

講義場面におけるテレビ会議装置を用いたキーボードの 連弾入力方式によるリアルタイム字幕提示システム

教育方法開発センター（聴覚障害系）小林 正 幸
教育方法開発センター（聴覚障害系）石 原 保 志
一般教育等（聴覚障害系）根 本 匡 文

客員研究員 西 川 俊
教育方法開発センター（聴覚障害系）内 野 権 次
筑波大学 奥 山 敏 雄

要旨：我々は、DOS/VパソコンのWindows NTに対応した日本語高速入力システム（Windows対応ステノワードPCシステム）と校正器を開発し、これらと字幕挿入装置を連動させ、1セット目のステノワードPCシステムで発話の内容を先ずひらがなで入力、表示し、その後かな・漢字変換を行い、かな漢字混じり文を提示し、2セット目のステノワードPCシステムで、1セット目で入力した確定文章の誤字、脱字の修正を行い、より正確な字幕を話者の映像と共にリアルタイムで提示する新型のWindows対応連弾入力方式RSVシステムを構築した。また、このシステムとテレビ会議装置を連動させ、電話回線のISDNを介して遠隔地で話者の音声を聞きながらキーボードの入力作業が可能な遠隔地Windows対応連弾入力方式RSVシステムを開発したので、このシステムの機能、特徴等について報告する。また、このシステムを本学聴覚部の非常勤講師が担当している講義場面で活用し、講義終了後、学生を対象とした質問紙による調査を行った。その調査結果についても述べる。

キーワード：リアルタイム、字幕、テレビ会議装置、ISDN、情報保障

1. はじめに

聴覚障害者を対象とする3年制の国立の筑波技術短期大学聴覚部一般講義室には、高速で文字が入力できるステノワードPCキーボードとパソコンを活用した日本語高速入力システム（ステノワードPCシステム）2セットと校正器、及び字幕挿入装置（VIP・4100R）により、1セット目のステノワードPCシステムで発話の内容を先ずひらがなで入力、表示し、その後かな・漢字変換を行い、かな漢字混じり文を提示し、2セット目のステノワードPCシステムで、1セット目で入力した確定文章の誤字、脱字の修正を行い、より正確な字幕を話者の映像と共にリアルタイムで提示する連弾入力方式RSV（Real-Time Captioning System by Stenoword and Video Information Processors）システム[1]が設置されているが、この旧型のシステムはMS・DOSの管理下で作動し、最新の日本語変換ソフトの利用や、パソコンからシリアルデータを伝送するRS・232Cパラメータ設定の変更ができない等の欠点があった。

そこで、我々はWindows NT Ver.4.0（又はWindows95）の環境下で、かな・漢字変換の精度が向上した日本語変換システム「ATOK 11」や市販されている豊富なソフトが利用でき、また様々な機器と接続できるようにRS・232Cのパラメータ設定の変更が可能なDOS/Vパソコンで動作するWindows対応ステノワードPCシステムと校正器を開発し、これらと字幕挿入装置を連動させ、新型のWindows対応連弾入力方式RSVシステムを構築した。また、このシステムとテレビ会議装置（PICSEND・RⅡ）

を連動させ、電話回線のISDN（INSネット64）を介して遠隔地で話者の音声を聞きながらキーボードの入力作業が可能な遠隔地Windows対応連弾入力方式RSVシステム[2]を開発した。

本論文では、この新システム（遠隔地Windows対応連弾入力方式RSVシステム）の機能、特徴等について報告する。また、この新システムを本学聴覚部の非常勤講師が担当している講義場面で活用し、講義終了後、学生を対象とした質問紙による調査を行った。その調査結果についても述べる。

2. Windows対応連弾入力方式RSVシステム

このシステムはWindows対応ステノワードPCシステム、校正器及び字幕挿入装置から構成されている。

2.1 Windows対応ステノワードPCシステムと字幕挿入装置

Windows対応ステノワードPCシステムと字幕挿入装置（VIP・4100R）の機能、特徴は次の通りである。

- （1）DOS/VパソコンのWindows NT、及びWindows 95に対応
- （2）RS・232Cパラメータの設定、変更が可能
- （3）一般のJISキーボードでも入力可能：ファンクションキーのF12を「デリミタ」キーに割付
- （4）日本語フロントエンドプロセッサ：ジャストシステム製のATOK 11
- （5）かな・漢字変換のモード：「スペース／句点」付き変換、「読点／句点」付き変換、「読点／句点」を「ス

ペース」付きで変換、「句点」を「改行」付きで変換

(6)「読み」の提示モード：「あり/なし」

(7)文字色：8色

(8)エッジ(文字の縁取り)の付加

(9)カーソル移動

(10)文章の編集

(11)VIP・4100Rとの接続：「あり/なし」

(12)文書保存

2.2 ステノワードPCキーボード

このキーボードの機能は次の通りである。また、このキーボードのキー配列を図1に示す。

(1)入力方法

図1の複数のキーを同時に押して、キーの組み合わせによって文字列を入力する。

(2)入力の種類

常用句入力

普段よく使用する単語を、図1の数個のキーを同時に押して入力する。例えば、「かんがえる」という単語の

入力は、JISキーボードではかな入力モードでキーを6回入力するが、ステノワードPCキーボードでは図1の「あ」「え」「い」「お」キーを同時に押すという1回のキー操作で入力する。

二音入力

「おい」等のような二音(2文字)を、図1の数個のキーを同時に押して入力する。例えば、「おい」は図1の「お」「い」キーを同時に押す。

単音入力

「あ」等のような単音(1文字)を入力する。例えば、「あ」は図1の「あ」キーを押す。

2.3 校正器

2台のキーボードと2台のパソコンとの間のデータの流れを制御する校正器の機能は次の通りである。

(1)DOS/VパソコンのWindows NTに対応

(2)キーボードのキーデータを一時的に蓄積するバッファメモリを内蔵(8Kバイトの容量)

B S	わ「A	ら(S	な-D	まR	やT	Y	U	I	O	P	
	濁Z	たX	はC	かV	さB	うH	あJ	おK	いL	-え	改行
	1	2	3	4	5	6	7	8	9	0	

、		。
S P A C E	無変換	変換
S P A C E		

図1 ステノワードPCキーボードのキー配列

3. 講義場面における遠隔地Windows対応連弾入力方式RSVシステム

このシステムの構成を図2に、講義室内に設置してあるシステムの写真を図3に示す。

このシステムの動作は、次の通りである。

(1)講義室で講義をしている非常勤講師の音声は、ワイヤレスピンマイク、ワイヤレスアンプ、テレビ会議装置(PICSEND・RⅡ)、ターミナルアダプタ、INSネット64を介して遠隔地(東京新宿のスピードワープ研究所)へ送出される。

(2)遠隔地にいる入力担当オペレータは(1)で送出された音声をスピーカで聞きながら、高速で文字が入力できる文字入力用Windows対応ステノワードPCシステムにリアルタイムで入力する。

(3)入力された文字は2台のキーボードと2台のパソコンとの間のデータの流れを制御する校正器を通り、文

字入力用パソコン(FMV-DESKPOWERMV205)と文字修正用パソコン(FMV-DESKPOWERTV307)のキーボード端子に送出され、それぞれのパソコンでかな・漢字変換される。

(4)修正担当オペレータは、文字修正用パソコンでかな・漢字変換された文章の誤字、脱字を判断した後、修正作業を行うかどうかを決める。

(a)修正作業をしない場合(誤字、脱字がない場合)

文字修正用パソコンは、このパソコンでかな・漢字変換した文章をRS-232Cポートから9600BPSのスピードで、文字コードとして、テレビ会議装置、ターミナルアダプタ、INSネット64を介して講義室の字幕挿入装置(VIP・4100R)へ送出する。

(b)修正作業をする場合(誤字、脱字がある場合)

修正担当オペレータは、文字修正用ステノワードPCキーボード(以下、文字修正用キーボードと略す。)の

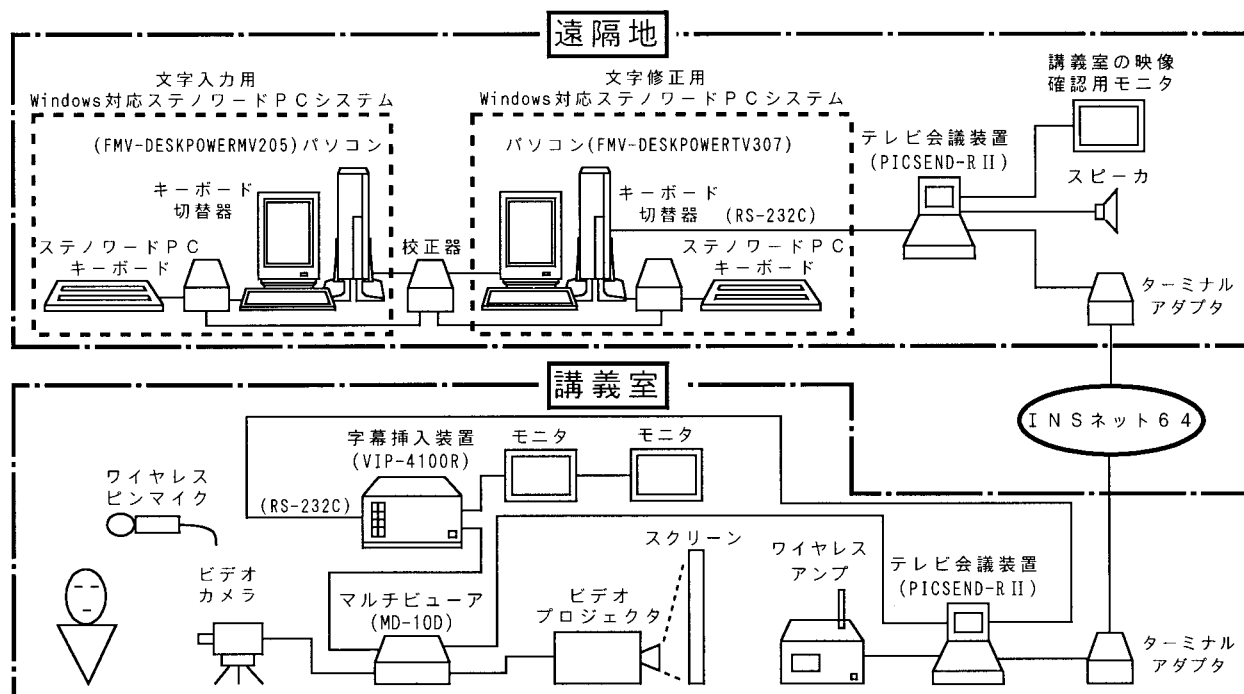


図2 システムの構成

「デリミタ」キーを押す。

校正器は、文字入力用パソコンで入力された文字コードを文字修正用パソコンへ送出することを一時停止する。

修正担当オペレータは、文字修正用Windows対応ステノワードPCシステムで誤字、脱字の修正を行う。修正された文字は、文字修正用パソコンのRS・232Cポートから9600BPSのスピードで、文字コードとして、テレビ会議装置、ターミナルアダプタ、INSネット64を介して講義室の字幕挿入装置へ送出する。

校正器は、修正担当オペレータの修正作業中に、入力担当オペレータが入力した文字コードを、バッファメモリに蓄積すると同時に文字入力用パソコンへ送出する。

文字入力用パソコンは、校正器が送出した文字コードをかな・漢字変換する。

修正担当オペレータは、修正作業終了後、文字修正用キーボードの「デリミタ」キーを押す。

校正器は、前記で蓄積したバッファメモリの文字コード（修正担当オペレータの修正作業中に、入力担当オペレータが入力した文字コード）を、入力担当オペレータが入力した文字入力速度と無関係に一定の送信速度で文字修正用パソコンへ送出する。

文字修正用パソコンは、前記の文字コード（修正担当オペレータの修正作業中に、入力担当オペレータが入力した文字コード）を受信し、かな・漢字変換を行い、

RS・232Cポートから9600BPSのスピードで、文字コードとして、テレビ会議装置、ターミナルアダプタ、INSネット64を介して講義室の字幕挿入装置へ送出する。（5）講義室に設置してある字幕挿入装置（VIP・4100R）は、前記（4）の文字コードを映像なしの字幕として、2台のモニターへ提示する。また、マルチビューア（MD・10D）は、ビデオカメラで撮影した非常勤講師（黒板の板書）の映像とモニターへ提示された字幕映像を、それぞれ縦方向へ1/2に圧縮した後、上下へ1画面に合成し、ビデオプロジェクタを介してスクリーンへ提示するとともに、テレビ会議装置、ターミナルアダプタ、INSネット64を介して、遠隔地に設置してある講義室の映像確認用モニターへ提示され、講義場面の把握やシステム全体が正常に動作しているかどうかを確認するために使用される。

4. 講義場面での活用

4.1 講義の方法

学外の非常勤講師が担当している講義は、口話と板書、及び講義資料のOHP提示で行い、20～30分程度で講義を区切り、学生はその間の板書等の講義内容をノートに記述する。

4.2 支援方法

遠隔地の東京新宿のスピードワープ研究所におい



図3 講義室内のシステム

て、聴覚部の講義室で講義をしている非常勤講師の音声聞きながらキー入力するオペレータ2名は、30分毎に入力作業と修正作業を交互に行った。

4.3 字幕提示形式

字幕は、かな・漢字変換する文章の「読み」をまず表示し、その後「、」「。」を入力することで、かな・漢字変換、確定を行い、入力した「、」「。」を自動的に空白文字として提示（分かち書き提示）するモードに設定した。また、字幕は要約せずに、できるだけ忠実に提示した。

図2の講義室内のモニター2台への字幕提示形式は次の通りであり、図4にモニターの字幕提示画面の写真を示す。

- (1) 表示行数：8行/画面
- (2) 表示文字数：20文字/行
- (3) 「読み」の文字色：黄
- (4) 「確定」の文字色：白
- (5) エッジ(文字の縁取り)：黒色の太いエッジを付加
- (6) 映像：なし

また、図2の講義室内のスクリーンへの字幕提示形式は、次の通りであり、図5にスクリーンの字幕提示画面の写真を示す。

- (1) 提示方法
非常勤講師（黒板の板書）の映像とモニターの字幕映像を、それぞれ縦方向へ1/2に圧縮した後、上下へ1画面に合成して提示
- (2) 表示行数：8行/スクリーン下半分
- (3) 表示文字数：20文字/行
- (4) 「読み」の文字色：黄

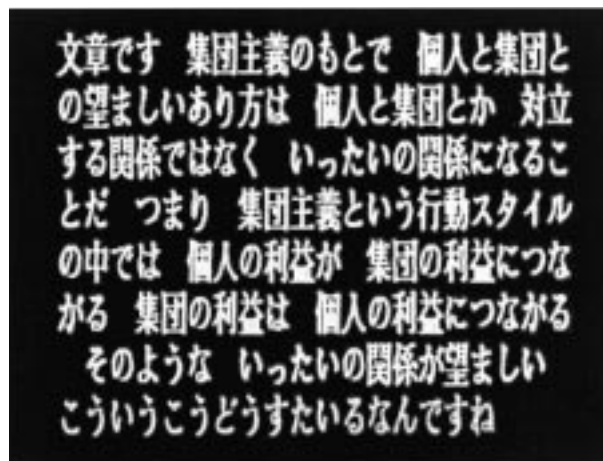


図4 モニターの字幕提示画面

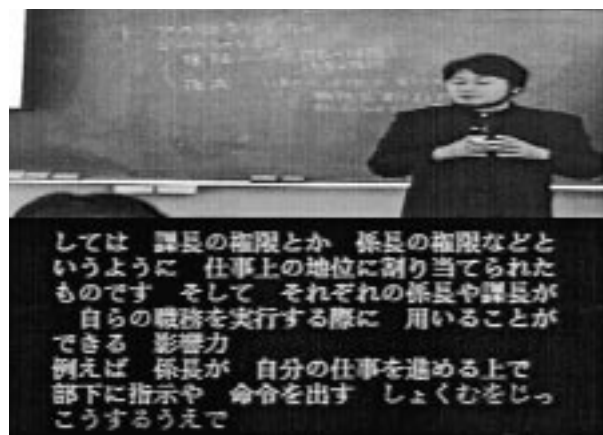


図5 スクリーンの字幕提示画面

- (5) 「確定」の文字色：白
- (6) エッジ(文字の縁取り)：黒色の太いエッジを付加

5. 質問紙調査

本学聴覚部の学外講師が担当している一般教育科目「総合Ⅱ(社会)」の平成10年度3学期の講義9回について、遠隔地Windows対応連弾入力方式RSVシステムによる情報保障を実施した。この情報保障を実施した3回目の講義で、リアルタイム字幕提示に関する質問紙調査を、講義終了後実施した。

- (1) 調査の内容
講義の内容を理解する上で、字幕がどの程度役に立ったかに関する意識(多肢選択)
字幕文章の誤りに関する意識(多肢選択)
字幕文章の理解度に関する意識(多肢選択)
字幕提示方法に関する意識(多肢選択)
字幕に関する意見(記述)

(2) 対象者

対象者は本学聴覚部の機械工学科、建築工学科、電子情報学科電子工学専攻、電子情報学科情報工学専攻の、本講義を受講した1学年34名の学生で全員重度及び最重度の聴覚障害者である。

6. 結果と考察

(1) 字幕への依存度

図6は、「講義の内容を理解する上で、字幕に頼りましたか。1つに をつけて下さい。 たいへん頼った 少し頼った あまり頼らなかった 全く頼らなかった」という質問に対する回答を集計したものである。「たいへん頼った」と回答した学生は34名中23名(67.6%)であり、60%以上の学生が講義内容を理解する上で字幕に依存しており、字幕が講義内容の理解で主要な情報受容メディアとして利用されていたことが明らかにされた。また、自由記述欄にも「字幕が無いと講義の内容が全く分からないと思う。」「先生の話が分かりにくかったが、字幕があって分かりやすかった。」「助かる。講義の流れ、内容、先生の言いたいことがわかる。」といった好意的な回答が見られた。

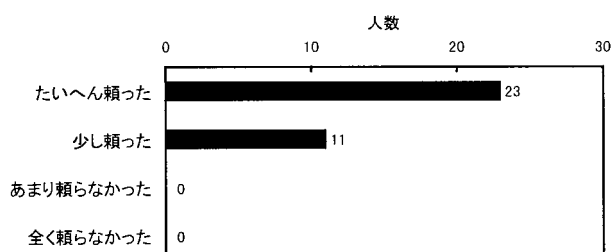


図6 字幕への依存度

(2) 字幕の誤変換

表1は、学外講師の音声と遠隔地Windows対応連弾入力方式RSVシステムで入力した字幕を録画したビデオテープから無作為に3分間を抽出し、発話内容と提示された字幕を比較照合して、発話内容がどれだけ正確に文節単位で字幕に変換されたかを表す正変換率を算出したものである。スピーチの文節数は発話した文節の総数、スピーチのスピードは1分間に発話した文節数、入力時誤変換文節数は文字入力用ステノワードPCシステムで誤変換された文節の総数、修正後誤変換文節数は文字入力用ステノワードPCシステムで誤変換された文節を文字修正用ステノワードPCシステムで修正した後の誤変換文節の総数、誤変換修正文節数は文字修正用ステノワードPCシステムでの誤変換文節の修正を表す文節数(入力時誤変換文節数 - 修正後誤変換文節数)、入力時正変換率は文字入力用ステノワードPCシステムで入力された字幕の正変換率、修正後正変換率は文字修正用ステノワードPCシステムで修正した後の最終的な正変換率を表したものである。

表1において、スピーチのスピードに関わらずサンプルD以外のサンプルA、B、Cについては入力時正変換率が97%程度以上であり、修正後正変換率についてはサンプルD以外は100%であった。また、誤変換修正文節数は1から3文節であり、修正後正変換率は修正前に対して、0.8~2.3%まで向上しており、講義全体では誤字、脱字の総文節数は23文節であった。

サンプルDが他のサンプルと比較して修正後正変換率が6%程度低下した要因は、日本的集団主義についての講義で学外講師が発話した「自律性」という単語を、オペレータが「自立性」と誤認識し、誤変換の修正作業を行わなかったことによるものである。この誤認識した「自立性」の間違いは修正後誤変換文節総数7文節中6文節であった。

表1 講義場面における字幕の正変換率(3分間を抽出)

サ ャ プ ル	学 外 講 師			
	A	B	C	D
スピーチの文節数(文節)	137	149	121	132
スピーチのスピード(文節/分)	45.7	49.7	40.3	44.0
入力時誤変換文節数(文節)	3	3	1	10
修正後誤変換文節数(文節)	0	0	0	7
誤変換修正文節数(文節)	3	3	1	3
入力時正変換率(%)	97.8	98.0	99.2	92.4
修正後正変換率(%)	100.0	100.0	100.0	94.7

図7は、表1のデータから、スピーチのスピードと入力時正変換率、及び修正後正変換率を表したグラフである。

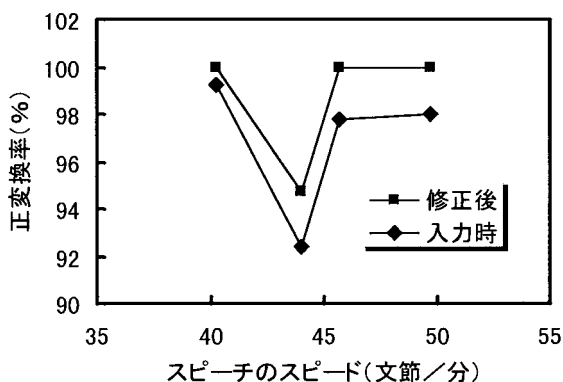


図7 スピーチのスピードと正変換率

(3) 字幕の誤りに対する評価と考察

図8は、「字幕の中に時々『誤り』がありました。この『誤り』は、講義の内容を理解する上で困りましたか。1つに をつけて下さい。たいへん困った 少し困った あまり困らなかった 全く困らなかった」という質問に対する回答を集計したものである。「あまり困らなかった」と「全く困らなかった」とした学生は、34名中21名(61.8%)であった。

表1において修正後正変換率が94%程度以上、また、講義全体では誤字、脱字の総文節数が23文節と非常に少ないにも関わらず、40%程度の学生が「たいへん困った」、「少し困った」と回答した理由については、次のようなことが言える。前述のように講義のキーワードである「自律性」を「自立性」と誤変換したことで、板書の内容と字幕が異なり学生が混乱したことによる。これは自由記述欄の回答である「自立性と自律性、似ていてとまどった。」からも推察できる。システムの不具合により12分間程度、最下行の字幕が発話内容と無関係な字幕が提示された。

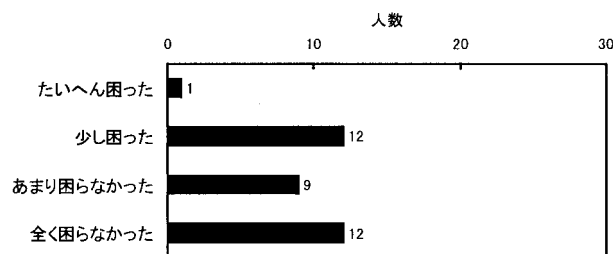


図8 字幕の誤りに関する意識

(4) 字幕の提示方法

図9は、「黒板に向かって右側にあるOHPスクリーン

には、上半分に先生(黒板)の映像、下半分に字幕が同時に提示されていました。テレビ画面に表示されている字幕のみの提示方法と比較して、どちらの提示方法が良いでしょうか。1つに をつけて下さい。先生の映像と字幕(OHPスクリーン) 字幕のみ(テレビ画面)」という質問に対する回答を集計したものである。「先生の映像と字幕」とした学生は34名中7名(20.6%)、「字幕のみ」は26名(76.5%)、「未回答」は1名(2.9%)であり、「字幕のみ」とした学生は、「先生の映像と字幕」と比較して56%程度上回ったが、20%程度の学生が「先生の映像と字幕」の方が良いと回答しており、学生個々に応じた字幕提示方法の選択肢が広がることを意味している。

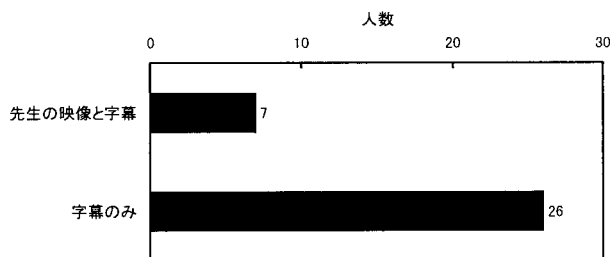


図9 字幕の提示方法に関する意識

(5) 字幕の文章

図10は、字幕文章の難易度に関する質問の回答を集計したものである。ほとんどの学生は、「よく理解できる」、「だいたい理解できる」としているが、「あまり理解できない」とした学生も2名いる。リアルタイム字幕提示は、基本的に、音声は全て文字化すべきであるという考えを持っているが、話者によってはスピーチのスピードに対して文字入力追従できない場合もあり、この場合の要約、意識の方法について、今後検討していく必要がある。

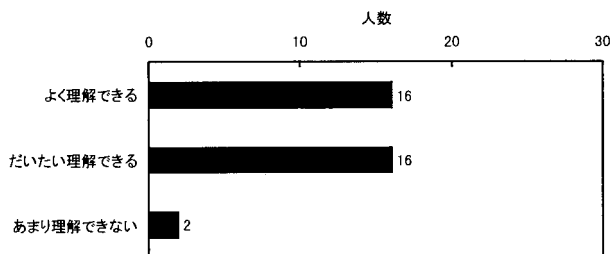


図10 字幕文章の難易度に関する意識

7. おわりに

今後は、校正器や字幕挿入装置(VIP-4100R)のような特殊な機器で構成するシステムではなく、一般的な機

器を用いた新規のキーボードの連弾入力方式によるリアルタイム字幕提示システムを開発する予定である。

参考文献

- [1] 小林正幸、西川俊、石原保志、高橋秀知：聴覚障害者のためのキーボードの連弾入力方式によるリアルタイム字幕提示システム、映像情報メディア学会誌、Vol.51, No.6, pp.886-895 (Jun.1997)
- [2] 小林正幸、石原保志、西川俊：聴覚障害者のための遠隔地でのキーボードの連弾入力によるリアルタイム字幕提示システム、ろう教育科学、Vol.40, No.3, pp.121-130 (Oct.1998)
- [3] 西川俊、高橋秀知、小林正幸、石原保志、柴田邦博：聴覚障害者のためのリアルタイム字幕提示システム、電子情報通信学会論文誌 D-II、Vol.J78-D-II, No.11, pp.1589-1597 (Nov.1995)
- [4] 小林正幸、西川俊、石原保志、高橋秀知：聴覚障害学生のためのリアルタイム字幕提示システム(3)、電子情報通信学会技術研究報告、ET97-74, Vol.97, No.397, pp.25-29 (Nov.1997)

Real-Time Captioning System Using Two Stenographic Systems And TV Conference Devices in the lecture

Masayuki Kobayashi (Research Center on Educational Media, Division for the Hearing Impaired)

Satoshi Nishikawa (Guest Researcher, Division for the Hearing Impaired)

Yasushi Ishihara (Research Center on Educational Media, Division for the Hearing Impaired)

Kenji Uchino (Research Center on Educational Media, Division for the Hearing Impaired)

Masafumi Nemoto (General Education, Division for the Hearing Impaired)

Toshio Okuyama (University of Tsukuba)

We have developed this system which can offer more accurate Japanese captioning than ever consisting of two stenographic systems (STENOWORD-PC SYSTEM) and a superimposer (VIP-4100R). One STENOWORD-PC SYSTEM operates for inputting Japanese words by speakers in Kana (Japanese character) and converting these Kana to Kanji (Chinese character). Another STENOWORD-PC SYSTEM operates for correcting errors with wrong or omitted letters. These systems are connected with TV conference devices, which make it possible to input speeches of speakers from a long distance through ISDN telephone network. The system configurations, functions and some problems to be solved and their solution are presented.

Key words: real-time, caption, TV conference device, ISDN, information support service