

The Hidden Music of Language

Juan Carlos David*

Universidad de Buenos Aires
Universidad Tecnológica Nacional
I.N.S.P. Joaquín V. González

"The laws of harmony are the same for painting and music." (Henri Ravel)

"The world is full of resonances. It constitutes a cosmos of things exerting a spiritual action. The dead matter is a living spirit." (Wassily Kandinsky, *"On the question of the form."*)

"In your works, you have realized what I, albeit in uncertain form, have so greatly longed for in music. The independent progress through their own destinies, the independent life of the individual voices in your compositions is exactly what I am trying to find in my paintings." (Letter to Schönberg written by Wassily Kandinsky, 1911, after the performance of Schönberg's second string quartet and the "Three piano pieces")

"Colour is a power which directly influences the soul. Colour is the keyboard, the eyes are the hammer, and the soul is the strings. The artist is the hand that plays, hitting one key or another, to cause vibrations in the soul." (Wassily Kandinsky, *"Concerning the spiritual in Art,"* Munich, 1911.)

Perception refers to interpretation of what we take in through our senses, how we make sense of what we see, hear, taste, touch, and smell. The five senses have been developed with only one objective, namely sensing vibration, they have been developed to decode the same phenomenon. All the senses express a sensation in the form of electrical impulses, once a stimulus activates the different impulses, receptor cells synapse with neurons and pass on electrical impulses to the different areas of the cerebral cortex and it is the brain, in turn, which interprets the sensations. Responses to stimuli like pressure, light or chemical composition take the form of separate perceptions like touch, vision or taste only when they reach the brain. Hearing: ears evolved to decode sound waves. Sight: the outcome of the frequency of the light waves as perceived by the eyes. Touch: the feeling of things having different textures. This has to do with the different rates of vibration that the body perceives as compared to the objects one feels. Taste: there are six qualities to taste represented by the following adjectives: sweet, sour, pungent, bitter, salty and astringent. These qualities are related to the vibration of the molecules in the food we try. The senses of taste and smell differ only in the way we perceive volatile and nonvolatile chemicals, given that the same chemoreceptors are responsible for gustatory and olfactory perceptions Smell: the nose is a sensor to measure the vibration of the molecules suspended in the air that we breathe. The flow of energy when the brain is at work, also takes the form of waves, which as such, can be measured by means of an electroencephalogram.

Talking to a much-admired colleague, I came to discover that according to her, the study of phonetics bounds up with mathematics. This assertion came as a surprise to me, for I have always regarded phonetics as part of the realm of music. I clearly recall how dramatically this idea influenced my view of the way the subject is to be taught. We, as teachers, should cater for differences inherent in our students: different intelligences, and different strategies they use when apprehending reality. Pondering upon this predicament, I, for my part concur with the fact that music is applied mathematics. Therefore, my colleague's perception and mine differed mainly in our approaches, one based upon reason and science for the most part, the other based

* Teacher of English INSP "J. V. González". Head Teacher of Phonology I Translation Course Universidad de Buenos Aires (UBA). Lecturer of Phonetics at the MA in Translation and Interpretation UBA. Lecturer of English Phonetics I and II ISP "J. V. González", IES en Lenguas Vivas "J. R. Fernández" and ENS en Lenguas Vivas "Sofía E. B. de Spangenberg." Lecturer of Phonetics and Lab II INSP Universidad Tecnológica Nacional. Former lecturer of Phonetics at Universidad de Morón, Universidad del Museo Social Argentino and LCB. Former investigator at Universidad del Salvador. Jazz musician.

upon the development of a “gut feeling,” a term that admittedly might sound vague, as if lacking scientific value. This is precisely my point.

How is music generally learned? Mostly, the approach comes down to drilling. Studying our musical scales, learning how to read music, playing a given score, etc. As for me, I am mostly self-taught, as far as music goes. This means that my first approach was intuitive and therefore more ear-oriented. Only after I had learned something by ear could I possibly play that particular piece, besides I am interested in jazz, which is based upon improvisation. Music, as language, is based upon rules. This eventually led me into learning the language of music in a more formal manner.

Why is it that naturally gifted students, albeit endowed with a good capacity for imitating the phonemes of the target language, fail when it comes to phonetically transcribing paragraphs or taking dictations? The age-old feud between Arts vs. Science seems to permeate our teaching tradition and the conspicuous winner seems to be the latter.

In my opinion, this is a good point of departure when choosing a path that might be of help to guide our students into the acquisition and subsequent production of the given phonemes in the target language. Could students of a second language both possess (acquire) and know (learn) the phonemes of the target language? We tend to fluctuate between an “inspired-static” pole and a “rational conscious” pole. Neither of them can truly overcome the other.

There is a tradition of thought that can be traced back to medieval times: the theory of the “ghost,” this one defined as the impression that is left on the soul, or the senses. This impression is then “impregnated” in a stanza (a term to be defined in the following paragraph) which in turn construes its own identity, and is fathomed and possessed in its entirety.

According to Giorgio Agamben in his book “Stanzas” the poets of the XIII century interpreted the word stanza as the essential nucleus of poetry, the place where it dwells. But how can we possibly capture the very essence of it? There has been a scission between poetry and philosophy, between the poetic word and the philosophical word. Plato declared there was a long-lasting feud between them. Poetry possesses its object without knowing it and philosophy knows it without actually possessing it.

The point is that ultimately, in order to judge whether a phoneme is properly produced by our students, we teachers need to rely on our ear, our experience in the target language and phonetic criteria. Vowels are more associated with emotion, especially involved with feeling, with our biological and sensory appreciation of what is going on and what is happening to us. Consonants are more associated with reason and can consequently be described with reference to it. They stand for formality, a draconian spirit amid the musicality of vowels. Therefore, it is our capacity for discrimination that will eventually help us conclude whether our students’ rendering should be judged as either proper or faulty.

We should start by examining what sound is. Sound is a pressure fluctuation in the air. This process starts by one particle triggering the particles about it, which in turn, trigger other particles which follow a longitudinal direction. This continuous flow of particles produces a wave which is propagated through the air at almost 350 metres per second. The number of pressure variations (cycles) per second is called “frequency of sound”. The unit of sound frequency is Hertz (Hz). We human beings have an audible range between 20 Hz and 20.000 Hz (20 kHz).

We know that when we produce a vowel, some frequencies are amplified due to the nature of our vocal tract and resonators. The whole process is triggered the moment we produce a voiced phoneme, with the production of a fundamental frequency (FO). This is the frequency at which the vocal folds vibrate normally. The range of vibration for male adults is between 80 and 200 Hz, for adult females, 150 to 300 Hz and for children 200 to 500 Hz. Some frequencies above the FO are amplified. We can draw a similarity here between what happens in our vocal tract and the harmonics which are naturally amplified due to the acoustic chamber of some musical instruments. We know that when we pluck a string we perceive the fundamental note plus the harmonics associated to it, through a process known as synergy, we perceive only one note, but this note is the sum total of harmonics. When we produce a vowel, some multiples of the fundamental frequency are going to be amplified due to the shape and size of our vocal tract and resonators. When we produce /i:/ the vocal folds are vibrating at about 100Hz (G#₂ in musical

terminology). The first amplified frequency is going to vibrate at about 280 Hz (C₃ in music), that is the 3rd Harmonic, roughly 3 times above the FO. Some time ago phoneticians decided to label these given frequencies “formants.” If we think about these formants as the product of audible frequencies, they might become part of the realm of music simply by relating a musical note to a given frequency. Thus, the first formant in /i:/, for a typical male speaker, could well be related to a C sharp (Do #), given that the first formant has its peak of intensity at 280 Hz. And C sharp produces a sine wave that vibrates at 277.18 Hz. The second formant will in turn be related to an E (Mi) and the third one to G# (So sharp). In musical terms, we can safely conclude that the resulting chord will be a minor C sharp. At this point, an aforementioned notion calls for discussion, namely synergy: i.e.: the combined power of a group of things when they are working together, which is greater than the total power achieved by each working separately. It is through this nature of synergy that we perceive a vowel or a chord, not as a sequence of individual sounds, but as the sum total of them. The word interval is another term we need to discuss. According to the Cambridge Advanced Dictionary, an interval is: MUSIC (specialized) the amount by which one note is higher or lower than another

An interval of a 5th (= the top note four notes higher than the bottom one)

In western music we divide the scale into twelve notes, we must remember we have seven notes, namely: A-B-C-D-E-F-G. (A is LA, B is Ti and so on). The separation between one note and the other is called a step. The interval between A and B forms an interval of a second, between A and C, a third and so on. We also have sharp and flat notes. The twelve notes are then:

A- A# (or Bb)-B-C-C# (or Db)-D-D# (or Eb)-E-F-F# (or Gb)-G-G# (or Ab)

Octaves go on forever above and below our musical scale, they are simply halvings or doublings of frequency. It is established that mid A vibrates at 440 cycles per second, thus its name A 440, we have a lower A at 220 Hz, and a higher one at 880 Hz, you can go up or down the different scales until the musical notes are no longer perceived as such. To provide just one example, one drum beat per second is what middle C sounds like 8 Octaves down.

Why is it then that musical intervals sound right? We might assume that they reflect the sounds of our own speech: they reflect the inherent timbre of the vowels we use. Musical scales just sound right because they match the frequency ratios that our brains are primed to detect.

The other topic that needs to be brought into consideration is that of pitch. If vowels get their timbre from the formants, which in turn are the result of the amplification of certain frequencies in our vocal tract, why is it then that the quality of the different vowels remains stable when we change the pitch of our voices, either through intonation when talking, or even, in more extreme cases, when singing?

The answer is that although it is true that the pitch of the voice changes, and the timbre of the vowels depends primarily upon the formants, the intervals remain unchanged. That means that when varying the frequency of one formant, namely F0 (by stretching or loosening our vocal folds) a change is triggered in the rest of the formants. The important point to consider here is that the intervals should remain unchanged. The process is similar to the one that occurs when playing a chord on a fretless instrument (a double bass, a cello, a viola, etc.) and you slide down your fingers along the fingerboard keeping the hand position for the chord unchanged. The result is that the notes are obviously going to change, while the intervals will remain unchanged. (We are then playing the same chord on different keys). At this point it is relevant to highlight the fact that very few people possess the gift of perfect pitch, or absolute pitch perception, but practically everyone is naturally endowed with relative pitch, that is to say, we are able to detect the “wrong” notes in a melody, but if the tune is transposed from its original key, we tend to perceive the melody as unchanged.

It is important to note that perfect pitch is much more frequent among speakers of tone languages. In Cantonese, for example, people use from seven to nine different pitch variations as a means of distinguishing words that share the same sequence of phonemes.

At this point I find it might be helpful to establish a parallelism between the phonemes in a certain language and the different divisions of notes on different musical scales. What we might perceive as a note being played out of tune, might sound perfect on musical scales that have an additional note within a half step. This is the case with some types of music, such as Indian music. We need to remember that a step in music is the largest difference in sound between two notes which are next to each other in the western musical scale, for example from C to D we have a step, but from C to C#, we find a half step.

When trying to play a small non-western piece I happen to have overheard, I discover I have not been "wired" to detect and play notes with a smaller division than a half-step. I find myself correcting, so to speak, the tune as I have to play it. The point is that the perception of the pitch of a given note is acquired, according to my view, in the same way as the phonemes of a given language are acquired. It is not surprising then that students find it very difficult to tell the difference between two foreign vowels that in our native language belong to the same phoneme.

The same can be stated as regards colours. According to Benjamin Whorf's Linguistic Relativity Hypothesis experiences can be limited by language. Certain experiences cannot be fully expressed, because a given language might not have a word to refer to a particular concept. This language barrier may in turn affect perception of colour. The Inuits or Eskimos can detect many hues of white, in fact they have 17 words for white as applied to different snow conditions, which are understood as different colours, a difference which may go unnoticed for those who are not used to living in their environment and do not share their language.

A possible approach that we might aim at using when presenting the phonemes of the target language to our students, is a holistic approach, one in which more than one sense can be used to apprehend the nuances and intricacies of the target vocalic phonemes our students are to acquire in the classes of phonetics. Wassily Kandinsky proposed that there is a parallel to be drawn between language and painting. Lines or drawings are then to be understood as consonants and colours as vowels. One method to understand how our senses apprehend reality rests on the fact that we have five independently functioning senses. This view is known as the modularity thesis. The other viewpoint sustains that we possess one integrated sensory organ with five sub-organs (the unitary thesis). The latter is the theory that has been in vogue of late. Furthermore, it is the one that serves as the basis to sustain my approach to relate vowels, music and colours. We must bear it in mind that ancient cultures relate vowels, music and vibration to the different chakras, because for them sound is music and sound, good or bad, influences us and has an impact on our energy systems. Moreover, sound, music, musical notes chanting or vowel production, can be used for chakra balancing.

Interest in colour hearing goes back to classic Greek philosophy, when it was disputed whether colour, like pitch, could be considered as a physical quality of music. Pythagoras discovered the mathematical order of musical harmony by relating the length of strings to successive octaves. This led to the idea that colours and sounds could be related too, following mathematical rules.

Some scientists claim that we might be "pre-wired" for music, I prefer the term designed, in the sense that we could have been born with a type of universal grammar for music. The additional meaning of design, best understood in the analogous poem by Robert Lee Frost presents the idea of the unavoidable, or destiny. We might therefore assume that we humans have been then endowed and designed with an innate gift to develop language and music.

The term *Gesamtkunstwerk* (total work of art), first used in philosophy in an essay in 1827 was afterwards adapted and re-interpreted by the German opera composer Richard Wagner in 1844 in two different essays. He proposed that visual, auditory and namely all the senses can be part of a gestalt experience in an opera. Already during the Renaissance, artists such as Michelangelo did not perceive any real division between the different forms of art or, in some cases, science.

The word "Gestalt" translates as "whole" or "form." The idea underlying this principle is that the whole is more than the sum of the parts. In other words, our brain needs to organize information into groups to better understand it. There are different principles: one is the principle of "Perceptual Constancy." This principle maintains that we can apprehend and interpret information without unneeded repetition. We cannot possibly reformulate an object every time we come across it as if it were the first time, every single time you hear a vowel or a musical interval we relate it to the previous notion our brain has developed.

All these theories pave the way for my proposal of choosing a certain hue or shade of colour to represent every vocalic phoneme of the target language. This is not wholly capricious, but rather based upon certain correspondences between notes and colours.

As we know, notes are measured in Hertz. For example, when we speak of La 440, what we mean, is that the sine wave produced by this note vibrates at 440 cycles per second. By multiplying this number 41 times, we obtain another sine wave which cannot possibly be captured by our ear. This sine wave is then part of the realm of colour, which results in a hue of yellow and orange. We must remember that human beings can perceive from 9 to 12 octaves of sound, but only one of colour.

Of late, different studies have been carried out so as to establish the relationship between vowels and colours, among them we can mention Wrembel, M. & Grzybowski, A. and their publication *Achievements and perspectives in SLA of speech: New Sounds 2010*. Frankfurt, and a recent investigation carried out by Rob Drummond from the Manchester Metropolitan University. I, for my part, related the main vocalic formants to musical notes, these resulting notes, were then related to a certain hue of colour, at that time, I did not know of the existence of these aforementioned studies that bear a great similarity with my findings and intuitions. A possible representation of the different relationships between vowels and colours, is the following one:

English vowel n°1: a shade of yellow and green. English vowel n°2: a shade of light green and grey. English vowel n°3: dark green. English vowel n°4: orange. English vowel n°5: red. English vowel n°6: blue. English vowel n°7: dark blue/ brown. English vowel n°8: light brown. English vowel n°9: dark purple. English vowel n°10: a shade of grey and yellow. English vowel n°11: grey. English vowel n°12: grey.

The grayness of the central vowels reminds me of another quote by Wassily Kandinsky:

Clarity is a tendency towards white, and obscurity is a tendency towards black. White and black form the second great contrast, which is static. White is a deep, absolute silence, full of possibility. Black is nothingness without possibility, an eternal silence without hope, and corresponds with death. Any other colour resonates strongly on its neighbors. The mixing of white with black leads to gray, which possesses no active force and whose tonality is near that of green. Gray corresponds to immobility without hope; it tends to despair when it becomes dark, regaining little hope when it lightens. (Wassily Kandinsky, "Concerning the spiritual in Art," Munich, 1911.)

How can we possibly apply these notions in a pronunciation class? We might use colour cards with examples illustrating the different vowels with the given colour related to it. Students who are more visual, might benefit from these additional clues when it comes to acquiring the different vocalic phonemes of the L2 they are to learn.

We have every reason to suppose that articulate speech is one of the latest, as it certainly is the highest, of the arts acquired by man, and as the instinctive power of producing musical notes and rhythms is developed low down in the animal series, it would be altogether opposed to the principle of evolution, if we were to admit that man's musical capacity has been developed from the tones used in impassioned speech. We must suppose that the rhythms and cadences of oratory are derived from previously developed musical powers. (Darwin, 1871, p. 12)

