Vol.2021-SLP-138 No.15
2021/10/19

音声の聞き取りやすさに関わる性別による音響特徴の違い
－ATR 音声データベースを用いた分析－

小林 まおり †　　倉片 憲治 ‡

早稲田大学人間科学学術院　〒359-1192 埼玉県所沢市三ヶ島 2-579-15

E-mail:　† maori-k@aoni.waseda.jp,　‡ kurakata.kenji@waseda.jp

あらまし　一般に，高齢者にとって女声は男声よりも聞き取りにくいとされている．しかし，話者の性別が聞き取りやすさに及ぼす影響については，特に日本語を用いた研究事例が少なく，聴取者の年齢や外乱も考慮した系統だてた検証が求められている．本研究では，その第１段階として ATR 音声データベースに収録された日本人話者の女声と男声の音響特徴を分析し，比較を行った．明瞭性・了解性に関連する音響特徴の性差に着目するため，１モーラあたりの発話長，変調スペクトル，基本周波数，フォルマント周波数，子音長，および子音部の COG を算出し，性別間で比較した．その結果，発話長と振幅変調スペクトルには明確な性差は認められなかった．一方，基本周波数やフォルマント周波数，GOE は女声の方が高く，子音長は男声の方が長かった．しかし，これらの音響特徴において聞き取りやすさの違いに影響しうる明確な性差は見出せなかった．今後，聴取実験を行い，その結果を音声の音響特徴や聴覚の加齢特性等と対応づけることで，音声の性別による聞き取りやすさの違いとそのメカニズムを検討する必要があると考えられる．

キーワード　音声の性別，聞き取りやすさ，音響特徴，ATR 日本語音声データベース

Gender Differences in Acoustic Features of Voices Affecting Speech Intelligibility
－Analyses of the ATR Japanese Speech Database－

Maori KOBAYASHI †　and　Kenji KURAKATA ‡

† Faculty of Human Sciences, Waseda University　2-579-15 Mikajima, Tokorozawa, Saitama, 359-1192 Japan

E-mail:　† maori-k@aoni.waseda.jp,　‡ kenji.kurakata@waseda.jp

Abstract　It is said that older people have difficulty in hearing female speeches compared to male speeches in general. However, it has not been examined systematically, especially for the Japanese language, whether the gender of voice has any effects on its intelligibility. In this study, we analyzed acoustic features of male and female speeches, as the first step, for investigating gender effects on speech intelligibility. Focusing on the influence on speech intelligibility, we extracted acoustic features such as the duration of mora, modulation spectrum, fundamental frequency (F0), formant frequency, the duration of consonant, and the center of gravity (COG) of constant. Results of analysis show that there were no significant differences in the duration of mora and modulation spectrum. The F0, formant frequency and COG were higher for female speeches whereas the male speeches had a longer consonant duration. However, we could not identify differences in acoustic features that could affect the intelligibility for older people. Psychoacoustic experiments should be conducted to examine the gender effects on speech intelligibility and underlying mechanism, in terms of age-related changes of hearing.

Keywords　Gender Effect，Speech Intelligibility，Aging Effect，Acoustic Features，the ATR Japanese Speech Database

1. はじめに

一度に不特定多数の人に情報を伝えられる音声アナウンスは，公共施設や商業施設などで必要不可欠な伝達手法である．雑音や残響といった外乱に対して頑健で，誰にとっても聞き取りやすい音声アナウンスが求められている．

聞き取りやすい音声とはどのような音声か．難聴者や高齢者など聞き取りが困難な人に向けた発話スタイルであるクリアスピーチ（clear speech）や，騒音環境下での発話スタイルのロンバード音声（Lombard speech），またアナウンサーや俳優など音声発話を職業にしている話者の発話スタイルは，聞き取りやすい音声として，よく知られている[1-10]．実際，これらの音声は，通常の会話音声に比べて，雑音下でも

ⓒ 2021 Information Processing Society of Japan

10~20 %程度了解度が高い[1-4]．こうした高い了解性を持つ音声と通常の会話音声を比較することで，雑音下でも頑健な知覚を可能にする音響特徴は明らかにされつつある．最近では，これらの音声の音響特徴を反映した音声合成の技術開発も進められている[7,8].

その一方で，男性と女性のどちらの音声が聞き取りやすいのか，といった素朴で身近な疑問は残されたままである．音声の性別に関しては「女性らしさ」や「男性らしさ」といった音声が伝える話者の個人性や声質に関する研究は盛んに行われてきた．それらに比較すると聞き取りやすさにおける話者の性差の検証は十分とは言えない[9, 15]．それにもかかわらず，音韻知覚に関する聴取実験ではどちらかの性別のみを使用する例が多く（男声のみを使用する研究が多い），音韻の知覚メカニズムを検証するうえで，この偏りを問題視する研究者もいる[9].

英語音声では，男声の方が聞き取りやすい[9]，音声の性別で聞き取りやすさは変わらない[11]とする報告はあるものの，近年では男声よりも女声が聞き取りやすいという知見が優勢である[12-14]．この知見をもとに公共空間の音声アナウンスが選択されることもある．しかし，なぜ女声の方が了解性が高いのか，その理由について明らかにされてはいない[9,15]．現在は，音声の音響特徴における性差の他に，言語社会学など多角的視点からの検証が進んでいる．また，外乱や聴取者の年齢の影響について十分に解明されているとはいえず，今後，メカニズムも含めてさらなる研究の発展が期待される．

日本語音声に関しては研究例自体が少なく，今のところ統一的な見解は認められない[16,17]．系統だった明瞭度・了解度試験による比較が十分に行われていないのが現状である．より聞き取りやすい音声を目指すうえでも，音声の性別の影響について，雑音や残響といった外乱要因や聴取者の年齢，聴力の観点から検証することが求められている．

そうした状況であるにもかかわらず，高齢者は「女声」が聞き取りにくいといったことが一般には指摘されている[18,19]．高齢者施設などでは女性職員に指導がなされることもある．この理由として，高齢者は高周波数帯域の聴力が低下することため，声の高い女性の声が聞き取りづらくなることが挙げられている[18,19]．しかし，声の高さに影響する基本周波数(F0)は大抵の場合 500 Hz 以下であり，加齢による聴力低下の影響を受けにくい低い帯域である．加えて，女声と男声で長時間スペクトラムはほとんど変わらない[20]．これらのことを踏まえると，単純に「声が高いために」女声を高齢者が聞き取りづらくなるとは考えにくい．

ただし，女声では基本周波数が高いために，音韻の手がかりとなる音響特徴が，男声よりも高周波数帯域にシフトすることが考えられる．しかし，音韻の手がかりとなる音響特徴が性別によってどの程度変化するのか，特に日本語音声については不明である．そこで，本項では，音声の性別による聞き取りやすさの違いを明らかにする検証の第 1 段階として，女声と男声の音響特徴を抽出し，比較を行った．

表 1. 分析対象音声の発話者の内訳　（人）

		女声	男声
年代	10 代	6	3
	20 代	17	20
	30 代	4	8
	40 代以上	8	3
	不明	2	3
出身地	東北・上越	2	2
	関東・甲信	10	6
	中部	4	3
	関西	14	18
	中国・四国	3	1
	九州	2	3
	不明	2	4

2. 分析対象音声

ATR 音声データベース C セット[23]を用いた．このデータベースには多数話者による同一発話内容の音声データが揃えられており，本研究の目的に適していると考えた．C セットに収録されている音声のうち，音素情報が付されている発話者男性 37 名，女性 37 名の音声を利用した．発話者の年齢や出身地の内訳を表 1 に示す．C セットには音素バランス文 150 文，重要語 520 語，音韻バランス単語 260 語が収録されているが，各音響特徴の分析で適した単語をそれぞれ用いた．各音声はサンプリング周波数 20 kHz，量子化ビット数 16 bit であった．

3. 音響分析

以下の音響特徴をそれぞれの性別で算出した．音響特徴については，先行研究[9,12-15,22]で音声の性別による違いがあると報告されているもの，従来の研究で音韻の手がかりとされている特徴[23-25]や了解性に影響するとされている音響特徴[1-8]の中から選択した．以下に，特定の調音法・調音位置に拠らない特徴と各調音法・調音位置の特徴に分けて示す．また，雑音下での音声明瞭度の推測値として，SII(Speech Intelligibility Index[26])を算出した．

3.1. 分節に渡る音響特徴

3.1.1. 発話時間長

クリアスピーチやロンバード音声では，通常会話に

比べて話速が遅いことが報告されている[1]．そこで，文章および重要語 520 単語を対象に，1 モーラあたりの発話時間長を性別ごとに算出し，比較した．

3.1.2. 振幅変調スペクトル

音声の振幅包絡は音韻知覚にとって重要であることは知られている．特に，変調周波数が 16 Hz 以下の変調周波数成分が音声了解度に影響する[23]．クリアスピーチでは，通常の会話音声に比べて低い変調周波数帯域の変調度が大きいことが知られている[3]．そこで，重要語 520 単語を対象に変調スペクトルを算出した．その際，全帯域とオクターブ帯域ごとに変調スペクトルを求めた．

変調スペクトルは下記の手続きで算出した．まず，各音声信号をバンドパスフィルタを用いて帯域分割する．次に，各帯域の音声信号の振幅包絡を Hilbert 変換によって算出する．そして，各帯域の振幅包絡線を離散フーリエ変換して，変調スペクトルを求めた．

3.1.3. 基本周波数

クリアスピーチやロンバード音声では会話音声に比べて平均基本周波数（F0）やその変動量が大きいことが報告されている[1-8]．また，女声と男声の違いとして，平均 F0 や F0 の変動量（最大 F0−最小 F0）の違いが報告されている[12-14]．

F0 は声帯振動を反映したものであり，一般に音韻知覚に直接影響することはないとされている．しかし，F0 の高さがフォルマントの形成に影響する可能性があることから，今回改めて測定することとした．

本項では，音素バランス単語 216 語を対象に F0 を抽出した．各単語の母音部分の F0 から，単語に平均基本周波数，および F0 変動量を算出した．また，先行研究の比較として，語頭に/a/がある単語を対象にして，/a/部の変動量および平均 F0 を併せて算出した．なお，F0 の抽出には STRAIGHT[27]を用いた．

3.2. 各調音･調音位置に関する音響特徴

3.2.1. 母音

母音の音響特徴としてフォルマント周波数を算出した．重要語 520 単語のうち語頭が/i/，/e/，/a/いずれかの母音の語に限定した．その当該母音部の時間中心点での第 1～第 4 フォルマント周波数（f1~f4）を算出した．フォルマント周波数の算出には Praat (ver. 6.1.53)を用いた．また，高母音・中母音・低母音の違いは f1 から F0 を引いた値と関連することから[24]，各母音で算出した．

3.2.2. 破裂音

破裂音の特徴として，破裂部分の時間長と重心（COG: Center of Gravity）を算出した．COG は破裂部分のスペクトル情報の指標とした．調音点によって声道長が変化することで，エネルギーが集中する帯域が異なるため，調音位置の手がかりとなる[25]．

重要語 520 単語のうち語頭が破裂音で，後続の母音が/a/のものを分析対象とした．単語リストには語頭が/p/の単語がなかったことから，/p/は分析対象から除外した．また，ラベル情報が破裂音以外を付されている単語も分析対象から除外した．破裂部分の時間長については，データに付されているセグメント情報を利用して算出した．COG については子音部が 30 ms 以上ある単語音声のみを対象とした．

3.2.3. 摩擦音

摩擦音の音響特徴として，破裂音と同様に子音部の発話長および COG を算出した．COG については，破裂音の分析と同様に，後続の母音は/a/に限定し，当該子音部が 30 ms 以下のものは分析対象から除外したところ，/s/,/z/のみとなった．

3.3. その他

3.3.1. SII (Speech Intelligibility Index)

SII は騒音下の音声了解・明瞭度の指標の一つであり，AI の後継と見なされている[28]．これは，音声信号とノイズをいくつかの周波数帯域に分け，それぞれの帯域での S/N と周波数帯域別の係数から帯域ごとの明瞭度･了解度指数を求め，それらの総和から明瞭度･了解度指数を求めるというものである．1 に近いほど明瞭度･了解度が高く，聞き取りやすさの目安になる．

本項では，音声信号に対して–5 dB のバブルノイズ[29]を騒音と仮定し，重要語 520 単語に対して SII を算出した．また分割帯域数は 32 とし，音声信号の提示レベルは 60 dB とした．

4. 結果

分析対象の音声から抽出した音響特徴について話者ごとに代表値を算出した．

4.1.1. 発話時間長

性別ごとに算出した 1 モーラあたりの平均発話時間長および SE を表 2 に示す．女性の方が 1 モーラあたりの発話長は長い傾向にあるが，両性別ともに文章では 150 ms 前後，単語では 190 ms 前後となり，性別に

表 2 文章および単語における 1 モーラあたりの発話長

(ms/mora)

	女声	男声
文章	151.5 (±4.45)	147.8 (±4.72)
重要語	193.5 (±5.02)	184.6 (±4.07)

(　)内は 95 %信頼区間を示す．

ⓒ 2021 Information Processing Society of Japan

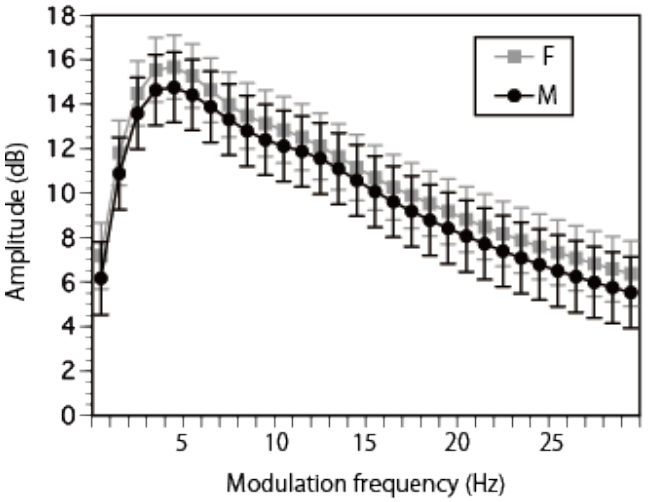


図 1 音声の性別による変調周波数の違い（エラーバーは 95%信頼区間を示す）

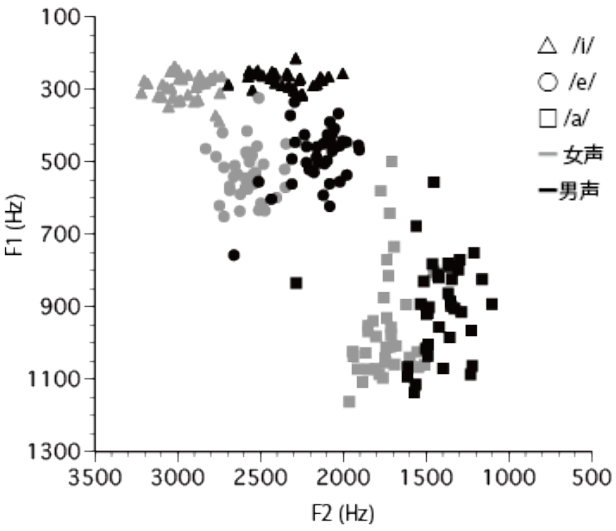


図 2 音声の性別によるフォルマントスペースの違い

よる大きな違いは認められなかった.

4.1.2. 振幅変調スペクトル

各性別の全帯域の平均変調スペクトルを図 1 に示す. オクターブ帯域ごとに変調スペクトルについては，全帯域と同様の傾向であったために割愛する．この図からわかるように，単語を分析対象とした場合，音声の性別によって変調スペクトルのパターンに違いは見られなかった．また，Amplitude の性差は最大でも 1 dB 程度であった.

4.1.3. 基本周波数

性別ごとに F0 と F0 変動量を表 3 に示す．先行研究と同様に, 女声の方が男声に比べて 110 Hz 程度高かった．また，変動量においても女声の方が大きく，先行研究と一致する結果であった[17]. ただし, 1 音節内での変動では線形スケールではやや女声の方が高かったが，対数スケールでみるとほぼ変わらなかった.

4.2. 各子音に関する音響特徴

4.2.1. 母音

表 3 平均 F0 および F0 変動量 (Hz)

	音響特徴	女声	男声
単語全体	平均 F0	252.0 (±8.0)	141.0 (±8.1)
	F0 変動量	146.3 (±11.8)	87.5 (±6.9)
/a/	平均 F0	240.1 (±16.4)	151.2 (±9.4)
	F0 変動量	27.5 (±2.6)	17.1 (±1.8)

( )内は 95 %信頼区間を示す

各話者につき，それぞれ母音の f1 および f2 にもとづき図 2 にプロットした. 横軸に f2, 縦軸に f1 を示す.

平均フォルマント周波数は従来研究とほぼ同等の値となった．また，母音空間は全体的に女声の方が広く，先行研究と一致していた[17].

次に, 各母音の f1 から基本周波数を引いたものを表 4 に示す．母音間の f1–F0 の違いは女声の方が大きかったことから，女声の方が母音間の差が強調されており，母音の違いを知覚しやすいと考えられる.

4.2.2. 破裂音

表 5 に破裂部分の時間長を示す．大半の破裂音において男声の方が子音長は長かった．1 モーラあたりの時間長は男声の方が短かったことから，CV 比は男声の方が大きいといえる.

次に，表 6 に破裂音の COG を示す．/t/と/k/の COG は男声よりも女声の方がそれぞれ高かった．/t/と/k/の COG の違いは女声の方が 100 Hz 程度大きかった.

4.2.3. 摩擦音

それぞれの子音の発話長を表 5 に, COG を表 6 に示す．発話長については，男声の方が 10 ms 程度長い傾向だった．また，/s/と/z/の発話長の差においても男声の方が大きいことがわかった.

表 4 各母音の f1 から F0 を引いた値 (Hz)

母音	女声	男声
/i/	62.8 (±12.4)	142.2 (±9.7)
/e/	227.5 (±15.7)	254.3 (±27.1)
/a/	74.1 (±48.4)	775.3 (±42.4)

( )内は 95 %信頼区間を示す

ⓒ 2021 Information Processing Society of Japan

表 5 各子音部の子音長 (ms)

子音		女声	男声
破裂音	/t/	25.6 (±2.5)	30.5 (±4.1)
	/k/	41.3 (±3.0)	49.0 (±4.5)
	/b/	11.1 (±2.0)	10.1 (±0.9)
	/d/	10.7 (±1.1)	12.4 (±1.1)
	/g/	14.4 (±1.0)	18.5 (±5.0)
摩擦音	/s/	110.4 (±5.1)	118.2 (±7.5)
	/z/	42.1 (±4.7)	54.6 (±5.3)

( )内は 95 %信頼区間を示す

表 6 各子音部での COG (Hz)

母音		女声	男声
破裂音	/t/	1829.3(±154.4)	1641.5 (±157.3)
	/k/	2022.5 (±94.2)	1759.7 (±108.9)
摩擦音	/s/	5817.4 (276.5)	5807.9 (±223.2)
	/z/	3925.3 (±677.7)	3320.1 (±536.5)

( )内は 95 %信頼区間を示す

どちらの性別においても，/k/の COG の方が/t/よりも高かった．この結果は先行研究[25]と一致する．また，どちらの子音においても，女声の COG は男声よりも 200 Hz 程度高かった．/t/と/k/の COG の差は女声の方が 80 Hz 程度大きかった．

4.3. その他

4.3.1. SII

SII は，女声 0.23 (95%信頼区間: ±0.01)，男声 0.24（±0.02）となった．健聴者を想定した SII では，性差は見られなかった．

5. 考察

本研究では，先行研究での知見をもとに，音韻の手がかりや了解性に影響する音響特徴に性差が見られるか，音響分析を行った．その結果，音響特徴によって性差の影響が異なることがわかった．

まず，女声と男声で発話長や変調スペクトルには大きな違いは認められなかった．話速に関しては先行研究でも変わらないという報告[17]があり，これらの結果と一致する．

次に，F0 やフォルマント周波数，COG は，女声の方が高い傾向が見られた．F0 やフォルマント周波数が女声の方が高いことは従来の知見と一致する．女声の方が了解性が高い要因として，母音空間が女声の方が広いことが考えられている[12, 17]．本研究においても，f1−F0 の値が母音間で大きく異なっており，女声の方が母音の違いを知覚しやすいことが予想される．また，F0 変動量も女声の方が大きかった．F0 自体が音韻の直接的な手がかりになるとは考えにくいが，F0 はアクセントを左右するため，変動量が大きいことによってアクセントの違いを知覚しやすいと考えられる．加えて，F0 が高いことによって COG も女声のほうが高かった．COG が重要な子音の手がかりかは議論の余地があるが[27]，この高周波数帯へのシフトが高齢者の聞き取りにとってどの程度影響するかは，今後検討する価値があるだろう．

子音部の時間長は，男声の方が長い傾向であった．1 モーラあたりの発話長に性差がないことを考慮すると，子音長対母音長比（CV 比）は男声の方が長いと推測できる．加えて，子音間の発話長の違いも男声の方が大きかった．破裂音や摩擦音については，子音の立ち上がりが長くなるのか，立ち下がりが長くなるかで異聴が変わるとされている[30]．そのため，子音部のどの部分の長さに性差があるか詳細に分析することで，聞き取りやすさへの影響を推測できるかもしれない．また，高齢者は時間分解能が低下すると言われている[31]．この解像度の低下と性別による子音長の違いが高齢者の聞き取りに影響するのかについても，さらに検証する必要がある．

また，SII については性差が見られず，了解度試験においても性差が見られないと推測される．しかし，SII は男声のみを使用した聴取実験の結果をもとに規格化されている．アメリカ英語音声を用いた研究では，SII の算出に用いるような周波数重要度関数の分布には性差があることが指摘されている[29]．そのため，そもそも女声に対応した指標ではないかもしれない．こうした SII 自体の問題点が結果に影響した可能性を否定できない．話者の性別が聞き取りやすさに及ぼす影響を検証することは，SII の問題解決につながると考えている．

最後に，今回の分析で扱った音響特徴は音韻に関与する音響特徴の一部でしかない．フォルマント遷移，アンチフォルマント，VOT など，検証すべき特徴は他にも多々ある．それらについても分析を行う必要があろう．今後は，音響分析の他に聴取実験を行い，異聴傾向などの結果と今回の音響特徴の対応について検討していく予定である．

6. まとめ

音声の性別が聞き取りやすさに及ぼす影響に関する検討の第 1 段階として，性別によって音声の音響特徴を比較した．それぞれの性別で 37 名の音声を対象と

ⓒ 2021 Information Processing Society of Japan

して，発話時間長，変調スペクトル，基本周波数，フォルマント周波数, COG, 子音長を算出し，比較した．その結果，発話時間長や変調スペクトルで性差は認められなかった．基本周波数やフォルマント周波数，COG については女声の方が高く，子音長は男声の方が長い傾向であった．

今後は，アンチフォルマント，フォルマント遷移，VOT などのその他の音響特徴についても検証する必要があろう．また，聴取者や外乱を操作した了解度実験を行い，その結果を音響特徴や聴覚の加齢特性の観点から対応づける．これらを通じて性別による聞き取りやすさの違いについて検討する予定である．

文　献

[1] R. M. Uchanski，Clear Speech, in the Handbook of speech perception, eds., D. B. Pisoni, R. P. Remez, pp. 207-235, Blackwell Publishing, MA, USA, 2005.

[2] M. A. Picheny, N. I. Durlach, L.D. Braida, “Speaking clearly for hard of hearing I : Intelligibility differences between clear and conversational speech,” J. Speech Hear. Res. 28, 96–103, 1985.

[3] J. Krause, L. D. Braida, “Acoustic properties of naturally produced clear speech at normal speaking rate,” J. Acoust. Soc. Am., 115, 362–378, 2004.

[4] M. Garmier, N. H. Bernardoni, “Spearing in noise: How does the Lombard effect improve acoustic contrasts between speech and ambient noise? Comput. Speech Lang., 28, 580–597, 2014.

[5] A. Castellano, J.M. Benedi, F. Casacubera, “An analysis of general acoustic-phonetic features for speech produced with the Lombard effect, Speech Commun., 20, 23–35, 2000.

[6] J. Junqua, “The Lombard reflex and its role on human listener and automatic speech recognizers, J. Acoust. Soc. Am., 93, 510–524, 1993.

[7] Y. Lu, M. Cooke, “The contribution of changes in F0 and spectral tilt to increased intelligibility of speech produced in noise,” Sp. Commun., 51, 1253–1262, 2009.

[8] T. Ngo, M. Akagi, P. Birkholtz, “Effects of articulatory and acoustic features on the intelligibility of speech in noise,” Sp. Commun., 117, 13–20, 2020.

[9] S. E. Yoho, A. B. Stephanie, T. S. Barrett, D. B. Whittaker, “Are there sex effects for speech intelligibility in American English? Examining the influence of talker, listener, and methodology,” Atten. Percept. Psychophys, 81, 558–570, 2019.

[10] D. R. McCloy, R. A. Wright, P. E. Souza, “Talker versus dialect effects on speech intelligibility: A symmetrical study,” Lang. Speech, 58, 371–386, 2015.

[11] R. W. Gengel, G. L. Kupperman, “Word discrimination in noise: Effect of different speakers,” Ear Hear, 1, 156–160, 1980.

[12] A.R. Bradlow, G. M. Torretta, D. B. Pisoni, “Intelligibility of normal sppech I: Global and fine-grained acoustic-phonetic talker characteristics,” Sp. Commun., 20, 255–272, 1996.

[13] D. Markham, V. Hazan, “The effect of talker-and listener-related factors on intelligibility for a real-word, open-set perception test,” J. Speech, Lang. Hear. Res., 47, 725–737, 2004.

[14] H. B. Known, “Gender difference in speech intelligibility using speech intelligibility test and acoustic analysis”, J. Adv Prosthodont, 2, 71–76, 2010.

[15] A.P. Simpson, “Phonetic differences between male and female speech,” Lang. Linguistics Compass, 3, 621–640, 2009.

[16] 関根, 稲葉, 森田, “高齢ドライバの聴覚情報獲得に関する基礎調査（第 2 報）−音声情報に対するラジオ聴取の影響について−, ” 自動車技術会論文集, 44, 149–155, 2013.

[17] 成田ら，”高齢者を対象とした合成音声の聞き取りやすさに関する研究,” 生活安全支援医療福祉工学系連合大会発表論文集, O1-3, 1–4, 2011.

[18] 平松類，老人の取扱説明書，SB クリエイティブ，2017.

[19] 亀井亜希子，高齢者・障害者へのマナーがわかる本，セルバ出版，2020.

[20] R. W. Benson, I.J. Hirsh, “Some variables in audio spectrometry,” J. Acoust. Soc. Am., 11, 278-288, 1940.

[21] 匂坂芳典，浦谷則好, “ATR 音声・言語データベース,” 日本音響学会誌，48, 12, 878-882, 1992.

[22] D. Byrd, “Relations of sex and dialect to reduction,” Speech Commun., 15, 39–54, 1994.

[23] R. Drullman, J. M. Fasten, R. Plomp, “Effect of temporal envelope smearing on speech perception,” J. Acoust. Soc. Am., 95, 1053–1064, 1994.

[24] R. Diel, K. Kluender, “On the objects of speech perception,” Ecol. Psychol., 1, 121–144, 1989.

[25] K. Stevens, B. Blumstein, S. “Invariant cues for phase of articulation in stop consonants,” J. Acoust. Soc. Am., 64, 1538–1368, 1978.

[26] ANSI S3.5-1997: American National Standard Methods for the Standards Institute, New York.

[27] H. Kawahara, I. Masuda-Katsuse, A. De Cheveigne, “Restructuring speech representations using pitch-adaptive time frequency smoothing and an instantaneous-frequency-based F0 extraction,” Speech Commun., 3, 187–207, 1999.

[28] 翁長博, 池田哲郎, 音声の聴取レベルが了解度に及ぼす影響に関する考察−Speech intelligibility index (SII)との対応について，日本建築学会環境系論文集, 74, 642, 931–936, 2009.

[29] A. Vargaa, H. J. M. Steeneken, “Assesment for audtomatic speech recognition: Ⅱ. NOISEX-92: A database and an experiment to study the effect of additive noise on speech recognition systems,” Speech Commun., 9, 357–363, 1990.

[30] K. R. Kluender, M. A. Walsh, “Amolitude rise time and the perception of the voiceless affricate/fricative distinction,” Percept. Psychophys., 51, 328–333, 1992.

[31] P. J. Fitzibbons, S. Gordon-Salant, “Auditory temporal processing in elderly listeners,” J. Am. Acad. Audiol., 7, 183–189, 1996.

[32] S. E. Yoho, E. W. Healy, C. L. Youngdahl, T. S. Barrett, F. Apoux, “Speech-material and talker effects in speech band importance,” J. Acoust. Soc. Am., 143, 1417–1426.

ⓒ 2021 Information Processing Society of Japan