Tanner, M. (2020). Back to the future of pronunciation teaching: Putting the zoom back into the 'zoom principle.' In O. Kang, S. Staples, K. Yaw, & K. Hirschi (Eds.), *Proceedings of the 11th Pronunciation in Second Language Learning and Teaching conference*, ISSN 2380-9566, Northern Arizona University, September 2019 (pp. 341–352). Ames, IA: Iowa State University.

BACK TO THE FUTURE OF PRONUNCIATION TEACHING: PUTTING THE ZOOM BACK INTO THE 'ZOOM PRINCIPLE'

Mark Tanner, Brigham Young University

While pronunciation is a key component in helping English language learners (ELLs) achieve fluency and accuracy in a second language, ESL teachers often lack confidence and competence in knowing how best to teach pronunciation, which causes it too often be neglected in ESL classrooms. This is unfortunate because research has shown that a focus on global pronunciation features can improve ELLs' intelligibility and comprehensibility. The 'zoom principle' was first written about over 25 years ago, but it has not been widely understood or utilized by teachers within today's methodology of communicative pronunciation teaching. While many textbooks emphasize the importance of rhythm, stress, and intonation and provide helpful exercises to practice these prosodic features, assisting teachers who may lack confidence or knowledge of pedagogical application is critical in having pronunciation be consistently integrated into the speaking and listening classroom. This teaching tip will take teachers on a ride back to the future of pronunciation teaching with tips on using the 'zoom principle' to provide a pedagogical framework for diagnosing learner difficulties and then systematically addressing these issues in contextualized communicative practice that goes beyond the word and sentence level.

INTRODUCTION

There are a number of challenges that teachers express when it comes to pronunciation teaching. These challenges include questions such as "How do I teach pronunciation if I've had no training?" (Foote, Holtby, & Derwing, 2011), "What methods and materials work in the classroom? And how do I include pronunciation work in my regular curriculum and lesson planning when I already have so much to cover?" (Sicola & Darcy, 2015). To help address these concerns, it can be helpful if teachers look back to a guiding principle that was advocated more than 25 years ago, the 'zoom principle' (Firth, 1992). This principle provides guidance to teachers who may lack confidence or training in teaching pronunciation or are concerned with knowing how to implement pronunciation instruction into a listening/speaking class. Firth (1992) suggests that the 'zoom principle' involves the ability to "adjust one's field of vision when it comes to pronunciation pedagogy" (p. 173). Previous to the advent of communicative pronunciation teaching, "older texts urged a close-up approach, teaching and testing a student's ability to discriminate and produce sound contrasts between minimal pairs, such as in the words 'bit' and 'bet' (p. 173). In the mid 1980s, pioneered by Judy Gilbert (2008) and others' work in pronunciation pedagogy and materials development, the attention to pronunciation features widened with a focus on suprasegmentals, namely the elements of stress, rhythm, intonation, thought groups, and pausing prior to addressing individual sound practice. Suprasegmentals help control the structure of information and are extremely important in the communication of meaning in spoken language (McNerney & Mendelsohn, 1992). Firth (1992) further argues that the focus has continued to widen with attention being given to general speaking characteristics that even native speakers of a language need to consider when giving an effective presentation (p. 173). These features could include speaking with sufficient

volume, using an appropriate rate of speech to ensure that articulation of words is clear, speaking with appropriate voice quality, maintaining eye contact with one's audience, being expressive with facial expressions and appropriate non-verbal gestures, and incorporating sufficient energy and enthusiasm for the topic being shared (Firth, 1992). To summarize then, Firth advocates that pronunciation instruction using the 'zoom principle' should proceed as follows:

> "A pronunciation syllabus should begin with the widest possible focus and move gradually in on specific problems, then move back to survey the overall effect with a focus on overall effectiveness of communication" (p. 173).

Given this advice, how does an ESL/EFL teacher who wants to address pronunciation concerns with their students move forward? The purpose of this teaching tip is to help the classroom teacher address four important stages that will guide them in addressing pronunciation problems produced by the students.

IMPLEMENTING THE ZOOM PRINCIPLE

Stage 1: Needs Assessment and Collection of Speech Samples

Implementing the 'zoom principle' in the classroom begins with conducting a needs assessment and collecting data from the students. Teachers should identify learner variables for their students that will give insights into the learners' motivation to improve their oral communication, and factors that could impact their intelligibility. These variables include the students' L1, age, level of proficiency, purpose for learning English, and their intelligibility goals. Along with this information, it is critical to collect a range of speech data. A variety of types of data needs to be collected to verify the distribution of speech errors that the students may be producing. Samples of speech data should include the following:

> 1) A controlled speech sample that has been specifically engineered to collect common examples of second language speakers' pronunciation difficulties. These paragraph-length pieces of read-aloud speech capture common difficulties that learners of English have with producing vowels, consonants, and suprasegmental features such as word and sentence stress, intonation, and pausing. Most current pronunciation texts provide samples of these types of oral readings (Grant, 2017). A speech sample that pronunciation teachers have used for a number of years is one developed by Prator and Robinett (1985) (see Appendix A).
>
> 2) A semi-controlled speech sample. Examples of this type of discourse might include having the students tell a story based on a series of picture prompts or narrate a short silent video clip that showcases a brief narrative event (See picture story sequence example in Appendix B). The picture story or video can be selected to elicit the pronunciation of words with particular sound contrasts that may be problematic for speakers from a particular native language such as /r/ and /l/ for Japanese speakers of English, or the use of prosodic features that may be difficult for learners in extended oral discourse such as the production of thought groups, or appropriate use of sentence stress or intonation.

3) A spontaneous speech sample. The speech data teachers collect needs to include at least two samples of spontaneous speech. This data can be elicited by having the student answer open-ended questions, such as "Tell me about a favorite vacation you took," or "Describe what you would do if you had a million dollars." These spontaneous speech samples help identify difficulties the student may have with organizing information or elements of fluency such as rate of speech and frequency of pausing.

Stage 2: Analysis of Learners' Speech

Once the teacher has captured data describing the students' communicative goals, the learner variables present within the group, and a variety of speech samples, stage 2 requires an analysis of this data. In their book, *Teaching American English Pronunciation*, Avery and Ehrlich (1992), provide a helpful chapter dealing with common pronunciation problems produced by nonnative speakers of English. Teachers of ESL/EFL learners need to carefully review the speech data with special attention to the following areas of common concern:

1) *General speaking characteristics*. The initial assessment of the speech data will identify the degree to which the L2 students may be struggling with features such as rate of speech, volume, fluency, voice quality, energy in their voice, and the use of appropriate eye contact while speaking.

2) *Suprasegmentals*. Within this area, teachers should note the frequency and distribution of prosodic errors that include word, phrase, and sentence level stress, rhythm, appropriate pausing, intonation, and linking (connecting words appropriately within sentences).

3) *Segmentals*. There are common difficulties that affect most second language learners of English when it comes to consonant and vowel production. With consonant production, common difficulties include producing word endings (-ed, and third person singular -s or plural -s), pronouncing final stops (e.g. voiceless stops - /p, t, k/) and voiced stops /b, d, g/), aspirating initial stops (e.g. /p, t, k/ as [p^h, t^h, k^h]), producing consonant clusters, contractions, and other problematic sounds such as /r/, voiced and voiceless /th/, and fricatives (e.g. /v, z, ʒ/). Vowels are another area that can cause difficulty for learners of English, such as the inability to produce distinctions between lax vowels and tense vowels, the production of schwa, which is particularly important in the production of destressed or unstressed vowels, and the tendency of some L2 speakers learning English to insert a vowel between consonant clusters (e.g. producing /pə liz/ for *please*) (Avery & Ehrlich, 1992).

Stage 3: Creating Instructional Tiers

Once the speech data has been analyzed and difficulties in the areas of general speaking characteristics, prosodic features, and segmentals have been identified, it is important to uncover additional information about these errors prior to creating a pedagogical framework. Cohen's (1975) article on error correction and the training of language teachers is still quite relevant in helping teachers determine what factors should be considered in addressing the seriousness of the errors the learners have produced. Cohen suggests addressing the following factors (pp. 414-416):

1. Basic information about the error (what was exactly said vs. what was meant)
2. The importance of the correction as it relates to:
   a. Intelligibility
   b. High frequency of the error
   c. High generality of rules
   d. Stigmatizing or irritating effects present
   e. Errors affecting a large percentage of students
   f. The pedagogical focus
3. Ease of correction
4. Characteristics of the students and their goals in learning the target language

The teacher should use the information gathered about the various errors produced by the learner and separate the types of difficulties into instructional tiers (See Figure 1). In the sample diagnostic data provided in Figure 1 for an intermediate-level Spanish speaker of English, the tiers help to systematically organize the problematic pronunciation features into levels of urgency and pedagogical focus. Tier 1 concerns should include those features where there is a definite impact on intelligibility, high frequency of the error, and specific instructional treatment can be provided that would help mitigate the error in a relatively short period of time. Items that would typically be included in Tier 1 would first include any general speaking characteristics that emerge from the speech data as problematic. Considerations may include a lack of eye contact when speaking, vocal quality that is monotone in nature, the loudness with which the person speaks (too loud or too soft), whether the person mumbles when speaking or covers their mouth, or uses a rate of speech that is much faster or slower than that produced by a highly proficient speaker of the target language.

Next, there may be particular prosodic features that appear to impede the speaker's intelligibility, such as inappropriate pausing and creation of thought groups, incorrect syllable, word, or phrasal stress, and/or improper pitch, and intonation patterns. If there are frequent reading miscues produced when the reader reads out loud, this behavior should also be targeted. Prosodic features have been shown in pronunciation research to be common areas of difficulty for non-native speakers which can dramatically influence intelligibility and comprehension (Hahn, 1994; Pickering, 2001). These features are ones that a teacher should consider providing targeted in-class practice for as a means of helping learners produce speech that is more rhythmically correct and intelligible. Several of the current pronunciation textbooks on the market (Gilbert, 2012; Grant, 2017; Miller, 2005; Reed & Michaud, 2005; Yoshida, 2016) provide targeted and contextualized speech tasks aimed at assisting learners in correcting these common areas of difficulty, thus boosting learners' overall intelligibility and self-confidence. There are also helpful instructional videos that teachers and learners can access that provide helpful guidance (See Cox, Henrichsen, Tanner & McMurry, 2019).

Tier 2 concerns would include common areas of difficulty in segmental use that impact overall intelligibility and would require a greater degree of targeted and sustained instructional treatment. Such difficulties as shown in the sample data (see Figure 1) might include difficulty distinguishing between tense and lax vowels in English, the production of consonant clusters, the production of appropriate word endings such as -ed (past tense), and plural or third-person singular -s. Other areas of concern to address in this second tier would include the inability of the L2 speaker to produce aspirated or final stops, the rounding of mid-vowels, and the improper production of the

schwa vowel. As stated, each of these areas will take more specific instructional focus and practice than features in Tier 1.

Finally, Tier 3 features might be those features that are less frequently produced, or which may be of high frequency, but do not significantly impact an individual learner's intelligibility. They may also be features that are uniquely tied to negative transfer from a learner's native language such as the production of the /b/ vs. /v/ distinction for Spanish speakers of English. This tier would also contain those items where the ease of correction might be quite difficult or take a considerable amount of time to change. In the speech difficulties given for a sample Spanish L2 speaker of English, Tier 3 features could include consonant sounds substituted for the 'th' sound, such as /s/, /z/, or /t/. Another example could be the production of a rolled or flapped /r/ sound instead of the retroflex /r/ typically produced by native English speakers. Given the discourse context, these example difficulties often do not keep a native speaker from interpreting what the non-native English speaker is saying. Other difficulties could include a lack of reduced vowels in unstressed syllables or using a substituted sound for affricates in English. Overcoming the production of these segmentals could take months of sustained practice to correct elements of negative transfer that could be due to differences between how sounds are produced in the native language and those produced in English.

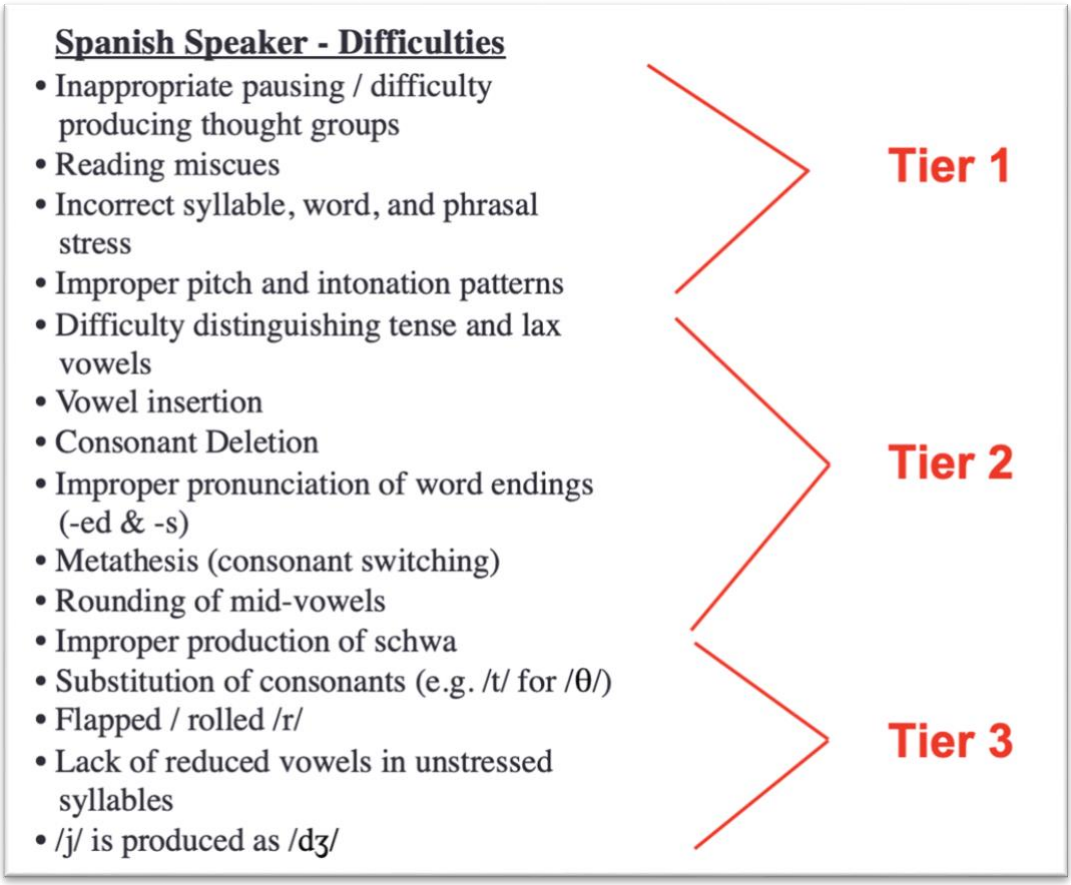

Figure 1. Example of three instructional tiers based on a speech analysis of an intermediate-level L2 Spanish speaker of English. *Note*: The pronunciation difficulties displayed in this example are not necessarily arranged in a hierarchical sequence within each tier. The tiers represent the immediacy, frequency of production, and ease of correction for the difficulties identified from the diagnostic data as it relates to the speaker's intelligibility.

Stage 4: Implementing the 'Zoom Principle': Classroom Instruction

Now that a tiered analysis has been constructed, revisiting the 'zoom principle' provides guidance for implementing classroom instruction. The zoom principle advocates that in incorporating pronunciation instruction into a classroom context, a teacher should start with a wide focus. In this wide focus, you have identified the global errors (errors that would typically fall into Tiers 1 and 2) that are regularly being produced by the students that may be impeding their intelligibility, are of high frequency, are common among most of the students in your class, and are features that you can address by providing target practice in the classroom (See Figure 2). As you develop your lesson plans, provide 10-15 minutes of instructed practice on a particular pronunciation difficulty that coincides with your three-tiered analysis, focusing on a feature in Tier 1 that is common to the majority of the students (e.g., producing correct thought groups through appropriate pause

insertion in the oral discourse). After targeted presentation and practice of this feature in the lesson, then move to contextualized practice of this feature before finishing the lesson. For example, if you had targeted the use of pausing to identify thought groups, contextualized practice could utilize a dialog script or the transcription of a portion of a TED talk given by a native English speaker or highly proficient non-native English speaker (Murphy, 2017, p. 16). Students listen to an audio recording or video recording of the discourse. While doing so, have them mark the transcript where the native speaker pauses while speaking. Once the passage has been marked, have the students rehearse saying the speech sample along with the native speaker model while focusing on grouping the words together in appropriate thought groups. Utilizing contextualized practice following the targeted practice helps the learner better internalize the pronunciation feature by embedding it in meaningful and communicative practice in larger segments of discourse (sentences and paragraphs).

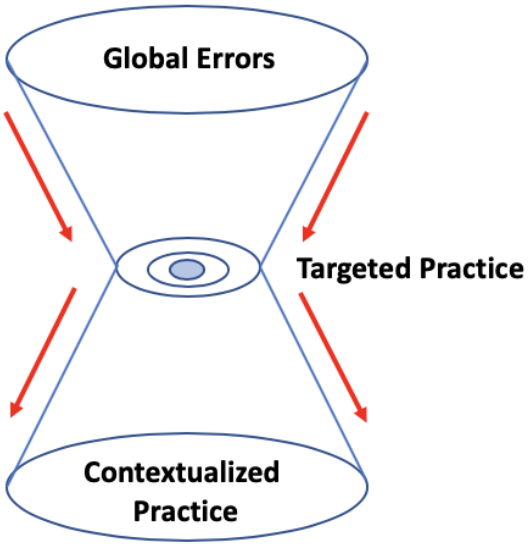

Figure 2. The 'zoom principle': moving from knowing global errors to focused, targeted practice in the classroom, to contextualized practice.

Sample Contextualized Practice Activities:

Teachers can incorporate a variety of different activities in the L2 classroom to engage learners in contextualized practice. These activities enable learners to transfer the knowledge gained in the targeted practice tasks to communicative practice activities that involve extended pieces of discourse beyond the word level. Several example activities are listed below with references that give additional explanation as to how these techniques can be integrated into a listening/speaking class:

- Dramatic monologues / Speeches or presentations (e.g. TED talks) using mirroring practice (Meyers, 2018)

- Jazz chants (Graham, 1978) incorporating Cued Pronunciation Reading (CPR) application (Tanner & Landon, 2009)
- Imitative conversations (Goodwin, 2005)
- Video voiceovers (Henrichsen, 2015)
- Reader's Theatre (Tanner & Chugg, 2018)

In all of these samples of scripted discourse, teachers can have learners transcribe speech and then mark it for those features that have been included in the focused, targeted practice. In the case of a Reader's Theatre script, students can listen to the conversation had by the various characters and then mark the script for pause location (indicating thought groups), word stress, and pitch changes at the end of each thought group (See Figure 3.)

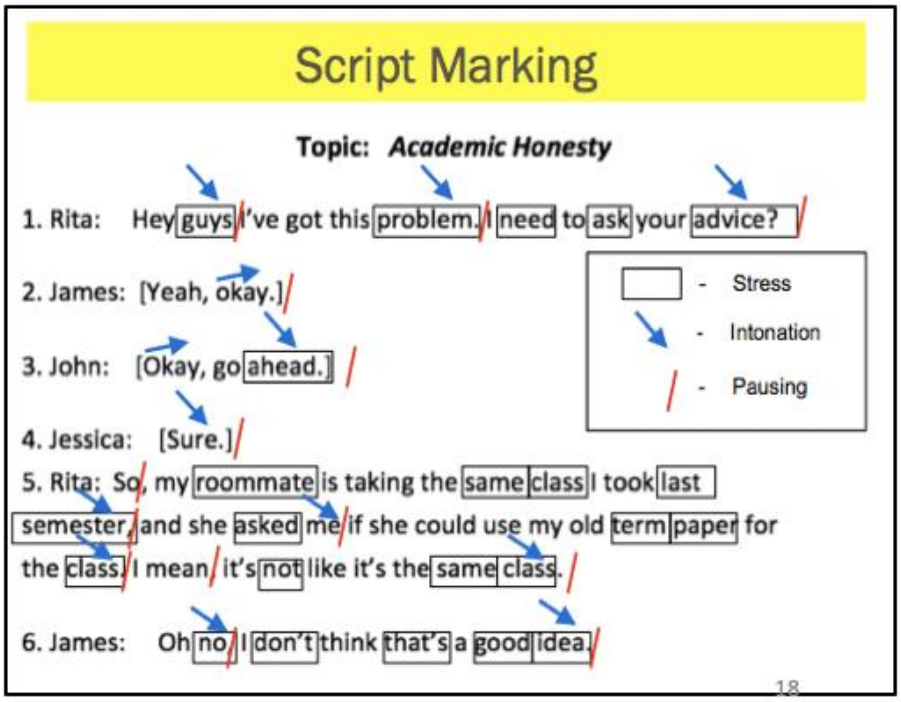

Figure 3. Example script marking of prosodic features for a portion of a reader's theater script on academic honesty.

CONCLUSION

When seeking to overcome the challenges that teachers often voice concerning the implementation of pronunciation instruction in the second language classroom, the 'zoom principle' (Firth, 1992) provides a helpful pedagogical framework. This framework guides teachers in how to diagnose learner difficulties and then systematically address these issues in focused, targeted practice, followed by contextualized communicative practice that goes beyond the word and sentence level using larger pieces of discourse. While the examples given in this article have focused on English pronunciation, the 'zoom principle' can help empower all second language teachers who desire to

help their students enhance their intelligibility and overall communicative ability in the target language.

ABOUT THE AUTHOR

Mark Tanner is an assistant professor in the Linguistics Department at Brigham Young University where he oversees the TESOL minor and certificate program and teaches in the MA TESOL program. His research focuses on second language pronunciation and pedagogy, fluency and comprehensibility, and teacher education. He has taught ESL and EFL for 41 years including stints in Thailand and Taiwan.

REFERENCES

Avery, P. & Ehrlich, S. (1992). Common pronunciation problems. In P. Avery & S. Ehrlich (Eds.), *Teaching American English Pronunciation* (pp. 95-109). Oxford: Oxford University Press.

Byrne, D. (1967). *Progressive picture compositions: Pupil's book*. Hong Kong: Longman.

Cohen, A. (1975). Error correction and the training of language teachers. *The Modern Language Journal, 59*(8), 414-422. https://doi.org/10.2307/325488

Cox, J., Henrichsen, L., Tanner, M., & McMurry, B. (2019). The needs analysis, design, development, and evaluation of the *English pronunciation guide: An ESL teachers' guide to pronunciation teaching using online resources. TESL-EJ, 22,* 1-24.

Firth, S. (1992). Pronunciation syllabus design: A question of focus. In P. Avery & S. Ehrlich (Eds.), *Teaching American English pronunciation* (pp. 173-183). Oxford: Oxford University Press.

Foote, J., Holtby, A., & Derwing, T. (2011). Survey of the teaching of pronunciation in adult ESL programs in Canada, 2010. *TESOL Journal Canada, 20*, 1-22. https://doi.org/10.18806/tesl.v29i1.1086

Gilbert, J. (2008). *Teaching pronunciation: Using the prosody pyramid*. Cambridge: Cambridge University Press.

Gilbert, J. (2012). *Clear speech: Pronunciation and listening comprehension in North American English* (4th ed.). New York: Cambridge University Press.

Goodwin, J. (2005). The power of context in teaching pronunciation. In J. Frodesen, C. Holten, & M. Celce-Murcia (Eds.), *The power of context in language teaching and learning* (pp. 225-236). Boston: Thomson Heinle.

Graham, C. (1978). *Jazz chants: Rhythms of American English for students of English as a second language*. New York: Oxford University Press.

Grant, L. (2017). *Well said: Pronunciation for clear communication* (4rd ed.). Boston, MA: National Geographic Learning.

Hahn, L. Primary stress and intelligibility: Research to motivate the teaching of suprasegmentals. *TESOL Quarterly, 38*, 201-223. https://doi.org/10.2307/3588378

Henrichsen, L. (2015). Video voiceovers for helpful, enjoyable pronunciation practice (Teaching Tip). In J. Levis, R. Mohammed, M. Qian, & Z. Zhou (Eds.), *Proceedings of the 6th Pronunciation in Second Language Learning and Teaching (PSLLT) conference*, ISSN 2380-9566, Santa Barbara, CA (pp. 270-276). Ames, IA: Iowa State University.

McNerney, M. & Mendelsohn, D. (1992). Suprasegmentals in the pronunciation class: Setting priorities. In P. Avery & S. Ehrlich (Eds.), *Teaching American English pronunciation* (pp. 1185-196). Oxford: Oxford University Press.

Meyers, C. (2018). Mirroring: A top-down approach to pronunciation instruction. Retrieved from https://www.pronunciationforteachers.com/uploads/6/0/5/9/60596853/teaching_techniques_mirroring_cmeyers.pdf

Miller, S. (2005). *Targeting pronunciation: Communicating clearly in English*. New York: Houghton Mifflin.

Murphy, J. (2017). Introduction. In J. Murphy (Ed.), *Teaching the pronunciation of English: Focus on whole courses* (pp. 1-29). Ann Arbor, MI: University of Michigan Press.

Pickering, L. (2001). The role of tone choice in improving ITA communication in the classroom. *TESOL Quarterly*, *35*, 233-253. https://doi.org/10.2307/3587647

Prator, C. & Robinett, B. (1985). *Manual of American English Pronunciation* (4th Ed.). New York: Holt, Rinehart & Winston.

Reed, M. & Michaud, C. (2005). *Sound concepts: An integrated pronunciation course*. New York: McGraw-Hill.

Sicola, L. & Darcy, I. (2015). Integrating pronunciation into the language classroom. In M. Reed & J. Levis (Eds.), *The handbook of English pronunciation* (pp. 471-487). West Sussex, UK: John Wiley & Sons.

Tanner, M. & Chugg, A. (2018). Empowering adult ELLS' fluency and pronunciation skills through readers theater. In J. Levis (Ed.), *Proceedings of the 9th Pronunciation in Second Language Learning and Teaching (PSLLT) conference*, ISSN 2380-9566, University of Utah, September 2017 (pp. 185-193). Ames, IA: Iowa State University.

Tanner, M. & Landon, M. (2009). The effects of computer-assisted pronunciation readings on ESL learners' use of pausing, stress, intonation, and overall comprehensibility. *Language Learning and Technology, 13*(3), 51-65.

Yoshida, M. (2016). *Beyond repeat after me: Teaching pronunciation to English learners.* Alexandria, VA: TESOL Press.

APPENDIX A – Accent Inventory

Accent Inventory
(From Prator and Robinett, 1985)

(1) When a student from another country comes to study in the United States, he has to find out for himself the answers to many questions, and he has many problems to think about. (2) Where should he live? (3) Would it be better if he looked for a private room off campus or if he stayed in a dormitory? (4) Should he spend all of his time just studying? (5) Shouldn't he try to take advantage of the many social and cultural activities which are offered? (6) At first it is not easy for him to be casual in dress, informal in manner, and confident in speech. (7) Little by little he learns what kind of clothing is usually worn here to be casually dressed for classes. (8) He also learns to choose the language and customs that are appropriate for informal situations. (9) Finally, he begins to feel sure of himself. (10) But let me tell you, my friend, this long-awaited feeling doesn't develop suddenly, does it. (11) All of this takes will power.

From: Prator, C. & Robinett, B. (1985). *Manual of American English pronunciation* (4th ed.). New York: Holt, Rinehart & Winston.

Task: After listening to the passage, create a list of pronunciation difficulties this speaker has in speaking English. (You may need to listen multiple times.)

APPENDIX B – Sample picture sequence. From Byrne, D. (1967). *Progressive picture compositions: Pupil's book*. Hong Kong: Longman.