NEURAL PHONOLOGY:[1] A MULTISENSORY, MULTICOGNITIVE APPROACH TO ITS ENHANCEMENT IN TEACHING PRONUNCIATION[2]

Edward Y. Odisho
e-odisho@neiu.edu
Northeastern Illinois University (Estados Unidos da América)

A smart teacher is the one who is prepared to learn from his/her students regardless of whether they err or excel

Abstract: As a reaction to the pedagogical needs of adult learners in mastering speech sounds and sound phenomena unfamiliar to their L1 phonology, there surfaced a pressing need for a radical change in the overall approach to teaching them. A transition had to be gradually implemented in the form of a multisensory (auditory, visual, tactile-kinesthetic) strategies to be propped up with a set of multicognitive ones (e.g. think, associate, analyze, synthesize, memorize) in lieu of the traditional audiolingual one. During the implementation of such strategies for a few years, a fully-fledged approach emerged and titled as: *Multisensory, Multicognitive Approach to Teaching Pronunciation* (MMA). Further classroom experience with MMA, led to the reconsideration of the linguistic domain of phonology for theoretical and applied purposes and recognizing it in two sections i.e., *linguistic phonology* vs. *neural phonology*. Below are a few reasons for introducing *neural phonology*. *First*, traditional phonology barely refers to the role of the brain in internalizing additional phonological constituents when targeting L2/FL as adults. *Second*, teaching phonology and the mastery of unfamiliar sound units/features to learners of L2/FL has been conventionally attempted in a *unisensory* (monosensory) manner, the auditory sense in this case, which tends to be less efficient and effective with many adults. *Third*, in real classroom situations there seems to be greater need to diversify and refine the sensory and cognitive strategies/techniques in teaching both sections of phonology, especially the *neural* one. *Fourth*, emphasizing the two sections further enhances the overall pedagogical and didactic worth of MMA.

Keywords: MMA; Linguistic Phonology; Phoneme; Neural Phonology; Neuroneme

1 *Neural Phonology* is coined and introduced not just to emphasize the role of the brain in cross-language teaching of L2/FL pronunciation to adults, but also to highlight the role of the *multisensory, multicognitive approach* to achieve that.

2 I deeply appreciate the feedback and advice of Professor Emerita Maria da Graça Lisboa Castro Pinto, University of Porto in the evolution of MMA throughout the last two decades.

1- Introductory Remarks

The primary purpose of this study is to enhance and better specify the domain and focus of the linguistic field traditionally known collectively as *phonology* especially when applied in cross-language pronunciation (Odisho, 2003; 2007 2014, 2016). This will eventually impact the overall approach not just to its understanding, but also to its application, above all in classroom situations and all relevant situations such as in acting, broadcasting, interrogation, comic impersonation and even in espionage etc…(Odisho, 2014). More formally 'phonology' has been identified as the study of the *sound system* of a given language including its basic segmental units known as 'phonemes' as well as its suprasegmental features such as stress, rhythm, tone, intonation etc... In case of second language teaching (L2) or foreign language teaching (FL) to adults, the most popular approach, especially in the case of 'phonemes', both vocalic and consonantal, has been through the *audiolingual* method popularly known as 'repeat-after-me' technique. According to this technique, the teacher demonstrates the targeted phoneme, to the best of his/her knowledge or ability and asks learners to repeat, individually or collectively (chorus), after the demonstration. Classroom experience has shown, undeniably, that many adult learners often repeat after themselves as a subconscious reflection of their native language (L1). In other words, they habitually tend to produce a sound that is in their L1 phonology which they erroneously assume is identical or, at least, similar to the targeted one. This is the source of accent be it *phonetic* or *phonological* (Odisho, 2016).

It is also well known that children and young adults do not usually have difficulty in readily and correctly internalizing sound units of any L2 or FL that they embark on learning or are exposed to as long as the exposure/input is exemplary and relatively extended.

This study reintroduces, for applied purposes, the traditional domain of phonology in a more refined manner as '*linguistic phonology*' vs. '*neural phonology*'. It follows that, for best intentions and purposes, each one of them should be strictly identified and defined. Hereby, '*linguistic phonology*' stands for the study of segmental and suprasegmental units of the sound system of a given language in theoretical or applied contexts, whereas '*neural phonology*' stands for the same linguistic units as neuronized (neurologically ingrained) in the brain of the learner or speaker. For the purpose of this study, the classical linguistic phoneme will bear two identities: a *neuroneme*[3] in the brain and an *allophone* in actual articulation (speech). The neuroneme is mental whereas the allophone is physical. The latter emerges as the phoneme is placed in the context of real speech. This distinction is extremely significant in applied situations, especially in adult L2/FL classroom situations.

3 Based on '*neuron*' following the example of *phoneme* for sound and *grapheme* for letter.

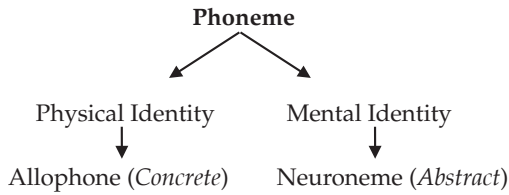


Figure 1. Physical vs. mental identities of Phoneme

There are four reasons for introducing *neural phonology* as opposed to *linguistic phonology*. *First*, traditional phonology barely refers to the role of the brain (i.e. mental orientation) in internalizing additional phonological constituents when targeting L2/FL as adults. *Second*, teaching phonology and the mastery of unfamiliar sound units/features to learners of L2/FL has been conventionally attempted in a *unisensory* (monosensory) manner i.e., the auditory sense in this case which tends to be less efficient and effective with many adults. One doesn't teach the ears how to acquire the sounds; rather, the ears simply serve as a conduit to coach the brain to neuronize them. *Third*, in real classroom situations there seems to be greater need to diversify and refine the sensory and cognitive strategies/techniques in teaching both domains of phonology, especially the *neural* one. *Fourth*, emphasizing the two domains justifies the overall teaching approach introduced in several previous publications (Odisho, 2003, 2007, 2014 and 2016) and designated as 'multisensory multicognitive'. The paragraph below was mentioned in the first chapter of the 2003 publication:

> In teaching pronunciation to adults and in L2/FL learning situations, the approach promoted here is a multisensory and multicognitive. 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 learn likewise. The approach implies that the so-called technique of 'ear training' [i.e., auditory learning or how to hear and recognize sounds] cannot solely do the job. Ear training should be supplemented by what is to be known, hereafter, as: a) eye training [visual learning of pronunciation through seeing and visualizing sound production and the accompanying dynamics of body and facial gestures; b) neuro-muscular training [tactile training or how to kinesthetically and proprioceptively sense and feel sound production and its dynamics]; and c) brain training [cognitive training] or how to psycholinguistically and consciously perceive, recognize, retrieve and produce the sounds and their underlying dynamics. (Odisho, 2003:6)

1.1 - Linguistic Phonology

Phonology is traditionally identified as the description, study and teaching of the sound system of a given language including its segmental components (i.e. vowels, consonantal etc…), its suprasegmental ones (i.e., stress, rhythm, tone, intonation etc…) and the dynamics that govern them. Because covering both

segmental and suprasegmental aspects will be too extensive a domain to deal with in a limited study such as this, the focus will mostly be confined to the study and teaching of segmental components with a touch on suprasegmentals when needed.

When teaching the segmental components, they are usually introduced as consonants and vowels. The consonants are described in terms of three features: *place of articulation*, *manner of articulation* and *voicing/voicelessness*. Traditionally, place of articulation refers to the location of a stricture beginning with lips and moving backwards along the vocal tract towards the glottis. Thus, we have bilabial, labio-dental, inter-dental, alveolar etc….Manner of articulation identifies consonants based on the type of strictures created along the vocal tract and the resulting airflow dynamics yielding plosives, fricatives, affricates, approximants etc…The third descriptive feature represents the status of the vocal chords accompanying the production of each consonant resulting in vibration (voiced) or its absence (voiceless).

As for vowels, they are described in terms of: *tongue height* (low: mid: high); *tongue section involved* (front-center-back); and labial shape (round: neutral: unround). In some more advanced descriptive approaches, the contrast of lax vs. tense is added which will be considered a very significant distinctive feature in this study.

One of the earliest methods of teaching L2/FL has been the audiolingual one. In case of L2/FL teaching, this method has two major drawbacks to be reiterated here. *First*, adult learners may fail to reproduce the targeted sound even though the demonstration by the instructor may be ideal simply because adult learners in this case are heavily influenced by their native language (L1) and may become less psycholinguistically sensitive to sounds unfamiliar to their native phonology. For example, the instructor may demonstrate an ideal Spanish tap 'r'= [ɾ], as in 'pero' but the native English speaker or speakers may produce a frictionless continuant 'r' [ɹ]. In other words, they reproduce after themselves under the bias of their L1 (English) phonology. *Second*, the instructor may *not* be a native speaker of the target language therefore his/her demonstration of the targeted sound may not be ideal due to L1 bias. Definitely, in such a scenario, learners will either impersonate the inaccurate demonstration of the instructor or reproduce what is deep-seated in their L1 (i.e., neural phonology). Such scenarios are very common throughout the world as the overwhelming majority of such teachers may lack the native or near-native phonological proficiency in their L2/FL. It is scenarios such as these that justify the initiation and application of *neural phonology*.

1.2 - Neural Phonology

Within the grand system of *economy in nature*, in general, and in human language in this particular instance, there are different sub-systems that contribute to building the overall unique system of speech generation. Without such sub-systems the load of language on the brain would simply be too much to endure. To illustrate, the human speech apparatus can hypothetically generate infinite

number of sounds; however, two questions immediately arise. *First*, does a human language need that many sounds to generate speech? *Second*, does the brain, which has thousands of other biological functions to handle, like to burden itself with thousands of minimally distinctive speech sounds? The answer to both questions is 'no'. In the *first instance*, the *generative design* of language requires only tens (hardly, if ever, in hundreds and impossible in thousands...) of sound units (phonemes) to be recycled again and again in a recurrent and generative manner. This highly economic blueprint is a universal feature of human language without which it will lose its limitless creativity. In the *second instance*, there is a salient tendency in most languages to manipulate sounds the articulatory maneuvers of which are predominantly easy to *produce* and easy to *perceive*. This implicitly limits the pool from which human speech sounds are selected.

The above two dynamics minimize articulatory ambiguity and enhance clarity. It is, therefore, not accidental that the majority of speech sounds are produced in the anterior half of the vocal tract in the form of bilabial, labio-dental, inter-dental, dental, alveolar, post-alveolar etc… sounds. Moreover, although in actual speech each phoneme may have a large number of different phonetic variations (allophones) for the same unit[4], the speaker *does not* recognize and store all those phonetic variations. In other words, allophones are not registered (neuronized) in the brain; rather, it only cognitively internalizes one abstraction for all the variants of a given sound unit commonly known as 'phoneme'. It is in the context of live speech that phonemes mold themselves thus yielding the suitable contextual variants (allophones). In light of the above explanation, if a speaker uses a phoneme that is not part of the phonological inventory of the adult listener, the latter will highly likely fail to recognize it and subsequently replace it with a sound that is phonetically the nearest one to it—*but it is not it*. This phonological failure occurs simply because the targeted phoneme, in this instance, is not neuronized in the brain of the listener. Stated differently, the targeted sound has not instituted itself yet as a *neuroneme* with the listener.

Let us consider yet another example to demonstrate the commanding principle of economy in language. When speaking, the brain fires instructions to construct a word, but the dynamic nature of speech, especially the mutual interactions of sounds in the flow of speech, creates different phonetic shades (allophonic variations) for each sound. The brain of the listener matches those allophonic variants to the abstract units in his/her *neural phonology*. To state the same fact in more accurate linguistic terms, in the context of actual speech flow, the speaker generates his/her *neuronemes* phonetically (i.e. allophonically) while the listener tries to identify them phonemically in terms of his/her L1 phonology. If a certain speaker utters a word which includes a sound that is not a unit in listener's native language the latter will either completely fail to recognize it or will mistakenly identify it as a sound that is similar to the one in his/her L1 mental inventory

4 For instance the /p/ phoneme in English may have several different allophones such as aspirated in <pit>, unaspirated in <spit>, lip-rounded in <pool> or lip-spread in <peel> etc…, but the brain neuronizes only one generic form known here as *neuroneme* or conventionally as 'phoneme'.

(neural phonology). At times, this mistaken identification is so obtrusive that it may even be transferred to the writing system (orthography). For instance, due to the absence of /v/phoneme in Spanish and /p/ phoneme in Arabic, spelling errors such as 'bery good' for 'very good' and 'baber' for 'paper' have been repeatedly observed by the writer.

The above prelude justifies the rationale for using 'neural phonology' synonymous with 'cognitive phoneme inventory'. Such justification is to highlight the significance of cognitive processing (neuronization) of any alien sound of the target language on equal footing with L1 inventory of phonemes (neuronemes). If the targeted L2/FL *phoneme* fails to become part of L1 *neural phonology* (cognitive inventory), the result is what has been formerly identified as *phonological accent* which leads to semantic change (Odisho, 2003; 2016).

2 - Conscious and Subconscious Brains

In the preceding sections, different pieces of evidence were knit together to highlight the economic premise of language as a reflection of the robust tendency towards scaling-down the burden of language on human brain. It is the massive data storage capacity of the brain through evolution that has accommodated language as an open-ended system capable of producing and recognizing endless meaningful stretches of speech. Nevertheless, one of the most salient attributes of the brain that relates to human language and empowers it to be such a rich, creative and sublime medium of communication will be discussed in this section. This attribute is the twin-nature of human brain as *conscious* and *subconscious*. It is, therefore, logical and substantiable to say that the human dichotomy of conscious-subconscious brains has been one of the main evolutionary developments that gradually emerged to manage, administer and execute millions of biological, social and cultural functions that humans have to successfully execute in order to survive healthily and rationally. One such fundamental function of the brain is language; indeed, without a highly potent and sophisticated brain there would be no language. It is true that human brain has billions of neurons and trillions of synapses, but if human language were to be absolutely conscious then it would need even more space for its additional neurons and synapses to be accommodated in a 'coconut-size shell' (i.e., skull) where contents are already packed to the brim. Without the dichotomy of conscious and subconscious brains, language would be too much of a mental drain exclusively on the conscious brain to be able to handle so smoothly and effortlessly. It is the conscious-subconscious dichotomy that accommodates language use as efficiently as it has been throughout human evolution.

Here again there is a distinct division of labor between the 'two brains' for the sake of economy in effort through securing maximum coordination and harmony in the execution of the myriad functions of the brain. The conscious brain is responsible for any action that it decides to initiate. Once it decides on a certain action, most of the requirements for completion are automatically delegated to the

subconscious brain. To cite an example, when engaged in an informal conversation with family members and friends, a sizeable percentage of the conversation is managed by the subconscious brain simply because one does not seriously engage in planning the contents of the conversation such as the selection of the needed vocabulary and monitoring the morphological and syntactical rules because the storage of lexicon and the rules of grammar are in the subconscious. However, unlike this informal conversation, if the person engages in delivering a formal speech orally, the conscious brain assumes a greater role to cater for targeted contents, careful selection of the needed lexicon and greater adherence to formal grammatical rules of morphology and syntax. It is because of the conscious role of the brain, the speaker might have more pauses, hesitations and repetitions than in a casual social conversation due to a covert conflict between the two attitudes of the brain.

For all crucial biological purposes, the subconscious brain never sleeps because it has to monitor all body functions that are basic for healthy survival: the heart beat, respiration, blood circulation, secretion of the necessary glands, digestion etc... Most important of all in this study, the subconscious brain is the sentinel of our normal continued existence especially inasmuch as language is concerned; it is the seat of the *long-term memory* unlike the conscious brain which usually handles *sensory memory* (for just a split second) and at best *short-term memory* (for a few seconds). With every impression, experience and event that a person intends to maintain for long term, it is the duty of the conscious brain to serve as a medium to transfer them to the subconscious brain followed by some reinforcement strategies[5].

In an earlier study of the strategies for teaching pronunciation, it was pointed out that language acquisition is a process of mental (cognitive) habit formation (Odisho, 2003). The term 'cognitive' was used to mean storing the '*linguistic habits*' in the brain and retrieving them instantaneously when needed. It is common knowledge to say that "whenever anything has been repeated a sufficient number of times to have become habitual, it becomes second nature, or rather a subconscious action (Larson, 1912). We all as babies, toddlers or young children have slowly and gradually learned how to grasp things with our fingers, balance our bodies, walk, run, ride a bicycle and perform countless number of functions which through constant and systematic repetition have been transformed into automatic, effortless and subconscious survival functions

The ease with which human beings use language as the most efficient social and cultural tool has only been possible through continuous mental transferences from the conscious to the subconscious. All the required basic constituent units of sounds, the rules that combine them in different formations and assign meaning to them are gradually transferred from the conscious brain to the subconscious especially in childhood (Carpenter, 2004).

5 The retention in the long-term memory is the result of anatomical or biochemical changes that occur in the brain (Tortora and Grabowski, 1996).

3 - Reconsidering Instructional Practices

Quality of instruction, its intensiveness and the approach implemented are often interrelated. It is not enough to be a conscientious instructor; what really matters is the methodology and/or approach to which one adheres and implements. To illustrate, during the first three decades of five of the teaching career of this writer, different methods have been used according to their availability in the field of language teaching. At times, an eclectic approach had been used combining different methods according to their suitability. It was only during the last two decades (1992-2009) of his career that he settled on what was progressively developed and labeled as '*multisensory multicognitive approach*' particularly geared towards *contrastive phonetics and phonology* teaching in adult classroom situations or one-on-one tutoring.

This fundamental turning-point to the 'multisensory multicognitive approach' gradually evolved in the context of teaching adult Spanish-speaking students who were registered for courses conducted in English. One such course was in contrastive pronunciation with focus primarily on English and Spanish. Days, months and years of teaching such courses convinced the writer that learners' primary difficulty was with the vowel system of English coupled with a few consonants such as [v] in 'very' and [j] in 'young' which sounded like 'bury'= [beri] and 'junk' = [ʤang], respectively. Initially, attempt was made to improve their pronunciation in the traditional technique of demonstrating the difficult sound and asking them to repeat thereafter. Usually, some learners failed to correctly reproduce the targeted sound; however, others managed to produce it correctly immediately after a model demonstration, but in a few minutes, hours and certainly days they slipped back into their Spanish version of the targeted sound. Frustration kicked in as an instructor and it was time to reconsider the approach in favor of a more effective one.

4 - Rationales for Introducing Multisensory Multicognitive Approach

As the traditional audiolingual method with Spanish-speaking students failed, there was no choice but to seriously seek for an alternative one. Since the brain is the seat of all human activities and functions, there was no better choice than seeking for the prime solution in cognitive studies with emphasis on human language from the perspective of psycholinguistics and neurolinguistics (Pinto, 2015). Seeking refuge in these two fields unveiled a treasure of knowledge to be intertwined with lengthy classroom experience.

Any study of human language internalization, let alone its teaching especially to adult L2/FL learners, should consider the manner in which normal children master their native language or any language they are immersed in as opposed to adults embarking on learning a second language. The two processes are known in psycholinguistic literature as *acquisition* vs. *learning*. Acquisition tends to be a subconscious, automatic and effortless process of internalizing a language. This is what a child typically does with the mother tongue or any language he/she is

exposed to, whereas adult learning of L2/FL tends to be more conscious, mechanical and effortful. Inasmuch as pronunciation acquisition is concerned, all what children need to do is ample exposure to speech in real-life contexts and situations. Perhaps, more interestingly, children are naturally able to accomplish the acquisition of the pronunciation of more than one language if they are amply exposed to.

In an earlier study of the strategies for teaching pronunciation, it was pointed out that language acquisition is a process of mental (cognitive) habit formation (Odisho, 2003) later refined and stabilized as *neuronization* (2016). The term 'cognitive' was used to mean storing the linguistic habits in the brain and retrieving them instantaneously when needed. It is common knowledge to say that "whenever anything has been repeated a sufficient number of times to have become habitual, it becomes second nature, or rather a subconscious action (Larson, 1912). At an early age in our life, we all transform most of our conscious acts into automatic, effortless and subconscious ones. Such transformations are the greatest relief that nature has ever bestowed upon human beings. It is this transfer of mental load from the conscious brain to the subconscious one that helps the former avoid early collapse under the pressure of too many requests and diversified commands for action. Without mental habit formations (neuronization), life would be too stressful and burdensome on the conscious brain. In our day-to-day life, when we say that life is becoming stressful we simply mean that we are engaging the conscious brain in responding to several problems simultaneously knowing that brain prefers to handle one problem at a time. Thus, normal survival of human beings without a powerful 'subconscious' brain is virtually impossible.

The ease with which human beings use language as the most efficient social and cultural tool has only been possible through continuous mental transferences from the conscious brain to the subconscious. All the required basic constituent units of sounds, the rules that combine them in different formations and assign meaning to them are gradually transferred from the conscious brain to the subconscious especially in childhood. Based on the fabulous significance of the transference of conscious acts (in this case speech acts) into subconscious ones, the 'multisensory and multicognitive approach' in teaching cross-language pronunciation was gradually developed and refined in description and application.

5 - MMA Emulating Motherese Style of Child Language Acquisition

Attempt was made to assist adult students' at internalizing (neuronizing) sounds alien to their L1 phonology by reinforcing their traditional learning strategies with those of child language acquisition approach the so-perfectly-called, *motherese* (Newport,1975). In certain respects, MMA is an attempt at emulating the motherese style of a child language acquisition, especially its sound system. Foremost among the salient prerequisites 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 necessary to acquire

all the skills and subskills of language. 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 (acquisition) that gradually becomes more subconscious, automatic and effortless. In contrast, 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 L2/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. MMA is more oriented toward teaching L2/FL to adults than to young children. At this stage, 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, in certain instances, be equally adept or even more adept than children.

5. 1 - Multisensory

In handling the skill of pronunciation and its exclusive association in the case of adults with unisensory (monosensory) modality, typically the auditory, 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 activities. An integration of auditory, visual and tactile/kinesthetic/proprioceptive sensory modalities is indispensable for the proper mastery of speech, in general, and pronunciation, in particular. The older the learners of L2/FL, the greater the need for the *motherese* style of teaching and acquiring new sounds. Students should learn to emulate what a baby does when not only carefully listening to Mom's speech, but also watchfully monitoring her mouth, her eyes, her facial gestures as well as other bodily acts. The manner in which babies emulate Mom's linguistic performance in both articulation and bodily gestures typically symbolizes a 'gestalt' approach to executing missions; in this case, *child language acquisition*.

In efficient and effective teaching of L2/FL, there has to be an interface between more than one sensory modality and allow the learners to manipulate the modalities severally or jointly. Hence, adult learners of L2/FL have to be prepared not just to listen to the sounds, but also, and equally importantly, to see and feel them in conjunction with the concomitant sensations and body gestures. This is why MMA is sturdily premised on a *multisensory*[6] *foundation.*

In light of the multisensory foundation, certain category of consonants, such as the bilabials [b, p, p^h], labio-dentals [f, v], interdentals [δ, θ] and even dentals and alveolars as well as many vowels should squarely be identified as *visible* sounds for the eyes to watch their concomitant facial and articulatory configurations. To cite a very relevant case in this regard is the teaching and distinguishing of sets of aspirated vs. unaspirated sounds that are so popular in the phonology of many languages across almost all known language families. Among such sets are the

6 "There has been an increasing interest in the study of the underlying brain mechanisms of multisensory integration both on an anatomical and functional level" (Paraskevopoulos & Herholz, 2013)

bilabials /p/ vs. /p^h/; alveolars /t/ vs. /t^h/; palatals /c/ vs. /c^h/ and velars /k/ vs. /k^h/. The best approach to teach such phonological contrasts is via the visual modality by placing a flimsy paper in front of the lips when the contrasts are performed. In the case of aspirated ones the flimsy paper is noticeably fluttered whereas with the unaspirated the fluttering is absent. Such a demonstration was extremely helpful for the Spanish-speaking learners of English in their transition from their unaspirated plosives to their English aspirated ones and the *vice versa*. Consider also the case of the absence of /f/ in Tagalog language of the Philippines and its replacement with /p/. The visual articulatory maneuvers in the production of these two sounds are so vivid to the eye and so kinaesthetically palpable that the correction process becomes easy and doable.

Another quite challenging phonetic/phonological articulatory phenomenon is that of retroflexion (tilting the tip of the tongue backwards) which is pervasive in languages such as Hindi and Urdu. Whether to help non-native speakers of such languages acquire it or to suppress it in case of their natives embarking on learning other non-retroflex languages is an extremely challenging articulatory maneuver to execute or undo. In the case of performing the retroflexion maneuver or suppressing it, the auditory and visual modalities are helpful, but less so compared to the kinaesthetic and proprioceptive ones. It is beneficial to advise learners to try to practice rolling the tip of the tongue backwards to execute retroflexion. Classroom experience has also shown that looking into the oral cavity with the help of a mirror convinces the learners whether the tongue-tip is actually curved backwards or not. Repeating such maneuvers consistently has proven helpful.

More importantly, many vowel sounds are better taught and learned by visually monitoring the lip configurations rather than by repeating after a model demonstration only. This is what is traditionally known as 'lip-reading' which in many situations people involved are advised to avoid. Typically, judges and lawyers in courts and coaches and players in sports arenas practice mouth-masking with their hands to avoid unintentionally revealing what they intend to say. Such mouth-masking is typically and extensively practiced in doubles tennis game to avoid revealing their intended instantaneous plan[7].

5.2 - Multicognitive

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 (neural phonology). Notice that in teaching pronunciation, the emphasis should be on imprinting and retaining the multisensory images of new sounds or sound phenomena rather than pursuing the route of mechanical imitation simply because the latter is the result of a *short-lived*

7 In fact, in English, no one dares not to mask his mouth if he, in anger, intends to use an expletive publicly that begins with a labio-dental fricative /f/.

unicognitive process, whereas retention is the result of a combination of multicognitive processes such as thinking, association, analysis, synthesis, comparison, contrast as well as memorization.

Although these cognitive activities may sound too abstract for some teachers, 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 cognitive endeavors on the part of learners to master the dynamics of the targeted sounds and reproduce them.

Conjointly, the multisensory and multicognitive modalities should be activated to help the *brain* with *perception* and *recognition* prior to the targeted *production*. It is a fact that adults embarking on L2/FL are highly likely to show insensitivity to alien sounds outside the realm of their L1 phonology. They experience what has been labeled as *psycholinguistic deafness* or insensitivity (Odisho, 2003; 2007; 2014) in perceiving and recognizing phonological contrasts in their targeted L2/FL language. Thus, any teaching of pronunciation should thoroughly follow the natural three-phase procedure of child sound acquisition in the sequence of *perception*, *recognition* and *production* (PRP). This three-phase procedure is highly consistent with brain's three-phase cognitive procedure of *registration*, *retention* and *retrieval* (Arnold, 1984; Levitt, 1981). It is such triplet sensory processes conjoined with their cognitive counterparts that tremendously help in transitioning alien L2/FL sounds to learners' L1 phonology (neural phonology) and enrich it. Indeed, it is practices like these that define, establish and enhance the domain of *neural phonology* and create L1phonology *plus*.

6 - MMA: a Key to Neural Phonology Consolidation

It has been hinted earlier on that the manner in which human brain functions is governed by the principle of economy in conscious effort. This principle should make it crystal-clear that any activity that is essential for physical, mental or social survival of humans must be *neuronized* to require minimum effort and time in order not to be a burden on the brain. The almost century-long argument as to whether the 'phoneme' is a concrete (physical) entity or an abstract (cognitive) one should not have arisen in the first place. In the brain, the phonemes are mental and highly limited in number[8] and identified in this study as *neuronemes* (Fig. 1). However, once articulated in real speech, they assume their physical identities according to contexts in which they occur. It is these physical identities that are

8 Most of the languages throughout the world have less than 40 phonemes or just slightly more. For more specific details consult: Ladefoged and Maddieson, 1995.

branded as *allophones* the number of which matches the contexts they occur in. It is very important to reiterate that neuronemes must be extremely limited in number (often with a range of 20-40), while allophones are too numerous to yield themselves to be counted readily.

In order to elaborate on the relationships of phoneme vs. allophone, on one hand, and the speaker vs. listener on the other hand, the foremost fact to be considered is whether the speaker and listener belong to the same language or to two different languages. If they are speakers of the same language then they have no difficulty in PRP of the same phonemes (neuronemes) and other speech components. Contrary to this, if speaking occurs in cross-language context then it is quite likely for one interlocutor to fail to PRP the neuronemes that are typical of one language only. This is because the neuronemes of a given L1 lead to the creation of a specific *neuronemic* (phonological) filter for that L1. Consequently, this L1-specific filter is likely to fail to PRP the L2 neuronemes that are alien to it. It is even likely to fail to PRP the neuronemes of L2 that happen to be identical with an *allophonic* variant in L1. Let us cite some concrete examples to illustrate the effect of language-specific *neuronemic filter*. A native speaker of English whose plosive phonemes /p, t, k/ are typically aspirated [p^h, t^h, k^h] fails to PRP their Spanish unaspirated counterpart phonemes [p, t, k] although English does have the unaspirated *allophonic* versions as in the s-initiated consonant clusters <spin>, <stick> and <skim>.

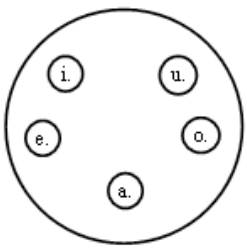
There are two significant inferences that are drawn from the above observations. *First*, it is the *neuronemic filter* of a given language that determines which phonemes of L2 will be difficult to PRP. This often happens by blocking the alien phonemes from passing through the *neuronemic filter* because it fails to recognize them. To perceive a new sound, the learner has to be able to switch his attention from the world of hearing (involuntary action) to the world of listening (voluntary action) as the former is a "sense while the latter is a skill… *listening is where hearing meets the brain*... listening to language is uniquely human" (Beck and Flexer, 2011).

7 - Pedagogical and Instructional Significance of Highlighting Neural Phonology

The emphasis on *neural phonology* as a key component of the multisensory multicognitive approach to teaching cross-language pronunciation gradually emerged as a reaction to the classical method of teaching pronunciation relying conventionally on the auditory sense. This exclusive reliance seems to have accounted for the initial failure of this writer in helping adult Spanish-speaking (Hispanics) students improve their pronunciation of some sounds and sound features of English. The initial instructions primarily based on the audiolingual method failed to help many students overcome their serious pronunciation problems most of which being phonological leading to semantic confusion such as English <sin> sounding as if it is <seen> and <scene> or the *vice versa* with latter

two sounding as <sin>.

In reality, the personal experience reflected in this study serves as an example of an instructor learning from the errors of his students to enhance and improve his own teaching approach and the strategies for its implementation. The most crucially important phonological problem discovered was related to the vowel systems. In a previous research work, it was concluded that the vowel systems in English and Spanish are drastically different in the number of vowels, their qualities and the dynamics that govern them as well as their overall rhythm types: English being stress-timed while Spanish being syllable-timed. All those differences led to naming the English vowel system as *centripetal* and Spanish as *centrifugal* (Odisho, 1992; see figs. 2 & 3).



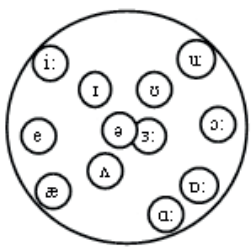


Figure 2. Spanish centrifugal vowel system.
Figure 3. English centripetal vowel system.

The centripetal system supposedly has greater diversity of vowels in both quality (impression) and quantity (length). There is also serious interaction between the primary stress and vocalic length. The centrifugal system, very much unlike the centripetal, has one degree of length (tenseness) that is half way between their short and long counterparts in the centripetal system. Thus, virtually, thousands of words in the form of minimal pairs in English are often confused such as in <pill>, <sin>, <kin>, <bit>, vs. <peel>, <seen> <keen>, <beat>, respectively. Minimal pairs such as the above are pronounced by Hispanics with one type of half-long vowel instead of [ɪ] for the former set and [i:] for the latter. In very many instances, such confusions are not just phonetically and/or phonologically confusing, but semantically as well; in fact, at times very embarrassingly. One such embarrassing instance was narrated in class by a Hispanic student. After winning a basketball game, she began yelling out what sounded for native English audience as: "We are the wieners...[i:]"[9] while the intended was: "We are the winners [ɪ]" This was a typical replacement of what sounded for the native speakers of English as an [i:] while the intended was an [ɪ].

Definitely, *neural phonology* stresses the significance of classroom implementation of *multisensory* strategies coupled with *multicognitive* ones aiming firstly at catering to different types of learners in classroom situations as well as

9 Obviously, in American slang, 'wiener' implies male sex organ.

reinforcing the targeted information through a variety of sensory and cognitive conduits. This is all aimed at enhancing the native L1 phonology domain by helping to acquire L2/FL *neuronemes* that are not part of L1 *neural phonology*.

The burgeoning of neural phonology in classroom situations evolved as a reaction to the needs of the learners. Their difficulties and failures compelled the instructor to rethink the overall approach to teaching cross-language pronunciation. This is an instructional scenario when the instructor learns from the learners and reacts to it. Therefore, the professional advice here is: *a smart teacher is the one who is prepared to learn from his/her students regardless of whether they err or excel*.

8 - Demonstrations to Reinforce Neural Phonology

Due to space limitations, only one example will be demonstrated in support of a multisensory and multicognitive reinforcement of L1 *neural phonology* when embarking on L2/FL learning. The selected example will be a vocalic one involving the English pair of [ɪ] as in <sin> = [sɪn] and [i:] as in <seen> = [si:n] vs. the Spanish vowel [i.] as in <sin> = [si.n] meaning <without>.

In dealing with vowels, quantity is generally meant to stand for length differences, whereas quality stands for different acoustic and/or auditory impressions (timbre) a vowel may have on the listener. However, length is a term that is somewhat controversial in that some phoneticians prefer to portray those differences in terms of *laxness* and *tenseness*– the former tends to be associated with shortness and the latter with length. To maintain a level of simplicity in handling the feature 'quantity', the dichotomy of short vs. long is preferred; however, this should not exclude the use of 'lax vs. tense' when and where necessary. In order to arrive at a relatively accurate identification of English vowels targeted here as in <sin> and <seen> versus the Spanish one in <sin> (without), Gimson's (1967) phonetic transcription is adopted. Accordingly, the English vowels are transcribed as [sɪn] and [si:n] vs. the Spanish [si.n]. Notice that the three vowels are different in both quantity and quality. Quantitatively, English [ɪ] is short and [i:] is long, whereas the single dot next to the Spanish vowel indicates 'half-length'. In other words, the Spanish vowel stands half-way between the English ones. Also, transcribing the English vowels with different symbols [ɪ] and [i] is meant to reflect a quality difference. Hence, the use of the same symbol [i] for English and Spanish reflects the shared quality, whereas [:] mark for English and [.] mark for Spanish signal the difference in quantity. As was hinted earlier on, it is because of this primary difference between the two vowel systems (centripetal vs. centrifugal) that Hispanic learners of English experience the most critical phonological problem with English vowels. They virtually squeeze the twelve English vowels into five thus generating not only momentous semantic problems, but also crucial interference with overall English rhythm which tends to be a *stress-timed* type as opposed to the *syllable-timed* of Spanish.

It is noteworthy pointing out that vowels, in general, do not yield themselves

readily to tactile/kinesthetic/proprioceptive (TKP) sensations as many consonants do; therefore, the instructor is primarily left with the auditory feedback with assistance from the visual feedback if the lips are distinctly involved such as different degrees of lip spreading vs. lip-rounding.

Below is a sketch of some of the recommended sensory and cognitive techniques used to enable learners overcome their problems. The exercises are planned to abide by the three phases of perception, recognition and production (PRP) rather than focussing immediately and exclusively on hearing/repeating as is the tradition with the majority of teachers of L2/FL throughout the world.

9 - Sensory Orientations

9.1 - Perception

This will include some tips for auditory and visual orientation to familiarize the learners with targeted vowels of L2/FL.

9.1.1 - Auditory Tips

a) Instructor carefully demonstrates the English vowels in the context of <sin> and <seen> as many times as deemed necessary.

b) Instructor inserts the Spanish word <sin> between the English ones and demonstrates the triplet thoroughly several times. Instructor has to be careful in demonstrating the difference as *precisely* and consistently as possible. Precision in demonstrating the differences is imperative because it may be tough to perform, especially by instructors with limited experience in practical phonetics.

c) Instructor *may* ask volunteers to repeat the demonstration in (a) and (b) to check the initial response from gifted and more sound-savvy learners. No correction is made if mispronounced since this is only a perception phase.

9.1.2 - Visual Tips

a) Instructor visually highlights the difference between the two English vowels and the Spanish one in the context of [sɪn] and [si:n] vs. the Spanish [si.n] using a short elastic band. The natural shape of the band should represent the English short vowel [ɪ]; with slight stretching, it should signal the Spanish vowel [i.]; and with more stretching, it should stand for the English vowel [i:] as visualized below:


= [ɪ], English short lax vowel

= [i.], Spanish half-long and somewhat tense vowel

= [i:], English long tense vowel

b) To connect the visual demonstration of the vocalic differences with the auditory ones, the instructor carefully demonstrates the three vowels several times with the elastic band in natural condition versus the other two conditions of mild stretching followed by somewhat greater one.

c) The above demonstration can be enacted somewhat differently by taking a

short step for [ɪ] →, a middle one for [i.] → and a longish one for [i:] →.

d) Instructor brings to the attention of the learners the difference in lip positions for the three vowels. There is hardly any sidewise stretching of the lips with [ɪ]; slight stretching with [i.] and more stretching with [i:]. The demonstration of each vowel is repeated three times in a row while pointing to the lips.

9.2 - Recognition Tips

a) Just to refresh the memory of the learners with the acoustic impressions of the three vowels [ɪ], [i.], [i:], the instructor demonstrates each vowel in context several times prior to initiating the recognition process.

b) Instructor designs an exercise in which each of the three vowels is repeated twice in context in *random* order: e.g.

[sɪt], [sɪt];
[si.t], [si.t];
[si:t], [si:t];
[si.t],[si.t];
[si:t],[si:t];
[sɪt], [sɪt] etc…

for a total of 12 pairs. Numbers the three vowels as: [sɪt] = #1; [si.t] = #2; [si:t] = #3. A work sheet has to be designed with 12 spaces and each time the instructor demonstrates a pair, learners mark it as 1, 2 or 3 on their work sheet. Work sheets are collected and checked to decide the degree of positive perception and recognition.

9.3 - Production

a) Once learners have had enough exposure to the three vowels and are fairly able to recognize the vowels at random sampling, it is time to initiate the production phase. However, just as a warm up, the instructor once again carefully demonstrates the three vowels in context.

b) Instructor asks for volunteers to produce each vowel *instantaneously* after his/her model demonstration in context. Oftentimes, it is quite likely for learners to succeed in accurate production due to *recency effect*. It is not unlikely for the immediacy effect to be a practice in 'hit or miss'; nevertheless, once the learner 'hits' the right production, the instructor has to ask the learner to keep repeating the targeted sound as many times as necessary. This repeated accurate production will certainly help the brain begin the process of impressing itself and neuronizing the targeted vowel i.e., enhance the *neural phonology* with the addition of a new *neuroneme*.

c) Instructor prepares a list of English monosyllabic[10] minimal pairs involving the targeted vowels with different consonants such as: <pill> vs. <peel>; <bit> vs. <beat>; <dip> vs. <deep> etc…It is extremely helpful to highlight the difference in meaning and associate it with pronunciation.

10 Easier for focus of attention than bisyllabic or multisyllabic words.

Finally, the instructor has always to be patient and bear in mind that some learners will take longer time to master a new *phoneme* that is alien to their L1 *neural phonology*.

9.3 - Cognitive Orientation

Field experience has shown that when the production of the sound is continuously repeated with very short pauses, such a demonstration may stimulate some thinking about the nature of the sound and leave a short-lived impression on the sensory memory.

a) With casual exposure, it is difficult to cognitively reflect on and retain an unfamiliar sound that is by definition alien to one's L1 *neural phonology*.

b) Obviously, the short-lived impression on the sensory memory may be wiped out in a split second, but with further exposure the impression could be tentatively shifted to short-term memory.

c) It is quite normal for the learner to forget the targeted L2 sound, but with continued exposure, it is likely to trigger a process of thinking about the nature of the new sound. This affords him/her the opportunity to internally compare and contrast the new sound with other sounds the learner already has in his/her L1 *neural phonology*.

10 - Conclusions

Teaching effective and efficient cross-language pronunciation to *adults* cannot be a casual practice on the part of the instructor; rather, it should be backed up with some knowledge in phonetics coupled with some field experience in teaching methodologies. Without such prerequisites, instructors are vulnerable to ineffective teaching away from the intended goals and objectives. It is such knowledge base that will enable an instructor to identify areas of phonetic and phonological similarities and differences between the targeted languages in light of which the overall approach will be designed and implemented.

For his practicals as a phonetician, the writer's orientation throughout his four years of graduate study at a reputable Department of Phonetics in 1970s was, unfortunately, exclusively focused, at the time, on the so-called 'ear-training' (auditory) in a wide variety of human speech sounds. However, *visual* and *tactile/kinaesthetic* orientations to human perception and recognition of sounds were barely considered, if ever. More importantly, no mention was made of *cognitive-orientation* (brain-training) in speech sounds *per se*.

Any instructor has to have an efficient approach to teaching cross-language pronunciation or at least has to gradually develop one. This writer familiarized himself with a few methods available in literature. Each one had positive and negative aspects which in both cases helped him throughout the years to gradually develop his own approach as is the case with MMA.

Further classroom implementation of MMA in teaching cross-language

pronunciation led to the emergence of the framework of *neural phonology* which necessitates diversified *sensory* input enhanced with *cognitive* input. Both jointly, add, modify and enhance the L1 phonological inventory beyond the exclusive reliance on listening/repeating especially in the case of *adult* learners. Learners of such ages, who, due to the perfect mastery of their L1 phonology, lose some of their sensitivity to unfamiliar sounds in the targeted phonologies of L2/FL. It is this condition that justifies the induction of *neural phonology* which promotes a troika of *listening*, *seeing* and *sensing* modalities functioning as three windows that stimulate the brain for more diversified action for retention. When those three sensory modalities are combined with other cognitive orientations in the form of thinking, associating, analyzing, synthesizing and even memorizing the gates of retention are widely open for adult L2/FL learners (Odisho, 2014; see also, Pinto, 2015). Indeed, the diversification of sensory and cognitive modalities of phonetic/phonological intake helps *adult learners* in their endeavor at cross-language enrichment of their L1 *neural phonology*.

MMA classroom application strongly supports the rationale for introducing *neural phonology* as a key area of focus for a more efficient approach to teaching cross-language pronunciation to *adults*. More importantly, *neural phonology* is introduced to attract the attention of *teachers* and *learners* to step beyond the traditional practice of 'repeat-after-me' which often fails to garner tangible improvement in teaching cross-language pronunciation especially to adult learners of L2/FL.

REFERENCES

Arnold, Magda B. (1984). *Memory and the brain*. Hillsdale, New Jersey: LEA Publishers.

Beck, Douglas L., and Flexer, Carol (2011). Listening is where hearing meets brain. *http://www. hearingreview. com/issues/articles/2011-02_02.asp.*

Carpenter, Harry (2004). *The Genie within: your subconscious mind*. San Diego: Anaphase II Publishing

Gimson, A. C. (1967). *An introduction to the pronunciation of English*. London: Arnold.

Handbook of the international phonetic association (1999). Cambridge: Cambridge University.

Larson, Christian D. (1912). *Your Forces and how to use them* (*http://www.sacred-texts. com/nth/yfhu/yfhu02.htm*).

Ladefoged, Peter and Maddieson, Ian (1995). *The sounds of the world languages*. Cambridge/Massachusetts: Blackwell.

Levitt, Robert A. (1981). *Physiological psychology*. New York: Holt, Rinehart & Winston.

Newport, Elissa (1975). *Motherese: The speech of mothers to young children*. Ph.D. dissertation, University of Pennsylvania.

Odisho, Edward Y. (1992). Transliterating English in Arabic. *Zeitschrift für arabische Linguistik*, Vol. 24, 21-34.

_____ (2003). *Techniques of teaching pronunciation in ESL, bilingual and foreign language classes*. München: Lincom-Europa.

_____ (2007/a). A Multisensory, Multicognitive Approach to Teaching Pronunciation. *Revista de Estudos Linguisticos da Universidade do Porto*, Vol. 2, 3-28.

_____ (2007/b). *Linguistic tips for Latino learners and teachers of English*. New Jersey: Gorgias Press.

____ (2014). *Pronunciation is in the brain not in the mouth: A Cognitive approach to teaching it*. Piscataway/New Jersey: Gorgias Press.

_____ (2016). The weight of phonological vs. phonetic accent in teaching pronunciation: Implications and applications, *Linguarum Arena*, Vol. 7, 31- 48.

Paraskevopoulos, Evangelos and Herholz, Sibylle (2013). Multisensory Integration and Neuroplasticity in Human Cerebral Cotex. *Translational Neuroscience*, Vol. 4, 337-348.

Pinto, Maria da Graça L. Castro (2015) *Pronunciation is in the Brain, not in the Mouth. A Cognitive Approach to Teaching it* (review article). *Linguarum Arena*, Vol. 6, Pp. 93-100.

Tortora, G. and Grabowski, S. (1996). *Principles of anatomy and physiology*,(8th ed.), New York: Harper Collins College Publishers.