

Linguistics 101 for Hearing Healthcare Professionals

Published March 10th, 2016

Marshall Chasin, AuD

The Speech Intelligibility Index (or SII) shows some interesting characteristics (Figure 1). The difference between 340 Hz and 3400 Hz (coincidentally the bandwidth of the telephone) is fairly similar for continuous discourse, whereas there is a high frequency bias for nonsense syllables. That is, whenever there is a context to speech, there is a greater reliance on lower frequency sounds.

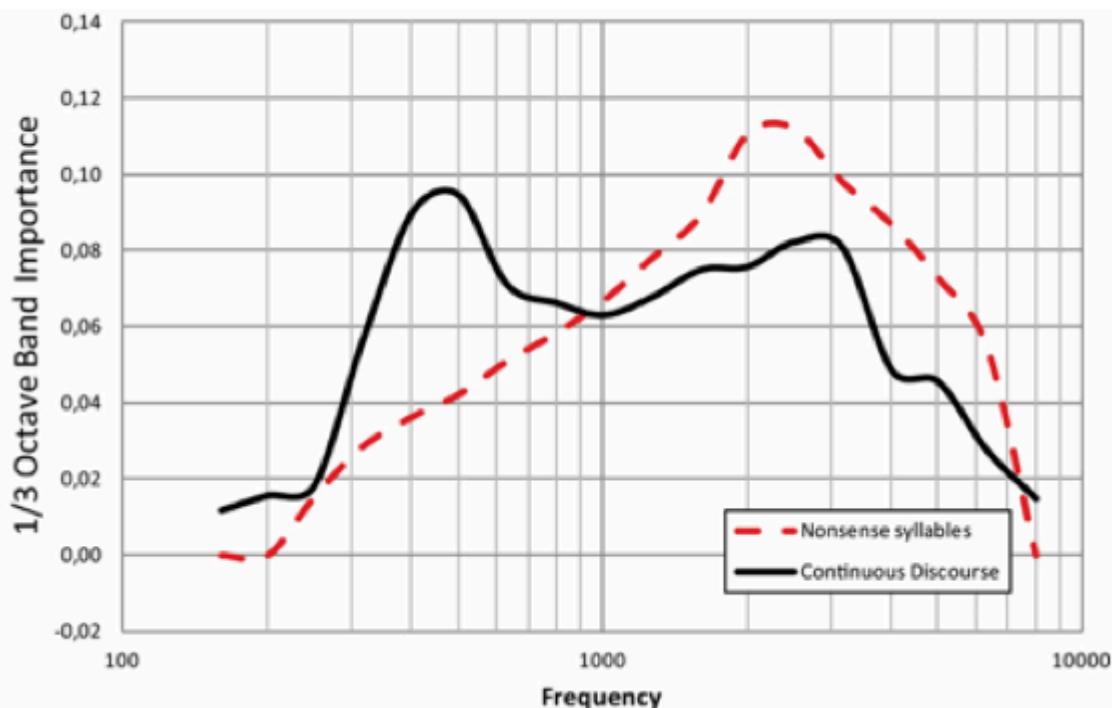


Figure 1. The Speech Intelligibility Index (SII) for non-sense syllables (dotted red) and for continuous discourse with context (solid black). Note the broader band, and lower frequency intelligibility contributions for continuous discourse. Interestingly enough, and certainly not a coincidence, the black solid line is also exactly the spectrum for the telephone.

Continuous discourse with context is much more often the case than is listening to nonsense syllables, so it stands to reason that the SII black-colored curve for continuous discourse in Figure 1 should be used to optimize speech communication rather than the dotted red curve that we mostly learn about in audiology classes. This is certainly not new material, as Drs Robert Sherbecoe and Gerald Studebaker studied this topic over a decade ago.¹

But let's take a step back and re-examine those consonants that have lower frequency cues. It would be erroneous to say that vowels are found mostly in the lower frequency region, and

consonants in the higher frequency region.

One of the first things to be learned in any first-year linguistics class is the division between sonorants and obstruents. Sonorants are those sounds that are “periodic” and have a well-defined formant resonant structure. These are the vowels, the nasals, and the liquids (/l/ and /r/). Despite being consonants, these nasals and liquids have most of their energy in the lower frequency regions. In many cases, it is not the exact frequency of the formants that is important, but the difference between the first (F1) and second (F2) formants that define the auditory cue.

And what about obstruents? These are the sounds that do not have a formant structure and are typically, but not always, characterized by high frequency hissing or noise in their energy spectrum. The high frequency obstruents are the sibilant fricatives (such as /s/, /z/, /sh/, /th/) and the affricates (such as, /ch/ and /j/). But obstruents can be low frequency as well, and these include the various stops (such as /p/, /t/, /k/).

It is true that the voiceless stops such as /p/, /t/, and /k/ have high frequency aspiration cues, but these are not typically found in average conversational speech, especially at the end of words. In a word such as /pot/, the final /t/ is unreleased and does not have any high-frequency cue, such as aspiration, associated with it. Phonetically this is marked with a small half square symbol called a diacritic, such as /t~/.

Table 1 provides a short list of some sounds with their various auditory cues. The low frequency obstruents have their auditory cues on how they affect the adjacent vowel. A voiced low frequency obstruent, such as the phoneme /d/, has its primary cue on the first and second formants, and also by making the vowel before or after it slightly longer. In languages such as English, this is actually a more important cue than in languages such as Russian; however, even in Russian, the length of the vowel still provides significant voicing information.

PHONEME	Formant F1 (F1)	Formant F2 (F2)	F2-F1 Difference	Adjacent Vowel Length	High Frequency
a	500 Hz	1500 Hz	1000 Hz	-	-
i	280 Hz	2200 Hz	1900 Hz	-	-
u	280 Hz	900 Hz	620 Hz		
s	-	-	-	-	3500 Hz +
b	Decrease F1	Decrease F2	-	lengthened	-
m	Wider F1	Wider F2	Minimal	lengthened	-

Table 1. This is a survey of only a few of the spectral cues that are available for any one phoneme, but it also depends on the location of that phoneme in a word and that word in a sentence. These are certainly not the only spectral cues for these sounds. For example, the phoneme /n/ has a loss of energy at 1700 Hz called an anti-formant. More information on this can be found in the author’s article “How to set hearing aids differently for different languages” (October 2008 Hearing Review) and “How much gain is required for soft level inputs in some non-English languages?” (September 2011 Hearing Review).

Sounds rarely have only important auditory cues in one frequency region. And when this is

combined with context and the rules of everyday continuous speech discourse, the various phonemes of a language are not always where you think they are.

Reference

1. Sherbecoe RL, Studebaker GA. Audibility-index functions for the Connected Speech Test. *Ear Hear.* 2002; 23:385–398.

Article reprinted with kind permission from *The Hearing Review*.