

## 音声駆動型身体引き込みキャラクタ InterActor

2V-03

## の話者音声特性への影響

檀原 龍正 南尾 勝彦

岡山県立大学大学院情報系工学研究科

渡辺 富夫 大久保 雅史

岡山県立大学情報工学部, CREST

## 1. はじめに

対面コミュニケーションでは、単に言葉だけでなく、顔みや身振りなどの身体によるリズムを共有して、互いに引き込むことでインタラクションしている。この身体性の共有が対話時の一体感を生み、対話相手との関わりを実感させている。情報機械を介してのコミュニケーションにおいても、円滑なインタラクションを実現するためには、このメカニズムの導入が不可欠である<sup>[1]</sup>。

著者らは既に、聞き手と話し手の両機能を有し、豊かなコミュニケーション動作を音声のみから自動生成する音声駆動型身体引き込みキャラクタ InterActorを開発して、コミュニケーション支援に有効であることを示した<sup>[2]</sup>。

本論文では、InterActorに語りかけたときの話し手の音声、とくに発話音声の時間パラメータに着目し、聞き手としてのInterActorが話し手の音声特性にどのような影響を与えているのかを定量的に解析している。

## 2. インタラクションシステムの概要

本システムのコンセプトを図1に示す。ここでPCのディスプレイ上に描画されるInterActorは、話し

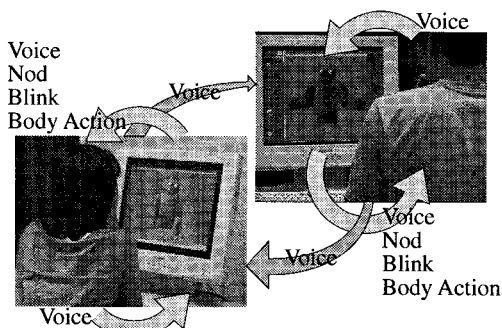


図1 身体的インタラクションシステムのコンセプト

手と聞き手の両機能を備えている。ユーザが InterActor に語りかけると、InterActor が聞き手として顔みや瞬きや身振りなど身体全体で反応することで円滑に話しかけることができる。その対話者の音声をネットワーク等を経由して対話相手方の InterActor に送信する。InterActor はその音声時系列に基づいて話し手としてのコミュニケーション動作をその音声に関連付けて生成し、音声と動作を同時に提示することで発話情報を効果的に対話相手に伝達することができる。この操作を相互に連続して行うことで、円滑なコミュニケーションが実現される。

## 3 発話音声の解析

## 3.1 実験方法

InterActorの音声に関連付いたコミュニケーション動作が、話し手の音声そのものに与える影響を検証するためにInterActorを用いた発話実験を行った。

留守番電話にメッセージを録音する場面を想定し、被験者に表1に示すメッセージをInterActorに対して語りかけさせ、その発話音声を記録した。その際、InterActorの状態として、反応がない場合と反応がある場合の2種類の条件を設定した。実験手順は、まず反応がない場合で行い、続けて反応がある場合で行った。これを1セットとして、1人の被験者に対して3セット繰り返し行った。被験者は学生26名である。

表1 メッセージ

もしも岡山県立大学の〇〇です。  
本日の論文の打ち合わせですが  
明日の3時に変更になりました。  
よろしくお願ひします。

\* 〇〇部分は被験者名

Effects of Speech Driven Embodied Entrainment Character "InterActor" on Talker's Speech Parameters  
Ryusei Danbara<sup>1</sup>, Katsuhiko Minamio<sup>1</sup>, Tomio Watanabe<sup>2,3</sup>, and Masashi Okubo<sup>2</sup>

\*1: Graduate School of System Engineering, Okayama Prefectural University

\*2: Faculty of Computer Science & System Engineering, Okayama Prefectural University

\*3: CREST, Japan Science and Technology Corporation  
111 Kuboki, Soja, Okayama, 719-1197, Japan

## 3.2 実験結果

## 3.2.1 発話時間

発話実験において、被験者が話し始めてから話し終わるまでの発話時間を計測した。図2(a)に26名の発話時間の平均値と標準偏差を各セットごとに示す。全てのセットにおいてInterActorの反応がある場合がない場合に比べて発話時間が長いことがわかる。各セットごとに対応関係のある平均値

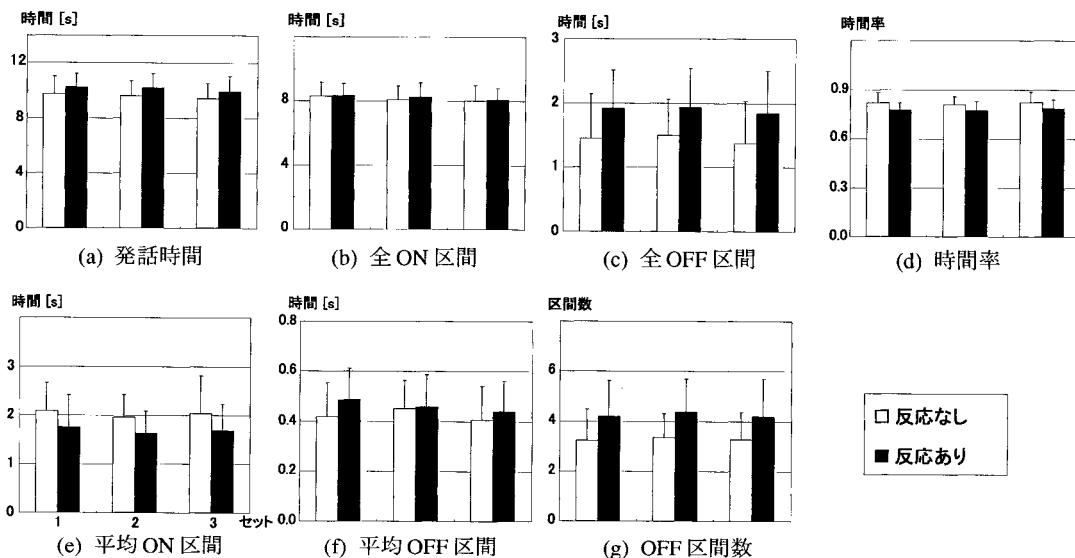


図2 発話音声の解析結果

の差の検定をした結果、全てのセットで有意水準1%で有意差が認められた。

次に、発話区間内におけるON区間(有音区間: talkspurt duration)のみを抽出したときの平均値と標準偏差を図2(b)に、OFF区間(無音区間: silence duration)のみを抽出したときの平均値と標準偏差を図2(c)に示す。ON区間に関しては、反応の有無による顕著な差は見られない。一方、OFF区間では反応がある場合で長くなっており、全てのセットで有意水準1%で有意差が認められた。

以上の結果より、InterActorの反応がある場合には、反応がない場合に比べてOFF区間が長くなることで、全体として発話時間が延びていることがわかる。

### 3.2.2 時間率

時間率は、発話区間内でのON区間の占める割合をさす代表的な発話の特徴パラメータで、平均ON区間/(平均ON区間+平均OFF区間)によって表現される<sup>1)</sup>。図2(d)に発話実験における時間率の平均値と標準偏差を示す。全てのセットにおいて、反応がある場合とない場合の間に有意水準1%で有意差があり、時間率の低下が認められた。また、平均ON区間の平均値と標準偏差を図2(e)に、平均OFF区間を図2(f)に示す。反応がある場合で平均ON区間が短くなっていることがわかる。全てのセットで、有意水準1%で有意差が認められた。一方、平均OFF区間では顕著な差は認められなかった。さらに、発話区間におけるOFF区間数の平均値と標準偏差を図2(g)に示す。全てのセットで、有意水準1%で有意差が認められ、反応がある場合にOFF区間の数が増えていると言える。

以上の結果より、発話時間が延びる要因として呼気段落区分でのOFF区間が大きく影響しており、その数が増えることによって発話時間が延びていることがわかる。これは、発話者がインタラクションを確認しながらInterActorに語りかけるためであると考えられる。

## 4 まとめ

本論文では、InterActorを用いた発話実験により、InterActorのコミュニケーション動作が発話者の音声特性に与える影響を解析した。その結果、OFF区間の増大とそれに伴い発話時間が延びることを示した。このようなOFF区間の変化を利用することで音声認識率を向上させるなど、コミュニケーション支援への応用が考えられる。

なお、本システム開発は、著者らが提唱している、心が通う身体的コミュニケーションシステムE-COSMIC (Embodied Communication System for Mind Connection) プロジェクトの一環で行われた。

## 参考文献

- [1] 黒川隆夫, 渡辺富夫: ノンバーバルコミュニケーションとインタフェース; ヒューマンインタフェース学会誌 Vol.3, No.2, pp.91-98 (2001).
- [2] 渡辺富夫, 大久保雅史, 中茂睦裕, 檀原龍正: InterActorを用いた発話音声に基づく身体的インタラクションシステム; ヒューマンインタフェース学会論文誌, Vol.2, No.2, pp.21-29 (2000).
- [3] 渡辺富夫: 人間-機械間の音声対話のための時間率適応化システム; 計測自動制御学会論文集, Vol.26, No.8, pp.902-907 (1990).