

科学研究費助成事業 研究成果報告書

令和 5 年 10 月 26 日現在

機関番号：32503

研究種目：基盤研究(C)（一般）

研究期間：2019～2022

課題番号：19K04927

研究課題名（和文）音声知覚特性を考慮した聞き取りやすく情報伝達に優れた放送方式に関する研究

研究課題名（英文）A study on a broadcasting method that is easy to understand and good transmitting information in consideration of speech perception characteristics

研究代表者

世木 秀明（SEKI, Hideaki）

千葉工業大学・情報科学部・准教授

研究者番号：60226636

交付決定額（研究期間全体）：（直接経費） 3,300,000円

研究成果の概要（和文）：聞き取りやすく情報伝達に優れた放送方式についてソフトウェア面から検討を行い、以下のことが明らかになった。

放送文には、文脈や聴取状況に応じたキーワード単語の使用が重要であり、伝えたい情報を想起しやすい知識情報として事前に提示することが理解度向上に繋がり、聞き取りやすく感じさせることも可能である。また、放送音声に加え、適切な文字やイラストなどの視覚情報の提示が放送文理解に役立つ。さらに、多くの高齢者や聴覚障害者では、背景雑音が放送文の聞き取りや理解に大きな影響を与えるが、句読点部分のポーズ時間の適切な設定や 2kHz～4kHz帯域の強調が聞き取りやすさ改善に大きく貢献する。

研究成果の学術的意義や社会的意義

公共空間では、平常時の明瞭な音声情報伝達はもとより、災害などの非常時には的確な情報伝達が重要である。従来から聞き取りやすい公共放送を行うための音環境条件や拡声システムの特性・性能などに関するハードウェア面に関する多くの検討が行われているが、聞き取りやすいアナウンス方法などに関するソフトウェア面からの検討はほとんど無い。

このような背景のもと、本研究では、若年健聴者だけでなく高齢者も含めた被験者を対象に聴取実験を行い、聞き取りやすく情報伝達に優れた公共放送や非常放送を行うためのソフトウェア面からの多くの知見を得ることができた。これらの知見は、学術的・社会的に大きな意義を持つと考える。

研究成果の概要（英文）：We studied the software side of a broadcasting method that is easy to understand and good transmitting information, and the following were shown.

(1) It is important to use keyword words in the message according to the contexts and listening situations. (2) Presenting the information to be given in the message as knowledge information that is easy to recall in advance leads to better comprehension and makes the message seem easier to listen. In addition, (3) the presentation of visual information such as appropriate text and illustrations in addition to the speech is useful for understanding the message. Furthermore, for many elderly and hearing-impaired listeners, (4) background noise greatly affects their ability to hear and understand broadcasts. Therefore, (5) Appropriate setting of pause time for punctuation parts and (6) emphasizing the bandwidth range of 2 kHz to 4 kHz will greatly contribute to improving ease of listening for the elderly and hearing impaired.

研究分野：複合領域

キーワード：公共放送 非常放送 聞き取りやすさ 理解度 心的辞書 高齢者

1. 研究開始当初の背景

公共空間では、平常時の明瞭な音声情報伝達はもとより、災害などの非常時には警報・避難誘導情報の確かな伝達が重要である。特に、東日本大震災時や北海道胆振東部地震時では、一層その重要性が認識された。このような問題を受け、公共空間において明瞭で確かな情報伝達を実現するために各種公共空間の音環境条件の改善、拡声システムの特長・性能改善などのハードウェア面に関する多くの検討や提案が行われている。しかし、若年健聴者だけでなく加齢による聴力低下や音声知覚特性が若年健聴者と異なっていると考えられる高齢者や聴覚障害者に対してどのようなアナウンス音声やアナウンス方法が聞き取りやすく情報伝達に優れているのかに関するソフトウェア面からの検討はほとんど無い。

このため、いくら聞き取りやすい音環境をハードウェア面から構築したとしても聴取者や聴取環境に適したアナウンス音声やアナウンス方法で情報提供が行わなければ十分な情報伝達は行えないと考えられる。特に高齢者や聴覚障害者に対しては聴覚特性を考慮したアナウンスが必要であると考えられる。

2. 研究の目的

本研究では、聞き取りやすく情報伝達に優れた公共放送や非常放送を行う方法についてソフトウェア面からの提案を行うために①放送文に使用される単語の音声単語親密度と理解度の関係、②知識情報の利用が可能な場合と困難な場合の理解度の違い、③視覚情報を付加した場合と付加しない場合の理解度の違い、④話速や句読点に対するポーズ時間が変化した場合の理解度の違い、⑤アナウンス音声に帯域強調を施した場合の聞き取りやすさや了解度に与える影響などについて若年健聴者だけでなく高齢者や聴覚障害者も対象にした聴取実験を行い、聴取者の音声知覚過程を考慮に入れて検討する。

3. 研究の方法

(1)放送文に使用される単語の音声単語親密度が放送文理解に与える影響に関する検討

音声合成プログラムを利用して通常の発話音声と緊急性を感じると考えられる発話音声の合成パラメータにより合成音声を作成し、これを使用した聴取実験を行い、放送文に使用される単語の音声単語親密度が放送文理解に与える影響に関して放送を聴取する環境や状況を考慮に入れて検討する。

(2)知識情報が放送文理解に与える影響に関する検討

放送文の前半で伝えたい内容を想起しやすい知識情報を与えることで放送文で伝えたい内容の理解が向上するのか、またその放送文を聞き取りやすく感じるのかについて聴取実験を行い検討する。

(3)視覚情報を併用した場合の理解度向上に関する検討

雑音環境下で提示する音声情報に加え、音声情報の理解を助けると考えられる文字やイラストを視覚情報として提示した場合、放送文の理解度向上に有効であるかについて聴取実験を行い検討する。

(4)話速や句読点に対するポーズ時間が文章理解に与える影響に関する検討

高齢者など聴力が低下した人に対して分かりやすい話し方は、ゆっくり話すということが有効な手段の一つであることが広く知られている。この理由として、ゆっくり話すことと構音動作がしっかり行えるため明瞭な発話が期待出来るだけでなく、母音の持続時間も長くなるので聞き取りやすくなるという利点がある。

しかし、ひとまとまりになっている単語をゆっくり話すことと逆に分かりにくくなるという意見に加え、ゆっくり話すよりも句読点に対するポーズ時間を長くするほうが分かりやすいという考え方や多くの高齢者は文章理解処理にかかる時間も 20 代の若年健聴者と比べ長くなっているという考えかたもある。このため、若年健聴者において聞き取りに問題のない話速で発話した場合でも、聴力の低下した高齢者や聴覚障害者にとっては聞き取りにくいことも予想される。

このような背景を踏まえ、話速やポーズ時間が文章理解に与える影響について聴取実験を行い検討する。

(5)放送音声の帯域強調が聞き取りやすさや了解度に与える影響に関する検討

オペラ歌手や NHK アナウンサーなど遠くまで明瞭に聞こえる音声や雑音下でも聞き取りやすい車掌のアナウンス音声は 2~4kHz 帯域の音声エネルギーが高くなる傾向が見られる。この現象はシンギングフォルマント (Singing Formant) として知られている。

このことから、シンギングフォルマントのような音響特性をアナウンス音声に付加することにより、雑音環境において放送内容の理解度や聞き取りやすさが向上するのかについて聴取実験を行い検討する。

4. 研究成果

(1)放送文に使用される単語の音声単語親密度が放送文理解に与える影響に関する検討

予備実験として緊急性を感じさせる音声の話速と基本周波数、提示音圧に関して音声合成プログラム VoiceText を利用して作成した合成音声を用いて調査したところ、最も緊急性を感じる話速は、約 9.0 モーラ/秒、基本周波数は男声の場合では約 165Hz、女声の場合では約 253Hz であると考えられた。また、提示音圧は緊急性を感じさせる合成パラメータとしての関与は低いと考えられた。この結果を考慮に入れて聴取実験に使用する実験用音声刺激を作成した。実験用音声刺激に使用する刺激材料は、非常放送文約 100 文を Web を利用して収集し、これから 20 文を選定した。これに加え、選定した文を理解するために重要

と考えられる単語をターゲット単語として選び、これと同じ意味を持ち、音声単語親密度が高い単語に置き換えた 26 文、合計 46 文を非常放送文とした。さらに非常放送文で使用したターゲット単語を用いて平常放送文を 86 文作成し、合計 132 文章を刺激材料とした。以下に刺激材料例を示す。

[刺激材料例]

刺激材料(選定した放送文)

- ・非常放送文:2 階の広場で、火災が発生しました。
 - ・平常放送文:冬は火災が起きやすい季節です。
- *下線で示す単語が文理解に重要と考えられる単語(ターゲット単語)

ターゲット単語を高親密度単語に置き換えた放送文

- ・非常放送文:2 階の広場で、火事が発生しました。
 - ・平常放送文:冬は火事が起きやすい季節です。
- *音声単語親密度:火災:5.688 火事:5.822

聴取実験に用いた実験用音声刺激は、音声合成プログラム VoiceText により女声の標準的な話速と基本周波数であると考えられる合成パラメータ(話速 7.5 モーラ/秒, 基本周波数 230Hz)と予備実験で得られた最も緊急性を感じやすいと考えられた話速 9.0 モーラ/秒, 基本周波数 253Hz で刺激材料を読み上げ、これらに同一ラウドネスレベルのマルチトーカーノイズを重畳させたものとした。聴取実験は、音声刺激を静かな部屋で被験者前方に設置したスピーカより至適レベル(約 70dB(A))で提示し、聞こえたとおりに筆記で回答させた。聴取実験結果の集計は、ターゲット単語の正答率を求めた。被験者は健康な聴力を持つ 20 代男女 17 名である。

図 1-1 に示す聴取実験結果から、平常放送文では音声単語親密度が高い単語で作成された音声刺激の正答率が高いが、非常放送文では非常放送文に用いられている音声単語親密度が低い単語を用いた音声刺激の方が有意に高い正答率を示した($p < 0.01$)。この結果から、音声単語親密度は聴取時に非常放送と感じる場合と平常放送とを感じる場合では異なっているのではないかと考えられた。

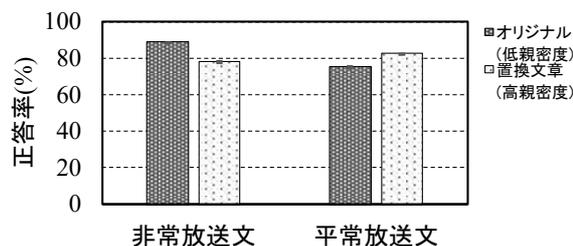


図 1-1 聴取実験結果

さらに、刺激材料を女声の標準的な話速と基本

周波数で読み上げた実験用音声刺激 132 文を用いた聴取実験を健康な聴力を持つ 20 代男女 14 名を対象に行った結果、非常放送文、平常放送文ともに音声単語親密度が高い単語を用いた実験用音声刺激の方が低い単語を用いた実験用音声刺激と比べ、高い正答率を示すことが観測された。

これらの結果から、同じ単語であっても非常放送と感じる場合と平常放送とを感じる場合では、音声単語親密度が変化する可能性があると考えられた。これより、聴取状況に応じたキーワード単語を使用することが聞き取りやすくわかりやすい放送を行うための一要因になると考えられた。

(2)知識情報が放送文理解に与える影響に関する検討

聴取実験で用いた刺激材料は、2 文による構成とし、第 1 文に第 2 文の内容を想起しやすい知識情報を含んだ文章、第 2 文に伝えたい内容を含んだ文章を組み合わせた 8 文章と第 1 文に第 2 文の知識情報を含まない文章、第 2 文に伝えたい内容を含んだ文章を組み合わせた 8 文章、合計 16 文章である。作成した刺激材料例を以下に示す。

[刺激材料の例]

知識情報あり文章

本日は、乾燥および強風のため、火災が発生しやすくなっております。火の取り扱いには十分ご注意ください。

知識情報なし文章

地域住民の皆様には、火の取り扱いには十分ご注意ください。

作成した刺激材料を音声合成プログラム Voice Text の男声により約 6.5 モーラ/秒、基本周波数 150Hz で読み上げた音声刺激材料を作成し、これと同一ラウドネスレベルのマルチトーカーノイズを重畳させたものを実験用音声刺激とした。聴取実験は、静かな部屋で至適レベル(約 70dB(A))で被験者前方に設置したスピーカにより実験用音声刺激を提示した。被験者には、放送文の第 2 文目の内容に関する質問に筆記で回答させた。さらに、提示した実験用音声刺激の聞き取りにくさを下記に示す 4 段階で評価させた。

[聞き取りにくさの 4 段階評価]

- 1.聞き取りにくくはない 2.やや、聞き取りにくい 3.かなり聞き取りにくい 4.非常に聞き取りにくい

被験者は、健康な聴力を持つ 20 代男女 25 名と加齢による聴力低下以外に聴力に異常を感じていない 65 歳から 77 歳の高齢者(平均年齢 72.7±3.9 歳)9 名で、4 分法による高齢者の平均聴力レベルは、16.9dB であった。

聴取実験の結果、20 代男女(若年健聴者)では、第 1 文に知識情報を含む実験用音声刺激と含まない実験用音声刺激の正答率を比較すると前者の正答率が有意に高くなることが観測された($p < 0.05$)。さらに、聞き取りにくさの評価でも第 1 文に知識情報を含む音声刺激を有意に「聞き取りにくくはない」と答える割合が多いことが観測された($p < 0.05$)。一方、高齢者では、図 2-1 に示すように正答率間に有意差は見られないものの若年健聴者を対象とした聴取実験結果と同様の傾向が観測された。また、

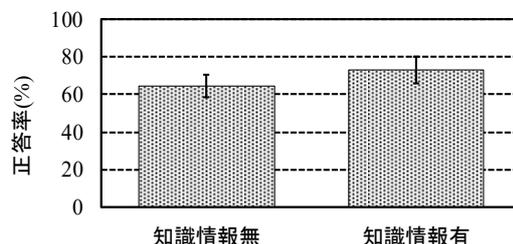


図 2-1 高齢者 9 名の聴取実験結果

聞き取りにくさの評価でも若年健聴者と同様に第 1 文に知識情報を含む実験用音声刺激を「聞き取りにくくはない」と評価する傾向が見られた。

これらの結果から、伝えたい内容が想起しやすい知識情報を事前に与えることで放送文の理解向上に加え、聞き取りやすい放送として感じさせることが可能であると考えられた。また、高齢者の実験結果で有意な差がみられなかった理由として、被験者数が少なかったことに加え、高齢者にとって雑音そのものが音声聴取に影響を与えており、これが主な原因となり放送文の理解が困難になっていることも考えられた。

(3)視覚情報を併用した場合の理解度向上に関する検討

聴取実験に使用する実験用音声刺激として公共施設などで使用されている放送文を参考に作成した 30 文章を音声合成プログラム VoiceText の男声により約 6.5 モーラ/秒、基本周波数 150Hz で読み上げたものに同一ラウドネスレベルのマルチトーカーノイズを重畳したものとした。さらに、放送文内容の理解を助けると考えられるイラストとキーワード文字を視覚刺激として用意した。また、高齢者を被験者とした聴取実験では、若年健聴者に対して行った実験用音声刺激のみの聴取実験で正答率が高かった 5 文章と音声にイラストを付与することで正答率に大幅な向上が認められた 5 文章の音声刺激材料に対してラウドネスレベルを 6dB 減衰させたマルチトーカーノイズを重畳したものを高齢者用の実験用音声刺激とした。

聴取実験は、静かな部屋で実験用音声刺激と同時に視覚刺激の提示後、被験者に放送文理解を調べる質問の答えと聞き取りにくさの 4 段階評価を筆記により回答させた。ここで、実験用音声刺激は至適レベル(約 70dB(A))で被験者前方 1.2m に設置したスピーカから提示し、視覚刺激は被験者前方 70cm に設置した 25"ディスプレイの中央にキーワード文字は 200 ポイントで白地に MS-ゴシック体の黒文字、イラストは 1,000 ドット×600 ドットで実験用音声刺激提示時間と同一になるように提示した。被験者は健康な聴力と視力を持つ 20 代男女(若年健聴者) 43 名と 65 歳から 77 歳の高齢者(平均年齢 72.7±3.9 歳) 9 名で、4 分法による高齢者の平均聴力レベルは、16.9dB であった。

若年健聴者を対象とした聴取実験結果から、実験用音声刺激のみの場合に比べ、キーワード文字やイラストを音声刺激と共に提示した場合の質問に対する正答率は有意な向上が見られたが($p < 0.01$)、視覚刺激の違いによる正答率間には有意差は認められなかった。また、聞き取りにくさの評価では、実験用音声刺激のみの場合に比べ、視覚刺激を提示した場合は「非常に聞き取りにくい」と評価する割合が低下し、「聞き取りにくくはない」、「やや聞き取りにくい」と評価する割合に大きな増加がみられた。

一方、高齢者の聴取実験結果では、図 3-1 に示すように若年健聴者と同様に音声刺激に加え視覚刺激を提示した場合の方が有意な正答率向上が観測された($p < 0.05$)。さらに、聞き取りにくさの評価では、視覚刺激を提示することで「非常に聞き取りにくい」という評価が大幅に減少し、「やや聞き取りにくい」という評価が大幅に増加することが観測された。しかし、若年健聴者で増加が見られた「聞き取りにくくはない」という評価にはほとんど変化が見られなかった。

以上の聴取実験結果より、放送文の理解を助けるキーワード文字やイラストなどの視覚情報の提示が放送文理解に大きく貢献することが確認された。さらに、高齢者では聞き取りにくさの評価結果から、背景雑音があること自体が放送文の聴取に大きな影響を与えていることも示唆された。

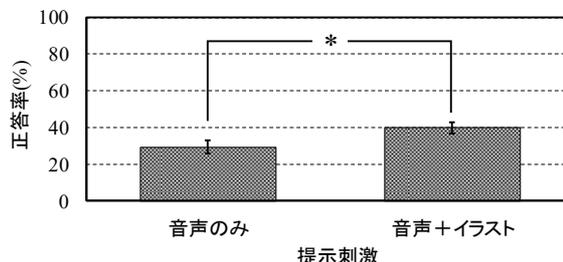


図 3-1 高齢者の聴取実験結果

(4)話速や句読点に対するポーズ時間が文章理解に与える影響に関する検討

一般に話速は発話された文章のモーラ数を発話時間で割った値が用いられているが、文章中に含まれる句読点に対するポーズ時間は考慮されていない。このため、図 4-1 に示すようにオリジナル音声の話速を遅くする場合、発話時間全体を伸張する方法と句読点部分のポーズ時間のみを伸張する方法が考えられるが、いずれも話速は同一となる。

また、文中の句読点に対応するポーズ時間はポーズ以前の音声内容を理解し、次の音声を聴取するための準備時間として使用されていると考えられている。このため、ポーズ時間が極端に短くなると文章理解の低下につながることも考えられる。

聴取実験に使用する音声刺激材料として第 1 文と第 2 文の間に意味的関係がある 13 文章を女性アナウンサーが約 6 モーラ/秒で読み上げた音声を図 4-1 に示すようにオリジナル音声に対して発話時間

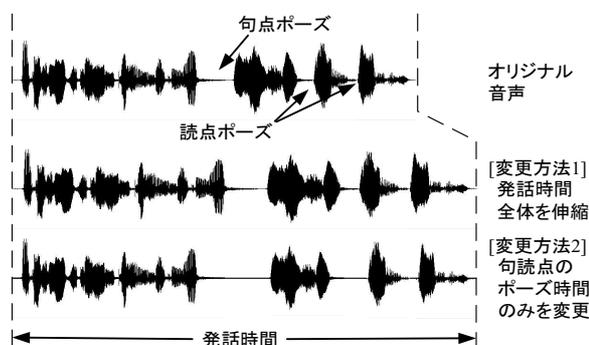


図 4-1 話速の変更方法

全体を伸縮することで 0.8 倍、1.2 倍、1.4 倍としたものと句読点に対するポーズ時間のみを変化させて発話時間を 0.8 倍、1.2 倍、1.4 倍に変化させたものを用意した。また、聴取実験で使用した実験用音声刺激は、若年健聴者には音声刺激材料に対してラウドネスレベルを 6dB 減衰させたマルチトーカーノイズを重畳したものと、高齢者には、若年健聴者の聴取実験結果で成績の良かった 8 文章の音声刺激材料に対してラウドネスレベ

ルを 12dB 減衰させたマルチトーカーノイズを重畳したものとした。被験者は、健康な聴力をもつ 20 代男女(若年健聴者)15 名と加齢による聴力低下以外に異常を感じていない 65 歳から 77 歳の高齢者 9 名(平均年齢 72.7±3.9 歳)で、4 分法による高齢者の平均聴力レベルは、16.9dB であった。聴取実験方法は、実験用音声刺激を被験者に至適レベル(約 70dB(A))で提示し、その内容に関する簡単な質問に筆記で回答させた。さらに、4 段階の聞き取りにくさの評価を行わせた。

聴取実験の結果、若年健聴者では話速変化方法によらず実験用音声刺激の内容に関する質問の正答率に有意な差は見られなかった。一方、高齢者では図 4-2 に示すように有意な差は見られないものの、発話時間が 1.2 倍以上ではポーズ時間のみを変更した実験用音声刺激の正答率の方が高くなる傾向が見られた。また、発話時間が 0.8 倍時では、発話時間全体を伸縮させた実験用音声刺激よりもポーズ時間を変更した実験用音声刺激の方が句読点に対するポーズ時間が短くなるのが原因で正答率が低下していると考えられた。

これらの結果から、若年健聴者では話速やポーズ時間が文章理解にほとんど影響を与えないが、高齢者では、句読点に対応するポーズ時間が文章理解処理に影響を与えるのではないかと考えられた。

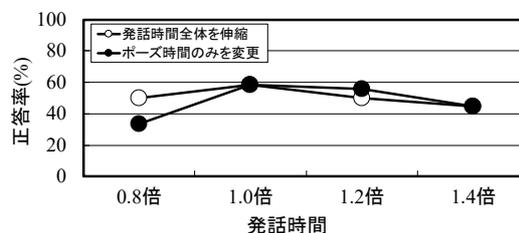


図 4-2 高齢者の質問に対する正答率

(5)放送音声の帯域強調が聞き取りやすさや了解度に与える影響に関する検討

若年健聴者を対象とした聴取実験 1 と若年健聴者に加え高齢者と聴覚障害者を対象とした聴取実験 2 を行い検討した。聴取実験 1 に使用する刺激材料として 4 連語、5 連語の文章を Voice Text の男声および、女声により読み上げた 60 文章の 2~4kHz の帯域を 0dB, +3dB, +6dB, +9dB 強調したものを用意した。ここで、刺激材料の話速は約 6.5 モーラ/秒、基本周波数は男声 150Hz, 女声 200Hz とした。聴取実験 1 に使用した文章例を以下に示す。

[文章例]

- ・4 連語:今朝、太郎が家で勉強をしました。
- ・5 連語:今朝、太郎が家で英語の勉強をしました。

*下線部分がターゲット単語

聴取実験 1 で使用した実験用音声刺激は、刺激材料に対してラウドネスレベルを 6dB 減じたマルチトーカーノイズを重畳させたものとした。聴取実験方法は、実験用音声刺激を静かな部屋で被験者前方に設置したスピーカから至適レベル(約 70dB(A))で提示し、聞こえたとおりに筆記で回答させた。集計は、ターゲット単語の正答率を求めた。さらに、4 段階の聞き取りにくさの調査を行った。被験者は聴力が健康な 20 代健聴者男女(若年健聴者) 10 名である。

聴取実験 1 から、帯域強調を大きくするほど正答率の上昇が観測された。帯域強調レベルの違いによる正答率の差について検定した結果、9dB 強調群とそれ以外の群間の正答率に有意な差が観測された ($p<0.01$)。また聞き取りにくさの評価も同様に帯域強調を大きくするほど評価値が低下する傾向が観測された。この結果から、2~4kHz の帯域を 9dB 程度強調した放送音声は、雑音環境下で聞き取りやすく理解度が向上すると考えられた。この結果を元に、若年健聴者 20 名に加え、加齢による聴力低下以外に聴力に異常を感じていない 67 歳から 76 歳の高齢者 10 名(平均年齢 72.4±3.8 歳)と聴覚障害者 1 名(22 歳)を対象とした聴取実験 2 を行った。ここで、高齢者の 4 分法による平均聴力レベルは、29.3dB であった。また、聴覚障害者 1 名の 4 分法による平均聴力レベルは 41.6dB であり、本人は両耳に補聴器を装着し、日常生活で聞こえに不自由を感じていないと答えている。

若年健聴者に対する聴取実験 2 で使用した実験用音声刺激は、聴取実験 1 で使用した刺激材料と 2~4kHz の帯域を 9dB 強調した 120 音声に対して同一ラウドネスレベルでマルチトーカーノイズを重畳させたものとした。一方、高齢者および、聴覚障害者に対しては聴取実験 1 で使用した実験用音声刺激の中で正答率が低いものから高いものまでを均等に男声、女声各 18 音声ずつ選定した音声とその音声に 9dB の帯域強調を行った音声の合計 72 音声に対してラウドネスレベルを 6dB 減じたマルチトーカーノイズを重畳させたものを実験用音声刺激とした。聴取実験方法は、聴取実験 1 と同一である。図 5-1 に示す聴取実験の結果から、若年健聴者では、女声、男声共に帯域強調を行った実験用音声刺激の方が有意に正答率が高くなることが観測された ($p<0.05$)。また、高齢者と聴覚障害者においても帯域強調による正答率の向上が観測された。これらの結果より、2~4kHz の帯域に対して 9dB 程度の強調を行うことは若年健聴者、高齢者、聴覚障害者を問わず聞き取りやすさの向上に有効であると考えられた。しかし、高齢者や聴覚障害者では、各人の聴力特性が大きく異なると考えられることから、帯域強調だけでなく話速など、他の要因にも左右される場合もあるのではないかと考えられた。

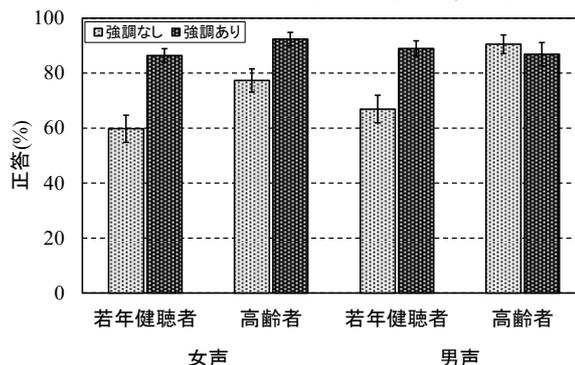


図 5-1 若年健聴者と高齢者の聴取実験結果

5. 主な発表論文等

〔雑誌論文〕 計0件

〔学会発表〕 計0件

〔図書〕 計1件

1. 著者名 竹内京子、稲田朋晃、金村久美、中村健太郎、峰松信明、世木秀明	4. 発行年 2023年
2. 出版社 メジカルビュー社	5. 総ページ数 152
3. 書名 : Crosslink 言語聴覚士テ キスト 音響・音声学	

〔産業財産権〕

〔その他〕

世木研究室 研究室紹介(研究内容) https://www.seki.cs.it-chiba.ac.jp/02.html
--

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
---------------------------	-----------------------	----

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
---------	---------