

## 拡声音のオーバーラップマスキング量SORと「聴き取りにくさ」の関係 — 拡声系(歪・発話速度)と音響系が変化する事例 —\*

☆中村進(上智大・理工), 栗栖清浩(TOA), 安啓一, 荒井隆行(上智大・理工)

### 1 はじめに

都市・建築空間において, 非常アナウンス等の音声情報を聴取者へ明瞭に伝えるために, 様々な場所で拡声システムが用いられる。そこで, 拡声システムによって伝達された音声の明瞭性を評価する必要がある。

従来, 明瞭性を評価する手法として *STI* (Speech Transmission Index)<sup>[1]</sup>が広く用いられているが, 拡声音の明瞭性を評価するならば, *pathware* (拡声系+音響系)<sup>[2]</sup>における評価が必要である。そこで栗栖らは, 拡声系に線形な処理が含まれる場合における *SOR* (Signal to Overlap-masking Ratio)<sup>[3]</sup>の評価を行った<sup>[4][5]</sup>。

本報告では, 音響系による明瞭性の変化に加え, 拡声系に非線形処理(歪・発話速度)が含まれる場合における明瞭性を *SOR* にて評価する。聴取実験によって得た「聴き取りにくさ<sup>[6][7]</sup>」の結果と *SOR* の関係を線形回帰式により表し, *SOR* の実用性を評価する。

### 2 評価法

#### 2.1 Signal to Overlap-masking Ratio

従来の *SOR*<sup>[3]</sup>は, 単音節明瞭度に対応させるため, 単音節のみに着目して算出していた。しかし, 本報告では拡声音全体を評価するため, 時間フレーム(150ms)ごとに音声と overlap-masking 量の比を求め, それらを平均した *SOR*<sup>[4][5]</sup>を用いた。

#### 2.2 聴き取りにくさ

音声の明瞭性を評価する方法として, 高親密度の単語を用いる「聴き取りにくさ(*LDR: Listening Difficulty Rating*)<sup>[6][7]</sup>」が提案された。本報告では, 予め決められた1文に対し「聴き取りにくさ」を評価した。刺激それぞれについて“1. 聴き取りにくくない, 2. やや聴き取りにくい, 3. かなり聴き取りにくい, 4. 非常に聴き取りにくい”という4段階評価を行い, 2~4を選ぶ人数の百分率を「聴き取りにくさ」とした。

### 3 聴取実験

#### 3.1 聴取者

聴取者は, 日本語を母語とする健聴者 34名(男性 31名, 女性 3名), 平均年齢 38.4歳(20代 8名, 30代 11名, 40代 8名, 50代 6名, 60代 1名)であった。

#### 3.2 刺激

2人の話者(男女)のアナウンス音源“職員の指示に従い, 落ち着いて避難してください”において, 発話速度, 零挿入, クリップ, 音響系, 音響調整の異なる全 160 刺激を用いた<sup>[4][5]</sup>。本報告では, 各音響系におけるクリップ処理と発話速度の違いについて調査した。

##### 3.2.1. 音響系

残響時間(RT)の異なる 5つ(RT = 0.2, 0.5, 0.9, 1.4, 2.6s)のインパルス応答(IR)を, 音声にそれぞれ畳み込むことによって, 5通りの音響系を模擬した。それぞれの IR は, ある体育館の IR (RT = 8.53s)における減衰カーブを調整し作成した。

##### 3.2.2. クリップ

音源の振幅を 50 倍に増幅した後, その振幅の絶対値が最大 1 を超えないように制限する非線形な処理を施した。これにより, 歪の生じた処理音を作成した。クリップ処理を施すか否かで, 各音響系において 2通りずつの音声を作成した。

##### 3.2.3. 発話速度

音響分析ソフトウェア Praat<sup>[8]</sup>の PSOLA アルゴリズムにより, 音源の時間長を 0.6, 1.0, 1.4 倍とし, 発話速度が異なる 3通りの処理音(男声: 4.7, 6.6, 11.0 mora/s, 女声: 4.8, 6.7, 11.3 mora/s)を作成した。

#### 3.3 実験手順

防音室にて, 耳覆い型ヘッドホンを通じて全 160 刺激をランダムに受聴し, 刺激ごとに「聴き取りにくさ」を評価した。

\* Signal to overlap-masking ratio of the transmitted speech and its listening difficulty: Cases of various combinations of sound system and acoustic path, by NAKAMURA, Susumu (Sophia Univ., Japan), KURISU, Kiyohiro (TOA Corp., Japan), YASU, Keiichi and ARAI, Takayuki (Sophia Univ., Japan).

### 3.4 結果・考察

#### 3.4.1. クリップ

実験から得られた「聴き取りにくさ」と *SOR* の関係を Fig. 1 (左) に示した。また、「聴き取りにくさ」を *z* 値<sup>[9]</sup>へと変換した結果を Fig. 1 (右) に示した。Fig. 1 (右) より、線形回帰式と決定係数  $R^2$  を求めた。

*SOR-LDR* において、決定係数の高い線形回帰式が得られた。このことから、拡声系で非線形のクリップ処理がある場合においても、*SOR* から「聴き取りにくさ」を予想できることが分かった。*SOR* は実際の拡声音を用いて測定するため、処理の線形性の如何に関わらず、拡声時に音声に変化したならばその違いに敏感に反応すると考えられる。

#### 3.4.2. 発話速度

男声、女声アナウンスにおける「聴き取りにくさ」と *SOR* の関係をそれぞれ Figs. 2, 3 に示した。「聴き取りにくさ」の *z* 値と *SOR* の関係から、男声女声それぞれにおいて、発話速度ごとに線形回帰式と決定係数を求めた。

発話速度が 6.6, 6.7 mora/s の時に比べ、4.7, 4.8 mora/s の方が聴き取りにくくなった理由としては、発話速度を遅くする際、母音部が伸長されすぎたために *overlap-masking* が増加したことが原因ではないかと考えられる。

また、発話速度が異なると線形回帰式の傾きが変化することから、発話速度を変数とした回帰式が比較的簡単な形で得られるものと考えられる。

### 4 おわりに

従来、*STI* では対象外であった拡声系に含まれる非線形処理による明瞭性の違いを、*SOR* によって評価することを試みた。ある特定の音源による結果ではあるが、拡声系(歪、発話速度)と音響系が変化する事例において、*SOR* と「聴き取りにくさ」の関係を線形回帰式で示すことができた。今後、様々な発話速度での結果を検証し、発話速度が異なる場合の *SOR* と「聴き取りにくさ」の回帰式を求め、考察することを行う。

#### 参考文献

- [1] IEC 60268-16, Edition 4.0, 2011.
- [2] D. Campbell, Syn-Aud-Con Newsletter, 35(1), 10-17, 2007.
- [3] T. Arai et al., AST, 28(6), 438-441, 2007.

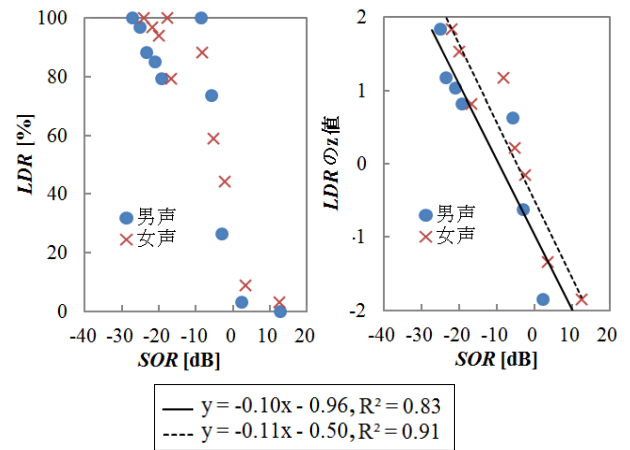


Fig. 1 *SOR-LDR* (話者×2, 残響×5, クリップ×2)

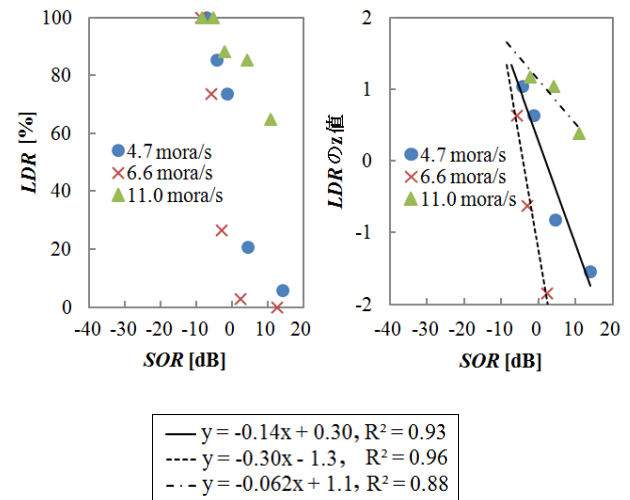


Fig. 2 *SOR-LDR* (男声, 残響×5, 発話速度×3)

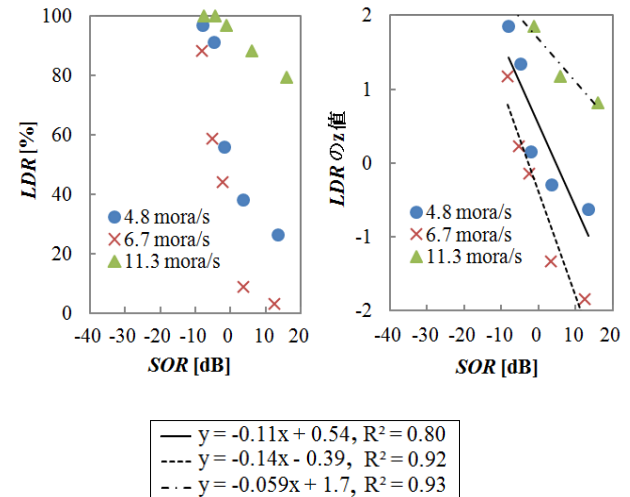


Fig. 3 *SOR-LDR* (女声, 残響×5, 発話速度×3)

- [4] 栗栖 他, 建音研資料, AA2011-47, 2011.
- [5] 栗栖 他, 音講論(春), 1-9-21, 2012.
- [6] M. Morimoto et al., JASA, 116(3), 1607-1613, 2004.
- [7] H. Sato et al., JASA, 63(5), 275-280, 2007.
- [8] P. Boersma and D. Weenink, Praat: <http://www.fon.hum.uva.nl/praat/>
- [9] 日本建築学会環境基準, AIJES-S0002-2011, 2011.