

科学研究費補助金研究成果報告書

平成 24 年 5 月 31 日現在

機関番号：32644
 研究種目：若手研究（B）
 研究期間：2009～2011
 課題番号：21700203
 研究課題名（和文） 残響下で高齢者や非母語話者に明瞭な録音音声・肉声による拡声音声の調査
 研究課題名（英文） Research on intelligible live/recorded speech announcements in reverberant environments for older adults and non-native listeners
 研究代表者
 程島 奈緒（HODOSHIMA NAO）
 東海大学・情報通信学部・講師
 研究者番号：40453609

研究成果の概要（和文）：雑音下で発話の特性が変化することを応用し、駅など残響がある公共空間で高齢者や非母語話者に聞きやすい音声案内の作成を目的とした。若年者と高齢者に対する聴取実験の結果、白色雑音/残響を聞きながら発話した音声は、静かな環境で発話した音声（現在の音声案内）よりも、雑音/残響下（信号対雑音比=-2.0 dB、残響時間 1.4～3.6 s）で明瞭度が有意に上昇した。これらの音声を公共空間に拡声することで明瞭な音声案内を提供し、「音声によるバリアフリー」を実現できる可能性を示した。

研究成果の概要（英文）：This study introduces an approach that makes speech announcements intelligible for older adults and non-native listeners in reverberant public spaces (e.g., train stations). This approach is based on the way we change how we speak depending in a noisy environment. Listening tests with young and older adults showed that speech spoken in noisy/reverberant environments was statistically intelligible than speech spoken in a quiet environment (i.e., normal speech announcements) with background noise/reverberation conditions (signal to noise ratios were -2 and 0 dB, and reverberation times were from 1.4 to 3.6 s). The results suggest that using speech spoken in noise/reverberation for public address systems makes speech announcements more intelligible in public spaces.

交付決定額

(金額単位：円)

	直接経費	間接経費	合計
2009 年度	1,000,000	300,000	1,300,000
2010 年度	800,000	240,000	1,040,000
2011 年度	600,000	180,000	780,000
総計	2,400,000	720,000	3,120,000

研究分野：音声情報処理

科研費の分科・細目：情報学・知覚情報処理・知能ロボティクス

キーワード：音声、明瞭度、高齢者、公共空間、残響、バリアフリー

1. 研究開始当初の背景

駅や空港などの公共空間では、残響（音が反射して響く現象）によって音声案内が聞き取りづらくなることがある。この残響による音声の明瞭度低下は、若年者よりも高齢者、

聴覚障害者、非母語話者の方が大きい[1他]。

公共空間の音声案内には地震や火災が発生した際の非常放送も含まれるが、公共空間における音声伝達に関する詳細な規定はない。日本では 2055 年に約 2.5 人に 1 人が高

齢者になると推計されていることや[2]、外国人人口の増加[3]をふまえると、公共空間での音声明瞭度の確保は重要な課題である。

2. 研究の目的

本研究の目的は、公共空間の利用者に明瞭な音声を提供するため、残響環境下でも聞き取りやすい音声案内を開発することである。

人は周囲の環境に応じて話し方を変化させる。雑音下での発話は、静かな環境での発話に比べて、音響的な特徴（発話の音響分析を行うと現れる特徴）が変化し、雑音下の正解率が上昇する[4 他]。雑音とはマスキングの特性が異なる残響下では、このような現象は検証されていない。

公共空間では残響が音声明瞭度を低下させる原因となるため、本研究ではこの現象を応用してより明瞭な音声案内を作成することができるのではないかと仮説を立てた。すなわち、残響下の発話においても、静かな環境での発話（即ち、通常の方法で作成された音声案内）と比べて音響的特徴が変化し、残響下で明瞭となるのかを調査した。その際に、

- ・公共空間で用いられる音声案内の形態（既に録音された音声を拡声する自動放送と、発話者が公共空間内で発話した音声を拡声するマイク放送）によって、提案する音声案内の明瞭度は変化するのか、
 - ・提案する音声案内は、聴取者（若年者、高齢者、非母語話者）によって明瞭度の改善度合いは変化するのか、
- も調査した。

3. 研究の方法

以下の手順で研究を行った。

(1) 音声の録音

発話者は東京方言話者 6 名であり、原音声はキャリア文に挿入された 4 モーラのターゲット語である。図 1 に録音環境を示す。音声は防音室においてマイク、マイクアンプ、オーディオインタフェースを介してコンピュータに録音した。発話条件は、静か (Q)、雑音 (N)、残響 (R1, R2) である。雑音/残響条件では、白色雑音または残響音[マイクに入力された音声に残響時間が 3.6 s (R1) と 12.3 s (R2) のインパルス応答を畳み込んだ音声]をクローズドヘッドホンから提示した。発話者には、発話した音声ヘッドホンから聞こえる雑音や響きを持つ公共空間内の聴取者に拡声されることを想像し、できるだけ明瞭に発話するよう指示した。発話者の耳元における提示音圧レベルは、4 名には平均 80dB、2 名には発話者の口元の発話レベルに対して -22 dB とした。録音音声には、雑音・残響のいずれも含まれていない。

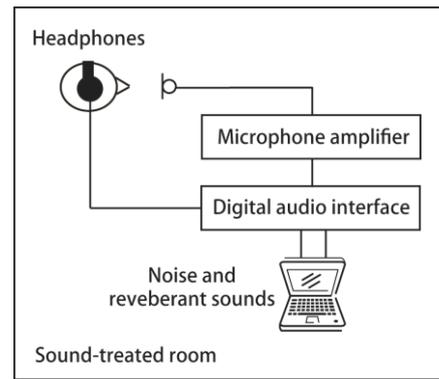


図 1 録音の環境

表 1 聴取条件

		Recording condition			
		Q	N	R1	QR1
Listening condition	N	Q_N	N_N		
	Ra	Q_Ra		R1_Ra	QR1_Ra
	Rb	Q_Rb		R1_Rb	QR1_Rb

(2) 刺激の作成

実験で使用する刺激は、Q、N、R1、QR（キャリア文は Q、ターゲット語は R を組み合わせた音声）条件下で録音された 2 名の発話に、白色雑音の付加もしくはインパルス応答の畳み込みを行った音声で、表 1 に示す 8 条件である。表 1 の記号の意味は、例えば Q_N は Q 条件下で録音した音声に、雑音を加えて聴取実験を行ったことを示す。信号対雑音比 (SNR) は若年者で -2 dB、高齢者で 0 dB、残響時間は若年者で R1 (Ra) と 2.6 s (Rb)、高齢者で 2.4 s (Ra) と 1.4 s (Rb) である。

(3) 音響分析

Q、N、R1、R2 条件下で録音した 4 名の音声のうち、ターゲット部の音響分析を行った。分析に用いた特徴量は、基本周波数 (F0)、第 1 フォルマント (F1)、第 2 フォルマント (F2)、発話レベル、子音対母音のインテンシティ比 (CVR)、時間長である。

(4) 聴取実験

実験参加者は若年者 32 名（男性 4 名、女性 28 名、平均年齢 23 才）と高齢者 32 名（男性 12 名、女性 20 名、平均年齢 75 才）であり、全員が日本語母語話者である。実験は参加者毎に防音室内で行った。刺激はコンピュータにオーディオインタフェースを介したヘッドホンから提示し、刺激の提示レベルは参加者の聞きやすいレベルに調整した。各試行では刺激を一度提示し、参加者は聞こえたターゲット語を回答用紙に記入した。

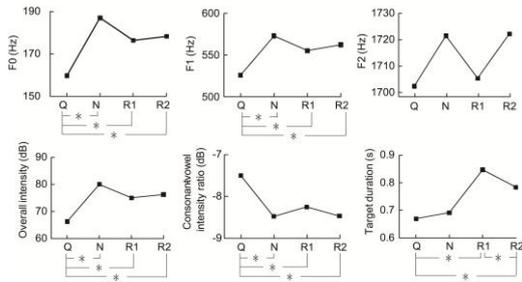


図2 各発話条件 (Q, N, R1, R2) におけるターゲットの音響分析 (F0, F1, F2, 発話レベル, CVR, 時間長) の結果。アスタリスクは5%水準の有意差を示す。

4. 研究成果

(1) 音響分析

図2に発話者4名の音響分析の結果を示す。結果より、以下の点が示された。

- ・残響下の発話は、静かな環境での発話と比べて音響的特徴が有意に異なった (F0, F1, 発話レベル、時間長の増加、CVRの減少)。
- ・音響的特徴の変化は雑音下と残響下では同一ではなく、発話の強調という面で同様のもの (F0, F1, 発話レベル、CVR) と、マスキングの違いにより異なるもの (時間長) がみられた。
- ・時間長のみ、R1とR2間で有意差がみられた。
- ・発話者間の音響的特徴の変動は、雑音下よりも残響下で増加した。これは、定常的な雑音よりも、非定常かつ反射音による遅延が存在する残響音に対しては、発話者の経験や意識によっても発話の適応方法が異なる可能性が示された。

(2) 音声知覚

図3, 4に各発話者 (S1, S2) 及び発話者の平均 (average) における、各聴取条件のモーラ正解率を示す。本研究での有意確率は5%とした。

雑音条件では、若年者・高齢者共に N_N は Q_N よりも正解率が有意に高くなった。この結果は先行研究[4]と一致するものであり、雑音下での発話は静かな場所での発話よりも、雑音下で聞いた時に明瞭度が高くなった。

残響条件では、若年者・高齢者共に R1_Ra は Q_Ra と QR1_Ra よりも正解率が有意に高く、R1_Rb は Q_Rb と QR1_Rb よりも正解率が有意に高くなった。これらの結果から、雑音と残響ではマスキングのメカニズムは異なるものの、雑音下に加え、残響下の発話も静かな環境での発話よりも明瞭になることが示された。よって残響音を聞きながら発話することにより、公共空間の音声案内がより明瞭になる可能性が示された。

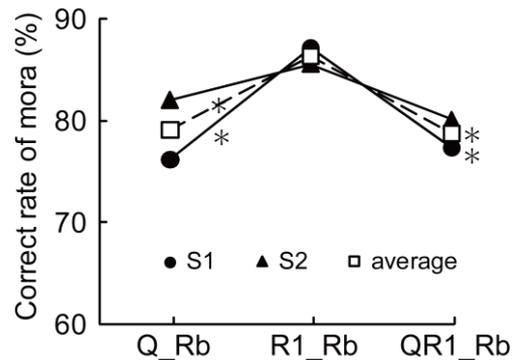
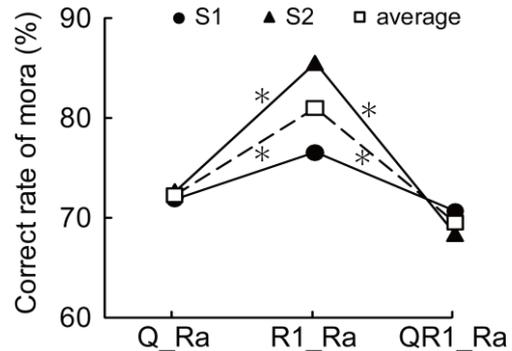
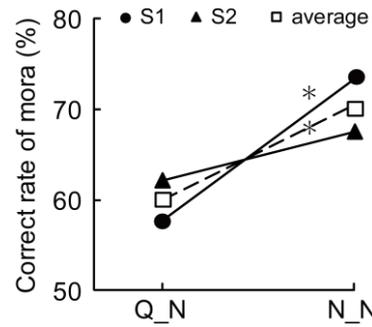


図3 各聴取条件 (上:N, 中央:Ra, 下:Rb)、各発話者 (S1:丸, S2:三角, 発話者の平均:四角) における若年者のモーラ正解率。SNR=-2 dB、残響時間はRaが3.6 s、Rbが2.6 s。アスタリスクは5%水準の有意差を示す。

また、発話と聴取が同じ残響環境 (マイク放送、若年者の R1_Ra 条件) のみならず、異なる場合 (自動放送、若年者の R1_Ra 以外の条件) においても、残響下の発話は静かな場所での発話よりも明瞭となることが示された。従って、音声案内を録音する際に使用する残響条件は、拡声する公共空間と同じである必要はないことも示された。

残響下の明瞭度は、先行する語句も残響下で発話をした方が明瞭になるのか調査した。若年者・高齢者共に Q_Ra と QR1_Ra 及び Q_Rb と QR1_Rb 間の正解率に有意差は認められなかった。このことから、ターゲット語が Q から R1 もしくは R2 に変化しても、ターゲット

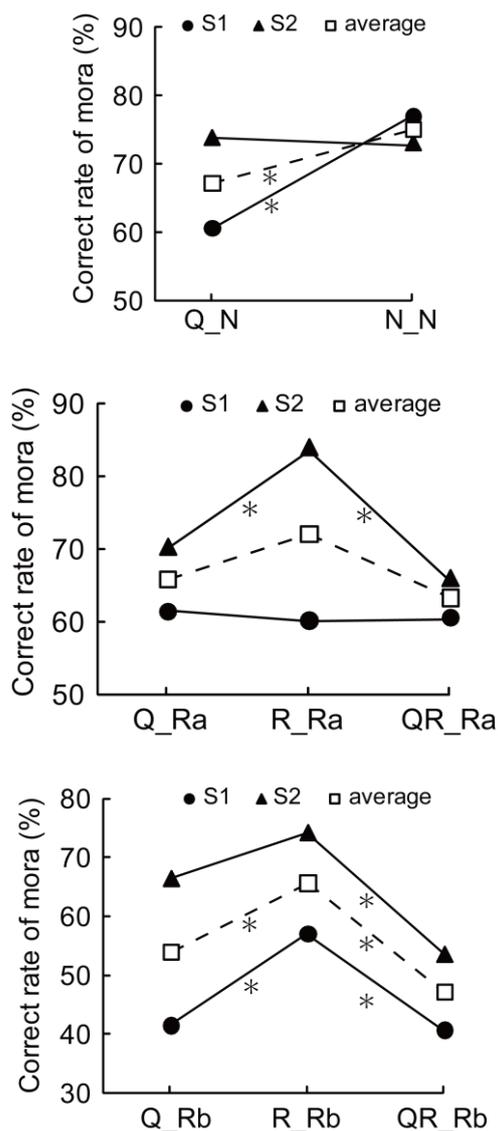


図4 各聴取条件(上:N, 中央:Ra, 下:Rb)、各発話者(S1:丸, S2:三角, 発話者の平均:四角)における高齢者のモーラ正解率。SNR=0 dB、残響時間はRaが2.4 s、Rbが1.4 s。アスタリスクは5%水準の有意差を示す。

語の了解度は有意に改善しないことが示された。一方、R1_RaとR1_RbはQR1_RaとQR1_Rbよりも各々正解率が有意に高くなった。つまり、ターゲットに先行する語句がQからR1/R2に変化すると、ターゲット語の了解度は向上することが示された。これは、4(1)章で述べた通り、残響下の発話は静かな環境での発話と比べて音響的特徴量が変わるため、先行するキャリア文から残響下で発話することで、残響によるマスキングが減少したと考えられる。

以上より、静かな環境での発話（即ち、通常の方法で作成された音声案内）と比べて、

残響音を聞きながら発話をすることで、公共空間で若年者と高齢者により明瞭な音声案内を提供することができる可能性が示された。その際、以下の点を示された。

- ・音声案内の作成に使用する残響条件は、拡声する公共空間と必ずしも同じである必要はなく、実用性を考慮すると有効であることが示された。
- ・自動放送とマイク放送を模擬した環境では、両環境において、提案手法により音声明瞭度は改善した。
- ・若年者と高齢者の両参加者群に対して、提案手法により音声明瞭度は改善した。ただし、有意差が得られる発話者や聴取条件は異なった。

(3) 今後の展望等

公共空間における聞き取りやすい音声の研究としては、これまでは建築音響的もしくは電気音響的なアプローチが主流であり、音声生成や知覚をベースとして音声案内自体を編集する手法は検討されてこなかった。本研究は、音響学、建築音響、電気音響、心理学、福祉などの分野を横断した学際的な研究を行い、「音声によるバリアフリー」の実現を目指した。

今後の検討事項としては、提案手法が非母語話者や聴覚障害者に対しても有効であるかの調査、発話者を増加させて発話者に独立である明瞭な音響的特徴の調査、公共空間の音声聴取により近い環境における提案手法の評価を行いたい。本研究の結果を受け、音声案内が明瞭となるような録音/合成環境の構築や、職員への発話訓練に応用することができると思われる。さらに、公共空間での音声伝達分野において、建築面におけるバリアフリー化と同様、音声のバリアフリー化（バリアフリー・ユニバーサルデザインのガイドライン化）が考えられる。これらの応用によって、若年者だけではなく残響の影響をより大きく受ける高齢者、聴覚障害者、非母語話者の音声明瞭度を改善することで、より多くの利用者にとって公共空間が安全で快適となることが期待される。

参考文献

- [1] A. K. Nablek and P. K. Robinson, "Monaural and binaural speech perception in reverberation for listeners of various ages," *J. Acoust. Soc. Am.*, 71, 1242-1248, 1982.
- [2] 内閣府, "高齢社会白書." 2012.
- [3] 総務局 "平成17年国勢調査 外国人に関する特別集計結果." 2008.
- [4] W. Van Summers *et al.*, "Effects of noise on speech production: Acoustics and perceptual analysis." *J. Acoust. Soc.*

- (2)研究分担者 なし
(3)連携研究者 なし

5. 主な発表論文等

(研究代表者、研究分担者及び連携研究者には下線)

[雑誌論文] (計 1 件)

- ① 程島奈緒, “「音声のバリアフリー」に向けて.” 東海大学情報通信学部紀要, 査読無, 3(2), 57-61, 2010.

[学会発表] (計 8 件)

- ① Nao Hodoshima and Takayuki Arai, “Making speech announcements intelligible in public spaces from a speech production view.” Acoustics 2012, 2012年5月15日, 香港(招待講演).
- ② 程島奈緒, 荒井隆行, 栗栖清浩, “高齢者に明瞭な拡声音声のための雑音・残響下の発話.” 日本音響学会秋季研究発表会, 2011年9月21日, 島根.
- ③ 程島奈緒, 荒井隆行, 栗栖清浩, “雑音・残響環境下の拡声音声に適した発話に関する検討.” 音声研究会, 2011年6月24日, 愛知.
- ④ 程島奈緒, 荒井隆行, 栗栖清浩, “雑音・残響下の発話による音声の明瞭度改善.” 日本音響学会秋季研究発表会, 2010年9月14日, 大阪.
- ⑤ Takayuki Arai and Nao Hodoshima, “Enhanced speech yielding higher intelligibility for all listeners and environments.” Interspeech, 2010年9月29日, 千葉.
- ⑥ Nao Hodoshima, Takayuki Arai and Kiyohiro Kurisu, “Intelligibility of speech spoken in noise and reverberation.” International Congress on Acoustics, 2010年8月23日, シドニー.
- ⑦ 程島奈緒, 荒井隆行, 栗栖清浩, “雑音・残響下における発話の音響的特徴の話者変動.” 音声研究会, 2009年11月26日, 静岡.
- ⑧ 程島奈緒, 荒井隆行, 栗栖清浩, “雑音・残響が発話に与える影響.” 日本音響学会秋季研究発表会, 2009年9月16日, 福島.

[その他]

ホームページ

<http://www.dm.u-tokai.ac.jp/~hodoshima/>

6. 研究組織

(1)研究代表者

程島 奈緒 (HODOSHIMA NAO)