

対話エージェントに対する人間の言語選択: 単語の種類と感情状態に着目した実験的検討

Word choice toward dialog agents: An experimental study on the effects of kinds of words and emotional state

星田 雅弘^{1*} 田村 昌彦¹ 林 勇吾^{1,2}
HOSHIDA Masahiro¹ TAMURA Masahiko¹ HAYASHI Yugo^{1,2}

¹ 立命館大学

¹ Ritsumeikan University

² カーネギーメロン大学

² Carnegie Mellon University

Abstract: 人間は対話エージェントとのコミュニケーションにおいて、相手が用いた言葉を繰り返して用いる傾向がある。これは、機械は言語理解能力が低いという先入観から、相手に合わせようと心掛けた言語選択であり、感情状態が影響していることが考えられる。本研究では、人間がエージェントに対して用いる単語の種類と感情状態の関係を実験的に検討した。仮説としては、心拍変動が大きいほどエージェントが用いた単語と距離の近い単語を用いると予想した。エージェントと画像を伝え合う課題を用いて、単語選択時における脈波の取得から心拍を推定した。分析の結果、単語の距離と心拍変動の相関が正と負の群に分けられ、単語選択に覚醒度が影響していることが示唆された。

1 はじめに

我々は他者とのコミュニケーションにおいて、適応的な言葉を用いながら会話を進めている。従来の研究では、会話において同調が起きることはコミュニケーションの成功とされてきた [1]。同調には、単語や文法構造といった言語の側面 [2] と、韻律などの非言語の側面 [3] がある。本発表では、言語の側面に着目する。

話者はある対象を伝えるとき、相手が用いた言葉を繰り返して用いることがあり、言語学の領域ではこの行為を言語同調と呼んでいる [2]。これにより、話者間で共通理解を作り上げることができる。

一方で、相手と異なる言葉を用いるという非言語同調によって会話を進めていくこともあり、共通カテゴリーの対象の区別が例として挙げられる。例えば、革靴を相手に伝えるとき、帽子と一緒にある場合では「靴」と表現すれば伝わるが、運動靴と一緒にある場合では「靴」と表現すると混同してしまう。そこで、「革靴」という異なる言葉を用いて伝えることができる。

このように、人間同士の会話では、言語同調と非言語同調を使い分けながら効率的に会話を進めている。

1.1 対話エージェントに対する言語同調

対話エージェントに対する言語同調についても研究が行われてきた [4]。インタラクションの相手が人間である場合よりも、エージェントである場合のほうが言語同調することが指摘されている [5, 6]。彼らによると、機械は人間より言語理解能力が低いため、エージェントが認識できるように人間がエージェントと同じ言葉を用いていると解釈されている。

このことから、人間-エージェント間では、非言語同調を必要とする場面においては、言語選択の困難さが懸念される。他の言葉を用いる場合、相手が理解可能な言葉を選ぶことになるが、エージェントに対しては、機械でも認識できる言葉を用いなければならない。

過去の研究では、エージェントに対して言語同調するかどうかという点に着目して検討されてきたが、非言語同調時における単語の種類については十分に検討されていない。言語同調しない場合でも、エージェントが用いた単語に近い単語を用いるという同調行動も十分に考えられる。そこで、共通カテゴリーの対象を区別してエージェントに伝える際に、人間が用いる単語の種類について検討する。具体的には、エージェントが用いた単語との距離に着目する。

*連絡先: 立命館大学大学院文学研究科
〒603-8346 京都府京都市北区等持院北町 56-1
E-mail: lt0610fs@ed.ritsumeikai.ac.jp

1.2 単語選択と感情状態

人間同士のコミュニケーションでは多様な言語表現が用いられているのに対し、エージェントとのコミュニケーションでは限定的な単語選択が求められ、機械でも認識できる単語を心掛けて用いることになる。そこで、エージェントが認識できるように人間は言語同調しようとする。先行研究 [6] では、共通カテゴリーに属さない 2 つの対象のうち、いずれかをエージェントに伝える課題が用いられていたため、言語同調すれば確実に伝わるという状況であった。

しかし、共通カテゴリーに属する 2 つの対象のうち、いずれかを伝えるという非言語同調を必要とする場面においては、エージェントが用いた単語以外で認識できる単語を用いることになる。その際、エージェントが認識できるように言語同調する場合 [5, 6] に比べ、異なる単語を考えることに集中するため、覚醒度が高くなることが考えられる。

そこで、本研究では、非言語同調を必要とする場面における、エージェントに対する単語選択と感情状態の関係を検討する。特に、感情状態としては覚醒度に着目して検討する。その指標のひとつとして心拍が挙げられ、人間は覚醒度が高まると心拍が上昇するとされている [7]。非言語同調を必要とする場面でのエージェントに対する単語選択においては、相手が認識できる単語を熟慮する場合に覚醒度が高まり、同調行動としてエージェントが用いた単語に近い単語を発言することが考えられる。

1.3 本研究の目的と仮説

本研究では、対話エージェントのコミュニケーションにおいて、人間が用いる単語の種類と感情状態の関係を検討する。具体的には、非言語同調を必要とする場面を設定し、人間のエージェントに対する単語選択と覚醒度の関係に着目する。単語選択に関しては、エージェントが用いた単語と人間の用いた単語の距離を指標とし、覚醒度は心拍変動から推定するものとした。

本研究の仮説としては、心拍変動が大きいほどエージェントが発言した単語に近い単語を発言すると予想した。これは、相手に合わせようと心掛けた単語選択であることを意味する。

2 方法

2.1 実験参加者

大学生 12 名 (男性 4 名, 女性 8 名, 平均年齢 19.5 歳) が参加した。ただし、データに欠損があった 3 名は分析の対象外とした。

2.2 実験課題

先行研究 [6] で用いられた課題を改変した。本実験の課題は、(1) 画像選択フェーズと (2) 言語入力フェーズで構成されている。図 1 は課題画面の例を示す。

(1) 画像選択フェーズでは、2 枚の画像が呈示され、エージェントが一方を単語で指定し、実験参加者は正しい画像を選択した。次に、エージェントが正誤についてのフィードバックを返した。

続く (2) 言語入力フェーズでは、直前の (1) で言及された画像と、共通カテゴリーの画像の 2 枚が呈示された。図 1 の場合、(1) ではコップの画像が「食器」と言及され、(2) ではコップとスプーン (食器の共通カテゴリー) の 2 枚が呈示された。そして、(1) で言及された画像 (図 1 ではコップ) が枠で囲まれ、実験参加者はその名称をエージェントに伝えた。実験参加者が単語を発言すると、エージェントは画像を選択するというフィードバックを返した。

(1) 画像選択フェーズ (2) と言語入力フェーズのセットが 30 回繰り返されると、実験課題は終了とした。

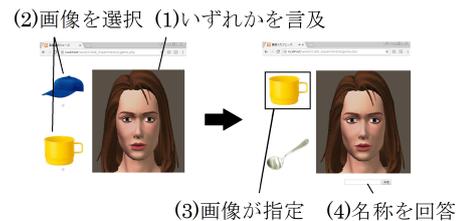


図 1: 課題画面の例 左: 画像選択フェーズ 右: 言語入力フェーズ

2.3 脈波

実験参加者の感情状態を測定するため、MaP1000/2000 シリーズ構成ユニット M-10d の携帯型脈波測定装置 Polypul (PCG) を指先に装着し、サンプリングレート 530Hz で脈波 (PCG) を取得し、心拍を推定した。また、単語選択時での感情状態を調べるため、言語入力フェーズにてエージェントがフィードバックを返す瞬間にマーキングを入れた。

2.4 手続き

図 2 は手続きのフローチャートを示す。実験参加者が実験室に入室し、実験課題の説明を行った。そして、携帯型脈波測定装置を指先に装着し、測定を開始してから実験課題を開始した。課題は音声発話で行い、画像選択フェーズと言語入力のセットが 30 回繰り返されると実験課題は終了となり、装置を外した。

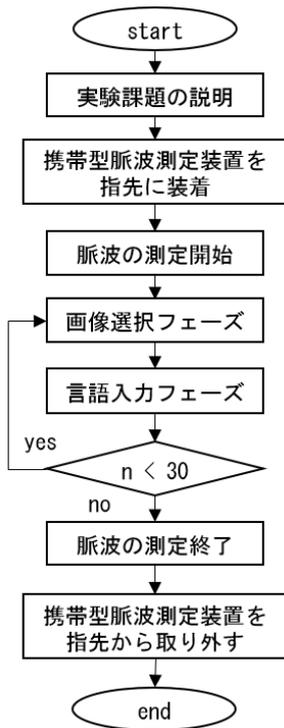


図 2: 手続きのフローチャート

3 結果と考察

3.1 単語選択

画像選択フェーズにてエージェントが発言した単語と、言語入力フェーズにて実験参加者が発言した単語の距離を分析した。分析には分類語彙表¹を用いた。これは、語を意味によって分類したシソーラス (類義語集) である。意味分類方式は、「分類番号」によって分類項目の体系的な位置づけを示す。「類」、「部門」、「中項目」、「分類項目」の4階層のカテゴリーとなっている。そして、「分類項目」の中をさらに分類する「段落番号」、「小段落番号」、「語番号」との計7つの階層を表す番号によって見出し、すなわち単語は一意となる。

ここで、エージェントと実験参加者が発言した見出しについて、「語番号」から共通の親ノードまでのエッジの数を「世代の距離」とし、両者が発言した単語の距離の指標とした。例えば、エージェントが食器 (1.4520 1 1 1)、実験参加者がコップ (1.4520 4 5 1) と発言した場合では、「分類項目」が共通の親ノードであり、「語番号」から遡るエッジ数は3であることから、「食器」と「コップ」の世代の距離は3となる。また、見出しが複数の番号をもつ場合もあるが、「世代の距離」が最短となる番号を採用した。この計算指標に基づき、各参加者の30試行ごとの「世代の距離」を算出した。

¹http://www.ninjal.ac.jp/publication/catalogue/kokken_mado/15/02/

3.2 感情状態

実験参加者の感情状態を測定するために、脈波 (PCG) を取得することによって心拍を推定した。そして、単語選択時における覚醒度を調べるために、マーキングに基づき、言語入力フェーズでの画像指定からエージェントのフィードバックまでの区間における脈波 (PCG) のSDを、心拍変動として算出した。

3.3 単語選択と感情状態

実験参加者の単語選択と感情状態 (覚醒度) の関係を調べるため、30試行における世代の距離と脈波 (PCG) のSDについて、スピアマンの相関係数を算出した。図3は実験参加者ごとの散布図と回帰直線を示す。

本研究の仮説としては、心拍変動が大きいほどエージェントが発言した単語に近い単語を発言すると予想した。すなわち、世代の距離と脈波 (PCG) のSDは負の相関を示すことになるが、分析結果では正と負の群に分かれた。ここでは、2つの群について議論する。

負の相関が示された群では、仮説に基づき、心拍変動が大きいほどエージェントが用いた単語に近い単語を発言していたことになる。これは、エージェントが認識できる単語を意識的に考えることによって覚醒度が高くなり、エージェントが用いた単語に近い単語を発言するという同調行動が現れたことが考えられる。

正の相関が示された群では、予想に反し、心拍変動が大きいほどエージェントが用いた単語から離れた単語を発言していたことになる。これは、多様な単語を熟慮することによって覚醒度が高くなり、エージェントが用いた単語とは離れた単語を発言していたことが考えられる。

世代の距離と脈波 (PCG) のSDの相関が正と負の群に分かれた理由として、個人特性が考えられる。先行研究 [6] では、言語同調のプロセスとして (1) 相手の要因と (2) 言語処理の要因について議論されており、人間よりもエージェントに対して言語同調することから、(1) 相手の要因が有力とされていた。本実験においても、負の相関が示された実験参加者は、エージェントが認識できるように相手と近い単語を心掛けて発言していたと考えられた。一方で、正の相関が示された実験参加者は、覚醒度が高いほど相手とは離れた単語を発言していたことから、多様な単語の熟慮として (2) 言語処理の要因が働いていたことが考えられる。

(1) 相手の要因か (2) 言語処理の要因が単語選択に影響するという点では、個人特性の影響が考えられる。例えば、協調性の高い実験参加者は、相手に合わせた単語選択を行うことが期待され、(1) 相手の要因のほうが単語選択に強く影響することが考えられる。

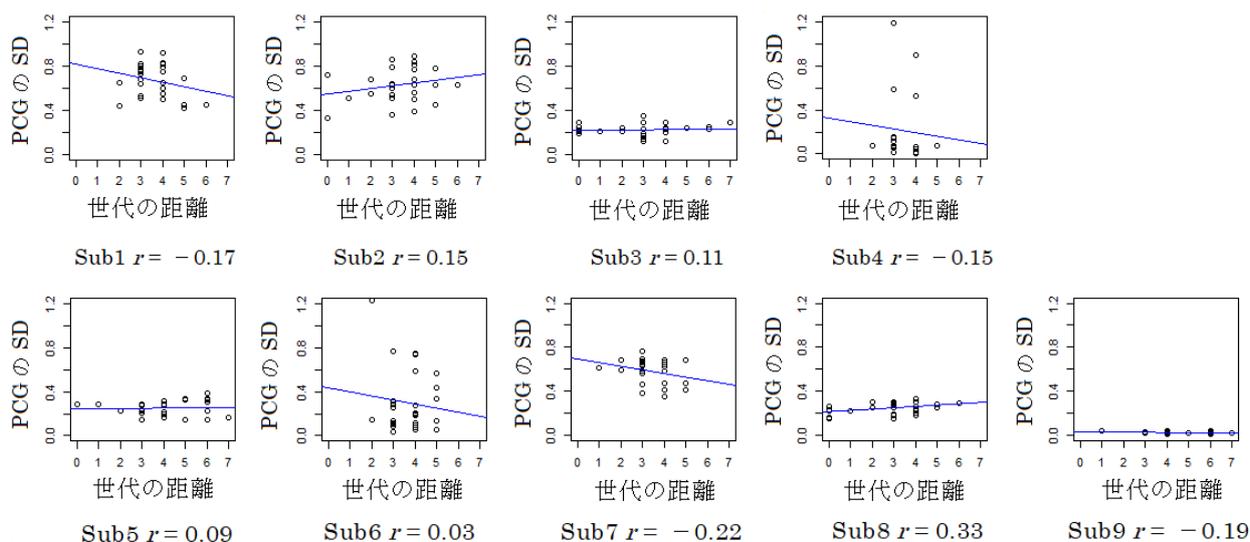


図 3: 実験参加者ごとの世代の距離と脈波 (PCG) の SD の散布図と回帰直線。

4 まとめ

本研究では、非言語同調を必要とする場面において、人間がエージェントに対して用いる単語の種類と感情状態の関係を検討した。人間はエージェントが認識できるように、相手と同じ言葉を心掛けて用いるという知見に基づき、心拍変動が大きいほどエージェントと近い単語を発言すると予想した。共通カテゴリーの対象を区別して伝える課題を行い、脈波の取得から心拍を推定した。実験結果から、心拍変動が大きいほど発言する単語がエージェントと近かった群と離れていた群に分けた。前者は相手の要因、後者は言語処理の要因が単語選択に影響することが示唆された。

本研究で得られた知見は、ユーザの感情状態から単語選択を推定するためのモデル構築に新たな指針をもたらすことが期待される。

謝辞

本研究は JSPS 科研費 16K00219 を受けたものです。ここに記して謝意を表す。

参考文献

[1] M.J. Pickering and S. Garrod, Toward a mechanistic psychology of dialogue, vol.27, Cambridge Univ Press, 2004.

[2] S.E. Brennan and H.H. Clark, “Conceptual pacts and lexical choice in conversation,” *Journal of Experimental Psychology: Learning, Memory, and Cognition*, vol.22, no.6, pp.1482–1493, 1996.

[3] P.S. Jennifer, “On phonetic convergence during conversational interaction,” *The Journal of the Acoustical Society of America*, vol.119, pp.2382–2393, 2006.

[4] S.E. Brennan and H.H. Clark, “Lexical entrainment in spontaneous dialog,” *Journal of Experimental Psychology: Learning, Memory, and Cognition*, vol.22, no.6, pp.1482–1493, 1996.

[5] H.P. Branigan, M.J. Pickering, J. Pearson, J.F. McLean, and C.I. Nass, Syntactic Alignment Between Computers and People: The Role of Belief about Mental States, *Proceedings of the 25th Annual Conference of the Cognitive Science Society*, 2003.

[6] H.P. Branigan, M.J. Pickering, J. Pearson, J.F. McLean, and A. Brown, “The role of beliefs in lexical alignment, evidence from dialogs with humans and computers,” *Cognition*, vol.121, pp.41–57, 2011.

[7] 菅原大地, 荒木友希子, 杉江 征, 覚醒度の異なるポジティブ感情がストレス抑制効果に与える影響, 第 4 巻, 北陸心理学会, 2015.