A Case for Mid-frequency Amplification?

Published September 3rd, 2014

Most individuals with hearing loss are affected mostly in the high frequencies. And, because consonants–those phonemes (distinctive speech sounds) that are believed to enable humans to differentiate one word from another–are high-frequency sounds, hearing aid fitting formulas universally call for high-frequency amplification. This continues to be a take-off of the old "audiogram fitting" approach (although with less gain, but still tending to mimic the audiogram), but those writing fitting formulas today would have a field day arguing this comparison.

Part of what I write about has to do with history related to hearing, and one part that never ceases to amaze me is that so much of what was written years ago, is still applicable today. This is especially true for topics related to psychoacoustics and acoustics. Hearing had significant contributors whose work continues to stand the test of time.

Many years ago I wrote about the significance of mid-frequencies when fitting hearing aids[1]. I was recently reminded of that article, re-read it, and found that the topic continues to intrigue me. Therefore, this post will look back to review what the "ancients" had to say about mid-frequencies, and to suggest again that we look to see if their information applies to current hearing aid fittings. Much comes from the original article.

Hearing Aid Fitting Mantra

The mantra of the late 1970s was to recommend high-frequency hearing aids not only for high-frequency losses, but even for flat losses[2][3]. A resultant problem was that acoustic feedback emerged as a major fitting concern. The feedback-reduction circuitry of today was not available then, with feedback managed by either employing a notch filter at the suspected feedback frequency region, or to roll off the high frequencies (in opposition to the intended goal).

Amplification of high frequencies is understandable, but it is a fact that hearing aid user satisfaction continues to be an issue despite the many significant hardware and algorithm improvements in hearing aids throughout the years. Is it possible that useful mid-frequency information is being ignored, even today? Many would agree that even judicious use of high-frequency amplification alone is not the answer. Might there be benefit in reviewing mid-frequency contributions again?

What is it about mid-frequencies that warrant attention, if anything?

What are mid-frequencies? For purposes of this post, mid-frequencies will be those between 500 and 2000 Hz, and certainly not above 3000 Hz.

Speech Perception

The spoken message exists in different forms in its progress from the mind of the speaker to the mind of the listener. These take the form of linguistic, physiologic, anatomic, and acoustic aspects

of speech production and perception. This post will not consider all of these, but limit discussion to some of the factors that are involved at the listener's end that are frequency related and have an impact on speech perception.

Success in recognizing and interpreting sounds depends on suprathreshold processing, which involves: intensity of the desired sounds in the ear, intensity of unwanted sounds (noise) in the ear, echoes, phase distortion, character of the speech material, reverberation, naturalness, loudness perception, spectral composition, temporal (time) factors, bandwidth, energy bands, transmissions, context, speaker, rate of speaking, and so forth.

It should be obvious that many things are important in the recognition and interpretation of speech – not just the frequency response! Still, in spite of this, and for a variety of reasons, high frequencies have been elevated to *the* position of significance in hearing aid fitting. What follows is an attempt to show that even time-accepted data might not support the almost blind devotion to high frequencies. What is even more interesting is that attempts to equate pure-tone hearing losses and word recognition scores seem not to provide much useful data.

Acoustic Characteristics of Speech

Frequency and energy considerations often are given substantial emphasis to justify high-frequency emphasis. These will be reviewed in terms of: (1) the average speech spectrum; (2) power, frequency, and intelligibility; (3) critical bands; (4) crossover frequency; (5) energy bands; and (6) intelligibility versus bandwidth.

Average speech spectra

The average speech spectra (energy vs. frequency) measured for men and women over a long

period of time is shown in Figure 1, after Dunn and White, 1940[4]. Little difference is shown between the male and female voices. It is also obvious that most spoken energy is below 1000 Hz, and certainly below 2000 Hz, with the greatest peak energy in speech very close to 400 or 500 Hz. This suggests that frequency information in the mid-frequencies is of some importance. It is recognized that this is also the frequency region of considerable interfering noise, making amplification of this region an extremely difficult undertaking, especially when both noise and speech utilize the same frequency areas in their spectra.
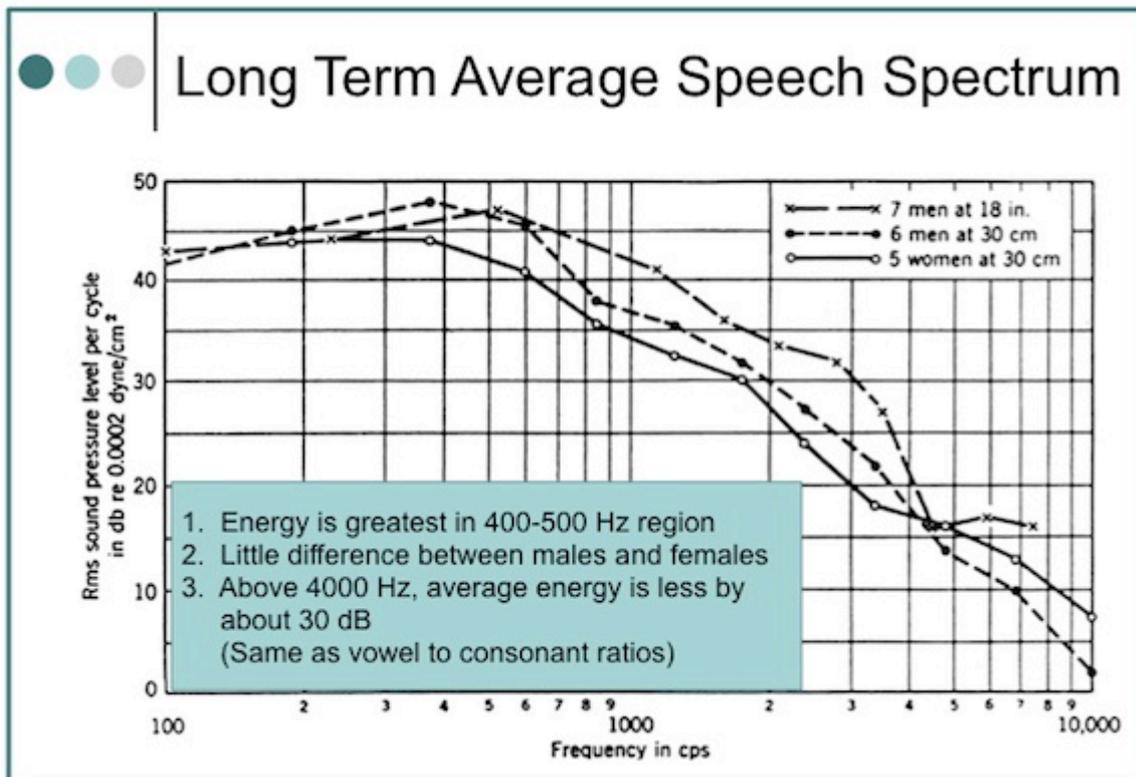
Figure 1. Average speech spectra. The overall levels, not indicated in the graph, were as follows: 18 inches in front of the seven male talkers, 75 dB SPL; 30 centimeters in front of the six male talkers, 76 dB SPL; and 30 centimeters in front of the five female talkers.

Power, frequency, and intelligibility

A. Power vs. intelligibility

The relationship of speech power to intelligibility is shown in Table 1 (adapted from Fletcher, 1929). These data have been referenced often to justify the significance of high frequencies for intelligibility. For example, from 1000 to 8000 Hz, the percent of speech power is only 5%, but contributes 60% to intelligibility. This is the usual presentation of these data. However, using the same data, 38% of the speech power and 70% of the speech intelligibility are present in the frequency range from 500 to 2000 Hz – definitely mid-frequency information by our definition.

Table 1. Frequency bands making equal (5%) Contributions to articulation index when all bands Are at their optimum levels. Composite data for Mens' and womens' voices.

Band	Frequency Limits (Hz)	Band	Frequency Limits (Hz)
1	250–375	11	1930–2140
2	373–505	12	2140–2355
3	505–645	13	2355–2600
4	645–795	14	2600–2900
5	795–955	15	2900–3255
6	955–1130	16	3255–3680
7	1130–1315	17	3680–4200
8	1315–1515	18	4200–4860
9	1515–1720	19	4860–5720
10	1720–1930	20	5720–7000

Going further, Table 1 shows the most important octave for the reception of speech to be from 500 to 1000 Hz. This contains 35% of the speech power and 35% of the speech intelligibility. Is speech power important? Hearing aid fitters would suggest that if sufficient speech power is not provided, the user does not as readily accept the hearing aids.

B. Vowel-to-consonant sound pressures

The loudest vowel [aw] as in "all," to the softest consonant (voiceless "th"), as in the word "thin," shows the vowel to have about 680 times greater sound pressure (ratio of 680:1), making it about 28 dB stronger (Figure 2). In the long-term average speech spectrum (Figure 1), 2000 Hz has less energy than 500 Hz by almost 35 dB. Initially, this would seem to imply that if normal conversation were to occur at 65 dB SPL, the high-frequency sounds above 2000 Hz would be softer by about 30 dB, or presented at 35 dB. The issue is complicated, however, when the intensity of a sound in free space follows an inverse-square law, which means that the intensity of a sound varies as the square of the distance. Following this law, when a listener moves to twice the distance from the sound source, the intensity will decease to one-quarter the original intensity. (If the distance is increased ten times, the intensity will be reduced to one one-hundredth the original intensity).
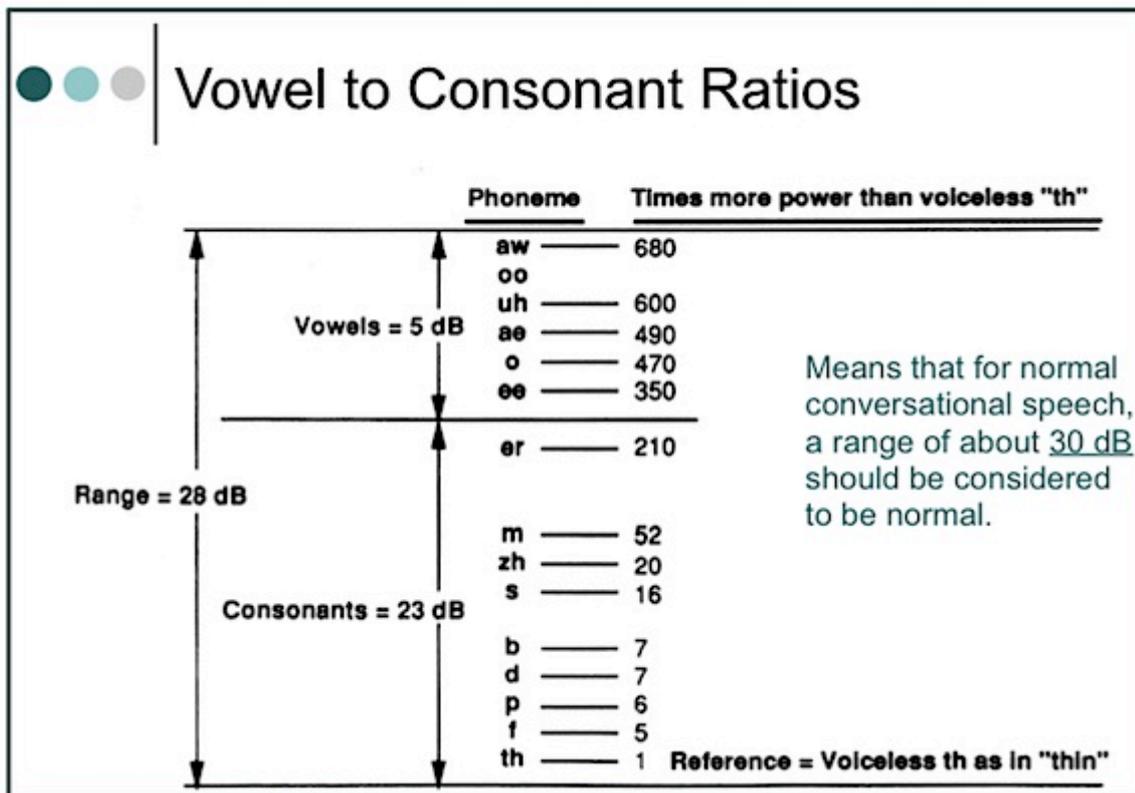
Figure 2. Ratio of vowel to consonant sound pressures. This means that for normal conversation, a range of about 30 dB should be normal. Most hearing aid fitting formulae use this range.

Fortunately, this is not what happens in a reverberant room (where most people listen), where different frequencies are emphasized (resonance) or de-emphasized, allowing high-frequency sounds to be heard more easily. Still, what is the real effect when a person has a high-frequency hearing loss? Can reverberation overcome the hearing loss as well as the overall decrease in high-frequency sound pressure, and if it can, what is the impact of increasing reverberation time on intelligibility? We know that it is not good.

Critical bands

Based on nonsense syllables

By definition, critical bands are frequency bands making equal (5%) contributions to intelligibility when all bands are at their optimum levels. Table 2 shows the critical bands developed by French and Steinberg for 20 bandwidths from 250 to 7000 Hz, and for a flat signal[5]. In these critical bands, the bands below 2000 Hz contribute a full 50% to intelligibility; and through 3000 Hz, a full 70%. It is suggested that if the bandwidths were not optimum, wider bands would be required for a given critical band. The question is, are all of the bandwidths at optimal levels in a hearing aid? In a narrow-band amplification system (what a hearing aid might be called), it might be suggested that the mid-frequencies are more likely to have optimum bandwidths, and thus contribute more to intelligibility. High frequencies are less likely to have optimum bandwidths, suggesting that a wider bandwidth is required to provide the 5% equal contribution that is shown in Table 2. However, wider bandwidths are not common in hearing aids, even today. Additionally, cochlear hearing impairment can have dramatic effects on the critical band by making it wider[6].

Critical Bands

- Frequency bands making equal (5%) contribution to intelligibility when all bands are at their optimum levels (nonsense syllables)

TABLE 2–3. *Frequency bands making equal (5%) contributions to articulation index when all bands are at their optimum levels: composite data for men's and women's voices*

Band	Frequency limits (Hz)	Band	Frequency limits (Hz)	
1	250–375	11	1,930–2,140	50%
2	375–505	12	2,140–2,355	
3	505–645	13	2,355–2,600	
4	645–795	14	2,600–2,900	
5	795–955	15	2,900–3,255	70%
6	955–1,130	16	3,255–3,680	
7	1,130–1,315	17	3,680–4,200	
8	1,315–1,515	18	4,200–4,860	
9	1,515–1,720	19	4,860–5,720	
10	1,720–1,930	20	5,720–7,000	

Adapted from French NR, Steinberg JC. Factors governing the intelligibility of speech sounds. *J Acoust Soc Am* 1947:19;90–119; with permission.

Critical Bands for Continuous Discourse

TABLE 2–4. *Cutoff frequencies of 20 equally contributing (5%) bands*

Band	Band limits (Hz)	Band	Band limits (Hz)	
1	150–300	11	1,189–1,407	
2	300–355	12	1,407–1,641	
3	355–408	13	1,641–1,915	70%
4	408–458	14	1,915–2,234	
5	458–515	15	2,234–2,568	
6	515–592	16	2,568–2,951	80%
7	592–697	17	2,951–3,390	
8	697–824	18	3,390–4,213	
9	824–990	19	4,213–5,361	
10	990–1189	20	5,361–8,000	

Adapted from Studebacker GA, Pavlovic CV, Sherbecoe RL. A frequency importance function for continuous discourse. *J Acoust Soc Am* 1987:81;1130–38; with permission.

Critical bands are different for continuous discourse

400 to 500 Hz = Frequency region of maximum importance

Based on continuous discourse

An assumption that critical bands are the same for all the speech materials has been shown not to be true[7]. Studebaker and his colleagues found that the frequency-importance function for continuous discourse (CD) was different than for nonsense syllables, and that lower frequencies were found to be relatively more important for the intelligibility of CD than for identification of nonsense syllable and other types of speech for which data are available, except for synthetic sentences.

These researchers also interpolated 20 bands of equal contribution (5%) for continuous discourse (Table 2). When compared with French and Steinberg's critical bands, they show greater intelligibility contribution at lower cutoff frequencies (70% below 2000 Hz and 80% below 3000 Hz). When plotted for one-third octave-band-importance function (importance/dB), they found the frequency region of maximum importance at 400 to 500 Hz, with a suggestion of bimodality with relatively less importance in the 800 to 1250 Hz region.

What is the implication for hearing aid amplification? It seems to suggest that greater concentration of the mid-frequencies, especially if amplified cleanly, may improve intelligibility as much, or perhaps even more, than unchallenged concentration on the high frequencies.

Crossover frequency

The crossover frequency is the frequency that divides the auditory frequency spectrum into two parts – each of which contributes equally to intelligibility. Figure 3 shows the crossover frequency for nonsense syllables to be 1930 Hz[8]. In summarizing the crossover frequency reported by several investigators, including themselves, Studebaker et al. reported lower crossover frequencies than did French and Steinberg. For continuous discourse, they reported a crossover frequency of 1189 Hz, and stated, "Although there are exceptions, crossover frequency generally decreases as the redundancy of the test material increases."
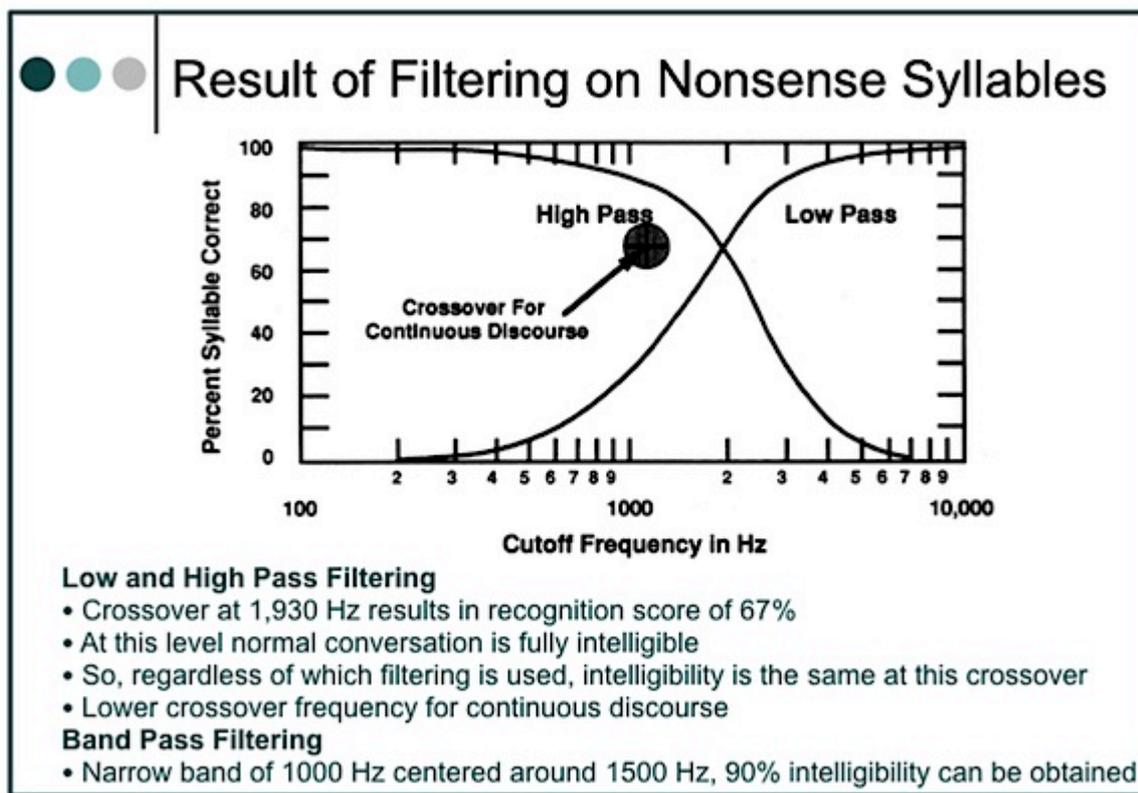

Figure 3. Crossover frequency and syllable discrimination (articulation) for high-pass and low-pass systems operating in quiet with optimal gain.

The data suggest that vocal communication is highly resistant to distortion, and that redundancy may shift the frequency of crossover even lower. Although conversation is intelligible with only the upper half of the speech spectrum, it is not unique because it can all be discarded and intelligibility can be maintained equally with only the lower half of the speech spectrum.

Important energy bands

Speech spectrograms have shown that F2 (the second formant or frequency energy region) of a phoneme is most important for intelligibility, both for normal and for sensorineural-impaired

listeners with sloping audiometric configurations[9]. Formants are characteristic of vowels. Figure 4 illustrates a sound spectrogram of a phrase and identifies the phonemes. Note the definite energy (formant) regions of the vowel sounds, with the formants labeled from the bottom of the graph upward.
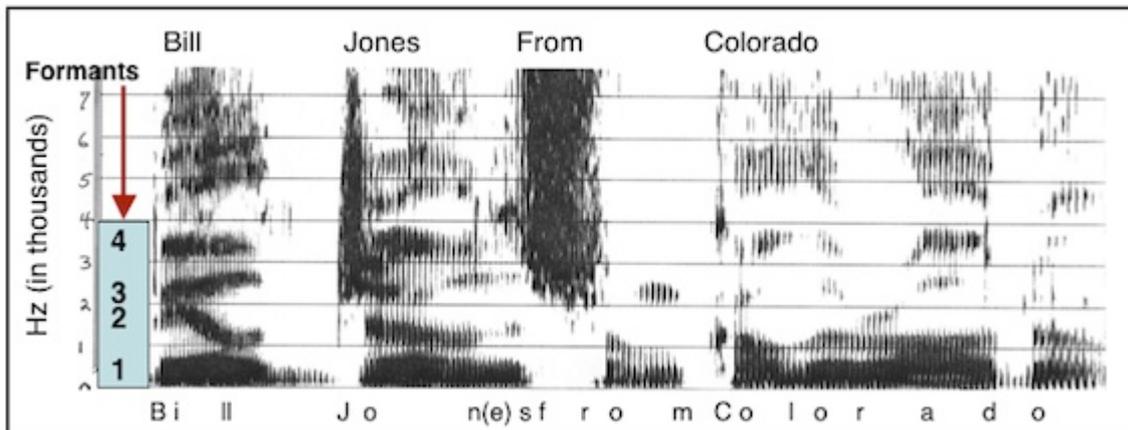

Figure 4. Speech spectrogram of the phrase.l.."Bill Jones from Colorado." The vertical frequency scale is in kilohertz (kHz), or thousands of Hertz, and time in milliseconds is along the horizontal scale. The dark bars in the vowels, below a frequency of about 3 kHz, are due to the peaks in the spectrum caused by the three lowest formants (resonances of the vocal tract). These formants change in frequency as the speech is articulated. Above these formants is another dark bar at about 3.5 k Hz; this is due to the fourth formant – its frequency does not change greatly with articulation, as do the frequencies of the three lower formants.

Table 3 shows the formant frequency ranges for various vowel sounds. Note that the range of F2 is from 880 to 2200 Hz. However, while F2 is considered very important, several authors recommend attenuating the first formant region to improve speech intelligibility of hearing impaired listeners[10][11].

Table 3. Average formant frequency values of 10 English pure vowels

Vowel Formants

	ee	I	e	ae	ah	aw	U	oo	A	er
First Formant Frequency										
Male:	270	390	530	660	730	570	440	300	640	490
Female:	310	430	610	860	850	590	470	370	760	500
Second Formant Frequency										
Male:	2290	1990	1840	1720	1090	840	1020	870	1190	1350
Female:	2790	2480	2330	2050	1220	920	1160	950	1400	1640
Third Formant Frequency										
Male:	3010	2550	2480	2410	2440	2410	2240	2240	2390	1690
Female:	3310	3070	2990	2850	2810	2710	2610	2670	2780	1960

Average formant frequency values of 10 English pure vowels as pronounced by a number of different speakers.

Footnotes

1. Staab, W.J. The significance of mid-frequencies in hearing aid selection. The Hearing Journal, June, 1988; 4:6-23.
2. Harford E: High frequency amplification for flat audiometric configurations. Paper presented at the Annual Convention of the American Speech-Language-Hearing Association, Chicago, IL, November 1977.
3. Rosenberg P, Rosenberg JR, Rosen RL, Meltzer CR: HFE regardless. Paper presented at the American-Speech-Language-Hearing Association, Chicago, IL, November 1977.
4. Licklider JCR, Miller GA: The perception of speech. Handbook of Experimental Psychology. New York, John Wiley & Sons, Inc. 1951, pp. 1040-1074.
5. French NR, Steinberg JC: Factors governing the intelligibility of speech sounds. J Acoust Soc Am 1947;19(1):90-119.
6. Hall JV, Tyler RS, Fernandes MA: The factors influencing the masking level difference in cochlear hearing-impaired and normal-hearing listeners. J Speech Hear Res 1984. 27:145-154.
7. Studebaker GA, Pavlovic CV, Sherbecoe RL: A frequency importance function for continuous discourse. J Acoust Soc Am 1987. 81 (4):1130-1138.
8. French NR, Steinberg JC. Factors governing the intelligibility of speech sounds. J Acoust Soc Am, 1947. 19(11:90-119.
9. Danaher E, Osberger J, Pickett.F. Discrimination of formant frequency transitions in synthetic vowels. J Speech Hear Res 1973. 16:439-451.
10. Martin E, Picket J. Sensorineural hearing loss and upward spread of masking. J Speech Hear Res 1970; 13:426 – 437.
11. Danaher E, Pickett J. Some masking effects produced by low frequency vowel formants in persons with sensorineural hearing loss. J Speech Hear Res 1975. 18:261-273.