

音声翻訳システムにおける韻律情報の活用*

◎秦 早穂子¹, △柏岡 秀紀¹², ニック・キャンベル¹²奈良先端科学技術大学院大学¹ ATR 音声言語通信研究所²

1. はじめに

音声翻訳システムは、音声認識・機械翻訳・音声合成の順で入力音声を目的言語へと変換し、変換結果を目的言語の音声で出力する。しかし文字列では肯定として解釈できる自然対話中の疑問文、たとえば「よくここに来るの?」、「毎日何時間も長電話する?」などの文章に関して、音声認識から認識結果の文字列を、翻訳部分に渡してしまうため平叙文として処理され、的確に翻訳結果を出力することができない。本稿ではそのような自然対話中の疑問文の韻律特徴を調べ疑問文と平叙文との相違を述べる。

これまでの研究では文末のF0値やパワーの変化などで文のタイプを決定する手法が提案されている。しかし対象となる文は語彙情報から平叙・疑問を判別できるものの分析が多く、同じ文字列で解釈が異なるデータに対しては報告されていない。ここでは同じ文字列の文章を疑問文・平叙文、それぞれに読み上げた音声データを収録しそれぞれのF0情報を分析した結果を報告する。

2. データ収録

例1のように、同じ文字列である文章の疑問文・平叙文を1セットとして30セット作成し、女性話者1名によって読み上げられた音声を収録した。

例1: 「そう思ったことない?」
「そう思ったことない。」

音声データはサンプリング周波数 48kHz、16bitでDATを用いて収録した。

*Prosody processing for a speech translation system
S.Hata(NAIST), H.Kashioka(NAIST/ATR), N.Campbell
(NAIST/ATR)

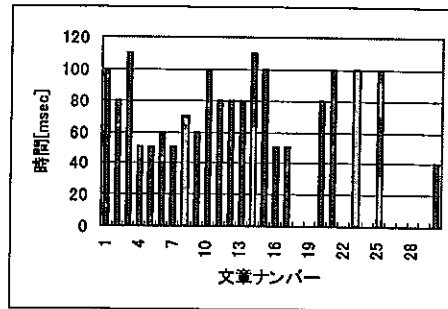


図1: 疑問文末のF0値が単調増加している区間の継続時間長

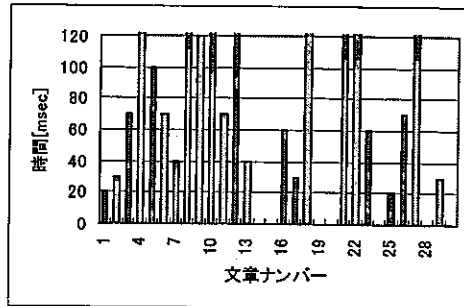


図2: 平叙文音声データの文末のF0値が単調減少、もしくは平坦である区間の区間長

3. 分析

30セット 60文の音声データを音声解析ツール wavesurfer に通し、それぞれのF0値を抽出し、視察によりハンドラベリングを行なった。本稿では特に文末からの2モーラ間のF0値に着目しその傾向を検証した。

図1に疑問文末のF0値が単調増加している区間の区間長を示す。図2に平叙文音声データの文末のF0値が単調減少、もしくは平坦である区間の区間長を示す。

従来、疑問文の認識、または平叙文の認識は文末のF0値の上昇・下降により判別されている。これらの結果より、文末のF0値の上昇・下

降の継続時間長の観点から認識できることが推察される。それぞれの閾値は以下のとおりである。

1. 文末のF0値が40msec以上上昇していれば疑問文として認識される。
2. 文末のF0値が20msec以上下降、もしくは平坦であれば平叙文として認識される。

なおこの図1, 2で棒グラフとして表されていない文章の文末におけるF0値の特徴は以下のとおりである。

3. 平叙文であるのに最後のF0値が約10~40msec間、上昇している。また上昇直前までの約40~160msec間下降している。
4. 疑問文であるのに最後のF0値が約20~30msec間、下降している。下降開始直前の約100msec間上昇している。

以上のことより疑問文と平叙文の違いを音響的に定義する場合、文末のF0値の上昇時間長、下降時間長を考慮することで認識できる。

1, 2の規則を用いると認識結果が約73%だったのに対し、3, 4を適応させると認識結果は80%に上がる。

3. 考察

図1, 2で棒グラフとして表されておらずまた上記の3, 4にも当てはまらないものが疑問文の中で数パターンみられた。そのパターンは以下のとおりである。

- A. 文末のF0値の変化の形が「N字型」になっているものが2文章含まれていた。しかし全体的には上昇している。
- B. 文末からの2モーラにおいてF0値が200msec間上昇したのち、120msec間に渡り下降し文章が終わっている。この文章と平叙文の文章全体のF0値を比較すると、文末の2モーラ間を除く文頭からのF0値のレンジ等には大きな違いは見られなかった。しかし最後の2モーラ間のF0値のレンジが疑問文の方は文頭の

F0値に対し高く、平叙文の方は低くなっているのが見られた。

- C. 疑問文として読み上げられたデータのF0情報と平叙文として読み上げられたデータのF0情報の高さ・形にはほとんど違いが見られなかったが、疑問文の方が最後の1モーラの長さが平叙文のそれと比べて約100msec長かった。

これらのパターンはデータ量としてごくわずかであるため、今後さらにデータを集め分析を行なう必要がある。また韻律情報以外の観点からも見ていく必要がある。

4. まとめ

同一表記の文字列で発話様式によって意味解釈がかわるもの、特に本稿では疑問文と平叙文との相違を文末2モーラのF0値上昇時間および下降・平坦時間の点から検証した。その結果、時間長の情報を利用することで80%程度の判別率が得られることがわかった。

今後、この結果を音声翻訳システムに反映させ機構の改善を行なっていく。

なお今回いくつかの例外に関して疑問文・平叙文のF0情報にさほど変化が見られないものの、確かに疑問と平叙で違うと認識できるものがあつた。これらに関しては語彙情報等の韻律以外の情報が関係していると見られる。そういったものに関しては、今後機械翻訳内で構文解析等を行い疑問であるか平叙であるかを判断させていく。

参考文献

- [1] 藤尾茂, ニック・キャンベル, 樋口宣男: "韻律を用いた発話タイプの識別", 日本音響学会講演論文集, pp.199-200 (1995-9)
- [2] 藤尾茂, ニック・キャンベル, 樋口宣男: "韻律を用いたテキスト非限定型発話アクト識別方法", 日本音響学会講演論文集, pp.245-246 (1996-3)
- [3] 前川喜久雄: "音声によるパラ言語情報の伝達: 言語学の立場から", 日本音響学会講演論文集, pp.381-384 (1997-9)
- [4] 田淵咲子, 甲斐朋子, 馬場良二: "文末「ノ」の音声的特徴" 日本音響学会聴覚研究会資料 H-99-120 (1999-12)