ESSAY 42

Verbotonal Worldwide

Carl W. Asp, Madeline Kline, and Kazunari J. Koike

With warm affection, the first author acknowledges the late Dr. Sadanand Singh as an inspiration and a special friend.

BREIF HISTORY

In 1939, Petar Guberina completed his dissertation at the Sorbonne University, Paris, which established how vocal pitch change (intonation) affected the meaning of a phrase, and discussed how spoken language affected the learning of written language. His study became the theoretical foundation of a practical strategy called the Structural-Global-Audio-Visual (SGAV) method for teaching normal hearing people to speak before reading a foreign language (Renard & Van Vlasselaer, 1976). More popularly used in Europe, the methodology was later expanded to teach people with hearing impairment how to speak their native language by learning to listen first. This methodology has been coined “The Verbotonal Method.”

In 1963, Guberina was appointed visiting Professor at Ohio State University, Columbus (OSU). For four years, 1963 to 1967, Carl Asp was his clinical research assistant on a federal OSU research grant with Professor John Black (Black, 1963). Professor Guberina taught Verbotonal theory to all PhD students including Carl Asp and Sadanand Singh. Following this exposure, Dr. Singh published the 1981, Verbotonal Method for Communication Problems (Asp, Guberina, & Pansini, 1981, revised 2011) with College-Hill Press, and in 2006, Verbotonal Speech Treatment (Asp, 2006, revised 2011) with Plural Publishing. Dr. Singh established himself as a well-known academician, researcher, publisher, and philanthropist (Dr. Sadanand Singh Fund) in the field of communication disorders.

Interaction with Professors Guberina and Black along with fellow student, Dr. Singh, inspired Dr. Asp to pursue further research in Verbotonal theoretical principles, and in 1967, he acquired a federal research grant as a certified professor (CCC-A and CCC-SLP) at the University of Tennessee, Knoxville (UTK) (Asp, 1967). This research grant became a cornerstone of spreading Verbotonal. For the past 70 plus years (1939 to present), Verbotonal has developed in many countries with various languages and cultures. For example, the UTK Web site (Verbotonal.utk.edu) identifies...
INTRODUCTION

What is Verbotonal? It’s an auditory-based strategy that maximizes the listening skills of children and adults with hearing impairment, while simultaneously developing intelligible spoken language through binaural listening (Asp, 2006, revised 2011). The strategy implies the clinician’s use of refined listening skills to analyze the client’s error(s) and correct those errors with use of various tools based on theoretical and scientific evidence.

Verbotonal has five areas of applications: (1) children with hearing impairment (peripheral or central), (2) adults with hearing impairment, (3) diagnostic therapy, (4) speech-language disorders (e.g., articulation, stuttering, aphasia, autism), and (5) foreign languages (e.g., English as a Second Language, ESL). The theory and tools are adjusted to the severity of the problem. The clinical setting may be in public schools, private or residential schools, clinics, or even home schooling. Verbotonal focuses on therapy by the trained teacher/clinician. Parents are, of course, involved through individualized parent education, in order to maximize what the child has learned in the therapy session.

Five basic principles of Verbotonal are described in the following paragraphs in an attempt to explain why and how it is applied to the aforementioned areas with focus on application to the hearing-impaired: (1) neuroplasticity of the human brain, (2) vibrotactile phase of listening, (3) listening through rhythm and intonation, (4) error analysis and correction, and (5) listening through spoken language.

NEUROPLASTICITY OF THE HUMAN BRAIN

The brain of all species constantly reorganizes new neural connections throughout life. This maturational process is often referred to as neuroplasticity. Because of neuroplasticity, children five years old and younger are more easily adapted to learn to speak two different languages, since their neural maturational process is still evolving and, in a sense, is optimal (Piaget, 1973) to different neural stimulation (e.g., foreign spoken language). After puberty, however, the child’s brain becomes an adult brain that is less plastic for learning a new language without systematic teaching approaches (e.g., SGAV, Rosetta Stone, etc.).

Verbotonal considers this concept of neuroplasticity a very important principle in the application of therapy to the hearing impaired, and looks for what is optimal stimulation to each individual brain (Asp, Gubr, & Pansini, 1981, revised 2011). Hearing impairment is primarily a peripheral disorder; however, the rewiring of the brain’s neural connectivity is considered to be the ultimate Verbotonal goal. For example, it has been observed that young children or adults with high frequency hearing losses can learn to perceive low and high tonality (pitch) words (e.g., bow vs. cease) through the most sensitive frequencies below 1000 Hz. We theorize that this clinical phenomenon is possible because audi-
tory perception within the brain has been rehabilitated (rewired).

Table 19–1 in Essay 19 illustrates body movements and speech modification based on the seven perceptual parameters for error analysis and correction of binaural listening through spoken language.

**VIBROTACTILE PHASE OF LISTENING**

Figure 42–1 shows the human brain receiving neural information from the vestibular, auditory and speech channels. With the adding of proprioception, the human brain goes through developmental neuroplasticity as the developing infant matures to childhood and then adolescence. As an infant, the vibrotactile channel (vestibular) is a more dominant sense than hearing (cochlear), whereas hearing becomes more dominant in the older child. Verbotonal considers that the utilization of this early dominant sense (vibrotactile phase) is a foundation toward the development of good sound awareness, which is the critical step for later development of good spoken language through listening.

Figure 42–2 shows the Verbotonal Listen Control Panel. It has five channels: wide-band, low-pass, low-peaking, high-pass, and high-peaking filters. The center frequency and slope of a filter is adjusted based on the client’s error(s) while he listens and feels speech.

For example, Figure 19–1 in Essay 19 shows a parent using a body movement, a speech-vibrating board, and wrist-vibrator with her son. Figure 19–2 shows the first author using a body movement with three children, each wearing a speech-vibrator and a headset. These two figures demonstrate vibrotactile input and body movement.

Neurologic research supports that the vestibular-otoliths respond to sounds (i.e., vibrations) from 2 to 1000 Hz. As both the vestibule and the cochlea comprise the infant’s inner ear, one may say that there is vestibular-cochlear listening. The perceptual frequency range of both organs overlaps from 20 to 1000 Hz, where the cochlea dominates above 1000 Hz, although the vestibule dominates.

---

**Figure 42–1.** The child’s brain receives simultaneous auditory, vestibular, and speech input, with proprioceptive feedback from body movements. Used with permission from Asp (2006).
below 20 Hz (infrasonics) (Asp, 2006, revised, 2011).

In order to facilitate this vibrotactile phase of listening, a device called the Verbotonal speech vibrator responds from 2 Hz to 1000 Hz. It can be attached to the infant’s wrist and/or a vibrating board that stimulates the infant’s brain through the vibrotactile sense. Figure 19–1 in Essay 19 shows a picture of a parent using body movement while her two-year-old son feels/hears her speech through a speech vibrator on his wrist and one on a vibrating board against his lower body. With the use of the more dominant vibrotactile sense, it becomes easier and optimal for the infant even with profound hearing

![Verbotonal five-channel control panel. Used with permission from Asp (2006).](image-url)
loss to become aware of sounds (vibrations). Furthermore, the infant feels his mother’s speech rhythms, as he did in his mother’s womb. This helps his babbling advance to rhythm and intonation patterns that become meaningful, as discussed in the next section. Table 19–1 in Essay 19 illustrates body movements and speech modification based on the seven perceptual parameters to correct errors.

LISTENING THROUGH RHYTHM AND INTONATION

Rhythm and intonation (suprasegmental or prosodic) is the foundation of both listening and spoken language. In a conversation, changes in rhythm and intonation and body language can affect the meaning of a phrase so that “no” may mean “yes” or “yes” may mean “no,” depending on how the rhythm and intonation pattern of the word is altered. Natural vocal-pitch changes in conversation go an octave lower and higher than the mean vocal pitch. For example, when we say, “I cannot go out; it’s raining” (Guberina, 1939), emotional vocal-pitch changes (intonation) alters the meaning (language) of this phrase (speech). The semicolon is a pause, created by the vocal pitch rising with tension to let the listener know there is more after the pause. Therefore, an upward inflection at the end of “go out” can mean “Yes, I dare go out,” as if the child is questioning the authority of the mother. According to Mehrabian (1969), verbal components (lexicon and syntax) are only 7% of human communication, whereas 93% (38% and 55%, respectively) are from suprasegmental components (e.g., rhythm and intonation) and nonverbal components (e.g., facial expressions, eye contact, gestures, etc.). Verbotonal considers that it is important to incorporate this observation into actual habilitative and rehabilitative therapies, emphasizing the development of good rhythm and intonation patterns in the early stage of spoken language.

In both adults and children, speech energy for rhythm and intonation typically lies below 300 Hz. Most children and adults with hearing impairment are observed to have hearing sensitivity below 300 Hz in the audiogram. Therefore, in order to ensure that children with profound deafness are able to perceive rhythm and intonation patterns, a Verbotonal training unit that can amplify the speech frequencies below 300 Hz effectively, with a wide frequency response bandwidth (2 to 20,000 Hz) that includes both extended lows (2 to 300 Hz) and extended highs (3000 to 20,000 Hz), has been designed and utilized for therapy. This is why most people who are deaf (90 dB) in the classical speech frequencies (300 to 3000 Hz), can learn to talk and listen with good rhythm and intonation and intelligible spoken language. Especially in the early stage of habilitation, Verbotonal training units may even include speech vibrators and headphones to create optimum listening to good rhythm and intonation, which ultimately helps to stimulate neural maturation in the brain. In this context, Verbotonal can be characterized as a multisensory strategy for rehabilitation of the brain.

In 1940, the AT&T Bell Research Lab determined an auditory memory span of eight syllables in 1.8 seconds, separated by a pause, was an essential part of a phone conversation. For example, a phone number 523 (pause) 0996 is easy for a normal listener to remember. The pause is an essential part of auditory memory. During a typical conversation, normal speech rate is five syllables per second. Each syllable has a vowel, with one or two consonants. The phoneme rate of a talker is 12 phonemes per second, 720 per minute, and 43,200 per hour. In a long conversation, the listener automatically “locks on” to the syllable stress and intonation patterns of the
talker, because the normal speech rate is too fast to “lock on” to all the phonemes. This is another reason why Verbotonal stresses listening through rhythm and intonation, that is, to enhance auditory memory to at least eight syllables.

**ERROR ANALYSIS AND CORRECTION**

American English has 25 consonants, and 15 vowels and diphthongs. Guberina (1972) proposed to use octave-band filters to determine what would be the most important frequency region for the perception of a particular phoneme (vowel or consonant). In his experiment, when he set the octave-band filter’s center frequency at 200 Hz, the listener heard the recorded vowel /i/ as the vowel /u/. When the center frequency was changed to 1000 Hz, the vowel /i/ was heard as the vowel /a/, and so forth. Therefore, he concluded that all phonemes are contained in one phoneme, because the perception of different phonemes (e.g., the vowels /i/, /e/, /a/, /o/, /u/) was possible from the same phoneme (e.g., the vowel /i/) depending on which filter frequency in which the phoneme was passed. Guberina further defined each of these various filter settings as an optimal octave for each vowel or consonant. In practice, this principle is a helpful tool in analyzing errors, that is, why a certain error occurs with an individual with a certain audiometric pattern. For example, it would explain why a child with only low frequency hearing below 300 Hz, will say the vowel /u/ when the vowel /i/ is presented as a stimulus. The vowel /i/ is filtered through the child’s hearing loss that is equivalent to a low pass filter, optimal for perceiving the vowel /u/.

Based on such research with optimal octaves, Guberina (1972) constructed a sequence of eight nonsense syllables, with their optimal octaves ranging from a low frequency to a high frequency continuum and defined them as *logatomes* (bru-bru, mu-mu, bu-bu, vo-vo, la-la, ke-ke, shi-shi, and si-si) (Asp, Guberina, & Pansini, 1981, p. 7). In practice, an abbreviated sequence of logatomes “mu-mu, la-la, si-si” has been used by Verbotonal clinicians to evaluate the detection and perception of these sounds through auditory training units, hearing aids, and cochlear implants.

Furthermore, in a paired-comparison experiment by Peterson and Asp (1971), normal-hearing listeners judged the unfiltered prevocalic consonant /s/ as higher in spectral pitch than /m/ and so forth. In 1979, Bessel and Asp, in a paired-comparison experiment, found that normal hearing listeners judged the unfiltered word “cease” as higher in pitch than the word “move.” In 1980, Bessel and Asp acoustically analyzed 30 words judged to be different in spectral pitch. Both the consonant and the word results support a low to mid to high spectral-pitch continuum. At the syllabic and word level, this spectral pitch was defined as tonality.

Similarly, words were separated into low, mid, and high tonality categories (Asp, 2006, revised 2011). For example, low, mid, and high tonality words would be, “Moon-Hat-Cease.” During diagnostic therapy, the appropriateness of a hearing aid’s frequency response can be diagnosed by analyzing tonality errors. At the same time, possible adjustment may be made to the frequency response and/or gain of the hearing aid. For example, the clinician presents one tonality word, and each time the client says what is heard, without visual cues. It is commonly observed that when the client has a high frequency hearing loss above 2000 Hz, a high tonality word such as “cease” can be mistakenly perceived as words from the lower tonality category, for example, “use” or “ooze.” Analysis of errors enables the clinician
to provide further auditory therapy with use of the Verbotonal training unit that has adjustable acoustic filters and slopes to enhance the correct perception of the test words and/or to adjust the settings of the hearing aid to optimize the perception of high tonality sounds.

One of the unique treatment tools Verbotonal uses in the correction of speech errors, either at the suprasegmental level (rhythm and intonation) or at the segmental level (phoneme, word, and sentence), is body movements. A detailed explanation of why and how body movements are utilized is available in Essay 19.

LISTENING THROUGH SPOKEN LANGUAGE

The ultimate goal of Verbotonal is to develop good, intelligible spoken language among children and adults with hearing impairment, so that they can communicate and be educated among normal hearing peers. As described above, Verbotonal initially utilizes multisensory input (vibrotactile, vision, audition, and proprioception in the form of body movements) for rehabilitation of the brain. In the process, however, audition (listening) must be established as the primary sensory input in learning spoken language.

Verbotonal evolved from Guberina’s observation of how people learn a foreign language. The process of teaching children with hearing impairment is almost analogous to how hearing children learn their native language. Infants begin with a babbling stage, where they begin to imitate the mother’s utterances, feel affectionate changes in rhythm and intonation, and see facial expressions and gestures. As they mature, they begin to form wordlike utterances, where the meaning of spoken language begins to be defined within various contextual cues. Progression of teaching various phonemic utterances also follows the developmental progression of normal hearing children (e.g., the acquisition of words like “mama” and “baby” occurs earlier than words like “sister” and “scissors”) (Asp, 2006, pp. 152–153). Interestingly, low tonality words are generally developed earlier than high tonality words. Just as slides and DVD are used in the teaching of foreign language courses (e.g., SGAV), the use of situational teaching with audiovisual materials is frequently incorporated into therapy sessions with hearing-impaired children.

SUMMARY

The five Verbotonal areas and five principles above comprise the Theory. Essay 19 describes the Treatment Tools, with emphasis on Verbotonal Body Movements.

Acknowledgments. The authors appreciate the editing by Mary Koike, Wayne Kline, and Jan Asp.

REFERENCES


