

ロボットとの対話がもたらす人の発話リズム変化

Effects of human-robot interaction on speech rhythm modulation in humans

川崎真弘^{1,2} 山田陽平¹ 牛久陽介¹ 宮内英里¹ 山口陽子^{1,2}

Masahiro KAWASAKI^{1,2}, Yohei YAMADA¹, Yosuke Ushiku¹, Eri MIYAUCHI¹, and
Yoko YAMAGUCHI^{1,2}

¹ 理化学研究所 BSI-トヨタ連携センター 脳リズムモデル連携ユニット

¹Rhythm-based Brain Computation Unit, RIKEN BSI-TOYOTA Collaboration Center

² 理化学研究所脳科学総合研究センター 創発知能ダイナミクスチーム

² Laboratory for Dynamics of Emergent Intelligence, RIKEN Brain Science Institute

Abstract: In our daily conversations, there exists a speech rhythm that is shared among speakers. This study attempted to investigate the difference between the speech rhythms of humans and robots and the effects of human-robot interaction on human-human speech rhythms so that we could further explore the possibility of introducing robots into human society. We conducted speech tasks where two human subjects or one human subject and one robot alternately pronounced alphabets (e.g., “A,” “B,” or “C,”). A human-human pair of subjects performed the task before and after a human-robot pair did, where the robot used five voices—electronic, male, female, the human subject’s partner’s voice, and the human subject’s voice. Forty subjects took part in the Electroencephalogram (EEG) experiments during the tasks. The results of subjective evaluations for the alternate speech tasks showed that subjects were more comfortable and felt a deeper sense of togetherness with other humans than with a robot, especially after the human-robot experiments were conducted. Along with subjective evaluations, the speech rhythms became synchronized and the EEG rhythms (i.e. the frontal theta activity) were more pronounced after human-robot experiments. These results suggested that human-robot interactions improve our speech rhythms.

1. はじめに

近年教育現場や職場など多くの場でロボットの導入が期待されている。このようなロボットには、従来の補助だけでなく、その存在やコミュニケーション力によって自律的かつ円滑な学習や労働を促すことも求められている。現在のロボットは、その姿かたち、会話に至るまで人間に近い状態を作り上げている。しかし、実際にロボットとのコミュニケーションがどのように我々の認知機能に影響を及ぼし、その効果が人とのコミュニケーションとどの程度違うのか、については不明である。この問題を解決するためには、いまだ不明瞭な社会的相互作用に関与する脳メカニズムを参考にする必要がある。

そこで本研究では、目には見えないが人間同士の発話コミュニケーションの中には確かに存在する「適切な間（ま）」に注目する。人間は時計のように正確に間を生み出してはおらず、このノイズが加わったリズムこそが、両者の行動を相互作用的に変化

させ、相手から思いやりを感じる信頼関係を生む可能性がある。事実、人間同士の発話コミュニケーションを取り扱った従来研究は発話のピッチや発話長に注目し、話者のしぐさや視線、呼吸の引き込みなどノンバーバル情報との相関関係を示してきた[1]。このような間は、人とロボットのコミュニケーションでも観測されるのか、そして人とロボットのコミュニケーションは人間同士の間を変化させるのか、について知ることが、ロボットを使った教育の有用性を説くうえで重要である。

このような他者との発話コミュニケーションには、ミラーニューロンシステムなどの社会性に関係する脳メカニズムが関係するだろう。ミラーニューロンとは、ある行動をとっている他者の行動を観測している時にあたかも自分が行動したように活動するニューロンのことで、人間にも存在すると考えられている[2]。このミラーニューロンは他者に共感した時の活動が報告されており、もしそうなら快適な対話コミュニケーションが実現されている時にもこのよ

うな活動が観測されるだろう。従来脳波研究では、2者が同時または交互にボタン押し（タッピング）をしている時の運動同期に関わる脳波同期を調べ、タイミングが一定の位相関係に引き込み合っていき、この共有されたリズムに関係する脳波リズムを特定した[3]。このような行動のリズムに相関した脳波リズムは、自然に協調し合う発話リズムでも見られるのだろうか？そして、人とロボットのコミュニケーションは脳波リズムに影響を及ぼすのだろうか？

そこで本研究では、人間とロボットの発話コミュニケーションが、その前後に行った人間同士の発話に与える影響を、発話間隔、主観的感情（アンケート結果）、脳活動から分析した。実験は、2者が交互にアルファベットを発話する課題を用いた。また脳活動の分析として、本研究は特に脳波リズムに注目した。近年の研究では、脳波リズムの同期現象と認知機能が高い関連性を持つことが報告されている[4]。

2. 実験

2. 1 被験者

2名の被験者を1組として実験を行った。20組40名の健常な被験者（右利き、 21.5 ± 0.5 歳、女性8名）が理化学研究所安全管理委員会承認の同意書記入の上、脳波計測実験に参加した。アンケート結果の記入を損じた1組を除く、19組38名のデータに対して解析を行った。

2. 2 交互発話課題

本研究では、人間と人間、または人間とロボットの2者がAからGまでのアルファベットを交互に発声し合う課題を行った。1被験者が「A」と発声した後、相手は「B」と発声し、被験者は続けて「C」と発声する。Gまで発声し終えたら続けてAから繰り返し発声する。この交互発話を70秒のセッションの間続けた。

人間と人間の交互発話のセッションでは、2被験者はシールドルーム内で約150cm離れて向かい合って座り、実験中は2者の間に置かれた透明な仕切りの中心点を注視することで視界内に相手の顔を含めた状態で課題を遂行した（図1A左）。被験者には、顎台に顎を乗せた状態で、手を机の上に乗せ体を動かさないように教示した。

1セッション終了後、各被験者はアンケートを記入した。アンケートは、「快適性」、「一体感」、「速さ」、「ペース」について5段階で評価した。快適性の項

目では、このセッションの発話が快適であれば5、不快であれば1を回答した。一体感の項目では、このセッションの発話中に、相手と自分の発話が同等に感じられるほど一体感を感じたら5、まったく感じなければ1を回答した。速さの項目では、このセッションの発話が速いと感じれば5、遅いと感じれば1を回答した。ペースの項目では、このセッションの発話が自分のペース、つまり相手が自分に合わせてくれたと感じた場合5、相手のペース、つまり自分が相手に合わせたと感じた場合1を回答した。

人間とロボットの交互発話のセッションでは、机の上に被験者から約50cm離れて置かれたPoCoBot（ビジネスデザイン社）を相手に交互発話課題を行った（図1A右）。PoCoBotは5種類の音声を発した：(a) 電子音で作られたロボット声、(b) ネイティブ英語を話す男性の声、(c) アナウンサーの女性の声、(d) 被験者自身の声、(e) 被験者の相手の声。被験者の声は実験前に記録したAからGまでの音声データを使用する。これらの各声について発話のペースを一定にした条件とノイズを入れた条件で行い、各被験者は合計10条件を行った。

各セッション後に行うアンケートの記入は、人間同士の場合と同じ「快適性」、「一体感」、「速さ」、「ペース」に加え、「人間性」についても回答を要求した。人間性の項目では、このセッションで発したロボットが人間らしいと感じた場合5、ロボットらしいと感じた場合1を回答した。

各被験者ペアは、最初に人間同士で先行話者を入れ替えて2セッション行った（ロボット実験前）。その後、被験者ごとにロボットを相手に5セッションずつ行った。5種類の声は被験者間でカウンターバランスをとり、一定ペースとノイズを含むペースは交互に行った。各被験者が10セッションを終えた後、再び人間同士で先行話者を入れ替えて2セッション行った（ロボット実験後）。1セッションは約70秒（AからGまでの交互発話が20セット以上達成される）、セッション間隔は少なくとも10秒とった。以上の計24セッションを脳波と音声データ、ビデオデータを計測しながら行った（図1B）。

2. 3 音声データ解析

音声データはリニアPCMレコーダPCM-D50（SONY社）を用いて、96.00kHz、16bitで録音した。ステレオで記憶した左右の音声データについて、それぞれ周波数解析によってノイズ除去を行った。音声と無声の分離を行うために、各被験者で実験前に記録していたAからGまでの音声データを用いて、フォルマント周波数を分析した。このフォルマント

データを用いて、2 者の発話長とインターバルを分離した（図 3）。発話の間に注目する本研究では、2 者の交互発話によって生まれる時間間隔を解析した。

2. 4 脳波データ解析

BrainAmp を用いて 2 名の脳波を同時に計測した。脳波は頭皮に置かれた 28 チャンネルのアクティブ電極で計測し、基準電位は両耳より計測した（サンプリングレート：1000Hz）。データ解析は作業記憶の脳活動を特定した先行研究を参考にした[4]。同時に眼球運動を計測し、独立成分分析を用いて瞬きや眼球運動と相関した因子を削除し、瞬きやサッケードなどのアーチファクトを除去した。これらの脳波データに対して、カレントソース解析を用いて、体積伝導の効果を減らし、空間解像度を上げた。

上記の前処理を施したデータ $s(t)$ に対して、Morlet 関数 $w(t, f)$ を用いたウェーブレット解析を行い、各時刻 (t) 、各周波数 (f) に対する振幅と位相の時系列分析を行った。振幅 $E(t, f)$ は、実験全体の平均をベースラインとして各周波数で比較を行った。

$$w(t, f) = (\sigma_f \sqrt{\pi})^{-1/2} \exp(-t^2 / 2\sigma_f^2) \exp(i2\pi f t)$$

$$E(t, f) = |w(t, f) \otimes s(t)|^2$$

3. 結果

3. 1 アンケート結果

3.1.1 快適さ

人間同士の発話に対する快適さは、ロボット実験前に比べ実験後で有意に高い値を示した（図 2A; $P < 0.01$ ）。また各ロボットの声に比べて、人間との交互発話は有意に快適さが高かった（ $P < 0.05$ ）。

人間対ロボットの交互発話に対して、ロボットの声の種類、発話リズムの 2 因子について 2 元配置分散分析を行った結果、声の種類に主効果が見られ（ $P < 0.04$ ）、発話リズムには主効果がなかった（ $P = 0.72$ ）。ロボット声が他の声に比べて快適さが低く、被験者自身の声が一番快適さが高い結果を示した。

A 人間 vs 人間



ロボット vs 人間



B

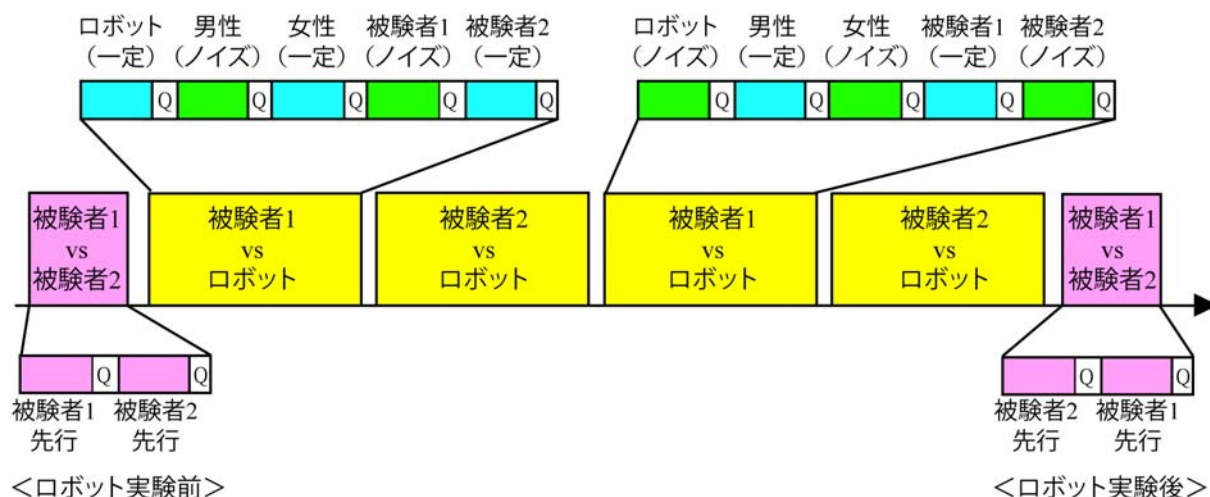


図 1 : A. 実験風景。人間同士（左図）と人間とロボット（右図）の交互発話課題。B. 実験の流れ図。最初に被験者同士で 2 セッション、その後 1 被験者 2 回ずつロボットを相手に 5 セッションを連続して行い、最後に被験者同士で 2 セッションを行った。1 セッション終わるごとにアンケート用紙の記入を行った（図の Q）。

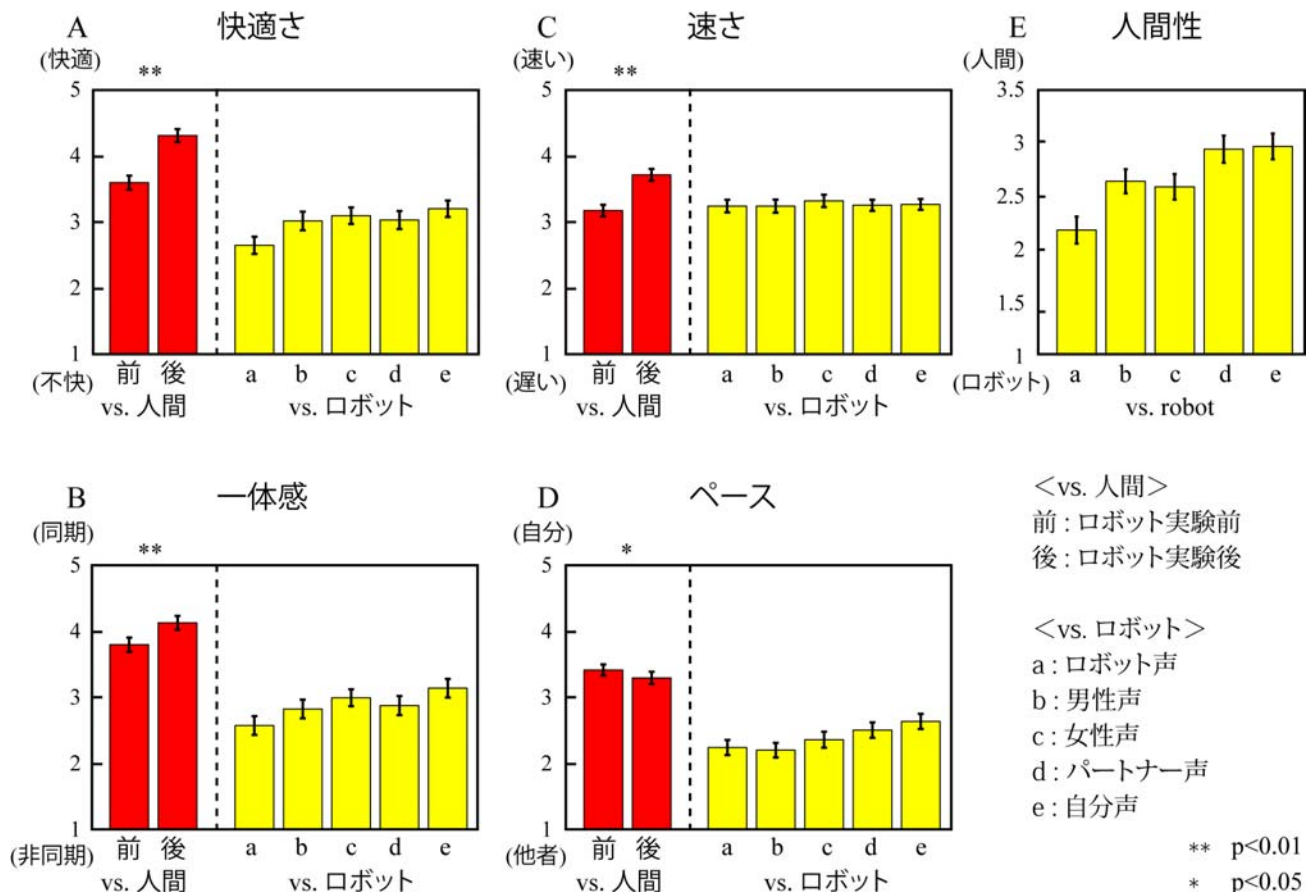


図 2: アンケートの結果。A. この交互発話を快適と感じたか、B. この交互発話で相手と一体感を感じたか、C. この交互発話は速かったか、D. この交互発話は自分が相手に合わせたか、相手が自分に合わせたか、E. この交互発話を行ったロボットに人間性を感じたか。ロボット実験を行う前後の人間同士の発話 2 条件 (紫色) と、a 電子音を使ったロボット声、b 男性の声、c 女性の声、d 被験者パートナーの声、e 自分自身の声、のロボット実験 5 条件に関するアンケート結果を示した。エラーバーは標準誤差を示した。

3.1.2 一体感

人間同士の交互発話に対する一体感は、ロボット実験前に比べ実験後で有意に高い値を示した (図 2B; $P < 0.01$)。また各ロボットの声に比べて、人間との交互発話は有意に一体感が高かった ($P < 0.05$)。

2 元配置分散分析より、声の種類に主効果があり ($P < 0.05$)、発話リズムには主効果がなかった ($P = 0.67$)。ロボット声が他の声に比べて一体感が低く、被験者自身の声が一番一体感が高い結果を示した。

3.1.3 速さ

人間同士の交互発話に対する速さは、ロボット実験前に比べ実験後で有意に高い値を示した (図 2C; $P < 0.01$)。またロボット実験前の速さに関する感想は、各ロボットの声に対するもの差がなかったものの、実験後では有意に速さの値が高くなる結果を示

した ($P < 0.05$)。

2 元配置分散分析より、声の種類 ($P = 0.97$)、発話リズム ($P = 0.33$) とともに主効果がなく、ロボットの声の違い、一定間隔かノイズを加えた間隔かによる違いは見られなかった。

3.1.4 ペース

人間同士の交互発話に対するペースは、ロボット実験前に比べ実験後で有意に低い値、つまり他者ペースである結果を示した (図 2D; $P < 0.05$)。また各ロボットの声に比べて、人間との交互発話のペースは有意に高い値、つまり自分ペースであった ($P < 0.05$)。

2 元配置分散分析より、声の種類に主効果があり ($P < 0.05$)、発話リズムには主効果がなかった ($P = 0.47$)。ロボット声が他の声に比べて他者ペースに感じる結果を示した。

3.1.5 人間性

2 元配置分散分析より、声の種類に主効果があり ($P < 0.01$)、発話リズムには主効果がなかった ($P = 0.44$)。特にロボット声は一番小さく、続いて男性声と女性声が小さい値を示した。被験者自身とパートナーの声は他の声に比べて大きかった。発話間隔が一定でもノイズを加えた場合でも差はなかった。

3. 2 音声データ解析結果

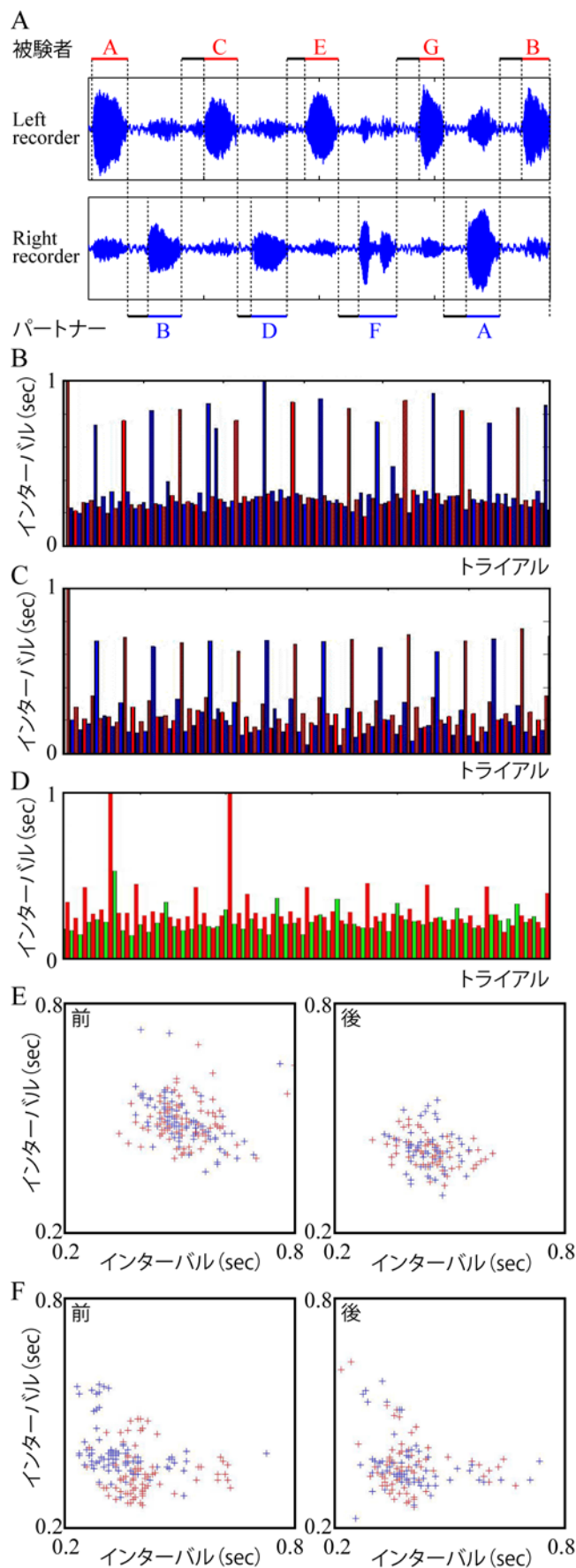
交互発話課題の中で、被験者の発した音声とその感覚の分離に成功した (図 3A)。これらのデータに対して、被験者の発声の間隔を分析した結果、2 者で間隔が一定になるパターン (例: 図 3B)、2 者で同じ揺らぎを共有するパターン (例: 図 3C)、2 者がそれぞれ固有の揺らぎをもつパターン (例: 図 3D)、が観測された。アンケート結果との対応付けを行った結果、間隔が一定のパターンと揺らぎを共有するパターンでは、快適さと一体感が高かった。一方で、2 者がそれぞれ固有の揺らぎを持つパターンでは、快適さと一体感が低く、代わりに自分ペースであると回答する割合が高かった。

発話間隔のリズムとして、あるトライアルと次のトライアルの 2 者の発話間隔をプロットし、関係性を調べた。その結果、ほとんどのペアがロボットとの発話課題を行う前後で、その関係性は分散が小さくなる (例: 図 3E) または一定になる (例: 図 3F) ように変化した。

3. 3 脳波データ解析結果

対話時の脳波データの各周波数帯域の振幅を算出し、時間平均した。本研究では多くの従来研究で認知機能や社会性に関係することが示されている前頭連合野の活動が顕著であったため、この脳部位に注目した[11, 12]。人間同士の交互発話実験の脳波データを周波数解析した結果、シータ波、アルファ波を含む 20Hz 以下の低周波のパワー値が高く、これはロボット実験の後の方が有意に高い値を示した (図 4A)。また 20Hz 以上のベータ波とガンマ波には実験前後の差が見られなかった。

図 3 : A. 音声データ解析例。音声部分 (赤・青線) と間隔の部分 (黒線) を切り分ける。B. 2 者で間隔が一定になるパターンの例。2 者の各トライアルにおける発話間隔を示した棒グラフ。C. 2 者で同じ揺らぎを共有するパターン。D. 2 者がそれぞれ固有の揺らぎをもつパターン。E, F. ロボット実験前後の n トライアルと $(n+1)$ トライアルの 2 者の発話間隔。



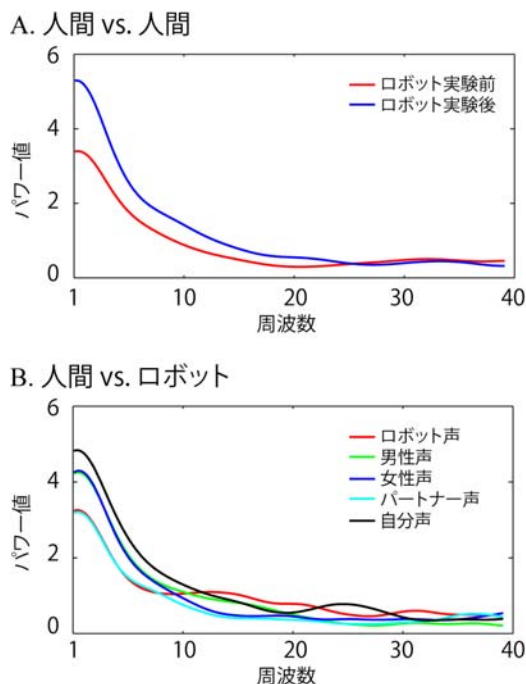


図 4：前頭連合野中心部（Fz）の脳波データを周波数解析した結果。A. ロボット実験前（赤色）後（青色）の人間同士の交互発話時の脳波データ。B. ロボット声（赤色）、男性声（緑色）、女性声（青色）、パートナー声（水色）、自分の声（黒色）のロボットに対する交互発話時の脳波データ。

人間とロボットの交互発話に関する脳波データは、自分の声、男性と女性の声、被験者パートナーとロボットの声、の順番で 8Hz 以下のシータ波のパワー値が大きい結果を示した（図 4B）。さらに自分の声を発するロボットとの交互発話に対しては、他の条件に比べて、20Hz から 30Hz のベータ波が増加することが分かった。

4. 考察

本研究はロボットとの発話をもたらす影響について、人と人、人とロボットの交互発話課題を用いて、両者の「間（ま）」の取り方の違い、それに対する主観的な感情、パフォーマンスの変化、脳波リズムの変化を、実験的に示した。主観的な感情は、ロボットに対するより、人間に対して行う交互発話の方が、快適さが高く、一体感を持ちやすく、自分のペースで発話できる結果を示した。さらに、これらの快適性と一体感は、ロボット実験を終えた後に行う交互発話でさらに高くなること、自分のペースと感じる割合が減ることがわかった。このような発話に関す

る主観的な感情は、音声データより解析した 2 者の発話の間隔の同期と相関していた。以上の結果はロボットとの交互発話が人間との交互発話に影響を及ぼしたことを示唆する。

交互発話課題時の脳波データ解析の結果、前頭連合野の活動が条件によって変化が見られた。人間同士の交互発話時には低周波数のパワー値が大きくなり、これはロボット実験を行った後にさらに増加する結果を示した。特に 8Hz 以下のシータ波は、快適さや一体感などの主観的なアンケート結果と同様にロボットの声に対しても変化した。前頭連合野のシータ波は従来研究によって認知機能だけでなく主観的な感情の変化にも関係することが示唆されている[5]。それゆえ本研究の結果は、ロボット実験が人間同士のコミュニケーションに関わる脳活動を変化させたことを示唆する。

今後の課題として他のアンケート結果と含めて、2 名の関係性や性別など、被験者個人の特徴に分類して解析を進める予定である。さらに 2 者の脳部位の同期解析や異なる脳波周波数間の解析を進めることによって、これらの行動のリズムに関する脳波リズムの詳細な脳ダイナミクスを示す必要がある。

謝辞

本研究は、科研費・新学術領域研究「人ロボット共生学」（課題番号：22118510）の補助により実施された。実験データの測定・解析に協力していただいた水野佑治さん、野口蓉子さん、塚越菜穂子さん、小泉有輝さん、ポコボット開発に協力していただいた高橋英之さんに深く感謝の意を示します。

参考文献

- [1] 山田誠二：人とロボットの＜間＞をデザインする、東京電機大学出版局、(2007)
- [2] Rizzolatti G, Craighero L.: The mirror-neuron system. Annual Review of Neuroscience. 27: pp.169-92, (2004)
- [3] Tognoli E, Lagarde J, DeGuzman GC, Kelso JA.: The phi complex as a neuromarker of human social coordination. Proc. Natl. Acad. Sci. U S A. 104(19): pp.8190-5, (2007)
- [4] Kawasaki, M. Kitajo, K. and Yamaguchi, Y.: Dynamic links between theta executive functions and alpha storage buffers in auditory and visual working memory. Eur. J. Neurosci., 31, pp.1683-1689, (2010)
- [5] Kawasaki, M. and Yamaguchi, Y.: Effects of subjective preference for colors on attention-related occipital theta oscillations. NeuroImage, 59: pp.808-814, (2012)