USING AUDIOVISUAL MATERIALS TO TEACH PRONUNCIATION

Carolin Jolitz, The Pennsylvania State University
Ines Martin, U.S. Naval Academy

Captioned video offers learners a combination of auditory and visual input and has therefore been suggested as a great tool to enhance second language (L2) learning. Recent research has also shown promising beneficial effects of captioned video on L2 pronunciation. This teaching tip is designed to demonstrate a step-by-step guide on how captioned video can be utilized as a springboard for controlled and communicative speaking activities and complement explicit pronunciation instruction in the L2 classroom. It also includes recommendations for how to incorporate captioned videos into homework assignments, using the educational video platform *EdPuzzle*.

INTRODUCTION

Although input exposure is crucial for acquiring a second language (L2) (Flege, 2009), the amount of input learners receive in the L2 classroom environment is often limited. In particular, in the classroom-based context, most L2 learners only receive a few hours of instruction per week, with the teacher being the primary source of L2 input (Muñoz, 2008). Learners living outside the target language country have even fewer opportunities to speak and listen to the target language. This scarcity of L2 input exposure can make it particularly challenging for L2 learners to improve their L2 speech skills (Flege, 2009; Saito & Hanzawa, 2018). Therefore, providing more opportunities for L2 learners to engage with the target language inside and outside the L2 classroom is vital. One way to achieve this is by incorporating audiovisual materials, such as videos, TV series, movies, or documentaries, into both classroom activities and homework assignments. The potential benefits of exposure to captioned video for various aspects of L2 learning have gained increasing attention in the Second Language Acquisition (SLA) research literature in the past decade (for recent reviews see Montero Perez, 2022; Muñoz, 2022). In captioned video, both the audio and on-screen text are presented in the L2. Recent research has also found beneficial effects of captioned video, such as fostering L2 pronunciation skills (Scheffler & Baranowska, 2023; Wisniewska & Mora, 2020). This teaching tip first describes these benefits based on theory and research findings. This brief overview will be followed by a step-by-step guide with detailed instructions on how to combine captioned videos with pronunciation instruction and practice, both in classroom activities and homework assignments.

The Benefits of Captioned Video for L2 Pronunciation Learning

As a primary theoretical basis to explain the benefits of captioned video for L2 learning, SLA research often refers to Mayer's Cognitive Theory of Multimedia Learning (CTML; Mayer, 2014, 2021). Drawing on Paivio's (1986) Dual Coding Theory, CTML posits that humans possess two separate, yet interacting channels to process visual and auditory information. While humans process visual information such as pictures, videos, illustrations, and on-screen text (i.e., captions, subtitles) through the visual channel, verbal information such as spoken words or sounds is processed through the auditory channel. When watching captioned videos, humans are simultaneously exposed to visual and auditory information, and both channels are activated (Mayer, 2014, 2021). This has multiple benefits for L2 learning in general, but also

specifically for L2 pronunciation learning. The simultaneous processing of auditory and visual information while watching captioned video supports learners in creating and strengthening the link between a word's written and auditory form (Wisniewska & Mora, 2020). This is particularly beneficial for L2 learners as it can aid them in mapping L2 orthography to phonological form (Mitterer & McQueen, 2009; Wisniewska & Mora, 2020). Additionally, the captions visualize the spoken input and facilitate the process of identifying and understanding individual words in the continuous speech stream (Charles & Trenkic, 2015). As learners attempt to match the written input to the auditory input in the captioned video, they can learn to adapt their perception to an unfamiliar accent and improve their L2 speech perception skills (Mitterer & McQueen, 2009).

Another potential benefit of captioned video for perception learning is the exposure to multiple speakers with different voices, as well as target sounds in a wide variety of phonetic environments (Hutchinson & Dmitrieva, 2022). Such varied input familiarizes learners with the range of variation in natural L2 speech and can help them generalize their perception to sounds produced by new speakers (Thomson, 2022). Especially in formal L2 classroom settings, where L2 learners are typically exposed to mainly the teacher's voice, such varied input is instrumental in preparing them to interact with a wide array of speakers in real-life situations. In this vein, combining perceptual training with awareness-raising (e.g., through explicit instruction) and oral production practice can further facilitate L2 pronunciation improvement (Sardegna & McGregor, 2022). Providing L2 learners with explicit information about the target sound and its characteristics not only draws their attention to the target L2 sound but also helps them recognize the specific features they should identify. This approach raises learners' awareness, enabling them to perceive and practice pronouncing the target sound in a more target-like manner (Saito & Plonsky, 2019; Sardegna & McGregor, 2022). Similar positive effects on L2 production have been found when explicit phonetic instruction is followed by exposure to variable speech through high variability phonetic training (Wiener et al., 2020) or exposure to podcasts (Fouz-González, 2019). These findings further support the idea that combining explicit pronunciation instruction with captioned video is beneficial for L2 learners' pronunciation development. In sum, captioned video offers learners both auditory and written L2 input and can serve as an excellent tool to provide perceptual training and help learners improve their L2 pronunciation skills, especially when combined with explicit information about the target sounds. Based on these prior research findings, our teaching tip combines explicit pronunciation practice with perceptual training and oral production practice.

HOW TO TEACH PRONUNCIATION USING AUDIOVISUAL MATERIALS: A STEP-BY-STEP APPROACH

After gaining a better understanding of the benefits of captioned video for L2 pronunciation and perception learning, this section presents step-by-step instructions for integrating captioned video into pronunciation instruction and practice (see also Martin & Jolitz, 2023). This step-by-step approach, illustrated in Figure 1, can be implemented gradually over multiple class sessions or in a single class and can be tailored to the instructor's learners and the learning environment. It is important to note that our steps are merely recommendations for how to integrate videos into pronunciation instruction, based on what we did in our classrooms and prior research suggestions.
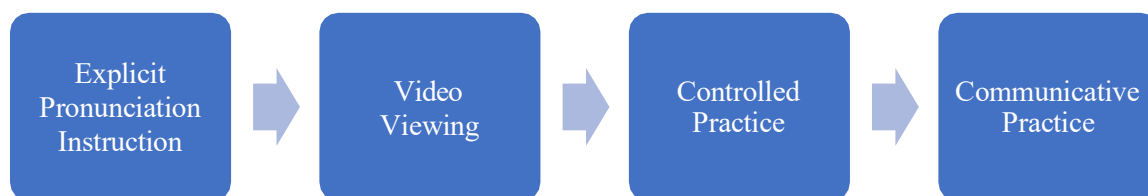
Figure 1. Step-by-step approach to teaching pronunciation using captioned video.

Step 1: Explicit Pronunciation Instruction

The first step entails providing explicit instruction on the target pronunciation feature. The goal is to draw the student's attention to the target pronunciation feature and help them better understand its characteristics (Schmidt, 1990). The instructor could, for instance, address why this particular sound is so challenging to pronounce, what to pay close attention to when

Figure 2. Example explicit instruction for German <ei> and <ie>.

pronouncing this sound, and why it is important to achieve intelligibility when pronouncing this sound. The instructor can use explicit explanations, diagrams, and/or vowel or consonant charts to introduce the pronunciation target. This can further be combined with brief awareness-raising exercises and isolated pronunciation practice activities. Students can, for example, be instructed to identify words based on written letters, followed by a short activity in which they practice saying these words and give each other feedback (see Figure 2 for an example activity).

Step 2: Video Viewing

In the second step, the instructor exposes the students to captioned videos to train their perception of the target pronunciation feature. This can be done using an educational video platform like *EdPuzzle* (see below for instructions), which allows for this step to be completed both at home and in the classroom. It is important that the instructor ensures beforehand that the video contains many words with the target sounds, preferably with the sound occurring in different positions within the word. To prevent students from having to split

their attention between focusing on content and pronunciation, we recommend splitting the viewing into two parts: First, students watch the video and focus on the content, then they re-watch the same video and focus on pronunciation. For the first viewing (focus on content), the instructor can provide the students with comprehension questions to guide their attention and help them understand unfamiliar words. The main goal of the first viewing is to help the students understand what is happening in the video. For the second viewing (focus on pronunciation), the instructor can provide the students with a word bank with target words containing the target sound on which they are to focus. Before they start watching the video again, the instructor can tell them to listen carefully to how these words are pronounced in the video.

Step 3: Controlled Practice

The objective of the third step is to provide the students with the opportunity to practice the targeted feature. To ensure that students do not get overwhelmed and that they can allocate

---

Controlled Practice: Pronunciation of <ie> and <ei>

Listen carefully to how these words are pronounced. Pay attention to the length of the vowel. Afterward, practice saying these words with your partner.

fliegt	mein	die	arbeitet	sie	deine

Figure 3. Example of controlled practice activity

their mental resources to pay attention to pronunciation, practice in this phase should take place in a guided or controlled manner. For example, the word bank introduced in Step 2 could be used to prompt students to practice saying the target words they encountered in the video on their own. After practicing the words from the word bank, instructors could assign a partner activity in which students provide feedback to each other on their pronunciation (see Figure 3 for an example of this activity). In addition to the word bank, instructors can provide a copy of the video transcript and prompt students to first mark all instances of the target feature and then practice saying these words with a partner, again encouraging them to provide feedback to each other.

Step 4: Communicative Practice

The goal of the fourth and final step is to allow the students to go beyond mere pronunciation practice by using the newly learned pronunciation features in authentic communication. For this phase, students should work in pairs or even in small groups of three or four learners. The instructor can provide students with prompts to guide them as they engage in meaningful conversations. Ideally, these communicative activities should still center around the videos, but the aim is now shifting from a strict focus on pronunciation to general communication practice. Some examples of such prompts would be (1) to have students do a role-play in which they continue a conversation between two or more characters that were introduced in the video, or (2) to have students discuss open-ended questions on the topic of the video. It can be expected that students will make more pronunciation errors in this phase than they did in the controlled

practice phase, seeing that the focus is no longer only on pronunciation and that they now also have to allocate mental resources to content in addition to form. Nevertheless, it is important to give them the opportunity to practice the newly acquired pronunciation skills in authentic communicative activities to prepare them to communicate in real-life contexts.

USING EDUCATIONAL VIDEO PLATFORMS TO TEACH PRONUNCIATION AS HOMEWORK

The use of captioned video for teaching and practicing pronunciation is not restricted to the classroom: it also lends itself to incorporation into homework assignments. A great resource that makes this objective feasible is the educational video platform *EdPuzzle* (https://edpuzzle.com/). *EdPuzzle* is freely available to instructors and students. It allows instructors to easily create interactive video lessons by offering many options for students to interact with the video. For example, instructors can add notes before, after, or at any point during the video, insert multiple-choice or open-ended questions, cut the video, record a voice-over for any part of the video, or upload pictures and links to outside resources. Additionally, *EdPuzzle* allows instructors to integrate the video lesson as an assignment into their Learning Management System (e.g., Canvas or Google Classroom), verify when and for how long students watched each video, and even prevent students from fast-forwarding or skipping parts of the video. Moreover, the playlist function is a great feature that lets students watch videos in a pre-determined order as they complete the assignment. This function is especially useful when the instructor wants to split the viewing into a "focus on content" and "focus on pronunciation" part: The instructor can add the videos in the desired order to the playlist, and *EdPuzzle* will guide them from one video to the next one. All of these features are helpful tools when assigning captioned videos as homework. To create a homework assignment as pronunciation practice with captioned video on *EdPuzzle*, we recommend the following steps: 1) Upload the video with embedded captions, 2) include explicit instruction or a refresher on the pronunciation rules for the targeted feature as a written or spoken note before the beginning of the video, and 3) insert multiple-choice or open-ended questions to raise awareness around the pronunciation target throughout the video. An example of what that could look like is provided in Figure 4. As the final step, we recommend concluding with a written or spoken note that provides instructions for a short practice activity after the video. For example, provide a word bank of pronunciation targets that students encountered in the video and have them practice saying these words out loud.
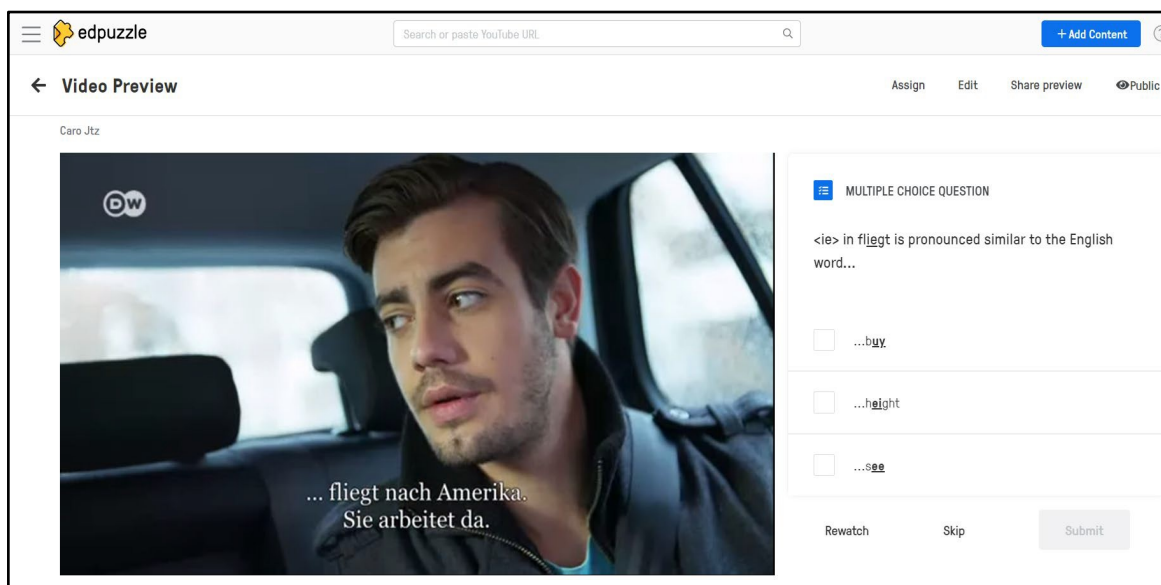
Figure 4. Example of a multiple-choice-question in a captioned video homework assignment on EdPuzzle, using video excerpts from "Nicos Weg" (Deutsche Welle Learn German, 2023).

These steps are just a starting point. In fact, there are many ways to create these homework assignments, and the videos can be adapted to instructors' preferences or students' needs. While using a new educational platform always comes with some growing pains and an up-front time investment, we would like to note that *EdPuzzle* is very user-friendly and that students generally enjoy the video activities. Therefore, we believe it is worth the time invested in creating these activities. When creating assignments specifically for the purpose of practicing pronunciation with captioned videos, one practical tip to keep in mind is to disable the Closed Captions (CC) feature in *EdPuzzle* (see Figure 5). This might seem counter-intuitive, but if the uploaded video already has captions embedded (that is, captions in the target language), automatically generated captions in addition to the embedded captions are counterproductive—especially because the algorithm does not always detect the language of the video and might try to caption, for example, a German video with English subtitles.
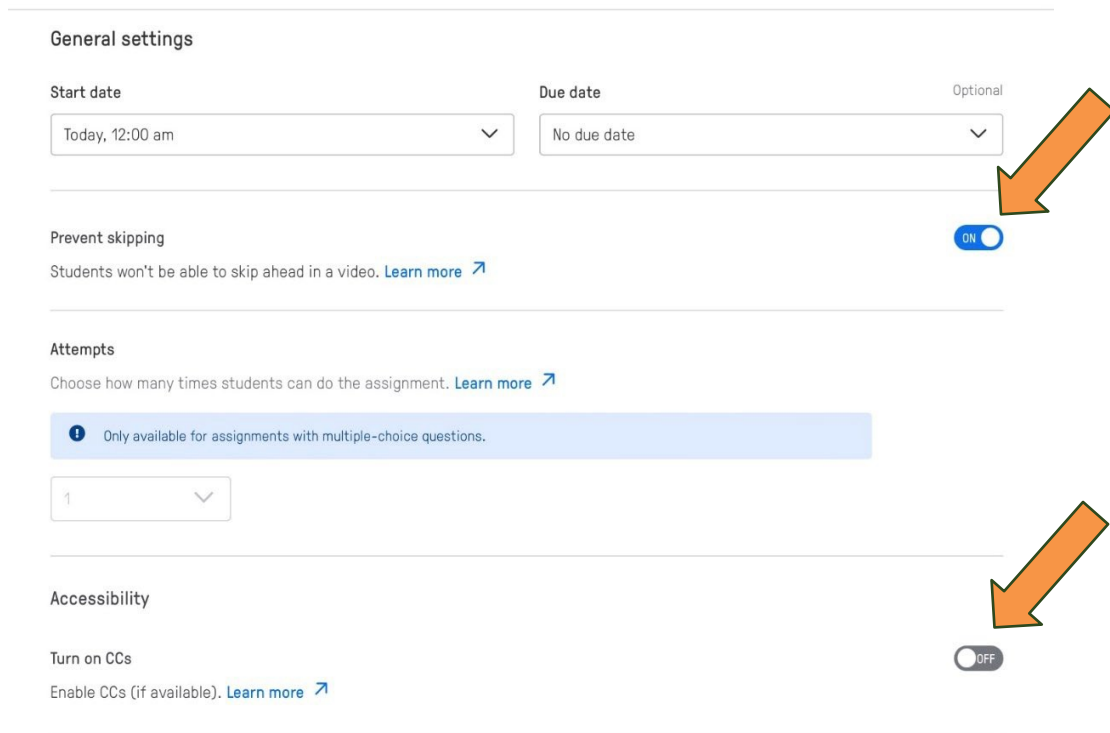
Figure 5. Practical tips for EdPuzzle settings specific to teaching with captioned video.

CONCLUSION

With the time and input limitations of instructed L2 classroom settings, audiovisual materials provide an excellent tool to expose students to L2 input. We hope this teaching tip inspires L2 instructors to incorporate more videos in their teaching and combine them with pronunciation instruction and practice. In the classroom, captioned video can be an effective resource to facilitate controlled pronunciation practice and stimulate guided and communicative activities. Watching captioned videos on educational video platforms like *EdPuzzle* makes pronunciation learning more interactive, keeps learners engaged, and provides them with ample exposure to L2 input, even outside the L2 classroom. We hope that future research will investigate the magnitude of instructional effects of captioned video and shed even more light on best practices when teaching with audiovisual materials.

ABOUT THE AUTHORS

Carolin Jolitz is a dual-title Ph.D. candidate in German Applied Linguistics and Language Science and a German instructor at the Pennsylvania State University. She earned her BA and MA in Educational Science and German Linguistics and Literature at the University of Kiel, Germany. Her dissertation research focuses on the potential of captioned and non-captioned videos to help beginner L2 German learners acquire incongruent grapheme-phoneme correspondences. Visit her homepage at www.carolinjolitz.weebly.com.

Contact Information:
Department of Germanic and Slavic Languages and Literatures
The Pennsylvania State University
253 Burrowes Building, University Park, 230 Old Coaly Way, PA 16802
814-867-5471, carolinjolitz@psu.edu

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. Visit her homepage at www.ines-martin.weebly.com to learn more about her research and teaching.

Contact Information:
Languages and Cultures Department
U.S. Naval Academy
105 Maryland Ave Carter Hall, Stop 327 Annapolis, MD 21402
410-293-6355
imartin@usna.edu

REFERENCES

Birulés-Muntané, J., & Soto-Faraco, S. (2016). Watching subtitled films can help learning foreign languages. *PLOS ONE*, *11*(6), e0158409. https://doi.org/10.1371/journal.pone.0158409

Charles, T. J., & Trenkic, D. (2015). Speech segmentation in a second language: The role of bimodal input. In Y. Gambier, A. Caimi, & C. Mariotti (Eds.), *Subtitles and language learning: Principles, strategies and practical experiences* (pp. 173–198). Peter Lang.

Deutsche Welle Learn German (2023). Nicos Weg. https://learngerman.dw.com/de/nicos-weg/c-36519687

Flege, J. E. (2009). Give input a chance! In T. Piske & M. Young-Scholten (Eds.), *Input matters in SLA* (pp. 175–190). Multilingual Matters. https://doi.org/10.21832/9781847691118-012

Fouz-González, J. (2019). Podcast-based pronunciation training: Enhancing FL learners' perception and production of fossilised segmental features. *ReCALL*, 31(2), 150–169. https://doi.org/10.1017/S0958344018000174

Hutchinson, A. E., & Dmitrieva, O. (2022). Exposure to speech via foreign film and its effects on non-native vowel production and perception. *Journal of Phonetics*, 95, 101189. https://doi.org/10.1016/j.wocn.2022.101189

Martin, I. A., & Jolitz, C. (2023). Teaching Perspectives – German. *Journal of Second Language Pronunciation, 9*(3), 390–397. https://doi.org/10.1075/jslp.23031.mar

Mayer, R. E. (2014). Cognitive Theory of Multimedia Learning. In R. E. Mayer (Ed.), *The Cambridge handbook of multimedia learning* (2nd ed., pp. 43–71). Cambridge University Press. https://doi.org/10.1017/CBO9781139547369.005

Mayer, R. E. (2021). Cognitive Theory of Multimedia Learning. In R. E. Mayer & L. Fiorella (Eds.), *The Cambridge handbook of multimedia learning* (3rd ed., pp. 57–72). Cambridge University Press. https://doi.org/10.1017/9781108894333.008

Mitterer, H., & McQueen, J. M. (2009). Foreign subtitles help but native-language subtitles harm foreign speech perception. *PLoS ONE*, *4*(11), e7785. https://doi.org/10.1371/journal.pone.0007785

Montero Perez, M. (2022). Second or foreign language learning through watching audio-visual input and the role of on-screen text. *Language Teaching*, *55*(2), 163–192. https://doi.org/10.1017/S0261444821000501

Muñoz, C. (2008). Symmetries and asymmetries of age effects in naturalistic and instructed L2 learning. *Applied Linguistics, 29*(4), 578–596.

https://doi.org/10.1093/applin/amm056

Muñoz, C. (2022). Audiovisual input in L2 learning. *Language, Interaction and Acquisition, 13*(1), 125–143. https://doi.org/10.1075/lia.22001.mun

Paivio, A. (1986). *Mental representations: A dual coding approach.* Oxford University Press. https://doi.org/10.1093/acprof:oso/9780195066661.001.0001

Schmidt, R. (1990). The role of consciousness in second language learning. *Applied Linguistics, 11*(2), 129–158. https://doi.org/10.1093%2Fapplin%2F11.2.129

Saito, K., & Hanzawa, K. (2018). The role of input in second language oral ability development in foreign language classrooms: A longitudinal study. *Language Teaching Research*, *22*(4), 398–417. https://doi.org/10.1177/1362168816679030

Saito, K., & Plonsky, L. (2019). Effects of second language pronunciation teaching revisited: A proposed measurement framework and meta-analysis. *Language Learning, 69*(3), 652–708. https://doi.org/10.1111/lang.12345

Sardegna, V., & McGregor, A. (2022). Classroom research for pronunciation. In J. M. Levis, T. M. Derwing, & S. Sonsaat-Hegelheimer (Eds.), *Second language pronunciation: Bridging the gap between research and teaching* (pp. 107–128). Wiley Blackwell.

Scheffler, P., & Baranowska, K. (2023). Learning pronunciation through television series. Language Learning & Technology*, 27*(1), 1–16. http://hdl.handle.net/10125/73520

Thomson, R. I. (2022). Perception in pronunciation training. In J. M. Levis, T. M. Derwing, & S. Sonsaat-Hegelheimer (Eds.), *Second language pronunciation: Bridging the gap between research and teaching* (pp. 42–60). Wiley Blackwell.

Wiener, S., Chan, M. K. M., & Ito, K. (2020). Do explicit instruction and high variability phonetic training improve nonnative speakers' mandarin tone productions? *The Modern Language Journal, 104*(1), 152–168. https://doi.org/10.1111/modl.12619

Wisniewska, N., & Mora, J. C. (2018). Pronunciation learning through captioned videos. In J. Levis, C. Nagle, & E. Todey (Eds.) *Proceedings of the 9th Pronunciation in Second Language Learning and Teaching Conference. Ames, IA: Iowa State University* (pp. 204– 215).

Wisniewska, N., & Mora, J. C. (2020). Can captioned video benefit second language pronunciation? *Studies in Second Language Acquisition*, *42*(3), 599–624. https://doi.org/10.1017/S0272263120000029