

[Software Review]

A Critical Review of ELSA: A Pronunciation App

Musa Nushi (M_nushi@sbu.ac.ir)
Shahid Beheshti University, Tehran, Iran

Mahsa Sadeghi (mahsa1995.sadeghi@gmail.com)
Shahid Beheshti University, Tehran, Iran

Application	ELSA
Publisher	ELSA
Product Type	Smartphone Application Software
Language(s)	English
Level	Any
Media format	APK/IPA
Operating Systems	Android/iOS
Hardware Requirements	Smartphone/Internet Connection
Supplementary Software	None
Price	Free (includes in-app purchases)

Abstract

This article reviews a mobile application called ELSA Speak, also known as ELSA, which has exclusively been designed for pronunciation development in EFL/ESL. Pronunciation is an important skill whose teaching is often neglected in second language teaching due to a number of reasons, including teachers' lack of confidence or training in teaching the skill, perceived lack of importance of pronunciation teaching and lack of sufficient and easy-to-use materials. Therefore, an application like ELSA can be beneficial and serve as a supplementary teaching aid for teachers. Furthermore, it can provide language learners with opportunity to enhance their pronunciation on their own and become more autonomous in learning the skill. It is worth noting that the feature that distinguishes ELSA from other pronunciation-development apps is its built-in speech recognition technology, which has made it rank among the world's top five artificial intelligence apps. However, ELSA is not flawless and can be improved in several respects.

Keywords: ELSA, mobile application, pronunciation, second language, technology

Introduction

It would not be an overstatement to say there is now technology to support almost every aspect of second language (L2) teaching and learning thanks to the rapidly emerging technologies such as artificial intelligence (AI) (Chapelle, & Sauro, 2017; Nushi & Eqbali, 2018; Pleschova, 2015). Pronunciation, as an integral component of L2 proficiency, has its own share of technologies (e.g., Otterwave, PRAAT, Say It Out, Sounds: The Pronunciation App, and English File Pronunciation). For instance, PRAAT, a computer software designed by Paul Boersma and David Weenink of the University of Amsterdam, can analyze, synthesize, and manipulate different aspects of speech (e.g., pitch, formant, intensity, and voice quality), and generate high-resolution images for phonetic analysis of human speech¹. These are welcoming developments because up until the recent past, teaching and learning pronunciation was not a priority for L2 educators' agenda, mainly due to their lack of confidence and knowledge about how to teach the skill and its components (Celce-Murcia et al., 2010; Dao et al., 2020; Derwing & Munro, 2005; Szyszka, 2016; Thomson, 2012). Derwing and Munro (2005) contend that "other aspects of pedagogy [e.g., grammar, vocabulary] receive extensive attention in teacher preparation courses and materials, but in many instances, L2 instructors are apparently left to teach themselves how to address pronunciation with their students" (p. 389). L2 learners also find the acquisition of the skill challenging because, as stated by Celce-Murcia et al. (2010), pronunciation engages cognitive, affective as well as psychomotor domains. Technology, however, may assist in this regard. Reinders and White (2011) believe that:

Technology has the potential to not only provide access to resources for learning in a superficial sense but also to offer increased affordances for autonomous learning. Opportunities for interaction, situated learning, and support for learning outside formal contexts, have greatly improved because of technology. (p. 1)

On the role of technology in teaching and learning L2 pronunciation, Pennington and Rogerson-Revell (2019) maintain that "technology holds great potential for pronunciation training, particularly in terms of maximizing opportunities for practice and exposure to spoken language" (p. 272). Among the technological innovations, mobile and handheld devices such as smartphones can provide language learners with a ubiquitous learning experience as learners can use them almost everywhere, at any time, and for any kind of learning, including learning L2 pronunciation. Mobile-Assisted Pronunciation Teaching (MAPT) (Nushi & Razdar, 2018) has in fact brought "pronunciation training to the learners' fingertips is through the use of their own mobile devices" (Fouz-González, 2020, p. 63).

Therefore, the focus of this article is to describe and evaluate English Language Speech Assistant (ELSA) Speak (ESLA for short), an application that has been developed to facilitate English pronunciation teaching and learning. The flagship feature of this app

¹ PRAAT requires technical training to operate.

is its built-in automatic speech recognition technology (ASR) and the way it provides learners with instant feedback on their pronunciation accuracy at both segmental and suprasegmental levels. Perhaps this is the reason the app website names it “the world’s smartest Artificial Intelligence pronunciation coach” (<https://elsaspeak.com/en/>). According to its developers, the state-of-the-art ASR technology is a top choice since it is ranked amongst the top five best artificial intelligence apps. The app provides learners with a detailed evaluation of their pronunciation of words and sounds giving them a clear picture of their strengths and areas of improvement.

Description

ELSA is an EFL/ESL pronunciation development app designed based on United States English, with over 1200 exercises to help learners sound like native speakers. The software is used by millions of language learners worldwide. The data on Apple Store and Google Play Store suggests more than 15 million installations of the app². To use this app, learners must first download it from Google Play (<https://play.google.com/store/apps>) or App Store (<https://apps.apple.com>) and install it on their Android or iOS device. When the user first opens the app, they are presented with a list of 30 languages from which they have to choose their native language. The reason for this initial language selection is that the app is capable of adapting to the learners’ mother tongue and will be able to provide them with more accurate feedback and identify their top pronunciation challenges if it knows their native language. There is also an ‘other’ option for those whose native language is not included. Selection of that option takes language learners to the ‘assessment test’ where their pronunciation is evaluated across all sounds including soft /s/, strong /sh/, aspirated sounds, and the unstressed schwa sounds. The app then gives a detailed score for each part and provides an overall pronunciation score to indicate how closely learners speak compared to the native model.

ELSA is pre-programmed in a way that offers users 10 minutes of daily practice, though they can increase or decrease the time when they sign up. They also have the chance to choose their best time of the day for practicing and the app will send them daily reminders. When the user reaches this point, s/he just has to tap ‘Get started’, decide to allow the app to send them notifications or not, and then sign up. It is worthwhile to mention that there is also a premium version of ELSA, called ELSA Pro, offering in-app purchases which provide you with a multitude of exercises and unlimited practice but needs paying subscription. However, the first time one signs up, the app offers them a seven-day free trial and unlimited access to the whole curriculum after which users have to choose one of the subscription plans to have continued access to the ELSA contents. Users can create their account and sign up either with Facebook or with an email account provided by different services including Hotmail, Gmail, and Outlook. ELSA’s lessons are divided into two categories of ‘skills’ and ‘topics’. In the ‘skills’ part, one has separate sounds and various aspects of English pronunciation skills to practice and work on

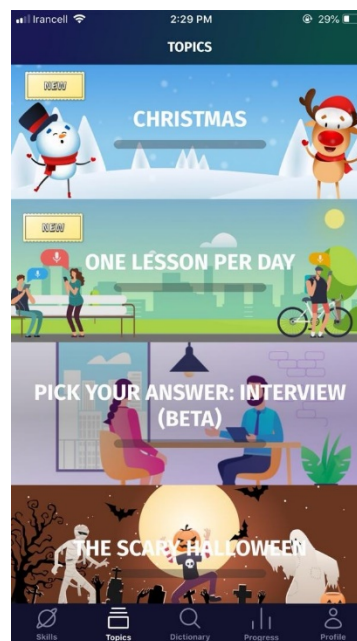
² Data accessed on April 20th, 2021.

independently (Figure 1), while the ‘topics’ part seeks to integrate pronunciation development lessons with different topics like Christmas, Halloween, music, movies, job interviews, and traveling (Figure 2).

Figure 1
Various pronunciation skills offered by ELSA



Figure 2
Topic samples offered by ELSA



Before learners choose any skill or topic to practice, the app offers them a way to learn faster (Figure 3). ELSA recommends learners take a five-minute quiz to find out what they need to learn and will provide them with a custom lesson plan. The quiz is the assessment test that the learners have the option to take, before or after they explore the app. Upon taking the quiz and after exploring the app for a while, the ‘learn faster’ option appears on the user’s screen. However, if one takes the test at the beginning, the scores they have achieved on each category appear on the side of each skill, making it easier for them to keep track of the troubled areas (Figure 4).

Figure 3
ELSA's offer for learning in a faster and more personalized way

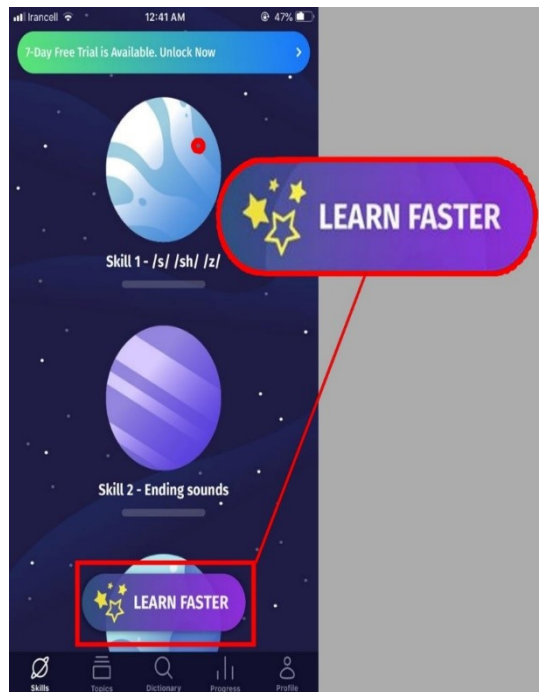


Figure 4
Placement test result for each skill



Whether one decides to take this quiz or not, they can choose any skill they wish to practice. When they do so, they will see a page like Figure 5. The question mark icon at the top right of the screen gives a detailed description of what the sounds are, what phonetic features they have (e.g., voiced or voiceless), and how to pronounce them correctly (e.g., the position of the lips, tongue and vocal cords) (Figure 6). The microphone icon in Figure 5 notifies the users to be ready to pronounce something, the headphone icon indicates that they are going to listen to something to check how well they can distinguish among two choices they have heard. The bar graph icon shows the lesson is focusing on stress patterns, and the chat bubble icon illustrates that the lesson is going to be in a conversation format. One more point needs to be mentioned and that is the difficulty level of the lessons. As seen in Figure 5, lessons are introduced as 'easy' but one can change the difficulty level to 'difficult' if they wish to, based on instructions that will be provided in a later section of this article.

Figure 5
Features of a lesson



Figure 6
The description provided on Schwa sound

Schwa: /ə/

The most common vowel in English is called schwa, and its symbol is /ə/ (sometimes people write it uh). This sound is only found in unstressed syllables. It's the first vowel in the words around and together. The word America has two schwas: one at the beginning and one at the end.

A common mistake is to make the schwa too strong and change it to another vowel. For example, some people accidentally pronounce the first vowel in around with a strong /eɪ/, so it sounds like ay-round. Similarly, some people make the mistake of saying /u/ instead of /ə/ for the first vowel in together, so it sounds like too-gether.

Since this is such a common sound, it's definitely worth practicing! Mastering the schwa will also help you with your stress skills (making some syllables stronger and some weaker).

When the learners select a lesson with a microphone icon next to it, it means that they are going to say out a word or a combination of words, so that the app can test their pronunciation (Figure 7). Whenever they feel ready to speak, they just have to tap the microphone icon at the bottom of the screen and repeat the word, phrase, or sentence that appears on the screen. As Figure 8 shows, the app scores the user's overall pronunciation in comparison to a native speaker's pronunciation. The software also scores the separate sounds one has produced. The green-colored letters indicate that the individual sounds have been pronounced perfectly, the yellow-colored ones mean that these sounds do not represent the native model and could have been pronounced better, and the red-colored ones mean that the sounds were pronounced inaccurately. A fascinating opportunity learners are presented with is that pressing each individual word also enables them to have a complete evaluation of their pronunciation of that word and its individual sounds, and the nearly correctly and incorrectly pronounced sounds are complemented with a detailed explanation of what learners need to do to produce those sounds (more) accurately (Figures 9 & 10). Of other notable features of this app is that one can hear what one has said by pressing the ear icon, and then save a lesson to the word bank for further practice by pressing the bookmark icon (Figure 8).

Figure 5
Inside a pronunciation development lesson

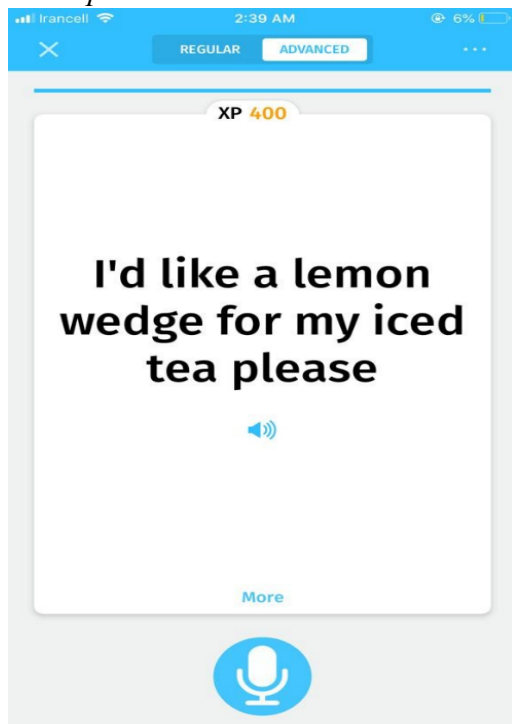


Figure 6
ELSA's pronunciation evaluation

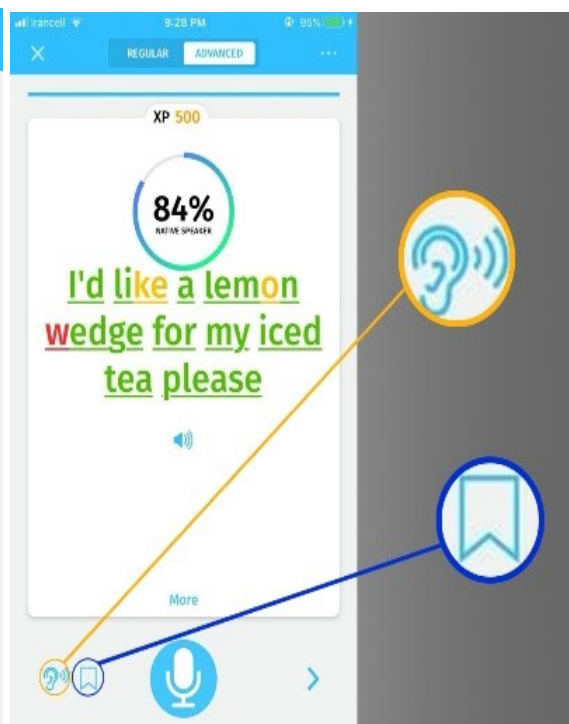


Figure 7
Detailed pronunciation evaluation of individual words and sounds

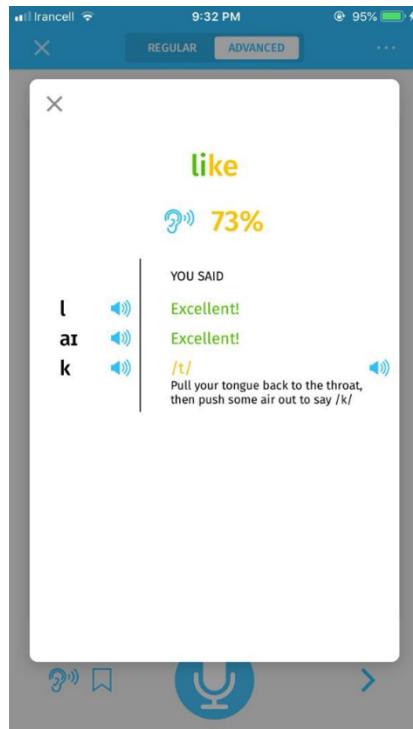
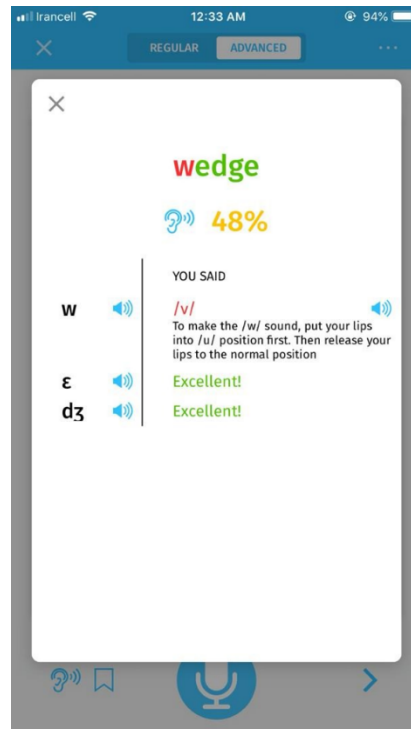
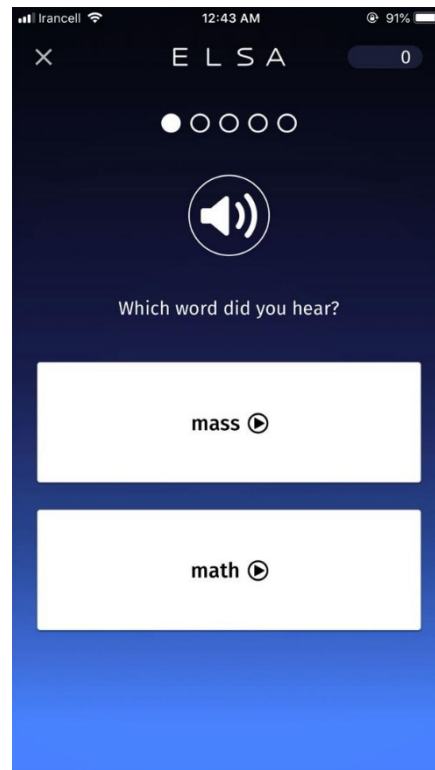


Figure 8
Detailed pronunciation evaluation of individual words and sounds



Selecting a lesson with a headphone icon next to it shifts the test from checking the learners' pronunciation ability to their ability to distinguish the sound they hear. As shown in Figure 11, after the learner has heard a word, the app presents him or her with two choices among which they have to choose the one they heard.

Figure 9
ELSA's sound recognition evaluation



Choosing a lesson with a bar graph icon next to it allows users to practice the stress patterns. A few of these words have a definition, an example, an occasional illustration if possible, and the International Phonetic Alphabet (IPA) attached to their 'more' option. The features and procedures are similar to the ones explained earlier, with the difference that here the users are scored based on their ability to place the stress on the correct syllable (Figures 12 & 13).

Figure 10
ELSA's stress pattern instruction

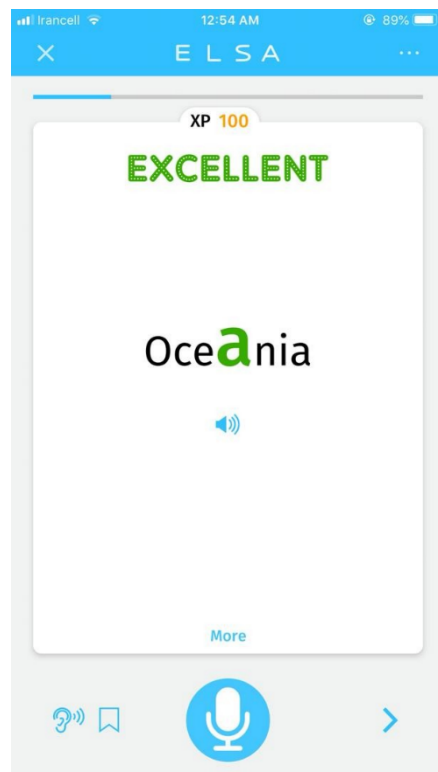
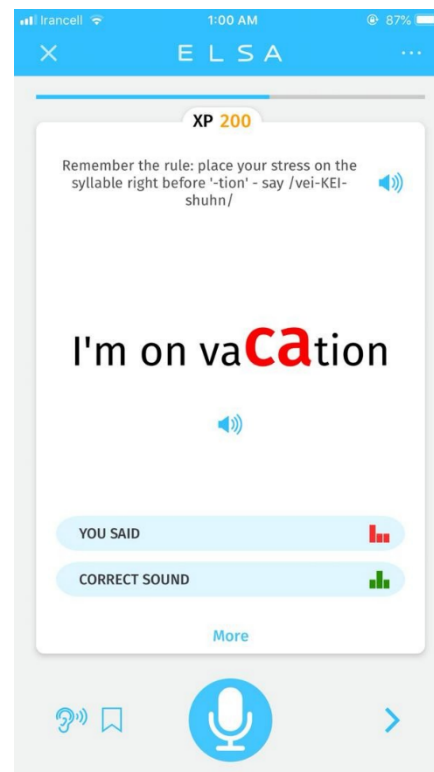
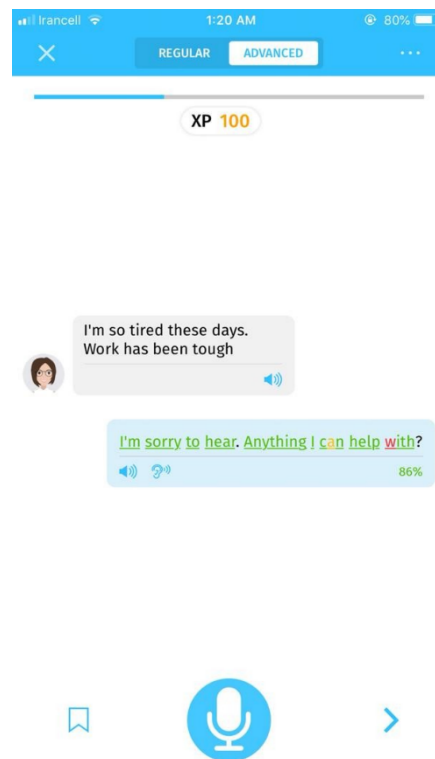


Figure 11
ELSA's stress pattern instruction



Selecting a lesson with a chat bubble icon means the learner's pronunciation exercise is going to be conducted in the form of a conversation between him or her and ELSA. The app starts the conversation with one statement or question, and continues it by exposing the learner to the sentence he or she has to say in reply, so the only thing one has to do is to read out the sentence presented to them. The app then checks their pronunciation. Procedures are again the same: an overall score, as well as a detailed scoring on individual words and sounds, are provided to the learners (Figure 14).

Figure 12
ELSA's evaluation of pronunciation in a lesson with a conversation format



As a final point, ELSA includes additional facilitative tools such as a ‘dictionary’ with a box into which users can type anything, from a single word to a complete sentence, that they want to say but are unsure as to how to pronounce them, and the app will pronounce the typed input for the users. Other features include a ‘progress’ part where learners can check their daily progress, find the words they have stored, the ones they have mastered and the ones which need more work (Figure 15), and a part called ‘profile’ in which one can manually change some options like the time of their daily practice and the difficulty level of the scoring system (Figure 16). Note should be made that the app offers only two difficulty levels (easy and difficult) with nothing in between.

Figure 13
Inside the progress part

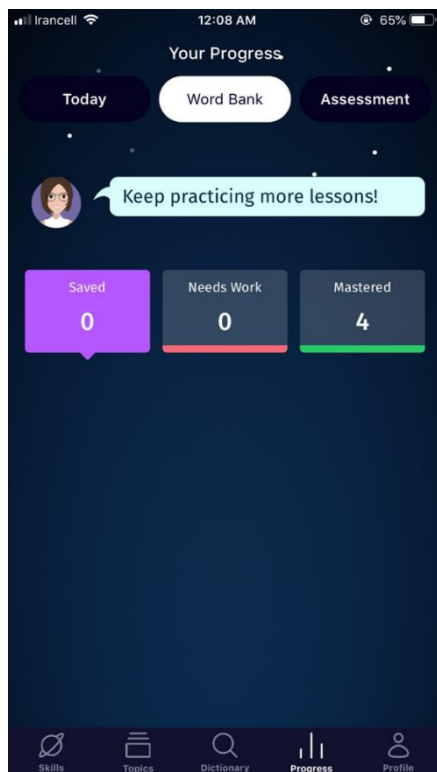
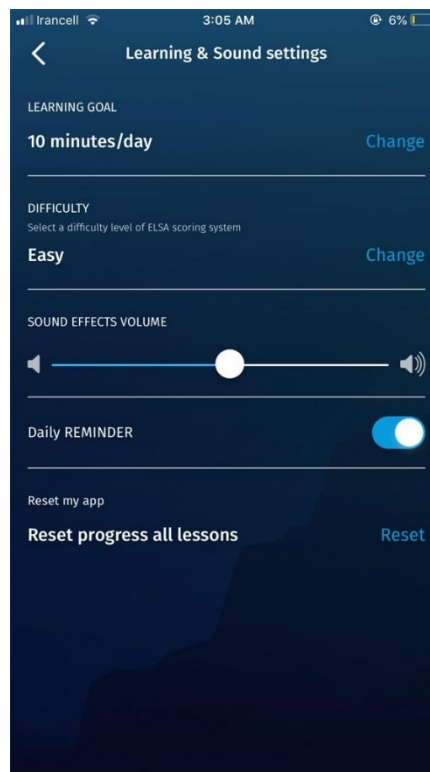


Figure 14
Inside the profile



ELSA also introduces a web widget called ‘ELSA Instant’ for the learners to know how a word is pronounced while surfing the net. One could use this feature from either Google Chrome web store or as an Add-on for Firefox. Users can simply right-click one word to hear a sample or to record themselves for instant feedback.

Evaluation

Language learning applications often have their own share of merits and demerits and ELSA is no exception. After an in-depth review of this application, we can claim that ELSA’s advantages outweigh its disadvantages. ELSA’s accurate speech recognition technology, in combination with the advanced AI implemented, helps users recognize the slightest of differences that exist between their pronunciation and the native model. The app also tries to cater to individual pronunciation needs by allowing them to select their native language in order to provide more effective instruction and feedback. Furthermore, ELSA users can adjust the difficulty level of the prescribed exercises. The application enjoys a number of advantages when compared with other pronunciation apps available on market such as *Otterwave*, *PRAAT*, *Say It Out*, *Sounds: The Pronunciation App*, and *English File Pronunciation*. Although these software programs can have their own potential contribution for certain learning purposes and contexts, they suffer a number of

limitations including the need for technical training or assistance to interpret the feedback offered (e.g., in the case of PRAAT), the lack of specific instructions on how to improve the problem areas detected (e.g., in the case of Say It Out), or the inability to provide feedback on connected speech (e.g., in the case of Sounds: the Pronunciation App).

Despite its merits, the app could benefit from a number of improvements. The first and most significant problem is that English is the only language supported by ELSA. There are many other languages in the world that are popular with language learners and it would be a great development if ELSA expanded its list of supported languages. The second problem is the app's reliance on the US English accent as the evaluation yardstick. This US-centric accent can constitute problems for users, especially for those interested in other varieties of English accent (e.g., British and Australian) as they cannot determine how their pronunciation compares with those varieties. Besides, depriving English language learners of exposure to other varieties of English language accents goes against the emerging paradigm of English as an International Language (EIL) (Crystal, 2012; McKay, 2018). English is no longer limited to one or even a few countries in the world. This recognition is evident in the international English exams such as IELTS and TOEFL which no longer base their evaluation on any specific English variety and accent. Educational Testing Service (ETS), for instance, has begun adding other native-speaker English accents (accents from the United Kingdom, New Zealand, or Australia) in the Listening and Speaking sections of the TOEFL iBT test since March 2013 to better reflect the variety of native English accents that students may be exposed to while studying abroad. It seems to be a pragmatic approach if ELSA developers provided separate options regarding different English accents among which students could have their pick of accents.

Second, while ELSA has taken some steps toward individualization of instruction (e.g., allowing users to choose their native language) to provide learning materials that best suit the user's needs, its curriculum remains quite prescribed and learners do not have much control over the content of the lessons. ELSA users are not given the facility of choosing activities that better meet their learning styles. ELSA's instructional approach is reminiscent of the behavioristic learning theory wherein the accurate production and comprehension of oral language with little or no emphasis on the context are demanded. However, that approach might be inevitable given the fact that pronunciation is mainly a psychomotor skill that requires plenty of practice to improve.

Next is the premium version of ELSA for which users have to pay fees. It has to be noted that ELSA's premium version does not offer features that the free version does not; the only difference is the larger number of exercises the premium version includes. The free content available on ELSA is relatively small and obtaining access to the premium version may not be affordable for all. ELSA can take some steps to ameliorate the situation; for instance, users can be asked to pay subscription fees based on the countries they live in and variant charges could be offered to learners from low-income, middle-income, and high-income countries. Another possible solution could be to use advertisements on the free version to compensate for the expenses, that is, learners can watch advertisements and then get access to the content available on the premium version.

This could allow learners in low-income countries or in countries where international payment services are not available to benefit from the services provided by ELSA.

Finally, users might appreciate the addition of pictorial aids (e.g., visemes or images and videos of real speakers) that visually demonstrate the mouth movements to further help accurate pronunciation. Google Dictionary, for instance, has newly added an innovative feature to the pronunciation of vocabulary to improve the educational quality of its pronunciation entries. This feature allows users not only to hear the words but also to see how they are pronounced graphically via visemes and generic facial images (Nushi & Moradi, 2020).

Despite the above-mentioned disadvantages, we would like to underline the fact that ELSA is a user-friendly application that can prove beneficial to both EFL/ESL teachers and learners who wish to teach or acquire the English pronunciation with a great rate of satisfaction.

Conclusion

The significance of pronunciation in learning an L2 is indubitable. As pointed out by Harmer (2001), although grammar and vocabulary are important elements of successful communication, they would be of no or little use if the speaker is not able to pronounce the language properly. Furthermore, intelligible pronunciation does not only include the ability to pronounce discrete sounds but also other types of competencies including stress, rhythm, linking, and assimilation (Celce-Murcia et al., 2010; Oxford & Scarcella, 1994). Pronunciation development is also a process that learners usually go through individually. For these reasons, well-designed instruction and an adequate amount of time need to be dedicated to pronunciation development.

Technological innovations like ELSA are a great aid in this regard. ELSA provides EFL/ESL learners with a ubiquitous learning environment wherein learners can learn and assess their pronunciation from anywhere and anytime provided that they have an internet connection. It also furnishes them with a range of tasks that focus on segmental and suprasegmental features of spoken English. Attention to the segmental and suprasegmental features of speech (e.g., intonation, connected speech, and sentence stress) is crucial since these features, particularly the suprasegmental features, tend to go unnoticed without explicit instruction and are more likely to fossilize in the L2 learners' interlanguage (Fouz-González, 2020). Teachers can also assign students these tasks as homework too, and as a result, less time would be needed to cover pronunciation lessons in the classroom. In other words, ELSA can serve as an 'assistant' to teachers. This also can assure teachers that pronunciation is not neglected in the learners' linguistic diet. Besides, teachers are not obliged to search for reliable materials as they can have access to a valid source with plenty of exercises. ELSA's reports and scoring system is so detailed and accurate that it can provide teachers with a complete assessment of their students' pronunciation proficiency. We recommend that teachers seize this opportunity to experience a more productive pronunciation development schedule.

Two main limitations confronted this review: first, the review was based on the two authors' evaluation of ELSA. To obtain a more accurate assessment, a more comprehensive analysis utilizing a variety of research methods, from case studies and surveys to experiments, and involving EFL/ESL teachers and students might be required. Second, the writers of this review did not have access to the in-app international payment services provided by ELSA and thus were unable to access the premium version. The limited access did not allow them to compare the free and paid versions of the app.

References

- Burston, J. (2013). Mobile-assisted language learning: A selected annotated bibliography of implementation studies 1994-2012. *Language Learning & Technology*, 17(3), 157-224. https://scholarspace.manoa.hawaii.edu/bitstream/10125/44344/17_03_burston.pdf
- Celce-Murcia, M., Brinton, D. M., & Goodwin, J., M., & Griner, B. (2010). *Teaching pronunciation: A Course book and reference guide*. Cambridge University Press.
- Chapelle, C. A., & Sauro, S. (Eds.) (2017). *The handbook of technology and second language teaching and learning*. Wiley-Blackwell. <https://doi.org/10.1002/9781118914069>
- Crystal, D. (2012). *English as a global language*. Cambridge University Press. <https://doi.org/10.1017/CBO9781139196970>
- Derwing, T. M., & Munro, M. J. (2005). Second language accent and pronunciation teaching: A research-based approach. *TESOL Quarterly*, 39, 379-397. <https://doi.org/10.2307/3588486>
- Dao, P., Nguyen, M. XNC, & Nguyen, C. (2020). Effect of pronunciation instruction on L2 learners' listening comprehension. *Journal of Second Language Pronunciation*, 7(1), 10-37. <https://doi.org/10.1075/jslp.19012.dao>
- Fouz-González, J. (2020). Using apps for pronunciation training: An empirical evaluation of the English File Pronunciation app. *Language Learning & Technology*, 24(1), 62-85. <https://doi.org/10125/44709>
- Harmer, J. (2001). *The practice of English language teaching*. Longman.
- McKay, S. L. (2018). English as an international language: What it is and what it means for pedagogy. *RELC Journal*, 49(1), 9-23. <https://doi.org/10.1177/0033688217738817>
- Nushi, M., & Eqbali, M. (2018). 50Languages: A mobile language learning application. *Teaching English with Technology*, 18(1), 93-104.
- Nushi M. & Moradi A. (2020). Google dictionary: A critical review. *Issues and Trends in Learning Technologies*, 8(1), 3-16 doi: https://doi.org/10.2458/azu_itlt_v8i1_nushi
- Nushi, M., & Razdar, M. (2018). Sounds: The pronunciation app. *Roshd Journal of Foreign Language Teaching*, 32(2), 24-31.
- Oxford, R.L., & Scarcella, R.C. (1994). Second language vocabulary learning among adult learners: State of the art in vocabulary instruction. *System*, 22(2), 231-243. [https://doi.org/10.1016/0346-251X\(94\)90059-0](https://doi.org/10.1016/0346-251X(94)90059-0)

- Pennington M. C., & Rogerson-Revell P. (2019). Using technology for pronunciation teaching, learning, and assessment. In M. C. Pennington & Rogerson- Revell (Eds.), *English pronunciation teaching and research. Research and practice in applied linguistics* (pp. 235-286). Palgrave Macmillan. https://doi.org/10.1057/978-1-137-47677-7_5
- Pleschova, G. (2015). Integrating technology into the classroom. In J. Ishiyama, W.J. Miller & E. Simon (Eds.), *Handbook on teaching and learning in political science and international relations* (pp. 291- 303). Edward Elgar. <http://irep.ntu.ac.uk/id/eprint/37341>
- Szyszka, M. (2016). English pronunciation teaching at different educational levels: Insights into teachers' perceptions and actions. *Research in Language*, 2, 165-180. doi: 10.1515/rela-2016-0007
- Thomson, R. I. (2012). Demystifying pronunciation research to inform practice. In H. M. McGarrell & R. Courchène (Eds.), *Special research symposium issue of CONTACT*, 38(2), 63-75.