

藤崎モデルに基づく対話音声の句末音調の操作が パラ言語情報の知覚に与える影響*

○森 大毅, 渡邊 諒馬 (宇都宮大)

1 はじめに

対話音声伝達するパラ言語情報と韻律的特徴との関係解明のためには、藤崎モデルパラメータの分析が有力な手段となり得る。これは、指令の生起という離散的な事象が語彙・構文といった言語情報に対応する一方で、指令の大きさや位置のパラメータという連続的な事象が話者のパラ言語メッセージや心理状態を反映して変化すると予想されるからである。我々はこれまで、対話音声合成におけるパラ言語情報の制御を目的として、対話音声から抽出した藤崎モデルパラメータとパラ言語情報との関係のモデル化を検討してきた。しかし、統制が不可能な自発音声に基づく分析では、あらゆるパラ言語情報空間を覆い尽くすサンプルは到底得られないため、各々の指令値を直接モデル化することは困難であった。

本論文では、分析合成音声の知覚実験からこの問題に迫る。具体的には、対話音声から抽出した藤崎モデルパラメータを操作することでパラ言語情報の知覚がどのように変化するかを調べる。今回は句末音調に着目した。句末音調は自発音声に特徴的な現象であり、特に音声対話においてはパラ言語情報の有力な担い手であると予想されたためである [1]。

2 句末音調とアクセント指令

宇都宮大学パラ言語情報研究向け音声対話データベース (UADB) [2] に収録されている発話のうち、手作業により藤崎モデルパラメータを付与した [3] 2605 発話を研究の対象とした。上昇調または上昇下降調の句末音調は、アクセント指令により表現されている。そのうち 459 発話 (17.6%) の発話末にはただ 1 つのアクセント指令が付与された。また、63 発話 (2.4%) の発話末には 2 つの連続するアクセント指令が付与された。これは、1 つのアクセント指令では f_0 軌跡を十分に再現できなかったためである。

Fig. 1(a)(b) に、それぞれ発話末が上昇調および

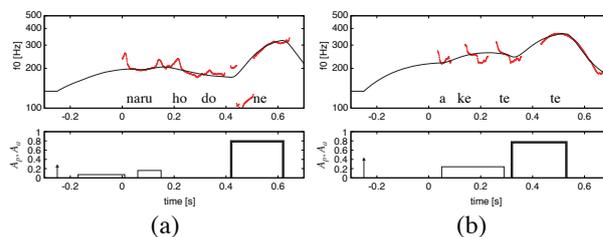


Fig. 1 (a) 上昇調を持つ発話例「なるほどね」
(b) 上昇下降調を持つ発話例「開けてて」

び上昇下降調の音声の f_0 軌跡および藤崎モデルで近似した f_0 曲線を示す。句末音調に対応するアクセント指令は太線で示されている。

3 句末音調の操作

パラ言語情報の知覚実験のため、句末音調に対応するアクセント指令の大きさ A_{aBPM} を可変、その他のパラメータを固定として、句末音調だけを操作した f_0 軌跡を生成し、STRAIGHT 合成により刺激を作成した。

対象とする発話は、句末音調にただ 1 つのアクセント指令が付与された上昇調を持つ 16 発話および上昇下降調を持つ 19 発話である。上昇調の発話末は全て終助詞「ね」、上昇下降調の発話末は格助詞「が」「で」「に」接続助詞「て」係助詞「は」「って」助動詞「で」のいずれかである。

A_{aBPM} の設定範囲は、変更前の値によらず 0.2, 0.4, 0.6, 0.8, 1.0, 1.2 の 6 通りとした。 $A_{aBPM} = 0.2$ は句末音調がほぼ消失した状態に相当する。

Fig. 2 に、上昇調の例 (Fig. 1(a)) および上昇下降調の例 (Fig. 1(b)) に対応する句末音調操作 f_0 軌跡を示す。

4 パラ言語情報知覚実験

上昇調実験用 (16 発話 \times 6 = 96 + ダミー - 28) および上昇下降調実験用 (19 発話 \times 6 = 114 + ダミー - 19) の 2 種類の呈示刺激セットを用意し、呈示順序はセット内でランダム化した。

被験者は音声研究室に所属する大学院生 4 名

*Boundary tone manipulation for conversational speech based on the command-response model and its effects on the perception of paralinguistic information.
by MORI, Hiroki, WATANABE, Ryoma (Utsunomiya University)

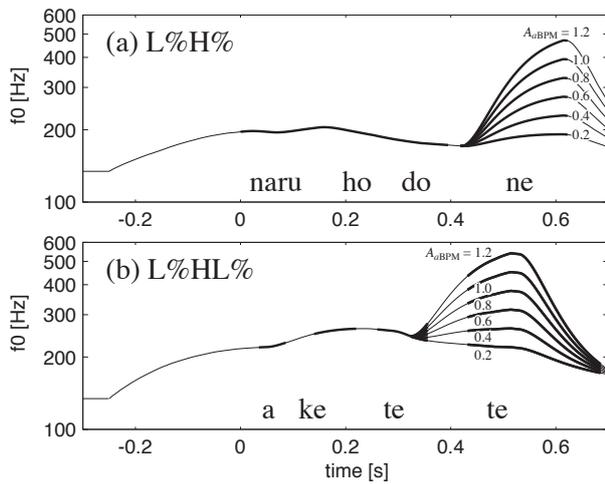


Fig. 2 句末音調を操作した f_0 軌跡の例 (a) 上昇調 (b) 上昇下降調

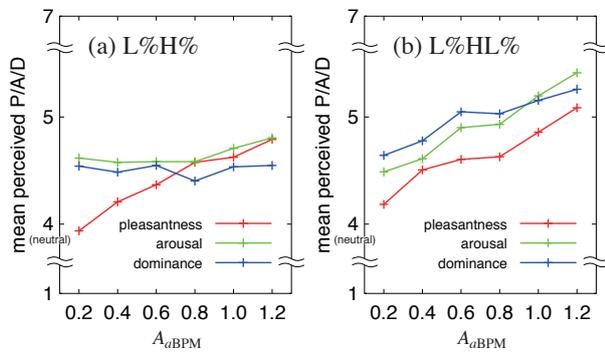


Fig. 3 合成音声から知覚されるパラ言語情報の平均評価値 (a) 上昇調 (b) 上昇下降調

および大学生5名(全員男性)である。被験者は、1音声ずつ聴取した後に、知覚されるパラ言語情報を以下の抽象次元について1から7の7段階で評価するよう教示された。

- 快-不快 (pleasantness)
- 覚醒-睡眠 (arousal)
- 支配-服従 (dominance)

A_{aBPM} の各条件における全発話の平均評価値を Fig. 3 に示す。図から、上昇調については、 A_{aBPM} を大きくすると、より快寄りに知覚される傾向が見られる。また上昇下降調については、 A_{aBPM} を大きくすると、より快寄り・覚醒寄り・支配寄りに知覚される傾向が見られる。発話および A_{aBPM} を要因とする分散分析の結果、上昇調実験では快-不快の次元で A_{aBPM} の主効果が有意 ($p < .01$) であり、交互作用は有意ではなかった ($p > .05$)。また、上昇下降調実験では全ての次元で A_{aBPM} の主効果が有意 ($p < .01$) であり、快-不快および覚醒-睡眠の次元で交互作用が有意

であった ($p < .05$)。

5 考察

発話末の句末音調の操作がもたらす韻律的特徴の変化は局所的であるにも関わらず、発話全体から知覚されるパラ言語情報には明確な変化が認められた。特に、有効な音響関連量が少なく音声合成の点では制御が難しい [4] 快-不快の知覚への影響は重要である。今回の結果は、上昇調の「ね」で終わる音声の合成については比較的簡単に快-不快の程度を操作できることを意味する。一方、覚醒-睡眠の知覚については発話全体の f_0 および強度が、支配-服従の知覚については言語情報がそれぞれ支配的であったことが、句末音調の操作の影響が認められなかった原因の可能性がある。

上昇下降調で終わる発話については、全ての次元で知覚されるパラ言語情報への影響が認められ、上昇調の場合と同様に句末音調の大きさが音声合成への応用の点で有効な音響関連量であると考えられる。快-不快および覚醒-睡眠の次元では発話との交互作用が認められた。著者の印象では、発話末以前の音声で覚醒寄りあるいは睡眠寄りに知覚される明確な特徴を持つ場合は、句末音調の操作に影響を受けない傾向があるように感じられた。

6 おわりに

発話末の句末音調を表現するアクセント指令の大きさの操作により、対話音声のパラ言語情報の知覚を変化させることができることを示した。

f_0 軌跡は、指令の大きさだけでなく生起位置にも影響を受ける。今後は、生起位置も含めた句末音調に関連するパラメータの予測/変換規則を定式化し、対話音声合成におけるパラ言語情報の制御に応用したい。

謝辞 本研究は JSPS 科研費 26280100 の助成を受けている。

参考文献

- [1] 岩田他, 信学技報, SP2012-77, 2012.
- [2] Mori et al., Speech Communication **53**, 36–50, 2011.
- [3] 渡邊他, 音講論 (春), 511–512, 2013.
- [4] 高橋他, 音講論 (秋), 355–356, 2013.