A Guide to Using a Mobile Application to Track Spoken Fluency

Heather Woodward

Joshua Rappeneker

Keiko Fujino

Abstract

In previous Journal of Foreign Language Education and Research (JFLER) articles, we wrote about designing a mobile application to measure students' monologic spoken fluency in classroom settings (Woodward & Gupta, 2022; Woodward & Rappeneker, 2023). Currently, instructors of Rikkyo University's English Discussion Class (EDC) do not have an objective and quantitative method for monitoring students' English spoken fluency development across the semester. Yet, the language-based goal of EDC is to develop it (Hurling, 2012). The purpose of this article is to provide an instructional guide on how to use the mobile application. After testing the app, it has undergone many modifications since 2022. Therefore, in addition to how to use the application, we discuss the features that have worked, the ones that have not, and the ones that we plan to add. We provide support for app-based choices with theory and research into the field of spoken fluency development. With this paper, we hope that the instructors who wish to incorporate the mobile app will be able to do so.

Keywords: *CALL, MALL, ESL, Speaking, Monologues*

Introduction

Assessing spoken fluency in second language (L2) speakers is recognized as a challenging task, acknowledged by many evaluators (Kang et al., 2019). Tavakoli (2016) and Tavakoli and Wright (2020) describe spoken fluency as a multifaceted construct influenced by various factors, including psycholinguistic aspects like anxiety and motivation, linguistic factors like L2 vocabulary and first language fluency, and sociolinguistic factors such as L2 exposure and group dynamics. Factors such as studying abroad, rehearsal time, task type, and topic familiarity can also impact spoken fluency (Wright, 2020).

These elements make spoken fluency difficult when assessing to generalize results. The language-based goal is to enhance L2 spoken fluency in English discussion class (Hurling,

2012), yet instructors currently lack a method to track students' progress in English spoken fluency throughout the semester. To tackle this challenge, we created a mobile application specifically tailored to assess students' monologic spoken fluency on topics familiar to them. Given the complex nature of spoken fluency, we have not determined the extent to which students' spoken fluency results are generalizable to other task types or to unfamiliar topics, so the results are best interpreted narrowly in regards to students' monologic speeches on familiar topics. According to Segalowitz and Freed (2004), spoken fluency can be improved significantly in non-immersion environments. Thus, despite its limitations, the app might be useful as a pedagogical tool for students to track their progress throughout the semester and to acquire greater metalinguistic knowledge of spoken fluency.

Literature Review

Spoken Fluency

Spoken fluency is narrowly defined as the effortless and continuous quality of speech, characterized by traits like fluidity and smoothness (Segalowitz, 2010; Tavakoli, 2016). Segalowitz (2010) delineates a multidimensional model of spoken fluency that comprises three interrelated components: cognitive, perceived, and utterance fluency. Cognitive fluency pertains to the mental processes involved in generating L2 speech, encompassing the time elapsed from conceptualization to articulation. This aspect is operationalized through reaction time measurements (Segalowitz & Segalowitz, 1993). Perceived fluency involves assessments of fluency made by listeners (Segalowitz, 2010; Segalowitz, 2016). Lastly, according to this model, utterance fluency entails observable, surface-level features of speed, breakdowns, and repairs. Speed is the rate of speech production; breakdown denotes a complete or partial interruption in the flow of speech; and repair occurs when speakers rectify mistakes or omissions in their intended message, often through repetition or self-correction. Utterance features of spoken fluency, along with listeners' perceived fluency judgments, can provide ways to measure dimensions of spoken fluency. One such utterance feature is untrimmed speech rate, a composite measure of fluency encompassing speed, breakdowns, and repairs, without the deduction of any disfluencies from the final calculation. According to Suzuki et al. (2021), composite measures such as speech rate are highly correlated with listener-based perceived spoken fluency.

Benefits of developing L2 spoken fluency include: a) intelligibility; b) effective communication in real-life situations; and c) fulfilling future aspirations. To begin with, a) intelligibility pertains to how easily listeners can understand the meaning of a spoken message. When speakers talk too slowly, it can hinder the comprehension and intelligibility of their message, as listeners may struggle to maintain attention (Lennon, 2000). In terms of b) effective communication in real-life scenarios, speakers who converse too slowly might encounter challenges in saving face (Lennon, 2000). While factors like vocabulary, complexity, and

accuracy contribute only minimally to human ratings of spoken performance, spoken fluency significantly influences scores on standardized speaking assessments (Ogawa, 2022). With regards to c) future objectives, standardized tests like IELTS can act as gatekeepers, facilitating learners in realizing their ambitions of studying or working abroad. While benefits to cultivating spoken fluency exist, activities crafted to enhance this skill are frequently overlooked in the classroom (Rossiter et al., 2010; Tavakoli & Hunter, 2018). In instances where educators do focus on speaking skills, they typically involve students in speaking exercises without introducing them to the particular metalinguistic aspects related to repair, breakdown, and speed (Tavakoli & Hunter, 2018).

3/2/1 Activity

According to Nation and Newton (2009), effective spoken fluency activities exhibit three characteristics: they (a) provide support to enhance students' performance, (b) encourage the use of already known language items, and (c) prioritize the emphasis on meaning over form. To meet the first requirement, time-constrained activities can be employed, where instructors set specific time limits for completing a task. For the second requirement, choose familiar topics, such as friends, family, or school life, and encourage students to incorporate their existing L2 knowledge. Finally, to fulfill the third requirement, students can be paired with a new listening partner each round to create a knowledge gap. The 4/3/2 activity introduced by Maurice (1983) aligns with all requirements. Speakers discuss a familiar topic for four minutes. After that, the speakers switch to new listening partners and repeat their monologues within three minutes. This process is repeated with speakers having only two minutes to repeat their monologues. The belief is that the decreasing time constraint and repeated practice helps speakers to increase their spoken fluency (Nation & Newton, 2009). For time-constraints or students with lower proficiency, the time can be modified to 3/2/1.

Mobile Application

Students can download the mobile application on their phones and use it both in and outside of the class. During EDC, while they participate in a modified 4/3/2 activity called 3/2/1, they can use the app. Figure 1 below shows the first few screens.

Figure 1

*Pages of the Mobile Application (From Rappeneker, 2023)*
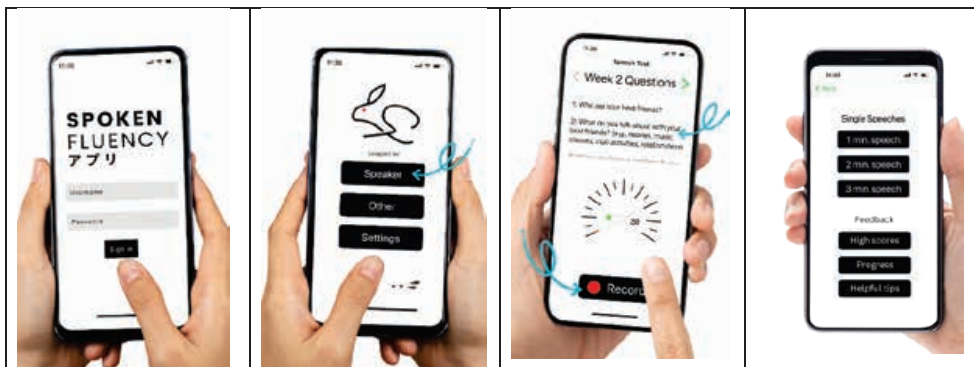
Upon logging into the app (first picture from the left), users see 'speaker,' 'other,' and 'settings' (second picture). If users click on 'speaker,' the application moves to the weekly questions page (third picture). After selecting 'other,' users can view the screen on the far right (fourth picture) the single speech sections, high scores, and graphs (fourth picture). On the third picture to the left, students are provided with a list of questions, derived from Sturges et al.'s (2023) "What's Your Opinion?" When students are using the mobile application, having these questions on the screen is convenient. At the bottom of the screen, students press the record button and respond to the questions. The speedometer represents the speakers' untrimmed speech rate (i.e., average syllables per minute). After students press the record button, a timer appears. The timer displays the seconds remaining before the recording ends. 180 seconds appears in the first round, 120 in the second round, and 60 in the third round. After the timer ends, the phone vibrates to alert speakers. Speakers see a page which instructs them to hand their phones to their listening partners. Once listeners have the speakers' phones, their partner rates spoken fluency.

Figure 2 below has screenshots of the listener-based evaluations of their partner's spoken fluency.

Figure 2

Listeners' Spoken Fluency and Comprehensibility (From Rappeneker, 2023)
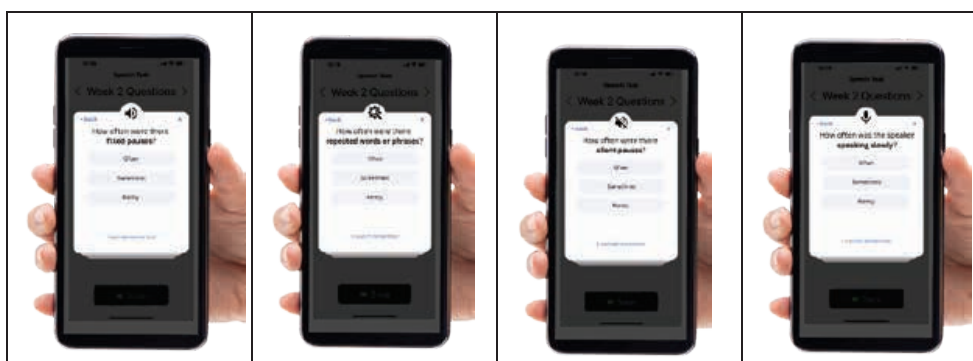
In the initial app design, there was only one question for fluency with six carrots (i.e., with short descriptors). The original question was, "How fluent was the speech?" This design had shortcomings, stemming from the multidimensional nature of spoken fluency and interrater reliability issues. One problem encountered was that raters might assign different overall fluency ratings based on their individual beliefs regarding one dimension's importance over another. The listener-based value differentiation was not clear to the speakers when only asked "How fluent was the speech?" and therefore, we found that from a pedagogical viewpoint, the utility of the feedback to speakers was limited. To account for this issue, we separated questions for spoken fluency to represent aspects of its multidimensionality (i.e., perceivable features of speed, breakdowns, and repairs). We also believe that listeners who evaluate the speeches also benefit from the experience of evaluating fluency as it might help to increase their awareness. In the original design, listeners would assign carrots, which corresponded to a level of fluency. In terms of changes, we removed the assigning of carrots so that listeners would focus primarily on describing the performance. Additionally, from observation, listeners seemed compelled to give speakers carrots as rewards rather than as a representation of the performance itself.

By observing videos of speakers at different L2 proficiencies, we devised a system for listeners to describe their speaking partners' performances. The questions are as follows:

- How often were there filled pauses?
  ➢ Often / Sometimes / Rarely / Cannot remember
- How often were there silent pauses?
  ➢ Often / Sometimes / Rarely / Cannot remember
- How often were there repeated words or phrases?
  ➢ Often / Sometimes / Rarely / Cannot remember
- How often did the speaker speak slowly?
  ➢ Often / Sometimes / Rarely / Cannot remember

Regarding the first two questions about filled pauses, these types of pauses are voiced with utterances such as 'uh' or 'um.' The next two questions about silent, or unfilled, pauses are unvoiced gaps. To help students identify pauses, the primary author created an instructional video. In the video, repeated words or phrases and speed of speech are also reviewed. We chose these aspects of breakdowns and repairs because they are perceivable aspects of performance and this information might motivate students to focus on the linguistic purpose of the 3/2/1 activity. In addition, the video also defines 'often,' 'sometimes,' and 'rarely' for the "How often…?" questions. Listeners who perceive many disfluencies select 'often' when they perceive less disfluencies and they select 'rarely' when they perceive fewer disfluencies. There is no 'none' option because to a certain extent, disfluencies are a natural characteristic of speech even for first language speakers, so the goal of trying to eliminate all disfluencies seems

unreasonable. In the video, listeners are instructed to choose 'rarely' even if they do not hear any disfluency.

Originally, we had added a comprehensibility question for listening partners. The reason we included the question was to encourage students to focus on both meaning and form throughout the speech. The initial question was as follows:

- How easy was it to understand the speaker?
 ➢ Easy / Often Easy / Sometimes Easy / Rarely Easy

There are two related types of questions regarding understanding: one pertains to comprehensibility, while the other relates to intelligibility (Munro & Derwing, 1995). According to Munro and Derwing (1997) comprehensibility is the listener-based judgment regarding the ease or difficulty of understanding utterances. On the other hand, intelligibility is a measurement of the extent to which the speaker's actual intent is understood. We do not have a way to test intelligibility in this situation, so the question we asked was regarding comprehensibility, regarding the extent to which listeners expended energy understanding the content of the speech. The problem with this question was that listeners could not focus simultaneously on comprehending the content of the speech and also spoken fluency. In addition to this problem, the comprehensibility question might be more of a reflection of the listeners' listening ability to understand the speech rather than the speakers' speaking ability. Students had been placed in classes based on their TOEIC test scores and therefore we believed that if they were roughly at the same proficiency, then listeners would be capable of making judgments pertaining to the speakers' ability rather than their own listening ability. Therefore, one of the future goals of the app is to train models to monitor different aspects of spoken fluency to mimic listeners' perceived fluency judgments so that listeners are free to use their attentional resources for comprehension.

We currently do not have a model that can mimic listeners' perceived fluency judgments. However, adding listener-based evaluation might increase students' metalinguistic awareness of spoken fluency and its dimensions of breakdowns, repairs, and speed. For example, if students know that pausing frequently to recall words decreases fluency, and then they receive feedback from the mobile application and peer feedback that they often pause, then they might feel compelled to seek out and utilize strategies to improve their performance. For this reason, we included a page about different strategies for improving spoken fluency that students will be directed to after reading automated comments about their performance in the feedback section. Figure 3 below includes representations of the survey, progress, and strategy pages.

Figure 3

Survey, Progress, and Strategy Pages (From Rappeneker, 2023)
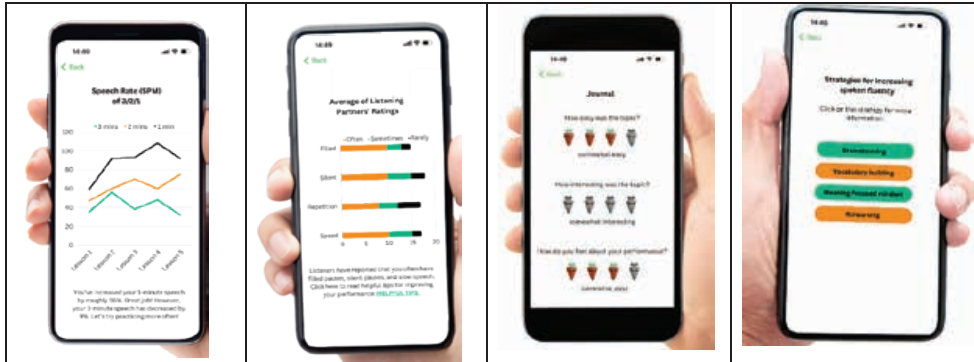
In the first photo from the left, in the future, application users will be able to view their speech rate for each of the three rounds of 3/2/1 across the semester. At the bottom of the page, the application will calculate the percentage of change between performances and offer feedback on strategies to improve their performance. In the second photo from the left, users will be able to see averages from their listening partners' spoken fluency evaluations as well as a summary of their performance generated by the application. In the third photo from the left, there are three survey questions regarding the last 3/2/1 performance. This information will also be displayed in the feedback section along with the results of their speech rate and listening partners' evaluations. The questions are:

• How easy was the topic?
• How interesting was it?
• How do you feel about your performance?

We included these questions because external factors such as topic familiarity and interest can affect spoken fluency performance. These questions are given to speakers after the three rounds. In the last photo from the left, users can view a page regarding different strategies that they can use to improve their spoken fluency performance. They can click on the button and view different activities and ideas under each category.

Concerning the final image, users are presented with an informative page delineating a spectrum of strategies aimed at enhancing their spoken fluency. They can interact with the content by clicking on a designated button, revealing an array of activities and ideas organized by category:

• Brainstorming
• Vocabulary building

• Meaning-focused mindset

• Rehearsing

To fortify students' monologic fluency on familiar topics, the application suggests that students engage in brainstorming activities. Techniques include creating word webs, providing a structure for ideas and thoughts, and employing the 5WH method (i.e., discussing the what, why, where, when, who, and how) to expand original ideas. Vocabulary building emerges as a strategy, wherein students can leverage flashcards featuring topic-specific keywords and employed phrases. Drawing from a repository of vocabulary utilized by previous students on the given topic facilitates integration during speech. In this section, the discourse on meaning-focused mindset underscores the significance of fluency over rigid adherence to grammatical perfection. Addressing the prevailing tendency, particularly in contexts such as Japan, where an undue emphasis on accuracy can impede intelligibility, students are encouraged through explanation on prioritizing effective communication over linguistic precision. The rehearsal section provides insights into the beneficial effects of pre-task rehearsals on spoken fluency. Directing students towards solo practice sessions and elucidating the positive impact of rehearsal on fluency development, this facet of the strategy suite empowers learners to refine their oral proficiency. In sum, this collection of strategies not only serves as a valuable resource for individuals seeking to enhance their spoken fluency, but also equips instructors with a pedagogical framework to guide students through explicit instruction and activities on the various ways to develop spoken fluency.

Upcoming Studies

We are currently undertaking two research studies. First, we have collected over 200 recordings of students' 1-minute monologues to determine the extent to which the three methods of data analysis can accurately determine students' untrimmed speech rate within the classroom context. The three potential methods are: Apple's speech-to-text, Whisper X's speech-to-text, and pitch detection. With this information, we hope to provide instructors and students with more information on the reliability of results as well as the ability to contextualize and interpret the results. In the second research study, we created training videos to help listeners evaluate their speaking partners' speeches. We are testing the extent to which listeners are able to rate a certain performance reliably. In other words, we would like to know whether we have created an evaluation system whereby students can trust the results from listener-based feedback as an authentic reflection of their spoken fluency ability.

Conclusion

In summary, the evaluation of spoken fluency in the context of L2 learning is recognized

as a complex and challenging endeavor, particularly within classroom settings. This complexity is compounded by factors such as noise, lack of privacy, and time constraints, making the collection of data for assessment a formidable task. To date, there is an absence of a MALL application designed to comprehensively track the various facets of L2 spoken fluency for students and provide customized feedback in educational environments. The current situation at Rikkyo University's EDC underscores the absence of a quantitative method for instructors to monitor students' spoken fluency development throughout the semester, despite the course's specific goal of enhancing spoken fluency. We address this gap by presenting an instructional guide for EDC instructors on how to effectively integrate the mobile application into their classes. This innovative approach has the potential to significantly enhance the fluency development of L2 learners within a structured educational context.

References

Hurling, S. (2012). Introduction to EDC. *New Directions in Teaching and Learning English Discussion*, *1*, 2–10.

Kang, O., Rubin, D., & Kermad, A. (2019). The effect of training and rater differences on oral proficiency assessment. *Language Testing*, *36*(4), 481–504. https://doi.org/10.1177/0265532219849522

Lennon, P. (2000). The lexical element in spoken second language fluency. In H. Riggenbach (Ed.), *Perspectives on fluency* (pp. 25–42). The University of Michigan Press.

Maurice, K. (1983). The fluency workshop. *TESOL Newsletter, 17*(4), 29.

Munro, M. J., & Derwing, T. M. (1995). Foreign accent, comprehensibility, and intelligibility in the speech of second language learners. *Language Learning*, *45*(1), 73–97. https://doi.org/10.1111/j.1467-1770.1995.tb00963.x

Munro, M. J., & Derwing, T. M. (1997). Accent, intelligibility, and comprehensibility: Evidence from four L1s. *Studies in second language acquisition*, *19*(1), 1–16. https://doi.org/10.1017/S0272263197001010

Nation, I. S. P., & Newton, J. (2009). *Teaching ESL/EFL Listening and Speaking.* Routledge.

Ogawa, C. (2022). CAF indices and human ratings of oral performances in an opinion-based monologue task. *Language Testing in Asia*, *12*, 4. https://doi.org/10.1186/s40468-022-00154-9

Rappeneker, J. (2023). English fluency (Version 1.0) [Mobile app]. TestFlight. https://testflight.apple.com/join/nJmHTaMh

Rossiter, M., Derwing, T., Manimtim, L., & Thomson, R. (2010). Oral Fluency: The Neglected Component in the Communicative Language Classroom. *The Canadian Moden Language Review–La Revue canadienne des langues vivantes*, *66*(4), 583–606. https://doi.org/10.3138/cmlr.66.4.583

Segalowitz, N. (2010). *Cognitive Bases of Second Language Fluency.* Routledge.

https://doi.org/10.4324/9780203851357

Segalowitz, N. (2016). Second language fluency and its underlying cognitive and social determinants. *International Review of Applied Linguistics in Language Teaching*, *54*(2), 79–95. https://doi.org/10.1515/iral-2016-9991

Segalowitz, N., & Freed, B. F. (2004). Context, contact, and cognition in oral fluency acquisition: Learning Spanish in at-home and study abroad contexts. *Studies in Second Language Acquisition*, *26*(2), 173–199. https://doi.org/10.1017/S0272263104262027

Segalowitz, N., & Segalowitz, S. J. (1993). Skilled performance, practice, and the differentiation of speed-up from automatization effects: Evidence from second language word recognition. *Applied Psycholinguistics*, *14*(3), 369–385. https://doi.org/10.1017/S0142716400010845

Sturges, J., Arthurson, D., Itoi, K., Koike, A., & Reid, S. (2023). *What's your opinion? Interactive skills for effective discussion – Book II* (3rd ed.). Rikkyo University.

Suzuki, S., Kormos, J., & Uchihara, T. (2021). The relationship between utterance and perceived fluency: A meta-analysis of correlational studies. *The Modern Language Journal*, *105*(2), 435–463. https://doi.org/10.1111/modl.12706

Tavakoli, P. (2016). Fluency in monologic and dialogic task performance: Challenges in defining and measuring L2 fluency. *International Review of Applied Linguistics in Language Teaching*, *54*(2), 133–150. https://doi.org/10.1515/iral-2016-9994

Tavakoli, P., & Hunter, A. M. (2018). Is fluency being 'neglected' in the classroom? Teacher understanding of fluency and related classroom practices. *Language Teaching Research*, *22*(3), 330–349. https://doi.org/10.1177/1362168817708462

Tavakoli, P., & Wright, C. (2020). *Second language speech fluency: From research to practice*. Cambridge University Press. https://doi.org/10.1017/9781108589109

Woodward, H., & Gupta, C. (2022). Designing a mobile application to track spoken fluency development. *Journal of Foreign Language Education and Research*, *3*, 160–174.

Woodward, H., & Rappeneker, J. (2023). Prompt-related challenges: Using a mobile application to measure spoken fluency. *Journal of Foreign Language Education and Research*, *4*, 1–27.

Wright, C. (2020). Effects of Task Type on L2 Mandarin Fluency Development. *Journal of Second Language Studies*, *3*(2), 157–179. https://doi.org/10.1075/jsls.00010.wri