

音声の変動量と「聴き取りにくさ」について  
— 残響環境下での様々なアナウンスの評価 —\*

○栗栖清浩 (TOA), 中村 進, 安 啓一, 荒井隆行 (上智大・理工)

## 1 はじめに

拡声システムによる音声の明瞭性評価に  $STI^{[1]}$  が用いられることがあるが、イコライザ等の調整によりすっきりと聞き取りやすい拡声音となっても、 $STI$  が殆ど変化せずその調整効果を数値で評価することが出来ないことがあった。そこで著者らは過去に提案した  $SOR$  (Signal to Overlap masking Ratio)  $^{[2]}$  に着目し、ある特定の音声を用いた聴取実験から、「聴き取りにくさ」の割合  $LDR$  (Listening Difficulty Rating)  $^{[3]}$  と  $SOR$  の関係式を得た。これにより拡声システムの調整で改善する  $LDR$  を、 $SOR$  の改善量として評価できることを示した $^{[4][5]}$ 。また様々な音声での  $LDR$  を調査し、 $SOR$  の改良について検討した $^{[6]}$ 。

今回、特定の音声だけでなく様々な音声において  $LDR$  を精度良く推定することを目的に、前報 $^{[6]}$ に引き続き聴取実験から  $LDR$  と  $SOR$  の関係を調べた。また、音声のラウドネス変動に着目したところ、変動量  $DLF$  (Depth of Loudness Fluctuation)  $^{[7]}$ により  $SOR$  とほぼ同程度の精度で  $LDR$  を推定できることが分かった。

## 2 聴取実験

前報 $^{[6]}$ と同じ100刺激(男声A:5文, 男声B:5文の計10文に10インパルス応答を畳込み), 同じ提示方法(防音室内で耳覆い型ヘッドホン使用)で、聴力に問題ない日本語を母語とする聴取者40名(20代:7名, 30代:12名, 40代:9名, 50代:10名, 60代:2名)が刺激の「聴き取りにくさ」を評価した。前報 $^{[6]}$ の10名と合わせて50名分のデータを用い、各刺激について  $SOR$  と  $Z_{LDR}$  ( $LDR$  の  $Z$  値) の関係を図示したのがFig. 1である。ここで確率密度関数 ( $LDR$ ) から分布関数 ( $Z$  値) への変換は、Williams の近似式 $^{[8]}$ を用いた。

10文で個別に回帰直線を計算すると決定

係数  $R^2 = 0.98$  となる文もみられたが、全刺激ではFig. 1にあるとおり  $R^2 = 0.92$  の回帰直線となった。かなり高い決定係数ではあるが、たとえば  $SOR = 0$  dB において  $LDR$  が約1SD (標準偏差) の範囲に分布している。 $LDR$  が1SD 異なると、音声伝送性能  $STP$  (Speech Transmission Performance)  $^{[9]}$  が1クラス異なる場合があることから、このばらつきが実用上問題ないか更に検討が必要だと思われる。

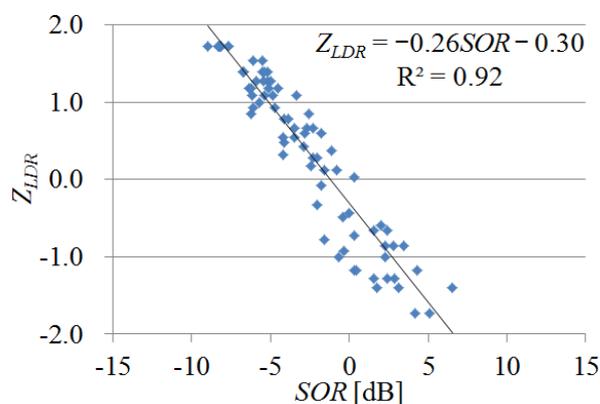


Fig. 1 Z-score of  $LDR$  vs.  $SOR$  and the regression line.

## 3 変動音解析

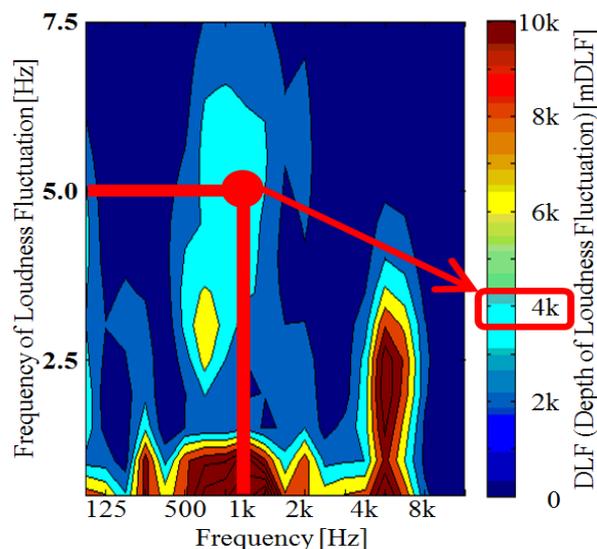


Fig. 2 An example of the DLF analysis

\* The relationship between the depth of loudness fluctuation and the listening difficulty ratings of various sentences under reverberant condition, by KURISU Kiyohiro (TOA Corp., Japan), NAKAMURA Susumu, YASU Keiichi and ARAI Takayuki (Sophia Univ., Japan).

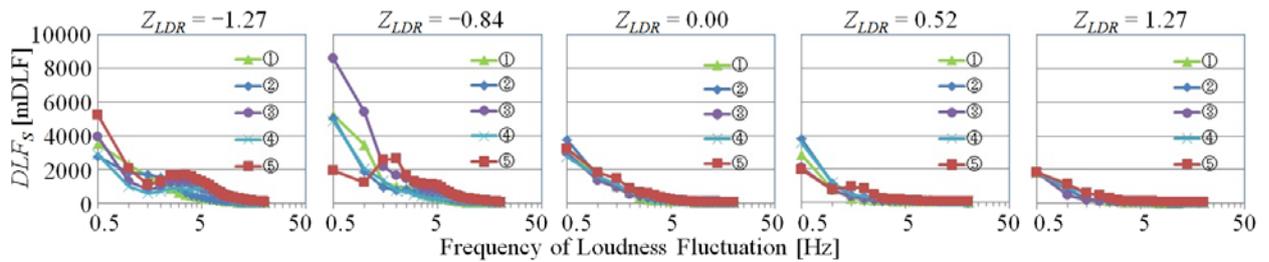


Fig. 3 Variation of the Depth of Loudness Fluctuation ( $DLF_S$ )

音声の明瞭性指標である  $STI^{[1]}$  は  $MTF$  (Modulation Transfer Function) を基に算出されるが、これは音源の AM 変調が受信点にどれだけ正確に伝搬しているかを数値化したものである。そこで刺激の変動に着目し、 $LDR$  が異なる刺激では変動にどのような違いがあるか調査した。

Fig. 2は市販ソフト<sup>[7]</sup>によるある刺激の変動音解析例である。図中の例は、1 kHz を中心周波数とする 1/3 oct. 帯域において、変動周波数 5.0 Hz で変動しているラウドネス変動量  $DLF$  が約 4000 mDLF であることを示しており、帯域が異なると、異なる変動周波数において大きく変動していることが分かる。

Fig. 3は聴取実験で用いた刺激を変動音解析した結果で、同じ  $LDR$  であった 5 刺激ずつの結果を示した。ここで縦軸は、変動周波数ごとに変動量  $DLF$  を加算 (Fig. 2の各変動周波数について横軸方向に数値を合計) した  $DLF_S$  で表示している。

$Z_{LDR} = 1.27$  のように  $LDR$  の  $Z$  値が大きく聴き取りにくい傾向の刺激では、変動量が広い変動周波数に渡って小さいが、 $Z_{LDR} = -1.27, -0.84$  のように聴き取りにくい傾向の刺激では、低い変動周波数において特徴的な変動量の盛り上がりが見られた。

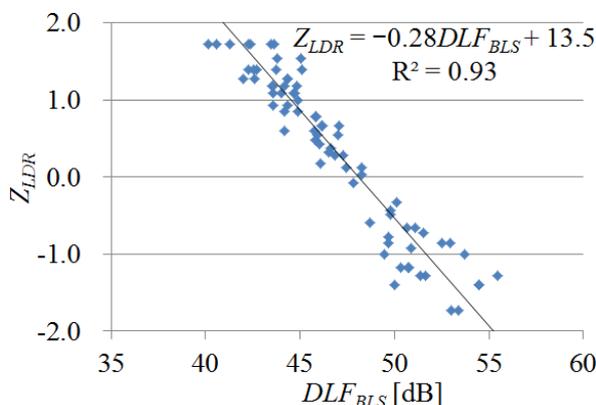


Fig. 4 Z-score of  $LDR$  vs.  $DLF_{BLS}$  and the regression line.

これらのことから変動周波数が 2.5 to 7.5

Hz の範囲の変動量を合計した  $DLF_{BLS}$  (変動帯域限定  $DLF$  和) と  $Z_{LDR}$  の関係を図示したのが Fig. 4である。尚、 $DLF_{BLS}$  は常用対数を 10 倍した dB で表示している。

$DLF_{BLS}$  を説明変数とし  $Z_{LDR}$  を目的変数として 10 文を個別に回帰分析すると、回帰直線の決定係数は 9 文において数%改善し、1 文において 0.4%悪化したが、全データでは  $SOR$  を説明変数とした時 (Fig. 1) と比べ、決定係数はほぼ同じであった。例えば  $DLF_{BLS} = 50$  dB において Fig. 1 と同様に  $LDR$  に 1 SD 程度のばらつきがあることが分かる。よって  $DLF_{BLS}$  による  $LDR$  の推定精度は、 $SOR$  による推定と同程度だと期待される。

ここで  $MTF$  は伝送系のみを数値評価したものであるが、刺激には伝送系だけでなく音源そのものの特性も含まれていることを指摘しておく。

#### 4 まとめ

聴取者 50 名による聴取実験結果から、 $LDR$  の  $Z$  値  $Z_{LDR}$  と  $SOR$  の間、または  $Z_{LDR}$  と変動量  $DLF_{BLS}$  の間は、ほぼ同精度の回帰式で関係づけられていることが分かった。

今回用いた刺激はいずれも男声であったので、女声や外国語音声であっても  $SOR$  または  $DLF_{BLS}$  で本報告と同様に  $LDR$  が推定できるか興味あるところである。

#### 参考文献

- [1] IEC 60268-16, Edition 4.0, 2011.
- [2] Arai T *et al.*, *AST*, 28(6), 438-441, 2007.
- [3] Morimoto M *et al.*, *JASA*, 116(3), 1607-1613, 2004.
- [4] 栗栖 他, 建音研資料, AA2011-47, 2011.
- [5] 栗栖 他, 音講論(春), 1225-1226, 2012.
- [6] 栗栖 他, 音講論(秋), 1251-1252, 2012.
- [7] 小野測器, OS-0272 変動音解析ユーザーズガイド.
- [8] 日本規格協会, 統計数値表 JSA-1972.
- [9] 日本建築学会, AIJES-S0002-2011.