Dickerson, W. B. (2011). Upstream destressing: Another step toward natural speech. In. J. Levis & K. LeVelle (Eds.). *Proceedings of the 2nd Pronunciation in Second Language Learning and Teaching Conference*, Sept. 2010. (pp. 70-81), Ames, IA: Iowa State University.

UPSTREAM DESTRESSING: ANOTHER STEP TOWARD NATURAL SPEECH

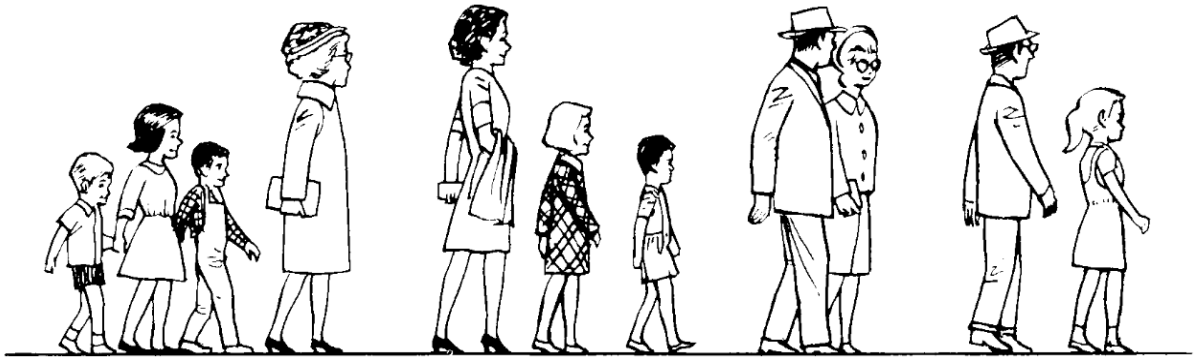
Wayne B. Dickerson, University of Illinois at Urbana-Champaign

English speech is known for its bumpiness—peaks and valleys of stress with one dominant peak variously called the discourse stress, primary phrase stress, sentence stress, nuclear stress, and tonic.

Less well known is the fact that native speakers selectively demote the peaks adjacent to the discourse stress to valley status for a variety of reasons. This is the phenomenon of destressing (or deaccenting).

This paper focuses on the least well-known part of the destressing story—upstream destressing. It is a systematic downgrading of the peak or peaks just before the discourse stress. We first describe what is involved phonetically in destressing, why this phenomenon occurs, and the rules governing its use. Then we examine the pedagogical implications of this phenomenon: How can we introduce the destressing behavior to linguistically naive learners? How can we help them predict when to destress peaks on their own? What does a destressed peak sound like? What kind of practice do they need so that this behavior will become a natural part of their oral performance?

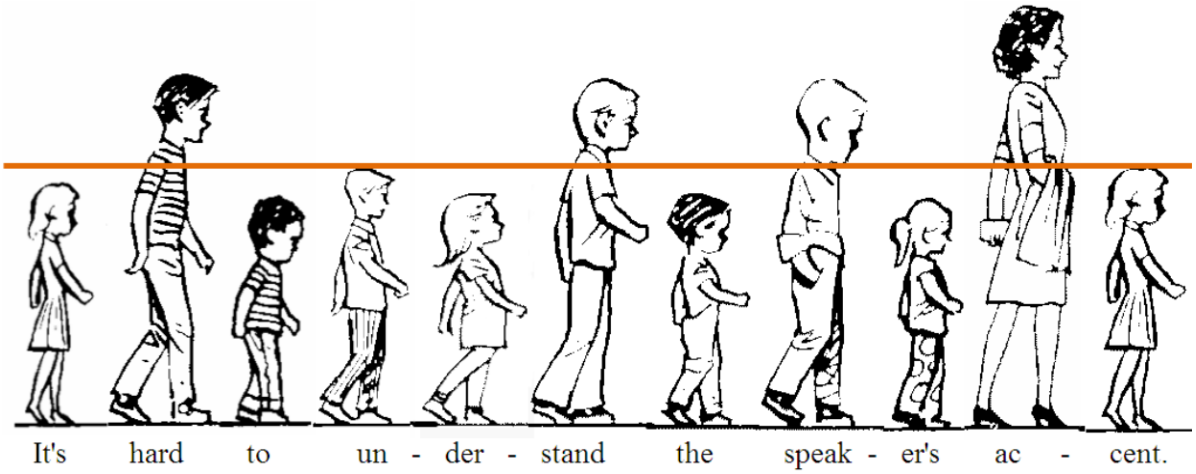
This paper centers on a stress behavior that is missing from the usual presentation of English rhythm and discourse stress in ESL pronunciation texts. The question is whether or not it should stay missing. As background, let's start with what is generally presented to ESL students about English rhythm and discourse stress. We begin with this famous picture of English rhythm from Prator & Robinett (1985, p. 29).



When the images of people are mapped on to a sentence like, *It's hard to understand the speaker's accent*, the graphic is sufficient to remind us of the key points.

First of all, English rhythm is really bumpy. We can describe this bumpiness using the landscape metaphor of peaks and valleys, by which we mean that some syllables (those above the line) have significantly greater stress—are louder, longer, and often higher pitched—than other

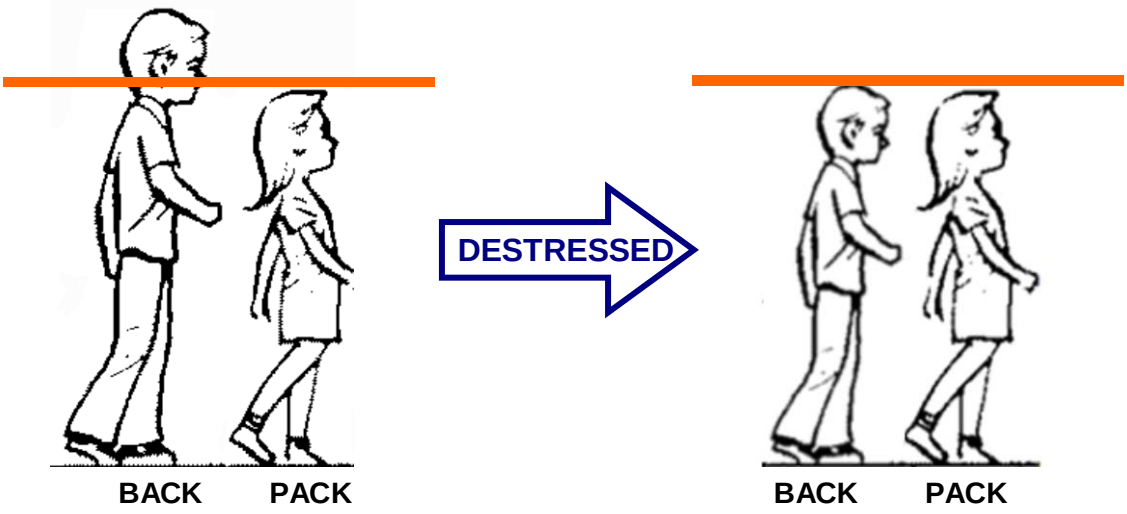
syllables (those below the line).



Second, each peak is the major stress of a content word. The syllables that populate the valleys are the non-major stressed vowels of content words and the vowels in function words—all below the line.

Third, a refinement distinguishes loud function words from soft function words. Loud function words are demonstrative pronouns, question words, and negative words and negative contractions—all of which behave like content words in that they carry a rhythmic peak. By contrast, soft function words—like conjunctions, articles, prepositions, pronouns, auxiliaries, modals, forms of the verb *to be*—belong in valleys.

Finally, although less commonly, pronunciation texts point out that any peak (content word or loud function word) coming after the primary or the nuclear stress in a phrase will be destressed or deaccented. That is, it will lose its rhythmic beat. One reason to destress words in that position is that those words contain information already mentioned in the discourse or assumed to be shared by the participants. They no longer need the listener's greatest attention. As an example, let's say that we've been talking about backpacks. In this context, I say, *I've already bóught her a backpack. Backpack*, which would carry a rhythmic peak on *back* if it were new information, is now destressed following the nucleus on *bought*. *Backpack* becomes part of a valley.

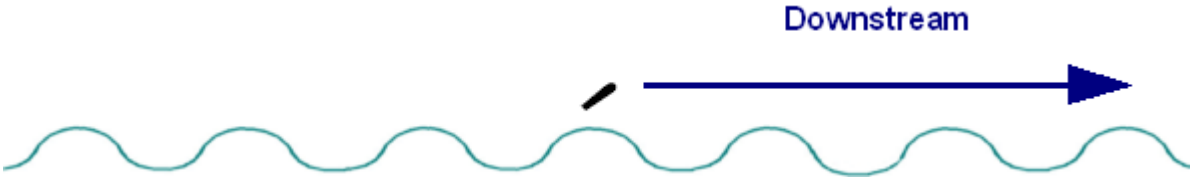


Since destressing is our topic, it is important to understand what this term refers to. First, destressing does not mean making a stressed vowel unstressed; when we destress the *back* of *backpack*, *back* does not become unstressed. Second, destressing does not affect vowel quality; it does not make a full vowel into a reduced vowel. The /æ/ in *back* of *backpack* is still /æ/ after being destressed; it does not become schwa. Third, destressing does not affect all stressed, full vowels. In *backpack*, both vowels are stressed and full, yet *pack* is unaffected by destressing. Destressing affects only the stressed, full vowel in the one syllable of a content word or loud function word that would ordinarily carry a rhythmic beat. In the case of *backpack*, that's the vowel in *back*. Fourth, the effect of destressing that particular vowel is to lower its stress so that it no longer carries a beat. Such stress lowering puts *back* into a rhythmic valley with *pack*.

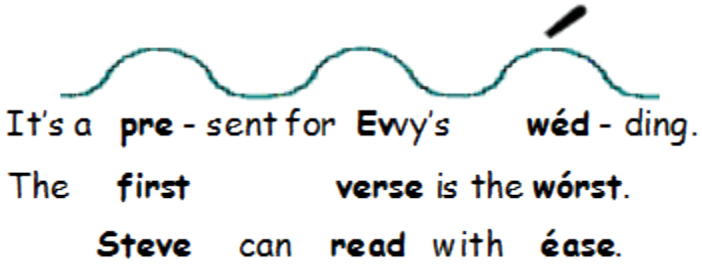
These basics about English rhythm and discourse stress are present in one way or another in many ESL pronunciation texts (e.g. Beisbier, 1995; Hahn & Dickerson, 1999; Prator & Robinett, 1985).

To go beyond this basic description, it will help to shift images. Let's borrow and extend the metaphor of the stream of speech, where the peaks in this stream represent content words and loud function words. The primary stress is one of these peaks.

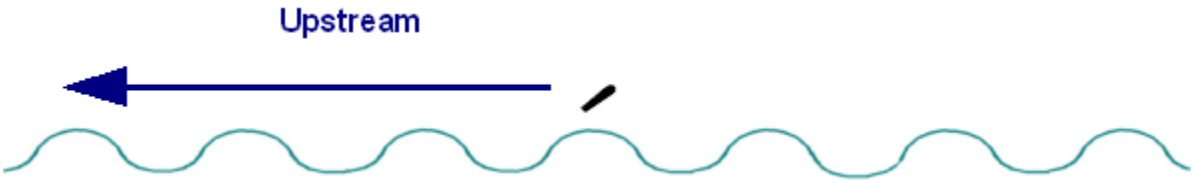
Continuing this metaphor, we can say that the destressing of *backpack* in *I've already bóught her a backpack* happens downstream from the primary stress—after the primary stress. Post-primary destressing is what we call **downstream destressing**. Bing (1980) and Bolinger (1986) discuss this notion at length because of its meaningfulness in discourse.



If we use this model for sentences with three peaks, they should easily fit into this diagram, contributing two rhythmic peaks up to the primary stress, as shown below.

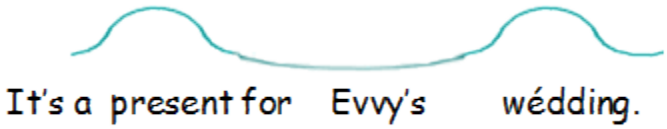


Here is where we leave the well-trodden path. What is much less well known and not represented in any ESL pronunciation texts I am acquainted with, is that the peak - valley alternation illustrated above is in fact quite rare in ordinary speech. That's because of another destressing phenomenon that happens before the nucleus, upstream from the primary stress. It's what we call **upstream destressing**.



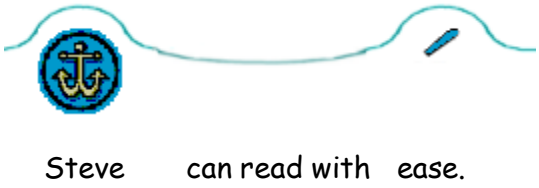
Since upstream destressing is the more specific focus of this paper and is not so well known, let's find out a bit more about it. Where does it occur?

First, like downstream destressing, upstream destressing affects only content words and loud function words—the kind of words that typically form rhythmic peaks. The effect of upstream destressing is to push these peaks into a valley, removing their rhythmic beat. In the following example, the first syllable of *Evvy* joins the preceding valley syllable, *for*, and the following valley syllable, *-vy*, to create an enlarged valley.



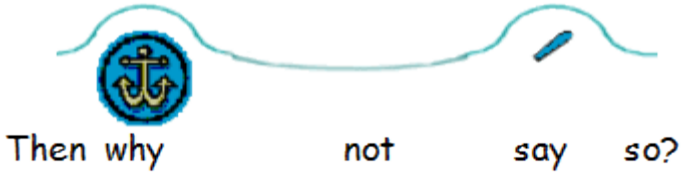
Second, it happens most commonly to the content word or loud function word nearest the primary stress on the left. It may happen further upstream, but only if it also happens nearest the primary stress.

Third, it doesn't happen if a content word or loud function word left of the primary stress is the first rhythmic beat of the phrase. So it wouldn't happen to *Steve* in *Steve can read*. But in the case of *Steve can read with ease*, it will happen—to *read*, right before the primary stress. The first rhythmic beat of a phrase acts as a kind of anchor in the phrase; that's *Steve*. In British parlance, *Steve* is called the onset of the phrase (Wells, 2006, p. 8). A peak stress in the anchor position is quite stable; it's not demoted.

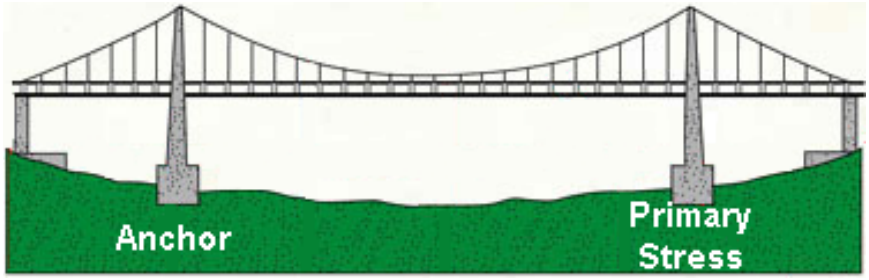


In short, the minimum conditions for upstream destressing to occur are three rhythmic stresses—an onset to the phrase, a primary stress, and a rhythmic stress between the two. It is this middle one that is ordinarily demoted to valley status.

Does the kind of rhythmic word between the anchor and the primary stress make a difference? Apparently not. Destressing seems to happen whether the middle peak is a noun, an adjective, a verb, or an adverb—a content word or a loud function word (as in the next example). *Why* is the anchor; the primary stress is on *say; not* is demoted.

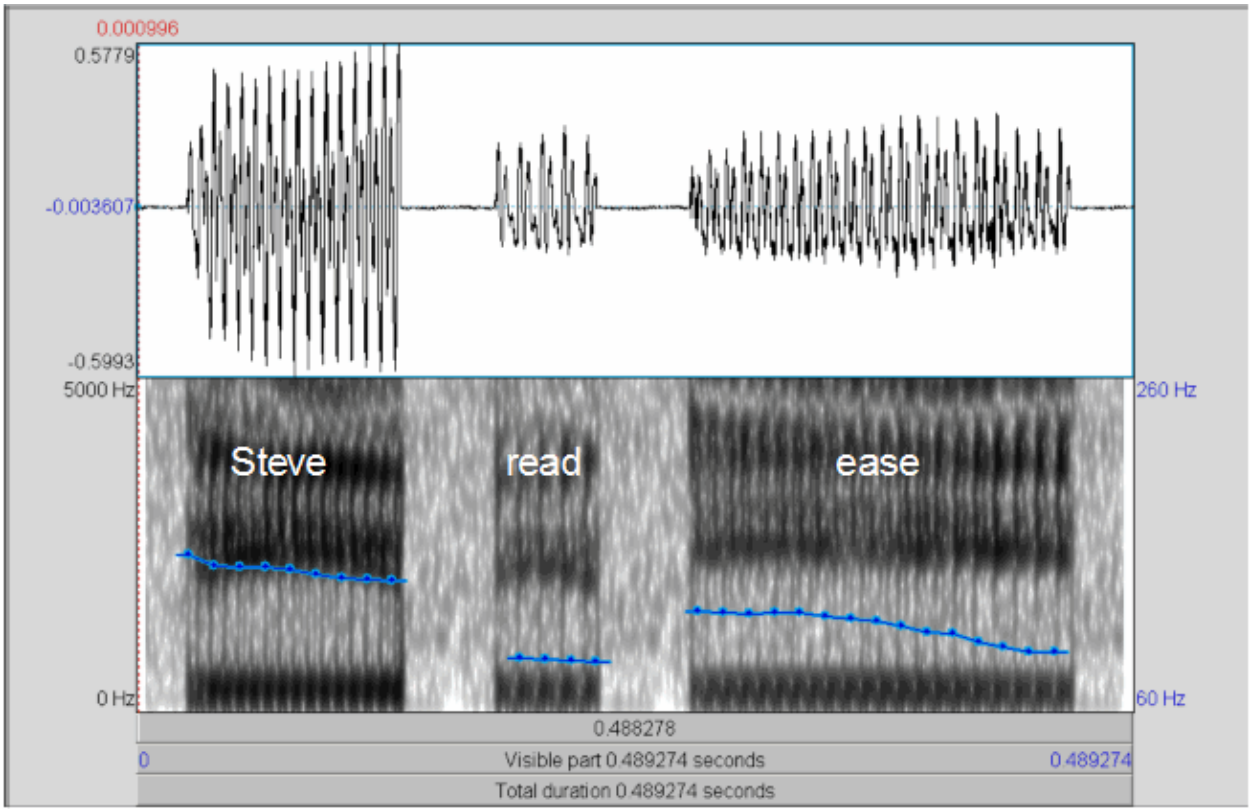


Referring to this overall pattern of anchor and primary stress ('positions of greatest impact'), Bolinger (1986, pp. 47-48) calls it a **suspension bridge**, where the 'two-accent shape' forms the 'two towers of the bridge'. Wong (1987, p. 56) picks up this apt metaphor in her guide to ESL pronunciation teaching, focusing on the towers without mentioning the destressing phenomenon between. Bolinger refers to the presence or absence of 'the sag in the suspension bridge' only as 'normal variants' (p. 47). Even though they do not give attention to what we call upstream destressing, Bolinger's metaphor is appropriate, as the graphics above and below illustrate.



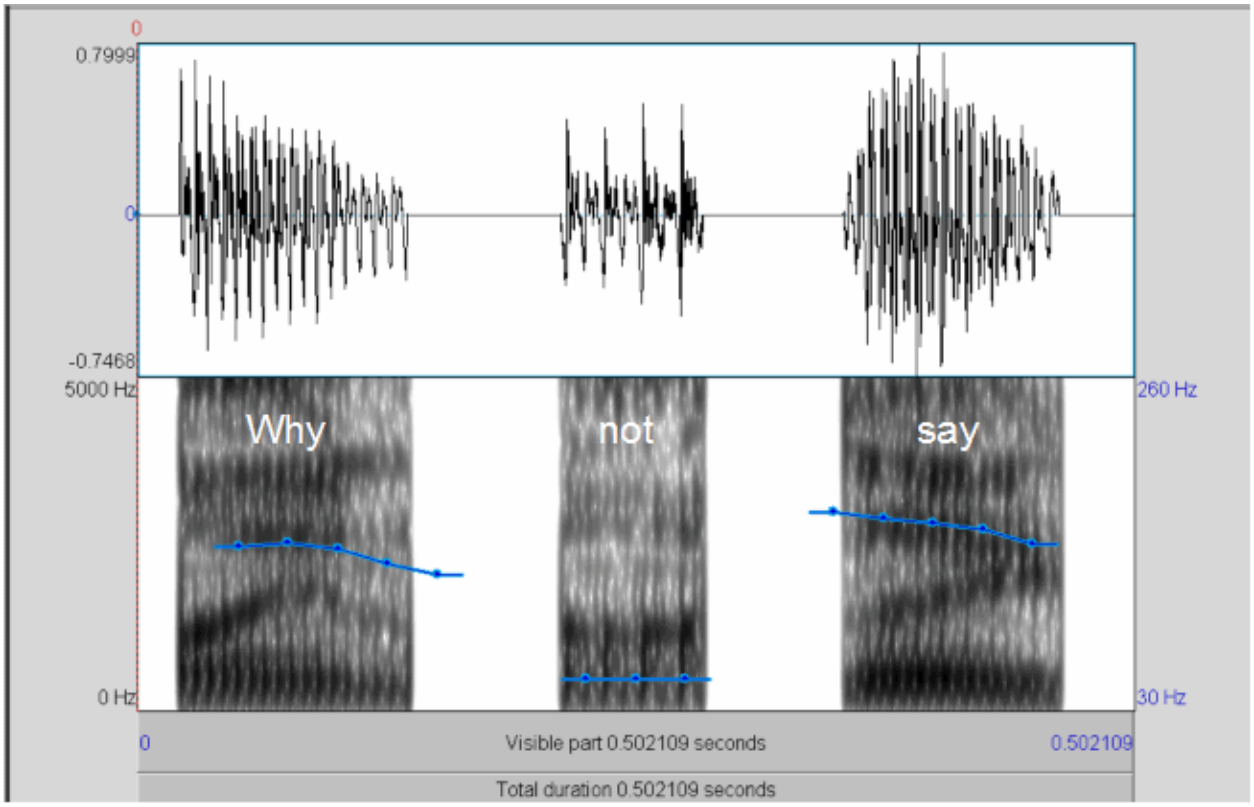
In contrast, Wells (2006) does take an interest in upstream destressing, saying, "The option to downgrade potential accents is a pervasive characteristic of English rhythm. It tends to operate whenever an accent is located between two other accents in the same IP [intonation phrase]" (pp. 229). Rather than appeal to a metaphor, Wells frames his discussion in terms of a rule he calls the Rule of Three.

What happens phonetically when we destress? As we know, stress in English is some composite of duration, pitch, and intensity. The consistent effect of destressing is to shorten the *duration* of the vowel. The following graphic displays the three [iy] vowels in *Steve can read with ease,* where the primary stress is on *ease*. With time along the bottom axis, the vowel in *read* is obviously much shorter than the vowel in *Steve*.



Depending on the intonation pattern, destressing may also drop the *pitch* of the word between the onset and nucleus as indicated by the blue lines at the bottom of the image. Furthermore, destressing reduces the *intensity* or loudness of the vowel. That is evident in the height of the wave form. There is much less energy in the [iy] of *read* than in the [iy] of *Steve*.

The same pattern is evident in *Why not say so?* where loud function words occupy the anchor and the position before the primary stress.



To understand the role of destressing, it may be helpful to compare and contrast upstream destressing with downstream destressing. While both are similar in affecting rhythmic peaks, they have quite different, non-overlapping functions otherwise. Upstream destressing has nothing *specific* to say about the message content. This is no doubt why Bolinger, who focuses on meaningful signals, dismisses the 'sag in the suspension bridge' as one of several 'normal variants' (p. 47). However, when it comes to enhancing the perception of the primary stress, upstream destressing trumps all. It serves as a marker or a listener alert: *Here comes a primary stress! Listen up!* It anticipates the upcoming nucleus by dropping the voice into a valley before the final peak. Furthermore, by expanding the valley before the primary stress, it maximizes the contrast between this valley and the primary peak so that the primary stress is strikingly highlighted.

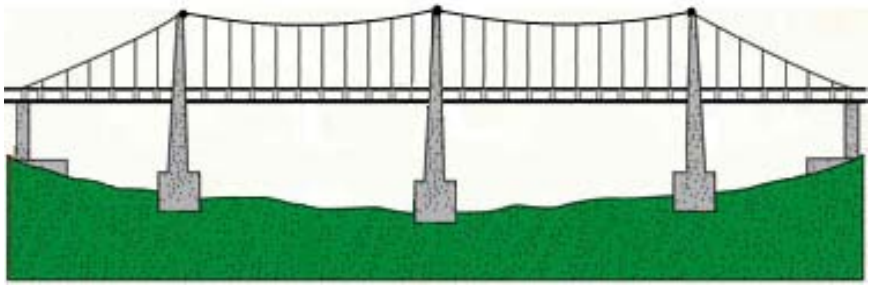
By contrast, downstream destressing, coming as it does after the primary stress, is not capable of serving as an alert. But as for carrying specific information about the relative importance of message content, downstream destressing is unsurpassed in this role. In no uncertain terms, its quiet, backgrounded delivery tells the listener: *We're not central. Focus elsewhere* (Bing, 1980, pp. 173-208; Bolinger, 1986, pp. 110-135).

Downstream destressing is clearly important to learners because of its message-bearing capabilities. But what about upstream destressing? Is there value in its alerting function and its role of making the primary stress stand out?

Let's put this question in pedagogical perspective. Pronunciation teachers know that it takes considerable effort to get students to make a remarkable primary stress. Our typical strategy is to

focus on the nuclear vowel. We emphasize a significant change in its duration, a marked change in its loudness, and a sudden change in its pitch. The approach is all about the nuclear vowel.

The problem with this approach is that students often don't make the primary stress stand out any more than other rhythmic peaks in the phrase, so there is nothing distinctive about any of the 'towers of the suspension bridge'. Or they tend to elevate the prominence of all the peaks, like making all the towers taller.



But what if one of the important ways to make a primary stress noticeable is to increase the contrast between the valley right before this major peak and the peak itself? That is, instead of focusing only on the character of the primary stressed vowel, we might help learners even more if we also show them how to demote the peak just before the primary stress. It may take such a double-barrel attack to help learners make the nucleus as prominent as native speakers do.

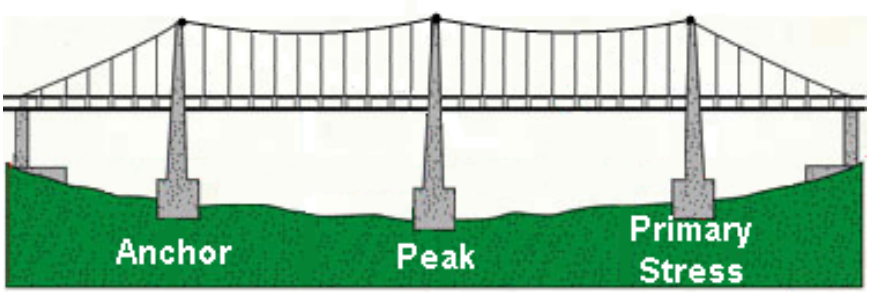
Furthermore, by giving attention to upstream destressing, students will not only foreground the primary stress and alert the listener to its imminent appearance, but their destressing will also deliberately remove potential competitors for attention—other peaks that could misdirect a listener's focus away from the central message.

Is there any evidence that the speech of learners is more intelligible and less distracting to listeners when learners demote peaks before the primary stress than when they don't? This is an important research question that awaits an answer. Anecdotally, however, when students in my pronunciation class do demote peaks as described, their speech sounds surprisingly like the English we are targeting.

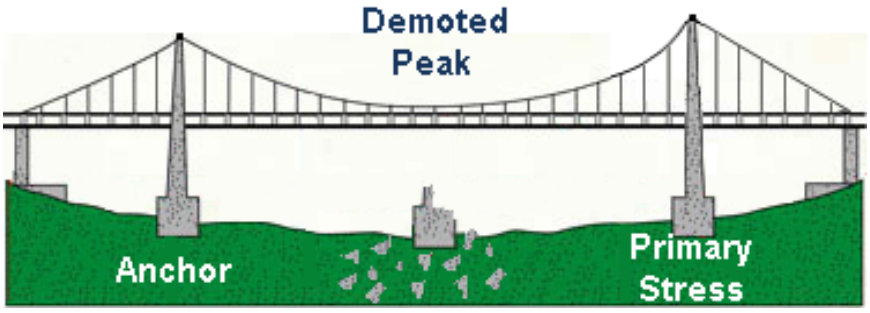
Until we have more definitive evidence, let us assume that there could be pedagogical value in teaching upstream destressing. How might this notion be introduced, predicted, pronounced, and practiced when the audience consists largely of linguistically unsophisticated learners of

English? As we shall see, the challenge is not so great as it might seem.

Introducing demoted peaks. Adult learners want to understand what we are trying to teach them and why it is important. To convey the concept of destressing, we use the graphic of the suspension bridge and identify the critical requirements for destressing: the anchor, the primary stress, and the intervening peak.

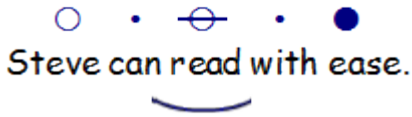


The objective of our practice is to make the primary stress stand out so that it communicates clearly to the listener. We do that by creating an extreme contrast between the valley before the primary stress and the primary stress itself. So we attack two vowels—the vowel of the peak before the primary stress and the vowel of the primary stress. The first, we make less prominent; the second, we make more prominent.



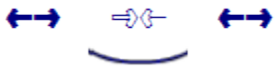
Predicting demoted peaks. To empower adult learners, we also show them how to predict upstream destressing on their own. We already help them predict rhythm by using big circles for peaks and dots for valleys, and we help them predict the location of the primary stress which we mark with a filled-in circle.

Instead of stopping here, we ask students to take another pass through the sentence to see if any upstream destressing is present. Using short sentences like the one above, we have students practice identifying the critical elements: Where is the primary stress? (*ease*) Is there an anchor? If so (*Steve*), is there a peak between the anchor and the primary stress? If so (*read*), what should we do with it? It gets a valley mark ‿. Learners become very good at determining where they should destress a peak.



Pronouncing demoted peaks. Once students find a peak to demote, how should they pronounce it? How do they make the middle tower crumble? The answer, of course, is to do the opposite of what we do with the primary stress. Using phrase sets, all with the same rhythm, we focus on each of the three dimensions of stress. (If we change the rhythm pattern with every practice item, learners will not get the feel for what is happening so easily.)

First, we focus on vowel duration. Judy Gilbert (2005, p. 18ff) hit on the technique of using rubber bands between the thumbs to have students feel the duration dimension. This kind of proprioceptive tool helps learners connect the stretching of the band with the lengthening of peaks and the relaxing of the band with the compression of valleys.



Betsy stayed at home

Try to come at five

Second, we attend to the intensity dimension. The technique for this dimension is more challenging to do, but it directly addresses the energy contrast. We simply whisper valley syllables. We start slowly and build up.

Lunch is served at noon

Bill's good at math

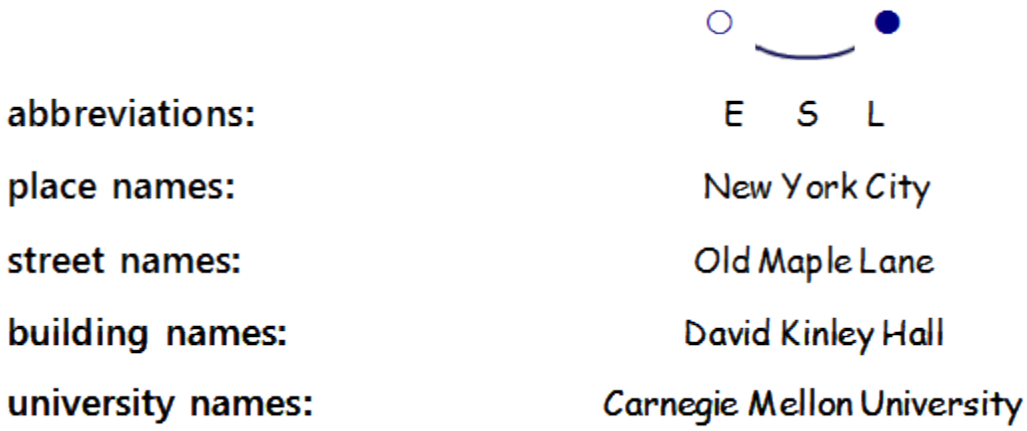
Third, the technique for the pitch dimension is to hum the word pitch, using a high pitch on the peaks and a low pitch on the valleys.

Rice is good with beans

Why didn't he come?

Practicing demoted peaks. Once students can predict where destressing should occur and what it should sound like, the principal challenge becomes giving them enough practice that demoted peaks begin to sound right to them. If Wells (2006, 229) is right, and upstream destressing is pervasive, then it takes no special materials to practice it on an on-going basis. Whenever we

give attention to the primary stress of phrases and sentences, we use the occasion to ask students to check for demoted peaks. For example, when we work on compound nouns that have phrasal stress, we routinely ask: Are there any demoted peaks? Since they can spot the anchor and predict the heavy stress of the construction, it is easy for them to locate an intervening peak to demote. Furthermore, when we give them opportunities to record sentences and phrases with demoted peaks, their practice begins to sound much more authentic than when they do not take this simple, extra step.



In summary, upstream destressing is a widely attested phenomenon that English speakers appear to use to highlight an upcoming primary stress. In ESL circles, however, the notion clashes with what is commonly taught, namely, that content words or loud function words create peaks right up to the primary stress. The fact is, they don't. And neither do ESL teachers I have listened to, including myself! Our natural inclination is to do in ESL classes what we do in ordinary speech. We use upstream destressing without noticing it. If our inclination is to sound natural, then why not use that inclination and explicitly help learners to follow it in the direction of more authentic discourse pronunciation?

ABOUT THE AUTHOR

Wayne Dickerson is Professor of Linguistics at the University of Illinois at Urbana-Champaign where he teaches pedagogical phonology in the MATESL program and ESL pronunciation for international students. He is certified as a Master Online Teacher and teaches the pedagogical phonology course online, as well. An applied phonologist, his principal interest is in the application of technical linguistic insights to practical teaching tasks, with an emphasis on English stress, rhythm, orthography, and materials development. His most recent books are *Stress in the Speech Stream: The Rhythm of Spoken English* (1989), published by University of Illinois Press, and (with co-author Laura D. Hahn) *Speechcraft: Discourse Pronunciation for Advanced Learners* (1999) and two workbooks, published by The University of Michigan Press.

REFERENCES

Beisbier, B. (1995). *Sounds Great! Intermediate pronunciation and speaking for learners of English, Book 2.* Boston, MA: Heinle & Heinle.

Bing, J. (1980). *Aspects of English prosody.* Bloomington, IN: Indiana University Linguistics Club.

Bolinger, D. (1986). *Intonation and its parts.* Stanford, CA: Stanford University Press.

Hahn, L. & Dickerson, W. (1999). *Speechcraft: Discourse pronunciation for advanced learners.* Ann Arbor, MI: The University of Michigan Press.

Gilbert, J. (2005). *Clear speech: Pronunciation and listening comprehension in North American English.* (3rd ed.) New York: Cambridge University Press.

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

Well, J. (2006). *English intonation: An introduction*. Cambridge, England: Cambridge University Press.

Wong, R. (1987). *Teaching pronunciation: Focus on English rhythm and intonation.* Englewood Cliffs, NJ: Prentice Hall Regents.