

# 体内伝導音認識システム構築に関する基礎的検討

北風裕教\* 石光俊介\*\* 神笠真穂\*\*\*

## Study for Constructing a Recognition System using the Body Conduction Speech

Hironori KITAKAZE, Shunsuke ISHIMITSU and Maho KAMIGASA

### Abstract

The effectiveness of the sentence recognition in body conduction was examined. In a pretest session, reception experiments took place in order to investigate if body conduction is actually effective for recognition systems. We investigated the difference between word-recognition rate and sentence-recognition rate. In the next part, we compared word-recognition rate with sentence-recognition rate with recognition systems which are on the market. In addition, we investigated the effect of recognition in females. Generally, female production signals are not easily perceived by detectors as females are relatively vulnerable to sentence-recognition. As a result, it was confirmed that the body conduction was effective for sentence-recognition as well as word-recognition.

Key words: Body Conduction, Muscular Conduction, Bone Conduction, Recognition System

### 1. 緒言

動力を伴う工場や船舶機関室内では信号より雑音の方が大きい  $0\text{dB SNR}$  以下の環境における作業が要求され、それに伴い通信システムを含む音声認識システムの早急の開発が必要となった。このような環境下での音声認識システムには雑音対策が必須となる。低 SNR 環境において音声そのものを対象とすると雑音も音声の一部であるとシステムは判断し認識率が著しく低下することから、これまでに低 SNR 環境での音声認識の研究はほとんどなされていなかった。

㈱Pioneer では低 SNR 環境の通信システムを、音声ではなく体内伝導音を用いて実現している。これは騒音が固体伝播する体内伝導信号に入り込めないという現象を利用したものである。本研究ではこの通信システムから認識システムの可能性を見出し、上唇左上部に設置した加速度検出器から体内伝導音を抽出することで単語認識率 90% 以上のシステムを構築することができるといった[1-6]。しかし実験は全て単語レベルで行っており、連続発声文章の検討は行ってはいないとい

う問題点があった。

したがって本研究では、体内伝導音の文章認識における有効性の検討を行った。まず予備実験として、実際に認識システムに体内伝導音が効果的であるのかを人間による聴取実験を通して行い、単語と連続発声文章の認識率の違いについて検討を行った。次に市販の認識システムを用いて単語と連続発声文章の認識率について比較検討を行った。更に加速度検出器で信号抽出が困難とされる女性の認識率への影響についても検討を行った。その結果、体内伝導音は単語認識だけではなく、文章認識においても非常に効果的であることが明らかとなった。

### 2. 体内伝導音

体内伝導音とは骨または筋肉を伝播する音で、固体伝播されている信号である。このため、空気伝播されている雑音はこの体内伝導音に入りこむことはできない。体内伝導音を解析した結果、骨導音・声帯信号・筋肉伝導信号の3種類があることが明らかとされている[4]。

しかしながら、骨は声帯で作られる有声音(体内で響く)は良く伝わるが、唇などで作られる破裂音、摩擦音のような無声音は非常に伝わりづらい。一方、このような無声音は体内で響かずに体外へ放出される。筋肉伝導では唇付近の運動および共振信号も採取することができるので、有声音のみならず無声音もよく伝えることができると考えられる(図1)。

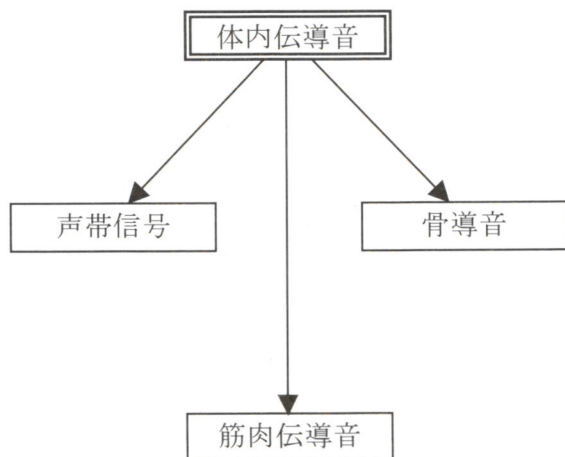


図1 体内伝導音

### 3. 聴取実験(予備実験)

体内伝導音の文章認識における有効性を調べるために、まず予備実験として体内伝導音を人が知覚できるか否かについて聴取実験を通して検討する。ここで聴取実験でも認識率が高ければ認識システムにおいても高い認識率が期待できる。

#### 3.1 実験準備

体内伝導音を用いた聴取実験として、予め男性2名(Mmn氏, Mhk氏), 女性2名(Fmk氏, Fnt氏)の体内伝導音(上唇左上部, 頬骨弓前部, 咽頭部下部)を抽出した。この部位で抽出することができる信号は、上唇左上部からは筋肉伝導音, 頬骨弓前部からは骨導音, 咽頭部下部からは声帯信号がそれぞれ抽出できる。抽出は加速度検出器(仕様: 表1)を用い、チャージアンプ(TEAC SA630)を介してレコーディングユニット(TEAC LX-10)に各12kHzサンプリングで同時記録する。図2に加速度検出器とその取り付け位置を示す。

抽出した認識語彙は電子協地名100の単語セットと機関長システムの文章セットである。電子協地名100とは日本全国の地名の中から音素の出現率が一樣になる様に集められた100個の地名のこ

とである。また機関長システムとは、船舶機関室で教育支援の為に開発が進められている音声認識システムであり、システムで準備されている76からなる命令文章である[7]。表2に地名百選で用いた単語と、機関長システムで用いた文章を示す。

被験者は聴覚障害のない20代(男性3名, 女性1名)で予備実験を行い簡単な練習を行っている。体内伝導音はヘッドホンを通して聞いてもらい、知覚できた内容を準備した回答用紙に直接記入してもらう。



図2 加速度検出器と取り付け位置

表1 加速度検出器 TEAC702FB仕様

感度	1.0±10% $mV/s^2$
横感度	5%以下
基底ノイズ	20 $\mu$
補償温度範囲	-50~110度
周波数応答(±3dB)	2~20,000Hz
共振周波数	約65kHz
使用最大加速度	3,600 $m/s^2$
耐衝撃性(peak)	50,000 $m/s^2$
出力インピーダンス	100 $\Omega$ 以下
重量	約2g
外形寸法	$\phi$ 7.9×11.9H
内臓プリアンプ駆動電源	電圧: 15V~25V

表2 地名百選(単語)と機関長システム(文章)

地名百選(単語)
八戸, 気仙沼, 行橋, 札幌, 北見, 恵庭, 横手
機関長システム(文章)
暖機用せいすいヒーター電源の押しボタンスイッチを押してください。
主空気タンクのドレインを排出して異常ないことを確認してください。
主機L Oスタンバイポンプ圧力を制御室の圧力計にて確認してください。
ターニングギアのかんごう状態を確認してからモーターをスタートしてください。

### 3.2 聴取実験結果(電子協地名 100)

表3に電子協地名100における聴取実験の結果を示す。単語認識においては頬骨弓前部と上唇左上部は90%以上の認識率であるのに対して、咽頭部下部は、50%程度の認識率しか得られなかった。上唇左上部に至っては96%と非常に高い認識率が得られ、音声と区別がないことが示唆される。咽頭部下部で間違いの多い地名には、“久居”や“珠洲”など摩擦音や破裂音を含む子音部の場合が多く、誤認識単語の80%以上を占めていた。

表3 電子協地名100(単語)の聴取実験結果

体内伝導音データ		電子協地名100(単語)		
		頬骨弓前部	上唇左上部	咽頭部下部
被験者	Mmn	95.25%	96.50%	63.75%
	Mhk	92.75%	96.75%	39.75%
男性3名	Fmk	88.00%	93.50%	45.75%
女性1名				
平均	Fnt	92.75%	97.75%	60.00%
単語認識 平均		92.19%	96.13%	52.31%

### 3.3 聴取実験結果(機関長システム)

表4に、機関長システムにおける聴取実験の結果を示す。認識結果の算出方法は、被験者により回答された文章を文節に区切り、その全文節に対する正答率とする。

文章認識は単語認識と比べて各部位全て平均認識率が向上した。また、単語認識では咽頭部下部が非常に認識率が低下したのに対し、機関長システムの文章では咽頭部下部は平均84%以上の結果が得られた。文章の方が認識率が高い理由として、劣化した体内伝導音を前後の文脈の意味から想像することで補っている可能性が考えられる。また単語認識結果と同様に上唇左上部が最も認識率が高かった。

表4 機関長システム(文章)の聴取実験結果

体内伝導音		機関長システム(文章)		
		頬骨弓前部	上唇左上部	咽頭部下部
被験者	Mmn	98.39%	98.95%	92.08%
	Mhk	97.59%	99.38%	85.03%
男性3名	Fmk	99.01%	99.44%	90.53%
女性1名				
平均	Fnt	98.58%	99.32%	69.24%
文章認識 平均		98.39%	99.27%	84.22%

## 4. 認識システムを用いた認識実験

次に市販の認識システム(I B M株製ViaVoice)を用いて単語と連続発声文章の認識率について比較検討を行う。

### 4.1 実験準備

体内伝導音を用いた認識実験の被験者として、男性4名、女性4名を採用した。被験者は予めエンロールによる話者適応を行い、認識率の向上を行っている。認識語彙には聴取実験同様、単語(電子協地名100)と連続発声文章(機関長システム)を用いる。機関長システムで用いられる専門用語は予めシステムに単語登録を行い利用できる状態である。電子協地名100はViaVoiceリスト認識を採用し、限定した100個の地名の中から最適な認識語彙が選ばれる方法を用いる。機関長システムはViaVoiceテキスト認識を採用し全登録単語の中から最適な語彙が選ばれる方法を用いる。加速度検出器の取り付け位置は、予備実験の結果から最も認識率が高いとされた上唇左上部を採用した。図3に実験風景を示す。



図3 ViaVoiceによる体内伝導音認識実験風景

### 4.2 認識システムによる結果(電子協地名100)

表5に電子協地名100(単語)の認識結果を示す。認識実験は被験者各3回行い平均を求めた。この結果、男性で平均95.9%、女性では平均96.3%の高い認識率を得ることができ、男性と女性間に認識率に違いがないことが明らかとなった。誤認識箇所(約4%)の検討を行なった結果、聴取実験でも認識が困難であった“久居”や“珠洲”など子音部(摩擦・破裂)において誤認識が集中した。

### 4.3 認識システムによる結果(機関長システム)

表6に機関長システム(文章)の認識結果を示す。認識結果の算出方法は、ViaVoiceにより認識された文章を文節に区切り、その全文節に対する正答率とする。認識実験は被験者各3回行い平均を求めた。この結果、男性で平均81.8%、女性で

は平均 85.0%の認識率が得られ、文章を用いた場合においても男性と女性との認識率に違いがないことが明らかとなった。単語認識と比較して認識率に平均 10%の低下が見られたが、80%以上の高い認識率が得られた。

表5 電子協地名 100 の ViaVoice 認識結果

地名 100		1回目	2回目	3回目	平均
男性	Mtt	97.0%	100%	93.0%	96.7%
	Mhk	95.0%	93.0%	98.0%	95.3%
	Mmn	97.0%	98.0%	98.0%	97.7%
	Msm	95.0%	93.0%	94.0%	94.0%
	平均	96.0%	96.0%	95.8%	95.9%
女性	Fsf	94.0%	95.0%	97.0%	95.3%
	Fmk	94.0%	97.0%	98.0%	96.3%
	Frm	100%	100%	99.0%	99.7%
	Fmh	91.0%	94.0%	97.0%	94.0%
	平均	94.8%	96.5%	97.8%	96.3%

表6 機関長システムの ViaVoice 認識結果

機関長		1回目	2回目	3回目	平均
男性	Mtt	70.1%	71.0%	64.9%	68.6%
	Mhk	86.4%	93.3%	96.3%	92.0%
	Mmn	80.5%	79.5%	81.2%	80.4%
	Msm	82.9%	86.9%	89.1%	86.3%
	平均	80.0%	82.7%	82.9%	81.8%
女性	Fsf	73.0%	78.7%	78.5%	76.7%
	Fmk	79.7%	86.1%	88.1%	84.7%
	Frm	86.9%	93.1%	94.3%	91.4%
	Fmh	87.9%	85.6%	88.6%	87.4%
	平均	81.9%	85.9%	87.4%	85.0%

## 5. 結言

体内伝導音の文章認識における有効性の検討を行った。聴取実験の結果、体内伝導音は単語認識と文章認識に違いがないことが明らかとなった。また、設置部位に関しては上唇左上部の認識率が高く、咽頭部下部位においては著しく認識率が低いことが明らかとなった。また認識システムによる

認識実験においても単語認識と文章認識にはそれ程影響がないことが示唆される。更に男性と女性においても認識率には影響がないことが明らかとなった。これらの聴取実験の結果より、加速度検出器を上唇左上部に取り付ければ有効的な認識システムが構築可能であると予想される。

## 謝辞

実験を手伝ってくれた情報工学科卒業生の田中寿宜氏、中野雅文氏に感謝致します。また、信号取り出し位置について有効な意見を頂いたパイオニア(株)好美氏に深く陳謝致します。

なお、本研究は科学技術研究補助金若手研究(B)(No.14750325)により遂行することができました。ここに感謝の意を表し、謝辞とさせていただきます。

## 参考文献

- [1]石光, 北風, 土伏, 石川, 高田, 齋藤, 柳川, 福島: 骨導音認識システム構築のための基礎的検討, 日本音響学会 2001 秋季研究発表会後援論文集
- [2]石光, 北風, 土伏, 高田, 石川: 骨導音を利用したノイズロバストな認識システム構築の検討, 日本機械学会中国支部 第 40 期総会・講演会 講演論文集 No.025-1
- [3]石光, 高家, 堀畑, 北風, 柳川: 適応フィルタを用いた骨導音明瞭度向上の基礎的検討, 日本音響学会 2002 年度春季研究発表会後援論文集 pp681-682,2002
- [4]石光, 北風, 土伏: 船用機関運転支援のための体内伝導音認識に関する研究, 第 67 回日本マリンエンジニアリング学会 学術講演会 pp79-82,2002
- [5]石光, 北風, 土伏, 柳川, 福島: 体内伝導音認識システム構築のための基礎的検討, 電子情報通信学会技術報告, EA2002-34 (2002) pp.25-28
- [6]北風, 神笠, 石光, 安田, 田中, 中野: 体内伝導音を利用した認識システム構築の検討, 日本機械学会 中国四国支部 第 41 期総会・講演会
- [7]松下他 “対話型船用機関運転支援システムの概念設計”, 日本マリンエンジニアリング学会誌第 36 巻 6 号(2001), 34