[Website Review]

A Review of the Website Pronounce.com

Simon Moxon (simon@sitamoxon.com)
Walailak University, Nakhon Si Thammarat, Thailand

Website	Pronounce.com
URL	https://www.getpronounce.com
Platform	Web-based and web-browser extension.
Language(s)*	English interface. British, American, Australian, Indian**, and Welsh*** accent coaches.
Levels	Beginner, Intermediate, Advanced, Native.
Media	HTML (Text, Audio, Image)
Hardware Requirements	Computer or smartphone. Internet access. Microphone and speakers (microphone headset preferable).
Pricing	Lite (up to three recordings per day): Free. Premium (unrestricted use): • 1-day free for each friend referral. • 7-day free trial, then US$19.99 per month. • 14-day free trial, then US$144 per year. More information regarding features and pricing can be found at https://www.getpronounce.com/pronounce-pricing
* The main web-based platform offers a selection of male and female voices for British, American, Australian, and Indian accents.	
** The browser extension version does not provide male speaker voices for the Indian accent.	
*** The Welsh voice is only available in the browser extension as a male speaker.	

Abstract

Artificial Intelligence and Automatic Speech Recognition technologies play a transformative role in EFL pedagogy, specifically in acquiring oral linguistic competencies like pronunciation and fluency. Traditional teaching methods often fall short of providing immediate, personalized feedback. In contrast, emerging technologies offer real-time, precise evaluations, significantly improving linguistic fluency and pronunciation while reducing cognitive load and alleviate speech-related anxiety. Web-based applications integrated with Automatic Speech Recognition and Automatic Pronunciation Assessment systems, are effective pedagogical tools offering unbiased, comprehensive feedback and further facilitate autonomous practice, which is essential for skill enhancement. Integrating such advanced tools is not just an option but a necessity for comprehensive linguistic development, preparing EFL learners for a future increasingly influenced

by technology. This paper highlights the features, strengths, and weaknesses of Pronounce (https://www.getpronounce.com), a web-based platform that employs these technologies for in-depth speech analysis and corrective feedback.

Introduction

Acquiring oral linguistic competencies, such as pronunciation and fluency, is a pivotal element in EFL pedagogy. Traditional instructional methodologies predominantly employ teacher-centric evaluative feedback, often lacking immediacy and individualization (Sun, 2023). However, the pedagogical landscape of EFL is experiencing a paradigm shift facilitated by advancements in Artificial Intelligence (AI) and Automatic Speech Recognition (ASR) technologies (Kholis, 2021; Son et al., 2023; Spring & Tabuchi, 2022).

These emergent technologies are engendering a transformative impact on the modality of feedback delivery in oral linguistic training. AI and ASR technologies afford instantaneous, precise evaluations and have evidenced considerable enhancements in linguistic fluency and pronunciation (Moxon, 2021; Zou et al., 2023a). Furthermore, AI-driven evaluative platforms are engineered to mitigate the cognitive load on learners, thereby optimizing the efficiency of the learning process. These platforms also exhibit the potential to ameliorate speech-related anxiety, a prevalent issue among EFL learners (Tai & Chen, 2020).

When integrated with ASR technology, Chatbots have demonstrated efficacy as pedagogical tools in EFL contexts. These chatbots offer the dual advantages of instantaneous feedback and eliminating human evaluator bias, thereby serving as a valuable adjunct to traditional instructional methods (Tai & Chen, 2020). Additionally, Automatic Pronunciation Assessment (APA) systems are gaining academic recognition for their comprehensive evaluative feedback on pronunciation and fluency, substantiated by statistical validation (Moxon, 2021). Web-based applications that employ speech recognition technology provide a conducive platform for autonomous reading-aloud practice, which is crucial for enhancing fluency and pronunciation (Rodríguez-Fuentes & Calle-Díaz, 2023).

In light of these technological advancements, it becomes imperative for EFL learners to integrate these innovative tools into their linguistic skill development regimen. Incorporating AI and ASR technologies augments the quality of EFL pedagogy and equips learners for a future increasingly dominated by such technological interfaces. Therefore, utilizing these advanced tools is not merely an elective but a requisite for holistic linguistic skill development. The following sections describe the functions and features of one such notable example, Pronounce, a web-based platform specializing in pronunciation evaluation and review.

Description of Features

Pronounce provides a user-friendly platform for enhancing English pronunciation via three ASR, APA, and AI-driven speaking tasks. These tasks, designed to engage users for at least five minutes daily, include reading text aloud, participating in open conversation, and interacting with an AI speech coach. Post-task, the platform analyzes the user's speech and generates a

comprehensive report highlighting areas of strength and pinpointing weaknesses. The report also offers corrective guidance and targeted content to correct linguistic errors.

Registration and Login

Access to Pronounce is via login credentials or a Google account. As shown in Figure 1, initial registration involves a brief survey asking for the user's native language, English proficiency level, and basic demographics. According to the website information, this data is used to tailor the feedback in the speech evaluations.
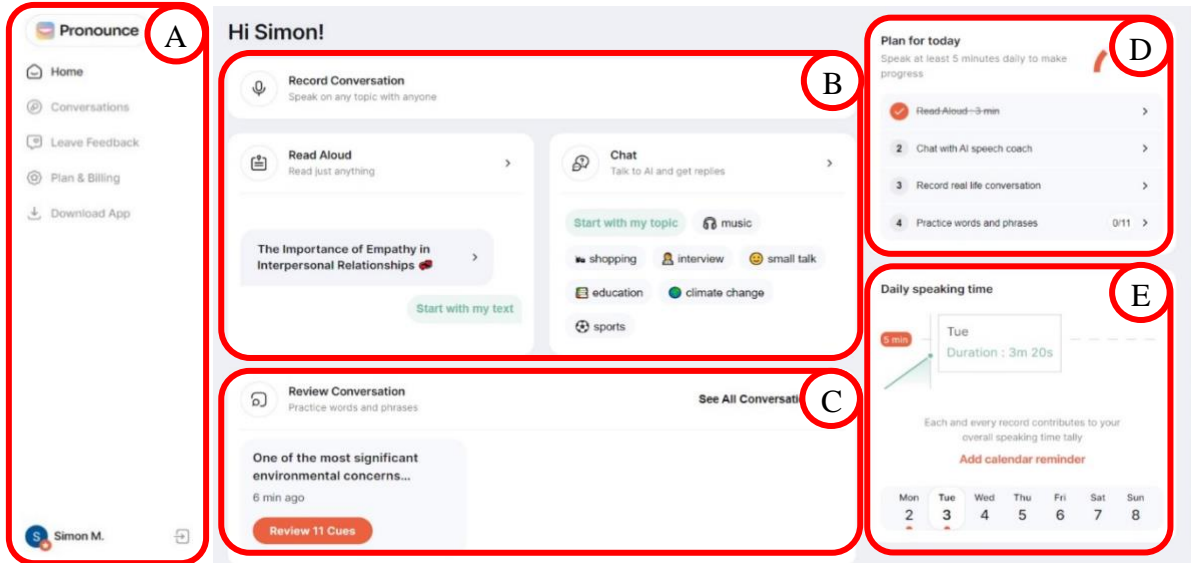
Figure 1
Registration Dialogue

Once logged in, the user is presented with the main dashboard page.

Dashboard

The dashboard page outlines daily speaking tasks and tracks progress, along with a historical record of past speaking evaluations. As illustrated in Figure 2, the layout is user-friendly and can be broken down into five specific sections:

A. Options for site navigation and logging out.
B. The trio of available speaking tasks.
C. Recent speaking evaluations.
D. The daily speaking plan.
E. Daily speaking time for the week and calendar reminders.

Figure 2
User Dashboard Interface

The website offers three ways to access speaking tasks: 1) Via the "Conversation" menu in the navigation panel (see Figure 2, item A). 2) Through the daily speaking plan (see Figure 2, item D). 3) Directly from the task menus (see Figure 2, item B).

Voice Capture

As depicted in Figure 3, the recording interface features a simple design for configuring audio devices and timing the recording, plus controls for pausing, submitting, or deleting recordings. While this layout is consistent for the "Read Aloud" and "Conversation" tasks, the "Chat" task interface differs in design and function.

Figure 3
Speech Recording Interface
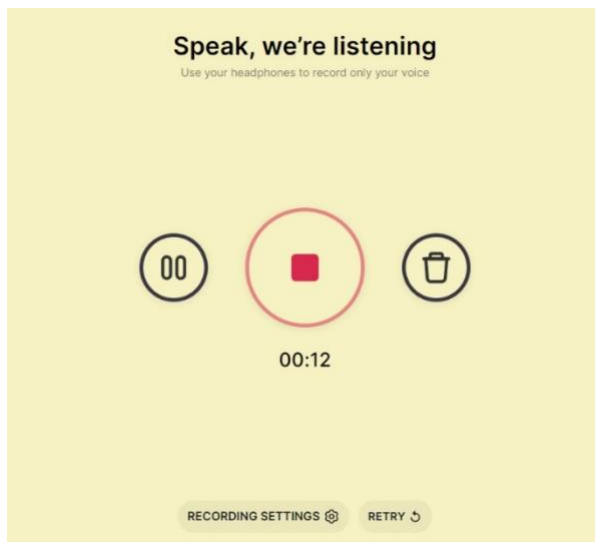
After submitting a speaking task, the first spoken sentence is transcribed and used for ease of identification in the speech history panel.
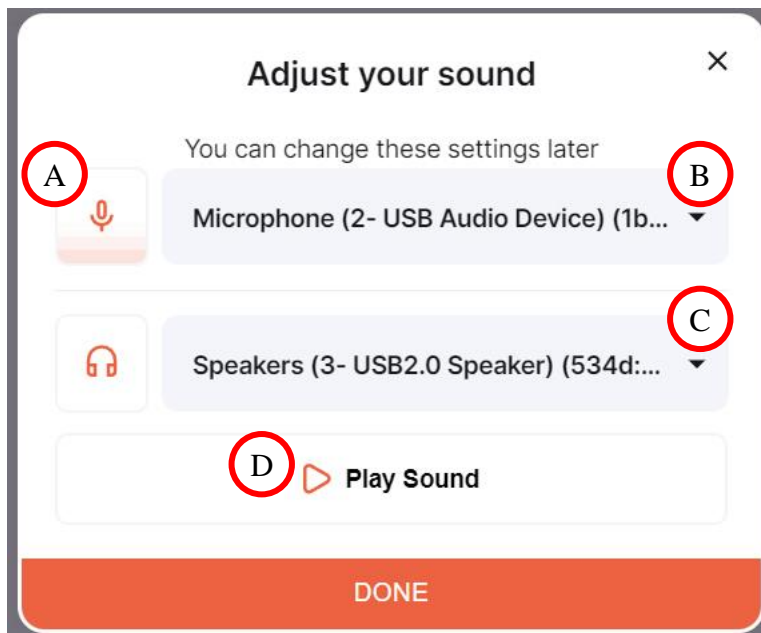
Configuration

The website offers two windows for configuring audio devices and the Speech Coach voice.

The audio device configuration window allows users to choose and test their recording and playback devices (see Figure 4) and features:

A. A microphone volume indicator.
B. Recording device options.
C. Playback device options.
D. Audio testing.

Figure 4
Audio Device Configuration

The configuration window for the Speech Coach voice, shown in Figure 5, allows users to choose the coach's nationality, gender, and speech rate. The interface consists of:

A. Access to the voice settings window.
B. Speaker nationality.
C. Speaker character selection.
D. Voice preview.
E. Rate of speech adjustment (0.5x to 2x).

Figure 5
Speech Coach Voice Configuration
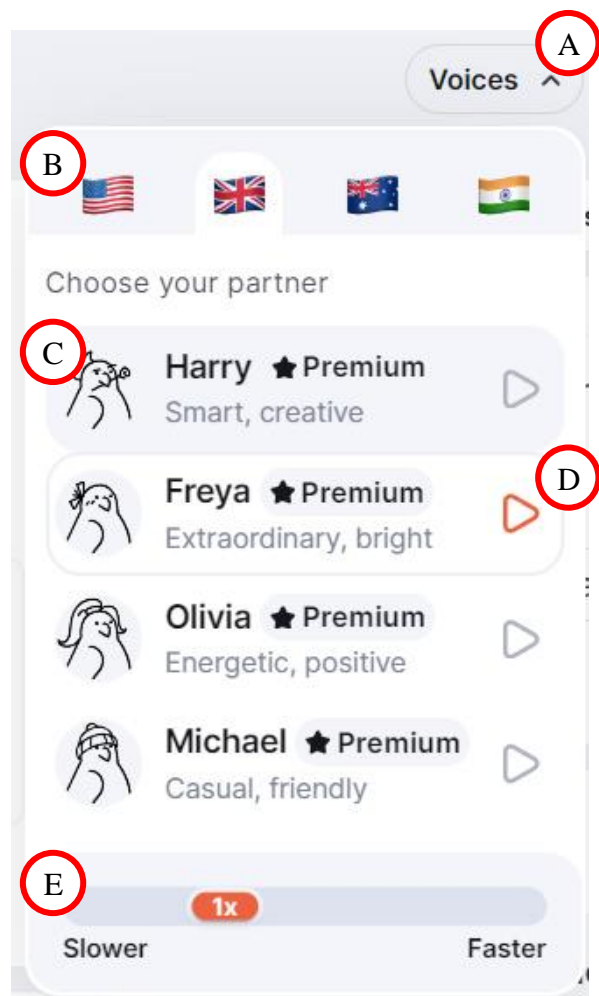
Note. The range of nationalities and characters differs between the main webpage and the browser plugin.

While all voices can be previewed, as Figure 5 shows, British English characters are only accessible to premium subscribers. This restriction applies to all but one American character, Blake, which, due to its low-quality robotic-sounding voice, offers a poor example of nativelike pronunciation. Users can adjust the speech rate, but rates above 1.25x may be too fast for comprehension, even for native speakers.

Configuration settings can be accessed within speaking tasks and the review screen but not from the main dashboard.

Daily Speaking Plan

The daily speaking plan (see Figure 2, item D) lists three hyperlinked tasks with completion indicators. A fourth option, "Practice words and phrases," appears after task evaluations flag pronunciation errors. Pronounce keeps a daily list of such mistakes accessible for review. As Figure 6 illustrates (post-task), the panel also shows completed speaking time (3

minutes) and a tally of reviewed mispronunciations (0/11), which updates after each successful review.

Figure 6
Daily Speaking Plan Panel
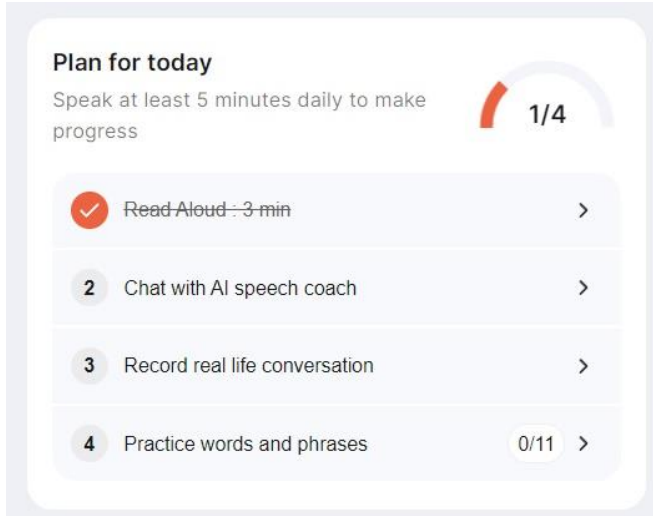
Daily Speaking Time and Reminder

The lower-right section of the dashboard features a panel for daily speaking time and reminders (see Figure 2, item E). It displays daily speaking achievements graphically (see Figure 7) and allows the creation of Google calendar reminders. In addition, users are sent an email reminder if they fail to meet daily targets.

Figure 7
*Daily Speaking Time Chart and Calendar*
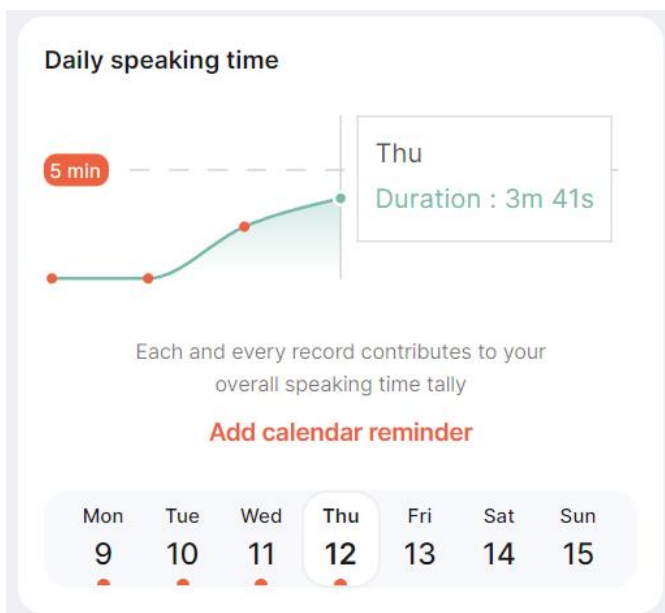
Read Aloud

The "Read Aloud" menu allows users to record speech from a set script, as displayed in Figure 8. The script, typically around 250 words, changes daily and can be overwritten by the user, although the method varies. From the dashboard, the option is "Start with my text" (see Figure 9), while from the "Conversations" menu, it is "Change text" (see Figure 8).
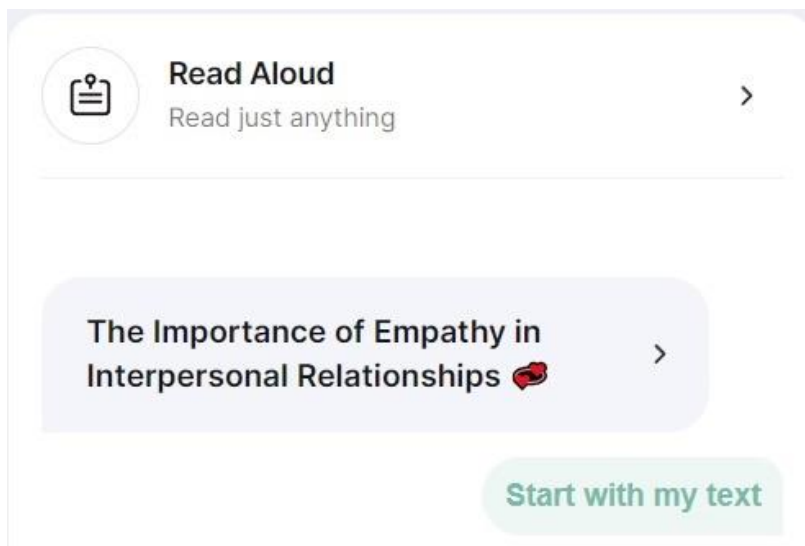
Figure 8
Read Aloud Recording Interface
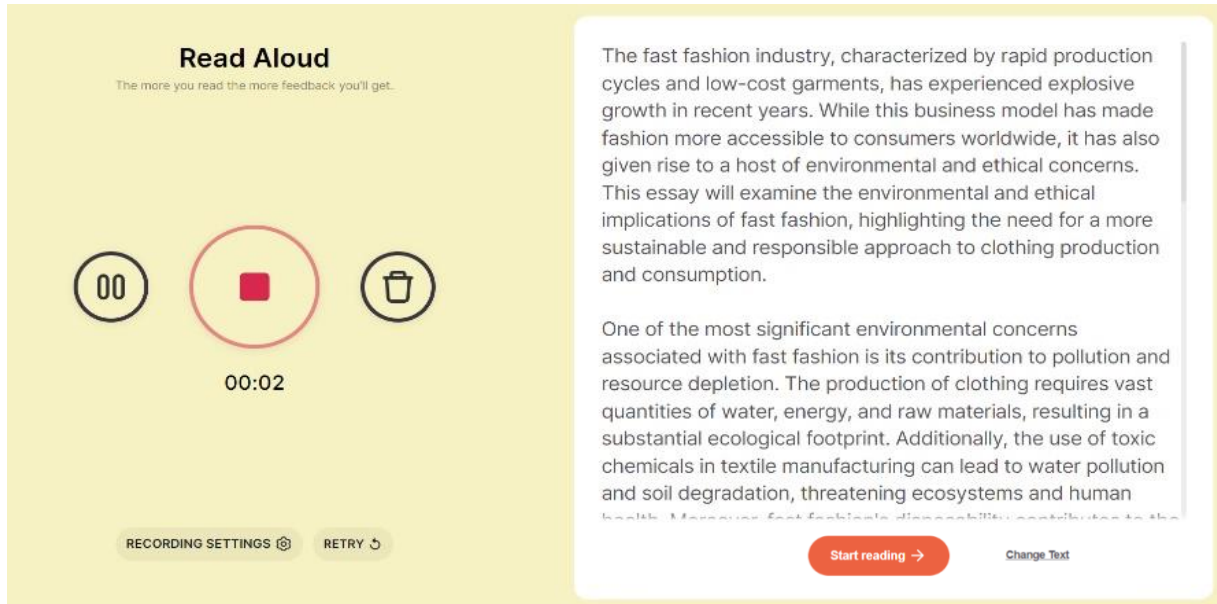
Note. The controls displayed below the text have been superimposed to illustrate the layout when accessed from the "Conversations" menu.

Figure 9
Read Aloud Daily Topic

The "Read Aloud" task offers an excellent platform for students to prepare for a reading test or classroom presentation. However, it should be noted that, unlike other ASR and APA systems, such as Microsoft Azure or Speechace, that compare submitted text and audio for evaluation (see Figure 10), Pronounce relies solely on ASR-transcribed information. Therefore, it does not evaluate speech accuracy, completeness, or fluency, even for the "Read Aloud" task.

Figure 10
*Microsoft Azure Pronunciation Response*

Accuracy Score	Fluency Score	Completeness Score	Pronunciation Score	Words Omitted	Extra Words Spoken
56	61	83	62	of	a, demonstration, all-talk-app

this[70]		is[94]	a[40]	demonstration[78]				of	all-talk-app[56]		
ðɪs[70]		ɪz[94]	ʌ[40]	dɛ[21]	mən[72]	streɪ[95]	ʃən[76]	-	ɔl[67]	tɔk[56]	æp[45]
ð	ɪ	s	ɪ	z	ʌ	d	ɛ	m	ə	n	s	t	r	eɪ	ʃ	ə	n	-	ɔ	l	t	ɔ	k	æ	p
79	97	59	93	95	40	0	36	84	64	60	73	100	100	100	92	79	58	-	78	48	42	81	42	45	45

Note. Microsoft Azure pronunciation response rendered in All-Talk. Adapted from *Pronunciation evaluation* [Image], by All-Talk-App, 2023 (https://www.all-talk-app.com/freespeech.php). Copyright 2023 by Simon Moxon.

Chat With AI

Pronounce's new "Chat" feature allows users to practice conversational English, addressing common EFL challenges like limited exposure to native speakers and communication anxiety. Users can choose from seven topics or create their own (see Figure 11). The set topics range from making small talk to climate change, which could easily be applied to users of all English-speaking levels.

Figure 11
Chat Topics

203
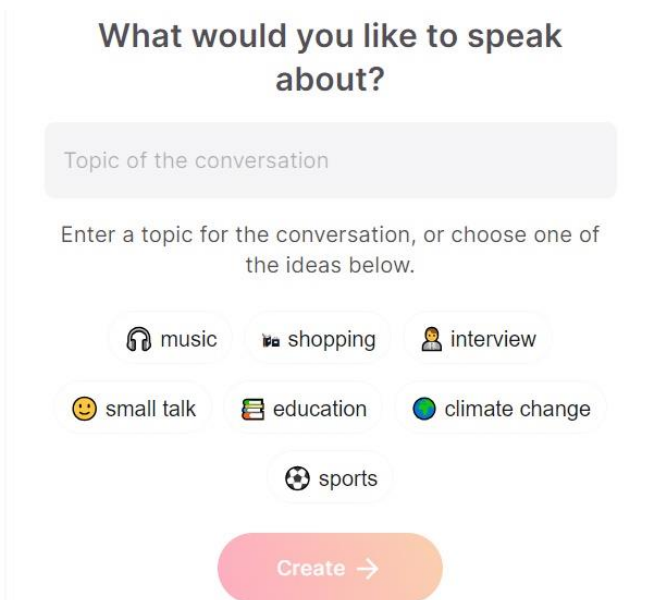
The "Chat" dialogue window resembles a standard messaging interface where speaker and responder messages appear on opposite sides (see Figure 12). Within the screen, users can:

    A. Configure Speech Coach voice settings.
    B. Adjust audio devices.
    C. Replay the speech.
    D. View the transcript.
    E. Request a suggested response.
    F. Record a response.
    G. End the chat session.

Figure 12
Chat Interface
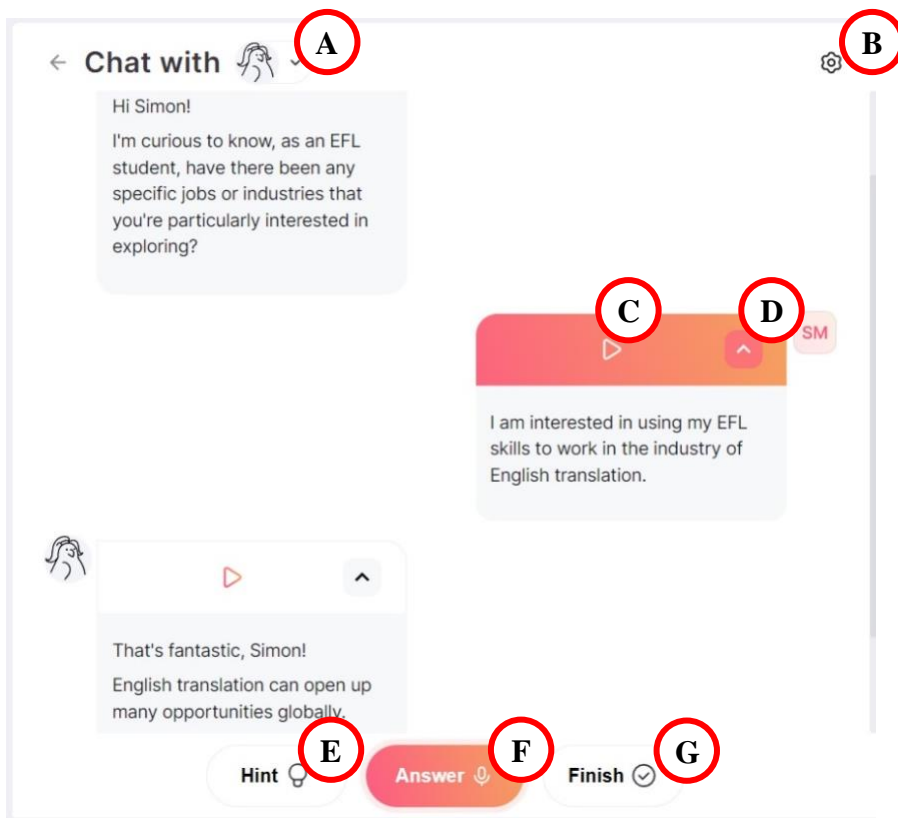
The AI Chatbot, known as the "Speech Coach," initiates the conversation with a greeting and a question. The coach can adapt to user requests, meaning one can express their speaking or comprehension limitations, prompting the coach to simplify its questions.

After the Speech Coach's prompt, users can prepare a response before manually starting the recording process (see Figure 12, item F). The site employs ASR for speech transcription, generating a dialogue box that can be expanded to replay the audio or view the text (see Figure 12, items C and D). The interface also provides a "Hint" option (Figure 11, item E), which generates a suggested response, as seen in Figure 13. However, Pronounce retains the transcribed speech, even when a hint is used. The user is prompted to end the conversation after every two or three minutes but may continue the conversation if they wish.

Figure 13
Chat Response Hint
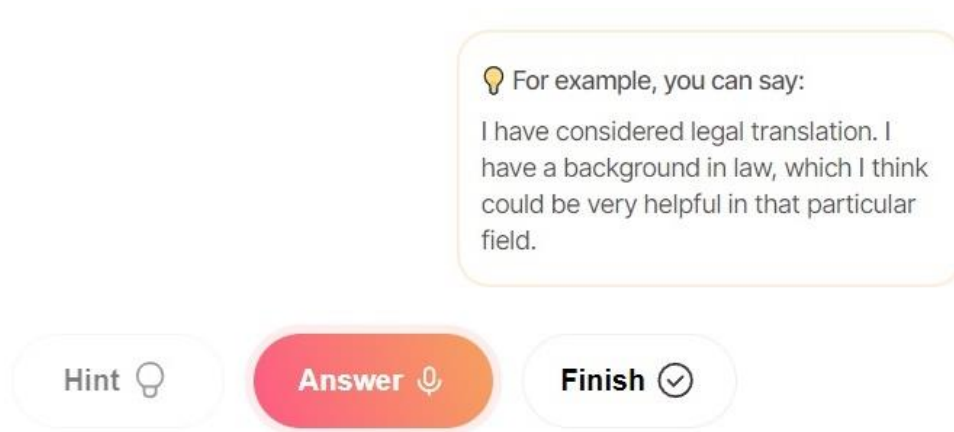
Within the "Chat" task, the user's responses are evaluated collectively, and feedback is provided accordingly.

Speech History

Speech history is accessible from two locations: the main list is available through the "Conversations" menu (see Figure 14), while as shown in Figure 2, item C, the three latest speaking attempts can be found at the bottom of the dashboard, which also includes a link to the main list.

The main list is chronological, showing icons for each task type, i.e., Conversation (microphone), Chat (speech bubbles), or Read Aloud (clipboard), and the first line of transcribed speech. Brief feedback, date, time, and a "Polished" score, which reflects the amount of practice needed, are also displayed. Tasks requiring review are highlighted in red; fully polished ones are in green.

Figure 14
Main Speech History List
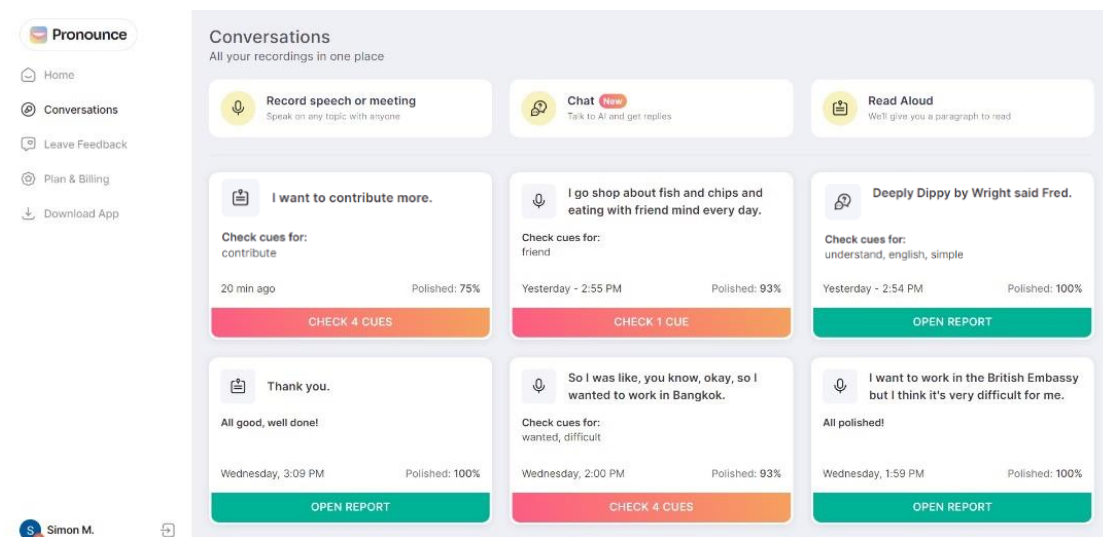
Clicking on any item in the speech history opens the feedback and correction dialogue for that task, offering detailed feedback and corrective guidance.

Feedback and Correction

Pronounce distinguishes itself with its feedback and remedial content. The user interface is uniform across speaking tasks, and while evaluation feedback could be improved, the remedial content is well-structured. Users can access feedback via three methods: 1) the "Review Conversation" panel (Figure 2, item C), 2) "Conversation History" (Figure 14), or 3) the "Practice words and phrases" option in the daily speaking plan (Figure 2, item D).

Figure 15 illustrates the layout of the evaluation feedback page, organized into eight sections:

- A. Task title and completion time.
- B. Navigation menu.
- C. Feedback for problematic phonemes.
- D. Feedback for mispronunciations.
- E. Remedial task progress.
- F. Words pronounced perfectly.
- G. Clarity feedback.
- H. Engagement feedback.

Figure 15
Evaluation Feedback Page Layout
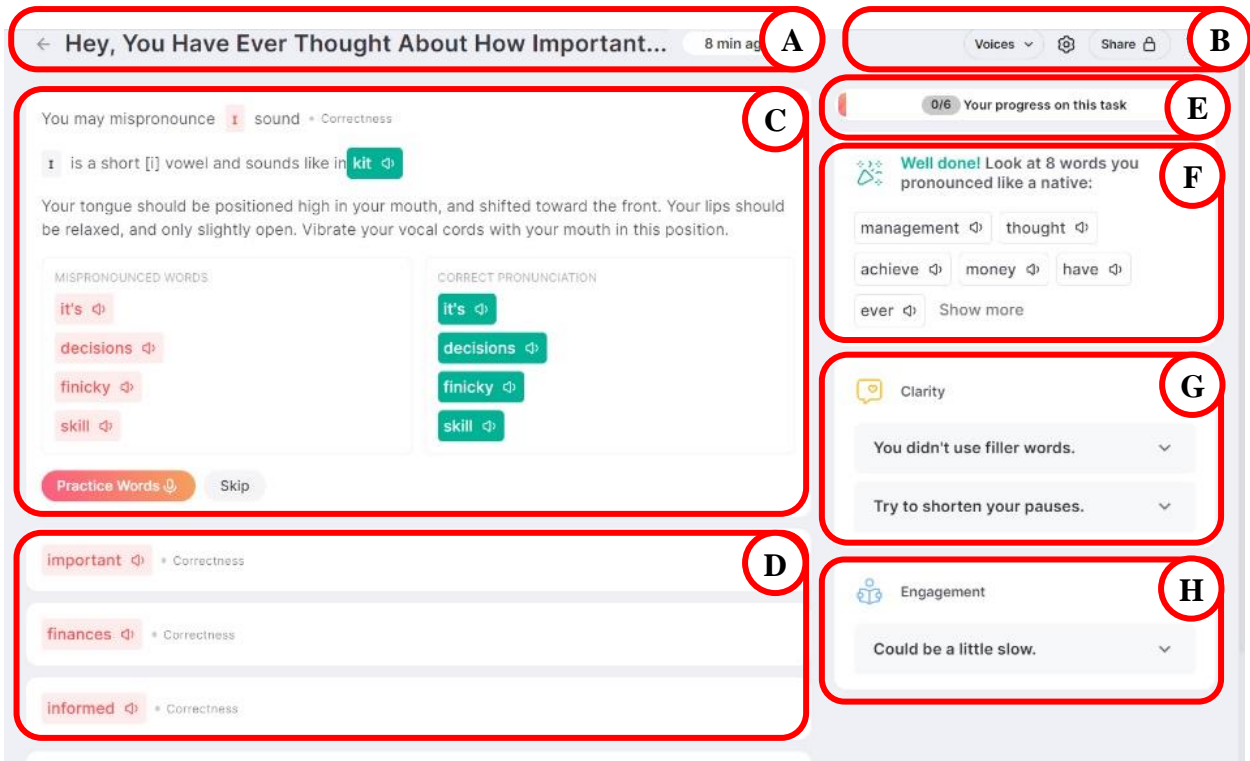
The navigation menu, depicted in Figure 16, allows users to configure the Speech Coach voice, set up audio devices, share feedback reports with other registered users, and delete task attempts.

Figure 16
Feedback Report Navigation Options

The feedback area (see Figure 15, items C and D) lists mispronounced phonetic sounds and words, leading to a remedial dialogue. The remedial dialogue explains the error (see Figure 17), followed by practicing the word in isolation (see Figure 18) and then in sentences (see Figure 19). Skipped words are marked as reviewed and removed from the list.

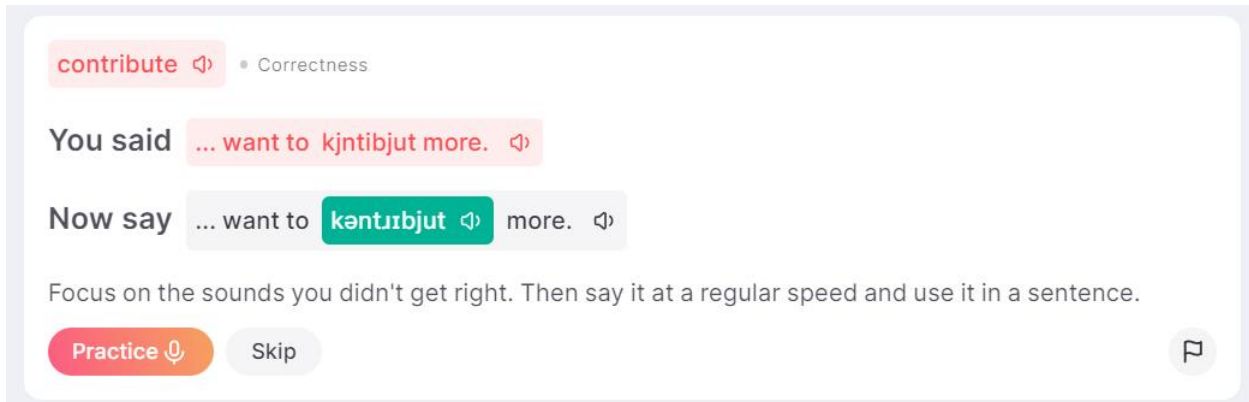**Figure 17**
Evaluation Feedback of a Mispronounced Word

Figure 18
Pronunciation Practice of Each Mispronounced Word
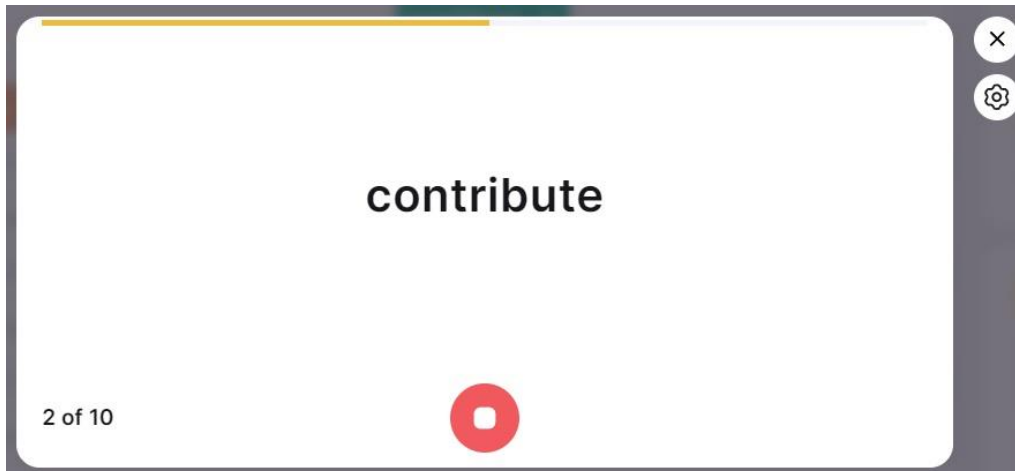
Figure 19

Reviewed Word in Practice Phrase

And, any additional plastics we add will **contribute** further.

1 of 3
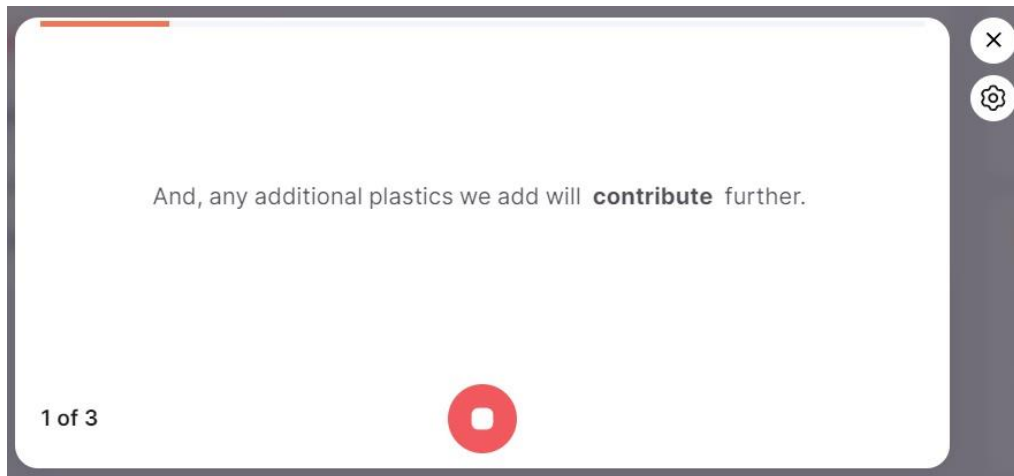
As users review or skip each problem word, it is removed from the review area, and the review progress bar updates, as illustrated in Figure 20. After completing the review, the corresponding speaking task in the conversation history panel will show as 100% polished.

Figure 20

Pronunciation Review Progress Bar

1/6   Your progress on this task
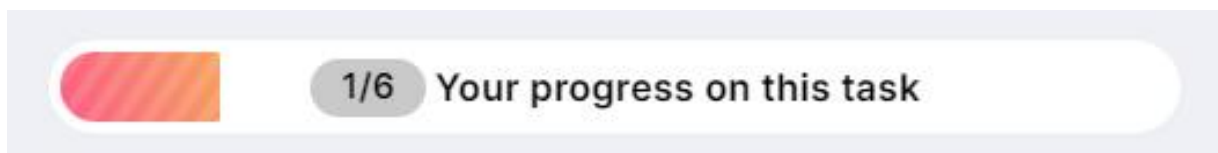
Besides pronunciation feedback, Pronounce offers suggestions for improving clarity, as shown in Figure 21. More generic feedback on clarity and engagement can also be found in panels G and H of the evaluation page (see Figure 15, items G and H), which give basic suggestions on using filler words, pauses, and pace.

Figure 21

Clarity Feedback for a Spoken Phrase

very ◁)   • Clarity

You said
I thought about working in the Thai Embassy as a correspondent or a translator, but it's very difficult for me.

It appears that **very** may be unnecessary in this sentence. Consider removing it.

Done
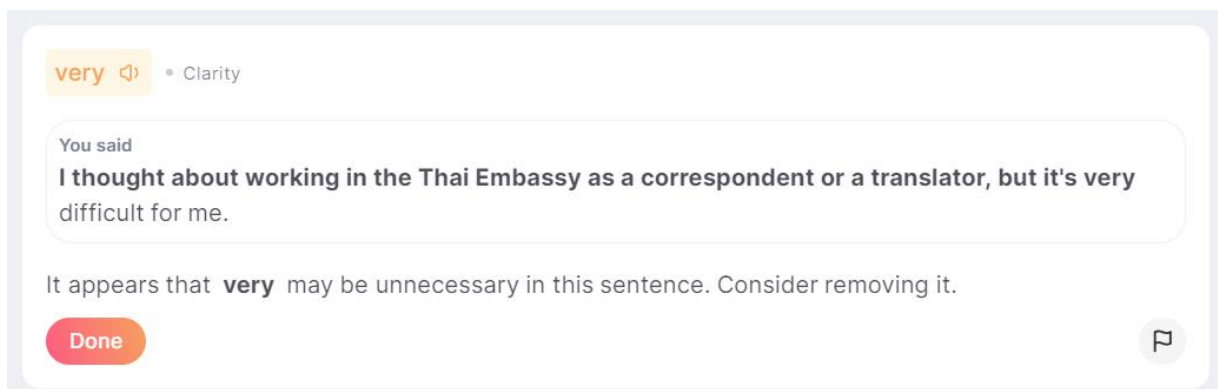
Evaluation Sharing

Akin to Google Drive, users can share evaluation reports with other registered users via a link (see Figure 22), providing a convenient means for EFL instructors to monitor student progress or address pronunciation issues. However, as Pronounce offers no means to create groups or administrator roles, instructors would need to rely on students to share their evaluation reports manually.

Figure 22
Feedback Sharing Link Creation

Evaluation

AI platforms offer several advantages over traditional face-to-face instruction, including consistent evaluation, autonomous learning features, and the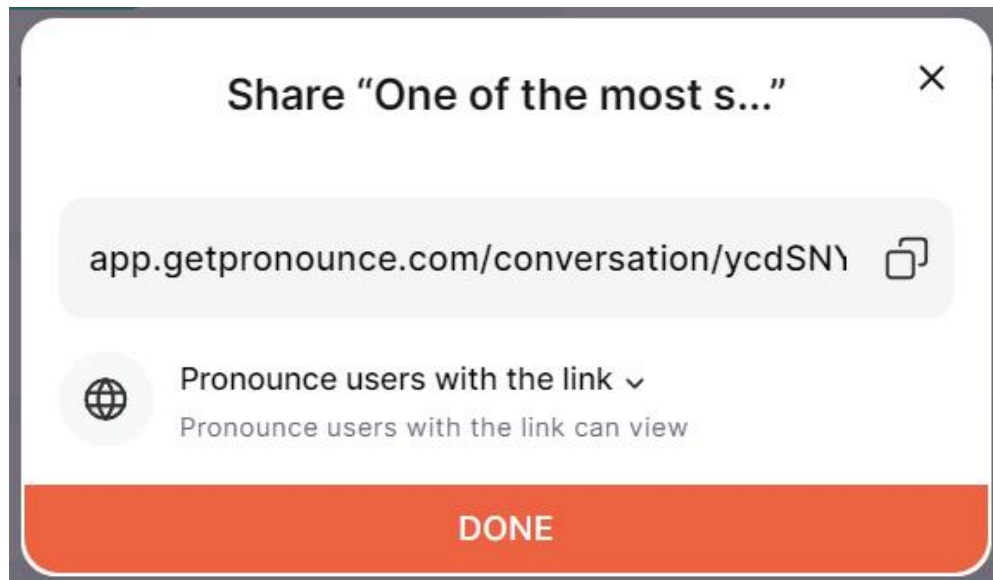 AI's unwavering patience. Pronounce is no exception, providing EFL learners with a convenient and stress-free way to practice both read-aloud and conversational skills while receiving immediate, impartial feedback.

The platform imposes minimal limitations for those using the free subscription, making it an excellent learning resource. Although users are limited to three daily recordings, they can still meet their speaking goals by extending their speaking time in each task. Furthermore, the platform's evaluation, review, and report-sharing features are fully accessible. However, a paid subscription is required to unlock extended speaking practice opportunities and access a broader range of natural-sounding Speech Coach characters, which may be cost-prohibitive for some students.

In terms of user experience, the platform's interface is straightforward but could benefit from improvements in clarity and consistency. Specifically, the recording triggers for each speaking task vary: "Read Aloud" starts automatically, "Conversation" features a three-second countdown, and "Chat" allows manual control. Additionally, the central "stop" button, shown in Figure 3, actually submits the recording for evaluation, which can lead to unintentional submissions.

Registration and Configuration

The user interfaces for the three-stage registration process, audio device setup, and Speech Coach configuration are generally intuitive but have room for improvement. Each proficiency level shown during registration features a "right-arrow" icon, suggesting clicking it will reveal more information or options. As the icon has no function, it is misleading and could lead to unintended selection. Moreover, there is no way to revise this information later.

Regarding audio device configuration, as shown in Figure 4, item A, the microphone volume indicator lacks distinct coloration and labeling, making it easily overlooked. The indicator also lacks recommended volume levels for optimal ASR performance. Moreover, the audio device and Speech Coach settings can only be accessed within speaking tasks or the review screen, not from the main dashboard. This seems like an oversight, as one might expect to configure these settings before initiating a task.

Speaking Tasks and Evaluation

Overall, the speaking task and evaluation interfaces are intuitive and engaging, effectively leveraging a daily time chart and Google reminders to motivate users to monitor their progress and achieve their goals. However, the platform has two significant shortcomings.

First, the language used in the speaking tasks and feedback is not customized to match the user's proficiency level, which can be daunting for beginners confronted with advanced vocabulary. Tests conducted using multiple accounts with different settings revealed no discernible variation in content or feedback.

Second, the evaluation relies solely on ASR, which has limitations. For instance, it does not assess reading accuracy, and poor pronunciation can lead to incorrect word recognition. For example, a mispronounced "next" could be wrongly identified as "neck" without any corrective feedback provided. This is particularly problematic for Thai EFL students, who, influenced by their native language, often soften or omit word-ending sounds in English—a nuance that the current evaluation method fails to capture.

Despite these drawbacks, Pronounce has a unique advantage over other platforms: its AI-powered Speaking Coach. This feature offers stress-free conversational practice and can be instructed to modify or rephrase questions as needed. However, the "Chat" interface has glitches, such as system crashes and, in rare cases, transcription into foreign languages. Issues with audio playback failure and ASR processing were observed more frequently. The system lacks a graceful recovery mechanism for these errors, making it impossible to submit the session for evaluation, which can discourage users after an extended interaction with the Speech Coach.

Review

The review process is thorough and generally effective, but it has limitations. Specifically, it concentrates on individual mispronounced words and phonemes, evaluating them in isolation rather than in the context of the entire speech. For instance, during the practice phrase stage shown in Figure 19, only the target word is assessed for correct pronunciation and placement. In contrast, adjacent words may be mispronounced, omitted, or replaced without detection. This compromises the exercise's effectiveness, as surrounding sounds can impact mispronunciation. Additionally, the lack of a comprehensive transcript and audio review for each task restricts instructional guidance and self-assessment. A potential remedy could be a complete

audio review accompanied by an annotated transcript, akin to the method employed by ELSA (Nushi & Sadeghi, 2021).

Sharing one's evaluation report offers the unique benefit of enabling targeted feedback from instructors in an asynchronous learning environment. However, testing revealed that shared users could also complete the review tasks on behalf of the original speaker, likely due to a platform glitch.

Lastly, while many EFL students aim for proficiency based on established scales like IELTS or CEFR, Pronounce's evaluations do not align with these benchmarks. A reporting format similar to the Speechace speaking test (Speechace, 2022) would be more informative.

Conclusion

Pronounce offers an interactive and captivating platform designed to elevate one's English pronunciation skills. While not without flaws, it has excellent potential as a self-guided learning tool for students and a progress monitoring system for educators. Despite being a relatively new platform, its unique "Chat" feature has the potential to boost speaking confidence. The premium version takes it up a notch with lifelike Speech Coach voices, although the cost may be a hurdle for some users. While the platform is compatible across various devices, its dependence on ASR for speech assessment is a limitation worth noting.

References

All-Talk-App. (2023). *Pronunciation evaluation* [Image]. Retrieved 7 October, 2023, from https://www.all-talk-app.com/freespeech.php

Kholis, A. (2021). Elsa Speak App: Automatic speech recognition (ASR) for supplementing English pronunciation skills. *Pedagogy: Journal of English Language Teaching, 9*(1), 01. https://doi.org/10.32332/joelt.v9i1.2723

McCrocklin, S. (2019). ASR-based dictation practice for second language pronunciation improvement. *Journal of Second Language Pronunciation, 5*(1): 98–118. https://doi.org/10.1075/jslp.16034.mcc

Moxon, S. (2021). Exploring the Effects of automated pronunciation Evaluation on L2 students in Thailand. *IAFOR Journal of Education: Language Learning in Education, 9*(3), 41-56. https://doi.org/10.22492/ije.9.3.03

Nushi, M., & Sadeghi, M. (2021). A critical review of ELSA: A pronunciation app. *Computer Assisted Language Learning Electronic Journal, 22*(3), 287-302. http://callej.org/journal/22-3/Nushi-Sadeghi2021.pdf

Rodríguez-Fuentes, R. A., & Calle-Díaz, L. (2023). Fluency development beyond speech rate: A study on the effects of read aloud protocols on EFL learners. *Ampersand, 11*, 100133. https://doi.org/10.1016/j.amper.2023.100133

Son, J., Ružić, N. K., & Philpott, A. (2023). Artificial intelligence technologies and applications for language learning and teaching. *Journal of China Computer-Assisted Language Learning*. https://doi.org/10.1515/jccall-2023-0015

Speechace. (2022, 21 June). *Rubrics-tabs* [Image]. https://www.speechace.com/speaking-test

Spring, R., & Tabuchi, R. (2022). The role of ASR training in EFL pronunciation improvement: An in-depth look at the impact of treatment length and guided practice on specific

pronunciation points. *Computer Assisted Language Learning Electronic Journal, 23*(3), 163–185. http://www.callej.org/journal/23-3/Spring-Tabuchi2022.pdf

Sun, W. (2023). The impact of automatic speech recognition technology on second language pronunciation and speaking skills of EFL learners: a mixed methods investigation. *Frontiers in Psychology, 14*. https://doi.org/10.3389/fpsyg.2023.1210187

Tai, T. Y., & Chen, H. H. (2020). The impact of Google Assistant on adolescent EFL learners' willingness to communicate. *Interactive Learning Environments, 31*(3), 1485–1502. https://doi.org/10.1080/10494820.2020.1841801

Zou, B., Du, Y., Wang, Z. H., Chen, J., & Zhang, W. (2023a). An investigation into artificial intelligence speech evaluation programs with automatic feedback for developing EFL learners' speaking skills. *SAGE Open, 13*(3). https://doi.org/10.1177/21582440231193818

Zou, B., Guan, X., Shao, Y., & Chen, P. (2023b). Supporting speaking practice by social network-based interaction in artificial intelligence (AI)-assisted language learning. *Sustainability, 15*(4), 2872. https://doi.org/10.3390/su15042872