PASAA

Pronunciation: The Big Picture

Donna M. Brinton

Author/Educational Consultant, Los Angeles, CA, USA

Corresponding author: brinton@humnet.ucla.edu

Article information	
Abstract	Numerous factors come into play when teaching pronunciation to second language (L2) learners of English. First, the learners may share the same first language (L1) or come from multiple L1 backgrounds. The process is further complicated by the various pronunciation issues needing to be covered, namely the segmental (vowel and consonant) and suprasegmental (stress, rhythm, and intonation) features of English. Finally, the task is colored by whether learners come from a language family more closely related to English (i.e., Indo-European languages) or from unrelated language families (i.e., non-Indo-European). This article focuses on the big picture of pronunciation teaching, examining those factors impacting students' acquisition of English pronunciation. To illustrate, two target populations—students from an Indo-European L1 background (German) and students from a non-Indo-European background (Thai)—are examined to showcase how their differing language backgrounds can impact L2 pronunciation acquisition. The analysis also highlights issues that arise depending on whether the learners' L1 is tonal or intonational and how this may facilitate or impede learners' acquisition of English intonation and sentence stress. Further, it examines how learners' L1 syllable structure may determine command of word stress in

	English. Finally, it examines the predictive power of the vowel and consonant inventories of these two languages compared to those of English and how L1 consonant clustering configurations may differ from those of the target language. The article concludes with suggested teaching strategies addressing both segmental and suprasegmental issues that teachers may encounter when teaching English pronunciation to their L2 learners.
Keywords	language family, grapheme, phoneme, segmental, suprasegmental, rhythm, stress, intonation
APA citation:	Brinton, D. M. (2025). Pronunciation: The big picture. *PASAA Journal, 71*, 1–21.

1. Introduction

As teachers of English learners, we aim to provide effective pronunciation instruction. Numerous factors come into play, as we may teach classes where our learners share the same first language (L1) or, alternatively, where they come from multiple L1 backgrounds. The process is further complicated by the territory that needs to be covered, as we must deal with both segmental (vowel and consonant) and suprasegmental (stress, rhythm, and intonation) pronunciation features. Finally, the task is colored by whether our learners come from a language family that is more closely related to English (e.g., the Indo-European family of languages such as the Germanic branch) or whether they come from language families that are not related (e.g., non-Indo-European languages).

In this article, we will focus on the big picture of pronunciation teaching and examine how the above-mentioned factors may impact learners' acquisition of English pronunciation. By way of illustration, we will look at two target populations—learners from an Indo-European language background (German) and those from a non-Indo-European background (Thai)—to understand how differing L1 linguistic characteristics can impact the teaching/learning process. In our

analysis, we will focus on how the tonal or intonational nature of the learners' L1 may impact their acquisition of English intonation and sentence stress. Further, we will examine whether words in the learners' L1 tend to be mono- or multisyllabic and how this may determine their command of word stress in English. Finally, we will look at the predictive power of the vowel and consonant inventories of these two languages compared to those of English and how consonant clustering configurations (the occurrence in a word of two or more consonants without an intervening vowel) may differ. We will conclude our examination of the big picture of pronunciation teaching with some suggested teaching strategies addressing both segmental and suprasegmental issues that teachers may encounter when teaching English pronunciation to their second language (L2) learners.

2. Literature Review

Before embarking on a more in-depth investigation of the issues, it is important to consider the following preliminary questions:

1.   What is pronunciation?
2.   When should it be taught?
3.   Who should teach it?
4.   Why is it important to teach?

Dictionary.com (n.d.) defines pronunciation as follows: (1) the act or manner of pronouncing *syllables*, *words*, and *phrases* with regard to the production of *sounds* and the placing of *stress*, *intonation*, etc.; and (2) an *accepted standard* of the *sound* and *stress patterns* of a syllable, word, etc. (italics added). As for when it should be taught, the consensus today is that it is a critical skill to teach learners at all stages of the learning process, including beginners of any age (Moreno Álvarez, 2023). However, an important caveat here is that teaching and learning strategies are highly dependent on the teaching context (Biazon Rocha, 2021) e.g., English as a second vs. foreign language, young vs. adult learners and thus teaching strategies may vary significantly depending on the learner population. The

current view as to who should teach pronunciation diverges sharply from the native speaker model espoused in earlier eras of pronunciation teaching. As noted by Brinton (2022, Deconstructing the Myths section, para. 2):

> …today's practice no longer dictates the need for a native speaker teacher to emulate, but rather for the teacher to be one that is highly intelligible, encouraging, and empathetic… [M]ost pronunciation experts today concur that the best pronunciation teachers (regardless of native or non-native speaker status) are those who have appropriate training, dedication, and are inspirational models for their students.

Finally, regarding the importance (or lack thereof) of addressing pronunciation in our curricula, Moreno Álvarez (2023) notes that not only is pronunciation essential for overall communication; it is also one of the first things that an interlocutor notices when meeting an L2 speaker of English and the result of poor pronunciation on the part of this speaker can lead to frustration and even total communication breakdown.

2.1 Key Elements

When we are speaking a second language and we are not understood, we are often at a loss to explain what is at issue. Consider a scene in a restaurant where communication between the customer and the waiter has broken down due to pronunciation issues. For example, when ordering, the customer might have said *I'd like to order the fresh fitch*, i.e.. substituting the wrong consonant /ʧ/ instead of /ʃ/at the end of the word *fish*. Or alternatively, he may have used the wrong stress or intonation pattern when the waiter arrived to take his order: *I DON'T want to order yet,* thus sounding perturbed or angry.

Considering the task of teaching English pronunciation to L2 learners, we need to attend to both the segmental (vowel and consonant) features as well as the suprasegmental (rhythm, stress, and intonation) features of the language (see

Figure 1). All of these features comprise the "big picture" that teachers need to attend.

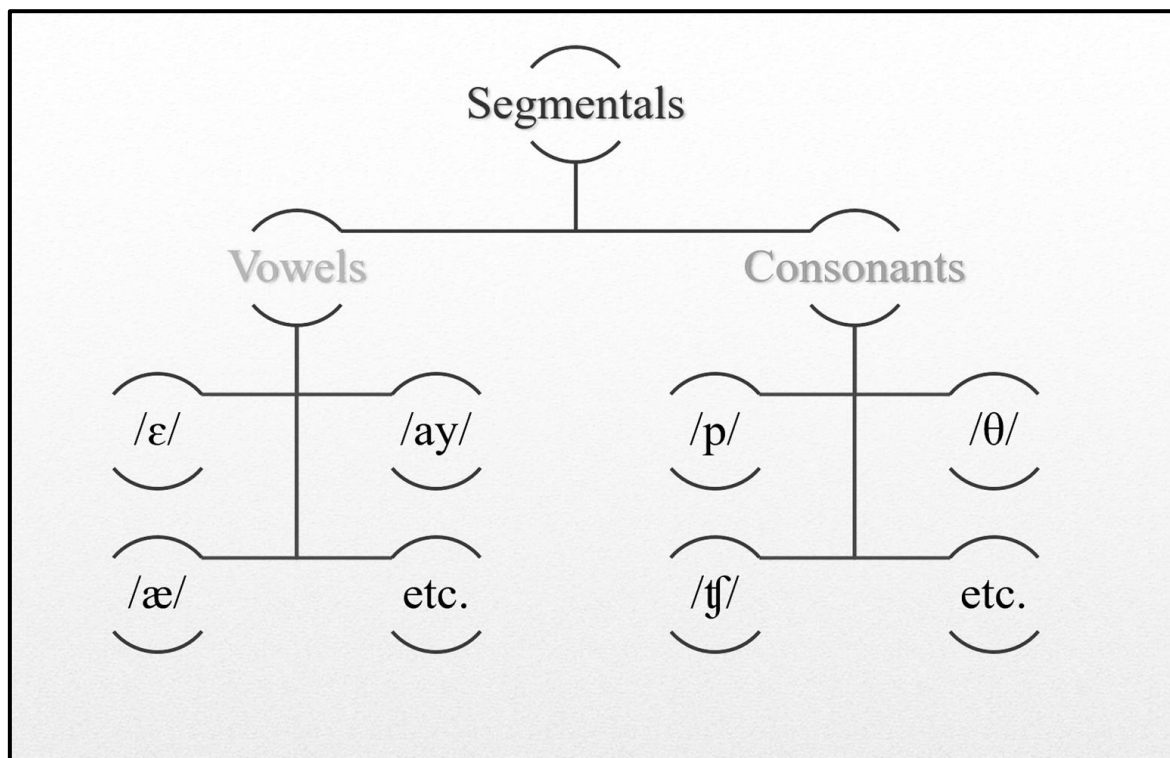**Figure 1**

The Sound Features of English
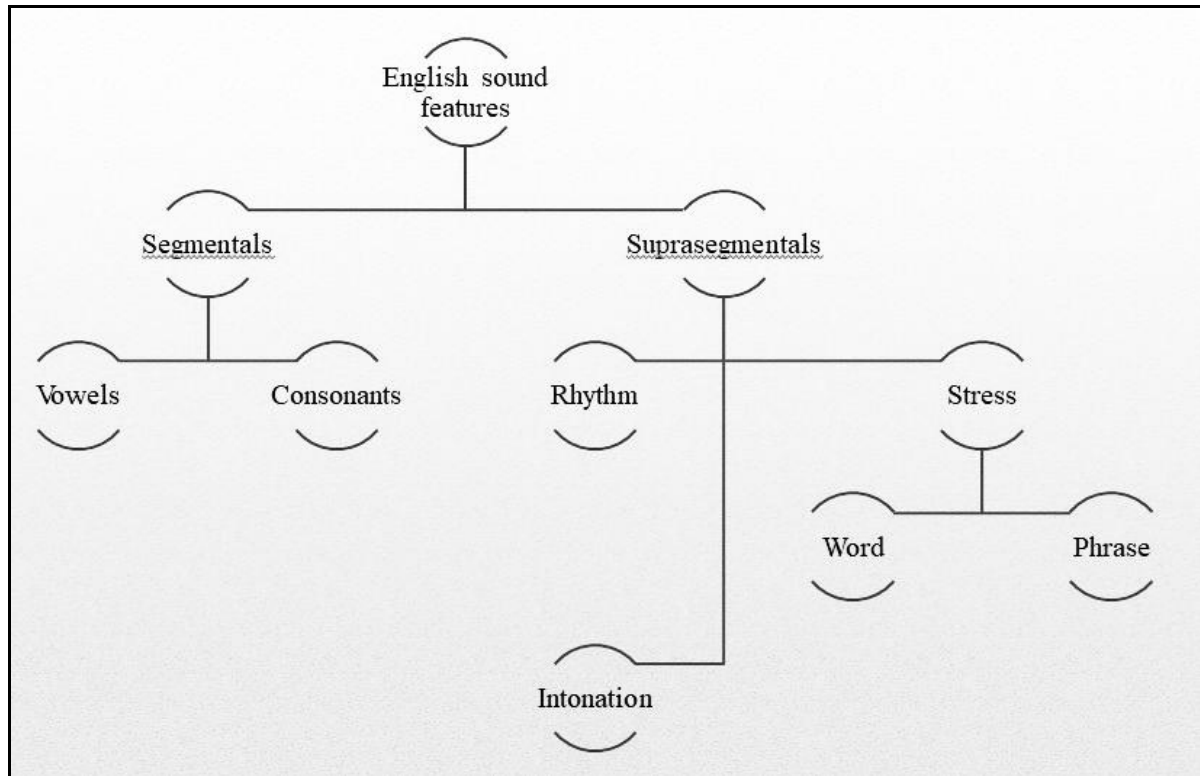
2.2 English Segmentals

The English segmentals are divided into two categories: vowels and consonants (see Figure 2). Vowels are typically classified as high, mid, or low and front, central, or back depending on where they are produced in the oral cavity; they are also characterized as tense vs. lax, with the former being produced with more muscle tension. All vowels are continuants, i.e., produced with continual vibration of the vocal cords, and are produced with either lip rounding or spreading (Celce-Murcia et al., 2010). Thus /iy/ is characterized as a high front tense vowel produced with extreme lip spreading, whereas /uw/ is defined as a high back tense vowel produced with extreme lip rounding. Finally, certain tense vowels of English belong to the separate category of diphthongs due to the gliding quality that accompanies their articulation.

As for consonants, they are typically characterized by voicing (whether the vocal cords vibrate or not during articulation), place of articulation (where the sound is made), and manner of articulation (whether the airflow is restricted or not). Thus, for example, /d/ is described as a voiced alveopalatal stop while /f/ is described as a voiceless bilabial fricative.

Figure 2

Segmentals

2.3 English Suprasegmentals

Transcending the segmental aspects of English are its suprasegmental features, which include stress, rhythm, and intonation (see Figure 3). Both word and sentence stress (commonly also referred to as prominence) are important features of English, as it is a stress-timed (as opposed to a syllable-timed) language, with alternating stressed and unstressed syllables. Typically, stress involves the speaker placing heavier emphasis and longer duration on a given syllable or phrasal element; it also frequently involves a higher pitch level on the stressed element. Intonation consists of the variation in the pitch pattern of utterances, with four levels of pitch typically distinguished. Depending on the

intonation contour (e.g., rising, falling, rising-falling), which extends over the entire utterance or thought group, differences in meaning can be communicated. Finally, the rhythm of English consists of a combination of unstressed, lightly stressed, and strongly stressed elements, giving English its characteristic rhythmic beat (Celce-Murcia et al., 2010).

Figure 3

Suprasegmentals
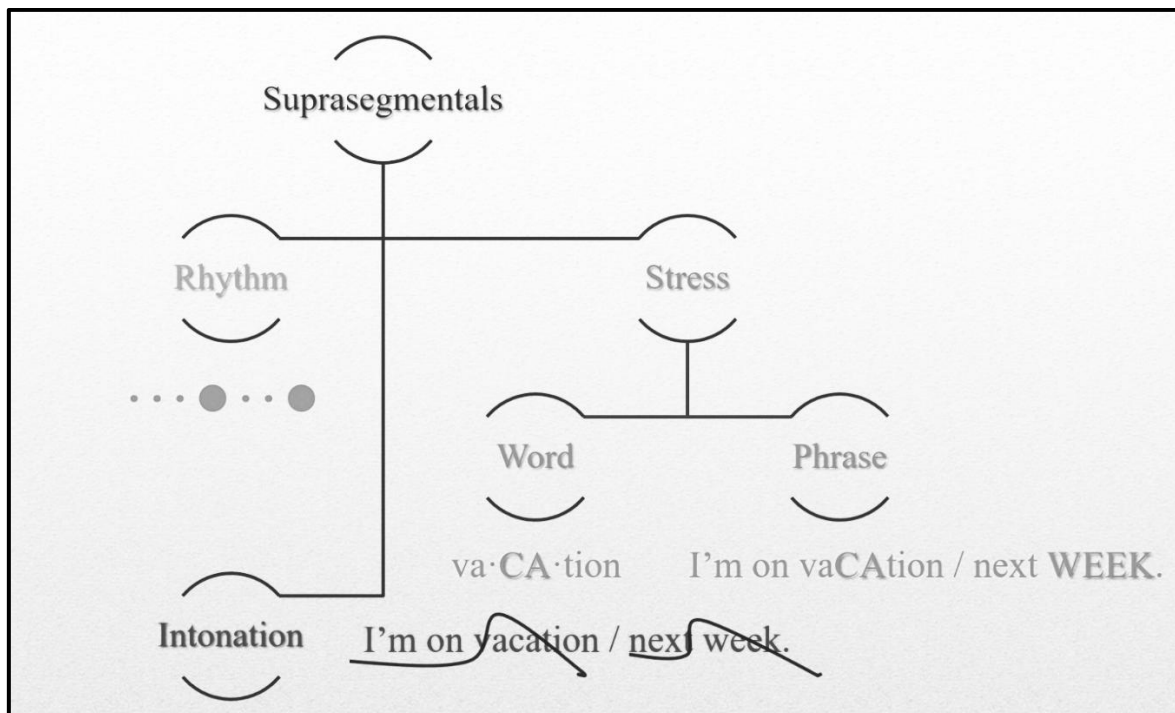
3. Pronunciation Challenges for L2 Learners of English

All languages of the world have their own unique origins and sound systems. This can sometimes lead to negative L1 transfer when learners are attempting to acquire L2 features such as the articulation of the new language's vowels and consonants or the intonation and rhythmic patterns of certain phrases or thought groups (Brinton & Pinweha, in press). In what follows, we will look at several challenges that teachers face when addressing pronunciation issues—most notably phoneme/grapheme correspondence in English, challenges due to L1 vs. L2 language family origins, and the respective vowel and consonant inventories of the learners' L1 and the target language.

3.1 Phoneme/grapheme Correspondence in English

The English spelling system is notoriously complex and can be frustrating for L2 learners attempting to pronounce unknown words (Brinton, 2014). The language consists of 26 written letters (or graphemes): five vowel letters (a, e, i, o, u) and 21 consonant letters (b, c, d, f, g, h, j, k, l, m, n, p, q, r, s, t, v, w, x, y, z). However, the picture becomes more complicated when we look at the pronunciation of these written symbols.

In terms of the vowels, the five written vowel letters a, e, i, o, and u represent 14 or 15 vowel sounds in American English, depending on the dialect in question; in BBC English, they represent 20 distinct vowel sounds. As may be obvious, this lack of a one-to-one correspondence between sound and symbol can negatively impact the teaching process. In terms of the consonants, there are 24 consonant sounds or phonemes, including several (ð, θ , ʃ, ʒ, ʧ, ʤ, ŋ) which do not have a one-to-one equivalent in the written system. Additionally, many of the consonant graphemes may have multiple possible pronunciations, as in the case of the letter c, which can be pronounced /k/ as in *cut*, /s/ as in *nice*, or /ʃ/ as in *ocean* (Wells, 2000), or the letter s, which can have 13 different pronunciations (Brinton, 2014). This same multiplicity of phoneme-to-grapheme correspondence holds true for many of the other English consonant graphemes, making sound/spelling correspondence a significant hurdle for English language learners.

3.2 The Impact of Language Family

A second factor that can influence the L2 learning process are the language families of the learners' L1 and the target language. It is a generally-accepted fact that learning a foreign language from the same language family as one's L1 is easier than the reverse case. This is typically attributed to the similarity of grammar, vocabulary, and syntax in the two languages. However, less commonly noted but also of significance in the acquisition process is that pronunciation features (such as intonation, word stress, and consonant clustering) also tend to be more similar in languages from a given language family. For instance, Dutch

speakers learning German pronunciation may have an easier time than Hindi speakers because both Dutch and German are Germanic languages with similarities in their intonation and word stress patterns as well as in their consonant clustering configurations, sentence prominence, and rhythm. To further elucidate this concept, let's examine the case of German and Thai L1 learners acquiring the suprasegmental features of English:

- German speakers learning English: To begin with, English and German share a number of features. First, both belong to the Germanic family of Indo-European languages. Patterns of word and sentence stress are quite similar in both languages. Additionally, the rhythmic and intonation patterns of the two languages are similar, as are consonant cluster configurations (Swan & Smith, 2001).

- Thai speakers learning English: Thai, on the other hand, shares few of these characteristics with English. A non-Indo-European language from the Tai family of languages, Thai tends to give equal weight and emphasis to each syllable. Thus its rhythmic patterns differ significantly from the stressed-unstressed syllable patterns of English. Intonation patterns in Thai also differ significantly from those of English, with the basic intonation pattern being a universal sharp up and down as opposed to the multitude of English intonation patterns. Additionally, as a monosyllabic language, Thai tends to stress all syllables more or less equally, rendering the pronunciation of multisyllabic words in English (with their characteristic stressed and unstressed syllable patterns) difficult. Finally, in English, consonant clusters occur frequently, both at the beginning and at the end of a syllable, which is not the case with Thai. As a result, Thai speakers may tend to drop consonants in both initial and final English cluster configurations (Swan & Smith, 2001).

3.3 Vowel and Consonant Inventories

A comparison of the consonant and vowel inventories of English and Thai yields the following results (Swan & Smith, 2001):

1. Of the 24 consonant phonemes of English, Thai is missing the following English consonants: /θ/, /ð/, /v/, /z/, /ʃ/, /ʒ/, /dʒ/, and /g/.

2. Similarly, regarding the vowel inventory, Thai is missing all the English diphthongs, namely /eɪ/, /aɪ, /ɔɪ/, and /aʊ/.

Given this situation, one can predict that both in terms of perception and production, Thai learners are likely to encounter difficulty with these English phonemes and may substitute other phonemes in their place, leading to confusion or misunderstanding.

4. Teaching Strategies

In this section we will look briefly at teaching strategies aimed at providing learners with assistance in producing the segmental and suprasegmental features of English. Given the space limitations, the intent is to highlight some of the more common classroom activities; for a more thorough overview readers are recommended to consult practical handbooks such as Celce-Murcia et al. (2010), Jones (2016), or Yoshida (2016).

4.1 Segmentals

In the description and analysis phase of the lesson, attention is typically paid to the articulatory features of the segmental(s) being taught. Here, typical practice when teaching consonant sounds includes the use of a sagittal diagram to indicate the position of the articulators. For vowels, a vowel chart indicating the position where the vowel is articulated within the oral cavity is common along with illustrations of lip position (rounded, spread, etc.); for consonants, a consonant chart indicating the point (e.g., bilabial, labio-dental, alveolar) and manner of articulation (voiced, voiceless) is used. In addition, many teachers use a variety of

"gadgets" or non-verbal tools (Gilbert, 1991) to illustrate articulatory features. These include realia or household objects (such as drinking straws to illustrate lip rounding in vowels, feathers to indicate aspirated vs. non-aspirated production of stop consonants, or rubber bands to illustrate vowel lengthening). Additionally, embodied pronunciation techniques are commonly used (Chan, 2018), such as having learners place both hands flat on their cheeks to feel the difference between tense and lax vowels or placing two fingers on their throat at the base of the neck in proximity to the vocal cords to feel the difference between voiced and voiceless consonants. In the listening and oral practice segments of the lesson, the teacher may introduce role play activities or other means of highlighting the pronunciation feature being taught via the use of songs, games, tongue twisters, and videos.

One of the most time-honored classroom practices for teaching segmentals is the minimal pair drill (see Figure 4). Minimal pair drilling consists of the teacher (or a proficient learner) asking learners to differentiate between two words that differ by only a single sound in the same position. Such drills can consist of either single words (such as *right* and *light*) or entire phrases (such as *It's not right* vs. *It's not light*) and can be used for both listening practice and oral production exercises.

Figure 4

*Minimal Pair Word and Sentence Drills*
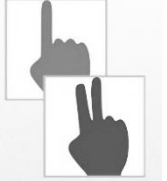
For a more complete inventory of possible segmental teaching techniques in these two phases of the lesson see Figure 5.

Figure 5

Additional Segmental Teaching Techniques

Description & analysis	L & S Practice
▪ Vowel quadrant	▪ Minimal pair words
▪ Lip diagrams	▪ Minimal pair sentences
▪ Consonant chart	▪ Tongue twisters
▪ Sagittal diagrams	▪ Songs & chants
▪ Realia (rubber bands, paper strips, etc.)	▪ Dialogs
	▪ Role play
▪ Use of gesture	▪ Story telling

Depending on the learners' L1, targeted work on consonant clusters may be called for—especially in the case where the learners' L1 has limited consonant clustering configurations. One useful technique is the consonant cluster pyramid, where learners begin by pronouncing a single initial or final consonant and then incrementally add additional consonants to the cluster (see Figure 6), listening first to the teacher model the correct pronunciation and then repeating the word and receiving teacher feedback. This exercise lends itself to pair or small group work as well.

Figure 6

Consonant Cluster Pyramid
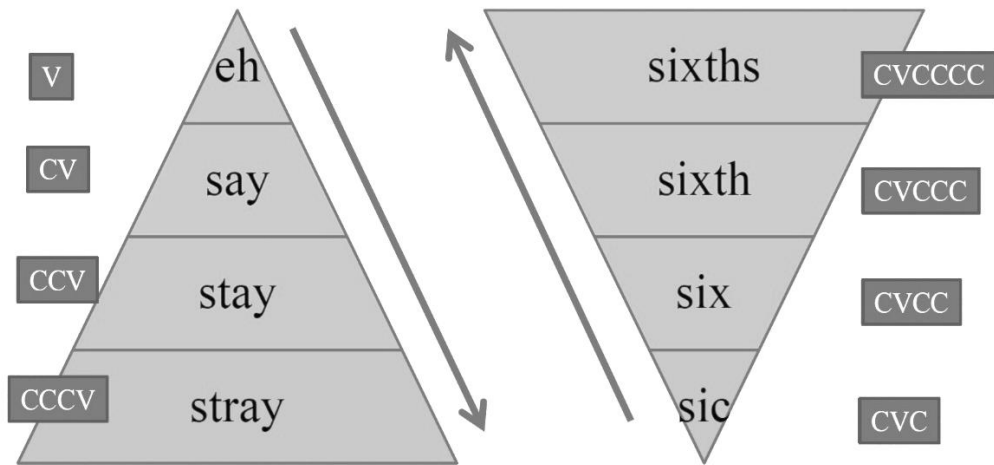
Once learners are more comfortable with pronouncing clusters, the teacher can introduce tongue twisters such as "Friendly Frank flips fine flapjacks" to practice the /fr/ and /fl/ initial clusters or "Brad's big blue bath brush broke" to practice the /br/ and /bl/ initial clusters. Similarly, to practice final clusters learners can be challenged to practice the tongue twister "Amidst the mists and coldest frosts he still insists he sees the ghosts." And for more advanced learners, teachers can easily construct a guided story-telling activity involving past tense clusters (such as "Joe's Visit to the Car Wash Yesterday) by creating a word bank (e.g., "check the tire pressure," vacuum the interior," "clean the windshield," etc.) that groups can use in creating and presenting their stories to the class as a whole (see Figure 7).

Figure 7

Guided Story-Telling Activity

4.2 Suprasegmentals

A wide variety of classroom techniques present themselves for the teaching of the English suprasegmentals, including embodied pronunciation techniques (Chan, 2018) such as using the fingers to indicate stressed syllables in a word, clapping or tapping out the rhythm of a nursery rhyme or chant, etc. (see Figure 8).
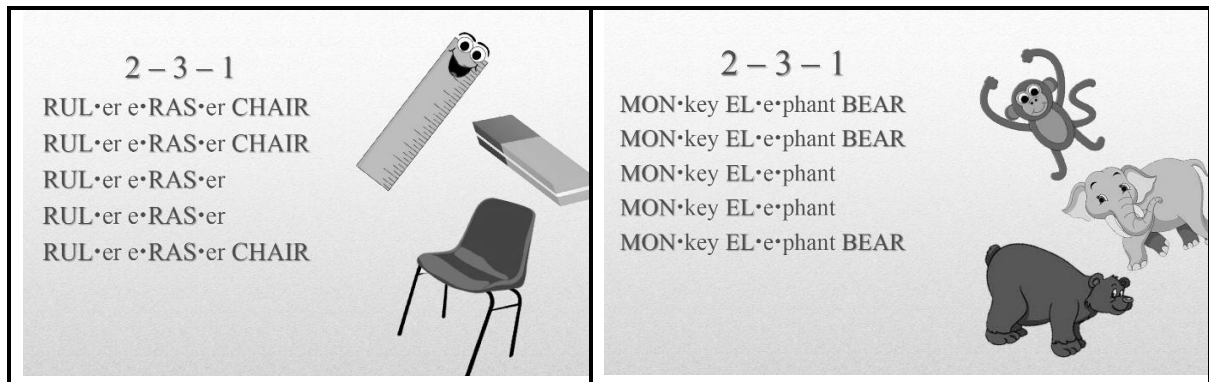
Figure 8

Suprasegmental Techniques
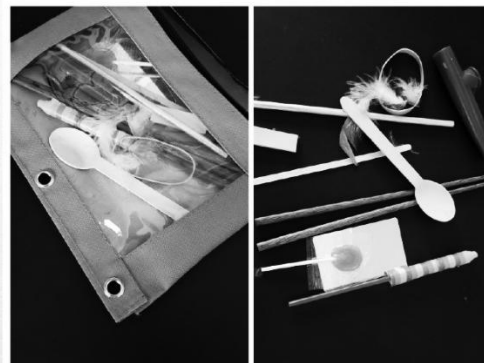
One engaging word stress activity that is geared to younger learners is the 3-2-1 jazz chant (Graham, 2006), where learners practice clapping out the word stress patterns of familiar words containing three, two, and one syllables (see Figure 9). Once familiar with the 3-2-1 rhythmic pattern, learners can be grouped and assigned to write their own 3-2-1 jazz chant on a given topic (e.g., capital cities, food, clothing, and natural catastrophes, etc.) and then perform it for their classmates.

Figure 9

3-2-1 Jazz Chants

Further work on stress can involve matching tasks where learners identify the rhythmic patterns of selected words (see Figure 10) as a way to help them transition to the parallel concept of sentence stress or prominence (see Figure 11).

Figure 10

Analysis of Word Stress

Task: Match the patterns in column 1 with the words in column 2.

Column 1	Column 2
1. ta-ta-TI-ta	a. reality
2. ta-TI-ta-ta	b. mysteriously
3. TI-ta-ta-ta	c. stimulus
4. TI-ta-ta	d. electricity
5. ta-ta-TI-ta-ta	e. arbitration
6. ta-TI-ta-ta-ta	f. permanently

Figure 11

Word vs. Sentence Stress

PATTERN	WORDS	PHRASES
· ●	to·NIGHT	It's LIGHT.
· ●	re·CEIVE	Please LEAVE.
· ●	Chi·NESE	Try CHEESE.
· · ●	vo·lun·TEER	It's a DEER.
· · ●	se·ven·TEEN	She's the QUEEN.
· · ●	un·der·STAND	Where's the BAND?
· ● ·	re·AC·tion	Where's JACK·son?
· ● ·	A·LAS·ka	I'll ASK her.

Following these activities, teachers can present students with congruent pattern drills such as those in Figure 12, which provide useful additional work with sentence stress and lend themselves to having students clap or tap out the rhythm while practicing the sentence patterns.

Figure 12

Congruent Pattern Drills

A final important means of communicating information in English is the use of intonation i.e., marking key elements of a phrase or thought group through the relative high or low nature of the accompanying pitch profile (e.g., rising, rising-falling pitch, etc.). Here, common classroom tasks in the listening phase involve the use of the hands, a rubber band, or a kazoo to emphasize the rising and falling nature of the intonation contours along with listening discrimination exercises that assist learners in distinguishing how differing intonation contours affect the speakers' meaning (see Figure 13).

Figure 13

Focused Listening Exercises for Intonation

In the production phase, other typical classroom practices involve having learners engage in focused reading practice, moving their hands in sync with the

rise and fall of a phrase's intonation, reciting rhythmic chants, singing songs, and watching videos.

5. Discussion

From the foregoing, one might conclude that speakers from language backgrounds that differ significantly from that of English are at a distinct disadvantage acquiring English pronunciation. This point of view, however, greatly exaggerates the difficulty faced by learners from such language backgrounds and ignores the many other factors (such as personal motivation, age of the learners, and amount of exposure to spoken English) that influence pronunciation acquisition. However, the difference in the consonant and vowel inventories of the learners' L1 and English do warrant serious attention in the instructional process, as do the differences in the suprasegmental characteristics of the two languages. It is therefore of great importance that teachers in training be acquainted with these differences and be provided with insight into practices that they can adopt to help learners overcome pronunciation hurdles in their quest to acquire an intelligible pronunciation of English.

Overall, there is general consensus that pronunciation pedagogy should involve the integration of listening and speaking and involve both perception and production activities. Further, to the extent possible, it should emphasize kinesthetic or whole-body involvement to assist learners in acquiring unfamiliar aspects of the L2 (Brinton & Pinweha, in press). In this regard, it is also of great importance that practices adopted in the classroom be evidence-based—i.e., that classroom pedagogy be informed by findings in the field of pronunciation research (Levis, 2017).

6. Conclusion

This article has presented the big picture of pronunciation teaching, emphasizing that there is much more involved than just the correct articulation of individual vowels and consonants. It has also detailed pronunciation challenges

that teachers and learners may face, such as the lack of a one-to-one sound/symbol correspondence in the English spelling system, the mismatch in L1 vs. L2 vowel and consonant inventories, and the differences in stress, rhythm, and intonation that may exist between the two languages.

Ultimately, pronunciation is inextricably entwined with both listening and speaking; accordingly, these two skills need to play an important role in the overall pronunciation curriculum. Equally important is the use of gestures, body movement, games, songs, chants, and other interactive activities to both engage learners and mirror the rhythmic patterns of the target language. As noted at the outset of this article, it is never too early in the language learning process to begin focusing on pronunciation.

7. About the Author

Donna M. Brinton is an author and educational consultant in the field of Teaching English as a Second or Foreign Language. She has worked as a teacher educator in over 40 countries and is the co-author and co-editor of numerous professional volumes, including *Teaching Pronunciation* (Cambridge University Press).

8. Acknowledgements

9. References

Biazon Rocha, A. P. (2021, August 16). *Understanding some essential terms and definitions in L2 pronunciation teaching and learning*. https://pronsig.iatefl.org/some-essential-terms-and-definitions-l2-pronunciation/

Brinton, D. M. (2014). Is English really crazy? Insights into grapheme/phoneme correspondence. *Speak Out*, *50*, 18–22.

Brinton, D. M. (2022, May). Teaching pronunciation: It doesn't take a native speaker. *The Mind Brain Ed Think Tank*. https://www.mindbrained.org/2022/05/teaching-pronunciation-it-doesnt-take-a-native-speaker/

Brinton, D. M., & Pinweha, S. (in press). Teaching pronunciation to young learners. In C. M. Goh, R. L. Sabnani, & W. Renandya (Eds.), *Developing second language oracy in young learners*. Springer.

Celce-Murcia, M., Brinton, D. M., & Goodwin, J. M. (with Griner, B.). (2010). *Teaching pronunciation: A course book and reference for teachers of English to speakers of other languages* (2nd ed.). Cambridge University Press.

Chan, M. J. (2018). Embodied pronunciation learning: Research and practice. *The CATESOL Journal, 30*(1), 47–68. https://doi.org/10.5070/b5.35964

Dictionary.com. (n.d.). *Dictionary.com*. https://www.dictionary.com/

Gilbert, J. (1991). Gadgets: Non-verbal tools for teaching pronunciation. In A. Brown (Ed.), *Teaching English pronunciation: A book of readings* (pp. 308–322). Routledge.

Graham, C. (2006). *Teaching jazz chants to young learners*. Office of English Language Programs, U.S. Department of State.

Jones, T. (Ed.). (2016). *Pronunciation in the classroom: The overlooked essential*. TESOL Press.

Levis, J. M. (2017). Evidence-based pronunciation teaching: A pedagogy for the future. *Journal of Second Language Pronunciation, 3*(1), 1–8. https://doi.org/10.1075/jslp.3.1.001evi

Moreno Álvarez, M. (2023, January 15). *The importance of teaching pronunciation from the very first lesson*. https://www.linkedin.com/pulse/importance-teaching-pronunciation-from-very-first-moreno-%C3%A1lvarez/

Swan, M., & Smith, B. (Eds.). (2001). *Learner English: A teacher's guide to interference and other problems* (2nd ed.). Cambridge University Press.

Wells, J. C. (2000). *Longman pronunciation dictionary*. Longman.

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