J Am Acad Audiol 19:639–650 (2008)

Use of Linear Frequency Transposition in Simulated Hearing Loss

DOI: 10.3766/jaaa.19.8.7

Petri Korhonen*
Francis Kuk*

Abstract

Purpose: To demonstrate the usability of new acoustic cues provided by a commercial hearing aid using linear frequency transposition. The importance of auditory training in realizing the benefit was evaluated.

Research Design: Quasi-experimental study design. All subjects were tested both with conventional amplification and with linear frequency transposition before and after directed training.

Study Sample: A total of nine normal hearing participants with simulated high frequency hearing loss at and above 1600 Hz participated in the study. All subjects were native English speakers ranging in age from 18 to 24 years.

Intervention: Identification of voiceless phonemes in CV, VC, and CVC context processed with and without linear frequency transposition was evaluated. Identification test was carried out four times. Between test trials, participants completed 15 minutes of self-paced training using transposed stimuli.

Results: Prior to any training, transposition did not improve phoneme identification scores. Training of 30 minutes improved the overall identification scores of the transposed stimuli over the nontransposed stimuli by 14.4%.

Conclusions: The results demonstrated that frequency transposition produces acoustic cues that normal hearing listeners with a simulated hearing loss at and above 1600 Hz may be trained to utilize.

Key Words: Frequency transposition, hearing aids, speech recognition

Abbreviations: AE = Audibility Extender; FFT = fast Fourier transformation

Sumario

Propósito: Demostrar la utilidad de las nuevas claves acústicas aportadas por un auxiliar auditivo comercial usando transposición lineal de la frecuencia. Se evalúa la importancia del adiestramiento auditivo para obtener el beneficio.

Diseño de la Investigación: Diseño de estudio cuasi-experimental. Todos los sujetos fueron evaluados tanto con una amplificación convencional como con una transposición lineal de frecuencia antes y después de un entrenamiento dirigido.

Muestra del Estudio: Participaron del estudio un total de nuevo sujetos normoyentes con una hipoacusia simulada en las altas frecuencias por encima de 1600 Hz. Todos los sujetos eran hablantes nativos del inglés en edades de 18 a 24 años.

Intervención: Se evaluó la identificación de fonemas no sonoros en contexto CV, VC y CVC, procesados con o sin transposición lineal de frecuencia. La prueba de identificación fue llevada a cabo cuatro veces. Entre pruebas, los participantes completaron un entrenamiento de 15 minutos a su propio ritmo para el uso de los estímulos de transposición.

*Office of Research in Clinical Amplification, Widex USA

Petri Korhonen, Widex ORCA, 2300 Cabot Dr., Suite 415, 60532, Lisle, IL; Phone: 630-245-0025; Fax: 630-245-0141; E-mail: pkorhonen@widexmail.com

The preliminary results of this study were originally presented at the American Auditory Society Annual Meeting, March 4–6, 2007, Scottsdale, AZ.

In the interest of full disclosure, it should be noted that the algorithm studied in this manuscript is manufactured by the Widex Hearing Aid Company and that the authors work at, but have no other financial interest in, the Widex Office of Research in Clinical Amplification.

Resultados: Antes de cualquier entrenamiento, la transposición no mejoró los puntajes de identificación de fonemas. Un entrenamiento de 30 minutos mejoró los puntajes globales de identificación de los estímulos transpuestos sobre los estímulos no transpuestos en 14.4%.

Conclusiones: Los resultados demuestran que la transposición de frecuencia produce claves acústicas para las que sujetos con audición normal con una hipoacusia simulada por encima de 1600 Hz puede ser entrenados a utilizar.

Palabras Clave: Transposición de frecuencia, auxiliares auditivos, reconocimiento del lenguaje

Abreviaturas: AE = Extensión de la audibilidad; FFT = transformación rápida de Fourier

For people with a precipitous high frequency hearing loss, the threshold of hearing may be so elevated such that conventional amplification provides minimal or no benefit. This might be the result of insufficient hearing aid gain before feedback occurs. In other cases, the high frequency region may be "dead" because of the complete depletion of inner hair cells. Acoustic stimulation of the dead regions may further decrease the already depressed speech understanding scores (Moore, 2004). Various signal processing techniques based on frequency lowering have been proposed in the past to restore the lost acoustic cues within the high frequencies (Braida et al, 1979). The rationale of frequency lowering is to render the information in the unaidable high frequency region audible in a lower frequency form so the aidable lower frequency fibers may decode the high frequency information effectively. The current study explored the usability of frequency transposed voiceless consonants using normal hearing individuals with a simulated hearing loss.

There are two main approaches in frequency lowering based on how the original frequencies are relocated in the new processed frequency domain: frequency shifting and frequency compression (Braida et al, 1979). *Frequency shifting* refers to a scheme in which all spectral components of the original signal are lowered linearly by a fixed displacement (say, 1000 Hz). The lowered frequencies that would be located at the "negative frequencies" after lowering are either low-pass filtered before shifting or are aliased to positive frequencies as a reversed spectrum. In the process, the pitch of the speech signal is lowered because all frequencies are lowered. The first studies on the effects of frequency shifting on speech intelligibility were reported by Fletcher (1953) on people with normal hearing, and Raymond and Proud (1962) on people with a high frequency hearing loss.

A variant of frequency shifting is *frequency transposition* (e.g., Johansson, 1961; Velmans, 1971). This is the scheme where only a selected portion of the original frequency spectrum is lowered. The lowered portion of the signal is combined with the unprocessed portion of the signal. The rationale is to retain as much of the original signal as possible. The pitch of the original signal may be better preserved depending on how much of the speech spectrum is transposed. This will be the case if only just the very high frequency spectrum is transposed, sparing the lower frequency region.

Frequency compression refers to a scheme in which the amount of lowering is proportional for each frequency so that the ratio between the original and compressed frequencies is defined (Fairbanks et al, 1954; David and McDonald, 1956). Typically, each frequency is lowered by a constant factor, but nonuniform frequency compression is also possible (Hicks et al, 1981; Aguilera-Muños et al, 1999). Similar to frequency shifting, frequency compression also alters the pitch of speech because the fundamental frequency (as well as all frequencies) is affected (Braida et al, 1979).

Using frequency lowering to achieve high frequency audibility is not a new idea (see Braida et al, 1979, for a review). Among the early pioneers, Johansson (1961, 1966) and Velmans (1971, 1974) reported varying degrees of success of a frequency transposition scheme. For example, Velmans (1973, 1975) reported an improvement in imitation with transposition. Subsequent studies reported improvements in auditory discrimination of high frequency consonants (Rees and Velmans, 1993). Unfortunately, for various reasons, none of the studies was conclusive on the conditions under which frequency transposition may be indicative.

A limiting factor in earlier studies of frequency lowering may be the ease at which lowering is achieved without confounding artifacts. This is because early attempts were designed to achieve frequency lowering based on existing technologies rather than to achieve the optimal signal processing results (Braida et al, 1979). While the high frequency information may have become audible, the methods also altered other aspects of sounds known to be important for speech communication, such as pitch, temporal structures, and so forth (Ladefoged, 1993). The introduction of digital signal processing (DSP) technology allows researchers to reexamine the possibility by designing frequency lowering schemes that meet the stated objectives without as much of the burden of technological limitations. It is possible that more recent improvement in DSP technology could yield a different outcome.

On the other hand, recent studies with frequency lowering have not yielded a better outcome. Simpson et al (2005) evaluated an experimental frequency compression hearing device that compressed the frequencies above a programmable cutoff frequency. Thus, high frequency sounds above the cutoff were lowered while those below were preserved. The digital device was tested on 17 experienced hearing aid users with moderate to severe sensorineural hearing losses and sloping audiograms. In a monosyllabic word recognition test, eight subjects showed a significant improvement in recognition scores, and one showed a decrease in scores. Simpson et al (2005) stated that differences in individual performances may have been a result of improper fitting parameters.

The dilemma in asserting the efficacy of frequency lowering is that one is never sure if the parametric settings on frequency lowering were optimal. Further aggravating the situation are the large individual differences that exist among potential subjects. Differences such as cognitive levels, the amount of distortion of the auditory system, the extent of potential cortical reorganization consequent to the hearing loss and subsequent hearing aid use (Palmer et al, 1998) could affect the outcome of the efficacy study even though the algorithm may have been truly effective. One must also address the individual factors in order to effectively realize or evaluate the potential benefit of the frequency lowering algorithm.

However, if the objective of the evaluation (of a frequency lowering algorithm) is to understand if, and how, the algorithm may provide useful speech cues, the use of hearing impaired listeners may not be critical. Indeed, the use of normal hearing listeners with a simulated hearing loss may be more desirable. This is because the impact of confounding variables such as the degree of hearing loss of the wearer, distortion and re-organization within the auditory system, and the optimal parametric settings for each listener may be minimized. Individual differences may still exist, but one can conclude with more certitude on the effectiveness of the algorithm. If the answer is affirmative, the evaluation may be extended to hearing impaired listeners. Hopefully, the studies with normal hearing individuals will also shed light on how to more effectively evaluate the algorithm in hearing impaired listeners. The use of a normal-hearing listener with a simulated hearing loss may be especially meaningful when evaluating a new signal processing algorithm.

Several studies on frequency lowering used individuals with normal hearing as subjects (e.g., Johansson and Sjögren, 1965; Takefuta and Swigart, 1968; Velmans, 1973; Reeder et al, 1975; Hicks et al, 1981; Fraga and Marotta, 2004). For example, Fraga and Marotta (2004) introduced an FFT (fast Fourier transformation)-based algorithm in which the frequencies above the highest energy band of the original speech signal were transposed to a fixed region along the slope of the hearing loss. Gain was applied to the transposed signal to compensate for the hearing loss. Twenty normal-hearing subjects with simulated high frequency hearing losses were evaluated on a phoneme identification task. The results of the study led the authors to hypothesize that despite the available extra cues, such cues may lead to confusion of the original signal unless training is also provided. Training may allow the subjects to develop the necessary associations between the percept associated with the new transposed acoustic cues and their underlying identification/meaning. Other researchers (Ahlström et al, 1968; Reed et al, 1985; McDermott and Dean, 2000) also included training in their evaluations on frequency lowering.

Recently, Widex introduced a linear frequency transposition algorithm called the Audibility Extender (AE). This algorithm linearly transposes frequencies above a user programmable start frequency to lower frequencies. It allows frequencies up to two octaves above the start frequency to be lowered linearly to one octave immediately below the start frequency. Because only the higher frequencies are lowered but the lower frequencies are spared, the pitch of the speech signal is preserved. This algorithm also differs from previous frequency compression schemes in that the newly lowered portion of the original signal occupies the same absolute bandwidth as it did before lowering. In addition, lowering is applied to all incoming signals above the start frequency regardless of voicing. This may minimize any discontinuities in the acoustic stream and reduce artifacts. Finally, the AE allows the clinician to adjust the parameters of the algorithm based on the hearing losses and preferences of the patients. This would ensure that the amount of frequency lowering is appropriate for each listener. Because of the uniqueness of this processing algorithm, we have conducted a series of studies aimed at understanding the efficacy of this algorithm for people with a hearing impairment. As a first step in this understanding, we conducted this preliminary study to evaluate if the algorithm provides additional cues for phonemes that will most likely benefit from transposition (i.e., voiceless consonants). In addition, we wanted to determine if directed training may be necessary to realize the usefulness of the transposed cues.

METHOD

Study Participants

A total of nine normal hearing listeners participated in the study. The hearing levels of all the participants were less than 10 dB HL in both ears at 500 Hz, 1000 Hz, 2000 Hz, and 4000 Hz. All were native English speakers ranging in age from 18 to 24 years. Six were females and three were males. Eight participants were recruited using an online advertisement. Seven of these study participants reported student (or part-time student) as their occupation. One participant reported musician as his occupation. One other participant was a staff audiologist at the research facility who was not familiar with the current study. None of the study participants reported skills in languages other than English. All participants signed an informed consent and were financially compensated for their participation. Using young, normal-hearing individuals allowed us to minimize variability in results that may occur from a difference in hearing levels and cognitive functioning among study participants. Obviously, differences in listener background and motivation could still affect the outcome of the study.

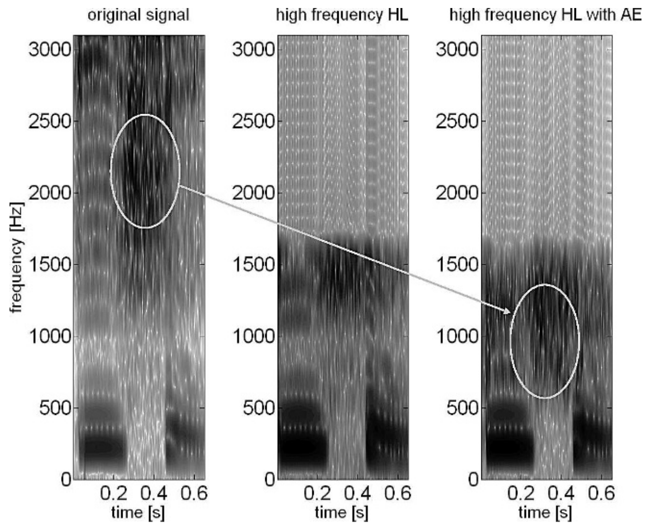

Figure 1. An example of linear frequency transposition for syllable /iʃi/: original signal *(left)*; signal with high-frequency hearing loss *(center)*; frequency transposed signal with high-frequency hearing loss.

Hearing Instrument

The Widex Inteo IN-9 digital mini behind-the-ear (BTE) hearing aid was used in the study. It is a 15 channel, wide dynamic range compression hearing aid that has an input dynamic range of 107 dB SPL and a frequency response from 100 Hz to 7450 Hz (ANSI S3.22-2003). It uses a sampling rate of 32 kHz with a 32-bit resolution. This hearing aid also has two distinctly different noise reduction algorithms that adjust its frequency characteristics in noise, a 15-channel fully adaptive directional microphone, and a multidirectional active feedback cancellation algorithm, as well as an algorithm that estimates the effective vent diameter and compensates for its effect. These features, however, were deactivated when the hearing aid was used to record the test stimuli in this study.

Linear Frequency Transposition—Audibility Extender (AE)

The linear frequency transposition algorithm on the Inteo hearing aid is an optional feature that can be activated if desired. The readers are referred to Andersen et al (2006) for a detailed description of its various stages of signal processing. Briefly, the AE algorithm first detects the most dominant spectral peak located in the *source octave*, or the octave immediately above a user programmable *start frequency*. Start frequency is set based on the degree and location of the slope of the hearing loss. The frequency components above the start frequency (or in the source octave) are transposed; the frequency components below start frequency are left unprocessed. Typically, the dominant peak is lowered by one octave, and the other frequencies at the source octave are lowered linearly by the same amount. The transposed signal is then band-pass filtered around the transposed peak (one octave bandwidth) to limit its masking effects on the original signal. Finally, the transposed sounds are amplified according to the hearing loss of the wearer at the transposed frequency and mixed with the original signal as the final output.

Unlike other frequency lowering methods reported in the past, the AE algorithm adapts to the spectrum of the input signal. At any given moment, the absolute amount of frequency lowering is directly related to the location of the dominant peak in the source octave. For a simple stimulus, this action guarantees that the harmonic relationship of the transposed and original signals is preserved. Figure 1 shows an example of the spectrogram for the syllable /iʃi/ when AE is used. High frequency cues of the original signal (left), which are not audible in the presence of a high frequency hearing loss above 1600 Hz (center), are made audible by lowering them to a lower frequency region (right).

The AE algorithm includes two programmable parameters: start frequency and AE gain. Ten start frequencies between 630 Hz and 6000 Hz at 1/3 octave intervals are available. The AE gain can be increased by 14 dB or decreased by 16 dB over its default position. The clinician can also select source region to include either one (Basic mode) or two octaves (Expanded mode) of transposition above the start frequency.

Test and Training Materials

To maximize the potential benefit of frequency transposition, we opted to use voiceless consonants

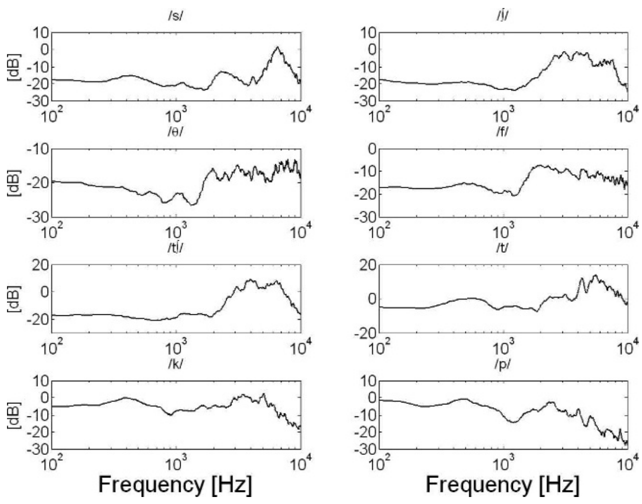

Figure 2. Average spectra of each phoneme used in the current study.

(/s/, /ʃ/, /θ/, /f/, /tʃ/, /t/, /p/, and /k/) as the test materials in this study because these sounds would most likely have dominant energy in the higher frequencies. Also, these phonemes are produced without voicing which would add lower energy component to the signal. Frequency transposition would potentially result in partial masking for sounds with acoustic cues at lower frequencies.

An identification test was created using these consonants presented in the initial, medial, and final positions with the vowel /i/ (i.e., CV, VCV, or VC). Thus, a total of 24 syllables were assessed. The syllables were spoken by a female native English speaker and recorded in a low noise level (22 dB SPL-A, slow, Quest model 1800 SLM [sound level meter]) double-walled IAC audiometric test booth (3 m × 3 m × 2 m). The sampling frequency of the 16-bit recordings was 44.1 kHz.

Spectral analysis of the source material was carried out to verify that these stimuli contained sufficient high frequency content to assess the benefits of frequency transposition (Figure 2). Averaged spectra of each phoneme were estimated using 1024-point FFT with Hamming windowing. Results were averaged in a 15 FFT point width window to aid visualization. Fricative /s/ has the highest spectral peak around 6300 Hz. Other phonemes have less pronounced peaks in their respective spectra, yet all phonemes have significant energy located at higher frequencies and rising spectral slope above 1600 Hz.

After the initial recordings, the stimuli were played back at 68 dB SPL-A through a single loudspeaker (KRK ST6 passive studio monitor) placed on-axis 1 m in front of an Inteo™ IN-9 hearing aid connected to a 2-cc coupler placed in the same audiometric test booth where the original recordings were made. The output of the hearing aid was recorded through a Quest™ 1800 sound level meter connected to a personal computer (HP Compaq DC5100 Intel® Pentium® 4) using an EchoGina24 soundcard, and Adobe™ Audition™ audio recording software.

The hearing aid was programmed two ways to process the played-back signals. One was with frequency transposition (*AE-on*) and one without transposition (*AE-off*). The start frequency of the transposition was set at 1600 Hz. This relatively low start frequency was chosen to minimize the use of mid frequency cues for identification of stimuli with a broad spectrum. For example, Figure 2 shows that the spectrum of /ʃ/ extends from 2600 Hz to 7700 Hz with the peak at around 3500 to 4000 Hz. A start frequency at 2500 Hz or higher could leave the /ʃ/ sound identifiable. Source region of the frequency transposition included sounds at two octave bands above the start frequency (i.e., expanded mode). In both programs the in-situ thresholds (sensogram) that were used to specify hearing aid gain were set to 20 dB HL at and below 1000 Hz, 50 dB HL at 1250 Hz, 70 dB HL at 1600 Hz, and 90 dB HL at and above 2000 Hz. The hearing aid microphone was set to the omnidirectional mode. Active feedback cancellation was deactivated. The Widex fitting software Compass (version 4.0) was used to program the hearing aid.

To simulate a hearing loss above 1600 Hz, the output recordings of the hearing aid were low-pass filtered offline with a FFT-filter created in Adobe Audition in order to realize steep transitions, extreme stop band attenuation, and linear phase. The FFT filter used in this study had attenuation of 0 dB for frequencies up to 1550 Hz, linearly increasing attenuation between 1550 Hz and 1650 Hz so that the attenuation was 90 dB at and above 1650 Hz. Blackman window with a length of 8196 samples was used. Finally, all the stimuli were equalized to the same maximum peak RMS (root mean square) level in a 50 msec sliding window.

Procedures

Testing was conducted in a large office where the background noise level was 47 dB-A (slow) (Quest Technologies headphone calibration system). Testing was conducted using a personal computer (HP Compaq DC5100 Intel Pentium 4 with 17″ LCD monitor). The testing software and associated graphical user interface were implemented using Matlab 6.5 (R13).

The study participants were seated approximately half a meter from the computer screen. The recorded stimuli were presented through headphones diotically (Sennheiser HD 580). Participants adjusted the presentation level to their comfortable listening level prior to testing by listening repeatedly to /ipi/ recorded without frequency transposition (AE-off) with one

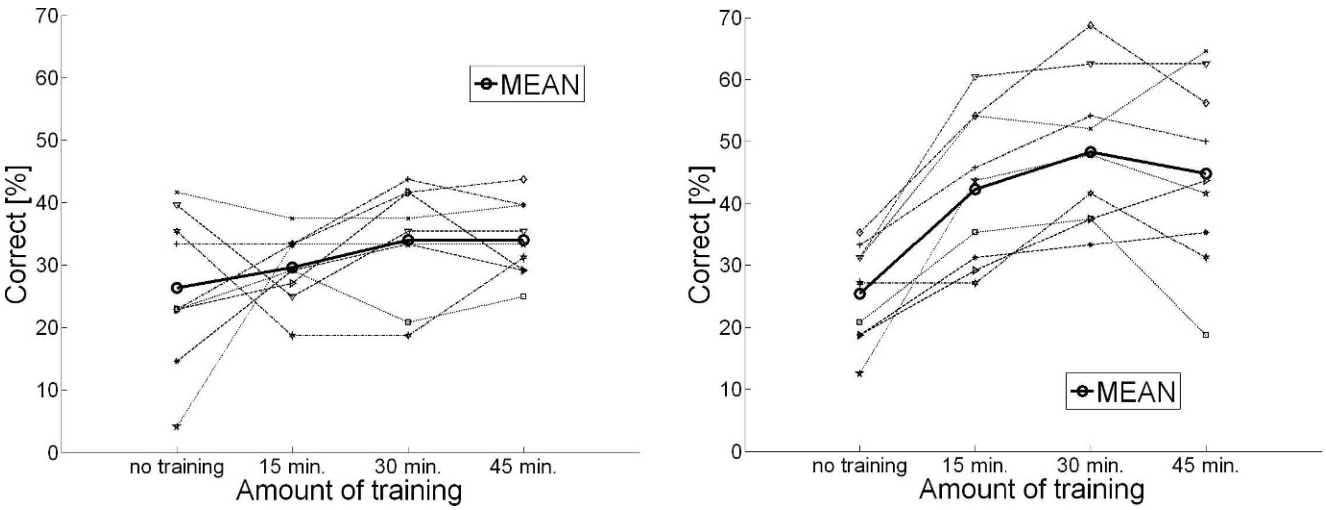

Figure 3. Identification scores for phonemes without *(left)* and with *(right)* frequency transposition at different time intervals.

second pause between repetitions. The selected A-weighted presentation levels measured in a headphone coupler ranged between 70 and 80 dB SPL with a mean of 77 dB SPL. Prior to testing, the study participants were familiarized with the unfiltered syllables. They were asked to verbally repeat each syllable and to click on the appropriate written form of the syllables that appeared on the computer monitor.

The stimuli recorded with frequency transposition (AE-on) and without transposition (AE-off) were mixed and presented together in random order during each trial. After each stimulus presentation, the study participant clicked on the corresponding button on the computer screen to identify the syllable. Each stimulus in the list was presented randomly twice. Thus, a total of 96 stimuli were presented during each trial (24 syllables recorded with AE-on and AE-off, each presented twice). A one-second pause was inserted between stimulus presentations. The listeners had no option of repeating the stimulus.

The phoneme identification test was conducted four times (trials) in a single session. The participants were informed of their errors at the end of each test so they could focus their attention on the most difficult sounds during the training trial that immediately followed each test trial. Each participant received three 15-minute, self-paced training sessions interleaved between test presentations (3×15 minutes = 45 minutes of training total). The same equipment setup that was used during testing was used for the training. The stimuli that were recorded with the AE-on condition were used as the training materials. The written versions of all the syllables were listed and shown on the computer screen. The study participants were allowed to listen to any of the syllables. They could select either a single syllable or two different syllables from the list in order to hear both syllables in a pairwise manner. The presentation level during the training was identical to that used during testing. A timer was also shown on the computer screen to enforce the 15-minute time limit. Each test trial lasted approximately 10 minutes. Participants were allowed to have short breaks between test trials and training sessions. Each listener spent approximately two hours to complete the testing and training.

RESULTS

A repeated measures analysis of variance was carried out with phoneme position, AE condition, and time of training as factors. Analysis revealed statistically significant effect for AE condition ($F[1,8] = 1.092$, $p < 0.005$) and for time of training ($F[3,24] = 15.871$, $p < 0.001$) as factors. However, phoneme position was not a statistically significant factor ($F[2,16] = 0.773$, $p = 0.478$). Thus, to streamline the presentation of the results, all subsequent reporting will be based on the results averaged over all three phoneme positions. Figure 3 shows the individual and averaged overall correct identification scores measured at the four intervals—no training, after 15, 30, and 45 minutes of training. Figure 3a shows the results from the AE-off condition, while Figure 3b shows the results from the AE-on condition.

Effect of AE Prior to Training

The mean identification scores for the two conditions before any training were 26.4% for the AE-off condition and 25.5% for the AE-on condition. The difference in identification scores between the AE conditions before training was not statistically significant as examined with a repeated measures ANOVA ($F[1,8] = 0.122$, $p = 0.738$).

Table 1. Immediate Effect of Using Linear Frequency Transposition without Training

Phoneme	*AE-off pretraining [%]*	*AE-on pretraining [%]*	*Improvement*	*Statistical significance p (paired t-test)*
/p/	16.6	31.5	14.9	0.05
/t/	25.9	37.0	11.1	0.15
/f/	14.8	20.4	5.6	0.5
/tʃ/	31.5	29.6	−1.9	0.7
/k/	42.6	38.9	−3.7	0.6
/θ/	25.9	18.5	−7.4	0.2
/ʃ/	33.3	24.1	−9.2	0.1
/s/	33.3	13.0	−20.3	0.1

Effect of AE with Training

Figure 3 also showed the identification scores after 15, 30, and 45 minutes of training. For the AE-off condition, the mean score ranged from 26.4% obtained without training to 34.0% obtained after 45 minutes of training. One-way ANOVA analysis for AE-off condition with amount of training as a factor shows that changes in performance between training intervals were not statistically significant ($F[3,32] = 1.669$, $p = 0.193$). With the AE-on condition the performance improved 16.9% from the initial 25.5% to 42.4% after 15 minutes of training. The mean identification score further improved to 48.4% after 30 minutes of training. An additional 15 minutes of training led to a decrease in mean identification score by an average of 3.5%. This may be explained by decreasing attention at the last test trial. One-way ANOVA shows that the improvement over time was statistically significant ($F[3,32] = 6.347$, $p < 0.005$).

The overall advantage in performance scores with AE-on over AE-off was statistically significant when analyzed using general linear model repeated measures analysis of variance with AE condition and amount of training as within-subject factors (AE condition: $F[1,8] = 18.009$, $p < 0.005$, $r = 0.8321$; amount of training: $F[1,8] = 22.353$, $p < 0.001$, $r = 0.8582$).

Phonemic Analysis

A set of stimulus-response (or confusion) matrices were constructed by combining the results of all participants in order to more closely analyze the error patterns at the phoneme level.

Immediate Effects of Linear Frequency Transposition (AE-on − AE-off at initial visit)

The immediate effect on phoneme identification when AE was introduced was listed in Table 1. Immediately, transposition improved the identification of /p/ without any training. Surprisingly, transposition negatively affected the identification of high frequency fricative sounds such as /s/ and /ʃ/. Intermediate effects (which were all statistically nonsignificant) were seen on other phonemes.

Error Patterns without Training

The stimulus-response matrices for AE-off and AE-on conditions obtained before training are presented in Figure 4a and b, respectively. The number of responses is shown in the cell for each stimulus-response pair.

a

Stimulus \ Response	tʃ	ʃ	f	s	θ	k	p	t
tʃ	17	4		1	6	11	5	10
ʃ	8	18	9	8	7	1	2	1
f	6	8	8	14	9	5	2	2
s	9	10	6	18	7	3		1
θ	11	8	10	9	14	1		1
k	4			1	4	23	6	16
p	7	1	3	1	3	18	9	12
t	8	1		2	1	21	7	14

b

Stimulus \ Response	tʃ	ʃ	f	s	θ	k	p	t
tʃ	16	4	3	4	5	14	4	4
ʃ	11	13	11	9	5	3	2	
f	7	4	11	10	9	3	6	4
s	6	15	12	7	9	5		
θ	9	9	11	8	10	3	3	1
k	4	2			4	21	7	16
p	5	1			3	15	17	13
t	6				1	17	10	20

Figure 4. *(a)* Stimulus-response matrix for AE-off condition before training. *(b)* Stimulus-response matrix for AE-on condition before training.

Table 2. Benefit of Frequency Transposition in Identification Scores in Comparison with Conventional Amplification

Phoneme	*AE-off pretraining [%]*	*Best score with AE [%]*	*Improvement [%]*	*Statistical significance p (paired t-test)*
/p/	16.6	70.4	53.8	0.005
/tʃ/	31.5	64.8	33.3	0.05
/ʃ/	33.3	59.3	26.0	0.06
/θ/	25.9	46.3	20.4	0.09
/f/	14.8	33.3	18.5	0.11
/t/	25.9	44.4	18.5	0.06
/k/	42.6	55.6	13.0	0.33
/s/	33.3	31.5	−1.8	0.9

Responses that fell on the diagonal of the matrix were correct responses, and all responses located off the diagonal were the number of times the specific stimulus-response pair were incorrect. Each stimulus was presented twice in each test trial in each of the three positions ($2 \times 3 = 6$), and matrices were combined from the results of nine participants resulting in $6 \times 9 = 54$ stimulus presentations for each row of the matrix.

The confusion matrix for AE-off without training reveals that phoneme /k/ was identified correctly most consistently (23/54: 42.6%). On the other hand, /t/ was misidentified as /k/ 38.9% of the time and correctly identified as /t/ 25.9% of the time. Identification scores for all other phonemes were between 14.8% and 33.3%.

The error pattern for AE-on before training was unsystematic. The accuracy of identification for all phonemes was below 40%. The results showed that fricatives (/ʃ/, /f/, /s/, /θ/) were correctly identified less frequently (13.0–24.1%) than stops (/k/, /p/, /t/) (31.5–38.9%). Fricatives like /ʃ/, /s/, and /θ/ were often incorrectly identified as /f/. With AE-off, there was no difference in the accuracy of identification between stops and fricatives before training.

Benefit of Using Linear Frequency Transposition

Table 2 summarizes the maximum improvements from frequency transposition (best performance with AE-on) from the control condition (AE-off before training). The best scores were obtained after 15 minutes of training for phoneme /k/; after 45 minutes for the phonemes /tʃ/ and /s/, and after 30 minutes for all other phonemes. There was a 33.3 to 53.8% improvement in phoneme scores for the phonemes /p/ and /tʃ/. For the phonemes /ʃ/, /θ/, /f/, /t/, the improvement was between 18.5 and 26.0%, but its statistical significance was marginal. For the phonemes /k/ and /s/, the change in performance was not statistically significant.

Error Pattern after 30 Minutes of Training

The previous results suggest the highest identification scores for most phonemes were obtained after 30 minutes of training. Thus, we restricted our analysis to data obtained at this time interval to examine the combined effect of transposition and training. Figure 5a and b show the stimulus-response matrices for AE-off and AE-on conditions obtained at the end of 30 minutes of training.

The identification scores for each phoneme ranged between 22.2 and 35.2% for the AE-off condition. An exception to this was the identification of the phoneme /k/, which was identified correctly 64.8% of

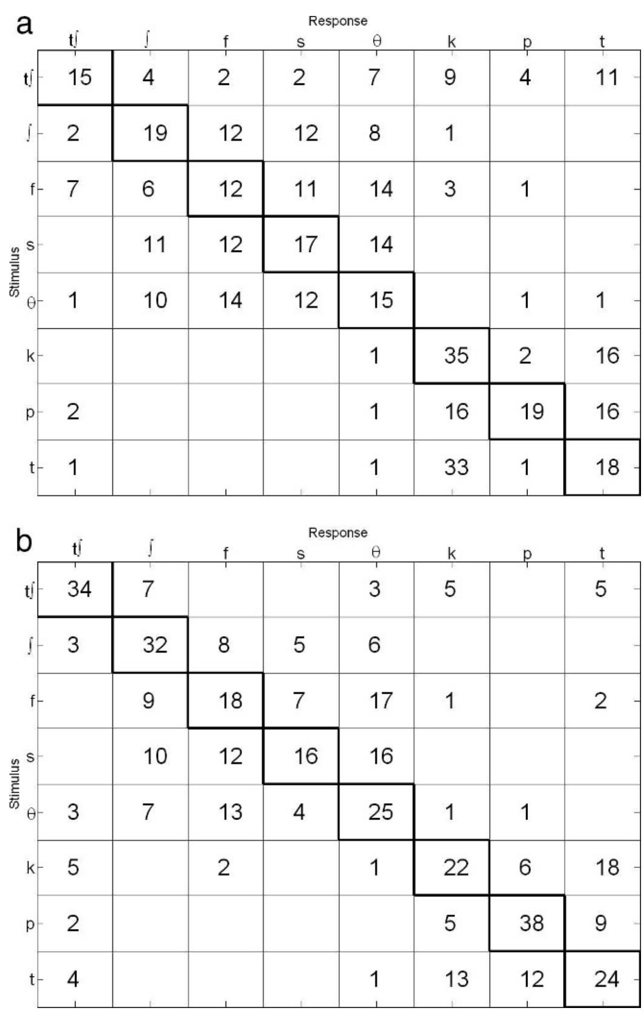

Figure 5. *(a)* Stimulus-response matrix for AE-off condition at the end of 30 minutes of training. *(b)* Stimulus-response matrix for AE-on condition at the end of 30 minutes of training.

Table 3. Phoneme Scores with "AE-on" before Training and after Various Intervals of Training

	AE-on				
Phoneme	*No training [%]*	*15 minutes of training [%]*	*30 minutes of training [%]*	*45 minutes of training [%]*	*Max. improvement [%]*
/p/	31.5	55.6	70.4	70.4	38.9
/tʃ/	29.6	50.0	63.0	64.8	35.2
/ʃ/	24.1	53.7	59.3	50.0	35.2
/θ/	18.5	38.9	46.3	35.2	27.8
/s/	13.0	24.1	29.6	31.5	18.5
/k/	38.9	55.6	40.7	38.9	16.7
/f/	20.4	22.2	33.3	31.5	13.0
/t/	37.0	42.6	44.4	44.4	7.4

the time. Figure 5b showed that AE-on plus training improved identification scores significantly. The phonemes /tʃ/, /ʃ/, and /p/ were identified especially well (/tʃ/: 34/54 or 63.0%; /ʃ/: 32/54 or 59.3%; /p/: 38/54 or 70.4%). The phonemes /f/ and /s/ were identified most poorly. The phonemes /f/ (18/54 or 33.3%) and /s/ (16/54 or 29.6%) were often misidentified as /θ/. The phoneme /f/ was identified as /θ/ 31.5% of the time (labio-dental as dental), and /s/ was identified as /θ/ 26.9% of the time (alveolar as dental). Identification scores for other phonemes with AE-on were /θ/: 25/54 or 46.3%; /k/: 22/54 or 40.7%; /p/: 38/54 or 70.4%; /t/: 24/54 or 44.4%.

Change in Phoneme Scores at Different Training Intervals

The effect of training on individual phoneme identification during the AE-on condition at different time intervals (before training, 15, 30, and 45 minutes of training) was summarized in Table 3. Results showed the greatest improvement for the phonemes /tʃ/, /ʃ/, and /p/, where over 35% of improvement was noted after 30 minutes of training. The identification score for the phoneme /k/ showed the greatest improvement after 15 minutes of training (55.6%), only to return to the initial level (38.9%) after 45 minutes of training. The phoneme /t/ improved the least (7.4%) with training.

DISCUSSION

The current study demonstrated that normal-hearing individuals with a simulated high frequency hearing loss at and above 1600 Hz can potentially benefit in voiceless consonant identification from frequency transposition. Such a benefit becomes apparent only after the listeners were self-trained to develop the associations between the new acoustic cues and the underlying meaning of the sounds. Although individuals showed a large variation in improvement, an overall improvement of 10 to 15% may be expected with voiceless consonants.

What Happens with Linear Frequency Transposition?

Frequency transposition lowers the high frequency information into a lower frequency form so the wearer may utilize them. Unfortunately, such information may not be immediately recognized by all wearers. For example, Table 1 shows that the identification of /p/ was immediately improved by almost 15% with transposition but the identification of /s/ was decreased by over 20% with transposition initially. Since the transposed sounds are added to the original sounds below 1600 Hz, the decreased score (as in the case of /s/) would suggest that the transposed signals could confuse the original signals. This may be the reason why no overall improvement with transposition was seen at the initial trial in this study. The same observation (of improvement in some phonemes and degradation in other phonemes) has also been reported by other investigators (Hicks et al 1981; Fraga and Marotta, 2004).

The observations made on phoneme identification with transposition may be explained by the interaction between the start frequency of transposition and the acoustic characteristics of each phoneme tested. For example, the most improvement was seen with the /p/ phoneme. Acoustically, this phoneme has the least energy and the lowest frequency compared to the other voiceless stops (Ladefoged, 1993). Given the start frequency used in this evaluation was at 1600 Hz, it is likely that the majority of the /p/ phoneme was transposed without interfering with the original signal. On the other hand, the /s/ phoneme has the most energy above 4000 Hz, yet the start frequency used in this study was set at 1600 Hz. Consequently, only a small fraction of the original /s/ phoneme was transposed. Furthermore, since the algorithm locks onto the frequency within the transposed region (source octave) with the highest energy peak for transposition, the risk of locking onto a frequency unrelated to the /s/ phoneme could be high since most of the /s/ spectrum is beyond

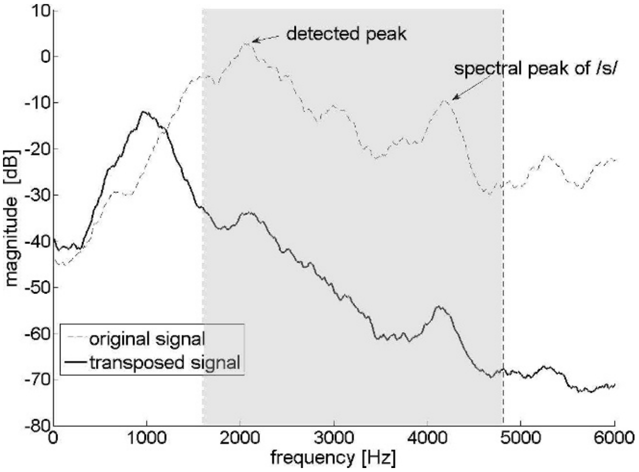

Figure 6. An example of incorrectly detected spectral peak when using relatively low start frequency (1600 Hz). The high-frequency spectral peak of phoneme /s/ is not detected because it is lower in level than other parts of the spectrum within the source region.

the source octave (see Figure 6). This may explain why the effect of transposition was negative with the /s/ phoneme. Even with training, the improvement on /s/ identification was slight. These observations suggest that when it comes to the actual use of frequency transposition, the choice of start frequency and how it relates to the spectra of the target sounds could be critically important (Kuk et al, 2006).

Importance of Training

Another finding of the current study that resonates with the impressions of many other researchers in the area (Johansson, 1966; Ahlström et al, 1968; Reed et al, 1985; Turner and Hurtig, 1999) is the importance of training in realizing the benefit of transposition. The new acoustic cues provided by frequency transposition may not always be immediately recognized by a listener. Indeed, some may even confuse the listeners. This is seen in Table 1 where the identification of /θ/, /ʃ/, and /s/ decreased by more than 5%. Luckily, with training on the transposed sounds, the identification of all the phonemes improved. Table 2 showed, with the exception of /s/, improvements over the initial AE-off condition that ranged from 13% to over 55% as a result of the combined training and use of transposition.

The extent of the improvement varied with the phonemes tested. However, most of the improvement may be seen within the first 30 minutes of training. Indeed, Table 3 shows that the averaged identification scores with AE-on improved from 25.5 to 48.4% after 30 minutes of training. This suggests that to evaluate the efficacy of a frequency transposition algorithm, it is important to allow the listeners to acclimate themselves to the new sounds so that they can learn to associate the new acoustic cues that the algorithm creates with the corresponding phonemes. It is important to note that the current study design of a 3 × 15 minutes training in a single session was chosen to answer the simple question of whether transposition results in usable cues. We do not suggest that the performance would plateau after 30 minutes, or that identification results would not improve further with additional visits.

Training improves not only the performance of the trained sounds but also that of the nontrained sounds. In this study, despite training the listeners only on the AE-on sounds, the performance with the AE-off condition improved from the initial 26.4 to 34.0% after 45 minutes of training. This improvement may be explained by procedural learning. However, it should be noted that the current transposition algorithm retains as much of the original acoustic information as possible. All the acoustic information that is available in the original signal below the start frequency was also available in the frequency transposed signal. Thus, even though listeners were trained on the frequency transposed sounds, they were also trained on the original sounds (and the associated cues) below 1600 Hz. Nonetheless, this observation highlights the importance of training in the hearing aid fitting/rehabilitation process.

Individual Differences among Study Participants

Although auditory training significantly improved the mean identification score with frequency transposed stimuli, there were large variations in individual results despite the use of young normal hearing listeners with identical simulated hearing losses. Their performance ranged from a minimum of 29.2% to a maximum of 43.8% in the AE-off condition and a minimum of 35.4% to a maximum of 68.8% in the AE-on condition. This represents a between-subject variation of 14.6% in the AE-off condition and 35.4% in the AE-on condition. This suggests that some listeners were more able than others in utilizing the transposed cues for identification. One possible factor could be the listeners' motivation to listen. Two of the best-performing and most enthusiastic participants reported "musician" and "speech pathology student" as their professions. They approached the auditory training in a more analytical and structured way than the other listeners. Clinically, this would suggest that a requirement to examine the efficacy of any signal processing algorithm, and specifically the frequency transposition algorithm, is not only to engage the listeners in training but also to identify the factors that would motivate the listeners during the training.

Limitations of the Study

Despite the encouraging findings, we cannot generalize the present findings to hearing impaired listeners or to speech materials not used in this study. Furthermore, the programmable parameters of the transposition algorithm were set identically for all the listeners. When a different start frequency is used, the acoustic cues provided by the transposition algorithm will be different. This will possibly create different types of errors than the ones reported in this study. Additionally, the current study only used voiceless consonant sounds as stimuli. Other phonemes, such as voiced consonants, vowels, and so on, were not evaluated. While frequency transposition may not affect the identification of these sounds directly, the transposed portion of the signal may cause frequency masking and result in more difficulty for the identification of lower frequencies sounds when they are included in the inventory of materials. Finally, the current study was limited in its use of only a single female speaker. The exact locations of spectral cues are speaker dependent. For example, Stelmachovicz et al (2001) showed that for hearing impaired subjects the optimum bandwidth for identification of unvoiced fricative /s/ was over 9 kHz when female or child speaker was used and between 4 and 5 kHz for male speakers. Under such speaker-dependent variations, the error patterns can be expected to differ from each other. An extension of the current study is underway to study the recognition of all vowels and consonants of American English. In addition, the use of listeners with various hearing loss configurations was also included in upcoming studies to examine the efficacy of such an algorithm.

Acknowledgments. The authors would like to thank Dr. Chi Lau for his advice in the statistical analysis of this research. Also, the authors would like to express gratitude to the reviewers for extensive and constructive criticism in the review of the manuscript.

REFERENCES

Ahlström KG, Risberg A, Lindhe V. (1968) *Experiment in Learning to Discriminate Frequency Transposed Speech*. Rpt. 36. Uppsala: Institute of Education, Uppsala University, Sweden.

Andersen H. (2006) *Audibility Extender—So the "Dead" (Region) May Hear, Integrated Signal Processing, A New Standard in Enhancing Hearing Aid Performance*. *Hear Rev* (Feb., Suppl.).

Braida L, Durlach I, Lippman P, Hicks B, Rabinowitz W, Reed R. (1979) Hearing aids—A review of past research of linear amplification, amplitude compression and frequency lowering. *ASHA Monogr* 19:1–114.

David EE, McDonald HS. (1956) Note on pitch-synchronous processing of speech. *J Acoust Soc Am* 28:1261–1266.

Fairbanks G, Everitt WL, Jaeger RP. (1954) Method for time or frequency compression-expansion of speech. *Trans IRE Prof Group Audio* 2(1):7–12.

Fletcher H. (1953) *Speech and Hearing in Communication*. Princeton: Van Nostrand.

Fraga FJ, Marotta AM. (2004) Frequency compression and frequency shifting for the hearing impaired. In: *Biomedical Engineering*. Tilg B, ed. IASTED (International Association of Science and Technology for Development), Track 417–808.

Hicks BL, Braida LD, Durlach NI. (1981) Pitch invariant frequency lowering with nonuniform spectral compression. In: Vol. 6 of *Acoustics, Speech, and Signal Processing, IEEE International Conference on ICASSP '81*. IEEE, 121–124.

Johansson B. (1961) A new coding amplifier system for the severely hard of hearing. ln: Vol. 2 of *Proceedings of the 3rd International Congress on Acoustics*. Amsterdam: Elsevier, 655–657.

Johansson B. (1966) The use of the transposer for the management of the deaf child. *Int J Audiol* 5:362–373.

Johansson B, Sjögren H. (1965) The effect on normal hearing subjects and profoundly deaf children by coding voiceless consonants using frequency transposition. In: *Proceedings of the 5th International Congress on Acoustics*. ICA, 1–4.

Kuk F, Korhonen P, Peeters H, Keenan D, Jessen A, Andersen H. (2006) Linear frequency transposition: extending the audibility of high frequency information. *Hear Rev* 13(10):42–48.

Ladefoged P. (1993) *A Course in Phonetics*. Fort Worth, TX: Harcourt Brace College Publishers.

McDermott HJ, Dean MR. (2000) Speech perception with steeply sloping hearing loss: effects of frequency transposition. *Br J Audiol* 34:353–361.

Moore B. (2004) Dead regions in the cochlea: conceptual foundations, diagnosis, and clinical applications. *Ear Hear* 25(2):98–116.

Muñoz CMA, Nelson PB, Rutledge JC, Gago A. (1999) Frequency lowering processing for listeners with significant hearing loss. In: Vol. 2 of *Proceedings of ICECS '99. The 6th IEEE International Conference on Electronics, Circuits and Systems, 1999*. IEEE, 741–744.

Palmer CV, Nelson CT, Lindley IV GA. (1998) The functionally and physiologically plastic adult auditory system. *J Acoust Soc Am* 103(4):1705–1721.

Raymond TH, Proud GO. (1962) Audiofrequency conversion. An aid for rehabilitation in neurosensory hearing loss. *Arch Otolaryngol* 76:436–446.

Reed CM, Schultz KI, Braida LD, Durlach NI. (1985) Discrimination and identification of frequency-lowered speech in listeners with high-frequency hearing impairment. *J Acoust Soc Am* 78(6):2139–2141.

Reeder K, Strong WJ, Palmer EP. (1975) Low-frequency formant speech code for the severely/profoundly hearing impaired. *J Acoust Soc Am* 58:S113.

Rees R, Velmans M. (1993) The effect of frequency transposition on the untrained auditory discrimination of congenitally deaf children. *Br J Audiol* 27:53–60.

Simpson A, Hersbach AA, McDermott HJ. (2005) Improvements in speech perception with an experimental nonlinear frequency compression hearing device. *Int J Audiol* 44:281–292.

Stelmachowicz PG, Pittman AL, Hoover BM, Lewis DE. (2001) Effect of stimulus bandwidth on the perception of /s/ in normal- and hearing-impaired children and adults. *J Acoust Soc Am* 110(4):2183–2190.

Takefuta Y, Swigart E. (1968) Intelligibility of speech signals spectrally compressed by a sampling-synthesizer technique. *IEEE Trans Audio Electroacoustics* AU-16(2):271–274.

Turner CW, Hurtig RR. (1999) Proportional frequency compression of speech for listeners with sensorineural hearing loss. *J Acoust Soc Am* 106(2):877.

Velmans M. (1971) Aids for Deaf Persons. U.K. Patent 1340105.

Velmans M. (1973) Speech imitation in simulated deafness, using visual cues and 'recoded' auditory information. *Lang Speech* 16: 224–236.

Velmans M. (1974) The design of speech recording devices for the deaf. *Br J Audiol* 8:1–5.

Velmans M. (1975) Effects of frequency "recoding" on the articulation learning of perceptively deaf children. *Lang Speech* 18:180–193.