

令和 6 年 4 月 30 日現在

機関番号：14501

研究種目：基盤研究(C)（一般）

研究期間：2021～2023

課題番号：21K04371

研究課題名（和文）言語バリアフリーと時間短縮を両立する多言語一斉放送システムの高度化に関する研究

研究課題名（英文）Study on advancement of multilingual simultaneous broadcast system that achieves both language barrier-free and time-saving

研究代表者

森本 政之（Morimoto, Masayuki）

神戸大学・工学研究科・名誉教授

研究者番号：10110800

交付決定額（研究期間全体）：（直接経費） 1,800,000円

研究成果の概要（和文）：本研究の目的は、言語バリアフリーと時間短縮を両立可能な多言語一斉放送システムの性能を、スピーカ配置等のパラメータの最適化により向上させることである。ターゲット音声として日本語単語、妨害音声として中国語、英語、韓国語の単語を用いた単語理解度試験を行い、以下を明らかにした。(1) 同時再生する言語数が2あるいは3の場合は最適化しなくても実用可能性があるが、4の場合はスピーカの分散配置、重要な情報を放送の後半に配置するといった方法による性能向上が必要である。(2) 周囲の騒音よりもターゲット音声を10dB程度大きくすることが望ましい。(3) 同時再生する言語数が3になると残響音の影響が大きくなる。

研究成果の学術的意義や社会的意義

従来、放送の重畳はタブーとされてきたため本研究で開発しようとするシステムが学術的に検討された事例は少ないが、複数の言語が重なって聞こえたとしてもスピーカ配置などを最適化すれば実用化の可能性があり、本研究により実際にその可能性が示された。このシステムが実用化できれば、日本語以外を第一言語とする人々まで含めた安全・安心な環境構築の一助となることが期待される。

研究成果の概要（英文）：The purpose of this study is to advance the performance of a multilingual simultaneous broadcast system that can achieve both language barrier-free and time-saving by optimizing parameters such as loudspeaker arrangement. We conducted word intelligibility tests using Japanese words as target speech and Chinese, English, and Korean words as disturbing speech, and found the following. (1) When the number of languages broadcast simultaneously is 2 or 3, there is a possibility of practical use without optimization, while when the number is 4, it is necessary to improve the performance by using multiple loudspeakers and placing important information in the latter part of the broadcast message. (2) It is recommended that the target speech be about 10 dB louder than the background noise. (3) When the number of languages broadcast simultaneously increases to 3, the effect of reverberation becomes significant.

研究分野：環境音響学

キーワード：言語バリアフリー 案内放送

1. 研究開始当初の背景

我が国の経済発展において、観光事業は重要な成長分野である。その成長を促進するために、特に大都市や観光地の公共空間では多言語対応が進んでおり、案内放送については、同内容の放送が異なる言語で繰り返されている。しかし、この方法では情報伝達の時間効率が極めて悪く、後から第一言語が放送された人々が逃げ遅れたり、空間の喧騒感が増したりする可能性がある。この問題を解決するために、研究代表者らは同時に多言語で放送するシステム（多言語一斉放送システム）について検討し、その実現可能性を示した（科研費 18H01597）。

2. 研究の目的

多言語一斉放送システムについて、スピーカ配置等のパラメータの最適化によって放送音声の了解度を向上させることにより、雑音や残響音といった妨害音がある音場にも適用できるように高度化することを本研究の最終的な目的とする。

3. 研究の方法

ターゲット音声として日本語のなじみの程度が高い単語、妨害音声として中国語、英語、韓国語の単語を用いる。これらの音声を聴取者に様々な条件で同時に提示し、聴取者がターゲット単語を正しく書き取れた割合（単語了解度）で評価を行う。図 1 に音声刺激の時間パターンを示す。連続した 5 つの単語（以下、5 連単語）を提示し、あらかじめ指定した時間的位置（1 番目～5 番目）の単語 1 つを回答するように指示した。外国語の単語は日本語単語と同じ長さになるように、単語を途中でフェードアウトして打ち切るか、複数の単語を連結した。単語の提示音圧レベルは全ての言語で同じとなるように揃えた。聴取者は正常聴力を持つ 20 歳代の大学生とした。

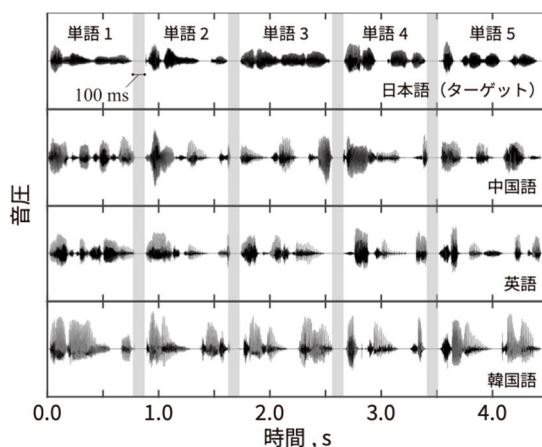


図 1 音声刺激の時間パターン

4. 研究成果

(1) テーマ 1：スピーカ配置による了解度の向上

図 2 に本研究で用いたスピーカ配置を示す。2 スピーカシステムは、それぞれのスピーカから 1 言語ずつ再生する 2 言語同時再生の条件も設定可能であったが、この条件については既往の研究（科研費 18H01597）で 90% 程度の高い単語了解度が得られることが分かっているため、本研究では検討対象に含めなかった。ターゲット音声は女性話者、妨害音声は男性話者とした。回答する単語の時間的位置は 1 番目、3 番目、5 番目の 3 条件とした。聴取者数は条件によって異なるが、6 から 8 名であった。

図 3 に実験結果を示す。左図は 1 番目、3 番目、5 番目の時間的位置の単語了解度の平均値である。2 言語同時再生（2-1）では 91.3% と高い単語了解度が得られた。3 言語同時再生の場合、単語了解度は 1 スピーカシステム（3-1）では 82.9% 程度とやや低いですが、3 スピーカシステム（3-3）では 90.6% と 2 言語同時再生と同程度まで上昇した。4 言語同時再生の場合、単語了解度は 1 スピーカシステム（4-1）では 66.0% と低かったが、2 スピーカシステムでは 79.1%、4 スピーカシステムでは 80.4% まで単語了解度が上昇した。2 スピーカと 4 スピーカシステムではほぼ差は見られず、スピーカを 4 つに増やす利点は確認できなかった。右図はターゲット単語の時間的位置ごとに算出した単語了解度である。2 言語同時再生（2-1）と 3 言語同時再生（3-3）ではターゲット単語の時間的位置の影響は見られなかったが、4 言語同時再生（4-2 と 4-4 をまとめて算出）では時間的位置が後ろになるほど単語了解度は上昇し、5 番目では 84.3% であった。

以上より、ターゲット音声と妨害音声を異なる空間的位置から再生することで、多言語一斉放送システムの性能を向上可能であることが確認された。また、4 言語同時再生の場合は、音声再生されてから時間が経過するほど了解度が上昇することが期待されるため、重要な情報を放送文の冒頭付近に配置しないなどの運用を行うことで、性能を実質的に向上可能であることが確認された。

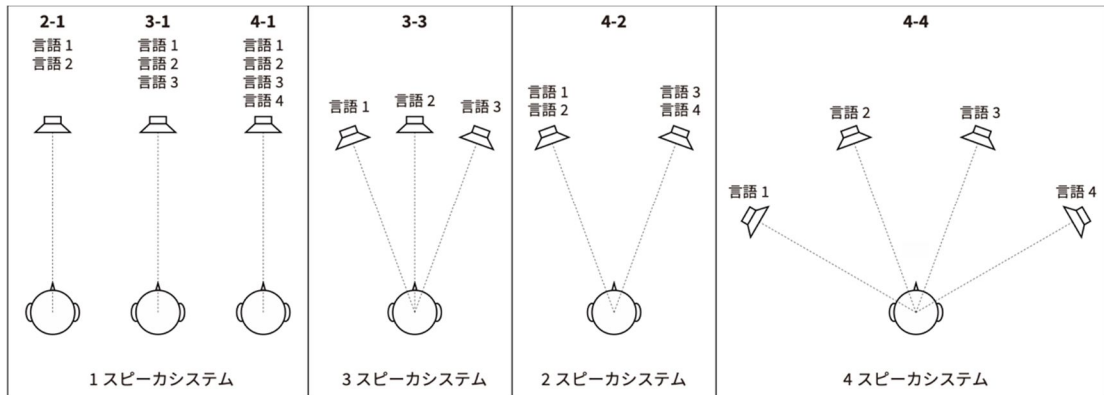


図2 テーマ1：実験に用いたスピーカ配置の模式図

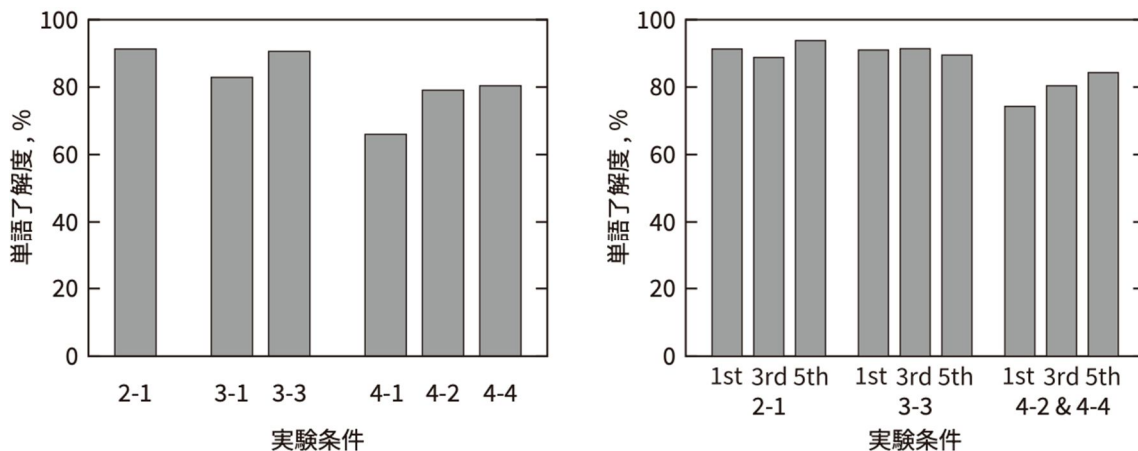


図3 テーマ1：スピーカシステムとターゲット単語の時間的位置が単語了解度に及ぼす影響

(2) テーマ2：視覚サインによる了解度の向上

図3の右図からも分かるように、様々な妨害音とターゲット音声を同時に聴取する場合、ターゲット音声聞き分けられるようになるまで、ある程度時間を要することが知られている。テーマ2では視覚サインを用いてターゲット音声の到来方向を事前に聴取者に伝えることで、最初に提示される単語の了解度が向上するかについて検討した。

VRゴーグルを用いて、複数のスピーカを模したオブジェクトを配置した仮想空間を視覚提示した。また、頭部伝達関数を試験用音源に畳み込むことによって、それらのオブジェクトから再生された音声をシミュレートして単語了解度試験を行った。図1に示すような5連単語ではなく、1単語のみを刺激として用いた。ターゲット音声と妨害音声はいずれも男性話者とした。聞きとった単語の回答は仮想空間内に表示されるキーボードを用いてタイピングさせた。図4に聴取者に視覚提示した仮想空間を示す。3言語同時再生を対象とし、1スピーカシステム(3-1)と3スピーカシステム(3-3)を用いた。また、全体から広がって聞こえるように高域減衰型の周波数特性を持つ定常雑音も付加した。音声および定常雑音はヘッドホンを用いて提示した。定常雑音の提示レベルはターゲット音声よりも7.5dB小さい値とした。図4のオブジェクトの上方に表示されている橙色の球が、ターゲット音声の到来方向を伝える視覚サインである。視覚サインを提示する条件では、音声提示の3秒前から視覚サインを表示し、その方向に顔を向けるように聴取者に指示した。実空間における聴取者の頭部運動は、VRゴーグルによるヘッドトラッキング機能により仮想空間内でも再現され、視覚および聴覚情報がリアルタイムで頭部運動に応じて更新される。聴取者数は12名であった。

学習効果の影響を相殺するために、視覚サインありの実験を先に実施する聴取者群(N=4)と視覚サインありの実験を後に実施する聴取者群(N=8)ごとに単語了解度を算出し、2群の単語了解度を平均した結果を図5に示す。なお、単語了解度の算出にあたり、条件によらず正答あ

るいは不正答となった単語（合計で 120 単語中 22 単語）は集計から除いた。

単語理解度は 45.8% から 52.9% であり、テーマ 1 の 3-1 および 3-3（図 3 左図参照）と比較して大きく低下した。テーマ 2 の実験では、ターゲット音声と妨害音声の話者の性別が等しいこと、暗騒音を付加したこと、シミュレートした音源をヘッドホン提示したこと、などが単語理解度の低下の原因と考えられる。ただし、単語理解度は 3-1 と比較して 3-3 の方が、視覚サインなしでは 7.1 ポイント、視覚サインありでは 3.6 ポイント上昇しており、テーマ 1 と同様の結果となったことから、シミュレーションについてはある程度の妥当性があったと考えられる。視覚サインの有無については、本研究の範囲では単語理解度の差は明確ではなかった。単語理解度が 50% 程度と厳しい聴取条件であったこと、1 単語のみの提示であったことなどの理由により、大脳中枢におけるマスキング解除が働きにくかったことが原因として考えられるが、これらの条件を変えた場合の検討については今後の課題とする。

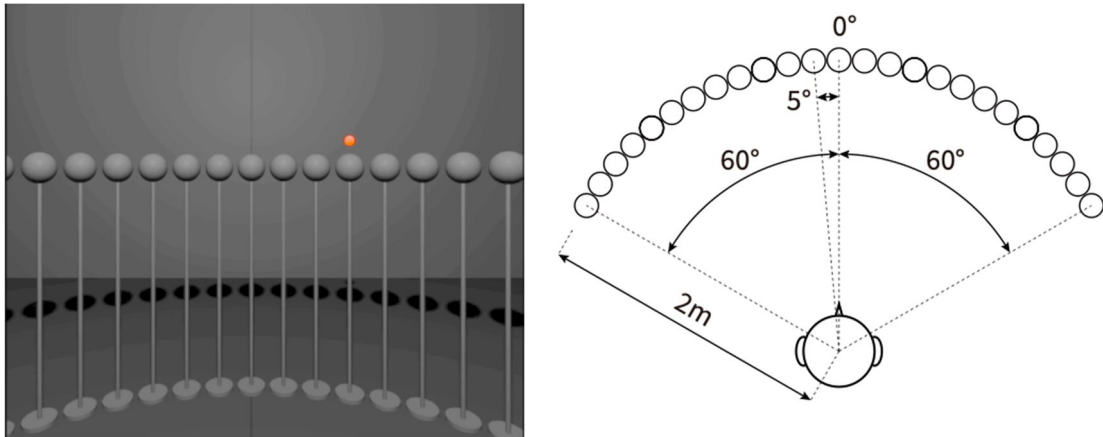


図 4 テーマ 2：実験に用いた仮想空間

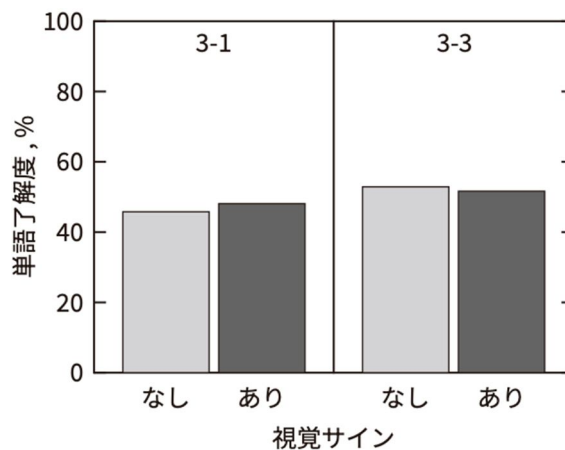


図 5 テーマ 2：視覚サインが単語理解度に及ぼす影響

(3) テーマ 3：雑音および残響音の影響

多言語一斉放送システムは、公共空間での運用を想定しているため、雑音と残響音による妨害の影響がどの程度であるかを把握しておく必要がある。テーマ 3 では 1 スピーカシステムを対象として暗騒音と残響音の影響を検討した。

雑音は公共空間において一般的な高域減衰型の周波数特性を持つ定常雑音とし、ターゲット音声とのレベル差（SN 比）を実験パラメータとした。雑音は聴取者から見て左 135°、左 20°、正面、右 20°、右 135° の 5 つのスピーカを用いて提示した。正面のスピーカはターゲット音声および妨害音声の再生にも用いた。したがって、雑音付加の条件では、音声は正面のスピーカから、雑音は全体から広がって聞こえる。残響音はホワイトノイズを指数減衰させることにより作成したインパルス応答を音声に畳み込むことで作成した。インパルス応答の指数減衰の傾きを変えることにより音声理解度の物理指標である C50 を制御して実験パラメータとした。残響音はターゲット音声および妨害音声に由来するものとし、雑音の場合と同じ 5 つのスピーカを用いて全体から広がって聞こえるように提示した。したがって、残響音付加の条件では、音声はターゲット音声と妨害音声のいずれも全体から広がって聞こえる。ターゲット音声および

び妨害音声はテーマ 1 と同様に 5 連単語とし、ターゲット音声は女性話者、妨害音声は男性話者とした。回答する単語の時間的位置は 1 番目、と 5 番目の 2 条件とした。聴取者数は雑音を付加した条件は 17 名、残響音を付加した条件は 3 名であった。

C50 はインパルス応答の初期部分 (~50ms) の総エネルギーを音声の聴取に有効なエネルギー、それ以降の部分の総エネルギーを有害なエネルギーとし、その比率をレベル表示した値である。つまり、SN 比と C50 は、妨害音として雑音と残響音を取り扱う点で異なるが、音声聴取に有効なエネルギーと有害なエネルギーの比率をレベル表示した値という点は共通である。ただし、多言語一斉放送システムの場合、妨害音声のエネルギーを SN 比および C50 における音声の聴取に有害なエネルギーとして加算する必要がある。例えば、雑音と残響音の両者を付加しない条件でも、2 言語同時の場合の SN 比は 0dB (ターゲット音声と妨害音声のエネルギーが等しい) となる。以下ではその加算を行った場合の SN 比および C50 を分析に用いる。

図 7 にテーマ 3 の実験結果を示す。横軸は SN 比あるいは C50 であり、妨害音声のエネルギーを妨害音のエネルギーに加算して求めた値である。全体として SN 比あるいは C50 が低下するほど単語了解度は低下した。また、SN 比と C50 が同程度の条件で比較した場合、2 言語同時 (2-1) の場合は雑音付加条件と残響音付加条件で大きな単語了解度の差は見られないが、3 言語同時 (3-1) の場合は、特に時間的位置が 1 番目の場合に残響音付加条件の方が明らかに単語了解度は低かった。

雑音付加条件の場合は妨害音声のエネルギーを考慮した SN 比で概ね単語了解度は予測可能であり、SN 比が低くなるほど単語了解度は低下する。したがって、周囲の暗騒音の音圧レベルは、妨害音声のエネルギーを考慮した SN 比が最大 (2 言語同時の場合 0dB, 3 言語同時の場合 -3dB) になるような値、具体的にはターゲット音声よりも 10dB 程度低い音圧レベル以下であることが望ましい。一方、残響音の影響は言語数が多くなると雑音の影響よりも大きくなると考えられる。これは、ターゲット音声を他の妨害音声から聞き分ける際に、音声波形の時間変動の情報が重要であることを示唆する。ただし、2 言語同時再生であれば残響音の影響は雑音の影響と同程度に抑えられるため、3 言語同時再生よりも妨害音に対して頑健であると考えられる。

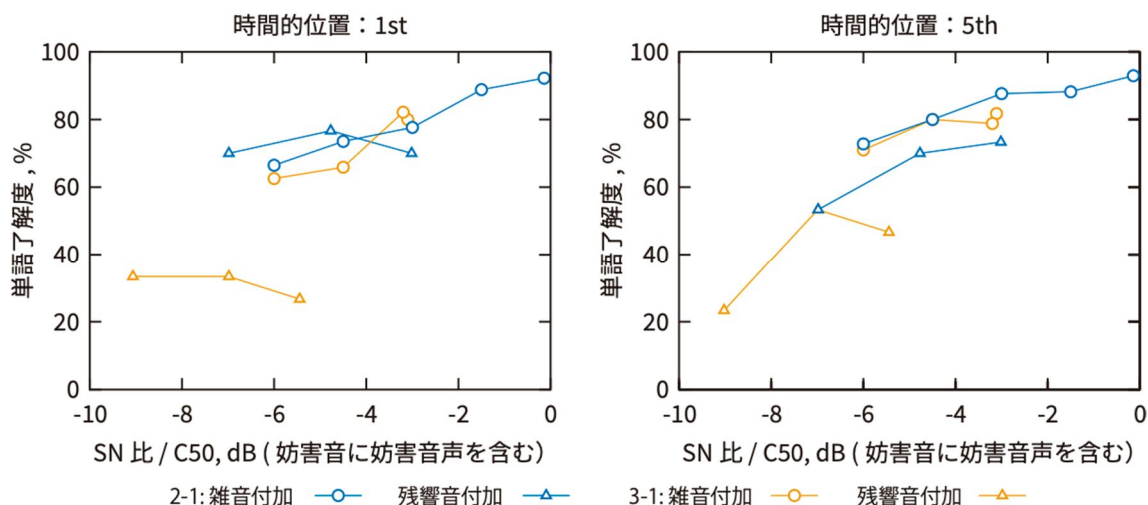


図 6 テーマ 3: 雑音および残響音が単語了解度に及ぼす影響

5. 主な発表論文等

〔雑誌論文〕 計0件

〔学会発表〕 計6件（うち招待講演 1件 / うち国際学会 1件）

1. 発表者名 Hayato Sato, Masayuki Morimoto, Hiroshi Sato
2. 発表標題 How many languages can be broadcast simultaneously as an emergency announcement?
3. 学会等名 INTER-NOISE and NOISE-CON Congress and Conference (招待講演) (国際学会)
4. 発表年 2023年

1. 発表者名 上羽 希佳, 佐藤 逸人, 森本 政之
2. 発表標題 多言語同時拡声による案内放送の了解度にSN比と言語数が及ぼす影響
3. 学会等名 日本建築学会大会
4. 発表年 2023年

1. 発表者名 上羽 希佳, 佐藤 逸人, 森本 政之
2. 発表標題 多言語同時拡声による案内放送の了解度に騒音が及ぼす影響
3. 学会等名 日本建築学会近畿支部研究発表会
4. 発表年 2023年

1. 発表者名 佐藤 逸人, 穴水 智也, 森本 政之, 佐藤 洋
2. 発表標題 目的音声の時間的位置と空間的位置に着目した多言語同時拡声による案内放送の最適化
3. 学会等名 日本音響学会建築音響研究会
4. 発表年 2022年

1. 発表者名 森本 政之, 佐藤 逸人, 穴水 智也, 佐藤 洋
2. 発表標題 多言語同時拡声による案内放送の了解度(1) 言語数の影響
3. 学会等名 日本音響学会秋季研究発表会
4. 発表年 2022年

1. 発表者名 佐藤 逸人, 穴水 智也, 森本 政之, 佐藤 洋
2. 発表標題 多言語同時拡声による案内放送の了解度(2) 目的単語の時間的位置と空間的位置の影響
3. 学会等名 日本音響学会秋季研究発表会
4. 発表年 2022年

〔図書〕 計0件

〔産業財産権〕

〔その他〕

-

6. 研究組織

	氏名 (ローマ字氏名) (研究者番号)	所属研究機関・部局・職 (機関番号)	備考
研究 分 担 者	佐藤 逸人 (Sato Hayato) (30346233)	神戸大学・工学研究科・准教授 (14501)	

7. 科研費を使用して開催した国際研究集会

〔国際研究集会〕 計0件

8. 本研究に関連して実施した国際共同研究の実施状況

共同研究相手国	相手方研究機関
---------	---------