

【原著】

音声出題システムにおいてコンピュータ合成音声を受験者の 解答時間および得点に与える影響

南谷和範，藤芳衛（大学入試センター入学者選抜研究機構）

文字の認知に困難を抱える障害者の受験を可能とすることを目的として，音声出題システムの開発を行っている。今回，当該音声出題システムにおいて，点字使用の視覚障害者と普通文字使用健常者を被験者群として，コンピュータ合成音声を受験者の解答時間および得点に与える影響を調べた。肉声による音声問題，コンピュータ合成音声による音声問題の間には，解答時間・得点いずれにおいても有意な差は見られなかった。

1 はじめに

文字の認知に困難を抱える障害者の受験を可能とする音声出題システムの開発を行っている。大学入試センター試験は開発当初から障害を有する受験者を始め，全ての受験者に公平に配慮して試験を設計する方針で開発されている（藤芳,2009）。具体的には，通常の文字冊子問題に加えて，重度視覚障害者用に点字冊子問題が，弱視者用に拡大文字冊子問題が用意されている（大学入試センター,2010）。

さらに，近年2次元コードを活用した「文書構造表方式の音声問題」を試作した（藤芳ほか,2009）。この方式では，通常のプリンタ印刷を用いて文書構造表を視認可能な表として印刷し，そこに視認を妨げない2次元コードを付与する。受験者は，専用のペン型音声読み上げ機能付きコードリーダーで2次元コードに対応付けられた音声を読み上げさせ，問題文を聴く。この方式は，重度読字障害者の受験での利用が期待できる。

ただし，この方式では普通文字の視認ができない重度視覚障害者には利用できない。日本の教育過程では，重度視覚障害者には点字の読み書き技能の指導が行われる。そのため，点字使用の技能を身につけた重度視覚障害者には点字冊子問題での出題が実施されている。

しかしながら，効率的な点字の読み書きの習得には最低でも3年程度の年月が必要である。それゆえ，中等教育段階で急激な視力の低下に見舞われ，普通文字の読み書きが不可能となるとともに，いまだ点字の読み書きを十分習得しえてない中途重度視覚障害者には，通常の文字冊子問題はもちろん，点字冊子問題の出題も，上記の文書構造表方式の音声問題出題も適用できない。

このような中途重度視覚障害者の試験実施に対応するために，「点字文書構造表方式の音声問題」を開発した（南谷ほか,2010）。この方式では，点字プリンタを用いて文書構造表を点字シンボルとして印刷し，そこに2次元コードを付与する。専用のペン型音声読み上げ機能付きコードリーダーで2次元コードに対応付けられた音声を読み上げ試験問題をアクセス可能とする点は，通常のプリンタ印刷を用いる文書構造表方式の音声問題と同様である。

現在，上述の通常印刷ないし点字印刷を用いる文書構造表方式音声問題において，コンピュータによる合成音声の利用可能性を研究している。その一環として，今回，当該音声問題を用いた出題システムにおいて，コンピュータ合成音声を受験者の解答時間および得点に与える影響を調べた。

2 問題収録に用いる音声の概要

今回、筆者らは、問題文読み上げ音声として、肉声に加えコンピュータ合成音声を用い、その被験者の回答に及ぼす影響を検証した。

2.1 コンピュータ合成音声による出題のメリット

コンピュータ合成音声による試験問題作成には肉声による作成と比較して、以下のようなメリットがある。

(1) リーダー(文字の認知に困難を抱える障害者のために文章を代読し読み上げる読み手を慣習的に「リーダー(reader)」と呼ぶ)が不要である。肉声問題の朗読は、読み上げの訓練を積んだ専門のリーダーが行わなくてはならないが、コンピュータ合成音声を用いる場合にはこうした人材を必要としないため、セキュリティの向上が期待できる。また、音声収録がリーダーの健康状態に左右されることがないため、作題計画が安定的に実行できる。後述のように読み上げ速度、ピッチ、イントネーションなどを数値化された値の操作で変更できるため、朗読者に対する感覚的な指示に依存しない出題者の意図に即した読み上げを実現することが容易である。

(2) 読み上げの速度、ピッチの調整がソフトウェアに与えるパラメータによって容易に変更可能であり、多様な音声の試験問題を作成できる。被験者にとって望ましい音声読み上げ速度やピッチは一定ではなく、これらの条件は柔軟に変更できるべきである。現在の音声出題システムでは、再生機の機能的制約の範囲で、3段階の選択を可能としている。

(3) コンピュータ上で柔軟なコンテンツ編集が可能である。前述のように数値化されたパラメータで音声を制御でき、読み上げ品質の一貫性、再現性が保証される。そのため、問題文章の差し替えのような編集作業が容易かつ迅速に実施できる。

2.2 問題文収録に用いた肉声とコンピュー

タ合成音声

2.2.1 肉声

肉声問題は、視覚障害者のための図書朗読サークルに所属する経験者に問題文の読み上げを依頼した。英語の読み上げは、イントネーションを付与しないいわゆるジャパニーズ・イングリッシュではなく、ネイティブスピーカーの発音・イントネーションを基準にした。イントネーションは若干抑制した朗読となった。

2.2.2 コンピュータ合成音声

日本語音声は、公的サービスでの採用例が多く肉声に近い読み上げと言われる(株)日立ビジネスソリューションのボイスソムリエに含まれる TTS エンジン(Text To Speech engine)を用いた。英語音声は(有)エクストラのスクリーンリーダー JAWS for Windows 日本語版に付属する RealSpeak Solo 4.0 TTS エンジン(Daniel 22KHz)を用いた。声質は、近年、英語圏の視覚障害者向け音声機器などの組み込み機器での採用が多く見られる Nuance 社の TTS と同等である。ネイティブスピーカーの発音に準じた英語合成音声となる。

3 評価実験

3.1 実験の概要

構造表方式コンピュータ合成音声問題を評価するために、文字冊子問題、構造表方式肉声音声問題、構造表方式コンピュータ合成音声問題の三つのテストメディアの要因が得点および回答速度に及ぼす効果を分析した。重度視覚障害者と健常者をそれぞれ被験者とする評価実験を行った。以下、構造表方式肉声音声問題を「肉声問題」、構造表方式コンピュータ合成音声問題を「合成音声問題」と略記する。また、構造表方式肉声音声問題と構造表方式コンピュータ合成音声問題を総称して「音声問題」と呼ぶ。

出題する問題の要因は、現在大学入試セン

ターで研究している総合試験問題の国語・数学・英語 3 教科のそれぞれ 3 セットずつである。手続きは試験時間を制限しない作業制限法である。音声の再生話速度は、通常の朗読速度の 1 倍速、1.5 倍速、2 倍速の 3 段階に任意に切り替え可能とした。文字冊子問題と音声問題の構造表は、重度視覚障害者用には点字印刷を、健常者用には通常印刷を用いた。試験の性質上、同じ問題を同じ被験者群にテスト・メディアを替えて出題できない。そのため、実験計画は、繰り返しのある 3×3 のグレコ・ラテン方格法である。表 1 に共通の実験計画のイメージを示す。

表 13×3 のグレコ・ラテン方格法の実験計画のイメージ

		被験者群		
		1 群	2 群	3 群
順序	1	文字冊子	肉声	合成音声
	2	肉声	合成音声	文字冊子
	3	合成音声	文字冊子	肉声

(1) 重度視覚障害者を被験者とする評価実験。被験者群の要因は、点字冊子問題と音声問題が使用可能な視覚特別支援学校の重度視覚障害高校生 5 名ずつの 3 群、計 15 名である（一部直近の卒業生を含む）。重度視覚障害者を被験者としたのは、前述のように中途重度視覚障害者が音声出題の適用対象として想定されるためである。テストメディアが得点および解答速度に及ぼす効果を評価するために、文字問題の出題が必要である。視覚障害被験者に出題する文字問題は、通常文字冊子問題を点字に翻訳（点訳）した点字冊子問題となる。よって、視覚障害被験者は、点字使用に重大な困難を抱える中途重度視覚障害者とはせず、点字の読み書きに習熟した重度視覚障害者とした。

テストメディアの要因は、点字冊子問題、肉声問題、合成音声問題の 3 水準である。

(2) 健常者を被験者とする評価実験。被験者

群の要因は、通常の冊子問題と音声問題が使用可能な都内公立高校に在学する高校生 7 名ずつの 3 群、計 21 名である。

テストメディアの要因は、通常の冊子問題、肉声問題、合成音声問題の 3 水準である。

3.2 実験結果

以下、得点と解答速度に対して及ぼす効果を、重度視覚障害者および健常者を被験者とする評価実験それぞれについて述べる。被験者群、テスト・メディア、問題、順序の 4 要因が及ぼす効果を略述した後、特に今回注目するテスト・メディアが及ぼす影響を詳述する。

3.2.1 得点分布

得点分布をテスト・メディア間で比較するため、テスト・メディア別得点分布の箱ひげ図と、Scheffe のテスト・メディア間の多重比較結果を図 1 に示す（次ページ）。「検定結果」の縦線は線で結ばれたテスト・メディア間に有意差がないことを示す。

以下、0 内は平均点（100 点満点）を示す。

(1) 重度視覚障害者を被験者とする評価実験

得点に対する被験者群、テスト・メディア、問題、順序の 4 要因の及ぼす効果を検討するため、分散分析を行った。

テスト・メディアの要因には、国語・英語・数学の 3 教科ともすべて有意な主効果が認められなかった。被験者群の要因には、国語・英語・数学の 3 教科ともすべて有意な主効果が認められなかった。問題の要因の主効果は、国語と英語は有意でなかったけれども、数学は有意であった。順序の要因の主効果は、すべて有意でなかった。

得点分布は、3 教科ともすべてテスト・メディア間に有意差は認められなかった。国語と英語では、点字冊子問題(国語 78.667、英語 68.718)、肉声問題(国語 73.333、英語 68.205)、合成音声問題(国語 71.556、英語 58.205)の順で、若干得点が下がる傾向が認め

られるけれども、有意差はすべて認められなかった。

点を教科間で比較すると、国語と英語の得点はほぼ同様であるけれども、数学(点字 52.51、肉声 40.50、合成音声 42.36)はかなり低いことが認められた。

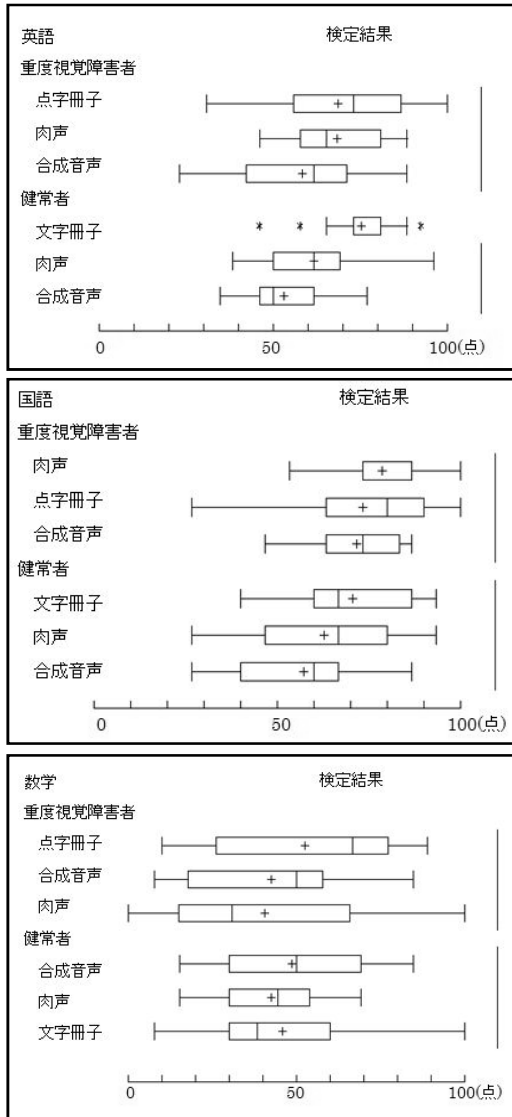


図1 テスト・メディア別得点分布の箱ひげ図と Scheffe の多重比較

(2) 健常者を被験者とする評価実験

得点に対する被験者群、テスト・メディア、問題、順序の4要因の及ぼす効果を検討するため、分散分析を行った。テスト・メディアの要因には、英語では有意な主効果が認められた。国語・数学では有意な主効果が認めら

れなかった。

被験者群の要因には、英語では有意な主効果が認められた。国語・数学では有意な主効果が認められなかった。問題の要因の主効果は、英語は有意でなかったけれども、国語と数学は有意であった。順序の要因の主効果は、英語と国語は有意でなかったけれども、数学は有意であった。

得点分布は、英語では、文字冊子方式と二つの音声問題の間で有意差が認められた。国語と数学ではテスト・メディア間に有意差は認められなかった。国語と英語では、文字冊子問題(国語 70.476、英語 75.275)、肉声問題(国語 62.857、英語 61.538)、合成音声問題(国語 57.143、英語 52.930)の順で、得点が下がる傾向が認められ、英語では文字冊子問題と肉声問題・合成音声問題の間に有意差が認められた。数学では、合成音声問題(48.551)、文字冊子問題(45.828)、肉声問題(42.637)の順で、若干得点が下がる傾向が認められたが、有意差は認められなかった。

得点を教科間で比較すると、国語と英語の得点はほぼ同様であるけれども、数学はかなり低いことが認められた。

3.2.2 回答速度

回答速度は文書量を回答所要時間で割って、単位時間あたりに処理可能な文書量として算出した。解答速度の分布をテスト・メディア間で比較するため、テスト・メディア別解答速度分布の箱ひげ図と、Scheffe のテスト・メディア間の多重比較結果を図2に示す(次ページ)。「検定結果」の縦線は線で結ばれたテスト・メディア間に有意差がないことを示す。

以下、0内は回答速度(1分あたりの処理文字数)の平均値を示す。

(1) 重度視覚障害者を被験者とする評価実験

回答速度に対する被験者群、テスト・メディア、問題、順序の4要因の及ぼす効果を検

討するため、分散分析を行った。テスト・メディアの要因には、英語では有意な主効果が認められた。国語・数学では有意な主効果が認められなかった。

被験者群の要因には、英語・数学では有意な主効果が認められなかった。国語では有意な主効果が認められた。問題の要因の主効果は、英語と数学は有意でなかったけれども、国語は有意であった。順序の要因の主効果は、英語と数学は有意でなかったけれども、国語は有意であった。

英語の回答速度は、点字問題(103.60)、肉声問題(74.26)、合成音声問題(73.08)の順に遅くなり、点字問題と二つの音声問題の間に有意な差が認められた。国語の回答速度は点字(134.07)、合成音声(129.44)、肉声(106.41)の順であり、メディア間に有意さは認められなかった。数学の回答速度は肉声(38.108)、点字(37.427)、合成音声(34.373)の順であり、メディア間に有意さは認められなかった。

(2) 健常者を被験者とする評価実験

回答速度に対する被験者群、テスト・メディア、問題、順序の4要因の及ぼす効果を検討するため、分散分析を行った。テスト・メディアの要因には、英語、国語、数学の3教科いずれでも有意な主効果が認められた。被験者群の要因には、英語、国語、数学の3教科いずれでも有意な主効果が認められなかった。問題の要因の主効果は、英語と数学は有意でなかったけれども、国語は有意であった。順序の要因の主効果は、国語と数学は有意でなかったけれども、英語は有意であった。

英語の回答速度は、文字問題(375.68)、肉声問題(104.02)、合成音声問題(99.05)の順に遅くなり、文字問題と二つの音声問題の間に有意な差が認められた。国語と数学の回答速度は文字問題(国語 394.14、数学 55.138)、合成音声問題(国語 168.83、数学 35.263)、肉声問題(国語 142.95、数学 34.801)の順に遅くなり、文字問題と二つの音声問題の間に有意な

差が認められた。3教科いずれにおいても、文字冊子問題と二つの音声問題の間で解答速度に有意差が認められ、二つの音声問題の解答速度は、文字問題に比べて遅い。

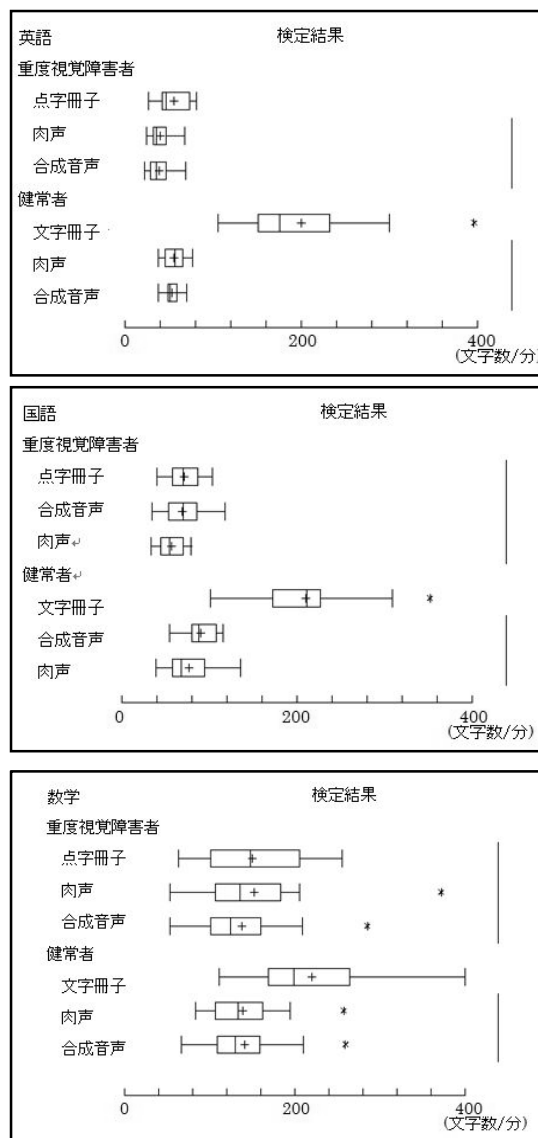


図2 テスト・メディア別解答速度分布の箱ひげ図と Scheffe の多重比較

4 結論と今後の課題

4.1 結論

総合試験問題における、合成音声問題および音声問題全般のテストメディアとしての特性について、以下を結論として導くことができる。

(1) 重度視覚障害者と健常者の得点、解答速度のすべてにおいて、肉声問題と合成音声問

題の間に有意差は認められない。合成音声による音声問題の出題を、肉声による音声問題の出題と同等の条件で行うべきではないことは示されなかった。

(2) 健常者の英語を例外として、肉声問題、合成音声問題で文字問題と同等の得点の取得が期待できる。重度視覚障害者に対しては、3教科いずれにおいても、素点をそのまま用いて採点ができる。

(3) 健常者の3教科すべて、重度視覚障害者の英語においては、文字冊子問題とは異なる回答時間を設定する必要がある。

4.2 今後の課題

上記結論を踏まえ、今後以下の研究課題が考えられる。

(1) 望ましい合成音声問題とそのオーサリング環境の研究。試験問題の音声出題に際して合成音声を用いることには2.2で示したような多くのメリットが存在する。今回の実験結果の検定力の限界に十分留意しつつ、音声合成技術の開発動向を注視し、望ましい音声出題方式の試験のあり方やオーサリング環境の研究が必要である。

(2) 英語試験問題の音声出題について。テストメディアとしての二つの音声問題は、重度視覚障害者の得点を除き、得点と解答速度に関して文字冊子問題との間に有意差が認められた。外国語の聴覚的な理解の特性を考慮しつつ、提示方法を検討する必要がある。

(3) 他試験への応用。他の試験問題、特に大学入試センター試験で同様の結果が得られるか検証する。同試験は、総合試験問題に比べ、難易度が高く、長文の分量も多い。このような大問形式の問題における音声出題が与える影響を分析しなくてはならない。

(4) 音声出題における試験時間延長の推定。今回の評価実験から、視覚障害者への試験実施に際して、肉声・合成音声に関わらず音声による出題では、点字冊子による出題と比較

して解答に同等ないしそれ以上の時間を要することが示唆されている。音声問題の回答に必要な時間延長を検証する。

(5) 視覚障害者の音声聞き取り能力についての研究。視覚障害者の音声聞き取り能力は、健常者よりも優れ、単位時間内により多くの情報を理解できるという認識が視覚障害関係者の間で広く共有されている(浅川,2005)が、このような認識と本研究が示唆する認識は親和的ではない。文章の長さや難易度、解答作業の介在の有無と言ったパラメータが、視覚障害者の音声聞き取り・理解に要する時間の増減に与える影響についての研究が必要である。

謝辞実験に協力いただいた皆様と適切な指摘をくださったレフリーの方に感謝します。

参考文献

- 浅川智恵子・高木啓伸・井野秀一・伊福部達(2005).「視覚障害者への音声提示における最適・最高速度」『ヒューマンインタフェース学会論文誌』7(1),105-111.
- 藤芳衛(2009).「テストのユニバーサルデザイン」『電子情報通信学会誌』92(12), 1022-1026.
- 藤芳衛・藤芳明生・青松利明(2010).「重度の読字障害者及び中途失明者の受験を可能にする文書構造表方式の音声問題の開発」『大学入試研究ジャーナル』, 20, 131-138.
- 南谷和範・藤芳衛・青松利明・原田清生(2010).「通常文字および点字の使用に困難を抱える中途重度視覚障害者の受験を可能とする音声出題システムの開発」『日本教育工学会第26回全国大会予稿集』.
- 大学入試センター 『平成23年度受験案内別冊』 <http://www.dnc.ac.jp/modules/center_exam/content0023.html> (2011年4月4日)