PERCEIVED SOUND QUALITY OF NOISY AND NOISE REDUCED SPEECH

Muhammad M Nabil, Tiancheng Hu, Hussnain Ali, Chin-Tuan Tan

Cochlear Implant Lab
Erik Jonsson School of Engineering & Computer Science, University of Texas at Dallas, Richardson, Texas, U.S.A.
(mxn150230, Tiancheng.Hu, hxa098020, Chin-Tuan.Tan)@utdallas.edu

Cochlear Implant Laboratory

1. Introduction

- Speech enhancement, mostly nonlinear signal processing, is usually implemented as the front-end processing in assistive hearing devices to aide listening under noisy environments.

- The residual noise and distortion introduced by different nonlinear signal processing techniques may degrade perceived sound quality of the output speech differently.

- Most speech enhancement techniques generally focus on improving signal to noise ratio to achieve a minimum physical difference between the original clean speech and the enhanced speech.

- It may not provide the necessary sound quality for speech perception with reduced cognitive-load and hence require another metric.

2. Method

- Four speech sentences, each comprised of two short phrases from a male and female speakers extracted from AzBios database [1], were calibrated at 75dB SPL and 65 dB SPL . (a total of 8 clean speech sentences)

- Cafeteria and Babble noises [2] were calibrated at 75dB SPL, 65 dB SPL and 55 dB SPL. (a total of 6 noises)

- The following speech enhancement techniques [2] were used to process the speech sentences with or without the noises added (a total of 48 noisy speech sentences)

  - 1) Iterative Wiener Filter
  - 2) Psychoacoustic Masking
  - 3) No processing (noisy and clean speech sentences)

- A total of 192 test stimuli for sound quality evaluation prepared through combination of speech level, noise level and speech enhancement scheme.

- Speech sentence at 75 dB SPL and noise at 65 dB SPL was considered a different listening condition, sentences at 65dB SPL and noise at 55 dB SPL (although both were at same SNR, i.e., 10dB)

- Participants were asked to rate the perceived sound quality of the stimuli in 2 separate trials.

Subject	Subject-1	Subject-2	Subject-3	Subject-4	Subject-5	Subject-6	Subject-7
Between trials	0.838	0.894	0.938	0.902	0.946	0.789	0.914
Between individual and group	0.953	0.952	0.901	0.964	0.933	0.815	0.956

Table. 1. Correlation on ratings between trials for each participant and average ratings between individual participant and the group.

3. Results

- The rate of degradation on the perceived quality of noisy speech as noise of increasing level was added, was less steeper than when they were enhanced. (Fig 1 and Table 2)

- For stimuli with same signal-to-noise ratio SNR (0dB: 75/65 dB SPL of speech and noise; and 10 dB: 75/65 dB SPL of speech and 65/55 of noise), the average rating of the noisy speech was higher before enhancement. The effect was greater with babble noise than with cafeteria noise.

- A more gradual rate of degradation on the perception of the enhanced speech was observed with higher speech level (75dBSPL)
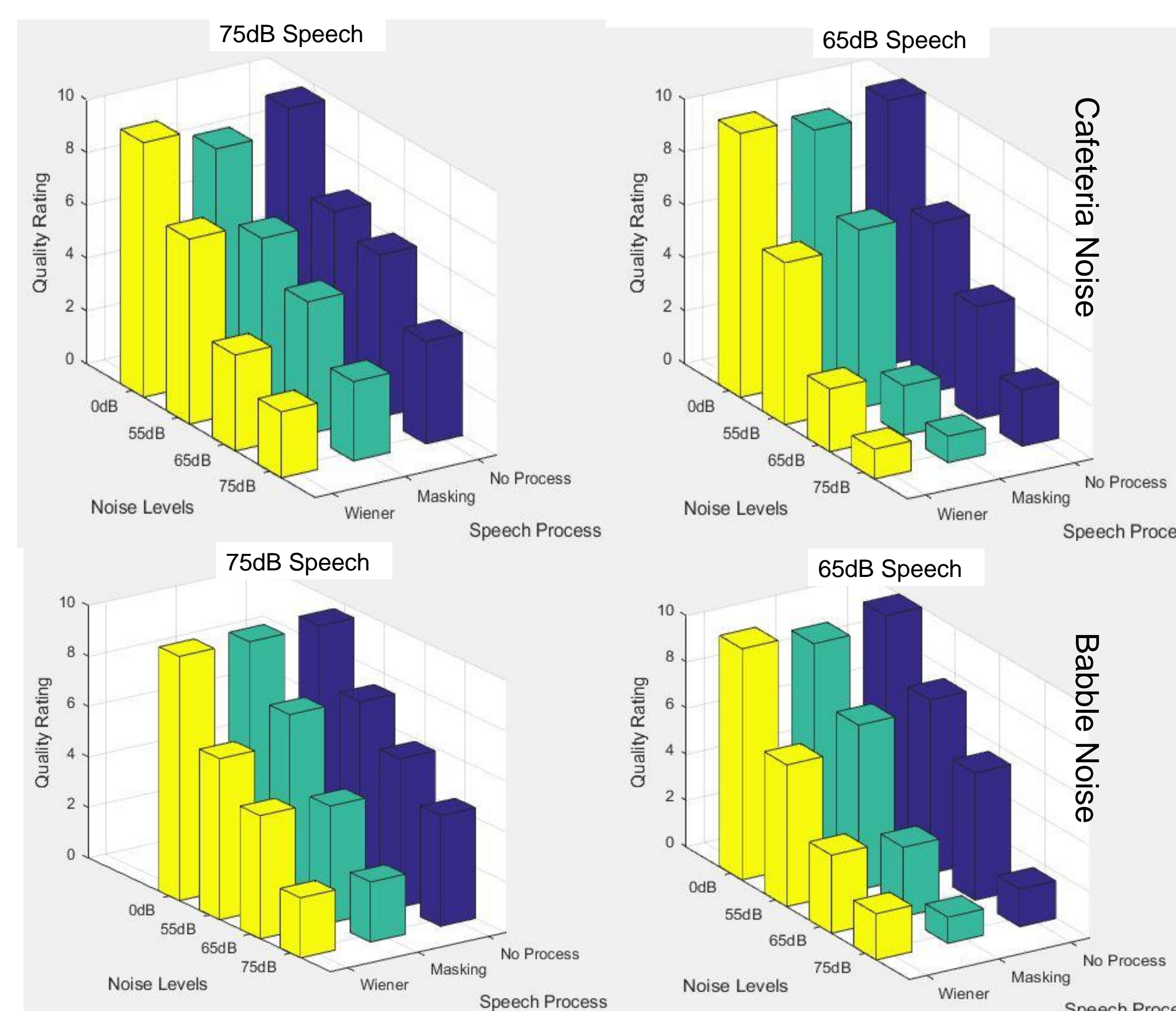

Fig. 1. Average user ratings for the three speech enhancement schemes for 75dB (left) and 65 dB (right) speech against different noise levels

Speech Level	75dB			65dB			
Noise Level	75dB	65dB	55dB	75dB	65dB	55dB	Cafeteria Noise
Average (No Process)	3.88	6.13	6.75	2.13	4.25	6.38	
Average (Wiener)	3.00	5.00	6.38	1.00	1.88	6.75	
Average (Masking)	2.50	3.63	7.00	1.13	2.38	6.13	
Improvement (Wiener)	-0.88	-1.13	--0.38	-1.13	-2.38	0.38	
Improvement (Masking)	-0.63	-1.29	0.29	-0.29	-0.46	-0.29	

Speech Level	75dB			65dB			
Noise Level	75dB	65dB	55dB	75dB	65dB	55dB	Babble Noise
No Process (A)	4.38	5.88	7.38	1.63	5.50	7.50	
Wiener (B)	2.38	4.63	7.50	1.13	3.00	7.13	
Masking (C)	2.38	4.88	6.38	2.00	3.38	6.13	
Difference (= B-A)	-2.00	-1.25	0.13	-0.50	-2.50	-0.38	
Difference (= C-A)	-0.67	-0.25	-0.71	4.42	-0.58	-0.79	

Table. 2. Average Ratings across all participants

Noise Levels	75dB	65dB	55dB	0dB	Correlation	Cafeteria Noise
No Process	1.75	1.88	0.38	-0.38		
Wiener	2.00	3.13	-0.38	-0.75	0.96	
Masking	1.38	1.25	0.88	-0.38	0.91	

Noise Levels	75dB	65dB	55dB	0dB	Correlation	Babble Noise
No Process	2.75	0.38	-0.13	-0.38		
Wiener	1.25	1.63	0.38	0.13	0.57	
Masking	0.38	1.50	0.25	-0.63	0.22	

Table. 3. Correlation between improvement in rating(65dB to 75dB increase in speech) for noisy speech and enhanced speech

5. Summary

- Quality perception in noisy and enhanced speech was dependent on the level of the target speech.

- Both remaining noise and nonlinear distortion in enhanced speech change the pattern of quality rating from that only degraded with additive noise.

- Nonlinear distortion prevents listeners from assessing the quality of the target speech more than the noise.

5. References

[1] A. J. Spahr, M. F. Dorman, L. M. Litvak, S. Van Wie, R. H. Gifford, P. C. Loizou, L. M. Loiselle, T. Oakes and S. Cook, "Development and validation of the AzBio sentence lists," Ear Hear., vol. 33, pp. 112-117, Jan-Feb, 2012

[2] P. C. Loizou, "Speech Enhancement: Theory and Practice," Databases, Second Edition. CRC Press, 2013