

重度障害者特殊スイッチ等の開発

Development of switches for the people with severe disabilities

— 音声認識スイッチの開発(第三報) —

— Developing a voice-recognition switch (report3) —

宇根 正美、尾田 継之
UNE Masami , ODA Tsuguyuki

キーワード：スイッチ、制御インターフェース、
音声認識

Keywords：Switch, Control Interface,
Speech-recognition

Abstract：

The voice recognition switch unit developed in last year was combined with the telephone and the electrically-driven bed, and it tried. Then, the operation was confirmed. The use of the personal computer was examined to add the guidance function for importance that to be confirmed further. Voice recognition software for the personal computer on the market was used, and the demonstration software which an environment control device was presumed was made, and the function of the marketing voice recognition software was evaluated by using it.

1. はじめに

重度の身体障害を持つ人が生活用具やコミュニケーション機器を扱うときには、残存する身体機能を生かすためにさまざまな補助具が使われる。特にパソコンなどの情報機器や電子化された機器を操作するためには、さまざまなソフトウェアやスイッチが必要に応じて使い分けられる。ただ、これらソフトウェアやスイッチは、さまざまな要因で十分に活用されていない場合が多く、その大きな要因として操作や個別の状況とのフィッティングの難しさが上げられる。

これを改善するために我々は音声認識を利用したスイッチの開発を行ってきた。本年度は平成11年度に製作した音声認識スイッチユニットを電動ベッドや音声認識電話と組み合わせて試行し、その操作性

を確認した。操作案内の提示機能の重要性が確認され、それを強化するために音声や画像などで操作案内しやすいパソコンの音声認識の利用を検討した。検討には市販音声認識ソフトを利用して環境制御装置を想定した評価用ソフトを制作し、それを用いて機能を評価した。

2. 音声認識スイッチユニットの応用

昨年度、ハードウェアを開発した音声認識スイッチユニット¹⁾を音声認識電話や電動ベッドと組み合わせてその操作性を確認した。組み合わせた音声認識電話はユニデン株式会社²⁾製のもので、これはボタンをひとつ押せば後は音声だけで操作できる音声認識電話器となっている。音声認識スイッチユニットと組み合わせることで、完全に音声だけで電話操作できるようになる。また、電動ベッドはパラマウントベッド株式会社製³⁾の電動ベッドで、手元リモコンに音声認識スイッチを組み付けて、音声でベッドを操作できるようにした。

昨年度開発した音声認識スイッチユニットの主な仕様は、表1のようになっている。音声認識は利用者毎に言葉を登録する特定話者認識タイプで、このスイッチユニットでは最大15件の言葉を登録することが可能になっている。なお、認識できる言葉は安定した発話を得られるものであればよいが、発話のばらつきを考慮すると精度の良い認識結果を得るにはなるべく似た発音の無い異なった発音の言葉を選ぶ必要がある。また、スイッチユニット内部のプログラムを書き換えることで組み合わせる機器に合わせてスイッチユニット動作を調整可能になっている。

表1 音声認識スイッチユニットの仕様

Table1 Specification of voice recognition switch

音声認識部	使用モジュール	VoiceDIRECT
	音声認識のタイプ	特定話者、単語
プログラム部	音声認識可能な語彙	最大15
	記述言語	BASIC
入力部	プログラムサイズ	500行
	音声信号	コンデンサマイク 感度-38dB~-42dB@600Hz~5kHz
	アナログ信号	0~5V, 2入力
	デジタル信号	Logic1入力
出力部	プログラム用	RS232C
		半導体リレーによる無接点出力 16出力
電源		2V~6V 範囲外の場合はアダプタ利用
大きさ		95×130×20
質量		172g(本体90g)

2.1 組み込みプログラムの調整

音声認識スイッチユニットは組み込む機器に応じてタイミングなどを調整するためにスイッチの組み込みプログラムを調整する必要がある。今回実験に利用した機器では以下の内容を作りこんだ。

(a) 音声認識電話の場合

音声認識スイッチユニットに要求されるのは、電話の音声認識機能をオン、オフさせる機能だけでよい。スイッチ動作としては、100ms 程度接点信号が「閉」になるようスイッチのプログラムを調整した。なお、操作する単語は「電話」「オフ」を使った。

(b) 電動ベッドの場合

今回利用した電動ベッドでは六つの操作ボタンがある。音声認識スイッチユニットとしてはその操作ボタンに対応する言葉で操作ボタンの信号を「閉」にしてベッドを動かし、別の言葉で操作ボタンの信号を「開」にしてベッドの動きを止めるようにした。操作する言葉は以下のように選んだ。

- 「あしあげ」「さげあし」: 足部動作起動
- 「たかさ」「ひくさ」: 高さ動作起動
- 「あたまあげ」「さげあたま」: 頭部動作起動
- 「オフ」: 動作停止

2.2 試用結果

(a) 音声認識電話の場合

音声認識スイッチを組み合わせた様子を図1に示す。試用してみると音声認識が中断することが大きな問題であることがわかった。これは開発当初からある程度は影響を予想していたが、音声認識が中断しているときは発光素子などで明示すればさほど問題にならないと考えていた。しかし、実際に機器と組み合わせて操作してみると想像以上に操作性が悪くなる。機械のタイミングに合わせて操作命令を発話するのはかなりストレスを感じた。

ただ、音声認識が中断する問題以外は安定して操作できた。操作する言葉が二つと少なく済み、操作のタイミングを計る必要がないため、落ち着いて安定した発話できたことが影響している。安定した発話であるため、音声認識も安定動作したと考えられる。



図1 音声認識スイッチと音声認識電話

Fig.1 Voice recognition switch and voice recognition telephone

(b) 電動ベッドの場合

音声認識スイッチを組み合わせた様子を図2に示す。



図2 音声認識スイッチと電動ベッド

Fig.2 Voice recognition switch and electrically-driven bed

電動ベッドでの試行では「とっさのときにスムーズな発話ができない、必要な言葉が思い出せない」問題が顕在化した。電動ベッドの動きに合わせて必要なタイミングで発話することが難しく、スムーズな発話ができない、あるいは焦って操作に必要な言葉が思い出せない状況になりやすかった。もちろん、

慣れの問題が大きいですが、少なくとも機器操作に慣れるまではスムーズな発話を促すために機器操作の案内機能が重要であることがわかってきた。

電話と組み合わせた試行で問題になった音声認識の中断は、電動ベッドではさらに大きな問題となった。認識が中断すると必要なタイミングで操作することが困難になるためである。ただ、これについては「ベッドの動きを遅くする」「短時間だけスイッチをオンにする」などでかなり改善できるとみられる。

2.3 音声認識スイッチ仕様の見直し

以上の試用結果から音声認識スイッチを図3のように簡素なものとして音声認識スイッチの仕様をまとめなおした。大幅に仕様を絞り込んだものになっているが、スイッチとして実用的に使うにはこの程度仕様を絞り込む必要があると考える。また、複雑な操作にはWing-Sk⁽⁴⁾のようなプログラブルな入力支援装置と組み合わせればよい。

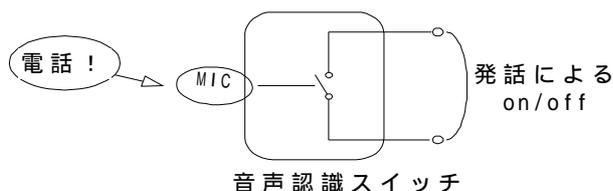


図3 簡素化した音声認識スイッチ

Fig.3 Simple voice recognition switch

このスイッチに組み込まれる機能は、以下のような簡素なものになる。

- (1) 連続音声認識
- (2) 特定話者の一つか二つの単語だけの認識
- (3) 一つか二つの接点出力

これは操作内容の案内機能を付けられない場合、なるべく簡単な内容にする必要があると考えた結果で、最新の音声認識IC⁽⁵⁾では連続認識も可能になったことを踏まえた仕様となっている。次年度以降このスイッチ開発に取り組みたい。

なお、環境制御装置などではスイッチ操作に音声認識を利用する場合、操作内容の案内機能を付けることはあまり問題にならないと考えられる。そこで操作内容の案内機能を付加しながら音声認識を行うにはパソコンが最も手頃と考え、最新の市販パソコン音声認識ソフトがどの程度機器操作に利用できるのか音声認識率を中心に確認を行った。

3. パソコンを利用した音声認識

3.1 評価対象の音声認識ソフト

評価には「IBM製 ViaVoice Ver8」「NEC製 Smar

tVoice i Ver4」「東芝製 LaLaVoice2001」の三種類の主立った市販音声認識ソフトを利用した。なお、これらのソフトはもともと健常者向けの口述筆記用（dictation）として製品化されたもので、健常者が口述筆記として利用した場合の評価は雑誌⁽⁶⁾などですで行われている。口述筆記に慣れた健常者の文章入力では90%以上の認識率となっており、文章入力としては概ね良好な結果となっている。

各々のソフトにはエンロール（音声認識精度を高めるための話者の特徴を抽出する作業）や認識しづらい単語の登録について以下の特徴がある。

(a) IBM 製 ViaVoice

エンロールには、音声認識精度の度合いによって「エンロール」「クイックエンロール」の2種類が用意されている。「エンロール」の場合15分程度以上時間を要するが、「クイックエンロール」の場合、1分程度で完了する。また、音声入力するときに認識しにくい単語を修正することが可能である。

(b) NEC 製 SmartVoice

音声認識精度の度合いによって「簡単エンロール」「標準エンロール」が用意されており、各々ViaVoiceと同程度の時間を必要とする。これもViaVoiceと同様に、認識しにくい単語を修正登録することができる。また、エンロールなしで使うことも可能になっている。

(c) 東芝製 LaLaVoice

今回評価した中で唯一まったくエンロールが不要なソフトである。ただ、音声認識処理に他のソフトより多少時間が掛かることから速度優先と精度優先が選べるようになっている。また、他のソフトと同様に固有名詞などの認識しにくい単語を修正登録できるようになっている。ただし、標準で登録されている一般的な名詞は登録できない。

3.2 評価方法

音声認識を環境制御装置などの機器操作の入力として使うには、口述筆記として使うより高い認識率が必要と考えられる。機器操作をスムーズに行うには誤認識による操作の繰り返しが少ない方が望ましいためである。

なお、音声認識では認識すべき単語を限定することで認識率を向上できることが経験的にわかっている。そこで認識率を高める工夫として認識すべき単語を限定した辞書を利用したときの認識率の向上を実験で確かめた。また、この他にも音声認識で機器操作するとき予想される発話速度の影響について確認を行った。

3.2.1 実験に利用したパソコン

確認実験には小型で周囲環境に合わせやすいIsaintsong社⁷⁾製のはがき大の小型パソコンを利用した(図4)。OSには安定性の高いWindows2000を使っている。また、音声入力には試行段階で認識率が高かったSmartVoiceの添付マイクを使った。



図4 はがき大の小型パソコン
Fig.4 Pocket size PC

3.2.2 確認に利用した単語

環境制御装置に使われそうな単語で、以下の単語を使って確認を行った。単語に制限を加える場合は下記の単語だけを登録した辞書を用い、制限を加えない場合は購入時添付の辞書をそのまま(あるいは認識しにくい単語を修正登録して)用いる。

ベッド、電灯、テレビ、エアコン、電話、つける、消す、音量、上げる、下げる、チャンネル、2(に)、4(よん)

12(じゅうに) 温度、風量、かける、切る

なお、確認は40才代男性と20才代女性一名ずつで行った。

3.3 単語限定の効果

単語を限定した場合の評価は図5のような簡単な評価用ソフトを制作して行った。

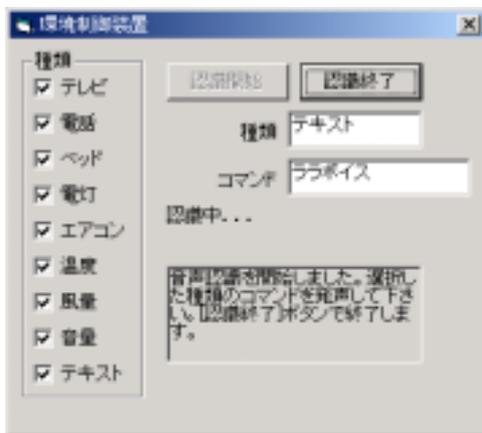


図5 評価ソフト画面(LaLaVoiceの場合)
fig.5 Evaluation soft screen shot(LaLaVoice)
単語限定した場合と限定しない場合の認識率を

表2、表3に示す。

表2 単語限定による効果
table2 Effect by the word limitation

音声認識ソフト	エンロールの種類	認識率
ViaVoice	クイックエンロール	94% (171/180)
	エンロール	96% (173/180)
SmartVoice	エンロールなし	99% (303/306)
	簡単エンロール	99% (303/306)
	標準エンロール	99% (305/306)
LaLaVoice	—	97% (456/468)

(注)認識率中の括弧書きは(成功数/試行回数)

表3 単語限定しない場合
table3 Without the word limitation

音声認識ソフト	エンロール種類/単語登録	認識率
ViaVoice	クイック/登録なし	58%(63/108)
	エンロール/登録なし	62%(67/108)
	クイック/登録あり	68%(37/54)
	エンロール/登録あり	79%(86/108)
SmartVoice	なし/登録なし	53%(58/108)
	なし/登録あり	72%(78/108)
	簡単/登録なし	57%(62/108)
	簡単/登録あり	75%(81/108)
	標準/登録なし	64%(70/108)
	標準/登録あり	67%(73/108)
LaLaVoice	精度優先/登録なし	46%(50/108)
	速度優先/登録なし	50%(54/108)
	速度優先/登録あり	71%(77/108)

(注)認識率中の括弧書きは(成功数/試行回数)

単語を限定しない場合は、最高でも79%の認識率しか得られていない。これに対して単語を限定すると94%以上の高い認識率が得られた。なお、単語制限をかけない場合の認識率が低いのは本来文章入力目的のものを単語認識だけで利用した結果と考えられる。文章であれば前後の文脈からも単語を推定できるが、単語だけでは判別できないためこのような結果になったと推測される。

3.4 発話速度の影響

焦っている状況などでは発話速度がかなり異なる可能性が考えられる。ついては、発話速度がどの程度認識率に影響するのか「テレビつける」という例文で確認してみた。

その結果を表4に示す。予想では早口のときの認識率は下がると考えていたが、各ソフトともほとんど認識率の低下は見られなかった。ただ、ゆっくり

とした発話の場合、LaLaVoiceの認識率が高くなる結果となった。

表4 発話速度の影響

table4 Influence of the utterance speed

「テレビつける」 を読む時間	ViaVoice	SmartVoice	LaLaVoice
約0.8s	100%(5/5)	100%(5/5)	100%(5/5)
約1s	100%(10/10)	100%(10/10)	100%(10/10)
約1.3s<通常>	0%(0/10)	100%(10/10)	100%(10/10)
約2.0s	0%(0/10)	0%(0/10)	100%(10/10)
約2.4s	0%(0/5)	0%(0/5)	100%(5/5)
約3.1s	0%(0/5)	0%(0/5)	0%(0/5)

(注)認識率中の括弧書きは(成功数/試行回数)

3.5 パソコン利用の音声認識のまとめ

(1) 単語限定の効果

辞書の単語を限られた単語に限定することで認識率を大きく改善できることを確認できた。単語入力の場合、単語を限定しないと79%以下の認識しかできなかったが、辞書の単語を限定することで最も悪い結果でも94%の認識率の向上が見られた。

(2) 単語読み上げ方の影響

発話速度が早くなる分にはあまり影響はない。むしろ発話速度が遅い場合、認識率が下がる傾向がある。また、複数の単語を続けて発話すると認識率が低下する傾向があった。しかし、単語の間に時間を置いて発話すると通常の単語認識として動作することになり、高い認識率を得ることができた。

(3) エンロールの影響

単語を限定した場合はエンロールを行っても認識率はあまり変わらなかった。口述筆記を行わない機器操作だけの音声認識であれば、エンロールは必要ないかもしれない。

4. 考察とまとめ

前年度から開発を進めてきた音声認識スイッチは、認識単語数を1、2個に限定すれば案内装置が不要となり、使いやすいものになると考えられる。これに連続音声認識機能が加われば、スイッチとして実用的な利用が可能と考える。来年度以降は今回新たにまとめ直した簡素化した音声認識スイッチとして開発を進めたい。手軽に使える安い音声認識スイッチとすることに重点を置く予定である。

また、パソコンの音声認識ソフトは単語を限定することで高い認識率が得られることがわかった。特に発話速度にあまり影響を受けないことやエンロールの必要性がないことは、機器操作する上で重要な点と考える。来年度以降の開発では音声認識も利用

できる環境制御装置として開発を進めたい。

なお、音声認識ソフトは、現在急速に発達している分野でその性能の発達が著しい。そのため、特定の音声認識ソフトだけを利用した装置では、機能の陳腐化が進みやすく、すぐに機器が使われなくなる可能性も高くなる。そこで、今後の開発では各メーカーの音声認識ソフトを交換できるように、ソフトウェアの内部インターフェースはできるだけ共通にして開発を進めたい。また、発話に不自由がない障害者の場合、環境制御装置だけの利用ではなく、口述筆記による文章入力の要求も強いと見られる。そのため、機器操作のための単語を限定した辞書と文章入力用の辞書を使い分ける方策についても検討したい。

参考文献

- 1) 宇根正美他：「重度障害者特殊スイッチ等の開発」, 福祉のまちづくり工学研究所報告集平成11年度版, pp.88-91 (2000).
- 2) ユニデン株式会社の URL
<http://www.uniden.co.jp/>
- 3) パラマウントベッド株式会社の URL
<http://www.paramount.co.jp/>
- 4) Wing-SK などの URL
<http://www5.wind.ne.jp/ja1syk/>
- 5) sensory 社の音声認識が紹介されている URL
<http://www.sensoryinc.com/html/products/technologies.html>
- 6) 音声認識ソフトの実力診断, 日経パソコン, 1/8号, pp128-137(2001).
- 7) saintsong 社の URL
<http://www.saintsong.com.tw/>