

**科学研究費助成事業 研究成果報告書**

平成 27 年 6 月 8 日現在

機関番号：12601

研究種目：若手研究(B)

研究期間：2012～2014

課題番号：24700588

研究課題名(和文) 発話障害者支援のための音声生成器に関する研究

研究課題名(英文) A speech-generating device for speech disorders

## 研究代表者

藪 謙一郎 (Yabu, Ken-ichiro)

東京大学・高齢社会総合研究機構・特任研究員

研究者番号：50626215

交付決定額(研究期間全体)：(直接経費) 2,000,000円

研究成果の概要(和文)：本研究では、まず操作盤面をなぞる指の位置やその動きに応じて、声を楽器のようにリアルタイム生成できる音声支援装置について改良を行なった。具体的には、入力速度を補うアルゴリズムを追加することで子音部分のホルマント遷移部分の急激な周波数変化を補うことが可能になった。また、標準的な音楽用ファイルの読み込みを可能として伴奏音と同期した歌声生成を可能とすることでエンタテインメント利用の有用性も示した。

さらに、本研究を基礎として開発されたスマートフォン用アプリを、神経・筋疾患を持つ患者に使用してもらい、その評価を通じてその有用性と問題点を整理した。

研究成果の概要(英文)：A speech-generating device that translates the motion and position of a user's finger or a stylus on the surface of the controller to speech sound in real-time like a musical instrument was improved. Concretely, rapidly changing resonant frequency in the formant transition of a consonant is complemented by a input speed complementing algorithm. In addition, application for use in entertainment was shown by the system which generates the singing voice in synchronization with musical accompaniment generated from a standard data file for electronic musical instruments. Moreover, a patient with neuromuscular disorder tried out a smartphone app that was developed on the basis of this research. The effectiveness and problem was extracted from the evaluation by the patient.

研究分野：福祉工学

キーワード：音声 支援機器 音声合成 ホルマント

1. 研究開始当初の背景

発話障害を支援する従来の音声機器として、テキストデータや単語を読み上げる方式が一般的である。これらは、正確に意味内容を伝えるという側面では優れている。一方で、文字等の入力のために発話開始に数秒の遅れが生じる、リズムの付加が困難、文字で表せない中間的な音の表現などができないといった課題がある。

代表者らは腹話術師や物真似鳥の発声原理の仮説に基づき、「舌の動き」を「指の動き」に置き換えて音声合成のパラメータを制御する新しい発話支援方式を提案<sup>①②</sup>してきた。これは、使用者の指の位置やその動きに応じた声を、楽器のようにリアルタイム生成できる支援装置である。

本課題の先行研究として、試作機を通じて、母音を基礎とした音声生成方式や、擬似子音を付加する方法、簡易的な抑揚付加の方法を提案し、開発を進めてきた。

具体的には、第1,第2ホルムント(F1, F2)を縦横軸として母音位置を表示した操作盤面上を、指やペンでなぞって音声を生成する。さらに、その速度や位置に変化をつけ子音音声を生成する。

音声の生成は Klatt ら<sup>③</sup>のホルムント合成器を応用し、第1,第2ホルムントを操作盤面からの入力位置で制御する。また、母音と同様のフィルタを通した雑音を適切なタイミングで付加することによって擬似子音を生成する。

ペンまたは指を適切に動かすことにより、より柔軟な声を生成できることを目指した方式であるが、子音らしい音を生成するためには、適切な軌跡を適切な速度でなぞって操作する必要がある。本課題は、この方式に改良を加え、初心者でも扱いやすくするとともに、当事者による試作機の評価から有用性と問題点を整理することを目指したものである。

2. 研究の目的

本課題は、開発中の発話支援方式について、センシングされた指の位置やその動きからより明瞭な音声を生成可能にするために、パラメータ変換手法を改良することと、当事者による試作機の評価を通じてその有用性と問題点を明確にすることを目的とした。

3. 研究の方法

本課題の音声生成器は、図1に示すような操作盤面をペンまたは指でなぞって操作するものである。縦軸と横軸とは、それぞれ第1,第2ホルムントに対応しており、図2に示すようなホルムント合成器により、音声生成される。図2中のFv1-Fv4とFc1-Fc4はそれぞれ母音用と子音用の共振フィルタを示す。また、音源としてはヒトの声からLPC逆フィルタによってホルムント成分を除去した音声を用いている。

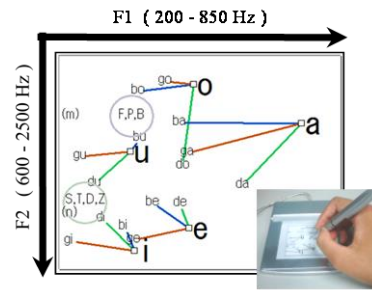


図1 音声生成器の操作盤面

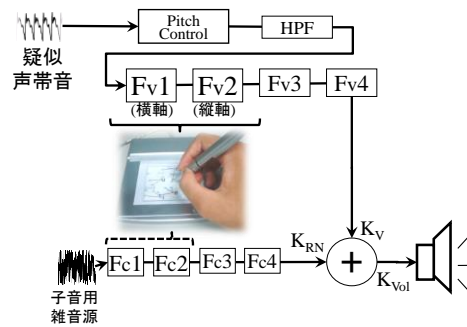


図2 音声生成の処理

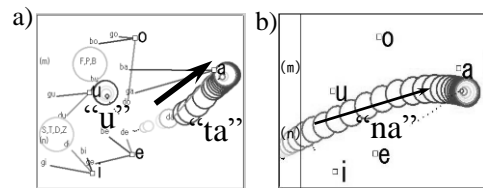


図3 「うた」と「な」生成時の入力軌跡

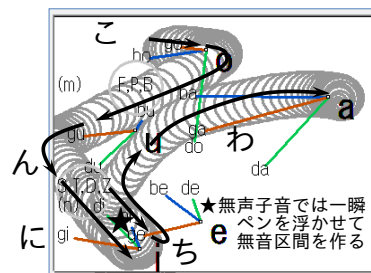


図4 「こんにちは」生成時の入力軌跡

操作盤面には、子音らしい音声を生成できるように、先行研究により求められた、「導線」が表示してある。図3や図4に示すように、導線を辿るようにペンや指を動かすと子音らしい音が生成される。

平成24年度から25年度には、試作機をWindows8のタブレット端末でタッチパネル操作をスムーズに行えるようにしたのち、操作方法習得の個人差を吸収できるような訓練手順や操作盤面上の補助機能追加を試みた。そして、歌声を生成できる楽器として

の利用を可能とする改良を加え、健常者を含む利用可能性を探った。

また、先行研究の成果を基に、北海道立総合研究機構の橋場参生氏と株式会社電制の協力により、スマートフォンアプリの開発を進め、一般の利用者が入手できるようになった。

平成26年度には、筋・神経系に疾患による発話障害者を対象とした場合に、言語的・非言語的な情報がどこまで表出できるかを、患者にスマートフォンアプリを利用してもらいながら、家族の意見を聞くことで調べ、考察した。また、操作盤面上の補助機能について、動作の安定性を高め、入力インターフェースの使用者への個別的な調整方法を検討した。

以上を踏まえて、対象となる障害の観点、明瞭さと即時性の観点、音の自由さの観点から、考察を加えて既存の音声支援機器とともに本課題の音声生成方式の位置づけを示し、今後の課題を整理した。

#### 4. 研究成果

##### (1) アプリ練習画面と試作機速度補助機能

先行研究の成果をもとに、本研究の音声生成方式で声を出せるスマートフォンアプリの調整を続け、一般の方でも利用できる形で平成25年4月にダウンロード可能となった。

アプリのメイン操作画面では、F1-F2の操作盤面に相当するエリアと、破裂音、摩擦音等を生成するための補助ボタンが設けられている。F1-F2エリアを指でタッチするとその瞬間に音が出力され、動かすとそれに対応して音も変化する。子音をだす場合は、例えば、「カ」を出したい場合には、「k/t」のボタンを押してすぐ、「g」-「a」の導線をなぞると、ノイズオンによる疑似的な子音が付加された音声生成される。さらに、本体を傾けることによって、声に抑揚をつけることも可能になっている。指を動かす速さや、速度変化によって、声のリズムや明瞭さが変化する。すなわち、高速に動かすほど高い子音らしさが得られる。

本方式では、指でタッチしている位置に対応するホルマント周波数がリアルタイムに生成されるため、指を動かす適切な軌跡をあらかじめ利用者が学習しておく必要がある。アプリでは、サンプル練習モードの機能が含まれており、50音をはじめ、あいさつ、曜日などの単語を補助線を見ながら練習することができるようにした(図6)。

並行して汎用PC上では、速度補完機能の改良と実験を進めた。スマートフォン向けアプリでは、破裂音、摩擦音等を生成するための補助ボタンが設けられたが、本研究では、出来る限りボタン数やタッチコマンドの操作を減らして単純化し、自由な声を出せること目指すため、汎用PCによる試作機では、子音を出す際にも、ペンの動きの速度や筆圧

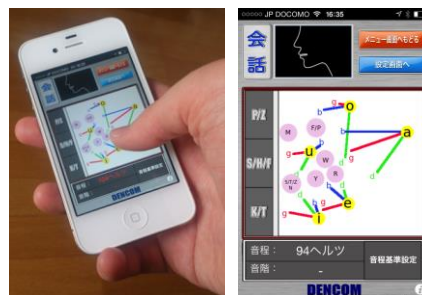


図5 スマートフォン向けアプリ



図6 アプリの練習モード画面

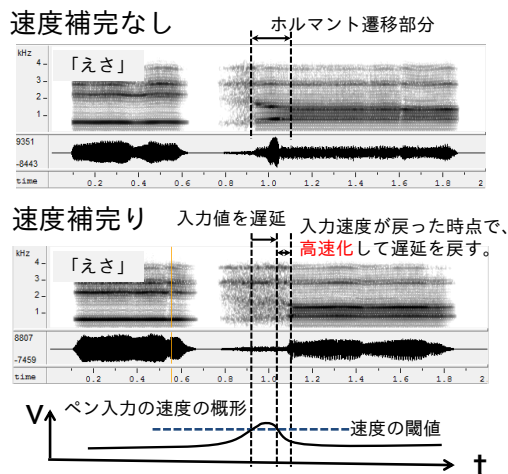


図7 速度補完機能の音声波形とスペクトログラム

の変化だけで操作できることを目指している。具体的には、無声子音については図4の「ち」に示されるように一瞬操作盤面からペンを離すと疑似子音が自動挿入されて無声子音が再現される。流音などを含む有声子音では、なるべく高速に操作することで再現をしている。有声、無声に関わらず、より子音らしい音を再現するには、高速に操作することが必要であった。

そのため、なぞる速度を補完する機能の付加が必要であったが、その際、新たなボタン等を追加せずに速度補完のタイミングをペンや指の入力合わせるアルゴリズムが問題になる。操作盤面の座標上に「子音補完領域」



を設けて、その領域を通過するとき子音の速度補完を行うということも考えられたが、子音から母音への接続と、母音から母音への接続とを両立できる設計が困難であった。

そこで、ボタンやコマンド操作を必要とせずに、ユーザの自然な操作で速度補完をする操作方法として、速度閾値による手法を考案した。すなわち、タッチ点の移動速度が閾値速度を超えた場合に、(1)超えた時点からの入力値を遅延させ入力値バッファに記憶させておき、(2)タッチ点の移動速度が閾値を下回った時点でバッファに記憶された入力値を高速に再生させるという方法である。閾値は、ユーザの慣れや全体てきな操作速度に応じて変更可能とした。

「えさ」を生成した時の速度補完ありの場合と無しの場合との生成音を図 7 に示す。タッチ点の移動速度が閾値以上になった場合に、以降の入力を無視させて音の変化を留め、閾値以下の速度になった時点で、前の入力座標から現在の入力座標までを約 40ms の時間で遷移させ速度補完が実現できている。

### (2) 歌声生成によるエンタテインメント利用

歌声の生成については、一般の音楽データとして標準化されている Standard Midi File (以下 SMF) を読み込み再生可能とし、データによる伴奏音と、指で操作して生成する歌声とを同期再生する方法で実現した。SMF 内の伴奏データは PC 上のソフトウェアシンセサイザに送信され、伴奏音声となり、主旋律データは音声生成器の原音再生の制御に用いられ、適切な周波数の原音を生成する。利用者は、主旋律の音程となった原音に操作盤面からの入力で歌詞をつけることによって歌声を生成する。また、主旋律の MIDI チャンネル選択、オクターブの調整などにより任意データを利用可能とした。これにより、市販の任意のカラオケデータ等からも歌声を生成でき、発話障害を持つ方々の娯楽やリハビリへの応用、健常者のためのエンタテインメント利用に効果的と考えている。

### (3) 重度の筋・神経系患者のアプリ利用

本課題の音声生成方式は、文字や言葉に頼らなくても非言語の形で音を出せる特徴がある。重症度の高い患者の場合には、スマートフォンによる本方式による音声生成アプリの非言語の特徴を活かせるのではないかと考えられた。

そこで、難病を持つ患者 1 名に試用してもらい、保護者からのご意見を伺った。実験参加者は、先天性の難病により、意識レベル、視力、聴力がいずれもやや低下していて、識字も困難な状態であり、普段のコミュニケーションは、独自の簡単な手話やわずかな声、うなずきなどで行っている。

実験の様子と保護者のご意見から、開発当初から期待されていた、文字や絵による意味情報を用いない非言語の音による新しいコ

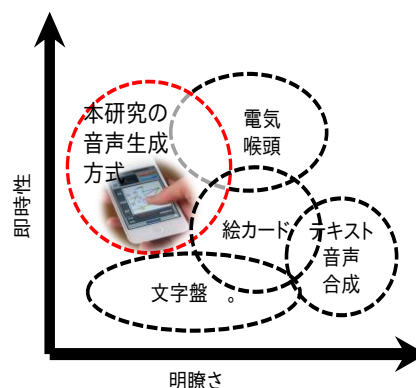


図 8 各音声支援機器の明瞭さと即時性の位置づけ

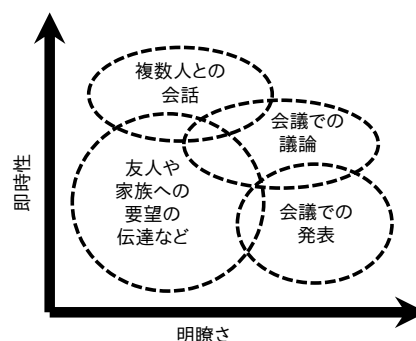


図 9 利用場面ごとに要求される明瞭さと即時性の違い

コミュニケーション手段の可能性に加えて、運動機能のリハビリや、音での他者との交流することによる精神的な活性化につながる可能性が期待できることが明らかになった。本音声生成器では、言葉を完全な形で出そうとすると相応の訓練が必要になるが、単に「あー」「うー」「いや」などの声を出すのであれば、触るだけで扱うことができる。重度の患者さんの場合、他者とのコミュニケーションが極度に少なく、座ったままほとんど外部からの刺激がなくなってしまう場合も多いため、本方式のように触るだけで音を出すことのできる音声機器が、このような精神的な刺激としてのツールの役割も担える可能性が示された。

### (4) 各音声支援機器と課題の整理

以上を踏まえて、対象となる障害の観点、明瞭さと即時性の観点、音の自由さの観点から、考察を加えて既存の音声支援機器とともに本課題の音声生成方式の位置づけを示し、今後の課題を整理した

#### ① 対象の障害について

本方式の操作盤面はヒトの「構音」の機能の再現を目指すものである。したがって筋・神経・脳の疾患で構音器官がうまく動かせなくなった患者や、器質性の異常があってもうまく話せない場合を対象としている。この点は、声帯の機能を代替する電気式人工喉頭は異

なる範囲となる。また、本方式は「文字」に頼らなくても声を出力できることから、文字認識が困難な方でもある程度利用できる。

### ② 明瞭さと即時性

現状では、本方式を用いて言葉をテキスト音声合成のように流暢かつ明瞭に伝えることは難しく、任意の単語を生成できるようになるには、ある程度の練習も必要となるが、その一方で指が画面に触れた瞬間に音が発せられることから、即時性は極めて高い。

現存の音声支援機器とともに、即時性と明瞭さの観点での本研究の音声生成器の位置づけを図 8 に示す。また図 9 には、想定される場面ごとに、必要とされる即時性と明瞭さについて概念的に図示する。

2つの図のような概念図を比較することで、場面ごとに適切な支援機器が推測される。

会議での発表などでは、なるべく明瞭に伝えることが必要であり、時間をかけて明瞭な声を準備できる音声合成が適切である。

複数人での雑談のような場合には、次々と話者や話題が移りかわっていくため、即時性が極めて重要となる。5秒でも間(ま)があれば、すぐに他の話者が割り込んで話しはじめて話題が変わるため、それ以前に一声を出さなければ、話に乗り遅れることとなる。

複数人であっても、ある程度話題が決まっていて、話者と親しい聞き手が注意深く耳を傾けられるような場合には、明瞭さよりも即時に円滑に伝えられることが求められる。また、日常の生活で他人に何か必要なものを要求したり聞かれた質問に答えたりする場合、周囲の状況や身体の状態からある程度内容が限られる。このような場合にも、明瞭さよりも即時性や円滑さが優先されると考えられる。明瞭さよりも即時性や円滑さが求められる場合には、発話に障害を持つ患者が、自らの発話が不明瞭でうまく伝わらないことがあっても自分の声で話すことを好む場合が多い。F1-F2による本方式は、おもに、その状況を補助するのにも有効であると考えている。

### ③ 音の自由さ

筋・神経系や脳に障害を持つ重症な患者の場合、全身に症状がありキー入力や文字の認識が困難である場合や、はっきりとした意味情報の表現が困難である場合がある。その場合、家族と本人との間だけで通じる独自のジェスチャーや声の表現によって、意思が伝達されることがある。

本方式の音声生成は、母音と母音との間の中間的な音を出すことができ、変化の速度を自由に変えられるため、自由度がかなり高い。また傾きで抑揚を制御することも可能である。このような自由度をうまく使うことで、「言葉」という型にとらわれない自由な使い方も可能である。また、単に音としての表現を楽しむという利用方法もできるであろう。

### (5) まとめ

本研究では、操作盤面をなぞる指の位置やその動きに応じて、声を楽器のようにリアルタイム生成できる支援装置について、改良を行なった上で、障害を持つ当事者に使用してもらい、有用性と問題点を整理した。また、一般の方も利用できる形で、スマートフォンアプリがダウンロード可能となった。

今後、汎用 PC の試作器で実装された速度補完機能等をアプリへ組み込み、改良を加えていくと同時に、試作器においては、よりヒトの発話に近い声を表出できるような方法を探り、重度の障害者でも自由に発話できる支援機器を目指していきたい。

### <引用文献>

- ① 橋場 参生, 藪 謙一郎, 井野 秀一, 泉 隆, 伊福部 達, “F1-F2 平面で操作する会話補助装置の提案と試作”, ヒューマンインタフェースシンポジウム 2005 論文集, pp. 585-588, 2005.
- ② 藪, 謙一郎 伊福部, 達 青村, 茂, “発話障害者支援のための音声合成器の基礎的設計”, 信学技報, .SP, 音声, 105, pp:59-64, 2006.
- ③ Klatt, D.H., “Software for a cascade/parallel formant synthesizer, Journal of the Acoustical Society of America, 67(3), pp. 971-995, 1980.

### 5. 主な発表論文等

(研究代表者、研究分担者及び連携研究者には下線)

[学会発表] (計 7 件)

- ① 藪 謙一郎, 伊福部 達, ペンの位置と動きで操作する音声生成器による本人の声を活かした発話補助方法の考察, 音響学会 2015 年春季研究発表会, 2015 年 03 月 16 日~18 日, 中央大学 (東京都) .
- ② 藪 謙一郎, 橋場 参生, 高瀬 雅由, 須貝 保徳, 伊福部 達, 指の位置と動きで操作するスマートフォン向け音声生成アプリの応用例の考察, 日本音響学会 2014 年秋季研究発表会, 2014 年 09 月 03 日~05 日, 北海学園大学 (北海道) .
- ③ 藪 謙一郎, 伊福部 達, タッチ点の動きで操作するリアルタイム音声生成方式の速度補完の考察, 音響学会 2014 年春季研究発表会, 2014 年 03 月 10 日~12 日, 日本大学 (東京都)
- ④ 橋場 参生, 藪 謙一郎, 高瀬 雅由, 須貝 保徳, 伊福部 達, 指の位置と動きで操作するスマートフォン向けリアルタイム音声生成アプリの開発, ヒューマンインタフェースシンポジウム 2013, 2013 年 09 月 10 日~13 日, 早稲田大学 (東京都) .

- ⑤ 橋場 参生, 発話困難者のためのリアルタイム音声生成アプリの開発, 北海道立総合研究機構 工業試験場 成果発表会, 2013年5月17日, ホテル札幌ガーデンパレス(北海道)
- ⑥ 藪 謙一郎, 伊福部 達, タッチ面上のなぞりで操作する音声生成器—メロディデータに基づく歌声生成機能, 電子情報通信学会2013年総合大会, 2013年03月19日~22日, 岐阜大学(岐阜県).
- ⑦ 藪 謙一郎, 伊福部 達, タッチ点の動きで操るリアルタイム音声生成器の開発—視覚障害者にも扱える楽器のような音声支援機器へ, ライフサポート学会視聴覚障害者バリアフリー技術研究会研究発表会 2012, 2012年11月02日, すみだ産業会館(東京都).

[研究報告] (計 1 件)

- ① 橋場 参生, 高瀬 雅由, 須貝 保徳, 藪 謙一郎, 伊福部 達, 発話困難者のためのリアルタイム音声生成アプリケーションの開発, 工業試工業試験場研究報告, 312, pp.105-108, 2013.

[その他]

- ① ゆびで話そうサポートページ(電制)  
<http://www.dencom.co.jp/yubihana/yubihana.htm>
- ② ダウンロードページ(iTunes store)  
<https://itunes.apple.com/jp/app/id626142603>

## 6. 研究組織

### (1)研究代表者

藪 謙一郎 (Ken-ichiro Yabu)  
東京大学 高齢社会総合研究機構  
特任研究員  
研究者番号 : 50626215