A Multisensory, Multicognitive Approach to Teaching Pronunciation

Edward Y. Odisho
e-odisho@neiu.edu
Northeastern Illinois University (USA)

ABSTRACT. The Multisensory, Multicognitive Approach (MMA) introduced in this paper is premised on the belief that the seat of language is in the brain prior to its physical manifestation in the form of speech being in the mouth. Hence, in teaching pronunciation, the identity of speech as a cognitive entity prior to being a physical one should be seriously considered in relevant language learning and teaching situations – more so in L2 situations than in L1. The traditional assumption that the ideal and the exclusive sensory modality of teaching pronunciation is the auditory modality is no longer acceptable because a holistic view of speech – in production, transmission and perception – manifests itself not only via the auditory sensory modality, but also equally significantly via the visual and tactile-kinesthetic sensory modalities. It is due to this fact that MMA is described as multisensory, a fact that determines the diversified auditory, visual and tactile-kinesthetic implementational techniques needed for effective and efficient teaching of pronunciation especially to adults. Equally importantly, the multicognitive nature of MMA requires the manipulation of diversified cognitive processes in the form of thinking, associating, analyzing, synthesizing, comparing, contrasting etc… for implementation. According to MMA, the teaching of pronunciation becomes more of a multi-faceted educational process than a mere repeat-after-me mechanical parroting of speech sounds. Such an approach requires more effort on the part of the instructor and learner and a stronger collaboration between them through the diversification of teaching and learning styles, respectively. Certainly, MMA requires more time to implement in classroom situations, but the time spent is worth it. MMA is no longer a single technique or drill that tackles one sound at a time; instead, it is a joint selection of cognitive and sensory techniques that are applied concurrently to facilitate the L2/FL mastery in a creative and generative manner similar to the process of child language acquisition.

KEY-WORDS: Teaching Pronunciation, MMA, L2, language learning, language acquisition.

1 – Outline of Principles and Techniques of MMA

The multisensory, multicognitive approach (henceforth, MMA) introduced in this paper has gradually evolved in this writer's classes during the last decade. In 2003, a more formal set of principles and techniques, on which the approach is premised, was published in the form of a book titled: *Techniques of teaching pronunciation in ESL, bilingual and foreign language classes* (Odisho 2003). MMA was further developed and illustrated in additional publications (Odisho 2004; 2005). The approach incorporates some of the latest theoretical and applied principles in linguistics coupled with input from psycholinguistics. The philosophical premise of this approach is that schools of knowledge do not always tend to be mutually exclusive in every respect; rather, in many other respects they may be complementary in mission and functions. To illustrate, if behaviorism promotes associative habit formation and structural linguistics implements that through repeated and regimented drills in language learning as is the case with the audio-lingual method, it should not imply that those practices of behaviorist psychology and structural linguistics are absolutely useless and incompatible with principles and practices promoted by cognitive psychology and transformational-generative linguistics with regard to human language acquisition. MMA strongly believes that for proper internalization of human language, associative habit-formation and drilling are still much needed, but they cannot be efficient and effective practices in reflecting the generative and creative nature of human language acquisition without the transformation of the physical habits into cognitive habits through their processing in the brain and retention in long-term memory for automatic and at will retrieval.

MMA is an attempt at duplicating the natural manner in which a normal child acquires language. Foremost among the salient features of child language acquisition are ample exposure to and rehearsal of language materials in authentic contexts in a holistic manner using as many sensory and cognitive modalities to necessary to acquire all the skills and subskills of which human language is constructed. Thus, children who grow up physically, cognitively and affectively under normal conditions, will have the benefit of the above experiences leading to a natural process of language internalization that is subconscious, automatic and effortless; it is this type of language internalization

that is referred to as acquisition. With age, adults begin to slowly lose their adeptness in the automatic and subconscious internalization of pronunciation. Consequently, the process of mastering the pronunciation of a second language (L2) or foreign language (FL) becomes increasingly more conscious, mechanical and effortful. It is here where the approach to teaching adults sets itself apart from the approach to teaching children. This approach is more oriented toward teaching L2/FL to adults than to young children. In the context of MMA, the distinction between the adeptness of children and adults to language acquisition is confined to the skill of pronunciation and not necessarily to other skills, such as morphology, syntax and lexicon in which adults may be equally adept or even more adept than children.

In the available literature, failure of adults to further improve their mastery of L2 beyond a certain limit has been named fossilization (Selinker 1972: 209-31). MMA rejects the term fossilization because it is too rigid of a process to describe a normally functioning brain. The rejection is justified based on the fact that systematic multisensory and multicognitive orientation helps all learners, regardless of age and aptitude for pronunciation, to improve their skills in the acquisition/learning of L2/FL pronunciation to different extents. Using a combination of diversified multisensory and multicognitive exercises and techniques, the learning process can continue, albeit slowly, but it will hardly cease completely as fossilization claims. In contrast, the slowness or provisional resistance in the acquisition/learning of L2 pronunciation by adults will be known as psycholinguistic deafness which does not imply total cessation of learning; rather, it keeps the doors of acquisition/learning of L2 pronunciation open depending on the approach to teaching it.

In the following sections, the multisensory and multicognitive nature MMA will be demonstrated. Teaching with a multisensory approach means the instructor has to approach the learner via more than one sensory modality and the learner should be prepared and encouraged to behave likewise. Similarly, the multicognitive aspect of the approach should encourage learners to try to attentively listen to sounds, retain an acoustic and auditory image of them and compare and contrast them with sounds already part of their psycholinguistic inventory using all the available cognitive processes such association, analysis, synthesis,

comparison, contrast, memorization, etc. MMA also calls for a move beyond the traditional understanding of the skill of pronunciation as the sole function of the auditory sensory modality and that its teaching is uniquely effected through 'ear training' as schematically illustrated in Figure 1 below:

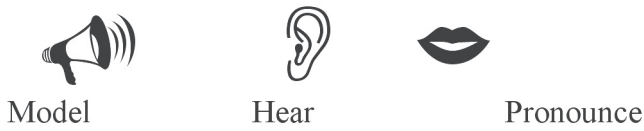


Figure 1 – Traditional monosensory teaching of pronunciation with exclusive reliance on the auditory sensory modality

The auditory sensory modality and ear-training alone often fail to teach pronunciation to adults because pronunciation as a linguistic skill is also conveyed through the visual sensory modality and the tactile-kinesthetic sensory modality. Consequently, ear-training should be supplemented by what is to be known, hereafter, as: a) eye-training (i.e., visual orientation of pronunciation through seeing and visualizing sound production and the accompanying dynamics of body and facial gestures); b) neuro-muscular training (i.e., tactile orientation or how to kinesthetically and proprioceptively sense and feel sound production and its dynamics; and c) brain training (cognitive orientation or how to psycholinguistically and consciously perceive, recognize, retrieve and produce the sounds and their underlying dynamics). This multisensory approach to teaching is schematically illustrated in Figure 2 below:

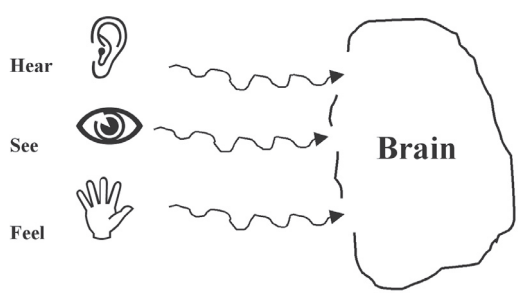


Figure 2 – Multisensory teaching of pronunciation with joint reliance on the auditory, visual and kinesthetic sensory modalities.

The triangular sensory modalities feed the brain with diversified input to reinforce the cognitive processing, internalization and retention of new sounds in long-term memory ready for instantaneous retrieval in the right context when needed.

2 – Elaboration on Principles and Techniques of MMA

MMA is based on a set of principles that serve as guidelines for further explication of the approach and the design and development of the techniques for its application. The following are the most relevant principles each of which will be demonstrated to the extent to which it is relevant to the implementation of MMA.

2.1 – Speech: A Cognitive Phenomenon

Human language is a code of communication that is a genetically determined cognitive potential before being a set of physical maneuvers which serve to activate the cognitive potential and reinforce it. Stated differently, language is in the brain before being in the mouth. This fact is a major premise on which the approach is developed. The instructor will often see that adults may experience serious difficulty in producing a new sound or sound unit to which they have never been exposed. This is a good example of the cognitive requirement for sound production meaning that the brain may need enough exposure time to the new sound to perceive and recognize it before being able to produce it appropriately. Therefore, any instruction in pronunciation should target both the cognitive potential for perception and recognition prior to the necessary physical maneuvers of production. If, for instance, an adult native speaker of English is asked to produce an unfamiliar sound which is not part of Standard English phonology, such the Arabic voiceless unaspirated uvular plosive [q][1] as in " قلب " /qalb/ (heart), or German voiceless uvular fricative [χ] as in "acht" /aχt/ (eight) or Spanish *trilled* [r] as in "perro" /pero/ (dog), and he/she, after continuous modeling by the instructor, fails to properly produce those sounds and persists in replacing the Arabic [q] with a [k] or [k^h], the German [χ] with a [k^h] and the Spanish [r] with a typical English or American *frictionless*

[1]For a thorough examination of this sound see Odisho (1977).

continuant (*approximant*) [ɹ] or [ɻ], then the whole situation indicates that the learner is psycholinguistically [*cognitively*] unable to perceive and recognize those L2 sounds, hence unable to produce them. This is a typical condition that is identified in this study as *psycholinguistic deafness* (as a substitute for fossilization); a condition that is characteristic of adults learning L2/FL. Psycholinguistic deafness in the teaching of pronunciation cannot be remedied without an approach and sets of techniques that enable the brain to cognitively perceive and recognize the new sounds and then fire the commands to the vocal organs to embark on a period of trial and error in executing the articulatory maneuvers needed for the production of the targeted sounds.

2.2 – Multisensory Approach to Speech & Pronunciation

In handling the skill of pronunciation, its exclusive association with the auditory sensory modality is a pervasive traditional bias. Pronunciation is not a solely audio-lingual activity; rather, it is the function of a much broader base of sensory and physical activities. An integration of auditory, visual and tactile/kinesthetic sensory modalities is indispensable for the proper mastery of speech, in general, and pronunciation, in particular. Serious consideration, for instance, should be given to the visual sensory modality in the form of facial and body gestures that are intertwined with the overall dynamics of speech production. It is because of this interface between more than one sensory modality that MMA is identified as *multisensory*. Hence, teachers and learners have to be prepared not just to hear and produce the sounds, but also, and equally importantly, to see and feel the sound in conjunction with the concomitant sensations and physical gestures in the context of authentic speech. In light of this principle, a certain category of consonantal sounds, such as the bilabial, labio-dental, interdentals and dentals should squarely be identified as *visible* sounds. Additionally, many vowel sounds are better taught and learned by visually monitoring the lip configurations rather than by repeating after a model only. It, therefore, hurts to see Spanish-speaking students, who have taken several courses in English, still struggling to distinguish minimal pairs such “ballet” vs. “valet”, “bowel” vs. “vowel”, “boat” vs. “vote”, among others, because of their failure to produce a [v] sound, which according to

MMA and the techniques implemented, the contrast between [v] and [b] is mastered in one session of no longer than an hour.

2.3 – Multicognitive Approach to Speech & Pronunciation

Due to the significant role of the brain in the acquisition/teaching of pronunciation, the need for the activation of the cognitive processes is inevitable. Learners have to be encouraged to try to attentively listen to sounds, retain them at least in their short-term memory and compare and contrast them with sounds that are already part of their psycholinguistic inventory. Notice that in teaching pronunciation, the emphasis should be on retention of the acoustic/auditory images of new sounds or sound phenomena rather than pursuing the route of mechanical imitation simply because the latter is the result of a monocognitive process, whereas retention is the result of a combination of multicognitive processes such as thinking, association, analysis, synthesis, comparison, contrast as well as memorization. The practice of thinking about sound production and its dynamics may proceed in the following manner: Why didn't I succeed in producing the sound at first trial? Why was my second attempt better than the first? What did I do differently? Why was his/her production more acceptable than mine? Where was the tip of my tongue when I tried? What will happen if I move my tongue slightly backward or forward? Is the tip of my tongue curled? Although these cognitive activities may sound too abstract for some teachers even to know about them, but in reality they do exist and their presence can be felt in different ways. Often when an instructor models a certain sound and then allows for a break before the reproduction session, many of the learners are already engaged in thinking of the reproduction. You can readily infer the thinking process through the facial and bodily gestures of the learners. For instance, you can easily see a learner moving his/her tongue inside the oral cavity to feel the place of articulation or to try to create a rounded configuration for the lips, or even to depress or elevate the jaw to secure the targeted degree of oral opening. These movements and gestures are all reflections of inner and mute endeavors on the part of learners to master the dynamics of the targeted sounds and reproduce them.

2.4 – Complementary Nature of Acquisition & Learning of Pronunciation

Teaching pronunciation should distinguish between the processes of acquisition and learning. Acquisition tends to be a subconscious, automatic and effortless process of internalizing a sound system, whereas learning tends to be more conscious, mechanical and effortful. The former tends to be primarily characteristic of normal children's mastery of the pronunciation of their L1 or even a given L2, whereas the latter tends to be primarily associated with the manner in which adults master pronunciation. Despite the difference between the two processes, acquisition and learning are *not mutually* exclusive in nature and function. Their nature and function are complementary and they depend on the age of the learners, extent of exposure and the conditions of exposure to the linguistic materials and the level of motivation. Generally speaking, research as well as life experience adduce ample evidence in the direction of more acquisition than learning in the case of children as opposed to more learning than acquisition in the case of adults. Hence, in the description of language internalization by children, the appropriate compound verb would be '*acquire-learn*' and the reversed order '*learn-acquire*' would be more appropriate for adults. However, the above two orientations in language/speech internalization should not, in any way, imply that adults are unable to attain a near-native or even native pronunciation. No doubt, those adults who have some degree of linguistic aptitude and a gift for language internalization will tend to handle languages with an '*acquire-learn*' strategy similar to children. However, even those adults who do not entertain a linguistic aptitude may easily enhance and sharpen their learning skills regardless of age if the conditions and techniques of learning/teaching are conducive enough to motivate them and activate all the sensory and cognitive processes needed for acquisition.

Consequently, the mere exposure to language materials through the traditional *repeat-after-me* technique may be far more functional and effective with children than with adults. With the latter, the mere exposure is not sufficient and, oftentimes, the above technique turns out to be useless because adults tend to repeat after themselves. In other words, adults may reproduce L2 or FL articulations in terms of their L1 or what is called the *interlanguage*. A considerably different approach

and far more different and diversified sets of techniques should be implemented with adults in L2 and FL language learning situations.

2.5 – Triangular Base of Pronunciation: Perception, Recognition & Production

Any teaching of pronunciation should thoroughly follow the three-stage procedure of sound acquisition, namely *perception, recognition* and *production* in the sequence indicated. The above triangular procedure is highly consistent with the three-stage procedure of *registration*, *retention* and *retrieval* in learning and with the three types of memories of *sensory*, *short-term* and *long-term* in which information is stored. In each case, the earlier stage serves as the gateway to the next and final stage. The transition to the final stage cannot be completed without continued rehearsal. Because the perception, recognition and production procedure plays a significant role in MMA, a brief clarification of the terminology is invaluable. *Perception* is used to denote the condition of feeling and sensing the presence of a given sound; *recognition* includes the condition of perception as well as the condition of being able to distinguish the given sound from others. As for *production*, it satisfies the above two conditions of perception and recognition in addition to the ability to retrieve the sound and reproduce it at will with different acceptable degrees of proficiency and accuracy.

The explanation above suffices to portray the functional and operational parallelism across the processes of *sound acquisition, general learning and memory* and the sequential stages through which they usually go. For instance, in order to perceive a sound one has to be exposed to it at least through the sensory memory; to have it registered, at least temporarily, it should be stored in the short memory; however, in order to retrieve and produce a sound, it has to be retained and consolidated in the long-term memory through rehearsal. Sequencing of stages is significant and bypassing a stage may negatively impact the outcomes. For instance, with casual and improper exposure to unfamiliar sounds, it is highly unlikely to succeed in producing them. A serious flaw in the traditional approach to the teaching of pronunciation is attributed to either insufficient dwelling on the perception and recognition stages or their total negligence. Those two conditions lead to an immediate jump to the production stage, a condition that is typically embodied in

the '*repeat-after-me*' technique of teaching pronunciation which may be so incompatible with the learning styles of adults

2.6 – Pronunciation & Feedback Mechanisms

The production of speech requires the simultaneous and coordinated use of respiratory, phonatory and articulatory mechanisms. The physical, aerodynamic and acoustic dynamics, movements and perturbations that result from the action of the mechanisms often yield multifarious sets of internal sensations of touch, pressure, movement, position etc., which constitute the kinesthetic and proprioceptive feedback control systems. For a more succinct summary of the nature and the function of those sensations see Daniloff (1973: 183). The important instructional fact that emerges as a result of the emphasis on diversified speech production feedback systems is that the auditory feedback system, should not exclusively dominate the approach to teaching pronunciation and that all types of feedback mechanisms, especially tactile/kinesthetic should be brought into play jointly in the form of different pronunciation teaching and learning techniques and activities.

2.7 – Phonetic & Phonological Aspects of Pronunciation

The natural acquisition of speech begins in a physical manner in the sense that children start dealing with sounds in the early stages by perceiving and recognizing them as physical experiences. Such an approach is a general phonetic one with no serious regard to the contrastive power of sounds that trigger semantic differences. A phonetic perception and recognition of sounds requires paying attention to all the features underlying those sounds. But by the nature of human language as a highly economic code of communication the focusing of attention on all the detailed features of the code (distinctive and non-distinctive) renders the code *uneconomical* and vulnerable to confusion. Gradually, children discover and identify the most distinctive features of each sound, abstract them and designate them for internalization as part of their phonological inventory. This exclusive abstraction of only the distinctive features of sounds is the reason behind the development of phonological habits of perception and recognition that are L1 bias. In other words, the speaker/listener automatically, and very subconsciously, perceives/recognizes the distinctive features and fails to do so with the

non-distinctive ones. In Werker's words, learners attend to only phonetic information that distinguishes meaning (Werker 1995: 99). The older a speaker-listener grows and the more he/she practices the native speech, the more phonologized the process of sound internalization becomes. It is those sound generalizations stored in long-term memory that are known as *phonemes*. This is why Werker assumes that if adults sometimes have difficulty discriminating nonnative phonemic contrasts, there must be, then, a decline across age in cross-language speech perception (Werker 1995: 89). Once the native sounds are phonologized, they are automatically and instantaneously perceived, recognized and produced. Automatic processes are the outcome of systematic rehearsal and the more they are rehearsed the less they require attention (Anderson 1980: 30). Some treat speech memory as motor memory and that we experience such speech memory when we try to learn a foreign language later in life. Because we have no motor memory traces available for the articulation of the foreign language sounds, we find ourselves almost unable[2] to form the vowels and consonants as they are pronounced in that language (Arnold 1984: 41-42). In a way, the exclusive focus of children on the internalization of their L1 phonological system results in promoting a bias to the sound units that are part of the system and hence become less sensitive to sounds which are not part of the system. Stated differently, the focus on the mastery of L1 phonology leads to the evolution of what could be called the 'protective shield' or the 'protective radar' to guard against the L2 phonologies which in this case amount to 'alien phonologies'. In order to enable those 'alien phonologies' to co-exist with the native phonology, the former have to be tamed and internalized through diversified sensory and cognitive modalities.

In light of the above explanation, the transformation from phonetic perception/recognition to a phonological one in L1 may be a major reason for the failure of adults in the perception/recognition of L2/FL sound contrasts which are absent in their L1 system. Consequently, any

[2] Generally, the above quotation is plausible, but the approach proposed here would replace the attribute '*unable*' with 'less able' since the main goal of MMA is to upgrade the level of learnability by a joint set of sensory and cognitive teaching techniques.

approach to teaching pronunciation in L2/FL situations should develop techniques that sharpen the sensitivity of learners, especially adults, to the perception/recognition of L2 phonologies. It is this sharpening of sensitivity that will enable adult learners to bypass the 'protective radar' of their L1 and succeed in perceiving, recognizing and producing the 'alien sounds' of L2.

2.8 – Phonetic and Phonological Accents

In teaching pronunciation, the distinction between sounds along the line of the phonetic and phonological contrast constitutes a major premise for the development of the overall approach to the teaching/learning of pronunciation. MMA places significant emphasis on the distinction between the nature of mispronunciation that occurs in the transition between L1 and L2/FL. If the mispronunciation of a given sound or feature does not trigger a semantic change (meaning) in the targeted language then such mispronunciation will be treated as phonetic and the outcome will be labeled as *phonetic accent* (Odisho 2003 :19). For instance, if a Hispanic learner of English mispronounces the approximant [ɹ, ɻ] of English as the tap or rolled [ɾ, r] of Spanish, the mispronunciation in this particular case is identified as phonetic accent simply because it does not cause any semantic change. If, however, the mispronunciation causes a change in meaning then it will be treated as *phonological accent* (Odisho 2003:20). For example, if a native speaker of English embarks on learning Spanish and encounters difficulty in pronouncing and distinguishing the tap [ɾ] and the rolled [r]'s of Spanish as in <pero> "but" vs. <perro> "dog" then the person has a phonological accent. The phonetic failure to distinguish between the two <r>'s results in a failure to signal the semantic difference between the two words. This example is, in fact, a typical case of a combined phonetic and phonological accent.

In any teaching of pronunciation, the priority should be directed to addressing the phonological accent. If, however, the purpose of a given course or the intention of a learner is to overcome all types of pronunciation difficulties and attain a native-like or near native-like proficiency in the pronunciation of L2/FL then both phonetic and phonological accents should be targeted.

2.9 – Pronunciation as a Generative Skill

Obviously, the term 'generative' is associated with Chomsky's theory of linguistics. The term is reused here with a somewhat different meaning though still somewhat related to the Chomskyan one. The generative nature of MMA implies that mastering the perception, recognition and production of one sound should facilitate the mastery of more than that one sound. In other words, developing a skill in one aspect/domain of pronunciation should serve as a key to enhance or generate a skill to master other aspects/domains of pronunciation. For instance, in English, mastering the production of a *schwa* does not only help with the mastery of the complicated vowel system of English, but it will also considerably facilitate the process of stress placement and the overall rhythmic performance. Also, learning how to kinesthetically and proprioceptively sense a tongue tip contact at the alveolar ridge should develop the skill of sensing any other contact of the tongue in the oral cavity. Even in the dynamics of sound production, mastering stress in a given word should pervade to other words and to the overall rhythm mastery in the targeted language or any other language for that matter.

2.10 – Pronunciation and Connection with your Students

Obviously, the instructor according to MMA needs to be conscious of the interactive connection between him/her and the learners. To establish this connection, the instructor should make sure of the following points. *Firstly*, he should make sure that the learners know what the theme/activity under demonstration is about. For instance, if the activity is about stress placement, he should make sure that learners know what stress and stress placement as phonetic phenomena are; never should the instructor assume that learners understand what is going on. For instance, if the instructor is teaching the students that in English certain words can function as both nouns ('contract… 'export…. 'content) and verbs (con'tract… ex'port …con'tent) based on the location of stress he should ascertain that learners can perceive and recognize the physical difference prior to asking them to actually signal the difference physically and master it cognitively. Many learners know this fact about English theoretically, but practically they fail to signal the difference. This author had suffered from this failure for long years until he was

practically oriented in the perception, recognition and production of stress and stress assignment.

3 – Demonstrating MMA Application

Due to limited space only three examples will be afforded for the application of MMA in learning situations related primarily to Hispanic learners of English.[3] The first two examples will deal with segmental sounds – a case for consonants in the form of [v] vs. [b], and a case for vowels in the form of lax vs. tense vowels. The third example will tackle the suprasegmental feature of stress assignment and stress perception, recognition and production for any learner of L2/FL.

3.1 – Teaching of [v] vs. [b]

No doubt, the reason for the substitution of [b] for [v], typically experienced by Hispanic learners of English, is attributed to the absence of [v] in Spanish phonology. However, in light of MMA teaching, a [v] sound should be an easily surmountable difficulty simply because it is a highly visible labial sound as [b] is. If one works with Hispanic students, he will readily notice that the mispronunciation is pervasive even among some students whose oral proficiency and fluency in English are very good. There are two ways to account for this situation: a) The mispronunciation has not received much attention from the instructor; b) The instructor did not follow some effective techniques in teaching it. To put it more bluntly, the instructor did not have the know-how of effective remediation of learners' mispronunciations. Most probably, he followed the '*repeat-after-me*' technique which may not necessarily be effective with adults due to psycholinguistic deafness. In what follows, some strategies are put forth to develop an effective procedure to overcome the problem. The strategies typically reflect different cognitive and sensory modalities for handling the problem.

3.1.1 – Sequence of Orientations

a) Cognitive Orientation: Prepare the learners mentally (cognitively) to recognize the existence of the problem[4] and its seriousness be-

[3] For more details, see Odisho (2007).

[4] Some learners do not even realize that they have a problem of replacing [v] with [b].

cause it leads to serious phonetic or phonological accent. The cognitive preparation requires the following steps:

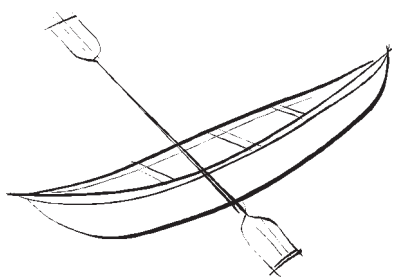
• Instruct learners to be ready to accept the problem and be willing to pay utmost attention to it.

• Tell them they will certainly manage the pronunciation.

• Tell them to watch your facial gestures, especially those of the mouth and recognize the difference in the pronunciation of [b] vs. [v]; the difference is drastic in both place and manner of articulation. In fact, to dramatize the postural difference in the articulation of the two sounds, you may call the [v] posture a '*dogface*' because when one assumes the posture, one looks like an angry dog ready to bark or bite. In contrast, you may call the [b] posture a '*tight-lip face*' since the lips have to come together tightly for the sound. The dramatization of the articulatory facial postures for the sounds oftentimes functions as a humorous, albeit robust and concrete, mnemonic to remind the learners of the required articulatory differences. Demonstrate the pronunciation of the sounds in selected minimal pairs of words for which the difference in meaning is easily noticeable and, perhaps, even funny or embarrassing, such as <vote> vs.<boat>, <vowel> vs. <bowel>, <valet> vs. <ballet> or <vending> vs. <bending>.

• Use colors and pictures or any other audio-visual to highlight the difference that results from substituting one sound for the other such in <vote> vs.<boat>, below:

vs.



• Ask learners to watch carefully your facial gestures, especially your mouth and lips, while you slowly and distinctly demonstrate the production of the two sounds. Stated differently, ask them to watch the dogface posture for [v] and the tight-lip-face posture for [b].

• While learners do all the above, carefully watch their facial gestures. If you notice that learners' facial gestures indicate attention and

seriousness in trying to impersonate you by moving the lips together for [b] or moving the lower lip to contact the upper teeth for [v] then you have to be sure that the learners are in a mode of thinking. In other words, they are trying to cognitively grasp the difference between the two sounds.

b) Auditory Orientation: Go back to the minimal pairs, number each member of the pair as #1 and #2 then produce each member of the pair and ask learners to identify the word as #1 or #2. Do this demonstration with your mouth covered with a piece of carton to prevent lip reading. Another major difference between the two sounds is that [v], being a fricative sound, is sustainable (can be prolonged), while [b], being a stop, is unsustainable (cannot be prolonged). If some learners still experience some difficulty in perceiving and recognizing the difference between the sounds, then go to the next step.

c) Visual Orientation: Remove the carton and pronounce the two sounds quite consciously while exaggerating the bilabial (upper & lower lips) posture for [b] and the labio-dental (lower lip and the upper teeth) posture for [v]. Put the learners in pairs facing each other and ask each member of the pair to perform the articulatory postures for the two sounds while the other learner is observing. Allow them to reverse their turns on this performance.

d) Kinesthetic/Proprioceptive Orientation: Ask the learners to carefully watch your demonstration of the two sounds with distinct performance of their articulatory postures. Stick with one of the sounds and repeat its articulatory posture then repeat its name. In other words, pronounce [və , və , və , və] followed by <Vee, Vee, Vee, Vee>. Repeat the demonstration with [bə , bə , bə , bə] followed by <Bee, Bee, Bee, Bee>. Ask them to impersonate what you have been doing with emphasis on the need to develop a kinesthetic and proprioceptive sensing of the articulatory contacts made for [v] and [b].

Obviously, there are additional follow-up exercises and demonstrations to reinforce the above procedures. However, the primary intention is to send as much diversified relevant input as possible to the brain via the auditory, visual and tactile-kinesthetic sensory modalities as illustrated in Figure 2, above, to help create the appropriate acoustic image in the brain.

3.2 – Teaching Lax (short) vs. Tense (long) Vowels of English

Those linguists who are familiar with contrastive aspects of the phonologies of English and Spanish know that the vowel systems of English and Spanish are maximally different. For the first time, in 1992, the English vowel system was identified as *centripetal*, wherein the vowels have lax (short) vs. tense (long) varieties with a strong tendency to move to the center of the vowel area where schwa [ə] is located, whereas the Spanish vowel system was identified as centrifugal in which the vowels tend to be tense with mid length and resist any movement to the center because of the absence of a schwa [ə] (Odisho 1992:26-7). These major differences between the two systems radically influence the syllable types and the assignment of stress; consequently, the teaching of vowels systems directly impacts the teaching of stress and stress assignment in English and Spanish.

Thus, the teaching of vowels tends to be the major problem of Hispanics in learning English. There are virtually thousands of pairs of words whose meaning is confused because of the failure to produce the targeted vowels; some such pairs can be very embarrassing, such as: <ship, bitch, chicks, shit> with [ɪ] vs. <sheep>, <beach>, <cheeks>, sheet> with [i:]. Think of the following real conversation with a Hispanic friend who had severe cold and facial muscle pain. He was asked: "How do you feel?" The answer was: "I am O.K., but my 'chicks' still hurt." Obviously, he meant his 'cheeks', but, unfortunately, it was rendered 'chicks'.

3.2.1 – What is the Problem?

The problem lies with the nature of vowels in the two systems with regard to vowel quantity and vowel quality – the former representing the length and the latter the acoustic impression of the vowel or the so-called color or timber of vowel. Let us, for instance, take the English pair <bid> vs. <bead> which is fairly satisfactorily transcribed phonetically as [bɪd] vs. [bi:d] indicating that the vowel elements are different in two respects: vowel quality in the form of [ɪ] vs. [i] as well as absence of length mark [:] in [bɪd] and its presence in [bi:d] for vowel quantity. For convenience, the term '*qualtity*'[5] is coined to represent a combination

[5] This is a blending of the terms 'quality' and 'quantity'.

of the features of quality and quantity. Descriptively and instructionally, the Spanish vowel will be transcribed as [i˙] with [i] indicating the quality of tenseness and the single dot indicating the medium or half-length (approximately, halfway between the English vowels [i:] and [ɪ]). The transcription is meant to signal the difference between [i˙] and [ɪ] in *qualtity* and only in quantity with [i:]. Consequently, because the Spanish vowel is half-way between the two English vowels, Hispanics replace the *qualtity* of the two English vowels with their single vowel *qualtity* which is the culprit for all vocalic confusions leading not only to phonetic accent, but also to serious phonological accent.

Obviously the features of quantity and quality are in many instances too intertwined together to be isolated and autonomously evaluated and described. Even though many authors and in many instances, handle the relationship of those two vowels as short vs. long, the relationship is too complex to be glossed over as short vs. long; it involves a feature of lax vs. tense accompanied by a difference in quality. This complex feature combination becomes an instructional reality when adult Hispanics embark on learning the vowels of English. The following strategies are suggested to handle such vocalic multiple-feature differences.

3.2.2 – Sequence of Orientations

a) Cognitive Orientation: Prepare the learners mentally (cognitively) to recognize the existence of the problem and its seriousness because it leads to serious phonetic and/or phonological accent. The cognitive preparation requires steps such as the following:

• Follow the same instructions in the first two bullets of cognitive orientation in 3.1, above.

• Tell learners to watch your facial gestures, especially of the shape of the mouth and lips. The lips are slightly more separated and the mouth is less spread sidewise for [ɪ] as opposed to [i:] for which the lips are less separated and the mouth is more spread sidewise.

• Demonstrate to learners a posture of relaxing the muscles as opposed to tensioning them. To achieve this, instructor puts his elbows on the table, relaxes his head between his hands and starts producing the [ɪ] vowel. Repeats those postures and asks learners to watch his facial gestures and impersonate them. It is the failure of Hispanic learners of

English to create a relaxed posture that leads them to render the English [ɪ] as [iˈ].

b) Auditory Orientation:

• To help learners with the perception of vowel quality and quantity, select several minimal pairs involving the English [i:] and [I] vowels and model their pronunciation with emphasis on both sound and meaning such as

[ɪ]	[i:]
<sick>	<seek>
<bit>	<beat>
<fill>	<feel>
<rich>	<reach>.

The selection of the minimal pairs should be conducted very carefully so as to avoid any unwanted interference from consonantal elements that are problematic for Hispanic learners such as the difficulty with [h], [v] or [z].

• To help learners with the recognition of the difference, number each member of the pair as #1 and #2 then produce each member and ask learners to identify it as #1 or #2.

c) Visual Orientation:

• Reproduce the above minimal pairs and ask learners to visually notice the difference in the mouth/lips/cheeks positions for the two vowels. Learners should notice more horizontally spread and more vertically approximated lips for [i:] as opposed to less horizontally spread as well as less vertically approximated lips for [ɪ].

• Produce the minimal pairs one by one placing your two index fingers vertically at the edges of your mouth. Narrow the distance between your fingers with the pronunciation of [ɪ], and widen it with [i:].

• Select a minimal pair and try to demonstrate schematically the qualitative and quantitative differences as reasonably as possible as in the schematic sketching below:

Word	**Schematic Representation**
Pill	
Peel	

Notice that the lesser length of the upper bold line for <pill> indicates a lesser vowel quantity (length) and its narrowness indicates its laxness, whereas the greater length of lower line indicates greater vowel quantity (length) and its thickness indicates its tenseness.

d) Tactile/Kinesthetic Orientation: This modality yields more feedback for the speaker than the listener a fact which implies that the instructor cannot be very helpful; however, he can ask learners to impersonate him in a different activity. The instructor asks learners to place their index finger in front of their mouth so as to touch the lips. While in this posture, learners try to repeat the pair <pill> vs. <peel> after the instructor. They will soon sense (feel) that with proper signaling of *qualtity* differences between the two vowels, there is more sidewise stretching of the lips and longer airflow with <peel> than with <pill>. Once they admit to sensing the difference, they should repeat the exercise several times.

3.3 – Comparative Teaching of English and Spanish Vowels

Instructional experience shows that the teaching of vowels, in general, is often far more challenging than the teaching of most consonants simply because vowels do not have firm anchorage places of articulation as most consonants do; in fact, vowels are formed with tongue and lip configurations that do not involve contacts. Thus, the teaching of vowels for L2/FL learners, especially those who are not distinctly gifted to sound impersonation, usually requires further steps and more practice and diversified exercises. Below are further exercises in the perception, recognition and production of vowel qualtity differences between Spanish and English vowels. The same pair of English vowels [ɪ] and [i:] will be used for demonstration with the additional insertion of the simulated Spanish vowel /i/, phonetically transcribed as [i^{l}].

In an experiment conducted with adult Hispanic learners of English at a beginning proficiency level, they were asked to pronounce minimal pairs such as the following: <did> vs. <deed>; <sit> vs. <seat>; <hit> vs. <heat>; <pill> vs. <peel> and <bid> vs. <bead>. The overwhelming common denominator in the rendition of those tokens was the failure to distinguish the vowel difference within each minimal pair. All those five pairs were reduced to five single renditions in the form of: [diˈd]; [siˈt]; [piˈl]; [hiˈt] and [biˈd].

The following strategies are suggested to handle such contrastive vocalic multiple-feature differences.

3.3.1 – Perception

• Model the triplet [sɪt], [si.t], [si:t] several times very carefully and as distinctly as possible.

• Ask learners to carefully watch your facial features especially of the lips and any skin and muscle contraction around your throat.

• Cite and pronounce other pairs of English words in which the meaning is very distinct such as: <kin> vs. <keen>; <bitten> vs. <beaten>; <tin> vs. <teen> and <fill> vs. <feel>. The semantic difference triggered by the sound difference will invoke thinking, more focused attention and better retention of the acoustic impression.

3.3.2 – Recognition

• Number the items of the triplet [sɪ t], [si.t], [si:t] as #1, #2, and #3.

• Record them randomly each repeated twice in, at least, fifteen to twenty attempts.

• Play the recordings back one attempt at a time with a few seconds of pause between each attempt and ask the learners to mark the items as 1, 2 or 3 on a specially prepared worksheet.

• Give the learners the key to the correct answers, ask them to identify the errors and notice the tokens which were with the highest percentage of inaccuracy. The results may be very significant for further design of exercises and drills.

• Ask learners to return all the worksheets of the first trial then ask them to prepare for a repetition of the exercise. Usually, the second and third trials are much better than the first one; more exposure creates

more acoustic familiarity and both lead to more confidence and better focus.

• Select a semantically appropriate minimal pair such as <will> vs. <wheel> mark the items as #1 and # 2, model them and ask learners to watch your mouth and identify them.

• Create carrier sentences with blanks in which the members of a given minimal pair such as <will> and <wheel> fit and ask learners to fill in with appropriate number such as the following:

A wheel (#2) is an important part of a car.

A will (#1) is something to be made public after death.

3.3.3 – Production

• Model the triplet [sɪt], [si.t], [si:t] very carefully and distinctly. Ask for volunteers to impersonate your production; gradually, all learners should be involved as individuals or groups.

• If some learners excel in the impersonation or production ask learners to repeat the performance.

• Allow those learners who excelled in performance to replace you in modeling. Preferably, learners should model while seated in their places among the students; this setting creates a more learner-friendly situation.

3.4 – Teaching of Stress Placement (Accentuation)

Generally speaking, the teaching of suprasegmental features (i.e., stress, rhythm, intonation) is equally challenging, if not more, to that of segmental features (consonants and vowels). Yet, they are granted lesser attention in classroom situations because many instructors do not have the know-how and experience to teach them. Take, for example, stress assignment of which very many L2/FL learners are not really aware; hence, it should be granted attention and MMA's multisensory and multicognitive orientations are capable of taking care of it. Fortunately, like any other human sound component, stress embodies and/or reveals itself in auditory, visual and tactile-kinesthetic sensory modalities which jointly help L2/FL learners cognitively internalize it as well as recognize and produce it.

An interesting technique of teaching the perception of stress is to practice it using nonsensical monosyllables such <la> or <ma>. Begin

with two syllables and keep increasing the number, but do not exceed four syllables because too many syllables will confuse the learner. Keep shifting the stress from one syllable to the other as demonstrated below.

'La La
La 'La

'La La La
La 'La La
La La 'La

One can then match those nonsense syllables with real words. For instance, <'La La> will match the noun <Import> which has the stress on the first syllable, whereas <La 'La> will match the <Import> which has stress on the second syllable. Likewise, <La 'La La> will match the adjective <Important>. Let us consider the following orientations to find out how to further reinforce the perception, recognition and production of stress.

a) Cognitive Orientation: Prepare the learners mentally to recognize the nature of stress as a physical phenomenon. After all, stress is the outcome of greater articulatory and aerodynamic effort on the part of the speaker. Show learners that when a syllable is accentuated (stressed) there are several physical gestures that indicate the stress. Tell learners to watch your facial gestures and body gestures, especially those of the hand, fingers and head (including hair for those who have long hair). Experience will help the instructor to identify those learners who have difficulty by simply looking at their faces or by simply asking them.

b) Auditory Orientation

• Through tapping or beating on something that yields good resonance, demonstrate to the learners three syllables <La 'La La> by three beats with the middle one being the most prominent. One can transform the beats visually in different forms such <im'portant>, <imPORtant> or even <· ● ·>

• To diversify the technique, take for example the words <insight> and <incite>, pronounce them emphatically to highlight the stressed syllable in each one of them and ask learners to identify the stressed syllable in each one of them. To dramatize the difference and capture the attention of the learners, grab an empty can and beat the rhythm of the two words on it. The auditory input of the beat is reinforced visually and both inputs reinforce the memory and the retention of the stress.

c) Visual Orientation

• One can demonstrate this in different ways such as coloring the stressed syllable, capitalizing and/or rendering it in bold type. Thus, <insight> will appear as <INsight>, or <**IN**sight> etc…, whereas <incite> will appear <inCITE> or <in**CITE**> etc.

• One can also render the difference in small and large dots or with short and long arrows for unstressed and stressed syllables, respectively as demonstrated below:

<insight> will look as ●· and  while <incite> will look as ·● and 

d) Tactile/Kinesthetic

Stress could be taught through this modality if the instructor is allowed to touch the learner, but since this gesture may have some social constraints, learners have to practice this modality on themselves by impersonating the instructor in the following manner.

• With the five fingers clustered together, the instructor taps a stronger beat on his chest followed by a weaker one for <insight> and reverses the beats for <incite> and asks learners to impersonate him. In actual fact, if learners succeed in this performance, this exercise tends to be very helpful because it is multisensory in nature; both instructor and learner can see, hear and feel the beats on the chest.

• Instructor can also take a large step followed by a small step for <insight> and reverse the steps for <incite>. This exercise also tends to be multisensory in nature.

Once the learners slowly improve their perception, recognition and production of stress all what they have to do is to pursue the practice until they transform their mechanical habits into cognitive one.

4 – Concluding Remarks

In short, teaching pronunciation is not a mechanical process in which isolated sounds keep moving to and fro between the mouth and the ear. To put it differently, it is not a process that is singularly based on the auditory channel. Teaching pronunciation is a far more sophisticated process in which the auditory input is reinforced by input from other sensory modalities. Without a multisensory approach coupled with cognitive orientation, it is extremely difficult to teach pronunciation to adult learners of L2/FL. Usually adults display considerable cognitive conditioning and bias to the native language phonetic and phonological rules and constraints and thus develop a kind of resistance to L2/FL acquisition/learning which was described earlier on as *psycholinguistic deafness* to replace the traditional *fossilization*. In sum, teaching pronunciation to adults seems much like a journey that transforms one from the concrete (physical) world of sounds to the abstract (mental) world of sounds where the brain has to be induced to accommodate the new sounds and enabled to fire the right commands for their targeted production.

References

Anderson, J. R. 1980. *Cognitive psychology and its implications*. San Francisco: W.H. Freeman & Company.

Arnold, M. B. 1984. *Memory and the brain*. Hillsdale NJ: Lawrence Erlbaum.

Daniloff, R. G. 1973. Normal articulation processes. In: F. D. Minifie; Th. J. Hixon; F. Williams (Eds.). *Normal aspects of speech, hearing and language*. Englewood Cliffs NJ: Prentice-Hall, 169-209.

Odisho, E. Y. 1977. Arabic /q/: a voiceless unaspirated uvular plosive. *Lingua*. **42**: 343-347.

Odisho, E. Y. 1992. Transliterating English in Arabic. *Zeitschrift für arabische Linguistik*. **24**: 21-34.

Odisho, E. Y. 2003. *Techniques of teaching pronunciation in ESL, bilingual and foreign language classes*. München: Lincom-Europa.

Odisho, E. Y. 2004. *A linguistic approach to the application and teaching of the English language*. New York: Edwin Mellen Press.

Odisho, E. Y. 2005. *Techniques of teaching comparative pronunciation in Arabic and English*. New Jersey: Gorgias Press.

Odisho, E. Y. 2007. *Linguistic tips for Latino teachers and learners of English*. New Jersey: Gorgias Press.

Selinker, L. 1972. Interlanguage. *International Review of Applied Linguistics*. **10**: 209-231.

Werker, J. F. 1995. Exploring developmental changes in cross-language speech perception. In: L. R. Gleitman; M. Liberman (Eds.). *An invitation to cognitive science*. Cambridge MA: The MIT Press, 87-106.