Binaural Objective Intelligibility Measurement and Hearing Aids

Nicolas Ellaham[1], Christian Giguère[2], and Wail Gueaieb[1]
[1]School of Information Technology and Engineering, University of Ottawa, 800 King Edward Ave., Ottawa, ON, K1N6N5
[2] School of Rehabilitation Sciences, University of Ottawa, 451 Smith Rd., Ottawa, ON, K1H8M5
nellaham@uottawa.ca, cgiguere@uottawa.ca, wgueaieb@site.uottawa.ca

1 INTRODUCTION

Assessment of speech intelligibility is an important area in the design of hearing devices, and consists of two complementary approaches: (1) the subjective measurement of intelligibility scores based on experimental listening tests, and (2) the objective prediction of speech intelligibility under different listening conditions. Objective measures of speech intelligibility date back to the Articulation Index (AI) [1]. Extensions of the AI have led to models of speech intelligibility for normal hearing and hearing-impaired subjects in a variety of listening conditions. Two such models which have gained popularity and become standard in the acoustic literature are the Speech Intelligibility Index (SII) [2] and the Speech Transmission Index (STI) [3], both designed for monaural listening conditions.

In recent years, much attention has focused on binaural hearing. Binaural objective measures incorporate a pre-processing stage to model binaural interactions, followed by intelligibility prediction using a monaural measure like the SII or STI. In [4], Beutelmann and Brand developed one such measure using the equalization cancellation (EC) model [5] and the monaural SII. In a more recent work [6], Wijngaarden and Drullman developed a binaural measure based on the STI. Such measures, however, are not designed to deal with complex nonlinearities. Digital processing inside a hearing aid often involves nonlinear operations such as clipping, compression, and those found in noise reduction algorithms. In this paper, we propose adding an initial stage to deal with such nonlinear processing in order to obtain a new binaural objective measure of speech intelligibility for use with digital hearing aids.

Figure 1 shows a schematic diagram of the proposed 3-stage binaural objective measurement system. Conceptual details of each stage are presented in the subsequent sections of this paper. Inputs to the overall system are the individual speech and noise signals received at the left and right ears (4 signals in total). The left and right hearing aids are enclosed within the first processing stage (HA blocks). The system's output is an objective prediction of speech intelligibility.

2 SIGNAL SEPARATION

The first stage of the binaural measure is based on a simple signal separation scheme used by Hagerman and Olofsson in a monaural study on the effects of noise reduction algorithms in hearing aids [7]. Its incorporation into the proposed model has a dual purpose: (1) it provides separate access to the speech and noise signals at the output of the left and right hearing aids, a necessary requirement for the subsequent stage in the model; and (2) it deals with nonlinear signal processing and complex signal and noise mixtures inside the hearing aid.

Signal separation is performed by presenting the speech and noise signals simultaneously to the hearing aid twice, the second time with the phase of the noise reversed. For each presentation, the inputs (a_{in}, b_{in}) and outputs (a_{out}, b_{out}) of the hearing aid are summarized in the following equations:

$$a_{in}(t) = u(t) + v(t) \qquad a_{out}(t) = u'(t) + v'(t) + e_1(t) \tag{1}$$

$$b_{in}(t) = u(t) - v(t) \qquad b_{out}(t) = u'(t) - v'(t) + e_2(t) \tag{2}$$

where u(t) and v(t) denote the input speech and noise signals respectively, u'(t) and v'(t) the output speech and noise signals respectively, and $e_1(t)$ and $e_2(t)$ denote error signals (e.g. internal noise, nonlinear interactions between speech and noise, and distortion due to the noise reduction algorithm). Assuming the error signals are sufficiently small, we can recover separate estimates of the speech c(t) and noise d(t) output signals by adding and subtracting the outputs of the hearing aid from both presentations:

$$c(t) = a_{out}(t) + b_{out}(t) = 2u'(t) + e_1(t) + e_2(t) \tag{3}$$

$$d(t) = a_{out}(t) - b_{out}(t) = 2v'(t) + e_1(t) - e_2(t) \tag{4}$$

This approach was used to study the impact of different noise reduction algorithms, with and without compression under monaural conditions [7]. In the present work, this scheme is applied to retrieve the separate speech and noise signals at the output of the left and right hearing aids (Figure 1). These will be the inputs to the next stage of the proposed binaural measure.

3 EQUALIZATION-CANCELLATION MODEL

The next stage is the central binaural processing element, and is based on the EC theory introduced by Durlach [5]. The EC theory models the way the human auditory system uses binaural cues to improve the perception of a "target" speech signal in the presence of a noise masker. The model transforms the total signal at one ear relative to the total signal at the other ear such that the masking components are equalized (*equalization process*), and then subtracts one from the other (*cancellation process*). If the equalization

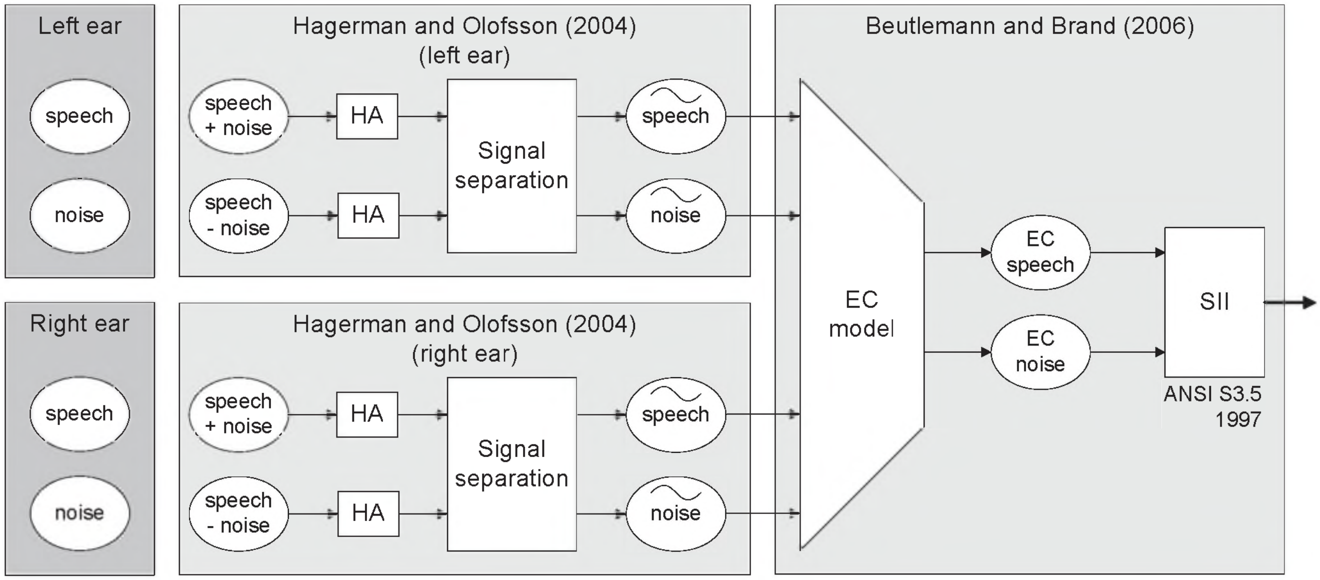

Figure 1: Diagram of the proposed binaural objective measure of speech intelligibility.

transform is perfect, and provided that interaural time and level differences for the target signals are different from those for the noise maskers, the latter will ideally cancel out due to destructive interference. If these interaural differences are identical, the EC processing will cancel out both target and masking components. Otherwise, the residual signal will theoretically have an improved SNR.

Since processing in the EC model is linear, speech and noise signals can be processed separately. This processing is performed independently in each band of the auditory filter, which is modeled using a Gammatone analysis filterbank. The EC model also includes a selection mechanism at the output of each band that selects the signal with the best SNR (left, right or EC-processed). Moreover, since the auditory system is not a perfect processor, the EC model includes various types of errors to account for human inaccuracies. Finally, the processed signals from each frequency band are combined at the output of this stage, to produce speech and noise signals to be used for intelligibility prediction.

4 INTELLIGIBILITY PREDICTION

In this work, prediction of speech intelligibility is done using the standard monaural SII. The SII is "a physical measure that is highly correlated with the intelligibility of speech as evaluated by speech perception tests given a group of talkers and listeners" [2]. It is essentially calculated as a weighted average of the amount of information available to the listener over a chosen number of frequency bands. The ANSI standard provides a detailed description of the SII calculation. Other standard measures could also have been used. The choice of the SII is primarily based on the model presented in [4], as well as the possible future extension of this model to use the coherence function in a manner similar to the work of Kates and Arehart with the Coherence-based SII [8].

5 CONCLUSION

This paper introduced a new objective measurement system for predicting speech intelligibility under binaural listening conditions. The new measure is designed for use with digital hearing aids, which often perform complex nonlinear signal processing. As binaural hearing aids become increasingly popular, the proposed measure will be useful to study the impact of emerging binaural noise reduction algorithms, similar to the monaural work presented in [7]. Our goal is also to use this measurement system to develop and test new algorithms for online learning of user control setting preferences in binaural trainable hearing aids.

REFERENCES

[1] ANSI-S3.5 (1969), "Methods for the calculation of the Articulation Index," American National Standards Institute, New York.

[2] ANSI-S3.5 (1997), "American National Standard: Methods for the calculation of the speech intelligibility index," American National Standards Institute, New York.

[3] IEC 60268-16 (1998), "Sound system equipment - Part 16: Objective rating of speech intelligibility by speech transmission index".

[4] Beutelmann R. & Brand T. (2006), "Prediction of speech intelligibility in spatial noise and reverberation for normal-hearing and hearing-impaired listeners," *J. Acoust. Soc. Am.*, 120 (1), pp. 331–342.

[5] Durlach NI. (1963), "Equalization and cancellation theory of binaural masking-level differences," J. Acoust. Soc. Am., 35 (8), pp. 1206–1218.

[6] Van Wijngaarden, SJ. & Drullman, R. (2008), "Binaural intelligibility prediction based on the speech transmission index", *J. Acoust. Soc. Am.*, 123 (6), 4514–4523.

[7] Hagerman B. & Olofsson A. (2004), "A method to measure the effect of noise reduction algorithms using simultaneous speech and noise", *Acust. Acta Acust.*, 90 (2), 356–361.

[8] Kates JM. & Arehart KH. (2005), "Coherence and the speech intelligibility Index", *J. Acoust. Soc. Am.*, 117, 2224-2237