

# Monosyllable intelligibility improved by spectral equalization of spoken sound presented in the water: Influence of the shape of voice spectrum

Kazuoki Kuramoto<sup>1,\*</sup>, Shin'ya Kuwahara<sup>1</sup>, Kensei Oimatsu<sup>1</sup> and Sizuma Yamaguchi<sup>2</sup>

<sup>1</sup>Japan Coast Guard Academy,  
5-1 Wakaba, Kure, 737-8512 Japan

<sup>2</sup>Faculty of Engineering, Yamaguchi University,  
2557 Tokiwadai, Ube, 755-8611 Japan

(Received 22 August 2005, Accepted for publication 24 August 2005)

**Keywords:** Underwater audibility, Monosyllable intelligibility, Voice spectrum, Vowel and consonant, Listening experiment  
**PACS number:** 43.71.Es [DOI: 10.1250/ast.27.56]

## 1. Introduction

When we think about a voice transmission underwater, frequency compensations in consideration of auditory characteristics between in water and in air are engineeringly important because the human audibility is greatly different in each medium. Until now, we have carried out the listening experiment in water tank and the intelligibility has been generally improved by giving the frequency compensation of auditory characteristics [1]. Moreover, the simulation that the quasi-voice underwater was given in air has been carried out in the air and the intelligibility fell down by the weighting based on frequency characteristics of the auditory sense in the water [2]. Then, it has been found that the simulation in the air was effective when the intelligibility in the water was considered.

In this study, we examine the relationship between the intelligibility and the voice spectrum from the viewpoint of spectra of 50 monosyllables (which consists of 5 independent vowels and 45 monosyllables connected by vowel and consonant) presented in the listening experiment. In addition, the difference between woman voice and man voice is also reported.

## 2. Experiments

It is known that the underwater audibility of human ear becomes relatively bad between 2 kHz and 6 kHz in comparison with the audibility in air. Figure 1 indicates the difference of the audibility of human ear between in water and in air, which derived from the equal loudness contour in each medium [1]. Therefore, the man can accurately listen a voice sound in the water by adding the frequency compensation to underwater loudspeaker as shown in Fig. 1, that is, improvement of intelligibility becomes possible with respect to the voice transmission underwater. As it was confirmed that the listening experiment for the intelligibility in water could be simulated in air from the previous report [2], the listening experiment of the present study was carried out in the air. In short, a quasi-voice underwater was reproduced in the air by adding the weight of frequency compensation with an equalizer (EQ), in which the sign of the vertical axis of Fig. 1 was inverted. We compared the two cases, one is that

the EQ compensation was carried out and the other is without EQ. And the intelligibility (correct hearing rate) and confusion rank table (results of ranking whether subjects listen in what kind of syllable) were obtained for each syllable. As a syllable material, the monosyllable (voice production person is a NHK man announcer) recorded on CD [3] for evaluation of the hearing aid adaptation was adopted. The syllable table has been composed of 50 monosyllables consists of 5 independent vowels and 45 syllables connected by consonant and vowel. Here, 8 tables (total 400 monosyllables) that randomly combined these 50 monosyllables were made to be the one set and were presented for each subject. The man voice was reproduced from the CD as it is and the woman voice was carefully recorded with a constant sound pressure level and a regular interval following these syllable tables. Subjects are 5 normal hearing men in 20-year-old generation throughout the present experiments.

On the other hand, we also carried out the frequency analyses of monosyllable voices presented in this study for each case with EQ and without EQ in order to examine the relationship between the intelligibility and the spectrum of the monosyllables.

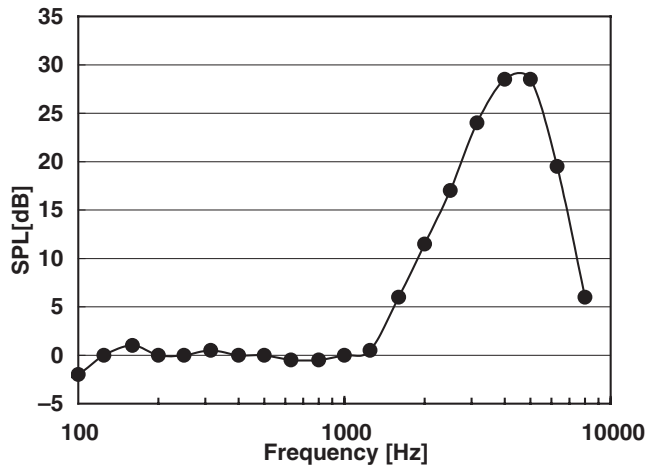
## 3. Results and discussion

The result of the intelligibility (correct hearing rate) of each monosyllable obtained by the listening experiment in the air is shown in Table 1. Both cases of the man voice and the woman voice, the representative results are shown respectively in the case without EQ in which the frequency compensation is not carried out and the other case with EQ in which the compensation was carried out. However, the vowel is conclusively shown as a column which analyzed all 50 syllables including independent 5 vowels, for example, "a" "ka" "sa" "ta" "na" ... are arranged in the "a-column."

### 3.1. On the vowel

From the results of arranging all syllables from "a-column" to "o-column," the intelligibility of the syllable of "i-column" with EQ has declined in both cases of the man voice and woman voice as compared with "a-column" or "o-column." Figure 2 shows the results of frequency analyses of independent vowel "a" and "i" in both cases with EQ and without EQ. It is found that the component of 2–6 kHz of the

\*e-mail: kuramoto@jcg.ac.jp



**Fig. 1** Frequency compensation curve between the audibility in water and in air.

vowel “*i*” was removed by means of the EQ compensation. This fact seems to be a cause of the lowering of the intelligibility in “*i*-column” shown in Table 1.

### 3.2. On the monosyllable including the consonant

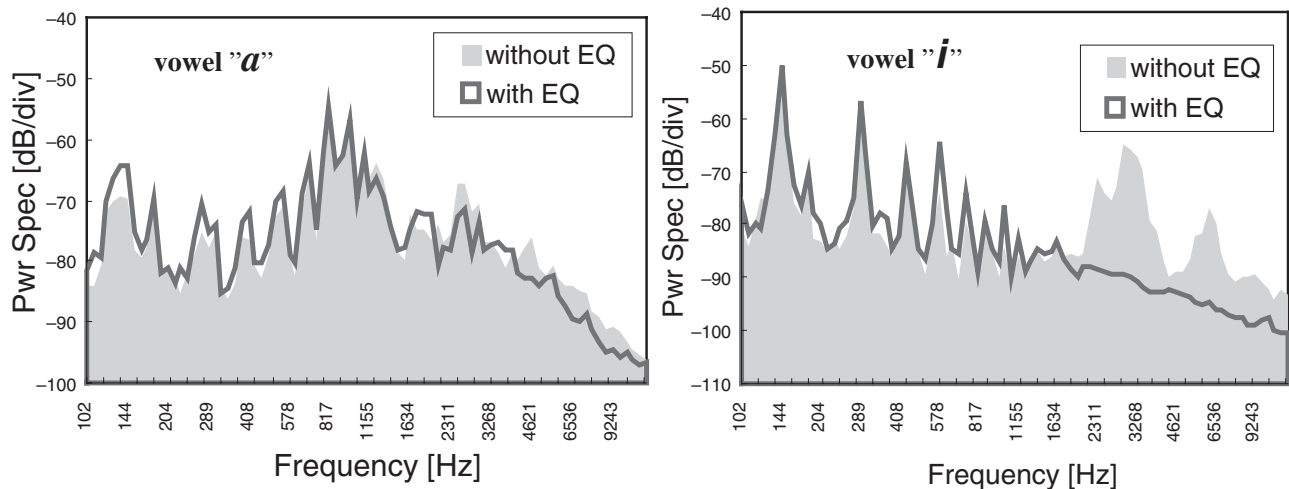
According to Table 1, the intelligibilities of “*ma*” and “*ya*,” which are connected with the vowel “*a*” and the consonant, hardly change without receiving the effect of the EQ compensation. On the other hand, the intelligibilities of “*shi*” and “*hi*,” which are connected with the vowel “*i*” and the consonant, largely lower by the EQ compensation. The frequency analyses of monosyllables of “*ma*” and “*shi*” are shown in Fig. 3. On the monosyllable of “*ma*,” the spectral shape is hardly affected by the EQ compensation as well as the vowel “*a*.” In the meantime, the monosyllable “*shi*” contains the high frequency component over 2 kHz because it consists of the vowel “*i*” including plosive consonant sound. This component was considerably removed by the EQ compensation and the intelligibility seems to become extremely bad.

**Table 1** Intelligibility (correct hearing rate) of each monosyllable obtained by the listening experiment in air.

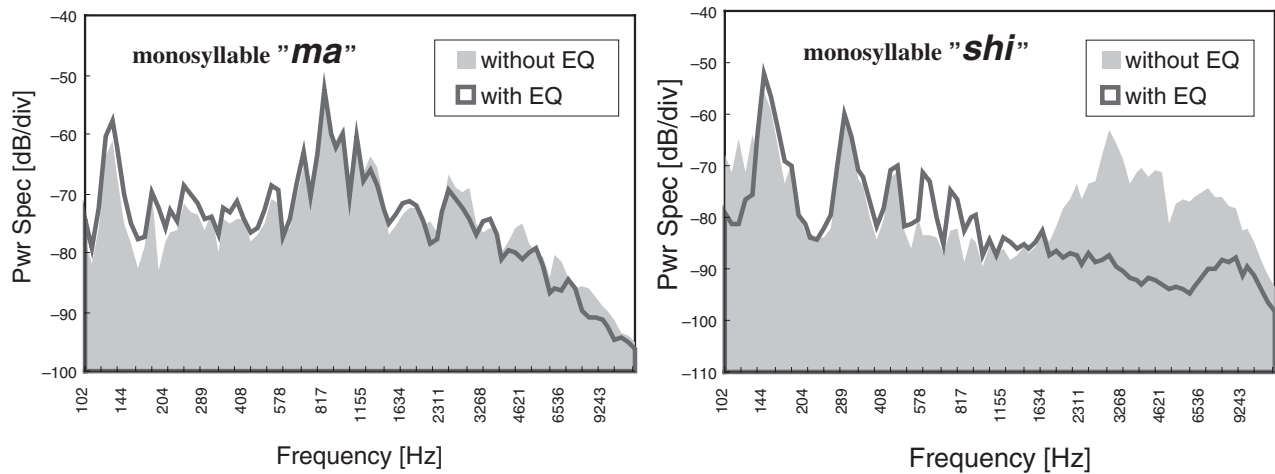
	Without EQ	With EQ
Man voice		
<i>a</i> -column	898/1040 = 0.863	845/1040 = 0.813
<i>i</i> -column	616/720 = 0.856	338/720 = 0.469
<i>u</i> -column	596/720 = 0.828	542/720 = 0.753
<i>e</i> -column	513/640 = 0.802	410/640 = 0.641
<i>o</i> -column	811/880 = 0.922	748/880 = 0.850
“ <i>ma</i> ”	80/80 = 1.0	80/80 = 1.0
“ <i>ya</i> ”	77/80 = 0.963	79/80 = 0.988
“ <i>shi</i> ”	79/80 = 0.988	28/80 = 0.350
“ <i>hi</i> ”	75/80 = 0.938	41/80 = 0.513
total	3,434/4,000 = 0.859	2,883/4,000 = 0.721
Woman voice		
<i>a</i> -column	441/520 = 0.848	374/520 = 0.713
<i>i</i> -column	308/360 = 0.856	127/360 = 0.353
<i>u</i> -column	300/360 = 0.833	168/360 = 0.467
<i>e</i> -column	271/320 = 0.847	170/320 = 0.531
<i>o</i> -column	373/440 = 0.848	295/440 = 0.670
“ <i>ma</i> ”	37/40 = 0.925	37/40 = 0.925
“ <i>ya</i> ”	39/40 = 0.975	35/40 = 0.875
“ <i>shi</i> ”	40/40 = 1.0	13/40 = 0.325
“ <i>hi</i> ”	34/40 = 0.850	16/40 = 0.400
total	1,693/2,000 = 0.847	1,134/2,000 = 0.567

### 3.3. On the difference between woman voice and man voice

We have obtained the results that the woman voice spectrum generally shifts to the high frequency side. And, it is also known that the *second formant frequency* deviates to the high frequency side in case of the woman voice. Therefore, it seems that the woman voice was received the effect of frequency compensation with EQ and the intelligibility generally lowers as compare with the man voice as shown in Table 1.



**Fig. 2** Frequency analyses of independent vowels “*a*” and “*i*” in both case with EQ and without EQ.



**Fig. 3** Frequency analyses of monosyllables of “ma” and “shi” in both case with EQ and without EQ.

#### 4. Conclusion

As a result of examining the relationship between the intelligibility and the voice spectrum of the syllable, it was confirmed that the intelligibility of the monosyllables containing the spectral composition with 2–6 kHz (monosyllable of “i-column” including the consonant, for example, “shi” “hi” “chi” “zi,” etc.) largely lowers in the water. Therefore, the frequency compensation in consideration of auditory characteristics in the water is indispensable for the purpose of the underwater voice transmission and accurate listening of the voice including the syllable of this type. Also, it was confirmed that the intelligibility of the woman voice lowered in the water because the voice spectrum shifts to the high frequency side as compare with the man voice.

#### References

- [1] S. Kuwahara, K. Oimatsu, K. Kuramoto and S. Yamaguchi, “Monosyllable articulation improved on an equalized frequency characteristics for underwater auditory,” *Proc. Spring Meet. Acoust. Soc. Jpn.*, pp. 1317–1318 (2003).
- [2] S. Kuwahara, K. Oimatsu, K. Kuramoto and S. Yamaguchi, “Monosyllable articulation improved on an equalized frequency characteristics for underwater auditory — Comparison of a simulation experiment in air with underwater results —,” *Proc. Spring Meet. Acoust. Soc. Jpn.*, pp. 1059–1060 (2004).
- [3] K. Yonemoto, “Characteristics of CD(TY-89) for assessment of a suitable hearing aid,” *J. Otolaryngol. Head Neck Surg.*, **11**, 1395–1401 (1995).