



ANDREA TOTH  
*City College of San Francisco*  
*Laney College, Oakland*

## What Not to Teach When Teaching Pronunciation

■ **The most common type of pronunciation exercises found in current ESL textbooks are those that focus on the voicing variation of past tense endings, /t/ and /d/, and the voicing variation of plural nouns and third-person singular verb endings, /s/ and /z/. While these voicing variations are a reality in English, knowing about them and practicing them do not help ESL students improve their pronunciation. The following article provides both linguistic and pedagogical arguments for excluding such exercises from ESL curricula because such exercises have a tendency to confound students, who often are already overwhelmed with the quantity of information they must master in the target language, and instructors who may have minimal training in the phonology of English and in the teaching of pronunciation.**

During the last 15 years, the integration of the teaching of pronunciation with other skill areas has become evident through the variety of ESL textbooks that have incorporated pronunciation exercises. Today, pronunciation drills and exercises are frequently found in textbooks that are directed toward grammar, listening/speaking, and idioms in addition to books aimed specifically for pronunciation. While it is encouraging to see that there is a greater emphasis on teaching and practicing pronunciation, we need to take a careful look at what is being presented in the textbooks, and consequent-

ly, what is being taught in the classrooms. Because most ESL instructors have minimal, if any, training in phonology and the teaching of pronunciation, they rely heavily on the exercises presented in the textbooks and tend not to evaluate them critically. Murphy (1997) concluded in his survey of Master of Arts in Teaching English to Speakers of Other Languages (MA TESOL) programs that the lack of adequate training in phonology and L2 phonology acquisition undermines the effectiveness of ESL instructors teaching pronunciation. To assure that the pronunciation exercises and drills practiced in classrooms are effective and worthwhile, they must be sound linguistically and pedagogically. It is this author's belief that some of the most prevalent exercises offer no advantage to our students, and may in fact intimidate not only the students who are trying to master a new phonological system, but also the ESL teacher who is trying to steer through unfamiliar phonological material.

Derwing and Rossiter (1992) report that of the ESL students who were able to identify items that were problematic for their pronunciation, the most commonly identified items were segmentals, including *th* and *l/r* and other individual vowel and consonant problems. Interestingly, the most commonly found pronunciation exercises in popular, commercial ESL textbooks are those that focus on the voicing variation of inflectional endings. Based on the frequency with which these exercises appear, there seems to be a determination in the ESL profession to have our students pronounce these inflectional endings as prescribed with little regard to the students' perceptions of what they need to improve, to the functional load, or to the level to which intelligibility is diminished if an item is mispronounced (Brown, as cited in Derwing & Rossiter, 2002). There are numerous arguments for excluding exercises focused on the voicing variations of inflectional endings from pronunciation curricula on both linguistic and pedagogical grounds. See the Appendix for a sample list of ESL textbooks on the market that incorpo-

rate pronunciation exercises involving voicing variations of inflectional endings.

In particular, the exercises focus on the [t] and [d] variation in past tense endings and the [s] and [z] variation in plural noun and third-person singular verb endings. In both cases, the voicing quality of the inflectional ending is determined by the voicing quality of the preceding phoneme. “Voicing” or “voiced” refers to the vibration of the vocal folds in the larynx. The only difference between /t/ and /d/ is the vibration in the vocal folds. For /t/, the vocal folds do not vibrate and it is therefore a voiceless phoneme, while for /d/, the vocal folds do vibrate and it is therefore a voiced phoneme. The position of the lips, tongue, and airflow are the same; only the voicing quality varies for these phonemes. Most consonants have voiced and voiceless counterparts (*p/b, f/v, s/z, t/d, k/g*), and these are often presented in minimal pair drills (Pennington, pp. 22-23).

The voicing variation of the inflectional endings is generally taught in the following way. In the case of past tense endings, *kissed* is pronounced [kɪst], with a /t/, because the past tense *ed* assimilates to the unvoiced /s/ preceding it. However, *bagged* is pronounced [bægd], with a /d/, because the past tense *ed* follows a voiced /g/.

Similarly, the final *s* in plural nouns and in third-person singular verb endings varies in voicing quality depending on the phoneme that precedes it. For instance, *bets* is pronounced [betz] because the plural *s* marker follows a voiceless /t/, while *beds* is pronounced [bedz] because the plural *s* marker follows a /d/, which is a voiced phoneme.

While these voicing assimilations are a reality of the English language, does knowing and practicing these assimilations help our students to become more intelligible? Are we serving our students in helping them to meet their goals by incorporating these types of exercises in the curriculum? The following are the linguistic and pedagogical arguments for excluding such exercises.

## Linguistic Arguments

1. *While it is important for ESL students to master the phonemic variations of English, there is little to gain from mastering the allophonic variations of English.* As stated earlier, the only difference between /t/ and /d/ or /s/ and /z/ is the feature of voicing quality. That is, the only difference between these pairs of sounds is the vibration of the vocal folds. These pairs of sounds can represent either phonemic or allophonic variations. A phonemic variation is one in which the variation produces a distinct word with a distinct meaning. Phonemic variations are frequently incorporated into minimal pair drills. No rules are associated with when the phoneme is voiced or voiceless. The voicing quality is determined strictly by the meaning of the word, as shown in Figure 1.

**Figure 1**  
**Voicing Quality Producing**  
**Phonemic Variation**

<b>(-Voice)</b>	<b>(+Voice)</b>
bat	bad
fat	fad
feet	feed
great	grade
face	phase
ice	eyes
rice	rise
race	raise
cease	seize

On the other hand, an allophonic variation is one in which some feature of a phoneme is influenced by its environment, often by a phoneme that precedes or follows it. The rules determining a particular variation of a phoneme can be complex and have no relation to meaning. For example, in English, a word final /t/ often transforms into a glottal [ʔ] or to a flapped *d* [ɾ] when followed by an inflectional ending, as shown in Figure 2. In this instance, the feature that is altered is the

point of articulation, not the voicing quality (Pennington, 1996, p. 59).

**Figure 2**  
**Allophonic Variation of Single Phoneme**

<i>Final /t/⇒</i>	<i>transforms to /ʔ/</i>
eat	eaten
bit	bitten
write	written
white	whiten
<i>Final /t/⇒</i>	<i>transforms to /r/</i>
hot	hotter
sweet	sweeter
neat	neater
sit	sitting
boat	boating

The [t]/[d] variation of past tense inflectional endings and the [s]/[z] variation of plural and third-person verb markers are allophonic variations. Specific complex rules determine the occurrence of a particular allophone. The variation does not result in a word with a distinct meaning. The amount of vibration of the vocal cords is influenced by the surrounding environment. Allophonic variations occur subconsciously. Native speakers are usually unaware of allophonic transformations but are very much aware of phonemic alterations because of the relationship to the meaning of the word. While learning the phonemic variations of English is fundamental to mastering the phonological system, is there any value for students to learn allophonic variations and assimilations?

2. *Another reason for excluding voicing assimilation rules for inflectional endings is that the rules are oversimplified and often not valid.* As stated, the rule for the past tense marker is that if the *ed* comes after a voiceless phoneme, it is pronounced as [t], but if it comes after a voiced phoneme, it is pronounced as [d]. If it comes after a /t/ or /d/, then it is pronounced as [ɪd]. This is a simple voicing assimilation rule. However, in connected speech, assimilation rules take precedence over the voicing rule. Blending, linking,

and palatalization are examples of such rules. Linking occurs when a regular past tense verb is followed by a word that begins with a vowel. The past tense /d/ is linked to the following word and the voicing assimilation rule does not apply.

In the following examples, the voicing assimilation rule would have the past tense marker pronounced as a [t], but the assimilation rule does not apply because it is replaced by a linking rule. The linking rule has the past tense marker assimilate in voice quality to the following, not the preceding, phoneme.

He parked on the street.	He <i>park don</i> the street.
She walked in the park.	She <i>walk din</i> the park.
Mary mopped up the floor.	Mary <i>mop dup</i> the floor.

In connected speech, sometimes a blending rule applies in place of the voicing assimilation rule. In blending, the past tense marker is blended with the following phoneme and assimilates in both voicing quality and point of articulation. As a result, the past tense is not pronounced as a [t] in the way that the assimilation rule predicts.

Julie parked the car.	Julie <i>parkthe</i> car.
Sam talked to her.	Sam <i>talkto</i> her.
Nancy baked some cookies.	Nancy <i>bakesome</i> cookies.

The voicing assimilation rule is in itself a complex rule and the application of the rule is even more confounded when related to linking and blending rules, as is the case in natural speech. Hewings and Goldstein (1999) remind pronunciation instructors that the variation of *s/z* and *t/d* is complex and should be only secondary to teaching the rhythm and stress patterns of these inflectional markers (pp. 88, 129).

3. *The voicing assimilation rule incorrectly predicts that a past tense /t/ would have the same qualities as a word initial /t/.* In many

Canadian and American dialects, where there is no linking or blending rule to apply, and when the past tense *ed* marker is preceded by a voiceless consonant, the past tense is pronounced as a weak, unaspirated /t/.

The dog barked.  
He drove but she walked.  
Suddenly they kissed.  
Frank was genuinely missed.  
The balloon popped.  
The plate broke because it was dropped.

The preceding sentences would sound very unnatural for most American and Canadian speakers if they were pronounced with the same [t] that is found in words such as *time*, *tooth*, *twist*. However, the voicing assimilation rule does not account for this difference and predicts the past tense marker to be pronounced the same as these word initial phonemes.

4. *Another argument for not teaching the voicing assimilation of inflectional endings is that voicing assimilation often occurs naturally in consonant clusters within word boundaries.* Voiced phonemes can be devoiced when they occur in a cluster of voiceless phonemes. In words such as *spray*, *string*, *street*, *spleen*, *splash*, the /t/ and /l/ may be pronounced as voiceless phonemes because the preceding phonemes are voiceless (Pennington, 1996, p. 57). In the case of a voiced phoneme the tendency is for voicing to carry over to a following voiceless phoneme. *Dogs*, *beds*, and *lads* are pronounced [dɔgz], [bɛdz], and [lædz] respectively because the vibration in the vocal folds from the initial voiced phoneme does not stop soon enough to prevent some vibration in the vocal folds during the onset of the next phoneme. Voicing assimilation in consonant clusters is a natural phenomenon and the past tense *ed* and third-person and plural *s* markers often result in consonant clusters. Voicing assimilation occurs in most languages (Pinker, 1994, p. 178). Teaching it serves little purpose since it is a natural physiological aspect of human sound production and there is no need for it to be transparent to our students.

## Pedagogical Arguments

1. *The rules for the voicing assimilation of inflectional endings are far too complex for someone to realistically master at a conscious level in extemporaneous speech.* The conditions for the assimilation of the past tense marker are different from the conditions for the plural and third-person verb marker. The rules are as shown in Figure 3.

**Figure 3**  
**Voicing Rules for Inflectional Endings**

### ***Regular plural nouns and simple present tense—third-person singular***

/s/ after voiceless sounds: p,t,k,f,th  
/z/ after voiced sounds:  
b,d,g,v,m,n,l,r,ng,th  
/ɪz/ after s,z,sh,ch,ge/dge sounds

### ***Simple past tense of regular verbs***

/t/ after voiceless sounds:  
p,t,k,f,s,sh,ch  
/d/ after voiced final sounds:  
b,g,v,l,r,m,n  
/ɪd/ after verbs ending in *t* or *d*

It is unrealistic to expect students to learn these conditions. In addition, the final result of mastering these conditions is of minor consequence in terms of their ability to communicate in the target language (Hewings & Goldstein, 1999, p. 132).

2. *The voicing assimilation of inflectional endings does not alter meaning in any way, and so the value of learning complex rules and practicing their application is questionable.* As discussed, voicing assimilation is an allophonic variation that does not alter meaning. Even native speakers are not aware of sound variations that are not linked to meaning (Pennington, 1996, p. 24). Learning without meaning is difficult. Students need to be able to associate semantic or syntactic meaning to what they learn. The amount of aspiration does not alter the meaning of a word (Haycraft, 1971, p. 110). Likewise, the amount of vibration of the vocal cords does not alter

the meaning of a word if the vibration is influenced by the surrounding environment. Variations in features do not change meaning and communication is based on the production of meaningful strings of sound.

3. *In the past there were doubts about the effectiveness and value of teaching phonemic minimal pairs (sheep versus cheap), where meaning is crucial. Why haven't there been doubts about the effectiveness and value of teaching allophonic variations (mobbed [mabd] versus mopped [mapt]), where there is no association to meaning?* The debate over the impact of teaching segmentals (phonemes) versus suprasegmentals in pronunciation classes has been longstanding, healthy, and vigorous. Morley (1998), Gilbert (1990), and Wong (1987) are among the contributors who have argued that pronunciation training must involve more than contrasting sounds and focusing on point and manner of articulation. The prosodic features, including rhythm, stress, and intonation, are critical components to intelligibility. The outcome of this debate has been a recognition that both are important in second language fluency (Celce-Murcia, Brinton, & Goodwin, 1996). Exercises that focus on voicing variations of inflectional endings are more confusing and more ineffective than the dreaded minimal pair drills ever were. Furthermore, they counteract the efforts to focus more on the prosody of English.

4. *By exposing students to complex phonological rules that do not convey meaning, we are more apt to intimidate and discourage them from articulating word final sounds than we are from building their confidence and encouraging them to attempt these sounds.* Many language groups, primarily Asian, drop or substitute glottal stops for word final consonants and consonant clusters (Gilbert, 2001, p. 176). It is a challenge for students from these language groups to master the pronunciation of word final consonants, and drills that focus attention on them are crucial so that students realize that word final sounds convey meaning in English. For example, when a student says [hɑ] or [hɑʔ], it could

mean *hard, heart, hot, hall, hog, hop*, etc. The meaning can sometimes be determined from context, but the student must come to the realization that word final sounds in English are an important part of intelligibility. We must encourage students to articulate word final consonants and consonant clusters, even if they are approximations of the sounds that a native speaker would make, by simplifying the task rather than confounding the matter with complex phonological rules.

5. *There are too many peculiarities and irregularities in the English spelling and phonological systems to address all of them.* Because the spelling system of English is very irregular, it is that much more difficult for nonnative speakers to master the phonological system. It is useful to point out some basic spelling rules to help with both spelling and pronunciation, but many of the irregularities are acquired rather than learned. For example, even native speakers are usually unaware that in words such as *is, was, does, these, those, whose, his, as*, the *s* is pronounced as [z]. Is there any value to discriminating between the *s* in *this*, pronounced with [s], and *these*, pronounced with [z]? We must decide which irregularities and peculiarities deserve attention. Additionally, our students, just as native speakers, simply acquire some of them. The voicing assimilation of inflectional endings falls in this category. We do a greater service for our students by ignoring it.

## Conclusion

Clearly, inflectional markers serve an important function in the English language. A more meaningful and effective approach to teaching the past tense marker and the plural noun and third-person singular marker is to focus on whether or not the inflectional ending is pronounced as an additional syllable (Hewings & Goldstein, 1999, p. 132). This approach involves few and simple rules and is consistent with reinforcing the rhythm and stress pattern of the language.

Pronunciation is perhaps the greatest challenge for the second language learner to mas-

ter, and focus on pronunciation should be a fundamental part of the second language curriculum. Through the years, pronunciation drills and exercises have matured and become more interactive and student-centered. They have also found their way into mainstream textbooks, so that they can be practiced often, recycled, and appreciated as a critical part of the communication process. However, it is important that we take a long and hard look at what we are teaching and what the value and outcome are for our students. There seems to have been a “jump on the bandwagon” approach to the inclusion of voicing variation exercises in textbooks and syllabi without any serious consideration as to the practicality of teaching, learning, or applying the concept.

The most popular ESL textbooks all incorporate drills and reference materials that focus on the voicing assimilation of inflectional endings. For both linguistic and pedagogical reasons, these drills should be excluded from classroom instruction and practice. From a linguistic aspect, they represent oversimplified phonological rules that convey no meaning and often are not valid in natural, fluid speech. From a pedagogical aspect, the rules intimidate and distract rather than encourage students to focus on meaningful phonological units. Writers, editors, and publishers must take responsibility for providing teachers with materials that are clear and useful in order to meet the goals of our students and exclude the superfluous and redundant materials. Teacher-trainer programs must provide more in-depth training in phonology so that ESL instructors have the confidence to critique the materials they use in the classrooms and to recognize the needs of their students and address them in effective ways.

### Author

*Andrea Toth has taught ESL in Japan, Canada, and the United States. She teaches at Laney College in Oakland and City College of San Francisco.*

### References

- Celce-Murcia, M., Brinton, D., & Goodwin, J. (1996). *Teaching pronunciation: A reference for teachers of English to speakers of other languages*. Cambridge, England: Cambridge University Press.
- Derwing, T. M., & Rossiter, M. (2002). ESL learners' perceptions of their pronunciation needs and strategies. *System*, 30, 155-156.
- Gilbert, J. (1990). Pronunciation: What should we be teaching? (ERIC Document Reproduction Service No. ED320443)
- Gilbert, J. (1996). Six pronunciation priorities for the beginning student. *The CATESOL Journal*, 13(1), 173-181.
- Haycraft, B. (1971). *The teaching of pronunciation*. Hong Kong: Longman.
- Hewings, M., & Goldstein, S. (1999). *Pronunciation plus: Practice through interaction teacher's manual*: Cambridge, England: Cambridge University Press.
- Morley, J. (Ed.). (1994). *Pronunciation pedagogy and theory: New views, new directions*. TESOL.
- Murphy, J. M. (1997). Phonology courses offered by MATESOL programs in the U.S. *TESOL Quarterly*, 31, 741-764.
- Pennington, M. C. (1996). *Phonology in English language teaching*. New York: Addison Wesley Longman.
- Pinker, S. (1994). *The language instinct*. New York: Harper Perennial.
- Wong, R. (1987). *Teaching pronunciation: Focus on English rhythm and intonation*. Englewood Cliffs, NJ: Prentice Hall.

**Appendix**  
**ESL Textbooks With Pronunciation**  
**Exercises That Involve Voicing**  
**Variations of Inflectional Endings**

**Grammar Textbooks**

- Azar, S. B. (1995). *Basic English grammar* (2nd ed.). Prentice Hall Regents, 49-50, 137, 180-182.
- Azar, S. B. (2002). *Understanding and using English grammar* (2nd ed.). Prentice Hall Regents, 20-21, 84-85.
- Azar, S. B. (2003). *Fundamentals of English grammar*. Longman, 28-29, 157-158.
- Elbaum, S. (2001). *Grammar in context 2* (3rd ed.). Heinle & Heinle, 451-453.
- Elbaum, S. (2000). *Grammar in context 3* (3rd ed.). Heinle & Heinle, 440-442, 445.
- Fuchs, M., Bonner, M., & Westheimer, M. (2000). *Focus on grammar: An intermediate course for reference and practice* (2nd ed.). Longman, A10.

**Listening/Speaking Textbooks**

- Benz, C., & Dworak, K. (2000). *Tapestry: Listening and speaking 1*. Heinle & Heinle, 189-190, 226.
- Blass, L. (2000). *Quest: Listening and speaking in the academic world, Book 1*. McGraw Hill, 158-160.

- Frazier, L. L., & Mills Frazier, R. (2004). *North star: Listening and speaking, Basic/low intermediate* (2nd ed.). Longman, 27-28.
- Maurer, J., & Schoenberg, I. (1998). *True colors 1: An EFL course for real communication*. Longman, 24, 60.
- Purpura, J., & Pinkley, D. (1999). *On target 1, Intermediate*. Longman, 8, 39.
- Richards, J. C. (1994). *Interchange: English for international communication*. Cambridge University Press, 3, 39, 90.
- Tanka, J., Most, P., & Baker, L. R. (2002). *Interactions 1: Listening/speaking* (4th ed.). McGraw-Hill, 7, 81-82.

**Idioms Textbooks**

- Kalkstein Fragioudas, H. (1997). *All clear, Advanced*. Heinle & Heinle, 126, 145.
- Kalkstein Fragioudas, H. (1993). *All clear: Idioms in context*. Heinle & Heinle, xix.

**Pronunciation Textbooks**

- Beisbier, B. (1994). *Sounds great, Book 1: Beginning pronunciation for speakers of English*. Heinle & Heinle, 84, 87.
- Hewings, M., & Goldstein, S. (1998). *Pronunciation plus: Practice through interaction*. Cambridge University Press, 122, 126.

Teaching pronunciation is of no less importance in the developing of reading and writing habits and skills, since writing (or what is written) is a graphic representation of sound sequences. In reading the visual images become acoustic images. These are combined with kinesthetic images, resulting in inner speech. Wrong pronunciation often leads to misunderstanding. For example, when a speaker or a reader replaces one phoneme with another he unintentionally uses quite a different word, in this way altering the sense of what he wanted to say. For example, white instead of wide, it instead of eat. Some teachers call it "accent modification", others "accent reduction", and others simply "pronunciation". It doesn't matter what you choose to call it, you are essentially trying to teach somebody English: both how to understand spoken English and how to be understood when speaking English. That's right, it's just English. I have studied several languages, and all of my teachers have, sometimes painfully, made me practice the different "hard to get" sounds. So why not in English class? Well that's what I would like to speak about today. Below I give some of the most important concepts to understand English pronunciation can be tricky, especially for ESL (English as a Second Language) learners. Start by introducing how to pronounce English vowels and consonants. Some sounds rarely occur in other... Tell them not to be discouraged if they have trouble, and explain that it takes time to train their mouth and tongue to move in unfamiliar ways. [3] X Research source. The sound "eth" is especially difficult, and students often substitute it with "et", "ef", or "es". Depending on a student's native language, other tough sounds could include "el", "er", and soft "eg" or "ej". Differentiating between "eb" and "ev" or "eb" and "ep" can also give students trouble.