PLAYING WITH PLAYBACK SPEED: PRACTICING L2 PRONUNCIATION WITH YOUGLISH

Ines Martin, U.S. Naval Academy

With recent advances in technology and multimedia materials that have transformed language learning by providing learners not only with the opportunity to access authentic language directly on their computers or phones, but also to adapt these materials to learners' specific needs, the role of manipulation of playback speed for listening comprehension and oral proficiency development has become the target of L2 research (Mahalingappa et al., 2024). This teaching tip demonstrates how to practice L2 pronunciation with the platform *YouGlish*, which allows for manipulation of playback speed by learners. As such, this teaching tip (1) briefly introduces research findings that suggest that using *YouGlish* for L2 pronunciation practice can be beneficial for learners, (2) familiarizes the reader with *YouGlish* and what it offers for instructors and learners, and (3) outlines ideas for pronunciation activities with *YouGlish* both on the segmental and suprasegmental level.

INTRODUCTION

Language learning is ever changing with new research and new learning resources influencing the language learning landscape. A major factor in language learning over the past two decades has been facilitated access to authentic language input, which is an important advancement, given that many foreign language learners have limited possibilities of direct interaction with speakers of the target language. That is, advances in technology and multimedia materials have transformed L2 learning by providing learners with never-ending opportunities to access authentic language input directly through computer and hand-held devices (Richards, 2015). These resources and tools often can be adapted to specific learner needs via use of captions (i.e., subtitled audio) and manipulation of playback speed. This teaching tip will present one of these resources: *YouGlish* (www.youglish.com). *YouGlish* indexes videos from YouTube and allows users to search for words or short phrases and listen to them while seeing captions for the video's audio. Seeing that on average 720,000 hours of video are uploaded to YouTube every day (Blogger, 2024), this provides unprecedented access to authentic materials for language learners. *YouGlish* is available for most commonly taught languages (see details below) and even for different dialects of languages. The indexing allows learners to easily listen to the targeted search term as produced by native (or nonnative) speakers in all publicly accessible YouTube videos. Moreover, it offers helpful tools for language learners, such as jumping back in 5-second increments or manipulating playback speed in 5% increments (e.g., reducing speed to 0.95 from 1.0). This teaching tip (1) briefly introduces research findings that suggest that using *YouGlish* for L2 pronunciation practice can be beneficial for learners, (2) familiarizes the reader with *YouGlish* and what it offers for instructors and learners, and (3) outlines ideas for pronunciation activities with *YouGlish* both on the segmental and suprasegmental level.

Martin, I. A. (2025). Playing with playback speed: Practicing L2 pronunciation with YouGlish. In J. M. Levis, M. Duris, S. Sonsaat-Hegelheimer, & I. Na (Eds.), *Proceedings of the 15th Pronunciation in Second Language Learning and Teaching Conference* (pp. 1-9). Iowa State University, September 2024. https://doi.org/10.31274/psllt.18440

This teaching tip (1) briefly introduces research findings that suggest that using *YouGlish* for L2 pronunciation practice can be beneficial for learners, (2) familiarizes the reader with *YouGlish* and what it offers for instructors and learners, and (3) outlines ideas for pronunciation activities with *YouGlish* both on the segmental and suprasegmental level. When considering research findings that inform teaching practice with *YouGlish*, two general areas of research are relevant: the possible impact that manipulation of playback speed has on listening comprehension and the use of shadowing as a technique for L2 pronunciation practice.

BACKGROUND

The Impact of Playback Speed

While to date, there is no research that specifically investigates how the manipulation of playback speed affects L2 pronunciation improvement, there are two related fields of research that could inform the debate around the impact of playback speed on L2 pronunciation. These fields are (1) research in adaptation to time-compressed speech and (2) the effects of playback speed on L2 listening comprehension more generally. As for the former, previous research has investigated perceptual adjustment to time-compressed speech (Banai & Lavner, 2012; Pallier et al., 1998; Sebastián-Gallés et al., 2000). The idea behind this line of research is that despite great variability in background noise, talker's voice, speech rate, or dialect, native speakers can usually still understand an utterance. How is it that speech perception remains so amazingly accurate under these changing circumstances? Apparently, the listener's processing apparatus has been devised to compensate for the input's instability. But what about rapidly delivered speech as compared to slower rate speech? Research on adjustment to time-compressed speech has investigated different questions around speech that is delivered at a much faster (i.e., time-compressed) rate and found that listeners' ability to understand improved after 10-15 sentences of hearing compressed speech because their perceptual system adapted to these fast speech rates (Pallier et al., 1998). Considering that the speech input is quite distorted through compression (as well as expansion, when thinking of slower speech rates), it is interesting that listeners have the ability to adapt. This is important to keep in mind when thinking about "playing with playback speed" for activities with L2 learners.

Another line of research has investigated how playback speed affects L2 listening comprehension more generally (Mahalingappa et al., 2024; Medina et al., 2020; Murphy et al., 2022). Insights from these studies might inform the role of speech rate in L2 pronunciation learning as well. First of all, one could ask why playback speed might even have an effect on L2 learning. The question can be answered by Cognitive Load Theory (Sweller, 1988). In somewhat simplified terms, when new information is being learned, this information is stored in working memory before being transferred to long-term memory. Long-term memory capacity is generally considered to be limitless, but working memory capacity is restricted and can vary from person to person. What is more, too much information in working memory can hinder transfer to long-term memory, thus blocking learning. It is further known that conversational speech can be rather fast and therefore, regular speech rates might be too fast for L2 learners. If we then apply this to playback speed, and input rate overwhelms our learners' cognitive resources, their comprehension of the input and learning outcomes might be impaired, even for small chunks of speech (Bloomfield et al., 2010; Murphy et al., 2022). Slowing down speech rate on audio or video material could therefore make learning more successful. Conversational speech ranges from 120 to 200 words per minute (wpm), but research suggests that slower rates (ranging from 97 to 155wpm) enhance comprehension amongst L2 learners, especially amongst lower proficiency learners in their first year of

language study (Mahalingappa et al., 2024). Again, the manipulation of playback speed has not yet been investigated for L2 pronunciation specifically. That is, we currently do not have research showing that the slower rates that benefit comprehension would also benefit learners' pronunciation, or what elements it would benefit, e.g., overall fluency or overall accuracy. Moreover, effects might not be the same for learners of all proficiencies. Nevertheless, in light of these findings of perceptual adaptation to compressed speech and listening comprehension more generally, it is reasonable to assume that manipulating playback speed can be beneficial for learners in shadowing activities for longer chunks of speech, as well as when simply listening to words containing the targeted sound over and over again. Importantly, learners can start any practice activity at a slower playback speed and slowly increase to normal or accelerated speech rate.

Shadowing in L2 Pronunciation

The shadowing technique was originally used to train simultaneous interpreters but has since found its way into L2 classrooms as a popular and effective technique for L2 pronunciation practice. During a shadowing task, the learner tracks heard speech (e.g., in the form of a video or audio recording) and repeats it as precisely as possible with only a very short delay. While more research is needed to fully understand how shadowing affects L2 pronunciation, initial research findings on the effectiveness of shadowing have been promising (Bovee & Stewart, 2009; Foote & McDonough, 2017; Martinsen et al. 2017; Micik & Rizaoglu, 2024). Bovee and Stewart's (2009) pilot study on the benefits and practicability of shadowing suggested a significant impact of the technique on general measures of L2 pronunciation. Foote and McDonough (2017) investigated the impact of shadowing on specific measures of L2 pronunciation, that is, comprehensibility, accentedness, and fluency. The researchers had ESL learners shadow dialogues from popular US sitcoms for eight weeks and tested learners in a pre-, mid-, and post-test by measures of native speaker ratings. Their results revealed significant improvements on comprehensibility and fluency measures, but not on accentedness. These findings were largely in line with Martinsen et al. (2017), suggesting that the shadowing technique used as a classroom or homework activity has the potential to significantly improve L2 pronunciation. What is more, both anecdotal evidence as well as research suggests that learners not only improve their L2 pronunciation but also genuinely enjoy shadowing activities.

When responding to survey questions about how much learners liked shadowing activities, 80% of learners in Bovee and Stewart's (2009) study responded that they thought the activities had educational value, and learners in Foote and McDonough's (2017) study rated the question "How much do you like the shadowing activity?" with a 7.6 out of 9 average (where 9 equaled a positive response). Moreover, Inceoglu (2022) investigated learners' perceptions of shadowing vs. ASR-based activities in more detail and found that 80% of her learners had a positive impression of the shadowing activities, mentioning pros of the technique such as "comparison with models, intonation and rhythm; speed; connecting spelling and pronunciation" (p. 180). Based on the responses to open-ended questions in Inceoglu's (2022) study, the author concludes that amongst other advantages, shadowing offers learners the opportunity to not only connect the oral input to their own L2 production, but also to connect it to written input (e.g., in the form of captions), thus strengthening lexical and phonological representations of the input. Nevertheless, it has to be mentioned that the original shadowing technique does not include written script, such as captions (Hamada, 2019). Overall, shadowing seems to be both an effective as well as an enjoyable technique for L2 pronunciation practice.

Previous Research on *YouGlish*

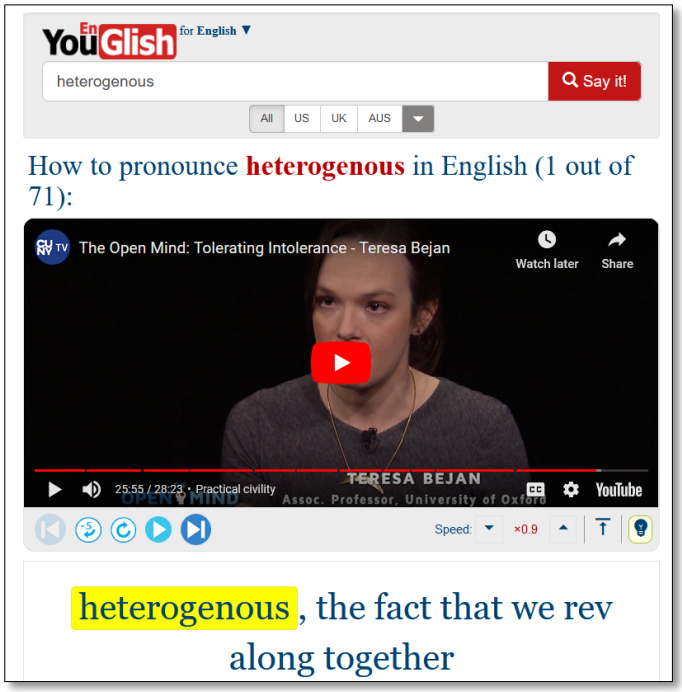
As is the case for most new tools that can be used for educational purposes, peer-reviewed research on *YouGlish* is still limited, but a study by Sardegna and Jarosz (2022) suggested that *YouGlish* can be used for educational purposes to improve learners' L2 pronunciation. In this study, the authors investigated whether a combination of explicit instruction and listening practice using *YouGlish* could help learners' prediction and perception skills for lexical stress assignment in polysyllabic words. They found that the L1 Polish learners in their study were indeed able to make improvements in their word-stress predictions and perceptions in L2 English. However, the authors caution that it is likely that a combination of explicit instruction and listening practice with *YouGlish* led to improvement, rather than one of these two elements on their own.

HOW YOUGLISH WORKS

Overview

YouGlish (www.youglish.com) is a free site that indexes authentic YouTube videos and allows users to search for words or short phrases in these videos. Figure 1 illustrates what this looks like. That is, the user can type the targeted term (here: "heterogenous") in the search bar. *YouGlish* then provides all indexed videos that contain this term (here: 71 videos). All videos are captioned, and the subtitles appear directly underneath the video. This is helpful to provide the learner with the context of the sentence in which the word is used. The algorithm jumps directly to a marker about 2 seconds before the word is uttered in the video. That means that learners do not have to search for the term in a lengthy video. Rather, they can listen to several iterations of the word in different videos by using the "next" button. If they want to hear the exact same iteration of the word, they can do so by using the "repeat" button.

Figure 1
Example of YouGlish interface



What Is Helpful to Know About *YouGlish*

1. *YouGlish* does not only index English videos but is available for most commonly as well as less commonly taught languages, including American Sign Language. As of 2024, it is offered for Arabic, Chinese, Dutch, English, French, German, Greek, Hebrew, Italian, Japanese, Korean, Polish, Portuguese, Russian, Spanish, Swedish, Thai, Turkish, Ukrainian, Vietnamese, and Sign Languages (US, UK, Ireland, Australian, New Zealand, and France). For most of these languages, one can choose between language-specific dialects as well. For example, choosing English further allows to choose between US, UK, or AUS English. *YouGlish* can therefore be used for pronunciation practice in a variety of L2 classrooms, not just ESL classrooms.

2. *YouGlish* allows for a fine-grained adjustment of playback speed. In the right-hand corner of the task bar, learners can reduce or increase playback speed. (See Figure 1, where playback speed has been decreased to 0.9). As discussed in the Background section, slowing down the playback speed of the video can lessen the cognitive burden on the learners and facilitate listening. They can start with a strongly decreased playback speed (e.g., 0.75) and slowly work their way up to regular speed (1.0). Slowing down the speed of the video is also helpful when applying the shadowing technique, that is, when learners not only listen to the pronunciation of the target word, but also practice repeating after the speaker.

3. *YouGlish* allows users to search not only for individual words, but for short phrases as well. This is particularly helpful in pronunciation teaching that targets suprasegmentals, for example, *liaisons* in French, where learners have to listen to whether two words are joined or separated in connected speech.

4. *YouGlish* uses authentic videos that have been uploaded to YouTube over the past decades. This has two major advantages: a) it introduces the learners to a wide range of voices and diversity of speakers, much more than one could ever bring to a classroom, and b) the authenticity of the videos is appealing to learners, as most authentic L2 materials inherently are.

5. *YouGlish* is free.

PRACTICING L2 PRONUNCIATION WITH YOUGLISH

Using YouGlish as a Pronunciation Dictionary

The quickest and easiest use for *YouGlish* is to employ it as a pronunciation dictionary—either in the classroom or at home. When students ask how to pronounce a word in class, instructors are usually quick to provide the answer. If the flow of the lesson allows for it, instead of the instructor providing the answer, this type of question is a great opportunity to introduce a variety of different voices in the classroom by playing 3-5 iterations of the word in *YouGlish* instead. In my own experience, learners respond really well to hearing the targeted words pronounced by several speakers. Of course, this might be disruptive to the flow of the lesson and takes up a little more time than when the instructor quickly models the pronunciation of a word, but whenever possible, it is worth the additional input. It also helps if the instructor gets used to having *YouGlish* open in the browser before class starts, so that they can immediately access it when necessary. Additionally, I like to at least show my students *YouGlish* as a

resource, so that they can use it during homework as well. My learners report that they prefer the authentic input to a computer-generated voice in other online dictionaries.

YouGlish Pronunciation Activities

Since *YouGlish* allows learners to search for both individual words *and* phrases, it can be used for pronunciation practice on the segmental level (i.e., on the level of individual consonants or vowels) as well as for practice on the suprasegmental level (for example, for stress assignment or features of connected speech). The details of the pronunciation activity will naturally vary a bit depending on what the target of the practice is, but the following outline can serve as an adaptable example for both. If the target is an individual sound, e.g., a specific vowel or consonant, such as German <ü> /y:/:

1. Give learners a list of five words containing the targeted sound. Ideally in word-initial, -medial, and -final position (if applicable).
2. Listening Practice:
    a. Ask learners to listen to five iterations of each word from the list. They will have then listened to the sound at least 25 times (i.e., 5 words x 5 speakers).
    b. Have them adjust playback speed while doing the listening activity. Suggest that they start at 0.8 and slowly increase the speed to 1.0 if they are comfortable with it.
    c. Remind them that all videos are captioned. That is, the students can see the targeted word spelled out in the context of the sentence below the video. This might be particularly helpful for languages with opaque orthography, such as French, where learners often erroneously produce silent final consonants (Inceoglu, 2022).
3. Speaking Practice:
    a. Ask the learners repeat the word as closely as possible after listening to it. They will practice the targeted sound at least 25 times (i.e., 5 words x 5 speakers, as in: speaker 1 word 1 – learner – speaker 2 word 1 – learner, etc.). Importantly, they will listen and repeat input from 25 different speakers because each video provides an occurrence of the word in an authentic context, and it is very unlikely that there are videos by the same speaker.
    b. Even in speaking practice it is helpful to play with the playback speed. It might be easier for the learner to repeat the word if it was played as a decreased playback speed.

If the target is a suprasegmental feature, e.g. *liaisons* in French, follow the same steps as outlined above, but instead of providing the learners with a list of words containing the targeted sounds, provide them with a list of short phrases that clearly demonstrate the targeted pronunciation feature. For suprasegmentals, it is even more important to adjust playback speed down and have learners listen to the examples at decreased speed but be careful not to adjust lower than .7 because around that point, pronunciation can start to be distorted too much.

Additional Pronunciation Activities

The following pronunciation activities can include an element of adjusting playback speed, but they do not have to. They are just meant to provide additional ideas of using *YouGlish* for pronunciation practice.

1. Minimal Pair Challenge

- Have learners look up words that are often confused (e.g. for English: “ship” vs. “sheep”, “bed” vs. “bad”) and listen to multiple examples.
- They should repeat each pronunciation and identify patterns in native speech.

2. Accent Exploration

- Learners search for the same word across different varieties within one language (e.g., in English between American, British, Australian, etc.).
- They compare and describe how pronunciation differs between varieties.

3. Stress and Intonation Analysis

- Learners look up a word and listen to how it is stressed within different sentences.
- They mark the stress patterns and practice saying the sentences with the correct intonation.

4. Pronunciation Journal

- Each learner keeps a journal of difficult words they have examined using *Youglish*, noting phonetic transcriptions, stress patterns, and example sentences.

5. Listening Comprehension Quiz

- One learner plays a clip with a word, and others have to guess what was said before seeing the transcript.

6. Context-Based Pronunciation

- Learners search for a word in different contexts and analyze how pronunciation varies depending on emphasis or speed of speech.

CONCLUSION

This teaching tip demonstrated how to practice L2 pronunciation with the platform *YouGlish*. This web-based tool allows learners to type in any word or phrase and see how it is pronounced in authentic YouTube videos. The application jumps directly to the use of the word in the videos, allowing learners to listen to several different people pronouncing the word. It is available for all commonly taught languages (including different dialects for some languages). Importantly, *YouGlish* also allows the adjustment of playback speed in increments of 5%. The tips for designing pronunciation activities using *YouGlish* on the segmental and suprasegmental level showed that *YouGlish* can be used in different ways to practice L2 pronunciation inside and outside the classroom.

ABOUT THE AUTHOR

Ines A. Martin (Ph.D., The Pennsylvania State University) is Associate Professor of German and Linguistics and German Program Director at the U.S. Naval Academy in Annapolis, MD. Her research in the field of Instructed Second Language Acquisition focuses on pronunciation teaching and learning with the goal of addressing the gap between research findings in pronunciation instruction and teaching practice.

Contact information
Dr. Ines Martin
Languages and Cultures Department
U.S. Naval Academy
105 Maryland Ave
Carter Hall, Stop 327
Annapolis, MD 21402
410-293-6355
imartin@usna.edu
www.ines-martin.weebly.com

REFERENCES

Banai, K., & Lavner, Y. (2012). Perceptual learning of time-compressed speech: More than rapid adaptation. *PLOS ONE, 7*(10), 1–9. https://doi.org/10.1371/journal.pone.0047099

Blogger, G. (2024, July 23). *Youtube Statistics 2024 (Demographics, users by country & more)*. Official GMI Blog. https://www.globalmediainsight.com/blog/youtube-users-statistics/#YouTube_Video_Statistics_2024

Bloomfield, A., Wayland, S. C., Rhoades, E., Blodgett, A., Linck, J., & Ross, S. (2010). What makes listening difficult? In *Factors affecting second language listening comprehension*. University of Maryland Center for Advanced Study of Language. https://apps.dtic.mil/sti/pdfs/ADA550176.pdf

Bovee, N., & Stewart, J. (2009). The utility of shadowing. In A. M. Stoke (Ed.), *JALT 2008 Conference Proceedings* (pp. 888–900). Tokyo: JALT. https://jalt-publications.org/recentpdf/proceedings/2008/E158.pdf

Foote, J., & McDonough, K. (2017). Using shadowing with mobile technology to improve L2 pronunciation. *Journal of Second Language Pronunciation, 3*(1), 34–56. https://doi.org/10.1075/jslp.3.1.02foo

Hamada, Y. (2019). Shadowing: What is it? How to use it. Where will it go? *RELC Journal, 50*(3), 386–393. https://doi.org/10.1177/0033688218771380

Inceoglu, S. (2022). Developing pronunciation learner autonomy with Automatic Speech Recognition and shadowing. In S. McCrocklin (Ed.), *Technological resources for second language pronunciation learning and teaching: Research-based approaches* (pp. 171–192). Lexington Books. https://doi.org/10.5771/9781666902303

Mahalingappa, L., Zong, J., & Polat, N. (2024). The impact of captioning and playback speed on listening comprehension of multilingual English learners at varying proficiency levels. *System, 120*, 103192. https://doi.org/10.1016/j.system.2023.103192

Martinsen, R., Montgomery, C., & Willardson, V. (2017). The effectiveness of video-based shadowing and tracking pronunciation exercises for foreign language learners. *Foreign Language Annals, 50*(4), 661–80. https://doi.org/10.1111/flan.12306

Medina, A., Socarras, G., & Krishnamurti, S. (2020). L2 Spanish listening comprehension: The role of speech rate, utterance length, and L2 oral proficiency. *The Modern Language Journal, 104*(2), 439–456. https://doi.org/10.1111/modl.12639

Micik, S, & Rizaoglu, F. (2024). The role of video-based shadowing practices in L2 oral proficiency development. *Journal of Language Education and Research, 10*(1), 166–190. https://doi.org/10.31464/jlere.1356763

Murphy, D. H., Hoover, K. M., Agadzhanyan, K., Kuehn, J. C., & Castel, A. D. (2022). Learning in double time: The effect of lecture video speed on immediate and delayed comprehension. *Applied Cognitive Psychology, 36*(1), 69–82. https://doi.org/10.1002/acp.3899

Pallier, C., Sebastián-Gallés, N., Dupoux, E., Christophe, A., & Mehler, J. (1998). Perceptual adjustment to time-compressed speech: a cross-linguistic study. *Memory & Cognition, 26*(4), 844–851. https://doi.org/10.3758/bf03211403

Richards, J. C. (2015). The changing face of language learning: Learning beyond the classroom. *RELC Journal, 45*(1), 5–22. https://doi.org/10.1177/0033688214561621

Sardegna, V. G., & Jarosz, A. (2022). Exploring how *YouGlish* supports learning English word stress: A perception study. In V. G. Sardegna & A. Jarosz (Eds.), *Theoretical and practical developments in English speech assessment, research, and training* (pp. 165–184). Springer. https://doi.org/10.1007/978-3-030-98218-8_10

Sebastián-Gallés, N., Dupoux, E., & Costa, A. (2000). Adaptation to time-compressed speech: Phonological determinants. *Perception & Psychophysics. 62*(4), 834–842. https://doi.org/10.3758/bf03206926

Sweller, J. (1988). Cognitive load during problem solving: Effects on learning. *Cognitive Science, 12*(2), 257–285. https://doi.org/10.1016/0364-0213(88)90023-7