Marc Picard

On Teaching the Pronunciation of Subphonemic Segments in English

Abstract

Although ESL pronunciation instruction has traditionally focused almost exclusively on phonemes, there have always been those who advocated devoting some attention to allophones. However, given the limited time and resources foreign-language teachers usually have to devote to pronunciation, they may have to be very selective in the subphonemic segments they introduce in the classroom. But on what basis can one choose which allophones to teach? In this study, I will attempt to show that there are two basic criteria which can serve to distinguish which of these may merit more consideration: transferability i.e., potential pronoumcability, and differential salience, i.e., the phonetic distance between a particular allophone and its corresponding phoneme. On this basis, it will be argued that if one is to consider teaching the pronunciation of any subphonemic segments in North American English (NAE), the highest priority should be given to flaps, especially those that are allophones of /t/. 

1. Introduction

Intelligible pronunciation is an essential component of communicative competence. (Morley 1991: 488)

In the area of pronunciation instruction, a lot of attention has traditionally been directed towards how best to impart prosodic and segmental information to students, but until recently there has been considerably less discussion on the content of an ideal pronunciation curriculum or program. Yet the issue of what should be taught in pronunciation classes is far from trivial in view of the limited time and resources foreign-language teachers usually have to devote to this particular concern. For instance, MacCarthy has stated that

[t]he teaching of pronunciation can occupy only a portion of class time for the teacher of a foreign language, whose time overall naturally has to be divided among the different areas of language study. (MacCarthy 1976: 212)
Similarly, in referring specifically to pronunciation instruction, Brown has pointed out that

[t]he ELT teacher (…) must often decide which features of language, on the one hand, are important and therefore merit precious class time and which, on the other hand, are relatively unimportant and may be overlooked until a more advanced stage. (Brown 1988: 593)

Traditionally, the phonological aspect of ESL has involved an almost exclusive preoccupation with phonemes, as Dickerson has so rightly noted:

[P]ronunciation instruction for years (…) consisted of a heavy emphasis on segmentals—the vowel and consonant sounds of English—and a somewhat lighter emphasis on suprasegmentals—stress and intonation (…) The content and presentation were so well defined that the myriad pronunciation textbooks on the market seemed to be cut from the same pattern (…) [I]t was not until recently that the crucial importance of rhythm for intelligibility began to be widely reflected in our textbooks. (Dickerson 1987: 11–12)

Although the prosodic features of English have been gradually receiving much more emphasis (cf. Gutknecht (1978), Chela de Rodriguez (1983), Dickerson (1987), Haycraft (1992), McNerney and Mendelsohn (1992), Gilbert (1994)), another area of phonological instruction, viz., that of allophonic or subphonemic variation, has continued to be almost totally neglected. Yet, over the years, a few researchers in applied linguistics and TESL have occasionally tried to bring attention to the potentialities, and even the necessities, of providing students with instruction in this area.

In Standwell’s view, for instance, “far from the phoneme being of any assistance to the language teacher, it is rather a red herring” since “in pronunciation teaching one does not teach phonemes, one teaches phones” (1973: 119). Although few have espoused such an extreme position, opinions have been clearly expressed to the effect that that allophones should be given some consideration. For one, Shen has claimed that “[a]llophones provide acoustical clues to the recognition of phonemes” so that “[b]y careful utilization, they can constitute an aid to a more satisfactory production of the foreign language” (1959: 18). In the same vein, Prator has opined that

[e]ven in a short course, if ability to speak English is an important objective, we should probably include attention to a few of the most important allophones that are in complementary distribution. (Prator 1971: 71)
More recently, Celce-Murcia et al. have made some interesting observations in this regard. Concerning the aspiration of /p t k/, for instance, they have seen fit to apprise ESL instructors of the following potential difficulties:

In many languages, initial voiceless stops are less strongly aspirated than in English, or are even unaspirated. Speakers of these languages may therefore tend to confuse initial /b, d, g/ in English with their own language’s unaspirated /p, t, k/ in this position. These learners may be misperceived by English native speakers as producing back instead of pack, or die instead of tie. In fact what they may be producing is an unaspirated /p/ or /t/ in place of the English aspirated counterparts. They may, of course, also have difficulty in differentiating such minimal word pairs. For these learners, aspiration can provide a valuable clue to perceiving and producing these words. (Celce-Murcia et al. 1996: 63)

Clearly, these phonologists are of the opinion that even though the aspiration of voiceless stops is not phonemic in English, failure to make ESL learners aware of it can lead to confusion. Native speakers may misinterpret some of their voiceless stops as voiced stops, and they themselves may have difficulty in discriminating between these two types of obstruents when they hear them.

There appears to be sound evidence, then, that there is more to teaching the pronunciation of English segmentals than simply concentrating on phonemes to the exclusion of their contextual phonetic variants, some of which may be articulatorily contrastive enough to cause confusion. But on what basis can one choose which allophones to teach? In this study, I will attempt to show that not all such sounds are created equal, as it were, and that there are two basic criteria which may serve to distinguish which of these may merit more attention: potential pronounceability or transferability, and differential salience, i.e., the phonetic distance between a particular allophone and its corresponding phoneme. On this basis, it will be argued that if one is to consider teaching the pronunciation of any subphonemic segments in North American English (NAE), the highest priority should be given to flaps, especially those that are allophones of /t/.

2. On teaching allophones

As was pointed out above, a major reason why allophones are so often neglected in pronunciation teaching is the widespread belief that phonemes are the only significant segments in language, the only ones that can make semantic differences. However, consider the following observation by Shen:
In the teaching of a foreign language, a comparison between the phonemic systems of both languages is essential. The purpose is to discover the phonemes that occur in the foreign language which do not occur in the native language. Such phonemes are generally accepted as the biggest learning load for the students and similarly the heaviest teaching problem for the instructor. But there are also structurally contrastive relations among the allophones of the two languages and between allophones of one language and phonemes of the other. Such relations must not be ignored either. (Shen 1959: 8)

Various other researchers in the area of L2 phonological acquisition have made comments along the same lines, i.e., that the phoneme is not the be-all and end-all of pronunciation teaching. Leather and James, for example, have noted that

it has become clear from a large volume of research over the past few decades that although the phoneme may be a useful construct in linguistic description, its status in the real-time processing of spoken language is problematic. (Leather & James 1996: 278)

As a consequence, then,

[p]erhaps a focus on the closer phonetic detail of F[o]reign L[anguage] sounds in acquisition (…) may be seen as a necessary corrective to previous more ‘coarse-grained’ contrastive analyses of the phonemes of first languages (L1s) and second languages (…). (James 1986: 225)

The whole issue of phonemes, allophones and intelligibility is summed up very astutely by Prator and Robinett:

[T]he safest solution for a teacher of ESL is (…) to regard unintelligibility not as the result of phonemic substitution, but as the cumulative effect of many little departures from the phonetic norms of the language. Many of these departures may be phonemic; others will be allophonic. But under certain circumstances, any abnormality of speech can contribute to unintelligibility. (Prator & Robinett 1985: xxii)

If we adopt the position that the teaching of pronunciation should involve more than simply imparting the phonemic contrasts of a language, the question then becomes how far one is willing to go, for there is virtually no limit to the phonetic details one can introduce to students. For example, some applied linguists have advocated the use of articulatory setting (also known as phonetic, phonatory or voice setting) in second-language instruction (cf. Ozga (1977), Erazmus (1982), Esling and Wong (1983), Esling (1987), Jenner (1990), Collins and Mees (1992a, 1992b)). The
concept of articulatory settings, which was first outlined in Honikman’s (1964) seminal article, can be defined as

[the overall tendency (…) to maintain the organs of speech in some particular configuration throughout speech, as reflected in such factors as the height of the velum, the degree of lip-rounding and the tension of the tongue and lips. (Trask 1996: 34)

In essence, the theory proposes that each language has a unique configuration of articulators accounting for or establishing the natural sounds of that language that give it phonological unity and differentiate it from other languages.

For example, here is how Mees and Collins (1992) propose that Danish speakers should modify their articulatory setting when learning NAE:

- adopt a generally tenser setting of the body of the tongue, with firmer closures and narrowings for stops and fricatives and use of the tongue-tip for alveolar consonants;
- avoid palatalization in favor of uvularization plus a bunched tongue-shape to produce /Š/ and provide r-coloring for r-adjacent segments;
- adopt semi-continuous nasalization;
- adopt a laxer lip-setting to facilitate weak rounding and protrusion for certain consonants;
- adopt a relaxed larynx setting for weaker glottalization and the avoidance of anterior voice.

Trying to implement these kinds of elaborate articulatory guidelines in an ordinary ESL classroom would be well nigh impossible given that students with different L1s would have to be given different sets of instructions. Moreover, one could easily run the risk of simply overwhelming them with indigestible and impracticable articulatory minutiae. In other words, too much attention to phonetic detail might be viewed as a waste of time given the widely recognized Critical Period Hypothesis. For if it is indeed the case that “after puberty it is nearly impossible to learn a second language and ‘pass for native’, especially in the area of phonology” (Major 1990: 14), that is, if we already know that “adults rarely attain native-like competence in an L2 phonology” (Young-Scholten 1992: 201), then the potential improvements in intelligibility that could be gained by teaching such fine distinctions of perception and production might not be worthwhile when measured against the time and effort involved. One cannot imagine any but
the most gifted students benefiting from the discrimination of such intricate sound distinctions.

Some sort of middle ground between this sort of exhaustive articulatory approach and one involving simple phonemic contrasts would therefore seem to be desirable. Such an approach, as was mentioned previously, would involve introducing students to allophonic variation. However, one might question whether all allophones deserve equal opportunity, so to speak, since there are surely some that are more frequent or distinctive than others, and so more apt to cause misperception when pronounced incorrectly.

For example, we know that “[t]he sound [f] sometimes occurs in English, as an allophone of /h/ in intervocalic position” (Laver 1994: 305), as in behind and ahead, and also that the labiodental nasal [m] is found in forms like triumph and comfort, but surely no one would want to attach as much importance to such phonetically inconspicuous and relatively uncommon phenomena as to the sort of ubiquitous flapping of word-medial and -final alveolar stops that is found in sequences like I did it Saturday or He credited it again.\(^1\) In the following sections, we will look at some possible criteria for ranking allophones in terms of the pedagogical attention they should be given.

3. The criterion of transferability

As was noted by Prator some thirty-five years ago, there exists an ever-present “problem of establishing a hierarchy of priorities for the teaching of pronunciation” (1971: 61). One reason for giving priority to certain allophones over others might be their superior potential pronounceability or transferability due to the fact that they are often found to be phonemic in many languages. This is based on the existence of

abundant evidence that the beginning learner seeking to impose phonetic structure on the L2 speech to which he is exposed makes perceptual reference to the phonetic categories of his L1. (Leather and James 1996: 274)

\(^1\) What should also be disregarded are those subphonemic segments that are the “universal consequences of inherent properties of the human speech-producing mechanism” (Anderson 1976: 340), such as the lengthening of vowels before voiced consonants, e.g., bit vs. bid, the rounding of consonants before back vowels (or vice versa), e.g., keel vs. cool, or the dentalization of alveolar consonants before dentals, e.g., tent vs. tenth.
Or, to look at it from another perspective, this potentiality stems from the frequent observation that “[n]otoriously, second language learners are unable to produce distinctions that do not function contrastively in their L1” (Busà 1992: 48). What this criterion entails, then, is that ESL instructors should normally achieve greater success in getting learners to pronounce a particular allophone correctly the more such an allophone is liable to be phonemic in their native language, though there are obviously limits to what can be done in this regard in heterogeneous L1 classes.

This is in line with one of Catford’s principles of pronunciation instruction, namely “the utilization of all sounds known to the students,” and the fact that “[t]eachers should (...) take advantage of articulatory possibilities of their students” (1987: 97). An example of this sort of technique as applied to flaps can be found in Cook’s pronunciation textbook wherein she gives the following suggestions:

If you speak any language—such as Spanish, Japanese, Italian, or Dutch among others—where your R touches behind the teeth, you are in luck with the American T. Just fix the association in your mind so that when you see a middle position T, you automatically give it your native R sound. Say, Beri bara bira (...) with your native accent. (Not if you are French, German, or Chinese!). (Cook 1991: 87)

Now, the most commonly adduced subphonemic segments in NAE are:

1. The flaps, i.e., the oral flap [r] and the nasal flap [r], which stem from /t d/ and /n/ respectively after a vowel or a central (and, for some speakers, a lateral) approximant, and before an unstressed vowel within words and any vowel at word boundaries, e.g., party [pʰáíri], get angry [ɡɛɹɛŋɡi], incredible [ɪnkʰɪɛɹɔbɛt], inspected it [ɪnsɪˈpektəɹɪt], banana [bənəˈɛɡə], an assistant [əɹɛsɪˈstənt]

2. The positional variants of /p t k/, viz., aspirated [pʰ tʰ kʰ], which occur at the beginning of words and stressed syllables, e.g., certain [pʰɛɹtʰˈɛyn], tomatoes [tʰəməˈeɪroʊz], correct [kʰəˈɛkt], and unreleased [pʰ tʰ kʰ] which are found in certain codas, e.g., opted [ˈɑːptəd], coat [kʰəˈoʊtʰ], acme [əˈkʰmɛi];

3. The velarized lateral [ɹ] which is found in syllable codas, e.g., fall [fɔːl], falter [fɔːtəl];

4. The syllabic nasal [n] which occurs when /n/ follows /t d/ in syllable codas, e.g., sweeten [swiːtnə], Sweden [swiːdən].
When we check for these various segments in Maddieson’s presentation of “the phoneme inventory of each of the carefully selected sample of 317 languages which comprise the UCLA Phonological Segment Inventory Database (UPSID)” (1984: 200), we find that unreleased stops, syllabic nasals and nasal flaps are not contrastive in any of them, velarized laterals are phonemic in only about 2%, and aspirated stops and oral flaps (which may also be described as taps, and as either dental or alveolar) are each found in approximately 25% of these languages. On this basis, then, there is no doubt that the criterion of potential pronounceability would strongly favor aspirated stops and oral flaps over any of the other NAE allophones.

4. The criterion of differential salience

The second criterion that should figure in the determination of which allophones should be prioritized is that of what can be termed differential salience, that is, the phonological distance that exists between a particular phone and its corresponding phoneme. The idea behind this concept is that a greater articulatory distance between two segments should normally make them more perceptually distinct and thus more apt to lead to unintelligibility when one is substituted for the other. One way to measure the dissimilarity of segments in complementary distribution is to compare them in terms of distinctive features. Among the high-prioritized segments that were established above, namely the aspirated stops and the oral flap, we see that the only feature that sets plain stops apart from their aspirated counterparts is [spread glottis] (or simply [spread]), whereas flaps are distinguished from alveolar stops by either two or three features, as shown below:

---

2 More generally, salience has been defined as “a property of a linguistic item or feature that makes it in some way perceptually and cognitively prominent” (Kerswill and Williams 2002: 81). The postulated phonological concomitant of this factor is that “[s]peakers are (...) more aware of variables whose variants are phonetically radically different” (Trudgill 1986: 11).

3 As defined by Halle and Stevens (1971), sounds that are produced with the vocal cords drawn apart such as aspirated and breathy-voiced or murmured consonants as well as voiceless vowels and glides are [+spread] while all other segments are [-spread].
Further evidence that differential salience plays a part in the recognition and noticeability of allophones can be drawn from various quarters. For instance, in Picard’s (2001) exhaustive study of the treatment of Flapping in NAE pronunciation textbooks—both student- and teacher-oriented—that were published over the last 35 years or so, a full 40% of the ones that mention [r] do so solely as a positional variant of /t/, that is to say, in the case where the articulatory difference is at its maximum. Recent pronunciation dictionaries such as Wells (1990) and Upton et al. (2001) are also of interest in this regard since the only allophone either one systematically transcribes is the flap emanating from /t/ (which the former writes as /t\^/ and the latter as /d/).

Overall, then, it would seem that oral flaps, and especially those that alternate with voiceless alveolar stops, as in write [r\^yt] and writer [r\^y\^r\^t], or hit [hitt] and hit it [h\^r\^t], should be given the highest consideration among the allophones of NAE. For anyone who might be wondering whether it is worthwhile to spend any time on these segments, the following observation by Celce-Murcia et al. would certainly seem to merit serious consideration:

Most learners are unaware of the flap allophone in NAE. This can mark their speech as foreign; it may also be the source of listening discrimination problems. For example, in place of the flap in phrases such as "cut it out" or "put it on," these learners may produce a fully articulated /t/, and they may hear /d/ as opposed to [r] in words like latter. Of course, those who have studied British English before encountering NAE may also be confused, since the presence of [r] is one of the most salient features differentiating NAE from British English dialects. For these learners, this difference in dialect will need elaboration (1996: 65).

---

4 For the whys and wherefores of the feature [vibrant], see Picard (1997).
5. On teaching flaps

One question that remains to be resolved is whether ESL instructors should have their students actually practice the pronunciation of flaps or simply make them aware of their existence in NAE. In other words, is the teaching of flaps to be approached in terms of their production, or should some training in their perception be considered sufficient? As it turns out, both points of view have been expressed in the literature.

Avery and Ehrlich, for example, give the following advice to ESL teachers:

You should not insist on having students pronounce flaps because using a /t/ where native speakers use a flap results in very little loss in comprehensibility. However, students should be given extensive practice in the recognition of flaps. They are very frequent in the spoken language and the ability to recognize words that contain flaps is very important in improving students' comprehension of natural speech. (Avery & Ehrlich 1992: 42–43)

This is echoed by Dauer who, having pointed out that Flapping “is one of the main differences between North American English and other varieties of English,” adds that although “[i]t is not necessary for a non-native speaker to pronounce /t/ in this way,” ESL students “need to be able to hear it in order to understand native speakers” (1993: 142).

On the other hand, if one is to judge by the following statement, Celce-Murcia et al. seem to be in favor of not only making ESL students in North America aware of Flapping but of also trying to get them to integrate flaps into their L2 phonological system:

The flap allophone of /t/ and /d/, being a distinct feature of NAE, would be a priority in situations where this variety is being taught. For example, students often complain that when ordering water in a restaurant, they are not understood and have to repeat their request. This is probably a function of their not producing the flap allophone of /t/ (my emphasis). Since this distinction is not present in many dictionaries (especially the small bilingual dictionaries), students never discover it. Anecdotes of this nature emphasize the critical need for an awareness of positional variation and a teaching agenda that addresses this need. (Celce-Murcia et al. 1996: 69)

More categorical is Wells who says:

Learners of English as a foreign language who take Am[erican] E[nglish] as their model are encouraged to use § [i. e., r] where appropriate. (Wells 1990: 703)
6. Conclusion

It seems apparent that a sizeable number of people involved in the phonological aspect of second language education do not fully subscribe to the notion of a phoneme-centered pronunciation universe. Many of them have evidently heeded Prator’s admonition that “teachers would do well to suspect that any departure from the phonetic norms of the language can have a negative effect on the intelligibility of speech” (1971: 61). Thus, in the aforementioned study of pronunciation textbooks by Picard (2001), for example, it was shown that about half of them contain some sort of information on Flapping, a fact that certainly seems to demonstrate an awareness on the part of the authors that ESL students should be familiar with this process. More significant, perhaps, is the fact that over 75% of the resource books designed specifically for teachers mention flaps, since it is quite possible that their influence might have some sort of trickle-down effect on those who will be designing and using ESL student-oriented pronunciation textbooks in the future.

All in all, then, the wisest course of action for anyone involved in teaching pronunciation in any capacity would seem to be that advocated by Prator:

In the absence of any consensus regarding the degree of accuracy to be sought in teaching pronunciation, most teachers will probably want to take a position somewhere between that of the champions of absolute allophonic accuracy and that of the methodologists who insist on no more than an ability to produce a rough approximation of phonemes. (Prator 1971: 64)

From the evidence that has been presented above, it would appear that certain NAE subphonemic segments should occupy a place in this sort of proposed phonetic middle ground, and that first among these should be the oral flap [r].

References


Prator, Clifford H. (1971) Phonetics vs. phonemics in the ESL classroom: When is allophonic accuracy important? TESOL Quarterly 5: 61–72

Shen, Yao (1959) Some allophones can be important. Language Learning 9: 7–18.
Standwell, G. B. J. (1973) The phoneme, a red herring for language teaching? Audio-
Trask, Robert Lawrence (1996) A Dictionary of Phonetics and Phonology. London: 
Routledge.
Young-Scholten, Martha (1992) The Subset Principle in Interlanguage Phonology: A 
Amsterdam: University of Amsterdam.

Contact information:

Marc Picard
Concordia University
Montréal, QC
Canada H2X 1E1

e-mail: picard(at-sign)alcor(dot)concordia(dot)ca