INTEGRATING DIGITAL TOOLS IN PRONUNCIATION INSTRUCTION: A REPERTOIRE APPROACH

DJ Kaiser, Webster University

Digital tools, such as computer- and mobile-based applications, can support English pronunciation learning (Kaiser, 2018; Neri et al., 2002). Focusing on one or two favorite tools may be too narrow in scope. This article advocates for pronunciation instructors to take a repertoire approach to support learners by drawing on a more diverse set of digital tools, promoting practice for a broader range of pronunciation features. This approach allows instructors and learners to address varied segmental and suprasegmental features, promoting exploration of authentic language and the variability and context-dependent nature of English pronunciation. The tools discussed include mobile devices, messaging apps, YouGlish, TED Talks, YouTube, the Corpus of Contemporary American English (COCA), and artificial intelligence tools.

INTRODUCTION

In the 1980s, Shulman (1986, 1987) differentiated content knowledge from pedagogical knowledge and defined their combination as "pedagogical content knowledge, that special amalgam of content and pedagogy that is uniquely the province of teachers, their own special form of professional understanding" (1987, p. 8). Three decades later, Mishra and Koehler (2006) built on Shulman's concept, adding the technological component with technological pedagogical content knowledge (TPCK).

As instructors of pronunciation, we have our content knowledge (phonetics, phonology, segmentals, suprasegmentals, etc.), our pedagogical knowledge (listen & repeat, communicative activities, physicalizations or haptics, gadgets, etc.), and technological knowledge (recorders, computer programs, mobile apps, etc.). The technological pedagogical content knowledge (TPCK) for pronunciation instruction requires that we match available digital tools to strong pedagogical uses for practicing and acquiring features of English pronunciation. The use of technology in pronunciation instruction dates back to at least the late nineteenth century with the first phonograph (invented in 1878), which Edison (1888) argued could "reproduce his utterance even more exact than a human imitator could" (p. 647). Technology for pronunciation instruction has evolved over more than a century, with rapid developments in the past few decades, particularly with computers and mobile devices. The increased availability of websites, computer programs, and mobile apps has made it a challenge for instructors and learners to identify the best digital resources. Within a TPACK framework, this requires aligning technological tools with pedagogical practices to address persistent challenges in pronunciation instruction (e.g., limited access to authentic materials, difficulty teaching suprasegmental features).

Selecting appropriate technology and digital tools can be a challenge. Some tools may have limited uses (often focused only on sound- or word-level features), while other tools can be expensive to purchase or use (increasingly requiring expensive subscriptions). This article

Kaiser, D. (2026). Integrating digital tools in pronunciation instruction: A repertoire approach. In P. Trofimovich, A. Bodea, T.-N. N. Le, R. Suzuki, & C. Teló (Eds.), *Proceedings of the 16th Pronunciation in Second Language Learning and Teaching Conference* (pp. 1–10). Concordia University, Montreal, Canada, October 2025. https://doi.org/10.31274/psllt.21298

focuses on widely available free (or freemium) digital tools that can be used by instructors and learners. Freemium tools are those tools that provide free use up to a limit (often a daily limit), after which the user must pay or subscribe. Even free tools can lead to disappointment when their use is limited (e.g., focused on one specific sound contrast or one specific pronunciation feature). This article applies the TPCK framework to Computer-Assisted Pronunciation Training, or CAPT (Neri et al., 2002) and Mobile-Assisted Pronunciation Training, or MAPT (Kaiser, 2018), taking a repertoire approach to digital tools.

This article adopts a repertoire approach to digital tools, viewing them not as isolated resources but as a flexible set of pedagogical options that instructors and learners can draw upon to address different pronunciation features, learning goals, and communicative contexts. A focused repertoire allows instructors to prepare learners to use these tools effectively and to become familiar with their affordances for pronunciation practice. The recommended digital tools in this article and their potential uses were compiled as part of a comprehensive review of pronunciation instructional materials (Kaiser, in press). This list of digital tools with suggested pedagogical uses demonstrates how a repertoire approach can be comprehensive (covering the full range of topics covered in an advanced pronunciation textbook) yet not require expensive download or subscription fees.

This repertoire approach can also help avoid the common pitfall of focusing CAPT and MAPT tools on segmentals or word-level pronunciation to the exclusion of suprasegmentals. An additional benefit of using digital tools that can address a broader range of pronunciation features (e.g., YouGlish) is that learners can become familiar with the tool through exploration of word-level segmentals, which are typically more manageable. Instructors can scaffold the use of digital tools with affordances that support suprasegmental practice. The careful sequencing of tasks using tools from this repertoire allows instructors to leverage the TPCK framework in aligning technological affordances with pedagogical goals. The digital tools and suggested uses in the next section provide an example of this repertoire approach (also see Kaiser, in press, for an example of this approach aligned to a full textbook).

DIGITAL TOOLS FOR PRONUNCIATION LEARNING

Mobile Devices

Dickerson (1983) advocates for covert rehearsal, where pronunciation learners practice pronunciation features and read practice items in private, often by recording their voice for their instructor to review (both for accountability and to provide formative feedback). Practice that once required cassette recorders or a computer with audio recording capabilities can now be done using the mobile devices that many learners carry in their pockets. Today's mobile devices provide options for both audio and video recording.

Suggested uses: Learners can record themselves reading the same passage at regular intervals (e.g., weekly, monthly). Many mobile devices will automatically provide a date and time stamp for each recording, making it easier to track recordings over time. It may even be possible to change the file name for recordings or to put them into folders. Audio recordings may be sufficient for practicing segmentals (consonants and vowels), but video recordings are

recommended for practice with suprasegmentals (e.g., rhythm, phrase stress) that incorporate the use of gestures.

Messaging Apps

The mobile devices that our learners carry also have access to messaging apps, such as WhatsApp, Telegram, WeChat, Line, and Viber. Most message apps allow users to record their voice and send audio messages. This affordance allows pronunciation instructors to develop activities that engage learners in exchanging pronunciation practice through audio recordings. Both one-on-one (pair work) and group chat options are possible, in addition to assigning learners to send the instructor audio files with their pronunciation practice (through covert rehearsal).

Suggested uses: Instructors could create a messaging group for their learners. Each week or after each class session, the instructor could select one or two exercises or activities for the learners to record and send to the messaging group. This could range from the reading of a list of words or sentences to more communicative tasks, such as describing the members of their family (consider how often /ð/ may occur for this task) or leaving a friend a voice message about planning for an upcoming party (consider how often questions and different intonation patterns may occur for this task). It may help to make explicit the pronunciation feature(s) in focus for each task.

YouGlish

YouGlish (https://youglish.com) allows users to search for words or phrases in multiple languages, including English, and play clips from thousands of videos to hear how speakers pronounce the searched word or phrase. YouGlish works on the freemium model, where users are allowed ten free searches per day. YouGlish offers tiered monthly subscriptions for additional daily searches. YouGlish allows you to search for words and phrases. There is also the option to select the gender of the speaker by adding ":m" for male speakers and ":f" for female speakers. YouGlish focuses on native English speakers and also allows users to select which varieties of English they want included in the video search results by using the options directly below the search bar. Options include the United States, the United Kingdom, Australia, Canada, Ireland, Scotland, and New Zealand. Users can control the speed of the video by increments of five percent. This speed control can be especially helpful for learners who need a little more processing time. The YouGlish website also allows you to record yourself practicing passages from the videos. Learners can listen to their own recordings and compare them to the speakers in the videos.

Suggested uses: Below are some suggested uses of YouGlish aligned to various features of English pronunciation.

- Single word practice – Provide learners with single words to search for, listen to, and practice with the focus on target segmentals, word stress, –s endings, or –ed endings.
- Contractions – Search for different contractions and listen to how various speakers pronounce them in context.

- Noun units – Search for different compound nouns (e.g., “blood pressure”), complex compound nouns (e.g., “blood pressure monitor”), or noun strings (e.g., “Memorial Hospital”) to see how various speakers put stress on these noun constructions.
- /h/-trimming – Search for words that commonly undergo /h/-trimming (e.g., “he,” “him,” “his,” “her,” “have,” “has,” “had”) to see when speakers pronounce the /h/ sound and when they trim it.
- Consonant clusters and potential trimming – Search for words with final consonant clusters ending in /t/ or /d/ (e.g., “wrapped,” “used”) to listen for when speakers trim the final /t/ or /d/ and analyze the phonological environment for the /t/ or /d/ sound.
- Syllabification – Search for words that may undergo /l/ and /n/ syllabification (e.g., “puddle,” “madden”) to listen for when speakers syllabify the /l/ or /n/ sounds.
- Palatalization – Search for pairs of words that can result in palatalization (e.g., “meet you,” “did you,” “miss you”) to see when speakers palatalize these word-final consonants.
- Linking – Search for pairs of words that can result in linking (e.g., “catch up,” “do it”) to see when speakers use linking.
- Intonation – Search for “?” to have YouGlish provide examples of phrases and sentences marked as interrogatives. It is also possible to search for specific question words along with the question mark (e.g., “who ?,” “what ?,” “where ?”) or common word combinations found in questions (e.g., “is it?,” “isn’t it?,” “did you?”).

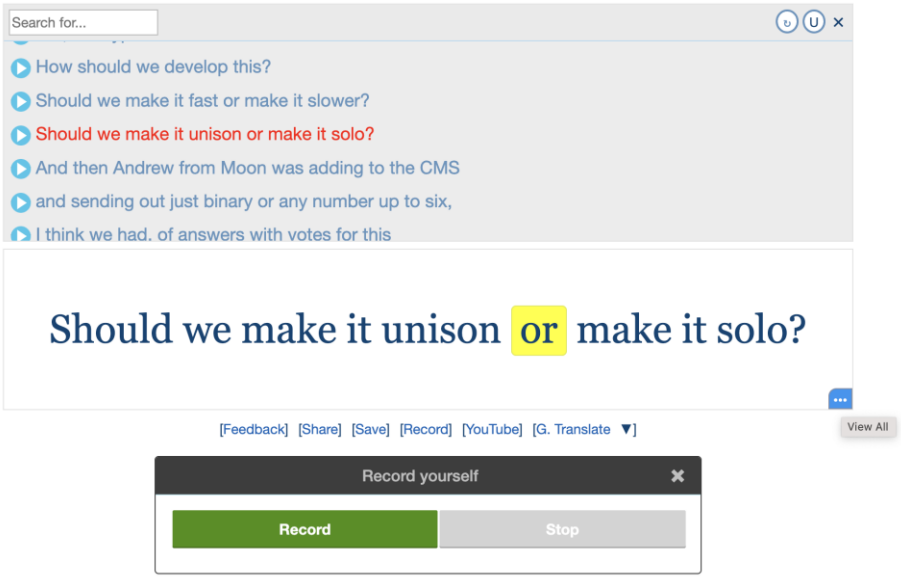
Note that YouGlish allows users to use negative prompts by including the “-” sign before words you wish to exclude. This can be especially helpful when trying to focus on specific kinds of questions. To increase the chances of choice questions in the search results, try the following prompt: “? -who -what -where -when -why -how.” This will exclude interrogatives that include a question word. If you want to focus on yes/no questions with “is it?” try the following prompt: “is it? -who -what -where -when -why -how.” This will exclude interrogatives that include “who is it,” “what is it,” etc. See Figures 1 and 2 for examples of a YouGlish search focused on choice questions. Notice the inclusion of “or” and “?” and the request for male speakers (with “:m”). The negative prompts exclude question words from the search results.

Figure 1

Search Results for a Very Specific Prompt

Figure 2

Transcript with the Recorder and Other Features for the Video in Figure 1



YouGlish also includes transcripts of each video (click the three dots in blue on the bottom right for a full transcript of the video). Underneath the transcript is an option to translate, which will

bring up portions of the transcript into one of four translators that users can choose from. This allows learners to translate portions of the video's transcript into more than one hundred different languages. There is also a "Save" feature, so learners can save a particular search for later review and practice. And finally, there is a "Share" feature, which allows a user to copy a link from their search results and send it to another person. This share feature allows instructors to share search results with their learners. Another application would be for instructors to assign their learners to explore YouGlish and send one or more links from their search results that focus on the target pronunciation feature.

TED Talks

TED Talks (https://www.ted.com/talks) is a free website that hosts thousands of educational and inspirational videos of numerous topics. Unlike YouGlish, TED Talks include both native and nonnative English speakers, representing a wider range of Englishes spoken in the world. English learners can greatly benefit from hearing and watching other nonnative English speakers model good pronunciation during public speaking situations. Similar to YouGlish, transcripts of videos are available. TED Talks allow you to control the speed but by increments of 25% (in contrast to YouGlish, which allows increments of 5%).

Suggested Uses: TED Talks can be particularly useful for observing how diverse speakers use suprasegmental features, such as pausing, rhythm, phrase stress, and intonation. Use of gestures, particularly beat gestures, can facilitate the auditory aspects of rhythm and stress (Biau & Soto-Faraco, 2013; Hardison, 2018; Hardison & Pennington, 2021). Instructors could assign learners to watch TED Talks and observe speakers' use of gestures with rhythm and stress (or any other target features). This activity could be coupled with an assignment for learners to record themselves speaking to model the same feature (e.g., gestures with rhythm and stress) on their mobile device.

YouTube

Similar to YouGlish and TED Talks, YouTube (https://www.youtube.com) provides access to thousands of videos. Videos on YouTube have not undergone the curation used to organize video content in YouGlish or TED Talks. While YouTube can provide a greater variety of video content, the quality and suitability for instruction can vary significantly.

Suggested uses: YouTube videos can be particularly useful for international teaching assistants (ITAs) who are working on their pronunciation skills as needed for classroom instruction. Learners can search for university lectures (e.g., "chemistry lecture") and identify individual or a series of videos that show professors or instructors delivering class lectures or leading discussions. ITAs may want to focus on features such as pauses, questions, rhythm, stress, gestures, and pronouncing key terms and phrases. Note that the use of YouTube may require more involvement from the instructor than YouGlish and TED Talks.

The Corpus of Contemporary American English

The Corpus of Contemporary American English (COCA) allows users to search English-language texts and transcripts spanning five main genres: spoken, fiction, popular magazines, newspapers, and academic journals (Davies, 2010). An advantage to COCA over other digital tools covered so far in this article is the ability to search by parts of words based on spelling patterns. This can be helpful when teaching word stress through rules based on orthography (Dickerson, 1978, 2018). COCA was used to verify and revise a set of word stress rules for the forthcoming *Practical Pronunciation: Pronunciation Strategies in Practice* (Kaiser, in press) and can be used to create lists for word stress practice, such as searching for "*ic" (for "democrátic" and "electrónic") or "*ia*" (for "espécially" or "associátion").

A particular advantage to COCA is that it can link search results to YouGlish. This means that while YouGlish does not allow for searching by spelling patterns or parts of words, COCA does allow this and will connect its search results to YouGlish for video content with specific words from the search results (click PRON/VIDEO above search results in COCA to activate the search results links to YouGlish results for the same words). Using COCA for these purposes can help instructors and learners identify words with a higher frequency for practice.

Suggested uses: The forthcoming textbook, *Practical Pronunciation: Pronunciation Strategies in Practice* (Kaiser, in press), includes the following for word stress rules, which align to the following search prompts that can be used in COCA:

- Word Stress Rule #1: *ic, *ical
- Word Stress Rule #2: *ate, *ated, *ating, *ator
- Word Stress Rule #3: *ia*, *io*, *ua*, *uo* (note the use of * before and after)
- Word Stress Rule #4: *y (but exclude –cy, –ry, and –ly words), *ied (but exclude –cied, –ried, and –lied words)

This same search process with COCA (linked to YouGlish) can also be used for strong prefixes, such as "anti–," "bio–," and "neuro–," by using the following search prompts: anti*, bio*, and neuro* (see Kaiser, in press, for a complete mini-lesson on these and more strong prefixes)

Artificial Intelligence

Artificial Intelligence (AI) has long had applications to pronunciation training through automated speech recognition (ASR) and text-to-speech (TTS). Many mobile and digital devices provide access to Siri, Gemini, ChatGPT, Alexa, and other AI tools that can transcribe what a user says. The accuracy of these tools is debatable (see Kaiser, 2018, 2024 for reviews of mobile apps that use ASR). A newer feature with AI tools is voice chat. At the time of writing this article, the "voice mode" from ChatGPT (GPT-5) and "Gemini Live" (Gemini 2.5 Flash) were two readily available voice chat tools on mobile devices. While these voice chat tools cannot yet provide accurate feedback on pronunciation, they can be an option for communicative practice.

Suggested uses: Provide learners with the prompt for a role-play activity. Each learner must read the prompt into the voice chat tool and then perform the role-play with the AI tool. A suggested beginning for the prompt is as follows: "I am practicing my English pronunciation; work with

me on the following role-play; you will be Person A and I will be Person B." After this initial framing prompt, then read the prompt for the role-play. This will allow learners to practice role-plays whenever they like and to repeat the role-play practice.

DISCUSSION

Integrating the Repertoire into Instruction

A repertoire of digital tools is more effectively developed gradually. In practice, no more than one or two digital tools should be introduced in a single class session. First, the instructor should model how to use the digital tool, then provide learners time to explore using the tool in pairs or small groups. Clearly defined out-of-class assignments can help learners develop autonomous skills using the digital tools in this repertoire. Within a repertoire approach, the goal is for learners to become familiar with the affordances of each tool and to use them independently to practice and explore both segmental and suprasegmental features of English pronunciation, including those that extend beyond classroom instruction.

Addressing Usage Limits

Many of the digital tools recommended in this article have daily limits (e.g., YouGlish, COCA, and voice chat with ChatGPT and Gemini). Spreading out practice and assignments with these digital tools across several days can help address these limitations. If instructors use digital tools as part of an assignment to do outside of class, they may want to provide additional time (more days) to complete the assignment. It may help to remind students about daily limits with some tools and not to wait until the final day to work on their assignment. Working with a variety of these tools can also help address daily limits. It is also helpful to recommend that learners engage with these tools for short periods of time each day.

Authentic Language

An advantage of many of these digital tools is their focus on authentic language. Video websites, such as YouGlish, TED Talks, and YouTube, focus on video content with authentic language. TED Talks focus on presentations, while YouGlish videos include many interviews. Many YouTube videos and some YouGlish videos may use scripted language, so instructors may need to be more selective with these tools. COCA uses all authentic language in its corpus. The distinction is that the majority of language found in COCA is written discourse, which can differ from authentic spoken language. While spoken language is included in COCA, this language is represented in a transcript format that often removes all aspects of pronunciation, such as linking, trimming, reductions, stress placement, etc. The advantage of integrating written discourse in pronunciation instruction (from a tool like COCA that can connect to YouGlish) is that many learners first encounter new words through the written form (e.g., in academic or technical readings).

The use of newer AI tools, particularly voice chat, is a new advance, allowing learners opportunities to produce their own authentic speech. Generative AI rarely provides the same output twice, meaning that learners can repeat the same role-play and should find themselves

required to produce new authentic language to address the responses from the voice chat tool. Voice chat from AI tools is a newer area that requires more attention from instructors to develop appropriate prompts. More research is needed in the use and effectiveness of voice chat for pronunciation practice.

CONCLUSION

With a repertoire approach, there is no one "best" digital tool for practicing pronunciation. If the goal is for learners to practice a broader range of pronunciation features and skills (especially beyond segmentals), then instructional time must be used effectively to integrate digital tools to meet these objectives. This requires pronunciation instructors to develop their technological pedagogical content knowledge (TPCK) in order to select and sequence appropriate digital tools with engaging tasks and activities. At the same time, focusing on free tools and limited use of freemium tools can meet the needs of price-sensitive learners. Relying on a repertoire of tools can also alleviate the limitations of freemium licenses by diversifying the tools used during class and for out-of-class practice.

This article demonstrates how a repertoire approach to digital tools for pronunciation instruction can promote learners' exploration of language in use beyond the sound and word levels. Pronunciation materials and textbooks often present pronunciation features as simplified rules with little variation. Using digital tools to explore these features can help learners appreciate the range of possibilities for pronunciation features, such as trimming, linking, palatalization, and stress placement. Ultimately, learners can use this repertoire of digital tools to analyze, question, and even critique many of the rules and patterns that they have learned through their pronunciation instruction, supporting their continued development beyond the classroom.

ABOUT THE AUTHOR

DJ Kaiser (djkaiser@webster.edu) is a professor and the director of TESOL at Webster University. Kaiser is a two-time Fulbright recipient and has completed three English Language Specialist projects for the U.S. Department of State, including a project to prepare English teachers in Brazil with the skills to use technology to teach English pronunciation. His textbook, *Practical Pronunciation: Pronunciation Strategies in Practice*, is scheduled for publication in the summer of 2026 by the University of Michigan Press.

REFERENCES

Biau, E., & Soto-Faraco, S. (2013). Beat gestures modulate auditory integration in speech perception. *Brain and Language*, *124*(2), 143–152. https://doi.org/10.1016/j.bandl.2012.10.008

Davies, M. (2010). The Corpus of Contemporary American English as the first reliable monitor corpus of English. *Literary and Linguistic Computing*, *25*(4), 447–464. https://doi.org/10.1093/llc/fqq018

Dickerson, W. B. (1978). English orthography: A guide to word stress and vowel quality. *IRAL: International Review of Applied Linguistics in Language Teaching*, *16*(2), 127–147. https://doi.org/10.1515/iral.1978.16.1-4.127

Dickerson, W. B. (1983). The role of formal rules in pronunciation instruction. In J. Handscombe (Ed.), *The question of control: Selected papers from the annual convention of Teachers of English to Speakers of Other Languages* (pp. 135–148). TESOL.

Dickerson, W. B. (2018). English orthography as a resource for learners of English. In O. Kang, R. I. Thomson, & J. M. Murphy (Eds.), *The Routledge handbook of contemporary English pronunciation* (pp. 169–185). Routledge.

Edison, T. A. (1888). The perfected phonograph. *The North American Review*, *146*(379), 641–650.

Hardison, D. M. (2018). Visualizing the acoustic and gestural beats of emphasis in multimodal discourse. *Journal of Second Language Pronunciation*, *4*(2), 232–259. https://doi.org/10.1075/jslp.17006.har

Hardison, D. M., & Pennington, M. C. (2021). Multimodal second-language communication: Research findings and pedagogical implications. *RELC Journal*, *52*(1), 62–76. https://doi.org/10.1177/0033688220966635

Kaiser, D. (2018). Mobile-assisted pronunciation training: The iPhone pronunciation app project. *IATEFL Pronunciation Special Interest Group Journal*, *58*, 38–52.

Kaiser, D. (2024). Artificial intelligence integration in three iOS pronunciation apps: ELSA Speak, Loora, and Vocal Image. *Journal of Second Language Pronunciation*, *10*(3), 404–426. https://doi.org/10.1075/jslp.24052.kai

Kaiser, D. (in press). *Practical pronunciation: Pronunciation strategies in practice*. The University of Michigan Press. https://doi.org/10.3998/mpub.12740856

Mishra, P., & Koehler, M. J. (2006). Technological pedagogical content knowledge: A framework for teacher knowledge. *Teachers College Record*, *108*(6), 1017–1054. https://doi.org/10.1111/j.1467-9620.2006.00684.x

Neri, A., Cucchiarini, C., Strik, H., & Boves, L. (2002). The pedagogy-technology interface in computer assisted pronunciation training. *Computer Assisted Language Learning*, *15*(5), 1–27.

Shulman, L. S. (1986). Those who understand: Knowledge growth in teaching. *Educational Researcher*, *15*(2), 4–14.

Shulman, L. S. (1987). Knowledge and teaching: Foundations of the new reform. *Harvard Educational Review*, *57*(1), 1–21.