

科学研究費助成事業（科学研究費補助金）研究成果報告書

平成25年 6月20日現在

機関番号：82616

研究種目：若手研究（B）

研究期間：2010～2012

課題番号：22700835

研究課題名（和文）音声・点字による試験問題出題に必要な字注・意味注の挿入を実現するシステムの開発

研究課題名（英文）Development of a software which semi-automatically generates notes to explain kanji characters for examinations using voice media or braille booklet

研究代表者

南谷 和範（MINATANI KAZUNORI）

独立行政法人大学入試センター・入学者選抜研究機構・准教授

研究者番号：90551474

研究成果の概要（和文）：テストメディアとして音声ないし点字を用いる場合に判別が困難な同音異字語を説明するための注記挿入を半自動で行うシステム（ソフトウェア）を開発した。また、研究の進捗過程でその必要性が確認された。

(1)『視覚障害者(児)のための漢字説明辞書』の編纂

(2)センター試験過去問中の字注の分析

(3)音声出題システムの基礎となるコンピュータ試験問題閲覧環境の開発

を行い、研究の目的達成度の総合的な向上に努めた。

研究成果の概要（英文）：In this research, a software which semi-automatically generates notes to explain kanji characters is developed.

Such characters which have same pronunciation could not be distinguishable when using voice media or braille booklet.

Also, three tasks are carried out:

(1) a compilation of a kanji dictionary for the blind,

(2) an analysis of sentences in exam questions in the National Center Test for University Admissions and

(3) development of a DAISY player that utilizes a braille display for document structure presentation and navigation.

交付決定額

(金額単位：円)

	直接経費	間接経費	合計
2010年度	900,000	270,000	1,170,000
2011年度	800,000	240,000	1,040,000
2012年度	1,200,000	360,000	1,560,000
総計	2,900,000	870,000	3,770,000

研究分野：総合領域

科研費の分科・細目：科学教育・教育工学

キーワード：視覚障害者・読字障害者・学力測定・字注／意味注・音声出題

1. 研究開始当初の背景

(1) 大学入試センターにおけるテストのユニバーサルデザインの研究。応募者が所属する大学入試センターでは、誰もが受験可能な試験を保障すべく、関連する研究を推進して

いる。ここ数年来は、障害受験者の試験時間延長倍率の研究に加え、読字障害を有する発達障害者や中途失明者のセンター入試受験を可能とする音声出題の研究が行われている。

現在は、試験実施に際してのセキュリティとコストを加味し、合成音声を用いた試験問題のオーサリングの研究を進めている。また、同センターでは重度視覚障害を有する受験者のために点字によるセンター試験出題を共通1次学力試験開始以来30年以上にわたり実施している。

(2) テストメディアとしての音声および点字の問題点。メディアとしての音声と点字には(「自立」と「自律」、「提起」と「定期」のように)同音異字語の判別が難しいという共通の問題が存在する(日本語点字は仮名分かち書きで日本語文章を表現するため、文章中に用いられている漢字を確認することができない)。このようなメディアを出題に利用する場合、漢字の判別不能ゆえの不利益を受験者が被ることになる。そのためセンター入試の点字試験問題作成に際しては、以前より問題作成に携わる特別問題部会において、字注、意味注の付与作業が行われている。字注とは「チュウ：注ぐの注」のような形式で漢字を1文字単位で説明する注記であり、意味注とは「イミ：ある表現に対応し、それによって指示される内容。」のように単語単位で意味内容を説明するものである。これらの注記は熟練した部会委員が経験的な知識に基づいて、

- ・注記の必要な場所、
- ・字注・意味注いずれを用いるか、
- ・注記の説明文、

を是々非々で判断し作成してきた。そのために、注記が必要な箇所が看過される可能性、注の使用の一貫性が保障できない(同じ文字や単語の説明に別様の字注・意味注が用いられる)というような問題が潜在している。

合成音声を用いた音声問題作成の実用化を視野に入れて検討するのであれば、字注・意味注の付与に習熟した人員の関与を前提とした体制を想定することは難しい。従来のようにマンパワーに頼ることなく、一定程度自動化された字注・意味注の付与手段を確立する必要がある。

2. 研究の目的

本研究は、音声および点字のような漢字を用いないテストメディアでは判別が困難な同音異字を説明するための注記の挿入に際して、

- ・注記が必要な箇所(単語)を抽出し、
- ・適切な注記(説明文)を生成

するシステムの開発を行う。具体的には以下の作業を自動化するシステムを構築する。

(1) 注記が必要な箇所(単語)の抽出を実現するために単語親密度を主要な指標とする、同音異字抽出システムを開発する。適切な注記(説明文)の生成を可能とするために

① 詳細読み辞書を活用した字注付与

② 国語辞書を用いた意味注候補文章提示を実現する。

受験者数と同音異字語の出現頻度を考慮し、国語、地歴公民(日本史、世界史、政治・経済)を本研究の第一義的な対象科目とする。なお、試験問題に要求される完全性に鑑み、担当者による校閲・確認の全面的な廃止を目標とはせず、本システムによる作題作業の効率化と問題品質の安定性保証を目標とする。

3. 研究の方法

以下、研究開始後の状況の変化を加味しつつ、研究期間内の進捗に即して講じた手段を記述する。

(1) コンテンツ(試験問題)の電子的な形態での供給・保持のための環境整備。

① コンピュータ上での試験問題データの扱いを可能とする格納形式の選定。当初、本研究では独自XML形式を策定し、これを用いて過去のセンター試験問題を記述・格納することを計画した。しかしながら、研究開始年度である22年度以後、電子書籍をめぐる技術的発展と社会的・経済的展開が本格化した。本研究も独自形式ではなく、こうした潮流に即した研究推進を要求された。国際的には共通フォーマットとしてのEPUBが着実に浸透する一方で、国内の電子書籍をめぐる状況は明確な方向性が現れない状況が続いた。また、電子教科書をはじめとする教育分野の個別の状況や試験問題に要求される独特の要件についても総合的な調査・検討を行った。国際学会での調査からは、視覚障害・読字障害分野において、DAISYフォーマットをベースにリッチ・コンテンツを整備する趨勢が確認されたが、後継企画策定中のため、現状どのバージョンを用いて試験問題を入力するか熟慮を要することも判明した(→DAISY 2.02を選択)。また同フォーマットでは規格上、テキスト情報が簡便に扱えないことを確認し、データ入力に用いるオーサリングツールの検証を慎重に進めた。将来的な他フォーマットへの移行の可能性も考慮し、SaveAsDAISYを選択した。

② 検証用データの入力。①に従い2000年以降のセンター試験過去問の入力、データベース化を行った。当初、国語、地歴、公民の試験問題の入力を進めた。しかしながら、たとえば地歴科目では中国の歴史上の国名である「しん」が同音異字語として頻出するなど、統計的な調査の対象として必ずしも有用でないことが判明した。そこで、検証作業におけるデータとしての有効性が高い国語とりわけ、その第1問・第2問を主要な課題として作業を進めた。

(2) システムに用いる辞書データの構築並びにその派生的成果物の編纂。

① システムの実装・評価に必要な辞書

データの編纂。同音異字語の説明は、字注と意味注を適宜使い分けて用いる必要がある。このうち、字注の自動生成においては、文字単位で意味を説明する辞書が必要である。これに類する辞書は、視覚障害児教育や視覚障害者の ICT 利用においていくつか整備されてきていることが確認された。これらの中には

- ・単語親密度に立脚した福祉工学からのアプローチを採用するもの
- ・視覚障害児国語教育の知見に基づくもの
- ・ユーザ(視覚障害者)自信がその実用性を重視して構築したもの

の3種が存在し本研究で活用すべき基礎データと判断した。

② 「視覚障害者(児)のための字注・漢字説明辞書(仮称)」としての編纂。交付申請書「研究の学術的な特色・独創的點」に指摘したように、視覚障害児漢字教育の方針は十分確立された状況にない。そのため、本研究で推進する課題、特にその字注に関連する部分に関しては、視覚障害者(児)に対する漢字指導・解説の広範な活動の中に位置づけることが可能かつ望ましい。そこで、①のシステムの実装・評価に必要な辞書データを本システム専用のデータに終止させることなく「視覚障害者(児)のための字注・漢字説明辞書(仮称)」として広く提供すべく編纂した。冊子形態での編集を一通り完了した。あわせて、検索性を確保した電子書籍形態(ePub フォーマット)での編集についても作業を行った。

(3) 課題システム(ソフトウェア)の実装とそれを応用したセンター試験点字問題過去問の分析。

① 主要課題であった、テストメディアとして音声ないし点字を用いる場合に判別が困難な同音異字語を説明するための注記挿入を半自動で行うシステム(ソフトウェア)を実装した。同システムでは、

- ・音声単語親密度を基準とした問題文からの注を付すべき単語の抽出、
- ・(2)①の辞書データを参照した字注文章の候補提示が実現された。

ソフトウェア開発は、日本語テキスト処理に優れるプログラミング言語 Ruby を用い、形態素解析モジュールとして Mecab を利用した(当初、Chasen の利用を想定したが柔軟性、拡張性に留意し変更)。動作検証には、(1)②の過去問データを用いた。

なお、当初、自動付与の対照として考慮していた「意味注」については、試験問題作成プロセスの調査の結果から、人力による国語辞書の検索が適切かつ十分という見解に到達し、課題から除いた。

② センター試験過去問中の字注の分析。①の実装過程で得られたソフトウェアルーチ

ンを用いて、センター試験国語過去問に対して作題担当者の判断で付された字注の分析を行った。なお、(1)②のデータは一般受験者が受験する試験問題のデータであり、日本語点字表記を用いる点字問題とは異なる。

(4) 音声出題システムの基礎となるコンピュータ試験問題閲覧環境の開発。(1)②のデータの効率的閲覧と視覚障害受験者への音声出題の将来的可能性を調査するためコンピュータ試験問題閲覧環境を実装した。点字ディスプレイを用いて文書構造のナビゲーションを行うユーザインターフェイスを導入し、その有用性をモニター実験で評価した。

4. 研究成果

以下、本研究の主な成果4点について①成果の内容、②国内外における位置づけとインパクト、③今後の展望に分けて記述する。

(1) 課題システムの開発

① テストメディアとして音声ないし点字を用いる場合に判別が困難な同音異字語を説明するための注記挿入を半自動で行うシステム(ソフトウェア)の開発。当初、計画した字注の挿入システムを開発した。音声単語親密度を基準とする注記付与箇所抽出は、

- ・より単語親密度の高い同音異字語を有するもの
- ・その単語の音声単語親密度が文字単語親密度より一定以上低いもの
- ・音声単語親密度が一定以下のものと言ったパラメータで柔軟に設定変更可能とした。これにより、試験の性質やシステム操作者の字注付与作業への習熟度に応じた動作を実現した。

② 当該システムに類する特別措置試験問題製作支援システムは前例がなく、これを具体的な実装を伴って提案・評価できた意義は大きいと判断する。

③ 本システムは、用いる辞書データや文章解析ルーチンを入れ替えることで、日本のみならず他の漢字圏における特別措置試験問題製作において一定の有用性が期待できる。そのため、これら諸外国における特別措置実施状況を含めた調査を行う。

(2) 『視覚障害者(児)のための漢字説明辞書』の編纂

① 当該システムが字注文章生成に用いる辞書データを『視覚障害者(児)のための漢字説明辞書』として編纂した。これにより同データの教育実践での利用も可能となり、学力試験と初等・中等教育の間での漢字説明方法の乖離を防ぐことができる。教育現場での要求に応えられるよう、冊子形態での編集を行った。あわせて、検索性を確保した電子書籍形態(ePub フォーマット)での編集についても作業を行った。

② 本辞書は、

- ・視覚障害者がパソコン上で日本語文章を入

力するに際して用いる「詳細読み」データ・視覚障害児国語教育の副読本資料を横断的に網羅しており、学力試験、教育、ICT利用と言った局面において、それぞれ別様の漢字の説明を使用するという現状を改善する。

③ 欧米では、試験特別措置の選択肢としてリーダー（朗読者）による試験問題の対面読み上げがしばしば利用される。日本国内で、こうした方式を導入するには試験問題中の語の恣意的な説明を回避する公平性と識別力の確保のための読み上げ方式が確立されなくてはならない。本辞書のような資料は、後の説明のガイドラインとなる。

(3) センター試験過去問中の字注の分析・評価。

① センター試験国語(第1問・第2問)過去問に対して作題担当者の判断で付された字注を音声単語親密度の基準を用いて分析した。その結果、当該点字冊子問題に付された注記はおおむね妥当という結論を得た。他方、注記が付された理由が判然としない語の存在も確認された。

② 字注に対して客観性を有する統一的な基準で分析・評価する試みは前例がなく重要と考える。

③ ①で指摘した注記が付された理由が判然としない語について、その注記付与が妥当性を欠くと軽々に判断することは差し控えるべきである。むしろ、理由の確認が困難であることの問題性を重視するのであれば、本分析・評価基準を指標として導入することが有望と予想される。

(4) 音声出題システムの基礎となるコンピュータ試験問題閲覧環境の開発。

①

- ・作成した過去問データの効率的閲覧
- ・視覚障害受験者への音声出題の将来的可能性の調査のためにコンピュータ試験問題閲覧環境を実装した。本環境は試験問題の内容そのものはコンピュータ合成音声で読み上げ、試験問題全体の構造提示と音声読み上げ位置の移動のために点字ディスプレイというデバイスを併用するユーザインターフェイスを採用した。これにより、音声読み上げのみでは厳しく制約される閲覧性の向上を実現した。

② センター試験をはじめ日本国内で広く採用されている大問形式の試験の音声出題に対しては、閲覧性の欠如が実用化を阻む重大な要因として存在する。本環境はこの問題を解消する1つの方法を具体的に示した。

③ 本環境は、作成した過去問データを閲覧するためにDAISY 2.02フォーマットに対応した。汎用的な電子教材・電子教科書閲覧環境としては、ePubあるいはHTMLに対応したブラウザとしての動作が有効である。既存のウェブブラウザをベースに本環境同様の

ユーザインターフェイスを実装した閲覧環境の実装が有望と予想される。

5. 主な発表論文等

(研究代表者、研究分担者及び連携研究者には下線)

[雑誌論文] (計6件)

- ① 南谷和範、視覚障害者への漢字説明と大学入試センター試験点字冊子問題への字注付与状況、大学入試研究ジャーナル、査読有、23巻、2013、135~142
- ② Kazunori Minatani、Development of a DAISY Player that Utilizes a Braille Display for Document Structure Presentation and Navigation、Proceedings of ICCHP 2012: LNCS、査読有、Vol.7382、2012、510~517
- ③ Kazunori Minatani、Tetsuya Watanabe、A Non-visual Interface for Tasks Requiring Rapid Recognition and Response、Proceedings of ICCHP 2012: LNCS、査読有、Vol.7383、2012、631~636
- ④ 南谷和範、藤芳衛、音声出題システムにおいてコンピュータ合成音声が受験者の解答時間および得点に与える影響、大学入試研究ジャーナル(22)、査読有、22巻、2012
- ⑤ 渡辺哲也、山口俊光、渡部謙、秋山城治、南谷和範、宮城愛美、太内進、視覚障害者用触地図自動作成システム TMACSの開発とその評価(福祉工学)、電子情報通信学会論文誌、査読有、J94-D(10)巻、2011、1652~1663
- ⑥ 藤芳衛、南谷和範、藤芳明生、青松利明、澤崎陽彦、読字障害者および重度の弱視者のための文字と音声のマルチモーダル問題の開発、大学入試研究ジャーナル、査読有、21巻、2011、181~190

[学会発表] (計8件)

- ① 南谷和範、文書構造提示・ナビゲーションに点字ディスプレイを用いるDAISYプレーヤーの開発、音声研究会・福祉情報工学研究会共催研究会、2012年9月28日、芝浦工業大学豊洲キャンパス
- ② 南谷和範、視覚障害者への漢字説明と大学入試センター試験点字冊子問題への字注付与状況 一定量の研究の試み、第7回全国大学入学者選抜研究連絡協議会、2012年6月1日、岡山コンベンションセンター
- ③ 南谷和範、重度視覚障害者の漢字使用技能の効率的習得を実現するための基礎的考察、日本特殊教育学会(第49回大会)、2011年9月23日、弘前大学

- ④ 南谷和範、工藤智行、試験問題を含む高度な文書の構造提示を点字ディスプレイで実現するリーダーソフトウェアの開発、電子情報通信学会技術研究報告(福祉情報工学(WIT))、2011年8月4日、京都大学杉浦ホール
- ⑤ 南谷和範、工藤智行、コンピュータによる試験問題の音声出題を実現するリーダーの開発、研究集会「科学情報の電子化・自動処理・アクセシビリティをめぐる諸問題」、2011年2月11日、筑波技術大学(茨城県)
- ⑥ 南谷和範、藤芳衛、青松利明、原田清生、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、2010年9月18日、金城学院大学(愛知県)
- ⑦ 澤崎陽彦、南谷和範、藤芳衛、青松利明、印刷物読書困難者のための文字と音声のマルチモーダル問題、日本教育工学会第26回全国大会、2010年9月18日、金城学院大学(愛知県)
- ⑧ 藤芳衛、南谷和範、青松利明、原田清生、重度弱視者及び読字障害者のための文字と音声のマルチモーダル問題の開発、平成22年度全国大学入学者選抜研究連絡協議会大会、2010年6月9日、北九州国際会議場(福岡県)

[図書] (計0件)

[産業財産権]

○出願状況 (計0件)

名称：
発明者：
権利者：
種類：
番号：
出願年月日：
国内外の別：

○取得状況 (計0件)

名称：
発明者：
権利者：
種類：
番号：
取得年月日：
国内外の別：

[その他]

ホームページ等

6. 研究組織

(1)研究代表者

南谷 和範 (MINATANI KAZUNORI)
独立法人大学入試センター・入学者選抜研究
機構・准教授
研究者番号：22700835

(2)研究分担者

研究者番号：

(3)連携研究者

研究者番号：