

聞くことに困難のある人がわかりやすい音声： 視覚刺激の付加により高次脳機能障がい者の理解は進むか

三谷 雅 純^{1)*}

Do added visual-stimuli aid in understanding persons with cognitive dysfunction who have difficulty hearing?

Masazumi MITANI^{1)*}

Abstract

A set of hearing experiments was conducted on 33 persons with cognitive dysfunction, as well as on 46 non-impaired persons, in order to search for an effective voice system that is suitable for broadcasting in ideal lifelong learning facilities. Digitally-recorded real voices by women were tested, as well as formant synthesis speech and concatenative synthesis speech. Non-impaired persons responded positive to the experiments in real-voices and in concatenative synthesis speech. The same persons responded negative in formant synthesis speech without visual signals, but responded positive if there were visual signals. The majority of the tested persons with non-serious damage responded positive to questions from all of the real-voices, formant and concatenative synthesis speech. The tested persons with serious damage were conspicuous in responding to the same questions, especially for the formant synthesis speech. These experiments indicate that real human voices are the easiest to understand in lifelong learning facilities in which disabled individuals use them frequently. Concatenative synthesis is desirable when using a synthetic system, but the staffs of the facilities should be careful when using formant synthesis, although it is more convenient than concatenative synthesis.

Key words: artificially-synthesized voice, cognitive, communication-disorder, dysfunction, language, lifelong learning facilities, natural voice.

はじめに

博物館や美術館、図書館、公民館など、生涯学習施設では施設内で流れる放送が重要な役割を担うことがある。日常的に放送が利用されるのは録音による展示解説やスケジュール案内などだが、高齢者や幼児、障がい者などを含む一般利用者を対象とする公共放送では、年齢

や難聴の有無などの肉体的条件と音響環境によって聞こえ方が大きく異なることがある。さらにこうした日常の利用以外にも、災害の発生を知らせ、避難をうながす緊急放送では格段の配慮が必要である。火事や地震などの災害時には「人によって放送が理解できない」ということがあってはならない。しかし、通常の放送は難聴者やろう者には聞き取りが難しく、また失語症者を含む高次脳

¹⁾ 兵庫県立大学 自然・環境科学研究所 〒669-1546 兵庫県三田市弥生が丘6丁目 Institute of Nature and Environmental Science, University of Hyogo; Yayoi-ga-oka 6, Sanda, Hyogo, 669-1546 Japan

併任：兵庫県立人と自然の博物館 コミュニケーション・デザイン研究ユニット／自然・環境マネジメント研究部 Study Unit of Communication Design/ Division of Nature and Environmental Management, Museum of Nature and Human Activities, Hyogo

* E-mail: mitani@hitohaku.jp

機能障がい者は言語音の認知が困難な場合がある(石合, 2012).

三谷(2013)はコミュニケーション障がい者と総称される人たちを再考し,その人たちに「読みやすい文章」や「表現方法」のより良いあり方を探った.この総説の中で「聞きやすさ/聞きにくさ」に関しては,聴覚障がいである難聴・ろうと,聴覚よりも聴覚刺激が脳に届いた後の高次脳機能の障がいである失語症は,困難の起こる機序の違いによって分けることが適当であることを示した.

同時に,三谷(2014)は1名の高次脳機能の障がい者を含む27名の男女に絵本や童謡を視聴してもらい,内容の理解しやすさを聴覚実験で調べた.この実験をさらに進め,一般来館者を高次脳機能障がい者とその他の人に分けて比較し,高次脳機能障がい者の視聴覚特性を抽出しようとするのが本研究である.

実験で比較したのは,人の肉声,フォルマント合成(formant synthesis)音声,波形接続型合成(concatenative synthesis)音声である.ここで言うフォルマント合成とは音声に似せた電氣的な信号を人工的に合成する技術で,あたかもロボットがしゃべっているように聞こえる.利用者にとってはもっとも簡便な音声合成システムで,テキスト文を入力するだけでしゃべらせたい合成音声が作成できる.一方,波形接続型合成では,モデルとなる人間が音や単語,成句などを吹き込み,それをデータベースにして,該当する文章にもっとも適した音の組み合わせを探して合成する.フォルマント合成よりも人間の肉声に近いものができると言われていた.ただし,データベースには一定以上の容量が必要である.肉声と波形接続型合成音声との聞きやすさの比較については,居村ほか(2008)を参照されたい.

なお「聞こえ」とは聴覚に関する概念であり,耳から神経を通して脳にとどく信号の伝達を言うが,本稿では「聞こえ」を拡大解釈して,おもに神経心理学的な聴覚認知(山鳥,1985)を論じる.聴覚と神経心理学的な聴覚認知は医学的には厳密に区別するべきだが,生涯学習施設のように多様な人を含む一般利用者を対象として考察する場合には,重複するところが大きいためである.

方 法

被験者と音源, 実験過程

2014年1月26日,および2014年10月から12月にかけて,高次脳機能障がい者を対象にした聴覚実験を行った.実験に際して協力を得た当事者団体を表1にまとめた.

実験には高次脳機能障がい当事者32名と,比較のための対照群として介助者と言語聴覚士合わせて20名(非

高次脳機能障がい者.以下,非障がい者と呼ぶ)が参加した.実験は,通常当事者団体が集会を行っている場所など,なるべく当事者が落ち着ける場所で行った.解析には2014年2月4日に行った実験(三谷,2014:障がい者1名,非障がい者26名)のデータも含めた.これによって障がい者33名,非障がい者46名のデータが解析できた.表2にすべての被験者の性・年齢を示した.

音源はデジタル図書『あなたのことが だーいすき』(ファン・ヘネヒテン,2006)を使用した.フォルマント合成の音声合成エンジンはドキュメントトーカ日本語音声合成エンジン(クリエートシステム開発株式会社)を使用した.音声には成人女声の人工音声「けいこ」を使った.波形接続型合成音声はAItalk(株式会社エーアイ)の女声「のぞみ」を使った.共に女声を使用したのは,一般利用者を想定した聴覚実験(三谷,2014)で,女性声が男性声よりもわかりやすいという評価を得たからである.表3に実験に使用した材料の音響特性を示した.合わせて他の質問内容も括弧に入れて示した.音声情報の一部はSUGI Speechanalyzerを使い,ソナグラムで視覚化した(図1).

マルチメディアDAISYは,音声とその部分のテキストや画像などが同時に出力されるデジタル図書で,さまざまな障がい者が利用しやすい書籍形態であるとされている(河村,2011).高次脳機能障がい者にも有効であることが期待されたため,この実験ではMicrosoft PowerPoint 2007を利用してマルチメディアDAISYと同様の形態の実験素材を作成した.具体的には被験者に録音した肉声と合成音声を聞いてもらい,同時に視覚情報としてスクリーン上に文章や絵を再生して,わかりやすさを主観的に答えてもらった.本論文では「聴覚実験」という用語を使っているが,厳密には視聴覚実験である.実験は複数の被験者に対して同時に行った.

実験過程は,スピーカから流れる音声の聞き取りとスクリーンに映る視覚情報がスムーズに被験者に伝わったかどうかを調べる予備的設問(デジタル録音した女性Aの肉声とフォルマント合成音声の女声)の後,『あなたのことが だーいすき』の一場面を朗読する(i)音声のみ,(ii)音声に挿絵による場面の視覚イメージを加え,漢字かな混じり文の文章で朗読するところが黄色のハイライトに変わるもの,(iii)音声と挿絵は同じだが,文章をひらがなで分かち書きし,朗読するところが黄色のハイライトに変わるものを聞いてもらった.『あなたのことが だーいすき』の音声は,①デジタル録音した女性(B)の肉声,②フォルマント合成音声の女声,③波形接続型合成音声の女声を使った.その後マイクで拡声した男性(三谷,60歳)の肉声が聞こえるかどうかを聞き,最後にこの実験では疲れたかどうかを聞いた.ス

表 1 聴覚実験の概略.

日付	協力団体	参加人数	会場
2014年			
1月26日	はなそう神戸の会	5	ことばの道研修室
10月15日	むつみ会	13	花と緑のまちづくりセンター会議室
11月 8日	若者と家族の会	10	人と自然の博物館会議室
12月13日	いなば会	24	尼崎市立障害者福祉会館
2月 4日	人と自然の博物館	27	人と自然の博物館セミナー室

表 2 被験者（高次脳機能障がい当事者と非高次脳機能障がい者）の性・年齢.

	(a) 軽度障がい者	(b) 中・重度障がい者	(c) 非障がい者
性			
男	9	13	5
女	7	4	41
年齢			
平均 ± SD	64.56 ± 10.97	60.12 ± 17.95	51.07 ± 14.56
最年少 - 最年長	27-74	21-83	24-81

表 3 設問と音声タイトル, 話者, 音声の長さ, 声の高低に対応するピッチ周波数. 音声タイトルで読み上げたのは, 出だしの一節のみである. 女性の肉声は, 職業アナウンサー A, ボランティア・アナウンサー B の 2 人がいる.

設問	音声タイトル	話者	長さ (sec)	ピッチ周波数 (Hz)
1	くんくくん おいしそう	肉声 (女性A)	20.33	165 ~ 393
2	くんくくん おいしそう	人工合成音声 (女声)	29.27	164 ~ 464
3	あなたのことが だーいすき (音声のみ)			
4	あなたのことが だーいすき (音声+漢字かな混じり文と絵)	肉声 (女性B)	29.09	160 ~ 558
5	あなたのことが だーいすき (音声+ひらがなと絵)			
6	(上のみっつの内で, いちばん聞きやすかったのはどれですか?)			
7	あなたのことが だーいすき (音声のみ)			
8	あなたのことが だーいすき (音声+漢字かな混じり文と絵)	フォルマント合成音声 (女声)	24.75	148 ~ 428
9	あなたのことが だーいすき (音声+ひらがなと絵)			
10	(上のみっつの内で, いちばん聞きやすかったのはどれですか?)			
11	あなたのことが だーいすき (音声のみ)			
12	あなたのことが だーいすき (音声+漢字かな混じり文と絵)	波形接続型合成音声 (女声)	17.32	172 ~ 347
13	あなたのことが だーいすき (音声+ひらがなと絵)			
14	(上のみっつの内で, いちばん聞きやすかったのはどれですか?)			
15	(三谷が言っていることは, わかりますか?)			
16	(この実験は, 疲れましたか?)			
17	(この実験に対するご意見などがありましたら, 何でもご自由にお書き下さい。)			

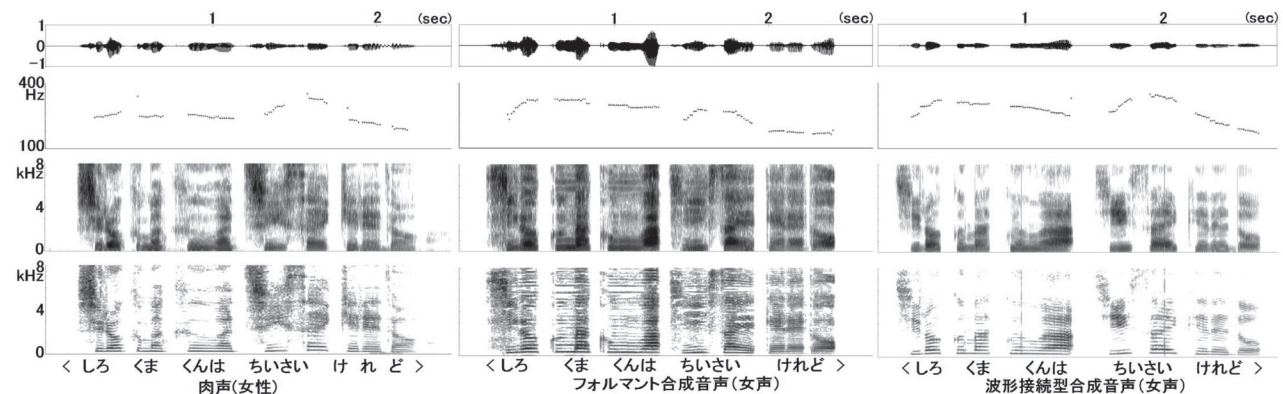


図 1 ソナグラムで表した音声. 〈くしろくまくんは ちいさい けれど〉というフレーズの, それぞれ肉声 (女性 B), フォルマント合成音声 (女声), 波形接続型合成肉声 (女声) を, 左から右に順に示した. それぞれのパターンは, 上から原波形, ピッチ, フォルマントがよく表れる広域帯スペクトログラム, 基底振動数が読み取りやすい狭域帯スペクトログラムを示す. スペクトログラムの帯域は, 人間の可聴域に合わせた 0 kHz から 8 kHz である.

ピーカはパワーアンプ内蔵のメディア・メイト (BOSE社) を使用した。

実験を始める前に、聞く場面では「1」ときどき言葉に詰まるが問題は感じない、「2」困難を感じる、あるいは、まったく認識できないに分けて、あらかじめ自分がどちらに当たるかを主観的に判断して答えてもらった。同様に話す場面についても、あらかじめ「1」「2」に分けて答えてもらった。スピーカ・チェックでは、聞こえるかどうかを「1」問題なしと「2」問題ありとして答えてもらった。これは被験者全員にスピーカの音が問題なく聞こえているかどうかを確かめるためである。位置によってスピーカの音が聞こえない／聞こえにくい被験者には席を替わってもらった。ただし、いずれの実験でも席を変りたいという希望はなかった。音圧は調整して音が割れないように注意した。

聞こえの理解を調べる設問では、音声聞きやすいか聞きにくいかを主観で判断し、①肉声、②フォルマント合成音声、③波形接続型合成音声のそれぞれについて、「1」わかる、「2」少しわかる、「3」わからないで答えてもらった。

回答はあらかじめ用意した回答用紙を配り、被験者には結果をそのつど記入してもらい、最後に回答用紙を回収した。実験に当たっては被験者に過度の負担を与えないように気を付けた。

実験は生涯学習施設の館内放送や緊急放送を模して行ったため、音声を聴くときの右耳・左耳 (左脳・右脳) の差異などは考慮しなかった。言語中枢は左脳に存在する人が多く右耳が優位であることが多いが、館内放送や緊急放送では、通常、左右を区別して入力することは少ないからである。

なお、以後被験者を (a) 軽度障がい者、(b) 中・重度障がい者、(c) 非障がい者に分けて結果を示し、煩雑さを避けるために、結果や議論ではそれぞれを「被験者グループ」あるいは単に「グループ」と呼ぶことがある。また必要に応じて (a) 軽度障がい者、(b) 中・重度障がい者、(c) 非障がい者を (a), (b), (c) と略す。

表4 高次脳機能障がいのある被験者が日常感じている「聞く側面」と「話す側面」の困難の程度を申告してもらった結果。実験に参加した高次脳機能障がいのある被験者は全体で33名いたが、回答のない人がいたために合計は合わない。

		聞く側面		話す側面	
問題が	少ない	16	少ない	11	
			多い	2	
	多い	17	少ない	8	
			多い	7	

結 果

実験の予備的な設問

実験時に聞いたアンケートから被験者となった高次脳機能障がい者の「聞く側面」と「話す側面」の困難を確かめておくと、「聞く側面に問題は少ない」と答えた被験者では、同時に「話す側面にも問題が少ない」と答えた人が多かった。しかし「聞く側面に問題がある」と答えた被験者は、「話す側面に問題がある」人と「問題が少ない」人が同程度いた (表4)。この内、今回の実験では「聞く側面」だけを問題にする。

マイク・チェックのための予備的な設問は、アナウンサー (女性A) とフォルマント合成音声の女声で、それぞれ阿部 (1994) の『くんくんくん おいしそう』を朗読する声を聞いてもらい行った (表3の設問1, 2。以下同様に表3を参照)。

肉声では被験者のグループ間で聞きやすさに違いが認められた (Kruskal-Wallis test, $p=.004$)。表5にBonferroniの修正による多重検定でペアごとに比較した結果を示す。非障がい者は中・重度の障がい者よりも有意に、「わかる」と答えた被験者が「困難を感じる」あるいは「まったく認識できない」と答えた被験者よりも多かった ($p=.011$)。フォルマント合成音声でも違いがあり (Kruskal-Wallis test, $p=.044$)、多重検定では軽度障がい者は中・重度の障がい者よりも有意に「わかる」と答えた被験者が多かった ($p=.039$)。

音声の評価

音声の質①肉声、②フォルマント合成音声、③波形接続型合成音声ごとに、(i) 音声だけ、(ii) 音声と絵、漢字かな混じり文、(iii) 音声と絵、ひらがな文を聞いてもらったときの、被験者のグループ (a), (b), (c) で人数の分布の差を見てみると、ほとんどの場合、わかりやすさに差が認められた (表6)。わかりやすい／わかりにくいと判断した被験者に差が検出できなかったのは、②フォルマント合成音声と③波形接続型音声で、ともに (i) 音声だけの場合であった (表6)。

表6から導かれる多重比較の結果を表7に示す。(a) 軽度障がい者と (c) 非障がい者の比較では、①肉声、②フォルマント合成音声、③波形接続型合成音声のいずれにおいても有意な差は認められなかった。

表7の女性の肉声に注目すると、(i) 音声だけでは軽度障がい者—中・重度障がい者では差は認められなかったが、「わかる」と答えた非障がい者は、中・重度障がい者よりも有意に多かった ($p=.002$)。視覚情報の加味された (ii) 音声と絵、漢字かな混じり文や (iii) 音声と絵、ひらがな文では軽度障がい者—中・重度障がい者、非障がい者—中・重度障がい者のいずれにも顕著な差が

表 5 スピーカー・チェックのために回答してもらった音声の聞きやすさの程度の差。数値は Kruskal-Wallis test の後の Bonferroni の修正による多重比較で得られた調整済み有意確率を表す。

	(a)軽度 - (c)非障がい者	(a)軽度 - (b)中・重度	(c)非障がい者 - (b)中・重度
女性の肉声	n.s.	n.s.	.011
フォルマント合成音声(女声)	n.s.	.039	n.s.

表 6 『あなたのことが だーいすき』の朗読を条件を変えて聞いてもらったときの結果。音声のみと文章（漢字仮名混じり文、ひらがなのみ）や絵を挿入した時、困難は感じない、困難がある、音声として認識できないのどれに当たるかを回答した例に、聞く困難が軽度の人、中程度～重度の人、非障がい者の間で差があるかどうかを Kruskal-Wallis test で検討した。

		N	漸近有為確率	有意水準
①女性の肉声	(i) 音声のみ	79	.013	< .05
	(ii) 音声に漢字仮名混じり文で書いた文章と絵を添付	79	< .001	< .01
	(iii) 音声に平仮名で書いた文章と絵を添付	79	< .001	< .01
②フォルマント合成音声(女声)	(i) 音声のみ	79	.068	n.s.
	(ii) 音声に漢字仮名混じり文で書いた文章と絵を添付	79	< .001	< .01
	(iii) 音声に平仮名で書いた文章と絵を添付	79	.014	< .05
③波形接続型合成音声(女声)	(i) 音声のみ	47	.404	n.s.
	(ii) 音声に漢字仮名混じり文で書いた文章と絵を添付	47	.005	< .01
	(iii) 音声に平仮名で書いた文章と絵を添付	47	.046	< .05

表 7 『被験者を軽度障がい者、中・重度障がい者、非障がい者の3グループに分けて比較した結果。数値は Kruskal-Wallis test の後の Bonferroni の修正による多重比較で得られた調整済み有意確率を表す。*：フォルマント合成音声と波形接続型合成音声で音声のみの場合は、Kruskal-Wallis test の全体でグループ間の有意な差異を示さなかったため、多重比較は実行できなかった。

		(a)軽度 - (c)非障がい者	(a)軽度 - (b)中・重度	(c)非障がい者 - (b)中・重度
①女性の肉声	(i) 音声のみ	n.s.	n.s.	.002
	(ii) 音声に漢字仮名混じり文で書いた文章と絵を添付	n.s.	.002	< .001
	(iii) 音声に平仮名で書いた文章と絵を添付	n.s.	.004	< .001
②フォルマント合成音声(女声)	(i) 音声のみ*	-	-	-
	(ii) 音声に漢字仮名混じり文で書いた文章と絵を添付	n.s.	.004	.001
	(iii) 音声に平仮名で書いた文章と絵を添付	n.s.	.022	.035
③波形接続型合成音声(女声)	(i) 音声のみ*	-	-	-
	(ii) 音声に漢字仮名混じり文で書いた文章と絵を添付	n.s.	n.s.	.006
	(iii) 音声に平仮名で書いた文章と絵を添付	n.s.	n.s.	n.s.

表 8 被験者をグループに分けて質の違う音声を聞かせたとき、聴覚刺激だけや視覚刺激が付加されたときでは、聞きやすさがどれくらい変わるかを見た結果。

		N	わかる	わからないとき がある	わからない
(a) 軽度障がい者					
	(i) 音声のみ	16	13	2	1
①女性の肉声	(ii) 音声に漢字仮名混じり 文で書いた文章と絵を添付	16	16	0	0
	(iii) 音声に平仮名で書いた 文章と絵を添付	16	15	1	0
	(i) 音声のみ	16	13	2	1
②フォルマント合成音声 (女声)	(ii) 音声に漢字仮名混じり 文で書いた文章と絵を添付	16	15	1	0
	(iii) 音声に平仮名で書いた 文章と絵を添付	16	15	1	0
	(i) 音声のみ	14	12	2	0
③波形接続型合成音声 (女声)	(ii) 音声に漢字仮名混じり 文で書いた文章と絵を添付	14	13	0	1
	(iii) 音声に平仮名で書いた 文章と絵を添付	14	14	0	0
	(i) 音声のみ	17	11	3	3
(b) 中・重度障がい者					
	(i) 音声のみ	17	11	3	3
①女性の肉声	(ii) 音声に漢字仮名混じり 文で書いた文章と絵を添付	17	12	3	2
	(iii) 音声に平仮名で書いた 文章と絵を添付	17	10	3	4
	(i) 音声のみ	17	7	8	2
②フォルマント合成音声 (女声)	(ii) 音声に漢字仮名混じり 文で書いた文章と絵を添付	17	9	5	3
	(iii) 音声に平仮名で書いた 文章と絵を添付	17	10	4	3
	(i) 音声のみ	15	11	2	2
③波形接続型合成音声 (女声)	(ii) 音声に漢字仮名混じり 文で書いた文章と絵を添付	15	9	5	1
	(iii) 音声に平仮名で書いた 文章と絵を添付	15	11	3	1
	(i) 音声のみ	46	43	3	0
(c) 非障がい者					
	(i) 音声のみ	46	43	3	0
①女性の肉声	(ii) 音声に漢字仮名混じり 文で書いた文章と絵を添付	46	46	0	0
	(iii) 音声に平仮名で書いた 文章と絵を添付	46	45	1	0
	(i) 音声のみ	46	29	15	2
②フォルマント合成音声 (女声)	(ii) 音声に漢字仮名混じり 文で書いた文章と絵を添付	46	42	4	0
	(iii) 音声に平仮名で書いた 文章と絵を添付	46	39	7	0
	(i) 音声のみ	18	16	2	0
③波形接続型合成音声 (女声)	(ii) 音声に漢字仮名混じり 文で書いた文章と絵を添付	18	18	0	0
	(iii) 音声に平仮名で書いた 文章と絵を添付	18	17	1	0

認められ、「わかる」と答えた軽度障がい者や非障がい者が中・重度障がい者に比べて有意に多かった [(ii) (a) —(b) : $p=0.002$, (c) —(b) : $p=0.004$; (iii) (a) —(b) , $p<.001$, (c) —(b) : $p<.001$].

表7の②フォルマント合成音声に注目すると、(ii) 音声と絵、漢字かな混じり文では軽度障がい者—中・重度障がい者、非障がい者—中・重度障がい者の間に顕著な差が表れ、いずれも「わかる」と答えた軽度障がい者、非障がい者が中・重度障がい者に比べて有意に多かった [(a) —(b) : $p=0.004$, (c) —(b) : $p=0.001$]. (iii) 音声と絵、ひらがな文でも、いずれも「わかる」と答えた軽度障がい者、非障がい者が中・重度障がい者に比べて有意に多かった [(a) —(b) : $p=0.022$, (c) —(b) : $p=0.035$]. ところが同じ合成音声であっても、③波形接続型合成音声では (ii) 音声と絵、漢字かな混じり文の場合、「わかる」と答えた非障がい者が中・重度障がい者に比べて有意に多かった ($p=0.006$) もの、その他の組み合わせでは差が認められなかった。なお②フォルマント合成音声と③波形接続型合成音声の (i) 音声だけの場合は、いずれもグループでも有意な差が検出できなかった (表6)。

次に被験者をグループに分けたとき、各人が (i) 音声だけ、(ii) 音声と絵、漢字かな混じり文、(iii) 音声と絵、ひらがな文を聞いた時どのように評価していたかを、グループごとに表8に示した。(a) 軽度障がい者の数値を見ると、いずれも「わかる」とした評価が多かった。ところが (b) 中・重度障がい者の数値では、①肉声と③波形接続型合成音声で「わかる」と判定した被験者が多かったが、②フォルマント合成音声を聞いてもらったときには「わからないときがある」や「わからない」が多かった。非障がい者ではいずれも理解できるとした被験者が多かったが、②フォルマント合成音声だけでは「わからないときがある」とした回答がめだった。ただし②フォルマント合成音声であっても、視覚情報があると「わかる」と判定した非障がい者が多数いた。

目の前にいる人の発話など

筆者が直接「わたしの言うことが理解できますか？」と問いかけると、ほとんどの被験者は「理解できる」と答えた (表9)。さらに「この実験で疲れませんか？」と問いかけると、軽度障がい者は大部分が「疲れなかった」と答えたが、中・重度障がい者は「疲れた」と回答した被験者が「疲れなかった」と回答した被験者よりも多かった。非障がい者の「疲れた」と回答した被験者は、軽度障がい者に比べても、やや多かった (表9)。

表9 直接の問いかけを理解できるかどうか、この実験で疲れたかという問いかけへの回答。

	(a)軽度障がい者	(b)中・重度障がい者	(c)非障がい者
三谷の話していることが			
理解できる	15	13	45
わからないところがある	1	3	1
わからない	0	1	0
疲れたか？			
疲れなかった	15	6	28
疲れた	1	10	5
無回答	0	1	13

議 論

言葉に関わる脳の領域はさまざまである (石合, 2012) ため、言葉にかかわる障がいの程度を軽いか重いかで分けて「軽度障がい者」や「中・重度障がい者」に分類することは、医学的に無理がある。そのために表4にまとめた回答も、神経心理学的な、あるいは脳科学的な実態はさまざまであると予想できる。しかしながら、社会参加を目的に生涯学習施設において高次脳機能障がい者に理解しやすい音声の質を問うとき、医学的な正確さを追い求めることには、あまり意味がない。科学的な正確さには欠けるが、恣意的に分けただけのグループに対しても、ある (合成) 音声かどの程度機能しているかを知ることは、社会教育学的には意味があるはずだからである。今回の実験と分析では神経心理学的、脳科学的な正確さは追わず、一般利用者に高次脳機能障がい者が含まれる場合を想定した議論を行う。

音声理解の鍵となる要素

結果で示した「音声の評価」の内、『『あなたのことがだーいすき』の朗読を条件を変えて聞いてもらったときの結果』(表6)のように「わかりやすい／わかりにくい」と被験者自身が判断をした3グループ (軽度障がい者、中・重度障がい者、非障がい者) の間で差異があるかどうかを Kruskal-Wallis test で検討した結果、差が検出されなかったのは、フォルマント合成音声と波形接続型音声を音声だけで示した場合であった。肉声はグループ間で差が認められたが、フォルマント合成音声と波形接続型音声では差が認められなかった。この結果は、肉声では非障がい者の聞こえ方と高次脳機能障がい者の聞こえ方に未知の差があり、一方合成音声では、非障がい者でも高次脳機能障がい者と同じように聞こえたのかもしれない。

そのことを確かめるために多重比較によって各グループ間の反応を比較したが、必ずしも合成音声の聞こえ方

が非障がい者と障がい者で同じというわけではなかった(表7)。フォルマント合成で作成した合成音声は軽度障がい者—中・重度障がい者と、非障がい者—中・重度障がい者の間に差が認められた。しかし波形接続型合成音声では、非障がい者と中・重度障がい者に顕著な差が認められたことを除いて、いずれも有意な差は検出できなかった。非障がい者や軽度障がい者には認められないこのような中・重度障がい者の反応は、フォルマント合成音声(図1)に対する中・重度障がい者の反応(表8)に表れていると想像できる。

遠藤ほか(2005)はヒト言語音の解読の鍵は、子音、子音と母音をつなぐフォルマントの遷移部となる過渡部、そして母音のフォルマント(図2)に認められることを、ソナグラム分析と脳の損傷部位を特定した高次脳機能障がい者への語音認知の正答率の比較から見付けた。中でも遠藤ほか(2005)は子音と過渡部の形態が重要であることを指摘している。子音と過渡部、中でも過渡部の形態を合成音声のみで再現することは、現在でも技術的に難しいだろう。このことが失語症者を含む高次脳機能障がい者にフォルマント合成音声の理解を難しくしている要因かもしれない。

なお軽度障がい者では肉声のみならず、フォルマント合成音声や波形接続型合成音声でも「わかる」とする回答が多かった(表8の軽度障がい者の項)。これはフォルマント合成音声聞いてもらったときの非障がい者で、音声だけだと「わからないときがある」や「わからない」が多いが、絵や文章を同時に見てもらおうと「わかる」が増える(表8の非障がい者の項)反応とは明らかに異なっている。この現象の意味は、今後、解明すべき課題である。

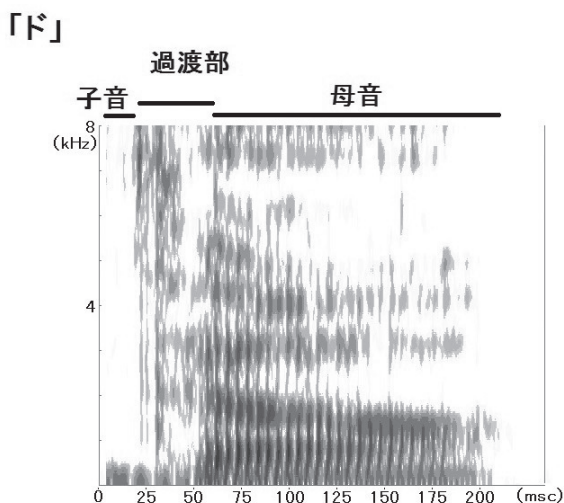


図2 母音「ド」をソナグラムで可視化したもの。母音は子音部、過渡部、母音部からなる。左から右に子音部、過渡部、母音部のパターンを示す。遠藤ほか(2005)を基に図を改変。

高次脳機能障がい者に理解しやすい公共放送

スピーカ・チェックのために行った予備的な回答(表5)を見ると、障がいの軽い人と非障がい者は、肉声やフォルマント合成音声では、音声の聞きやすさに差がなかったが、肉声では中・重度の障がい者と非障がい者の間で、フォルマント合成音声には軽度の障がい者と中・重度の障がい者の間で差が検出できた。この実態は不明だが、少なくとも中・重度の障がい者に重要な放送を聞いてもらう必要がある場合には、非障がい者や障がいの軽い人とは異なる配慮が必要であることを示す。

高次脳機能障がい者に理解しやすい音声は、障がいの程度によらず人の肉声(表8の録音した肉声、表9の「直接の問いかけを理解できるかどうか」の回答)であった。しかし、直接、人が録音して肉声を展示解説に使ったり緊急避難放送に利用することは、可能ではあるものの、あらゆるレパートリーを常に再生可能であるようにしておく必要がある。そのため常に更新していくことが望ましい展示解説では、経費と手間の両面から難しいかもしれない。

その時、人の肉声に次いで推奨されるのは波形接続型合成音声であった(表7, 8)。波形接続型合成音声であれば、高次脳機能障がいの有無や重い／軽いに関わりなく、比較的、理解しやすい(表7, 8)。生涯学習施設で合成音声を使用しなければならない時は、波形接続型合成音声の活用を検討するべきである。

今回、検討した音声の中でもっとも利用に簡便な形態がフォルマント合成技術である。フォルマント合成音声の制作ではワード・プロセッサのようなテキスト入力方式が使える、その応用例は視覚障がい者や読字障がい者が利用することの多いスクリーン・リーダーなどに見られるように、汎用性はきわめて高い。しかし、今回の実験では、重い高次脳機能障がい者には理解できないことの多い合成技術であるとされた(表7, 8)。フォルマント合成方式は非障がい者においてさえも音声だけでは理解することが難しいことがある。一般に障がい者を含む一般利用者の利用には、特に慎重であるべきである。

謝 辞

本研究は日本学術振興会 平成25～27年度 科学研究費助成事業(学術研究助成基金助成金) 基盤研究(C)「DAISYを応用したコミュニケーション障がい者にもわかりやすい展示解説技術の開発」(課題番号25350404, 代表者 三谷雅純)、およびJR西日本あんしん社会財団 平成27年度研究助成「公共放送で使われる人工合成音声聞き取りにくいコミュニケーション障がい者の実態把握」(助成番号15R007, 代表者 三谷雅純)から研究費の助成を受けた。関西テレビCSR推

進局から技術的な援助を受けた。

浜口美幸さん、大本紗緒里さん、安田綾穂さん、松田沙耶香さん、谷口春菜さん、美濃杏奈さん、瀬良裕子さん、植山貴子さん、伊藤むつみさん、田口加代子さん、林今日子さん、南恵美子さん、小野真弓さん、両角明子さん、氏丸淳子さん、後藤順子さん、西本正子さん、隅野光代さん、田中博士さん、平岡重光さん、黒田健一さん、岡本久仁子さん、今泉カツ子さん、新井秀福さん、大窪むつみさん、西之原郁子さん、竹中幸子さん、金森静枝さん、佐々木典子さん、本田 實さん、長谷川政美さん、川口昭彦さん、川東 透さん、川本道德さん、尾崎美由紀さん、波多野優さん、波多野スミ江さん、稲美 修さん、藤井伸彦さん、橋本祐三子さん、篠村幸子さん、杉本和彦さん、北園福吉さん、木下邦昭さん、木下晶人さん、榎本友二さん、谷尾利之さん、前田 実さん、福本修三さん、近澤智康さん、朝倉英里さん、朝倉ひろみさん、樽居幸信さん、樽居一美さん、南部恒夫さん、南部秀子さん、谷口博明さん、谷口セツ子さん、福島 實さん、福島十三子さん、仲島成好さん、仲島清美さん、大森清和さん、大森雅子さん、橋本祐三代さん、福本喜代子さん、近澤千哉さん、木原誉子さん、香山加代子さん、川本京子さん、松岡佐和子さん、浅田美穂さん、寺田博子さん、伊東めぐみさん、石塚君予さん、安居和輝さん、安居道子さん、田中加代子さん、田中昌明さん、宮前周司さん、森田誠二さん、山本悠美子さん、新見彰平さん、水川賢一さん、武田直子さん、山家健盛さん、中西祥子さん、広谷浩子さん、田口公則さん、東 ゆかりさん、南 和代さんと、お名前を挙げられなかった、この聴覚実験でお世話になった全ての皆さんに感謝する。

要 旨

高次脳機能障がい者にわかりやすい音声を探るために、高次脳機能障がい者 33 名、非障がい者 46 名の参加を得て聴覚実験を行った。実験に使ったのは女性の肉声、フォルマント合成音声、波形接続型合成音声であった。非障がい者は人の肉声や波形接続型合成音声に対して「わかる」とする回答が多かったが、聴覚情報だけのフォルマント合成音声では「わからない」とする回答が増えた。しかし、聴覚情報に視覚情報を加味すると「わかる」被験者が増えた。障がいの軽い人では、肉声、フォルマント合成音声、波形接続型合成音声のいずれでも

「わかる」とする回答が多かったが、障がいの中・重度の人では「わからない」とする回答が多かった。中でも中・重度の人ではフォルマント合成音声がわからない被験者がめだった。この実験の結果から、本来、障がい者を含む多様な人が利用する生涯学習施設では、人の肉声を直接、あるいは録音したもので放送することが一番わかりやすいが、合成音声を使う場合には波形接続型合成であることが望ましい。生涯学習施設の職員にとってフォルマント合成は波形接続型合成よりも簡便だが、使用は慎重であるべきだ。

文 献

- 阿部知暁(1994) くんくんくん おいしそう。こどものとも 年中向き 190, 福音館書店, 東京, 27 p.
- 遠藤邦彦・阿部晶子・津野田聡子・柳 治雄・衣佐原 均(2005) 言語音を解説する神経機構：音のどこに解説の手がかりは書き込まれるか？ 高次脳機能研究, **25**, 165-178. (2015年7月3日閲覧) [https://www.jstage.jst.go.jp/article/hbfr/25/2/25_2_165/_pdf]
- ヒド・ファン・ヘネヒテン(2006) マルチメディア DAISY 図書 あなたのことが だーいすき。ひしき あきらこ (訳), 日本障害者リハビリテーション協会, 東京, CD-ROM (原本出版社 フレーベル館, 東京, 26 p.)
- 居村太介・村上仁一・池原悟(2008) 波形接続型音声合成のフレーズへの適用. 言語処理学会第14回年次大会発表論文集 965-968. (2015年10月27日閲覧) [http://unicorn.ike.tottori-u.ac.jp/2007/s042011/paper/NLP_2008_03/PD3-1.pdf]
- 石合純夫(2012) 高次脳機能障害 第2版. 医歯薬出版株式会社, 東京, 281 p.
- 河村 宏(2011) デジタル・インクルージョンを支える DAISY と EPUB. 情報管理, **54**, 305-315. (2015年6月9日閲覧) [https://www.jstage.jst.go.jp/article/johokanri/54/6/54_6_305/_pdf]
- 三谷雅純(2013) 生涯学習施設は言葉やコミュニケーションに障がいを持つ人とどう向き合うべきか：総説. 人と自然 Humans and Nature, **24**, 33-44. (2015年7月3日閲覧) [http://www.hitohaku.jp/publication/r-bulletin/No24_04.pdf]
- 三谷雅純(2014) 生涯学習施設の館内放送はどうあるべきか：聴覚実験による肉声と人工合成音声の聞きやすさの比較. 人と自然 Humans and Nature, **25**, 63-74. (2015年7月3日閲覧) [http://www.hitohaku.jp/publication/r-bulletin/No25_02.pdf]
- 山鳥 重(1985) 神経心理学入門. 医学書院, 東京, 410 p.

(2015年 7月29日受付)

(2015年 11月25日受理)

