

博士學位論文

タイピング駆動型身体引き込み
キャラクタチャットシステムの研究

平成29年3月

服部 憲治

岡山県立大学大学院

情報系工学研究科

タイピング駆動型身体引き込み キャラクタチャットシステムの研究

目次

第1章 序論	1
1.1 本研究の背景と目的	1
1.2 本研究の位置付け	2
1.2.1 テキスト情報の提示効果	3
1.2.2 吹き出しなどテキスト情報提示方法の検討	3
1.2.3 キャラクタなど別モダリティによる表現方法の検討	4
1.3 本論文の構成	5
参考文献	7
第2章 iRT を用いたチャットシステム	11
2.1 緒言	11
2.2 iRT に基づく身体的コミュニケーション支援	11
2.3 音声とタイピングの関係	12
2.4 InterChat のコンセプト	14
2.5 InterChat のインタラクションモデル	16
2.5.1 聞き手動作モデル	16
2.5.2 話し手動作モデル	18
2.5.3 文字列に対応した動作モデル	19
2.6 プロトタイプの開発	19
2.7 結言	21
参考文献	22
第3章 InterChat のシステム評価	23
3.1 緒言	23

目次

3.2	身体的引き込み動作の評価	23
3.2.1	実験システム	23
3.2.2	実験方法	25
3.2.3	実験結果	26
3.3	InterChat の評価	28
3.3.1	実験システム	28
3.3.2	実験方法	30
3.3.3	実験結果	31
3.4	考察	33
3.5	結言	34
	参考文献	35
第4章	テキストのリアルタイム入力状態表示手法	37
4.1	緒言	37
4.2	テキスト入力状態表示の検討	38
4.2.1	テキスト入力状態表示機能	38
4.2.2	評価実験	39
4.2.3	実験結果	42
4.3	入力文字の表示タイミングの検討	43
4.3.1	確定文字同期表示機能	44
4.3.2	評価実験	45
4.3.3	実験結果	46
4.4	考察	47
4.5	結言	50
	参考文献	51
第5章	情動表現の即時表示手法	53
5.1	緒言	53
5.2	キャラクタチャットシステム	53
5.2.1	対話時の画面表示	53
5.2.2	情動表現の即時表示手法	54

目次

5.3	チャットコミュニケーション実験	56
5.3.1	実験方法	56
5.3.2	実験結果	57
5.4	考察	60
5.5	結言	60
第6章	テキストおよび情動表現同期表示手法	61
6.1	緒言	61
6.2	テキストおよび情動表現同期表示の検討	61
6.2.1	テキスト入力確定時情動表現機能	61
6.2.2	実験方法	63
6.2.3	実験結果	64
6.3	考察	67
6.4	結言	67
第7章	結論	69
7.1	本研究のまとめ	69
7.2	今後の展望	71
	謝辞	73
	本論文に関する研究業績	75
	原著論文	75
	口頭発表	75

第1章

序論

1.1 本研究の背景と目的

情報技術やネットワークの発展により、遠隔コミュニケーションを支援するインタフェースとしてメールやチャットといった文字によるコミュニケーションが広く利用されている。とくにチャットは、インターネットを通して即座にメッセージを伝え合うことができるために盛んに使用されており、そのリアルタイム性からオンラインゲーム等にも導入され、自分の代役キャラクタを通して複数の人と同じ時空間を共有したコミュニケーションが実現されている [1.1].

しかし、コミュニケーション手段の幅が広がる中、文字のみによるコミュニケーションでは、対話者のノンバーバル（非言語的）情報が欠落してしまうために思いが伝わりにくい。また、ノンバーバル情報の欠落が感情の行き違いを招き、非難・中傷行為といったフレーミングが起りやすいといった問題もある [1.2, 3]. そのため、文字情報では表現できない思いを効果的に伝える手段として、アニメーションを使った絵文字やキャラクタを用いたチャットシステムの研究が行われている [1.4, 5, 6]. これらの多くは、ノンバーバル要素を付与することで、視覚表現により感情や意図を補足し、誤解を防ぐ試みである。

一方、人の対面コミュニケーションでは、単に言葉だけでなく、音声に対するうなずきや身振り・手振りが相互に同調して、対話者同士が互いに引き込み合うことでコミュニケーションしている。この身体的リズムの引き込みによるコミュニケーションは幼児期にも観察される人間の本質的なコミュニケーションであり、身体的リズムの引き込みが対話時の一体感を生むことで、対話相手とのかかわりを実感させている [1.7]. これまでに、人の対面コミュニケーション時の身体的リズムの引き込みに着目し、会話音声のみを入力としてキャラクタの豊かなコミュニケーション動作を自動生成するインタロボット技術 iRT が研究開発されている [1.8]. iRT はコミュニケーション時の発

話音声と身体動作と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、身体的リズムの引き込みによりインタラクションを円滑にして、コミュニケーションを支援する技術である。既に、音声駆動型身体的インタラクションロボット InterRobot や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、iRT の有効性が示されている [1.9, 10]。iRT が音声リズムのみからキャラクタの身体動作を自動生成して身体的インタラクションを実現していることから、チャットにおいてもタイピングのリズム的特徴からキャラクタの身体動作を表現することで、円滑な身体的インタラクションを実現できるものと期待される。

そこで本研究では、タイピングの入力がリズム的であることに着目し、タイピング入力を音声入力に見立て iRT を音声入力インタフェースからタイピング入力インタフェースへと拡張することで、タイピング入力からキャラクタの身体動作を自動生成するとともに、文字列に対応する意味のある動作を自動生成するタイピング駆動型身体引き込みキャラクタチャットシステム InterChat のコンセプトを提案する。対話者の代役となるキャラクタを介したチャットコミュニケーションにおいても、iRT のメカニズムを導入することで、身体的リズムの引き込みによる一体感のあるチャットコミュニケーションを実現できる。

1.2 本研究の位置付け

従来、テキストを用いたコミュニケーションメディアにおいては、限られたモダリティの中でそのメディアが持つコントロール可能性、編集可能性に応じてデザインされた発話が行われている [1.11]。しかしテキストのみではノンバーバル情報が欠落するという問題点から、送信者のノンバーバル情報を補足するためには、顔文字やアスキーアート、アイコン画像などの視覚表現によって送信者が意図的に補足する手法と、送信者に関する入力中の何らかの計測データに基づいて非意図的に予測・推定あるいは変換して補足する手法とが考えられる。

これらのコミュニケーション手段について、スクロールなどテキスト情報の提示効果、吹き出しなどを利用したテキスト情報提示方法の検討、また、本研究のようにキャラクタへの身体動作や表情変化など別モダリティとして付加する手法など、活用可能性に関する検討が行われており、以下に本研究との関連について述べる。

1.2 本研究の位置付け

1.2.1 テキスト情報の提示効果

テキスト情報を提示する上で、縦スクロール表示における文章の読み速度として、1行に20文字表示した場合に、1秒間当たりで平均的に10文字程度の読み速度が快適速度との報告がある [1.12]。一般的に人の発話速度は300～400字/分であることに對して、キーボードによる文字入力速度は中級者で150～250字/分、上級者でも400～500字/分と想定される [1.13]。さらにローマ字入力では文字を母音と子音の組み合わせによって入力することから、キー入力は音声入力の2倍程度、またはそれ以上の時間がかかると考えられる。またライブ型チャットでは20文字以下の発言が約89.8%を占め、とくに10～20文字の発言が多いとの報告もあり [1.14]、より対話的な文字入力特性を持つと考えられる。

対話的な応答表示の待ち時間として、入力後に相手の応答を許容できる時間は平均で1分51秒との報告があるが [1.15]、入力時のアウェアネス状況や疑似応答の伝達によって応答許容時間の延長が期待できる結果が得られるなど、システムからの自動応答についての効果が確認されている [1.16]。本研究ではタイピング入力情報を利用した身体情報による自動応答システムを実現しているが、身体情報の提示との関連を含め、テキスト提示タイミングについて検討する。

1.2.2 吹き出しなどテキスト情報提示方法の検討

テキスト情報を用いたコミュニケーション手段としてメールシステムに着目し、編集過程に対する時間、打鍵速さ、修正割合など編集過程のノンバーバル情報をテキスト情報と同時に送受信する取り組みも行われている [1.17]。自動取得されたテキスト入力時の情報も同時に送受信されるため、受信者はテキストだけでなくその編集状態を把握しながらメールの内容を解釈することができ、送信者の感情推定に對して機能することが確認されている。このテキストの編集状態提示について、チャットコミュニケーションにおいてリアルタイムなインタラクション支援のために利用する提案も行われており、会話履歴を表示するメインウィンドウに對する編集中のテキストを表示するサブウィンドウを用いた発話タイミング共有システム [1.18]、あるいはテキスト情報の動的スクロール表示によって対話の流れを示し、ターンテイキングを支援するシステムなどが研究開発されている [1.19, 20]。

また Wang らは、生体信号の計測によって感情モデルを推定し、テキスト情報の提示を変化させ印象付けを行うシステムを提案している [1.21]。小俣らも生体信号からの感情推定によって、テキスト情報を提示する吹き出しの形状や色情報を変化させその印象評価を行っている [1.22]。皮膚伝導などの生体信号を用いることで、対話者のコントロールが難しい情報からテキスト情報の提示方法を変化させることは、単なるテキスト情報による編集可能なコミュニケーション手段のカテゴリから外れた、新しいコミュニケーション環境を提案していると言える。

さらに小寺らは、チャットコミュニケーションにおいて、吹き出しの形状変化とアバタの表情変化を対話者自身が意図的に選択提示し、統合的に表現するシステムの検討を行っている [1.23]。このようにテキスト情報の提示方法についての印象評価が検討されており、キャラクターの身体情報など別モダリティとの関係を用いた研究について次節に述べる。

1.2.3 キャラクタなど別モダリティによる表現方法の検討

身体情報を持たない状態であるテキストのみのコミュニケーションにおいて、身体情報を組み合わせて提示することによりテキスト自体が持つ意味とは異なった印象を与えることが可能となる。仮想環境においてキャラクターを用いたチャットシステムは以前から数多く提案されており、入力されたテキスト情報を利用してキャラクターの動作画像や表情を変更するシステムや [1.24]、キャラクターのジェスチャや表情をより効果的に利用するために、メッセージ作成時にキャラクター動作などの非言語表現を選択することで、モダリティ提示するシステムなどが提案されている [1.25, 26, 27]。また、キャラクターの有無や視野角による距離知覚が与える影響の検討 [1.28] や、拡張現実感を利用したアバタチャットシステムによるモダリティ提示 [1.29] なども行われている。これらは言語的な情報を活用する目的でキャラクターなど空間表現を用いて別モダリティとして提示している。本研究では言語的情報である入力文字列と、非言語的情報であるタイピングリズムに着目し、キャラクター表現を行うことで対話感を向上させる検討を行う。

ここで山田らは、打鍵振動によって対話状況を伝達する TangibleChat を開発し、発話者の打鍵情報を対話相手に振動として提示することで、対面対話のモダリティとも異なる新たなコミュニケーション環境を提案している [1.30]。本研究とはタイピング情

1.3 本論文の構成

報を利用するという観点で共通しており，発言のしやすさ，タイミングのとりやすさといった評価項目に対して効果が確認されている。

1.3 本論文の構成

本論文は本章を含め，全7章により構成されている。本章を除いた2章以降の概要を以下に述べる。

第2章では，人間の発話音声リズムとキーボード入力におけるタイピング入力リズムの類似性を示し，タイピング入力に基づく身体引き込みキャラクタチャットシステムを提案している。発話音声からキャラクタの豊かなコミュニケーション動作を自動生成するiRTモデルを応用し，タイピング入力から聞き手動作，話し手動作，および入力された文字列に対応した動作それぞれを生成するコンセプトと動作モデルについて述べ，システムのプロトタイプを開発している。

第3章では，2章で開発したシステムのプロトタイプを基に，仮想空間内にシステム使用者の代役となるキャラクタを配置し，互いのキャラクタの身体動作によって，使用者間で身体的リズムを共有しながらチャットを行うチャットコミュニケーションシステム InterChat の開発およびシステム評価を行っている。まず身体的インタラクションの観点から，InterChat から文字列に対応した動作モデル機能を省略し，話し手動作モデル，聞き手動作モデルのみを導入したシステムを用いたコミュニケーション実験により，身体的引き込み動作提示の有効性を示している。また文字列に対応した動作モデルを導入したシステムを含めて従来のテキストチャットと比較する評価実験を行い，聞き手動作モデル，話し手動作モデル，および文字列に対応した動作モデルを用いたシステムが高く評価され，開発したシステムの有効性を示している。

第4章では，自動生成された身体的引き込み動作をより効果的に受信者に提示するために，テキスト入力状態と身体的引き込み動作を同期表示する手法として，入力中の情報全てを表示するリアルタイム表示手法，入力中の文字列を中点(·)に変換して表示する中点表示手法を開発し，3章でのメッセージ表示手法である送信時表示との比較検討を行っている。その結果，入力中の状態は相手の入力メッセージが中点に変換されて表示される中点表示手法が評価されている。さらにテキストの編集可能性を維持しつつより同期表示する手法として，メッセージ確定時に確定部分のみが中点からテ

第1章 序論

キスト表示に切り替わる中点逐次確定表示，変換過程を見せず変換確定時に相手に表示される変換過程隠し表示を開発し，比較検討した結果，中点逐次確定表示手法がより高く評価されている。

第5章では，それまでテキストに対応した情動表現がメッセージの送信後に提示されていたことによる問題点を克服するため，新たに確定時動作，入力時動作の2つの提示手法を開発している。確定時動作では，入力中のメッセージが確定される度に認識を行い，特定の文字列が文中にあった場合はその文字列に対応した動作を提示し，入力時動作では，メッセージが入力される度に認識を行い，特定の文字列が文中にあった場合はその文字列に対応した動作を提示する。3つの手法を用いた評価実験において，入力の度に文章の認識を行い，動作を提示する入力時動作手法が高く評価されている。

第6章では，InterChatの統合的なシステム評価として，4章でのテキスト表示および5章の情動表現表示それぞれの着眼点から提案された手法のうち，最も高く評価された2つの手法，さらに入力状態をテキストで表示する際に最も好まれた中点表示(·)を使用して，テキストの編集可能性を維持しつつ，テキストとキャラクタの情動表現を同期して表示する手法を新に開発し，3つの手法を用いて評価実験を行っている。

第7章では，本研究で得られた成果をまとめると共に，今後の発展性について述べている。

参考文献

- [1.1] 例えば Second Life (<http://secondlife.com/> : 2016/12/20 アクセス) など.
- [1.2] 江下 雅之 : ネットワーク社会の深層構造 「薄口」の人間関係へ, 培風館 (1999).
- [1.3] 和田 悟, 近藤 佐保子 : 文科系のための情報学シリーズ インターネットコミュニケーション デジタルライフのおとし穴, 中公新書 (2000).
- [1.4] 神田 智子 : 感性コミュニケーションツール「ペタろう」の開発と分析, ヒューマンインタフェース論文誌, Vol. 8, No. 1, pp. 101–108 (2006).
- [1.5] 宗森 純, 大野 純佳, 吉野 孝 : 絵文字チャットによるコミュニケーションの提案と評価, 情報処理学会論文誌, Vol. 47, No. 7, pp. 2071–2080 (2006).
- [1.6] Soo-Hyun Park, Seung-Hyun Ji, Dong-Sung Ryu and Hwan-Gue Cho : A new cognition-based chat system for avatar agents in virtual space, VRCAI '08 Proceedings of The 7th ACM SIGGRAPH International Conference on Virtual-Reality Continuum and Its Applications in Industry, No.13 (2008).
- [1.7] 渡辺 富夫 : 身体的コミュニケーションにおける引き込みと身体性—心が通う身体的コミュニケーションシステム E-COSMIC の開発を通して—, ベビーサイエンス, Vol. 2, pp. 4–12 (2003).
- [1.8] Tomio Watanabe : Human-entrained Embodied Interaction and Communication Technology, Emotional Engineering, Springer, pp. 161–177 (2011).
- [1.9] 渡辺 富夫, 大久保 雅史, 小川 浩基 : 発話音声に基づく身体的インタラクションロボットシステム, 日本機械学会論文集 (C編), Vol. 66, No. 648, pp. 251–258 (2000).
- [1.10] 檀原 龍正, 渡辺 富夫, 大久保 雅史 : 音声駆動型身体引き込みキャラクタ InterActor が発話音声に与える効果, 日本機械学会論文集 (C編), Vol. 71, No. 712, pp. 152–159 (2005).

第1章 序論

- [1.11] 原田 悦子：人の視点からみた人工物研究：対話における「使いやすさ」とは（認知科学モノグラフ⑥），共立出版（1997）.
- [1.12] 石井 亮登，森田 ひろみ：縦スクロール表示された文章の快適な読み速度と眼球運動，情報処理学会論文誌，Vol. 54, No. 6, pp. 1784–1793（2013）.
- [1.13] e-typing <http://www.e-typing.ne.jp/>（2016/10/20 アクセス）
- [1.14] 山田 裕之：動画を用いたWEBセミナーコンテンツに付随したテキストチャットの特徴分析，日本感性工学会論文誌，Vol. 14, No. 1, pp. 223–229（2015）.
- [1.15] 宮部 真衣，吉野 孝：リアルタイム遠隔テキストコミュニケーションにおける対人許容応答時間の評価，情報処理学会論文誌，Vol. 50, No. 3, pp. 1214–1223（2009）.
- [1.16] 村田 和義，川口 修，渋谷 雄，倉本 到，辻野 嘉宏：チャット参加者の応答期待時間の延長を目指した情報提示とそのタイミング，ヒューマンインタフェース学会論文誌，Vol. 8, No. 3, pp. 423–433（2006）.
- [1.17] 角野 清久，西本 一志：言外の情報としての編集過程情報を伝えるメールシステムの提案と評価，情報処理学会論文誌，Vol. 50, No. 1, pp. 254–267（2009）.
- [1.18] Kanayo Ogura, Takeshi Masuda, and Masato Ishizaki: Building a New Internet Chat System for Sharing Timing Information, Proceedings of the 4th SIGdial Workshop on Discourse and Dialogue, pp. 97–104（2003）.
- [1.19] Tara Rosenberger Shankar, Max VanKleek, Antonio Vicente, and Brian K. Smith: Fugue: A Computer Mediated Conversational System that Supports Turn Negotiation, Proceedings of the 33rd Hawaii International Conference on System Sciences-Volume 3, p. 3035（2000）.
- [1.20] 山田 祐士，竹内勇剛：非交替型チャットシステムの開発と社会的な対話のダイナミクスの解析，電子情報通信学会，信学技報，HCS2002-50, pp. 19–24（2003）.
- [1.21] Hua Wang, Helmut Prendinger, Mitsuru Ishizuka, and Takeo Igarashi: Affective Communication in Online Chat Using Physiological Sensors and Animated Text, Journal of Human Interface Society, Vol. 7, No. 1, pp. 39–45（2005）.

参考文献

- [1.22] 小俣 昌樹, 内藤 雄也, 今宮 淳美: 生体信号に連動する吹き出しオンラインチャットの設計と評価, ヒューマンインタフェース学会論文誌, Vol. 10, No. 2, pp. 179–189 (2008).
- [1.23] 小寺 貴士, 北森 慶太, 大久保 雅史: チャットコミュニケーションにおける吹き出しとアバターによる感情表現手法の提案, ヒューマンインタフェース学会研究報告集, Vol. 16, No. 3, pp. 231–234 (2014).
- [1.24] David Kurlander, Tim Skelly, and David Salesin: Comic Chat, Proceedings of SIGGRAPH 96, pp. 225–236 (1996).
- [1.25] Toru Takahashi, Christoph Bartneckb, Yasuhiro Katagiri, and Noriko H. Arai: TelMeA–Expressive avatars in asynchronous communications, International Journal of Human-Computer Studies, 62, pp. 193–209 (2005).
- [1.26] 伊藤 冬子, 佐々木 康成, 廣安 知之, 三木 光範: アバターを利用した主観アノテーションのための感情表現とその一貫性に関する検討, 知能と情報 (日本知能情報ファジィ学会誌), Vol. 20, No. 4, pp. 487–499 (2008).
- [1.27] 楠見 孝, 米田 英嗣, 小島 隆次: アバターの感情表出機能によるマルチユーザ仮想空間コミュニケーション・システムの改良, 日本教育工学会論文誌, Vol. 31, No. 4, pp. 415–424 (2008).
- [1.28] Soo-Hyun Park, Seung-Hyun Ji, Dong-Sung Ryu, and Hwan-Gue Cho: A New Cognition-based Chat System for Avatar Agents in Virtual Space, Proceedings of The 7th ACM SIGGRAPH International Conference on Virtual-Reality Continuum and Its Applications in Industry, No. 13 (2008).
- [1.29] Kei Tanaka, Dai Hasegawa, Martin J. Dürst, and Hiroshi Sakuta: Avatar Chat: A Prototype of a Multi-Channel Pseudo Real-time Communication System, Proceedings of the 28th Annual Conference on Computer Animation and Social Agents (CASA2015), pp. 25–28 (2015).

第1章 序論

- [1.30] 山田 裕子, 平野 貴幸, 西本 一志 : TangibleChat : 打鍵振動の伝達によるキーボードチャットにおける対話状況ウェアネス伝達の試み, 情報処理学会論文誌, Vol. 44, No. 5, pp. 1392-1403 (2003).

第2章

iRTを用いたチャットシステム

2.1 緒言

本研究の目的は、テキストによるバーバル（言語的）情報のみで行われるチャットコミュニケーションにおいて、言語以外のノンバーバル情報を活用することで、より互いのインタラクションを支援し思いを伝え合うことのできるシステムの開発である。チャットコミュニケーションにおいて、最低限のシステム構成で取得できるノンバーバル情報としてキーボード入力におけるタイピングリズム情報が挙げられる。このタイピングリズム情報を利用してリアルタイムに相手のテキスト入力状況が捉えられることで、互いのインタラクションを支援し、より豊かなコミュニケーション環境が構築できる。

本章では、タイピングの入力がリズム的であることに着目し、発話音声とキーボード入力におけるタイピング入力リズムの類似性を示し、発話音声からキャラクタの豊かなコミュニケーション動作を自動生成する身体的コミュニケーション技術をタイピング入力へと応用展開するシステムを提案している。

2.2 iRTに基づく身体的コミュニケーション支援

会話音声のみを入力としてキャラクタの豊かなコミュニケーション動作を自動生成するインタロボット技術iRT (InterRobotTechnology) は、身体的リズムの引き込みによりインタラクションを円滑にして、コミュニケーションを支援する技術である [2.1]. このiRTを用いて、うなずきや身振り、手振りなどの身体性を、物理メディアであるロボット (InterRobot) [2.2] や電子メディアのキャラクタ (InterActor) [2.3] などの情報機械に導入することで、遠隔地間においても互いの身体性を共有し対話者相互に相

第2章 iRT を用いたチャットシステム

手とのかかわりを知覚できる身体的コミュニケーションシステムを開発し、コミュニケーション支援での有効性を示している。

図2.1はiRTをCGキャラクタに導入したInterActorで、発話音声に基づいてうなずきや身振り・手振りなどのコミュニケーション動作を自動生成し、話し手や聞き手として動作する。図2.2はInterActorを用いた身体的インタラクションシステムの使用風景である。遠隔コミュニケーションにおいては、対話者各々の音声のみをネットワークを経由して相互に送信し、発話音声に基づく身体動作をInterActorに表現させることで、キャラクタの身体動作による引き込みで身体性が共有され、対話者は対面コミュニケーションの一体感を実感し円滑なコミュニケーションを行うことができる [2.1]。

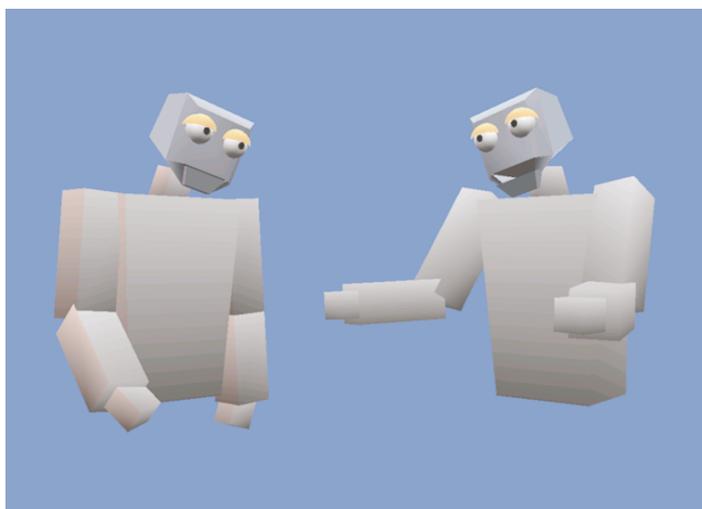


図 2.1 InterActor

2.3 音声とタイピングの関係

前述のiRTは音声に含まれる情報のうち、リズムと振幅のみを使用してコミュニケーション動作を生成している。図2.3（上）は「今日は学食で唐揚げを食べようかな」という音声の時系列変化の一例である。InterActorではこの音声のリズム、つまり音声の呼気段落区分でのON-OFFパターンからうなずきを予測し、さらに、うなずきの予測値から体の各部位を動作させてコミュニケーション動作を生成している。この呼気段

2.3 音声とタイピングの関係

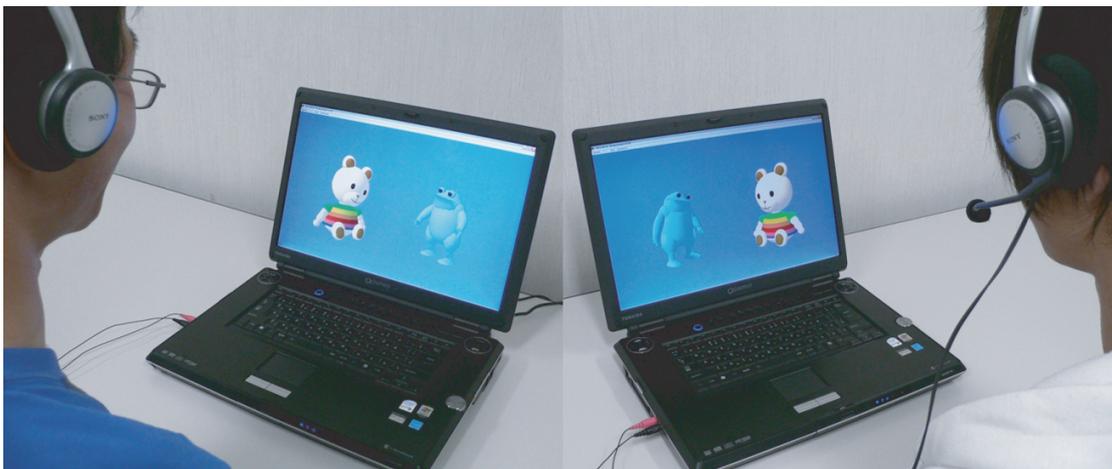


図 2.2 InterActor による遠隔コミュニケーション

落区分は文章中の意味的に区切る部分や、発話者の思考中に出現する。タイピングにおいても音声会話の場合と同様に、文章中の意味的に区切る部分や、入力者の思考中に音声の呼気段落区分に相当する区分が存在すると考えられる。そこで、予備実験として音声と同じ文章をタイピングし、30Hz でキーの入力状態をサンプリングし計測した。タイピングと発話音声の関係としては、一般に中級者のタイピングは1分間あたり150～250字程度と言われ、また、アナウンサーは1分間あたり300～400字程度の速さで原稿を読んでいると言われていることから、計測したキー入力状態からON区間を伸ばすハングオーバー処理を、音声の場合の165msに対して2倍の330msとし、タイピングのON-OFF区間を計測し音声の場合と比較した[2.5]。その結果、時間軸を伸縮することで音声の場合と非常に類似することがわかった。図2.3（下）に計測例を示す。iRTの身体的インタラクションは、入力に対する反応のタイミングが重要であり、タイピングにおいても音声と類似するON-OFFパターンの結果が得られたことは、タイピングのキー入力データを音声のデータに見立てiRTの入力とすることで、タイピングからもInterActorのようなコミュニケーション動作の生成が可能であると考えられる。

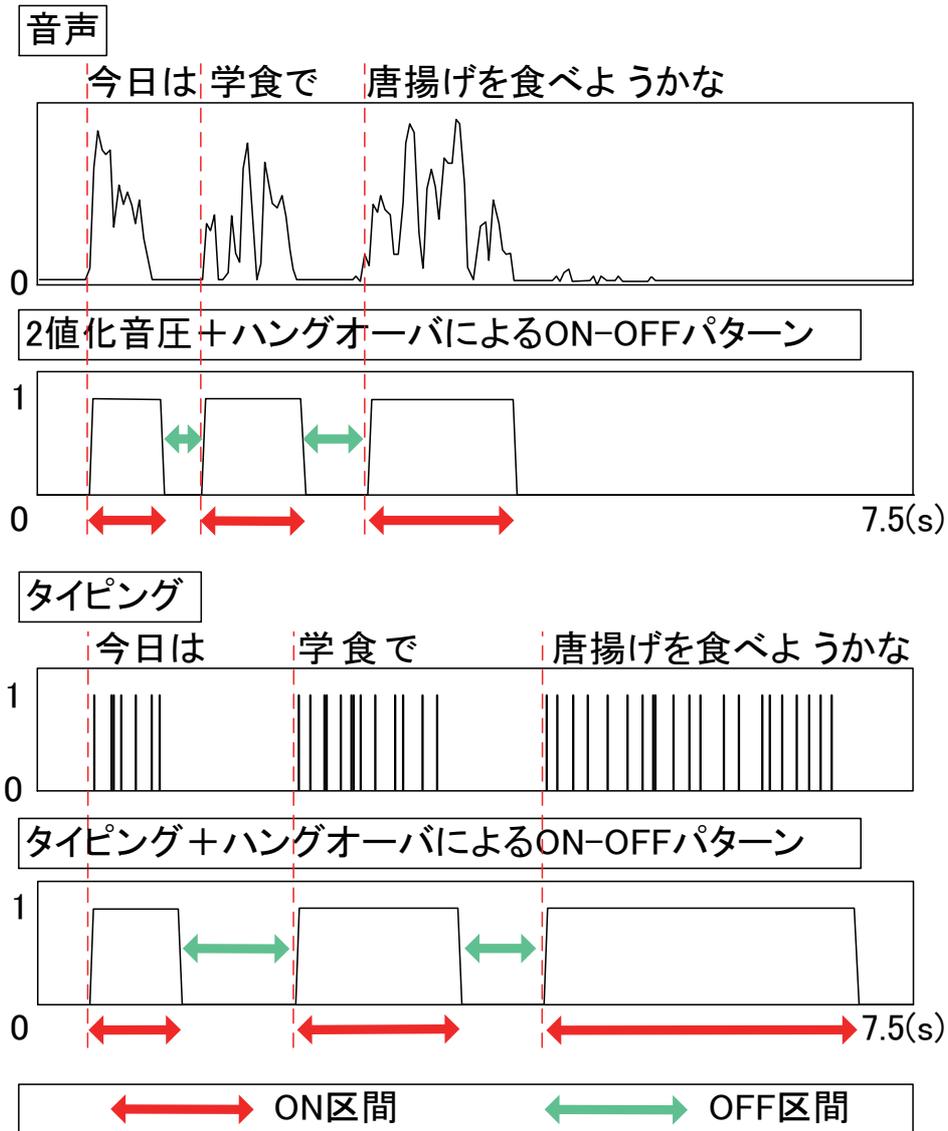


図 2.3 音声とタイピングの類似性

2.4 InterChat のコンセプト

チャットコミュニケーションを支援するために、キャラクターチャットに iRT を導入したタイピング駆動型身体引き込みキャラクターチャットシステム InterChat を提案する [2.6].

本システムのコンセプトを図 2.4 に示す. InterChat は、システム使用者が代役キャラクターを通じて仮想空間に入り込み、同一空間でチャットを行うことができるシステム

2.4 InterChat のコンセプト

である。タイピングに基づき、使用者の代役キャラクターは話し手としてのコミュニケーション動作を生成し、メッセージとともに他の参加者に提示する。代役キャラクターが身振り・手振りを伴って情報を伝えることで、効果的に情報を伝達することが可能である。同時に、他の参加者の代役キャラクターが聞き手としてのコミュニケーション動作を生成し提示することで、参加者は仮想空間内のインタラクションを確認することができる。これを参加者が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。また、文字列を扱うため、メッセージを即座に認識し、バーバル情報から文字列に対応した動作を生成することが可能であり、身体的引き込み動作と重畳合成することで、さらなるコミュニケーション支援効果が期待できる。

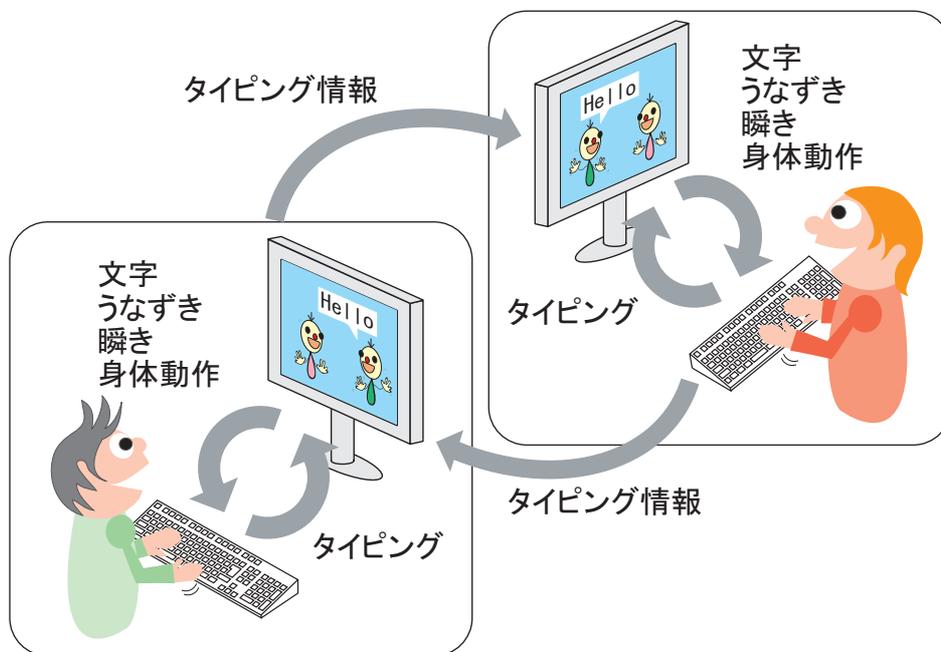


図 2.4 InterChat のコンセプト

2.5 InterChat のインタラクションモデル

2.5.1 聞き手動作モデル

キャラクタのタイピングリズムに基づくコミュニケーション動作の自動生成モデルには、これまでに iRT の研究開発において提案されている音声の ON-OFF の時系列からうなずき、瞬き、身体動作のタイミングを推定する MA(Moving-Average) モデルを使用する [2.1]. タイピングを音声入力に見立て、キーの入力状態を 30Hz で取得したキー入力の有無のデータを音声の ON-OFF データに対応させ、MA モデルへの入力データとした。MA モデルでは音声の場合と同様に、タイピングを連続した入力状態とみなすために ON 信号を伸ばすハングオーバー処理をした ON 区間 $T[i]$ と無入力状態の OFF 区間 $S[i]$ の 1 対をユニットとして扱い、音声の呼気段落区分の ON-OFF 区間およびユニットに対応するようにした。(図 2.5). ハングオーバー処理は、タイピング時

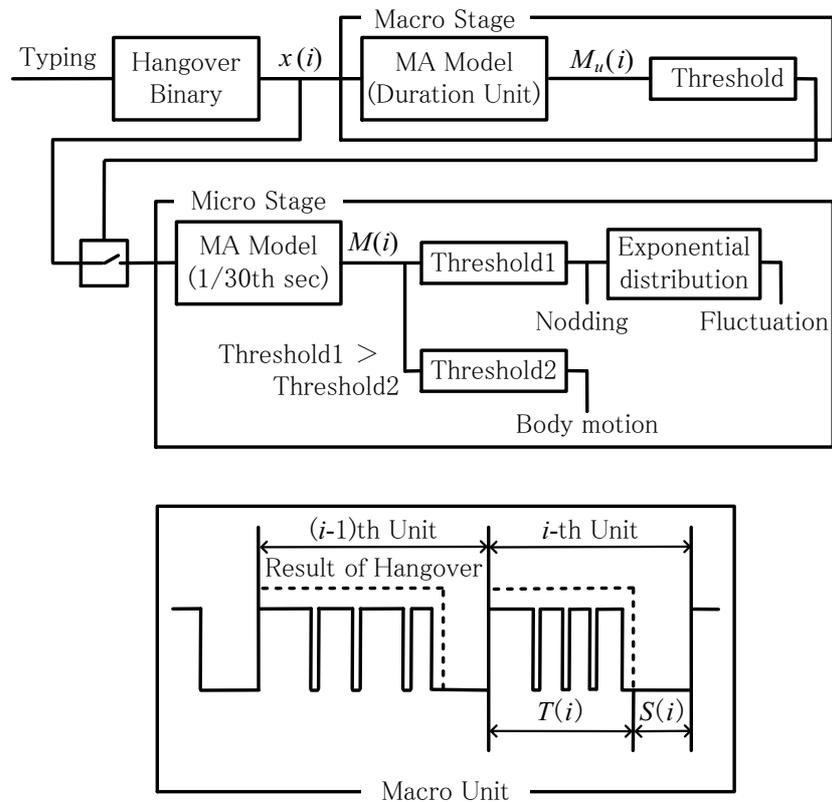


図 2.5 聞き手動作モデル

2.5 InterChat のインタラクションモデル

にキーストローク等で発生する短時間の無入力状態による ON-OFF パターンの断片化を除去するために、予備実験で得た結果より、音声では 165ms であったハングオーバー処理を 330ms で施している。このモデルは、マクロ層とマイクロ層からなる階層モデルを用いてうなずきの予測を行う。マクロ層では、タイピングの ON-OFF 区間からなる i 番目のユニットにうなずきの開始が存在するかを、 $i-1$ 番目のユニット以前のユニットの時間率 $R(i)$ ((2.1) 式, ユニット区間で ON 区間の占める割合) の線形結合で表される MA モデル ((2.2) 式) で 6 個のパラメータ $a(j)$ を用いて予測する [2.2]。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、処理はマイクロ層に移る。マイクロ層では、そのユニット区間に対して (2.3) 式の MA モデルで、60 個のパラメータ $b(j)$ を用いて過去 2 秒間のタイピングの ON-OFF データから、うなずきの開始時点を推定する [2.2]。

$$R(i) = \frac{T(i)}{T(i) + S(i)} \quad (2.1)$$

$$M_u(i) = \sum_{j=1}^J a(j) R(i-j) + u(i) \quad (2.2)$$

$a(j)$: 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

$$M(i) = \sum_{j=1}^K b(j) x(i-j) + w(i) \quad (2.3)$$

$b(j)$: 予測係数

$x(i)$: タイピングデータ

$w(i)$: ノイズ

予測値が閾値を越えた場合にキャラクタがうなずき動作を行う。また、InterActor[2.3]と同様に、うなずきと関連付けるため、うなずき以外の身体動作の推定にもうなずきの予測値を用い、うなずきよりも低い閾値で各部位（頭部、胴部、右肘、左肘）のうちいずれかを選択して動作させている。瞬きについては、対面コミュニケーション時における瞬き特性に基づいて、うなずきと同時に瞬きをさせ、それを起点にして次の瞬きまでの時間間隔を指数分布させている [2.4]。

2.5.2 話し手動作モデル

話し手の動作モデルについても同様に，対面コミュニケーション時の音声と身体動作の特性から，タイピングの ON-OFF パターンから身体全体の動作を予測し，キャラクターの各部位を動作させた．

加えて，日常会話における身体動作は，話の盛り上がりに伴って大きくなることから [2.2]，iRT では音声の振幅から生成していた手振り動作（腕を広げる，振る，交互につき出すなど），身振り動作（身をのり出す，上体を反るなど）といった大きな身体動作を生成するために，タイピングの盛り上がりを示すタイピング活性量 $U(i)$ ((2.4) 式) を定義し，音声での手振り動作，身振り動作生成モデルの入力に対応させた． $U(i)$ はキーを継続して入力するほど大きな値を取り，閾値を超えた入力があった場合，予め設定した複数パターンの一つを選択して動作させた．図 2.6 は実際にタイピングを行った際の $U(i)$ の変化量である．継続したタイピングに $U(i)$ が対応しているのがわかる．

$$U(i) = \sum_{j=1}^J (J + 1 - j)x(i - j) \quad (2.4)$$

J : タイピング活性量定数

$x(i)$: タイピングデータ

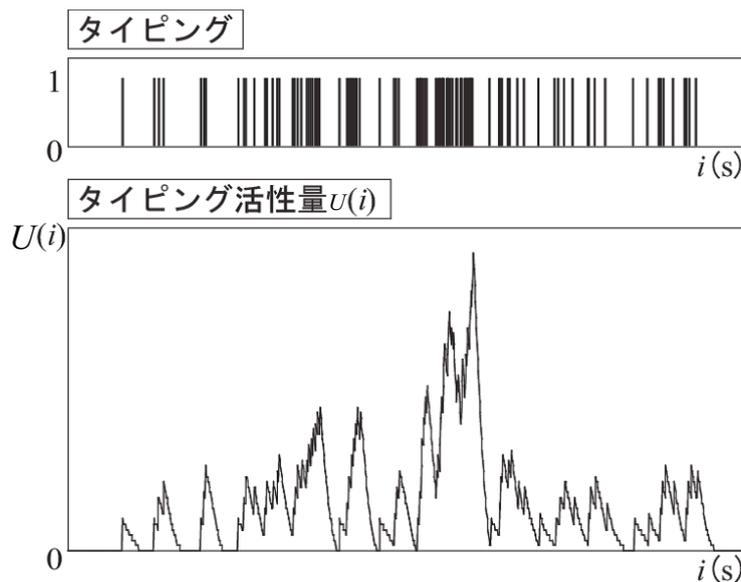


図 2.6 タイピング活性量

2.5.3 文字列に対応した動作モデル

メッセージ中に含まれる「こんにちは」などの特定の文字列を認識し、「頭を下げる」などの対応した意味動作をキャラクタの話し手動作に関係付けた。具体的には、30Hzでキー入力状態を取得する際に、タイピングデータをログデータとして取得し、予め定義されている文字列とテキスト照合した。照合が一致した場合に、その文字列に対応して登録されている動作を身体的引き込み動作と重畳合成した。話し手の動作モデルによる自然なコミュニケーション動作を保ったまま、話の内容に沿った動作の生成や感情・感嘆や疑問などを表すオブジェクトをキャラクタの頭上に表示することで、対話者の思いを効果的に表現することができる。現在 352 種類の文字列とそれに対応した 105 種類の動作を導入しており、文字列に対応した動作をデータベース化することで簡単に動作を追加することができる。親しみやすい CG キャラクタを使用して対応させた一例を図 2.7 に示す。

2.6 プロトタイプの開発

本研究で開発した InterChat は、仮想空間内にチャット参加者の代役となるキャラクタを配置し、互いのキャラクタの身体動作によって使用者間で身体的リズムを共有しながらチャットを行うコミュニケーションシステムである。概略を図 2.8 に示す。

システムは参加者毎に PC、ディスプレイ、キーボードによって構成され、互いにネットワークで接続されている。PC 同士の接続は Microsoft DirectPlay を用いて peer-to-peer の接続を行うことで、サーバを介さず使用者間で直接タイピング情報を送受信しており、通信による遅れを最小にし、リアルタイムにリズムを共有することができる。仮想空間は Microsoft DirectX9.0b を用いて作成し、家具などを配置して親しみやすい雰囲気のある部屋を作成した (図 2.8)。キャラクタは参加者が増えるたびに追加され、テーブルを囲むように向かい合わせて配置することで複数の参加者でコミュニケーションができるようにした。自分の画面では、タイピング時、自分の代役キャラクタの上に入力中メッセージおよび変換中の文字列を吹き出しとして表示することで話し手であることを明示し、そのキャラクタが身振り手振りを交えて話しているように見せている。また、タイピングによる ON-OFF データは即座に他の参加者の PC へと送信され、話し手動作を生成することで使用者の入力状態を身体動作で表現する。この時、タイ

文字列に対応した動作(一例)



	ごめんなさい すいません おねがいします		ガッカリ ガッカリ しょぼーん		寒い さむっ ブルブル
	いやです ちがいます 無理です		ひゃあ 怖い ゾクゾク		ぱちぱち パチパチ 拍手
	やあ おはよう こんにちは		よし! ガッツ 頑張るぞ		ハート キュン 好き
	やったー うれしい わーい		うああ どうしよう 困った		音符 ♪ ルンルン
	うるうる 泣ける つらい		南無三 なんまいだ ナムナム		ZZZ 眠い 寝る
	こまる こまった 残念		さあ? やれやれ おやおや		タラーリ 汗 たらー
	ジャンプ とおっ トゥッ		悩む 参ったな 照れる		ひらめいた 思いついた ピコーン
	さようなら バイバイ またね		暑い あつつ あちー		ムカムカ イライラ 腹立つ

図 2.7 文字列に対応した動作の例

ピング活性量によって生成したキャラクターの腕部動作により使用者の気持ちの盛り上がりも伝わる。また、メッセージは入力を確認したときに送信され、他の参加者の画面に確定した文字列を吹き出しとして表示する(図 2.9)。

2.7 結言

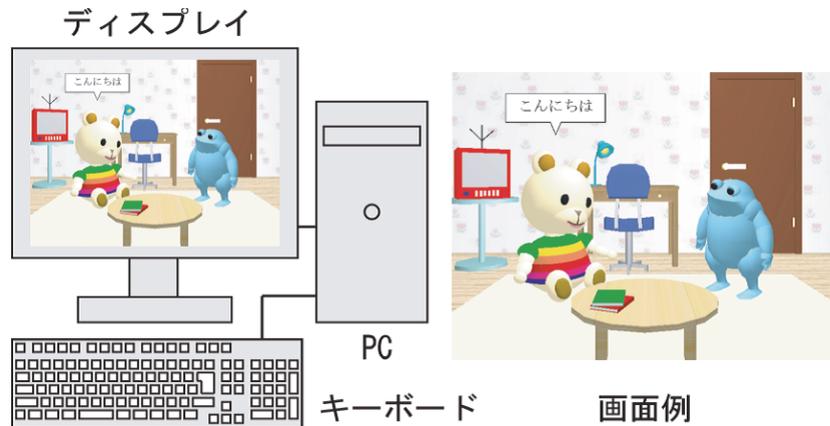


図 2.8 InterChat の概略

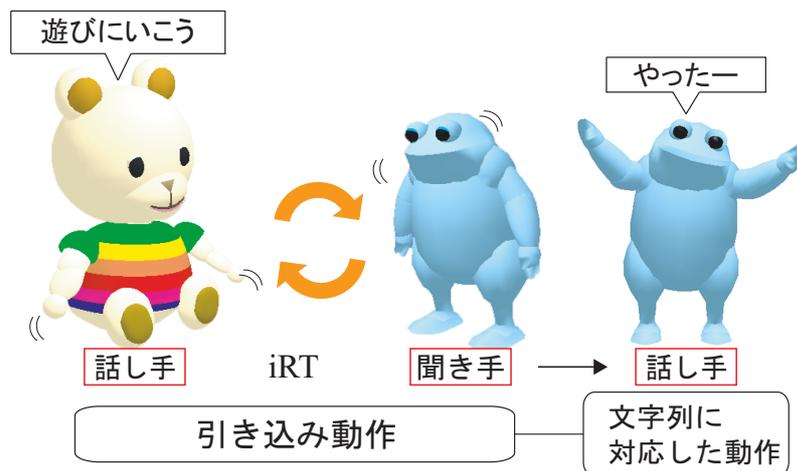


図 2.9 InterChat の動作イメージ

2.7 結言

本章では、タイピングのキー入力データを音声のデータに見立てることで、タイピング入力からコミュニケーション動作を自動生成するタイピング駆動型身体引き込みキャラクターチャットシステム InterChat のコンセプトを提案し、システムのプロトタイプを開発した。キャラクター動作として、話し手動作モデル、聞き手動作モデル、文字列に対応した動作モデルを定義し、親しみやすい CG キャラクターを使用して動作確認を行った。

参考文献

- [2.1] Tomio Watanabe : Human-entrained Embodied Interaction and Communication Technology, *Emotional Engineering*, Springer, pp. 161–177 (2011).
- [2.2] 渡辺 富夫, 大久保 雅史, 小川 浩基 : 発話音声に基づく身体的インタラクションロボットシステム, *日本機械学会論文集 (C編)*, Vol. 66, No. 648, pp. 251–258 (2000).
- [2.3] 檀原 龍正, 渡辺 富夫, 大久保 雅史: 音声駆動型身体引き込みキャラクタ InterActor が発話音声に与える効果, *日本機械学会論文集 (C編)*, Vol. 71, No. 712, pp. 152–159 (2005).
- [2.4] Watanabe, T. and Yuuki, N. , A Voice Reaction System with a Visualized Response Equivalent to Nodding, *Advance in Human Factors/Ergonomics*, 12A, pp. 396-403 (1989).
- [2.5] e-typing <http://www.e-typing.ne.jp/> (2016/10/20 アクセス)
- [2.6] 田淵 健太, 渡辺 富夫, 山本 倫也 : タイピング駆動型身体引き込みキャラクタチャットシステム InterChat の開発, *ヒューマンインタフェース学会研究報告集*, Vol. 9, No. 2, pp. 35–40 (2007).

第3章

InterChat のシステム評価

3.1 緒言

前章で構築したチャットコミュニケーションシステムを、身体的引き込み動作生成の有効性、および情動表現を含めたシステムの有効性評価を行う目的で、2つのコミュニケーション実験を行う。まず、iRT をタイピングに応用した身体的引き込み動作の評価を行い、さらに文字列に対応した情動表現を含めた InterChat の評価を行う。

3.2 身体的引き込み動作の評価

3.2.1 実験システム

本章ではまず、タイピングからのコミュニケーション動作生成モデルの有効性を検討するために、図 3.1 に示すような、A: テキストチャット（一般的なチャットを想定したテキストのみのチャットシステム）、B: キャラクタチャット（キャラクタは表示されるが動作を行わないチャットシステム）、C: InterChat(-)（キャラクタが iRT を導入した身体的引き込み動作のみを行うチャットシステム）、の3種類の異なるモードのチャットシステムを開発し、評価実験を行う。各モード毎の特徴を表 3.1 に示す。

A のテキストチャットは、インスタントメッセージを想定したもので、タイピングにより入力ウィンドウに文字を入力したのち、確定された文字が会話ログ画面に表示されることでチャット参加者が対話を行うチャットシステムである。また、主にユーザの識別を目的として、システムでは2Dのキャラクタの静止画を表示する機能や、参加者が入力中であることを相手に通知する機能を設けている。A のシステムはチャットを行う行為そのものを比較することを目的として作成した。

第3章 InterChat のシステム評価

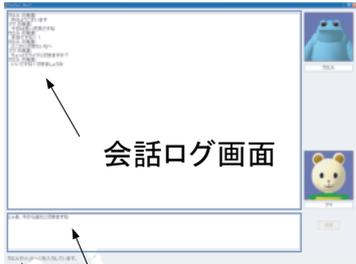
A: テキストチャット	B: キャラクタチャット	C: InterChat(-)
 <p>会話ログ画面</p> <p>入力ウィンドウ 相手の入力状態を表示</p>	 <p>キャラクターは動作しない (吹き出しは出る)</p>	 <p>引き込み動作のみ</p>

図 3.1 チャットモード

表 3.1 実験におけるモードの違い

	A: テキストチャット	B: キャラクタチャット	C: InterChat(-)
想定したもの	インスタント メッセージ	オンラインゲーム	— (iRTを導入)
会話ログ	○	×	×
キャラクターの動作	— (2Dの静止画)	×	○
入力状態を表示	○	×	○ (身体動作で提示)
制限した機能	絵文字	エモーション機能	文字列に対応した 動作モデル

Bのキャラクターチャットは、オンラインゲームなどで利用されるチャット機能を想定したもので、身体を持った自分の代役キャラクターを仮想空間内に配置し、各チャット参加者の入力文字を代役キャラクター上の吹き出しに表示して対話を行うチャットシステムである。ただし、タイピング中にキャラクターの身体動作は行わず、相手に入力状態も伝えない。Bのシステムにおいては、タイピングのリズムとは無関係な動作を導入することで、キャラクターの導入と、身体的引き込み動作の導入とを比較することも可能

3.2 身体的引き込み動作の評価

になるが、キャラクターがタイミングを無視したランダムな動作や遅延した動作を行うことは明らかに低い評価となることが想定されることから、キャラクターは動作させないこととした。

CのInterChat(-)は前述のとおり、InterChatから文字列に対応した動作モデル機能を省略し、話し手動作モデル、聞き手動作モデルのみを導入したInterChatである。文字列に対応した動作モデルを省略することで、動作生成モデルの有効性のみを比較することを目的とした。

各モードでは、ショートカットキーによる単語入力や、絵文字機能は省略し、タイピングによる機能の比較が行えるようなシステムとした。

3.2.2 実験方法

開発したシステムを使用して、別々の部屋に分かれた2人1組でチャットでコミュニケーション実験を行った。実験には、デスクトップPCと21.3型ディスプレイを用いた。被験者は18～22歳の友人関係である男女20組40人（同性各10組）で、普段からパソコンを使用しタイピングにある程度慣れた学生である。図3.2に実験風景を示す。

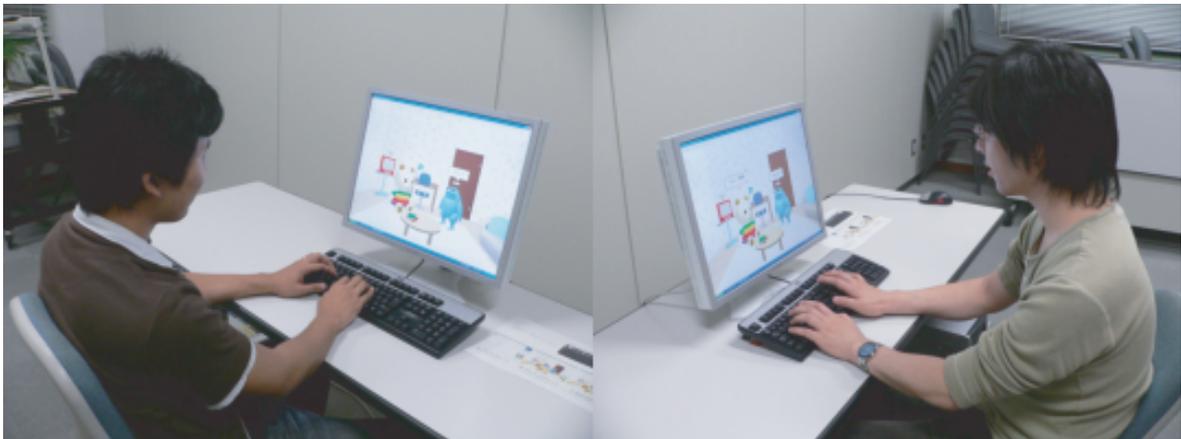


図 3.2 実験風景

実験の手順を以下に示す。

手順 1:

被験者に、各モードの違いと操作法を説明し、各モード毎に3分間試用させる。

手順 2:

3種類のモードから2種類のモードをランダムにカウンターバランスをとって選び、各1分間使用させ、一対比較をさせる（計 ${}_3C_2 = 3$ 回）。

手順 3:

被験者に3種類のモードについて3分間使用させた後、7項目による7段階評価をさせる。

一対比較は、総合評価として「総合的によい」という観点からどちらがよいかを選ばせた。7段階評価は、身体的コミュニケーション支援の観点から「楽しさ」、「好き」、「チャット（対話）しやすさ」、「使用したい」、「一体感」、「思いが通う」、「盛り上がる」の7項目について行った。A～Cのモードをランダムな順で使用させ、評価をアンケート用紙に答えさせた。被験者同士は同じモードを使用してチャットさせた。

3.2.3 実験結果

一対比較の結果を表3.2に示す。表中の数字は各モードの勝敗、つまり各行のモードを選択した回数を表している。この結果に対して被験者による評価を定量的に評価するために、以下に示す Bradley-Terry モデルを想定した [3.1].

$$P_{ij} = \frac{\pi_i}{\pi_i + \pi_j} \quad (3.1)$$

$$\sum_i \pi_i = \text{const.} (= 100) \quad (3.2)$$

π_i : i の強さの量

P_{ij} : i が j に勝つ確率

これにより、一対比較に基づく評価を一義的に定めることができる。強さ π を最尤推定した結果を図3.3に示す。有意水準5%でまず適合度検定を行い、さらに尤度比検定を行ってモデルの整合性を検定した結果、モデルは棄却されず、強さ π の妥当性が保証された。InterChat(-) が最も高く評価され、次いでテキストチャット、キャラクタチャットの順に評価された。Bradley-Terry モデルのもとで、InterChat(-) はテキストチャットに比べ、2倍以上の評価を得ており、キャラクタチャットに比べ9倍以上の評価を得た。

3.2 身体的引き込み動作の評価

表 3.2 一対比較の結果

	A	B	C	合計
A		30	14	44
B	10		2	12
C	26	38		64

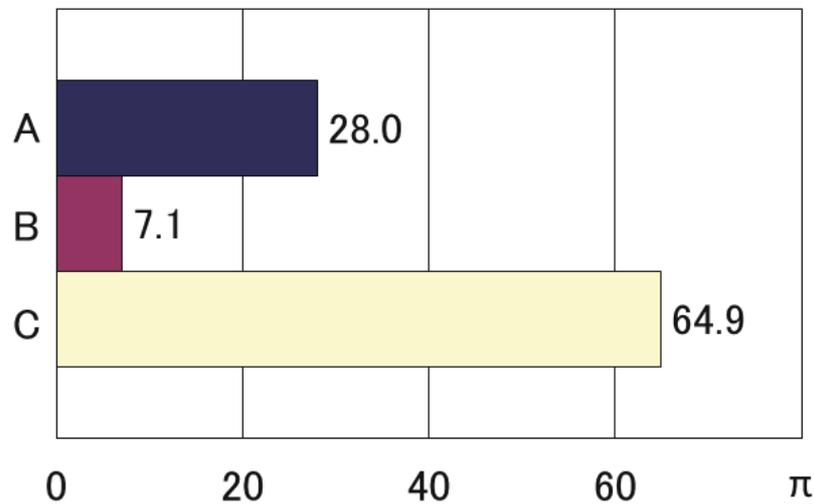


図 3.3 Bradley-Terry モデルによる好ましき π

7段階評価の結果を図 3.4 に示す。図では、7段階評価の平均値および標準偏差の結果と、それぞれの項目について Friedman 検定 [3.1] を行った結果を表している。「楽しさ」、「一体感」、「盛り上がる」の各項目では C と A, B の間に有意水準 1% で有意差が認められ、C の InterChat(-) が最も高く評価されている。「好き」、「チャット（対話）しやすさ」、「使用したい」の各項目では C と B, A と B の間に有意水準 1% で有意差が認められ、テキストチャットと InterChat(-) がキャラクタチャットより高く評価されている。「思いが通う」では C と A の間に有意水準 5%, C と B の間に有意水準 1% で有意差が認められ、C の InterChat(-) が最も高く評価されている。また、A と B の間に有意水準 1% で有意差が認められ、A のテキストチャットが B のキャラクタチャットより

第3章 InterChat のシステム評価

高く評価されており、キャラクタ表示ではなく相手の入力状態が把握できる機能が評価されている。

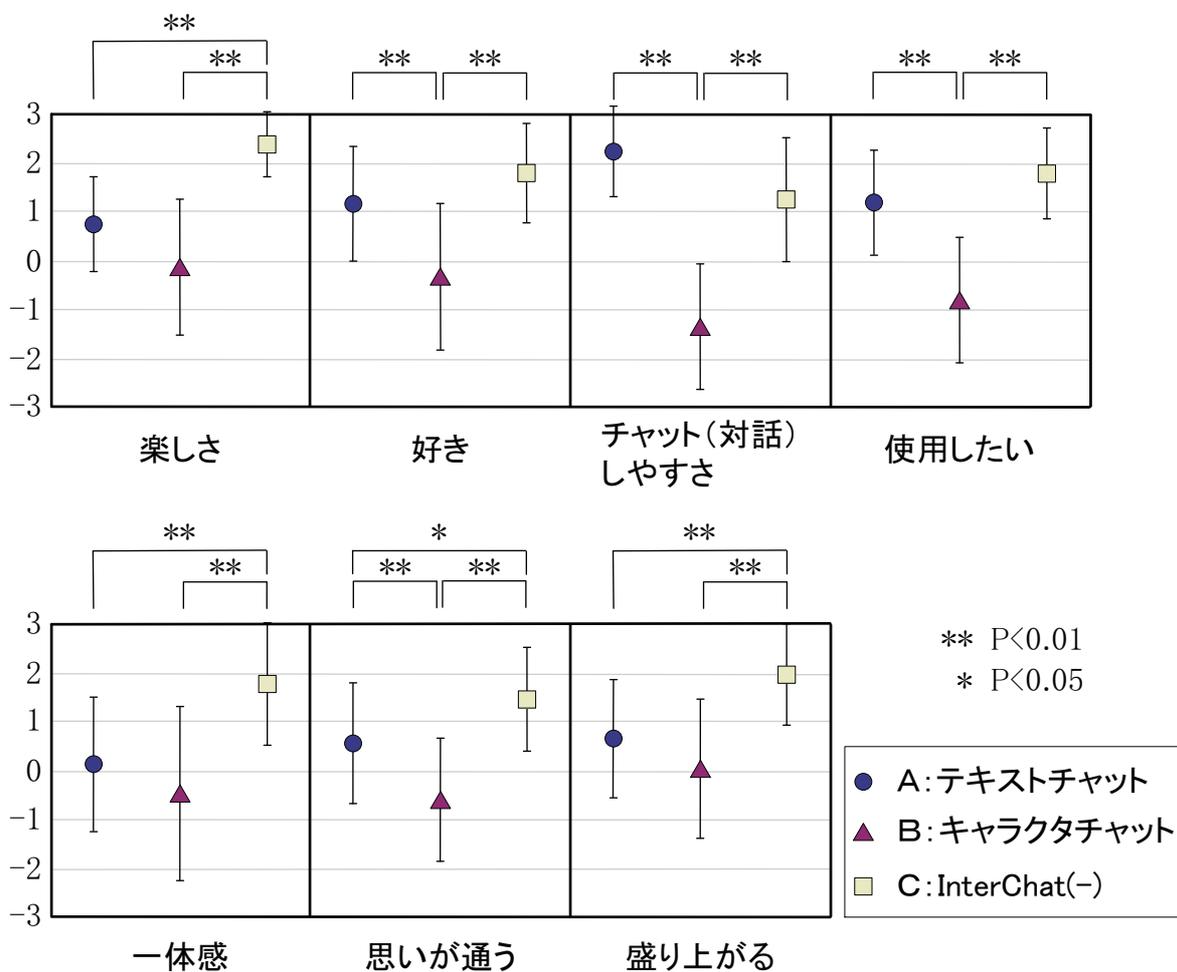


図 3.4 7段階評価

3.3 InterChat の評価

3.3.1 実験システム

本節では、文字列に対応した動作モデルを導入した InterChat の有効性を検証するために、図 3.5 に示すような、A: テキストチャット, C: InterChat(-), D: InterChat の

3.3 InterChat の評価

3種類の異なるモードのチャットシステムを用いた評価実験を行う。DのInterChatは話し手動作モデル、聞き手動作モデル、文字列に対応した動作モデルのすべての動作モデルを導入したInterChatである。C、Dは身体動作により相手に入力状態を提示する。また、本章の実験では、実験中の利用者のキー入力状態を時系列ログとして取得した。音声対話において、音声の重なりと場の盛り上がりの関係性が報告されており[3.2, 3]、タイピングを行うチャットコミュニケーションにおいては、キーの入力回数や相手とのキー入力の重なりが音声の場合と同様に、場の盛り上がりを示すのではないかと考え、キー入力状態の分析を行った。本節の評価実験において、文字列に対応した動作モデルはデータベースに登録されている文字が入力されるまでは動作の確認が行えない。そのため被験者にモード毎の違いを明確に説明し、文字列に対応した動作モデルを認識させるため、手順1、手順2において、各モードの使用時間を前章より1分間長くして実験を行った。

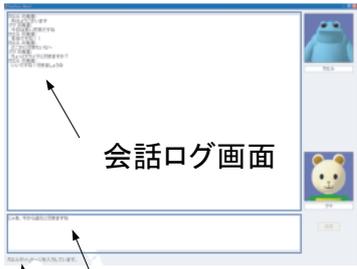
A: テキストチャット	C: InterChat(-)	D: InterChat
 <p>会話ログ画面</p> <p>入力ウィンドウ 相手の入力状態を表示</p>	 <p>引き込み動作のみ</p>	 <p>引き込み動作モデル + 文字列に対応した動作モデル</p>

図 3.5 チャットモード

前節での評価実験では、ディスプレイ（21.3型）の視野角度が広がったため、テキストチャットでは「余白部分が目立ち全画面表示でチャットを行うには違和感がある」といった意見や、キャラクタチャットやInterChat(-)では「視野範囲が広いため両者のキャラクタを見にくい」という意見があった。そのため、違和感なくチャットを行わせるために、被験者の視野角を狭め一目でチャットを行うことができるように、ノートPC (lenovo ThinkPad T60) を用いた。また、「InterChatにも会話ログ画面があった方がよい」といった意見や、「半角入力にも対応してほしい」といった機能面での要望があったため、会話ログ画面を追加し（図3.5）、半角入力・ファンクションキーなどのあら

第3章 InterChat のシステム評価

ゆるキー入力にも対応できる機能を追加した。会話ログ画面の位置は、研究室内の被験者に会話ログ画面配置の予備実験を行い、位置と大きさを自由に変化させ、チャットを行いやすい会話ログ画面を決定させた。その結果、画面上の配置分布として一番多かった画面上中央部に会話ログ画面を配置し、大きさは平均をとった。

3.3.2 実験方法

開発したシステムを使用して、別々の部屋に分かれた2人1組でチャットでコミュニケーション実験を行った。実験には、ノートPCを用いた。被験者は18~23歳の友人関係である男女20組40人（同性各10組）で、普段からパソコンを使用しタイピングにある程度慣れた学生である。

実験の手順を以下に示す。

手順1:

被験者に、各モードの違いと操作法を説明し、各モード毎に4分間試用させる。

手順2:

3種類のモードから2種類のモードをランダムにカウンターバランスをとって選び、各2分間使用させ、一対比較をさせる（計 ${}_3C_2 = 3$ 回）。

手順3:

被験者に3種類のモードについて3分間使用させた後、7項目による7段階評価をさせる。

一対比較は、総合評価として「総合的によい」という観点からどちらがよいかを選ばせた。7段階評価は、身体的コミュニケーション支援の観点から「楽しさ」、「好き」、「チャット（対話）しやすさ」、「使用したい」、「一体感」、「思いが通う」、「盛り上がる」の7項目について行った。A, C, Dのモードをランダムな順で使用させ、評価をアンケート用紙に答えさせた。被験者同士は同じモードを使用してチャットさせた。

3.3 InterChat の評価

3.3.3 実験結果

一対比較の結果を表 3.3 に示す。表中の数字は各行のモードを選択した回数を表している。この結果に対して被験者による評価を定量的に評価するために、Bradley-Terry モデルを想定した。一対比較の結果から最尤推定した好ましさ π (合計 100) の値を図 3.6 に示す。D の InterChat が最も高く評価され、次いで C の InterChat(-)、A のテキストチャットの順に評価された。Bradley-Terry モデルのもとで、D は他のモードに比べ極めて高い評価を得た。

7 段階評価の平均値および標準偏差の結果と Friedman 検定を行った結果を図 3.7 に示す。「チャット (対話) しやすさ」の項目を除いた全ての項目で、D と C、A の間に有意水準 1% で有意差が認められ、D の InterChat が最も高く評価された。

表 3.3 一対比較の結果

	A	C	D	合計
A		17	3	20
C	23		2	25
D	37	38		75

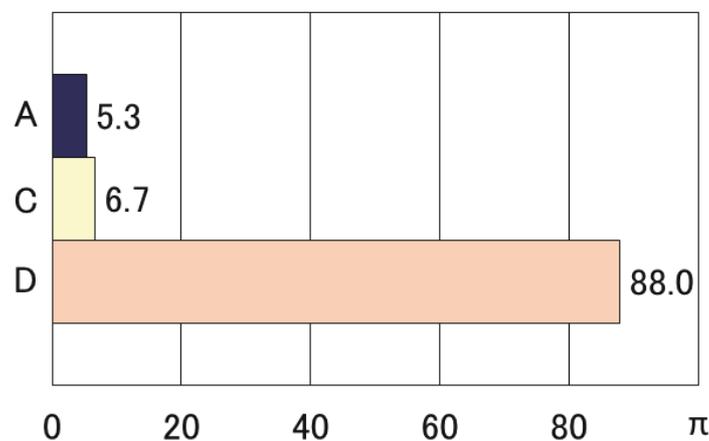


図 3.6 Bradley-Terry モデルによる好ましさ π

第3章 InterChat のシステム評価

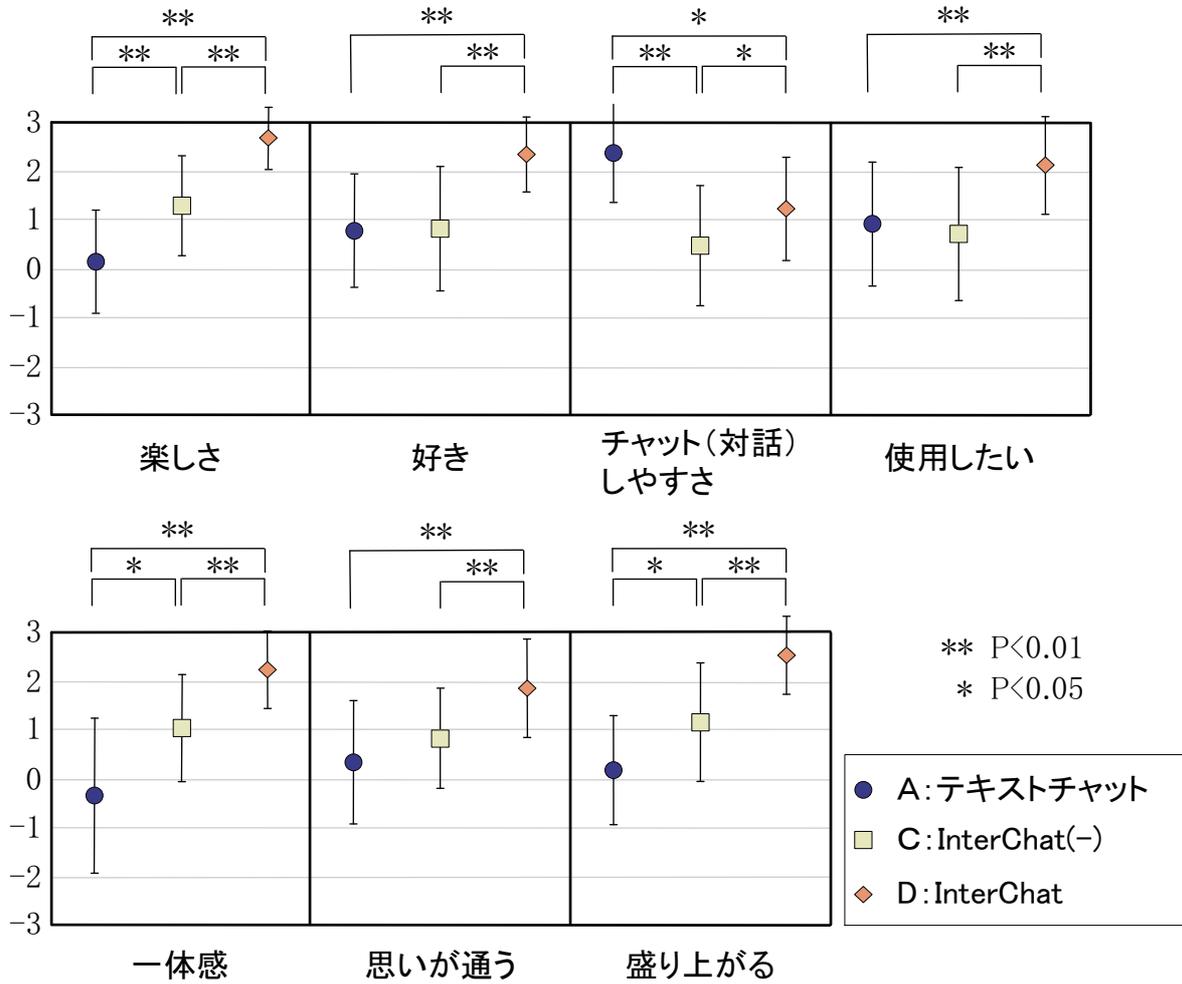


図 3.7 7段階評価

また、7段階評価時の各モード3分間のキー入力回数と、相手と入力重なった回数の平均値と標準偏差、一要因の分散分析の結果を図3.8に示す。キーの入力回数では、CとAの間に有意水準5%で有意差が認められ、CのInterChat(-)はAのテキストチャットよりもキー入力回数が多く、また、相手と入力重なった回数では、CとA、DとAの間に有意水準1%で有意差が認められ、DのInterChatとCのInterChat(-)はAのテキストチャットよりも相手と入力重なった回数が多かった。

3.4 考察

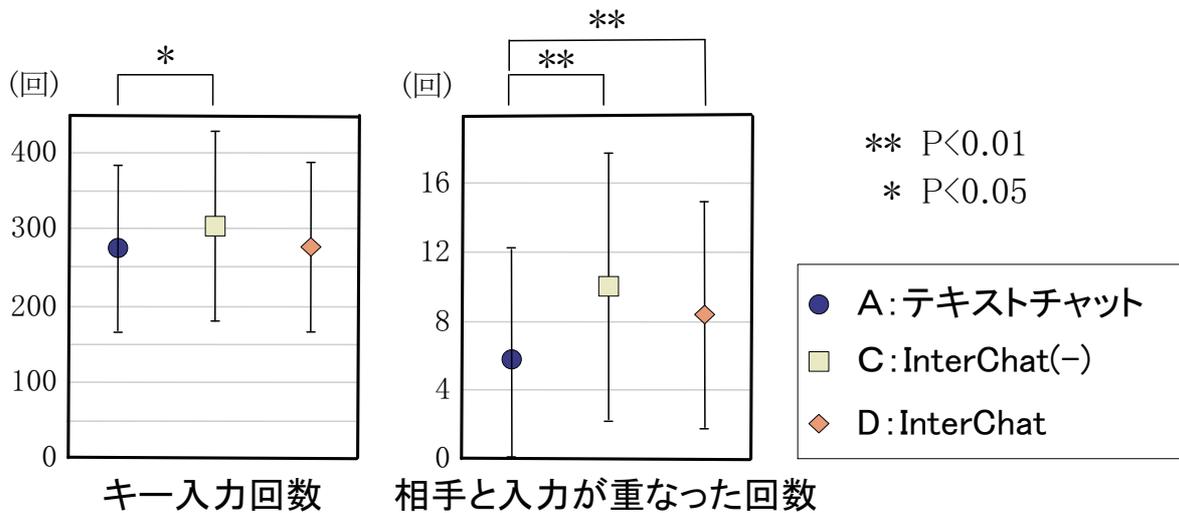


図 3.8 タイピング情報の分析結果

3.4 考察

チャットコミュニケーション実験の結果、一対比較において InterChat が他のモードに比べて高く評価された。また、7段階評価においても、InterChat がすべての項目において肯定的な評価が得られ、とくに「楽しさ」、「一体感」、「思いが通う」、「盛り上がる」の項目で他のモードよりも高く評価された。これは、キャラクターチャットと InterChat(-) との評価より、タイピングリズムより身体的引き込み動作をリアルタイムに自動生成することでキャラクターがただ存在しているだけでは共有し難い身体性を共有できているからと考えられる。このように、チャットなどタイピングを利用したメディアにおいても iRT を導入して身体性を共有することにより、一体感を深め、円滑なコミュニケーションを行うことができると考えられる。一方で、「チャット (対話) しやすさ」の項目でテキストチャットが高く評価されており、文字のみによるシンプルな構成が評価された可能性がある。

また、InterChat はテキストチャットに比べ相手とキー入力重なった回数が多かった。前述のとおり、音声の重なりと場の盛り上がりに関係性がみられることと同様に、キー入力の重なりが、InterChat の場の盛り上がりの評価につながったものと考えられる。

これらの結果より、InterChat はチャットコミュニケーションを支援するシステムとし

て有効であることがわかる。ただし、本実験ではチャット相手の被験者はお互いに友人であり、チャット内容是对立意見を言い合うような場面を想定していないので、チャットの内容や対話相手との関係性を含めた結果を示しているわけではない。

3.5 結言

本章では、身体的リズムの引き込みに着目し、キャラクタチャットでのコミュニケーションにおいて、チャット参加者のタイピングに基づいてキャラクタの身体性を共有するキャラクタチャットシステムを開発し、評価を行った。

まず InterChat の身体的引き込み動作に着目し、新たにテキストチャット、キャラクタチャットを開発し、これら3種類の異なるモードを比較検討することで InterChat の特徴を評価した。官能評価結果より、一対比較による総合評価において InterChat が高く評価され、また7段階評価により InterChat の身体的引き込み動作が楽しさや一体感、思いが通う、盛り上がるなどの項目でとくに高く評価された。また文字列に対応した動作を含む InterChat を、改めてテキストチャットおよび身体的引き込み動作のみのシステムと比較評価した。その結果、一対比較で InterChat が非常に高く評価され、また7段階評価でも多くの項目で高く評価されるなど InterChat の特徴を明らかにし、チャットコミュニケーションでの有効性を示した。

参考文献

- [3.1] 広津 千尋：実験データの解析—分散分析を超えて—，共立出版 (1992).
- [3.2] 伊藤 秀樹，重野 真也，西本 卓也，荒木 雅弘，新美 康永：対話における雰囲気
の分析，音声言語情報処理, Vol. 2002, No. 10, pp. 103–108 (2002).
- [3.3] 西村 良太，北岡 教英，中川 聖一：対話における韻律変化・タイミングのモデル
化と音声対話システムへの適用，人工知能学会言語・音声理解と対話処理研究
会, Vol. 48, pp. 37–42 (2006).

第4章

テキストのリアルタイム入力状態表示手法

4.1 緒言

チャットを題材としたオンラインテキストコミュニケーションについて、以前からそのコントロール可能性、編集可能性、および時間構造といった特徴に着目して検討がされており、ノンバーバル情報が伝達されにくい状況下でのコミュニケーションが議論されている [4.1, 2, 3]. リアルタイムでのテキストコミュニケーションにおいて送受信されるテキスト情報は、たとえオンライン上であってもコントロール可能で、編集可能であり、バーバル情報による送信者の意図的な情報として扱われる。しかしテキスト情報の送信タイミングや、編集過程などの情報も、受信者への提示方法によっては非意図的な情報として間接的に伝わる場合があり、これらの影響を考慮しておく必要がある [4.4, 5]. またチャットコミュニケーションにおいて、テキストの編集状態提示をリアルタイムなインタラクション支援のために利用する提案 [4.6, 7, 8] から、テキストの編集過程をより早く提示し、リアルタイムにテキスト情報を共有する効果についても検討が必要である。

前章までに開発評価してきたシステムは、タイピング入力から iRT により身体的引き込み動作を生成し、また送信されたテキスト情報に基づいて、予め定義された文字列と一致した場合に既定動作を重畳合成するシステムである。しかしこの手法では、タイピング中に身体的引き込み動作を生成することで身体的インタラクションを実現していたが、テキスト情報自体は従来のチャットと同様に全ての入力が確定した後に送信していたため、受信者には自動生成された身体的引き込み動作と送信されたテキスト情報が時間的に切り離されて伝達されていた。テキスト情報の送受信がチャットシステムの本来の目的だとすれば、テキストの入力状態をテキスト情報から得ることがより直感的であり、キャラクタも同期して動作すればより高い効果が得られる可能性がある。

本章では、自動生成された身体的引き込み動作をより効果的に受信者に提示するために、テキスト入力状態と身体的引き込み動作を同期表示することで、テキスト情報とキャラクタ動作を同時に観察できるシステムを開発する。そのため、送信者のテキスト入力状態とキャラクタの身体動作を受信者に同期表示する手法について、従来手法である入力中はテキスト情報を表示せず送信時に全て表示する送信時表示手法に加え、入力中の文字列を中点 (·) に変換して表示することで入力状態を提示する中点表示手法、および変換中も含めた全てのテキスト情報をリアルタイムで表示するリアルタイム表示手法について、2つのコミュニケーション実験によってその効果を検討する。

4.2 テキスト入力状態表示の検討

4.2.1 テキスト入力状態表示機能

本研究で開発した InterChat は、仮想空間内にチャット参加者の代役となるキャラクタを2体配置し、互いのキャラクタの身体動作によって使用者間で身体的リズムを共有しながらチャットを行うコミュニケーションシステムである。しかしこれまでのシステムでは、タイピングによる入力状態をキャラクタの身体動作のみで提示しており、同じモダリティとしてのテキスト情報そのものは送信後に相手側に表示されていた。そのため図4.1に示すように、相手側のメッセージの表示方法が異なる3つのチャットモードを開発し、キャラクタの身体動作とテキスト情報を提示する効果について検討を行う。

Aの送信時表示は、従来の InterChat での表示方法であり、メッセージ送信時に全てのテキスト情報をキャラクタの頭上に吹き出しとして表示する。送信者の入力状態はキャラクタの身体動作のみで伝達される。Bの中点表示は、入力中の文字列を中点 (·) に変換して身体動作と同期表示する。メッセージ送信時に入力中の中点がテキスト情報に変換されて表示される。半角1文字と半角中点1つが対応しており、全角文字入力時には中点が2つずつ増える。変換による文字数の減少や、BackSpace 使用時には文字数に対応して中点の数も減少する。また、変換中の文字列に対応する中点は灰色で表示される。これにより、入力時のリズムや、勢い、文字数、修正状況などの入力状態をキャラクタの身体動作と同期して伝達することが可能となる。Cのリアルタイム表示は、変換の様子なども含めた全てのテキスト情報を身体動作と同期してリアルタ

4.2 テキスト入力状態表示の検討

イムでキャラクタの頭上に吹き出しとして表示する．入力中や変換中の確定前文字列は灰色で表示される．図 4.2 に 3 つの表示方法の動作の一例を示す．また，システム使用中は 30Hz でキー押下の ON-OFF データやタイピング活性量，入力状態などといったタイピングデータを取得し保存する機能を持つ．

また，チャット中の入力では文節で変換することや Enter キーでメッセージを送信することが多くその検出も容易である．そこで，聞き手の動作において変換開始時の Space キーや，送信時の Enter キーの押下を検出し，メッセージの区切りで確実にうなずき反応を行わせるよう改良を加えた．また，チャットでは「はい」や「w」などの極めて短いメッセージを送信する場面が多くみられるため，話し手の動作においてハンズオーバー処理後の ON-OFF データを 8/30 秒毎に検出し，ON ならば体の各部位を動作させることで短い入力に対する身体動作の生成を可能にした．

4.2.2 評価実験

各モードを使用してチャットコミュニケーションを行った．本実験はチャットコミュニケーション時のテキスト入力状態に対して，従来の InterChat が持つキャラクタ動作

	A:送信時表示	B:中点表示	C:リアルタイム表示
受信前			
送信 ↓			
受信後			

図 4.1 実験で使用した 3 モード

第 4 章 テキストのリアルタイム入力状態表示手法

	相手 入力キー	g	a	m	Back Space	n	n	b	a
入力情報	送信時表示								
	中点表示	⋯	⋯	⋯⋯	⋯	⋯⋯	⋯⋯	⋯⋯⋯	⋯⋯⋯
	リアルタイム 表示	g	が	がm	が	がn	がん	がんb	がんぼ
身体動作	話し手								
	聞き手								

r	u	変換 Space	確定 Enter	z	o	確定 Enter	送信 Enter
							頑張るぞ
⋯⋯⋯	⋯⋯⋯	⋯⋯⋯	⋯⋯⋯	⋯⋯⋯	⋯⋯⋯	⋯⋯⋯	頑張るぞ
がんばる	がんばる	頑張る	頑張る	頑張るz	頑張るぞ	頑張るぞ	頑張るぞ
							
							

図 4.2 メッセージ表示方法

4.2 テキスト入力状態表示の検討

(A)による理解に加え、同じモダリティであるテキスト情報の中点表示(B),あるいは全てのテキスト情報の表示(C)による理解が与えるコミュニケーション効果について検討する。Cについては変換前のテキストも全て表示されるため、テキストの編集可能性が低く、場合によっては低く評価される可能性がある。

実験のタイムテーブルを図4.3に示す。被験者は18~23歳の男女学生15組30人(同性同士)で、普段からパソコンを使用している学生である。

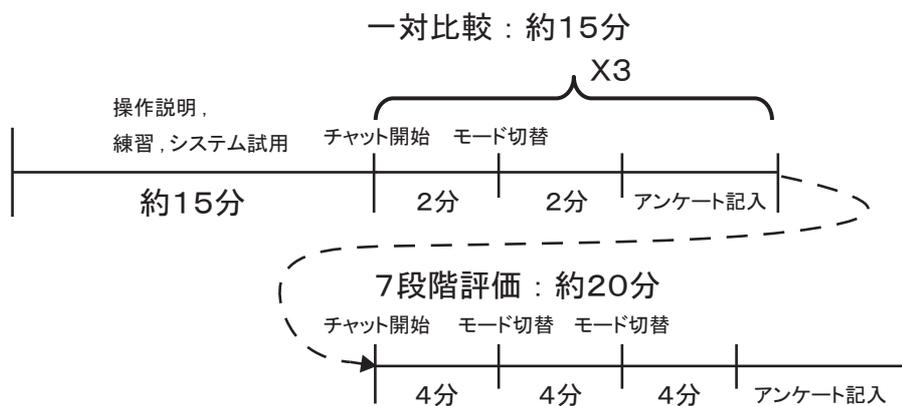


図 4.3 実験のタイムテーブル

キーボードと画面を自然な形で近づけるためノート PC を使用した。被験者 2 人 1 組で別々の部屋に分かれ、用意した 1 組のノート PC を LAN ケーブルで接続し使用した。被験者には実験開始時に椅子およびノート PC の位置とディスプレイの傾きを自由に調整させ、モードの切り替えは全て実験者が行った。実験中のタイピングデータはログデータとして保存した。図 4.4 に実験風景を示す。

実験はまず、最初に各モードの特徴と操作方法の説明を被験者に行った。操作方法に慣れるように、各モード 3 分ずつ使用させた。この間およそ 15 分程度である。その後、2 つの評価法を用いて比較させた。最後に実験に関する感想を自由記述形式で記入させた。

評価法 1: 一対比較

A~C のモードから、提示順による評価への影響あるいはシステムへの慣れによる対話への影響を抑えるため、ランダムに 2 つを一対比較させ、これを計 3 ($=_3C_2$) 回繰り返した。まず、1 つのモードでチャットをした後に別のモードによるチャットを行い、

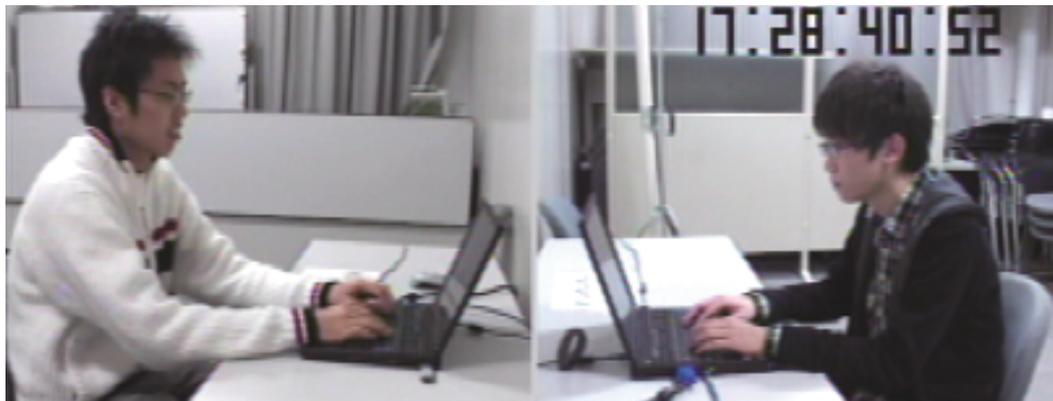


図 4.4 実験風景

2モードを使用した後に「総合的に良い」という観点からどちらがよかったかアンケート用紙に記入させた。対話者同士は同じモードを使用してチャットし、1つの比較につき対話時間を4分（1つのモードにつき2分）とした。

評価法 2: 7段階評価

A～Cのモードを「楽しさ」、「好き」、「チャット（対話）しやすさ」、「使用したい」、「一体感」、「思いが通う」、「盛り上がる」の7項目について評価させた。各モードをランダムな順番で使用し、その後、各モードがどの程度その項目に当てはまるかを7段階（中立0）でアンケート用紙に記入させた。対話者同士は同じモードを使用してチャットし、1つのモードにつき対話時間を4分とした。

4.2.3 実験結果

一対比較の結果を表4.1に示す。表中の数字は各行のモードを選択した回数を表している。この結果に対して被験者による評価を定量的に評価するために、Bradley-Terryモデルを想定した(式3.1, 3.2)。これにより一対比較に基づく評価を一義的に定めることができる。一対比較の結果から最尤推定した π の値の比較を図4.5に示す。Bの中間表示が最も高く評価され、次いでCのリアルタイム表示、Aの送信時表示の順に評価された。次に7段階評価の結果およびFriedman検定を行った結果を図4.6に示す。Bの中間表示はAの送信時表示に対して全ての項目で高く評価され、「チャット（対話）しやすさ」「一体感」「思いが通う」の項目では有意水準1%で、「好き」「使用したい」「盛

4.3 入力文字の表示タイミングの検討

り上がる」の項目では有意水準5%で有意差が認められた。また、Cのリアルタイム表示はAの送信時表示に対して「楽しさ」「一体感」「思いが通う」「盛り上がる」の項目で有意水準1%で有意差が認められたが、「好き」「チャット（対話）しやすさ」「使用したい」の項目では同程度の評価であることが分かった。

表 4.1 一対比較の結果

	A	B	C	合計
A		5	14	19
B	25		19	44
C	16	11		27

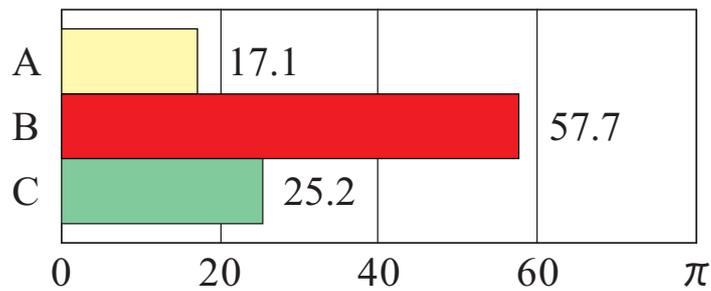


図 4.5 Bradley-Terry モデルによる好ましさ π

4.3 入力文字の表示タイミングの検討

前節の実験結果より、一対比較では中点表示がより高く評価された結果が得られたが、一方で7段階評価ではリアルタイム表示も評価項目によっては高く評価される傾向にあることが分かった。中点表示では、テキストの編集可能性および入力状態と身体動作の同期表示が評価されたと考えられ、またリアルタイム表示の評価では、メッセージ入力情報と身体動作の同期性に対して全ての入力メッセージが伝わることのライブ感が評価された可能性がある。しかし編集可能性が失われることで変換過程や入力ミスといった未確定の入力メッセージが表示されてしまうことで一部の項目で評価

第4章 テキストのリアルタイム入力状態表示手法

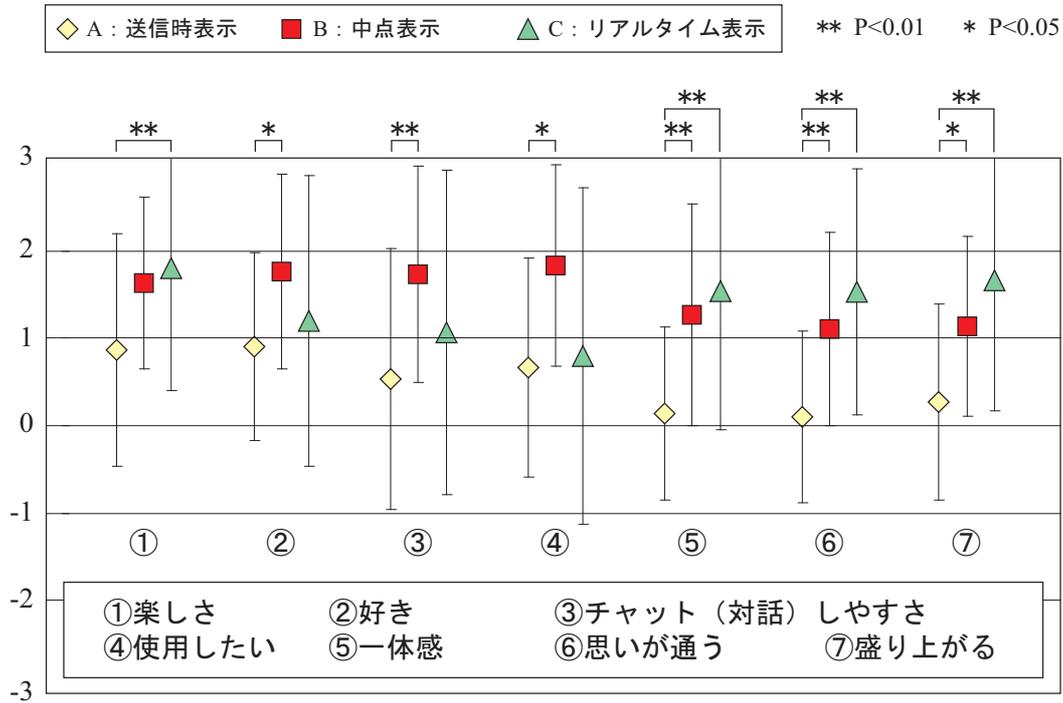


図 4.6 7段階評価の結果

を下げたと考えられる。そこで、入力情報における確定文字情報のみを参加者間で同期する機能を開発し、効果を検討する。

4.3.1 確定文字同期表示機能

前節の実験で最も高く評価された「中点表示」に加え、新たに「中点逐次確定表示」、
「変換過程隠し表示」の2つの表示手法を開発した(図4.7)。

Iの中点表示は、前章の「中点表示(Bモード)」と同じ表示方法である。IIの中点逐次確定表示は、中点表示を発展させたもので、入力中の状態は中点表示と同様に相手の入力メッセージが中点(·)に変換されて表示されるが、メッセージ確定時に確定部分のみが中点(·)からテキスト表示に切り替わる。これにより中点表示よりも早く相手に内容を伝え、身体動作との同期性を確保しつつ入力過程の秘匿も行っている。IIIの変換過程隠し表示は入力中のメッセージがそのまま相手に表示される。ただし、入力文字変換時はその過程を相手に見せず、変換が確定した時に変換文字に切り替わり相手に表示することで、送信相手に変換過程を見せないようにしている。

4.3 入力文字の表示タイミングの検討

状態	入力	変換	確定	入力	確定	送信
相手 入力文字	がんばる	頑張る	頑張る	頑張る！！	頑張る！！	頑張る！！
受信側	中点表示	頑張る！！
	中点逐次 確定表示	頑張る.....	頑張る！！	頑張る！！
	変換過程 隠し表示	がんばる	がんばる	頑張る！！	頑張る！！	頑張る！！
身体動作						

図 4.7 メッセージ表示過程

4.3.2 評価実験

被験者 2 人 1 組で別々の部屋に分かれ、相手側メッセージの表示方法が異なる I~III の各モードでチャットを行った (図 4.8)。前節の実験と同様に、まず被験者に各モードを 3 分間試用させ、その後各 2 分間ずつの一対比較をさせた。次に、3 種類のモードについて 4 分間使用させた後、7 項目について 7 段階 (中立 0) で官能評価をさせた。I~III のモードの比較順はランダムとし、被験者同士は同一のモードでチャットを行った。被験者は前節の実験とは異なる 18~25 歳の男女学生 12 組 24 人 (同性同士) で、普段からパソコンを使用している学生である。



図 4.8 実験で使用した 3 モード

4.3.3 実験結果

一対比較の結果を表 4.2 に示す。表中の数字は各行のモードを選択した回数を表している。この結果に対して被験者による評価を定量的に評価するために、Bradley-Terry モデルを想定した (式 3.1, 3.2)。一対比較の結果から最尤推定した好ましさ π の値を図 4.9 に示す。II のモードが最も高く評価され、次いで III, I の順に評価された。7段階

表 4.2 一対比較結果

	I	II	III	合計
I		6	8	14
II	18		13	31
III	16	11		27

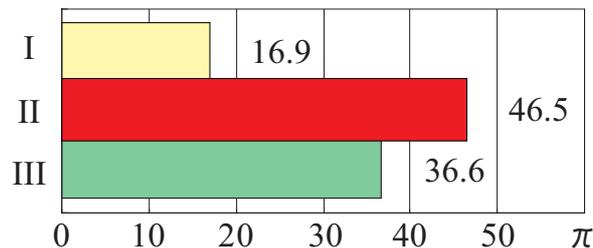


図 4.9 Bradley-Terry モデルによる好ましさ π

評価の結果を図 4.10 に示す。II の中点逐次確定表示は全ての項目で高く評価され、とくに、「楽しさ」「使用したい」「一体感」「思いが通う」「盛り上がる」の項目では、I に比べて有意水準 1% で有意差が認められた。また、III の変換過程隠し表示は「楽しさ」「使用したい」「一体感」「思いが通う」「盛り上がる」の項目で、I に比べて有意水準 1% で有意差が認められた。

また、2章で確認された音声とタイピングリズムの類似性に基づいて、会話の盛り上がるの観点からタイピング量について分析した。7段階評価時の各モード4分間の平均無言時間（対話者がお互いに入力していない一区分あたりの時間）、無言回数（対話者がお互いに入力しなかった回数）、一回確定送信回数（最初に確定した文字列をそ

4.4 考察

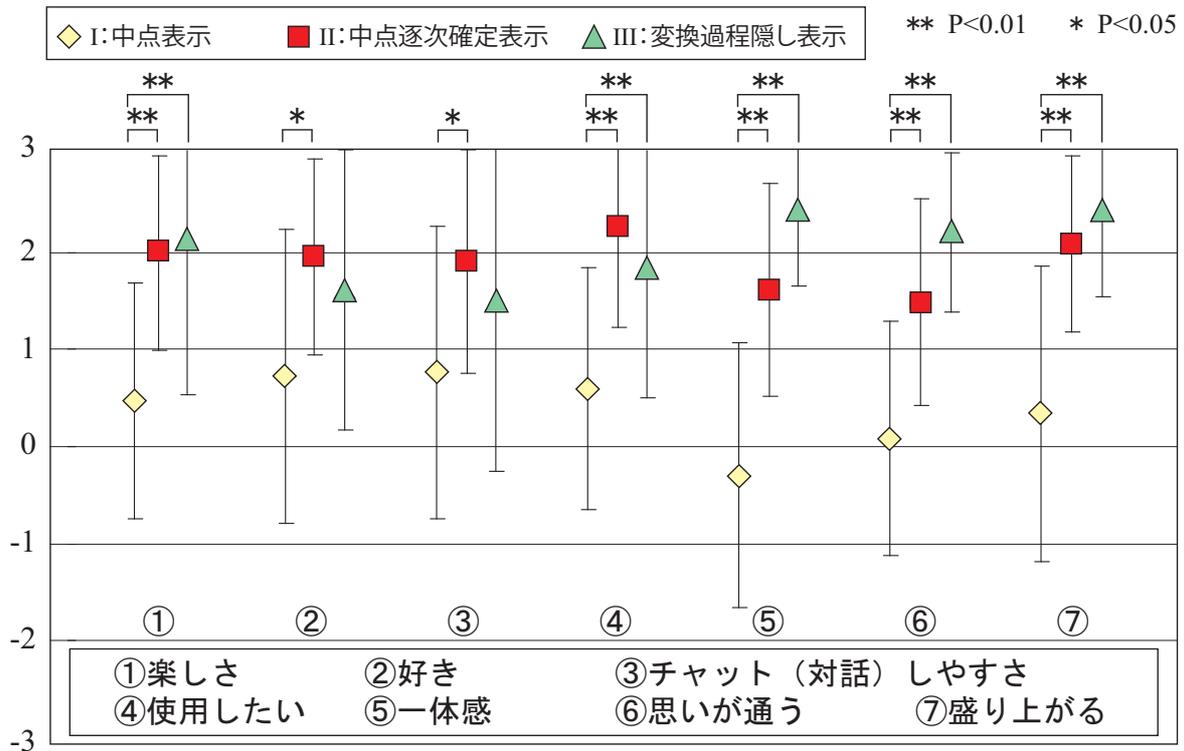


図 4.10 7段階評価の結果

のまま送信した回数), BackSpace回数(BackSpaceキーを押下した回数)の平均値と標準偏差, 一要因の分散分析による結果を図4.11に示す. 平均無言時間はIIIが最も短く, Iに比べて有意水準1%, IIに比べて有意水準5%で有意差が認められた. 無言回数はIIIが最も多く, IIに比べて有意水準5%で有意差が認められた. 一回確定送信回数はIIIが最も多く, IIに比べて有意水準5%で有意差が認められた. BackSpace回数はII, I, IIIの順に多く, IIはIIIに比べて有意水準1%, IはIIIに比べて有意水準5%で有意差が認められた. これら以外に, タイピング速度や送信回数, 無言時間の合計などのタイピング情報を分析したが, とくに有意差は認められなかった.

4.4 考察

本章では, 対話者の入力状態をキャラクタの身体動作だけでなく, テキスト入力状態を表示することで互いの発話状態を捉え, より効果的に対話空間を利用する手法の

第4章 テキストのリアルタイム入力状態表示手法

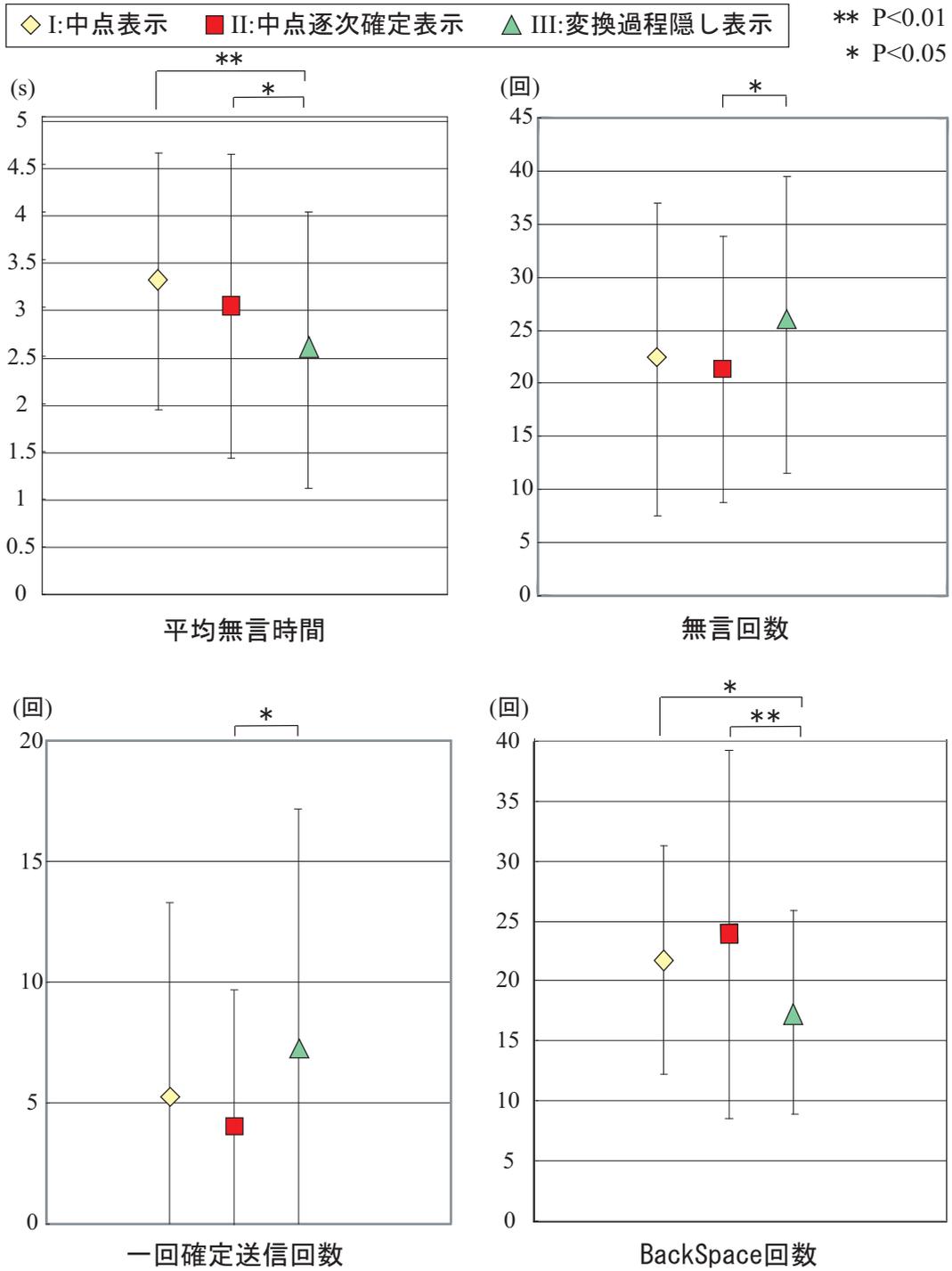


図 4.11 タイピング情報の分析結果

4.4 考察

検討を目的としている。4.3節では従来の提示方法である送信時表示，入力状態を全て中点に変換した中点表示，入力テキストを全て提示するリアルタイム表示の3モードを比較した。結果的に中点表示とリアルタイム表示が好まれる傾向にあり，4.4節において両機能の特徴を活かした中点逐次確定表示，変換過程隠し表示を提案し詳細に検討した。実験の結果，一対比較においてIIの中点逐次確定表示は他のモードに比べて高く評価された。さらに，7段階評価においても，Iの中点表示に比べて全ての項目で高く評価された。これは，メッセージを確定してから送信するというチャットの性質を保ちつつ，相手のメッセージ作成の様子をキャラクタの身体動作とともに確認できることにより，実用性と一体感がともに向上した結果だと考えられる。

一方，IIIの変換過程隠し表示は，7段階評価において，Iの中点表示と比較して，「楽しさ」「使用したい」「一体感」「思いが通う」「盛り上がる」の項目で高く評価された。またIIの中点逐次確定表示と比較して，「一体感」「思いが通う」「盛り上がる」の項目で高く評価され，リアルタイム性が活かされた機能が高く評価された。さらに，自由記述アンケートにおいて，「本当に話しているみたい」「途中でも読め，早く返信ができる」といった意見も得られた。これは，メッセージをリアルタイムで伝達することにより，身体動作とテキスト表示の同期性が高まり身体的リズムの共有が強まることで，メッセージ内容の把握が円滑になり，より一体感のあるチャットコミュニケーションが実現できたと考えられる。

また，タイピング情報の分析結果より，無言時間の合計に差はなく，IIIが無言回数が多いため平均無言時間が短くなり，III，II，Iの順に短くなったことが分かる。

さらに，IIIが一回確定送信回数が多いことから，文を短く区切って，より早く相手とかかわろうとする傾向がみられる。ここでBackSpace回数はIIIが最も少なく，入力ミスも含め，全てリアルタイムで相手に伝わっているため，編集過程を気にせずに送信している。IIは確定しなければ相手に表示されないため，テキストが編集される回数が増加したものと考えられる。よって，よりリアルタイム性を重視する場合はIIIの変換過程隠し表示で，正確性を考慮する場合はIIの中点逐次確定表示がより適切に利用できると考えられる。

とくに本研究での評価実験は既知の友人同士を対象として行われており，テキストメディアの持つ編集可能性に着目した場合，より正確な情報伝達を行うべき相手とのコミュニケーションにおいては異なった結果となる可能性も高く，さらに詳細な検討

が必要である。

4.5 結言

本章では、仮想空間におけるテキストとCGキャラクタを用いたチャットコミュニケーションに対して、対話者のタイピング入力に基づいて、テキスト入力状態と自動生成された身体的引き込み動作を同期表示し、テキスト情報とキャラクタ動作を同時に観察できるシステムを開発した。またメッセージ送信者のテキスト入力状態とキャラクタの身体動作を受信者に同期表示する手法について、従来手法である入力中はテキスト情報を表示せず送信時に全て表示する送信時表示に加え、入力中の文字列を中点 (·) に変換して表示することで入力状態を提示する中点表示、および、変換中も含めた全てのテキスト情報をリアルタイムで表示するリアルタイム表示に着目して検討した。その結果、中点表示において、入力中の状態は中点表示と同様に相手の入力メッセージが中点 (·) に変換されて表示され、さらにメッセージ確定時に確定部分のみが中点 (·) からテキスト表示に切り替わる中点逐次確定表示手法が高く評価された。よってキャラクタの身体的引き込み動作によってテキスト入力状態を把握しつつ、同じモダリティであるテキストも同期表示することでより高いコミュニケーション効果が得られることが確認された。

参考文献

- [4.1] 原田 悦子：人の視点からみた人工物研究：対話における「使いやすさ」とは（認知科学モノグラフ⑥），共立出版（1997）.
- [4.2] 細馬 宏通：チャットは何を前提にしているかーチャットの時間構造と音声会話の時間構造（身体性とコンピュータ 第6部5章），共立出版，pp. 338–349（2001）.
- [4.3] 杉谷 陽子：メールはなぜ「話しやすい」のか?: CMC(Computer-Mediated Communication)における自己呈示効力感の上昇，社会心理学研究，Vol. 22, No. 3, pp. 234–244（2007）.
- [4.4] 角野 清久，西本 一志：言外の情報としての編集過程情報を伝えるメールシステムの提案と評価，情報処理学会論文誌，Vol. 50, No. 1, pp. 254–267（2009）.
- [4.5] Hua Wang, Helmut Prendinger, Mitsuru Ishizuka, and Takeo Igarashi: Affective Communication in Online Chat Using Physiological Sensors and Animated Text, Journal of Human Interface Society, Vol. 7, No. 1, pp. 39–45（2005）.
- [4.6] Kanayo Ogura, Takeshi Masuda, and Masato Ishizaki: Building a New Internet Chat System for Sharing Timing Information, Proceedings of the 4th SIGdial Workshop on Discourse and Dialogue, pp. 97–104（2003）.
- [4.7] Tara Rosenberger Shankar, Max VanKleek, Antonio Vicente, and Brian K. Smith: Fugue: A Computer Mediated Conversational System that Supports Turn Negotiation, Proceedings of the 33rd Hawaii International Conference on System Sciences-Volume 3, p. 3035（2000）.
- [4.8] 山田 祐士，竹内勇剛：非交替型チャットシステムの開発と社会的な対話のダイナミクスの解析，電子情報通信学会，信学技報，HCS2002-50, pp. 19–24（2003）.

第5章

情動表現の即時表示手法

5.1 緒言

4章ではタイピング情報に基づいて自動生成された身体的引き込み動作に対して、好ましいテキスト情報提示手法の分析評価を行った。ここでテキストに対応した情動表現はメッセージの送信後にテキストが認識されて初めて対応する動作が提示されており、発話者の入力時の感情とキャラクタの情動表現との間に時間的なずれが生じている。また送信メッセージに対応のある単語が複数含まれていた場合には、最初に検出された単語に対応する動作しか提示されない、などの問題点があった。そこで本章では、テキストの表示とテキストに対応したキャラクタ動作の提示時間の差を縮め、より高いコミュニケーション支援効果を目指した手法を開発する。またこの手法を用いてチャットコミュニケーション実験を行い、テキストに対応した情動表現の好ましい提示タイミングを分析評価することで、コミュニケーション支援における開発システムの効果を検討する。

5.2 キャラクタチャットシステム

5.2.1 対話時の画面表示

メッセージ入力時、タイピングの ON-OFF データやメッセージ内容はネットワークでつながった他の参加者と常時共有されている。これにより、他の参加者が入力のリズムや内容といった入力情報と、キャラクタの身体動作の時間的關係を保って確認することが可能となる。本手法では、テキストに対応する情動表現を、より早く相手に表示させる観点から、テキスト表示については前章で一体感や場の盛り上がりに対しての有効性が示された「メッセージを入力の際に送信し、他の参加者の画面に吹き出

第5章 情動表現の即時表示手法

しとして同期表示する」リアルタイム表示手法を採用する。メッセージ送信側の入力状態と表示される画面の関係の例を図5.1に示す。共有されたメッセージは各参加者のPCに表示され、全てのシステム使用者が同様の入力状態を確認することができる。InterChatの使用風景を図5.2に示す。

状態	入力	入力	入力	変換	確定	送信
送信側 入力文字	が	がん	がんばる	頑張る	頑張る	頑張る
受信側 表示画面						

図 5.1 メッセージ表示手法

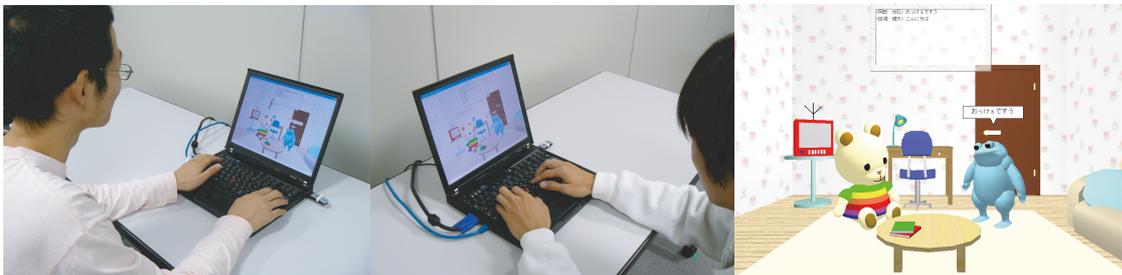


図 5.2 InterChat の使用風景

5.2.2 情動表現の即時表示手法

文字列に対応した動作を提示する方法を検討するために、動作の提示タイミングに着目し、従来のInterChatでの動作提示手法として使用されていた、メッセージの送信時に文字列に対応した動作を提示する「通常動作」に加え、新たに「確定時動作」、「入力時動作」の2つの提示手法を開発した。

確定時動作では、入力中のメッセージが確定される度にそれまで入力していた文章の認識を行い、特定の文字列が文中にあった場合はその文字列に対応した動作を提示

状態 表示文字	入力 さむ	入力 さむい	入力 さむいけど、	変換 寒いけど、	確定 寒いけど、	入力 寒いけど、	入力 寒いけど、	入力 寒いけど、	変換 寒いけど、	確定 寒いけど、	送信 寒いけど、
通常											
確定時											
入力時											

身体動作

図 5.3 動作提示手法

する。入力時動作では、メッセージが入力される度にそれまで入力していた文章の認識を行い、特定の文字列が文中にあった場合はその文字列に対応した動作を提示する。これらの3つの動作提示手法の比較を図5.3に示す。

5.3 チャットコミュニケーション実験

5.3.1 実験方法

本章では、好ましい情動表現の提示タイミングを検証するために、被験者2人1組で別々の部屋に分かれ、「A: 通常動作」、「B: 確定時動作」、「C: 入力時動作」の3つのモードを使用してチャットコミュニケーションを行った。被験者には最初に、システム全体と各モードの特徴を説明した。その後別々の部屋に分かれ、システムに慣れるために各モードを3分ずつ使用させてから、2つの評価法を用いて比較させた。最後に実験に関する感想を自由記述形式で記入させた。被験者は18～24歳の男女学生12組24人（同性同士）で、普段からパソコンを使用しておりタイピングにある程度慣れた学生である。

評価法1: 一対比較

A～Cのモードからランダムに2つを一対比較させ、これを計3 ($=_3C_2$) 回繰り返した。まず、1つのモードでチャットをした後に別のモードによるチャットを行い、2モードを使用した後に「総合的によい」という観点からどちらがよかったかアンケート用紙に記入させた。対話者同士は同じモードを使用してチャットし、1つの比較につき対話時間を4分（1つのモードにつき2分）とした。

評価法2: 7段階評価

A～Cのモードを「楽しさ」、「好み」、「チャットしやすさ」、「使用したい」、「一体感」、「思いが通う」、「盛り上がる」の7項目について評価させた。各モードをランダムな順番で使用し、その後、各モードがどの程度その項目に当てはまるかを7段階（中立0）でアンケート用紙に記入させた。対話者同士は同じモードを使用してチャットし、1つのモードにつき対話時間を3分とした。

5.3.2 実験結果

実験中のタイピング行動について分析を行った結果を図 5.4 に示す。その結果、タイピング行動のうち、文字の削除であるキーボードの BackSpace 回数、編集集中の入力文字を確定した回数、およびメッセージを送信した回数いずれも各モードで有意差は認められなかった。これは、情動表現の提示タイミングがメッセージの編集に影響していないことを示している。

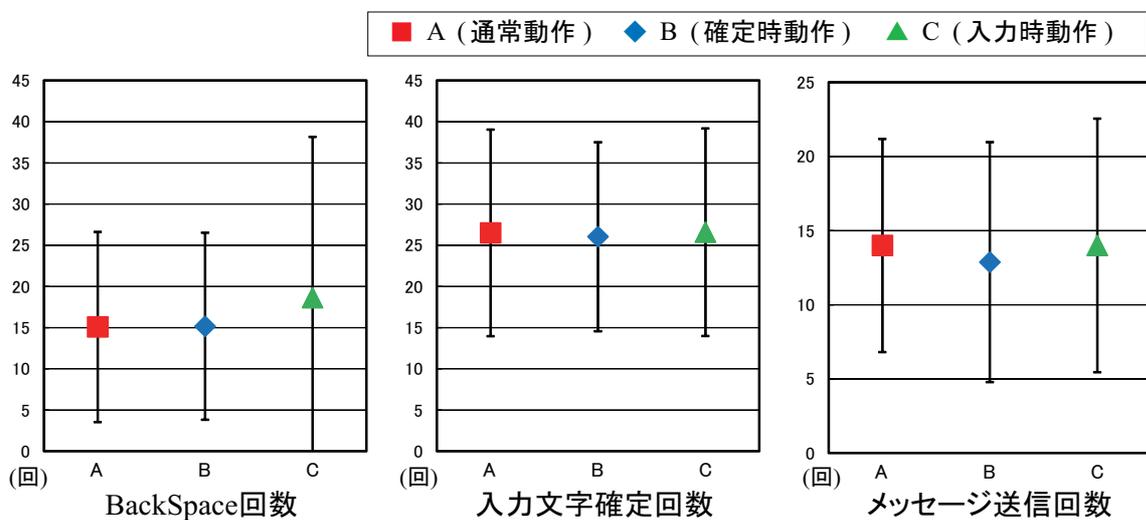


図 5.4 タイピング行動分析結果

次に一対比較の結果を表 5.1 に示す。表中の数字は各行のモードを選択した回数を表している。

表 5.1 一対比較の結果

	A	B	C	合計
A		11	7	18
B	13		10	23
C	17	14		31

この結果に対して被験者による評価を定量的に評価するために、Bradley-Terry モデル

第5章 情動表現の即時表示手法

を想定した(式3.1, 3.2). 結果から最尤推定した π の値の比較を図5.6に示す. Cの入力時動作が最も高く評価され, 次いでBの確定時動作, Aの通常動作の順に評価された.

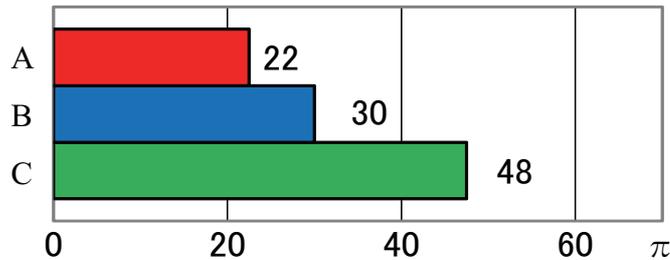


図 5.6 Bradley-Terry モデルによる好ましさ π

7段階評価の平均値および標準偏差の結果と Friedman 検定を行った結果を図5.7に示す. Cの入力時動作は全ての項目で高く評価され, 「一体感」, 「思いが通う」の項目では, Aの通常動作に比べて有意水準5%で有意差が認められた. また, Bの確定時動作と比べた場合も「使用したい」, 「一体感」の項目において有意水準5%で有意差が認められた.

また, 自由記述式のアンケートで得られた意見を表5.2に示す.

表 5.2 システムに対する意見

- ・相手との一体感があって楽しかった
- ・動きがあると, 感情が伝わりやすいと思った
- ・思っていた動作と違うと混乱してしまう
- ・一部のアクションが若干変に感じる
- ・Cだと入力を間違えた時でも動いてしまう
- ・一文打ちきってから変換するので, CやAの方がよかった
- ・Enterを押してから動作した方(A)が使いやすいと感じた
- ・Cだと自分の動作が確認しづらかった

ここで, 図5.8に誤認識による情動表現が発生した組数を調査した結果を示す. 実験者が前後の文脈から不適切な情動表現を抽出したもので, 略式な挨拶表現として追加

5.3 チャットコミュニケーション実験

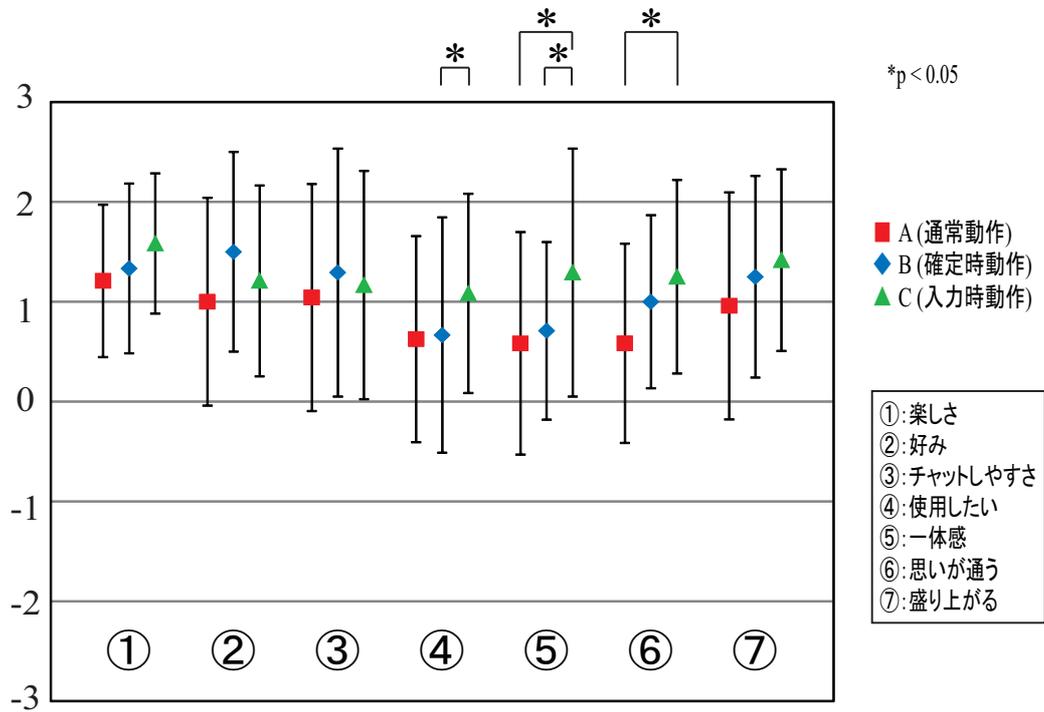


図 5.7 7段階評価

されていた「うい」という動作が「こういう」などの入力に対して反応し、誤認識として多く見られた。

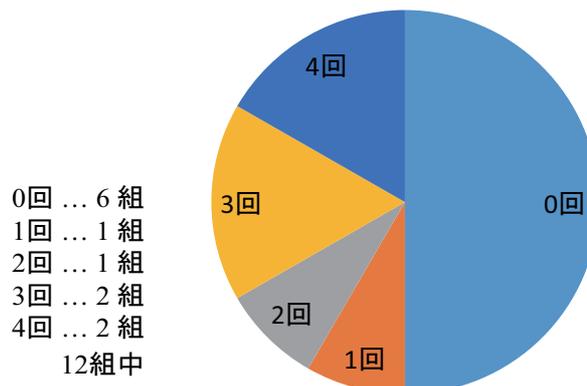


図 5.8 誤認識による情動表現が発生した組数

5.4 考察

実験の結果、一対比較においてCの入力時動作は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、「使用したい」、「一体感」、「思いが通う」の項目で高く評価された。これは、発話者の入力時の感情とキャラクターの情動表現との間の提示時間の差を小さくすることで、使用者とキャラクターとの一体感が向上し、より豊かな感情表現が可能となった結果だと考えられる。Bの確定時動作は、Cの入力時動作ほど高い評価は得られなかったものの、7段階評価の「好み」、「チャットしやすさ」の項目において高く評価された。これは、文章を確定するまで動作を行わせないことで、打ち間違いなどにより意図しない動作が提示されることを抑えられる編集可能性が評価されたためと考えられる。これらの結果により、情動表現を即時表示する本システムはチャットコミュニケーションの支援に有効であることがわかる。また、自由記述アンケートから得られた意見において、Cの入力時動作に対して「入力を間違えた時でも動いてしまう」「自分の動作が確認しづらかった」など、否定的な意見が得られた。しかしながら、一対比較および7段階評価においてCの入力時動作が高く評価されたことから、情動表現の提示を行うキャラクターチャットシステムにおいては、メッセージの表示と情動表現の提示時間の差を小さくすることが好ましいと考えられる。一方で、会話のコンテキストにおける使用者の意図と動作の一致率とシステム評価の関係についてはさらに詳細に検討する必要がある。

5.5 結言

本章では、テキストの表示とテキストに対応したキャラクター動作の提示時間の差を縮め、情動表現を即時提示する手法を開発した。開発した手法を用いてチャットコミュニケーション実験を行い、提示タイミングの評価を行った。その結果、入力の度に文章の認識を行い動作を提示する手法が高く評価され、テキスト表示とキャラクター動作の同期表示効果を確認した。

第6章

テキストおよび情動表現同期表示手法

6.1 緒言

本研究におけるキャラクタチャットシステムの開発について、テキスト情報あるいはキャラクタの身体動作の提示手法評価に対して、4章では対話相手に表示されるテキスト情報に対する正確性、編集可能性を考慮し、キャラクタ動作を相手の入力アウェアネスを含めた身体的なインタラクション把握のために用いた提示手法によって評価した。また5章では、情動表現の提示タイミングの効果を評価するために、相手に表示されるテキスト情報の編集可能性を考慮せず、入力中のメッセージを全て表示する提示手法で評価した。これらは異なる観点で開発提案された手法であり、別々の評価実験によって比較検討されたものである。

そこで本章では、InterChatの統合的な評価を目的としてこの両者の比較検討を行うとともに、両者の特性を活かし、テキストの入力状態を中点表示してキャラクタの身体動作を自動生成する手法を新たに開発する。中点表示によって編集集中の文字を秘匿することでテキストの編集可能性を維持しつつ、メッセージの最終的な確定送信ではなく各単語の変換確定時に文字認識を行い、テキストとキャラクタの情動表現を同期して表示する手法を開発する。新たに開発する手法を用いたシステムの評価実験を行うことで、より使用者に好まれるテキストおよび情動表現の提示手法を検討する。

6.2 テキストおよび情動表現同期表示の検討

6.2.1 テキスト入力確定時情動表現機能

前章までに開発したテキストおよび身体動作、情動表現手法および本章で開発したテキストおよび情動表現同期表示手法におけるメッセージ表示過程を図6.1に示す。

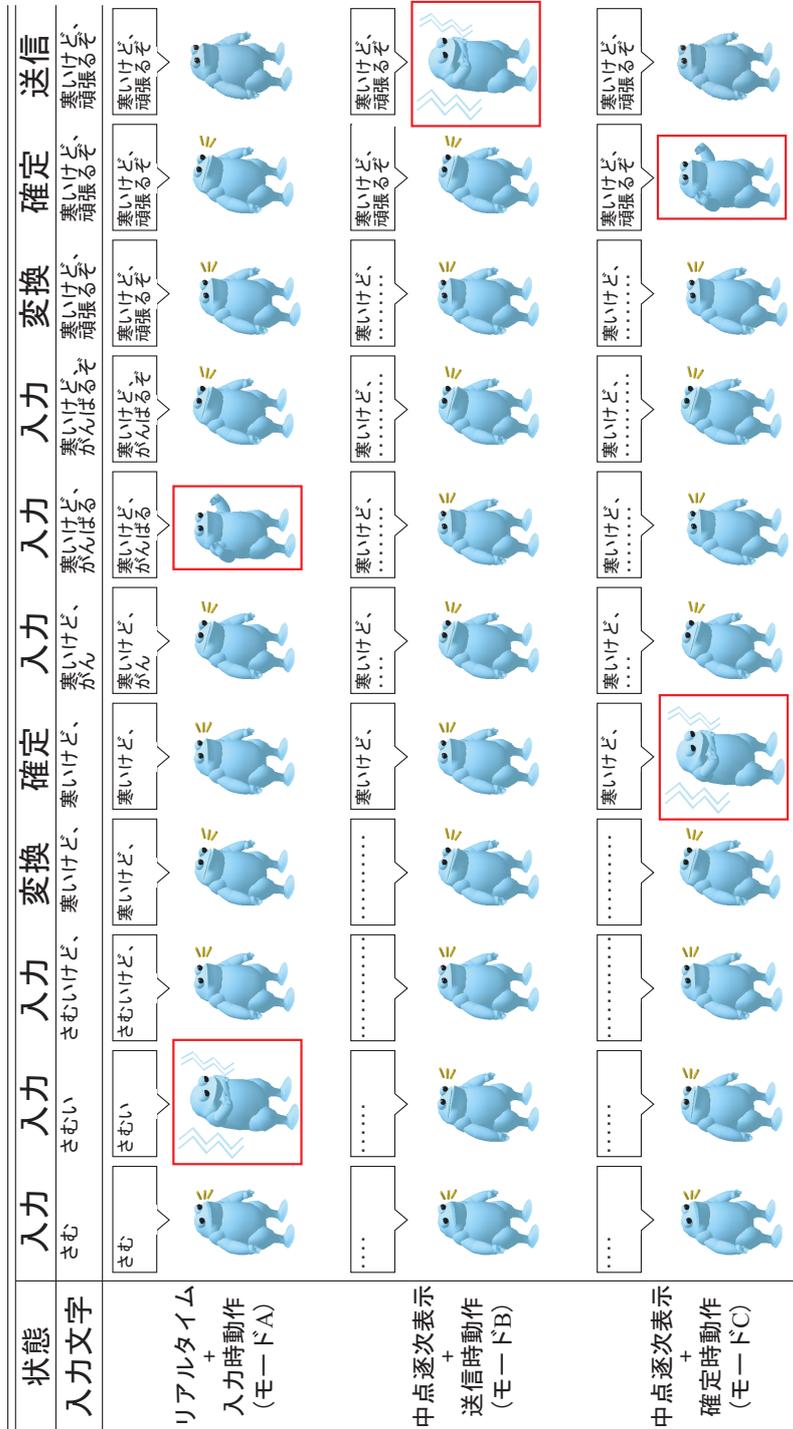


図 6.1 メッセージ表示過程

6.2 テキストおよび情動表現同期表示の検討

リアルタイム表示手法（モードA）は，作成中のメッセージと情動表現を時間的に最も早く対話者に提示できかつ，入力されたメッセージと情動表現の表示時間の遅延が最も少ないモードである．中点逐次送信時手法（モードB）は，入力中はテキストを中点表示（ \cdot ）し，入力が確定した時点で対話相手はメッセージとして確認できるが，情動表現はメッセージ送信時に行われる手法である．本章で開発した中点逐次確定時手法（モードC）は，テキスト表示はモードBと同様に中点表示であるが，情動表現も確定時に同時に行われる手法である．入力動作と情動表現との表示時間の遅延はモードAよりも大きくなるが，モードAと同様に対話相手がテキスト確認時点で情動表現も同期して確認できる．

また，本章におけるシステム構築の際，前章において誤認識として観察された略式な挨拶表現である「うい」という動作を削除するなど，システム内の情動表現を一部見直している．

6.2.2 実験方法

本章では，好ましいテキストおよび情動表現の提示タイミングを検証するために，被験者2人1組で別々の部屋に分かれ，図6.1の「A: リアルタイム表示手法」，「B: 中点逐次送信時手法」，「C: 中点逐次確定手法」の3つのモードを使用してチャットコミュニケーションを行った．被験者には最初に，システム全体と各モードの特徴を説明した．その後別々の部屋に分かれ，システムに慣れるために各モードを3分ずつ使用してから，2つの評価法を用いて比較させた．最後に実験に関する感想を自由記述形式で記入させた．被験者は18～24歳の男女学生12組24人（同性同士）で，普段からパソコンを使用しておりタイピングにある程度慣れた学生である．

評価法1: 一対比較

A～Cのモードからランダムに2つを一対比較させ，これを計3 ($=_3C_2$) 回繰り返した．まず，1つのモードでチャットをした後に別のモードによるチャットを行い，2モードを使用した後に「総合的によい」という観点からどちらがよかったかアンケート用紙に記入させた．対話者同士は同じモードを使用してチャットし，1つの比較につき対話時間を4分（1つのモードにつき2分）とした．

評価法 2: 7段階評価

A～Cのモードを「楽しさ」、「好み」、「チャットしやすさ」、「使用したい」、「一体感」、「思いが通う」、「盛り上がる」の7項目について評価させた。各モードをランダムな順番で使用し、その後、各モードがどの程度その項目に当てはまるかを7段階（中立0）でアンケート用紙に記入させた。対話者同士は同じモードを使用してチャットし、1つのモードにつき対話時間を3分とした。

6.2.3 実験結果

まず、図6.2に誤認識による情動表現が発生した組数を調査した結果を示す。実験者が前後の文脈から不適切な情動表現を抽出した結果、システム改良により前章での誤認識数から減少したことが分かる。

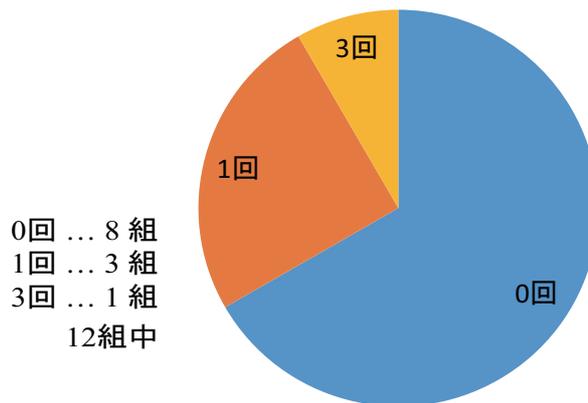


図 6.2 誤認識による情動表現が発生した組数

次に一対比較の結果を表6.1に示す。表中の数字は各行のモードを選択した回数を表している。この結果に対して被験者による評価を定量的に評価するために、Bradley-Terryモデルを想定した(式3.1, 3.2)。結果から最尤推定した π の値の比較を図6.3に示す。Aのリアルタイム表示手法が最も高く評価され、Bの midpoint 逐次送信時手法、Cの midpoint 逐次確定手法が同程度と評価された。

7段階評価の平均値および標準偏差の結果とFriedman検定を行った結果を図6.4に示す。全モードの全項目で肯定的な評価が得られているが、一対比較同様、Aのリアルタイム表示がとくに高く評価されている。AモードはBの midpoint 逐次送信時手法に比べ

6.2 テキストおよび情動表現同期表示の検討

表 6.1 一対比較の結果

	A	B	C	合計
A		15	16	31
B	9		12	21
C	8	12		20

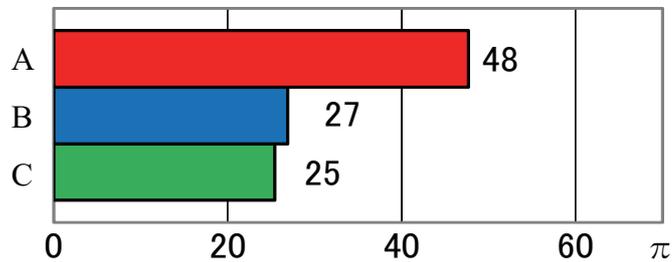


図 6.3 Bradley-Terry モデルによる好ましき π

て、「楽しさ」の項目で有意水準1%で、「思いが通う」の項目で有意水準5%で有意差が認められた。また、Cの midpoint 逐次確定手法に比べて「楽しさ」の項目で有意水準5%で有意差が認められた。

また、タイピング入力行動分析として、7段階評価時の各モード3分間のメッセージ送信回数（メッセージを送信した回数）、入力文字確定回数（メッセージの作成中に変換を確定した回数）、BackSpace回数、の平均値と標準偏差、一要因の分散分析による結果を図6.5に示す。

メッセージ送信回数はAモードが最も多く、AはB、Cに比べてそれぞれ有意水準5%で有意差が認められた。入力文字確定回数はAモードが最も多く、AはCに比べてそれぞれ有意水準1%で有意差が認められた。BackSpace回数はCモードが最も多く、CはAに比べて有意水準5%で有意差が認められた。

第6章 テキストおよび情動表現同期表示手法

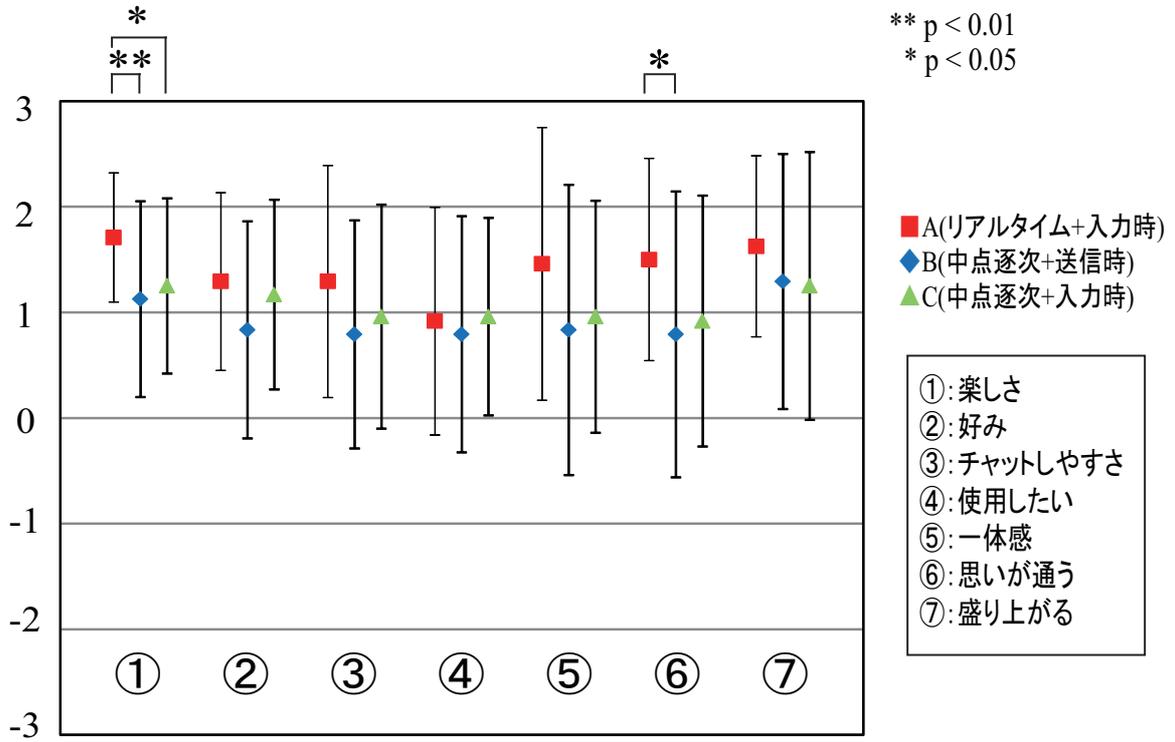


図 6.4 7段階評価結果

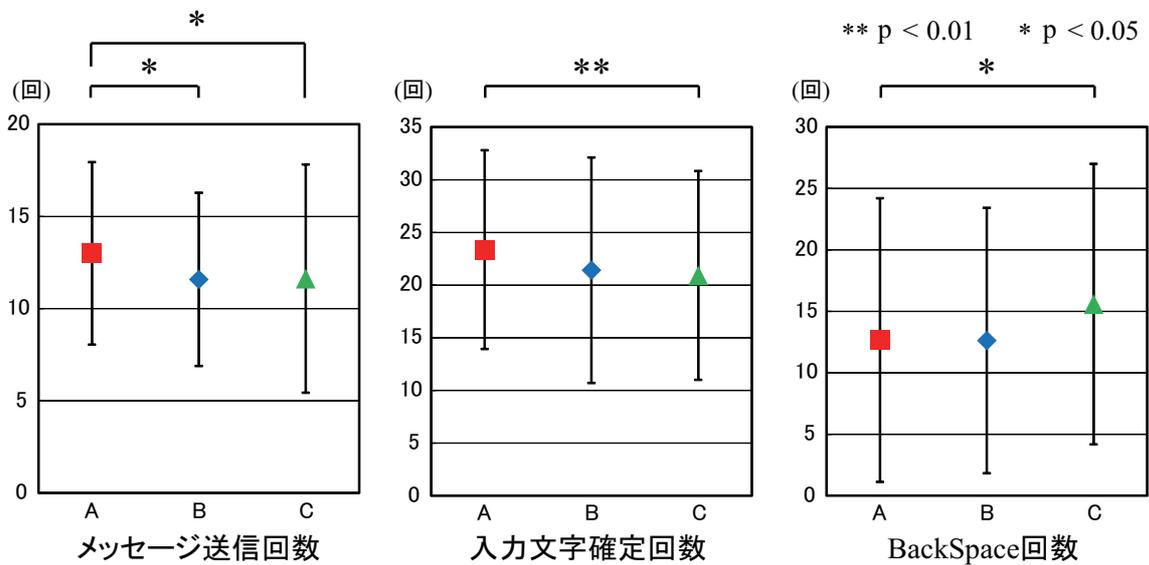


図 6.5 タイピング行動分析結果

6.3 考察

本章で開発した「C: 中点逐次確定手法」は、テキスト入力時に送信までの編集可能性を持ちつつできるだけ早く相手に表示し、情動表現を同期表示できることが特徴である。前章までに開発されてきた「A: リアルタイム表示手法」、「B: 中点逐次送信時手法」と比較評価した結果、一対比較および7段階による官能評価において、メッセージと情動表現を時間的に最も早く提示でき、メッセージと情動表現の表示時間の遅延が最も少ないAモードが評価される結果となった。ここでタイピング入力の行動分析結果から、メッセージ送信回数、および入力文字確定回数がAモードで最も多くなっており、より短めで多くのメッセージ交換が行われている。しかしBackSpace回数はCモードが最も多く、テキストの編集可能性によってメッセージの送信前にテキストを修正する行動が観察されるなど、システムが効果的に使用されていることが分かる。

これらの理由として、本実験での被験者は既知の友人同士2人1組であり、テキストの編集ができないリアルタイム表示手法において、メッセージの作成ミスがあったとしても単純なミスとして扱われ、それまでに形成された人間関係に大きく影響を与えないという信頼感が存在する可能性は高い。初対面の相手や業務利用など、テキストの編集可能性が重視される場合は異なる評価となる可能性も高く、さらに検討が必要と考えられる。

6.4 結言

本章では、InterChatの統合的なシステム評価として入力状態をテキストで表示する際に最も好まれた中点表示(·)を使用して、テキストの編集可能性を維持しつつ、テキストとキャラクタの情動表現を同期して表示する手法を開発した。システムを用いたコミュニケーション実験によって、タイピング入力とテキストあるいは情動表現を含む身体動作の提示タイミングについて分析評価を行った結果、親密な相手とのチャットコミュニケーションでは、より早くテキストと身体動作を表示できるインタラクションシステムが好まれる結果が確認された。

第7章

結論

7.1 本研究のまとめ

本研究は、テキストを用いたコミュニケーション手段において、タイピングの入力がリズム的であることに着目し、タイピングを音声入力に見立て iRT (InterRobot Technology : 発話音声から豊かなコミュニケーション動作を自動生成するインタロボット技術) を音声入力からタイピング入力へと拡張することで、タイピングからキャラクターの身体動作を自動生成するとともに、文字列に対応する意味のある動作を自動生成するタイピング駆動型身体引き込みキャラクターチャットシステム InterChat のコンセプトを提案し、システムを開発した。さらにタイピング入力からの身体動作の自動生成、および文字列に対応した意味のある動作の自動生成について、身体的インタラクションの観点から評価実験を行い、システム改良を行うとともに、チャットコミュニケーションにおける InterChat の有効性を示した。

以下、本論文における各章の成果をまとめる。

第1章では、本論文の背景と目的および本論文の位置付けを示し、あわせて論文の概要について述べた。

第2章では、人間の発話音声とキーボード入力におけるタイピング入力リズムの類似性を示し、タイピング入力に基づく身体引き込みキャラクターチャットシステムのコンセプトを提案した。発話音声からキャラクターの豊かなコミュニケーション動作を自動生成する iRT モデルを応用し、タイピング入力から聞き手動作、話し手動作、および入力された文字列に対応した動作それぞれを生成するモデルを定義し、プロトタイプシステムを開発した。

第3章では、開発した InterChat のシステム評価を行った。まず身体的インタラクションの観点から、InterChat から文字列に対応した動作モデル機能を省略し、話し手動作モデル、聞き手動作モデルのみを導入したシステムを用いたコミュニケーション

第7章 結論

実験により、身体的引き込み動作提示の有効性を示した。また文字列に対応した動作モデルを導入したシステムを含めて従来のテキストチャットと比較する評価実験を行い、聞き手動作、話し手動作、および文字列に対応した動作モデルを用いたシステムが高く評価され、開発した InterChat の有効性を示した。

第4章では、InterChat におけるテキストの表示方法に着目し、対話者の入力状態を、キャラクターの身体動作だけでなく同期的な文字表示によって、対話相手が把握できる効果を分析評価した。入力中の情報全てを表示するリアルタイム表示、入力中の文字列を中点（・）に変換して表示する手法を開発し、メッセージの送信時表示との比較検討を行った。その結果、入力中の状態は相手の入力メッセージが中点に変換されて表示され、さらにメッセージ確定時に確定部分のみが中点からテキスト表示に切り替わる中点逐次確定表示手法が高く評価された。

第5章では、それまでテキストに対応した情動表現がメッセージの送信後に提示されることによる問題点を克服するため、新たに「確定時動作」、「入力時動作」の2つの提示手法を開発した。確定時動作では、入力中のメッセージが確定される度に認識を行い、特定の文字列が文中にあった場合はその文字列に対応した動作を提示し、入力時動作では、メッセージが入力される度に認識を行い、特定の文字列が文中にあった場合はその文字列に対応した動作を提示する。この新たな2つの提示手法と従来の手法との3つの手法を用いた評価実験において、入力の度に文章の認識を行い、動作を提示する入力時動作手法が高く評価された。

第6章では、InterChat の統合的なシステム評価として4章でのテキスト表示および5章の情動表現表示それぞれの着眼点から提案された手法のうち、最も高く評価された2つの手法、さらに入力状態をテキストで表示する際に最も好まれた中点表示（・）を使用して、テキストの編集可能性を維持しつつ、テキストとキャラクターの情動表現を同期して表示する手法を開発し、3つの手法を用いて評価実験を行った。その結果、5章で提案された入力の度に文章の認識を行い、動作を提示するリアルタイム表示手法が高く評価され、親密な相手とのコミュニケーションのような、短い文章で送受信が繰り返される場合、テキストおよび身体動作表現をより早く提示できる手法が好まれる結果が得られた。

7.2 今後の展望

リアルタイムでのテキストを用いた情報共有は、今後情報コミュニケーション環境が大いに向上し、リッチなメディア利用が可能になったとしても、手軽さや簡便さなどを背景に、遠隔コミュニケーションに対して一定の役割を担い続けることが予想される。しかし一方で、身体性の欠落や断片的な情報伝達などによって思いが伝わりにくいメディアであることも認識しておく必要がある。

今回の開発システムを利用することで、タイピング入力による身体リズム共有によって、よりインタラクティブなコミュニケーション環境を実現できる。本論文ではより一体感を感じられる、あるいは心地よいコミュニケーション環境の実現のため、既知の友人同士の親密な関係におけるシステム評価を行ったが、テキストメディアの持つ編集可能性に着目した場合、より正確な情報伝達を行うべき相手とのコミュニケーションにおいては異なった結果となる可能性も十分にある。本論文の評価においても、誤った情報の伝達を好まないといった意見や編集可能性を評価する結果が見受けられた。これはチャットが本来持つ編集可能性といった特性が好まれる状況が存在することを示しており、企業活動においても電話によるオンラインサポート業務がオンラインチャットに移行されるケースや、就職活動の面接にチャットを導入するといったニュースも聞かれるなど、状況に応じたシステムを開発していくことでさらなる応用が期待できる。

また、本論文で検討したキャラクタの身体動作とテキスト表示の同期性やテキスト表示と情動表現の同期性は受信者が情報を受信する際の問題点であり、このことは送信者が情報を生成した時間と受信者が情報を受信した時間が時間的に切り離されている、オフラインや非同期でのコミュニケーションにおいて応用できることを示している。今後はオンラインアンケートを回答してもらうといった場合の非同期分野での活用、また、近年進歩の著しい音声認識による音声のテキスト化といった分野においても、本論文での知見を応用し発展させ、よりインタラクティブなコミュニケーション環境の実現が期待される。

謝辞

本研究は、著者が岡山県立大学大学院情報系工学研究科に社会人学生として在籍していた2011年から6年間にわたり、岡山県立大学情報工学部教授、渡辺富夫博士のご指導のもと進めてきたものです。本研究を行う上で、渡辺富夫博士には、多くの時間を費やして貴重な御意見および終始熱意に溢れた御指導、御鞭撻を賜りました。改めてここに深く感謝の意を表します。

また、岡山県立大学情報工学部教授、忻欣博士並びに佐藤洋一郎博士、そして岡山県立大学情報工学部准教授、山崎大河博士には、本論文に関して数多くの有益な御助言を頂きました。ここに深く感謝の意を表します。

関西学院大学理工学部人間システム工学科教授、山本倫也博士には、論文作成に関する有益な御助言を頂きました。ここに深く感謝の意を表します。

岡山県立大学情報工学部准教授、石井裕博士には、研究から論文作成に至るまで、細部にわたる有益な御助言、ご支援を頂きました。ここに深く感謝の意を表します。

岡山県立大学大学院情報系工学研究科を御修了された岩佐厚郎さんには、本研究の一部を御協力頂きました。深く感謝の意を表します。

また、渡辺研究室の皆さんにはシステム開発や実験の補助などにおいて数多くの御協力を頂きました。心より感謝の意を表します。

渡辺研究室秘書として様々な雑務に快く対応していただきました岡正子秘書に心より感謝の意を表します。

最後に、著者を温かく見守り、研究に没頭する生活を可能にしてくれた家族に感謝します。

本論文に関する研究業績

原著論文

- [1] 服部 憲治, 渡辺 富夫, 山本 倫也: タイピング駆動型身体引き込みキャラクタチャットシステム InterChat, ヒューマンインタフェース学会論文誌, Vol. 15, No. 4, pp. 389–398. (2013).
- [2] 服部 憲治, 渡辺 富夫, 石井 裕: タイピング駆動型身体引き込みキャラクタチャットシステムにおけるテキストの実時間入力状態表示手法, ヒューマンインタフェース学会論文誌, 条件付採録 (2017).

口頭発表

- [1] 服部 憲治, 岩佐 厚郎, 渡辺 富夫, 石井 裕: タイピング駆動型身体引き込みキャラクタチャットシステムにおけるテキストおよび情動表現同期表示手法の開発, 電子情報通信学会ヒューマンコミュニケーション基礎研究会, Vol. 116, No. 436, pp. 89–90, 2017–01.
- [2] 藤井 亮, 四方 拓, 服部 憲治, 渡辺 富夫, 石井 裕: 音声駆動型身体的引き込みキャラクタシステムにおける音声認識による動作・情動表現提示の評価, 第14回計測自動制御学会システムインテグレーション部門講演会 (SI2013) 論文集, pp. 1945–1948, 2013–12.
- [3] 岩佐 厚郎, 服部 憲治, 渡辺 富夫, 石井 裕: タイピング駆動型身体的引き込みキャラクタチャットシステムにおける情動表現提示タイミングの評価, HAI シンポジウム 2016, G-1, pp. 1–6, 2016–12.