

# The Acoustic Properties of Vowels: A Tool for Improving Articulation and Comprehension of English

By Candalene J. McCombs

Associate Professor, Linguistics Department, Canisius College

## **Abstract**

Correct pronunciation is often a later step in the process of teaching English as a second language. However, a focus on the correct articulation of vowels can significantly improve listening and comprehension skills as well as articulatory skills. Vowels and consonants differ in their acoustic properties. Unlike consonants, vowel sounds are produced with very little obstruction of airflow, resulting in a difference in the way they sound. Vowels are more sonorous, or acoustically powerful, than consonants, thus we perceive them as both longer lasting and louder than consonants. The greater sonority of vowels also permits them to form the basis of syllables.

Introducing the phonetic properties of vowels is relatively easy. Second language teachers can train students to listen for vowel distinctions and teach the articulatory properties of vowels, strengthening students' listening and articulatory skills. Vowels form the nuclei of syllables, thus clarity in vowel sounds helps native speakers better understand foreign speakers. The focus on vowel sounds also supports instruction in the stress patterns of English, allowing students to more easily recognize individual words within sentences. This approach works particularly well with adult speakers who need to be clearly understood in professional settings.

## **Introduction**

This article draws upon two years of intensive English language training in a special project for two groups of white male speakers of Spanish from Colombia at various levels of proficiency in English and varying levels of education, from university freshman to graduate level. The students exhibited a range of proficiencies, from no knowledge of English, to a level of competency adequate for reading comprehension appropriate to the undergraduate level in an American university, but with very limited oral competency. The fact that both groups were composed only of males is significant only in relation to relevant affective concerns relating to the acquisition of English. Each male had been recruited to study in a seminary with the intent to become a Catholic parish priest in service to Catholic churches in the United States, though not necessarily in churches in which the dominant population is comprised of native speakers of Spanish. Competence in English was absolutely required. Each group of potential seminarians was given one

year to learn English and to reach a level of proficiency sufficient for study in either an undergraduate program or a graduate level seminary.

In the one year period, the Colombian students were to participate in their own intensive English program essentially segregated from the general student body. They were housed together in a group and lived in a seminary, attending English classes weekdays on campus and participating on weekends in a broad spectrum of service activities in a variety of parishes where the members were primarily native speakers of English. At the end of one year, they needed to be ready to enter an undergraduate degree program or the seminary, depending upon their individual educational credentials. Those who could not meet that deadline would return to Colombia to pursue their vocations in the seminaries of their homeland. It is important to note that learning English held no consequences for becoming a priest; however, whether the vocation was to be carried out in the United States or in Colombia was dependent upon acquisition of English.

Early in the program, by the end of four weeks of instruction, it was clear that all the students needed to work on pronunciation, regardless of the various levels of their proficiency. Two students with excellent comprehension could communicate in English, but their speech was so strongly accented that their interactions with native speakers other than the English as a foreign language (EFL) team of instructors were frustrating, and, as might be expected, their confidence was eroded, leading to avoidance behaviors. Though their motivation in the classroom remained high, the students were reluctant to engage in conversation in the settings where they were expected to provide volunteer service. At the same time, the students with little proficiency avoided engaging in even

minimal exchanges with the native speakers with whom they had contact in service work settings.

The complexity of the second language acquisition context was daunting. Ultimately, these students had to be able to address a congregation of native speakers of English, both in situations demanding public speaking and in more intimate exchanges that could be emotionally charged, demanding sensitive and clear communication. Post-pubescent learners of any second language confront the problem of interference across all the systems of the first language (L1) with the second or foreign language (L2). However, correct pronunciation of the sounds and sound combinations in a second language is significantly more difficult for post-pubescent learners to learn.

The time pressure of the one year program added to the urgency to address the problem of L1 interference. An additional complicating factor was the varied levels of proficiency among the participants. This contextual information is significant and relevant because the strategy for addressing the problem emerged through exploration and experimentation driven by practical necessity. A compelling concern in developing a program for this group of students was the sociolinguistic context in which the group would have to interact. Labov, as well as others, have shown through their research that listeners make very quick judgments about speakers based upon the pronunciation of a few words.<sup>1</sup> Additionally, native speakers are often uncomfortable in conversations with nonnative speakers whose speech is unintelligible. They shift topics, speak slowly and loudly, and they typically end uncomfortable conversations with nonnative speakers as quickly as possible.<sup>2</sup>

---

<sup>1</sup> William Labov, *Sociolinguistic Patterns* (Philadelphia: University of Pennsylvania Press, 1972).

<sup>2</sup> Gillian Brown and George Yule, *Discourse Analysis* (New York: Cambridge University Press, 1983).

Interaction between the students and members of the various parishes where they engaged in service activities provided a proving ground and a source of immediate feedback on our language training program. Feedback from church officials revealed that accented speech in and of itself was not a problem. Intelligibility of the accented speech was, however, requisite. That distinction allowed for a subtle shift in thinking about the pronunciation problem. The task became a matter of clarifying speech for the listener with a new focus on what could be more easily distinguished by the listener of speech.

This program was in development in 1998-99 at a time when a noticeable shift was occurring in attitudes toward the issue of pronunciation. Joan Morley, in her Georgetown University Round Table discussion of second language speech and pronunciation clearly identified and mapped out strategies for incorporating pronunciation in second language teaching, noting as well that to teach pronunciation one must possess “knowledge of phonetics and phonology and applied phonetics and knowledge of discourse phonology and the structure of oral discourse.”<sup>3</sup> She also stated that the “treatment of accent can no longer be trivialized, no longer relegated to a peripheral corner of the second language field.”<sup>4</sup> Her observation that there was a growing body of research in second language acquisition phonology provided hope for development of effective instructional materials in the future, but in fact confirmed a concern that most of the available materials for teaching pronunciation were unlikely to produce the results needed. Labov’s findings that native speakers based highly specific decisions about nonnative speakers based on relatively small amounts of information, for

---

<sup>3</sup> Joan Morley, “Second Language Speech/Pronunciation, Acquisition, Instruction, Standards, Variation, and Accents,” James E. Alatis et al., *Georgetown University Round Table on Languages and Linguistics* (Washington, D.C. : Georgetown University Press, 1996), 143.

<sup>4</sup> *Ibid.*, 144.

example, simply the way a few words are pronounced, led to a listener-centered rather than learner-centered instructional design for the pronunciation component of the program.<sup>5</sup>

Initially, the decision to work intensively on the vowel system of English was based upon several factors. First, only one member of the EFL team had formal training in phonetics, with detailed knowledge of place and manner of articulation of consonants and the parameters of vowel sounds, but it was evident that all five members of the EFL instructional team, and a number of tutors as well, would need to work together to provide the level of attention needed to work on pronunciation. Vowels represented a smaller more manageable class of sounds than consonants. All instructors and tutors could easily learn vowel differentiation to support instruction in the vowel system of English.

The fact that vowels are syllabic sounds was the point of departure. It is a well known fact that listeners hear vowels more strongly than they hear consonants. This understanding launched a search of the sound qualities of vowels and led to a strategy for intrusive instruction in pronunciation. The developmental sequences of interlanguage in second language acquisition, phonology in particular, discussed by Major and others, also provided some insight for what the internalization of the phonology might resemble as a multi-staged process in second language, but a practical strategy for aggressively addressing pronunciation through vowels was drawn from the research on acoustic differences between vowels and consonants.<sup>6</sup>

---

<sup>5</sup> William Labov, *Sociolinguistic Patterns* (Philadelphia: University of Pennsylvania Press, 1972).

<sup>6</sup> Roy C. Major, "A Model for Interlanguage Phonology," in G. Ioup and S. H. Weinberger, eds., *Interlanguage Phonology: The Acquisition of a Second Language Sound System* (New York: Newbury House/Harper and Row, 1987), 101-125; Susan M. Gass and Larry Selinker, *Second Language*

There are many materials available for teaching pronunciation, but the methods are directed toward the articulatory properties of sounds, the manner and place of articulation and voicing in the production of consonants, and tongue height, tenseness or laxness, and rounding in the production of vowels, presenting an extensive and complicated set of parameters to teach. For this reason, the focus drastically shifted to the acoustic properties of sound, and it was the research in acoustic phonetics and its application in speech synthesis that provided the necessary insights. Sounds that are perceived by the hearer to be more forceful became the dominant theme in the instruction of pronunciation.

The development of artificial speech has led to an expansion of knowledge about speech perception. Speech depends, of course, upon the vocal organs. The lips and tongue play an obvious part in the articulation of speech sounds; however, the vocal organs, or those parts of the body that are involved with speech production also include the lungs, trachea, larynx, (which includes the vocal cords), the pharynx, the nose, the jaw, the mouth, which includes not only the lips and tongue, but the teeth, and the soft and hard palate. All of these features contribute to the production of the sounds, or phonemes, of a language. The process of speech involves the intentional vibration of a stream of air that is segmented into a series of puffs that vary in frequency. The frequency of speech sounds plays a significant role in speech perception.<sup>7</sup>

Aside from the articulatory aspects of speech sound production, the research on speech synthesis has revealed information about how the characteristics of the flow of air

---

*Acquisition: An Introductory Course* (Hillsdale, NJ: Lawrence Erlbaum Associates, Publishers, 1994), 139-189.

<sup>7</sup> Peter B. Denes and Elliot N. Pinson, *The Speech Chain: The Physics and Biology of Spoken Language* (New York: W. H. Freeman and Company, 1993), 17-78.

through the vocal tract influence the perception of individual speech sounds. The research reveals that speech perception involves a multitude of factors including both stored information and the acoustic speech wave. The expectations of a listener influence perception. Thus, the acoustic information is combined with the articulatory features of sounds as well as other linguistic aspects such as syntax and semantics, and the subject matter, and even the circumstantial factors, such as the identity of the speaker and hearer, all contribute to speech perception. It is important to note that acoustic cues are one of many sources of information that contribute to speech perception. Obviously, then, a focus on acoustic cues is just one of many useful tools in second language acquisition.<sup>8</sup>

Vowels and consonants differ at the most elemental level of speech sound production, which is the obstruction or restriction of the stream of air, the series of puffs that differ in frequency, which results in speech. Consonants require obstruction or restriction of airflow while vowels require minimal if any restriction of airflow. Additionally, all vowels are voiced, requiring vibration of the vocal cords. Vowels are also syllabic, meaning that they are peaks of sonority, or loudness, forming the nucleus of a syllable.

Sound spectrograms have provided visual evidence for the strength of vowel sounds. A spectrogram of a vowel shows changes in the peaks of resonance as the shape of the vocal tract changes during the production of a vowel. Such peaks of acoustic energy are called formants, and most vowels have four to five obvious peaks of energy. Peter Denes and Eliot Pinson explain that technology and experimentation with natural and synthetic speech have revealed that although speech perception is based to some degree upon the acoustic features of a sound wave, a listener's expectations, informed by

---

<sup>8</sup> Ibid., 153-183.

knowledge of the language, the speaker, and even the topic, play a strong role identifying speech sounds.<sup>9</sup>

Experiments with combinations of vowels and consonants also showed the effects of coarticulation, which is the overlap of articulatory properties as speech sounds are produced in rapid succession. These experiments demonstrated that the acoustic energy involved in vowel production affects the contiguous consonants. It is also important to note that vowel sounds affect consonant sounds differently depending upon the strength, that is, the amount of acoustic energy present, of the consonant, which means some consonants are more strongly affected by adjacent vowels than others.<sup>10</sup> A discussion of the specific differences caused by coarticulation is beyond the scope of this article, but the knowledge that such variations exist serves to support the importance of teaching the vowel system of the target language in second language instruction programs.

In the articulation of vowels, the passage of the air stream is nearly unobstructed. Traditionally, vowel sounds are classified according to the highest point of the tongue and the position of the lips, notably whether the lips are rounded or unrounded. Vowels are also classified as tense or lax (or in some systems tense and non-tense), depending upon the amount of acoustical energy needed to produce the sound. In terms of articulatory properties, tense vowel sounds require a relatively strong muscular effort involving a greater movement of the vocal tract away from the position of rest. In acoustic terms, a tense vowel involves a strong spread of acoustic energy. Since all vowels are produced with vocal cord vibration, they are among the sounds classified as voiced. Sounds that do not involve vocal cord vibration are referred to as voiceless

---

<sup>9</sup> Ibid., 164-167.

<sup>10</sup> Ibid., 167-174.

sounds. The description of the tongue position is not always easy to understand, partly because the tip, sides and main body of the tongue can move independently. Thus, with respect to vowel production, the tradition is to describe the highest, or lowest, part of the main body of the tongue. Further, the tongue positions are described in relationship to the positions assumed in the cardinal vowels.

Daniel Jones first proposed the set of cardinal vowels. The cardinal vowels represent a set of evenly spaced reference points on a vowel chart, such as the one in Figure 1, in the usual quadrilateral shape. Theoretically, the cardinal vowels are independent of any actual language. They represent spaced points within a continuum of possible qualities. Cardinal vowels are arbitrary, but, in fact, some languages actually have some of these exact vowels. Most real language vowel sounds, however, are

#### Cardinal Vowels

Front	Central	Back
High	• <b>i</b>	• <b>u</b>
Mid-High	• <b>e</b>	• <b>o</b>
Mid-Low	• <b>ɛ</b>	• □
Low	• <b>a</b>	• <b>ɑ</b>

Figure 1. This scheme shows the vowel quadrilateral typically used in mapping the tongue positions for the eight cardinal vowels.

somewhat different in quality from the cardinal vowels. Ladefoged states, for example, that cardinal vowel 1 is produced with the lips spread and the tongue position is as high and far forward as possible without causing friction. He notes that this cardinal vowel is like the English vowel [i], but it is more extreme than the English [i]. However, the

symbol used for this cardinal vowel 1 is still [i], the same as the high front English vowel [i]. Identically transcribed vowels from different languages may actually differ in quality, so the sounds are not actually the same, despite being identically transcribed. Still, the point of reference is useful. The cardinal vowels allow for a description of sounds in reference to specific points within the range of possibilities, allowing phoneticians to describe the sounds as higher or lower than a specific point of reference, forward or backward of a fixed point, or in the middle of the space between front and back or high and low vowels. Thus, a traditional vowel chart provides a visual representation of sound values. Though there are several problems associated with the concept of the cardinal vowels, the system has allowed for the vowels of many languages to be described in relatively accurate acoustic terms despite the focus on tongue height.<sup>11</sup>

The standard vowel chart in Figure 2 with the addition of key words used in this program shows the vowel system of English. Most native speakers are taught that English has the five vowels a, e, i, o, u and often y and w. These are simply the graphemes, or letters, used in orthography, or spelling. All the students who had some familiarity with English had also learned that English has this set of vowels, an especially misleading bit of false knowledge that invites and supports interference from L1 for native speakers of Spanish. In fact, English has 11-15 vowels depending on the dialect while Spanish has the five vowels shown in Figure 3. In a phonemic or phonetic description, a distinct symbol is provided for each vowel. Vowels, then, can be identified and differentiated, and any language can provide known reference points. In teaching any language as a

---

<sup>11</sup> Peter Ladefoged, *A Course in Phonetics*, 3<sup>rd</sup> ed. (Fort Worth: Harcourt Brace College Publishers, 1993), 217-225.



identify individual sounds and sound sequences that are not English. Accented speech is easily recognized even when the accent is subtle. In their introductory text, Fromkin, Rodman and Hyams provide the example of several terms, all nonsense terms that do not exist in English, but as they point out, it is very easy for a native speaker of English to identify “blick” as a possible word, while “bnick” is not.<sup>12</sup> Native speakers know that the voiced bilabial stop [b] in the initial position cannot be immediately followed by the velar nasal stop [ŋ], though they are unlikely to be able to state this information in a rule. Native speakers’ knowledge of the sounds and sequences of sounds provides glosses for similar sounds and sequences in L2. At the same time, native speaker knowledge of the phonology of L1 can also act as a barrier to the acquisition of new sounds and/or sound sequences.

Listener’s expectations play a considerable role in the complex process of speech perception, for example, the first two or three formants in a vowel sound strongly affect the perception of the sound. It is also true, however, that varied combinations of formant frequencies can be perceived as the same vowel.<sup>13</sup> The concept of the cardinal vowels actually demonstrates this reality. The high front vowel [i] in American English is perceptibly different from what may be transcribed phonetically as the same vowel in another language. Phoneticians working with the classification of differences prior to spectrograms and digital processing did an amazing job identifying articulatory properties that actually correspond to acoustic values. This perceptual glitch—perceiving a variety

---

<sup>12</sup> Victoria Fromkin, Robert Rodman, and Nina Hyams, *An Introduction to Language*, 7<sup>th</sup> ed. (Boston: Thomson Heinle, 2003).

<sup>13</sup> Peter B. Denes and Elliot N. Pinson, *The Speech Chain: The Physics and Biology of Spoken Language* (New York: W. H. Freeman and Company, 1993), 17-78.

of acoustic values as the same vowel—suggests in part why it is easier to teach sounds that are not in the L1 inventory and supports, logically, the reason to do so.

The need for correct pronunciation extends far beyond the sociolinguistic context. A native speaker has a vast amount of unconscious stored knowledge that interacts with the acoustic cues as well as knowledge of the topic, the speaker, essentially the whole pragmatic context. The stored unconscious knowledge of the linguistic system in effect filters the acoustic information. The phonological rules establish expectations for sound sequences. The concept of interference or transfer of linguistic variables from L1 to L2 has a long history, stretching from the investigations of Prague School linguists such as Nikolai Trubetzkoy in the 1930s, to Uriel Weinreich's discussion of interference in the 1950s. At about the same time, the development of the contrastive analysis movement was in progress. Based upon the premise that comparing and contrasting languages could provide specific predictable points of error in second language acquisition, contrastive analysis presumed that transfer of knowledge from the first language could negatively and/or positively affect the learning and production skills in second language acquisition. Clearly, this idea is not unique or new to contrastive analysis. Since second language is obviously filtered through the first language, one might expect that the influence of the first language could support second language learning in areas of structural similarities or thwart language learning in areas of dissimilarity. For the native speaker of Spanish learning English, for example, the common pronunciation error of [ʃip], which sounds like "sheep" for ship is not caused by a transfer of knowledge; rather, there is a systematic difference in the sound systems of Spanish and English. Spanish does not

contain the vowel [I], so the perceptual capacity for [I] is simply not available to the native speaker of Spanish.

Production of speech sounds is, of course, related to the perception of sounds. It is widely known that infants are able to make fine phonetic distinctions. However, as children grow older, they lose their ability to differentiate a large number of sounds, distinguishing instead the sound categories characteristic of the language, or languages, they are acquiring. The perceptual ability to distinguish fine phonetic detail significantly diminishes by adulthood. Thus, the L2 learner has an inventory of sounds based on the first language, which provides a kind of map of phonetic categories. Some of those categories may exist in the L2 language, but others may not. Consequently, the L2 learner's pronunciation is manifested in part by the learner's perception of sounds in the first language.

The research on phonological development in L2 learners pointed to several problems that are likely to arise. Flege's research on phonological development revealed that the sounds in L1 and L2 that are similar present the greatest possibility for interference, which is probably not the likely assumption.<sup>14</sup> However, since the cardinal vowels are essentially guide posts for identifying, for example, a variety of sounds that fall within the framework for the symbol [i], this phenomenon makes perfect sense, supporting the notion that novel sounds in L2 are less likely to provide a context for transfer or interference from the L1 phonological system.

The problem of perception in the phonological aspect of second language learning presents an array of possibilities. For example, a sound or sound combination in the L1

---

<sup>14</sup> J. E. Flege, "Second Language Speech Learning: Theory, Findings, and Problem." in *Speech Perception and Linguistic Experience: Issues in Cross-Linguistic Research*, ed. W. Strange (Timonium, MD: York Press, 1995), 233-277.

may also be present in the L2 language, in which case the learner can articulate the sound with the same or very similar proficiency as a native speaker. It is also possible that a sound or a combination of sounds in the L2 language may not be present in the L1 language, and consequently the learner may not perceive the sound or sound combination and will be unlikely to articulate the sound at all. Still another very likely possibility is that the absence of perception of the sound will be overlooked in an instructional setting because the actual production of the sound is the focus. It is difficult to assess whether the mispronunciation is related to production or perception, even with considerable knowledge of the phonology of L1 and L2. With no knowledge of phonology, it is unlikely that an instructor would even consider the perception-production link. Often L1 sequential constraints affect production as well so that another possibility is the correct pronunciation in a phonetic environment where L1 has the same sequential constraints and mispronunciation in a sequence of sounds not present in L1. Sometimes, students' pronunciation may be a matter of luck, which can be frustrating for an instructor who may not know the phonology of a learner's first language.

English has an extensive system of vowels, particularly by comparison with Spanish. Thus, it is very likely that many adult native speakers of Spanish learning English perceive fewer than half the vowels of American English. However, despite the normal reduction in phonological discriminatory abilities, adults can still be taught to perceive new categories of sounds.<sup>15</sup>

---

<sup>15</sup> Roy C. Majors, *Foreign Accent: The Ontogeny and Phylogeny of Second Language Phonology* (New Jersey: Lawrence Erlbaum Associates, Publishers, 2001), 52-62.

A number of studies of vowel perception demonstrate that vowels are perceived in relation to the categories extant in the first language.<sup>16</sup> However, without training, few native speakers of any language can identify the articulatory features of the sounds of their language, vowels or consonants. Identifying or placing the vowels of L1 and L2 in relation to the cardinal vowels on a chart provides a visual as well as auditory and articulatory context for exploring the similarities and differences between and among the vowels in L1 and L2 languages, allowing the learner to both see and hear the different sound values in the two languages.

In *A Course in Phonetics*, Ladefoged describes several practical exercises for locating vowels. He proposes these exercises primarily as a way to demonstrate that vowel sounds form a continuum. Unlike consonants, there are no clear boundaries between vowel sounds. Categories of consonants are more distinctive, not allowing, for example, a sound that is halfway between voiced and voiceless. Ladefoged suggests beginning with the vowel [æ] as in “bad.” Native speakers can move through the following front vowels [æ, ε, e, i] ([ε] as in bet, [e] as in bait, [i] as in beet) though he notes that passing through these sounds from [æ] to [i] may require some practice, stating each sound, “slurring” from one vowel to another.<sup>17</sup> The next step is to reverse the process, moving from [i] to [æ]. To feel the difference from front to back he suggests moving from [æ] to [a] as in “father.” Some phonetic charts use the key word “bomb” as speakers of American English would pronounce it, but the students persisted in rounding the vowel, coming closer to [o]. The word “father” produced the best results. Native speakers of American English are likely to go directly from [æ] to [a], though there are

---

<sup>16</sup> Ibid., 52-62.

<sup>17</sup> Peter Ladefoged, *A Course in Phonetics*, 3<sup>rd</sup> ed. (Fort Worth: Harcourt Brace College Publishers, 1993), 78.

possible vowels between [æ] and [a]. Many European students who have learned the British English dialect stop almost midway between [æ] and [a]. Finally, moving from [a] up to [u] reveals the range of back vowels [a, ɒ, o, u, ʊ] ([a] as in father, [ɒ] as in fore, [o] as in boat, [ʊ] as in put, and [u] as in boot). The mid central vowel [ə] as in “Rosa” occurs in unstressed syllables, and is, in fact, the most used vowel in English, a useful thing to know for foreign speakers, particularly in light of our orthography. The vowel [ʌ] as in “bud” is sometimes difficult to distinguish from [ə], even for native speakers. However, [ʌ] is more likely to occur in a syllable where some stress occurs. It is important to note that the vowel qualities high, low, front, back, and mid and central are actually auditory values showing how vowels sound in relation to each other. Thus, the auditory quality of the sounds changes as there is movement from one vowel to another. These features correspond to acoustic properties rather than to actual tongue position.

Vowel sounds represent peaks of sonority, or the inherent loudness of a sound. If the pitch, stress, and length are equal, for example, the vowel sound [a] as in barn is more sonorous than [i] as in bean. Likewise, both vowel sounds are more sonorous than the consonant sounds in the two words.<sup>18</sup> This quality is evident in the context of a syllable. Vowels or diphthongs form the nuclei of syllables; that is, vowels are usually the central element in a syllable. A syllable consists of a single sound formed by a vowel or a diphthong alone, or sometimes a syllabic consonant alone, or a combination of a vowel or diphthong with one or more consonants. A syllable is a phonological unit that contains a vowel as its nucleus.<sup>19</sup> Students sometimes use “syllable” and “morpheme” interchangeably, which is inaccurate. Though it is true that some syllables are actually

---

<sup>18</sup> Peter H. Matthews, *The Concise Oxford Dictionary of Linguistics* (Oxford: Oxford University Press, 1997), 346.

<sup>19</sup> Mehmet Yavaş, *Applied English Phonology* (Malden, MA: Blackwell Publishing, 2006), 127-149.

morphemes, and have meaning, it is an error to assume that all syllables are also morphemes. Syllables are phonological units. The issue of word structure, or morphology, is not a focus of this article, but for the sake of clarity it is necessary to inform students, and perhaps members of an instructional team, of the difference between syllables and morphemes. The concept of a syllable is also important because the coarticulation factor of sound production plays an important role in how a listener actually hears a set of speech sounds.

Coarticulation of sounds is a natural phonetic force. Speakers do not independently articulate individual sounds in a word, or more importantly, in a sentence. Thus, the sequencing of sounds results in assimilations and dropped segments, as in a contraction, for example. The difference between individual sounds is reduced to a minimum. Ease of articulation, which reduces the articulatory effort, is highly dependent upon coarticulation, in which adjacent sound segments are affected by each other. Ease of articulation works as long as listeners are able to easily distinguish the words. Ladefoged notes that articulation is acceptable to listeners as long as there is “sufficient perceptual separation.”<sup>20</sup> The sounds that affect the meaning of a word need to be distinct from each other.

The desired outcome of training EFL students in the differing values of vowels is improvement in pronunciation skills. Once students have acquired the perceptual categories, that is, they are able to perceive the vowels of the target language, they can begin the task of adding these sound values to their own inventory of sounds, thus lessening to some degree the influence of their first language in the production of the

---

<sup>20</sup> Peter Ladefoged, *A Course in Phonetics*, 3<sup>rd</sup> ed. (Fort Worth: Harcourt Brace College Publishers, 1993), 268.

vowel sounds of the target language. It is important to note, however, that there may be differences in those sound qualities. For example, in the discussion of the cardinal vowels, it was noted that the cardinal vowel [i] actually represents a range of sounds that are high front sounds. Thus, even once students perceive all the vowel sounds, they may or may not pronounce them exactly as a native speaker, but at least they will be less likely, with substantial practice, to make substitutions such as [i] for [ɪ], creating errors such as [ʃip], or “sheep” in orthographic form, for [ʃɪp], or “ship.” An important additional benefit can also occur because sounds affect each other in phonetic environments. Consequently, because of the coarticulation factor, some consonants will be heard more accurately when the vowel is pronounced in a manner consistent with listeners’ expectations. Clarification of the vowel system can lead to clarification of other sounds as a natural consequence.

The students reported that training in the vowel system strengthened their listening skills, allowing for better comprehension of both individual words and sentences. Students were easily able to master aural and oral word articulation exercises. They showed significant and rapid improvement in both auditory and articulatory tasks that involved vowel contrasts in the context of word pairs, averaging 90 percent accuracy in comprehension. Similar performance in sentence level comprehension and in protracted discourse predictably took much longer because syntactic and lexical skills had such an important role, and the students were still learning syntax and acquiring vocabulary. In accordance with the typical performance in all language acquisition, comprehension skills were stronger than production skills. In reference to hearing vowel sounds, “perceptual” is actually a more accurate descriptor than “comprehension.” Just

as comprehension precedes production, perception of the sounds is a prerequisite in the comprehension of individual words or even sentences.

In service and other activities relating to their everyday life living in a seminary, students were frequently required to recite prayers or to read liturgical texts aloud, providing a set of pragmatically oriented samples of English to use for pronunciation practice. It is important to note that such text was written in language that was often archaic. It was suitable for pronunciation practice, but it evoked many questions that required explanation of historical change in the English language. However, it was of paramount importance to include such samples for this particular group of students. Initially, in preparing for oral presentations, students transcribed only the vowel sounds in prayers and various required recitations. In reading practices, students mastered the vowel sounds and could deliver readings with clearly distinguishable vowels. However, in free speech immediately after reading, pronunciation was still likely to contain some errors caused by vowel substitutions.

Students were also asked to do prepared talks, delivering five and ten minute presentations on topics relating to moral or spiritual issues, much as they might have to do in seminary life. Again, vowels were transcribed in students' prepared notes, but the students were required to speak rather than to read. These exercises were conducted in large rooms from a lectern, sometimes with an audience of instructors and tutors seated in the back of the room using a complicated rating system to provide feedback to students on areas of pronunciation that needed improvement. However, in this context, many factors affected the auditory quality of the speeches. Also, the audience's familiarity with the speech patterns of the students was a factor that could not be eliminated. Predictably,

some of the students were introverted and some extroverted, which affected performance. Some spoke loudly most of the time, while others always spoke softly. The detailed rating system required listeners to check specific errors such as consonant and vowel substitutions, stress and rhythm, and intonation, but it failed miserably. It was perceived by listeners as too complicated. A holistic 1-10 scoring procedure worked better for listeners, but did not provide significant feedback on phonetic detail. However, as listeners' holistic perception of students' performance rose from low scores of two or three to seven, eight and beyond, both the speakers and listeners could acknowledge change and improvement. The one team member with expertise in phonology provided individual assessments with specific phonetic indications for follow-up practice to improve performance for each student.

The number of students in this pilot program is too small to provide statistically significant outcomes. However, the most significant and telling outcome was the acceptability of the students' English skills at the end of the program as judged by the members of the community the students had hoped to enter. All the students except one completed the intensive English program. All students were given the opportunity to pursue an undergraduate program of study, if needed, and all students were judged acceptable in terms of English skills. As might be expected, however, not all the students in the English program actually entered the seminary. Of the students who entered the seminary, one student needed further individual instruction and practice in pronunciation, particularly in the stress system.

The results in this special program led to the institution of a phonetics course for the regular international student population that attends the college. The focus in the

course is the vowel sounds of English. The phonological effects of ethnic influences that have provided some pronunciations distinctive to the geographic area are also addressed. The target outcome in this course is also clarity of speech, not the elimination of an accent. The population in this course consists of undergraduate and graduate students, primarily native speakers of Japanese but often including native speakers of French, from France, Quebec, and Morocco as well as other Middle Eastern and African countries; Spanish from Spain or Mexico; Polish, Ukrainian, Russian, German, Arabic, Thai, Chinese and Korean. Occasionally, native speakers of English are also enrolled in this course to minimize dialect differences that impede communication.

Students report that intensive training in vowel perception and production is effective in helping them to understand classroom interaction and to become more confident in their ability to engage in conversation with native speakers. Since the majority of international students in the institution whose native language is not English are study abroad exchange students or MBA students, rather than matriculated undergraduate students, and spend only a semester or a year in the United States, the faster they become integrated within the student body in particular, the more fulfilling their experience.

#### Conclusion

It is important to note that any one factor or procedure that may change the rate of the second language acquisition process and/or the level of proficiency does not suggest in any way how second languages are learned. Too often pedagogical approaches emerge, based on relatively narrow assumptions about the nature of language or the nature of learning in general. Shifts in theory result in new trends, which are often

conventionalized and become, in effect, a rigid template for instruction. A template, by nature, obscures at least a part of the total entity. There is not yet a complete and coherent theory that presents a unified explanation of second language acquisition. Chomsky's well known observation was that ultimately abstract mental representations that generate language are fundamental to its very nature.<sup>21</sup> What is important here in this approach to teaching pronunciation is confirmation of research outcomes through practical application of the principles discovered. First, perception of speech sounds is fundamental to all oral language acquisition. Second, the L1 sound categories affect the development of the L2 phonological system in both positive and negative ways. Third, the perception and production of sounds novel to the L2 learner are easier to achieve than the perception and production of L2 sounds that are very similar to L1 sounds.

It is clear that the L2 learner, regardless of age, can be taught to discriminate the acoustic boundaries of sounds, and since that is so, then providing specific instruction in the perception and production of the sounds of the target language provides one more piece in the complex interlocking puzzle that confronts every L2 learner, particularly the post-pubescent learner who does not possess the same facility for language learning that a young child does. As a matter of policy, rudimentary knowledge of phonetics should be included in English as a second or foreign language teacher education programs.

Finally, the value of providing detailed instruction and intense practice in the perception and production of vowels is actually not so much in the positive outcomes experienced among two small groups of students as in the fact that the whole process, from the initial recognition of challenges to the end result of the acquisition of English, reflects the necessary interaction between theory and practice. The information provided

---

<sup>21</sup> Noam Chomsky, *Rules and Representations* (Oxford: Basil Blackwell, 1980).

by researchers in the physical, biological, social, and psychological aspects of second language acquisition established a foundation for interpretation through reasoned application. A continuous iterative loop in which theory informs practice which informs theory is the cornerstone of science. Intrusive instruction in vowel sounds supports second language acquisition, but the answer to the question of exactly how a learner acquires new acoustic boundaries that result in the requisite mental representations will advance our knowledge and ultimately further inform and reform practice in second language acquisition instruction.

## References

- Brown, G. and G. Yule. 1983. *Discourse Analysis* New York: Cambridge University Press.
- Chomsky, N. 1980. *Rules and Representations*. Oxford: Basil Blackwell.
- Denes, P. and E. Pinson. 1993. *The Speech Chain*, 2<sup>nd</sup> ed. New York: W. H. Freeman and Company.
- Flege, J. E. 1987. A critical period for learning to pronounce foreign languages? *Linguistics* 8: 162-77.
- Flege, J. E. 1995. Second language speech learning: theory, findings, and problems. In W. Strange, ed., *Speech Perception and Linguistic Experience: Issues in Cross Linguistic Research*. Timonium, MD: York Press, 233-277.
- Fromkin, V. and R. Rodman, N. Hyams. 2003. *An Introduction to Language*, 7th ed. Boston: Thomson/Heinle.
- Gass, S. M. and L. Selinker. 1994. *Second Language Acquisition: An Introductory Course*. Hillsdale, NJ: Lawrence Erlbaum Associates, Publishers.
- Labov, W. 1972. *Sociolinguistic Patterns*. Philadelphia: University of Pennsylvania Press.
- Ladefoged, P. 1993. *A Course in Phonetics*, 3<sup>rd</sup> ed. Forth Worth: Harcourt Brace College Publishers.
- Major, R. 1987. *A model of interlanguage phonology*, in G. Ioup and S. Weinberger, eds., *Interlanguage Phonology: The Acquisition of a Second Language Sound System*. New York: Newbury House/Harper Row, 101-124.
- Major, R. 2001. *Foreign Accent: The Ontogeny and Phylogeny of Second Language Acquisition*. Mahwah, NJ: Lawrence Erlbaum Associates, Publishers.
- Matthews, P. H. 1997. *The Concise Oxford Dictionary of Linguistics*. Oxford: Oxford University Press.
- Morley, Joan. 1996. Second language speech/pronunciation, acquisition, instruction, standards, variation, and accent, in James E. Alatis, Carolyn A. Strahle, Maggie Ronkin, and Brent Gallenbergr, eds.

*Georgetown University Round Table on Languages and Linguistics.* Washington, D.C.:  
Georgetown University Press, 140-160.

Yavaş, M. 2006. *Applied English Phonology*. Malden, MA: Blackwell Publishing.

Published by the Forum on Public Policy

Copyright © The Forum on Public Policy. All Rights Reserved. 2006.