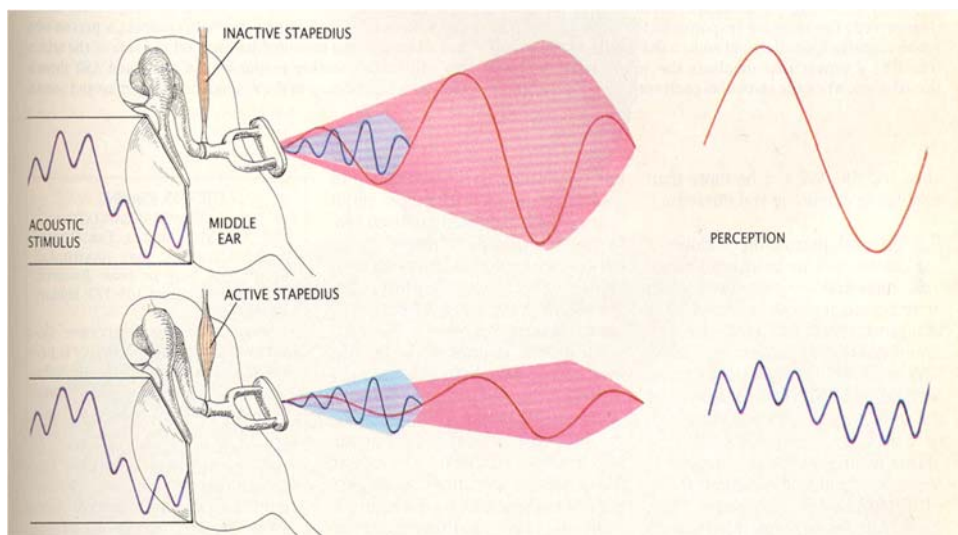


## Neurophysiological Background for the Safe & Sound Protocol

The Safe & Sound Protocol (SSP) is an innovative intervention designed to improve an individual's social communication behaviors by reducing hearing sensitivities (i.e., hyperacusis) and improving the ability to process human speech. The intervention is based on the Polyvagal Theory (Porges, 1995, 1997, 1998, 2001, 2003, 2007) and is designed to enhance the processing of human speech by dampening the masking and interference of background sounds. The intervention is theoretically driven by scientific evidence relating the regulation of the middle ear muscles to: 1) dampen background sounds and improve perception of human speech; 2) neuroanatomical and neurophysiological circuits controlling facial expressions, vocal intonations and gestures; and 3) neural circuits regulating behavioral state. [See Porges and Lewis (2010) for scientific basis of intervention and see Porges et al. (2013, 2014) for peer-reviewed publications of application of the intervention to children with autism spectrum disorders.]

The middle ear muscles actively dampen low frequency background sounds and facilitate the ability to hear and to understand human speech (see Borg & Counter, 1989; Zwislock, 2002). When the middle ear muscles do not contract appropriately, individuals tend to have sound sensitivities and difficulties in understanding speech in noisy environments. This apparent problem has an adaptive advantage and functionally amplifies the ability to hear very low frequency sounds (from 20 - 100 Hz, whereas human speech is in the 500 - 4,000 Hz range), which through the evolutionary history of mammals has been associated with danger and predators. Thus, if middle ear muscles do not appropriately contract, low frequency sounds (below the level of human speech) will be perceived as loud, even when others, whose middle ear muscles contract, can barely hear them. Illness, fever, and aging also reduce the function of the middle ear muscles and result in associated difficulties of understanding human voice in noisy environments. In addition, psychological states of fear also adaptively shift the function of middle ear muscles to promote states of hypervigilance in anticipation of a predator, while compromising the ability to process the meaning of human speech.

Figure 1 illustrates how the stapedius (the primary middle ear muscle) attenuates low frequency sounds to



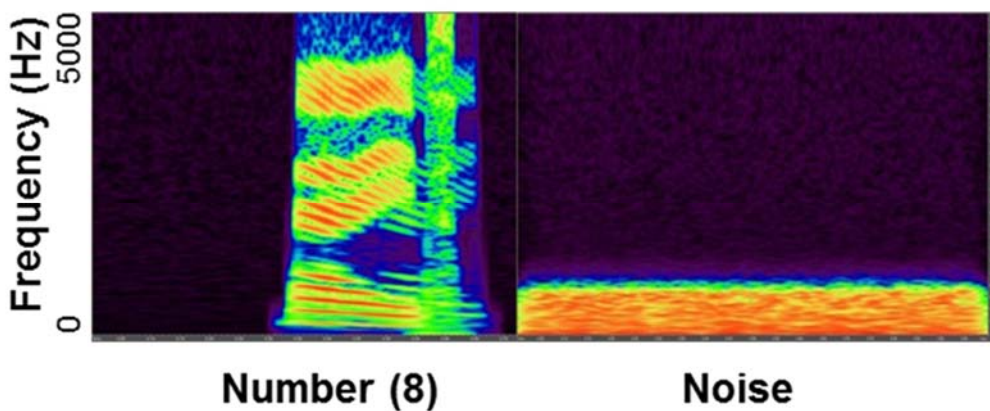
facilitate the processing of human speech. The pattern on the left side of the figure illustrates an acoustic signal representing the composite of speech (high frequency and low amplitude) superimposed on background noise (low frequency and high amplitude). The complex acoustic information impinges on the eardrum and is initially filtered by middle ear structures prior to being transmitted through the inner

ear on its journey to the brain, where it is perceived.

The top part of the figure illustrates how speech is lost in background sounds when the stapedius is inactive. An

inactive stapedius releases the tension on the eardrum (similar to loosening the tension on a kettledrum) and the low frequency sounds of the environment, instead of being attenuated by the middle ear structures, are transmitted to the auditory cortex via the auditory nerve from the inner ear. In this case, the information that reaches the cortex is dominated by low frequency sounds and human speech is masked. In this example, the loud low frequency sound is processed and detected, but the features of the softer higher frequency sounds of human speech are lost in the background sounds. In the bottom part of the figure, the stapedius actively functions to dampen the impact of low frequency sounds. As the middle ear muscles are activated, the rigidity of the ossicle chain in the middle ear increases and the eardrum becomes tighter. Metaphorically, this is similar to striking a kettledrum. When the head of the drum is tightly pulled, the sounds have a higher pitch and when the tension of the head of the drum is lessened the sounds have a lower pitch. Thus, similar to the mechanics of a kettledrum, the middle ear muscles have a capacity to ‘dampen’ low frequency sounds (i.e., background noises) to enhance the processing of speech. The SSP was developed to rehabilitate the function of the middle ear muscles by providing a sequence of acoustic stimuli designed to enhance the neural regulation of the middle ear muscles.

The acoustic environment is frequently dominated by acoustic energy in frequencies lower than human speech. Figure 2 illustrates the frequency representation on the left of a word (the number “eight”) and background



sounds on the right. Note the overlap of the low frequencies and the unique information in the frequency band above 800 Hz. This

Figure SEQ Figure \\* ARABIC 2: Frequency

figure emphasizes the importance of detecting higher frequency content in speech. Understanding speech is dependent on processing not only the

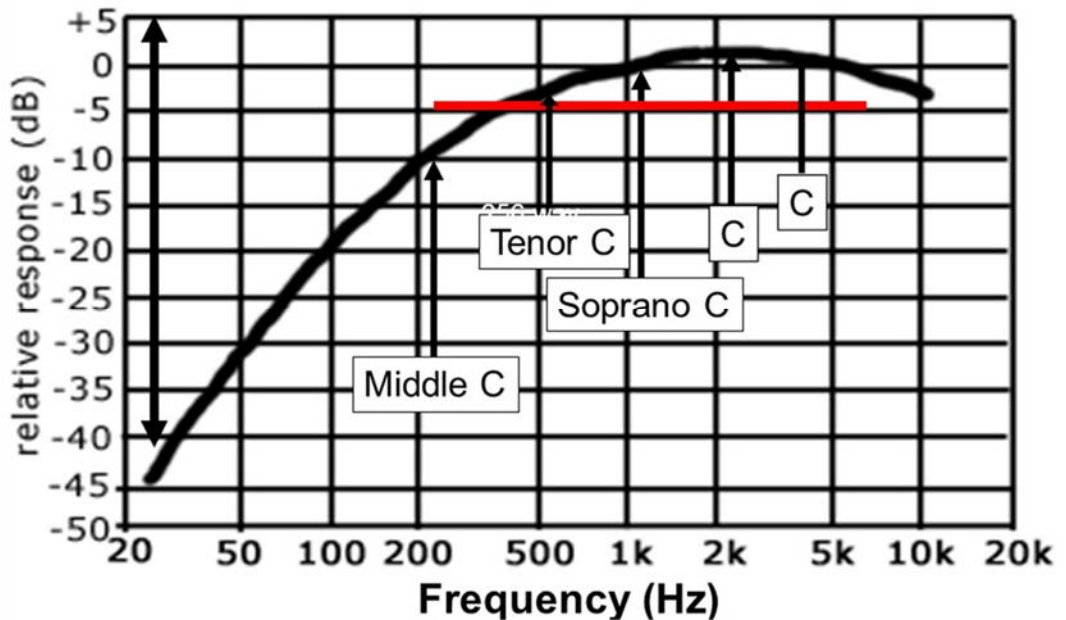
lower frequencies of vowels, but also the higher frequencies of consonants. Thus, when middle ear muscles are NOT functioning appropriately the subjective experience is hearing speech without being able to decipher the words, especially when consonants (i.e., higher frequency sounds) are at the end of the word.

Normal ears in mammals evolved to optimize the processing of vocalizations. Based on the transfer function of middle ear structures, mammals have a *frequency band of perceptual advantage* (see Porges & Lewis, 2010). In humans, this coincides with the frequencies that have been used to define the index of articulation (see Kryter, 1962) or the speech intelligibility index (see American National Standards Institute, ANSI, 1997). In Figure 3, these frequencies are emphasized by the red line. Figure 3 illustrates the A-weighted decibel scale, which attempts to describe relative loudness to a normal human ear. The Y-axis represents the difference between the acoustic energy (i.e., sound pressure level) and the perceived loudness. Note that the frequency band described by the red line is close to a value of ‘0’ on the Y-axis. This means that normal ears do not attenuate sounds within this frequency band. However, the acoustic energy of the sounds (i.e., sound pressure level) at low frequencies are greatly attenuated. In Figure 3 low ‘bass’ frequencies are attenuated up to 45 dB relative to the

sound level of voice. This attenuation is due, in part, to the physics of the ear and the function of the middle ear muscles. Thus, individuals with poor middle ear muscle tone will experience low frequency sounds as louder (and often painful) than individuals with optimal functioning middle ear muscles.

Figure 4 provides an illustration of how typical adults judge the relative loudness of tone at specific frequencies. The data were generated in Dr. Porges' laboratory and illustrate that a 30 Hz tone at 100 dB is judged to be the equivalent loudness to a 1000 Hz tone at 60 db. In both Figure 3 and 4, note that relative 'flatness' in the frequency band between about 500 Hz and 4,000 Hz. This is the band in which virtually all the information related to social communication, including speech, occurs.

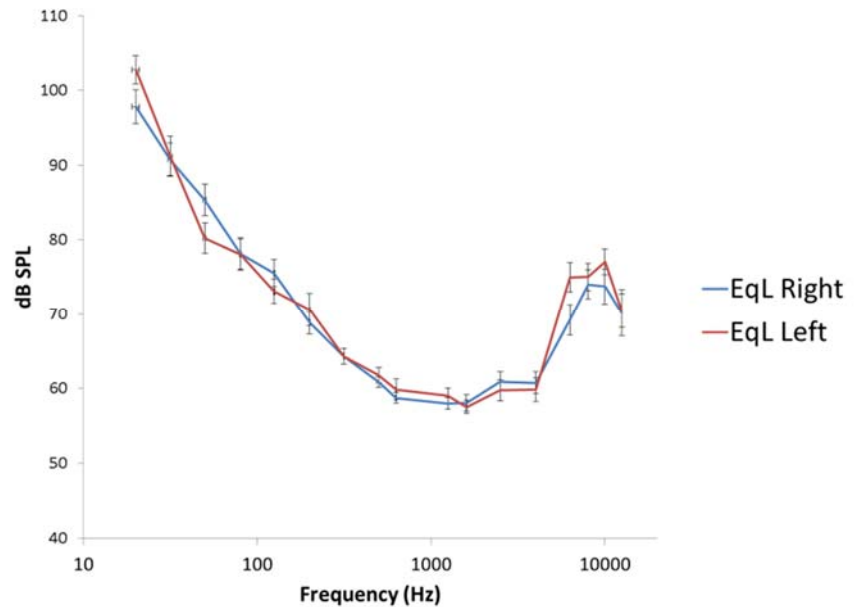
Our personal experiences confirm that there are individual and maturational differences in the neural regulation of the middle ear muscles. These differences contribute to difficulties in understanding human speech in noisy environments.



We can understand these differences in middle ear function by observing the "natural" experiment that occurs in many families with typically developing children. Typically developing adolescents will adjust the bass of a car radio to deal with the attenuation of low frequency sounds when music has a vocal track (i.e., voice). Many of us have experienced the booming of the car bass and the difficulties we have hearing vocal tracks or understanding when someone in the car is talking to us. With typically developing adolescents, the middle ear muscles are so responsive to human voice that when the vocal track is played, their middle ear muscles contract and greatly reduce the amplitude of the lower frequencies in the music (consistent with the Figures above). To compensate for this attenuation, the child will normally turn the bass control up to amplify the low frequencies in the music. For many of us, when we are monitoring our children while driving, we are in physiological states that support our hypervigilance (this physiological state downregulates the neural tone to the middle ear muscles to be prepared for danger and not conversation). Once the bass is turned up, the low frequencies trigger our hypervigilance as our body associates the low frequencies with a potential threat. The loud low frequency sounds prepare us to detect predator (signaled by low frequency sounds) and we can no longer hear the vocal track or individuals talking. Thus, when we are bothered by the loud low frequency sounds and the child is not, we are experiencing a moderate auditory hypersensitivity to low frequencies that interferes with our ability to perceive and understand human voice. There are individuals in whom

this system is greatly compromised, so that vacuum cleaners, ventilation systems, escalators, traffic, airplanes and even elevators can be perceived as being loud and painful.

The ability to recruit the middle ear muscles to contract and dampen background sounds is compromised with aging. Also, the neural tone to these muscles is greatly decreased in fearful situations, during high fever, or even during states of physiological or psychological stress. This decrease in muscle tone enhances the detection of low frequency sounds, sounds often associated with predators. However, this 'adaptive' stress response makes it difficult to perceive and understand human speech. Thus, under stressful conditions we have difficulty in listening to the specific words of others, but have no difficulty in focusing on the low frequency sounds that may accompany an intruder or a potential threat. Most people experience this phenomenon in "unsafe"



environments. For example, during walks into unfamiliar and potentially dangerous neighborhoods, our listening shifts from our companion's voice to the low frequency sounds of footsteps or traffic. Similarly, when left alone in a house, individuals often report sounds such as creaking of floor boards or shutters that are seldom heard in the presence of friends. In addition, aging, illness, medication, trauma, and experience may also compromise the function of middle ear muscles.

The Polyvagal Theory focuses on how function and structure changed in the vertebrate autonomic nervous system during evolution. The theory is named for the vagus, a major cranial nerve that regulates bodily state. As a function of evolution, humans and other mammals have a "new" vagal pathway that links the regulation of bodily state to the control of the muscles of the face and head including the middle ear muscles. These pathways regulating body state, facial gesture, listening (i.e., middle ear muscles), and vocal communication function collectively as a Social Engagement System. Because the Social Engagement System is an integrated system, interventions influencing one component of this system (e.g., middle ear muscles) may impact on the other components.

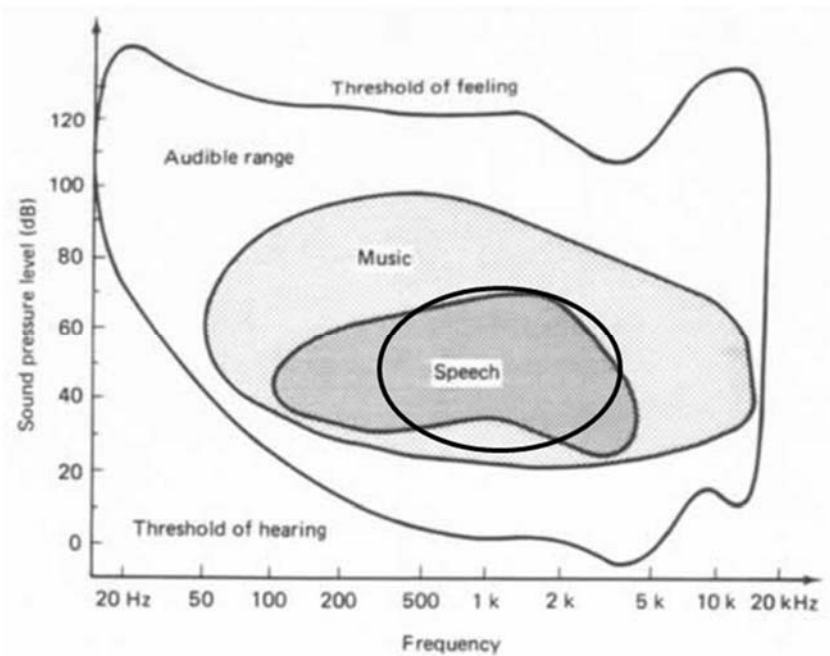
The middle ear muscles are the smallest muscles in the body. These muscles control the rigidity in relation to each other's mobility of small bones in the middle ear. The rigidity of these structures determines the intensity of low frequency information reaching the inner ear and the brain. Difficulties in regulating the middle ear muscles increase the loudness (i.e., energy) of low frequency sounds that reach the inner ear and overwhelm the relatively soft sounds that characterize the human voice. This potentially results in auditory hypersensitivities (especially to low frequency sounds) and difficulties in understanding human voice in noisy environments (see Borg & Counter, 1989).

Physical exercise and other rehabilitative strategies provide a useful metaphor to explain how the SSP works. With the appropriate monitoring by a coach, training exercises may be matched to fitness level. If the training exercises are too intense, fatigue and injury may occur. If the exercise demands are appropriate, neural regulation of the muscles will improve and the muscles will become stronger and more flexible. Since the middle ear muscles are primarily fast twitch and fatigue easily, care must be taken in designing the SSP. A great deal of research has gone into the SSP design to insure that the daily sessions do not result in inappropriate fatigue.

Other everyday experiences provide additional insights relating the exercise metaphor to listening. Musicians, vocalists, and actors develop skills to monitor the sounds from their instruments or voices. They learn “listening” techniques to focus on their own voices or musical instruments, even in a background of other sounds. They learn to identify specific auditory signals by concentrating on other voices and instruments. Listening provides an important venue to extract critical information regarding the world. Our success in adapting to the continuously changing and challenging acoustic environment mandates an ability to rapidly extract accurate information, even in environments characterized by loud background sounds.

In Figure 5 (next page) the acoustic environment is mapped on two dimensions, pitch (frequency) and loudness (sound pressure level). The Figure is included to emphasize the overlap in acoustic features of speech and music. Note that music has frequencies that extend beyond speech and may be louder than speech. Moreover, not all speech provides signals of safety. Speech, as well as music, can signal danger and life threat. Our nervous system universally detects high frequency shrill cries as alerts for danger and this is mimicked in music. In addition, our nervous system detects monotonic low frequencies as sounds of threat. This may at times be conveyed by male voices that can terrorize a young child. In both examples, loudness ramps up the bodily response.

A segment of the frequencies of human speech defined as the band of perceptual advantage (see Porges & Lewis, 2010) is capable of triggering bodily states of safety when modulated to mimic the prosodic features of a mother’s calming voice. These features characterize a mother singing a lullaby. **The circle in the figure illustrates the approximate frequencies and loudness of the stimuli used in the SSP.**



The information we extract from our environments is broader than the explicit physical features. Successful people also accurately evaluate the emotions and intentions of others. Thus, a child’s developmental trajectory is, in part, dependent upon the ability to detect relevant social cues from human speech, gestures, and facial expressions. These skills are, in part, dependent on our ability to dampen irrelevant and often noisy background

sounds and to accurately detect feelings and intentions of others and to communicate feelings and intentions to others with socially appropriate cues and signals (e.g., facial expressions, head tilt, eye movements, vocalizations and body posture.)

The Polyvagal Theory provides the neurophysiological insight to train the neural regulation of the middle ear muscles. Based on the theory, this training should optimize not only verbal communication, but should also improve the regulation of affect through facial expressions, intonation of voice (i.e, prosody), gesture, and autonomic state. The SSP, informed by the Polyvagal Theory, was developed to use a complex program of acoustic stimulation to exercise and systematically challenge the neural regulation of the middle ear muscles.