McMaster University

CORRECTING FOR IMPAIRED AUDITORY NERVE FIBER GROUP DELAYS COULD IMPROVE SOME FEATURES OF THE SPATIOTEMPORAL DISCHARGE PATTERN

Timothy J. Zeyl and Ian C. Bruce

Department of Electrical & Computer Engineering, McMaster University, Hamilton, ON, Canada

zeyltj@mcmaster.ca and ibruce@ieee.org

AUDITORY ENGINEERING LABORATORY

Abstract

Auditory nerve fibers in an ear with outer hair cell damage can be conceptualized as filters having a broadened frequency response area, a shallower phase response and a shorter group delay with respect to a healthy fiber, particularly at low stimulus presentation levels. As well, the presence of inner hair cell damage requires increased stimulus presentation levels for restoration of fiber discharge rates, which results in broad auditory filters with shallow phase response and short group delay. As a consequence, the discharge times in the impaired ear in response to a tone stimulus are more coincident across a population of fibers with a range of characteristic frequencies. This behavior resembles the spatiotemporal response pattern in a healthy auditory periphery in response to loud stimuli and has been postulated as a potential correlate to loudness recruitment. The present study evaluates the potential for correction of the altered phase response in the neural firing pattern of the impaired ear by a hearing aid. We implement a version of the spatiotemporal pattern correction scheme presented by Shi et al. [1], which measures the instantaneous difference in group delay between a bank of model healthy and impaired auditory nerve fibers and inserts the corresponding delays into an analysis-synthesis gammatone filterbank in the hearing aid. Human testing of the processing scheme showed that listeners preferred unprocessed sounds over processed sounds and that no systematic improvement in speech intelligibility was provided by the processed speech [1, 2].

We evaluate this processing scheme with a computational model of the auditory periphery (Zilany & Bruce, [3, 4]) in response to a synthesized vowel for a mild and a moderate-to-severe high-frequency sloping hearing loss, both with mixed hair cell damage. Analysis indicates that there are some technical and conceptual problems associated with the processing scheme that need to be addressed. These include: i) a possible non-flat frequency response through the analysis-synthesis filterbank due to time-varying changes in the relative temporal alignment of filterbank channels, ii) group delay corrections that are based on potentially incorrect frequencies due to the spread of synchrony in auditory nerve responses, and iii) modulations of frequency in the processed signal created by the insertion of delays resulting in the presence of abnormal frequencies of auditory nerve synchronization. Despite these issues, evaluation with an error metric derived from auditory nerve response cross-correlations shows that this processing scheme has the potential to improve performance at some sound pressure levels if the technical limitations are addressed sufficiently.

I. INTRODUCTION

- Individuals with sensorineural hearing loss exhibit a decreased dynamic range of sound levels at which sounds are both audible and tolerable. This results in a rapid growth of the perception of loudness termed 'loudness recruitment.' [5]
- The mechanism of loudness recruitment is not yet known. Carney [6], suggests that a loss of compressive non-linearity in the basilar membrane, which results in changes in the temporal response pattern in the auditory nerve, may lead to a steeper rate-level curve at higher levels of auditory processing.
- Spatiotemporal Pattern Correction (SPC) is a processing scheme that has been proposed by Shi et. al. [1] to compensate for erroneous group delays in impaired auditory filters. Correcting group delays could lead to more normal temporal response patterns in the auditory nerve.
- SPC did not perform well in human testing and may be based on incorrect frequencies because the impaired ear exhibits an upward spread of synchrony to low frequencies.

II. METHODS

A. Auditory Models

- The auditory-periphery model used in this study (Fig. 1) was that of Zilany and Bruce [3,4]. This phenomenological model describes the cat auditory pathway from the middle ear through to the auditory nerve.

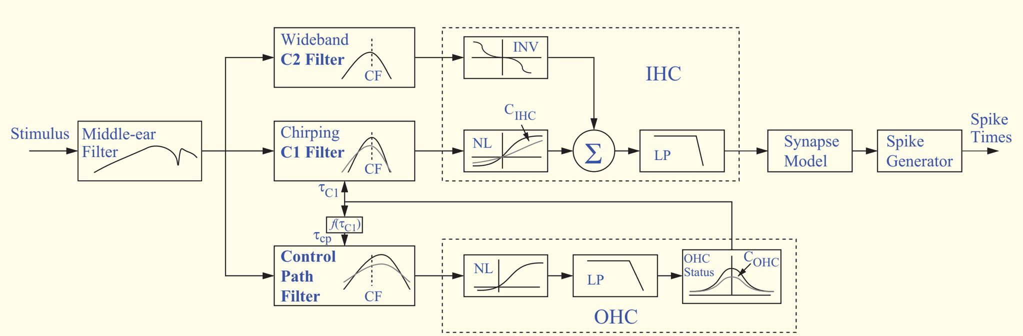

Figure 1: *Zilany and Bruce cat auditory nerve model [3, 4].*

B. SPC - Spatiotemporal Pattern Correction

- SPC measures the difference in group delay through a bank of model healthy and impaired auditory filters and inserts this difference into the signal as a temporal delay via an analysis-synthesis filterbank.

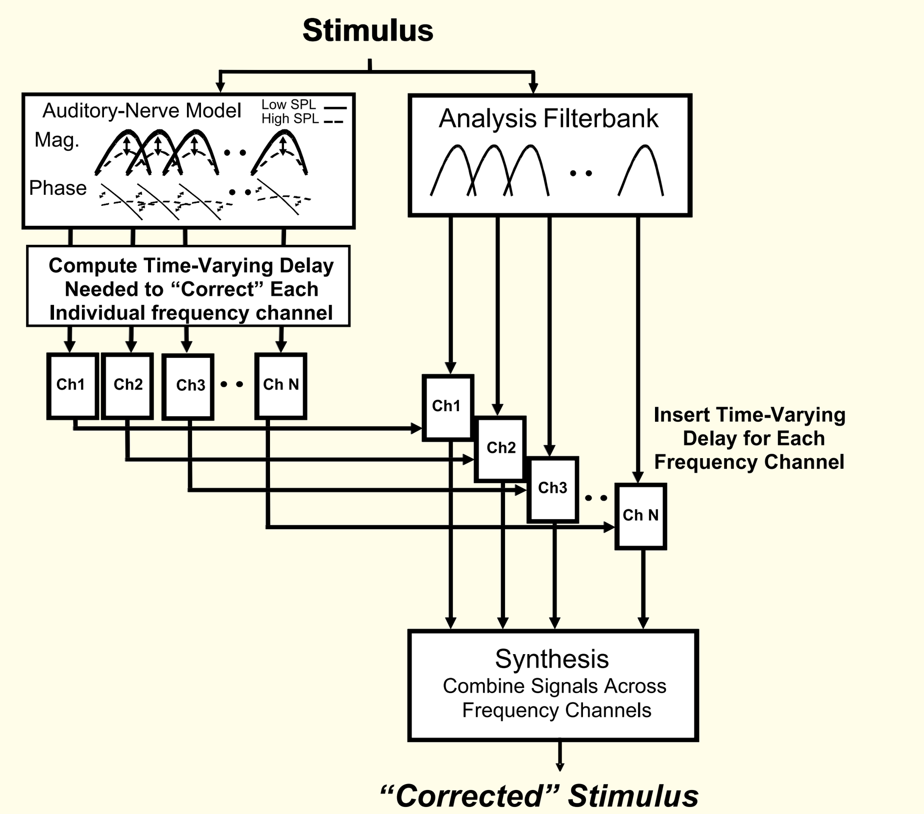

Figure 2: *Block diagram of the SPC processing scheme. (Reprinted from Figure 2 of Shi et. al., 2006 [1])*

Updates to SPC

- SPC has been updated to use the Zilany and Bruce auditory model (Fig. 1) to prescribe insertion delays.
- The times at which the insertion delays prescribed by SPC are applied are themselves delayed to account for the difference in latencies between the modeling path and the processing path, as depicted for one channel in Fig. 3.

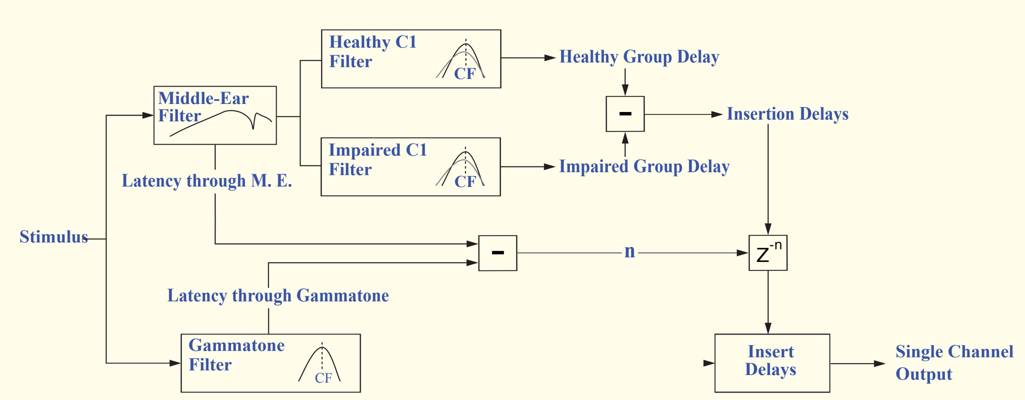

Figure 3: *Prescription and insertion of time-varying delays for one frequency channel in the SPC filterbank*

C. Hearing Loss Profiles

- Two hearing loss profiles are evaluated in this study, one mild hearing loss and one moderate-to-severe hearing loss, as shown in Fig. 4.
- The impairment in the auditory model is set so that 2/3 of the threshold shift is due to outer hair cell damage, and the remaining 1/3 is due to inner hair cell damage.

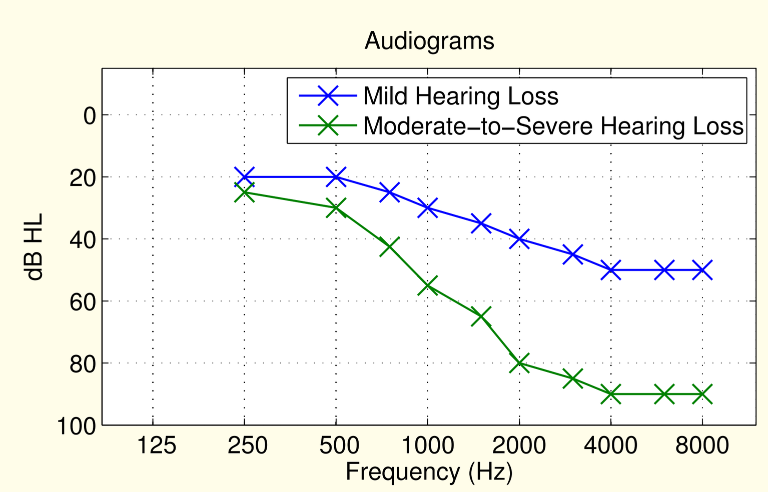

Figure 4: *Audiograms used in the evaluation of the SPC processing scheme*

D. Stimuli

- We evaluated the SPC scheme with the synthesized vowel /ɛ/, which has the spectral envelope shown in Fig. 5.

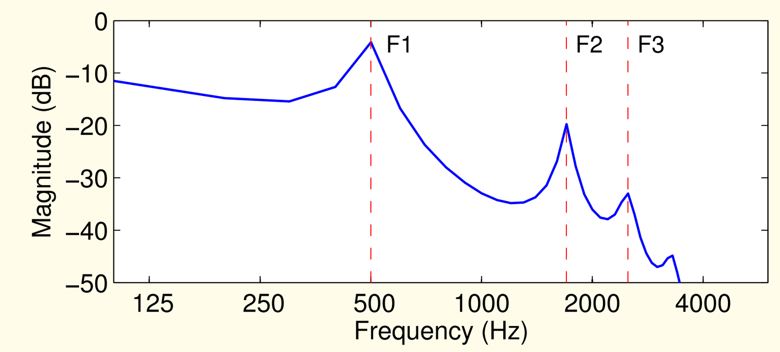

Figure 5: *Spectral envelope of the synthesized vowel /ɛ/. The vowel's first three formants are: F1 = 500 Hz, F2 = 1700 Hz, and F3 = 2500 Hz.*

E. Error Metric

- We developed a metric that compares the cross-correlation of the neural firing patterns of adjacent auditory nerve fibers for model healthy and impaired ears. This method is resistant to errors due to overall hearing-aid processing delays.
- Information about the relative phase of firing between fibers is captured by the lag at which the cross-correlation is maximum (or 'peak-lag'). This information can be used to analyze the abnormal relative AN phase response that has been observed with sensorineural impairment and associated with loudness recruitment.

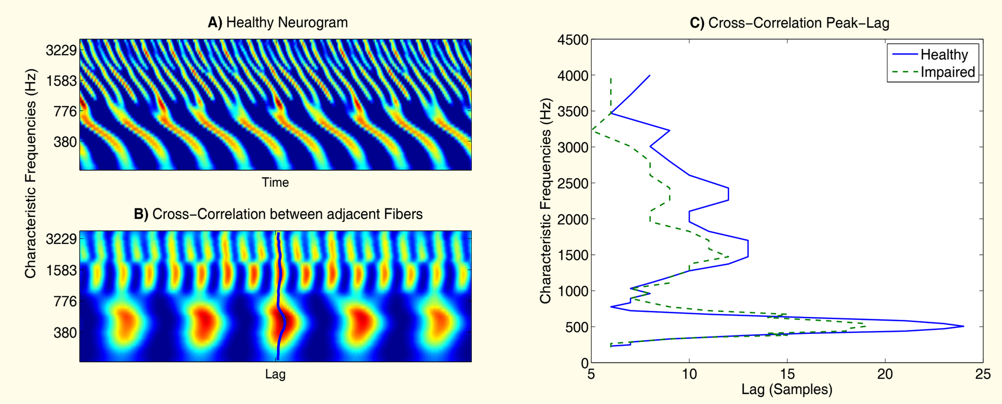

Figure 6: *Development of cross-correlation error metric. A) The spatiotemporal response pattern (or neurogram) of the healthy auditory model in response to a 3-tone stimulus. B) The cross-correlation between adjacent fibers in the neurogram. The maximum value of each cross-correlation is outlined by the superimposed blue line. The lag at which the cross-correlation is maximum is termed the 'peak-lag.' C) The peak-lag for each neighboring fiber pair is shown at their approximate CF for the healthy model and the mildly-impaired model in response to the stimulus presented at 60 dB SPL.*

III. RESULTS

A. Non-Flat Frequency Response

- Analysis showed that insertion of time-varying delays caused the gain-frequency response of the SPC filterbank to become non-flat and time variant.
- This is likely caused by destructive interference in a signal component that passes through neighboring analysis filters with different phase responses at the component frequency because of the time-varying delay.

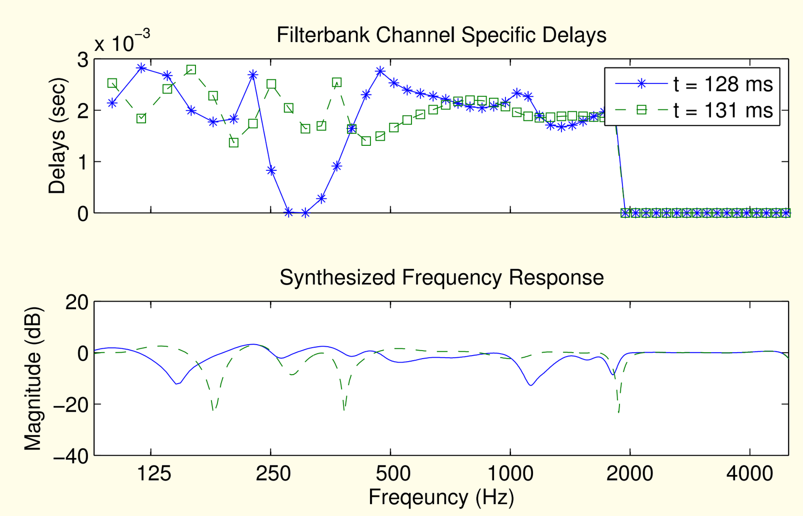

Figure 7: *Top: The delays prescribed by SPC for each channel in the filterbank at two time-points. Channel center frequencies are indicated with markers. Bottom: Gain-frequency response of the filterbank at the time-points given in the top panel.*

B. Frequency Modulations

- The insertion of a time-varying delay causes frequency modulations in each channel of the filterbank, as can be seen in the left side of Fig. 8.
- This results in high frequency noise that is easily identifiable in the time-domain waveform on the right side of Fig. 8. This noise is audible to listeners with normal hearing.

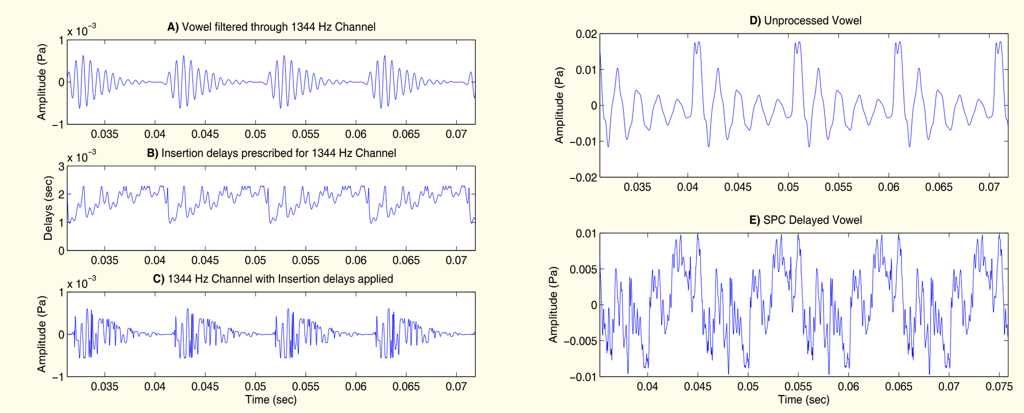

Figure 8: *Left: A) Vowel filtered through the 1344 Hz SPC filterbank channel. B) Time-varying insertion delays prescribed for this channel. C) Result of applying insertion delays to the filtered vowel. Right: D) Temporal waveform of the vowel /ɛ/. E) Temporal waveform of the SPC-processed vowel.*

C. Group Delay evaluated at potentially incorrect frequencies

- The group delay found in the healthy and impaired auditory filters used to prescribe the SPC insertion delays may be based on incorrect frequencies for two reasons:
  1. Synchrony in the spatiotemporal response of impaired auditory nerve fibers may be disrupted due to lack of synchronization to higher formants and an upward spread of synchrony to lower formants. This results in corresponding healthy and impaired auditory filters synchronizing to different frequencies.
  2. The auditory filter group delay is evaluated at its center frequency (which corresponds to the AN fiber's characteristic frequency) to prescribe the SPC insertion delays. However, the band-pass nature of auditory filters means that the response frequency is not always equal to the filter's center frequency.

D. Neurograms

- The effects of SPC on the spatiotemporal response pattern (a.k.a. neurogram) to the synthesized vowel are demonstrated in Fig. 9.
- The neurograms here depict the probability of spiking in the auditory nerve before refractoriness, with red indicating a high probability of spiking and blue indicating a low probability of spiking.
- SPC is shown to alter the unaided neurogram significantly. Synchrony is diminished, but still present, and additional frequency components arise in the response.

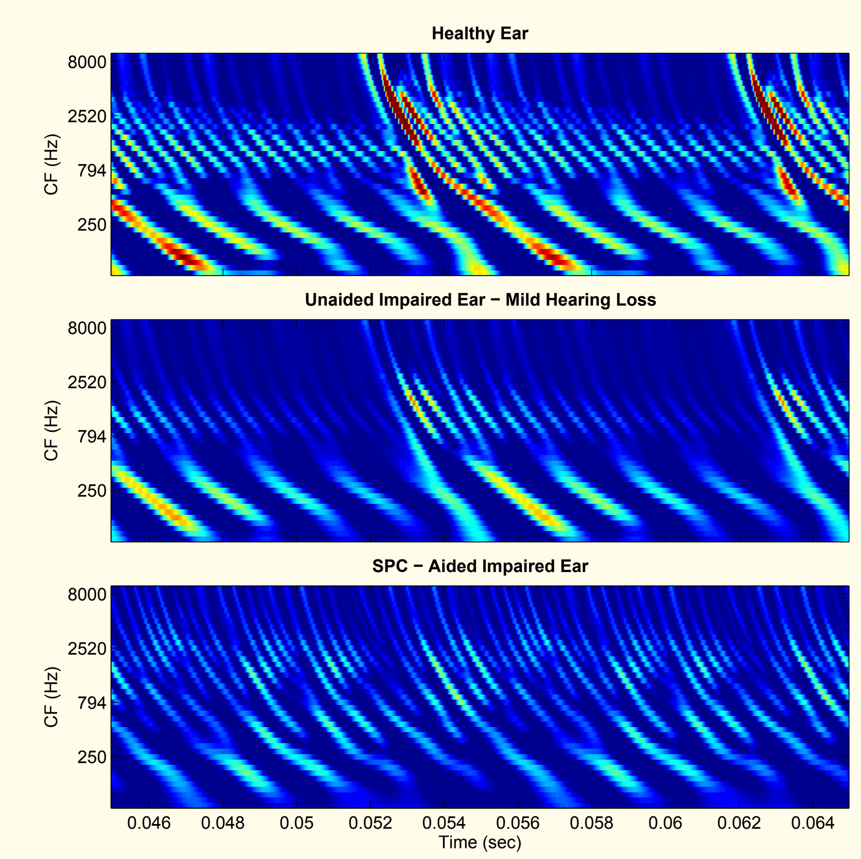

Figure 9: *Healthy, impaired and SPC-aided neurograms in response to the synthesized vowel presented at 60 dB SPL. The x-axis represents time and the y-axis represents individual auditory nerve fibers, arranged in order of characteristic frequency (CF). The colors represent the probability of action potentials in terms of spikes/sec, where blue is a low spiking rate and red is a high spiking rate.*

E. Peak-Lag Error

- Peak-lag error is taken as the difference between healthy and impaired peak-lags, each summed across all characteristic frequencies evaluated. Errors are normalized by the 2-norm of the impaired case.
- SPC decreases the peak-lag error at low SPLs for both hearing loss types evaluated in response to a synthesized vowel.
- To determine if prior restoration of formant synchrony could improve SPC performance, the algorithm was tested on the vowel pre-processed with the spectral-enhancement scheme proposed by Harte et al. [7], Multiband and Improved Contrast-Enhanced Frequency Shaping (MICEFS). This improved the peak-lag error for the moderate-to-severe loss only.

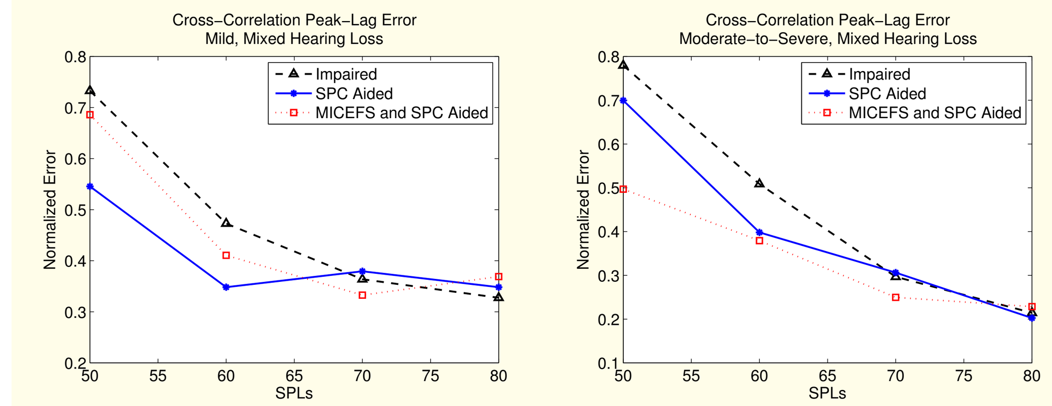

Figure 10: *Peak-lag error shown for the impaired and SPC-aided vowel response as a function of stimulus level. The left panel shows the mild hearing loss and the right panel shows the moderate-to-severe loss.*

IV. CONCLUSIONS

- Several problems with the SPC scheme were found including a fluctuating gain-frequency response, frequency modulations present in the processed signal, and group delay calculations based on potentially incorrect frequencies. These limitations likely contribute to the disrupted synchrony that is present in the processed spatiotemporal response.
- Despite disrupted synchrony, SPC processing led to reduced peak-lag error in the spatiotemporal response at low SPLs. This could indicate that the phase relationships between some AN fibers in the impaired ear are improved.
- Restoration of formant synchrony before processing with SPC may reduce the error caused by group delay evaluation at incorrect frequencies. This has been tested using MICEFS [7] to restore formant synchrony and results in improved peak-lag error for the moderate-to-severe hearing loss.
- In addition to these findings, we attempted to reduce high frequency noise present in the SPC filtered signal by low-pass filtering the insertion delays. This led to greater disruption of synchrony and poorer peak-lag error restoration. This suggests that noise produced by frequency modulations may be an unavoidable side-effect of reducing peak-lag error. Furthermore, this high frequency noise may not be audible to listeners with hearing impairment.

V. ACKNOWLEDGMENTS

This work was supported by NSERC Discovery Grant 261736 (Bruce) and an NSERC Canadian Graduate Scholarship (Zeyl).

References

[1] L.-F. Shi, L. H. Carney, and K. A. Doherty, "Correction of the peripheral spatiotemporal response pattern: a potential new signal-processing strategy," *J Speech Lang Hear Res*, vol. 49, no. 4, pp. 848–55, Aug. 2006.

[2] L. Calandruccio, K. A. Doherty, L. H. Carney, and H. N. Kikkeri, "Perception of temporally processed speech by listeners with hearing impairment," *Ear Hear*, vol. 28, no. 4, pp. 512–523, Aug. 2007.

[3] M. S. A. Zilany and I. C. Bruce, "Modeling auditory-nerve responses for high sound pressure levels in the normal and impaired auditory periphery," *J Acoust Soc Am*, vol. 120, no. 3, pp. 1446–1466, Sep. 2006.

[4] ——, "Representation of the vowel (epsilon) in normal and impaired auditory nerve fibers: Model predictions of responses in cats," *J Acoust Soc Am*, vol. 122, no. 1, pp. 402–417, Jul. 2007.

[5] B. C. J. Moore, *An Introduction to the Psychology of Hearing*, 5th ed. Academic Press, 2003.

[6] L. H. Carney, "Spatiotemporal encoding of sound level: models for normal encoding and recruitment of loudness," *Hear Res*, vol. 76, no. 1-2, pp. 31–44, Jun. 1994.

[7] N. Harte, S. U. Ansari, and I. Bruce, "Exploiting voicing cues for contrast enhanced frequency shaping of speech for impaired listeners," in *Proceedings of the 31st IEEE International Conference on Acoustics, Speech, and Signal Processing (ICASSP2006)*, vol. 5, Piscataway, NJ, 2006, pp. V–137–V–140.