

た単語が何であるかを記入する方法で行った。音声の提示は、汎用 PC に接続されたアームレスヘッドホン Victor HP-AL202 を通して行われた。

聴取者には、実験の前に提示音声が日本語の 4 音節の単語であることを教示し、単語が推測可能な場合には、聞こえたままの文字を記入するのではなく、その単語を答えるように指示した。

聴取実験の順序効果の影響を考慮し、生成された 36 語の単語を、A 群と B 群とに分け、A 群の鼻音あり生成音と B 群の乱流音なし生成音をペアとして音声群 I とし、B 群の乱流音あり生成音と A 群の乱流音なし生成音をペアとして音声群 II の音声とした。音声群 I と II は、それぞれ全く同じ単語が含まれ、鼻音処理の有無が入れ替わった群である。聴取実験では、2 名の聴取者には音声群 I の実験の後に音声群 II の実験を行い、残り 2 名聴取者には、音声群 II の実験のあとに音声群 I の実験を行った。

4.3. 結果

図 10 に聴取実験の結果を示す。また、聴取者 4 人の単語理解度の平均値と、先行の 100 語による実験[5]の結果の中のうち今回の実験で用いた 36 単語の 6 名の聴取者の理解度の平均値を、図 11 に示す。

聴取者 4 人のうち、3 名が鼻音なし音声よりも鼻音なし音声のほうが理解度が高かった。4 人の平均値は、鼻音なし音声で 45.8%、鼻音あり音声で 47.2%と、わずかに鼻音あり音声のほうが良い結果となった。また、これらは先行実験の 17.2%の理解度と比較して、高い結果となった。

4.4. 考察

鼻音なし音声と鼻音あり音声との理解度の比較から、鼻音処理の付加の効果がある程度見られた。また、実験では、先行実験で理解度が低かった単語を採用したにも関わらず、先行実験に比べて理解度が倍以上の結果となった。鼻音なし音声でも理解度が高かったのは、生成時のペンの軌跡が、鼻音領域の設定によって変化したためと考えられ、操作面上における鼻音領域の設定が妥当なものであることが示唆される。

聴取者 4 で鼻音なし音声のほうが理解度の高い結果となった理由として、実験の順序効果が考えられる。鼻音の有無を入れ替えた 2 回の実験のうち、いずれの被験者も前半の実験で正当したものは、鼻音の有無に関わらず後半にも正当する傾向にあった。被験者 4 は、後半に行われた鼻音なし音声の単語に正当数が増えたため、前後半合計の鼻音なし音声理解度がやや好成績となった。図 12 に、鼻音の有無に関わらず、前半の実験と後半の実験とで比較した図を示す。前半の実験の平均は 43.7%であったのに対し、後半は 49.3%と、高

かった。このことから、聴取者の慣れが生じることが示唆される。

了解者が 0 名だった 5 語を見ると、アクセントの影響が強いと思われる単語が含まれており、聴取者からもアクセントが合致していない音は聞き取りづらいという意見が合ったため、今後の課題として挙げられる。

5. まとめ

本研究では、発話障害者の日常会話を支援する音声生成支援機器として、ペンタブレットやタッチパッドで入力された座標値を音声生成パラメータへ直接対応させた、リアルタイムの音声生成器を提案・開発している。これまでに、生成音声の明瞭度を徐々に改善するべく、改良を加えている。本稿では特に、鼻子音に関する実験について報告し、ある程度の効果が見られたことを示した。音声の明瞭度と出せる音の自由さは、トレードオフとなる側面があるが、今後も、できるかぎり単純な操作で、多様な音を思い通りに生成できるように、改良を加えていきたい。

文 献

- [1] 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(聴覚・音声・言語とその障害,一般)",信学技報, .SP, 音声, 105, pp:59-64, 2006.
- [2] ジャック・ライアルズ著, 今富撰子, 荒井隆行, 菅原勉監訳/新谷敬人, 北川裕子, 石原健訳: 音声知覚の基礎.海文堂出版, 東京, 2003
- [3] 藪 謙一郎, 伊福部 達, 青村 茂, ポインティングデバイスを利用した音声生成方式-発話障害者のための支援機器として-, 日本保健科学学会誌, 12(1), pp.49-57, 2009
- [4] 藪 謙一郎, 青村 茂, 伊福部 達, ポインティング・デバイスで操作する発話支援インタフェース, ヒューマンインタフェース学会誌, Vol.11 No.4, (135-146) 2009
- [5] 発話障害者支援のための連続タッチ平面で操作する音声生成器--子音改善のための基礎的検討, 藪 謙一郎, 伊福部 達, 信学技報, 110(221), 41-46, 2010
- [6] 難聴者のための単語理解度試験用単語リスト <http://www.ais.riec.tohoku.ac.jp/lab/wordlist/indexj.html>
- [7] ピー・アクティブ P-Active : XP-3300A ポケットサイズペンタブレット <http://www.p-active.com/product/pt/3300a.htm>
- [8] 伊福部, 達, 橋場参生, 松島純一. 母音の自然性における「波形ゆらぎ」の役割. 日本音響学会誌 47, no. 12: 903-910., 1991
- [9] Klatt,D.H.: Software for a cascade/parallel formant synthesizer: Journal of the Acoustical Society of America, pages 971-995, volume 67, number 3, March 1980.
- [10] 坂本修一, 鈴木陽一, 天野成昭, 近藤公久: 親密度と単語の音韻バランスを統制した単語理解度試験用リストの構築: 東北大学電通談話会記録第, 69(2), pp.21-34, 1998.

「なぞり」の入力による発話障害者支援のためのリアルタイム音声生成器 — 障害音と鼻音付加の操作方法の検討 —*

◎ 藪 謙一郎, 伊福部 達 (東大)

1 はじめに

ヒトは発話する際、構音器官の筋肉を巧みに操って、リズムや間(ま)や抑揚などを付加した場に適した声を、リアルタイムに出すことができる。それによって、1対1または1対多の話者間で対話的なコミュニケーションを円滑に行うことが可能となっている。

発話障害を持つ場合、筆談や絵カード、音声合成による支援機器などを用いてコミュニケーションの機能を代替する。どのような代替手段が良いかは、個々の状況や障害の程度に応じて選択される。

TTSの音声合成による支援機器は、そのうちで最も代表的なものであり、かつ有効な手段となっている。一方でTTSには、出力される合成音声の話者や発話様式が限られてしまい、個性のある表現や多様な感情表現がしづらいという問題も残されている。^[1] また、日常の対話では、入力に要する時間も大きな問題となる。例えば、複数人の対話中では、話し出しに数秒の間があるだけでも、別な話題が立ち上がるのに十分な時間となり、話そうとしたことを話し出せなくなる場合がある。

そこで我々は、ヒトが構音器官を操るのと同じように、手や指などの身体の一部の動きで操作でき、より自由な音声生成されるような発話音声支援機器が必要であると考えている。そして、ユーザ自身がペンや指で平面を押したりなぞったりして、楽器のようにリアルタイムで音声を生成できる音声生成器を提案し^[1] PC上で動作する試作器の開発をしてきた。

本稿では、その音声生成器の試作器の概要を述べた上で、新たに取り組んでいる、子音の付加とそれを制御するための簡易的な方法の実装について述べる。

2 提案している音声生成器の概要

2.1 入力用操作面と母音の生成

母音はヒトの発話で最も基本となる。ヒトは、ホルマント周波数の違い、特に、第1ホルマント周波数(F1)と第2ホルマント周波数(F2)の違いに応じて母音を知覚できることが知られている。^[3] また、そのホルマント周波数は、母音を発声する時の舌の位

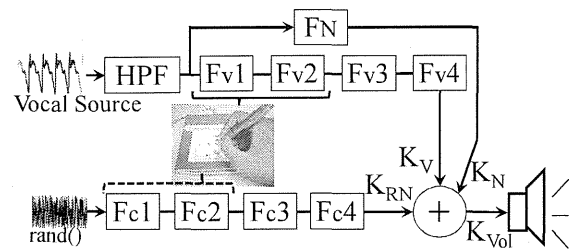


Fig. 1 Block diagram of the synthesizer.

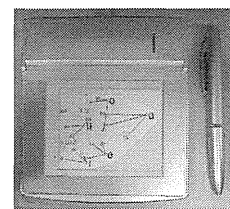


Fig. 2 A input device to control the voice synthesizer. (P-Active XP-3300A)

置に関係している。^[4]

そこで本研究では、使用者が制御するパラメータとして、F1とF2の2つのホルマント周波数を採用した。F1を入力操作面上の横軸に、F2を縦軸に割り当て、入力操作面上をペンや指で指したりなぞったりすると、その点に対応するF1、F2を持つ音声連続的にリアルタイムに出力される。この方式は、構音器官(舌)の動きに相応したものを、ペンの動きによって表現して、音声を生成することを意識したものである。

音声生成処理のブロック図をFig. 1に示す。図中のFv1~Fv4が母音を特徴づけるためのデジタル

Table 1 Formant frequencies of vowels

母音	F1	F2
/a/	760	1200
/i/	370	2340
/u/	360	1450
/e/	500	2140
/o/	510	850

* A real-time speech synthesizer to support speech disabilities by tracing input
-A deliberation of method to control obstruent and nasal consonants-
by YABU, Ken-ichiro, IFUKUBE, Tohru (The University of Tokyo)

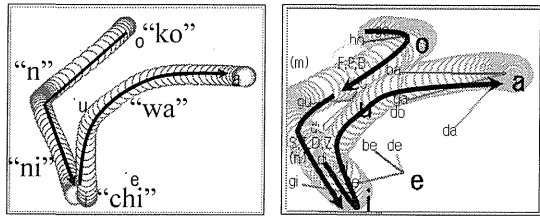


Fig. 3 Movement of the pointer to generate "kon-nichiwa".

共振器である。また、HPFは高周波数域を調整するハイパスフィルタである。各共振器は、Klattらが用いたもの^[5]と同様のものを用いた。母音用の原音波形としては、話者の振幅やピッチの揺らぎを維持するため、筆者の約3秒間のあーという録音音声にLPC逆フィルタ処理したものを用いた。図中の $F_{c1} \sim F_{c4}$ と F_N は、後述の阻害子音と鼻子音のための共振器である。

入力デバイスとして、Fig. 2に示す小型のペンタブレット^[6]を用いた。ペンタブレットの座標の横軸には F_1 の200~850 Hz、縦軸には F_2 の600~2500 Hzを割り当て、画面上と操作面上には、Table. 1の周波数の位置に対応する場所へ母音の位置を表示させた。また、ペンの筆圧を音量へ割り当て、入力操作面上へペンを置いた瞬間にその点の座標に対応する F_1, F_2 を持った声連続的に即座に発せられるようにした。ペンを放せば、その瞬間に声が停止する。

以上のような音声生成器を用いて音声の生成を行ったところ、Fig. 3の左図に示すように、ペンで単語の母音を辿るだけでも、ほぼ誰が聞いても「おはよう」「こんにちは」等の簡単な単語を知覚できる程度の音声を生成することができた。^[7] また、無声子音が存在する箇所ではペンを瞬間的に面から放し、100~200 ms程度の無音区間を挿入したり、母音の前に、ある角度へ速度を持たせて面をなぞると、より子音らしく聞こえることが分かった。

2.2 子音のための「導線」

子音を生成するための複雑な処理を導入していなくても、子音を含む簡単な単語を生成できた要因として、子音に存在するホルマント遷移が挙げられる。

そこで、子音に近い音を出せるホルマント遷移を、使用者がペンの軌跡で再現可能にするため、Fig. 4に示すような「導線」を提案した。それぞれの子音に対応した「導線」に沿って、母音の方向へ向けて一定の速度でペンを動かすと、子音のホルマント遷移に近い音声を出力することができる。各導線は成人男性の発話音声の分析と、合成音声の聴取実験から得られたものである。

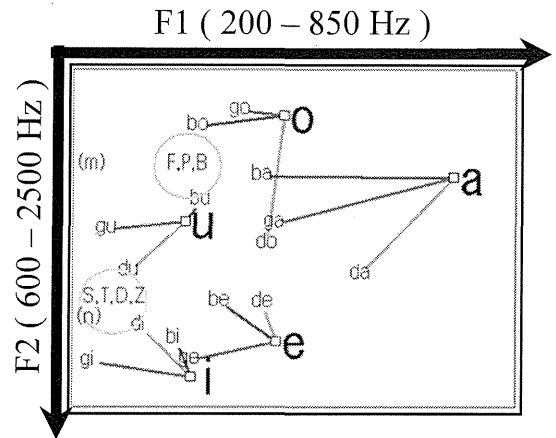


Fig. 4 Guiding lines for each formant transition.

また、全ての子音について表示すると複雑になってしまうため、ホルマント遷移の軌跡が似ている、同じ構音位置の子音では同じ「導線」を共通に使用することとして表示させた。すなわち、/k/と/g/、/s/と/z/と/t/と/d/、/p/と/b/では、同じ導線をたどる。ただし、/h/は構音位置が声門であり、ホルマント遷移がほとんどないため、表示していない。また、前述の場合と同様に語中にある無声子音の場合には、ペンを一瞬だけ操作面から離し、約100~200msの無音区間を子音の前に付けると、子音に近い音声となる。

先に行われた、坂本らによる単語リスト^[8]の中から文脈無く提示された100語による4モーラ単語の生成・聴取実験では、1~2割程度の単語理解度を得ることができた^[9]。しかし、特に無声子音を多く含んだ単語は理解度が低く、課題となっていた。

3 子音生成のための検討

3.1 擬似子音 (阻害音) 生成のための操作方法

そこで、Fig. 1の $F_{c1} \sim F_{c4}$ で示される乱流音生成部を加えた。原音は乱数関数によって生成される雑音とし、母音生成部と同様のフィルタにより簡易的な乱流の音を生成した。そして、上述の「導線」を用いた場合と同じ操作で、擬似子音である乱流音を生成できるようにした。さらに、ペンが操作面から離されてから300msの間は乱流音モードとなり、乱流音モードの間にペンが置かれた場合には、乱流音が生成されるようにした。すなわち、語頭でない無声子音の場合には、一瞬ペンを離しすぐに置くと乱流音が生成され、擬似子音が付加される。語頭で乱流音を生成させたい場合には、一瞬だけ弱くペンを置いて話したあとに、ペンを置くと、乱流音を生成することができる (Fig. 5)。また、この方式では、ペンを離してから置くまでのタイミングにより、乱流音部分の長さ

を調整することができ、破擦音や破裂音に近づけることが可能である。

この方式による4モーラの100語での音声生成・聴取実験では、乱流音なしで約2割程度だったものが、乱流音ありで約3割の了解度となった。新しく加えた乱流音付加による効果が、ある程度あったといえる。了解度が低かった単語を見ると、「ま」や「な」等の鼻音を含む単語や、有声子音を多く含んだ単語が多く見られた。^[10]

3.2 鼻子音生成モデルと数値制御による予備実験

そこで、本稿では新たに鼻子音生成のために加えた処理について述べる。まず、鼻音の生成部分として、Fig. 1の F_N に示される、共振器を加えた。

3.2.1 パラメータの取得と聴取実験

まず、鼻子音/m/, /n/を生成可能な制御パラメータを得るため、成人男性1名の音声の周波数を分析し、必要な周波数特徴量を求めた。そして、それをFig. 6に示すような、線形的に遷移するようなデータとして、音声生成器へ数値として与え音声生成を行い、正しく聞こえる音を生成できるかを調べた。実験者による音の生成と聴取の繰り返しにより各係数の値とタイミングを定め、共振器は共振周波数250Hzの共振器と、 $Fv4$ と同じ共振器を直列につないだものを用い、Fig. 6のようなデータを各モーラについて得た。その際、 $Fv1 \sim Fv4$ については、各鼻子音と同じ構音位置の「導線」に使用したのと同じ周波数変化となるようにした。

最後に、得られた制御パラメータによって単モーラ音声を生じ、聴取実験によりその明瞭さを確認した。

聴取実験は、聴取者8名に対して、鼻子音以外のモーラとの混合のランダム提示によって行った。ヘッドホンまたはイヤホンから提示された音声は何のモーラであるかを、PC画面に表示されたテキストボックスへローマ字で記入させた。

3.2.2 結果と考察

実験の結果、/ma/, /mi/, /mu/, /me/, /mo/については平均で63%、/na/, /ni/, /nu/, /ne/, /no/については、平均で53%の正答率を得られた。この結果から、/m/についてはその他の両唇構音の子音(/p/, /b/)、/n/についてはその他の歯茎構音の子音(/s/, /z/, 他)と同様の $F1$, $F2$ の遷移の特徴に、鼻子音用の共振器による音を加えるだけでも、鼻子音を簡易的に生成できることが示唆される。そこで、実際にペンタブレットの動きによって、ペンの動きだけで鼻子音を生成可能とするように、音声生成器へ実装した。

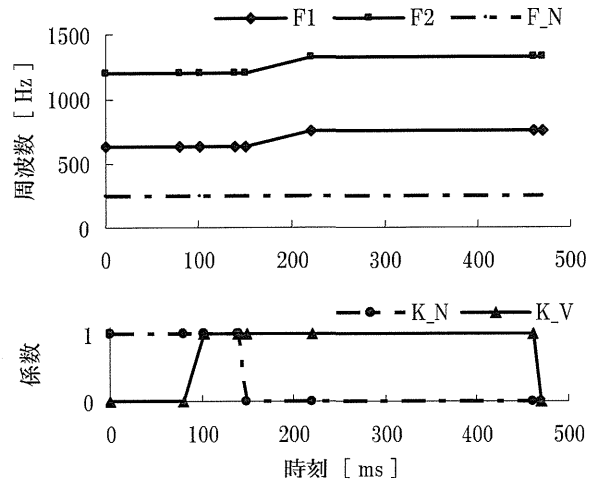


Fig. 6 Resonant frequencies and factors for "na".

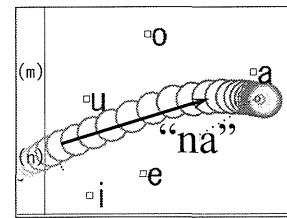


Fig. 7 Movement of the pen to generate "na".

4 音声生成器への実装

一般に、第1ホルマント周波数について着目すると、狭母音は $F1$ が低く、広母音は $F1$ が高い傾向がある。一方で、鼻子音/m/と/n/は、両唇や舌で閉鎖を行って発音する。そこで、/m/, /n/について狭母音の延長上にあるものと捕らえ、Fig. 7に示すように、操作面上の左端($F1$ が極端に低い部分)に鼻音領域を定めた。この領域へペンの座標がくると、鼻音用共振器が機能するように設定した。ここで、/m/と/n/の違いは $F2$ に対応する上下の位置となっている。

Fig. 8に、実際にペンによる「なぞり」で生成された"na"の音声波形とスペクトログラムを示す。生成された音声は、ヒトが発話した"na"の周波数変化と似たものとなっている。同様にして、/ma/, /mi/, /mu/, /me/, /mo/と、/na/, /ni/, /nu/, /ne/, /no/も生成することができた。今後、これらの生成音を用いて、ペンにの動きの操作で生成された単語音声を対象として、生成・聴取実験を行い、評価していきたい。

5 まとめ

本研究では、発話障害者支援のために、構音器官の動きをペンや指の動きで代替しているような感覚で

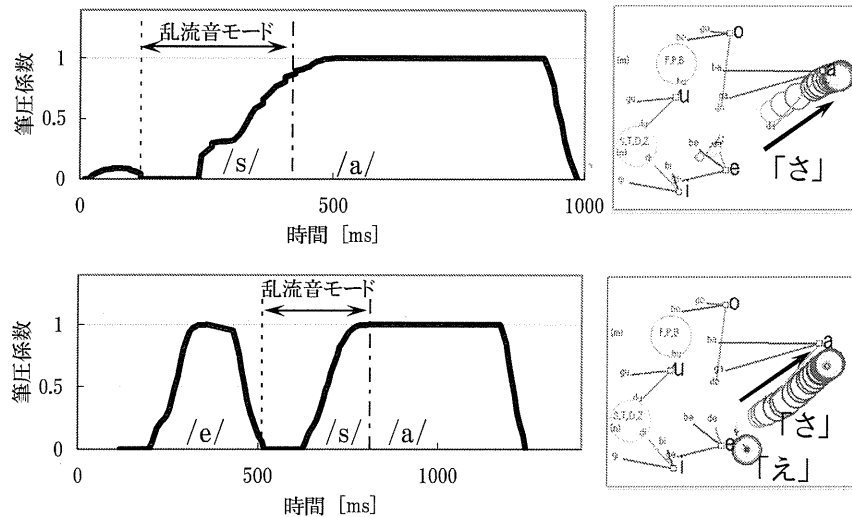


Fig. 5 Pen pressure and trajectory to generate consonant sound.

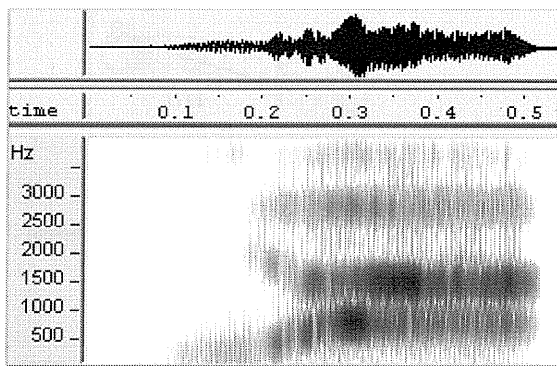


Fig. 8 A waveform and spectrogram of "na" generated by the pen movement.

扱える、音声生成器として、ホルマント合成方式を基礎とした音声生成器を提案・開発している。

本稿では、これまでの経過を紹介するとともに、母音に加えてより子音に近い音声を生成できるように、新しく加えた障害子音や鼻子音の生成手法について述べた。

音声の明瞭さや自然さだけでなく、多様な発話様式をリアルタイムに自由にコントロールできるような音声生成器として、今後も改良・評価を行っていきたい。

謝辞 本研究の一部は、平成 22 年度厚生労働科学研究費補助金 (障害者対策総合研究事業 身体・知的等障害分野 H22 - 身体・知的 - 一般 - 003) によるものである。

参考文献

- [1] 藪, 伊福部, 青村, "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, 電子情報通信学会, 105(686), 59-64, 2006.
- [2] 小林, 音響学会誌, 67(1), 15-16, 2011.
- [3] ジャック・ライアルズ, "音声知覚の基礎". 海文堂出版, 東京, 2003
- [4] Fant, "Acoustic Theory of Speech Production", Walter de Gruyter, 1970.
- [5] Klatt, Software for a cascade/parallel formant Synthesizer, Journal of the Acoustical Society of America, 67(3), 971-995, 1980.
- [6] ピー・アクティブ P-Active : XP-3300A ポケットサイズペンタブレット <http://www.p-active.com/product/pt/3300a.htm>
- [7] 藪, 伊福部, 青村, "ポインティングデバイスを利用した音声生成方式-発話障害者のための支援機器として-", 日本保健科学学会誌, 12(1), 49-57, 2009
- [8] 坂本, 鈴木, 天野, 近藤, "親密度と単語の音韻バランスを統制した単語理解度試験用リストの構築", 東北大学電通談話会記録第, 69(2), 21-34, 2000
- [9] 藪, 青村, 伊福部, "ポインティング・デバイスで操作する発話支援インタフェース", ヒューマンインタフェース学会誌, .11(4), 135-146, 2009
- [10] 藪, 伊福部, "発話障害者支援のための連続タッチ平面で操作する音声生成器-子音改善のための基礎的検討", 信学技報告, 音声, 110(220), 41-46, 2010.

構音機能障害者のための 音声生成器の抑揚制御方式に関する基礎的検討

藪 謙一郎[†] 伊福部 達[‡]

[†] 東京大学先端科学技術研究センター 〒153-8904 東京都目黒区駒場 4-6-1

[‡] 東京大学高齢社会総合研究機構 〒113-8656 東京都文京区本郷 7-3-1

E-mail: {yabu,ifukube}@human.iog.u-tokyo.ac.jp

あらまし 発話障害者を支援する音声生成支援機器として、我々は、使用者が音声を楽器のようにコントロールできるリアルタイムの音声生成器の提案・開発をしている。この方式は、ペンや指の指し位置の座標を音声生成パラメータへ直接対応させているもので、自由な間(ま)やリズムなどの非言語的な情報が表現可能となる。先の研究により、徐々に音韻の明瞭さが高められてきたが、片手で音韻と抑揚を同時に制御する方法については未だ導入されていなかった。本稿では新たに、(1)小指の握力を押圧センサから取得し抑揚制御パラメータとする方法と、(2)使用者の声を頸部に取り付けた皮膚伝導マイクロホンによって取得しホルマント合成の原音に用いる方法とを導入し、試作器による簡易的な実験を行ったので、報告する。

キーワード ホルマント、音声合成、発話障害者、抑揚

Basic study on a pitch-control method of a speech synthesis device for articulation disorders

Ken-ichiro YABU[†] Tohru IFUKUBE[‡]

[†] Research Center for Advanced Science and Technology, The University of Tokyo

4-6-1 Komaba, Meguro-ku, Tokyo 153-8904 Japan

[‡] Institute of Gerontology, The University of Tokyo 7-3-1 Hongo Bunkyo-ku Tokyo Japan

E-mail: {yabu,ifukube}@human.iog.u-tokyo.ac.jp

Abstract As an assistive device for speech disability persons, we have proposed and developed a real-time speech synthesis device by which a user can control sound like playing a musical instrument. In this method, coordinates of a position pointed by a finger or pen is directly corresponding the speech synthesis parameters, the device enables expressions of nonverbal information such as time intervals or rhythms. Intelligibility of voice has been gradually progressed in previous study, but the method to control both phonological and intonation change by using only one hand was not introduced. Therefore, two method was introduced: one is that (1) intonation control parameter is input from a pressing force by grip strength of little finger, and another is that (2) source sounds of formant synthesis is provided from a voice of a user by skin conduction microphone fitted on neck. In this paper, we report about a basic experiment of them by prototype device.

Keyword Formant, Voice Synthesis, Articulation Disorder, Intonation

1. はじめに

発話の障害を補うための音声支援機器として、文字や単語の入力から音声を生成する、TTSを基本とした機器がよく用いられる。これらは、使用者が自分の要求や意思を確実に伝える上で、有力な手段となっている。

一方で、ヒトが発話する音声には、文字通りの言語の情報の以外にも、間(ま)や話すテンポや抑揚、リズムなどの表現が含まれている。それらによって、感情やユーモアの表現など、文字では表せない非言語的な付加情報を持たせることができ、日常生活の中で他者との関係を円滑に送るために重要な役割を持っている。

また、複数人との会話の中では、声ならではの応答スピードも重要な役割を持っている。例えば、4~5人の会話の中で5~6秒も間があれば、誰かが次の話題を持ちかけるのに十分な時間となり、その前までに話し始めなければ前の話題に関する発言が難しくなる場合も多い。

筆者らは、音声支援機器でこのような多彩な表現を可能とするためには、ヒトが構音器官を巧みに操って自由に音声を出すのと同じように、身体のどこかの部位の動きと出力音声とが連動するような方式が必要であると考えている。

そこで、文字や単語に頼らずに、指やペンで面をな

ぞる動きで、楽器のように使用者自身がリアルタイムに音声をコントロールしながら音声を生成できる音声生成器を提案している^[1]。

この音声生成器では、ヒトの母音知覚に必要な第1ホルムントと第2ホルムントの周波数(以下、 $F1, F2$)^[2]を、音声生成器の操作面の縦軸と横軸とに割り当て、使用者が指やペンを操作面に置いた瞬間に、その位置に対応したホルムント周波数を持った音を発する。

この方法では、操作面に指やペンを置いた瞬間に声が出せ、自由なリズムをつけることも可能となる。また、母音と母音の間の中間的な音を連続的に変化させて出すことも可能であり、多様な音韻が表現できることが期待される。

先までの研究では、母音音声の出力をベースとした本方式を用いて、疑似子音音声を生成する手法を提案し、改良を進めてきた^[3-5]。

一方、声の高さに関しては、MIDIデータのシーケンス制御による歌声メロディの生成や、タッチパッドによる押圧を抑揚に変える方法、タッチパッドの操作と逆の手で抑揚制御をする方法などが試されてきたが、話し声に近い抑揚を、片手だけで音韻と同時にコントロールすることは困難であった。

本稿では、タッチパネルを人差し指で操作しながら、(1)人差し指と独立に動かし易い小指の握力抑揚をコントロールする方式と、(2)タッチパネルの操作と同時に使用者が声を出し、その声の抑揚をそのまま用いる方式との、2種類の抑揚制御方式を新たに提案・導入し、簡易的な音声生成実験とその聴取実験を行ったので、報告する。なお、後者の方式は構音機能のみに主な障害をもち、声帯音の発声には障害を持たない場合を想定したものである。そのような話者にとっては、もっとも自然な抑揚付加方法であるといえる。

2. 音声生成器の原理と機能

2.1. 基本原理(母音)

音声生成器のブロック図を図1に示す。本研究では、ホルムント合成方式を採用している。音声生成の処理は、声帯振動による音を再現する母音処理部と、摩擦音・破裂音・破擦音などに含まれる雑音成分(以下、乱流音と呼ぶ)を処理する乱流音処理部から成る。

図1の上部の母音生成部では、原音となる音声信号が、各ホルムントに対応する $Fv1$ から $Fv4$ のデジタル共振フィルタを通して特徴付けられ出力される。

これらのフィルタのうち、 $Fv1$ と $Fv2$ の共振周波数が、図2に示す使用者の操作面からの入力によってリアルタイムに制御される。その他のフィルタの特性は固定とした。本稿の研究では、デバイスとして図3に示すタブレットPC(ONKYO-TW217A)を用いた。その仕様を表1に示す。画面には“a, e, i, o, u”の文字が、

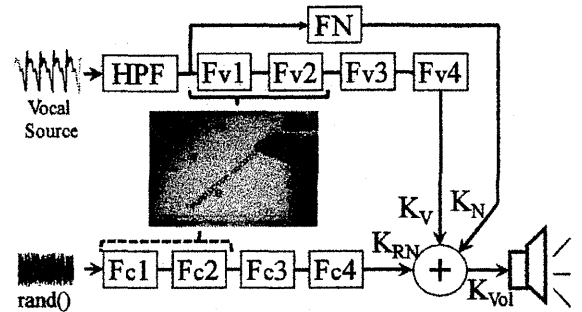


図1 音声生成器のブロック図

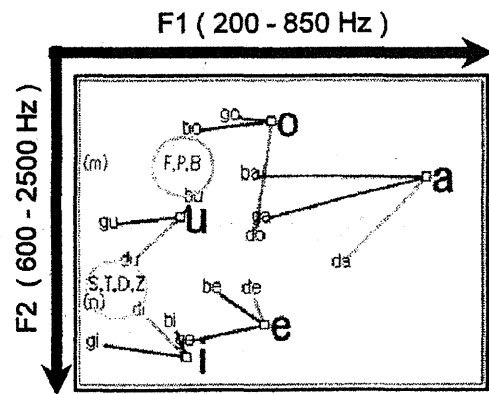


図2 音声生成器の操作盤面

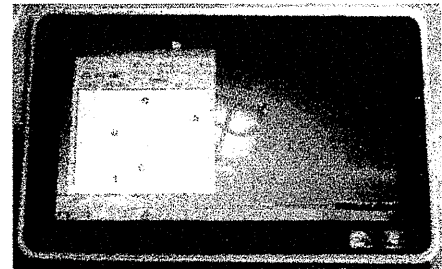


図3 タブレットPC上の動作画面

表1 タブレットPCの仕様^[8]

本体仕様	TW217A5
OS	Windows® 7 Home Premium
内蔵ディスプレイ	タッチパネル付 10.1型ワイド 最大 1,024×600ドット
キーボード	なし
ポインティングデバイス	静電容量式タッチパネル (マルチタッチ対応)
寸法	274(×)173×18.5 mm
質量	約 850g
バッテリー	動作時間 約 6.4時間

各々のホルマント周波数と対応した位置に表示され、使用者がその位置をタッチすると同時にその音が出来、タッチしながら位置を動かすと、連続して変化する音が出来される。指を離すと音声は停止する。

ここで、図1中のHPFと示されているものは、高い周波数成分を強調するために、設けた緩やかなハイパスフィルタである。また、FNは、後述の鼻音生成のためのフィルタで、係数KNで振幅調整されてから合成される。

母音生成部の原音波形には、声の振幅やピッチのゆらぎを再現するため、筆者の「あー」という約3秒間の声をサンプリング周波数16kHzで録音し、30次のLPC逆フィルタ処理で補正した擬似声帯音声を用いている。

図1下部の乱流音処理部は、2.3節に述べる擬似子音の付加に用いており、音源が乱数発生関数によって生成された雑音である以外は、母音処理部とほぼ同じ構成である。

なお、共振・反共振器フィルタについては、Klattらと同様のフィルタ^[7]を使用した。HPFには、反共振フィルタの反共振周波数をゼロとしたものを用いた。

2.2. 子音への拡張と導線の導入

先の研究により、「おはよう」、「こんにちは」等の簡単な単語であれば、2.1節で述べた操作面中の、母音だけをペンでたどるだけでも、十分に認識可能な音声を生成することができる。さらに、動かし方を工夫すると、より子音に近づいた音声を生成することができる。これは、ヒトの発話音声の子音開始部分に存在しているホルマント周波数の急激な変化(ホルマント遷移)が再現されるためであると考えられ、本研究では、ヒトが発話した音声を分析し、操作面上において各々の子音生成に対応する軌跡を表示させた、図2に線で示される「導線」を導入している^[4]。なお、全ての子音について各々に対応する奇跡を表示させると複雑化しすぎる。一方で、音声学的な分類で、発話の際の舌の位置が同じ子音では、ホルマント遷移も似ている。そこで、同じ構音位置に属する子音では同じ導線を用いるものとして、導線の数を削減している。

2.3. 乱流音と鼻音フィルタによる疑似子音

導線を用いた音声生成では、母音を基礎とする子音に近い音を出すことができるが、明瞭さを増すためには子音のための雑音成分や、鼻音成分の付加が不可欠となる。そこで、図1下部に示されるランダム雑音を用いた乱流音源と、鼻音用フィルタFNを設け、疑似子音を付加できるようにした^[5]。

この時の操作は、無声子音の付加の部分で、一瞬操作盤面から手を離し、300ms以内に操作盤面に手を戻

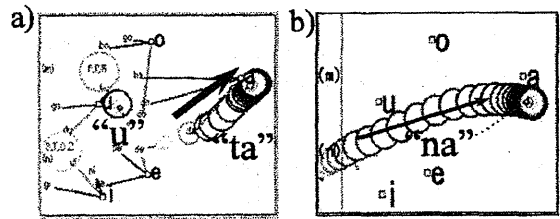


図4 子音生成時の入力軌跡

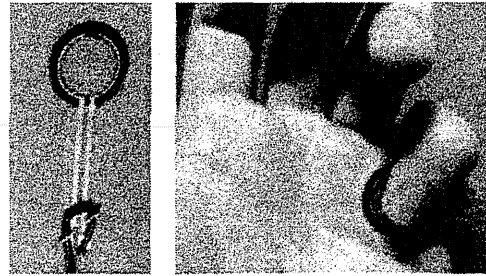


図5 押圧センサFSRと取り付け位置

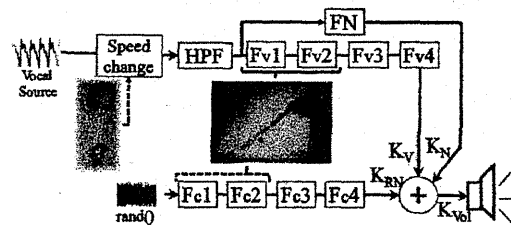


図6 押圧センサによる抑揚制御

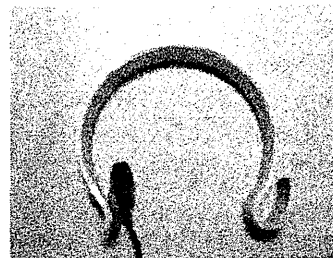


図7 皮膚伝導マイクと試作固定具

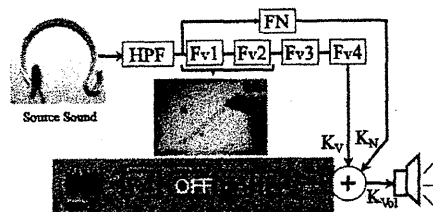


図8 声を用いた抑揚制御

すと乱流音が自動的に付加される。また、鼻音フィルタについては、図 4(b)に示されるように、操作盤面左端に鼻音フィルタ動作領域を設け、その位置から母音の方向へポイント位置を動かすと、鼻音が生成される。

3. 抑揚制御方式の導入

以上が本研究の音声生成器の音韻の制御方式である。これらの方式により、音声の明瞭さが徐々に向上されてきた。一方で、多様な表現を実現するためには、声のリズムや声の高さ（抑揚）をリアルタイムにコントロールできることが必要である。

リズムについては、操作するタイミングを自由にコントロールすることによって、本方式による制御が可能となっている。声の高さについては、MIDI データを用いた音程制御による方法や、押圧センサを用いた抑揚制御方法が試みられてきた。しかし、片手だけを用いた音韻と抑揚の同時操作は未だ十分に試みられていなかった。

そのひとつの理由として、従来の、音韻を制御するのと同じ指の指先の押圧を用いる方法では、音韻制御と独立に抑揚を思い通りに制御することが困難だったことがあげられる。

そこで、本稿では新たに、タッチパネルを操作する指とは独立に動きやすいと考えられる小指の握力を用いた方法と、使用者の実際の声を用いた方法との 2 種類の方法を試み、簡単な実験を行った。

3.1. 小指握力による抑揚制御

本研究の音声生成方式では、操作面をタッチする場所を素早い動きで移動したり、100~200 ms の単位で一瞬操作面から指を離したりして、音韻を制御する。これらの操作により、操作面を押す力や、タッチする指にかかる力が自然に変化してしまうが、抑揚制御は、この音韻の制御と無関係に制御可能でなくてはならない。そこで、本研究では操作面をタッチする人差し指からもっとも離れた、小指の握力によって抑揚を制御する方式を試みた。

図 5 に抑揚の制御に用いた押圧センサと取り付け位置を示す。押圧センサとして、INTERLINK ELECTRONICS 社製、FSR センサ(FSR402)を用い、小指の握力によってこの押圧センサを押すことによって、抑揚が制御される。このセンサは、導電ゴムフィルム的一种で、加えられた押圧に応じて抵抗値が減少する。本研究では、固定抵抗と FSR を組み合わせ、抵抗値の変化を電圧として、USB 方式の入力装置からリアルタイムに取得できるようにした。

使用者はまず最大の握力かけ、握力無付加時と最大握力時との間で、1 オクターブの声を制御可能にした。

3.2. 使用者の声による抑揚制御

本研究の音声生成方式では、ホルマント合成方式を

採用しているため、声帯音を模擬した原音と、音韻を与える共鳴の制御を分離できる。実際のヒトの発話においても、抑揚や声の音量は声帯音によって出され、音韻については構音器官で特徴づけられるため、構音器官のみに障害を持つ場合もある。そのような患者では、自分の正常な声帯から出される抑揚や強弱によって発話できることが有用である。

そこで、本研究では新たな試みとして、図 8 に示すように、皮膚伝導マイクロホンによって頸部から取得した音声波形をホルマント合成の原音として用い、その音に対してタッチパッドからの入力で制御されるホルマントの特徴づけを行う手法を導入した。

まず、マイクロホンの固定具として、橋場らのハンズフリー人工喉頭の研究^[9]で用いられた固定具と同様にして、図 7 に示すような固定具を試作した。固定具のサイズは、幅約 20mm、直径 95mm である。マイクロホンには、皮膚伝導マイクロホン 817(糊プリモ)を用いた。マイクロホンの信号は USB 音声入出力デバイス UAB-350 (SONY(株))を介して音声生成器が動作するタブレット PC へ入力し、そのままホルマント合成の原音として用いた。図 9 に皮膚伝導マイクロホンから得られた音声の波形とスペクトログラムを示す。

4. 実験

前述した音声生成器の抑揚付加によって、抑揚のついた音声を十分に制御可能であるのかどうかと、抑揚付加によって音声の明瞭度が変化するのかどうかを調べるため、音声の生成実験と聴取実験とを行った。

4.1. 音声の生成実験

生成は、本インタフェースの開発に関わり使用方法を熟知している 1 名(以下、生成者と呼ぶ)が行った。生成者は、左耳 50dBHL、右耳 30dBHL 程度の聴力レベルであり、器質性および運動性の構音障害を持つ。手や指の運動機能は正常である。

対象とする単語は、先の研究^[4]で用いた 100 語の単語のうち、特に正答率が低かった単語 26 語を使用した。

生成者は、表記された対象の単語を見ながら、(1)抑揚無しの場合、(2)小指握力による抑揚制御の場合、(3)自身の声を用いる場合との 3 通りで、音声生成を行った。それぞれの単語の生成回数は自由とし、生成者自身が納得できる音声を生成できた時に、次の単語へ移るようにした。なお、2.3 節で述べた疑似子音の付加機能については、(1)、(2)では有効、(3)の時には無効とした。生成実験(3)では、生成者は口を半開きの状態にして動かさずに、声を出すと同時にタッチパネルを操作した。また、子音部において、相応する/h/の摩擦音を発声した。なお、生成作業中の音声は、イヤホンで生成者へ提示すると同時に、PC 上で動作する音声波形エディタ Wavesurfer によって録音された。

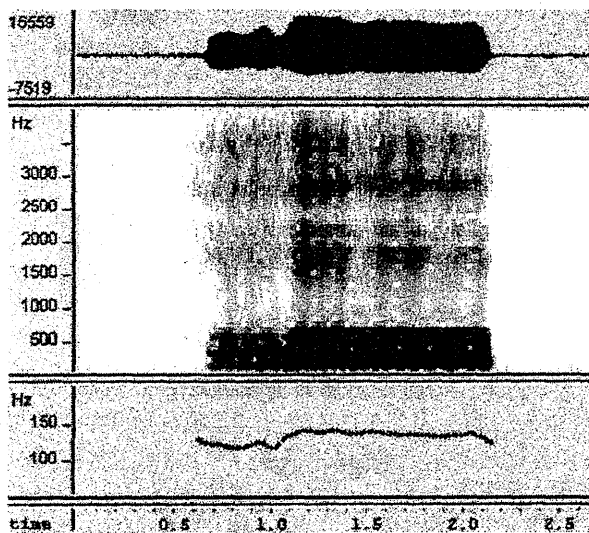


図 9 皮膚伝導マイクからの取得音声
(上から、波形、スペクトログラム、ピッチ、時間)

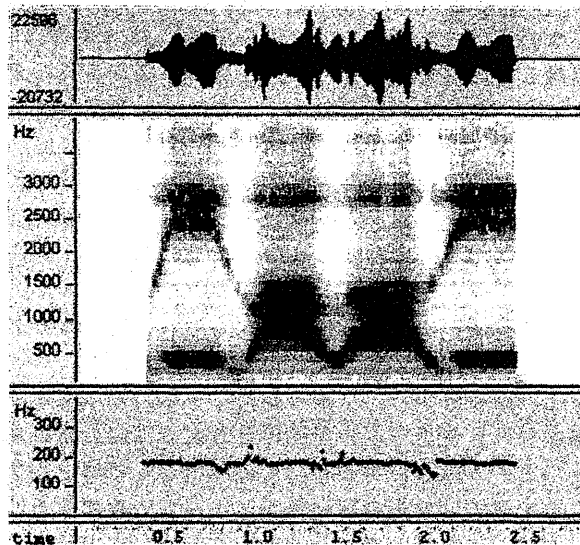


図 10 抑揚制御無しの「みまわり」の出力音声
(上から、波形、スペクトログラム、ピッチ、時間)

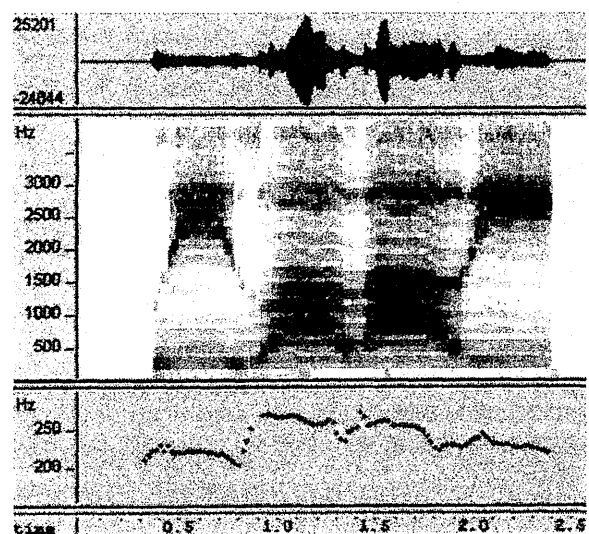


図 11 小指の握力による抑揚制御の音声
(上から、波形、スペクトログラム、ピッチ、時間)

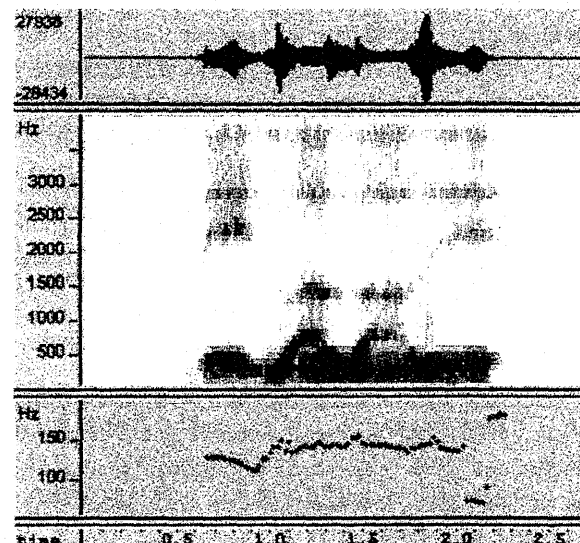


図 12 皮膚伝導マイクを使った音韻付加音声
(上から、波形、スペクトログラム、ピッチ、時間)

4.2. 聴取実験

聴取実験は、録音された生成音 1 語ずつを聴取者へランダムに提示し、聴取者自身が PC のキー入力によって、聞こえた単語が何であるかを記入する方法で行った。聴取者は全員、聴力に関する大きな病気をしたことがなく正常な聴力を持つ 20 代から 30 代の男性 4 名とした。

音声の提示は、汎用 PC に接続されたアームレスヘッドホン Victor HP-AL202 を通して行われた。聴取者には、実験の前に提示音声日本語の 4 音節の単語であることを教示し、単語が推測可能な場合には、聞こえたままの文字を記入するのではなく、その単語を答えるように指示した。

4.3. 実験結果と考察

音声生成実験では、生成者により 3 つのパターンで所望の音声生成された。いずれの単語でも、単語内での抑揚の位置は、通常の発話される単語と同様の出力がされた。小指圧力による制御では抑揚は強め、使用者の声による方式ではやや弱めの抑揚であった。生成された音声の例を図 10-12 に示す。音声生成の実験から、2 つの方式によって、単語における抑揚コントロールが可能であることが示された。ただし、特に小指圧力による方法の難易度には、個人差があると考えられるため、今後、複数の生成者による検証の必要がある。

図 13 に聴取実験の結果を示す。抑揚の有無による正答数の違いをみると、3人の被験者のうち、1名では、抑揚無しに比べて抑揚を付加した場合のほうが高い正答率となった。しかし、他の2人では大きな差が見られなかった。この結果から、音韻の明瞭さについては、全体としては抑揚の有無の影響は少なかった。ただし、本研究で用いた単語リストは、アクセントの位置が 0 型もしくは 4 型 (LHHH) に限られていたため、抑揚の有無の影響を受けにくかった可能性もある。また、文脈がある場合では、抑揚も聞き取りの手掛かりになる可能性がある。

また、聴取者から、小指の握力による制御の音声では明るい声、使用者の声をういた方式では低く暗い声であるという感想があった。この理由としては、小指の握力制御では抑揚の制御幅が広がったために明るく感じられ、使用者の声ではゆっくりと低い声で音声生成され暗く感じられたことが挙げられる。

5. まとめ

本研究では、操作盤面上をペンや指でなぞることによって、楽器のようにリアルタイムに音声を生成する方式を提案し、開発を進めている。先の研究では、おもに音韻を制御する方式に焦点をあて、導線の導入や疑似子音の付加の方式を扱ってきた。一方で、声の高さに関することについては、MIDI シーケンサを用いた歌声の生成や、タッチパッドの押圧を用いた方法、操作盤面を操作するのと逆の手で押圧センサを扱う方法等が検討されてきたが、片手で抑揚と音韻との両方を操作する方法は検討が不十分であった。

本稿では新たに、小指握力による押圧センサの操作によって抑揚を制御する方法と、使用者自身の声を皮膚伝導マイクロホンを通じて取得しホルマント合成の原音とする方法との2種類を導入し、簡易的な音声の生成実験と聴取実験とを行った。

音声生成実験の結果から、通常抑揚と同じ抑揚を各単語に付けられたことから、この2つの方式において音韻と独立した抑揚の制御が可能であることが示された。また、聴取実験から、抑揚のある音声のほうがわずかに、単語を了解しやすくなる可能性が示された。

ただし、音声生成の際の慣れや使いやすさには個人差があると思われるため、今後の検証が必要である。また、多様なアクセント型による音声の伝わりやすさや、文脈を持った場合の影響についても調べる必要があると考えており、引き続き検討を進めていきたい。

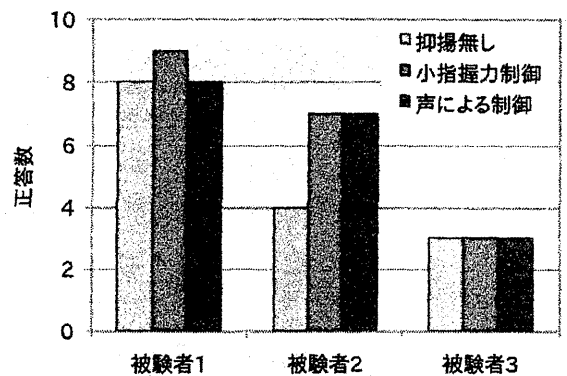


図 13 聴取実験結果

謝 辞

本研究の一部は、平成 23 年度厚生労働科学研究費補助金(障害者対策総合研究事業身体・知的等障害分野 H2 2 - 身体・知的 - 一般 - 003)によるものである。

文 献

- [1] 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(聴覚・音声・言語とその障害, 一般)", 信学技報, SP, 音声, 105, pp:59-64, 2006.
- [2] ジャック・ライアルズ著, 今富瑛子, 荒井隆行, 菅原勉監訳/新谷敬人, 北川裕子, 石原健訳: 音声知覚の基礎. 海文堂出版, 東京, 2003
- [3] 藪謙一郎, 伊福部達, 青村茂, ポインティングデバイスを利用した音声生成方式-発話障害者のための支援機器として-, 日本保健科学学会誌, 12(1), pp.49-57, 2009
- [4] 藪謙一郎, 青村茂, 伊福部達, ポインティングデバイスで操作する発話支援インタフェース, ヒューマンインタフェース学会誌, Vol.11 No.4, (135-146) 2009
- [5] 藪謙一郎, 伊福部達, 「なぞり」の入力による発話障害者支援のためのリアルタイム音声生成器 - 阻害音と鼻音付加の操作方法の検討 -, 日本音響学会, pp.1623-1626, 2011
- [6] Fant, G., Acoustic Theory of Speech Production, Walter de Gruyter, (1970) pp.111-112
- [7] Klatt, D.H.: Software for a cascade/parallel formant synthesizer: Journal of the Acoustical Society of America, pages 971-995, volume 67, number 3, March 1980.
- [8] ONKYO 製品情報: TW217 <http://www.jp.onkyo.com/pc/personalmobile/tw217/spec.htm>
- [9] 橋場 参生, 須貝 保徳, 泉 隆, 井野 秀一, 伊福部 達, "喉頭摘出者の発声を支援するウェアラブル人工喉頭の開発と評価", ヒューマンインタフェース学会論文誌 9(2), 163-172, 2007

電気人工喉頭のための拡声器に関する一考察

藪 謙一郎[†] 上田 一貴[†] 稲永潔文[‡] 伊福部 達^{††}

[†] 東京大学先端科学技術研究センター 〒153-8904 東京都目黒区駒場 4-6-1

[‡] (株)サザン音響 〒248-0027 神奈川県鎌倉市笛田 5-34-18

^{††} 東京大学高齢社会総合研究機構 〒113-8656 東京都文京区本郷 7-3-1

E-mail: [†] {yabu, ueda}@human.iog.u-tokyo.ac.jp, [‡] inanaga@human.iog.u-tokyo.ac.jp

^{††} ifukube@human.iog.u-tokyo.ac.jp

あらまし 喉頭摘出患者は声帯が使えなくなるため、声が出せなくなる。その代替手段の一つに電気人工喉頭がある。電気人工喉頭はスイッチを押すだけで、メンテナンスや練習を要さず、すぐに使えるという利点があるが、機器からの直接音が雑音となる問題や大きな声を出せないと言ったことが、課題として挙げられる。その問題の低減手段として、拡声器を用いる手段が考えられる。本研究では、電気人工喉頭使用者のための拡声器の開発を行い、これまで調べられていなかった人工喉頭による発声の明瞭さについて定量的に調べると同時に、マイクによる明瞭さの違いや、増幅音量による明瞭さの違いを調べるための予備実験を行なった。

キーワード 人工喉頭, 声, 拡声器, 了解度

A basic study of amplification equipment for an electro-larynx

Ken-ichiro YABU[†] Kazutaka UEDA[†] Kiyofumi INANAGA[‡] Tohru IFUKUBE^{††}

[†] Research Center for Advanced Science and Technology, The University of Tokyo

4-6-1 Komaba, Meguro-ku, Tokyo 153-8904 Japan

[‡] Southern Acoustics Co.,Ltd. 5-34-18, Fueda, Kamakura-shi, Kanagawa, 248-0027, Japan

^{††} Institute of Gerontology, The University of Tokyo 7-3-1 Hongo Bunkyo-ku Tokyo 113-8656 Japan

E-mail: [†] {yabu, ueda}@human.iog.u-tokyo.ac.jp, [‡] inanaga@human.iog.u-tokyo.ac.jp

^{††} ifukube@human.iog.u-tokyo.ac.jp

Abstract Laryngectomized patient cannot vocalize because of unavailability of their vocal cord. An electro-larynx device is one of an alternative device for them. It has the advantage that it can use immediately by simply pressing a button switch without daily maintenance and practices. But, there are problem that a sound which directly comes from a device is noisy, and that a user cannot speak with a loud voice. It is considered that these problems will be reduced by the use of a loudspeaker system. The authors developed a loudspeaker system for an electro-larynx device user. In this paper, we describe about a preliminary experimentation to investigate a intelligibleness by use of different microphones or gain of a loudspeaker system, and intelligibleness of a sound which is spoken with an electro-larynx in quantitative form.

Keyword Electro-larynx, Voice, Loudspeaker system, Intelligibleness

1. はじめに

ヒトが日常生活を送る上で、「声」は最も頻繁に使われる基本的な意思疎通手段である。そのため、声を出す機能は、単に自分の要求や意思を伝えるだけでなく、人間関係や社会参加を円滑にするために、重要な役割を担っている。

喉頭がん等により喉頭摘出手術を受けた患者は、発声に必要な声帯の機能を失ってしまうことから、声を出すことができなくなる。ただし、喉頭摘出だけの場合には、声帯による声の源音生成機能のみが失われ、源音に言葉の特徴を付ける構音の機能については、残存する。したがって、何らかの方法で声帯の代替とな

る原音を生成できれば、残存する構音器官の機能を使って言葉を発声することができる。このような、声帯音を代替する手段で実用されているものとしては、人工喉頭や食道発声法などがある。

電気人工喉頭は、図1に示すように、頸部に機器を当てた状態でボタンスイッチを押すと、声帯の代替となる振動音が声道内に伝わって発声できる装置である。特別なメンテナンスを必要とせず、すぐに使用できるため、初心者でも使いやすい利点がある。一方で、頸部に当てた装置から直接外へ漏れだす音が雑音となったり、大きな声が出せないために聞き取りづらかったりすることが課題の一つとなっている。

声の大きさや雑音の問題を軽減する方法として、拡声器を用いて声の音量を増幅する方法が考えられる。この場合、使用者が拡声器をポケットへ装着したり、首から下げたりして携帯して使用することになる。しかし、マイクと出力スピーカとの距離が近いため、市販される汎用的なマイクによる拡声器では、音量を上げるとハウリングを生じてしまって十分な音量を得ることが困難であった。

そこで著者らは、使用者の至近距離にスピーカがある場合でも、ハウリングを抑えたまま十分な拡声効果を得られる拡声器の開発を試み、拡声器の試作を行った。試作拡声器は、図2に示すように双指向の特性を持つマイクを用いて、雑音源が位置する方向への感度がほぼゼロとなるように配置して、ハウリングマージンを上げ、同時に人工喉頭の直接音を抑える工夫を施したものである。

本研究では、拡声器による声の聞き取り易さへの効果を調べる予備実験として、発話・聴取実験を行ったので報告する。

2. 目的

本研究は、人工喉頭を使用して単語を読み上げた場合の、①拡声器を用いない場合、②無指向マイク(汎用マイクを想定)の拡声器Ⅰを用いた場合、③双指向マイク(改良マイクを想定)の拡声器Ⅱを用いた場合との3種類について、比較し、拡声器の効果を調べることを目的とする。また同時に、これまで定量的に調べられていなかった、電気人工喉頭で発話される声の明瞭さについて、文脈の無い単語音声の了解度で、調べることが目的とする。以上の目的で、本稿では予備実験として、話者1名、聴取者1名で、発話・聴取実験を行った結果を報告する。

3. 実験方法

前節で述べたように、本研究では、人工喉頭で単語を読み上げた声について、拡声器を用いない場合(条件Ⅰ)、無指向性マイクの拡声器Ⅰを用いる場合(条件Ⅱ)、改良型である双指向性マイクの拡声器Ⅱを用いる場合(条件Ⅲ)を、発話・聴取実験で比較した。その際、話者が話した声の質を3条件で一致させる必要がある。そのため、あらかじめ条件1～3のマイクからの音声を同時に収録した上で、聴取実験で各条件をシミュレートした音量で再生・提示する方法をとった。

3.1. 人工喉頭音声の録音

3.1.1. 実験装置

音声の録音装置を図4に示す。実験用のあご台に3種類のマイクを固定した。話者はマイクに向かって、人工喉頭で単語を読み上げ、その声を録音した。その

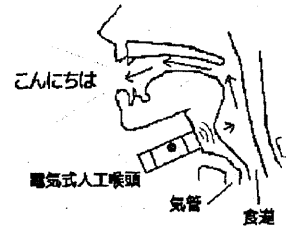


図1 電気人工喉頭による発話([3]より)

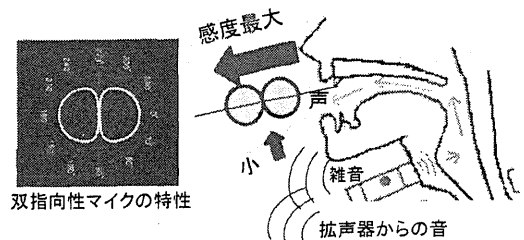


図2 マイクの指向性と雑音源の関係

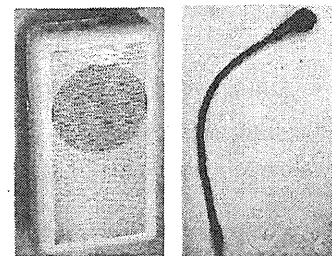


図3 電気人工喉頭のための拡声器 (協力：㈱サザン音響)

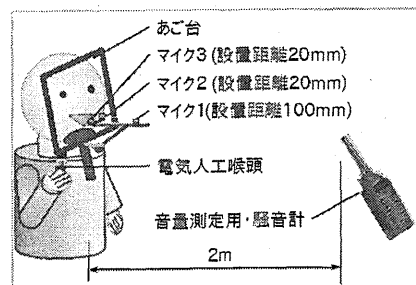


図4 録音用マイクの設置位置

際、話者の口からマイクの距離を一定にするために、額をあご台上部へ当てて発話させた。

まず、マイク1として口元から約100mmの位置に無指向性マイクを設置した。このマイクは、拡声器を使わない場合の音を得るためのもので、人工喉頭からの直接音を含んだ音声を得られる。次に、マイク2とし

て、マイク 1 と同じ無指向性マイクを口元から約 20mm の位置に設置した。これは、拡声器 I のマイク音声を意図している。最後に、マイク 3 として、双指向性マイクを口元から約 20mm の距離に設置した。これは、改良型の拡声器 II のマイクを想定したものである。

これらのマイクは、プリアンプ(AT-MA2: audio-technica)と USB オーディオインタフェース(UA-101: Roland)を介して、汎用パソコンにより 44.1 kHz 16bit で多チャンネル録音した

なお、実験に用いた電気人工喉頭は、憐電製の「ユアトーン II ゆらぎ」を、ノーマルモード、音量“5”、ピッチ“5”の設定で使用した。

3.1.2. 録音手順

話者は人工喉頭の使い方を熟知する 30 歳代の健康男性 1 名とし、録音は防音室内で行なわれた。初めに、話者が人工喉頭を使って「あー」という声を出し、それを録音すると同時にその音量を 2m 離れた位置で測定した。その結果、約 60 dBA であった。

読み上げる単語として、坂本らによる難聴者のための単語理解度試験用単語リスト^[4]の中から、親密度 7.0 ~ 5.5 の単語 50 語を用いた。これらは、全てアクセント型が 0 型又は 4 型(LHHH)の、4 モーラ音声からなる単語音声である。

話者は、録音の前に 1 時間程度、単語リストを見ながら聞き取り易い発音で話せるように練習をしたのち、前節で述べたようにマイクに向かって、単語リストに記載されている単語を順に読み上げ、その音声を録音した。読み上げは 10 語ごとに、数秒から十数秒の休憩をはさみながら行なった。

3.2. 聴取実験

3.2.1. スピーカ設置位置と音量の設定

聴取実験は、録音された音声を聴取者にランダムに提示して行なった。提示するスピーカの種類と位置を、図 5 に示す。スピーカ A(BOSE Companion 2 series II)は、マイク 1 の録音音声を再生するためのもので、拡声器無しの場合の音声を再現するものである。また、スピーカ 2 は、マイク 2 またはマイク 3 の録音音声を再生するためのもので、拡声器から出される音を再現する。

すなわち、聴取実験の条件として、マイク 1 の音をスピーカ A から再生する条件 1、それに追加してマイク 2 の音声をスピーカ B から再生する条件 2、マイク 3 の音声をスピーカ B から再生する条件 3 の、3 条件とした。(表 1) それぞれの音は全て同期して同時に再生される。

スピーカ A の音量設定は、録音時の音量とほぼ同値となるように、前節で録音された「あー」の音声再生時に 2m 離れた位置で 58~60dBA とする音量を設定し

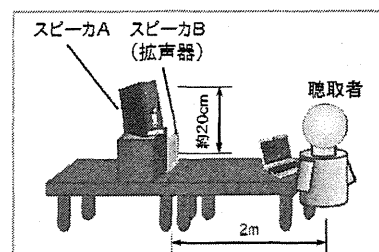


図 5 聴取実験

表 1 想定する拡声器と実験条件

拡声器使用の想定		条件 1	条件 2	条件 3
		無	有(I)	有(II)
録音・提示	マイク 1 スピーカ A	○	○	○
	マイク 2 スピーカ B	×	○	×
	マイク 3 スピーカ B	×	×	○
混合音量[dBA]		58-60	63-65	68-70

た。

また、スピーカ B の音量は、条件 2 と条件 3 とで異なる値を設定した。前節の「あー」を再生時にスピーカ A とスピーカ B との混合音声が、距離 2m の位置でそれぞれ、条件 2 で 63~65dBA, 条件 3 で 68~70dBA とするよう設定した。このスピーカ B の音量設定は、各マイクの拡声器でリアルタイムの拡声動作時に、スピーカがマイクの下 20cm の位置でマイクの方向に向けて置かれた条件で、ハウリングを起こさない範囲の最大のゲインを再現した音量である。

3.2.2. 聴取実験

聴取者は 30 歳代の健康男性 1 名とした。実験は静かな部屋で行なわれた。聴取者は、スピーカから 2m の位置に耳が来るように椅子に座り、一語ずつ提示音声を聴取した。

提示音声は、録音された 50 語について、条件 1、条件 2、条件 3 のもとで再生される 150 音声を、ランダムに提示するものとした。すなわち、聴取者はランダム化された 150 刺激の中で、同じ語を 3 回、異なる条件で聴くことになる。ただし、ランダム化の際には、同じ語が連続しないようにした。

聴取者には、提示される音が全て 4 モーラから成る意味のある単語であることを伝え、聞こえた単語が何であるかを聴取者の前に置かれたノートパソコンへキーボードからひらがなでテキスト入力をさせた。なるべく意味のある単語を記入することとし、どうしてもわからない場合は、わかる範囲で聞こえた音を記入するように教示した。音声は 5 秒おきに最大 3 回再生さ

れるように設定し、少ない回数で入力できた場合は、聴取者が入力をした時点で次の後へ移るようにした。初めの4語は「あいうえお」「かきくけこ」「さしすせそ」「たちつと」の音声を提示し、聴取者は聴取や文字入力の練習を行った。

4. 結果および考察

図6に結果を示す。50語中、完全に正答された単語数は、条件1では16語(32%)、条件2では20語(40%)、条件3では18語(36%)であった。150語全ての回答に要した時間は、約24分30秒で、1語あたり9.8秒であった。

また、単語に含まれる4モーラのうち、部分的に正答していた場合の単語数を、正答したモーラごとにカウントすると、図7のようになった。

拡声器無しの場合である条件1と、拡声器を用いた場合である条件2、条件3を比較すると、拡声器を用いる条件のほうが拡声器無しの条件よりも正答数が多かった。被験者数が1名であるので断定はできないが、これら結果から、ある程度の拡声器の効果が伺える。また、1モーラも正答しなかった単語数を比較すると、条件1に比べて条件3では約半数になっていることから、拡声器とマイクの効果によって、特に聴きづらかった音が、聞きやすくなったものと考えられる。

今回の実験では、条件3の音量を、ハウリングを起こしにくいマイク3で、スピーカ位置を20cmに設定した状態でハウリングを起こさずに最大増幅できる音量に設定した。そのため、静かな部屋ではやや煩く感じる程度の音量であった。単語の正当数が条件2よりも条件3のほうが低い結果となったのは、このことが影響したと思われる。音量の増幅は、本来は環境雑音がある時に効果を発揮するものと考えられ、今後の実験では、環境雑音の条件も入れた実験を行う必要があると考えている。

5. まとめ

本研究では、電気人工喉頭による発話で課題となっている、声の音量の問題と、人工喉頭からの直接音の雑音の問題について、適切な拡声器を開発することで問題を低減することを提案した。

具体的には、双指向性マイクを接話マイクに似た方法で用いることで、人工喉頭からの直接音によるノイズを低減すると同時に、マイクとスピーカが至近距離にある場合でもハウリングが起こりにくいような拡声器を開発した。その結果、スピーカをマイクに向けて、マイクの下20cmの距離に置いた状態でも、ハウリングを起こさずに、約60dBAの声を約70dBAにまで増幅できた。

本稿では、拡声器の有無による声の明瞭さの違いを調べるため、4モーラの単語音声を発話・聴取実験を

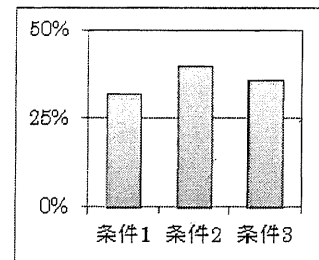


図6 正答した単語数

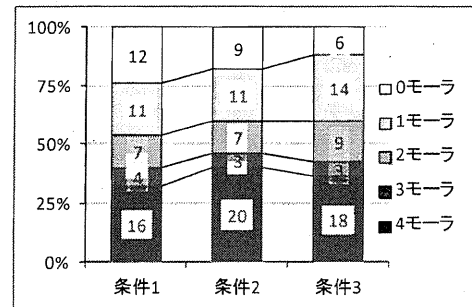


図7 正答したモーラ数ごとの語数

行った結果を報告した。実験の結果から、拡声器の効果がある程度あるものと考えられるが、今後、環境雑音等の条件を加える等の検討を行ったうえで、被験者数を増やした実験を要すると考えている。

また、人工喉頭は日常的に使用されるものであるため、マイクの固定方法や設置位置を、使用者の動きを妨げずに使えるように設定することや、脱着がスムーズに行えることも考慮に入れる必要がある。特にこの電気人工喉頭は、少ない練習ですぐに使えることを利点としていることから、そのような形状的な面での使いやすさにも配慮しながら、引き続き改良を加えていきたい。

6. 謝辞

本研究の一部は、平成23年度厚生労働科学研究費補助金(障害者対策総合研究事業身体・知的等障害分野H22-身体・知的-一般-003)によるものである。

文献

- [1] 伊福部 達, 八幡 英子, 橋場 参生, 佐々木 忠之, "電気人工喉頭の音質改善のための基礎的研究", 国立身体障害者リハビリテーションセンター研究紀要 9, 75-80, 1988
- [2] 橋場 参生, 須貝 保徳, 泉 隆, 井野 秀一, 伊福部 達, "喉頭摘出者の発声を支援するウェアラブル人工喉頭の開発と評価", ヒューマンインタフェース学会論文誌 9(2), 163-172, 2007
- [3] DENCOR 株式会社 電制 * 電気式人工喉頭 * ユアトーンII ゆらぎ <http://www.dencom.co.jp/product/yourtone/yt2.html>
- [4] 坂本, 鈴木, 天野, 近藤.: "親密度と単語の音韻バランスを統制した単語理解度試験用リストの構築", 東北大学電通談話会記録第, 69(2), 21-34, 2000.

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 -

A speech synthesis device controlled by traced position on a touch control surface
- Singing voice production based on a melody data -

藪 謙一郎¹
Ken-ichiro Yabu

伊福部 達²
Tohru Ifukube

東京大学 先端科学技術研究センター¹
Research Center for Advanced Science and Technology, The University of Tokyo.^{*1}

東京大学 高齢社会総合研究機構²
Institute of Gerontology, The University of Tokyo.^{*2}

1. 背景と目的

発話障害者の音声支援機器の方式として、テキスト音声合成方式が主流となっており、発話したい意味内容を正確に伝えるという観点でとても有力な手段となっている。

一方で、音声合成に必要な文字入力に時間を要するという課題や、リズムやタイミング、抑揚といった非言語的な情報の伝達が難しいという課題が残っている。

これらの課題の一手段として、我々は図 1 に示すような操作盤面上を、ペンタブレットやタッチ画面等でなぞることで、なぞり位置と動きを直接的にリアルタイムに音声出力する音声生成器を提案している^[1]。これは、ヒトの舌や顎の動きを、指やペンによって代替した方式で声を生成させるという発想に基づいたものである。

本発表では、この方式による音韻の制御を使って、ソフトウェアシンセサイザによる伴奏と同期して、主旋律の音声を出力できるように改良を加え、発話障害者でも楽しみながら使用できる音声生成方式を提案する。

2. システムの概要と歌声生成

音声生成器の処理の概要を図 2 に示す。音声の生成は Klatt らのホルマント合成器^[2]を基礎としており、第 1、第 2 ホルマ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疑似子音を生成する。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マント成分を除去した音声を用いている。

本発表では、新たに Standard Midi File (以下 SMF) を読み込み再生可能とした。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、伴奏音声となる。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。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生成機能を実現した。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。

3. 結論と展望

本稿では、本研究の音声生成方式と SMF の再生との統合によるリアルタイム歌声生成方式を提案した。

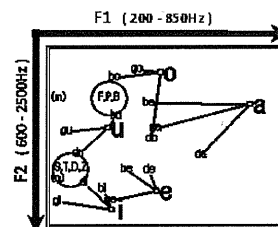


図 1 操作盤面

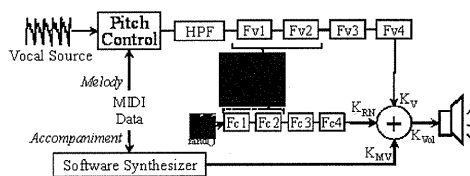


図 2 音声生成処理

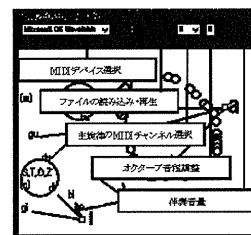


図 3 演奏データの読込と制御

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。

一方で、音韻の明瞭度や声質の変化、会話内での抑揚については課題が残るため、引き続き改良を加えていく。

文献

- [1] 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(聴覚・音声・言語とその障害、一般)", 信学技報, .SP, 音声, 105, pp:59-64, 2006.
- [2] Klatt, D.H.: Software for a cascade/parallel formant synthesizer: Journal of the Acoustical Society of America, pages 971-995, volume 67, number 3, March 1980.

電気式人工喉頭のための拡声器付きウェアラブル装着具の試作*

◎藪 謙一郎, 伊福部 達 (東大)

1 はじめに

喉頭摘出手術等で声帯を失った患者が声で意思疎通をする手段の一つに、電気式人工喉頭がある。電気式人工喉頭は、Fig. 1のように頸部へ押し当てた加振器から振動音を声道へ送り込み、生体音の代替として用いるものである。喉頭摘出患者は、多くの場合、声帯から上の構音器官の機能が残されているため、送り込まれた振動音と同時に声道を動かして会話が可能となる。ボタンを押すだけで簡単に操作ができ初心者でも扱いやすい。また、非侵襲でメンテナンスもほとんど不要であるという利点がある。

一方で、現状の手持ち型の装置では、持ち手側の手が使えなくなり、電話や会話中のメモ書きなどに不自由が生じる。このことから、使用者の患者や販売代理店などから、ウェアラブルな電気式人工喉頭を開発してほしいという意見が以前から寄せられている。

そこで著者らの研究班では、加振器を頸部に押し当てる固定具を用いた、ウェアラブル式の人工喉頭の開発に取り組んでいる。その際、加振器を頸部に押し当てる圧が弱かったり、押し当て位置がずれたりすると、加振器からの直接音が漏れ出して、声道から出された声が聞き取りづらくなる課題があった。また、手持ち式の場合と同様に、出せる音量に限りがあって大きな声を出せないため、騒音下での会話が難しいという課題もある。

これらの課題の軽減のため、話された声をマイク・スピーカで拡声するというアプローチをとり、先の研究ではマイクとスピーカの位置関係に着目してハウリングを軽減した専用拡声器についての検討を行い、一定の音量を得られた [4]。しかしながら、使用実験から拡声器と人工喉頭装置という2つの装置への増加や、マイクや加振器へのケーブル管理の煩雑さ、電源スイッチの複雑化の問題が浮上ってきている。そこで拡声装置と人工喉頭装着具とを一体化した装着具が必要であると考え、その試作に取り組んでいる。本研究では、マイクとスピーカによる拡声器と一体化したウェアラブル人工喉頭装着具の試作について、具体例を提示する。また、試作器を通して、実用に至るまでの技術的な課題を抽出し、今後に必要な改良点を提示することを目的とする。

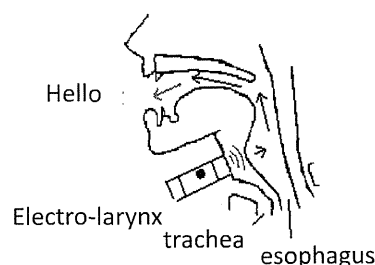


Fig. 1 Principle of electro-larynx.[1]

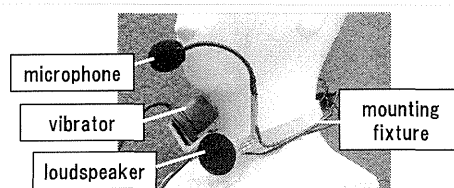


Fig. 2 A composition of a mounting fixture.

2 装着具の試作

2.1 全体構成

拡声器一体型の人工喉頭装着具を Fig. 2 に示す。装着具のアーム部分は、先の研究 [3] で用いられた熱成形樹脂を用いた。装着具の前方右側にはウェアラブル専用の加振器を取り付けた左側の先端にはマイクとスピーカを取り付けて拡声器としている。また、拡声器の増幅回路は後方に取り付けられており、マイクとスピーカまでのケーブルはアームに沿うように取り付け使用時に邪魔にならないように配慮している。

2.2 マイクとスピーカについて

拡声器のためのマイクロホンには、先の研究 [4] で用いられたものと同等のものを用いた。すなわち、Fig. 3 に示すように双指向の特性を持たせることにより、感度最大となる方向に対して直角方向の感度をゼロにすることで、下方に設置されたスピーカからの音声を排除するという考えに基づいたものである。これによって、ハウリングマージンを上げると同時に人工喉頭の直接音に対する感度を抑える。

また、マイクの固定についてはマイクを一時的に口元の位置からすぐに外せようとし、使用時にすぐに同じ場所に戻せるようにするため、位置微調整のフ

* A preproduction prototype of wearable mounting fixture with microphone amplifier for an electro-larynx. by YABU Ken-ichiro, IFUKUBE Tohru. (The University of Tokyo.)

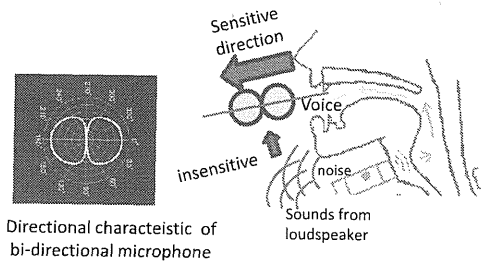


Fig. 3 Target sounds and directional characteristics of a microphone [4]

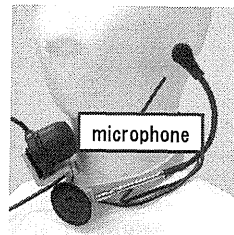


Fig. 4 Folded microphone arm.

レキシブルアームとは別に、関節を設けた (Fig. 4)。

スピーカについては、重量と大きさの観点から、東京コーン紙製作所製のφ40マイクロスピーカを用いた。ただし、重量と音質がトレードオフとなる傾向があるため、スピーカの選定には今後の調査が必要と考えている。

3 実験と考察

健常男性1名による予備的な装着実験を行った。実験では振動子の振動によって出された声が、拡声器により比較的良好に音量が拡声された。一方で、装着感の面では、マイクロホンやスピーカが話者の目の前の至近距離にあるため、首を前後に振るなどの動作にやや不自由だと意見があった。マイクを使う場合と使わない場合とがあるため、拡声ユニットを脱着可能にしてあると良いのではないかと意見があった。

マイクやスピーカの付近に手を近づけた時や、マイクの角度調整が不適当だった際、また、装着具自体を机の上に置いた際に、ハウリングを起こすことがあった。これは、スピーカからの反射音がマイクに入り込んで起こったと考えられる。ハウリング音は発生したときに即時に停止しなければ実用上の大きな問題となる。対策として例えば、拡声器の動作と振動子の動作とを連動させ、発話している間だけ拡声器が動作するといった方法をとれば、不必要なハウリングに悩まされることが少なくなると同時に、省電力化にもつながるだろう。

4 結び

拡声器と一体化したウェアラブル人工喉頭の具体例を、試作器を通して提案した。試作器による予備実験から反射音によるとみられる一時的なハウリングが観察されたが、装置全体としては拡声器と人工喉頭本体とを一体化させることにより、ケーブルやスイッチが最小限に抑えられ、従来の拡声器と人工喉頭との2つの機器を扱う場合に比べて、扱いやすくなったものと考えている。今後は、拡声器と振動子との動作の連動機能の追加や、振動子スイッチの操作性、マイクロホンの脱着等の改良を進め、患者による装着実験を進めていきたい。また、長時間使用した場合の装着感や使い勝手についてはさらなる調査が必要であろう。

謝辞 本研究は、平成24年度厚生労働科学研究費補助金(障害者対策総合研究事業身体・知的等障害分野H22-身体・知的-一般-003)によるものである。

また、研究の遂行にあたり(株)電制の須貝保徳氏及び、(株)サザン音響の稲永潔文氏に技術的なご協力を多大に頂いた。御礼申し上げます。

参考文献

- [1] DENCOM 株式会社 電制 * 電気式人工喉頭 * ユアトーンゆらぎ <http://www.dencom.co.jp/product/yourtone/yt2.html>
- [2] 伊福部 達, 八幡 英子, 橋場 参生, 佐々木 忠之, "電気人工喉頭の音質改善のための基礎的研究", 国立身体障害者リハビリテーションセンター研究紀要 9, 75-80, 1988
- [3] 橋場 参生, 須貝 保徳, 泉 隆, 井野 秀一, 伊福部 達, "喉頭摘出者の発声を支援するウェアラブル人工喉頭の開発と評価", ヒューマンインタフェース学会論文誌 9(2), 163-172, 2007.
- [4] 藪 謙一郎, 上田 一貴, 稲永潔文, 伊福部 達, "電気人工喉頭のための拡声器に関する一考察", 信学技報, vol. 111, no. 471, SP2011-161, pp. 35-38, 2012.

福祉・介護テクノロジーにおける価値観の変容

伊福部 達氏
 東京大学・高齢社会総合研究機構
 〒113-8656
 東京都文京区本郷7-3-1
 東京大学工学部8号館701
 E-MAIL ifukube@human.iog.u-tokyo.ac.jp

1 はじめに

私は福祉工学という分野を歩き続けて40数年になりますが、この間に、日本人の価値観は大きく変わり、また、多様化していることをつくづく感じます。一昔前は、「金持ちになること」、「偉くなること」、「長生きすること」を目標としていた人が多かったのに対して、最近は「いかに楽しいか」、「やりがいがあるか」、「いかに快適な生活を送るか」ということに価値を置く人たちが明らかに増えています。そして、快適な生活を支援するためのテクノロジーの1つとして、当事者だけでなく産業界でも福祉工学に期待を寄せるようになってきています。ここでは、私が歩んできた福祉工

学の道を振り返りながら、これからは必要とされる福祉機器について私見を述べたいと思います。

2 福祉工学の道へ

私は、学生時代は電子工学科に所属していましたが、卒業論文を書くために「メデイカルエレクトロニクス(ME)部門」という研究室に配属されました。当時のME部門は、電子工学を医学に応用することを目指しており、研究の主流は生命維持を目指した人工臓器でした。特に呼吸・循環器系が働くなつた人々を人工心肺で助ける研究が多くを占めていました。試作した人工心肺が予想した通りに働くかをチェックするため、いつもイヌやヒツジなどの動物たちが実験台になっていました。そして、当然ながら動物たちが息を引き取るまで実験は続きました。

一方、学科では電子工学の基礎となる物理や数学という理路整然とした理論を徹底的に叩き込まれました。その講義を聴いたあと、イヌの胸を切り開き人工呼吸

器を接続する様子を横目にしながら、また、胸を開くときに大量に噴出する血の臭いを嗅ぎながら、卒論のための研究に頭を切り替えなければならなかったのです。学生は実験ばかりでなく医学部で生理学などの基礎医学を受講するよう指導されました。医学部の講義では「人の命は金には変えられない」と、命の尊さを教えられ、一方、電子工学科では、これからは製造業が日本の経済を引っ張っていくのだと鼓舞され、「金になるものを作る」ことが我々の使命だとして胸に刻み込まれていたのです。

私の頭の中には、このような相入れない2つの価値観が混在することになり、それらを融合させた新しい価値観を探し続けることになりました。その結果、心臓や肺などの臓器ではなく、社会生活を送る上で不可欠な「感じる」、「考える」、「動く」を技術で支援する医療工学の道を選ぶことにしました。弱ったり失ったりした「感覚」、「脳」、「運動」を支援する技術は医療に貢献しますし、そこか

ら優れた人工の感覚、脳、運動の機能が生まれれば、それらは「センサ」「コンピュータ」「アクチュエータ」などの先端技術に寄与すると考えたのです。当時の私の頭では、このように二つの価値観を強引に融合させることしかできませんでした。結果的には、現在求められている「快適な生活」や「生きがいのある生活」を支援する福祉工学への道に結びついていったと思います。

3 福祉機器産業の宿命

しかし、「感覚」や「運動」でも特に「脳」が関与した機能を支援する福祉機器の産業化は未だに暗中模索の部分が多いのが実情です。一般に、福祉・介護機器を設計する場合、車椅子やベッドにしても、抛り所となる力学や化学などのサイエンスがあります。それに対して、例えば、音声の聞き取りが低下したのを補うための機器を設計するとすると、その抛り所となるサイエンスは未知の部分の多い聴覚と脳の科学になります。特に感覚や脳の一部の機能が

失われたり弱ったりすると、生体特有の「可塑性」により、それらを代償する機能が働きますので、問題は単純ではありません。首尾よく製品ができたとしても、今度はそれを必要とする人の数、すなわちマーケットが小さいために高価なものになり、結局、当事者の手の届かないものになってしまいます。大企業ならば採算が合わなくても自力で価格の壁

を乗り越えられるかもしれませんが、中小企業はそのような余力はありません。そもそもマーケットが小さいことを知りながら製品化を進める大企業は少ないでしょうし、また、福祉機器の開発には基礎研究が必要となればなおさらは敬遠されることになるでしょう(図1)。

そのために、福祉機器とりわけ生体機能を支援する機器は、開発

本格的な取り組みが行えなかった理由

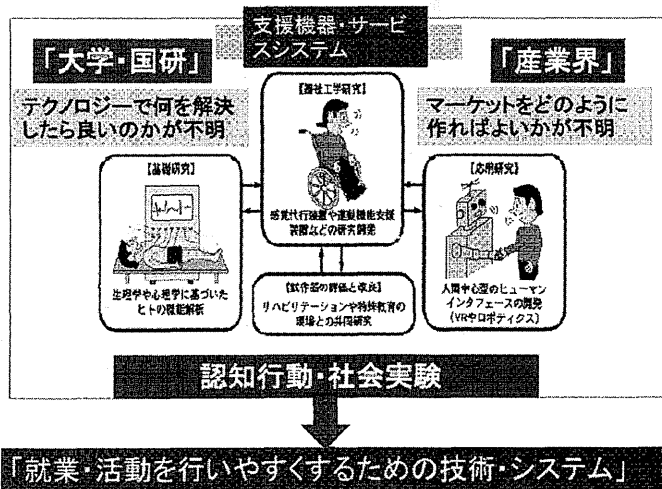


図1 福祉機器開発の問題点

4 高齢社会を豊かにする

科学・技術・システム

この40年の間に、日本の超高齢化はますます現実のものとなってきました。福祉工学は何らかの形で高齢社会で必要な課題に応えるよう求められています。その課題は、社会的にみると高齢化による「労働者人口の減少」と「社会保障費の増加」に起因します。また、個人としてみると、高齢になっても「できる限り働きたい」、障害が生じて「できる限り自立して生活したい」という人たちが増えていることによります。時代は、「生きがい」や「快適な生活」を送りたいという価値観に重心が置かれるようになってきているのです。

4-1 高齢者の社会参加を促す

そのような中で、2010年に科学技術振興機構(JST)の