International Journal of Innovative Research in Advanced Engineering (IJIRAE) **ISSN: 2349-2163**
Volume 7, Issue 03 (March 2020) **www.ijirae.com**
Special Issue – 6th International National Level Conference- “MEEMIC-2020”

PERCEPTUAL AND OBJECTIVE ASSESSMENT OF ENVELOPE ENHANCEMENT FOR CHILDREN WITH AUDITORY PROCESSING DISORDER

Prof.V. Loganathan.,[1], George Max Berlise.S[2], Manoj.G[3], Jegan.D[4]
Assistant Professor[1], UG Scholar, [2],[3],[4]
Department of Medical Electronics Engineering,
Sengunthar College of Engineering, Tiruchengode, Tamil Nadu, INDIA
[1]loguvaithi@gmail.com, [2],maxgeorge578@gmail.com,[3]manothrajammal999@gmail.com,
[4]jegan0367@gmail.com

Manuscript History
Number: **IJIRAE/RS/Vol.07/Issue03/Special Issue/01.MRAESCE10090**
Received: 15, February 2020
Final Correction: 27, February 2020
Final Accepted: 10, March 2020
Published: **14, March 2020**
Editor: Dr.A.Arul Lawrence selvakumar, Chief Editor, IJIRAE, AM Publications, India

Abstract— This paper evaluated the performance of an envelope enhancement (EE) algorithm subjectively by children with auditory processing disorder (APD), and objectively through computational models. Speech intelligibility data was collected from children with APD, for unprocessed and envelope- enhanced speech in the presence of stationary and non-stationary background noise at different signal to noise ratios (SNRs), both with and without noise reduction (NR) algorithms as a front-end to the EE algorithm. Furthermore, intrusive and non-intrusive objective speech intelligibility metrics were derived to predict the perceptual impact of this EE algorithm. Subjective data for stationary noise conditions revealed that the combination of NR and EE algorithms significantly improved the speech intelligibility scores at poor SNRs. In contrast, the same combination was ineffective in improving speech intelligibility in non-stationary noise conditions. Taken together, subjective results suggest that exaggerating the envelope cues improves speech identification scores for children with APD. However, the benefit obtained varies depending upon the type and level of the background noise. Both intrusive and non-intrusive objective speech intelligibility estimators exhibited good correlation with the subjective data, with the intrusive metric demonstrating better generalization capabilities. Implications of these results for hearing aid applications for children with APD is discussed.

Keywords— Envelope enhancement, Auditory processing disorder, Speech Intelligibility, Hearing Aid Speech Perception Index, Temporal processing.

INTRODUCTION

Some children who have a normal pure-tone audiogram experience speech perception difficulties in the presence of background noise, especially for poorer Signal to Noise Ratio (SNR) conditions. It is estimated that about 3-5% of children suffer from auditory processing disorder (APD) [1]. APD can be considered as a distinct hearing disorder due to the deficits in temporal, spectral, or binaural processing [2]. Irrespective of the aetiology of the APD, children suspected with APD experience listening difficulties which directly impact their ability to learn and communicate in a classroom, especially in the presence of background noise. Conventional assistive hearing technology provides insufficient speech perception improvement for children with APD. As an example, Kuk [3] reported hearing aid benefit data from fourteen children diagnosed with APD, which showed that only some participants demonstrated improved performance with hearing aids.

Researchers have also investigated the potential effectiveness of remote microphone (RM) assistive listening devices for APD wherein the teacher in a classroom wears a wireless microphone, and the speech signal picked up by the microphone is transmitted to the child's hearing aid receiver (e.g.[4]and[5]). A recent systematic review by Reynolds et al. [6] reported consistently positive results with RM systems but offered moderate support for the use of RM systems for children with APD. As such, better performing signal processing algorithms within both RM and fully wearable devices are desirable to better serve this clinical population. This paper focuses on such algorithms when implemented in a fully wearable device.

Speech envelope enhancement (EE) algorithms, which can enhance the temporal cues by exaggerating the speech modulation pattern, are attractive candidates for this purpose, and evidence exists that such techniques can be effective in improving the speech perception by children with APD. The existing speech EE algorithms can be categorized into two groups: dynamic and static. In the dynamic EE algorithm, the enhancement is based on both the instantaneous and minimum amplitudes of the input speech envelope [7]. Research by Narne and colleagues ([7], [8], and [9]) demonstrated the effectiveness of the dynamic EE for subjects afflicted with the Auditory Neuropathy Spectrum Disorder (ANSD) in terms of improved speech perception in the presence of background noise. In a similar vein, Moshgelani et al. [10] showed that the dynamic EE improves speech perception by APD children, as long as the speech envelope is enhanced prior to contamination by background noise.

In contrast to dynamic EE, the static EE applies a fixed modulation boost over a restricted range of acoustic frequencies. Shetty and Kooknoor ([11] and [12]) have demonstrated the benefits of the static envelope enhancement (SEE) algorithm on phrase and consonant identification by individuals with normal hearing, sensorineural hearing loss, and ANSD. While these results are promising, there currently is a lack of knowledge on the effectiveness of SEE algorithms for children with APD, especially when applied to longer sentences rather than short segments of speech (i.e., consonants, vowels).

Additionally, the impact of background noise must be carefully considered in assessing the EE algorithm effectiveness. Previous studies by Narne and colleagues ([7], [8], and [9]) and Shetty and Kooknoor ([11] and [12]) added background noise only after the speech envelope was enhanced. However, in a typical hearing aid application, the background noise is mixed in with the desired speech signal prior to enhancement, which may impact the EE algorithm and limit its perceptual benefit. Indeed, our recent studies benchmarking EE algorithm performance using instrumental measures have revealed a degraded performance with both dynamic and static EE algorithms when background noise was added prior to envelope enhancement, with the static EE more robust than the dynamic EE ([13] and [10]). Furthermore, we have demonstrated the utility of incorporating a noise reduction algorithm prior to EE in some background noise conditions (*viz.* stationary noise and SNR < -3 dB) ([13] and [10]). However, these results have not been validated with subjective data from children with APD.

While subjective speech perception data is invaluable and can be considered as gold standard, it is also time- and resource- consuming. It is economical and efficient to employ objective instrumental measures ([14], [15], [16], and [17]) that incorporate computational models to predict the subjective speech perception data, for further development and benchmarking the EE algorithms across different noisy conditions. A good, validated objective metric that correlates well with subjective data can reduce the need for further subjective assessment of the EE algorithms.

The aim of this study, therefore, is to investigate the performance of the static EE algorithm in enhancing speech perception by children with APD, in different types and levels of background noise, and to develop objective indices predictive of the children's speech intelligibility performance. In particular, the answers to the following research questions were sought: (a) How effective is the static EE algorithm for children with APD in hearing aid applications? (b) What is the benefit of applying different noise reduction algorithms as a front- end to the static EE? and (c) Can objective speech intelligibility estimators be developed to predict the perceptual impact of EE algorithms?

METHOD

A. Static EE algorithm

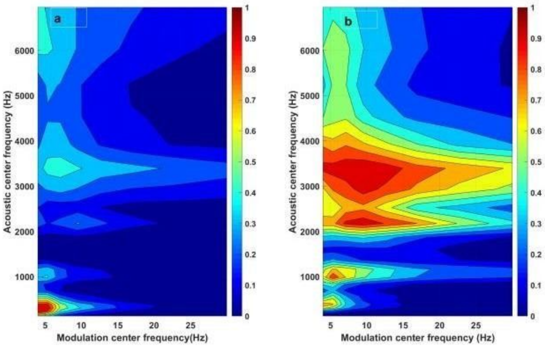
The static EE algorithm investigated in this paper was based on the description given in Praat software (Phonetic Sciences, University of Amsterdam, Netherlands), and implemented in MATLAB 2017a. The discrete-time input speech signal, $x(n)$, is first transformed to the spectral domain by applying the fast Fourier transform (FFT). The signal spectrum between 100- 8000 Hz is segmented into 22 critical bands, with each band spanning one Bark. The transformation between the frequency, f, in Hz and the Bark scale (b) is shown in Equation 1:

$$b = 7 * \sinh{-1} (f) / 650 \qquad (1)$$

The intensity envelope extracted from each of the Bark filter bank outputs is subsequently filtered by a band pass filter, whose frequency response, *H(f)*, is given in Equation 2:

$$(f) = e{-}(\alpha \mathrm{f} \ )2) \ \mathrm{ffact} \ - e(-(\alpha \mathrm{f} \ \mathrm{fclow}) \quad (2)$$

where $\alpha = \sqrt{ln 2}$, and fslow = 3 Hz, and ffast = 30 Hz represent the modulation frequency range for enhancement. The filtered envelope in each band is scaled by the enhancement factor, which is computed as: Fig. 1 allows for the visualization of the impact of the static EE algorithm. The plots show the distribution of modulation energy as a function of the modulation frequency and acoustic frequency. It can be observed from Fig. 1.b that the static EE boosts the modulation energy significantly from 4-10 Hz, especially for acoustic channels around 1000 Hz. It is well-known that slow temporal envelope modulations from 4-10 Hz provide useful cues for speech perception [14]. It is evident from Fig. 1.b that the static EE algorithm exaggerates these useful cues.



(a) original speech, (b) static EE.

Fig. 1. Modulation spectrograms for the original and enhanced speech stimulus.

Noise reduction (NR) algorithms

This study investigated the performance of two different NR algorithms as a front-end to static EE: (a) the log Minimum Mean Square Error (log MMSE) NR algorithm [18], and (b) the single microphone NR algorithm from the open source Master Hearing Aid (MHA) framework [19], labelled as the MHA NR algorithm. The log MMSE algorithm belongs to the statistical- model-based noise reduction techniques and is an example of algorithms based on maximum likelihood estimation. The MHA NR algorithm is an example of Wiener filtering algorithms that work in the short-time Fourier transform (STFT) domain. The main difference between statistical- model-based methods and Wiener filter models is that in the latter, the goal is to estimate the complex spectrum of the speech while in the former, the focus is on finding an estimation for the magnitude of speech spectrum [18].

Musical noise is typically associated with NR algorithms, and smoothing the maximum likelihood estimate of the SNR is vital for the suppression of musical noise [20]. The log MMSE NR algorithm was chosen as it generates less musical noise [18] and was also shown to be judged better in subjective experiments [21] than other NR algorithms based on subspace processing and spectral subtraction. Similarly, the MHA NR algorithm applies temporal smoothing in the cepstral domain for the maximum likelihood estimate of the SNR. In contrast to temporal smoothing in the frequency domain, temporal smoothing in the cepstral domain allows smoothing only for coefficients that correspond to noise. As a result, less musical noise is generated compared to the other techniques that use the well-known decision-directed approach for a priori SNR estimation [20].

Objective speech intelligibility metrics Several objective indices have been proposed, examined, and reviewed for assistive hearing device applications [22]. However, there is little evidence on the validity of these metrics with subjective speech intelligibility data collected from APD subjects. In the current research study, two objective speech intelligibility metrics, the Hearing Aid Speech Perception Index (HASPI) and the Modulation Spectrum Area (ModA) were investigated. The HASPI was chosen due to the fact that it showed an excellent correlation with subjective scores collected from individuals with sensorineural hearing loss ([16] and [22]). Mod A was chosen since it highly correlates with subjective scores collected from cochlear implant users and individuals with sensorineural hearing loss ([17] and [22]).

It is pertinent to note that HASPI is an example of an *intrusive* objective metric as it estimates speech intelligibility by comparing the processed speech with the corresponding clean speech, while Mod A is a *non-intrusive* metric as it estimates speech intelligibility solely based on the processed speech. HASPI [16] predicts speech perception by utilizing an auditory model. The middle ear transfer function, auditory filterbank, outer hair cell dynamic range compression, two tone suppression, and the temporal firing rate activity of inner hair cells are the building blocks of HASPI auditory model [16]. The auditory model compares the clean (reference speech) and the processed (degraded speech) inputs in terms of their temporal envelope and fine structure, and computes the envelope cepstral correlation (*C*), and three-level (*a* , *a* , *and a*) fine structure, and the American Academy of Audiology [26] that included the administration of a thorough case history and listening surveys or questionnaires; the behavioural evaluation of auditory abilities (temporal resolution, auditory discrimination, auditory pattern recognition, performance with degraded acoustic signals, performance with competing signals) and the objective evaluation of the integrity of the auditory nervous system through electrophysiologic measures. According to ASHA best practice guidelines [24] and [25], performance on a minimum of two measures falling at least two standard deviations below age *low Mid High* mean was required to receive an identification of covariance between the reference and processed speech signal. Note that the *alow, aMid, and aHigh* values are computed for 16 ms speech segments whose intensities fall in the lowest third, middle third, and upper third of the cumulative speech intensity histogram [16].

Subjective data collection

Subjective data were collected in two separate experiments. In experiment #1, the static EE algorithm was evaluated subjectively in the presence of SSN by two different groups of participants: ten children with APD in the age range of 8.1 to 13.5 years (mean: 10.6 years), and ten typically developing children in the age range of 8.4 to 17.4 years (mean: 11.8 years). In the second experiment, the performance of the static EE in the presence of MTBN was evaluated subjectively by a new APD group of ten participants in the age range of 7.9 to 15 years (mean: 9.8 years), who were distinct from the APD group that participated in the first experiment. Children suspected of APD were referred to H.A. Leeper Speech and Hearing clinic at Western University because their parents or teachers expressed concerns about their listening abilities. Pure tone hearing thresholds and middle ear function were assessed according to the American Speech and Hearing Association (ASHA) best practice guidelines [24]. All participants demonstrated hearing thresholds falling within the normal range (< 20 dB HL) at octave frequencies from 250-8000 Hz in both ears. Middle ear function also fell within the normal range in both ears for all participants. The auditory processing assessment battery was designed to meet ASHA best practice guidelines [24] and [25] auditory processing disorder. All children who participated as part of the clinical population in this study met the criteria for identification as having auditory processing disorder (see [27] for further information on the assessment of children for auditory processing disorder). It should be noted that typically developing children had no developmental, academic, or listening concerns. The subjective data collection protocol was approved by the Western University Health Sciences Research Ethics Board. After an explanation of the research study and its protocol, both parents/guardians and child participants signed an informed consent form prior to participation.

RESULTS

Subjective results – Experiment #1

The averaged speech intelligibility scores along with their 95% confidence intervals of the mean for children with APD and normal hearing children are illustrated in Figs. 3 and 4 respectively. Note that the speech intelligibility score was normalized to have a maximum value of 1 (representing 100%- word recognition The "UP", "SEE", "logMMSESEE", and "MHASEE" conditions represent the (1) the unprocessed, (2) static EE by itself, (3) combination of logMMSE NR and static EE, and (4) combination of MHA NR and static EE processing conditions at different SNRs respectively.

CONCLUSIONS

Subjective and objective experiments were conducted to evaluate the performance of an envelope enhancement algorithm, with and without a noise reduction front-end, for hearing aid applications. Subjective speech intelligibility scores were obtained from children with auditory processing disorders using unprocessed and enhanced speech stimuli in the presence of stationary and nonstationary background noise at different SNRs. Computational objective speech intelligibility metrics were derived from the same stimuli used for subjective experiments and correlated with behavioural data. The salient results from this study include:

(a) The static EE by itself or in combination with NR algorithms is not beneficial in improving the speech intelligibility regardless of the subjective group and processing conditions,
(b) The incorporation of the MHA NR algorithm as a front-end to the static SEE is beneficial when the background noise is stationary and when the SNR is poor, while the application of the logMMSE NR algorithm prior to the static EE is not beneficial regardless of the noise type and processing condition

(c) A modified version of the hearing aid speech perception index (MHASPI) exhibited good correlation with subjective data and displayed generalization capabilities. Results from this paper can potentially guide the choice and activation of the static EE in assistive hearing device applications, and in the development and benchmarking of new envelope enhancement algorithms.

REFERENCES

1. J. R. Lucker, “The History of Auditory Processing Disorders in Children,” in Auditory Processing Disorders: Assessment, Management and Treatment, D. Geffner and D. R. Swain, Eds. 2013, pp. 33–58.
2. S. Flanagan, T. Zorila, S. Yannis, and B. C. J. Moor, “Speech Processing to Improve the Perception of Speech in Background Noise for Children with Auditory Processing Disorder and Typically Developing Peers,” Trends Hear., vol. 22, pp. 1–8, 2018.
3. F. Kuk, “Hearing Aids for Children with Auditory Processing Disorders?,” Semin. Hear., vol. 32, no. 2, pp. 189–195, 2011.
4. C. Schafer et al., “A Critical Review of Remote-Microphone Technology for Children with Normal Hearing and Auditory Differences,” J. Educ. Audiol., vol. 20, pp. 1–11, 2014.
5. G. Rance, “Auditory Neuropathy/Dys- synchrony and Its Perceptual Consequences,” Trends Hear., vol. 9, no. 1, pp. 1–43, 2005.S. Reynolds, H. M. Kuhaneck, and B. Pfeiffer, “Systematic Review of the Effectivness of Frequency Modulation Devices in Improving Academic Outcomes in Children with Auditory Processing
6. Raja, G P& Mangai, S 2018, ‘Investigation On Optimization, Prioritizing and Weight Allocation Techniques for Load Balancing and Controlling Multimedia Traffic in Wireless Mesh Network’, International Journal of Business Information Systems, SCOPUS Indexed Journal (Inderscience) - (P ISSN No: 1746-0972). Published Online: 10th Feb 2020, DOI: 10.1504/IJBIS.2020.105161.IF: 0.72.
7. Geetha. E & Nagarajan. C , 2018, ‘Induction Motor Fault Detection and classification Using Current Signature Analysis Technique’ Emerging Devices and Smart Systems, (IEEE Xplore) NSPEC Accession Number:18302693,DOI:10.1109/ICEDSS.2018.8544272, pp:48 – 52
8. Dr.R.Satish Kumar and Dr.K.Umadevi ” Finite Element Analysis of an Exterior Rotor Permanent Magnet Brushless DC Motor for Torque Improvement by a Novel Peak Torque Excitation Technique”, International journal of Innovative research in Advanced Engineering, Vol. 1, 2014, pp. 1-6 (Impact factor 1.311).
9. Dr.R.Satish Kumar and Dr.K.Umadevi ” A Novel peak torque Excitation Technique for Torque Improvement in Exterior Rotor Permanent Magnet Brushless DC Motor”, International journal of Innovative research in Advanced Engineering, Vol. 1, 2014, pp. 227-236 (Impact factor 1.311).
10. Dr.R.Satish Kumar and Dr.K.Umadevi ” Torque Improvement for an Exterior Rotor PermanentMagnet Brushless DC Motor”, International journal of Innovative research in Advanced Engineering, Vol. 1, 2014, pp. 1-5 (Impact factor 1.311).
11. Dr.R.Satish Kumar and Dr.K.Umadevi ” Novel Technique for Measurements of Dielectric Properties and Microwave Heating of In-Shell Eggs without Explosions in Microwave Oven for Pasteurization”, International journal of Innovative research in Advanced Engineering, Vol. 2, , 2015, pp. 69-77 (Impact factor 1.311).