

談話標識と話題語を用いた重要文抽出手法のCSJの学会講演における評価

北出 祐 南條 浩輝 河原 達也

京都大学大学院 情報学研究科 知能情報学専攻

〒606-8501 京都市左京区吉田二本松町

e-mail: {kitade, nanjo}@ar.media.kyoto-u.ac.jp, kawahara@i.kyoto-u.ac.jp

あらまし 談話標識と話題語の統計量を用いた重要文抽出手法を日本語話し言葉コーパス(CSJ)の学会講演に適用した結果を報告する。CSJの一部の講演には全文の50%および10%を目安に重要文が人手によりタグ付与されている。作業者間のタグの不一致も考慮して4通りの正解セットを作成し、人間の抽出精度と提案手法を比較した。その結果、談話標識に基づく手法の有効性、および話題語の統計量と統合する相乗効果を確認した。また音声認識結果に対しても頑健であることが示された。人間の抽出精度との比較では、10%抽出では大きな差があったが、50%抽出では約10%の低下にとどまった。

キーワード 講演音声, 重要文抽出, 談話標識, 教師なし学習

Automatic Extraction of Key Sentences from CSJ Presentations using Discourse Markers and Topic Words

Tasuku Kitade Hiroaki Nanjo Tatsuya Kawahara

Graduate School of Informatics, Kyoto University

Yoshida-Nihonmatsu-cho, Sakyo-ku, Kyoto 606-8501 Japan

e-mail: {kitade, nanjo}@ar.media.kyoto-u.ac.jp, kawahara@i.kyoto-u.ac.jp

Abstract Automatic extraction of key sentences from the CSJ lecture audio archives is addressed for the purpose of generating indexing and summary. The proposed method focuses on characteristic expressions used in initial utterances of sections, which are defined as discourse markers and derived in an unsupervised manner based on word statistics. The statistics of the presumed discourse markers are then used to define the importance of the sentences. It is also combined with the conventional tf-idf measure of topic words. The proposed method is evaluated with 19 oral presentations of the CSJ (Corpus of Spontaneous Japanese). Experimental results show that the combination of statistics of discourse markers and topic words are effective, and the method is robust against ASR errors. Comparison with human performance is also reported.

key words spontaneous presentation speech, automatic key sentence extraction, discourse marker, unsupervised training

1 緒論

近年の計算機性能の向上やメディア処理技術の進展に伴い、音声や映像などのマルチメディアのデジタルアーカイブを構築することが可能になってきている。しかし、音声アーカイブはテキストと異なり、そのままでは目的とする情報を得ることや短時間で全体の内容を把握することが非常に困難である。そのため、インデックスや要約などの二次情報をアーカイブに付与することが必要不可欠である。これらの情報を人手で付与するには膨大なコストを要するため、自動的に付与する技術が望まれている。

テキストを対象とした重要文抽出や要約の研究は、これまで数多く行われている。新聞記事や論文を対象にした研究においては、セクション構造を用いた位置情報が有効であることが知られている。これは、それぞれの記事やセクション、パラグラフの冒頭にその話題を説明する重要文が記されていることが多いためである。そのようにして集められた重要文をもとに要約を作成することができる。しかし講演などの音声ではパラグラフやセクション境界が明確ではなく、このようなアプローチをそのまま適用することは困難であった。これに対して、我々はセクション冒頭に頻出する談話標識の情報を用いてセクション境界を推定することで、このアプローチを(講演)音声に適用する方法を提案した [1]。

日本語の話し言葉を対象とした重要文抽出の研究は他の研究機関でも行われているが [2][3][4]、評価データや評価方法が一定ではなく、比較が困難であった。これに対して「話し言葉工学」プロジェクトにより整備が進められている「日本語話し言葉コーパス(CSJ)」[5][6]では、その一部に対して重要文の抽出作業が行われており、3名による重要文タグが付与されている。そこで本研究では、この重要文セットに対して、我々の提案する重要文の自動抽出手法 [1] を適用する。特に、作業員間のタグの不一致も考慮して、人間の重要文選択の精度と比較できる枠組みを設定し、評価を行う。

2 コーパスと重要文抽出処理の概要

2.1 日本語話し言葉コーパス(CSJ)における重要文セット

CSJは主に学会講演と模擬講演からなるコーパスであり、音声データと人手による書き起こしテキストから構成される。CSJの講演のうち、コアとよばれる一部の講演には複数の作業員による重要文タグが付与されている。各講演には3名の作業員が割り当てられており、全文の50%、および10%を目安に重要文を選択している。作業員は主に言語学の研究者であり、学会講演の発表スタイルに精通しているが、必ずしもそれぞれの講演の分野の専門家とは限らない。

本稿では、2003年8月時点で重要文タグが付与されたCSJの学会講演19講演を用いて実験・評価を行う。

2.2 処理の概要

重要文抽出の処理の概要を図1に示す。

まず、講演の音声を認識してテキスト(音声認識結果)を生成する。次に、話し言葉では文の単位が明示的でないため、統計的な手法を用いてテキストを文に分割する。さらに分割された各文に対して重要度を求める。最後に各文の重要度に基づいて、指定した抽出率に達するまで重要文を抽出する。

これらの処理を以下に説明する。

2.2.1 音声認識システム

本研究では、以下に示す音響・言語モデルと音声認識エンジン Julius rev.3.3p3 で構成した音声認識システムを用いる。

音響モデルには、CSJの2496講演(486時間)で学習した状態共有型 triphone モデル(3000状態混合)を用いる。言語モデルには、CSJの2592講演(6.7M形態素)から学習した単語 3-gram モデル(語彙サイズ: 24437)を用いた [7]。

形態素単位として、国立国語研究所で定義された短単位 [8] を使用し、形態素解析システムは、通信総合研究所で最大エントロピー法によりCSJのコアを用いて統計的に学習されたものを用いる [9]。

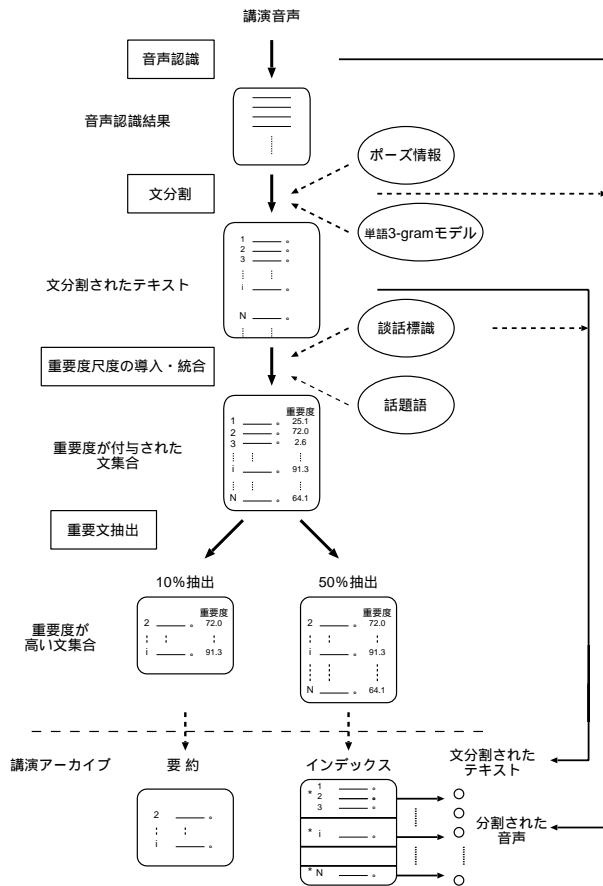


図 1: 重要文抽出の処理の概要

2.2.2 文単位への分割

日本語の話し言葉においては文の定義・境界が曖昧であるため、文を定義する必要がある。実際に CSJ の講演の書き起こしにも句点は含まれない「話し言葉工学」プロジェクトでは、原則として、

- (1) 「～です」「～ます」「～ない」「～た」「～と (と文末)」のような文末表現
- (2) 「～けれども」「～が」「～まして」「～でして」「～し」といった節末表現
- (3) 節末表現+接続詞

の箇所のうち機能的に文に区切られると判断できる箇所で文に区切ると定められている。本研究でもこれを文と定義する。

しかし上記のうち典型的な文末表現以外は、文末以外でも用いられるため、パターンマッチングをすることで規則的に句点を挿入するのは難しい。

ポーズ長も含めた前後の表現を考慮して統計的機械翻訳の枠組みにより句読点挿入を行う手法 [10] を

採用する。

これは、式 (1) で表されるように、話し言葉 (書き起こし) の文字列 X に対し、それを最もよく説明する書き言葉 (句読点を含む) の文字列 Y を求める方法である。

$$\max_Y P(Y|X) = \max_Y P(Y)P(X|Y) \quad (1)$$

ここで言語モデルのゆう度 $P(Y)$ の計算には、句点を含むテキストコーパスで学習した言語モデルを用いる。また、変換モデル $P(X|Y)$ には、ポーズ長とその前後の表現に依存するモデルを用いた。具体的には、話し言葉に特有の文末・文頭パターン「～と」「～ない」「～た」及び「節末表現+接続詞」の箇所に対しては、平均値以上の長さのポーズがある場合にのみ挿入しうる ($P(X|Y) = 1$) とした。それ以外の文末表現の箇所はポーズがなくても挿入しうるとした。

2.2.3 重要文抽出

区切られた各文に対して、3章で述べる談話標識に基づく重要度尺度および話題語に基づく重要度尺度を統合して重要度を付与する。その上で、重要度の高い順に指定された割合の文を抽出する。

CSJ の講演の書き起こしはその一部に人手により 50% と 10% を目安に抽出された重要文セットがある。したがって、50%抽出および 10%抽出を行い、人間の精度と比較する。

本研究では、この重要文抽出手法を用いて講演アーカイブへの要約やインデックスの付与を目指している。

3 談話標識に基づく重要文抽出

3.1 学会講演における談話構造

一般に学会講演は、談話構造が比較的明確である。すなわち、話題を問わず、大きく緒論、本論、結論に分けることが可能である。さらに、緒論については背景と目的に、本論については手法の説明と実験およびその評価に分けられることが多い。これらは講演においては 1~数枚程度のスライドに対応するまとまった話題の単位に該当する。本稿ではこれをセクションと定義する。これらのセクションの冒頭

では、次に話すことを述べた典型的な表現(例:「次に本研究の手法を説明いたします」、「今回の評価実験としまして...」)を含む文が多く用いられる。これは発表者が次に話すことを簡潔に伝え、聴衆の注意を引こうとするためと考えられる。このことは、セクションの冒頭部分では、重要な発話がなされることが多いことを示している。

このようなセクション境界情報の利用はテキストベースの重要文抽出では有効であることが知られている。ただし、テキストからの重要文抽出では改行や字下げなどの明示的な情報からセクション境界の検出が容易であるのに対して、音声ではそれらに相当するものは存在しないため、単純に適用できない。本研究では、セクションの冒頭に頻出する典型的な表現(=談話標識)に着目し、セクション境界を自動で検出することによって、重要文を抽出する。従来の研究では、談話標識は言語的な分析に基づいて人手により設定されるのが一般的であったが[11]、本手法では、談話標識をコーパスから教師なし学習により選定している。

3.2 談話標識に基づく重要度尺度

談話標識の学習は長谷川らの手法[12]を用いて次の1)~3)のように行う。この談話標識の学習は自動で行なわれ、人手によるタグ付けは一切必要ない。

1) 学習データである各講演の平均ポーズ長以上の箇所をセクション境界候補とする。これは、話題の転換点がスライドを変えるとといった作業や息継ぎにより他の箇所よりも比較的長めのポーズが挿入されるという仮定に基づいている。

2) 得られたセクション冒頭の1文を抽出し、セクション冒頭の文集合を作る。

3) セクション冒頭に特徴的に出現する単語 m を談話標識とし、以下の式(2)で統計値 DM_m を求める。

$$DM_m = wf_m * \log\left(\frac{N_s}{sf_m}\right) \quad (2)$$

単語頻度 wf_m は、セクション冒頭の文集合において単語 m が出現する回数であり、文頻度 sf_m は、学習セットの全ての文(総数 N_s)で単語 m が出現する文の数である。セクションの冒頭、すなわち話題の転換点で頻出し、それ以外の箇所ではあまり出現しない談話標識 m に対しては、この値 DM は大きな値となる。

表 1: 学会講演から選定された談話標識(上位 30)

こと	もの	次	結果	研究	語
方	実験	場合	説明	訳	先程
音声	モデル	問題	学習	ところ	今
データ	中	情報	時	実際	評価
文	単語	認識	システム	関係	発表
例	音	為	意味	方法	後

本研究では、CSJの学会講演 930 講演を用いて統計値を学習した。予備実験の結果をもとに、ここではすべての名詞(24421種類)に対して、この量を求めた。

表 1 に学会講演から選定された統計量 DM の大きいものの例を示す。「次」や「結果」、「研究」、「実験」、「評価」などの学会講演のセクションの先頭に頻出する単語が多く含まれていることがわかる。

この談話標識の統計量 DM を利用して、各文 s_j に重要度を付与する。各文に出現する全ての談話標識の統計量の合計(式(3))はセクション境界らしさ(=重要度)を表すので、この評価値が高いものを重要文として抽出する。

$$S_{DM}(j) = \sum_{m_i \in s_j} DM_{m_i} \quad (3)$$

3.3 話題語に基づく重要度尺度との統合

次に、話題語の統計量を用いた重要度について述べる。これは、話題と関連のある単語(話題語)は各講演において繰り返し出現すると仮定して、話題語を多く含む文を抽出するものである。

話題語の統計量としては、式(4)で定義される単語(名詞)の $tf*idf$ 値を用いる。

$$KW_m = tf_m * \log\left(\frac{N_d}{df_m}\right) \quad (4)$$

tf_m は名詞 m の当該講演内での出現回数を、 df_m はコーパス(全講演数 N_d)のうち、名詞 m が出現した講演数を表し、 N_d を df 値で除したものが idf 値である。当該講演を特徴付ける語、すなわち話題語はこの KW の値が大きくなる。ここでは単純に単語そのものを用いるのではなく、当該講演において3回以上出現する名詞列を複合語として扱う。これは、より話題を明確に示す語を定義するためである。例えば、「音声」「認識」という2単語が連続して出現した場合、これらを複合名詞「音声認識」の1単語

表 2: 重要文抽出における作業者間の一致度

	任意の 2 名	3 名
50%抽出 (全文比率)	75.5%(37.8%)	62.7%(31.3%)
10%抽出 (全文比率)	46.6%(4.7%)	30.8%(3.1%)

とみなすことで「音声」や「認識」を単独で用いる場合に比べて、話題をより明確に特徴付ける。

統計量の学習には、談話標識の統計量を求めたものと同一の CSJ の学会講演 930 講演を用いる。

話題語に基づく重要度は各文 s_j に含まれる名詞の $tf*idf$ 値の合計 (式 (5)) と定義する。

$$S_{KW}(j) = \sum_{m_i \in s_j} KW_{m_i} \quad (5)$$

次に、この談話標識に基づく重要度と話題語に基づく重要度とを統合する。具体的には、それぞれの手法により得られる各文の重要度 ($S_{KW}(j)$, $S_{DM}(j)$) の重みつき幾何平均を、新たな重要度とする (式 (6))。

$$S_{KW}(j)^\alpha * S_{DM}(j)^{(1-\alpha)} \quad (6)$$

本実験では、この重み係数 α の値を事後的に 0.6 に決定したが、この値を極端な値に設定した場合を除いて、最終的な重要文抽出の結果に大きな差はない。

4 評価実験

本研究では、2.1 節で述べた重要文タグが付与された CSJ の学会講演 19 講演を用いて、重要文の自動抽出の評価を行う。

まず、作業者間での重要文抽出の一致度を調べた。結果を表 2 に示す。全文の 50% を抽出した場合は比較的高い一致率がみられたが、10% 抽出では一致率が半分に満たず、主観性が高いことがわかる。

このような作業者間の不一致も考慮して、本研究では人間の抽出精度と比較できる形で正解セットを作成した。

まず、50% 抽出の重要文については、以下の 3 通りが考えられる。

50-1: 50% 抽出において任意の 1 名が重要文として抽出した文 (全文比率 50.2%)

50-2AND: 50% 抽出において任意の 2 名が共に重要文として抽出した文 (全文比率 38.1%)

表 3: 人間の重要文抽出の精度

抽出率	正解セット	再現率	適合率	F 値
50%	50-1	75.3%	75.2%	0.752
	50-2AND	82.3%	62.3%	0.709
10%	10-1	45.9%	45.1%	0.455
	10-2OR	39.8%	60.4%	0.483

50-2OR: 50% 抽出において任意の 2 名のうちいずれかが重要文として抽出した文 (全文比率 62.6%)

このうち、正解セット 50-2OR については正解数が全文の 6 割を越えるためインデキシングには適さないと考え、正解セットから除外した。

また 10% 抽出の正解セットとしては以下の 3 通りが考えられる。

10-1: 10% 抽出において任意の 1 名が重要文として抽出した文 (全文比率 10.2%)

10-2AND: 10% 抽出において任意の 2 名が共に重要文として抽出した文 (全文比率 4.7%)

10-2OR: 10% 抽出において任意の 2 名のうちいずれかが重要文として抽出した文 (全文比率 15.7%)

このうち、正解セット 10-2AND については人間の一致率が非常に低く、文の数が極端に少ないため、正解セットから除外した。

評価に用いる 4 種類の正解セットには 1 講演につき 3 通りの正解が存在するため、それぞれに対する再現率 (recall), 適合率 (precision), F 値 (F-measure) を算出し、その平均で評価を行う。ここで、F 値は以下の式 (7) で定義する。

$$F\text{-measure} = \frac{2 * recall * precision}{recall + precision} \quad (7)$$

この枠組みにおいて人間の重要文抽出の精度を推定することができる。正解セット 50-1, 10-1 については、1 名による正解重要文に対して他の 2 名それぞれが抽出したものと照合することで抽出精度を算出しその平均を求めた。正解セット 50-2AND, 10-2OR については、2 名による正解重要文に対して他の 1 人が抽出したものを照合しその平均精度を求めている。結果を表 3 に示す。人間の抽出精度は、50% 抽出で F 値約 0.75, 10% 抽出では F 値約 0.45 であった。

表 4: 各手法による書き起こしからの重要文抽出(正解セット 50-1(正解 50.2%), 抽出 50%)

	文分割	再現率	適合率	F 値
DM	人手	66.1%	66.0%	0.661
KW	人手	67.4%	67.3%	0.673
KW+DM	人手	69.1%	69.0%	0.690
KW+DM	自動	70.7%	62.4%	0.663
KW+DM	自動	72.6%	60.3%	0.658
人間	人手	75.3%	75.2%	0.752

DM: 談話標識, KW: 話題語, : 音声認識結果

表 5: 各手法による書き起こしからの重要文抽出(正解セット 50-2AND(正解 38.1%), 抽出 50%)

	文分割	再現率	適合率	F 値
DM	人手	70.8%	53.7%	0.611
KW	人手	71.6%	54.3%	0.618
KW+DM	人手	74.1%	56.1%	0.639
KW+DM	自動	74.5%	49.9%	0.598
KW+DM	自動	76.0%	47.7%	0.586
人間	人手	82.3%	62.3%	0.709

DM: 談話標識, KW: 話題語, : 音声認識結果

4.1 50%重要文抽出

まず 50%の重要文抽出(インデキシング)の評価を行った。談話標識に基づく提案手法(DM)と話題語に基づく方法(KW), およびこれらを統合した手法(DM+KW)を用いて書き起こしからの重要文抽出を行った。正解セット 50-1(正解数 3469, 総文数 6912), 50-2AND(正解数 2633, 総文数 6912)を用いた場合の抽出結果をそれぞれ表 4, 5 に示す。

談話標識に基づく手法は, 従来の手法である話題語に基づく方法といずれもほぼ同精度の結果を得た。また 2つの手法を統合することにより精度が改善している。2つの正解セットの再現率及び適合率はともに, 人間の抽出精度に比べて 10%程度の低下であった。

正解セット 50-1, 50-2AND それぞれに対して抽出率を変化させたときの重要文抽出の精度を図 2, 3 に示す。検索においては, 正しいインデックスが検出されていないと検索はできないが, 誤って検出された部分はスキップすればよいことから, 再現率を重視する必要がある。人間の再現率と同程度の精度を得るためには, とともに 60%程度の抽出が必要であった。

次に, 音声認識結果に対して適用を行った。音声認識結果に対しては, 2.2.2 節で述べた手法で文に分

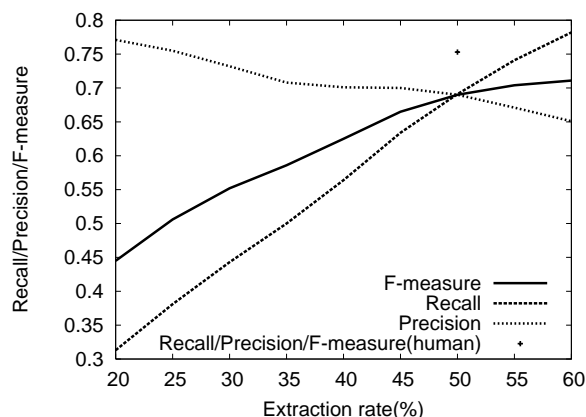


図 2: 抽出率による重要文抽出精度の変化(手法: DM+KW, 正解セット: 50-1, 書き起こし)

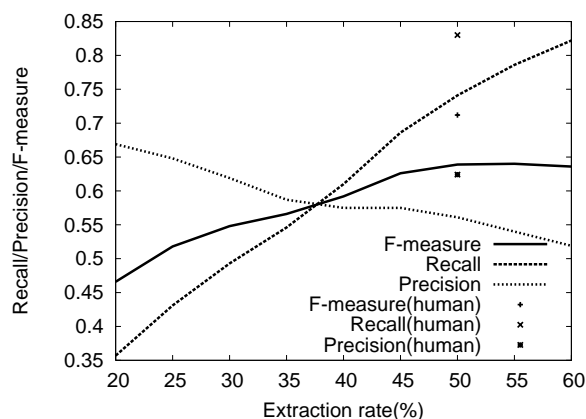


図 3: 抽出率による重要文抽出精度の変化(手法: DM+KW, 正解セット: 50-2AND, 書き起こし)

割を行なうが, この結果は人手によって文分割された書き起こしの区切りの位置と必ずしも一致しないため, 評価の際には対応する(複数の)人手の文区切りにおける文を選択したとみなして再現率や適合率の計算を行う。

音声認識結果への適用の前に, 文分割精度が重要文抽出に与える影響について調べた。書き起こしに対して自動文分割をした結果も表 4, 5(印)に示す。人手による文分割した場合と比較すると, 再現率はほぼ変わらないものの, 適合率がかなり低下している。これは正しく句点が挿入されなかったために長い文が生成され, 重要文の前後の非重要文も抽出したためと考えられる。

音声認識結果から重要文を抽出した場合は, 書き起こしに対して自動で文区切りを行った場合と比べて, 精度の差は小さく, 提案手法は認識誤りに対し

表 6: 各手法による書き起こしからの重要文抽出(正解セット 10-1(正解 10.2%), 抽出 10%)

	文分割	再現率	適合率	F 値
DM	人手	24.8%	24.4%	0.246
KW	人手	24.9%	24.5%	0.247
KW+DM	人手	25.2%	24.8%	0.250
KW+DM	自動	31.6%	22.3%	0.261
KW+DM	自動	30.2%	20.0%	0.240
人間	人手	45.9%	45.1%	0.455

DM: 談話標識, KW: 話題語, : 音声認識結果

表 7: 各手法による書き起こしからの重要文抽出(正解セット 10-2OR(正解 15.7%), 抽出 20%)

	文分割	再現率	適合率	F 値
DM	人手	39.2%	30.1%	0.341
KW	人手	39.2%	30.1%	0.341
KW+DM	人手	43.2%	33.2%	0.375
KW+DM	自動	47.0%	27.7%	0.348
KW+DM	自動	48.0%	26.5%	0.341
人間 *	人手	39.8%	60.4%	0.480

*: 抽出率 10%

DM: 談話標識, KW: 話題語, : 音声認識結果

て頑健であることが示された。

4.2 10%重要文抽出

次に 10%の重要文抽出(要約)の評価を行った。正解セット 10-1(正解数 702, 総文数 6912), 10-2OR(正解数 1082, 総文数 6912)を用いて, 重要文抽出を行った結果をそれぞれ表 6, 7 に示す。なお 10-2OR では正解の重要文の比率が 15.7%であるので, 20%抽出した。ここでも談話標識に基づく手法と話題語に基づく手法の統合の効果が確認された。

10%を抽出した場合, 人間同士でも高い一致率がみられず, 人間の推定精度も 0.5 以下であるが, 本手法で得られた結果はさらに低い結果となった, このように重要文の数が少ない場合に対するアプローチについては検討の必要がある。

正解セット 10-1, 10-2OR それぞれに対して抽出率を変化させたときの重要文抽出の結果を図 4, 5 に示す。ここでも 50%と同様の傾向が見られ, 人間の精度と同等の再現率を得るためには人間より 10%程度多くの文を抽出する必要があった。

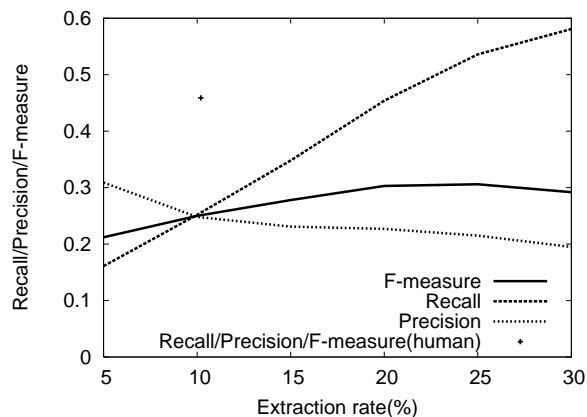


図 4: 抽出率による重要文抽出精度の変化(手法: DM+KW, 正解セット: 10-1, 書き起こし)

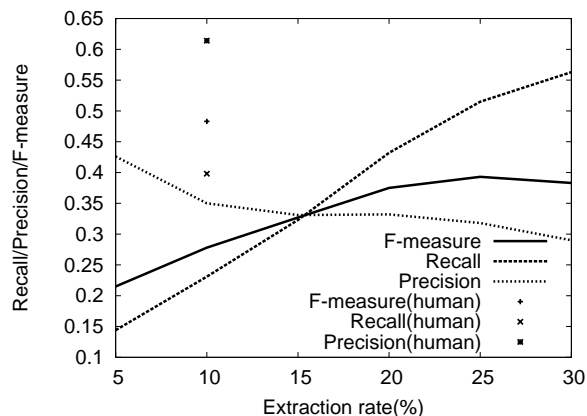


図 5: 抽出率による重要文抽出精度の変化(手法: DM+KW, 正解セット: 10-2OR, 書き起こし)

音声認識結果への適用を行った結果も表 6, 7 (印) に示している。書き起こしの場合と比べて精度の大きな差が見られず, また自動文分割の影響も小さかった。

5 アーカイブシステム

このように自動抽出された重要文のインデックスを用いて図 6 に示すような講演アーカイブのブラウザを作成することができる。これらのタグは XML で記述されており, XSLT で変換して, Web ブラウザで見ることができる。また, 講演音声を認識した結果がセクションに区切られており, 区切られたセクション単位で音声を聞くことができる。

ここでは, 50%抽出の結果をセクション単位のイ

ンデックスとして、10%抽出の結果を要約として用いることを想定している。

6 結論

CSJの重要文タグが付与されている講演セットを用いて重要文抽出の実験・評価を行った。重要文の自動抽出には談話標識と話題語の統計量を統合した手法を用いた。談話標識に基づく手法の有効性および話題語に基づく手法との統合の相乗効果、および認識誤りに対しても頑健であることを確認した。50%抽出においては人間の抽出精度に比べて10%程度低いものの大きな差はなく、インデキシングの手法として有効であることがわかった。しかし10%抽出においては人間の抽出精度との差は大きく、改善の余地が大きいことがわかった。

参考文献

- [1] 北出祐, 南條浩輝, 河原達也, 奥乃博. 談話標識と話題語に基づく統計的尺度による講演からの重要文抽出. 情処学研報, 2003-SLP-46-2, 2003.
- [2] 伊藤山彦, 松本賢司, 谷田泰郎, 柏岡秀紀, 田中英輝. 講演文を対象にした重要文抽出実験. 「話し言葉の科学と工学」ワークショップ予稿集, pp157-164, 2001.
- [3] 笠原力弥, 山下洋一. 講演音声における重要文と韻律的特徴の関係. 情処学研報, 2001-SLP-35-5, 2001.
- [4] 井上章, 三上貴由, 山下洋一. 連続音声認識による言語情報と韻律情報を利用した講演音声の重要文抽出. 情処学研報, 2003-SLP-49-14, 2003.
- [5] 古井貞熙, 前川喜久雄, 井佐原均. 科学技術振興調整費開放的融合研究推進制度 - 大規模コーパスに基づく「話し言葉工学」の構築 -. 日本音響学会誌, Vol.56, No.11, pp.752-755, 2000.
- [6] 前川喜久雄. 言語研究における自発音声. 音講論, 1-3-10, 春季 2001.
- [7] T.Kawahara, H.Nanjo, T.Shinozaki, and S.Furui. Benchmark test for speech recognition using the Corpus of Spontaneous Japanese. In *Proc. ISCA & IEEE Workshop on Spontaneous Speech Processing and Recognition*, pp. 135-138, 2003.
- [8] 小椋秀樹. 話し言葉コーパスの単位認定基準について. 「話し言葉の科学と工学」ワークショップ講演予稿集, pp. 21-28, 2001.
- [9] 内元清貴, 井佐原均. 話し言葉コーパスの形態素解析. 「話し言葉の科学と工学」ワークショップ講演予稿集, pp. 33-38, 2002.
- [10] 河原達也, 下岡和也. 統計的手法を用いた講演書き起こしの自動整形. 音講論, 3-9-10, 秋季 2002.
- [11] 小林聡, 吉川裕規, 中川聖一. 表層情報と韻律情報を利用した講演音声の要約. 情処学研報, 2002-SLP-43-7, 2002.
- [12] 長谷川将宏, 秋田祐哉, 河原達也. 談話標識の抽出に基づいた講演音声の自動インデキシング. 情処学論, Vol. 43, No. 7, pp. 2222-2229, 2002.

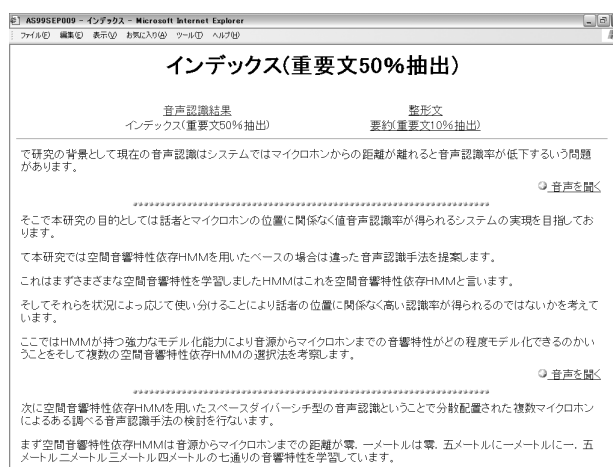


図 6: 講演アーカイブ表示例