

残響下での発話変化を応用した公共空間での音声明瞭度改善 -話者による明瞭度の違い-*

◎程島奈緒（東海大），荒井隆行（上智大），栗栖清浩（TOA）

1 はじめに

音声の生成と知覚には密接な関係があり，聴取する雑音に応じて発話やその明瞭度は変化する（Lombard 効果）^[1-3]．具体的には，雑音下の発話の方が静音下の発話に比べて，単語の時間長・インテンシティ・基本周波数・第一フォルマント周波数などが増加し^[2,3]，雑音下で若年者の単語理解度が上昇した^[1-3]．

聴取する音が残響音でも，Lombard 効果と同様に残響下で音声明瞭度の改善が得られている．残響下の発話の方が静音下の発話に比べて，先行研究^[2,3]と同様の音響的な特徴が増加し^[4]，静音下での発話を静音下で聞くよりも，残響下/雑音・残響下での発話を残響下/雑音・残響下で聞く方が若年者・高齢者・非母語話者の単語理解度が上昇した^[5-7]．一方，残響下の発話において音声明瞭度が改善しない話者もみられた^[8,9]．

本研究の目的は，雑音や残響が存在する公共空間において，雑音や残響音を聞きながら発話した音声を拡声することにより音源系で明瞭な音声案内を作成することである．本研究では，先行研究^[5-9]で明らかにされなかった(1) 残響下の発話が静音下の発話よりも明瞭になる条件は話者に依存するのか，(2) 話者にフィードバックする残響音のレベルによって発話された音声の明瞭度が変化するか，について聴取実験による検討を行った．

2 聴取実験

参加者は日本語を母語とし，日常の聞こえに支障がないと回答した若年者 25 名（男性 20 名，女性 5 名，平均年齢 22 歳）である．

原音声は，東京方言話者 10 名（参加者とは異なる男女各 5 名，平均年齢 23 歳）が，キャリア文に挿入された 4 モーラのターゲット語（FW03^[10]より親密度が 2.5~4.0 の 50 語）を 3 発話条件 [静か (Q)，残響 (R1, R2)]

下で発話した音声である．R1・R2 条件では，直接音を除いたインパルス応答（残響時間 3.6 s）を発話に畳み込んだ残響音を，ヘッドホン（SENNHEISER, HDA200; dynamic, closed circumaural type）から話者に提示しながら発話させた．残響音のフィードバックレベルは，発話レベルに対して -22 dB (R1) / -10 dB (R2) である．音声はマイクロホン（SHURE, KSM141），オーディオインタフェース（TASCAM, US-144MKII）を介してコンピュータに録音した後，キャリア文とターゲット語のインテンシティ比を揃え，話者・発話条件毎に同一のキャリア文にターゲット語を挿入した．この音声に録音時と同じインパルス応答を畳み込み，音圧レベルを正規化し刺激とした．

聴取実験は，参加者毎に防音室内で行った．2 回の練習の後，48 刺激（3 発話条件×2 話者×8 ターゲット語）をランダムな順でヘッドホン（STAX, SR-303; electrostatic, open circumaural type）から参加者に聞きとりやすい音圧レベルで提示した．話者/ターゲット語と発話条件の組み合わせは，参加者間でカウンタバランスをとった．

3 結果・考察

図 1 に話者毎 (M: 男性話者, F: 女性話者)・発話条件毎のモーラ正解率を示す．混合要因の 2 元配置の分散分析及び多重比較の結果，話者の効果と発話条件による主効果は共に有意であった ($p < 0.01$)．先行研究^[5-7]と同様に残響下の発話 (R1/R2) は静音下での発話 (Q) よりも，モーラ正解率は有意に高くなった ($p < 0.05$)．別の話者では残響下で発話された音声は静音下と比べて音響的な特徴が変化したことから^[4]，R1・R2 条件では音響的な特徴の変化により明瞭度が上昇したと考えられる．残響のフィードバックレベルが発話レベルに

* Improving speech intelligibility by reverberant-induced speech at public spaces: Effect of speaker difference on speech intelligibility, by HODOSHIMA Nao (Tokai University), ARAI Takayuki (Sophia University) and KURISU Kiyohiro (TOA).

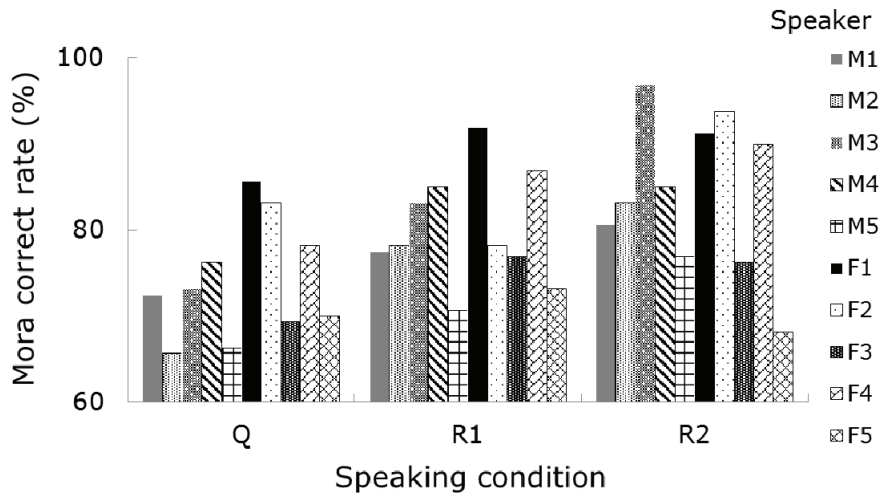


Fig. 1. Mean percent correct of mora for each speaking condition (Q: quiet, R1/R2: reverberation) and speaker (M: male, F: female).

対して-22 dB と-10 dB で、モーラ正解率の平均に有意差はみられなかったが、R1・R2 条件間で正解率が変化している話者もいることからさらなる分析が必要である。

話者条件において、F1 は M1, M2, M5, F3, F5 に対して、F2 は M5, F5 に対して、F4 は M5, F5 に対してモーラ正解率が有意に高くなった ($p < 0.05$)。また、Q 条件に比べて R1, R2 条件の正解率が高くなった話者は、時間長が長くなり基本周波数が高くなる傾向が見られた。Lombard 効果^[3]と同様に、残響音を聴取しながらの発話方法は話者によって異なり、全ての話者に対して音声明瞭度が改善するわけではないことが示された。

4 おわりに

残響下で発話された音声の方が静音下で発話された音声よりも、残響下で若年者に対するモーラ正解率が有意に高くなることを聴取実験により示した。さらに、残響下の発話によって音声明瞭度が改善するかは、話者（の発話方法）に依存すること、本研究の条件内では残響のフィードバックレベル（-22 dB と -10 dB）を変えてもモーラ正解率は変化しないことが今回明らかになった。

本結果から、残響を聞きながら発話をした音声を公共空間で拡声する方が、従来の静音下で発話/合成された音声よりも明瞭となる可能性が示された。雑音や残響下では高齢者や非母語話者の方が若年の母語話者よりも一

般的に音声明瞭度が低下することから、今回の話者が高齢者や非母語話者に対しても明瞭であるのかの検討、提案法が有効である話者や発話条件を検討する必要がある。残響下の発話に現れるどの音響的な特徴が音声明瞭度に関連しているかの調査も行い、話者への発話訓練や聞き取りやすい合成音声への応用も検討したい。

謝辞

本研究は東海大学総合研究機構の助成を受けて行われた。

参考文献

- [1] Lane and Tranel, J. Speech Hear. Res., 14, 677-709, 1971.
- [2] Junqua, J. Acoust. Soc. Am., 93, 510-524, 1993.
- [3] Summers *et al.*, J. Acoust. Soc. Am., 84, 917-928, 1988.
- [4] 程島他, 電子情報通信学会技術報告, SP2009-69, 43-48, 2009.
- [5] Hodoshima *et al.*, Proc. ICA, 2010.
- [6] Hodoshima *et al.*, Proc. Interspeech, 2012.
- [7] 島上他, 音講論 (春), 439-440, 2012.
- [8] 井下田他, 音講論 (春), 643-644, 2014.
- [9] 井下田他, 音講論 (秋), 493-494, 2014.
- [10] 天野他, 音声資源コンソーシアム, 2006.