

修士論文

ユーザとの信頼関係構築のためのロボット対話における ペーシング評価

中村 卓矢

2019 年 2 月 1 日

奈良先端科学技術大学院大学
情報科学研究科 情報科学専攻

本論文は奈良先端科学技術大学院大学情報科学研究科に
修士（工学）授与の要件として提出した修士論文である。

中村 卓矢

審査委員：

萩田 紀博 教授 （主指導教員）

加藤 博一 教授 （副指導教員）

神原 誠之 准教授 （副指導教員）

ユーザとの信頼関係構築のためのロボット対話における ペーシング評価*

中村 卓矢

内容梗概

本研究では，対話ロボットとのインタラクションを通してユーザとの信頼関係構築を目的として，ペーシングがユーザに与える影響を主観的・客観的に評価する．ペーシングとは，対人における信頼関係構築で用いられる，相手の話し方や状態，呼吸などのペースに合わせる技術である．これにより，相手の警戒心を取り安心感を与えるだけでなく，相手の自分への肯定感や重要感を満たし，信頼関係を構築できると考えられている．そこで，このペーシング技術をロボットに適用し，ペーシングシステムとして実装した．実験では，ペーシングを適用したロボットと対話をした際の印象評価を，主観的・客観的指標を用いて行う．その結果，主観的評価においてペーシングを適用したロボットとの対話の方が適用しなかったロボットとの対話と比較して，「親しみやすいか」「対話は楽しいか」「繋がりを感じたか」「自分に同調していたか」「継続的に利用したいと思ったか」「話を聞いている感じがするか」の各項目に有意な差が見られた．また，客観的評価においてはペーシングを適用したロボットの方が適用しなかったロボットと比較して，LF/HF の被験者間平均値に有意な差が見られた．これらの結果から，ロボットにペーシングを適用したインタラクションはユーザとロボットの信頼関係を構築する上で有効であることが示された．

キーワード

ヒューマンロボットインタラクション，信頼関係構築，音声合成，ペーシング

*奈良先端科学技術大学院大学 情報科学研究科 情報科学専攻 修士論文，2019 年 2 月 1 日．

Effect of Pacing in Building a Trust Relationship between an User and a Robot*

Takuya Nakamura

Abstract

This study aims to clarify an effect of pacing in building a trust relationship between a human and a robot. Pacing is a technique that matches the pace among person in terms of speaking, state, breating, etc. It is shown that it may not only relax a human's guard and give ease to the human, but also create affirmative responses and actions so that it may build the trust relationship among humans. This study implement the pacing system that applied the pacing technique to a robot. Subjective evaluations over 8 subjects are made when the robot with or without pacing function interacts with the subject. The experient results for the comparative studies of pacing show that factors with the friendly and enjoyable interaction, the strong connection, the synchronization, the continuous use of interaction, and the active hearing are statistically dominant in effect of pacing. Objective experiments with LF/HF measurements are also made for 8 subjects when the robot with or without pacing function interacts with a subject. From the result, the mean value in LF/HF among subjects for the robot with pacing is higher than the robot without pacing. These results shows that pacing techniques in human-robot interaction is effective to build a trust relationship between an user and a robot.

Keywords:

Human-Robot Interection, Building Trust Relationship, Speech Synthesis, Pac-

*Master's Thesis, Department of Information Science, Graduate School of Information Science, Nara Institute of Science and Technology, February 1, 2019.

ing

目次

図目次	v
第 1 章 はじめに	1
第 2 章 信頼関係構築に関する関連研究	3
2.1 人と人における信頼関係構築	3
2.2 人とロボットにおける信頼関係構築	4
2.3 ペーシングを用いた信頼関係構築	5
2.4 本研究の位置づけ	7
第 3 章 ロボットへのペーシングの実装	9
3.1 ペーシングシステム	9
3.2 Julius, SPTK による韻律情報推定	10
3.3 韻律情報から音声合成パラメータへの変換	11
3.3.1 発話速度の変換式の導出	13
3.3.2 ピッチの変換式の導出	14
3.3.3 音量の変換式の導出	15
3.4 VoiceText による音声合成	17
3.5 ペーシングシステムの精度評価	18
3.5.1 精度の主観的評価	18
3.5.2 精度の客観的評価	20
第 4 章 ペーシングを用いた対話時のロボットに対する印象評価実験	22
4.1 対話シナリオの評価	22
4.1.1 実験手法	22
4.1.2 実験結果	24
4.2 印象評価実験	25
4.2.1 実験手法	26
アンケートによる主観的評価	28

	LF/HF による客観的評価	30
4.2.2	実験結果	31
	アンケートによる主観的評価結果	32
	LF/HF による客観的評価結果	34
4.3	考察	36
第 5 章	終わりに	39
	謝辞	48
	参考文献	49
	発表リスト	52

図目次

2.1	ペーシングを適用したロボットとの対話の様子	7
3.1	ペーシングのシステムフロー	10
3.2	音声合成パラメータへの変換式導出作業の様子	13
3.3	発話速度に関する推定された値と音声合成パラメータの関係	14
3.4	ピッチに関する推定された値と音声合成パラメータの関係	15
3.5	音量に関する推定された値と音声合成パラメータの関係	16
3.6	ピッチの音声合成パラメータとその変化に伴い音量が変化する倍 率の関係	17
3.7	ペーシングシステム精度の主観的評価グラフ	20
4.1	対話シナリオの一例	23
4.2	アフェクト・グリッド	24
4.3	対話シナリオの評価結果	25
4.4	ロボットとの対話の様子	26
4.5	会話文に対する読み方の指示例	27
4.6	ペーシング非適用時のパラメータの決定手法	28
4.7	心電図の波形	31
4.8	心拍センサ（左），解析用ソフトウェア（右）	31
4.9	アンケート評価結果：質問 1-6	33
4.10	アンケート評価結果：質問 7「どちらのロボット（ペーシング適用 時・非適用時）が好きか」	34
4.11	LF/HF による評価結果	35
4.12	ラッセルの円環モデル	37

第 1 章 はじめに

ライフスタイルの変化に伴い、年齢問わず独居世帯の数が増加している [1]。それによって、日常生活における対話が減少している。日常対話が不足すると、脳機能の低下や、生活環境の悪化によりうつ病や認知症の原因になると考えられている。そのため、このような日常対話が不足している人の発話機会を増やすための対策が重要な課題となっている。

このような課題の対策として、日常的に対話ロボットを利用してコミュニケーションを活性化させる取り組みは従来から数多く研究されてきた。神田ら [2] は、ロボットが社会的な存在として人々に受け入れられるためには人とロボットが友好的な関係を築く必要があると述べている。このように、日常的に対話ロボットを利用してもらうためには人とロボットの友好的な関係が重要であるとする。

そのため我々は、ロボットとの対話をユーザにとって好ましいものにするために、心理学や介護福祉分野で、人同士のインタラクションにおいて円滑な関係を築くために用いられている同調行動に注目した。同調行動には、相手の無意識的な行動パターンを鏡に映したように真似する技術である「ミラーリング」、相手の発言内容のキーワードを繰り返し用いる技術である「バックトラッキング」、相手の話し方や状態、呼吸などのペースに合わせる技術である「ペーシング」などがある [3][4]。その中でも今回は、「ペーシング」に着目した。これにより、相手の警戒心を解き安心感を与えることができるだけでなく、相手の自分への肯定感や重要感を満たし、信頼関係を構築することができると考えられている。西村ら [5] は、対話の中で 2 話者間の基本周波数が同調している対話音声で第三者が聞いた時、親しさがあったり、盛り上がっているといった印象を受けたと述べている。しかし、実際にペーシングの技術をロボットによって実現し、その効果を検証した研究は少ない。

この課題に対し Nishimura ら [6] は、ペーシングをロボットに適用することによって、ロボットとユーザの対話時における信頼関係構築を目指した。しかし、実際にロボットと対話をしたときに、ペーシングを適用した対話とペーシングを適用しなかった対話で被験者の評価にほとんど差が見られなかった。この原因として、「ペーシングシステムによって作られた音声とそうでない音声の差が分かりづらかった」ということが挙げられている。また、ペーシングの評価が被験者の主観

のみで行われている。そもそも、臨床心理学の分野においてさえ、ペーシングを客観的に評価したものは少ない。

そこで、本研究ではペーシングを適用したロボットとの対話を通して、ペーシングそのものを主観的・客観的に評価することができる枠組みの構築を行う。そして、先行研究をもとにペーシング適用時と非適用時の差が被験者の主観レベルで認識できるような実験系を構築した。

以下、本論文は2章で対話を通じたユーザとロボットの信頼関係構築を目的として、人と人における信頼関係構築、人とロボットにおける信頼関係構築についてそれぞれ詳述し、本研究の位置づけを述べる。3章では、今回用いるペーシングシステムの実装とその評価について述べる。4章では、実装したペーシングシステムを適用したロボットがユーザと対話した時の印象評価を行う。そして、5章で本研究のまとめと今後の展望について述べる。

第 2 章 信頼関係構築に関する関連研究

本章では、本研究に関する従来研究について紹介する。まず、人と人の信頼関係構築に関する心理学的な関連研究を紹介する。次にヒューマンロボットインタラクション分野におけるロボットと人の信頼関係構築に関する研究を紹介する。そして、最後にそれらを踏まえた本研究の位置づけを示す。

2.1 人と人における信頼関係構築

対人社会心理学の分野において、同調傾向 (congruence, convergence, symmetry) は多くの研究者によって、円滑なコミュニケーションの指標としてみなされてきた。同調傾向とは、相互作用相手との間でコミュニケーション行動が連動し、パターンが類似化していくことを指す。さらには、同調傾向はラポールをもたらしたりポジティブな対人印象をもたらしたりすることが示されてきた。このことは、同調傾向がコミュニケーションにおいて何らかの重要な役割を果たしていることを示唆している [7]。

認知心理学においては、同調行動 (synchronous actions) は反応の 4 つの次元に影響を及ぼすとされている。4 つの次元には、

- 社会性のある行動 (prosocial behavior)
- 認知されている社会的結合 (perceived social bonding)
- 社会的認識 (social cognition)
- ポジティブな感情 (positive affect)

がある。また、同調 (synchrony) が人に影響するメカニズムとして、

- 同調が自己と他者の境界をぼやかせる
- グループとしての共通認識から社会的結合は生まれる
- 同調が人々の情緒的感性にも影響を与える

の 3 つが挙げられている [8]。

精神療法においては、患者とセラピストの同盟 (alliance) 関係が重要であると

言われている．同盟は患者とセラピストの脳のカップリングに基づいている．そして，運動の同調（synchrony）が脳のカップリングを促進する．Interpersonal Synchrony model によると時間経過における同調のレベルは，

- レベル 1：運動の同調（数百ミリ秒－10 秒）
- レベル 2：社会的認知の促進により同盟を構成（10 秒－1 時間）
- レベル 3：同盟による治癒効果が患者の自己規制能力の向上に繋がる（数週間－数年）

のように分類することができる [9]．

このように，使用されている用語に多少の違いはあるものの，様々な心理学分野において同調行動は人と人の信頼関係構築に有効であるということが示されている．

2.2 人とロボットにおける信頼関係構築

ヒューマンロボットインタラクション分野において，人とロボットのインタラクションによる信頼関係構築に関する研究を紹介する．

山野ら [10] は人とロボットのインタラクションにおいて，人の情動に同調させた際に同調しない場合よりもインタラクション時間が長くなり，ポジティブな印象となることを示した．さらに人の心理状態が快適な方向に変化したことを示した．

小林ら [11] は，高齢者の”話し相手”になる対話インターフェイスの開発に取り組んだ．対話インターフェイスにはユーザに様々な話題を提供する「話題提示モード」とユーザの話の聴き手となる「傾聴モード」があり，ここでは「傾聴モード」における発話文生成に関する検討を行っている．話題辞書を生成し，話題辞書中の単語に返答することで，音声認識成功時にユーザの発話に対して自然な発話が 67% でき，日本語として適切な質問文が 83% できたことが確認されている．この傾聴モードにおける対話戦略の 1 つとして反復相槌（バックトラッキング）が用いられているが，その影響についての評価は行われていない．

Huang ら [12] は RapportAgent と呼ばれるエージェントに対し，信頼関係を構築するための 3 要素（明るさ，自然な注視，強調）に基づいた対話エージェントを

開発した。エージェントと人間の2者間の対話において、人間の発話の有無、視線、頷き、笑顔等をリアルタイムに分析して自らの反応をパターン化できる。この結果、相槌、話者交代等の動作にいくつかの動作パターンが設定され、信頼関係の構築に貢献していることが示された。このロボットはミラーリングによる同調的なインタラクションが実装されている。

以上から、ヒューマンロボットインタラクション分野においても、心理学的分野で対人に使用される技術を用いることで信頼関係構築に有効であることが分かる。

2.3 ペーシングを用いた信頼関係構築

西村ら [5] は、国立国語研究所から提供されている「日本語話し言葉コーパス」(Corpus of Spontaneous Japanese; CSJ)[13]を用いて、対話をしている2人の基本周波数(F0)の相関を調べた。そして、CSJコーパスのインタビュー形式の対話音声を実際に4名の被験者に聞いてもらい、その印象を以下の各項目について5段階のアンケートを取った。

- 相手との親しさ
 - 親しみを持っている(5 4 3 2 1) 遠慮している
- 盛り上がり
 - 良い(5 4 3 2 1) 盛り上がってない
- 相手の意見に同意しながら対話
 - 同意しながら対話(5 4 3 2 1) 意見を戦わせながら対話
- 立場の違い
 - 目上(5 4 3 2 1) 目下
- インタビュアーのフランクさ
 - 気を使っている(5 4 3 2 1) くだけている
- インタビュイーのフランクさ
 - 気を使っている(5 4 3 2 1) くだけている
- インタビュアーの表現
 - 敬語ばかり(5 4 3 2 1) 敬語を使っていない
- インタビュイーの表現

- 敬語ばかり（5 4 3 2 1）敬語を使っていない

「立場の違い」については、インタビュアー側から見て相手が目上か目下かというところを評価して値を付けている。また、「フランクさ」は言い方の違いを示すものであり、声の調子などから受ける印象を評価しており、「表現」は文字上でどのようなになっているかを示しており、語彙的に敬語を用いているかどうかを評価している。この評価値と対話音声における対話中の 2 人の F0 相関値の相関を表 2.1 に示す。したがって、F0 が 2 話者間で同調している対話は、親しさがあったり、盛り上がったり、同意して対話が行われている印象を受けるとということが示されている。また、F0 相関値が高いとくだけた対話の印象になると言える。

また、Nishimura ら [6] は、ロボットにペーシングを適用したペーシングシステムを実装し、それを適用したロボットとの対話時の印象を評価している。図 2.1 はその時のロボットとの対話の様子である。ここでは、人の発話から「ピッチ」と「音量」の値を推定し、その値を音声の合成に反映させることでペーシングを実現している。また、ロボットに関してはテレビ視聴型の雑談ロボットを用いており、テレビを見ている際にその番組に関するソーシャルメディア上のコメントを発話してくれる。実験内容としてはロボットにペーシングを適用した対話と適用しなかった対話をそれぞれ 5 分間行い、それぞれの対話後にアンケートに答えてもらうというものである。しかし、アンケート 6 項目のうち 5 項目においてほとんど差が見られな

表 2.1 アンケート評価値と F0 相関値との相関値

アンケート項目	相関値
親しさ	0.348
盛り上がり	0.350
同意・否定	0.279
立場	-0.282
インタビュアーフランクさ	-0.283
インタビュイーフランクさ	-0.266
インタビュアーの表現	-0.340
インタビュイーの表現	-0.182



図 2.1 ペーシングを適用したロボットとの対話の様子

いという結果になった。この理由として、「対話時間が短すぎた」「ペーシングを適用した対話とそうでなかった対話の差が分かりにくかった」ということが挙げられている。また、評価方法はアンケートを用いた主観的評価のみとなっており、ペーシングを行った時の人の印象を客観的に評価したものは未だ見られていない。

2.4 本研究の位置づけ

本研究では、ペーシングによるロボットとユーザが対話を行った時の信頼関係構築を目指す。さらに、ペーシングがユーザに与える影響を主観的・客観的に評価する。今回用いるペーシングシステムにおいては 2.3 節で述べた「ペーシングを適用

した対話とそうでなかった対話の差が分かりにくかった」という問題を解決するために「発話速度」「ピッチ」「音量」の3つのパラメータを用いる。これにより、より人の声に近い音声を実現することができ、ペーシングを非適用時のロボットの音声との違いが分かりやすくなると考えた。また、今回のシステムにおいて対象とする対話の内容は事前に設定したシナリオに従うものとし、対話時間は特に指定しない。さらに、対話はユーザの発話区間を検出し、自動で進んでいく。そして、ペーシングシステムの評価を主観的・客観的に行い、本システムの妥当性を示す。最後に、それを適用したロボットとユーザの対話を行い主観的・客観的指標を用いて印象評価し、ロボットとユーザの信頼関係構築を示すとともに、ペーシングがユーザに与える影響について評価する。

第 3 章 ロボットへのペーシングの実装

本章では，ペーシング技術のロボットへの実装について説明する．まず，ロボットにおけるペーシングのシステム構成，その実現のために必要な式の導出について説明する．そして，完成したペーシングシステムの精度を主観的・客観的に評価する．ここでのペーシングシステムの精度とは人が発話した音声を用いて，ロボットがどれくらい類似した音声を発話できるかを指す．

3.1 ペーシングシステム

本研究で用いたペーシングのシステムフローを図 3.1 に示す．本システムの構成にあたって，大語彙連続音声認識エンジンである Julius[14]，音声信号処理ツールである SPTK[15]，音声合成ソフトウェアである VoiceText[16] の 3 つを用いた．ペーシングを実現するための流れとしては，まずユーザの発話を Julius で認識し，発話速度を推定，SPTK を用いてピッチ，音量を推定する．これら 3 つの値を，今回用いる韻律情報とする．そして，それぞれの値を VoiceText 固有の音声合成パラメータの値に変換し，VoiceText を用いて音声合成を行う．次節で，求めた韻律情報から音声合成パラメータへの変換について詳述する．

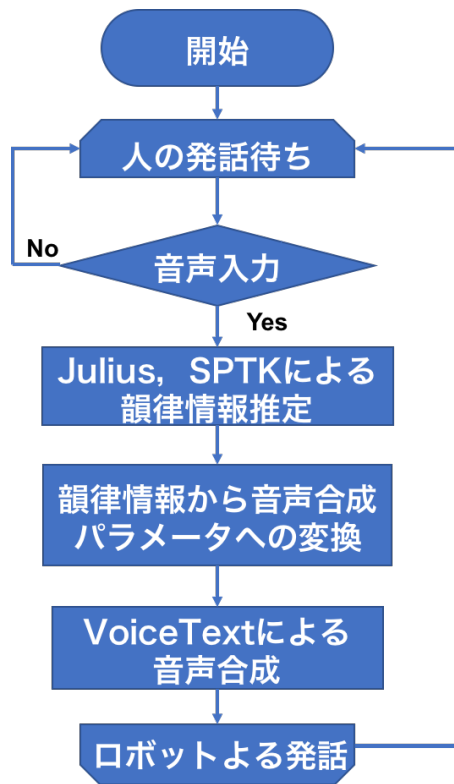


図 3.1 ペーシングのシステムフロー

3.2 Julius, SPTK による韻律情報推定

ここでは, Julius, SPTK を用いて人の韻律情報を推定する手法について説明する. 今回, 人の韻律情報として用いる「発話速度」「ピッチ」「音量」について, それぞれ具体的な推定手法について説明する.

- 発話速度の推定 対話者の発話速度推定は大語彙連続音声認識エンジンである Julius を用いて行う. ここでは, Julius の音声認識処理の前処理で得られる発話区間長でモーラ数を除算して得られる, 1 秒あたりの平均モーラ数を発

話速度を表す特徴量として用いた。モーラとは、日本語における「拍」のことで、各モーラは時間的に等価である。基本的には「(子音 +) 母音」の構造を持つが、例外として、母音を持たない撥音「ん」や、音そのものにかかる促音「っ」がある [17]。

- ピッチの推定 対話者の発話ピッチ推定には音声信号処理ツールである SPTK を用いて行う。発話者の発話から取得した wave ファイルの基本周波数 (F0) を抽出し、その対数領域で平均値を求めることで発話者の平均ピッチを推定する。そこから 10[msec] 毎に分析窓をシフトして F0 を出力する。ここから得られた F0 値から無声音区間を取り除いた F0' の平均値を人のピッチとして推定する。
- 音量の推定 対話者の発話音量推定にはピッチ推定と同様、SPTK を用いる。ここでは簡単のために音声区間のパワー項の平均値を用い、これを人の音量として推定する。

以上で得られる情報から、人の声に関する韻律情報を推定することができる。

3.3 韻律情報から音声合成パラメータへの変換

ここでは、Julius, SPTK を用いて推定した人の韻律情報から VoiveText の音声合成パラメータへの変換式を求める作業について説明する。図 3.2 はその作業の様子を示している。本研究において、ロボットは日本サードパーティ株式会社の NAO[18] を用いた。ロボットの選択理由として、見た目が中性的であるということが挙げられる。これは、より厳密なペーシングを行うためには、男性と女性の発する声域に差があるため、人とロボット見た目と声の違和感を少なくする必要があると考えたからである。また、入力にはソニー株式会社のエレクトリックコンデンサーマイク [19] を用いた。なお、VoiceText の音声合成パラメータの値の制限は発話速度は 50~400(%), ピッチ, 音量は 50~200(%) となっている。変換式の導出作業で発話させた文章は、

- あらゆる現実をすべて自分のほうへねじまげたのだ
- テレビゲームやパソコンでゲームをして遊ぶ

- 救急車が十分に動けず救助作業が遅れている
- 老人ホームの場合は健康器具やひざ掛けだ
- 嬉しいはずがゆっくり寝てもいられない

の 5 種類を用いた [20]. これら 5 種類の文章は様々な音素をバランスよく含んでおり, 様々な発話に対応する必要があるペーシングシステムに適していると考えた. また, 導出作業は今回用いる音声合成パラメータである発話速度, ピッチ, 音量においてそれぞれ行った. 以下にその導出作業を具体的に示す.

1. VoiceText の該当する音声合成パラメータを 50 (発話速度は 70) に設定 (変換式を導出するパラメータ以外は 100 に固定) し, 合成音声を出力する
2. マイクで入力した音声を Julius で認識, 韻律情報を推定 (発話速度: Julius, ピッチ・ボリューム: SPTK) する
3. 推定した韻律情報とその時の音声合成パラメータを記録する.
4. 上記の作業を既に示した 5 種類の文章それぞれにおいて 3 回ずつ行う.
5. 該当する音声合成パラメータを 10 変化させ, 上記の作業を行う.
6. 以上の作業を該当するパラメータが 200 (発話速度は 150) になるまで行う.

以上の作業を行い, 記録した音声合成パラメータと韻律情報の関係性を用いて変換式の導出を行う.



図 3.2 音声合成パラメータへの変換式導出作業の様子

3.3.1 発話速度の変換式の導出

発話速度の変換式の導出のためには，推定した発話速度と VoiceText の発話速度パラメータの関係性を利用した．ここで用いる発話速度とは，発話速度 = モーラ数/発話時間 と定義する．図 3.3 に示すように，人の発話から求めた発話速度を x 軸に取り，VoiceText の発話速度パラメータの値を y 軸に取った時，変換式は線形回帰を用いて，

$$y = 10.843x + 8.7307 \quad (3.3.1)$$

となった．また，決定係数は $R^2 = 0.90864$ であった．

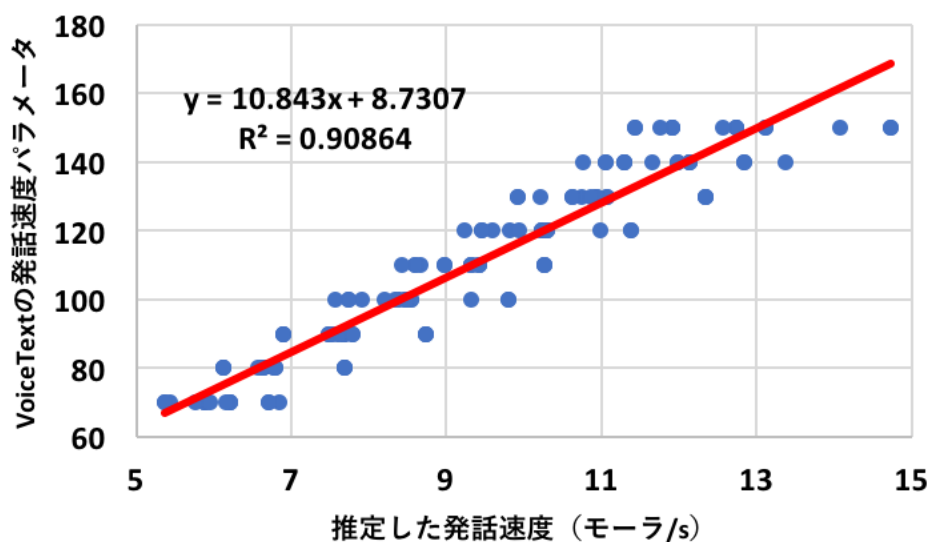


図 3.3 発話速度に関する推定された値と音声合成パラメータの関係

3.3.2 ピッチの変換式の導出

ピッチの変換式の導出のためには，推定したピッチと VoiceText のピッチパラメータの関係性を利用した．なお，人間の知覚では音の高さは基本周波数の対数軸上で等間隔に知覚されるため，推定したピッチには対数を取る処理を施している．そのため，VoiceText のピッチパラメータに関しても事前に常用対数を取った値を用いた．図 3.4 に示すように，人の発話から求めたピッチを x 軸に取り，常用対数を取った VoiceText のピッチパラメータの値を y 軸に取った時，変換式は線形回帰を用いて，

$$y = 1.0065x - 0.1826 \quad (3.3.2)$$

となった．また，決定係数は $R^2 = 0.99301$ となった．なお，音声合成の際には求めた値の真数を音声合成パラメータとして用いた．

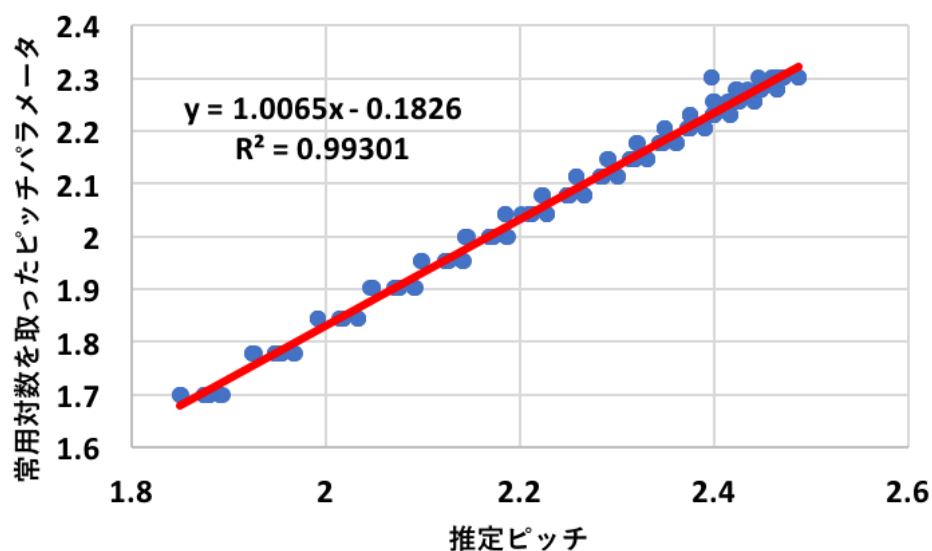


図 3.4 ピッチに関する推定された値と音声合成パラメータの関係

3.3.3 音量の変換式の導出

音量の変換式の導出のためには、推定した音量と VoiceText の音量パラメータの関係性を利用した。なお、人間の知覚では音量も音の高さと同様に対数化した上で等間隔に知覚されるため、推定した音量には対数を取る処理を施している。そのため、VoiceText の音量パラメータに関しても事前に常用対数を取った値を用いた。図 3.5 に示すように、人の発話から求めた音量を x 軸に取り、常用対数を取った VoiceText の音量パラメータの値を y 軸に取った時、変換式は線形回帰を用いて、

$$y = 0.2472x - 2.9784 \quad (3.3.3)$$

となった。また、決定係数は $R^2 = 0.97594$ となった。なお、音声合成の際には求めた値の真数を音声合成パラメータとして用いた。

しかし、VoiceText の特性上、ピッチが変化することで、それに伴い音量が変化

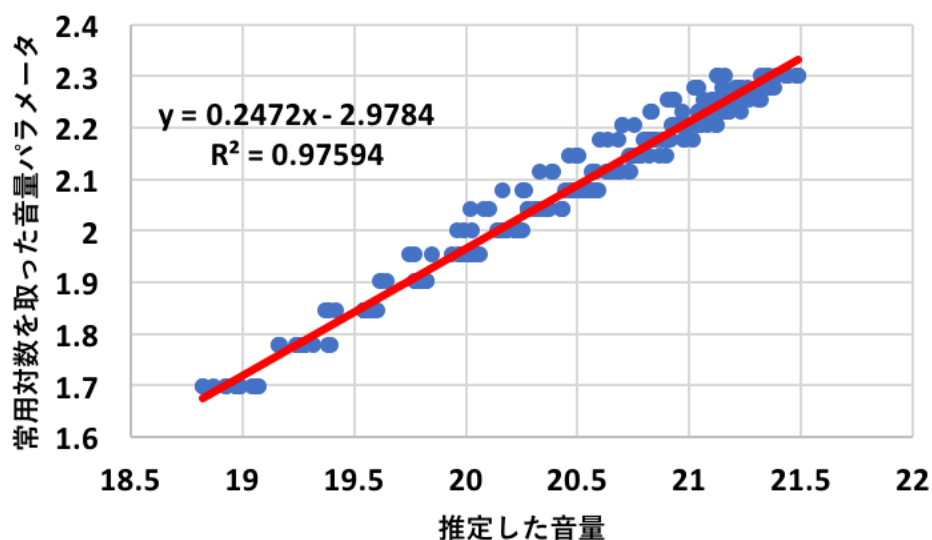


図 3.5 音量に関する推定された値と音声合成パラメータの関係

することが判明したため、ピッチの変化に伴い音量が変化する際の倍率を求める式の導出を行った。そして、実際のペーシングシステムでは推定したピッチから求めた倍率で音量パラメータの値を割るような処理を加えた。式の導出作業としては、基準となる音量を固定し、ピッチを徐々に変化させることで、基準となる音量に対する算出された音量の倍率を求めた。そして、ピッチとこの倍率との関係性から倍率を求める式を導出した。図 3.6 に示すように、VoiceText のピッチパラメータを x 軸に取り、ピッチの変化に伴う音量の変化の倍率を y 軸に取った時、その倍率を求める式は非線形回帰を用いて、

$$y = 0.0000001x^3 - 0.00005x^2 + 0.0122x + 0.2149 \quad (3.3.4)$$

となった。また、決定変数は $R^2 = 0.99601$ となった。

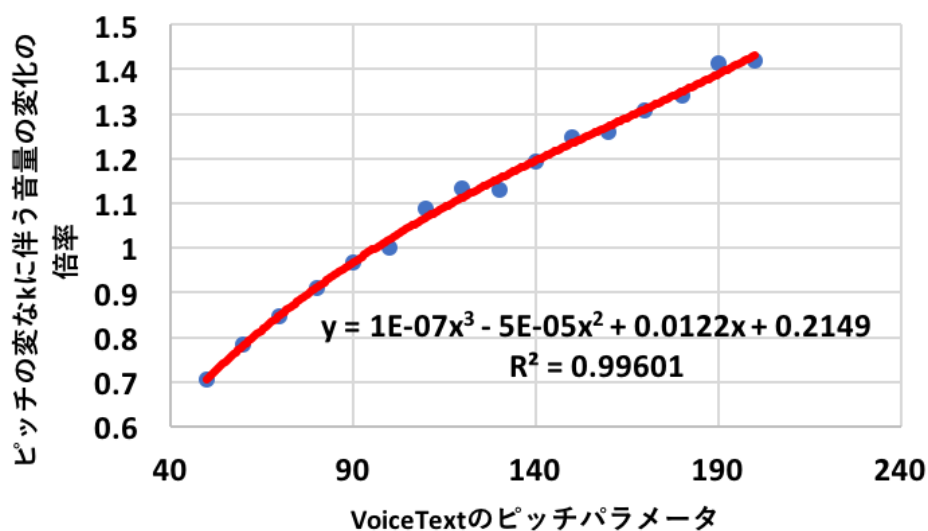


図 3.6 ピッチの音声合成パラメータとその変化に伴い音量が変化する倍率の関係

3.4 VoiceText による音声合成

ここでは、得られた VoiceText の音声合成パラメータを用いて音声合成を行う。今回、HOYA 株式会社が提供している音声合成ソフトウェアである VoiceText シリーズの中でも、Web サービスやスマートフォンアプリの開発用として提供されるウェブサービスである「Voice Text Web API」を用いた。VoiceText の設定可能パラメータとしては、話者、エンジンタイプ、ピッチ、速度、音量、感情/感情レベルがある。その中でも、今回は話者と、人の韻律情報から取得可能であるピッチ、速度、音量をパラメータとして用いた。話者パラメータについては対話者の性別に合わせて事前に設定をした上で、人の韻律情報をもとに得られた音声合成パラメータを用いて音声合成を行う。

3.5 ペーシングシステムの精度評価

本節では、構成したペーシングシステムの精度評価を行う。具体的には、人の声とペーシングシステムを用いて生成した合成音声の類似度に関して、主観的・客観的にそれぞれ評価を行う。その評価手法に関して、評価の準備も含めた一連の流れを以下に示す。

1. 様々な韻律特徴 (fast, slow, high, low, big, small, neutral) を持った 7 種類の人の音声を用意する。発話内容は「あらゆる現実をすべて自分のほうへねじまげたのだ」とした。
2. 各音声から、ペーシングシステムを用いて音声合成パラメータを取得する。発話内容は人の音声と同じものを用いた。
3. 上記の試行を 5 回ずつ繰り返し、各音声ごとに平均を取ったものを各音声における人の韻律情報とする。
4. 取得した人の韻律情報を音声合成パラメータとして、VoiceText を用いてロボットの合成音声を 7 種類生成する。
5. 人の音声と同様に、この各合成音声からもペーシングシステムを用いて音声合成パラメータを取得する。
6. 人の音声と同様に、5 回の試行の平均をロボットの韻律情報とする。

ここでは、この一連の試行から生成した 7 組の人の声とロボットの合成音声について主観的評価を用いて評価し、7 組の人とロボットの韻律情報を用いて客観的評価を行う。

3.5.1 精度の主観的評価

ペーシングシステム精度の主観的評価においては、人の声とロボットの合成音声の類似度に関するアンケートに回答してもらう。アンケートは 7 組の音声のペアに対して、4 つの評価項目（発話速度、ピッチ、音量、全体）を設け、それぞれ 7 段階のリッカート尺度（1：全く異なる-7：完全に一致している）で評価してもらう。なお、「全体」の評価項目とは、音声を聞いたときの「全体」的な印象を指す。また、アンケートの回答者は 20 代の男性 8 名である。表 3.1 に各音声の種類ごとのアン

ケートスコアの平均を示す．図 3.7 はそれをグラフ化したものである．

結果として，全ての種類の音声において発話速度のスコアが最も高く，ピッチのスコアが最も低かった．また，音量に関しても発話速度に比べてスコアが低い結果となった．これは，発話速度に比べてピッチと音量はわずかな差を人間の耳で判別するのが難しいからであると考えられる．また，被験者によっては，人の声と合成音声の声質の違いをピッチの違いと捉えた人もいたため，このような結果になったと考えられる．しかし，全体として見たときには，全ての項目においてこのペーシングシステムとして用いるのに十分なスコアであった．

表 3.1 ペーシングシステム精度の主観的評価値

韻律特徴 \ 評価項目	発話速度	ピッチ	音量	全体
fast	6.00	3.50	5.75	4.38
slow	6.13	4.63	5.13	4.88
high	5.88	4.50	4.88	4.38
low	6.25	4.63	5.75	4.88
big	6.38	4.38	5.13	4.75
small	6.13	4.38	4.88	4.63
neutral	6.00	4.63	5.00	4.86

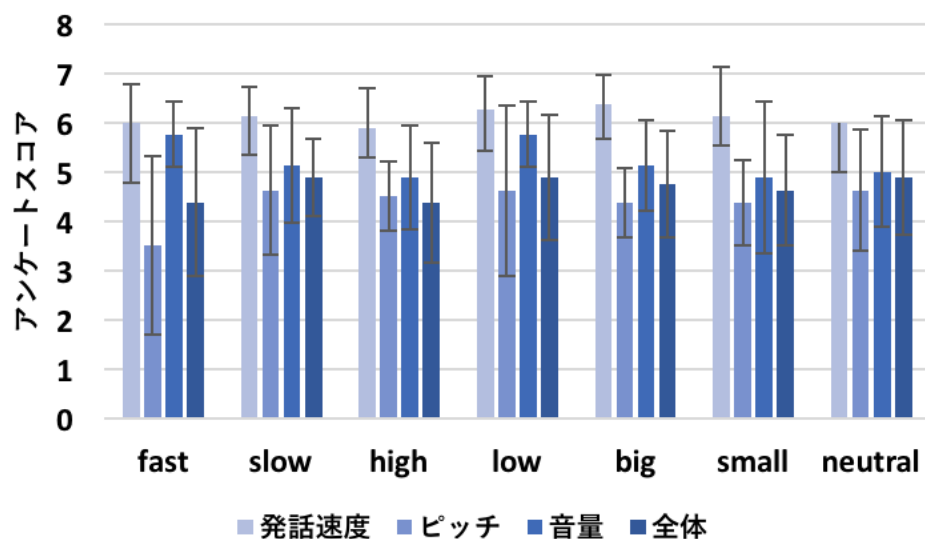


図 3.7 ペーシングシステム精度の主観的評価グラフ

3.5.2 精度の客観的評価

ペーシングシステム精度の客観的評価においては、人の声とロボットの合成音声のそれぞれの韻律情報の値のコサイン類似度と相関係数を求めた。表 3.2 に客観的評価の結果を示す。

結果として、全てのサンプルにおいて十分に高いコサイン類似度、相関係数を示した。

表 3.2 ペーシングシステム精度の客観的評価値

サンプル	話者	発話速度	ピッチ	音量	コサイン 類似度	相関係数
fast	人	124.24	98.16	101.02	0.9999	0.9954
	ロボット	125.00	98.05	103.59		
slow	人	83.43	98.74	100.53	0.9997	0.9944
	ロボット	81.11	98.93	103.68		
high	人	98.53	123.85	102.76	0.9998	0.9799
	ロボット	97.30	123.96	106.77		
low	人	102.03	79.13	98.75	0.9999	0.9969
	ロボット	98.53	78.95	97.33		
big	人	101.58	98.78	124.22	0.9994	0.9905
	ロボット	97.71	99.06	129.51		
small	人	101.58	98.81	73.82	0.9998	0.9936
	ロボット	98.53	99.14	75.46		
neutral	人	97.71	99.33	99.97	0.9999	0.8868
	ロボット	97.71.	99.03	102.18		

第 4 章 ペーシングを用いた対話時のロボットに対する 印象評価実験

本実験では、ペーシングを適用したロボットとの対話時の印象を主観的・客観的指標を用いて評価する。評価指標として主観的評価にはアンケート、客観的評価には心拍変動解析から取得できる LF/HF を用いた。実験の概要としては、被験者にロボットと対話を行ってもらい、その際に LF/HF を測定し、対話後にアンケートに答えてもらう。対話内容は事前に準備したシナリオ対話を行ってもらい、それを、ペーシング適用時と、非適用時で行い、比較する。

4.1 対話シナリオの評価

本実験の予備実験として、感情を喚起するシナリオを用意したという主張をサポートするため、用いる対話シナリオがどのような感情を喚起するのか評価を行う。ここでは、予備実験の実験手法、実験結果について述べる。

4.1.1 実験手法

実験では、20 代の男性 7 名にロボットと人の対話シナリオを読んでもらい、対話終了後にどのような感情になるかアンケートによる評価を行った。対話シナリオは 14 種類用意した。図 4.1 にその中の一例を示す（全てのシナリオについては付録を参照）。また、アンケートとしてアフェクト・グリッド [21] と呼ばれる感情の評定に使用される評価手法を用いた。アフェクト・グリッドとは、図 4.2 に示すように感情の地図のようなものである。図の真ん中は、感情的に neutral（中性）、つまりポジティブでもなくネガティブでもなく平常的な日常の感情を表す。図の右半分は、快感情を表し、右に行くほど快となる。逆に左半分は不快感情を表す。上下は覚醒の程度を表す。ポジティブ、ネガティブとは独立に、どれだけ覚醒し活性化していると感じるかを示す。上に行くほど覚醒的となり、下の方は眠気が強いということになる。右上は、エキサイト、喜びといった感情を表す。逆に左下は、抑うつ、悲しみ、憂うつといった感情を表す。左上は、ストレス、緊張を表す。逆に右

口： こんにちは！
被： こんにちは！
口： 好きな食べ物は何ですか？
被： 私，ラーメンがすごい好きなんですよ。
口： 私もラーメンは大好きです！何ラーメンが好きなんですか？
被： 醤油ラーメンが好きです。
口： 私も醤油ラーメンはよく食べに行きますよ！。
被： え，今度一緒に食べに行きませんか？
口： 本当ですか！私も誰かと行きたいと思っていたところなんですよ！
被： 実は，美味しいお店を知ってるんですよ。
口： どこにあるんですか？教えてください。
被： 大阪にあるんで，今度紹介しますね！
口： ありがとうございます！楽しみにしてます！
被： はい，私も楽しみです！

口＝ロボット，被＝被験者

図 4.1 対話シナリオの一例

下は，静穏，リラックスを表す．実験では，各シナリオを読んだ後に対話全体として最も当てはまる感情の位置に印を記入してもらった．

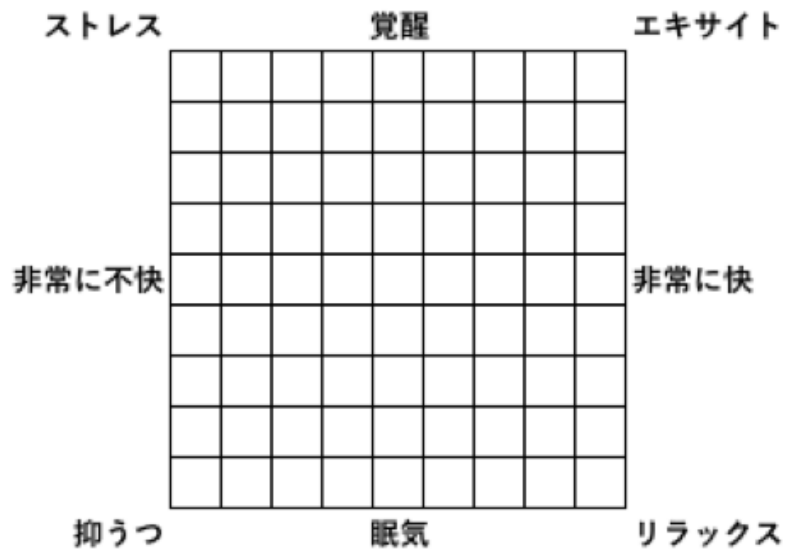


図 4.2 アフェクト・グリッド

4.1.2 実験結果

アフェクト・グリッドを用いた評価を図 4.3 に示す．なお，今回は覚醒-眠気の軸において真ん中をスコア 0 として上に行くほどプラスの評価（最高値：4），下に行くほどマイナスの評価（最低値：-4）とした．また，非常に快-非常に不快の軸においても真ん中をスコア 0 として右に行くほどプラスの評価（最高値：4），左に行くほどマイナスの評価（最低値：-4）とした．この結果より，用意した 14 種類の対話シナリオのうち 12 種類において，覚醒-眠気の軸においてプラスの評価を示した．また，この 12 種類のシナリオにおいては，非常に快-非常に不快の軸においてもプラスの評価を示した．なお，今回の実験で行う LF/HF を用いた評価では，人の覚醒の程度を測ることができる．さらに，一部のアンケート項目における評価から，人が感じる快感情の程度を見ることができる．そこで，覚醒度と快感情を喚起する対話シナリオを用いた時に，ペーシングを用いた対話によってユーザの覚醒度と快

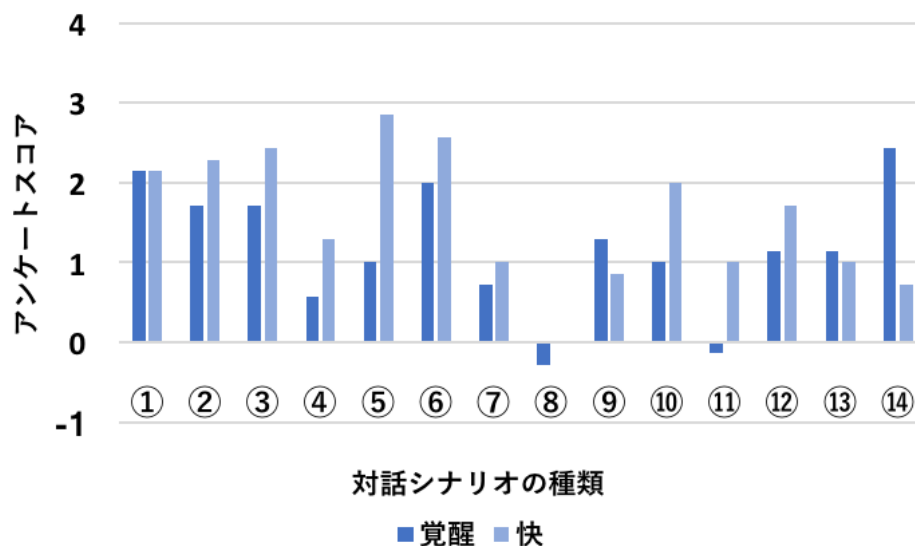


図 4.3 対話シナリオの評価結果

感情がそれぞれどのように変化するか調査する．そのため，本実験では選定した 12 種類の対話シナリオを用いて，人とロボットの対話を行うこととする．

4.2 印象評価実験

本実験では，4.1 節で選定した対話シナリオを用いて，被験者に実際にロボットと対話をしてもらいその際の印象をアンケートと LF/HF を用いて評価した．ロボットとの対話の様子を図 4.4 に示す．ここでは，実験手法について詳しく述べた上で，実験結果，考察について述べる．



図 4.4 ロボットとの対話の様子

4.2.1 実験手法

実験では、20 代の男性 8 名に 4.1 節で選定した 12 種類の対話シナリオを用いてロボットと対話してもらう。12 種類の対話を 3 種類ずつ、4 つのまとまりに分割し、1 つのまとまりを 1 回の対話とする。それぞれの対話に対して、ペーシングを適用した対話と適用していない対話を行ってもらうため、被験者には全部で 8 回の対話を行ってもらうことになる。対話の順番として、カウンターバランスを取るために被験者 8 人のうち半分の 4 人の被験者には、ペーシング適用→ペーシング非適用→非適用→適用→適用→... の順番で対話を行い、残りの 4 人に被験者にはその逆の、ペーシング非適用→ペーシング適用→適用→非適用→非適用→... の順番で対話を行った。また、シナリオには、図 4.5 に示すように、テンションをあげて読むところは赤字でテンションを下げて読むところは青字で示されており、その他の文章に関しても感情を込めて読むように指示した。

ペーシング非適用時のロボットの音声合成パラメータの決定は、各被験者の平

口： こんにちは！
 被： こんにちは！
 口： 好きな食べ物は何ですか？
 被： 私，ラーメンがすごい好きなんですよ。
 口： 私もラーメンは大好きです！何ラーメンが好きなんですか？
 被： 醤油ラーメンが好きです。
 口： 私も醤油ラーメンはよく食べに行きますよ！。
 被： え，今度一緒に食べに行きませんか？
 口： 本当ですか！私も誰かと行きたいと思っていたところなんですよ！
 被： 実は，美味しいお店を知ってるんですよ（内緒話みたいに）。
 口： どこにあるんですか？教えてください。
 被： 大阪にあるんで，今度紹介しますね！
 口： ありがとうございます！楽しみにしてます！
 被： はい，私も楽しみです！

口＝ロボット，被＝被験者

図 4.5 会話文に対する読み方の指示例

均的な各音声合成パラメータに対して逆の値を取るようにした。その手法を図 4.6 に示す。例えば，図に示すように，ロボットのある音声合成パラメータについて，ペーシング適用時の場合，図の赤い丸のような推移をする発話をしたとする。この時，ペーシング非適用時の場合は図の青い丸に示すように被験者内平均との差について，絶対値が同じで正負が逆になるような値を取るようにした。

被験者内の平均的な音声合成パラメータを取得するためには，被験者に実験本番

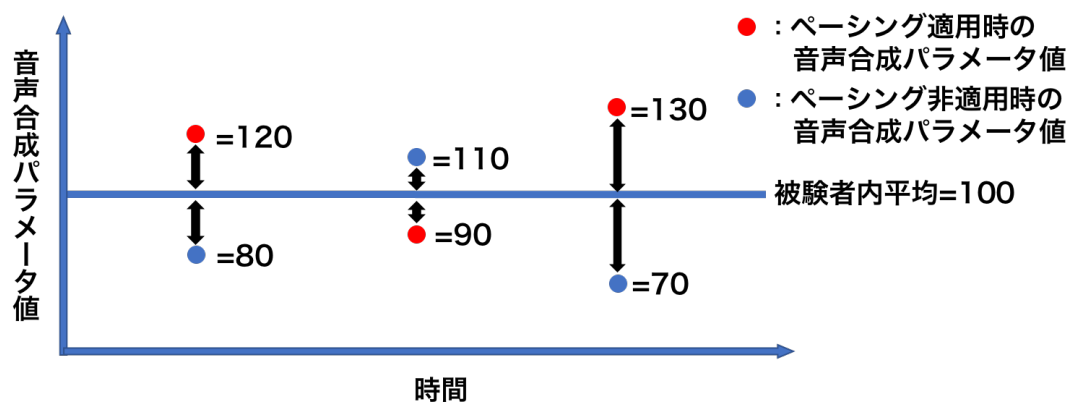


図 4.6 ペーシング非適用時のパラメータの決定手法

の対話を行ってもらう前に、「対話の練習」と称しロボットと対話を行ってもらう。この時の、対話内容は 4.1 節において、実験のためのシナリオとして選ばなかった 2 シナリオを合わせたものとした。この際、対話相手としてペーシングを適用したロボットを用いた。この対話から取得した被験者の各音声合成パラメータの平均を取り、それを被験者の平均的な音声合成パラメータとした。表 4.1 に各被験者の平均的な音声合成パラメータを示す。

アンケートによる主観的評価

ユーザがロボットと対話をした時の印象についての主観的評価として、アンケートによる評価を行った。アンケートの内容は先行研究を参考に、信頼関係構築に関する質問を複数行い、ペーシング適用時と非適用時、各条件での評価値を比較する。表 4.2 にアンケート項目を示す。質問 1 に関しては、柿井 [22] による関連研究において、人同士の信頼関係構築を示す質問として利用されていたものを用いた。質問

2, 3, 4 に関しては Tarr ら [23] によるダンスにおける同調が社会的結合に与える影響について調査した研究において利用していた質問を利用した。質問 5 に関しては、ロボットへの信頼性は継続利用に繋がるため [24], 項目の 1 つとした。質問 6, 7 に関しては今回試験的に用いた。質問 1-6 に関しては、7 段階のリッカート尺度（1：全くそう思わない-7：非常にそう思う）で質問を行った。質問 7 に関しては、同一対話シナリオにおける 2 回のうち、どちらのロボットが好きであったか選択してもらった。そのため、1 回目の対話では回答せず、2 回目の対話後のみ回答してもらった。

表 4.1 各被験者の平均的な音声合成パラメータ

	発話速度	ピッチ	音量
被験者 1	121	99	67
被験者 2	114	103	171
被験者 3	118	82	167
被験者 4	112	60	138
被験者 5	110	77	119
被験者 6	120	80	109
被験者 7	121	92	100
被験者 8	114	84	156

表 4.2 ロボットの印象についてのアンケート項目

1. ロボットは親しみやすかったですか？
2. ロボットと話していて楽しかったですか？
3. ロボットとどれくらい繋がりを感しましたか？
4. ロボットはどれくらい自分に同調していましたか？
5. ロボットを継続的に利用したいと感じましたか？
6. ロボットはあなたの話を聞いている感じがしましたか？
7. 2 回の対話のうちどちらのロボットが好きでしたか？（2 回目の対話後のみ回答）

LF/HF による客観的評価

ユーザがロボットと対話をした時の印象についての客観的評価の1つとして心拍変動解析から得られる LF/HF を用いた。図 4.7 のような心電図の R 波と R 波の間隔 (R-R 間隔) を周波数解析することで、自律神経系指標の 2 つのパワースペクトル成分を抽出できる [25]。その中でも、パワースペクトルの低周波成分の領域 (LF:0.05-0.15Hz)、高周波成分の領域 (HF:0.15-0.40Hz) の強度の積分値を用いる。HF 成分は、副交感神経が活性化している場合にのみに現れる。LF 成分は、交感神経が活性化している場合と副交感神経が活性化している場合の両方に現れる。そのため、交感神経の活動指標を求めるためには、一般的に LF と HF の比をとり、LF/HF を交感神経の活動として取り扱っている。LF/HF の値の取得に使用した心拍センサ、解析用ソフトウェアは図 4.8 に示す。人が活性状態にあるときには交感神経の活性により、LF/HF の値が上昇し、不活性状態にあるときには副交感神経の活性により、HF の値が上昇する。なお、今回用いる対話シナリオは 4.1 節で行った評価より、覚醒度を上昇させることが分かっている。そのため、そこで、覚醒度を上昇させる対話シナリオを用いた時に、ペーシングを用いた対話によってユーザの覚醒度がどのように変化するのか調査する。

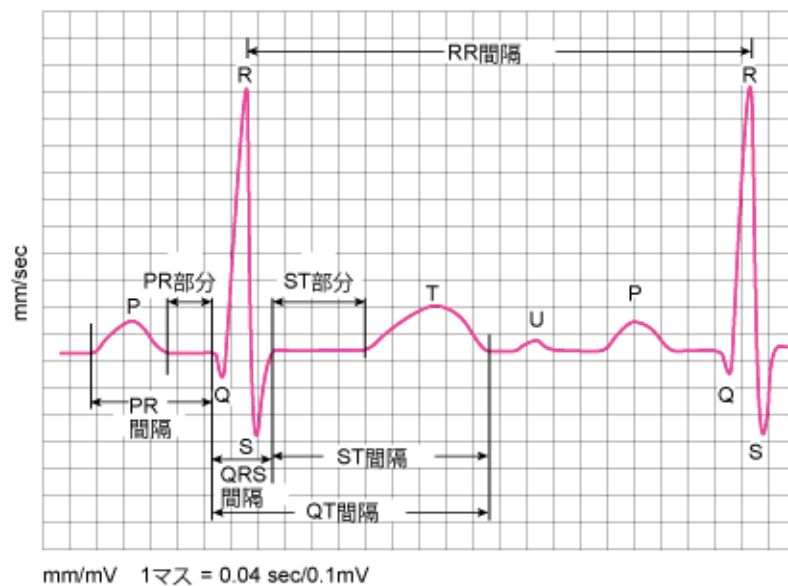


図 4.7 心電図の波形

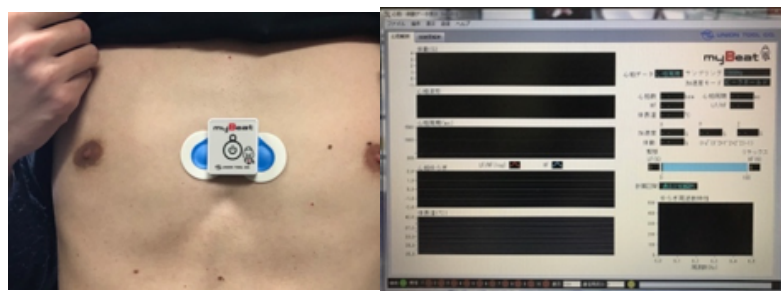


図 4.8 心拍センサ（左），解析用ソフトウェア（右）

4.2.2 実験結果

被験者には，ペーシングを適用したロボットと適用していないロボットそれぞれと対話を行ってもらった．その際の印象について，アンケートによる主観的評価結

果と LF/HF による客観的評価結果を示す。

アンケートによる主観的評価結果

アンケートによる主観的評価結果を表 4.3, 表 4.4 に示す。表 4.3 に質問 1-6 の回答の平均, 標準偏差, 及び p 値を示す。なお, ここでの平均は各被験者でシナリオ間での平均値を取り, その値について被験者間での平均値を取ったものである。そして, それをグラフ化したものが図 4.9 である。結果として, 各項目においてペーシングを適用した対話時の方が評価値が高く, 有意水準 5% で有意差が見られた。表 4.4 には質問 7 への回答結果とシナリオ間における平均を示す。そして, それをグラフ化したものが図 4.10 である。結果として, 全てのシナリオにおいてペーシングを適用したロボットの方が好きという意見を得た。また, シナリオ間での平均値においてもペーシングを適用したロボットの方を好きという意見が多く, 有意水準 5% で有意差が見られた ($p = 0.0032 < 0.05$)。

表 4.3 アンケートスコア (質問 1-6) の平均と標準偏差及び p 値

	ペーシング非適用		ペーシング適用		p 値
	平均	標準偏差	平均	標準偏差	
親しみやすいか	3.09	0.78	4.97	0.73	0.0014
楽しいか	3.22	0.83	4.84	0.87	0.0037
繋がりを感じたか	2.94	0.68	4.75	0.59	0.0052
自分に同調していたか	3.03	0.77	5.16	1.01	0.0004
継続利用意欲があるか	2.78	1.13	4.34	0.77	0.0126
話を聞いている感じがするか	3.91	0.99	5.47	0.88	0.0051

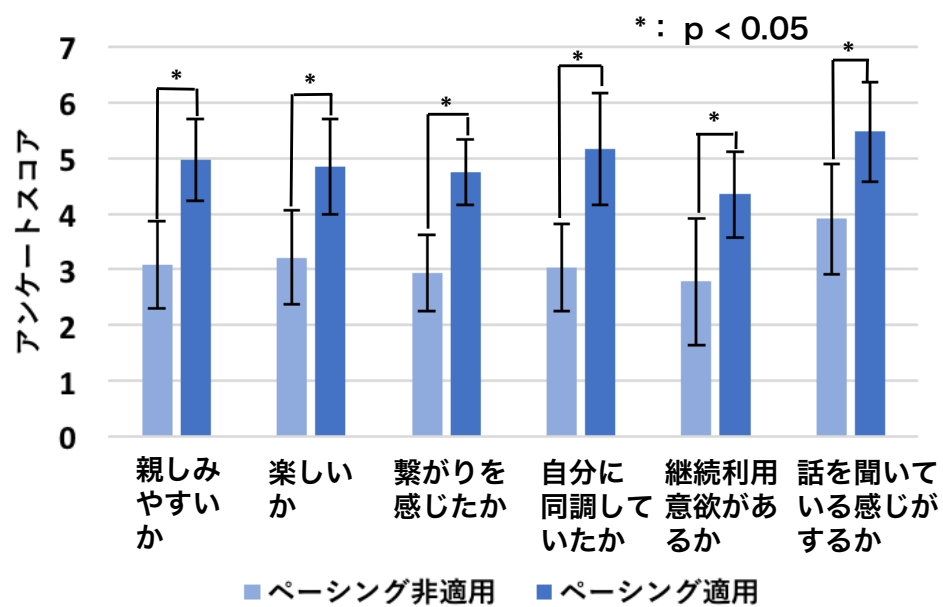


図 4.9 アンケート評価結果：質問 1-6

表 4.4 質問 7 への回答結果と全シナリオにおける平均

	ペーシング非適用	ペーシング適用
シナリオ 1	1 人	7 人
シナリオ 2	2 人	6 人
シナリオ 3	2 人	6 人
シナリオ 4	1 人	7 人
平均	1.5 人	6.5 人
標準偏差	0.5	0.5

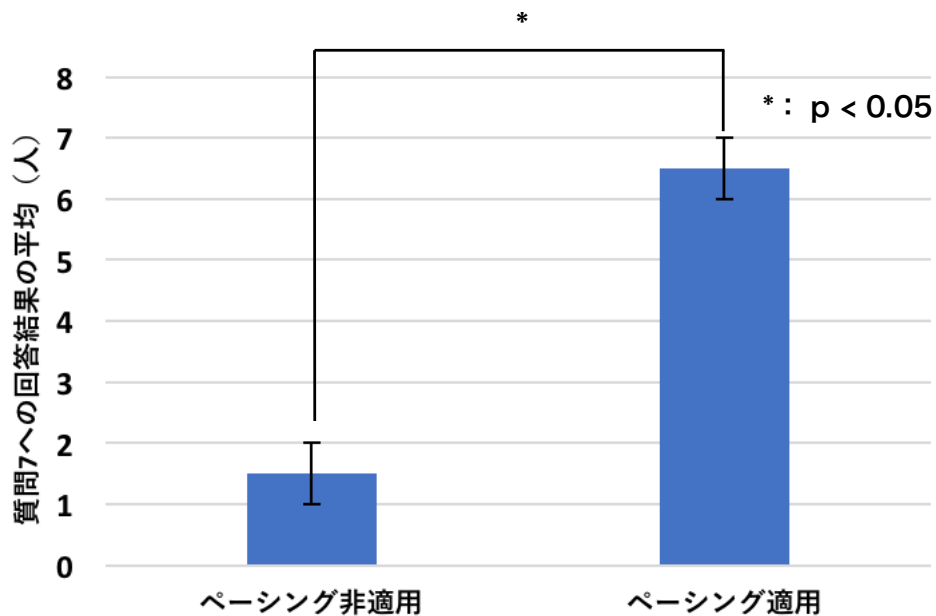


図 4.10 アンケート評価結果：質問 7「どちらのロボット（ペーシング適用時・非適用時）が好きか」

LF/HF による客観的評価結果

LF/HF の値の平均と標準偏差及び p 値を表 4.5 に示す．ここでの平均とは，各被験者ごとに LF/HF の時系列データの平均を取り，その値について被験者間での平均を取ったものである．また，それをグラフ化したものが図 4.11 である．結果として，ペーシング適用時とペーシング非適用時の被験者間での LF/HF の平均値において，有意水準 5% で有意差が見られた．

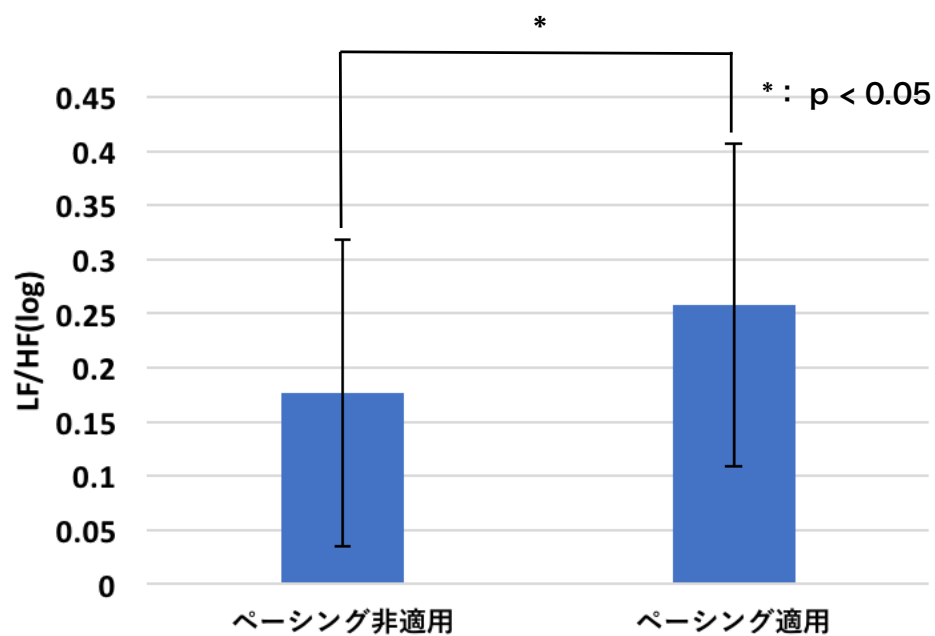


図 4.11 LF/HF による評価結果

表 4.5 LF/HF の値の平均と標準偏差及び p 値

ペーシング非適用		ペーシング適用		
平均	標準偏差	平均	標準偏差	p 値
0.18	0.14	0.26	0.15	0.004

4.3 考察

表 4.3, 図 4.9 に示されるように, アンケート (質問 1-6) による主観的評価の結果より, ペーシングを適用したロボットとの対話の方が適用しなかったロボットと比較して, 「親しみやすいか」「対話は楽しいか」「繋がりを感じたか」「自分に同調していたか」「継続的に利用したいと思ったか」「話を聞いている感じがするか」の各項目において有意差が見られた. これらの評価項目は関連研究などで信頼関係構築などを示すための評価項目として利用されていたものである. また, 表 4.4, 図 4.10 に示されるように, 「2 回の対話 (ペーシング適用時, 非適用時) のうちどちらのロボットが好きか」という質問への回答より, ペーシングを適用したロボットの方がペーシング非適用のロボットに比べて好感を持たれることが分かった. よって, これらの印象評価結果よりペーシングを適用した対話が信頼関係構築に貢献することが示された.

次に, 表 4.5, 図 4.11 に示されるように, LF/HF を用いた客観的評価の結果より, ペーシングを適用したロボットとの対話とペーシング非適用時の対話において, LF/HF の値においても有意差が認められた. この結果より, ペーシングを適用したとロボットとの対話では非適用時の対話に比べて, 覚醒度が増幅していることが分かる. また, アンケート項目である「対話は楽しいか」という項目における評価より, ペーシングを適用した対話では非適用時に比べて「楽しい」といった印象を受けることが分かる. これは, 図 4.12 に示すラッセルの円環モデルより, 「快」感情を喚起することが分かる. これらより, ペーシングを適用した対話では非適用時に比べて図 4.2 のアフェクト・グリッドにおける「快」と「覚醒」においてプラスの値を示す領域の感情を喚起することが分かる. 今回使用した対話シナリオは, 事前に 4.1 節で行った評価より, 「快」と「覚醒」がプラスの値を示すことが分かっている. これらの結果より, ペーシングを適用した対話の方が, 非適用時の対話に比べて, 対話シナリオが喚起するであろう感情を増幅させていることが分かる.

また, 実験の際のアンケートへの自由記述において,

- ロボットの声色も高くなったり楽しんでくれている気がした.

といった意見も見受けられた. これは, 被験者が想像しているロボットの感情と被験者自身の感情が類似しており, 共感を引き起こしているのではないかと考えられ

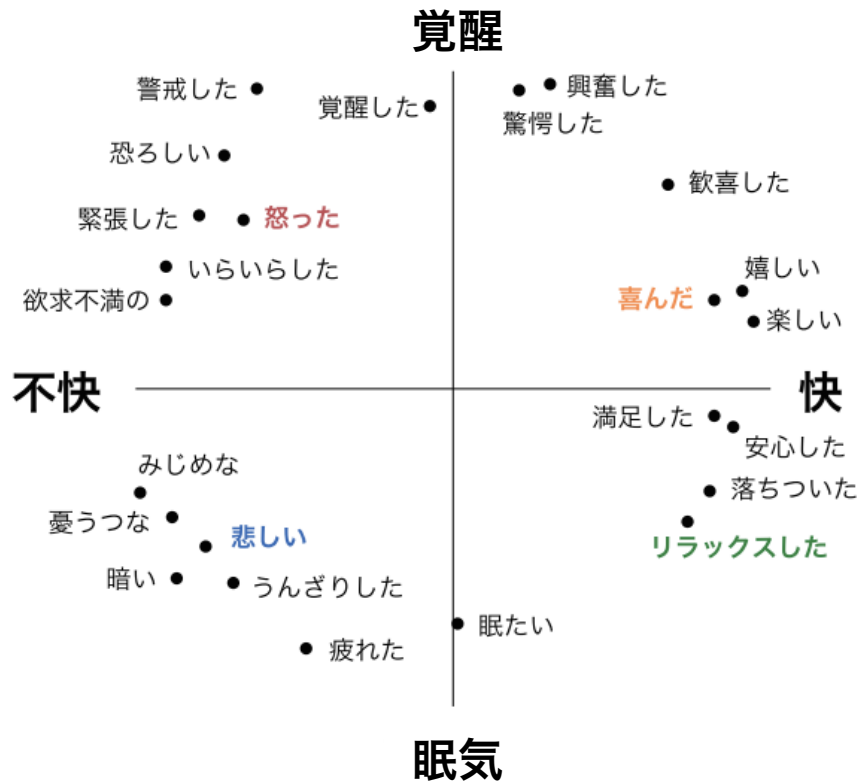


図 4.12 ラッセルの円環モデル

る。さらに、

- ロボットも乗ってきてくれると話を聞いて同じ感情を共有しているように感じた。
- こちらの調子に合わせてロボットも同じ感情を共有して話しているように感じた。
- ロボットが話し手の言うことをただ聞くだけでなく、感情まで理解しようとしてくれているように感じた。

といった意見からも被験者がロボットに対して共感を示していることが分かる．長岡ら [7] によると信頼関係は共感的態度で接することで構築されるとされており，これらの結果からもペーシングを適用した対話が信頼関係構築に貢献することが示された．

第 5 章 おわりに

本研究では、ロボットとユーザの信頼関係構築を目指して、対話ロボットにペーシングを適用したペーシングシステムを実装し、それを適用したロボットとの対話における印象を主観的・客観的指標を用いて評価した。今回実装したペーシングシステムでは、推定した人の韻律情報を音声合成パラメータに変換する必要がある。そのため、推定した人の韻律情報を音声合成パラメータに変換する式の導出を行った。さらに、実装したペーシングシステムが、実際に使用するにあたって妥当であるか検証を行うために、人の声とそれをもとに作成した合成音声のペアを 7 種類用意し、その類似度を主観的・客観的に評価した。主観的評価ではアンケートによる 7 段階評価、客観的評価ではコサイン類似度、相関係数を用いた評価を行った。その結果、主観的評価において、各音声において 4 以上のスコアとなった。また、客観的評価においても各音声において、コサイン類似度、相関係数においてほぼ 0.9 以上の値を示した。そのため、本研究ではこのペーシングシステムを用いて実験を行うこととした。本実験を行う前に、まず、予備実験としてロボットとユーザが行う対話シナリオがどのような感情を喚起するのか、アフェクト・グリッドを用いて評価を行った。この結果、今回の実験では覚醒度を高め、快感情を喚起する対話シナリオを用いることにした。そして、実際に被験者にロボットと対話を行ってもらい、対話中の LF/HF の値を取得し、対話後に信頼関係構築に関するアンケート調査を行った。

アンケートの結果、「親しみやすいか」「対話は楽しいか」「繋がりを感じたか」「自分に同調していたか」「継続的に利用したいと思ったか」「話を聞いている感じがするか」の各項目において、ペーシングを適用したロボットとの会話を行った時の方が高い評価値を示し、有意な差も認められた。これらの項目は、関連研究において信頼関係構築などを示すために用いられていたものである。この結果より、ペーシングを用いたロボットとの対話は、信頼関係構築に貢献することが示された。

LF/HF を用いた評価においては、ペーシングを適用したロボットとの対話時と適用しなかったロボットとの対話時の LF/HF の値において有意な差が認められた。この結果は、ペーシング非適用時の対話に比べて覚醒度が増幅していることを示している。また、主観的評価の結果よりペーシングを適用した対話では非適用時

に比べて、被験者が「楽しい」といった「快」感情を喚起するような印象を受けることが分かった。これは、予備実験で示した対話シナリオが喚起するであろう覚醒度と快感情を増幅させており、ペーシングによって対話シナリオが喚起するであろう感情が増幅することが示された。

さらに、アンケートにおける自由記述と上記の結果より、ロボットに対して共感を示す被験者が見られた。共感は信頼関係に繋がると言われており、この結果からもペーシングを適用した対話が信頼関係の構築に貢献することが示された。

今回の実験による結果から、主観・客観的評価を用いた印象評価よりペーシングを用いたロボットとの対話が信頼関係構築に貢献することが示された。また、ペーシングの効果として、対話シナリオが喚起する人の感情を増幅させるという効果も示された。今後は、今回用いることができなかった、快感情を測ることができるような客観的指標を用いて、今回の結果と組み合わせることで、さらに強力な根拠のもとに信頼関係構築を示すことができるのではないかと考えられる。また、今回とは別の感情の条件においても、ペーシングによって感情の増幅を引き起こすことができるのか網羅的に調べる必要がある。

付録

練習用

・1 気候について

ナ：こんにちは！

被：こんにちは！

ナ：最近、寒くなってきましたね。

被：そうですね。特に夜はすごい冷えますね。

ナ：そのせいで最近風邪気味なんですよ。

被：えー、大丈夫ですか？

ナ：喉が痛い程度なんで全然大丈夫です。

被：じゃあ、まだ安心ですね！

ナ：この時期は毎年風邪引くんですよ。

被：私もよく風邪引くんで気をつけます。

ナ：そうですね。お互いに気をつけましょう。

被：はい、お大事になさってください。

ナ：治ったらまた遊びましょう！

被：もちろんです！

・2 一人暮らしについて

ナ：こんにちは！

被：こんにちは！

ナ：今は一人暮らしですか？

被：はい、そうですよ。

ナ：一人暮らし寂しくないですか？

被：もう、長いこと一人暮らしなんで慣れましたね。

ナ：実家には帰ったりしてないんですか？

被：そうですね、最近は全然帰れていないんですよ。

ナ：そうなんですよ。忙しそうですもんね。

被：そうですね、なかなか難しいですね。

ナ：次はいつ実家に帰る予定なんですか？

被：年末には久しぶりに帰省しますよ！

ナ：良かったじゃないですか！年末が楽しみです！

被：はい、楽しみです！

本番用

・1 自己紹介

ナ：こんにちは！初めまして！

被：初めまして！

ナ：私の名前はナオです。よろしくお願いします。

被：私は〇〇（自分の名前）です。

ナ：出身はどこですか？

被：私は名古屋出身です！

ナ：そうなんですね！私は大阪出身です！

被：あー、じゃあ、ちょっと遠いですね。

ナ：そうですね。なかなか大阪には来ないんですか？

被：そういえば、来月大阪に行く予定ありますよ！

ナ：そうなんですか！じゃあ、その時遊びましょう！

被：良いですね。大阪の美味しいお店教えてください！

ナ：分かりました！任せてください！

被：ありがとうございます！

・2 好きな食べ物

ナ：こんにちは！

被：こんにちは！

ナ：好きな食べ物とかありますか？

被：私、ラーメンがすごい好きなんですよ。

ナ：私もラーメンは大好きです！何ラーメンが好きなんですか？

被：醤油ラーメンが好きです。

ナ：私も醤油ラーメンはよく食べに行きますよ！

被：え、今度一緒に食べに行きませんか？

ナ：本当ですか！私も誰かと行きたいと思っていたところなんですよ！

被：実は、美味しいお店を知っているんですよ。（内緒話っぽく）

ナ：どこにあるんですか？教えてください。

被：大阪にあるんで、今度紹介しますね！

ナ：ありがとうございます！楽しみにしてます！

被：はい、私も楽しみです！

・3 映画の誘い

ナ：こんにちは！

被：こんにちは！

ナ：最近、休みの日は何をしているんですか？

被：最近は予定がなくて暇なんですよ。

ナ：じゃあ、今度、映画を見に行きませんか？

被：良いですね！行きましょ！

ナ：今週の日曜とかはどうですか？

被：うーん、今週の日曜は予定が入っているんですよ。

ナ：そうなんですね。

被：来週の土曜日とかはどうですか？

ナ：行けますよ！じゃあ、来週の土曜日にしましょう。

被：はい！昼ごはんでも食べて映画を見に行きましょ。

ナ：分かりました！じゃあ、12時に駅前集合にしましょう。

被：分かりました！

・4 料理について

ナ：こんにちは！

被：こんにちは！

ナ：今は一人暮らしですか？

被：そうですよ。

ナ：自炊とかするんですか？

被：最近、忙しくてできてないですね。

ナ：前は自炊していたんですか？

被：少し前までは毎日していましたよ。

ナ：すごいですね。私、料理苦手で外食が多いんですよ。

被：練習すれば誰でもできるようになりますよ！

ナ：じゃあ、もし良かったら今度何か料理教えてもらってもいいですか？

被：もちろんです！一緒に作って食べましょ！

ナ：良いですね！楽しみにしています！

被：はい、私も楽しみです！

・5 ご飯の誘い

ナ：こんにちは！

被：こんにちは！

ナ：今度、一緒にご飯に行きませんか？

被：何食べに行きますか？

ナ：カレーとかどうですか？

被：あー、私、辛い料理苦手なんですよ。

ナ：そうなんですね。じゃあ、辛いものにしませんか！

被：そうしてくれたらありがたいです。

ナ：パスタとかどうですか？

被：パスタ良いですね！

ナ：前回行ってみて美味しかったお店があるんですよ！

被：それは期待大ですね！

ナ：私のほうで予約しておきますね！

被：ありがとうございます！

・6 寝不足の原因

ナ：こんにちは！

被：こんにちは。

ナ：眠そうですね。昨日、夜更ししてたんですか？

被：はい、今、とても眠いです。

ナ：何してたんですか？

被：今日の午前中にあったテストの勉強をしました。

ナ：大変ですね。無事に終わりましたか？

被：お陰様でなんとかいけました！

ナ：それはめでたいですね！今日でテストは終わりなんですか？

被：はい！今日まででした！

ナ：じゃあ、今日の夜は飲みにも行きますか！

被：はい！行きましょう！

ナ：また、後で連絡しますね！

被：わかりました！

・7 好きなスポーツ

ナ：こんにちは！

被：こんにちは！

ナ：好きなスポーツは何ですか？

被：サッカーが好きです。

ナ：だったら今度一緒に試合を見に行きませんか？

被：あー、どちらかというと実際にやる方が好きなんですよ。

ナ：そうなんですか。試合観戦には行ったことは無いんですか？

被：そうですね、行ったことはないです。

ナ：じゃあ、一回行ってみませんか？楽しいですよ！

被：確かにそうですね！行きたいです！

ナ：来週の土曜日に試合があるので、見に行きましょう！

被：はい、楽しみにしています！

ナ：また、詳しいことは連絡しますね！

被：はい、お願いします！

・8 夏休みの出来事

ナ：こんにちは！

被：こんにちは！

ナ：夏休みは何をしてたんですか？

被：夏休みは旅行に行っていましたよ。

ナ：どこに行ってたんですか？

被：沖縄に行ってきました！

ナ：沖縄良いですねー。何したんですか？

被：色々、観光名所を見てきましたよ！

ナ：海には入らなかったんですか？

被：天気が悪くて、海には入れなかったんですよ。

ナ：それは残念ですね。でも、天気は仕方がないですよ。

被：次行くときは絶対、海に行きます！

ナ：沖縄の海は本当に綺麗なんでオススメです！

被：すごい楽しみです！

・9 サウナについて

ナ：こんにちは！

被：こんにちは。

ナ：最近、疲れ気味ですね。

被：仕事が忙しくてなかなか疲れが取れないんですよ。

ナ：だったら、サウナとかどうですか？

被：サウナ苦手なんですけど、何が良いんですか？

ナ：たくさん汗をかいたら、疲れも吹き飛びますよ！

被：アツくてなかなか我慢出来ないんですよ。

ナ：正しいサウナの入り方を覚えたら絶対にハマりますよ！

被：そうなんですね！是非、教えてください！

ナ：もちろんです！

被：もし、良かったら一緒に行ってくださいませんか？

ナ：是非行きましょう！オススメの銭湯紹介しますね！

被：ありがとうございます！

・10 アルバイトについて

ナ：こんにちは！

被：こんにちは！

ナ：最近、節約してるらしいですね。

被：お金がないんですよ。

ナ：アルバイトとかしていないんですか？

被：時間が無いんで、したくてもできないんですよ。

ナ：だったらオススメのバイトがありますよ！

被：え、教えてください！

ナ：家庭教師のバイトなんですけど、時間短くて、時給も高いんでオススメですよ！

被：どこで見つけてきたんですか？

ナ：知り合いの家が家庭教師を探していたんですよ！

被：なるほど！詳しく教えてもらっても良いですか？

ナ：はい！分かり次第連絡しますね！

被：ありがとうございます！

・ 11 年末年始について

ナ：こんにちは！

被：こんにちは！

ナ：年末年始は何してたんですか？

被：美味しいものを食べてずっと寝てました。

ナ：たまにはそういうのも良いですね！

被：ただ、そのせいで少し太ってしまいました。

ナ：年末年始はそうなりますよね。私も太っちゃいました。

被：運動しないとイケないですね。

ナ：そうなんですよ。

被：じゃあ、一緒に運動しましょう！

ナ：良いですね！しましょう！

被：毎日夕方一緒に走りましょう！

ナ：そうしましょう！

被：早速明日から始めましょう！

・ 12 無くした財布について

ナ：こんにちは！

被：こんにちは。

ナ：元気ないですけど、どうしたんですか？

被：実は、昨日財布を無くしてしまったんですよね。

ナ：それは大変ですね。どこらへんで無くしたんですか？

被：多分、駅前で落としたんだと思うんですが、定かではないですね。

ナ：そういえば、昨日友達が財布拾ったって言ってましたよ！

被：本当ですか！もしかしたら、僕のもشれないです！

ナ：駅前の交番に届けたって言ってたんで行ってみたらどうですか？

被：そうですね、いつ頃拾ったって言ってましたか？

ナ：確か、夕方頃って言ってた気がします！

被：分かりました！今から行ってみます！

ナ：見つかったら良いですね！

被：はい、教えてくださってありがとうございます。

謝辞

研究活動全般にわたり，指導教官の萩田紀博教授には幅広い視野で多くの助言をいただきました．ここに感謝の意を表します．丁寧かつ熱心なご指導，ご鞭撻をいただいた神原誠之准教授に深く感謝致します．副指導教官としての的確なご指摘をしてくださった加藤博一教授に熱く御礼申し上げます．本研究において，心理学的な観点からの的確なアドバイスをくださった，京都大学こころの未来研究センター佐藤弥准教授には心から感謝致します．本研究に関わる出張手続きなど，多くの事務手続きを一手に引き受けてくださった南あずさ秘書に深くお礼申し上げます．また，実験に参加してくださった被験者の方々に感謝申し上げます．そして，本研究だけでなく大学院生活で多岐にわたりお世話になった西村祥吾先輩や，澤邊太志先輩に心から感謝致します．その他，共に研究活動を歩み，日々切磋琢磨した環境知能学研究室の皆様にも心から感謝申し上げます．最後に，今までお世話になった両親に深く感謝致します．

参考文献

- [1] 内閣府, “高齢者の生活環境,” in, 2016. 平成 28 年版高齢社会白書
- [2] 神田崇行, 佐藤留美, 才脇直樹, and 石黒浩, “対話型ロボットによる小学校での長期相互作用の試み,” ヒューマンインターフェース学会誌, vol. 7, no. 1, pp. 27-37, 2005.
- [3] 永野浩二, “人間の存在 (presense) を体験する授業の試み,” 追手門学院大学教育研究所紀要, no. 30, pp. 25-40, 2012.
- [4] 加藤雄士, “プロセス・モデルのコーチングに関する一考察: 日本における「ニューコード NLP コーチング」の可能性について,” 産研論集, no. 44, pp. 67-78, 2017.
- [5] 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.
- [6] S. Nishimura, H. Kawanami, M. Kanbara, and N. Hagita, "Empathetic speech synthesis applied to a chat robot to obtain the users confidence," Presented by poster in Late Breaking Results Poster Sessions, 2017 IEEE International Conference on Robotics and Automation, Singapore, 2017.
- [7] 長岡千賀, “対人コミュニケーションにおける非言語行動の 2 者相互影響に関する研究,” 対人社会心理学研究, vol. 6, pp. 101-112, 2006.
- [8] R. Mogan, R. Fischer, and J. A. Bulbulia, "To be in synchrony or not? A meta-analysis of synchrony's effects on behavior, perception, cognition and affect," Journal of Experimental Social Psychology, vol. 72, pp. 13-20, 2017.
- [9] S. L. Koole, and W. Tschacher, "Synchrony in psychotherapy: A review and an integrative framework for the therapeutic alliance," Frontiers in psychology, vol. 7, p. 862, 2016.
- [10] 山野美咲, 薄井達也, 橋本稔, “情動同調に基づく人間とロボットのインタラクション手法の提案,” HAI シンポジウム, 2008.
- [11] 小林優佳, 山本大介, 土井美和子, “高齢者対話インターフェースユーザの聞

- き手になる音声対話インターフェース," 情報処理学会-シンポジウム論文集, vol.109, no.224, pp.799-802, 2011.
- [12] L. Huang, L. P. Morency, and J. Gratch, "Virtual Rapport 2.0," International Workshop on Intelligent Virtual Agents, pp.68-79, 2011.
 - [13] K. Maekawa, K. Hanae, S. Furui, and H. Isahara, "Spontaneous speech corpus of Japanese," Proceedings of the Second International Conference of Language Resource and Evaluation, vol. 2, pp.947-952, 2000.
 - [14] A. Lee and T. Kawahara, "Recent development of open-source speech recognition engine julius," APSIPA, pp.131-137, 2009.
 - [15] SPTK working group, "Examples for Using Speech Signal Processing Toolkit Ver. 3.9," 2015.
 - [16] HOYA 株式会社, "VoiceText Web API," <http://voicetext.jp/products/vt-webapi/>
 - [17] 熊田政信, 小林武夫, 小崎寛子, 新美成二, "けいれん性発声障害の新しい評価法: モーラ法," 音声言語医学, vol. 38, no. 2, pp. 176-181, 1997.
 - [18] JTP 日本サード・パーティ株式会社, "NAO ヒューマノイドロボット-JTP 日本サード・パーティ株式会社," <https://www.jtp.co.jp/services/robotics/nao/>.
 - [19] ソニー株式会社, "ECM-674/9X|業務用オーディオ|法人のお客様|ソニー," https://www.sony.jp/pro-audio/products/ECM-674_9X/.
 - [20] 国立情報学研究所, "音声資源コンソーシアム," <http://research.nii.ac.jp/src/ATR503.html>
 - [21] J. A. Russell, A. Weiss, and G. A. Mendelsohn, "Affect grid: a single-item scale of pleasure and arousal.," Journal of personality and social psychology, vol. 57, no. 3, p. 493, 1989.
 - [22] 柿井俊昭, "双方向型 TV を用いたマルチメディア・カウンセリングの基礎的研究," 心理学研究, vol. 68, no. 1, pp. 9-16, 1997.
 - [23] B. Tarr, J. Launay, Dunbar, and R. IM. Dunbar, "Silent disco: dancing in synchrony leads to elevated pain thresholds and social closeness," Evolution and Human Behavior, vol. 37, no. 5, pp. 243-349, 2016.
 - [24] 中川佳弥子, 篠沢一彦, 松村礼央, 石黒浩, and 萩田紀博, "ヘルスケアロボット

へのパーソナリティ付与による説得効果," 第 9 回情報科学技術フォーラム講演論文集 (FIT2010), pp. 89-92, 2010.

- [25] 松本佳昭, 森信彰, 三田尻涼 江鐘偉, "心拍揺らぎによる精神的ストレス評価法に関する研究," ライフサポート, vol. 22, no. 3, pp. 105-111, 2010.

発表リスト

- [1] 中村卓矢, 西村祥吾, 川波弘道, 神原誠之, 萩田紀博, "対話意欲向上を目的としたロボット対話における同調的音声合成に関する検討", 信学技報, vol, 118, no. 94, pp. 39-42, 2018.6.
- [2] 中村卓矢, 西村祥吾, 川波弘道, 神原誠之, 萩田紀博, "ユーザとの信頼関係構築を目的としたロボット対話におけるペーシングの評価", 信学技報, vol, 118, no. 184, pp. 33-38, 2018.8.
- [3] 中村卓矢, 西村祥吾, 川波弘道, 神原誠之, 萩田紀博, "ユーザとの信頼関係構築を目的としたロボット対話における同調的音声合成の評価", 情報科学技術フォーラム, J-002, 2018.9.
- [4] 西村祥吾, 中村卓矢, 川波弘道, 神原誠之, 萩田紀博, "ユーザとの信頼関係構築を目的とした対話ロボットのためのユーザの情動に同調する対話手法", 情報科学技術フォーラム, J-001, 2018.9.