

発話蓄積を表現するシャボン膜によるアバタを介した コミュニケーションシステム

A Communication System

via Avatars with Soap Film Expressing Utterance Accumulation

桂木 章吾[†] 石井 裕^{††} 渡辺 富夫^{††}

Shogo Katsuragi[†] Yutaka Ishii^{††} Tomio Watanabe^{††}

[†] 岡山県立大学大学院 情報系工学研究科 ^{††} 岡山県立大学 情報工学部

1 はじめに

近年の情報機器の高性能化に伴い、人とコミュニケーションを図ることを目的に、様々な音声インタラクションシステムやコミュニケーションシステムの研究開発が進められている [1, 2, 3]. 対話者の身体性の共有に着目し、共創を支援するシステム開発も行われており [4], 人とのインタラクションを行うコミュニケーションシステムの研究開発において、身体的なかかわりは重要な役割を果たすと考えられる.

実際、人間同士のコミュニケーションでは、バーバル情報だけでなく、音声に対するうなずきや身振り・手振りなどのノンバーバル情報が相互に時系列的に同調して引き込み合うことで、一体感を生み出している [5, 12]. この一体感により、場が盛り上がり、共感を深めていくことで親密なコミュニケーションが形成される. しかし、人との会話の際に、コミュニケーション導入後の当事者同士の会話を継続することが困難な場合がある. このコミュニケーションにおける緊張状態は、円滑な会話や一体感形成の妨げとなる [6]. この緊張・不安状態は、他者への意識といった他者との関係が原因で生じる場合がある [7, 8]. さらに、近年利用が拡大されている遠隔コミュニケーションにおいて、空間的に分離された状態では、コミュニケーション場の共有が困難になる可能性から [9, 10], 対話者同士が対話空間共有・対話者相互の身体的インタラクションを把握しやすい提示手法を検討する必要がある [11].

著者らは、これまでに会話音声と身体動作が同期する身体引き込みに着目し、インタロボット技術 iRT (InterRobot Technology) を開発してきた [12]. iRT は、音声と身体動作の関係をモデル化することで発話音声からコミュニケーション動作を自動生成し、身体引き込みによりインタラクションを円滑にし、コミュニケーションを支援するものである. すでに、発話音声からコミュニケーション動作を自動生成する音声駆動型身体引き込みキャラクタ InterActor に実装することで、システムの有効性を示している [13]. また、話し手から聞き手への発話伝達・蓄積表現を有する InterActor を用いた自由発話実験を行った結果、発話伝達・蓄積表現および語りかけの区切り表現を導入することによって、ユーザの発話意欲増大に有効であ

ることを示している [14].

本研究では、視覚的な空間共有感提示のために、話し手と聞き手の機能を有し、各ユーザのアバタとなる InterActor を包み込む、シャボン玉を模した CG オブジェクトを構築している. また、アバタを介した音声通話において、この CG オブジェクトによる発話蓄積表現を用いた発話促進および一体感形成支援を行うコミュニケーション支援システムを開発している. さらに、システムを用いた発話促進および一体感形成効果について、自由対話と役割対話の2種類の実験を行い、官能評価から検討を行っている.

2 コンセプト

本システムのコンセプトを図1に示す. 話し手の発話の伝達とコミュニケーション場への蓄積を可視化させることで、話し手は聞き手とともにコミュニケーション場を作り上げていることが認識できる. ここで話し手と聞き手によって作り上げる、コミュニケーション空間を「共存在空間」と定義する. 話し手の発話および聞き手のうなずきに伴って両者間に存在する閉空間が膨張し、両者を包み込むことで共存在空間となり、一体感が感じられる. さらに、本システムには対話の区切りを設ける [14]. システム側から対話の区切りを提示することで、ユーザは繰り返し両者で共存在空間を作り上げていく感覚が得られ、その都度ユーザは信頼関係の強化を感じることができると考えられる. 以上より、共存在空間を構築することで一体感が生まれ、緊張感の緩和や発話促進およびコミュニケーション支援につながる.

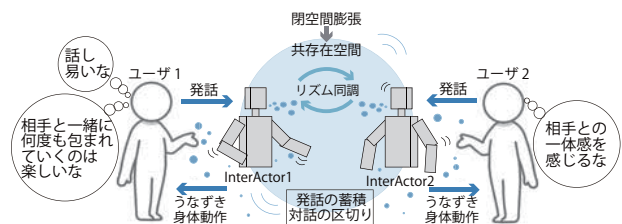


図 1: コンセプト

3 システム開発

3.1 システム概要

本システムは仮想空間を Unity(2017.1.0) で構築しており、描画速度は 30fps である。インタラクションモデルは、音声の ON-OFF パターンに基づくうなずき反応モデルおよび腕部・胴部を動作させる新規モデルを導入している。音声データ $V(i)$ は 8bit, 44.1kHz でサンプリングし、閾値で二値化するとともに、音節間の短時間の無音区間による発話の断片化を除去するために 133msec でハングオーバー処理を施している。

先行研究により、対面コミュニケーション時における相手と自己の身体性の共有を投影することと、呼吸情報の提示が有効であることが検証されている [15]。発話時の音声の ON-OFF (バーストポーズ) の区切りは呼吸段落区分と呼ばれ、呼吸と密接な関係があることから、ユーザの発話状態をハングオーバー処理後の ON-OFF に対応した呼気・吸気の 2 種類で視覚化する。

インタラクションモデルの構築のために、人の対面コミュニケーションでの音声と身体動作の関係を解析した結果、うなずきと頭部の動きには強い相関があることから、聞き手の身体動作はうなずきを主体としていることが明らかになった [16]。その解析結果に基づき、本システムの両キャラクターは応答としてうなずきを行う。

本システムの使用画面を図 2 に示す。発話促進につながる対話空間および一体感形成支援のための空間共有を実現するために、本研究ではシャボン玉オブジェクトを採用した。本システムのユーザは 2 名であり、ユーザ間での音声通話を可能としている。

本システムでは閉空間を「シャボン膜」、発話の伝達および対話を積み重ねた度合いの視覚提示を「バブル」として表現する。シャボン膜は 2 者間中央上空に存在し、発話およびうなずきに伴って膨張する。シャボン膜は両者間の距離を直径とした大きさで膨張を停止する。その後、本システムでは対話の区切りを表現するため、両者が「上昇・到達」する。ユーザらは、各シャボン膜へ向けて上昇・到達することで対話の区切

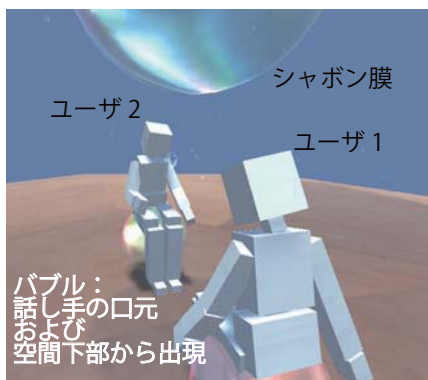


図 2: システム使用画面

りを認識可能となる。両者が上昇し、膨張後のシャボン膜の横に到達した時、シャボン膜が膨張し、両者を包み込むことで一体感を共有できる。すなわち、共存空間を構築することで一体感が生まれ、緊張感の緩和や発話促進およびコミュニケーション支援につながる。シャボン膜の状態の遷移を図 3 に示す。両者がシャボン膜に包み込まれた後、上空にある次の小さなシャボン膜の膨張に移る。また、仮想空間下部から上昇するバブルの量が増え、対話を積み重ねた度合いを視覚提示する。ユーザらは以上を繰り返し体験する。

3.2 発話音声に基づくアバタの動作生成

3.2.1 うなずき予測モデル

本システムにおける両キャラクターは、iRT のうなずき予測モデル [12] に従ってうなずき動作を行う。うなずき予測モデルを図 4 に示す。うなずきの予測モデルはマクロ層とミクロ層からなる階層モデルである。マクロ層では音声の呼吸段落区分での ON-OFF 区間からなる $[i]$ 番目のユニット区間にうなずきの開始が存在するかを予測する。 $[i-1]$ ユニット以前のユニット時間率 $R(i)$ (ユニット区間での ON 区間の占める割合, (2) 式) の線形結合で表される (1) 式の MA(Moving-Average) モデルで、 J (定数: 6) 個のパラメータ $a(j)$ を用いて予測する。

$$M_u(i) = \sum_{j=1}^J a(j) R(i-j) + u(i) \quad (1)$$

$$R(i) = \frac{T(i)}{T(i) + S(i)} \quad (2)$$

$a(j)$: 予測係数

$T(i)$: i 番目ユニットでの ON 区間

$S(i)$: i 番目ユニットでの OFF 区間

$u(i)$: ノイズ (ここでは 0)

$$M(i) = \sum_{k=1}^K b(k) V(i-k) + w(i) \quad (3)$$

$b(k)$: 予測係数

$V(i)$: 音声 ON-OFF データ

$w(i)$: ノイズ (ここでは 0)

予測値 $M_u(i)$ がある閾値を越えて、うなずきが存在すると予測された場合には、処理はミクロ層に移る。ミクロ層では過去 2 秒間の音声の ON-OFF データ (30Hz, 60 個) を入力とし、(3) 式の MA モデルで、 K (定数: 60) 個のパラメータ $b(k)$ を用いてうなずきの開始時点を推定する。予測値が閾値を越えた場合には両キャラクターをうなずかせる。

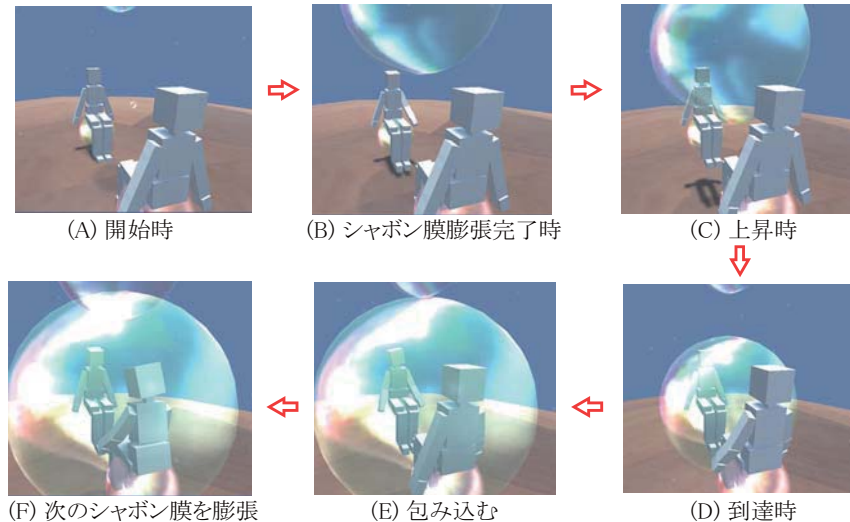


図 3: シャボン膜の状態遷移

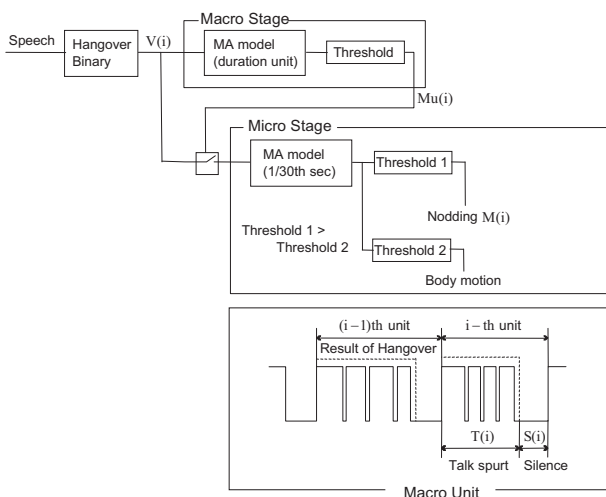


図 4: うなずき予測モデル

3.2.2 各アバタの動作生成

両キャラクターの腕部・胴部に関する動作モデルは、音声の ON-OFF 変化に基づく動作モデルを新規に開発し、導入している。

発話音声に基づくキャラクター動作について、まず話し手動作を説明する。画面手前に位置するキャラクターが、ユーザに対応した話し手キャラクターとなる。ハングオーバー処理後の呼気段落区分での音声 ON-OFF の変化に伴って話し手キャラクターが動作する。ユーザの音声 ON 区間において話し手キャラクターは徐々に前傾しつつ、徐々に腕部を身体の外側へ押し出すように動き、音声 OFF 区間においては徐々に後傾し、徐々に腕部を身体側へ引き戻すように動くことで、音声 ON-OFF と対応させている。ユーザが発話中、話し手キャラクターの口元からシャボン膜へ向かってバブルが発生し、

発話の伝達を示す。

次に聞き手動作を説明する。聞き手キャラクターは、相手ユーザの音声に伴って、聞き手動作を自動生成する。さらに相手ユーザの音声 ON 区間において聞き手キャラクターは徐々に前傾し、音声 OFF 区間では徐々に後傾することで、話し手の動作と同調させている。また、うなずきの予測値に基づく身体動作モデルを用いて、うなずき動作および腕部・腕部を動作させる。

以上より、ユーザらの音声 ON-OFF に伴って両キャラクターが動作することにより、両者の発話状態を把握可能となり、リズム同調が促進される。画面手前に位置するキャラクターがユーザ、奥側に位置するキャラクターが対話相手に対応している。両キャラクターの腕部・胴部の動作タイミングはユーザの音声 ON-OFF 変化によって決定されるが、両キャラクターの腕部動作の角度・速度は違和感の無い範囲でランダムに決定される。また、両キャラクター腕部動作において、初期角度および設定した到達角度を超えて前傾・後傾することはない。

3.3 システム構築

本システムはオプション的にモード選択が可能である。アバタの動作生成におけるモード選択について説明する。本システムは音声通話を用いた双方向的なコミュニケーション支援システムであるが、対話相手を存在させず、傾聴システムへの語りかけを目的とした、「傾聴」モードも選択可能である [17]。このモードでは、相手キャラクターは聞き手動作のみ行う。

各アバタ動作およびシャボン膜を組み合わせたシステムを構築し、開発システムを学生 5 名程度に試用させたところ、シャボン膜が膨張し、両キャラクターを包み込む過程において、一連の視覚表現がユーザに圧迫感やパーソナルスペースの拡大を意識させてしまい、インタラクションを阻害する可能性があるとの意見があった。そこで、両キャラクターがシャボン膜に包み込まれることによるコミュニケーション支援効果を検証



図 5: システムを使用した実験の様子

する必要がある。

4 シャボン膜による包み込み表現の評価実験

4.1 実験方法

システムの有効性を検証するために、各実験参加者へ、以下の2モードを使用させ、自由対話と役割対話で評価実験を行った。

- A: シャボン膜は一定の大きさで膨張を停止した後、両者を包み込まない
- B: シャボン膜は一定の大きさで膨張を停止した後、両者を包み込む

実験の様子を図5に示す。実験参加者はノートPC前でヘッドフォンを装着させて話させた。実験者は説明時を除いて室外で待機した。実験参加者には最初に、システム全体と各モードの特徴を説明した。モード説明の提示順序はカウンタバランスを取って順序を変更した。その後、実際に対話させることで各モードを体験させた。システムに慣れさせるため、モードごとに3分ずつ使用させた。その後、2種類の評価方法を用いて比較させた。実験参加者は、同性友人同士の18～24歳の学生24名(男性:12名、女性:12名)であった。

1つ目の評価方法は一対比較実験である。自由対話実験を行い、A・Bのモードから2つを一対比較させた。提示順序についてもカウンタバランスをとって順序を変更した。まず、1つのモードで対話を行った後に別のモードによる対話を行い、2つのモードを使用した後に二者択一させた。1つの比較につき発話時間を4分(1モードにつき2分)とした。

2つ目の評価方法は7段階評価実験である。まず自由対話実験を行い、A・Bモードを、コミュニケーション機能に関する一般的な項目として「楽しさ」、「対話しやすさ」、「場の盛り上がり」、またシステムを介した受話、応答反応による安寧の感覚に関する項目として「安心感」、「一体感」、発話促進に関する項目として「発話意欲」の6項目について評価させた。各モー

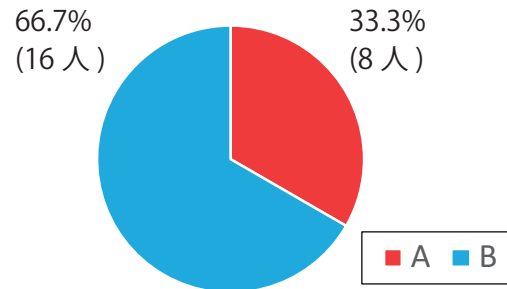


図 6: 自由対話実験における一対比較結果

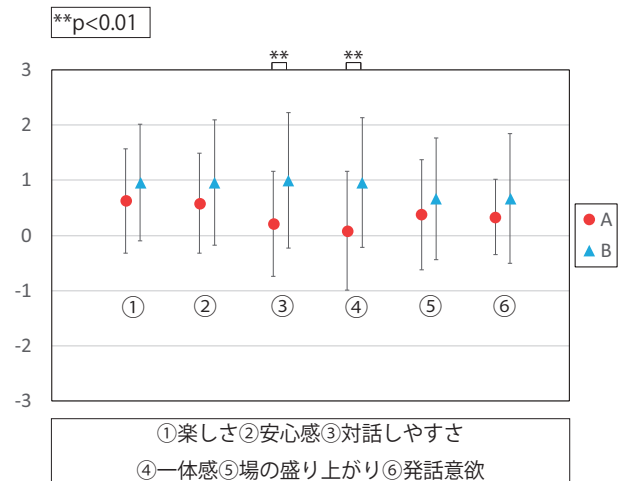


図 7: 自由対話実験における7段階評価結果

ドを使用した後、どの程度その項目に当てはまるかを「全く感じない」から「非常に感じる」までの7段階のリッカート尺度(中立0)の官能評価でアンケート用紙に記入させた。また、各モードに対する意見を自由記述形式で記入させた。提示順序はカウンタバランスを取り、各モードにつき対話時間を2分とした。

次に、話し手と聞き手に分かれた役割対話実験を行い、自由対話実験と同じくA・Bモードを2分ずつ使用させた。こちらもモード終了毎に7段階評価アンケートを行い、上記の6項目の他に、話し手側には「相手に話が伝わったと感じたか」、聞き手側には「相手の話を理解できたか」という項目を加え計7項目を調査した。話し手の準備時間を4分間設けた後、本実験を開始した。本実験では話し手の話す内容についての指示は出していない。また、話し手・聞き手の役割交代は行っていない。

4.2 実験結果

一対比較の結果を図6に示す。Bモードが高く評価されていることがわかる。

自由対話実験の7段階評価アンケートによる6項目の評価結果を図7に示す。Wilcoxonの符号順位検定によって評価した結果、「対話しやすさ」、「一体感」の2項目において、有意水準1%で有意差が認められ、Bモードが高く評価された。

役割対話実験の7段階評価アンケートによる聞き手

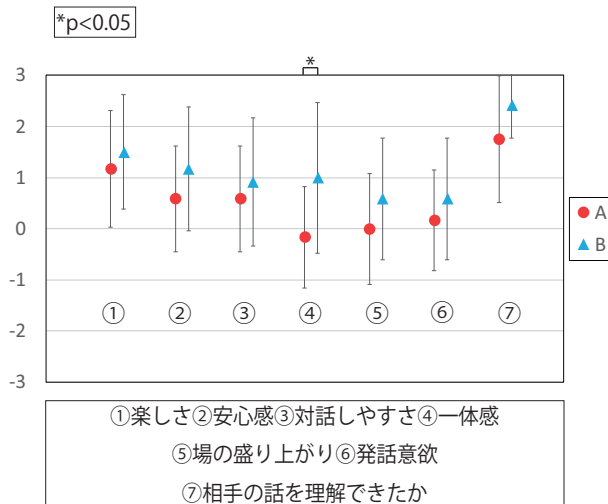


図 8: 役割対話実験における聞き手の 7 段階評価結果

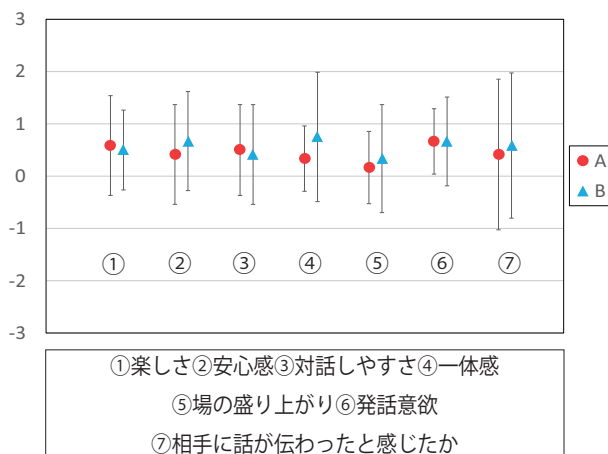


図 9: 役割対話実験における話し手の 7 段階評価結果

側と話し手側の 7 項目評価結果をそれぞれ図 8, 図 9 に示す。Wilcoxon の符号順位検定によって評価した結果、聞き手側の「一体感」の項目において有意水準 5% で有意差が認められ、B モードが高く評価された。

また、自由記述欄に寄せられたシステムに対する意見を表 1 に示す。各モードの特徴が反映されているのがわかる。

4.3 考察

実験の結果、B モードが A モードに比べて高く評価された。7 段階評価においても、「対話しやすさ」、「一体感」の項目で高く評価された。これらの結果は自由記述欄の意見より、両キャラクターが包み込まれることによって、心理的なつながりの強化を視覚表現したためであると考えられる。逆に、両者を包み込むことなく両者間にシャボン膜が存在する A モードの自由記述欄では、「相手との間にシャボン玉があって邪魔に感じた」との意見があった。よって、仮想空間上での対話において、両者を俯瞰的に見える状況であったとして

表 1: システムに対する意見

モード	内容
A	・視覚的に少し相手との距離を感じる ・相手との間にシャボン玉があって邪魔に感じた ・対話しやすさ、盛り上がりには欠ける ・あまり一体感が得られなかった ・安心感を持った ・Bモードに比べ対話を楽しめた。リズム感のようなものがあった
B	・包まれると安心感がある ・シャボン玉の中に居ると相手と同じ場所に居る気持ちになって話しやすかった。相槌も打ちやすかった ・シャボン玉が膨らんで包まれる方が、見ていて自分の心理状態を表しているようで良かった ・話が進み盛り上がっていくとシャボン玉が膨らみ、高さも上昇するので、目で見て盛り上がっていると感じ、話しやすくなった ・(役割対話:聞き手側)ただ聞いているだけでも視覚的に一緒になっているので、一体感があり好き ・(役割対話:聞き手側)話し手側の盛り上がりを感じられ、聞いている側も楽しく、もっと話を聞きたいと思った ・(役割対話:話し手側)一方的に話していても2人がシャボン玉に包まれると一体感を感じ取れた ・Aモードとの違いがわからなかった

も、両者間に何かしらのオブジェクトが存在することは、対話の障害とみなされる可能性がある。これは、現実空間においても両者間について存在するとリズム同調を阻害することから [18], オブジェクトが相互インタラクションを阻害する壁のような役割を果たすためであると考えられる。

本実験では、シャボン膜に包み込まれることによる圧迫感やパーソナルスペースの拡大について言及した意見は見られなかった。よって、3.3 節で問題視した、シャボン膜膨張によるインタラクション阻害感覚は、シャボン膜が両者を包み込まない程度に大きくなり、対話相手とのインタラクションを視覚的に遮るときに発生するものと考えられる。

B モードを高く評価しなかった実験参加者の内 4 名は自由記述欄において、「2 つのモードの差がわからなかった」と述べている。よって、単純な膨張だけでは包み込む表現としては不十分である可能性があるため、今後の検討が必要である。

5 おわりに

本研究では、発話促進および一体感形成支援のための視覚的な空間共有感提示を実現するため、各ユーザに対応した InterActor を包み込む、シャボン膜による発話蓄積表現を用いたコミュニケーション支援システムを開発した。システムを用いた発話促進および一体感形成効果について、自由対話と役割対話の 2 種類の

実験を行い、官能評価から検討を行った。

参考文献

- [1] J. Cassell, "Embodied conversational interface agents", *Communications of the ACM*, Vol.43, No.4, pp.70–78, 2000.
- [2] 川本真一, 下平博, 新田恒雄, 西本卓也, 中村哲, 伊藤克亘, 森島繁生, 四倉達夫, 甲斐充彦, 李晃伸, 山下洋一, 小林隆 夫, 徳田恵一, 広瀬啓吉, 峯松信明, 山田篤, 伝康晴, 宇津呂武仁, 嵯峨山茂樹, "カスタマイズ性を考慮した擬人化 音声対話ソフトウェアツールキットの設計", *情報処理学会論文誌*, Vol.43, No.7, pp.2249–2263, 2002.
- [3] 鹿野清宏, CincarekTobias, 川波弘道, 西村竜一, 李晃伸, "声情報案内システム「たけまるくん」および「キタちゃん」の開発 (特別企画「音声認識デベロッパーズフォーラム」)", *情報処理学会研究報告*, 2006-SLP-63, pp.33–38, 2006.
- [4] 三輪敬之, "共創表現のダイナミクスー実践, 理論, システム技術ー", *共創学*, Vol.1, No.1, pp.23–30, 2019.
- [5] W. S. Condon, L. W. Sander, "Neonate movement is synchronized with adult speech", *Science*, Vol.183, pp.99–101, 1974.
- [6] 原田朋枝, 島田修, "社会的スキルの自己評価と対人不安との関連", *川崎医療福祉学会誌*, Vol.12, No.1, pp.75–81, 2002.
- [7] 伊藤由美, 丹野義彦, "対人不安についての素因ストレスモデルの検証ー公的自己意識は対人不安の発生にどう関与するのか", *パーソナリティ研究*, Vol.12, No.1, pp.32–33, 2003.
- [8] 遠藤由美, "自己紹介場面での緊張と透明性錯覚", *実験社会心理学研究*, Vol.46, No.1, pp.53–62, 2007.
- [9] 清水博, 久米是志, 三輪敬之, 三宅美博, "場と共創", NTT 出版, 2000.
- [10] 渡辺富夫, "心が通う身体的コミュニケーション", *日本機械学会誌*, Vol.121, No.1195, pp.14–17, 2018.
- [11] 尾上聡, 山本健太, 田中一晶, 中西英之, "遠隔対話者の身体動作の提示による音声コミュニケーションの円滑化", *情報処理学会論文誌*, Vol.54, No.4, pp.1462–1469, 2013.
- [12] T. Watanabe, "Human-entrained Embodied Interaction and Communication Technology", *Emotional Engineering*, Springer, pp.161–177, 2011.
- [13] R. Danbara, T. Watanabe, M. Okubo, "Effects of Speech-Driven Embodied Entrainment Character "InterActor" on Talker's Speech", *Transactions of the Japan Society of Mechanical Engineers, Series C*, Vol.71, No.721, pp.152–159, 2005.
- [14] 桂木章吾, 石井裕, 渡辺富夫, "話し手と聞き手を表す身体引き込みキャラクタを用いた発話伝達・蓄積表現システム", *ヒューマンインタフェース学会論文誌*, Vol.21, No.3, pp.293–302, 2019.
- [15] 渡辺富夫, 荻久保雅道, 石井裕, "身体的バーチャルコミュニケーションシステムにおける呼吸の視覚化と評価", *ヒューマンインタフェース学会論文誌*, Vol.3, No.4, pp.319–326, 2001.
- [16] T. Watanabe, M. Okubo, M. Nakashige, R. Danbara, "InterActor : Speech-Driven Embodied Interactive Actor", *International Journal of Human-Computer Interaction*, Vol.17, No.1, pp.43–60, 2004.
- [17] 桂木章吾, 石井裕, 渡辺富夫, "シャボン玉オブジェクトによる発話蓄積表現を用いた発話促進システムの開発", *ヒューマンインタフェースシンポジウム論文集*, pp.235–239, 2019.
- [18] T. Watanabe, M. Okubo, "Evaluation of the entrainment between a speaker's burst-pause of speech and respiration and a listener's respiration in face-to-face communication", *Proc. of 6th IEEE International Workshop on Robot and Human Communication*, pp.392–397, 1997.