

発音の視覚化のための予備検討

筑波技術大学 産業技術学部 産業情報学科¹⁾, 静岡県立大学 経営情報学部 経営情報学科²⁾

東京医科歯科大学大学院 医歯学総合研究科口腔機能再構築学系³⁾

田中哲男¹⁾ 岡崎彰夫¹⁾ 松浦 博²⁾ 近藤 洋²⁾

秀島雅之³⁾ 和田淳一郎³⁾ 安藤智宏³⁾ 犬飼周佑³⁾ 五十嵐順正³⁾

要旨: 発音を学習する聴覚障害者にとって、自分の発音についてのフィードバック情報を得ることは重要であるが、従来はこのフィードバック情報は、教師としての人間を介してしか与えられなかった。音声認識手法によって聴覚障害者の発音を音声セグメント系列に変換する実験を行い、得られた系列が人間による聞き取り内容を反映していることを確認した。この音声セグメントラベルを新しい視覚的フィードバック情報として学習者に提供できると考えられる。

キーワード: 音声認識、発音学習、視覚化

1. はじめに

人は聞くことを通して発音や言語表現を学んでいくと考えられている [1]。しかし早期から聴力に障害を持つ者は発音能力の獲得においても不利な立場に置かれている。このため聴覚障害児の発音の解析 [2]、発音の指導法 [1] について多くの研究がなされ、発音指導を支援する機器の研究も古くからなされている [3]。

これらの研究では、学習者の発音を聞き取り、または機器の出力を読み取って、学習者にフィードバック情報を与えるのは、教師（母親の場合もある）としての人間の役割である。

一方、松浦 [4] らは発声された音声を解析して統合音声ラベルを抽出するプログラム VoiceAnalyzer を開発し、義歯装着者の発音を解析することにより、補綴治療効果の評価 [5] に役立てている。本稿は、VoiceAnalyzer が作り出した音声セグメントラベル列が発話者の発音特性を十分に表しているか否かの検討を行うことを目的とする。

まず、聴覚障害者が発声した音声サンプルを検査者が聴取し、どのように聞き取ったかを記録した。次に同じ音声を VoiceAnalyzer が解析して音声セグメントラベル列に変換し、この記号列が検査者による聞き取りと同じ特性を有しているか否かを調べた。検査者による聞き取りと同じ特性が音声セグメントラベル列に表現されていれば、音声セグメントラベルをフィードバック情報として利用することの可能性が確かめられたことになる。

2. 実験

4人の発話者の協力を得て、それぞれ100語ずつの発音をしてもらい、4人の検査者による聞き取り結果を記録した。

実験に用いた単語は板橋 [3] が「単語水準の発音技能の評価に関する研究」のために作成した単語リストを用いた。板橋の単語リストは5語からなる群が17群、4語からなる群が3群で計97単語が用意されている。4単語しかない3群について報告者がそれぞれの群に1単語ずつ追加し100単語とした。

これらの単語を発話者ごとに単語の順をランダムに変えて提示し1回ずつ発音してもらった。VoiceAnalyzer の録音機能を用いて各発話者の発音を記録した。発音に際しては、

すいどう 水道

のように、平仮名表記と漢字表記を並置したリストを発話者に提示した。

この録音音声を4人の検査者が聴取しどのような音声として聞き取ったかを記録した。この表記法は検査者間では統一していない。通常の日本語単語として聞き取った時にはその単語が記録され、すこし異なる発声として聞こえた時には、検査者が感じたままの表現で記録されている。また併せて聞き取った音声の明瞭度を5段階で評価した。(5:明瞭である、4:正確に聞き取れる、3:聞き取れる、2:聞き取りづらい、1:聞き取れない、で評価)の5段階である。この評価もあくまで検査者個人の主観が大きい。

検査者4人のうち1人は単語リストを作成した者であり、ある程度どのような単語が発せられたかの記憶を持っている。他の3人はどのような単語が発声されたかの予備知識をまったく持っていない。また検査者のうち2人は日常発話者と接触した経験を持った者であり、他の2人は全く面識を持っておらず、発音を聞いたこともない。

表1 実験に用いた単語のリスト
発声者にはこのリストに示した平仮名表記と漢字表記とを示した。

群	番号	単語	群	番号	単語	
A	1	かご	籠	6	いま	今
	2	かこ	過去	7	いま	居間
	3	かっこ	「」	8	しば	芝
	4	かっこう	滑降	9	ひま	暇
	5	かこう	加工	10	しま	島
C	11	ふび	不備	16	ゆり	百合
	12	うみ	海	17	うり	瓜
	13	ゆみ	弓	18	くり	栗
	14	ゆび	指	19	つり	釣り
	15	ふみ	文	20	むり	無理
E	21	しみ	染み	26	いし	石
	22	しゅみ	趣味	27	いっち	一致
	23	しゅび	守備	28	いち	一
	24	すみ	墨	29	しち	七
	25	すみ	炭	30	しっち	湿地
G	31	そり	ソリ	36	ゆうかん	夕刊
	32	おり	檻	37	しゅうかん	習慣
	33	とり	鳥	38	ちゅうかん	中間
	34	そん	損	39	しんかん	瞬間
	35	ほん	本	40	しんかん	新刊
I	41	ようか	八日	46	たんご	単語
	42	しょうか	消火	47	たんご	団子
	43	しょうが	生姜	48	たんぼ	田圃
	44	ようが	洋画	49	たらこ	タラコ
	45	よっか	四日	50	たっこ	抱っこ
K	51	すいどう	水道	56	さんみやく	山脈
	52	すいそう	水槽	57	さんびやく	三百
	53	すいとう	水筒	58	さんちやく	三着
	54	ふうとう	封筒	59	さんぎやく	三脚
	55	ふいとう	沸騰	60	さんじやく	三尺
M	61	ざる	猿	66	すな	砂
	62	ざる	笹	67	つな	綱
	63	はる	春	68	ふた	蓋
	64	はる	貼る	69	うた	歌
	65	たる	樽	70	ふな	鮪
O	71	ばす	バス	76	げき	劇
	72	ばす	バス	77	せき	席
	73	ます	升	78	せき	咳
	74	ばつ	罰	79	えき	駅
	75	まつ	松	80	てき	敵
Q	81	うで	腕	86	けた	下駄
	82	ふで	筆	87	へた	下手
	83	ふね	船	88	えだ	枝
	84	うね	歌	89	けた	桁
	85	ふえ	笛	90	へら	へら
S	91	ばっじ	バッジ	96	はり	針
	92	まっち	マッチ	97	はり	梁
	93	まーち	マーチ	98	あり	鱈
	94	ばんち	パンチ	99	やり	槍
	95	ばんち	番地	100	じゃり	砂利

程度明瞭に聞き取れたかを5段階評価であらわした。表には評価者4人の明瞭度の合計を示した。4人全員が「明瞭である」と評価した場合に、明瞭度計は20である。

表2 聞き取り結果の例(群P)

課題単語がどのように評価者に聞き取られたかを示す。

話者	課題単語	評価者A	評価者B	評価者C	評価者D	明瞭度計	
1	76	げき 劇	どう	でっ	で	でっ?	5
	77	せき 席	たてい	ててい	?	でき	9
	78	せき 咳	せっま	せっ	?	せっ?	6
	79	えき 駅	えち	えてい	えき	えき?	7
	80	てき 敵	ててい	ててい	?	たてい??	8
2	76	げき 劇	ぐえか	げっか	ぎたっ	げっかあ	4
	77	せき 席	せか	せか	せけ	せ???	5
	78	せき 咳	せく	せーか	せいけい	せーか	3
	79	えき 駅	えっく	えく	えく	えっく?	5
	80	てき 敵	どっ(か)	どあか	でく	てっく??	4
3	76	げき 劇	げき	けーき	げき	けーき	15
	77	せき 席	げき	てき	へき	てき	10
	78	せき 咳	げき	げき	げき	げき	9
	79	えき 駅	いき	えき	えき	えき	18
	80	てき 敵	でき(げき)	てき	てき	げき	14
4	76	げき 劇	げき	げき	げき	げき	17
	77	せき 席	せき	せき	せき	せき	19
	78	せき 咳	せき	せき	せき	せき	19
	79	えき 駅	えき	えき	えき	えき	19
	80	てき 敵	てき	てき	てき	てき	19

表3 聞き取りの一致数

評価者の聞き取り結果と課題単語とが一致していた発声の数

発話者	評価者ごとの一致数				平均
	A	B	C	D	
1	40	50	67	53	53
2	21	21	29	6	19
3	41	61	48	57	52
4	92	85	91	86	89

聞き取り評価結果は、発話者に大きく依存する。表3に、「一致数」すなわち課題単語どおりに評価者が聞き取った単語数を示す。発話者4の発声の9割が聞き取られているのに対して、発話者2の場合にはそれが2割である。発話者1と発話者3は発声のうち約半数が聞き取られている。

聞き取り結果は評価者によっても異なるが大きな差異はないと考えられる。

表4には、すべての単語群について群ごとの「一致数」を示す。実験に参加した発話者の人数が4と少数であるので、この表から一般的に「苦手な発音」を結論付けることは避けるべきである。個人ごとの発音の特性が表れているといえる。

2.1 単語リスト

用いた単語の一覧を表1に示す。この単語リストは [3] に拠るが、D群20「むり無理」、E群25「すみ炭」、H群40「しんかん新刊」の3単語はこの実験で追加した。

2.2 聞き取り評価

群Pの単語を例として4人の発話者ごとの聞き取り評価の結果を表2に示す。それぞれの欄は評価者がどのような音として聞き取ったかを表している。また評価者はどの

表4 群ごとの一致数
課題と同じ単語として聞き取られた発声の数

群	発話者			
	1	2	3	4
A	11	6	19	19
B	17	3	16	18
C	18	4	19	18
D	17	1	8	20
E	7	7	8	16
F	6	2	11	18
G	11	4	4	16
H	16	5	13	19
I	16	1	12	19
J	15	6	15	14
K	8	0	6	20
L	9	5	5	20
M	7	2	2	10
N	8	11	14	15
O	1	2	2	19
P	1	0	7	20
Q	14	4	13	19
R	8	7	11	20
S	5	2	12	15
T	16	5	10	19

それでも若干の共通性が見られる群を拾い上げると、群O、群M、群Pが目立つ。

群Oは、発話者1、2、3の3者が共通に苦手としている。この群に属する単語はばす（パス）、ばす（バス）、ます（升）、ばつ（罰）、まつ（松）である。発話者1と発話者3は共通に /s/ および /ts/ の発音を苦手としている。発話者2はさらに /p/、/b/、/m/ の発音も苦手としているようである。

群Mは、発話者2、発話者3が苦手としているほか、全体的に高いスコアである発話者4もこの群Mのスコアは低い。発話者1もかなり低いスコアとなっている。この群に属する単語は、さる（猿）、ざる（筥）、はる（春）、はる（貼る）、たる（樽）である。発話者2、発話者3両方とも /ru/ を /gu/ と発音しているようである。発話者4は /ru/ が /du/ と発音しているようである。発話者1はこの群の単語に対して /sa/ を /ha/、/za/ を /ta/ または /da/、/ha/ を /a/ と発音している。

群Pは、発話者1と発話者2が苦手としている。この群に属する単語は、げき（劇）、せき（席）、せき（咳）、えき（駅）、てき（敵）である。発話者1の /ki/ の音はほ

とんど聞き取れず、語頭の /g/ が /d/、/s/ が /t/ と発音されているようである。発話者2については、第1音節は聞き取られているが、第2音節では /ki/ の /k/ は発せられているもののそれに続く /i/ が聞き取れず、/a/ や /u/ に変化しているように思われる。

3. 指標としての音声セグメントラベル

この節では VoiceAnalyzer が作り出した音声セグメントラベル列を検討する。健聴者である評価者が聞き取った発声の特徴が、音声セグメントラベル列の中に対応してあらわれているかどうかを検証することが目的である。

下に音声セグメントラベル列の例を示す。この例はある健聴者の「じゃり砂利」の発音を解析した結果である。音声セグメントは発音を 8 ms ごとに区切って213種類に分類したものであり、空白記号で区切られた2文字で一つの音声セグメントを表すラベル [4] とする。

```
QQ QQ QQ [ZZ ZZ ZZ ZZ JI CC CI CI] IA
IA IA IA IA YA IA IA IA EE EE AE AI
ER ER ER ER ED [RR RY RI RI RI RI RI]
II II IQ IQ IQ IQ IQ IQ IH IQ QQ QQ
```

音声セグメントには持続性セグメント、子音セグメント、境界セグメントの3種類がある。持続性セグメントは母音、撥音、あるいは子音が持続したセグメントを表すもので、たとえばラベルとして II（母音「い」）、RR（ら行の子音）、##（「ひ」の子音）などがある。QQは無音部を表す。子音セグメントは子音から母音への渡りを含む部分を表現するもの「さ」はSA、「ち」はCIなどと表す。境界セグメントは、母音から別の母音や子音の渡りを表すセグメントである。ラベルIAは母音「い」から「あ」への渡りを表し、IQは母音「い」から無音への渡りを表している。

中括弧 [] で囲まれた音声セグメントラベル列は自動的に抽出した子音領域である。子音領域とは子音および母音への渡りを示し、「じゃり砂利」の場合は通常 /ja/ と /ri/ を表す2か所が抽出される。

上の例に現れるセグメントラベルを簡単に説明すると、無音部を表すQQに続いて語頭の [ZZ ZZ ZZ ZZ JI CC CI CI] が「じゃ」の子音に相当し、母音および次の子音「り」への渡りIA、EE、AE、AI、ERなどを経て子音領域「り」[RR RY RI RI RI RI RI]に続いている。その後は「い」から無音への境界であるので、II IQ、に続いて無音部QQが最後に現れる。

3.1 正しく聞きとれた発音

まず4人の評価者が明瞭に正しく聞き取ることができた

単語の発音が、音声ラベル列としてどのように解析されたかを検討する。

「じゃり」

発話者1の「じゃり(砂利)」(T群単語番号100)は、4人の明瞭度合計は19で全員が「じゃり」と聞き取っている。この発声に対する音声セグメントラベル列は

```
[TA BZ QQ CC ZZ KI KY JY JY JY JY CY
YO YA YA] IA OO AA AA AA OZ AR AR OZ
[RR RI RI RI RI] II II IG II II US
US UZ II II II IN UN IQ IQ IQ
```

である。評価者の聞き取りと音声セグメントラベルがほぼ一致している。最初に挙げた健聴者の「じゃり」との微妙な違いは興味深い、評価者はどちらも「じゃり」の発音であると認識する。これは、日本語単語についての言語知識が影響しており、多少の音響的な相違は言語知識によって補われていると考えられる。

「ふな」

発話者2の「ふな(鮒)」(N群単語番号70)は4人の明瞭度合計は10であるものの評価者全員が「ふな」と聞き取っている。この発声に対する音声セグメントラベル列は、

```
[KU QQ CC SS SS QQ SS I4 I4 I4 I4 I4 I4
I4 QQ SS QQ SS PU FF SS ZU TI] U4 U4 U4
QQ U4 U4 I4 FF U4 U4 IQ [CC QQ ZZ ZZ ZU
ZU QU QU QU] UZ UZ UZ UZ UG GO GO GO
UU UU UU [WA WA WA GA] OS OI OI OH OH
ON OH OH OH [HY ## ## US ## FF KY N9 NA
NA NA NA NA] AA AS AS AS AS AD AD DD
SS SS QQ U4 U4 U4 U4 HI KI UB UB IQ IQ
IQ XX QQ QQ M9 XX XX XH QQ QQ QQ QQ
```

である。1番目の子音領域は声のパワーが小さく息継ぎであり単語発話の部分ではない。2番目の子音領域に注目すると[CC QQ ZZ ZZ ZU ZU QU QU QU]であり、有声であるが摩擦性の音が出ている。4番目の子音領域[HY ## ## US ## FF KY N9 NA NA NA NA NA]でNAがしっかりと出ていることから評価者の耳に「ふな」と聞こえたと思われる。なお、3番目の子音領域[WA WA WA GA]は母音の領域内の変化が誤って子音領域として検出されたものである。

「ふみ」

発話者3の「ふみ(文)」(C群単語番号15)は、4人の評価者の明瞭度合計は19である。この発声に対する音声セグメントラベル列は、

```
[SS QQ SS SS SS SS SS SS SS SS ZO ZO
```

```
ZU GU ZU ZU ZU] UU UU UU UU UU UG UG
UG XX OG [N9 N9 N9 M9 M9 NY MI MI MI
MI MI] II II II II II II IQ
```

である。ここでは最初の子音領域は[SS QQ SS SS SS SS SS SS SS ZO ZO ZU GU ZU ZU ZU]となっており、「ふ」の/f/が捉えられていないように思えるが摩擦性の/s/が検出されている。また、「み」の/m/は2番目の子音領域で[N9 N9 N9 M9 M9 NY MI MI MI MI MI]でM9→MIとして検出されている。なお、M9はマ行の子音定常部を、N9はナ行の子音定常部を表している。

「さんみゃく」

発話者4の「さんみゃく(山脈)」(L群単語番号56)は、4人の評価者の明瞭度合計は19である。この発声に対する音声セグメントラベル列は、

```
[BZ QQ CC QQ QQ CC SS SS SS SS $$ $$
$$ SS CC SO SA SY HH HH HH HH HH HH]
IQ IQ IQ AX ON OM AX OX OX XX QQ XX
UZ US [ZU ZU XX M9 RI ## ## ## MY MY
YA YA YA YA] AS AA AD AD AG UQ UQ UU
QQ QQ QQ QQ QQ [CC CC ZZ KU JI JI JI
GU] UU UI II IG IG IG IG IG IG UU IQ
IQ IQ IQ IQ
```

である。ここで、\$\$は「し」の子音定常部、XXは「ん」で、AX、OXは「あ」から「ん」、「お」から「ん」への境界である。最初の子音領域はSS→\$\$→SS→SA→AXで、続く母音領域はOX、XXとなっている。2番目の子音領域はMY→YAとなり、続く母音領域でAA、最後の子音領域ではKU→UUと変化している。「ん」はAX、OX、XXで検出されていて、「みゃく」の/mya/はMY、YAで検出され、/k/はKUとしてわずかに検出されている。また、最後の母音/u/は濁っているため、一部GU、IGなどとして検出されている。なお、2番目の子音領域の##は「ひ」の子音定常部であるが、「ん」から「みゃ」への遷移の途中の発声で現れてしまったものと思われる。

セグメント系列からはこの発音は「さんみゃく」と言っているようにも解釈できそうだが、「さんみゃ」までを聞き取った人間にとっては最後の音節は「く」しかあり得ないようである。

この4つの例では、単語を構成する主要な音が音声セグメントラベル列の中にある程度明瞭にとらえられており、評価者の聞き取りと一致していると考えられる。

3.2 聞き取れなかった発音

つぎに評価者が聞き取ることができなかった発音をとりあげる。

「せき」 その1

発話者1の「せき(席)」(P群単語番号77)を評価者は「たてい」「ててい」「?」「てき」と聞きとっている。明瞭度合計は9である。この発声に対する音声セグメントラベル列は、

```
[CC KY KE TO DE DE QE] IR EQ EQ EQ
EQ UQ UQ OO QQ QQ QQ QQ AA [CC ZZ
KY DU TI TI DU] II II UI UI UI EY II
II IQ IQ IN IM IQ QQ M9 UG UG IQ [SS
SS OY SS SS] UU UX UX QQ [QY OY QQ
HY KY] UN IQ EE IQ QQ QQ [BZ BZ QQ
SS SS SS SS] OO I4 QQ QQ QQ QQ
```

である。5個の子音領域があるが3つ目以降は息つぎによって生じたもので、パワーが小さく無視できる。最初の子音領域 [CC KY KE TO DE DE QE]、2番目の子音領域 [CC ZZ KY DU TI TI DU] に「せ」や「き」に相当する音が無く、それぞれ、TO、DE および TI が出ており評価者の印象を裏付けている。

「せき」 その2

発話者3の「せき(席)」(P群単語番号77)の明瞭度評価は10である。評価者はこの発声を「けき」「てき」「へき」「てき」と聞きとっている。この発声に対する音声セグメントラベル列は

```
SS QQ QQ QQ QQ QQ [CC ZZ KY KY KY KE
KE GE KE II II GI] II II EQ EQ IQ IQ
IQ IG IQ IQ QQ QQ QQ QQ QQ QQ QQ
[CC CC KY KY KI KI KI $$ KI KI KI
GI] II II IG IX IX IQ IG IG IQ IG IG
BZ BZ EX QQ QQ
```

である。最初の子音領域 [CC ZZ KY KY KY KE KE GE KE II II GI] では KY、KE が存在している。これが「け」「て」「へ」として聞こえた理由であろう。第2音節の「き」は2番目の子音領域にかなり明瞭に KI KI KI KI として存在していて評価者にもはっきり捉えられている。

「くり」

発話者2の「くり(栗)」(D群単語番号18)の明瞭度合計は4である。評価者はこの発声を「こり」「ごり」「きこり」「ごり」と聞きとっている。この発声に対する音声セグメントラベル列は

```
QQ QQ [KA KU KU KU PU KU] UU AA AA
XX XX XX XX OA OA RA QQ AH AH AA AA
AA AA AA AA AR AR AY AY AY AY OI OI
AY OE OE AY AY AY EX EX I4 EE EG EQ
EQ EQ QQ QQ
```

である。

最初の子音領域で [KA KU KU KU PU KU] と /ku/ がはっきり存在しているものの、そのうしろに OA、AA などの母音が存在しているために、「き」と聞き取れずに、「こ」「ご」と聞こえたのではないかと思われる。「り」に関する子音領域は RR、RI が出力されてないため、全く抽出されていないが、/r/ や /i/ は、AR、OI としてかろうじて検出されている。

「ざる」

発話者4の「ざる(筈)」(M群単語番号62)の明瞭度合計は11である。評価者はこの発声を「ざる」または「ざどう」と聞きとっている。この発声に対する音声セグメントラベルは

```
QQ QQ [CC BZ CC TO ZO ZO TO KA ZO
ZO] OE OO OO OO OE OE AU AU AS OO OD
OQ OQ OZ [ZZ DD GY DU JI CU ZU ZU]
UI UU UU UR UR II UQ IG IQ IQ UQ IQ
IQ QQ
```

である。最初の子音領域 [CC BZ CC TO ZO ZO TO KA ZO ZO] では ZO が出ており、次の母音では OO や AU が出ているので、きれいな「あ」ではないか全体では「ざ」と聞こえる。次の子音領域では [ZZ DD GY DU JI CU ZU ZU] と「る」の発音が無く、DU が存在していることがわかる。これが「どう」と聞こえた理由であろう。

評価者が課題単語通りには聞き取れなかった単語については、評価者の聞き取りを裏付けるような音声セグメントラベル列が得られていることがわかった。

評価者が正しく聞き取った単語、聞き取れなかった単語の音声セグメントラベル列を検討した。いずれの場合についても、評価者の聞き取り結果を裏付ける音声セグメントが捉えられていると考える。

4. 今後の課題

聴覚障害者が発声した音声信号から VoiceAnalyzer が抽出した音声セグメント列が示す特性について検討した。聴覚障害者の発声は個人差が大きいものの、それぞれの特徴をあらわす音声セグメントラベル列が得られているものと考ええる。

この音声セグメントラベル列を利用して、発音に対するフィードバック情報を与えるには、解決すべき課題がいくつかある。

第一の課題は VoiceAnalyzer の認識精度の問題である。これについては、松浦ら [4] が7名の健聴者の54音節発声

データの子音領域を用いて評価を行っている。適正なラベルの出現率の高い音節と低い音節とがあり、被験語、被験音を設定する際に、その点を配慮して選択することが重要なことを述べている。本研究のポイントは聴覚障害者の苦手な発音を指導するための新たなフィードバック情報が音声セグメントラベルから得られるかどうかにある。模範的な発音においても音声セグメントラベルの出力の揺れはありと考えられるが、その場合は許容範囲とする必要がある。

第二の課題は模範的な発声をどのようにあらわすかという問題である。アナウンサーなどの発音をあらかじめ解析して教師音声セグメントとする方法、学習者の家族の発音を模範とする方法、発音辞典などから作成する方法などが考えられる。

第三の課題は、模範的な音声セグメントラベル列と学習者の音声セグメントラベル列との違いの検出方法と違いの提示方法である。単に両者のセグメントラベルを提示するだけでも学習者にとって情報を与えることにはなるが、適切に違う場所を指摘し、適切な方法で提示することが必要である。音声セグメントに対応した声道断面図 [6] のアニメーションを提示することも一つの方法である。

また、健聴者でも起こる、言語知識によって補完される発音の揺れと Voice Analyzer の精度との関係を多くの

健聴者および聴覚障害者の発話データを用いて確認し、その結果を上述の提示方法に織り込む必要がある。

これらの課題に対処したうえで、適切なフィードバック情報を発話者に与える方法を検討していく。

参考文献

- [1] 板橋安人：聴覚障害児の「発音・発語」学習，聾教育研究会，2006.
- [2] 板橋安人：聴覚障害児の発音技能の形成に関する研究，風間書房，1999.
- [3] 佐久間麗子，間根山祥行：発音・発語学習の基礎と実践（基礎編）第3章 発音訓練器，聴覚障害教育工学，Vol.22, No.2, pp75-110, 1999
- [4] 松浦 他：チェアサイドで使用可能な発語評価のための音声認識の開発，情報処理学会論文誌，Vol.46, No.5, May, 2005
- [5] Inukai, Hideshima et. al.: Analysis of the Relationship between the Incisal Overjet in a Maxillary Denture and Phonetic Function Using a Speech Recognition System, Prosthodontic Research & Practice, 5(3), July, 2006
- [6] 磯村一弘：音声を教える，日本語教授シリーズ第2巻，国際交流基金，2009

A Preliminary Study on the Visualization of Pronunciation for Hearing-Impaired Persons

TANAKA Akio¹⁾, OKAZAKI Akio¹⁾, MATSUURA Hiroshi²⁾, KONDO Hiroshi²⁾,
HIDESHIMA Masayuki³⁾, WADA Junichiro³⁾, ANDO Tomohiro³⁾, INUKAI Shusuke³⁾, IGARASHI Yoshimasa³⁾

¹⁾ Faculty of Industrial Technology, Tsukuba University of Technology

²⁾ School of Administration and Informatics, University of Shizuoka

³⁾ Removable Prosthodontics, Department of Masticatory Function Rehabilitation, Division of Oral Health
Sciences, Graduate School, Tokyo Medical and Dental University

Abstract: It is very important for pronunciation learners to obtain adequate feedback information on their speech. For hearing-impaired persons, however, feedback until now has been limited to that obtained from their teachers. Our phonetic segment engine transforms uttered speech to a sequence of phonetic segment labels for every 8 ms frame shift. We applied the engine to the pronunciation of hearing-impaired persons and concluded that the sequence of phonetic segment labels reflects the features of both the correct and incorrect pronunciation of the speakers. This report describes the experiment and discusses how the sequence of phonetic segment labels shows the features of the pronunciations. We believe that the phonetic segment labels can be a source for visualizing the pronunciation of hearing-impaired persons.

Keywords: Phonetic recognition, Learning pronunciation, Feedback, Visualization