EFFECT OF FAST AND SLOW COMPRESSION ON VOICE ONSET TIME IN HEARING AID PROCESSED SPEECH

[1]**Ajish K. Abraham**

Abstract

Compression introduces distortion in hearing aid processed speech, when compared with linear amplification. Static parameters of compression namely compression ratio, compression threshold, attack time (AT), release time (RT) as well as the number of compression channels of a hearing aid may be the potential sources of compression induced distortion. The role of hearing aid compression parameters on speech perception and sound quality of hearing aids has been investigated in several studies. In some studies, it has been reported that, long ATs/ RTs leads to improvement in intelligibility of speech sounds. However, further research is required to find out the factors responsible for this. The present study therefore, investigated the influence of slow and fast compression time constants on transformation of voice onset time (VOT) in hearing aid processed speech. The objectives of the study were (1) to measure the Attack Time and Release Time of four behind the ear Digital Hearing Aids programmed in fast and slow AT/RT settings and (2) to investigate the effect of slow and fast AT/RT on VOT transformation. Six plosives in CV syllables categorized as: - unvoiced – velar /k/, dental /t/, and bilabial /p/, voiced – velar /g/, dental/d/ and bilabial /b/ produced by an adult (native Kannada) male speaker at a comfortable loudness level were recorded and used as stimulus for the study. AT and RT values of all four hearing aids in slow and fast compression setting were evaluated as per the test procedure specified by American National Standards Institute (ANSI) S3.22 standard. VOT values were measured for the input stimuli and the digitized hearing aid processed speech samples by using speech analysis software and the transformation in VOT was estimated. Results showed a reduction in VOT for voiced and unvoiced plosives at fast compression time constant setting. However, no changes in VOT was noticed at slow AT/RT setting, for unvoiced plosives, but significant difference was observed for voiced plosives.

Keywords: *Compression Speed, Attack Time, Release Time, Speech intelligibility, Temporal Distortion*

Speech includes a wide and dynamic range of intensity levels, ranging from low intensity consonants (for eg: /f/) to high intensity vowels (for eg: /i/). Intensity levels also vary from soft speech (low intensity) to loud speech (high intensity). To bring all these intensity levels within the dynamic range of hearing of a person with hearing impairment, most hearing aids offer some form of compression in which gain is automatically adjusted on the basis of the intensity of the input signals. Hearing aids with wide dynamic range compression (WDRC) feature adjust their gain automatically, based on the measured level of the input signal. This may introduce distortions to input signals, which would be much higher than which is introduced in hearing aids with linear amplification. Static parameters of compression namely compression ratio, compression threshold, attack time (AT), release time (RT) as well as number of compression channels in hearing aid are all potential sources of compression induced distortion. As reported by Souza, 2002, the speed at which a hearing aid alters its gain according to changes in input sound levels, is considered as a major parameter of a compression hearing aid. As defined by American National Standards Institute (ANSI), attack time is the time interval between the abrupt increase in the input sound level and the point when the output sound level from the hearing aid stabilizes to within + 2dB of the elevated steady state level. Similarly, release time is the time interval between abrupt reduction in the steady state input signal level after the hearing aid amplifier has reached the steady state output under elevated input conditions, and the point at which the output sound pressure level from the hearing instrument stabilizes within +2dB of the lower steady state level (Vonlanthen & Arndt, 2007).

The influence of hearing aid compression parameters on speech perception and sound quality of hearing aids has been investigated in several studies done by Gatehouse, Naylor & Elberling, (2006); Hansen, (2002); Moore, Stainsby, Alcantara, & Kuhnel, (2004); Neuman, Bakke, Mackersie, Hellman, & Levitt, (1998) (Quoted by Shi & Doherty, 2008). A literature review by Gatehouse et al., 2006), reveals that

[1]Reader in Electronics & Acoustics, All India Institute of Speech and Hearing (AIISH), Mysore-06, Email: ajish68@aiishmysore.in

some researchers (e.g., Moore et al., 2004) have found no effect of time constants, while others have found fast-acting compression to be superior to slow acting compression (e.g., Nabelek & Robinette, 1977), while yet others have found slow-acting compression to be superior to fast-acting compression (e.g., Hansen, 2002). Ole Hau & Andersen (2012) commented that this discrepancy is because, these studies were designed to investigate different outcomes. Some of the studies focused on the effect of time constant on speech intelligibility (e.g., Nabelek et.al., Nabelek & Robinette, 1977; Moore et al., 2004), while others (Neuman et al.; 1998; Hansen, 2002) focused on subjectively perceived sound quality. Gatehouse et. al. (2006) attributed this discrepancy to the differing signal processing strategies and compression algorithms incorporated in the hearing aids used for these studies.

Fast AT (≤ 10 ms) and fast RT (≤ 150 ms) helps the compression system to react to changes in sound pressure level between syllables, resulting in increased temporal distortion in comparison with slow AT/RT (Shi & Doherty, 2008). Stone and Moore (2007) reported that envelop distortion of running speech is more with fast compression when compared with slow compression (AT is ≥ 350 ms, RT ≥1080 ms). Shi et al., (2008) reported that with slow AT/RT, real ear aided response is better than fast AT/RT and linear amplification.

To stress the need for more research to find out which AT/RT settings are ideal for improved intelligibility for listeners, Shi et al., (2008) has analyzed the following studies. In an effort to find the combined effect of a fast AT/RT and high compression ratio (CR) on intelligibility of vowels (VCV syllables), Van Tasell and Trine (1996) conducted an experiment and found that the combination resulted in low intelligibility. Studies conducted by Moore & Glasberg, (1988); Moore, Glassberg, & Stone, (1991); Stone & Moore, (2004) have also shown that speech intelligibility increases with long ATs/RTs. But Gatehoue et al., (2006) and Jenstad & Souza, (2005) reported that there is no consistent increase in intelligibility of phonemes with long RTs. According to Jenstad et al., (2005) some plosives and affricates were reported to be better intelligible with a lesser RT (12 ms) in comparison with a longer RT (800 ms). Studies by Moore et al., (2004) and Stone et al., (1999) shows that there is no significant increase in speech intelligibility in case of slow AT/RT over fast or vice versa.

An investigation is required to find out the effect of different compression time constants on the temporal parameters of speech. Along with this, the factors responsible for the contradictory results need to be investigated. A recent study conducted by the author on hearing aid processed speech has revealed that the voice onset time gets transformed significantly while the speech signal passes through the signal processing stages in a digital hearing aid. VOT is defined as the time difference between the onset of burst and the onset of voicing in stop consonants (Kent & Read, 1995). VOT forms an important cue to signify voicing contrast in word-initial stop consonants. For unvoiced stops, the burst release takes place before the onset of glottal activity and for voiced stops it follows the onset of glottal activity. Shobha, Thomas and Subba Rao, (2009) studied the effect of VOT modification on speech perception in noise. They have reported that, for no noise presentations, VOT modification in voiced and voiceless stops reported a benefit up to plus 4% for few conditions. Nataraja, Savithri, Sreedevi and Sangeetha (2000) evaluated the improvement in speech intelligibility after transforming the temporal parameters in the speech of hearing impaired children. Results showed that VOT transformation was the one with maximum effect and its enhancement has led to significant increase in speech intelligibility. Hence VOT transformation in slow and fast AT/RT settings can be a potential factor influencing speech intelligibility, which could throw some light on the conflicting results reported in the previous studies. Hence the effect of slow and fast AT/RT on VOT modification needs to be investigated.

Gatehouse et al., (2006), reveals that, in all studies of compression, it is important to ensure that test conditions reflect a relevant diversity of signal characteristics. This is because of the fact that the characteristics of the test signals used may influence the results of evaluation of AT/RT. This is especially true as the object of present study is time constants, since varying these can fundamentally alter the nature of the nonlinearity. Hence there is a need to measure the AT & RT of all the hearing aids used in the study as per the protocol specified by American National Standard Institute (ANSI).

The present study therefore, investigated the influence of slow and fast compression time constants on transformation of VOT in hearing aid processed speech. The objectives of the study were (1) to measure the Attack Time and Release Time of four Digital Hearing Aids configured in fast and slow settings of compression speed and

(2) to investigate the effect of slow and fast AT/RT on VOT transformation.

Method

a) Experiment 1- Measurement of AT and RT

Stimuli: A 1,600 Hz pure tone signal at 55dB SPL was used as test signal. This signal was modulated by a square envelope pulse increasing the input level by 25dB. The pulse length was kept at least five times longer than the attack time being measured. Intervals between two consecutive pulses were kept at five times longer than the recovery time being measured (Vonlanthen et al., 2007).

Instrumentation: The setup used for measurement is shown in figure 1. Test signal, generated by a Matlab program in the computer as a wave file is downloaded to the front end of B & K pulse analyzer system. The hardware unit of the pulse analyzer system delivers the stimulus through B & K 2716C precision power amplifier and a leveled response loudspeaker to the hearing aid under test kept in the hearing aid test chamber (B & K 4232). Four behind the ear (BTE) digital hearing aids were used for testing. These hearing aids had the option to select two AT/RT settings:- slow and fast. All hearing aids were programmed with a flat gain curve for mild hearing loss. The output is acquired from all 4 hearing aids for each of the setting by B & K 4189 condenser microphone that is placed inside a 2cc coupler and are recorded by BZ 7226 recording setup and stored on the PC for analysis.

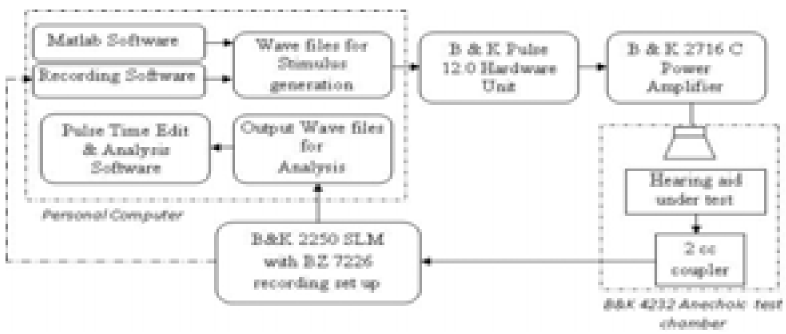

Figure 1*: Instrumentation for stimulus delivery and AT/RT/VOT Measurement*

Test environment: All the tests were carried out in an air conditioned sound treated room. The ambient noise levels were within permissible limits.

Measurement: AT/RT values were measured as specified by ANSI S3.22 standard from the digitized hearing aid output by using a speech analysis software (B & K Pulse Time Edit & Analysis). AT was measured as the duration between the abrupt increase in the input signal level and the point when the output sound pressure level from the hearing aid stabilizes to within ± 2 dB of the elevated steady state level, as illustrated in figure 2. RT was measured as the duration between the abrupt reduction in the steady state input signal level and the point when the output sound pressure level from the hearing aid stabilizes to within ± 2 dB of the lower steady state level. (Vonlanthen et al., 2007)

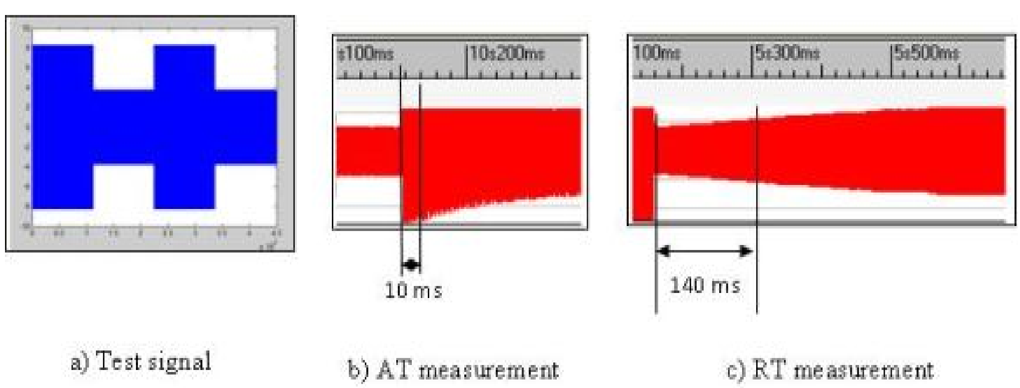

Figure 2: *Illustration of measurement of AT and RT*

Reliability measurements: In order to ensure the consistency and reliability in VOT measurements, intra-judge reliability was assessed. For this, 30% of the measured values were selected (on random basis) and reevaluated by the same investigator. Absolute percent errors were used to indicate the intra-judge reliability of measurement.

Results showed an absolute percent error of 3.92% between the measurements done by the primary investigator, which indicate that the primary investigator was consistent and reliable in AT/RT measurements.

b) Experiment 2 - VOT Measurement

Stimuli: Six CV syllables, including /ka/, /pa/, /ta/, /ga/, /da/ and /ba/ were used for the study. The stimuli comprised of six stop consonants in the initial position categorized as:- unvoiced – velar /k/, dental /t/, and bilabial /p/, voiced – velar /g/, dental /d/ and bilabial /b/

Recording: The utterance of each stimulus produced by an adult (native Kannada) male speaker at a comfortable loudness level was recorded. Speech samples were acquired by a condenser microphone (B & K 4189) that was kept at a distance of 6" from the mouth of the speaker at an angle of 30^0 azimuth. Speech samples were immediately digitized at a sampling frequency of 22 kHz and 16 bits sample by using a sound recording set up (BZ7226, Bruel & Kjaer, Denmark) facilitated in a precision Type I sound level meter (B & K 2250). Wave files were then transferred on to a PC for delivery and analysis.

Instrumentation: The setup used for measurement and analysis was the same as in experiment1 (figure 1). Recorded stimuli, stored on to the computer as a wave file is downloaded to the front end of B & K pulse analyzer system. These hearing aids had the option to select two AT/RT settings:- slow and fast. All hearing aids were programmed with a flat gain curve for mild hearing loss. The output for each of the stimuli were acquired from all 4 hearing aids, recorded by BZ 7226 recording setup and were stored on to the PC for analysis.

Acoustic Analyses: VOT was measured from the digitized hearing aid processed speech samples (Figure 3) for both slow and fast AT/RT setting, by using a speech analysis software (B & K Pulse Time Edit & Analysis). The difference in time between the release of a complete articulatory constriction and the onset of quasi-periodical vocal fold vibration was noted as VOT. VOT values were directly measured from the waveform with the help of two cursors, the first one indicating the release of the burst, and the second one at the first distinguishable period of onset of the vowel (Lisker & Abramson, 1964). VOT (in milliseconds) was noted as the time interval between the two cursors as illustrated in figure 3. Spectral analysis using Pulse Analyzer (B & K) was done to locate these points whenever the burst or vowel onset was difficult to be located from the waveform.

Statistical analysis: Statistical analysis was carried out using commercially available SPSS (Version 18) software. Wilcoxon signed ranks test was administered to find the effect of slow and fast AT/RT on transformation of VOT.

Results and Discussion

Measurement of AT/RT: The results indicated that the hearing aids used for the study were having their AT ≤ 10 ms and RT ≤ 150 ms (Shi et al., 2008) in fast setting. In slow setting all the hearing aids showed an AT ≥350 ms and RT ≥1080 ms (Stone et al., 2007). As suggested by Gatehouse et al., (2006), the measurement was done in such a way that the validity of the test results were ensured. Thus these results have consolidated that all the hearing aids used for the study were having their AT & RT in slow and fast category at the programmed settings and the test conditions. The results also rule out the possibility of change in AT/RT due to the type of signal processing strategy or compression algorithm. Table 1 shows the AT and RT of all the four hearing aids used for the study, in both slow and fast setting.

Table 1: *Measured AT / RT values of hearing aids used for study*.

Hearing aids	Slow		Fast	
	AT (ms)	RT (ms)	AT (ms)	RT (ms)
Hearing aid 1	380	1150	10	140
Hearing aid 2	500	1200	10	100
Hearing aid 3	480	1120	9	105
Hearing aid 4	420	1180	10	90

Transformation of VOT in slow and fast AT/RT settings: VOT reduced by 7.33 ms and by 6.58 ms in unvoiced and voiced plosives, respectively under fast AT/RT settings. The reduction was longer for unvoiced plosives compared to voiced plosives. Further VOT reduced to a greater effect for velar unvoiced plosives compared to other places of articulation. It was interesting to note that VOT reduction was least in velar voiced plosive. Hearing aid 3 was better than the other hearing aids. VOT reduced by 6.50 ms and 0.83 ms for unvoiced and voiced plosives respectively in slow AT/RT setting. Also, VOT reduced to a greater extent for velar unvoiced plosive compared to others. Further, VOT transformation was negligible for voiced plosives. Performance of Hearing aid 2 was better compared to other hearing aids. Table 2 shows the reduction in VOT in fast and slow AT/RT settings and table 3 shows the mean, media and SD of reduction in VOT.

Table 2*: Reduction in VOT (ms) for plosives in fast and slow AT/RT setting*

Stimuli /Hearing aid		1	2	3	4
		Fast AT / RT setting			
		Change in VOT (ms)			
Unvoiced	/ka/	15	15	12	15
	/ta/	3	3	2	5
	/pa/	5	5	3	5
Voiced	/ga/	5	5	5	4
	/da/	10	5	10	5
	/ba/	10	10	5	5
		Slow AT / RT setting			
Unvoiced	/ka/	10	8	12	15
	/ta/	3	3	3	5
	/pa/	5	5	4	5
Voiced	/ga/	0	2	0	2
	/da/	0	0	4	2
	/ba/	0	0	0	0

Table 3: *Mean, Median and SD of reduction in VOT (ms) for fast and slow AT/RT setting*

	Fast AT / RT setting		Slow AT / RT setting	
	Unvoiced	Voiced	Unvoiced	Voiced
Mean	7.33	6.58	6.50	0.83
Median	5.00	5.00	5.00	5.00
SD	5.26	2.53	3.91	1.33

The results revealed a high SD in both conditions. Wilcoxon Signed Ranks Test was administered to ascertain the effect of slow/ fast RT on transformation in VOT. Results revealed no significant difference between conditions (fast and slow) for unvoiced plosives {$|z|=0.736$, $p>0.05$} and significant difference between conditions for voiced plosives {$|z| = 3.089$, $p<0.01$}.

Improvement in speech intelligibility was reported from previous studies conducted by Moore et al., (1988), Moore, Glassberg & Stone (1991), Stone & Moore (2007) when ATs/RTs were slow. Results of the present study support these findings as slow AT/RT settings did not bring any transformation in VOT of voiced plosives, and hence account for the improved speech intelligibility. Nataraja et al., (2000) reported that transformation of VOT would bring changes in speech intelligibility and hence a reduction in VOT would result in reduced intelligibility. Fast AT/RT setting showed reduction in VOT for both voiced and unvoiced plosives and this accounts for reduced intelligibility. This also explains the reported finding of Stone et al., (2007) that there is higher envelop distortion of running speech with fast compression than with slow compression

Results of the present study also show no significant difference between conditions for unvoiced plosives. Possible reason for this must be the fact that in unvoiced plosives compression gets activated with the onset of voicing only and hence it may not affect the VOT. This also accounts for the results of the study conducted by Moore et al., (2004) where it is reported that there is no effect of time constant on speech intelligibility. These results also validate the results of the study conducted by Jerlvall and Lindblad (1978) that, no notable difference in speech discrimination could be observed for the various combinations of attack and release times for the S/N = 60dB condition, both for the normal subjects as well as for the persons with hearing impairment.

Another reason might be because of the difference in VOT in Kannada and English. English is a two-way language (with voiced and unvoiced plosives) and Kannada is a four-way language (with unaspirated voiced, murmered voiced, unaspirated unvoiced and aspirated unvoiced). Further an unvoiced plosive in the word-initial position in English is aspirated by rule. Hence VOT is longer for unvoiced plosive in English compared to Kannada.

Thus the difference in VOT transformation of voiced and unvoiced plosives in fast and slow AT/RT settings can be one reason for the conflicting results obtained in the previous studies conducted on the effect of compression time constants on speech perception. The conflicting results may be because of the difference in speech stimuli (voiced vs. unvoiced) used in these studies. Hence results of this study figure out VOT transformation as a potential temporal parameter responsible for distortion in hearing aids with compression.

Conclusions

The present study measured attack time and release time of the hearing aids used for the study and estimated the transformation of VOT in hearing aid processed speech for slow and fast compression time constant setting. Results indicated a reduction in VOT for unvoiced plosives in both slow and fast AT/RT setting. For voiced plosives, reduction in VOT was noticed in fast AT/RT setting whereas there was no reduction in slow AT/RT setting. This accounts for reported degradation in speech intelligibility in fast AT/RT settings as reported in the previous studies. The extent of degradation in VOT in fast AT/RT setting can be used as an indicator for specifying the ideal AT/RT setting in fast mode without producing any distortion in hearing aid processed speech. Thus VOT transformation can be used as a factor to find out what constitutes the best compression system in hearing aids.

Acknowledgements

The author thanks Director, AIISH, Mysore for permitting to carry out the study. The author also thanks Dr. M.S. Vasanthalakshmi, Lecturer in Biostatistics, Department of Speech Language Pathology, for the help in statistical analysis. Sincere thanks to G. Raghavendra, Software Engineer and M.S. Ravishankar, Junior Technical Officer, Department of Electronics for the immense support.

References

Gatehouse, Stuart., Naylor, Graham., Elberling, Class. (2006). Linear and nonlinear hearing aid fittings – 1. Patterns of benefit. *International Journal of Audiology, 4*5, 130-152.

Hansen, M. (2002). Effects of multi-channel compression time constants on subjectively perceived sound quality and speech intelligibility. *Ear & Hearing, 23(4)*, 369-380.

Jenstad, L. M., & Souza, P.E. (2005). Quantifying the effect of compression hearing aid release time on speech acoustics and intelligibility. *Journal of speech, language, and hearing research, 48*, 651 – 657.

Jerlvall, L.B., Lindblad, A. C. (1978). The influence of attack time and release time on speech intelligibility. A study of the effects of AGC on normal hearing and hearing impaired subjects. *Scandinavian audiology supplementum, 6*, 341-353.

Kent, R.D., Charles Read. (1995). *The Acoustic Analysis of Speech*. Delhi: AITBS Publishers and Distributors.

Lisker, L., Abramson, A. (1964). A Cross-language study of voicing in initial stops: Acoustical measurements, *Word 20*, 384-422.

Moore, B. C. J., Thomas H. Stainsby, Jose I. Alcantara, Volker Kuhnel. (2004). The effect on speech intelligibility of varying compression time constants in a digital hearing aid. *International Journal of Audiology 43*, 399-409.

Moore, B.C.J., Glasberg, B.R., Stone, M.A., (1991). Optimization of a slow-acting automatic gain control system for use in hearing aids. *British Journal of Audiology, 25*, 171–182.

Moore, B.C.J., & Glasberg, B.R. (1988). A comparison of 4 methods of implementing AGC in hearing aids. *British Journal of Audiology 43*, 399-409.

Nabelek, I.V., & Robinette, L.N. (1977). A comparison of hearing aids with amplitude compression. *Audiology, 16(1)*, 73-76.

Nataraja, N.P., Savithri, S.R., Sreedevi, N., and Sangeetha, M. (2000). *Transformation of the speech of the Hearing Impaired*, Project at AIISH, Mysore funded by Department of Science and Technology, New Delhi.

Neuman, A.C., Bakke, M.H., Mackersie, C., Hellman, S., & Levitt, H. (1998). The effect of compression ratio and release time on the categorical rating of sound quality. *Journal of the Acoustical Society of America, 103(5)*, 2273-2281.

Ole Hau, & Hanne Pernille Andersen. (2012). Hearing aid compression:- Effects of speech, ratio and channel bandwidth on perceived sound quality. *http://www.audiologyonline.com*

Souza, Pameela E. (2002) Effects of Compression on Speech Acoustics, intelligibility, and Sound Quality. *Trends in Amplification*, *6*, 131-164.

Shi, Lu-Feng., Karen A. Doherty. (2008). Subjective and objective effects of fast and slow compression on the perception of reverberant speech in listeners with hearing loss. *Journal of speech, language, and hearing research, 51*, 1328 – 1340.

Shobha, N.H., T.G. Thomas & K. Subbarao. (2009). Effect of consonants duration modifications on speech perception in noise. *International Journal of Electronics Engineering, 1(2)*, 179-184.

Stone, M.A. & Moore, B.C.J. (2007). Quantifying the effects of fast-acting compression that affect intelligibility in a competing speech task. *The Journal of the Acoustical Society of America, 121*, 1654-1664.

Stone, M.A., Moore, B.C.J. Alcantara, J.L.,& Glasberg, B.R. (1999). Comparison of different forms of compression using wearable digital hearing aids. *The Journal of the Acoustical Society of America*, *106*, 3603-3619.

Van Tasell, D.J. & Trine, T.D. (1996). Effects of single-band syllabic amplitude compression on temporal speech information in nonsense syllables and in sentences. *Journal of Speech and Hearing Research*, *39*, 912-922.

Vonlanthen, Andi., Arndt, Horst. (2007). *Hearing Instrument Technology for the Hearing Healthcare Professional – 3rd Edn.* USA: Thomson Delmar Learning.