

◎正木 芽衣子 (奈良先端大) △柏岡 秀紀 (奈良先端大/ATR)
ニック・キャンベル (奈良先端大/ATR/CREST)

1 はじめに

近年、音声言語処理技術の進展にともない、単に言語情報を伝達するだけでなく、話者の意図や感情などのパラ・非言語情報をも伝達可能な合成音声が求められている。そのために、発話速度の制御も重要な課題となる。発話速度の制御を行うものとして、話速変換装置がある。話速変換装置を用いて発話速度をゆっくりとすると、聞き手の理解を助けるが明瞭度が損われ[1]、また、主観的なピッチがより低く認知され、認知的なバイアスが生じる[2]ことが報告されている。本稿では、音声合成における音韻継続長の予測モデルを、異なる発話速度に対応させることを目標として、発話速度の異なるデータ間の分析を試みる。

発話速度は、ポーズと密接な関係があることが知られている。ポーズが聞き手の理解に重要な役割を果たしており、人は呼吸と脳裏にある文法とを合わせながら話していると言われている[3]。また、日本語の音韻継続長において品詞の影響があり、自立語、特に情報を多く担っている箇所における母音長の伸張傾向が顕著であることが知られている[4]。以上のことから本稿では、音韻継続時間長の伸縮率に注目し、品詞と発話速度に応じた音韻継続時間長の関係を明かにするために、速い、普通、遅いの3つの発話速度のデータを収録し分析した結果について報告する。

2 データ収録

女性話者3名により、絵本『百万回生きたねこ』[5]を3つの異なる発話速度で収録した。発話収録に用いたテキストは絵本を読みやすくするために漢字を用いて書き直したものである(2241文字)。収録前に練習をおこない、指定した時間で読めるようになってから収録を行った。時間指定なし(以後この発話を普通発話と呼ぶ)、遅い、速いの順に収録した。音読しやすいようテキストはA4横置き、縦書きとし、4ページに収まるよう1ページ26行×28文字(728文字)で設定した。遅いときは原稿1ページあたり3分、速いときは原稿1ページあたり1分半を目安に発話するよう指示した。聞き手がわかりやすいように話すようにという指示も行った。読み間違っただけは中断せずにそのまま読み続けた。そのため今回は1箇所でも誰かが誤った場合、話者間、発話速度間での比較が容易なように、同じ文を全話者、全発話速度の発話から削除した。分析対象は、その結果106文から82文(1679モーラ)となった。

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、16kHzにダウンサンプリングした後、音素セグメンテーションキットJulius Segmentation Kitを用いて音素アライメントを行った。ポーズを全て除去したあとの各発話の継続時間長は表1のようになった。

*Differences in speech timing according to position in the utterance, Meiko Masaki(NAIST), Hideki Kashioka(NAIST, ATR), Nick Campbell(NAIST, ATR, CREST)

表1: ポーズ・誤り文除去後の各発話の継続時間長 [sec]
() は1秒あたりのモーラ数 [mora/sec]

	速い	普通	遅い
話者 A	182.91 (9.27)	220.74 (7.60)	271.16 (6.19)
話者 B	171.45 (9.79)	206.46 (8.13)	276.85 (6.06)
話者 C	195.22 (8.60)	218.06 (7.69)	279.51 (6.00)

3 分析

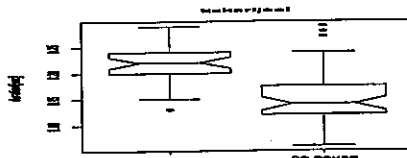


図1: 話者 A による遅い発話における係助詞の発話時間長 左から直後にポーズあり・ポーズなし

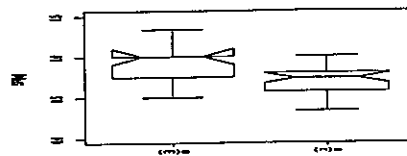


図2: 話者 A による助詞-係助詞と助詞-連体化における母音 o の直前子音の速い発話と普通発話の伸縮率 左から SR(m), SR(n)

分析に利用した品詞情報は、発話テキストを日本語形態素解析システム茶筌¹で解析した出力結果による。この品詞区分に基づいて分類した結果、937語、37種の品詞を得た。どの発話速度においても、文末・ポーズの前の音素は他に比べ継続時間長が長くなる。(図1)そこで以下の分析では取り除く。

単語の継続時間長は構成音素等により異なる。そこで発話速度の違いによる単語内の継続時間長変化を調べるため、伸縮率を求めた。文内での位置が等しく発話速度が異なるある音素/A/同士の伸縮率 SR は以下のように定義される。

$$SR(A) = \frac{\text{遅い・速い発話内の音素/A/の継続時間長}}{\text{普通発話における音素/A/の継続時間長}}$$

¹ ChaSen version 2.1 for Windows

表 2: 話者 A, B, C 間での差異 SR() 内の" は前後の音素を表す

	話者 A	話者 B	話者 C
SR([mo]) と SR([no]) の有意差	なし	なし	なし
SR('m'o) と SR('n'o) の有意差	SR('m'o) > SR('n'o)	SR('m'o) > SR('n'o)	SR('m'o) < SR('n'otextbfo)
SR(m'o') と SR(n'o') の有意差	SR(m'o') < SR(m'o')	なし	なし
全体での SR(m) と SR(n) の有意差	SR(m) < SR(n)	SR(m) < SR(n)	SR(m) > SR(n)

各音素で計算した伸縮率を各品詞ごとに比較したところ、速い発話における母音/o/において、助詞一係助詞と助詞一連体化で3話者に有意な差がみられた。(図2) 話者A,Bは助詞一係助詞のSR(/o/)が大きく、話者Cは助詞一連体化のSR(/o/)のほうが大きかった。各品詞における/o/を含む単語の内訳は助詞一係助詞では[mō], 助詞一連体化では[no]だけであった。さらに話者Aの速い発話について、助詞一係助詞[mō]の/m/, 助詞一連体化[no]の/n/についても調べたところ(図3)となった。

話者Aにおいて助詞一係助詞[mō], 助詞一連体化[no]の単語単位における速い発話の伸縮率SR([mo]), SR([no])を調べたところ、(図4)となった。

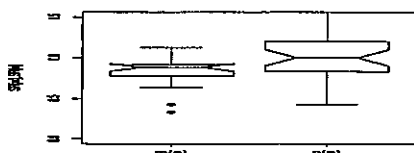


図 3: 話者 A による助詞一係助詞と助詞一連体化における母音 o の直前子音の速い発話と普通発話の伸縮率 左から SR(m), SR(n)

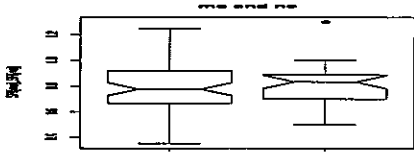


図 4: 話者 A による速い発話における助詞一係助詞[mō], 助詞一連体化[no]の伸縮率 左から SR(mo)・SR(no)

4 考察

以上の結果をまとめると表1のようになる。速い発話における伸縮率SR(/o/)において、助詞一係助詞と助詞一連体化間に有意な差が見られた。助詞一係助詞と助詞一連体化間で明らかに有意な差があることは、速い発話において、普通発話の全体を一律に縮めたものではなく、発話時間が個々に変化していることを示している。だが、話者A, Bと話者Cでは結果が逆転しており、一概に助詞一係助詞の母音/o/が縮みやすいとはいえない。A, Bは東京方言で朗読し、Cは関西方言で朗読しているためその影響がでている可能性がある。話者Cにおいて全体での/m/, /n/の伸縮率

SR(/m/)とSR(/n/)に有意な差がありながら、助詞一係助詞と助詞一連体化の/o/直前の子音に差がみられなかったことから、品詞間で何らかの影響を受けていると考えられる。また、助詞一係助詞と助詞一連体化間で単語単位で有意な差が見られないことから、単語内において音素の継続時間長の構成比率が変化していると思われる。このような傾向が見られることは、発話速度の制御において文法的な役割を用いて継続時間長を制御しているといえる。それゆえ、より自然な音声合成するためには文法的な働きを考慮すべきである。また速い発話をわかりやすく話すための技術力の違いや、話者の個性があらわれている可能性がある。今後データ数を増やし、各音韻ごとの比較が必要である。また、発話速度に変化を起させる要因として品詞だけではなく、句構造や意味情報も含めて分析する必要がある。

5 まとめ

助詞一係助詞と助詞一連体化間において母音/o/の伸縮率SR(/o/)に有意な差がみられた。助詞一係助詞と助詞一連体化間で単語単位で比較すると有意な差は見られなかった。単語の伸縮率に差がなくても単語内で各音素の伸縮率が変化している場合がある。このことから、発話速度を大局的に変化させる場合、一律に伸縮させるのではなく、個々に伸縮させる必要があると思われる。

今後、ポーズや音韻による影響も考慮した、異なる発話速度に対応した継続時間長の予測モデルを構築したい。

謝辞

本研究の一部は科学技術振興事業団戦略的基礎研究推進事業(CREST/JST)の援助により行われた。

参考文献

- [1] 宮坂栄一, "話すスピードと明瞭度", 言語, 大修館書店, No.9, pp.36-39(1999).
- [2] 内田照久, "音声の発話速度の制御がピッチ感及び話者の性格印象に与える影響", 日本音響学会誌, 56巻6号, pp.396-405(2000).
- [3] 杉藤美代子, "談話分析・発話とポーズ", 日本人の声-日本語音声の研究1, 和泉書院, pp.263-277(1994).
- [4] 海木延佳, 武田一哉, 勾坂芳典, "言語情報を利用した継続時間長の制御", 電子情報通信学会論文誌(A), J75-A, 3, pp.467-473(1992-03).
- [5] 佐野洋子:百万回生きたねこ, 講談社, (1977).