

点訳機能を有する講演原稿受付システムによる 手頃な点字資料作成方式の検討

A Reasonable Generation of Braille Documents by a Manuscript-receiving System with a Built-in Braille Transcriber

安念克洋[†] 稲川友宏^{†,‡} 岡本明^{††,‡} 河井良浩^{†††,‡} 坂根裕[†] 竹林洋一[†]
Katsuhiro Annen[†], Tomohiro Haraikawa^{†,‡}, Akira Okamoto^{††,‡}, Yoshihiro Kawai^{†††,‡},
Yutaka Sakane[†] and Yoichi Takebayashi[†]

1. はじめに

研究発表会等への視覚障害者参加の補助となる資料の点訳は、点訳者や印刷の手配、費用負担の困難から、一部の研究発表会で限定的に行われているにすぎない。本発表では、発表申込システムに機械点訳を組み込み、予稿回収と同時に全文の点字粗訳を自動生成する手法を、現行著作権法や工数の制約を考慮した1つの現実解として提案する。

当該システムは、“手頃な情報保障”を目的とした自動システムであるが、任意校正の機会をかなへ変換後の読み原稿ファイル、点訳後の点字ファイルの2つの時点において設けることで、情報保障への関心がある発表者は点字に明るくなくとも協力できるよう設計した。また、将来的な学会および著者の了解のもとテキストファイルも配布できるように配慮した。本稿では、このシステムの具体的な構成と、運用に際して生じた諸問題について報告する。

2. 現行制度内の現実的なデータ提供方法

本稿は、これまで点訳者による点訳と点字印刷物の会場配布などを行ってきた研究会等における情報保障を縮小し、あるいは電子化で置き換えようとする試みではなく、それらの継続に加えて、なんらかの情報保障が得られる範囲をより広い研究分野に広げていこうという提案である。

これまで情報保障が行われてこなかった学会・研究会に対して情報保障の裾野を広げていこうと考えると、講演をすべて点字印刷して準備することが困難なため、必然的に電子データでの事前配布を選択せざるを得ない。

2.1 電子データの形式に関する検討

ところで、視覚障害者への情報保障というと真っ先に点字が思い浮かぶ。が、会場で印刷物として入手するのではなければ、ピンドィスプレイや点字電子手帳の所持者以外には意味を持たない点字データは扱いにくい。電子データになってしまっても相変わらず点字のみを考えるのは情報保障のやり方としては筋が悪く、(講演者の音声との同時聴取にやや難はあるが) PC単体でも合成音声として読めるテキストデータを提供可能なシステムを設計しなければならない。

しかしながら、情報保障への理解が充分でない現時点では、副次的利用の虞からテキストデータの配布に主催者の協力が得られないケースがままあり、点字データによる提供に実施の容易性で分がある。公表された出版物の点字データの複製や公衆送信は著作権法第37条でそもそも認められており、主催者の了解を直ちに得やすいからである。

以上の検討から、広範囲の情報保障を浸透させるためのデータ提供方法としては、機械点訳により生成した点字データをオンライン提供し、テキストデータの提供については、主催者の現在の事情により選択可能としたうえで、将来的には配布の方向で理解を求めていくのが、即実性のある案として現状では最善であると考える。

3. 点訳機能つき講演原稿受付システムの設計

以下、本稿では、講演原稿受付システムにテキスト抽出機能と機械点訳機能を組み込み、ワープロ文書として入稿された予稿をテキストデータおよび点字データとしてダウンロード可能にするシステムについてその実装例を示す。

講演原稿受付システムと一体化する形で設計した理由は主に4点ある。(1) 主催者側に入稿管理コスト低減を訴求して導入促進をはかること。(2) 著者の講演原稿入稿と同時に点訳を生成することで、(誤読・誤訳は含まれるが)点字資料の配布という最低限の情報保障を実現できること。(3) 入稿フォーム上で即時に点訳を生成することで、著者の協力によるインタラクティブな訳質向上が期待できること。(4) 講演登録フォームの内容を点訳に生かせることである。(1), (2)は自明であるので、残る問題点について述べたあと、(3), (4)について説明する。

3.1 問題の分析

機械点訳では図表や数式の扱いと誤訳が問題になり、図表についてはテキストファイルや点字ファイルでは解決が困難である。しかし、点字ファイルを生成しておくことで、点字使用者自身が完訳すべきデータかどうかを判断したり、点訳者に下訳として自由に渡したりすることができる。この点でも、著作権法第37条で複製等が保護されている点字ファイルは、現時点での取り扱いは断然容易である。

誤訳は、とくに著者名をはじめ難しい読みの固有名詞などの誤読により多く発生する。これらは点訳者が主催者に照会するなどして対応できていた部分であるが、機械点訳では正しく訳出することが難しい。これは訳出時に読みを確定しなければならない点字ならではの問題である(図1)。

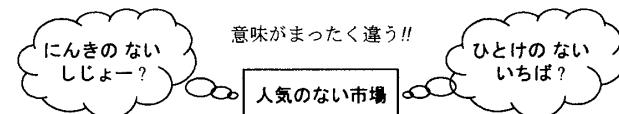


図1. どう読む?「人気のない市場」

原稿の著者はもちろん、原稿に現れる漢字の読みをすべて知っている。しかし、点字エディタによる校正まで著者に期待することには無理がある。点字や点字エディタに関わりのない著者でも、情報保障への関心がありさえすれば

[†]静岡大学 [‡]筑波技術短期大学 ^{††}産業技術総合研究所

[‡]電子情報通信学会 HCG AMAI ワーキンググループ

誤訳低減に協力できるような仕掛けが必要である。また、著者に関心がまったくない場合でも、タイトルや著者名などの最低限の重要事項は正しい読みで訳出する必要がある。

3.1 システムの設計と実装

この2つの問題に対し、以下のように対応した。

(3) インタラクティブな訳質向上: 講演者に点字ファイルの誤訳を直接修正してもらうのではなく、中間ファイルとして「読み原稿ファイル」を生成することにした。すなわち、講演原稿を入稿すると同時にテキストファイル、読み原稿ファイル、点字ファイルの3つをその場で同時に自動生成する(図2)。読み原稿ファイルはひらがなと英数字からなるテキストファイルであり、点字や点字エディタに詳しくなくとも修正が可能である。これに任意で修正を施し再登録すると、点字ファイルが再作成されるようになる。これを講演原稿受付システム上でリアルタイムに行えるようにし、著者と協力してのインタラクティブな訳質向上を狙う。

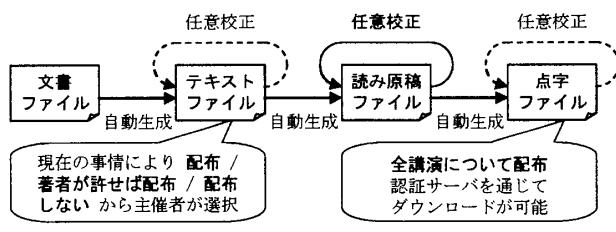


図2. 点訳機能つき講演原稿受付システムの流れ

(4) 講演登録フォームの活用: 講演者に情報保障への関心がなくとも、データは最低限の正確さを保証しなければならない。そこで、重要なデータである著者名等については、たとえ任意校正がまったく行われなくとも、講演申込時にwebフォームで取得したタイトルおよび著者名のふりがなをもとに正しい読みの表紙を生成し、自動結合する。

システムはVine Linux 2.6上にApache+Perlで実装した。入稿形式としては手はじめにMicrosoft Word文書としている。プレインテキストへの変換は、wvWare^[1]を用い、プレインテキストから読み原稿ファイル、読み原稿ファイル

から点字ファイルへの変換は、日本語英語自動点訳ソフトウェアEXTRA^[2]を用いた。今回EXTRAはLinux版のものが調達できず、dosemu+freedesuによりLinux上に仮想MS-DOSマシンを構築し、フリーのMS-DOS版EXTRAを実験的に試用させていただいている。実際の動作のようすを図3に示す。

4. まとめ

どこに出かけても最低限の情報保障が得られる“ユビキタスサービスとしての情報保障”を広く実現するための“手頃な情報保障”的構想と、それに基づく点字資料作成方式の概要について報告した。

“手頃な情報保障”を目的とした当該システムの限界は、a) 図表や数式の扱い、b) レイアウトの扱い、c) 誤訳の扱いの3点ある。これらについて、b)についてはワープロ文書のスタイル解析で章・節の行空けや字下げを正しく処理しようと改良を試みており、c)については(3), (4)に述べた解決法で一応の目的を達成した。

本システムの一番の特長は、講演原稿受付システムと一体化することで読みファイルのインタラクティブな任意校正の機会を設け、点字に詳しくない講演者でも誤訳の低減に貢献できるようにしたことである。これは誰でも容易に思いつくことであり、特に奇抜なアイディアでもない。

福祉社会の前提をなす“ユビキタスサービスとしての情報保障”的には、情報保障に关心があればだれでも貢献できるような仕掛けが必要である。ユビキタスサービスによる福祉社会実現のために、より“手頃な”情報保障方式が多数提案され、優れたものが発展していくことを期待したい。

参考文献

- [1] wvWare, library for converting Word documents
<http://wvware.sourceforge.net/>
- [2] 石川准: 日本語英語自動点訳プログラムEXTRA for DOS Ver.3.72 DOS/V版, <http://fiji.u-shizouka-ken.ac.jp/~ishikawa/devpgm.htm>

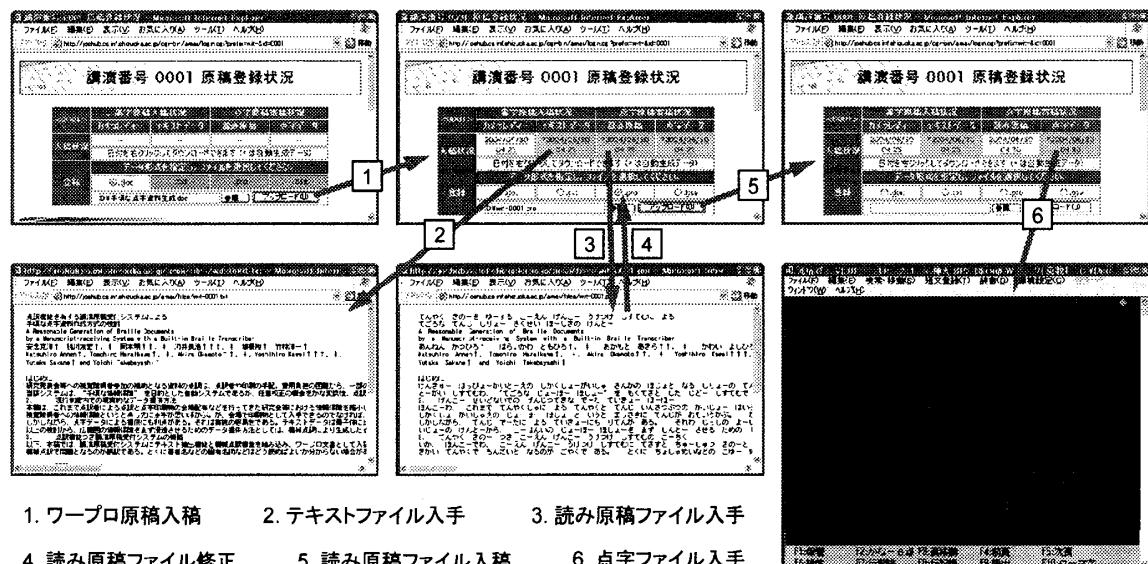


図3. 提案システムによる入稿および任意校正のようす