

# **Using the Voice as Means of Monitoring and Training the Level of Human Discourse**

by Dr. Ernesto Sh. Korenman<sup>1</sup> Prof. Yehuda Kahane<sup>2</sup> Dr. Liora Weinbach<sup>3</sup>

The human speaking ability distinguishes human beings from other creatures.

The human speech is modulated by the functional interactions of several centers in the brain. On the one hand, neural centers from the “ancient” brain whose activity support a “polished” version of “natural selection” expressing the hereditary weight of the species. This type of speech is mechanical, automatic and unconscious, and in humans serves to sustain the ego and its needs. On the other hand, there is another type of speech modulated by the frontal lobe activity of the neo-cortex. This type of speech is of higher degree of quality since it reflects the executive, analytical and integrative functionality of its neural source. Its output is empathic, creative and well adapted to highly dynamic environmental signals. An important role of education is to train people to use this type of speaking ability that progresses cooperation and integration rather than rivalry and conflict.

Communication is only one of the functions of speech. Our hearing capabilities, supplemented by an effective cognitive analysis, can frequently provide further information: we can identify interlocutors only from the voice, (even through the phone); we can know the quality of their feelings while talking, a hidden intention behind an external wording façade, or even realizing we are hearing a lie.... This is because, in addition to the literal content of the communication, the voice can frequently reveal the Psycho- Emotional- Mental state of the communicator.

The above discussed abilities have fascinated developers with the possibility to mimic them technologically. Amongst these attempts are: developments in voice personalized identification, voice content recognition, voice lying detection, voice-mediated diagnosis and treatment, etc.

Even with some degree of error, these methods and apparatus have managed, sometimes by means of machine learning algorithms), to provide some of the above extra information. This is frequently done based on pre-defined formant frequencies data and from their expected changes in amplitude and/or frequency as defined from the meta-analysis of large body of data from whole populations under similar conditions. Depending on the implementation, the starting point is usually a sample of a person voice, (talking, reading, singing or from sustained vowel emissions) of several seconds to, (usually), a minute or more of length. Variation of the voice signature during the transition between words or the presence of diphthongs and/or other voice transitions are also frequently used.

---

<sup>1</sup> Biophysicist, Psycho-physiologist, Bio/Neurofeedback Expert. CTO Uni-Talk Ltd. eskorenes@gmail.com

<sup>2</sup> Prof. Emeritus, Tel Aviv University. WAAS Fellow

<sup>3</sup> Head of the Unit for Ethics in Thought, Language and Action at the International Center for Health, Law and Ethics at Haifa University

Each individual voice produces specific resonant peaks of frequency in their voice signature, based primarily on their unique phonation apparatus (vocal cords & oral-pharyngeal cavities). This stable feature can be slightly modified according to the subtle application of barriers like the tongue, lips, etc.

The tongue, the lips and the whole swallowing vocal apparatus occupy a large and disproportionate proportion, (>20%), of the brain neuro-homunculus: the representation of the body parts in the brain cortex. These three structures are involved in vital activities such as feeding and communication. They richly interconnect with autonomic (unconscious) and somatic, (conscious) nervous centers that control and modulate general activity and behavioral expression. It has been said that “*what comes out of the mouth*”<sup>4</sup> is as important as, if not more important than, what comes into the mouth.

As with instrumental music, the “highest quality of the voice” depends not only on the mechanical parts of the vocal apparatus, (the instrument); but also depends on the ability of the performer, (the speaker), to generate the most sublime mental-emotional expressions (as “artistic sentic music forms” are played by Master instrumental soloists).

The “Switch my Speech” (SMS) technique developed by Dr. Liora Weinbach has been designed for the speaker to switch from unconscious to conscious speech, (see separate presentations by Dr. Weinbach and Prof. Kahane for details). Its training and evaluation require a tool to feedback information to the trainee to self-assess his/her performance.

For this purpose, a mobile phone-based application, (“I SMS”), has been developed. This application includes recording, tutorial and evaluation modules that teach the methodology and, (in addition), runs an algorithm able to accurately qualify and quantify statistically significant improvements in the quality of the voice.

A very short, (0.25sec), and stable vowel pronunciation can define a very accurately the personal voice signature and also characterize subtle changes in the voice recording disconnected from its signature. These changes properly defined, parametrized and processed enable our system to qualify and quantify four types of Harmony:

1. Mathematical Harmony: From statistical parameters from the Whole Spectrum Analysis, including fundamental and formant frequencies peaks and all their harmonics
2. Musical Harmony: From the degree of Musical Syntonisation with external musical scales, and from the Musical Syntonisation with the dominant frequency in the voice (internal music).
3. Natural Harmony: From the Attunement with Environmental Adaptive Cues, (which are known to affect brain and heart functionality), and from the Presence of Naturally Balanced Proportionality
4. Negentropic Harmony: From the Entropic Quantitative Analysis of spectral data.

---

<sup>4</sup> See Matthew 15:11 King James Version

Measurements of vocal performance before and after the application of SMS to resolve emotional or mental challenges in children and adults have shown large quantitative improvements in the quality of voice after SMS implementation (research still in progress).

The algorithm from, I SMS has been even able to detect changes in the quality of the voice after just thinking shortly, (few seconds), on polar concepts; for instance: “Conflict” vs “Calmness”. For instance, the application registered a huge, (264%), and highly significant, ( $p < 0.001$ ), increased of Overall Harmony by thinking “Calmness” compared to “Conflict”.

Once established, the processed data result can be used as a feedback marker to inform the speaker online on changes in the voice quality before and after SMS training or before and after any other type of behavioral or cognitive manipulation.

Results from the use of this application in bigger groups or even entire populations enable the assessment of their collective quality of speech, hence, providing for the first time a practical tool to train, monitor quantify and assess the quality of discourse of a whole group entity. This is to achieve universal social gain, (without discrimination of race, color, nationality, faith, background, culture, economical position, social status, political persuasion, etc.); and through the empowerment of every single individual that express his/her willingness to communicate consciously.

We trust that such an entrainment will help us all to improve the quality of communication and cooperation in intra-individual, inter-individual, group & social interactions and revert the present existential crisis in which we are all immerse into a more empathic and cooperative Human viable format because *“Death and life are in the power of the tongue”*<sup>5</sup>.

---

<sup>5</sup> Proverbs 18:21 American Standard Version (ASV)