

# 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 —

辰 巳 浩 隆      樋 口 恭 子      小 出   武  
米 谷 裕 之      辻   一 起 子      米 田   護  
大 西 明 雄      谷 岡 款 相      中 井 智 加  
稗 田 具 美      岩 見 江 利 華

**抄録:** 近年, 高齢難聴患者へのコミュニケーション手段のひとつ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。このシステムが機能する重要なポイントとして, ソフトの音声認識率が高いことがあげられる。

そこで, 音声入力文字表示システムの確立の一端として, 成人6名を対象に, 歯科医療面接で用いる文章に対する様々な音声認識ソフトとマイクの音声認識率を検討した。

今回, ソフトはDragon Naturally Speaking Ver. 11 (DNS), AmiVoice<sup>®</sup> SP2 (AS) および AmiVoice<sup>®</sup> Ex7 (AE) を, マイクは各ソフト専用ヘッドセットの2種の有線マイクと1種の無線マイクを使用した。

その結果, ソフト別の平均誤認識率はDNSが $15.4 \pm 5.3\%$ , ASが $16 \pm 4.7\%$ , AEが $5.3 \pm 3.6\%$ でAEが最も低かった。DNSとAEおよびASとAE間で統計学的な差が認められた ( $p < 0.05$ )。また, DNSとASのマイクは無線よりも有線マイクが有効であるのに対して, AEは差が認められなかった。

以上の結果から, 音声入力文字表示システムに使用するソフトはAEが最適であること, およびそのマイクは有線と無線のいずれも有用であることが示唆された。

**キーワード:** 音声認識 高齢難聴患者 歯科医療面接 有線マイク 無線マイク

## 緒 言

将来, 超高齢社会と高度情報化社会により, 今まで以上に高齢難聴者が急増し, 社会的に問題視されることが予想される。事実, 医療の場においても高齢難聴患者が増加し, 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ことが少なくない<sup>1)</sup>。一般に, 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には, 手話と文字によるものがあり, 文字によるコミュニケーション支援は既存方法である筆談と字幕システムがある<sup>2,3)</sup>。そのうち, 筆談は従来, 多くの医療機関で行われ, 筆記用具と紙だけの低コストという利点を有するが, 筆記に時間を要することや筆談による意志疎通には質的および量的に限界があるなどの欠点がある<sup>4)</sup>。一方, 字幕システム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) はコストが少しかかるが, 誤認識などの点をクリアできれば, 話し手の音声を即時にわかりやす

く読みやすい文字に表示でき, 臨場感あふれる形で伝えることが可能である<sup>3,5,6)</sup>。現在, 様々な音声認識ソフトが市販されているが, ソフトの認識性能やユーザー側の発話の韻律的要因から完全に音声を認識できるソフトは見当たらない。それゆえ, このシステムを確立させるためには, 音声誤認識率 (以下, 誤認識率と略す) が低いソフトとマイクを使用することが重要である。

本研究は歯科医療面接時の高齢難聴者とのコミュニケーションを円滑に図るための表示システムとしての音声認識ソフトの有用性の検討と歯科医療面接時の表示システムに用いる最適な音声認識方法を明らかにすることを目的とした。

そこで, まず様々なソフトとマイクを使用し, 初診時の歯科医療面接で用いる文章に対するソフトの誤認識率から最も有用なソフトとマイクの組み合わせを検討した。

## 対象および方法

### 1. 表示システムの概要

表示システムの概要を図1に示す。表示システムは、音声認識ソフトにより音声を瞬時に文字へ変換させ、画面に表示するシステムである。

### 2. 被験者, Operating System (OS), 音声認識ソフトおよびマイク

本実験における被験者は平均年齢25.5歳で、歯の欠損がなく、顎口腔、声および聴覚器官に異常がない個性正常咬合を有する健常有歯顎者の成人6名（男性4名と女性2名）とした。OS、音声認識ソフトおよびマイクを表1に示す。OSはWindows 8.1 Pro Update 64bit (Microsoft Corporation, WA, USA)を、ハードウェアはLet's note CF-RZ4 (パナソニック株式会社, 大阪)を使用した。音声認識ソフトはDragon Naturally Speaking Ver. 11 (ニュアンス・コミュニケーションズ・ジャパン, 東京; 以下, DNSと略す), AmiVoice<sup>®</sup> SP2 (アドバンスト・メディア社製, 東京; 以下, ASと略す) および AmiVoice<sup>®</sup> Ex7 (ア



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

図1 音声入力文字表示システム

ドバンスト・メディア社製, 東京; 以下, AEと略す)を用いた。また, DNSは有線マイクのニュアンス・ドラゴン USB ヘッドセットマイク (ニュアンス・コミュニケーションズ・ジャパン, 東京) とゼンハイザー SC 230 USB (Sennheiser electronic GmbH & Co. KG, Deutschland) および無線マイクのプラントロニクス Voyager Legend UC (Plantronics, Inc., CA, USA)を, ASとAEは有線マイクの Parrott TalkPro USB 100 (VXI corporation, Rollinsford, NH, USA) と VXI TalkPro UC1 (VXI corporation, Rollinsford, NH, USA) および無線マイクの Jabra PRO 935 (GN Netcom, Denmark) を使用した。

3. 歯科医療面接の音読用シナリオと区切った文節例  
音声認識率を検索するための音読に用いた歯科医療面接の音読用シナリオを表2に, 区切った文節を表3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。また, 本実験で使用したソフトは, 通常3語以上の前後文脈における連携の特徴で認識する連続音声認識である<sup>7,8)</sup> ため, 今回, シナリオを任意に区切り各文節の音声認識を検索した。

### 4. 誤認識率の測定法

ソフトとマイクの組み合わせを表4に示す。ソフトとマイクの組み合わせは, DNSとニュアンス・ドラゴン USB ヘッドセットマイク (以下, DNS-Nと略す), DNSとゼンハイザー SC 230 USB (以下, DNS-Sと略す), DNSとプラントロニクス Voyager Legend UC (以下, DNS-Pと略す), ASとParrott TalkPro USB 100 (以下, AS-100と略す), ASとVXI TalkPro UC1 (以下, AS-1と略す), ASとJabra PRO 935 (以下, AS-Jと略す), AEとParrott TalkPro USB 100 (以下, AE-100と略す), AEとVXI TalkPro UC1

表1 本実験で用いた OS, ソフトおよびマイク

<u>Operating System</u>
Windows 8.1 Pro Update 64bit (CPU: Core™, M-5Y10c Memory: 4GB)
<u>Software</u>
Dragon Naturally Speaking Ver. 11 (DNS)
AmiVoice <sup>®</sup> SP2 (AS)
AmiVoice <sup>®</sup> Ex7 (AE)
<u>Microphone</u>
・ Dragon Naturally Speaking 11
Nuance Dragon USB Headset Microphone (Wired)
Sennheiser SC 230 USB (Wired)
Plantronics Voyager Legend UC (Wireless)
・ AmiVoice <sup>®</sup>
Parrott TalkPro USB 100 (Wired)
Parrott TalkPro UC1 (Wired)
Jabra PRO 935 (Wireless)

表 2 歯科医療面接の音読用文章

---

(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いただ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いただきます

---

(計 20 文)

表 3 音読用文章の区切り

---

(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 84 文節)

(以下, AE-1 と略す) および AE と Jabra PRO 935 (以下, AE-J と略す) の 9 組とした。

誤認識率の測定は, システム機能である単語登録と音響学習および DNS のボキャブラリの個人用カスタマイズと認識結果の修正学習のシステムと有線マイク使用時における事前のトレーニングを活用せずに, ソフトを無作為に使用し, 静寂な個室で各被験者に対して計 5 回の音声入力をした。誤認識率は, 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 (図 2)。

#### 5. 統計学的分析

各群間の平均誤認識率は, 統計ソフト<sup>9)</sup> (ystat2013.xls, 東京) を使い, 有意水準 0.05 として Bonferroni-Holm 検定で判定した。

#### 6. 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。

表 4 ソフトとヘッドセットマイクの組み合わせ (9組)

Dragon Naturally Speaking Ver. 11 (DNS)
+ Nuance Dragon USB Headset Microphone (DNS-N)
+ Sennheiser SC 230 USB (DNS-S)
+ Plantronics Voyager Legend UC (DNS-P)
AmiVoice® SP2 (AS)
+ Parrott TalkPro USB 100 (AS-100)
+ Parrott TalkPro UC1 (AS-1)
+ Jabra PRO 935 (AS-J)
AmiVoice® Ex7 (AE)
+ Parrott TalkPro USB 100 (AE-100)
+ Parrott TalkPro UC1 (AE-1)
+ Jabra PRO 935 (AE-J)

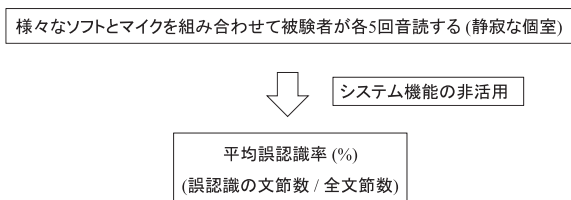


図 2 音声誤認識率の測定法

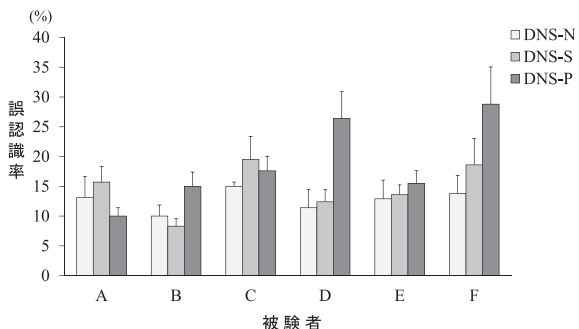


図 3 各被験者別の平均誤認識率 (DNS)

結果

1. 各被験者別の平均誤認識率

DNSにおける被験者別の平均誤認識率を図3に示す。全被験者における平均誤認識率の範囲は、8.3～28.8%で広く高い値を示した。ASにおける被験者別の平均誤認識率を図4に示すように平均誤認識率の範囲は、8.3～25.5%でDNSと同様に広く高い値を示した。一方、AEにおける被験者別の平均誤認識率を図5に示す。平均誤認識率の範囲は、1～14.8%で広い値を示したが、すべての被験者においてDNSやASよりも全体的に低い値であった。また、被験者B, D, EおよびFは、同じマイクを使用しているにもかかわらず、ASとAEはマイク間の誤認識パターンが異なった。

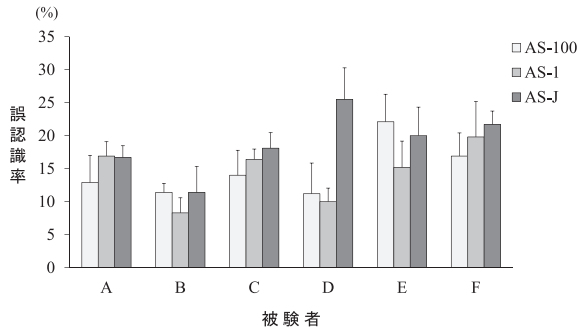


図 4 各被験者別の平均誤認識率 (AS)

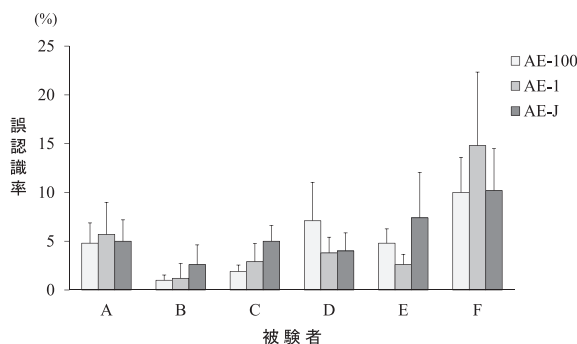


図 5 各被験者別の平均誤認識率 (AE)

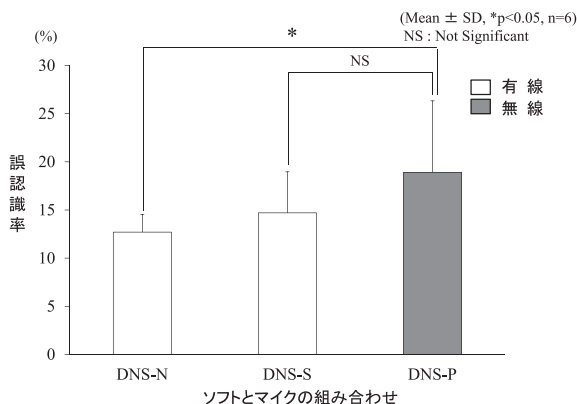


図 6 ソフトとマイクとの組み合わせ別の平均誤認識率 (DNS)

2. ソフトとマイクの組み合わせ別の平均誤認識率

DNSにおけるソフトとマイクの組み合わせ別の平均誤認識率を図6に示す。DNS-Nが12.7%, DNS-Sが14.7%, DNS-Pが18.9%であった。有線マイクの方が無線マイクよりも誤認識率は低く、DNS-NとDNS-Pとの間では統計学的な有意差 ( $p < 0.05$ ) が認められた。また、ASはAS-100が14.8%, AS-1が14.4%, AS-Jが18.9%であった。DNSと同様に、有線マイクの方が無線マイクよりも誤認識率は低く、AS-1とAS-Jの間では統計学的な有意差 ( $p < 0.05$ )

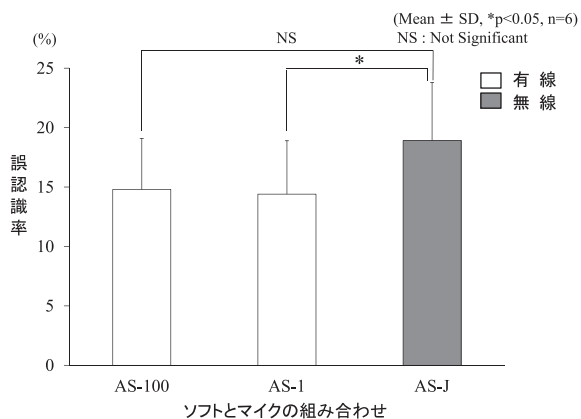


図 7 ソフトとマイクとの組み合わせ別の平均誤認識率 (AS)

が認められた (図7)。一方, AE は AE-100 が 4.9%, AE-1 が 5.2%, AE-J が 5.7% であった。有線と無線マイクとの間で誤認識率の有意差は認められなかった (図8)。

3. ソフト別の平均誤認識率

ソフト別の平均誤認識率を図9に示す。DNSが15.4%, ASが16%, AEが5.3%であった。AEはDNSやASよりも平均誤認識率が低く, AEとDNSおよびAEとAS間で統計学的な有意差 ( $p < 0.05$ ) が認められた。

4. 高い誤認識率が認められた認識困難語

各ソフト別の高い誤認識率を示した認識困難語を表5に示す。DNSとASは「腫れましたか」が「晴れましたか」, 「治した」が「直した」, 「診させて」が「見させて」, 「菌の」が「あの, その」, 「化膿止めや」が

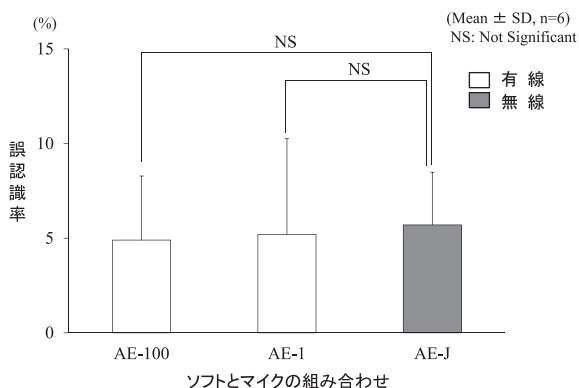


図 8 ソフトとマイクとの組み合わせ別の平均誤認識率 (AE)

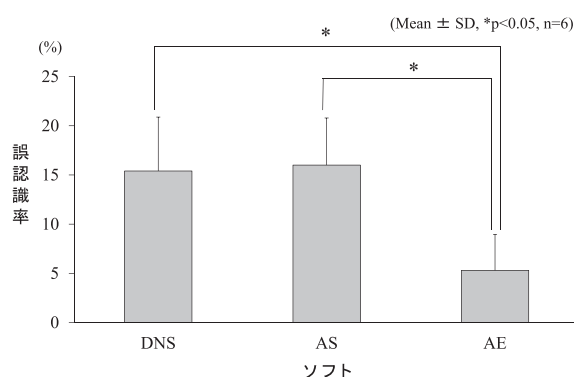


図 9 ソフト別の平均誤認識率

表 5 高い誤認識率を示した認識困難語

認識困難語 (誤認識率)	誤認識語 (例)	推測される原因
<u>DNS</u>		
腫れましたか (100%)	晴れましたか	同音異義
治した (97.8%)	直した	同音異義
診させて (94.4%)	見させて	同音異義
菌の (84.4%)	あの, その	システム
化膿止めや (78.9%)	和度目	システム
食べ物や (61.1%)	喰べ物	システム
<u>AS</u>		
腫れましたか (100%)	晴れましたか	同音異義
診させて (98.9%)	見させて	同音異義
治した (96.7%)	直した	同音異義
化膿止めや (95.6%)	狩野どめ, かのどめ	発音
菌の (84.4%)	はな	発音
<u>AE</u>		
腫れましたか (25.6%)	されましたか	発音
言い忘れた (23.3%)	言わせた, 言わせたこと	発音
ありましたか (15.6%)	ありました	発音



「和度目, 狩野どめ, かのどめ」, 「食べ物や」が「喰べ物」, 「治した」が「直した」に誤変換され, 高い誤認識率が認められた。一方, AE では「腫れましたか」が「されましたか」, 「言い忘れた」が「言わせた, 言わせたこと」, 「ありましたか」が「ありました」に誤変換され, 高い誤認識率が認められた。

## 考 察

近年, 音声認識ソフトを用いた音声認識技術はカーナビゲーション, コールセンターおよび聴覚障害者向け字幕呈示システムなどがあり, 我々の社会生活において様々な場面で利用されている<sup>10-12)</sup>。また, 医療領域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>13-16)</sup>。現在, 市販されている音声認識ソフトは多種多様であるが, なかでも DNS, AS および AE は, 一般人が購入し簡易に使用できる代表的なソフトである。DNS の特徴は, 話し手の声の特徴や話し方の習癖などを分析し事前登録するエンロールが必要であること, および汎用性が高く日常会話における一般語の認識が高いが, 医療用語の認識は低く登録が必要であることがあげられる<sup>3)</sup>。一方, AS や AE の特徴は, 不特定話者を対象にしているためエンロールの必要がなく即時に使えること, および色々な医療分野に特化したソフトがあり医療用語の認識は強いが, 自由登録単語数が少なく日常会話に用いる場合はやや不便であることがあげられる<sup>3, 5, 6, 17)</sup>。このように, ソフトにより長所と短所があり, その使用目的によりソフトを選択することが重要である。

一般に, 歯科医療面接におけるコミュニケーションは, 歯科に関連した言語と日常会話に関する一般語の両方が用いられる<sup>18)</sup>。

そこで, 本研究は歯科医療面接時の表示システムに用いる最適な音声認識方法を明らかにするため, DNS, AS および AE と様々なマイクを使用し, 初診時の歯科医療面接で用いる文章に対するソフトの誤認識率から最も有用なソフトとマイクの組み合わせを検討した。

その結果, 各被験者別の平均誤認識率において DNS, AS および AE の平均誤認識率の範囲は幅広い値を示したが, すべての被験者において AE は DNS や AS よりも全体的に低い値であった。また, 被験者 B, D, E および F は, AS と AE はマイク間の誤認識パターンが異なった。このことから, AE の誤認識率が全被験者に対して一定した低い値を示すこと, および同じ AmiVoice<sup>®</sup> でも AS と AE ではマイクとの相性に違いがあることが示唆された。

橋爪ら<sup>19)</sup>は AmiVoice に Bluetooth の無線ヘッド

セットマイクを試用し, 使い勝手が悪く認識率も低く実用に耐えないことを報告している。また, 三好ら<sup>20)</sup>は指向性マイク, 鼻覆い型および口覆い型マスクマイクを用い, 市販の音声認識ソフト 2 種類における読み上げ用文章に対する音声認識精度を測定した結果, 2 種類の音声認識ソフトともにすべての被験者において指向性マイク, 鼻覆い型マスクマイク, 口覆い型マスクマイクの順に認識精度が低下する傾向であることを明らかにしている。

このように, マイクのタイプにより誤認識率が異なり, またソフトとマイクとの相性に違いがあることが考えられたため, 本実験ではソフトとマイクの組み合わせ別の平均誤認識率を検索した。その結果, DNS と AS は有線マイクの方が無線マイクよりも誤認識率は低かったのに対して, AE は有線と無線マイクとの間で誤認識率の有意差は認められなかった。

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トポイントなどの問題があげられる<sup>21)</sup>。本実験では, すべての被験者が各ソフトについて同じシステム環境と使用場所で音声を入力した。また, 実験結果で DNS と AS における無線マイクの誤認識率が高かったのに対して, AE における無線マイクの誤認識率は低かった。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ト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。

松島ら<sup>22)</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, 赤外線を用いた無線マイクは, すべての騒音環境下において有線マイクと同様の言語認識率を有することを報告し, 赤外線を用いた無線マイクによる音声入力, 現在一般的に用いられている有線マイクと同様に有用であることを明らかにしている。本実験で用いた無線マイクは Bluetooth であり赤外線とは音声の転送方法は異なるが, AE は松島ら<sup>22)</sup>の報告と同様に無線マイクの有用性が示された。一方, DNS と AS は有線マイクの方が無線マイクよりも有意に誤認識率が低かったことから, 本実験で用いた無線マイクの Jabra PRO 935 は AE と相性が良く, 本研究では AE のみが無線マイクに有効であることが示唆された。

著者ら<sup>5)</sup>は, 難聴患者への音声入力文字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 AE と有線マイクを組み合わせで誤認識率を検討している。その結果, AE は単語登録と音響学習のシステムを活用した場合, 平均誤認識率が  $0.4 \pm 0.3\%$

であり、書き言葉に対して有用であることを示唆している。また、著者ら<sup>6)</sup>は話し言葉に対する AE の誤認識率も検討している。その結果、AE は単語登録と音響学習のシステムを活用した場合、平均誤認識率の範囲が 0.4 ~ 2.6% であり、話し言葉にも有用であることを報告している。さらに、栗岡ら<sup>3)</sup>は医療コミュニケーション言語を音声認識ソフトで認識させる場合の問題点を抽出するために、音声認識ソフト Dragon Naturally Speaking 2005 (ScanSoft) を用いて、日常診療用語の認識精度と誤変換について検討している。その結果、音声認識ソフト Dragon Naturally Speaking 2005 (ScanSoft) を用いた雛形文書の誤変換率は平均 3.6% であり、日常診療においても十分実用に供しうることを明らかにしている。

本実験におけるソフト別平均誤認識率の結果は、AE は DNS や AS よりも平均誤認識率が低く、AE と DNS および AE と AS 間で統計学的な有意差 ( $p < 0.05$ ) が認められた。したがって、AmiVoice の単語登録と音響学習や Dragon Naturally Speaking のボキャブラリの個人用カスタマイズと認識結果の修正学習のシステムを活用しない場合、音声入力文字表示システムに使用するソフトは AE が最適であることが示唆される。

しかし、今回、誤認識率が高かった AS と DNS も単語登録と音響学習やボキャブラリの個人用カスタマイズと認識結果の修正学習のシステムを活用した場合、使用するにつれて徐々に誤認識率が低くなることが報告されている<sup>23, 24)</sup>。したがって、ソフトの使用回数を増加することにより、日常臨床に実用可能なレベルに達すると思われるが、そのレベルに達するまでは単語登録や音響学習に時間を要する。それゆえ、購入直後より誤認識率が低い AE の方が AS や DNS よりも即時に使用できる有用なソフトであると考えられる。

また、本研究結果での DNS の誤認識率は栗岡ら<sup>3)</sup>の誤認識率の結果よりも高い値を示した。その理由として、本実験では音声入力前のエンロールを行わなかったことがあげられる。今回、エンロールを行わずに音声入力した理由は、臨床では即時に使うことなどの利便性が優先されるためである。

本実験で使用したソフトの認識困難語の原因として同音異義、システムおよび発音が推測された。そのうち、DNS と AS の認識困難語は、「腫れましたか」が「晴れましたか」や「治した」が「直した」などの同音異義が原因によるものが多かった。これは栗岡ら<sup>3)</sup>の報告である DNS の誤変換の原因に同音異義が多い結果と同様であった。DNS と AS は、日常用語に強いソフトであることから、本実験でも専門用語の多くが同音異義語の日常用語に変換されることは考えられ

る結果であろう。一方、AE の認識困難語は、「腫れましたか」が「されましたか」や「言い忘れた」が「言わせた」などの発音が原因によるものが多かった。AE は専門用語に強いソフトであることから、専門用語の誤変換は少なく、多くは発音によるものが認められたことも考えられる結果であろう。以前に、著者ら<sup>5)</sup>は初診時の歯科医療面接シナリオの書き言葉に対する AE の認識困難語における原因を検討している。その結果、原因として発音、システムおよび同音異義をあげ、なかでも発音が半数以上を占めていることを明らかにしている。本実験結果も認識困難語の原因として発音が多く認められ、話し手の声の韻律的要素が大きく関与していると考えられた。それゆえ、DNS と AS はシステム機能を活用し専門用語を単語登録すること、および AE では話し手がボイストレーニングを行うことが、より一層高い音声認識率を得る重要な要因であると考えられる。

以上のことから、表示システムに使用するソフトは、AE が最適であること、および AE に使用するマイクは有線と無線との間で差がなく、両方とも有用であることが示唆された。

今後、音声認識に関するユーザー側の要因や診療室での使用における環境の影響を検討し、さらに臨床の場で活用できる最適な表示システムを確立させる必要があると考えられる。

## 結 論

表示システムの確立の一端として、成人 6 名を対象に、歯科医療面接で用いる文章に対する様々な音声認識ソフトとマイクの認識率を検討した。その結果、表示システムに使用するソフトは AE が最適であること、およびそのマイクは有線と無線のいずれも有用であることが考えられた。

## 謝辞ならびに利益相反

本論文について利益相反はありません。

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 8 回日本総合歯科学会総会・学術大会 (平成 27 年 11 月 21 日、東京都) 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) 長嶋祐二. 視聴覚障害者のコミュニケーション支援. 電子情報通信会誌 2006; 89: 231-235.
- 3) 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.

- 4) 北原照代, 埜田和史, 西山勝夫. 聴覚障害者の受療に関する医療機関側の問題 ～医療機関を対象とした面接調査の分析～. 社医研 2001; 19: 45-55.
- 5) 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.
- 6) 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.
- 7) 河原達也. ここまできた音声認識技術. 情報処理 2000; 41: 1-4.
- 8) 小野木雄三. 連続音声認識を利用したレポートシステム. 新医療 28: 109-111.
- 9) 山崎信也. なるほど統計学とおどろき Excel® 統計処理. 改訂第7版. 東京: 医学図書出版; 2013. 78-125.
- 10) 角谷直子, 北岡教英, 中川聖一. 音声による地名入力タスクにおける誤認識時の言い直し発話の検出と認識. 情報処理学会研究報告音声言語情報処理 2002; 65: 17-22.
- 11) 古井貞熙. 音声認識の動向 [ I ] —話し言葉音声認識—. 電子情報通信会誌 2006; 89: 746-751.
- 12) 中野聡子, 牧原 功, 金澤貴之, 中野泰志, 新井哲也, 他. 音声認識技術を用いた聴覚障害者向け字幕呈示システムの課題—話し言葉の性質が字幕の読みに与える影響—. 電子情報通信会誌 D 2007; J90-D: 808-814.
- 13) 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.
- 14) 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.
- 15) 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.
- 16) 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.
- 17) 陣内自治. 一般病院における電子カルテの実際. JOHNS 2014; 30: 1699-1704.
- 18) 木尾哲朗, 大住伴子, 富永和宏, 飯塚 勝, 寺下正道, 他. 医療コミュニケーションのための患者の歯科用語理解度調査. 日歯医学会誌 2006; 22: 138-144.
- 19) 橋爪 崇, 宗近次朗, 池田真也, 笹森寛人, 後閑武彦. AmiVoice をちょっと便利に使う方法. Rad Fan 2011; 9: 44-45.
- 20) 三好茂樹, 黒木速人, 河野純大, 白澤麻弓, 石原保志, 他. 音声認識技術を利用した字幕作成担当者のための支援技術とそのシステム開発. 筑波技術大学テクノレポート 2007; 14: 145-152.
- 21) 石橋 透. ワイヤレスマイクの混信とデット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.
- 22) 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.
- 23) 岩崎 康. 汎用音声認識ソフトドラゴンスピーチ 2005. Rad Fan 2006; 4: 81-83.
- 24) 浮州龍太郎, 樋橋民生. AmiVoice —大学病院の7年間, 個人使用の4年間での雑感—. Rad Fan 2009; 7: 56-57.

#### 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  
大阪歯科大学総合診療・診断科  
TEL 06-6910-1066 FAX 06-6910-1026  
E-mail: tatsumi@cc.osaka-dent.ac.jp



Establishment of communication support for hearing-impaired  
elderly patients using a speech recognition system;  
Rate of speech recognition with different software programs and microphones

Hiroataka Tatsumi, Kyoko Higuchi, Takeshi Koide,  
Hiroyuki Kometani, Ikiko Tsuji, Mamoru Komeda,  
Akio Ohnishi, Tadasuke Tanioka, Chika Nakai,  
Kumi Hieda and Erika Iwami

Department of Interdisciplinary Dentistry and Oral Diagnosis, Osaka Dental University

**Abstract** : Medical speech display systems that use speech recognition software are recently being used as one method of communicating with hearing-impaired elderly patients. A key requirement of these systems is the ability to quickly process and accurately recognize speech.

To establish a medical speech display system that uses speech recognition, we investigated the rate of speech recognition by six dentists with different speech recognition software programs and microphones using lines of text from a dental interview.

The software programs used were Dragon Naturally Speaking Ver. 11 (DNS), AmiVoice<sup>®</sup> SP2 (AS), and AmiVoice<sup>®</sup> Ex7 (AE), and the microphones were two wired headsets and one wireless headset designed to be used with the software.

The mean recognition error rates for DNS, AS, and AE were  $15.4 \pm 5.3\%$ ,  $16 \pm 4.7\%$ , and  $5.3 \pm 3.6\%$ , respectively; the lowest rate being demonstrated by AE. There were statistically significant differences between DNS and AE and between AS and AE ( $p < 0.05$ ). The wired microphones were more effective than the wireless microphone for DNS and AS, but there was no difference in effectiveness for AE.

These results indicate that AE is the most suitable software program for use with medical speech display systems, and that both wireless and wired microphones can be used for that program.

**Key words** : speech recognition, hearing-impaired elderly patient, dental interview, wired microphone, wireless microphone