

## 講演音声における重要文と韻律的特徴の関係

笠原 力弥 山下 洋一

立命館大学理工学部情報学科

〒525-8577 滋賀県草津市野路東1-1-1

Email : rikiya@slp.cs.ritsumeai.ac.jp ,yama@cs.ritsumeai.ac.jp

概要:本研究では、講演・講義といった音声データを対象とした要約の自動生成を行うために、重要文と韻律的特徴の関係についての分析を行った。文の重要度は、人手による要約実験から決定し、韻律的特徴として、正規化音素長、正規化パワー、ピッチの3つのパラメータを用いた。重要文、非重要文の特徴分析を文単位、単語単位で行い、韻律的特徴の時間変化を定性的に分析するとともに、文の重要度と韻律的特徴の相関を求めた。文単位での分析では、パワーに関して文の重要度とのやや強い相関が見られた。単語単位で分析し、自律語の韻律的特徴に注目することにより、パワーだけでなく、音素長、ピッチにおいても文の重要度との相関が見られた。

キーワード: 要約、講演音声、韻律的特徴、重要文

## The Relation between Important Sentences Lecture Speech And Prosodic Features

Rikiya Kasahara and Yoichi Yamashita

Email : rikiya@slp.cs.ritsumeai.ac.jp ,yama@cs.ritsumeai.ac.jp

Department of Computer Science, Ritsumeikan University  
1-1-1Noji-Higashi, Kusatsu-shi, Shiga, 525-8577 Japan

Abstract : This paper describes characteristics of prosodic features in important sentences in lecture speech for automatic summarization. The importance of each sentence in lecture talk is measured by a subjective experiment of summarization. We analyzed relation between the sentence importance and three prosodic parameters, duration, power, and pitch, after qualitatively analyzing trajectory of the prosodic parameters. In the sentence level analysis, the power was relative to the sentence importance. In the word level analysis based on content words, the duration and the pitch as well as the power had much relation to the sentence importance.

Keywords: summary, lecture speech, prosodic features, important sentences

## 1. はじめに

近年、電子化されたテキストが世の中に溢れ、情報洪水という言葉が使われるようになってきた。しかし、残念ながら、我々の情報処理能力は、たとえ処理しなければならない情報が増えても、それほど向上しない。そのため、これからは、自動要約技術などにより、読み手が読むテキストの量を制御できることが求められる。現在、テキストにおける要約技術[1]がかなり進んでおり、いくつかの市販ソフトウェアが販売されているほどである。しかし、講演や講義など長時間の話の内容(発話)を要約するという技術はあまり進んでおらず[2]、今後はこのような音声データに対する要約技術が期待されている。

要約は利用目的から、原文(テキスト)の適切性を判断するなど、原文を参照する前の段階で用いるような要約(indicative)、原文の代わり(字幕など)として用いるような要約(informative)に分かれている。本研究では、情報検索などに役立つような講演・講義といった音声データを対象とした indicative な要約の自動生成を目的とする。

一般的に「要約」とは、伝達すべき情報を凝縮し、それをより短い形で表すことである。そして、人間は文章を「理解」、「再構成」して、要約の「文章生成」を行っている。しかし、コンピュータでこうした処理を行うのは容易なことではない。そこで、多くの要約に関する研究では、「再構成」から「文章生成」を行う処理を省略し、要約を「理解」≡「重要部分の同定」としている。本研究でもこの考えに基づき講演から重要な文を抜き出し、それを出現順に並べて文を抽出(抄録)することを要約[3]とする。上記のような要約の生成は次の3つの工程からなる。(1)「基となる講演音声全体を文単位に分ける」(2)「1文ごとに重要度を

与える」(3)「重要度の高かった文を抜き出す」このうち、(1)の文単位は既知として扱い、文単位への分割は本研究の対象外とする。本研究としては特に(2)の重要度を与える部分に注目して研究を行う。(2)は文字テキストの場合、重要度を定める際に、テキスト中のキーワードの出現頻度やタイトルなどの情報を利用して重要度を与えている。音声を対象とした要約では、キーワードの出現頻度などの言語情報だけでなく、音声の韻律的特徴を利用することが重要文を決める際に有効である可能性がある。そこで、本報告では、人手で決定された重要文、非重要文を基に、文単位、単語単位についての韻律情報分析を行った。

## 2. 人手による要約実験

重要文、非重要文を得るために、まず、手による要約実験を行った。

### 2. 1 実験方法・結果

NHKの論説番組(約10分間番組、1日分、内容は、「茨城県東海村」JOC臨海事故)の音声を手により書き起こし、それを文単位に分割した。まず、被験者に論説番組のビデオを見て内容を理解してもらった上で、書き起こしテキストから重要だと思われる文(重要文)と重要だと思わない文(非重要文)を抜き出してもらった。被験者数は14人である。文の重要度を

$$\text{文の重要度} = \text{重要文の数} - \text{非重要文の数} \quad (1)$$

で定義し、重要度が7以上の文を重要文、-6以下の文を非重要文とした。その結果を図1に示す。横軸は文番号、縦軸は重要度を示す。

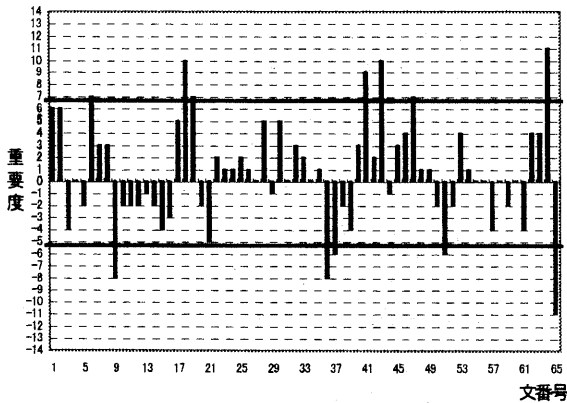


図1：人手による要約実験結果

### 3. 韻律的特徴の分析

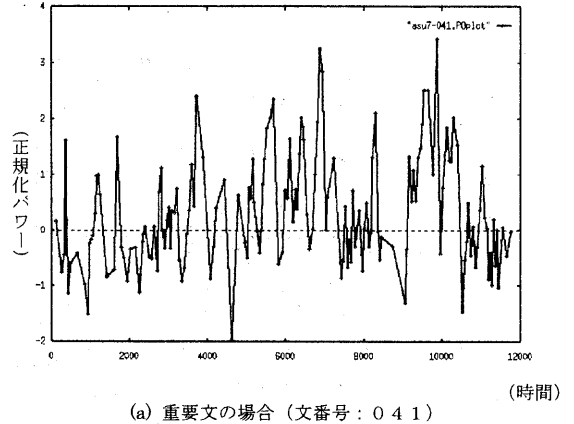
#### 3. 1 文単位での分析

##### 3. 1. 1 分析パラメータ

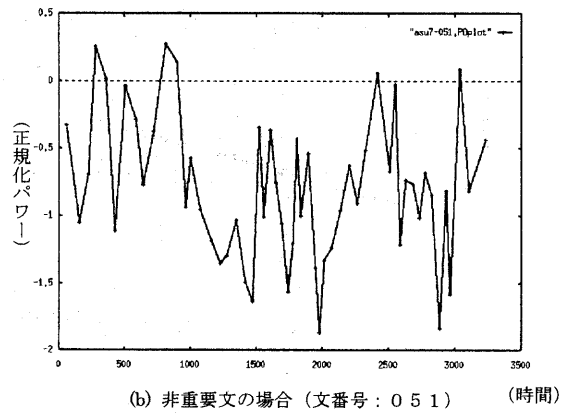
3つの韻律的特徴(時間長、パワー、ピッチ)に対して、重要文における韻律的特徴の分析をおこなった。まず、自動ラベリングを行い、そこで得られる結果から、音素ごとの正規化時間長、音素中心の正規化パワーを式(2)から求めた[4]。

$$\text{正規化パラメータ} = \frac{\left[ \begin{array}{l} \text{長さ} \\ \text{パワー} \end{array} \right] - \text{その音素の平均}}{\text{標準偏差}} \quad (2)$$

また、ピッチは各音素中心の値を用いる。韻律的特徴の時間変化の例として、重要文、非重要文の中で、最も平均パワーが高かった文と低かった文を1文ずつ、それぞれ図2(a)、(b)に示す。これらの図を見ると、重要文には大きな変化がないが、非重要文についてパワーが弱いという傾向が見られた。



(a) 重要文の場合(文番号: 041)



(b) 非重要文の場合(文番号: 051)

図2：正規化パワーの時間変化

#### 3. 1. 2 重要度と相関

正規化時間長、正規化パワー、ピッチに対してのパラメータと人手による要約実験から得た文の重要度との相関図を図3～5にそれぞれ示した。横軸が文の重要度、縦軸は1文単位の各韻律パラメータの平均値である。また、1文単位の各韻律パラメータ平均値に関して重要文、非重要文、その他の3つのカテゴリにおける平均を求めた結果をそれぞれについて表1に示す。

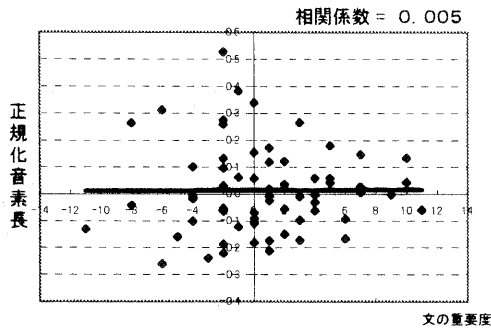


図3：正規化音素長と文の重要度

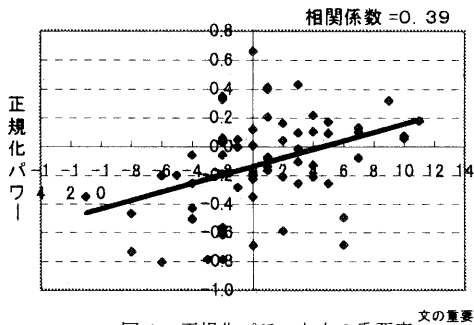


図4：正規化パワーと文の重要度

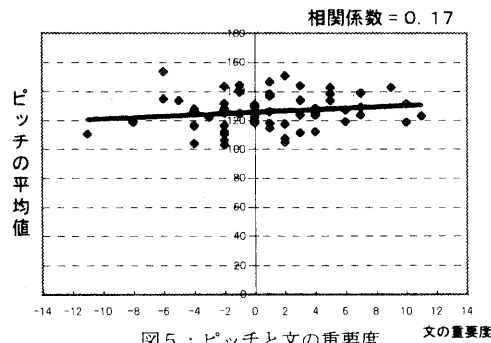


図5：ピッチと文の重要度

表1：正規化パラメータ平均値の比較

	音素長	パワー	ピッチ
重要文の平均	0.041	0.110	129.7
その他の平均	0.0083	-0.114	125.2
非重要文の平均	0.027	-0.512	127.4

図3～5における相関図をみると、パワーについてのみ、文の重要度とやや強い相関が見られたが、その他の韻律的特徴については相関が見られなかった。また、表1のように、重要文、非重要文の平均値をとってみると、これも、パワーについてのみ重要文と非重要文の間で差がでていることがわかった。

### 3. 2 単語単位での分析

#### 3. 2. 1 分析パラメータ

次に、単語単位に韻律的特徴を分析し、文の重要度との関係を調べる。ここでは、式(2)で求めた音素ごとの正規化時間長、音素中心のパワーを利用して式(3)を使って単語単位の韻律パラメータを求める(この場合もピッチについては正規化をしていない)。

$$\text{単語の正規化} \begin{bmatrix} \text{音素長} \\ \text{パワー} \end{bmatrix} = \frac{\text{正規化} \begin{bmatrix} \text{音素長} \\ \text{パワー} \end{bmatrix} \text{の和}}{\text{単語の音素数}} \quad (3)$$

さらに、形態素解析された品詞の中で、助詞、助動詞、フィラーを除く、自律語の韻律パラメータのうち、1文の中で最も大きい値を文の韻律パラメータとして使用した。

#### 3. 2. 2 単語単位による重要度との相関

単語単位で求めた文の正規化時間長、正規化パワー、ピッチと文の重要度との相関図を図6～8に示す。横軸は文の重要度であり、縦軸は1文の中での自律語の韻律パラメータの最大値を示す。また、重要文、非重要文、その他の文、3つのカテゴリにおけるこれらの韻律パラメータの平均を求めた結果を表2に示す。

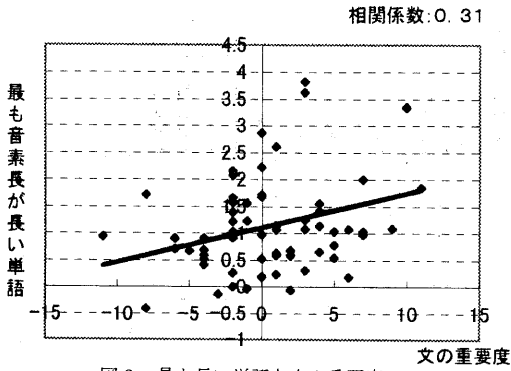


図6：最も長い単語と文の重要度

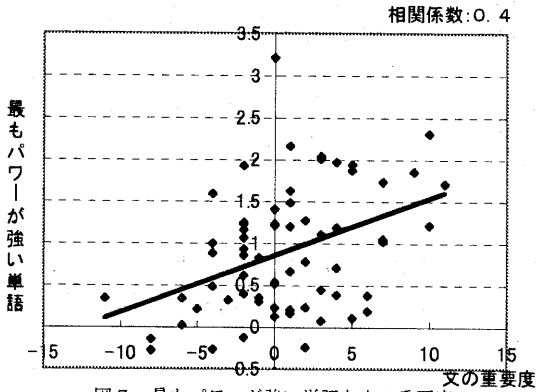


図7：最もパワーが強い単語と文の重要度

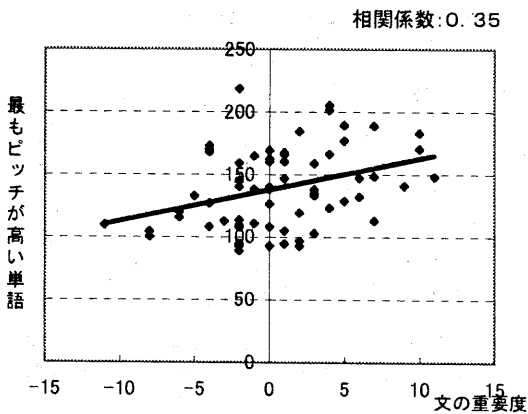


図8：最もピッチが高い単語と文の重要度

表2：単語単位で求めた韻律パラメータの比較

	音素長	パワー	ピッチ
重要文の平均	1.94	1.55	156
その他の平均	1.06	0.89	139
非重要文の平均	0.76	0.005	109

文単位の韻律パラメータでは、パワーにしか相関が見られなかったが、図6～8における相関図をみると、1文の中にある最も値の高い単語（自律語）に注目することで、音素長、ピッチにも相関が見られるようになった。また、表2で、単語単位で重要文、非重要文の平均値をとって見ても、パワーについてのみ重要文と非重要文の間で差がでているだけでなく、こちらも、音素長、ピッチについて重要文、非重要文の間で差がでていることがわかった。

### 3. 2. 3 単語単位による定性的分析

単語単位における正規化時間長、音素中心のパワーおよびピッチの時間変化を定性的に分析を行った。ここでの分析では、全ての単語を扱っており、助詞、助動詞の変化についてもわかるように示した。図2で扱った文単位での平均パワーが高かった文、低かった文と同じ文の単語単位での正規化パワーの変化を、それぞれ図9の(a)、(b)示す。縦軸は式(3)で求めた単語単位の韻律パラメータで、横軸は文中の単語である。

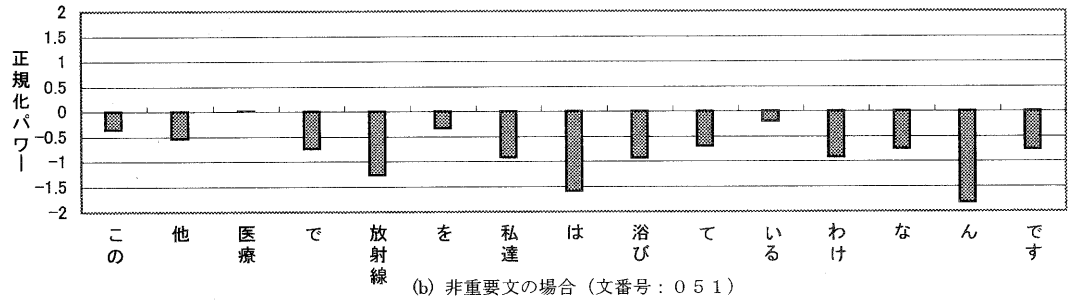
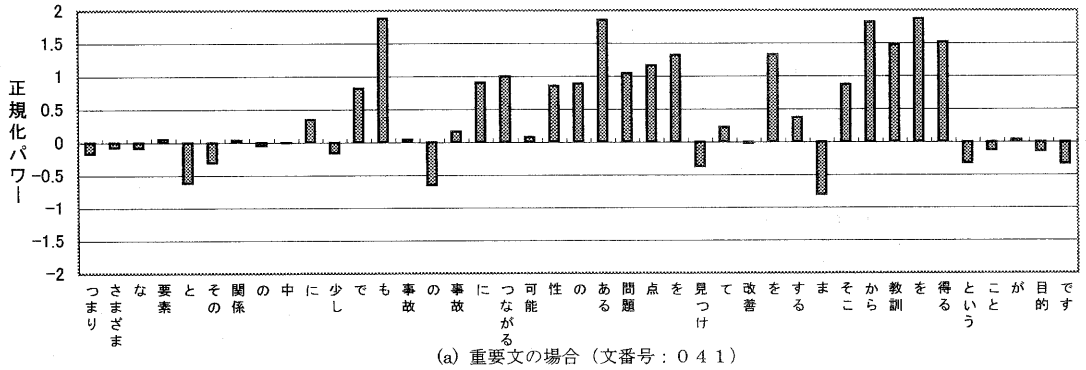


図9：単語単位による正規化パワーの時間変化

単語単位での正規化パラメータの図(時間長、パワー、ピッチ)から以下のことがわかった。

- ・重要文に含まれる単語について、話題に関連のあると考えられる単語、または、強調したいと考えられる単語が連続している部分に、音素長が長く、パワーが強く、ピッチが高いといった傾向が見られた。

- ・非重要文に含まれる単語については、話題に関連のない単語が多く、重要文の中にある単語とは逆に、どの単語も、音素長が短く、パワーが弱く、ピッチも低いといった傾向が見られた。

## 5. まとめ

本報告では、重要文と韻律的特徴の関係についての分析を文単位、単語単位で行った。文単位では、あまり多くの特徴を得ることができなかったが、単語単位での分析を行ったことで、特徴が見られるようになってきた。今後は、引き続き、単語単位についての分析を行っていく

とともに、今回は、1話者によるデータしか扱っていないので、データ量も増やしていく予定である。

## 謝辞

本研究の遂行には、文部省科研費(特定領域研究(B)(2),No.12132203)の支援を受けた。

## 参考文献

- [1]奥村 学、難波 英嗣：テキスト自動要約に関する研究動向、自然言語処理、Vol.6 No.6 pp.1-26 (1999).
- [2]川崎剛史、川俣真人、山本幹雄、板橋秀一：韻律情報を考慮した音声要約一方法、日本音響学会春季研究発表会講演論文集、pp.239-240(2000.3).
- [3]佐藤 理史、奥村 学：電脳文章要約術—計算機はいかにしてテキストを要約するか—、情報処理、Vol.40、No.2、pp.157-161(1999).
- [4]藤尾 茂、ニック キャンベル、樋口 宜男：韻律を用いたテキスト非限定型発話アクト識別方法、日本音響学会春季研究発表会講演論文集、pp.245-246(1996.3).