A NEW ALGORITHM FOR FREQUENCY TRANSPOSING IN DIGITAL HEARING AIDS

¹Ajish, K. Abraham, & ²Surya, D.

Abstract

Frequency transposition is a signal processing strategy of transposing high frequency acoustic content in a speech signal to a low frequency range. This technique has been widely discussed for a long period of time as the conventional amplification strategies are not beneficial to those suffering from the specific condition of high frequency hearing loss. But still, this strategy as a corrective option, never gained a good momentum or interest amongst audiologists as most of the technological solutions developed have got one or the other drawback. The present study was aimed at developing a new algorithm exploring frequency transposition and testing its effectiveness both objectively and subjectively. The developed algorithm involves transposing the significant high frequency acoustic content based on an adaptable threshold, to a specific region in low frequency band, frame by frame. The algorithm also incorporates a compression strategy to adjust the transposed speech content to the hearing loss pattern of the client. The new algorithm eliminates most of the limitations of the existing technologies. Objective evaluation is done with the help of three signal processing software and subjective evaluation with five subjects of normal hearing and ten subjects of moderate to severe sloping hearing loss who are native Kannada speakers within the age range of 20 to 80. 100% detection score is obtained for all the five normal hearing subjects and in high frequency hearing impaired subjects a significant 22% increase is achieved with transposition. The score obtained sufficiently prove that the transposed words have satisfactory sound quality and transposition has increased the speech cues for detection of words. The results hold good for further research in this area and offers a potential rehabilitation technology for persons with high frequency hearing impairment.

Key words: Algorithm, High frequency hearing loss, Frequency lowering, Frequency compression, Frequency transposition.

Introduction

Audible spectrum extends from 20 Hz to 20 kHz with frequencies up to 8 kHz adequate for normal speech perception. Any degree of impairment of the ability to perceive sound is referred as hearing loss and is defined as the difference between the hearing threshold of the impaired person and that of persons with normal hearing. High frequency hearing loss is a condition in which a person has normal or near normal hearing in the low frequencies up to 1000Hz, but it falls like a ski slope in the higher frequencies. High frequency hearing loss is not an acceptable loss, as speech is reasonably intelligent only with the presence of high frequency content (Berlin, 1982).

Among all types of hearing disorders, occurrence of high frequency hearing loss in persons with hearing impairment is more than 60%. Reasons for this can be many; it can be of genetic factors, due to aging, through infectious agents or due to the retardation of cochlear hair cells by continuous exposure to high energy sounds. High frequency hearing loss is a sensory neural hearing loss where high frequency hair cells are not competent of responding to vibrations or the hair cells may be completely absent (dead region).

Persons with high frequency hearing impairment find difficulty in adjusting to the natural environment as our world is full of high frequency sounds whether it be telephone ringing, approaching train or bus, birds etc. High frequency sounds are vital cues for recognising words, awareness of environment, speech comprehension, enjoying music and gives sharp features of naturalness and clarity to the sound. Energy for many voiceless consonants (For eg: /s/, /th/, /f/, /sh/, /t/) which are critical for speech understanding lies in this frequency region. Hence, it is necessary that the high frequency sounds are made audible.

Traditional correction technique for hearing loss is amplification. But this does not work well for persons with high frequency hearing loss. The main reason is that, it is not possible to provide high gain in high frequency region as it would result in low maximum power output, limited

¹Ajish K. Abraham, Professor of Electronics, All India Institute of Speech and Hearing (AIISH), Mysore, E-mail: ajish68@aiishmysore.in, & ²Surya, D., M. Tech. Scholar, Department of Biomedical Engineering, Manipal Institute of Technology, Manipal, E-mail: surya23das@gmail.com

bandwidth and an increased chance of acoustic feed back before desired gain (Kuk, Keenan, Korhonen & Lau, 2009) is reached. The situation is further worsened by the poor frequency response of the hearing aid receiver above 4000 Hz, as the receiver is the transducer at the final end, which converts electrical signal to acoustic form and delivers it to the ear canal. Moreover, persons suffering from hearing impairments due to dead cochlear region do not get any benefit from high gain. Sometime high gain provides more negative effect as because of loudness recruitment, they can't tolerate increment in loudness and in some other cases, high gain result in pain. McDermott & Dean (2000) quotes that when sounds are presented at frequencies where sensitivity loss is maximum, it leads to anomalous qualities such as abnormal pitch or timbre. Thus persons with sloping high frequency hearing loss never use their hearing aids or use them irregularly. Cochlear Implant is another choice for rehabilitation for individuals with severe to profound loss at frequencies above 1 kHz. (Simpson, 2009). Even though many of such individuals fitted with Cochlear Implant showed improvements in identification scores, the enormous cost involved prevents common man from availing this option. Hence, alternative approaches which shift components from inaudible high frequency areas of speech to audible low frequency areas are needed to tackle this issue.

Frequency transposition, the signal processing strategy with which high frequency components of speech are shifted to lower frequencies, is one such approach. Several attempts were made in this direction over the years. (Ling, 1968; Beasley et al., 1976; Reed et al., 1985; Posen et al., 1993). As summarized by Katz (2009), most of these attempts were not very successful because of limitations such as changes in pitch, variations in the temporal nature of acoustic signal and loss of naturalness in sound quality. Another approach to tackle the high frequency hearing loss was the use of proportional frequency compression. This addressed some of the issues with frequency transposition as this technique preserved the normal frequency relations between the fundamental frequency and the formants. In this approach, the entire range of speech frequencies is compressed to a lower range and the spectrum is thinned down. The first hearing aid with this type of signal processing was introduced in 1991 (Katz, 2009) and subsequently followed by several manufacturers. Third approach in this direction is through bringing

down all frequency components in the speech signal by a fixed ratio which is called frequency shifting. A comparison of all these three approaches has been carried out by Simpson (2009).

Studies conducted on the performance of these devices in comparison with traditional hearing aids have shown mixed results. Study conducted by Parent et al (1998) on four adults reported improvement in speech perception for two of the four subjects. Another study by McDermott et al (1999) on five adult subjects reported small improvements in speech perception for two subjects. No significant difference was found in the study conducted by McDermott and Knight (2001) when they compared the speech perception scores between traditional hearing aids and hearing aids with frequency compression.

One reason pointed out by McDermott et al (1999) for the disappointing results is the fact that the frequency transposition did not preserve important features depicting the spectral pattern of the original speech. Another reason for insignificant differences between the conventional aid and the one with transposition may be due to the fact that the high frequency signals when transposed will result in the masking of low frequency components of speech (McDermott et al., 1999). Another drawback of the existing techniques is that, after transposition sound quality is reduced and low frequency speech It was reported by content gets distorted. Aguileru, Nelson, Rutledge & Gaga (1999) that, transpositions that involve shifting or modifying fundamental frequency and formant frequency leads to poor speech discrimination scores. They also reported that, the poor discrimination scores may be due to alteration of ratio of fundamental and formant frequency, as this ratio is an invariant cue for vowel and speech perception. Thus, there is a need to develop a new algorithm for frequency transposition which preserve the significant details of the spectral shape, retains the ratio of fundamental and formant frequencies as well as prevents masking of low frequency components of speech.

The objectives of the study were 1) to develop a new algorithm for shifting information from inaudible high frequency areas of speech to audible low frequency areas overcoming the limitations of existing ones and 2) to evaluate the effectiveness of the developed algorithm through subjective and objective evaluation.

Method

Algorithm for frequency transposition: The incoming speech signal is segmented to 32 msec frames and convolved with hamming window. Spectral domain processing is done with the help of Discrete Fourier Transform as it gives a good approximation of speech signal in spectral domain with a reasonably good frequency resolution (Burrus & Parks, 1985). Discrete Cosine Transform (DCT) and Reeds Equation are used for compression and Inverse Discrete Fourier Transform (IDFT) is used for transforming the signal back. Sampling frequency of 16 kHz and overlap factor of 25% are used in the algorithm. The algorithm first executes a check to find out whether there is any speech content in the 2 kHz to 8 kHz range of the speech signal. This is done by comparing the amplitudes in this frequency range with a set threshold. If more than 30% (arrived empirically) of the samples are above the threshold, then only transposition is done. The threshold value is arrived in real time for every frame on the basis of the sample values of the previous frame. Once significant amount of speech content is identified in the high frequency region, then the transposition begins by accentuating the high frequency content to a compact array. This array is then added to low frequency region specifically from 2 kHz onwards. The method is illustrated in the figure 1.



Figure 1: Method used for frequency transposition.

The transposed array can be further accentuated by a scaling element (based on subjective preference) before getting added to low frequency region from 2 kHz onwards. In this algorithm, the need of frequency compression is not for frequency lowering but more for fitting the transposed acoustic content to individual requirement. Fitting is important as the start frequency of the person with hearing impairment need not always be 2 kHz or above. Two compression methods are incorporated in the algorithm design, linear and non linear. Option of selecting between the two is again a subjective element. Linear compression is done with the help of DCT. The dimension of DCT coefficient is modified, if required, by the simple technique of zero padding. Zero padding can result in compression of speech signal at required extend but with the disadvantage of modifying pitch periods. Nonlinear compression is done with help of Reeds equation which is widely used in various speech compression algorithms. In the algorithm, parameters chosen for optimization of the equation are sampling frequency of 16000 Hz, K as 2 and therefore value of wrapping parameter is 1/3 (0.333).

Recording of speech material: Ten words are selected from the phonetically balanced (PB) high frequency word list. It is a list of monosyllabic words selectively chosen so that they approximate the relative frequency of phoneme occurrence in each language (Browman & Goldstein, 1995). The selected high frequency test words (Appendix 1) were uttered by an adult Kannada speaking female speaker (age 35) into a condenser microphone placed at 15 cm from the speaker. The words were recorded through a precision sound level meter (B & K 2250) with sound recording facility (B & K 7226) in a sound treated room. Recorded samples were digitized at a sampling frequency of 22 kHz and 16 bits per sample.

Participants: Ten persons (S1-S10) with moderate to severe sloping hearing loss and within an age range of 20 to 80 were recruited for perception test. Their audiograms are shown in figure 2. Five normally hearing participants who were native Kannada speakers with no experience of using any hearing aids also participated in the study. The hearing impaired participants were not having any experience of using frequency transposition hearing aids.

Outside the candidacy criteria, three non-native/ non-Kannada speakers belonging to the same age group as in the candidacy criteria with moderate to severe high frequency hearing loss were considered for study as a separate group. Along with them, three children in the age range of 8 to 14 who are native Kannada speakers with moderate to severe sloping hearing loss were also included as another separate group.



Figure 2: Threshold levels (Left ear: x Right ear: o) of each participant

Speech material for the study: The ten recorded PB words were processed by the algorithm to prepare the transposed version. In a random order the sets of transposed and un-transposed words were mixed and a new list of twenty words was generated and stored as wave files in the computer.

Procedure: The test words were delivered from the computer through a precision power amplifier (B & K 2716C) through a headphone with a constant gain for frequencies up to 10 kHz. The participants were made to listen to the sounds in a sound proof room from the headphone. The stimuli were presented at 40 dB SL for each participant. Level adjustment was done with the gain control of B & K 2716C power amplifier (flat frequency response from 20 Hz – 20 kHz). The participants were made to hear the test words one by one and were instructed to repeat back as well as write down what they have heard as soon as they perceive it. The responses were recorded instantly using a binary scoring pattern. Correct identification of the score was awarded with 1 and otherwise 0. Finally a total score of transposed and un-transposed speech is made and the effectiveness of algorithm was measured by calculating the percentage increase in the score.

Analyses: The following analyses were carried out: a) Objective analyses of the transposed and un-transposed words using PRAAT software, b) measurement of speech identification score for transposed and un-transposed words. SPSS software was used to carry out the statistical analysis. Paired 'T' test was carried out to find out the significance of improvement in intelligibility of transposed words. As 'N' was only ten, nonparametric test was also administered

Results and Discussion

Performance of the developed algorithm: The algorithm uses a new approach of transposition. In the developed algorithm, the edge frequency or cut off frequency is taken as 2 kHz with region from 125 Hz - 2 kHz kept unmodified during transposition. Selection of scale factor is an additional feature included in the algorithm to enhance the performance. This factor is kept as a subjective element based on the patient's hearing loss pattern. Increasing the scale factor linearly scales the generated array for transposition. It build the effect of adding noise, since high frequency sounds are mostly fricatives and has a noise like characteristics, doing so can potentially add more speech cues.

The compact array generated after extracting the high frequency content is added from 2 kHz. Through this, the main advantage is over all distortion that can occur due to transposition of the speech signal will be least, with 125 Hz - 2 kHz unaffected. Since transposition is done always in the frequency range of 2 kHz - 8 kHz, in long term, the listener will be more comfortable in identifying the same speech sound.

Linear frequency compression, one of the existing techniques, which works by lowering all frequencies contained in the speech signal by a constant factor, keeps the ratio between fundamental frequency and formant frequencies unchanged (Turner & Cummings, 1999). For vowel intelligibility these ratios are significant, but this technique has the serious drawback that the speech will become unnatural due to lowering of pitch. The technique used in our algorithm is different, as the frequency compression is done only for the transposed speech segment whereas the frequency region upto 2 KHz is untouched. Accordingly the developed algorithm overcomes the drawback of unnatural sound quality associated with linear frequency compression.

Nonlinear frequency compression also involves lowering of all frequency components but by different factors at different frequencies, the factor going high at higher frequencies. Thus the frequency ratios are not preserved, which may reduce the speech perception. In the developed algorithm, the frequency range from 0 to 2 KHZ is unaffected. For the reason that all the main vowel formants and fundamental frequency are concentrated in the 125 Hz - 2 kHz range, (Raphael, Borden & Harris, 2011) the ratio between fundamental frequency and formant frequency is unchanged. Accordingly the algorithm overcomes the drawback of both linear and nonlinear frequency compression, preserving the advantages of both.



Figure 3: Time domain waveform and spectrogram plot of Kannada word 'sashe'.

Objective evaluation: Objective evaluation of the performance of the algorithm is done with the help of three signal processing software namely Matlab, Praat and CSL. The time and frequency domain plot along with spectrogram (figure 3) and LTA plot (figure 4) of transposed sounds with un-transposed sounds clearly indicate that there is no significant speech component in high frequency region of the transposed sound. Analysis of the LTA plot can also prove that frequency transposition has taken place in transposed sound provided scaling value is sufficiently increased.

Subjective evaluation: Subjective evaluation is done in two phases: 1) pilot test with five listeners with normal hearing and 2) intelligibility tests in listeners with moderate to severe sloping hearing loss. Perceptual evaluation of five normal hearing native Kannada speakers showed a 100% intelligibility score for transposed as well as untransposed words. It could be inferred clearly that after significant spectral modification there is still a good amount of speech cues in transposed version.

The results obtained for intelligibility tests for listeners with hearing impairment are tabulated in table 1. A mean increase of 22% (Table 2) is recorded for transposed sound over un-transposed sound. In other words, it could be concluded that from the list of ten words, two more transposed words were intelligible compared to untransposed words. Paired 'T' test showed a statistically significant mean improvement of 22% (t (9) = 8.820, p<0.001). As 'N' is only 10, Wilcoxon's signed rank test was employed which also showed a significant improvement in intelligibility (|Z| = 2.842, p<0.05).



Figure 4: Long term average spectrum of Kannada word 'sashe'

The data is obtained for an untrained adult population without the use of hearing aids in a soundproof room. So it could be expected that the result will be more positive for trained candidates for long term use of frequency transposition hearing aids in a natural environment (Simpson, 2009). As reported by Simpson (2009), among the existing technologies, the maximum statistically significant improvement of 20% in speech identification was observed with the transposition technique devised by Robinson, Baer and Moore (2007). For the nonlinear frequency compression technique developed by Simpson, Hersbach and Mcdermott (2006) the reported improvement is 6%. The developed algorithm showed a statistically significant mean improvement of 22% in the intelligibility score, thus proving that the performance of the newly developed algorithm is better than the existing technologies.

Table 1: Intelligibility score obtained for hearingimpaired participants

	Intelligibility score of un- transposed	Intelligibility score of transposed	Percentage increase in score
Dontinin ont 1	words	words	20
Participant 1	/	9	20
Participant 2	7	10	30
Participant 3	6	9	30
Participant 4	8	9	10
Participant 5	7	9	20
Participant 6	6	8	20
Participant 7	7	8	10
Participant 8	6	9	30
Participant 9	6	9	30
Participant 10	7	9	20

Overcoming the limitations of existing techniques: As pointed out by Simpson (2009), masking of useful low frequency content is one of the disadvantages of existing methods of transposition. This algorithm overcomes this limitation as the frequency region from 125 Hz to

2 kHz, where most of the useful low frequency information is present, is kept unmodified.

Table 2: Mean & SD of Intelligibility scoreobtained for hearing impaired participants

	Intelligibility score of untransposed words (N = 10)	Intelligibility score of transposed words (N=10)	Percentage increase in score (N = 10)
Mean	6.7	8.9	22
Standard deviation	0.675	0.568	7.89

Another drawback of the existing method was that it may transpose undesirable high frequency noise (Simpson, 2009). This is taken care to a certain extent in the new algorithm by fixing an amplitude threshold to decide whether transposition is required or not. A serious limitation of the frequency compression technique was that it does not preserve the harmonic relationship between formants (Simpson, 2009). This drawback is also eliminated in the new algorithm as all the main vowel formants lie in the unmodified region from 125 Hz to 2 kHz. Limitation of the existing frequency shifting algorithm was the lowering of the pitch which leads to unnatural sound quality (Simpson, 2009). As the new algorithm does not shift the fundamental and formant frequencies, the pitch remains unchanged.

Limitations of the new algorithm: Shortcomings were observed while testing children as well as non Kannada speaking adult population. The scores obtained for these two populations shown in figure 5 were not satisfactory in comparison with the participants chosen through the described candidacy criteria. Even though the hearing loss pattern is moderate to severe sloping hearing loss, the candidate could not find or perceive any additional speech cue in transposed sound than

that in the un-transposed sound. And hence the number of correctly identified words is less with almost similar score in transposed and untransposed sound. This unsatisfactory performance is due to the fact that, in both these populations the words selected for testing were unfamiliar. In children, especially with high frequency hearing loss, language development is a slow process and they may not have acquired the high frequency phonemes and the words containing them. In adult population, even though they have acquired high frequency phonemes, they were not familiar with Kannada words containing these phonemes. The algorithm is designed not for perceiving high frequency sounds as high frequency sounds, but for detecting the presence of occurrence of high frequency phoneme in a word through a substitute in low frequency speech sounds. In other words, through transposition algorithm, the user listens to a low frequency speech sound which is super imposed by features of high frequency sound.



Figure 5: Recognition scores for non-Kannada speaking adults (N1-N3) and Kannada speaking children (C1-C3)

Positive result obtained for adult native Kannada speakers is due to the auditory processing mechanism in the brain, as the test words are familiar. Auditory processing in brain includes a large amount of pattern matching processes to make sense out of them through the speech sounds that were heard before and hardwired in memory. So when transposed speech sounds are given as stimuli, the memory of the listener could have helped in judging the stimuli and came up with a better understanding of the situation. This cannot be expected from a population of non Kannada speakers and children.

Effect of transposition on children could be studied only after long term training with usage of frequency transposition hearing aids. Minimum ten to twenty hours of listening exposure to transposed speech is required (Simpson, 2009). Training should be given for children in using the transposed high frequency speech cues in low frequency sound and speech production of proper high frequency sound when they are perceived. Children have an advantage of powerful and developing brain which can remodel to include the new challenges. So it could be expected that performance of children and adult could be improved with sufficient training.

Conclusion

Modified frequency transposition algorithm has been developed which overcomes most of the drawbacks of the existing techniques. The effectiveness of the algorithm has been tested through objective and subjective methods, which showed a positive result. Because of less number of participants included in the study, generalization of results is not possible. However, the transposed speech has shown satisfactory sound quality with minimal level of distortion and the participants with high frequency hearing loss were able to perceive transposed sounds, 22% more than un-transposed original sound. To prove it conclusively, a testing methodology with sufficient training should be adopted in a larger population including children and listeners who speak other languages. Natural sound quality, preserving harmonic relationship between frequency components, good Signal to Noise Ratio (SNR) and less overlap of frequency components are features looked in a good transposition algorithm, all of which have been achieved in the new algorithm. The results hold good for further research in this area and offers a potential rehabilitation technology for persons with high frequency hearing impairment.

References

- Aguileru Mutioz, C. M., Peggy Nelson, B., Janet C., Rutledge & Gago, A. (1999). Frequency lowering processing for listeners with significant hearing loss. *Proceedings of ICECS 1999, the 6th IEEE International Conference on Electronics, Circuits and Systems, Volume: 2*
- Beasley, D. S., Mosher, N. L., Orchik D. J. (1976). Use of frequency shifted / time-compressed speech with hearing-impaired children. Audiology, 15 (5):395-406.
- Berlin, C. I., (1982). Ultra-audiometric hearing in the hearing impaired and the use of upward-shifting translating hearing aids. *The Volta Review*, 84(7), 352-363.
- Browman, C.P., & Goldstein, L.M. (1995). Dynamics and articulatory phonology. In T. van Gelder & R. F. Port (Eds.), Mind as Motion (pp.175-193). Cambridge MA: MIT Press
- Burrus, C. S. S., & Parks, T. W. (1985). DFT/FFT and Convolution Algorithms and Implementation: Theory and Implementation. John Wiley & Sons, Inc. New York, NY, USA
- Katz, J. (2009). Handbook of Clinical Audiology Sixth Edition. USA: Lippincott Williams & Wilkins

- Kuk, F., Keenan, D., Korhonen, P., & Lau, C. (2009). Efficacy of linear frequency transposition on consonant identification in quiet and in noise. *Journal of American Academy of Audiology*, 20(8), 465-479.
- Ling, D. (1968). There experiments on frequency transposition. *American Annals of Deaf, 113*, 283-294.
- McDermott, H. J. & Dean, M. R. (2000). Speech perception with steeply sloping hearing loss; effects of frequency transposition. *British Journal* of Audiology, 34, 353-361.
- McDermott, H. J., Dorkos, V. P., Dean, M. R., Ching, T. Y. (1999). Improvements in speech perception with sue of the AVR TranSonic frequency transposing hearing aid. *Journal of Speech Language and Hearing Research*, 42, 1323-1335.
- McDermott, H. J. & Knight, M. R. (2001). Preliminary results with the AVR impact frequency transposing hearing aid. *Journal of American Academy of Audiology*, *12*, 121-127.
- Parent, T. C., Chmiel, R., Jerger, J. (1998). Comparison of performance with frequency transposition hearing aids and conventional hearing aids. *Journal of American Academy of Audiology*, 9, 67-77.
- Posen, M. P., Reed, C. M., Braida, L. D. (1993). Intelligibility of frequency lowered speech produced by a channel vocoder. *Journal of Rehabilitative Research and Development*, 30, 26-38.
- Raphael, L. J., Borden, G. J., Harris K. S. (2011). Speech Science Primer - Sixth Edition. USA: Lippincott Williams & Wilkins

- Reed, C. M., Schultz, K. I., Braida, L. D., Durlach, N. I. (1985). Discrimination and identification of frequency lowered speech in listeners with high frequency hearing impairment. *Journal of Acoustical Society of American*, 78, 2139-2141.
- Robinson, J. D, Baer, T, & Moore, B. C. (2007). Using transposition to improve consonant discrimination and detection for listeners with severe high frequency hearing loss. *International Journal of Audiology*, 46, 293-308.
- Simpson, A., Hersbatch, A. A, & McDermott, H. J. (2006). Frequency compression outcomes for listeners with steeply sloping audiograms. *International Journal of Audiology*, 45, 619-629.
- Simpson, A. (2009). Frequency-Lowering Devices for Managing High Frequency Hearing Loss: A Review. *Trends in Amplification*, 13: 87.
- Turner C. W. & Cummings K. J. (1999). Speech audibility for listeners with high frequency hearing loss. *American Journal of Audiology*, 8, 47-56.

Acknowledgement

The authors thank Director, AIISH, Mysore for permitting to carry out the study. The authors also thank Dr. M.S. Vasanthalakshmi, Lecturer in Biostatistics, Department of Speech Language Pathology, for the help in statistical analysis. Sincere thanks to Dr. N. Sreedevi, Reader, Department of Speech Language Sciences and Mrs. V.T. Kalaiselvi, Junior Technical Officer, Department of Electronics for the immense support.

Anne	xure 1
List of Phonetically Balanced Words (Ka	nnada) used for Speech Intelligibility test

Sl. No.	Word	
1.	isța	
2.	kuĺĺi	
3.	t∫amat∫a	
4.	vit∫ara	
5.	irali	
6.	t∫akra	
7.	sikku	
8.	hasivu	
9.	keiĺu	
10.	serisi	