

ESL pronunciation: a "larger" view¹

William Acton²
Nagoya University of Commerce

1. Pronunciation research and methodology

I had considered other titles such as: "ESL pronunciation meets Frankenstein" or "All who teach here, abandon pronunciation!" Each captures something of the spirit of what follows. For most contemporary language instructors, effective pronunciation teaching implies relatively small classes, a good deal of teacher training, easy access to native speakers, and "high" technology. Here we will consider how, with little or none of the preceding conditions met, pronunciation instruction might be done well in large classes.

In the last few years there has been a most interesting convergence of ideas concerning pronunciation work. On the one hand we have those who have been teaching in large classes (defined as a class of forty or more); on the other hand are the various "communicative" language teaching methodologists and friends. In the case of the former, we have a group of professionals who have been trying to figure out what is possible in large classes, experimenting with a wide range of formats. In the latter, are those who, for other reasons, have chosen to eliminate as theoretically unacceptable many pronunciation-based classroom procedures. What is most interesting is that, from different perspectives, the two groups seem to have arrived at a very similar inventory of acceptable pronunciation-teaching procedures.

There has been a parallel development in research in the acquisition of second language phonology. We are at a point in the evolution of the field where explanations of learner variability have come to be more and more multi-variate and multi-dimensional. Whereas earlier we could talk in terms of two or possibly three variables as accounting for most of the variance in acquisition of sound systems, it is now clear that from a research perspective the sound "system" as a distinct construct may be of little value to us. In other words, what impacts language learning in the broadest sense affects the acquisition of sounds.

The basic premise of this paper is that much of what is *possible and effective* in the large class setting is also *best* for the communicative classroom. The reason for pursuing this line of argument here is first to attempt to create a workable model of large class pronunciation teaching, and then (and most importantly) to develop from it a better model of general pronunciation teaching. Again, the idea is to move in a straightforward, deductive manner to persuade the reader that

¹Caveat emptor: One of the "pleasures" of contributing to a collection of conference presentations is that one can in good conscience allow the "spirit" of the oral presentation to survive the transition to the page. I have done that here, staying with the style and tenor of the original. Consequently, the reader should try to visualize a receptive, delightful group of like-minded scholars and researchers listening and responding throughout. Likewise, I will not burden the reader with references unnecessarily. . . .

²I would especially like to thank Corrine Cope for her many insightful comments on earlier drafts of this paper.

what is good pronunciation method for the large class "gander" is equally valid for the modern communicative language teaching "goose."

We begin with a dual purpose, somewhat truncated, somewhat irreverent, account of the "history" of ESL pronunciation teaching and research. That discussion will serve both to make clear how I view the development of the field and to identify what appears to be the range of possibilities when it comes to models and theories of how pronunciation is acquired, produced and taught. We will briefly consider ten such models or schools or "theories" which have enjoyed, or still do enjoy, some acceptance in the profession. In general, each position is either (a) true, in part, (b) true for some limited population of students, or (c) true for some yet-to-be discovered group of language learners.

2. Ten schools of thought on pronunciation

I. The A-historic School

WOpronunciationRD

II. The Structure-first School

CA = *maybe*

III. The "Bottom up" School

individual sounds, then
morphology, then
vocabulary, then
sentences, then
maybe conversation

IV. The Naturalists

"Why bother?
Not necessary.
Don't bother!"

V. The Neo-naturalists

"Listen first; then *maybe* pronounce later."

VI. The Affect-first School

"I love the way you say that."

VII. The Personality-plus School

"My mother can't pronounce that either!"

VIII. The Neuro-pre-puberty School

"It's just too late!"

IX. The Communicative School

"Did you mean to say that?"

X. The Top-down School

sort of conversation, then
sort of sentences, then
sort of vocabulary, then
sort of morphology, then
sort of individual sounds

The A-historic School (I) represents the common sense notion that the pronunciation of a word is just part of the word, like the meaning or its usage. From that perspective pronunciation is learned by repeating or writing the word down as the rest of the word is learned. There is no particular consideration of the

sound system of the language or the specific qualities of any sound that is especially troublesome.

There is good evidence that children and adults in "natural," non-school environments learn at least some of what they do that way, at different stages in the process. From the standpoint of remembering pronunciation, we certainly do access much information lexically, that is by "going to" the word that we know has a particular sound to recall how to pronounce it. The concept of lexically conditioned memory and learning are central to current second language learning theory.

The Structure-First School (II), too, has a very common-sense ring to it. The "CA=*maybe*" suggests the idea that by doing some kind of contrastive analysis you can predict what will be hard for learners to learn in a second language. In reality, because of a multitude of factors, often the best we can do is simply guess "*maybe*". There are always good reasons to do at least a brief contrastive work-up of the sound systems of languages and their learners, particularly when there is no generally agreed upon list of potential problems; but that analysis will probably never be as useful as that consensus arrived at by years of enlightened teaching experience.

The "Bottom up" School (III) is an extension of the Structure-first School. What this position implies is that it is generally preferable to build up language skills and, in this case, pronunciation, from the smallest unit to the largest. That, too, seems eminently reasonable, but experience alone has taught the profession that not much of language is learned that way. Presenting language material in a semi-systematic way, from the smallest unit to the largest -- as part of an overall program -- may be less problematic; but requiring mastery of one level before allowing learners to go "up to the next" has proven to be highly injurious to the learning.

The Naturalists (IV) have appeared in several forms this century. One group made very telling critiques of the Structure-first School, and another group had earlier been proponents of the "WOPronunciationRD" model. Their primary argument is a powerful one: in natural settings we do not focus very much on structure but rather work at communicating, ignoring form and strict accuracy most of the time. It leads in several directions. We will return to some of those implications later.

Among the Neo-naturalists (V) I include what is today termed the "Natural Approach." They have taken part of the orthodox naturalist position and used it very effectively in arguing for a preliminary "listening" stage in acquisition and instruction. Their position on pronunciation is parallel: that premature speaking, for various reasons, is counterproductive. They argue, with some preliminary evidence, that even pronunciation will be better if the initial listening period is allowed. The kind of listening proscribed is not skill-based listening but rather task-based, action-oriented activities.

The Affect-first school (VI) represents a wide range of language teachers and researchers who see attitude or emotion or unconscious states as being the most important variable (or at least the first to be considered) in learning or using pronunciation, or language. The role of emotional state in any performance art is taken for granted; in language learning it is far less universally accepted or understood. Although we know that "nerves" can affect pronunciation and diction, it has not been established just how, in the long term, emotion influences

the learning of pronunciation. The problem with strongly affect-based explanations is the tendency to isolate one variable at the expense of others which may be closely related to affect and emotion.

The Personality-plus school (VII) has a nice ring of authenticity about it. Common sense tells us that certain personality types are better at learning the pronunciation of a language than are others. Empirically, there seems as yet no convincing way to show that, however. Aside from the problem of finding reliable, valid personality measures that can be correlated with language learning, and especially pronunciation, it is extremely difficult for researchers to agree on what standard or learning phase of pronunciation to use. Furthermore, some models of personality are static and others are dynamic. Learning a language may, for some, be an occasion for the personality to change in some aspects, to grow or reveal otherwise hidden potentials. That seems to be especially the case with pronunciation and accent.

It is a wonderfully complex task, trying to predict based on personality whether or not one's pronunciation will, in the end, be better or worse than that of some "other" personality type. In other words, personality has little predictive validity today in language research. Personality discussions related to pronunciation learning tend to have a decidedly deterministic bent. Personality is most often called up as an excuse for failure, rather than as an explanation for success. From a cultural perspective, general considerations of personality may be more useful. For example, the Japanese "student" may come to the second language learning process with a very different, culturally-based, perspective on the appropriateness of volunteering answers and avoiding public-speech errors than does the typical American student. Granted, that is more a cultural question; hence, the "-plus" in the name of this school of thought.

The Neuro-pre-puberty school (VIII) has much in common with the Personality-plus school. This group points to what may be biological determinants of pronunciation authenticity. Here, too, despite the very common-sense basis of the idea that puberty is the watershed for successful language learning, especially that of the pronunciation and the accent, the empirical evidence is not there yet to support such as position. It is easy to make interesting non-empirical explanations for children's obvious "edge" in acquiring native-like pronunciation, but no one has yet succeeded in identifying the true source of that difference. The pre- vs. post- puberty distinction is also used as an excuse for failure.

That there are differences between the "results" of learning as a child and learning as a teenager or adult is not the issue. But what it is in the complex of human personal and inter-personal variables that accounts for that is still a mystery. I suspect that the answer will emerge in more complex models than we have today, models which will allow us to feed in more specific information about the individual learner, the specific task and the immediate social context -- along with the neurological data on structures in the brain.

The Communicative school (IX), like the others, focuses on one dimension of the process of learning pronunciation. What a member of this group might maintain is that learning the sounds of the language must be done in the context of real communication, as it occurs in the outside world. Then in the process of communicating and getting feedback on the relative accuracy of our pronunciation, we gradually begin to approximate the necessary or "correct" pronunciation.

That is a powerful and often seductive argument. In the best of all possible worlds and classrooms, that may be exactly how it happens. That certainly agrees with our observations of our own learning at times, and it may be a more valid assertion regarding children than adults; but when it comes to operationalizing that notion, it is another matter. Perhaps the best approach to this position is to assume that it is correct, and that we can agree on when such a communicative event occurs, such that the feedback is available and used to alter pronunciation.

The last "school", the Top-down (X) approach, like the Bottom-up school, does not have many strong proponents; but the idea underlying it shows up frequently. The extreme position would be to claim that pronunciation (or language) is best learned from the top-down, that we learn to converse before we learn words. From first language acquisition there is a great deal of evidence that children acquire the rhythm, stress and intonation of a language long before they spend much time on most of the vowels and consonants.

Several colorful but largely ineffectual attempts were made to operationalize the top-down approach in foreign language teaching. What those early second language researchers and methodologists who argued for the same process succeeded in doing was convince many that adult language teaching should at the very least, proceed on several levels at once. To this point in time there is little empirical evidence that initial work on the melody and the rhythm of the second language ultimately results in better, more accurate speakers. Nonetheless, for several reasons, it makes sense to many of us; and we rely on experience and the parallel with child language learning in this area to justify our design decisions.

Before we turn to the application of these ten perspectives to large class pronunciation teaching, and pronunciation teaching in general, let me clarify why I have treated those "schools" as I have. Because we have difficulty in establishing a strong research base for a method does not mean that we abandon the project. On the contrary, as we "experiment" in the construction of methods, the successful methods (using whatever explanations we as a profession will accept at that point) provide us with strong hypotheses for further research.

In Part III I will sketch out a model of key features of large class pronunciation, allowing each "school" to contribute, if only as apologist for what I am proposing.

3. Large-class pronunciation instruction

First, let us consider briefly a "priority" or two which each of the schools would probably have for our curriculum:

I. A-historic - Pronunciation should be taught focusing on the correct expression and expressiveness of words and sentences.

II. Structure-first - Identify and prioritize those sounds and processes that are most problematic in initial learning. Provide for various kinds of corrective feedback.

III. Bottom up - Work on vowels, consonants and word stress from the beginning. Present clear frameworks to students, including charts, etc.

IV. Naturalists - Don't focus on pronunciation unless absolutely necessary, unless it interferes with communication.

V. Neo-naturalist - Do a great deal of listening, some of it with a pronunciation-focus.

VI. Affect-first - Make sure that the classroom is as positive, "stress free," and open to free expression of emotion as possible.

VII. Personality-plus - Tailor the class to the culture of the students so that they can ultimately become expressive in the language, being able to "act" like a native speaker.

VIII. Neuro-pre-puberty - Allow for adult-style strategies of learning.

IX. Communicative - Students must engage in communicative activities, working in groups to develop better pronunciation; take responsibility for learning.

X. Top-down - From the outset, focus on teaching suprasegmentals (rhythm, stress and intonation), only later getting to individual sounds.

Taken as a whole, this is quite a "manifesto;" but in fact these principles form a reasonable base for successful pronunciation work in large classes. In the following I will briefly describe some of the classroom procedures that the manifesto³ would require:

A. Identification

1. Schemas and charts and explanations will always be available, but not always formally presented (II, VIII, X).
2. There will be clear statements of achievable objectives, etc. (VIII).

B. Practice

1. Extensive use of pair and group work, both communicative and form-oriented (IX, II).
2. Homework will be closely tied to individual needs, required and "checked" (VI, VII).
3. The ultimate goals for the practice will be self-monitoring, self-initiated pronunciation work (V, IX).

C. Correction

1. Students correct each other in pair work (e.g., stress placement, vowel identification, expressiveness, gesture) (II, IX).
2. Students correct themselves in lab or homework. Focus on aspects other than vowel or consonant articulatory problems (II, V, VIII, X).
3. Requirement for high quality audio and video materials and models for correction and feedback (II, III, V, X).

D. Suprasegmentals

1. Attention to rhythmic structure of any text being used, including various schema, dictations, small group work, etc. (I, VIII, X).
2. Stress placement can be attended to in any text (II, VII).
3. In any conversation exercise, intonation contours or juncture can be the focus (II, IV, X).

E. Discourse and conversation

³In some cases the descriptions of the techniques listed in this section are obviously too short to be of much use to the inexperienced instructor. In addition, the connection or relevance of that application to the particular "school" may not be all that obvious.

1. Overall, pronunciation work must be embedded within some kind of discourse or conversation-based activity (I, IV, V, VI, IX).
2. Focus should always be on the context of the sound, not simply on the sound itself (I, IV, IX, X).
3. Dialect differences should be part of the materials and receive some conscious attention and discussion (II, VII, IX).

F. Expression as pronunciation

1. All speech in class must be emotionally expressive. Any utterance by instructor or student can require a re-transmittal due to lack of emotional "meaning" or expression (I, II, III, IV, VII, IX).
2. The model of the instructor and/or the recorded material must be expressive. That does not mean that the mode must be excessively accurate, only alive and animated (I, II, VI, IX).

G. The Kinesthetic Connection⁴

1. In large classes, if most of the other conditions are met, it is possible to do drama of various kinds. Ultimately, drama activities may be the most effective device for improving pronunciation.
2. As part of drama work, body movement and gesture can (and must, as far as I am concerned) become the basic carrier of the words and sounds. If language is said expressively, with real emotion, the body can not help but move, the speaker becomes almost incapable of gesture-less expression. (In my large class work, gesture, along with expressive speech is a *quid pro quo*. In other words even the simplest of expressions or dialogues must be uttered with emotion and movement.)

4. Conclusion

Were one to take the large-class model outlined above and use it in a "small class" what would happen? First, it would necessitate that less time be spent on those areas requiring more one-on-one instructor-to-student instruction. Second, more would have to be done in groups. Third, and most important, what counts as "pronunciation" would be redefined. That has, in fact, already happened. With the coming of the Neo-naturalists and the Communicativists in North America and Europe, instruction in pronunciation was greatly de-emphasized.

In retrospect, the abandonment of what then had earlier passed as pronunciation teaching was not all bad. As is often the case, however, too much was discarded, particularly the idea that attention to the sound system was essential. When the relatively meaningless drills and repetition exercises of the Structure-first school were de-emphasized, there was at the time nothing convincing to replace them. The most recognizable candidate, the top-down approach, emphasizing attention to suprasegmentals and a *laissez faire* attitude toward individual sounds, was tried and found lacking. (There were a number of reasons for that, but basically, I think, it had the same shortcomings as the bottom-up method; namely, it was overly simplistic and lacked a coherent overall methodology.)

⁴This category is my own personal "favorite" but does not necessarily represent the current "state of the art" in large-class pronunciation teaching. (See Acton 1992 and 1993 for review of general kinesthetic approaches to pronunciation teaching.)

What you see in successful large class instruction, in pronunciation or in other subjects, is the integration of language with communication and the whole person. If students can be brought to the point where they (1) are able to work in groups together, creating expressive, communicative language, (2) can comment appropriately (tactfully, and unemotionally) on each other's speech production, and (3) are "responsible" enough to do related homework and some "self-starting" of the process, significant improvement in pronunciation is inevitable.

References⁵

- Acton, W. 1993. forthcoming. Marching to a different drama: LABO in the ESL university classroom, Bulletin of the Faculty of Commerce, Nagoya University of Commerce and Business Administration, 37(1).
- Acton, W. 1992. A kinesthetic approach to teaching English stress and rhythm, Bulletin of the Faculty of Commerce, Nagoya University of Commerce and Business Administration, 36(2): 245-274.
- Acton, W. 1984. Changing fossilized pronunciation, TESOL Quarterly 8(1):71-85.
- Brown, A. (ed.) 1991. Teaching English pronunciation: a book of readings. London: Routledge.
- Gilbert, J. 1984. Clear Speech: pronunciation and listening comprehension in American English. Cambridge: Cambridge University Press.
- Lessac, A. 1981. Body Wisdom, New York: Drama Book Specialists.
- Maley, A. and Duff, A. 1984. Drama Techniques in Language Learning. Cambridge: Cambridge University Press.
- Morley, J. 1989. Current perspectives on teaching pronunciation. Washington, DC.: TESOL.
- Stern, S. 1983. On the psycholinguistic reality of drama in the classroom. Paper presented at the 1983 TESOL Convention.
- Wong, R. 1988. Teaching English rhythm. New York: Prentice Hall.

⁵This presentation is broad in scope, a first approximation of a model for large class pronunciation teaching. As such, the bibliography could be almost endless. With the exception of the references to my own work, the various "schools" I allude to are the "standard" paradigms described in most current ESL methods texts. Consequently, I include only a few survey works that cover the general issues discussed. The ideas outlined here will be developed further in a symposium entitled "Large Class Pronunciation Teaching" at the 1993 TESOL Convention in Atlanta.