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・蓄積表現の効果

桂木　章吾[*1]　　渡辺　富夫[*2]　　石井　裕[*2]

Effects of Utterances Transmission / Accumulation Expressions by Using Embodied Entrainment Characters Expressing Each Function of Speaker and Listener

Shogo Katsuragi[*1], Tomio Watanabe[*2] and Yutaka Ishii[*2]

Abstract – **We have developed a speech-driven embodied entrainment computer graphics character called "InterActor" which generates automatically communicative motions from the rhythm of speech, and have shown the effectiveness of the system for supporting embodied communication. However, even though conventional InterActor activates embodied interaction, it does not have expressions accompanying time transition such as accumulation of utterances and delimitation of speech. If you can confirm the achievement of the transition action, the motivation for that action will be strengthened. In this paper, we develop a system with embodied entrainment characters that visually presents the utterances transmission and accumulation accompanying the time transition from the speaker to the listener to motivate and promote the utterances longer. In addition, in the evaluation experiment for the effects of utterances transmission / accumulation expressions, the mode in which the utterances transmission / accumulation expressions were added to the InterActor corresponding to the conventional speaker and listener was highly evaluated.**

Keywords : **Nonverbal information, Embodied interaction, Communication support**

1.　はじめに

人の対面コミュニケーションでは，発話音声とうなずきや身振り手振り等の身体的リズムが相互に共有され，引き込み合うことで円滑にコミュニケーションが行われている．著者らは先行研究として，発話音声と身体動作の相関を利用し，発話音声からうなずきや身振り手振り等の身体引き込み動作を自動生成するインタロボット技術，iRT(InterRobot Technology) を開発している [1]．また，仮想空間上に聞き手の身体動作モデルを備えた，音声駆動型身体引き込み CG キャラクタ InterActor を開発し，その有効性を示している [2]．音声インタフェースとしての InterActor は，引き込み反応によってかかわりを実感することができる．しかし，これまでの InterActor には発話の蓄積や区切りといった時間経過に伴う表現が導入されていなかった．自らの行動に対する達成度が確認できれば，その行動に対する内発的動機が高められる．すなわち InterActor が発話蓄積量をユーザに示しつつ，対話の到達目標となる区切りを提示することは，ユーザの発話意欲増大に重要であると考えられることから，著者らは発話の伝達・蓄積や区切りといった時間経過に伴う表現を導入したシステムを開発した [3]．

本研究では，開発システムにおける発話の伝達・蓄積表現の効果を検証するために，話し手から聞き手への発話の伝達および，聞き手への蓄積それぞれを表現する身体引き込みキャラクタを新たに構築し，内容を制限しない自由対話による評価実験を行うことでシステムの有効性を示している．

2.　コンセプト

本研究のコンセプトを図 1 に示す．人は発話に考えや思いを乗せて対話している．ユーザの発話の流入と蓄積を可視化させることで，ユーザの話を聞き手キャ

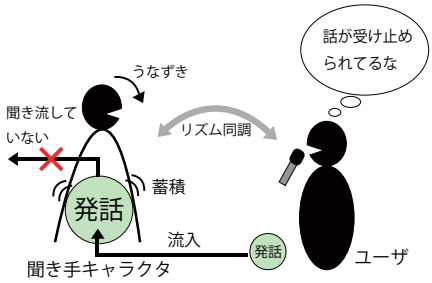


図 1　コンセプト
Fig. 1　Concept.

*1: 岡山県立大学大学院　情報系工学研究科
*2: 岡山県立大学　情報工学部
*1: Graduate School of Systems Engineering, Okayama Prefectural University
*2: Faculty of Computer Science and Systems Engineering, Okayama Prefectural University

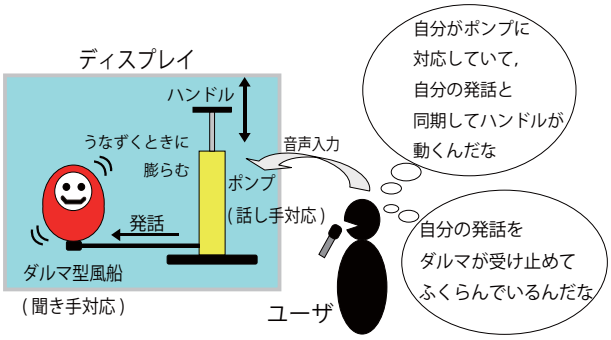


図 2　ダルマとポンプのコンセプト
Fig. 2　Concept of Dharuma and Pump.

ラクタがきちんと受け止めていることが提示される．ユーザの発話状態を視覚化するとともに，リズム同調を表す「うなずき」のタイミングで聞き手キャラクタに蓄積される様子を示すことによって，ユーザは自分の話を聞き流さずに受け止められている感覚が得られ，発話意欲が増大し，発話が促進される．また，ユーザが発話の伝達の総量を把握可能にすることで，対話中に自らが聞き手とどれだけ対話したかが認識でき，ユーザは対話に対する達成感が得られ，発話意欲が増大し，発話が促進される．以上より，発話の伝達を視覚的に提示する．

本研究ではユーザの発話の伝達を視覚化・抽象化するため，ユーザを空気入れ (ポンプ)，聞き手キャラクタをダルマ型風船 (ダルマ) に見立てた 3D キャラクタを用いる．ダルマとポンプのコンセプトを図 2 に示す．

3.　システム構成

3.1　システム概要

本システムは仮想空間を Unity で構築している．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 [4]．発話時の音声の ON-OFF (バーストポーズ) の区切りは呼気段落区分と呼ばれ、呼吸と密接な関係があることから，ユーザの発話状態を，発話・非発話に対応した呼気・吸気の 2 種類で視覚化する．

本システムでは聞き手キャラクタを風船のように膨らむオブジェクトとして表現したが，聞き手に iRT を実装するため，球形で，身体動作となる「揺れ」と「うなずき」を行っても違和感のない存在として，転んでも起き上がる「ダルマ」のイメージを聞き手キャラクタとして採用した．ユーザをポンプに見立てることで，ユーザの発話を取得・移行させる役割を持つ．発話の移行は，ポンプからダルマへ伸びたホースを移動する空気の球という形で表現している．以上より，ユーザがポンプに自己投影し，ダルマとかかわっている感覚

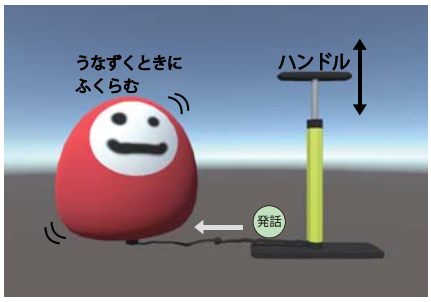


図 3　システム使用画面
Fig. 3　Screen of the system.

図 4　システム使用場面
Fig. 4　Example of the scene using the system.

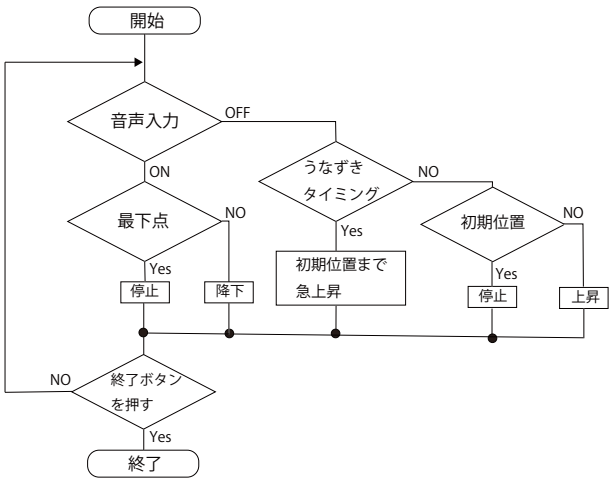
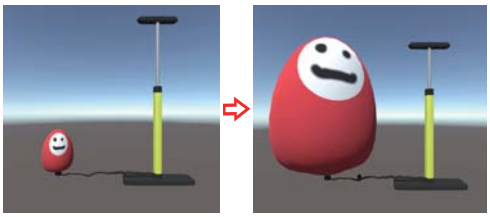


図 5　ハンドル動作のフロー図
Fig. 5　Flow diagram of handle operation.

が得られる．ポンプが風船に空気を入れるように，発話の流入を示し，ダルマがうなずきの累積回数に応じて，流入した発話を吸収し膨らむことで，発話の伝達と蓄積を視覚的にユーザに提示する．本システムの使用画面と使用場面をそれぞれ図 3，図 4 に示す．

話し手に対応したキャラクタであるポンプのハンドル動作について説明する．本システムでは音声 ON/OFF の変化に伴ってポンプの状態が注入，排出の 2 種類に分類される．使用者が発話するとポンプのハンドルは最下点に到達するまで降下し，発話をやめるとハンドルは初期位置までゆっくりと上昇する．さらに，ダルマのうなずきに同期してハンドルは初期位



(A) 通常時　　(B) 膨張時

図 6　ダルマの傾聴状態の遷移
Fig. 6　Transition of Dharuma's listening state.

置まで急上昇する．ハンドルの動作フロー図を図 5 に示す．また，ホースを移動する空気の球は，ハンドル降下時にポンプからダルマへ移行するよう対応付けられている．

ダルマは傾聴，待機の 2 種類の状態を持つ．とくに傾聴時には通常の傾聴動作の他に，膨張を行う．ダルマの傾聴状態の遷移を図 6 に示す．使用者がマイクに話しかけると，傾聴状態のダルマは揺れやうなずきを行い，うなずき時に膨張する．

ダルマ傾聴時のうなずきモデルを図 7 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気段落区分での ON-OFF 区間からなる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．予測には $[i-1]$ 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) モデルを用いる．

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

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

$a(j)$: 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
$u(i)$: 雑音

$$M(i) = \sum_{j=1}^{K} b(j)\, V(i-j) + w(i) \tag{3}$$

$b(j)$: 予測係数
$V(i)$: 音声データ
$w(i)$: 雑音

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．ミクロ層では音声の ON-OFF データ（30Hz，60 個）を入力とし，(3) 式を用いて MA モデルでうなずきの

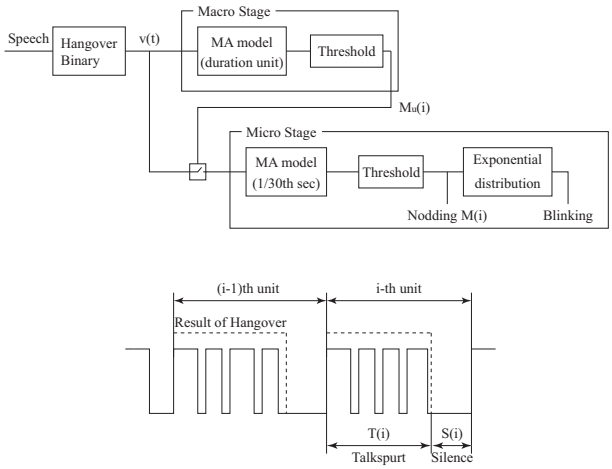


図 7　うなずきモデル
Fig. 7　Nodding model.

開始時点を推定する．予測値が閾値を越えた場合にはダルマをうなずかせる．

4.　発話促進システム評価実験

4.1　実験方法

システムに対してより継続的に話したくなるかを検証するために，実験参加者一人ずつに「A: ポンプとダルマで話し手と聞き手を表現する InterActor のモード」,「B: A モードに加え，話し手から聞き手への発話の移行を空気の球で視覚提示するモード」,「C: B モードに加え，聞き手となるダルマがハンドル降下時に少し膨らみ，うなずき時に大きく膨らむモード」の 3 モードを使用して自由対話実験を行った．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ーバランスを取って順序を変更した．その後，あらかじめ用意したシナリオを一度読み上げさせることで，システムへの発話と，どのような話し方をすればいいのかの感覚をつかませた．実験参加者には，独り言でありながらも友達に語り掛けるような感覚で話し，敬語であろうとなかろうとどちらでも構わないと伝えた．システムに慣れさせるため，モードごとに 3 分ずつ使用させながら，実験本番で話す内容を考えさせた．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 18〜23 歳の男女学生 24 名であった．

1 つ目の評価方法は一対比較実験である．A〜C のモードから 2 つを一対比較させ，これを計 3 回 $(= {}_3C_2)$ 繰り返した．提示順序についてもカウンターバランスをとり，評価への影響を排除するために，総当たりにより順序を変更した．まず，1 つのモードでシステムに対する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「発話意欲」という観点からどちらにより継続的に話したくなったかアン

表 1　一対比較結果
Table 1　Result of pair comparison.

	A	B	C	Total
A		7	3	10
B	17		5	22
C	21	19		40

ケート用紙に記入させた．1 つの比較につき対話時間を 2 分 (1 モードにつき 1 分) とした．

2 つ目の評価方法は 7 段階評価実験である．A〜C のモードを「好み」,「楽しさ」,「話しやすさ」,「一体感」,「安心感」,「思いの伝達感」,「システムを使用したいか」の 7 項目について評価させた．各モードを使用した後，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，各モードに対する率直な意見を自由記述形式で記入させた．提示順序はカウンターバランスを取って順序を変更した．1 つのモードにつき対話時間を 1 分とした．

4.2　実験結果

一対比較の結果を表 1 に示す．表中の数字は一対比較において各行のモードを選定した実験参加者の数を表している．この結果に対して実験参加者の評価を定量的に評価するため，Bradley-Terry モデルを想定した (式 4,5).

$$P_{ij} = \frac{\pi_i}{\pi_i + \pi_j} \tag{4}$$

$$\sum_i \pi_i = const.(= 100) \tag{5}$$

π_i: i の強さの量, P_{ij}: i が j に勝つ確率

π_i は 3 種類のモードの強さ（3 モードの合計 100）を表し，このモデルを想定することにより一対比較に基づく評価を定量的に評価することができる．Bradley-Terry モデルにより推定された各モードの強さの比較を図 8 に示す．その結果，C モードが最も高く評価され，次いで B，A の順に評価された．

次にアンケートによる 7 項目の評価結果を図 9 に示す．帰無仮説をモード間に差は無いとして，Wilcoxon の符号順位検定によって評価した．B,C モード間では「楽しさ」の項目において，有意水準 1%で有意差が認められた．また，A,C モード間では「好み」,「一体感」の 2 項目において有意水準 5%,「思いの伝達感」,「システムを使用したいか」の 2 項目において有意水準 1%,「楽しさ」の 1 項目において有意水準 0.1%で有意差が認められ，帰無仮説が棄却され，C モードが高く評価された．

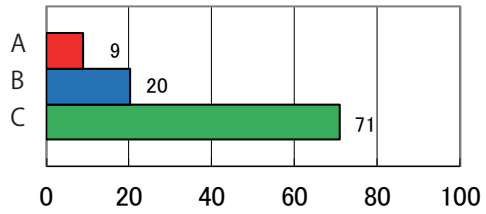


図 8　Bradley-Terry モデルによる好ましさ π_i
Fig. 8　Preference π_i based on the Bradley-Terry model.

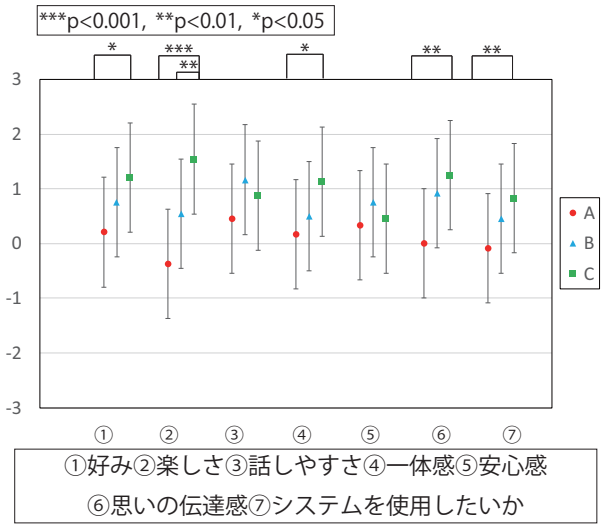


図 9　7 段階評価結果
Fig. 9　Result of seven points bipolar rating.

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 2 に示す．

4.3　考察

実験の結果，C モードが他のモードに比べて高く評価された．7 段階評価においても,「楽しさ」,「思いの伝達感」,「システムを使用したいか」の項目で高く評価された．これは，話し手から聞き手への発話伝達および蓄積を視覚提示することによって，ユーザはシステムに対してより継続的に話したくなることを示している．しかし,「話しやすさ」と「安心感」の項目では，各モード間に有意差は認められなかった．これは，自由記述欄の意見にもあるように，C モードにおいて，初期は話せば膨らむことに対してユーザが楽しさを感じるものの，ある程度まで話すと巨大化したダルマに対して圧迫感のようなものを感じ，話が途切れ途切れになってしまうからだと考えられる．ダルマがある程度まで大きくなった時に，どのような対話の区切りを入れ，その後の対話を継続させるかを今後検討する必要がある．

表 2　システムに対する意見
Table 2　Coments for the system.

モード	内容
A	・静かな聞き手のほうが好きだが，話を聞いてもらってる感はあまり感じられなかった ・うなずいているから話を聞いている感じは伝わった． ・空気が移動しないし膨らまないから達成感の様なものが得られなかった ・他モードから移ると空気が動かないため，話を聞いてもらっていないように感じる
B	・空気が入っていくことで対象に自分の言葉が伝わっている印象を受けた ・ちょうど良い感じ．落ち着いているけど聞いてくれているような感じ ・空気が伝わっていても何も起こらないのが気になった
C	・話せば話すほど膨らむので，話を聞いてもらえている感じがした ・話すのが少しだけ楽しくなった ・人形が大きくなっていくリアクションがとても面白く感じたので好印象を出した．他モードと比較して一番抵抗が無かった ・膨らむのでとても話した感があった ・このモードが個人的に好きだが，大きくなっていくのが気になって，「話しやすさ」と「安心感」が他のモードと比較すると無いように感じた ・だんだんと大きくなってくる時点では良いが，大きくなりすぎると「もう話を聞きたくない」と思われている気がして言葉に詰まってしまった

5.　おわりに

本研究では，ユーザの発話意欲を増大させ，発話を促進させるために，話し手から聞き手への時間経過に伴う発話伝達および蓄積を視覚提示する身体引き込みキャラクタを用いた発話促進システムを提案し，空気入れと風船に見立てたポンプとダルマを用いたプロトタイプを開発した．さらに 3 つのモードを用いた評価実験において，従来の話し手と聞き手に対応した InterActor に加えて発話伝達および蓄積表現を備えた C モードが高く評価され，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・蓄積表現の効果が示された．

謝辞

本研究は科学研究費 (16K00278) の助成を受けたものである．

参考文献

[1] Watanabe,T.: Human-entrained Embodied Interaction and Communication Technology, Emotional Engineering, Springer, pp.161–177 (2011).
[2] Danbara,R., Watanabe,T., Okubo,M.: 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).
[3] 桂木，渡辺，石井：思いの伝達を表現する身体引き込みキャラクタ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学会研究報告集 Vol. 20 No. 3, pp.189–192 (2018).
[4] 渡辺, 荻久保, 石井: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.105–112 (2001).