



Teaching Pronunciation: Simplicity Is the Key

by [Judy B. Gilbert](#)

When teaching pronunciation, simplicity is best because if a lesson is too complicated, students get discouraged—and a discouraged student is harder to teach. For instance, when explaining the difference between the final sounds of *had* and *have*, it is going the hard way to use the terms *plosive* and *fricative*. It's much easier to explain with everyday terms like *stop sound* and *continuing sound*, because the basic words are more apt to be familiar and in fact have been used by distinguished linguists at least as far back as 1967 (Abercrombie). It's especially easy if you can reinforce the ideas with hand gestures, for instance a hand moving in front of you (for continuing sound) and a hand held up flat (meaning “stop”). Little signs, too, can serve as icons representing the difference (see Figure 1).



Figure 1. Icons representing final end sounds.

The Music Signals That Govern English Pronunciation

The most efficient way to lighten the general work load is to present a simple overall picture of the prosodic system of spoken English. More than most languages, English depends on prosodic (musical) signals to call the listener's attention to the most important words, and also to the separation of thought groups. All languages must have a way to highlight the most important word in a sentence, but different languages do this differently, and if the speaker's native-language system is transferred to English, it seriously affects intelligibility. Here are a few examples of ways to express contrastive emphasis; capital letters represent the important/most stressed syllable in the English version.

English/Japanese (post particle *-ga*)

THIS is my bag. / *Kore-ga watashino kaban-desu.*

English/German (use of *doch*)

You DID forget it! / *Jetzt hast du es doch vergessen!*

Because the signaling system overpowers everything else in spoken English (and can change individual sounds beyond the way they are presented in a dictionary), you can simplify your

goals by beginning with the prosody system. Many people assume that rhythm and melody (suprasegmentals/prosody) are ornaments on top of the sounds, and therefore of lesser importance. This is backward. If the learner tries to make an English sound correctly, but speaks with the rhythm of the first instead of the second language, he or she will have a hard time getting the timing right. In English, the timing of a sound is at the heart of its clarity. That is why the prosody system must come first: Time and energy spent working on individual English sounds without attention to correct rhythm is simply inefficient.

The Prosodic System

The rhythm of English is controlled by the prosodic system. Figure 2 shows a simple pyramid graphic visually explaining the hierarchy of suprasegmental (prosodic) signals.



Figure 2. Prosody pyramid based on the work of Bolinger (1989), Cutler (2015), Derwing and Munro (2015), and Gilbert (2012).

The foundation of the pyramid system is a thought group (a short sentence or clause): How do you spell “easy”?

1. A thought group has one focus word: *easy*
2. A focus word has one most stressed syllable: *ea*
3. The vowel in this syllable /iː/ is the peak of information for the whole thought group (short sentence, clause, or phrase)

The system requires the peak vowel to be strongly highlighted, by being extra clear, extra long (rhythm), and with a noticeable pitch change (melody). Contrast must be provided by making the neighboring vowel extra short and hard to notice. The technical term for this sound is *schwa*:

/ə/

Schwa is a sort of “ghost vowel.” This is a key stumbling block for learners because it doesn’t appear in the written language, and also because it is hard to notice. But schwa is the most

common vowel sound in spoken English. This is why English rhythm is irregular; the peak vowel is extra long and most of the other vowels are extra short.

If you help your students absorb a threshold command of the prosody pyramid system (including the contrastive function of schwa), they will have a strong base to learn other aspects of English pronunciation (see Figure 3).




Stressed vowels Long, clear	Unstressed vowels Short, clear	Reduced vowels Very short and unclear
 a e i o u	 a e i o u	 ə ə ə ə ə
Very easy to hear	Easy to hear	Hard to hear

Figure 3. The English system of contrastive clarity.

English calls attention to the focus words by three basic signals: extra clarity, extra length, and a marked pitch change. Focus words convey new information (what hasn't already been talked about). See Figure 4 for an example of the focus changing as a conversation continues. It is crucial that learners are able to hear/process those signals.

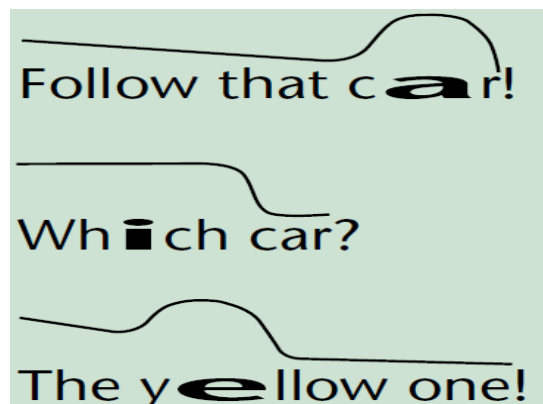


Figure 4. Example of conversation with focus changing.

Many years ago at a TESOL convention, I was impressed with Earl Stevick's remark that "People are often better at pronouncing a foreign language when they are making a joke about those other people." And I've often noticed how people casually express dislike for a regional accent other than their own, such as one from Chicago, or Brooklyn, or Texas, or make such

accents a source of comedy. I concluded that Stevick meant that people who are making jokes are *allowing* themselves to take on a foreign persona.

There are various reasons why students may be uneasy about sounding “foreign” to themselves, and this makes it hard for them to try on the new way of speaking. Our job is to help them allow themselves to role-play a new self.

To help overcome reluctance, Grant (2010) suggested:

Changing your pronunciation, especially stress and rhythm, involves changes in breathing, facial expression, and movement. As a result, when you speak English, you might feel less...like yourself....View English pronunciation like a jacket that you can put on and take off, depending on who your listener is. (p. 93)

I like to think of this approach as *accent addition*, which is a lot more psychologically positive than *accent reduction*.

NOTE: All of the images in this article are used with permission and come from Gilbert, J. (2012). *Clear speech: Pronunciation and listening comprehension in North American English* (4th ed.). New York, NY: Cambridge University Press.

References

- Abercrombie, D. (1967). *Elements of general phonetics*. Edinburgh, United Kingdom: Edinburgh University Press.
- Bolinger, D. (1989). *Intonation and its uses: Melody in grammar and discourse*. Palo Alto, CA: Stanford University Press.
- Cutler, A. (2015) *Native listening: Language experience and the recognition of spoken words*. Cambridge, MA: MIT Press.
- Derwing, T. M., & Munro, M. J. (2015). *Pronunciation fundamentals: Evidence-based perspectives for L2 teaching and research*. Philadelphia, PA: John Benjamins.
- Gilbert, J. (2012). *Clear speech: Pronunciation and listening comprehension in North American English* (4th ed.). New York, NY: Cambridge University Press.
- Grant, L. (2010). *Well said* (3rd ed.). Boston, MA: Cengage Learning.

Judy B. Gilbert has an MA in linguistics from the University of California at Davis, with special study in acoustic phonetics at the University of California, Berkeley. She is the author of *Clear Speech From the Start (2nd edition)* and *Clear Speech (4th edition)*, both from Cambridge University Press, 2012.