. Through observation, and thematic analysis of in-depth interviews, we gained valuable insights into how the tool can be refined to improve both the user experience and the user interface.**

***Index Terms*—pronunciation, reading aloud, text annotation, web application**

I. Introduction

Intelligent computer-assisted language learning (iCALL) applications draw on natural language processing to produce powerful tools that are able to tailor their output to help individual language learners [1]. Developers may base their tools on firm theoretical foundations and create practical pedagogic tools. However, problems may arise from the mismatch of expectations, resulting in dissatisfied users. One way to identify potential issues is to conduct usability tests.

This paper reports the results of usability testing conducted on a web application developed to help Japanese computer science majors read their presentation scripts more easily and appropriately. Scripts are submitted and annotated with symbols and colour coding. Users can select the aspect of pronunciation that they want to be annotated, such as pausing and intonation. Annotations are designed to be intuitive, but for some aspects, learners need to refer to a key to understand the symbols and colorization. Learners are expected to be familiar with aspects of pronunciation that are taught in Japanese high schools, such as intonation and word stress, which are covered in detail in government-approved textbooks [2]. Learners may, however, be unfamiliar with aspects of pronunciation, such as the impact of content and grammar words on rhythm, and linkages occurring at the junctions between words [3].

Although the annotation accuracy of the Pronunciation Scaffolder has already been established [4], the user experience (UX) and the usability of the user interface (UI) has not been formally evaluated. Version 3.0 was written in Elm language, which compiles to JavaScript, the advantage of which is zero runtime errors; but the downside is the lack of student developers willing to improve the codebase. Therefore, to ensure a steady supply of developers; Version 4.0 will be written in Python, one of the most popular programming languages. Given that the codebase will be completely rewritten, this is an opportune moment to identify aspects of UI and UX that could be improved. By identifying how users actually use the web app, and how they feel about it, opportunities for improvement will be revealed, which will inform the development of Version 4.0 of the Pronunciation Scaffolder. Online tools that are user-friendly and fulfil user expectations are likely to generate positive word of mouth and lead to increases in the user base, helping more learners of English deliver presentations with more appropriate pronunciation.

The remainder of the paper is organized as follows. Section II describes the extant literature on usability testing, contextualizing this research. The Pronunciation Scaffolder is introduced in Section III. Section IV describes the method used. The results are given in Section V. Section VI concludes the paper with an outline of future work to be undertaken to solve or ameliorate the usability issues discovered.

II. Usability testing

Accuracy, efficiency and satisfaction [5] are main factors that impact usability. Usability for pedagogic software applications focuses on the ease with which learners are able to achieve their goals using the target application while user experience is a more holistic metric encompassing the user pathway and user feelings. The user pathway [6] refers to the route that users select when navigating. On a simple website, this may be reflected in the breadcrumb trial. The duration and sequence of user actions taken to achieve a particular goal helps developers understand the user experience.

Usability evaluation methods may be categorized into four broad groups: user testing methods (e.g. think-aloud protocol), inspection methods (e.g. cognitive walkthrough), inquiry methods (e.g. focus groups) and analytical methods (e.g. task environment analysis) [7]. Usability [8] comprises a number of facets, such as effectiveness and efficiency, consistency, design and layout. Usability testing may be categorised based on moderation, proximity and purpose. Table I shows two variables for each of the three parameters, creating six potential combinations.

TABLE I
Three parameters of usability testing

Moderation	Proximity	Purpose
Moderated	In person	Explorative
Unmoderated	Remote	Comparative

979-8-3503-2422-8/23/$31.00 ©2023 IEEE
DOI 10.1109/IIAI-AAI59060.2023.00016

Unmoderated remote usability testing is the most scalable option since no budget is needed to remunerate moderators while moderated in person usabiilty testing can potentially harness more data. As the presence of a moderator may influence behaviour, moderators may need to take specific actions to ameliorate this effect. The choice between explorative and comparative approaches centres on whether potential choices have been identified. For example, when deciding whether to release a mobile app or a web app, a comparative approach makes sense. However, adopting a more grounded approach to usability testing [9] provides developers with the opportunity to discover hitherto unknown issues.

III. Pronunciation Scaffolder

The first release of this web app was developed specifically to annotate presentation scripts written by Japanese computer science majors who needed to present their capstone project report in English as a graduation requirement. Students tended to deliver their presentation in a monotonic speech with little use of pausing, intonation, and rhythm. To help them read aloud more appropriately, the Pronunciation Scaffolder annotates the script using a combination of colours and symbols. Students are recommended to check their scripts using online error checkers, including Grammarly and a specialist error checker for computer science [10] to ensure the accuracy and appropriacy of their scripts prior to using the Pronunciation Scaffolder.

Users are able to select one or more pronunciation features to be annotated. The current release, version 3.0 allows users to annotate nine pronunciation features, which are shown in Table II. The first four functions are the core functions, which based on pilot studies make the most noticeable impact on the intelligibility and appropriacy of the delivery of the presentations given by Japanese learners of English. The three sound-focused functions help readers who may be confused about voiced and unvoiced sounds. The final two functions are designed for advanced learners who want to better understand the pronunciation of connected speech by focusing on the specific sounds that are added, omitted or altered at word junctures.

TABLE II
Nine aspects of pronunciation

Function	Components
`Pausing`	using short, medium and long pauses
`Intonation`	using falling and rising intonation
`Content Word`	emphasizing content words
`Word Stress`	placing stress on the appropriate syllable
`ed` `Sound`	pronouncing -ed suffix appropriately
`th` `Sound`	pronouncing voiced and voiceless sounds
`s` `Sound`	pronouncing -s suffix appropriately
`Consonant Links`	using elision and linking
`Vowel Links`	inserting linking sounds

Figure 1 shows a screenshot of the viewport of the Pronunciation Scaffolder 3.0 on a mobile device.

Table III shows the different version releases of the Pronunciation Scaffolder to date. Version 1.0 was released under the name *Script Annotator*. This version, which was released in 2017, was only able to show the annotation of one function at a time.

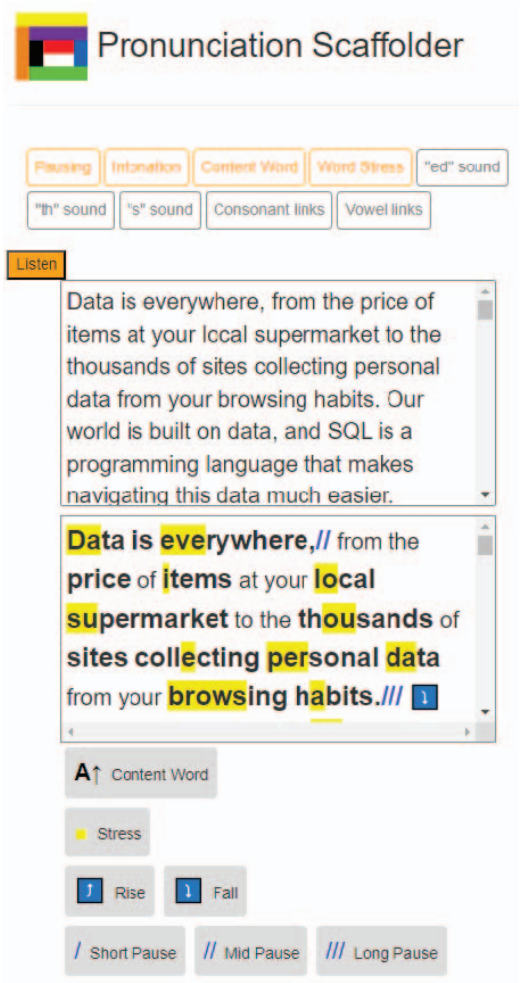

Fig. 1. Screenshot of Pronunciation Scaffolder 3.0

TABLE III
Different versions of the Pronunciation Scaffolder

Version	Description
Ver 1.0	Displayed functions separately
Ver 1.1	Displayed functions simultaneously
Ver 2.0	Improved accuracy of functions
Ver 3.0	Replaced `Word Stress` function

Based on user demand, in the next release (Ver 1.1) users could highlight any combination of functions simultaneously. Given the complexity of displaying nine levels of annotation at the same time, it was expected that users would only highlight two or three functions simultaneously. The name was changed to *Pronunciation Scaffolder* from Version 2.0 to make it easier to find using Google and other search engines. In Version 3.0 a new `Word Stress` function was integrated that is able

to show the primary stress for approximately $65,000$ words. Secondary stress in polysyllabic words was ignored to reduce the cognitive load. This function works by matching words in a tailormade dictionary, which was created by synchronizing words that occurred in both a syllable dictionary and the pronouncing dictionary.

IV. Method

Usability studies typically use small samples [11], with effective usability studies normally utilizing sample sizes of between five to eight participants [12]. Such sample sizes are considered valid due to 85% of problems being detected by five users, especially amongst homogeneous users [13]. A call for participation was circulated at a Japanese university amongst first-year computer science majors. As an incentive, participants would receive an Amazon gift card. Initially, 11 students registered interest to participate; six were successfully scheduled and tested. All participants had similar educational backgrounds: completing all their formal education in Japan and undertaking this study in the final quarter of their first year of study.

As previously noted, user experience with pedagogic software applications encompasses user feelings. Accordingly, in this study, usability is concerned with users' perception of the interface; measures which are subjective [14]. In line with this study's explorative purpose and in order to capture rich aspects of subjective measures of usability, a qualitative approach was adopted.

Data was collected from two sources: screen recordings of participants using the tool in a laboratory setting, and a semi-structured interview which immediately followed the test. A moderated laboratory setting was utilized for three reasons: (1) it allowed for control over the test [15], (2) ease of conducting the semi-structured interview [16], (3) it allowed for the inclusion of ethnographic observations to be discussed during the interview. The protocol used for testing is given in Table IV.

Testing was divided into three phrases, which were conducted as follows: (1) Preparation: purpose of study was explained, followed by an introduction on how to use the tool. Identical text was input into the tool for all participants. (2) Usability test: tool was used for between 18 to 20 minutes. Exact times differed due to the researcher not stopping participants mid-way through the script. Screen recordings were collected, including participant video and audio (see Fig. 2). The first author was present. (3) Semi-structured interview: immediately following test, interview was conducted. Participants responded in both English and Japanese to ensure their perceptions were accurately captured.

Interview transcripts were imported into NVivo 14 for coding. Participant experiences described in the interviews were triangulated with screen recordings. For example, statements on the usage frequency of certain functions and their influence on behavior were checked with screen recordings to ensure interview data was valid. Coding was inductive, allowing themes to emerge naturally rather than imposing a preset framework through which to analyse the data.

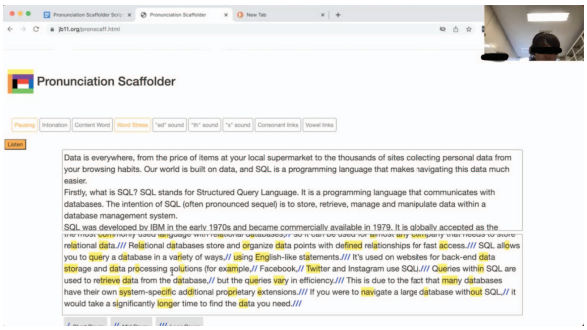

Fig. 2. Screen Recording Example

TABLE IV
Usability Test Protocol

Phase	Steps
Preparation	study explained; explain tool; identical text input for all participants.
Test	use tool for 18 to 20 minutes; collect screen recordings (including participant video and audio) (see Fig. 2). One researcher was present.
Interview	conduct semi-structured interview immediately after test.

V. Results

The themes emerging from the data are as follows: graphical user interface, usage patterns, satisfaction, improvements, and bugs. Typically, measures of usability are categorised as one of the three aspects of usability as outlined by ISO 9241:11-2018: effectiveness, efficiency, and satisfaction [17]. Effectiveness refers to the accuracy and completeness of users realizing specific goals; efficiency is concerned with the resources users use to achieve goals; satisfaction refers to positive attitudes towards users of the tool. However, the themes emerging from this data-set were not always able to be cleanly ascribed to one of these three aspects of usability; themes often related to all three. Therefore, results present each emerging theme, followed by a discussion of its relationship with the three aspects of usability. Given the uniqueness and niche for this tool, comparison to other software is not included.

A. Graphical User Interface

User experiences of the graphical user interface (GUI) varied according to the function(s) being used. The `Intonation`, `Content Word`, and `Word Stress` functions were described as being intuitive and easy to read within the output box. Thus, their annotations, rarely, if at all, caused participants to refer to the explanation of the annotations located below the output box. This resulted in satisfactory GUI experience for these functions.

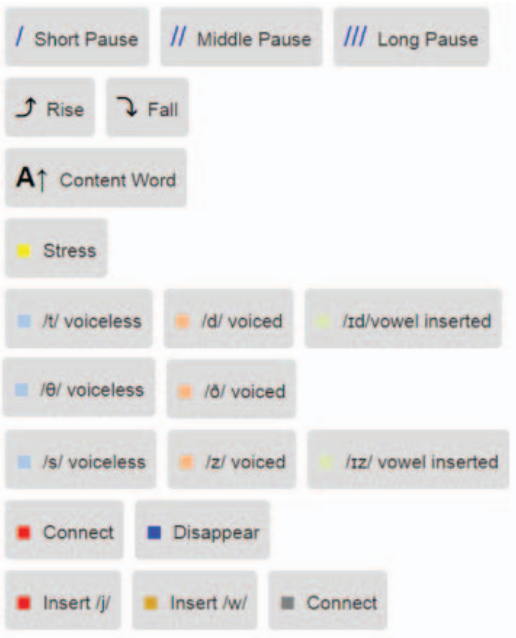

Fig. 3. Key for annotations displayed below output

For the more intricate functions, such as `Consonant Links` and `Vowel Links`, however, users found the GUI less user friendly. User comments regarding this included:

> "I have to check these one by one." (Participant D)
> "I have to go down and check." (Participant F)

When users use terms such as "check" and "go down", they are referring to the processes of changing visual focus from the annotated text to the key underneath the text, where explanations of symbols of the functions currently in use are displayed. Fig. 1 shows the location of these explanations. The explanations for all functions are shown in Fig. 3. These comments indicate that the more intricate functions required more effort due to the eye movement required to check the key for the meaning of annotation symbols.

A final point on GUI experience is Participant B noted that he felt he required the assistance of several functions due to a lack of proficiency in English. Participant B commented that individually each function's annotations are relatively easy to read. However, when using multiple functions, he found the annotated text in the output box too complicated to read in an efficient manner. Accordingly, he did not use more than four of the nine available functions at any one time.

The issues related to GUI mainly affected the efficiency of the tool. The more intricate functions displayed annotations which required additional effort to use due to the eye movement required to find the explanation. In the case of a user feeling that only four functions could be used simultaneously, the tool's effectiveness was reduced due to the user not utilizing all of the assistance available.

B. Usage Patterns

Overall, a strong positive correlation was found between the percentage of time for which functions were used during the test and the perceived usefulness of the functions. While usage patterns varied accordingly to each participant, Table V shows an overall trend for four of the nine functions being used with higher frequency than the remaining five.

TABLE V
PERCENTAGE OF TEST TIME THAT FUNCTIONS WERE UTILIZED BY PARTICIPANTS A TO F[a]

Function	A	B	C	D	E	F
`Pausing`	98	95	68	89	99	30
`Intonation`	98	74	97	71	25	96
`Content Word`	45	53	93	33	20	94
`Word Stress`	97	73	75	92	98	50
`ed` `Sound`	0	0.6	1	5	0	0
`th` `Sound`	0	0.6	0	2	0	0
`s` `Sound`	0	0.6	30	6	0	0
`Consonant Links`	0	13	27	52	68	95
`Vowel Links`	0	0.1	0	0.7	6	35

[a] Given that the duration of each test slightly differed, percentage of total time rather than exact duration is used.

The most frequently used functions were: `Word Stress`, `Pausing`, `Intonation`, and `Content Word`. One possible reason for this trend is, as previously noted, the annotations these functions output were described as intuitive, and thus easy to read and understand. These functions were also most commonly described as being useful by all participants. Another possible reason for this trend is that participants educational background resulted in these aspects of pronunciation being features they are most familiar with. The high usage of these functions and screen recordings showing the incorporation of these annotations into practising reading the script positively affected effectiveness by increasing the accuracy and completeness of users achieving their goal of giving a presentation in English.

Usage patterns varied for `Consonant Links`, with Participant A not using the function at all and Participant F using it for the majority of the test. Participant descriptions of `Consonant Links` include:

> " I was very happy that this (`Consonant Links`) is included...to be honest, I often hear speech connected like this when talking to proficient English speakers." (Participant B)

> " `Consonant Links` was one of the especially useful function." (Participant D)

> "this connected [sic] consonant links, the links thing that I never (consciously) paid attention in my life." (Participant F)

Despite using the function for only 13% of test time, Participant B described the tool as an important aspect of pronunciation during his interview. However, as previously

noted, Participant B found using more than four functions simultaneously difficult. Accordingly, his infrequent usage was likely due to readability rather than not valuing the function. Participants C and D both described `Consonant Links` as useful, an opinion which is evidenced in their moderate usage of the function. Participant C explained her moderate usage did not reflect the value she placed on the function. Due to her lack of knowledge on how to connect speech, she limited its usage. Despite Participant E using the function for 68% of test time, he did not identify the function as being useful. This contradiction indicates Participant E turned the function on but did not pay much attention to linking consonants; an interpretation evidenced in his screen recording. Participant F used the function for 95% of test time and displayed a strong knowledge of connected speech during his interview. The importance Participant F placed on connected speech is further evidenced in using `Vowel Links` for 35% of test time; significantly higher than all other participants.

As shown in Table V, with the exception of Participant C, the `"ed" Sound`, `"th" Sound`, and `"s" Sound` functions were rarely used. Three participants did not use these functions at all. Excluding Participant C's usage of `"s" Sound`, these functions tended to be turned on briefly and turned off without participants actually using the annotations to improve their production of the script. The tendency of users not realizing the benefits of these functions impacts the effectiveness of the tool due to a reduction in users' ability to accurately and completely achieve their goals.

Three participants sought assistance from an online dictionary. One participant used a dictionary for 18% of test time, while the other two participants used a dictionary for less than 10% of test time. All three participants used a dictionary to check the pronunciation of unknown words. This check was performed by looking up the word in an online dictionary and using the dictionary's text-to-speech (TTS) function. It may seem that this issue arose due to the script being prepared for users rather than written by users themselves as would occur in a natural environment. However, participants noted that the same issue occurs during regular classes due to machine translation (MT) being commonly used to assist them to prepare presentations; i.e. the MT output unknown vocabulary. The need to seek assistance external to the tool reduces efficiency due to increasing the effort needed to achieve user goals.

C. Satisfaction and Improvements

All users expressed their overall experiences with the tool in positive terms. Some comments made by Participant F are representative of all users when he stated "by seeing this (the output box) it illustrates how it (the presentation) should be". Additionally, all participants indicated that through using the tool, they were able to pay attention to aspects of pronunciation which had been previously unknown to them. The unknown aspects naturally differed according to each participant, but were all related to the tool's functions. The most frequently mentioned aspects of pronunciation users became more cognizant of while reading the text are: pausing, word stress, and consonant links. User accounts indicate the effectiveness of the tool to be high because it was perceived as being able to facilitate assistance for users to achieve their goal of making a presentation in English.

Despite the overall satisfaction of the tool, users provide useful insights into suggestions for further improvements to the tool to increase usability. The improvements suggested fell into three categories: understanding annotations, editing current functions, and additional functions. Despite `Pausing` being one of the most frequently used functions, users identified a need for more information on how to discern the length of the annotated pauses. Currently, the length of pauses is indicated as follows: one, two, or three forward slashes, with additional slashes indicating a longer pause (see Fig. 1).

Users indicated that they felt they needed a more objective measure on how long each of these types of pauses should be. However, research on pausing [18] shows that the duration of any pause is dependent on multiple factors related to both the speaker and the text, including rate of speaking, location of pause, i.e. intra- or inter-sentential, and the number of words before and after intra-sentential pauses. It is possible to put a rule-of-thumb (although rather arbitrary) guide of 500 milliseconds per backslash and so the recommended length of a short pause would be 500 milliseconds while a long pause would be 1500 milliseconds.

Additionally, learners noted that additional explanations on how to interpret the `Word Stress` and `Content Word` functions would be useful. Participant C best summarized this suggestion by noting that with these two functions, there are two kinds of emphasis the speaker needs to be cognizant of. However, there is no assistance for the user to understand how the emphasis differs for syllabic stress within a word and the stress placed on content words within a sentence.

Participants identified three suggestions to improve the current functions of the tool. Currently, the `Intonation` function only indicates a rise or fall at the end of a clause. It was suggested that the `Intonation` function would be improved if it displayed intonation throughout the whole sentence rather than only at the end of a clause. A second suggestion was an expansion of the `Pausing` function. Currently, the function indicates the insertion of a pause only after a comma or full stop. As noted by a participant, clearly spoken text often requires pauses in places where there is no comma, especially when dealing with long clauses. A final suggestion made by many participants was related to the issues identified with the tool's GUI; the `Consonant Links` and `Vowel Links` functions would improve if more intuitive symbols were used to annotate a text.

An additional function participants suggested be added to the tool was the ability to select specific sections of the text for TTS. Currently, the tool has a TTS function which allows users to listen to the script as a whole, but not specific sections. Nevertheless, participants indicated they felt the tool would benefit from having TTS for two purposes. The first was to confirm the pronunciation of individual words in

isolation. This function may seem unnecessary at first blush, given that in natural settings participants would input the text they write. However, interviews revealed participants typically created scripts for the presentations completed in class by using MT; resulting in unknown words being used in their presentations. The second TTS function suggested for the tool was TTS which could highlight the consonant and vowel link annotations. In other words, the ability to hear how sounds are joined or omitted when using the `Consonant Links` and `Vowel Links` functions.

D. Software bugs

Testing revealed three bugs which resulted in inaccurate annotations being displayed. The first of these occurs with the word Microsoft when the Word Stress function is used. The `Word Stress` function highlights the stressed syllable of multi-syllabic words in yellow. Rather than highlight the primary stress on the first syllable, a coding error caused the word *Microsoft* to be displayed as `Mound: #FF0">icrosoft.`

The second issue users found was the use of the `Intonation` function. At the end of one sentence in the script, the intonation arrows point downwards, correctly indicating that intonation should fall at the end of the sentence. However, the downward arrow is immediately followed by an upward arrow, indicating intonation should rise. Further investigation revealed this occurs whenever a sentence begins with the conjunctions *and*, *but*, or *so*.

Finally, one participant expressed their inability to understand why the annotated text displayed the instruction to elide the final consonant "t" in the verbal phrase *built on*. Further investigation revealed a coding error causing annotations to display the symbol indicating elision rather than connection for words with a final consonant "t" and followed by words beginning with the vowels "o" or "i".

While these bugs did not significantly reduce satisfaction with the tool, they caused confusion, increasing mental effort, and thus reducing efficiency. In the event incorrect annotations are acted upon, effectiveness is reduced due to incorrect pronunciation of sections of the script.

VI. Conclusion

We conducted a laboratory study to explore the usability of Pronunciation Scaffolder 3.0 with Japanese university students majoring in computer science. Results show participants were satisfied with the tool. Its effectiveness was high, with all participants describing it as significantly enhancing their preparation for reading a script.

To enhance the user experience, in addition to solving the reported software bugs, four areas of improvement were identified. First, more detailed explanations of annotations and a change in their positioning was noted. Rather than explanations being located below the output box, through the use of an event manager, textual explanations of the functionality of each button could be displayed on hover. This explanation could incorporate a hyperlink to a short explanatory video for users who would like more detailed information. Second, participants suggested extending the `Intonation` function to indicate intonation patterns for complete sentences rather than phrase endings only. The caveat with this suggestion, however, is the significant increase in the complexity of the annotation. Third, the need for a more objective measure for the length of pausing was identified. Finally, the provision of TTS functionality to allow users to listen to sections of text (e.g. word junctures and unknown vocabulary) would enhance usability.

References

[1] J. Blake, "Intelligent call: Using pattern matching to learn english," in *New Technological Applications for Foreign and Second Language Learning and Teaching*, M. Kruk and M. Peterson, Eds. IGI Global, 2020, pp. 1–23.

[2] J. Sugimoto and Y. Uchida, "How pronunciation is taught in English textbooks published in Japan," *Seishin Studies*, vol. 130, pp. 3–35, 2018.

[3] M. Maeda, "Some problems with word-linking in the English of Japanese learners," *Gengo Kenkyu*, vol. 33, pp. 1–20, no date.

[4] T. Kondo, J. Inoue, and J. Blake, "Pronunciation scaffolder: Annotation accuracy," in *Proceedings of the International Symposium on Applied Phonetics*, 2018, pp. 84–87.

[5] J. Casas, M.-O. Tricot, O. Abou Khaled, E. Mugellini, and P. Cudré-Mauroux, "Trends & methods in chatbot evaluation," in *Companion Publication of the 2020 International Conference on Multimodal Interaction*, 2020, pp. 280–286.

[6] M. Hall, A. Mazarakis, M. Chorley, and S. Caton, "Editorial of the special issue on following user pathways: Key contributions and future directions in cross-platform social media research," *International Journal of Human–Computer Interaction*, vol. 34, no. 10, pp. 895–912, 2018.

[7] P. Weichbroth, "Usability of mobile applications: a systematic literature study," *Ieee Access*, vol. 8, pp. 55 563–55 577, 2020.

[8] C. M. Barnum, *Usability testing essentials: Ready, set... test!* Morgan Kaufmann, 2020.

[9] Z. Hussain, W. Slany, and A. Holzinger, "Investigating agile user-centered design in practice: a grounded theory perspective," in *HCI and Usability for e-Inclusion: 5th Symposium of the Workgroup Human-Computer Interaction and Usability Engineering of the Austrian Computer Society*. Springer, 2009, pp. 279–289.

[10] J. Blake, "Genre-specific error detection with multimodal feedback," *RELC Journal*, vol. 51, no. 1, pp. 179–187, 2020.

[11] Q. Wei, Z. Chang, and Q. Cheng, "Usability study of the mobile library app: an example from chongqing university," *Library Hi Tech*, vol. 33, no. 3, pp. 340–355, 2015.

[12] K. D. Pendell and M. S. Bowman, "Usability study of a library's mobile website: An example from portland state university," *Information technology and libraries*, vol. 31, no. 2, pp. 45–62, 2012.

[13] J. Nielsen and T. K. Landauer, "A mathematical model of the finding of usability problems," in *Proceedings of the INTERACT'93 and CHI'93 conference on Human factors in computing systems*, 1993, pp. 206–213.

[14] K. Hornbæk, "Current practice in measuring usability: Challenges to usability studies and research," *International journal of human-computer studies*, vol. 64, no. 2, pp. 79–102, 2006.

[15] D. Zhang and B. Adipat, "Challenges, methodologies, and issues in the usability testing of mobile applications," *International journal of human-computer interaction*, vol. 18, no. 3, pp. 293–308, 2005.

[16] A. Naiakshina, A. Danilova, C. Tiefenau, M. Herzog, S. Dechand, and M. Smith, "Why do developers get password storage wrong? a qualitative usability study," in *Proceedings of the 2017 ACM SIGSAC Conference on Computer and Communications Security*, 2017, pp. 311–328.

[17] International Organization for Standardization, "Ergonomics of human-system interaction — part 11: Usability: Definitions and concepts (ISO 9241:11-2018)," 2018.

[18] P. Šturm and J. Volín, "Occurrence and duration of pauses in relation to speech tempo and structural organization in two speech genres," *Languages*, vol. 8, no. 1, p. 23, 2023.