Imber, B., Maynard, C., Parker, M. (2017) Using Praat to increase intelligibility through visual feedback. In M. O'Brien & J. Levis (Eds). *Proceedings of the 8th Pronunciation in Second Language Learning and Teaching Conference*, ISSN 2380-9566, Calgary, AB, August 2016 (pp. 195-213). Ames, IA: Iowa State University.

TEACHING TIP

USING PRAAT TO INCREASE INTELLIGIBILITY THROUGH VISUAL FEEDBACK

Brenda Imber, PhD, University of Michigan
Carson Maynard, MA, University of Michigan
Maria Parker, MS, Duke University

In our oral skills courses, we use various feedback tools to help students learn how to analyze their speech production and how to become more actively involved in their own improvement. We have found that one of the most effective ways to develop this learner autonomy, especially in pronunciation, is with feedback that is visually demonstrable at the point of occurrence. In our experience, visual feedback has been crucial for helping learners both notice and analyze their errors. In this teaching tip, we explain how we use Praat as a visual-feedback tool for guided feedback practice in the classroom and for self-reflection homework.

INTRODUCTION

As pronunciation instructors, our goal is not to transform students into "accent-free" speakers but rather to help them improve control over their production of consonants, vowels, and prosodic elements to increase intelligibility – or, as we say, to strive for "approximation, not perfection" (Abercrombie, 1991 [1945]; Celce-Murcia, Brinton & Goodwin, 1996; Harmer, 1993; Jenkins, 1998; Kenworthy, 1987; Morley, 1991; Munro & Derwing, 1995; Neri, Cucchiarini & Strik, 2006; Seferoğlu, 2005). Many divergences from a native speaker (NS) standard—in grammar, word choice, fluency, and accentedness—do not cause communication difficulties; indeed, like many pronunciation instructors, we encourage our students to embrace these non-critical differences, both as a way of signaling to NSs their non-native speaker (NNS) status (which can provide social capital in the event of a speech mistake that could be considered serious or offensive if made by a NS) and as a form of multilingual and cultural pride. What actually constitutes an error depends to some degree on the individual student; learners from the same L1 background may share similar divergences, but the degree to which each student's intelligibility is negatively impacted will depend largely on the frequency (Munro & Derwing, 2006) and markedness (Santos, 1987) of their particular divergences, among other factors (Levis, 2007). Jenkins (2000, 2002) proposed a set of highly-salient features that require the greatest attention from both students and teachers, called the Lingua Franca Core (LFC), including categories such as high-frequency commonly conflated consonant sounds, coda consonant deletion, and voice quality, while sidelining less critical features such as /θ/ and /ð/, dark [ɬ], or vowel reduction, which, if misarticulated, do not typically cause significant comprehension difficulties for the listener. We generally follow Jenkins's recommendations, though with the caveat that, particularly given our students' ESL/EAP environment, and because many of them are training to become teachers, their self-confidence is also a significant factor for learning. So, if students wish to align themselves more closely with the standard for certain features that they feel, rightly or wrongly, are impacting their intelligibility, we find it worthwhile to first discuss with them the

relative importance of those features for their particular context. Then, if they still feel strongly about it, we help them finesse their control of those features.

In this effort, we continually look for ways to effectively incorporate software, multimedia websites, and other technology into instruction. In our courses, students aim to approximate academic US English, so the resources we offer in this paper are tailored to that audience; however, if your students are targeting a different variety of English or if you teach a language other than English, you will of course want to use models or resources in that dialect or language.

Error correction relies on learner awareness, and it has long been agreed that the nature and timing of feedback can increase this awareness (Annett, 1969; Imber, Maynard & Meechan, 2012; Kulik & Kulik, 1988; Panova & Lyster, 2002). Ideally, then, we should provide feedback on students' pronunciation strengths and weaknesses when they occur, but without interrupting the flow of speech. In the absence of an oral equivalent of the "track changes" feature in Microsoft *Word*, this remains a challenging prospect.

Fortunately, the increasing ubiquity and variety of software designed for language learners is beginning to level the playing field. In our opinion, computer-assisted visual feedback is one of the most promising avenues for providing point-of-occurrence comments to help students improve intelligibility in their target dialect.

In pronunciation research specifically, visual feedback of various types has long been touted as a method of teaching intonation (Anderson, 1960; Cranen, Weltens, de Bot & van Rossum, 1984; de Bot, 1980, 1983; Hardison, 2004; Hermes, 1998; Hirata, 2004; James, 1979; Levis & Pickering, 2004; Ruellot, 2007; Seferoğlu, 2005; Taniguchi & Abberton, 1999; van Wieringen, 1994; Weltens & de Bot, 1984), perhaps because it cannot be seen by the naked eye in the same way that the articulation of segmentals is often observable (e.g., the position of the tongue, lip rounding). However, a non-expert can interpret a pitch contour representing intonation more intuitively than a spectrogram, making visual feedback a more natural fit for teaching intonation.

Other researchers – notably Brett (2004), Gonet (2001), Lambacher (1999), Landahl & Ziolkowski (1995), and Molholt (1990), – have successfully used computer-assisted visual feedback to teach segmentals, as well. Few, though, had advocated for using visual feedback to teach global pronunciation skills until Derwing, Munro & Wiebe (1998), who argued that students could be trained to improve their speaking rate, intonation, rhythm, vocal projection (volume and clarity), word stress, and sentence stress through visual feedback. Other studies have confirmed that visual feedback has led to improved perception and production of both segmentals and suprasegmentals (Hirata, 2004; Seferoğlu, 2005), particularly when students were able to visually compare their own output to a model speech sample (Hermes, 1998; James, 1979).

For visual feedback to be effective, it must include instructor monitoring, so that the student receives concrete feedback to help him or her interpret the acoustic display. As de Bot and Mailfert (1982) observe, "it became clear that a student could not just be placed in front of a display and told to learn correct intonation" (p. 72). Moreover, a number of researchers (Chun, 1989; Hirata, 2004; Neri, Cucchiarini, Strik & Boves, 2002; Pennington, 1999) agree that simply providing an acoustic display is unhelpful; the huge amount of raw acoustical data is

overwhelming, and it provides no guidance to students on how to determine which features warrant focus, how to modify their speech to approximate a model, or how closely aligned with the model their speech must be in order to be considered "acceptable." Watson and Kewley-Port (1989) agree: "The critical information necessary to modify erroneous speech productions should be presented as a salient feature of the display, rather than as one that must be sought in the midst of a complex array of irrelevant detail" (p. 37). Thus, instructors must help their students recognize which features of their acoustic display are relevant to their errors, and which differences do not affect listener understanding, so that they can eventually learn to study a display and pinpoint their personal key problem areas, specifically those that impact their intelligibility.

Of the various visual-feedback programs available, our favorite, *Praat* (Boersma & Weenink, 2017), was not originally intended for pedagogical purposes. In a nutshell, *Praat* is a robust speech analysis program developed by phoneticians for phoneticians who study features such as vowel duration, voice onset time, and formant frequencies. However, it can also be highly accessible to pronunciation instructors—even those with little experience—and, ultimately, to students. Despite the program's complexity (and the potentially intimidating level of detail in its text-heavy interface), what drew us to *Praat* is that it also provides an extremely useful visual depiction of a student's speech. Readers who have used *Audacity*[1] (Mazzoni & Dannenberg, 1999) may note similarities, although in our opinion, while *Audacity* is initially more intuitive and user-friendly, we have found *Praat* to be more useful for our purposes because it provides a number of features simultaneously. The standard *Praat* display (*Figure 1*) shows:

- a **waveform**, which illustrates volume (i.e., intensity) and the presence or absence of linking (i.e., connecting words without producing breaks between them) and pausing;
- a **spectrogram**, for pinpointing segmentals and word boundaries[2];
- a **pitch contour**, for measuring pitch range and identifying stressed syllables, which together contribute to intonation;
- and **timing bars**, for evaluating speaking time and vowel duration.

[1] This is another free speech recording and analysis software, which is available at http://www.audacityteam.org/.
[2] Though linking in speech makes it difficult (if not impossible) for a listener to distinguish discrete word boundaries, *Praat*'s visual display allows a user to at least roughly identify these delineations – though there can be significant overlap in the case of coarticulated phonemes – which can be useful, for example, in isolating a specific word or phrase that a student wishes to practice.

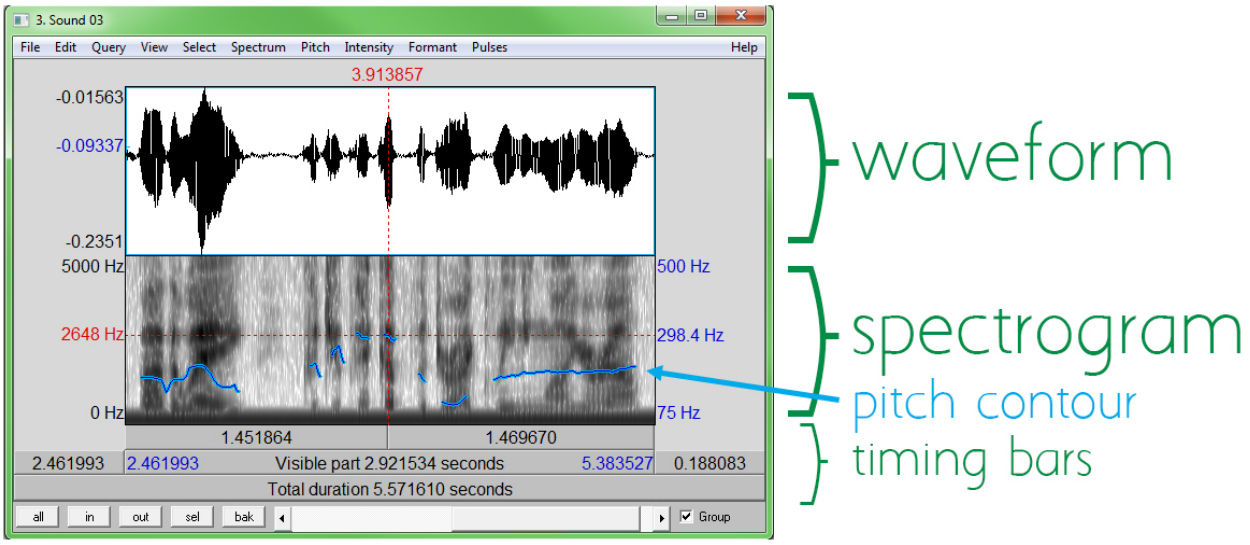

Figure 1. Praat display.

Students can analyze similarities and differences within these features when they compare their recording with the same utterance recorded by a native speaker. As we mentioned earlier, for non-experts (i.e., our students), spectrograms are less intuitive to interpret than the waveform, pitch contour and timing bars, but with guidance and practice, all four features can become useful resources for student improvement. A few examples are provided in figures 2-5 below.

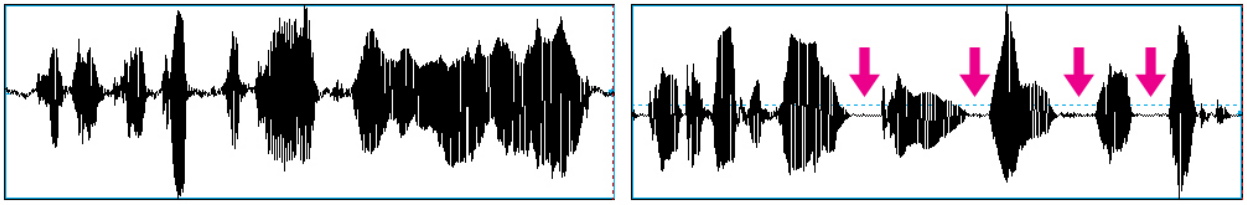

Figure 2. Linking comparison

Figure 2 shows part of the *Praat* waveform from Figure 1, for the phrase "Could you pick up the dry-cleaning while you're out?" The model, on the left, illustrates a well-linked phrase: even when the volume decreases between syllables, the waveform never completely flattens out (though it comes close during the unreleased [pˀ] at the end of "up"). In contrast, the student recording, on the right, has four distinct un-linked breaks after the words *dry*, *cleaning*, *while* and *you're*, indicated with arrows.

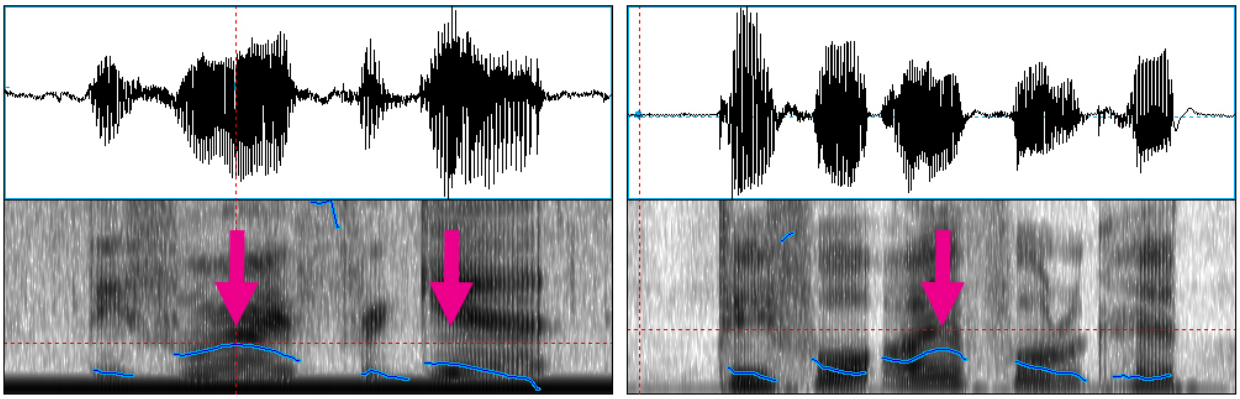

Figure 3. Stress comparison

Figure 3 shows a *Praat* display for the phrase "Before I forget…" The model recording, on the left, shows primary and secondary stress on the syllables *Before I forget*, indicated with arrows; the student recording, on the right, shows only primary stress on a different syllable: *Before I forget*. We typically ask students first to count how many stressed syllables they find in their recording vs the model, and then listen to each one to determine whether they've stressed the same syllables as the model. (This is most straightforward for formulaic expressions like the one above, where stress is relatively fixed. For other utterances, where a shift in stress might make no difference, create a difference in connotation, or sound unnatural, we discuss the appropriacy on a case-by-case basis.)

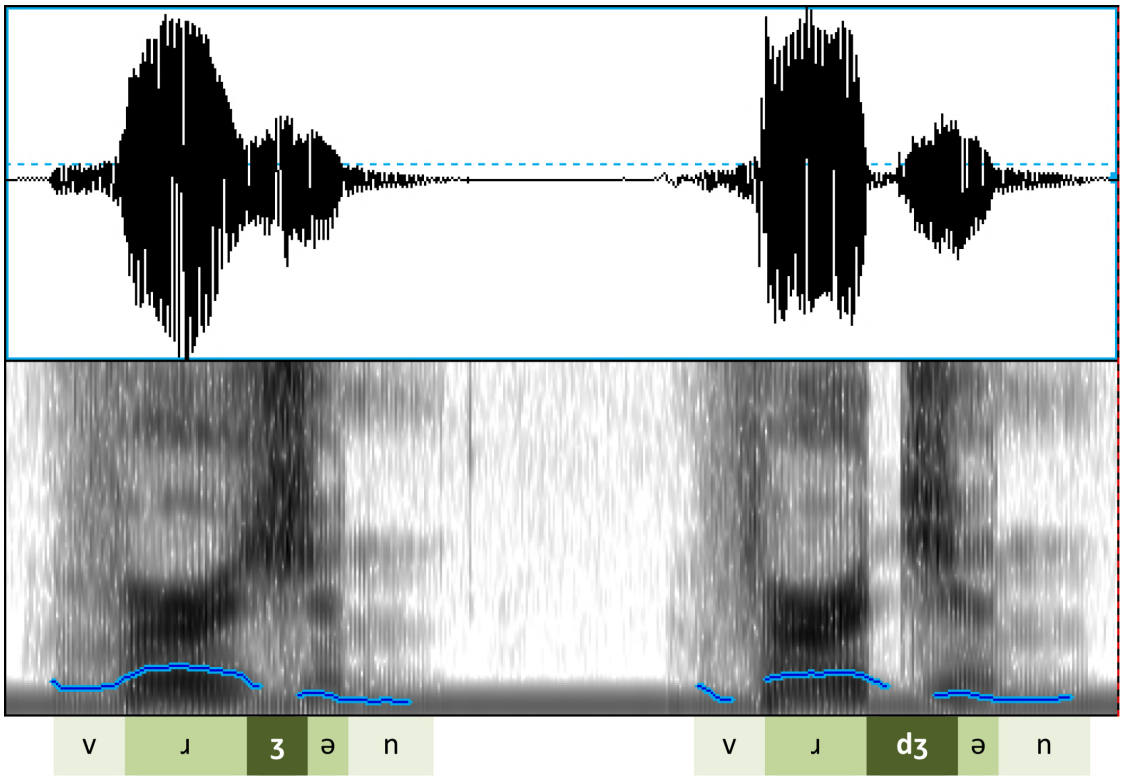

Figure 4. Consonant comparison

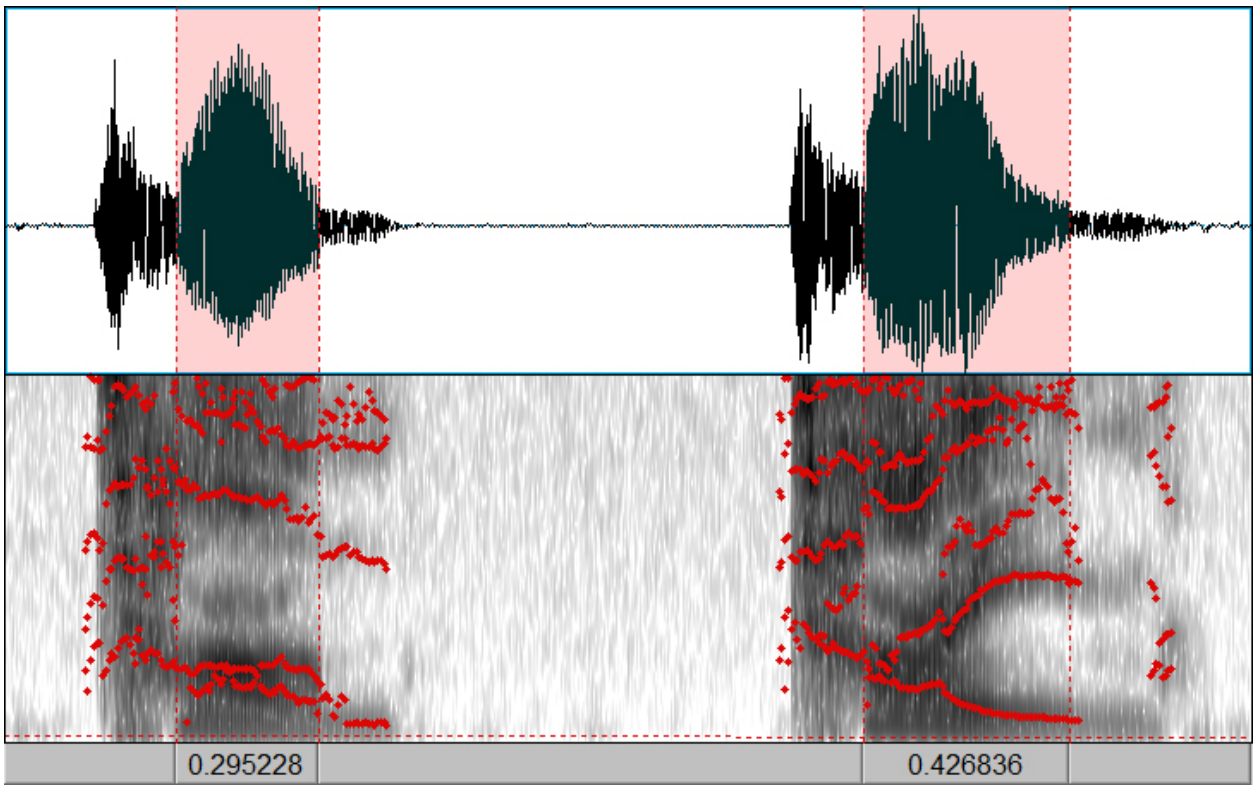

Figure 5. Vowel comparison

Figures 4 and 5 illustrate instances in which the *Praat* spectrogram can be helpful. Because teaching students to interpret spectrograms is labor-intensive and may not reap commensurate benefits, we prefer to use the spectrogram as a resource only when it clearly illustrates a sound difference that is causing significant problems. For example, in Figure 4, by showing the break caused by a stop consonant, *Praat* makes it evident whether a student is pronouncing fricative /ʒ/ (in *version*, left) or affricate /dʒ/ (in *virgin*, right). Similarly, our students often struggle not to monophthongize diphthongs, especially before nasals, which can create comprehension difficulties for pairs like *pain* and *pen*, or *ice* and… ahem. Figure 5 (a spliced image, with formant markers overlaid) illustrates how *Praat* can help students recognize whether they're pronouncing a shorter monophthong (as in *Tom*, left) or a longer diphthong (as in *time*, right, which shows a clearly shifting vowel). The spectrogram can also help to illustrate somewhat broader issues, like dropped final consonant clusters, deleted syllables, or epenthetic vowels. For more fine-grained distinctions, such as a conflation of /i/ and /ɪ/ in *gene* versus *gin*, /θ/ and /s/ in *math* versus *mass*, or /r/ and /l/ in *right* versus *light*, the spectrographic analysis becomes more cumbersome than we feel is worth our students' time and effort.

As always, we remind students that they need not match the model exactly, merely work to approximate the specific features that might cause confusion (in our experience, the most common features needing attention are linking, dropped syllables, epenthetic vowels, pitch, and stress placement). For example, baseline pitch is highly individualistic, correlating largely to

vocal tract length, which can vary significantly by gender (Childers & Wu, 1991; Rendall, Kollias, Ney & Lloyd, 2005). We find that our students tend to produce relatively flat intonation, so, rather than trying to match the model speaker's exact pitch, we encourage students to match the pitch *range* (as discussed in Appendix 1), which helps them more closely approximate native-speaker intonation.

Before computer-assisted visual feedback became readily available, self-studying students more commonly listened to audiocassettes or CDs in a language lab, having to rely on their intuition to guess whether they were pronouncing a word or phrase correctly. The opportunity to work with audio-visual depictions of their errors or strengths without interrupting the natural flow of speech creates a hands-on experience for the learner, and being able to see their speech displayed on the monitor gives students an immediate and quantifiable way to assess their pronunciation. Best of all, *Praat* is free to download, so students can use it outside the classroom at no expense.

It should be mentioned that, compared to the common visual shorthand of Microsoft Office, for example, *Praat* has a sharper learning curve for both instructors and students. As can be seen from Figure 1, the visual displays are packed full of information, so students require training in order to get comfortable using the program and even more training to learn to analyze their speech appropriately.

DISCUSSION

First, we provide an hour-long orientation at the start of the course, to introduce students to the specific functions and features that will be relevant and helpful. In order to provide further scaffolding, we encourage frequent usage of *Praat* both in and out of class, so that students become comfortable using it.

Our ultimate goal is for students to be able to use *Praat* for autonomous self-study, so in the weeks following the orientation, we help prepare students to "fly solo" by getting them gradually accustomed to using the program. During a ten-week course, we typically meet students once or twice a week for an hour. Throughout these confidence-building weeks, the three most helpful factors are instructor support, guided practice time, and an accountability procedure.

First, **instructor support** is essential because as noted, the displays are information-dense. We need to train students in what to look for, how to interpret what they're seeing, and then how to reflect on what modifications they might make.

Second, during **guided practice time**, the instructors facilitate the process of analyzing and providing feedback. To start, however, we must create a "sound bank" of model sentences with which the students can practice. We ask each student to invent their own sentence, incorporating formulaic expressions (from the Verbal Stratagems website in Table 1 below), terms from their field or from daily life, and words they have trouble pronouncing. We record a model version of each sentence, and the student then records their own version before downloading the model and comparing the two, as described above. This activity not only enables the whole class to get involved, but also generates a set of model sentences which students can use for later practice at home.

We continue with *instructor-led* analysis, in which we sit down with each student to help them compare and analyze their recordings. (For this reason, we like to hold our classes in a computer lab, so that we can spend at least the last 10 to 15 minutes working with students.)

In time, we transition to *pair or group-led* analysis, in which students analyze and discuss each other's recordings before reporting out to the class. To analyze their recordings, we can either move individually from student to student, or ask a particularly brave student to share their recording for the whole class to view and discuss.

Once students are comfortable and successful at analyzing without our help, we move to *independent* analysis, encouraging them to practice at home on their own. Because instructor modeling is critical for student improvement, we felt it was vital to make sure that students understood the software, had a chance to practice it with the instructors, and were comfortable using it before we asked them to practice at home and self-monitor.

Finally, these independent homework assignments are the core of our **accountability procedure**. We assign a set of formulaic expressions that students must expand into full sentences, as in our guided practice time activity, which they then record and submit. We urge students to *micro-practice,* i.e., practice of about 2-3 minutes, 5 times per day, because this spaced approach is far more effective for improving recall and building muscle memory than longer but less frequent sessions (deWinstanley & Bjork, 2002; Maley, 2017; Rohrer & Pashler, 2010; Wiklund-Hörnqvist, Jonsson & Nyberg, 2013). Additionally, breaking up the time into shorter chunks is not only more manageable in a busy schedule, but it also helps students to regard their practice time as motivational and fun rather than exhausting and boring.

We frequently found that students were engaged and enthusiastic about *Praat* in class, but still felt unsure about how exactly to use it without an instructor present, so we created a "Raising Awareness" activity (Appendix 1). Its purpose is two-fold: initially to train our students to use *Praat*, and ultimately as a step-by-step guide they can use for working independently.

In order to practice at home, students will need model sound files of their target dialect for comparison. Ideally, we recommend that you record the model files for your students yourself. If you do not have time, you lack a good recording environment, you want students to hear a different voice, you are not a speaker of your students' target dialect, or if your students want further practice materials after your course has ended, you could point your students to the short sample list of reliable .wav archives (of primarily US and UK English) provided in Table 1, where your students can download sound files for *Praat* that have been captured, for the most part, from TV and film. However, do alert students to avoid sound files that contain background noise, which muddies the acoustic display.

Table 1

WAV archives

Selected Internet Resources	
Praat	***http://www.praat.org/***
The Daily .WAV	*http://www.dailywav.com/*
The Frogstar Huge WAV Archive	*http://www.frogstar.com/wav/*
MovieWAVs	*http://www.moviewavs.com/*
WAVCentral	*http://wavcentral.com/*
WAVSource	*http://www.wavsource.com/*
UM ELI Verbal Stratagems (no WAV files)	*http://www.umich.edu/~flcourse/verbstrat/*

Figures 6, 7, and 8 below illustrate the potential for student improvement using *Praat*. The student whose acoustic displays are presented, LS, enrolled in a seven-week *Praat* workshop (as described in Imber, Maynard, Ohlrogge & Chien, 2009) for one hour a week. The workshop consisted of brief lessons on the topics from Appendix 1 followed by time for the students to practice with *Praat* in class. They were also required to keep a homework log to track how frequently they practiced on their own. Figure 6 shows the *Praat* display for one of the model sentences they were given: "By the way, could you pick up the dry cleaning while you're out?"

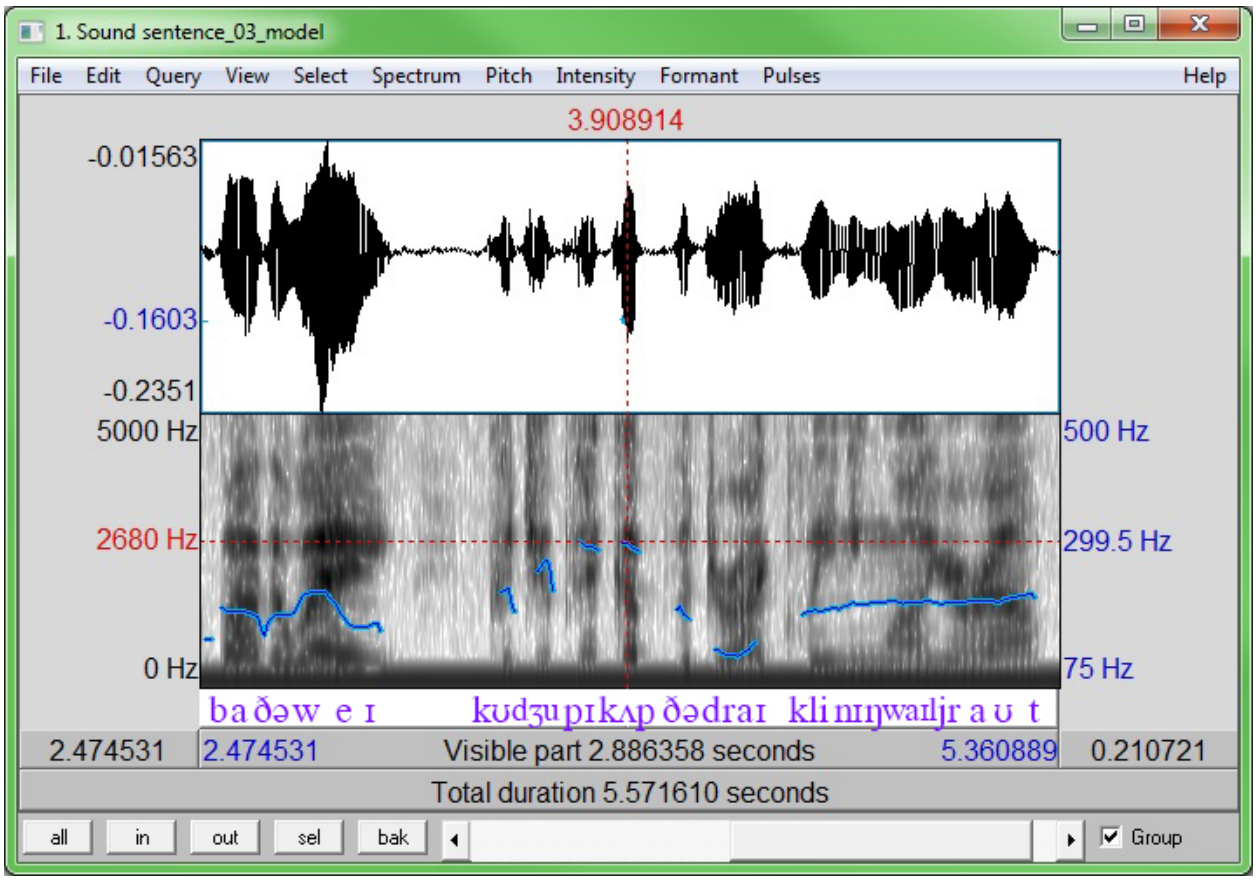

Figure 6. Model sentence

It is important to note that, as seen in Figure 6, the pitch contour does not display for voiceless sounds (e.g., Anderson-Hsieh, 1993), so the pitch contour may show breaks even if no pause is present in the recording. If learners have a particular need for pitch or intonation improvement, you may prefer to use model sentences with mainly voiced sounds. Another important point to note is that, as seen mildly in Figure 6 and dramatically in Figure 8, extraneous noises in the audio can produce a misleading swerve in the pitch contour, sometimes even causing it to jump to the extreme upper or lower limit of the spectrogram. While this is not a frequent issue, or a major drawback since the veracity of any apparent pitch changes can be easily double-checked by listening to the audio, it can be minimized by making your recordings in a location as soundproof and free of background noise as possible. Our location, evidently, was not: these sentences were recorded in a university computer lab.

As described in Appendix 1, students practicing with a model sentence are asked to compare their own speed, pausing and linking, pitch, stress, intonation, and pronunciation of segmentals to the model, making note of any significant similarities or differences. While students are welcome to work on any or all of the features that catch their attention, we encourage them to focus on, at most, their top 3 or 4 problem areas (assuming they have that many). Not only does this make their task more manageable, it also helps them concentrate their efforts on features that

affect their intelligibility, rather than spending their time on "accent reduction" of non-critical differences.

LS recorded this sentence, as part of a 6-sentence diagnostic set, twice: during the first-day orientation session (Figure 7) and again on the last day of the workshop (Figure 8). During the intervening six weeks, LS practiced with an additional 30 sentences, but was not exposed to the six diagnostic sentences again during that time period, nor did she ever view or listen to the model sound files for those sentences.

Each of the six diagnostic sentences, and the thirty subsequent practice sentences, were composed of two parts: a formulaic expression, followed by a colloquial English sentence. With the colloquial English sentence, we hoped to avoid giving students a sentence they might have seen before. Our hope was that, after using visual feedback to practice with other colloquial English sentences for six weeks, the students would have gotten a feel for the "flavor" of features like English-language linking and intonation. We used formulaic expressions to introduce these sentences because formulaic speech is much more "fixed" in terms of pausing, linking and intonation, as well as grammar; therefore, these phrases provided a less ambiguous target than non-formulaic utterances, and gave students an advantage in improving overall intelligibility.

The needs of any given student will of course be individualistic, largely influenced by their L1 and the degree to which any divergence from the model interferes with their intelligibility. A major benefit of *Praat* is that it facilitates this individualistic approach to error feedback. In the case of LS (Figure 7), her phonology, though noticeably accented—for example, [zə vei] for *the way*, [klinɪŋk] for *cleaning*—did not pose problematic barriers to her intelligibility. More salient were the amount of time it took her to produce the utterance and her intonation. LS's speed of 5.51 seconds (shown at **1** in Figure 7) was nearly twice as long as the model, at 2.89 seconds. The source of this discrepancy was largely to be found in her intonation (**2**), which was characterized by consistent rise-fall stress on almost every word ("**Could** you **pick** up the **dry clea**ning **while you're out**?"), even on the final word of the question (**3**), with frequent brief pauses.

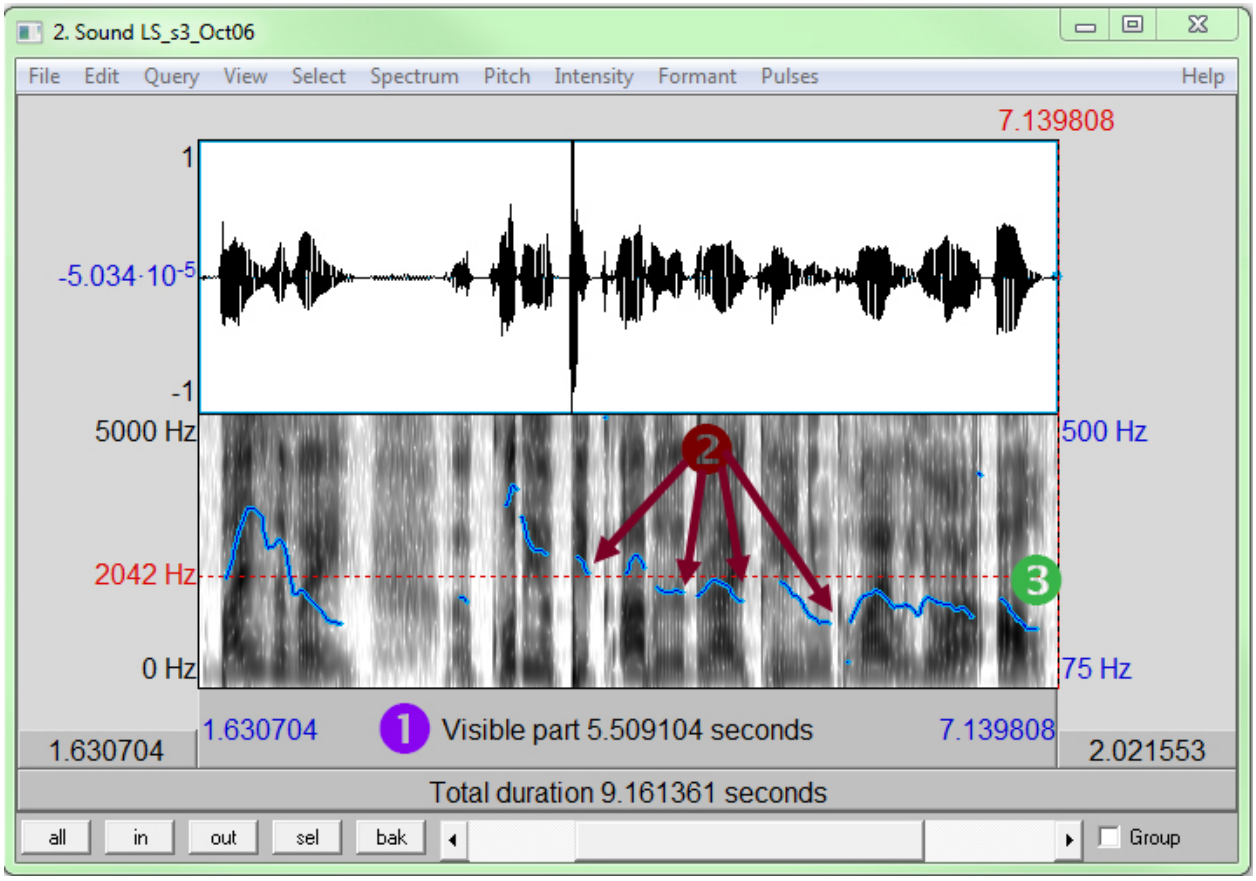

Figure 7. LS initial diagnostic

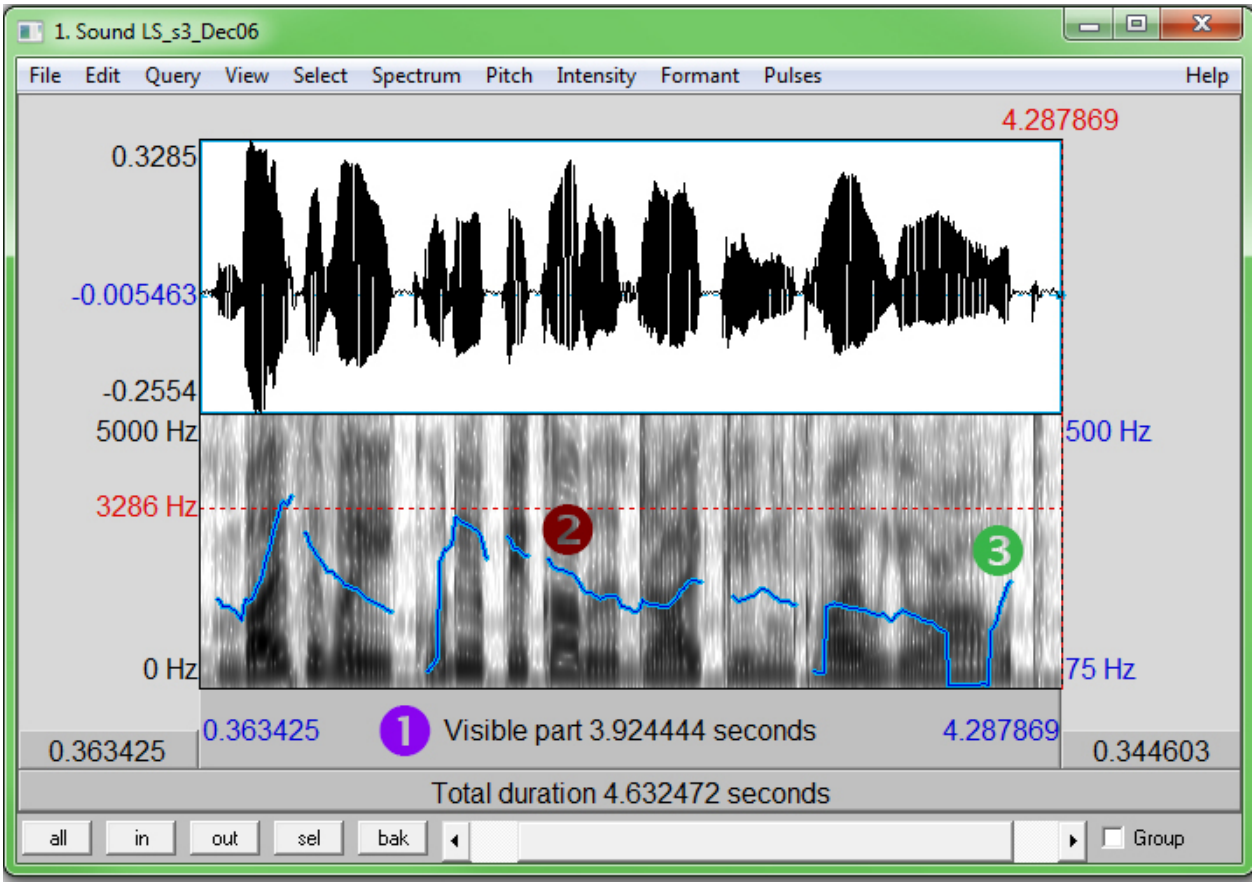

Figure 8. LS final diagnostic

From the visual display alone, we can see that after six weeks of practice, **(1)** LS's pace more closely approximates the model sentence (now 3.92 seconds, down from 5.51), and her intonation pattern has improved, **(2)** becoming smoother (stressing only the key words, and eliminating the frequent pauses), and **(3)** with phrase-final rising (rather than falling) intonation appropriate for yes/no questions. We would expect this more native-like intonation, in particular, to improve LS's intelligibility, given that research has demonstrated that unfamiliar (or unvarying) intonation contours can slow down or decrease comprehension (Braun, Dainora & Ernestus, 2011; Hillenbrand, 2003; Holub, 2010).

LS's progress was notably dramatic, even among her workshop peers. As with most pronunciation instruction, students typically see results with *Praat* after a longer period of regular practice. In general, compared with the control group, we saw mild improvement after six weeks of practice with *Praat*. Of course, for more pronounced improvement, most students will need a longer timespan, and we cannot overstate the importance of frequency of use.

Students should also note that it is difficult to improve segmentals with *Praat* in a self-study context, unless they can read spectrograms; otherwise, without visual support, they are entirely reliant on perception. Even in the classroom, students require a fair amount of instructor support for comparison and analysis of segmentals. The waveforms and pitch contour, however, are dramatically informative about the suprasegmentals – linking, stress, pitch and intonation – because these elements are relatively easy to interpret in waveforms and spectrograms, making it

much easier for students to compare and analyze prosody. But it is important to reiterate that not every marker of foreign accent or language control affects intelligibility (Derwing & Munro, 1997); only those divergences that interfere with comprehension need to be targeted for improvement. Although student goals tend to align with their intelligibility needs, if the instructor feels that a student has chosen a pronunciation goal that does not interfere with listener comprehension, this goal can be discussed with the student.

Finally, language level plays a role in how (or whether) students benefit from *Praat*. It is important that students' language skills be strong enough both to understand the instructions and to navigate the program, but not so advanced that they have little room for significant improvement. For example, at the University of Michigan, we have seen more pronounced results with visiting international scholars, who are working or studying in the USA temporarily, than with international graduate students, who had to meet a minimum TOEFL score for admission to a four-year degree program. Still, even this more advanced group often has considerable room for improvement, especially those who are aiming to become teaching assistants.

CONCLUSION

Both empirical research, which showed a moderate effect on fluency and intonation after seven weeks (Imber et al., 2009), and anecdotal comments from our students suggest that the use of visual feedback enhances their language-learning experience, not only in terms of measurable improvement but also self-confidence. As one of our students remarked, "If I can see it, I can say it, I can practice." Students appreciate being able to see differences, such as the "melody" of English, between their speech and the model, and value the ability to practice as much as they want without bothering the "speaker." While some training is necessary, both to accustom students to using the program and to focus only on the errors that are significantly impacting their intelligibility, we have found it to be time well spent, because quality visual feedback empowers students to self-assess their speaking skills with more certainty, less guesswork, and greater impact on their comprehensibility. Overall, the incorporation of a visual feedback tool such as *Praat* can have a positive impact on the language learning classroom for students interested in improving pronunciation and increasing intelligibility.

ABOUT THE AUTHORS

Brenda Imber, PhD, has taught at the University of Michigan English Language Institute (ELI) for over 25 years. Carson Maynard, MA, has worked at the ELI since 2003. Their professional interests include pronunciation, teacher training, and integrating technology into language learning.

Maria Parker, MS, has taught at Duke University English for International Students for over 25 years, and served as program director from 2000 – 2015. Parker and Imber co-authored *Integrated Lessons: Pronunciation & Grammar* (University of Michigan Press, 1993).

REFERENCES

Abercrombie, D. (1991 [1949]). Teaching pronunciation. In A. Brown (Ed.), *Teaching English Pronunciation: A Book of Readings* (87 – 95). London: Routledge.

Anderson, F. (1960). An experimental pitch indicator for training deaf scholars. *Journal of the Acoustical Society of America*, 32 (18), 1065 – 1074.

Anderson-Hsieh, J. (1993). Interpreting visual feedback on suprasegmentals in computer assisted pronunciation instruction. *CALICO Journal*, 11 (4), 5 – 22.

Annett, J. (1969). *Feedback and human behaviour: The effects of knowledge of results, incentives and reinforcements on learning and performance.* Oxford, England: Penguin Books.

Boersma, P., & Weenink, D. (2017). Praat: doing phonetics by computer (Version 6.0.25) [computer program]. Available from http://www.praat.org/

Braun, B., Dainora, A., & Ernestus, M. (2011). An unfamiliar intonation contour slows down online speech comprehension. *Language and Cognitive Processes*, 26 (3), 350 – 375.

Brett, D. (2004). Computer generated feedback on vowel production by learners of English as a second language. *ReCALL: Journal of Eurocall*, 16 (1), 103 – 113.

Celce-Murcia, M., Brinton, D. M., & Goodwin, J. M. (1996). *Teaching Pronunciation*. Cambridge, England: Cambridge University Press.

Childers, D. G., & Wu, K. (1991). Gender recognition from speech. Part II: Fine analysis. *The Journal of the Acoustical Society of America*, 90, 1841 – 1856.

Chun, D. M. (1989). Teaching tone and intonation with microcomputers. *Computer Assisted Language Instruction Consortium (CALICO) Journal*, 7 (1), 21 – 46.

Cranen, B., Weltens, B., de Bot, K., & van Rossum, N. (1984). An aid in language teaching: the visualization of pitch. *System*, 12 (1), 25 – 29.

de Bot, K. (1980). The role of feedback and feedforward in the teaching of pronunciation – an overview. *System*, 8, 35 – 47.

de Bot, K., & Mailfert, K. (1982). The Teaching of Intonation: Fundamental Research and Classroom Applications. *TESOL Quarterly*, 16 (1), 71 – 77.

de Bot, K. (1983). Visual Feedback of Intonation I: Effectiveness and Induced Practice Behavior. *Language and Speech*, 26 (4), 331 – 350.

Derwing, T. M., & Munro, M. J. (1997). Accent, Intelligibility, and Comprehensibility: Evidence from Four L1s. *Studies in Second Language Acquisition*, 19 (1), 1 – 16.

Derwing, T. M., Munro, M. J., and Wiebe, G. (1998). Evidence in Favor of a Broad Framework for Pronunciation Instruction. *Language Learning*, 48 (3), 393 – 410.

deWinstanley, P. A., & Bjork, R. A. (2002). Successful lecturing: Presenting. information in ways that engage effective processing. In Halpern, D. F., & Hakel, M. D. (Eds.), *Applying the Science of Learning to University Teaching and Beyond* (19 – 31). San Francisco, CA: Jossey-Bass.

Elliott, H. (2016). The Daily .WAV. Retrieved from http://www.dailywav.com/

Goebel, J. (2016). Frogstar.com: The Audio Collections. Retrieved from http://www.frogstar.com/wav/

Gonet, W. (2001). Obstruent Voicing in English and Polish: A Pedagogical Perspective. *International Journal of English Studies*, 1 (1), 73 – 92.

Hardison, D. M. (2004). Generalization of Computer-Assisted Prosody Training: Quantitative and Qualitative Findings. *Language Learning & Technology*, 8 (1), 34 – 52.

Harmer, J. (1993). *The practice of English language teaching*. New York, NY: Longman.

Hermes, D. J. (1998). Auditory and visual similarity of pitch contours. *Journal of Speech, Language and Hearing Research*, 41 (1), 63 – 72.

Hillenbrand, J. M. (2003). Some effects of intonation contour on sentence intelligibility. *The Journal of the Acoustical Society of America*, 114 (4), 2338 – 2338.

Hirata, Y. (2004). Computer assisted pronunciation training for native English speakers learning Japanese pitch and durational contrasts. *Computer Assisted Language Learning*, 17 (3 – 4), 357 – 376.

Holub, E. (2010). Does intonation matter? The impact of monotony on listener comprehension. *The Interpreters' Newsletter*, 15, 117 – 126.

Imber, B. P., Maynard, C., Ohlrogge, A., & Chien, Y-H. (2009). Using visual feedback to improve L2 pronunciation. In Katsampoxaki-Hodgetts, K. (Ed.), *Options & Practices of LSP Practitioners Conference Proceedings* (408 – 441). Heraklion, Crete: University of Crete Publications.

Imber, B. P., Maynard, C., & Meechan, P. (2012). An approach to video annotation for oral skills feedback. In Pattison, T. (Ed.), *IATEFL 2011 Brighton Conference Selections* (212 – 215). Eynsham, UK: Information Press.

Imber, B. P., Maynard, C., Parker, M., Reilly, M., Ohlrogge, A., Chien, Y-H.. Meechan, P., Cameron, P., & Beals, J. (2017). Verbal Stratagems. Retrieved from http://www.umich.edu/~flcourse/verbstrat/

James, E. F. (1979). Intonation through visualization. In Hollien, H. & Hollien, P. (Eds.), *Current Issues in the Phonetic Sciences* Vol. IV, Amsterdam Studies in the Theory and History of Linguistic Science (295 – 301). Amsterdam: John Benjamins.

Jenkins, J. (1998). Which pronunciation norms and models for English as an international language? *ELT Journal*, 52, 119 – 126.

Jenkins, J. (2000). *The Phonology of English as an International Language: New Models, New Norms, New Goals.* Hong Kong: Oxford University Press.

Jenkins, J. (2002). A sociolinguistically-based, empirically-researched pronunciation syllabus for teaching English as an international language. Applied Linguistics, 23, 83-103.

Kenworthy, J. (1987). *Teaching English pronunciation*. New York, NY: Longman.

Kulik, J. & Kulik, C. (1988). Timing of feedback and verbal learning. *Review of Educational Research*, 58 (1), 79 – 97.

Lambacher, S. (1999). A CALL tool for improving second language acquisition of English consonants by Japanese learners. *Computer Assisted Language Learning*, 12 (2), 137 – 156.

Landahl, K. L. & Ziolkowski, M. S. (1995). Discovering phonetic units: Is a picture worth a thousand words? *Papers from the Regional Meetings, Chicago Linguistic Society*, 31 (1), 294 – 316.

Levis, J. (2007, March). *Guidelines for promoting intelligibility.* Paper presented at the meeting of International TESOL, Seattle, WA.

Levis, J., & Pickering, L. (2004). Teaching intonation in discourse using speech visualization technology. *System*, 32, 505 – 524.

Maley, A. (2017, April). *Is what we know, what we do?* Paper presented at the meeting of International Association of Teachers of English as a Foreign Language (IATEFL), Glasgow, UK.

Mazzoni, D., & Dannenberg, R. (1999). Audacity (Version 2.1.3) [computer program]. Available from http://www.audacityteam.org/

Molholt, G. (1990). Spectrographic analysis and patterns in pronunciation. *Computers and the Humanities*, 24, 81 – 92.

Morley, J. (1991). The pronunciation component in teaching English to speakers of other languages. *TESOL Quarterly*, 25, 481 – 519.

Munro, M. J. & Derwing, T. M. (1995). Foreign accent, comprehensibility, and intelligibility in the speech of second language learners. *Language Learning*, 45, 73 – 97.

Munro, M. J. & Derwing, T. M. (2006). The functional load principle in ESL pronunciation instruction: An exploratory study. *System*, 34, 520 – 531.

Neri, A., Cucchiarini, C., Strik, H., & Boves, L. (2002). The pedagogy-technology interface in computer assisted pronunciation training. *Computer Assisted Language Learning*, 15 (5), 441 – 467.

Neri, A., Cucchiarini, C., & Strik, H. (2006). Selecting segmental errors in non-native Dutch for optimal pronunciation training. *International Review of Applied Linguistics in Language Teaching (IRAL)*, 44 (4), 357 – 404.

Panova, I., & Lyster, R. (2002). Patterns of corrective feedback and uptake in an adult ESL classroom. *TESOL Quarterly*, 36 (4), 573–595.

Pennington, M. C. (1999). Computer-aided pronunciation pedagogy: Promise, limitations, directions. *Computer Assisted Language Learning*, 12 (5), 427 – 440.

Rendall, D., Kollias, S., Ney, C., & Lloyd, P. (2005). Pitch (F0) and formant profiles of human vowels and vowel-like baboon grunts: The role of vocalizer body size and voice-acoustic allometry. *The Journal of the Acoustical Society of America*, 117, 944 – 955.

Rohrer, D., & Pashler, H. (2010). Recent research on human learning challenges conventional instructional strategies. *Educational Researcher*, 39, 406 – 412.

Ruellot, V. M. (2007). French pronunciation learning and computer-mediated visual feedback. *Dissertation Abstracts International, A: The Humanities and Social Sciences*, 67 (07), 2556.

Santos, T. (1987). Markedness theory and error evaluation: an experimental study. *Applied Linguistics*, 8 (3), 207 – 218.

Seferoğlu, G. (2005). Improving students' pronunciation through accent reduction software. *British Journal of Educational Technology*, 36 (2), 303 – 316.

Taniguchi, M., & Abberton, E. (1999). Effect of interactive visual feedback on the improvement of English intonation of Japanese EFL learners. *Speech, Hearing and Language*, 11, 76 – 89.

van Wieringen, M., & Abberton, E. (1994). The use of computerized representations in L2 acquisition of intonation: a pilot study. *Speech, Hearing and Language*, 8, 245 – 258.

Watson, C. S., & Kewley-Port, D. (1989). Advances in computer-based speech training: Aids for the profoundly hearing impaired. In McGarr, N. S. (Ed.), *The Volta Review*, 91, 29 – 45.

WavSource.com: Free Wav Files and Sound Bites. (2016). Retrieved from http://www.wavsource.com/

Weltens, B., & de Bot, K. (1984). The visualization of pitch contours: some aspects of its effectiveness in teaching foreign intonation. *Speech Communication*, 3, 157 – 163.

West, G. (2016). The MovieWavs Page. Retrieved from http://www.moviewavs.com/

White, B. (2016). Free Movie and TV Clips at Wav Central. Retrieved from http://wavcentral.com/

Wiklund-Hörnqvist, C., Jonsson, B., & Nyberg, L. (2013). Strengthening concept learning by repeated testing. *Scandinavian Journal of Psychology*, 55, 10 – 16.

Appendix 1: "Raising Awareness" activity

Guided Practice Exercise: Raising Awareness

1. **Load** Praat.
2. **Open** two files:
   i. Any model sound file
   ii. Your own recording of the same sound file
3. **Line them up** on your screen so that the model file is above your sound file.
4. Under *View > Show Analyses*, make sure that the **spectrogram** and **intonation line** are showing, and "formants" and "pulses" are **un-checked**. You can also increase the "Longest analysis" setting if the file is too long to display a spectrogram.
5. **Compare** your file to the model, and **look for** the following features:

i. Speed

- ☐ Is your sentence significantly longer or shorter than the model?
- ☐ If so, can you find any places where your speed was much faster or much slower?
- ☐ Can you tell what might be causing the difference? (If not, think about this as you analyze.)

Look for: *vowel sound duration, dropped words*.

ii. Pausing and Linking

- ☐ Do you have the same number of pauses (breaks) as the model?
- ☐ Are they roughly the same length?
- ☐ Are they in the same place? i.e., are your thought groups (chunks) different from the model?
- ☐ Do any of your pauses sever the links between words, or is your linking similar to the model?

Look for: *micro-pauses, dropped final consonants (some, door, feel, etc), vowel insertion, voiced pauses ("big-uh dog"), fillers (uh, um, well, etc)*.

iii. Stress, Pitch and Intonation

- ☐ How many words did you stress, compared to the model?
- ☐ Are there any phrases where your stress is on a different word?
- ☐ Are there any words where your stress is on the wrong syllable?
- ☐ How do your highest and lowest pitch compare to the model?
- ☐ Do you have a wide pitch range, or does your pitch stay "flat"?
- ☐ Do you see any phrases where your intonation was very different from the model?

Remember that **the *blue line*** shows **pitch, stress and intonation**.
If you click on **the peak or valley** of a syllable, **the *red number*** on the right (Hz) shows **how high or low** your pitch is.

iv. Pronunciation

- ☐ Are there any places where your words sound very different from, or similar to, the model?

Look for: *problematic sounds (r/l, th, sh/ch/zh, etc)*.