

# 音声認識システムを用いた高齢難聴患者への コミュニケーション支援の確立 —騒音環境下における音声認識率について—

辰 巳 浩 隆      樋 口 恭 子      米 谷 裕 之  
辻 一 起 子      米 田 護      大 西 明 雄  
谷 岡 款 相      紺 井 拡 隆

**抄録：**近年，高齢難聴患者へのコミュニケーション手段に音声認識ソフトを用いた音声入力文字表示システムがある。このシステムが有効に機能するために，ソフトの音声認識率が高いことが重要である。

そこで，本表示システムの確立の一端として，騒音環境下における音声認識率を検討した。音読用文章を作成し文節に区切ったのち，被験者の成人8名に対し，騒音環境下での音声認識率を検索した。静寂な個室で有線と無線マイクを用いて歯科診療室の騒音である歯科用タービンとバキュームの混合騒音（70dBと80dB）を流した場合と流さなかった場合（無騒音）について音声認識率を比較した。ソフトはAmiVoice<sup>®</sup> Ex7を，マイクはAmiVoice<sup>®</sup>専用である有線と無線マイクを用いた。

その結果，有線マイクの場合，無騒音の平均誤認識率は3.5±2.0%，70dBは10.6±7.5%，80dBは26.1±13.6%で無騒音と混合騒音間で統計学的な差が認められた（ $p < 0.05$ ）。無線マイクも同様に，無騒音の平均誤認識率は8.8±4.4%，70dBは25.9±10.8%，80dBは41.6±11.2%で無騒音と混合騒音間で統計学的な差が認められた（ $p < 0.05$ ）。

以上，騒音が小さい環境下で有線マイクを使用した場合，音声認識率が高かったことから，本表示システムは有線マイクを用いて小さな騒音環境下で活用することが有効であると考えられる。

**キーワード：**音声認識 高齢難聴患者 騒音環境 有線マイク 無線マイク

## 緒 言

昨今の超高齢社会により，年々高齢難聴者が増加し，社会的な対応が求められている。医療の場においても高齢難聴患者が増え，医療面接時に医師と患者間でコミュニケーションギャップを起こす頻度が高くなってきている<sup>1)</sup>。その対策のひとつとして，高齢難聴患者を含む聴覚障がい者のコミュニケーション支援の手段である音声入力文字表示システム（以下，表示システムと略す）が報告されている<sup>2)</sup>。表示システムは，話し手の音声即時にわかりやすく読みやすい文字に表示し，臨場感あふれる形で伝達できることからコミュニケーション手段として有用であることが示唆されている<sup>2-4)</sup>。このシステムが有効に機能するためには，ソフトの音声認識率が高いことが不可欠である。

著者ら<sup>3)</sup>は，今までに難聴患者への表示システムの確立の一端として，書き言葉に対する音声認識ソフトのAmiVoice<sup>®</sup> Ex Dentalの音声認識率を検討している。その結果，AmiVoice<sup>®</sup> Ex Dentalが単語登録と音響学習の活用により，書き言葉に対して有用であることを報告している。また，話し言葉に対するAmi

Voice<sup>®</sup> Ex Dentalの音声認識率についても検討した結果，AmiVoice<sup>®</sup> Ex Dentalは単語登録と音響学習の活用により，話し言葉に対しても有用であることを報告している<sup>4)</sup>。さらに，著者ら<sup>5)</sup>は表示システムをより確立させるために，市販の音声認識ソフトであるAmiVoice<sup>®</sup> SP，Dragon SpeechおよびAmiVoice<sup>®</sup> Ex Dentalと各ソフト専用ヘッドセットの2種の有線マイクおよび1種の無線マイクを使用し，それぞれの相性による認識率を検討している。その結果，ソフトはAmiVoice<sup>®</sup> Ex Dentalが最適であること，マイクは有線と無線のいずれも有用であることを示唆している。

このように，表示システムのシステム側の観点からソフトはAmiVoice<sup>®</sup> Ex Dentalを用いシステム機能である単語登録と音響学習を活用することにより，表示システムが有効であることを明らかにしている。

一方，表示システムの音声認識率を低下させる要因として，主に話し手の韻律的要素や騒音などのユーザーや使用環境に関する因子が挙げられる<sup>3,4,6,7)</sup>。なかでも，騒音は音声認識の入力部において音声認識率を低下させる阻害要因として大きな問題である<sup>6,7)</sup>。

今回，表示システムの確立の一端として，使用環境

の観点から音声認識率の影響を検索するために、歯科診療室を想定した騒音環境下における音声誤認識率を測定した。

対象および方法

1. 被験者

被験者は 25 ~ 28 歳 (平均年齢 26.3 歳) で、歯の欠損がなく、顎口腔、発声および聴覚器官に異常が認められない個性正常咬合を有する健常有歯顎者の成人 8 名 (男性 4 名と女性 4 名) とした。

2. 音声認識ソフトと使用環境

オペレーティングシステムは Windows 8.1 Pro Update 64bit (Microsoft Corporation, WA, USA), ハードウェアは Let's note CF-R8 (パナソニック株式会社, 大阪), 音声認識ソフトは AmiVoice<sup>®</sup> Ex Dental (アドバンスト・メディア社製, 東京), およびヘッドセットマイクは AmiVoice<sup>®</sup>専用である有線の Parrott TalkPro USB 100 (VXI corporation, Rollinsford, NH, USA) (以下, 有線と略す) と無線の Jabra PRO 935 (GN Netcom A/S, Copenhagen, Denmark) (以下, 無線と略す) を使用した。

3. 表示システムと騒音

表示システムは、音声認識ソフトにより音声を文字に変換させ、画面に表示するシステムである。

本研究で使用した騒音は、歯科治療ユニット GC EOM-α II (株式会社ジーシー, 東京) の歯科用タービン (KAVO Super-Torque Lux 650B, 株式会社ジーシー, 東京) とバキュームの混合音を Linear PCM Recorder PCM-M10 (ソニー株式会社, 東京) で録音し、それをパーソナル・アクティブ・スピーカーの

Fostex PA3 (フォスター電機株式会社, 東京) で発生させた音とした。騒音の種類は、混合音を流さずマイク先端部付近でデジタル騒音計 Center 322 (株式会社佐藤商事, 神奈川) で測定した 35.5 ~ 39.3dB となる音 (以下, 無騒音と略す), 混合音が被験者のマイク先端部付近で 69.3 ~ 75.1dB となる音 (以下, 70dB と略す), および 79.6 ~ 86.2dB となる音 (以下, 80dB と略す) とした (図 1)。また、音源の設定は各被験者を机の上に設置したスピーカーに向き合って座わらせ、スピーカーの高さを床から 85cm, スピーカーからマイク集音部までの距離を 75cm および被験者の口先とマイク集音部との直線距離を 5cm とした。

4. 音読用歯科説明文

音声誤認識率を検索するための音読に用いた歯科説明文と区切った文節を表 1 に示す。「おはようございます」から「次に口の中を診させていただきます」までの計 20 文を作成したのち、84 文節に区切った。

5. 音声入力時における韻律的特徴の設定

音声入力時における韻律的特徴の設定は、AmiVoice<sup>®</sup> Ex Dental の取扱説明書に準じた。すなわち、各被験者に対し、普段発音しているように自然に発声すること、単語ごとに区切るのではなく文章で発声す

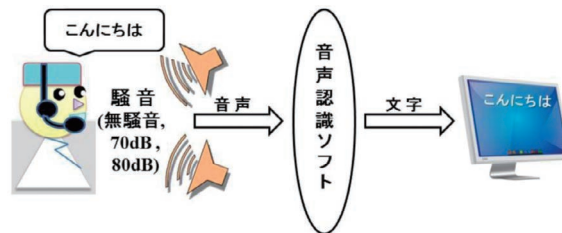


図 1 音声入力文字表示システムと騒音

表 1 音読用文章と区切り

(1) おはよう / ございます
(2) これから / 治療する / 上で / 必要な / ことを / 質問させて / 頂きます / よろしい / でしょうか
(3) 今日は / どうされましたか
(4) 痛い / ところは / どこですか
(5) どのような / 痛みですか
(6) いつ頃 / から / 腫れましたか
(7) 歯茎の / 腫れは / ありますか
(8) 歯の / ぐらつきは / ありますか
(9) その歯は / 以前に / 治した / ことが / ありますか
(10) どのような / 治療を / しましたか
(11) 化膿止めや / 痛み止めの / 薬を / 飲みましたか
(12) 必要が / あれば / エックス線 / 写真を / 撮ります
(13) 今までに / 大きな / 病気に / かかった / ことは / ありますか
(14) 薬の / 副作用は / ありますか
(15) 今までに / 歯の / 麻酔を / したことが / ありますか
(16) 食べ物や / 金属に / 対する / アレルギーは / ありますか
(17) 歯を / 抜いた時に / 血が / 止まりにくい / ことは / ありましたか
(18) 治療に / 対して / 何か / 希望は / ありますか
(19) 他に / 何か / 言い忘れた / ことは / ありませんか
(20) 次に / 口の / 中を / 診させて / 頂きます

(計 20 文, 84 文節)

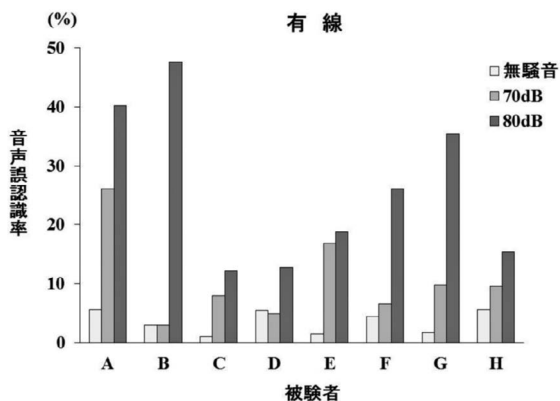


図 2 各被験者別の平均誤認識率 (有線マイク)

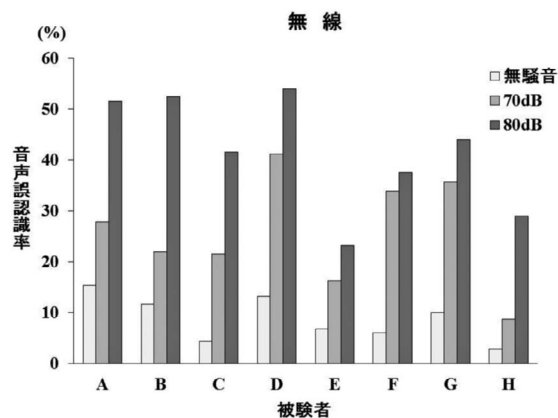


図 3 各被験者別の平均誤認識率 (無線マイク)

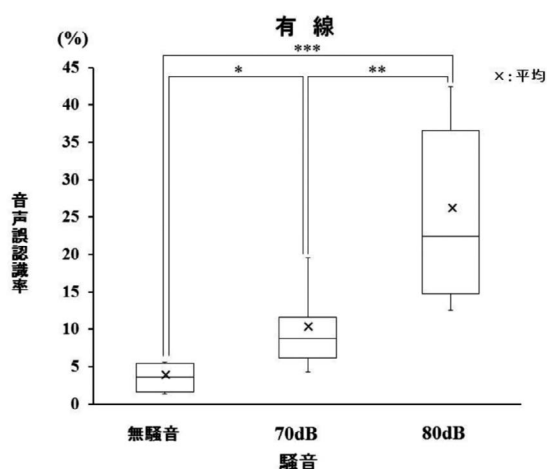


図 4 騒音別の平均誤認識率 (有線マイク)

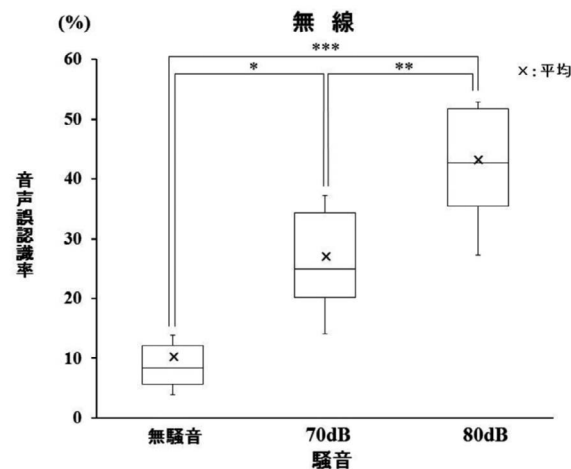


図 5 騒音別の平均誤認識率 (無線マイク)

ること、ソフトが指定した適切な音量で明瞭に発声すること、および抑揚はつけずに淡々と発声することを事前に説明した。

6. 誤認識率の測定法

誤認識率の測定法は、無騒音下の静かな個室で各被験者に対し、システム機能である単語登録と音響学習を活用せずに計5回の音声入力をした。ついで、同様に70dBと80dBの騒音下において各被験者に対し、計5回の音声入力をした。また、各騒音下における有線と無線マイクを用いた詳細な音声入力の順序は、初日にまず無騒音下での有線マイク、ついで同無線マイク、2日目に70dBでの有線マイク、ついで同無線マイク、そして3日目に80dBでの有線マイク、ついで同無線マイクとした。

誤認識率は各音声入力時に誤認識が認められた文節数を全文節数で割った値とした。

7. 統計学的分析

有線と無線の各騒音環境下における平均誤認識率の有意差は、有意水準0.05としてFriedman検定で分析した。

8. 倫理的配慮

本研究は大阪歯科大学 医の倫理委員会の承諾 (大歯医倫第110846号) を得て行った。

結 果

1. 全被験者における平均誤認識率の範囲

有線における被験者別の平均誤認識率の範囲は、無騒音で1.0～5.6%、70dBで4.9～26.1%、および80dBで12.2～47.6%であった(図2)。

一方、無線における被験者別の平均誤認識率の範囲は、無騒音で2.9～15.4%、70dBで8.8～41.2%、および80dBで23.2～53.9%であった(図3)。

騒音環境下において、有線および無線ともに誤認識率の範囲は広く高い値を示した。また、全体的に有線よりも無線の方が高い値を示した。

2. 騒音別の平均誤認識率

有線における騒音別の平均誤認識率を図4に示す。無騒音の平均誤認識率は3.5±2.0%で中央値は3.7%、70dBの平均値は10.6±7.5%で中央値は8.8%、80dBの平均値は26.1±13.6%で中央値は22.5%であり、無

表 2 高い誤認識率を示した認識困難文節

有線		無線	
誤認識文節 (誤認識率)	誤認識文節 (例)	誤認識文節 (誤認識率)	誤認識文節 (例)
<b>無騒音</b>			
希望は	(45%) 希望	希望は	(85%) 希望
化膿止めや	(20%) 可能な眼や	化膿止めや	(35%) 可能な目や
治療を	(12.5%) 治療	どうされましたか	(30%) ドされましたか
ありましたか	(12.5%) あります	その歯は	(25%) その歯を
治した	(10%) 何をした	血が	(25%) 位置が
薬を	(10%) 薬は	ありましたか	(22.5%) ありましたか
血が	(10%) 位置が	ことを	(20%) こと
<b>70dB</b>			
希望は	(60%) 希望	希望は	(95%) 希望
ありましたか	(47.5%) ありましたか	化膿止めや	(85%) 可能な面あり
血が	(47.5%) 歯が	はありますか	(80%) がありますか
飲みましたか	(37.5%) 飲みましたか	血が	(72.5%) 歯が
はありますか	(35%) がありますか	その歯は	(67.5%) この歯は
腫れましたか	(30%) 生えましたか	ありましたか	(62.5%) ありましたか
化膿止めや	(30%) 可能な面あり	どうされましたか	(55%) のされましたか
<b>80dB</b>			
希望は	(87.5%) 希望	希望は	(90%) 希望
はありますか	(77.5%) はありますか	その歯は	(87.5%) この歯は
ありましたか	(77.5%) ありますが	ありますか	(87.5%) ありますが
腫れましたか	(70%) 腫れましたか	痛みですか	(85%) 痛みですが
血が	(67.5%) 歯が	はありますか	(85%) はありますか
どうされましたか	(62.5%) どうされましたか	血が	(82.5%) 歯が
飲みましたか	(62.5%) 飲みましたか	どうされましたか	(80%) どうされましたか

騒音と各騒音間で統計学的な差が認められた ( $p < 0.05$ )。また、無線における騒音別の平均誤認識率を図5に示す。無騒音の平均誤認識率は  $8.8 \pm 4.4\%$  で中央値は  $8.4\%$ 、70dBの平均値は  $25.9 \pm 10.8\%$  で中央値は  $24.9\%$ 、80dBの平均値は  $41.6 \pm 11.2\%$  で中央値は  $42.7\%$  であり、有線の場合と同様に、無騒音と各騒音間で統計学的な差が認められた ( $p < 0.05$ )。

### 3. 認識困難文節

各騒音下での有線と無線において誤認識率が高かった認識困難文節を表2に示す。有線および無線ともに全騒音下で「希望が」の誤認識が最も多かった。

また、騒音環境下では無騒音よりも、有線では「飲みましたか」と「腫れましたか」、無線では「飲みましたか」と「痛みですか」などの特定の文節および疑問文に多くの誤認識がみられた。

## 考 察

現在、音声認識ソフトを用いた音声認識技術は、製造業全般、流通業、電力、鉄道、船舶、住宅など社会生活の様々な場面で活用されている<sup>8)</sup>。医療領域の分野では放射線科の読影レポートの作成、電子カルテの入力システムおよび頸髄損傷患者のコンピューターアクセスビリティの向上などに用いられている<sup>9-12)</sup>。一般に、その技術は音声の音響的多様性、言語的多様性

および音声信号の集音環境に影響されることが報告されている<sup>7)</sup>。特に、集音環境での阻害因子のひとつである騒音は、音声認識精度の劣化に大きく影響を及ぼし、特殊なマイクの使用や集音環境が限定された車中のカーナビゲーションなどを除いて、音響的に音声認識性能が劣化し易い場面では実用化に至っていないのが現状である<sup>7)</sup>。また、歯科診療室も、歯科用タービン、ハンドピースおよびバキューム音や術者と患者の会話などの騒音が日常的に発生し<sup>13)</sup>、集音環境として音響的に音声認識性能が劣化し易い場所である。

そこで、表示システムの確立の一端として、使用環境の影響を検索するために、歯科診療室で発生する騒音のうち、歯科用タービンとバキュームとの混合音下での音声認識率を測定した。

従来から歯科領域における特有な騒音のうち、歯科用タービンの騒音は聴力損失以外にも、耳鳴り、心理的影響および生理的影響などの悪影響を与えることが報告されている<sup>14)</sup>。そのうち、心理的影響と生理的影響については、話し手の精神的に悪影響を与え、音声の韻律的因子に変化が生じて認識率が低下する可能性が考えられる。その歯科用タービンは一般的に歯科用バキュームと併用することが多いことから、本研究で用いた騒音は、歯科用タービンとバキュームとの混合音とした。

その結果, 有線および無線における被験者別の平均誤認識率の範囲は, 騒音環境下において, 誤認識率の範囲は広く高い値を示し, 全体的に有線よりも無線の方が高い値であった。

これまでの他の報告<sup>15-17)</sup>で話し手の感情, 心理および精神状態は, 音声に影響を及ぼし声質などを変化させることが明らかにされている。本研究では被験者Hの無騒音および70dBにおける有線マイクの誤認識率(5.6と9.5%)が, 無線マイク(2.9と8.8%)よりも高い値を示した。これは, 今回, 各騒音下の音声入力順序においてすべての日で無線マイクよりも有線マイクを先に用いて入力した。その結果, 被験者Hが有線マイクでの入力時に緊張や不安などで精神的に不安定となったことや入力の慣れが十分でなかったことにより, スムーズに発声できず韻律的特徴に乱れが生じたことが原因として考えられる。今後, 音声入力回数や被験者数を増やし, 入力順序を色々と変えて検討する必要がある。

騒音別の平均誤認識率では, 有線における無騒音の平均誤認識率は $3.5 \pm 2.0\%$ で中央値は3.7%, 70dBの平均値は $10.6 \pm 7.5\%$ で中央値は8.8%, 80dBの平均値は $26.1 \pm 13.6\%$ で中央値は22.5%であり, 無騒音と各騒音間で統計学的な差が認められた。一方, 無線における無騒音の平均誤認識率は $8.8 \pm 4.4\%$ で中央値は8.4%, 70dBの平均値は $25.9 \pm 10.8\%$ で中央値は24.9%, 80dBの平均値は $41.6 \pm 11.2\%$ で中央値は42.7%であり, 有線の場合と同様に, 無騒音と各騒音間で統計学的な差が認められた。しかし, 各被験者別の無騒音と騒音下における有線および無線マイクの誤認識率では, 被験者Bの無騒音と70dBにおける誤認識率が同じ値(2.9%)を, 被験者Dの無騒音における誤認識率(5.4%)が70dB(4.9%)よりも高い値を示した。本研究における各被験者の全入力パターンのうち, 無騒音の有線マイクによる入力が最初であった。その結果, 被験者Hの場合と同様に緊張や不安などで精神的に不安定となったことや入りに少し慣れていなかったことにより, 両被験者はスムーズに発声できず韻律的特徴に乱れが生じたことが原因と考えられる。前述のように, 音声入力回数, 被験者数および入力順序を考慮する必要がある。

松島ら<sup>18)</sup>は, 音声認識ソフトのVia Voice Pro Ver. 10 (IBM) に対して48dB, 60dB, 70dBおよび80dBの騒音環境下での赤外線を用いた無線マイクと有線マイクとの言語認識率を検討している。その結果, 各騒音レベル48dB, 60dB, 70dBおよび80dBにおける言語認識率は, 赤外線システムで83.3%, 83.5%, 69.8%, 52.3%, 有線システムで83.7%, 81.6%, 69.5%, 43.5%であり, いずれも70dBおよび80dB下で, 言語認識率は低下したものの, 両システムに有意差は認められ

なかったことを報告している。本結果も, 騒音が大きくなるにつれて, 認識率が低くなるという松島ら<sup>18)</sup>と同様の結果を示した。

この理由は, 音声信号の集音環境による影響<sup>7)</sup>とロンバード効果<sup>19)</sup>の2つの要因が考えられる。音声信号の集音環境の要因として, 本研究で使用したマイクはノイズキャンセルマイクでまわりの雑音を低減するマイクであるが, 現実には静寂な場合と比較して認識対象とする音声に多少の騒音混入などの信号変化が加わるため, 音声認識の低下が起こったと考えられる。もう一つは, 周囲の騒音が大きい環境にいと, 話し手が自然と声を張り上げ, 静寂な環境で発声された音声と比較してインテンシティが大きくなるだけでなく, 基本周波数やフォルトマン周波数が高くなるなどの音響的特徴が変化するロンバード効果が影響したことが考えられる。すなわち, ロンバード効果により話し手の音声は通常よりもかなり大きくなり, ソフトが正常に音声を把握できず音声認識率の低下が起こったと考えられる。

本研究では無線マイクよりも有線マイクの音声認識率の方が高かった。通常, 無線マイクが有線マイクよりも機能が劣る原因として, 主として無線マイクの周波数の混信や使用場所による電波のデッドポイントなどの問題があげられる<sup>20)</sup>。本研究では, すべての被験者が同じシステム環境と使用場所の同じ条件で音声を入力した。このことから, 今回の有線と無線マイクにおける誤認識率の相違は, 無線周波数の混信や使用場所による電波のデッドポイントの問題よりも, ソフトとマイクとの相性が主な原因であると考えられる。

また, 本実験の認識率は, 松島ら<sup>18)</sup>の報告である70dBおよび80dBの騒音環境下で赤外線の無線マイクおよび有線マイクとVia Voice Pro Ver. 10 (IBM)の言語認識率よりも全体的に高かった。これは, Via Voice Pro Ver. 10 (IBM) とAmiVoice<sup>®</sup> Ex Dentalとのソフト, および本研究で用いた無線マイクのBluetoothと松島ら<sup>18)</sup>の赤外線との音声転送方法の違いによるものと考えられる。したがって, 表示システムにはAmiVoice<sup>®</sup> Ex DentalとParrott TalkPro USB 100あるいはJabra PRO 935との組み合わせの方がより有効であることが示唆された。

河原<sup>21)</sup>は, 生放送の字幕付与における音声認識レベルの認識率は95%以上が要件とされているが, 認識率が90%を上回ると誤りがあってもほとんど理解に支障がなく, 「間違い探し」のレベルとなると報告している。また, 米国国防総省高等研究計画局(Defense Advanced Research Projects Agency: DARPA) 音声認識プロジェクトはテレビニュース番組への字幕付与におけるアナウンサーの読み上げ部分は, 90%程度の認識率が得られていること<sup>21)</sup>, および

本間<sup>22)</sup>は雑音環境下でのニュース字幕システムにおけるNHKアナウンサーの音声認識率は、88.2～95.1%であることを報告している。したがって、音声誤認識率が約10%以下の場合、聞き手にとってほぼ支障がなく実用的な音声認識レベルであると考えられる。本研究では、有線の無騒音における平均誤認識率の $3.5 \pm 2.0\%$ 、70dBの $10.6 \pm 7.5\%$ 、および無線の無騒音における平均誤認識率の $8.8 \pm 4.4\%$ であった。約10%以下の音声誤認識率が実用レベルとした場合、有線と無線の無騒音、および有線マイクの70dBが許容範囲であり、その他は実用レベルではないと示唆される。それゆえ、騒音環境の観点から、表示システムは有線マイクを用い可能な限り小さな騒音環境下で活用することが重要であると考えられる。

今回、有線および無線における騒音の有無にかかわらず、全般にわたって「希望は」、「化膿止めや」、「ことはありましたか」ならびに「血が」などの文節に高い誤認識率が認められた。例えば「希望は」は「希望」や「希望が」に、「化膿止めや」は「可能な面あり」や「化膿止め」に、「ことはありましたか」は「ことはありましたか」や「ことはあります」に、「血が」は「位置が」や「歯が」などに誤認識された。この結果は、以前に著者ら<sup>3-5)</sup>が報告した認識困難文節と同様の結果であり、原因として、ソフトのシステムや話し手の発音が考えられた。また、有線および無線において、騒音環境下では無騒音よりも、疑問文に多くの誤認識がみられ、騒音が大きくなるにつれて誤認識疑問文が多く認められた。例えば、「飲みましたか」が「飲みましたが」や「飲みました」、「腫れましたか」が「腫れましたが」や「腫れました」、「痛みですか」が「痛みですが」や「痛みです」などの誤認識がみられた。すなわち、文末における「か」が「が」になる誤認識、および文末の「か」が抜ける現象が認められた。これは、騒音が妨害音として働き、認識精度が劣化し認識困難が生じていると示唆される。

以上の結果より、騒音が小さい環境下で有線を使用した場合において誤認識率が低くなることが明らかとなった。したがって、使用環境の観点から、表示システムは有線を使用して可能な限り小さな騒音環境下で活用することが有効であると考えられる。

今後、さらに表示システムを確立させるためには、話し手の音声の様々な韻律的要素と音声認識率との関連を詳細に検討する必要がある。

## 結 論

表示システムの確立の一端として、成人8名を対象に、騒音環境下での歯科説明文に対する音声認識率を検索した。その結果、使用環境の観点から、表示システムは有線を使用して可能な限り小さな騒音環境下で

活用することが重要であると考えられた。

## 利益相反

本論文について開示すべき利益相反関係にある企業などはありません。

## 謝 辞

本研究の一部は、JSPS科学研究費補助金 挑戦的萌芽研究(15K15779)の助成により実施した。

本論文の要旨は、第10回日本総合歯科学会総会・学術大会(平成29年11月4日、新潟市)において発表した。

## 文 献

- 1) 横尾美希, 原 祥子. 急性期病院に入院している難聴高齢者の難聴に由来する体験. 老年看 2011; 16: 66-74.
- 2) 栗岡成人, 有田清三郎, 武田隆久. 難聴患者の診療時における音声入力文字表示システム —診療コミュニケーション改善のために—. バイオ・ファジィ・システム会誌 2009; 11: 7-13.
- 3) Tatsumi H, Arita S, Koide T, Matsumoto K, Nagame S, et al. A Study of a Medical Speech Display System Using Voice Recognition Software for Hearing-impaired Patients. —Rate of Speech Recognition—. Jpn J Oral Diag/Oral Med 2011; 24: 255-262.
- 4) Tatsumi H, Arita S, Koide T, Matsumoto K, Nagame S, et al. A Study of a Medical Speech Display System Using Voice Recognition Software for Hearing-impaired Patients. —Rate of Speech Recognition for Spoken Language—. J Jpn Assoc Oral Rehabil 2012; 25: 1-10.
- 5) 辰巳浩隆, 樋口恭子, 小出 武, 米谷裕之, 辻 一起子, 他. 音声認識システムを用いた高齢難聴患者へのコミュニケーション支援の確立 —様々なソフトとマイクの音声認識率について—. 日総歯誌 2016; 8: 33-41.
- 6) 平井優司, 長井美和, 横井英人, 澤田秀之, オントロジーに対応した電子カルテ入力支援システムと医療現場での音声入力実験. 医療情報学 2012; 32: 73-81.
- 7) 西光雅弘, 堀 智織. 現在の音声認識にできること・できないこと. 日本音響学会誌 2015; 71: 158-163.
- 8) 塚田 聡. 耐騒音音声認識装置 VoiceDo. NEC 技研 2010; 63: 62-64.
- 9) Ramaswamy MR, Chaljub G, Esch O, Fanning DD, vanSonnenberg E. Continuous speech recognition in MR imaging reporting: advantages, disadvantages, and impact. AJR 2000; 174: 617-622.
- 10) 大山嘉将, 河辺謙治, 水口和夫, 多田卓仁, 岡村光英, 他. 市販音声入力ソフトウェアを用いた所見入力システムの可能性の検討. Radiol Fronti 2001; 4: 129-131.
- 11) 徐 和廷, 坂部長正. 電子カルテへの音声入力ツール応用研究 —誤認識医学用語修正インターフェイスの開発—. 交通医 2002; 56: 99-108.
- 12) 松下 功, 大島淳一, 影近謙治. 頸髄損傷患者に対する音声認識ソフトの使用経験. 臨床リハ 2003; 12: 564-566.
- 13) 山田朋美. 歯科医療機器から発生する超高周波音 —歯科切削騒音による不快感の低減をめざして—. 日本音響学会誌 2009; 65: 52-57.
- 14) 笠井 岳. 歯科用エアタービンハンドピースから発生す

- る騒音の高周波成分. 日大口腔科学 1994; 20: 129-141.
- 15) 小林 聡, 北澤茂良. 日本語の自然対話音声におけるパラ言語的特徴の検討. 日本音響学会誌 2000; 56: 467-476.
- 16) 森山 剛, 茂 美穂. 音声感情を含むことによる声質変化の解析. 東京工芸大学工学部紀要 2011; 34: 58-64.
- 17) 森 大毅. 話し言葉が伝えるもの. 国語プロジェクトレビュー 2014; 4: 183-190.
- 18) 松島俊介, 寺井親則, 岡本 健. 赤外線無線を使った電子カルテ音声入力システムの開発. INNERVISION 2003; 18: 32.
- 19) Lane H and Tranel B. The Lombard sign and the role of hearing in speech. J Speech Hear Res 1971; 14: 677-709.
- 20) 石橋 透. ワイヤレスマイクの混信とデッドポイントの回避について: 周波数帯の確保と教育実践. 尚美学園大  
学芸術情報研究 2014; 23: 1-21.
- 21) 河原達也. 話し言葉の音声認識の進展 — 議会の会議録作成から講演・講義の字幕付与へ —. メディア教育研究 2012; 9: S1-S8.
- 22) 本間真一. 生字幕制作のための音声認識. NHK 技研 2010; 122: 25-31.

#### 著者への連絡先

辰巳 浩隆  
〒540-0008 大阪府大阪市中央区大手前 1-5-17  
大阪歯科大学附属病院 口腔診断・総合診療科  
TEL 06-6910-1066 FAX 06-6910-1026

E-mail: tatsumi@cc.osaka-dent.ac.jp

## Communication support for elderly patients with hearing impairment using a voice recognition system —Voice recognition rate in a noisy environment

Hiroataka Tatsumi, Kyoko Higuchi, Hiroyuki Kometani,  
Ikiko Tsuji, Mamoru Komeda, Akio Ohnishi,  
Tadasuke Tanioka and Hiroataka Kon'i

Department of Oral Diagnosis and Interdisciplinary Dentistry, Osaka Dental University

**Abstract** : In recent years, as a means of communication with elderly patients who have hearing impairment, there are systems that display characters with voice input using voice recognition software. For these systems to function effectively, it is important that the voice recognition rate of the software is high.

As one part of establishing this system, we investigated the voice recognition rate in a noisy environment. Texts for reading aloud were prepared and phrases were delineated. The voice recognition rate in a noisy environment was then examined with 8 dentists as subjects. Using wired and wireless microphones in a quiet, private room, the voice recognition rate was compared in cases with mixed noise (70 dB and 80 dB) from a dental turbine and vacuum played in the room, which are common noises in a dental office, and in cases without background noise. The software used was AmiVoice<sup>®</sup> Ex7, with the AmiVoice<sup>®</sup> accessory wired and wireless microphones.

With the wired microphone, the mean false recognition rate was  $3.5 \pm 2.0\%$  without background noise,  $10.6 \pm 7.5\%$  with 70 dB noise, and  $26.1 \pm 13.6\%$  with 80 dB noise, showing a statistical difference between the voice recognition rates without background noise and with mixed noise ( $p < 0.05$ ). Similarly, with the wireless microphone, the mean false recognition rate was  $8.8 \pm 4.4\%$  without background noise,  $25.9 \pm 10.8\%$  with 70 dB noise, and  $41.6 \pm 11.2\%$  with 80 dB noise, showing a statistical difference between the voice recognition rates without background noise and with mixed noise ( $p < 0.05$ ).

The above results indicate that the voice recognition rate is high when a wired microphone is used in an environment with little noise, and the use of this system is thought to be beneficial when a wired microphone is used in an environment with little noise.

**Key words** : speech recognition, hearing-impaired elderly patient, noisy environment, wired microphone, wireless microphone