

対話者の口癖を模倣するロボットが与える印象の調査

Investigating Human Impression towards Robot that Mimics Utterances

安田 元樹 *¹

Motoki Yasuda

田中 文英 *²

Fumihide Tanaka

*¹筑波大学 大学院 システム情報工学研究科

Graduate School of Systems and Information Engineering, University of Tsukuba

*²筑波大学 システム情報系 知能機能工学域

Faculty of Engineering, Information and Systems, University of Tsukuba

Researches demonstrated that the chameleon effect leads to providing a good impression to humans during verbal mimicry. However, it is believed that verbal mimicry not always gives good impression and few study investigated this dual tendency in verbal mimicry which can lead to both positive and negative impression. We consider that favorable words to be a factor expected to not always lead to good impression, then investigated the difference of impression towards between a robot which mimics the utterance of a speaker and a robot which mimics the utterance of a speaker while emphasizing his/her favorable words. Results showed no significant difference in the impression between the two conditions. However, it was suggested that emphasis of his/her favorable words might not lead to providing a good impression to the participants.

1. はじめに

社会的相互作用において人が無意識に他者のしぐさや癖、振る舞いを模倣し、また模倣された人は模倣した人に対して良い印象や好感を持つことをカメレオン効果と呼ぶ [Chartrand 99]. Kulesza らは、言語模倣の先行研究で提起される模倣条件の曖昧性に対して、注意深く条件分けをし、言語模倣による印象の調査をした。実験を行うことで、模倣条件においては、言葉の繰り返しが人の向社会性を向上させるが、それらが繰り返される順序は関係なく、また、非模倣条件においては、応答がない条件は簡単な応答の条件と同じであると報告している [Kulesza 14].

人がエージェントなどによって模倣された際にも模倣効果があるかどうかの研究がされている。Suzuki らは、ハム音を合成することによって人間の音韻的な特徴を模倣するアニメーションキャラクターを用いて、模倣率を変えながら対話した実験参加者に評価アンケートを取ることで、模倣により好印象が与えられたと報告している [Suzuki 03]. しかし、模倣によって常に好印象を与えるとは限らないと我々は考える。例として、相手の動作や言動を全て模倣することや、滑舌の悪い人に対してその人の滑舌の悪さを模倣することなどが挙げられる。これらの模倣は模倣をされた人にとって、模倣されても好印象につながらないデリケートな要素まで模倣をしてしまったからであると考えられる。先行研究においても、Bailenson らによって、模倣をするスクリーンエージェントは、模倣されている人が模倣していることに気が付いた場合、模倣されていることに気がつかない場合と比べて、評価が低かったと報告されている [Bailenson 08]. このように、模倣しても常に好印象につながらない模倣要素があると考えられる。しかし、その模倣要素が何であるかを調査した研究は無い。

そこで本研究は、常に好印象につながらないと思われられる模倣要素を調査することを目的とする。本稿では口癖に着目して、Kulesza らの先行研究の模倣条件 [Kulesza 14] か

ら、話者の発話をそのまま繰り返して模倣する条件を模倣条件、話者の発話をそのまま繰り返して模倣しながら話者の口癖を強調する条件を強調模倣条件とし、この2つの条件間で話者が抱いた印象の違いを調査する。Shimada らの先行研究によって、人と人のコミュニケーションについて研究する際、スクリーンエージェントでは人間と会話する臨場感が少ないため、人間的な外観のアンドロイドを用いた場合と比べて人が抱く印象が変わることが報告されている [Shimada 08]. このことを考慮し、本研究ではロボットを用いて実験を行う。

2. 口癖検出手法

大野らは、話者の発話から形態素を抽出し、全ての形態素から口癖らしさを示す口癖スコアを求め、その値が閾値以上である形態素を口癖として検出するシステムを開発した [大野 10]. 2人の作業者が推測した口癖の積集合を正解データとすると、正解データに対する、システムが検出した口癖と正解データが一致した口癖の総数の割合を再現率とし、また、システムが検出した口癖の総数に対する、システムが検出した口癖と正解データが一致した口癖の総数の割合を適合率とした時、再現率と適合率の調和平均である F 値を計算したところ 37.56%であった。この値は必ずしも高い値では無いが、正解データ作成時の作業者間で再現率、適合率をとり、F 値を計算すると 41.02%であり、ある程度の検出性能を有していると報告されている。

本研究は、この手法を参考にした。以下に、大野らが提案した手法を説明する。

2.1 口癖の特徴と口癖スコア

大野らは、口癖には4つの特徴を備えていると定義し、各特徴に対してスコア $FP_i (1 \leq i \leq 4)$ をそれぞれ定めた。スコアの値が高ければ高いほど、その特徴の度合いが高まる計算式になっている。4つ全てのスコアの積を口癖スコア $FP = FP_1 \times FP_2 \times FP_3 \times FP_4$ とする。ただし、 $w_\alpha (1 \leq \alpha \leq N_w)$ はある形態素を、 $s_\beta (1 \leq \beta \leq N_s)$ は話者を、 $t_\gamma (1 \leq \gamma \leq N_t)$ はトピックを、 N_w は口癖フレーズ候補の総数を、 N_s は話者の総数を、 N_t はトピックの総種類数をそれぞれ表している。

連絡先: 安田 元樹, 筑波大学システム情報工学研究科, 〒305-8573 茨城県つくば市天王台 1-1-1, yasuda@ftl.iit.tsukuba.ac.jp

ある話者 $s_{\beta'}$ に発話された, ある形態素 $w_{\alpha'}$ に対する各スコアの計算式 $FP_2(w_{\alpha'})$ を以下に説明する.

特徴 1: その口癖を発する話者はそれを頻繁に発する

特徴 1 を表すスコア FP_1 は, 全話者に対して, 話者 $s_{\beta'}$ が形態素 $w_{\alpha'}$ を発話した頻度の割合とした. しかしこの場合, 割合だけで算出すると, 1 人の話者だけが発話している場合, その頻度に関係なく, 一定値 ($= 1$) となるため, 頻度との積をとった $tf(w_{\alpha'}, s_{\beta'})$ は, 話者 $s_{\beta'}$ が形態素 $w_{\alpha'}$ を発話した頻度を示す.

$$FP_1(w_{\alpha'}) = tf(w_{\alpha'}, s_{\beta'}) \times \frac{tf(w_{\alpha'}, s_{\beta'})}{\sum_{\beta=1}^{N_s} tf(w_{\alpha'}, s_{\beta})} \quad (1)$$

特徴 2: その口癖を発する話者は偏る

特徴 2 を表すスコア FP_2 は, 形態素列 $w_{\alpha'}$ を発話する話者が誰であるかのエントロピーの逆数とした. 形態素列 $w_{\alpha'}$ を発話する話者が偏ると, 形態素列 $w_{\alpha'}$ を発話する話者が誰であるかのエントロピーが小さくなるためである.

$$\begin{aligned} FP_2(w_{\alpha'}) &= \frac{1}{H(S|w_{\alpha'}) + 1} \\ &= \frac{1}{-\sum_{\beta=1}^{N_s} P(S_{\beta}|w_{\alpha'}) \log P(S_{\beta}|w_{\alpha'}) + 1} \end{aligned} \quad (2)$$

特徴 3: その口癖はトピックに関係なく発せられる

特徴 3 を表すスコア FP_3 は, 形態素列 $w_{\alpha'}$ が発話されたときのトピックが何であるかのエントロピーとした. 形態素列 $w_{\alpha'}$ がトピックに関係なく発話されるのであれば, 形態素列 $w_{\alpha'}$ が発話されたときのトピックが何であるかのエントロピーは大きくなるためである.

$$\begin{aligned} FP_3(w_{\alpha'}) &= H(T|w_{\alpha'}) + 1 \\ &= -\sum_{\gamma=1}^{N_t} P(t_{\gamma}|w_{\alpha'}) \log P(t_{\gamma}|w_{\alpha'}) + 1 \end{aligned} \quad (3)$$

特徴 4: 単位性 (ひとかたまりのフレーズとして見なせる度合い) が強い

特徴 4 を表すスコア FP_4 は, 形態素列 $w_{\alpha'}$ の左側に接続する形態素が何であるかのエントロピーと, 形態素列 $w_{\alpha'}$ の右側に接続する形態素が何であるかのエントロピーの積とした. 左右に接続する形態素のエントロピーが小さいほど, 特定の形態素が接続しやすく単位性が弱いことを意味し, 逆にエントロピーが大きいほど, 様々な形態素が接続しやすいことになり, 強い単位性を持つと考えられるためである. ただし, L を形態素列 $w_{\alpha'}$ の左側に接続する形態素, R を形態素列 $w_{\alpha'}$ の右側に接続する形態素, l_x を形態素列 $w_{\alpha'}$ の左側に接続する x 個目の形態素, r_x を形態素列 $w_{\alpha'}$ の右側に接続する x 個目の形態素とする.

$$\begin{aligned} FP_4(w_{\alpha'}) &= \sqrt{H(L|w_{\alpha'}) \times H(R|w_{\alpha'})} \\ &= \sqrt{-\sum_x P(l_x|w_{\alpha'}) \log P(l_x|w_{\alpha'})} \\ &\quad \times \sqrt{-\sum_x P(r_x|w_{\alpha'}) \log P(r_x|w_{\alpha'})} \end{aligned} \quad (4)$$

2.2 口癖スコア算出

テキストファイルとして保存された実験参加者全員の発話内容を形態素解析エンジン MeCab を用いて形態素解析する. そこで得られた形態素を用いて話者の口癖を検出した.

3. ロボットの説明

本章では, 本研究で用いたロボットについて説明する.

3.1 ロボットの模倣条件

実験で会話するロボットに, 非模倣条件, 模倣条件, 強調模倣条件の 3 条件を設けた. それぞれ条件の内容を以下に示す.

非模倣条件

ロボットは実験参加者に対して質問をする. 実験参加者はその質問に答える. 話者の発話が正しく認識された場合は, ロボットは何も発話をせず, 次の質問に移る.

模倣条件

ロボットは実験参加者に対して質問をする. 実験参加者はその質問に答える. 話者の発話が正しく認識された場合は, ロボットは話者の発話を繰り返し発話し, 次の質問に移る.

強調模倣条件

ロボットは実験参加者に対して質問をする. 実験参加者はその質問に答える. 話者の発話が正しく認識された場合は, ロボットは話者の発話を繰り返しながら話者の口癖を強調して発話し, 次の質問に移る.

3.2 ロボットに実装した機能

本研究では, SoftBank Robotics 社のヒューマノイドロボット NAO を用いた. SoftBank Robotics 社ホームページ上で配布されている公式モジュールの naoqi SDK for python を用いて NAO の制御を行った. 実験を進めていく上で必要な音声認識機能, 発話の模倣機能, 口癖の強調機能を以下のようにしてロボットに実装した. 今回の実験では, ロボット操作者が仕切りの後ろで操作し, 音声認識結果を確認しながら会話を進めた. 会話中にロボットが全く動作しないと話者は不自然に感じると考えられるため, 発話の際に任意の動作をさせる Naoqi API の ALAnimatedSpeech モジュールを用いて, ロボットが発話するとき, 頭や腕などを動作させている.

3.2.1 音声認識機能

GoogleCloudSpeechAPI を用いてストリーミング音声認識をする. 据え置きマイクを用いて音声認識した. ロボットは質問を発話した後に音声認識を始める. 音声認識開始時にロボットは, 目の色が白色から緑色に変化させ, 音声認識終了時に, 目の色が白色に戻ることで, 話者が話し始めるタイミングを分かりやすくしている. 実験参加者はロボットの目の色が変化したことを確認してから, 質問に対して答える.

本稿で実装した音声認識機能では, 音声为正しく認識されないまま会話が進むことを避けるために, 話者が発話途中で音声認識が途切れた場合など, 正しく音声認識が行われなかった場合に, ロボットは, 「聞き取れなかったから, もう一度教えて?」と発話し, 話者に聞き返すこととした. 口癖スコアを算出するときは, 正しく音声認識されたもののみを取り扱う.

3.2.2 発話の模倣機能

話者の発話を音声認識して得られたテキストファイルの内容をそのまま発話させた. また, 会話の流れとして, ロボットが話者の発話をそのまま繰り返すだけで次の質問に移ると不自然

であると考えられるので、話者の発話を繰り返した後に、「なんだね～」と付け足した。

3.2.3 口癖の強調機能

口癖の強調手法は、口癖として検出された形態素を発話する際、前後に間を挿入し、少し高い声で発話させることとした。これは、Naoqi API の ALAnimatedSpeech モジュールで設定できる、声の高さを調整するパラメータを、通常は 100 であるところを 125 とし、少し高い声で発話させ、またその形態素の前後に 50[msec] の間を挿入することで実現している。

それぞれの形態素に対して、あらかじめ用意していたその話者の口癖リストと照らし合わせ、その結果、口癖である形態素がある場合、その形態素をロボットは強調して発話する。模倣条件と同じく、会話の流れとして、ロボットが話者の発話をそのまま繰り返すだけで次の質問に移ると違和感があると考えられるので、発話を繰り返した後に、「なんだね～」と付け足した。

4. 実験

4.1 目的

本実験では、常に好印象につながるとは限らないと予想される模倣要素として口癖に着目し、模倣条件のロボットと会話した時と、強調模倣条件のロボットと会話した時に、実験参加者が抱くロボットの印象の違いを調査することを目的とする。

4.2 実験の手順

本実験では筑波大学田中文英研究室に所属する 20 代の日本入大学生 8 人 (男性 5 人, 女性 3 人) に参加して頂いた。

まず、口癖を模倣するためには、相手の発話から口癖スコアを求めなければならないが、口癖スコアを求めるためにはそれぞれの話者の発話量がある程度必要となる。そこで、実験は発話データ取得フェーズと模倣フェーズに分けて行う。

4.2.1 実験室環境

実験環境を図 1 に示す。実験参加者は椅子に座り、ロボットは机の上に立った状態で会話をする。実験者はロボットの背後に仕切りを設置し、その後ろで PC を操作した。また、記録のために実験中はビデオカメラで会話の様子を撮影した。

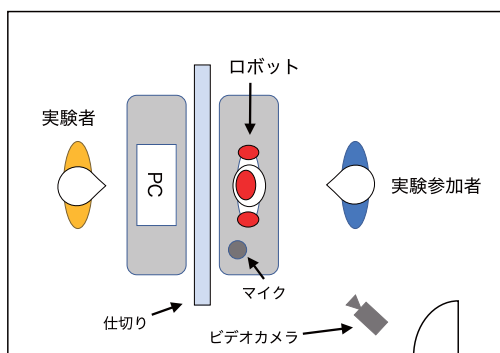


図 1: 実験室環境

4.2.2 発話データ取得フェーズ

発話データ取得フェーズは、非模倣条件のロボットと会話してもらい、それぞれの話者の発話データを取得する。また、話者から多くの発話を取得するために、実験者は実験参加者に対して、「ロボットの質問に答える際、たくさん喋りながら答えてください」とお願いした。

まず、イントロダクションを行ってから会話実験を行い、アンケートやインタビューなどで印象を調べた。

イントロダクション

本実験の全体の流れを説明し、実験者がロボットとの会話を実演し、会話の手順を説明した。

非模倣条件の会話実験

3.4 節で述べた非模倣条件のロボットが、話者に対して好みや思い出などを尋ねる 27 問の質問をした。話者はその質問に答えた。

アンケート&インタビュー

会話実験を行なった後、ロボットに対して抱いた印象を調査するためにアンケートに答えてもらう。アンケートは Bartneck らの The Godspeed Questionnaire Series[Bartneck 09] を用いた。アンケート項目の並びはランダムに並び替えた。実験の感想やアンケート項目では表すことのできない印象を、自由記述に書いてもらい、また、その後のインタビューにて調査した。

4.2.3 模倣フェーズ

次に模倣フェーズでは、模倣条件と強調模倣条件の 2 条件を 4 人ずつに割り振って実験を行った。発話データ取得フェーズで得た発話データから算出した口癖スコアのリストを参照し、実験者がより口癖らしいものが出ていると判断した実験参加者を強調模倣条件に割り振った。

イントロダクション

本実験の全体の流れを説明し、実験者がロボットとの会話を実演し、会話の手順を説明した。

模倣条件の会話実験

3.4 節で述べた模倣条件または強調模倣条件のロボットが、発話データ取得フェーズと同じ 27 問の質問を話者にして、話者はその質問に答える。

アンケート&インタビュー

会話実験を行なった後、ロボットに対して抱いた印象を調査するため、発話データ取得フェーズで用いたのと同じアンケートに答えてもらう。また、実験の感想やアンケート項目では表すことのできない印象を自由記述に書いてもらい、その後のインタビューにて調査した。

5. 結果

模倣条件、強調模倣条件それぞれの条件のロボットと会話した後に実験参加者にロボットに対して抱いた印象をアンケートにて調査した。その結果を図 2 に示す。

6. 考察

模倣条件と強調模倣条件のアンケート評価結果は、図 2 の通りである。グラフから好ましきなどにおける評価の差が大きいことが分かる。そこで、発話データ取得フェーズでは非模倣条件のロボットの印象をアンケートによって評価しており、このアンケートの評価結果から、模倣条件と強調模倣条件における、模倣によるアンケート評価の変化量を調べた。

表 1 の通り、口癖を強調することによって評価が上がるはずの人間らしさはあまり上がらず、また、好ましきの評価がより下がっていることが分かる。このことから、口癖の強調が好印

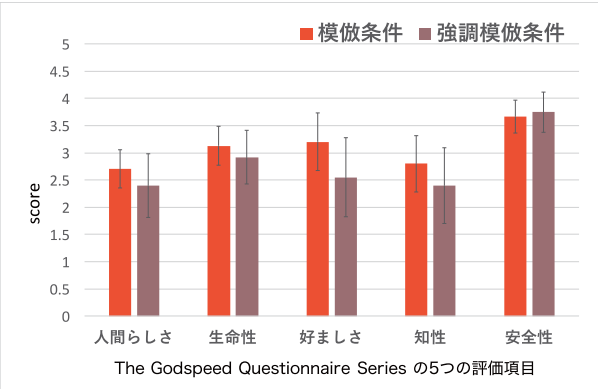


図 2: 模倣条件と強調模倣条件のロボットに対する印象の違い

表 1: 模倣条件と強調模倣条件における、模倣によるアンケート評価の変化量

	人間らしさ	生命性	好ましさ	知性	安全性
模倣条件 (A)	0.55	0.38	-0.1	-0.2	0
強調模倣条件 (B)	0.4	0.38	-0.3	-0.2	-0.83
口癖強調による印象の違い (B-A)	-0.15	0	-0.2	0	-0.83

象につながらない可能性が示唆される。ここで模倣条件も強調模倣条件でも好ましさの評価が下がる傾向にある理由の1つとして自由記述にて、「なんだね～」の単調性に飽きたという記述が複数あったことが考えられる。

インタビューにて、口癖の強調に気がついたかどうか尋ねたところ、強調模倣条件の実験参加者全員が気が付かなかったと回答しており、また、自由記述にて、「イントネーションがおかしくてロボットがふざけているのかと思った」との回答もあった。これらの原因として、今回検出された口癖に助詞が多く、口癖強調によって発話が途切れ途切れになり、また不自然なイントネーションとなったことが考えられる。したがって、今回の口癖強調手法によって実験参加者は口癖を強調するロボットではなく、ただ不自然な発音をするロボットと認識した可能性があり、口癖強調による印象の違いが効果的に出なかった可能性が考えられる。

7. おわりに

本研究は、常に好印象につながるとは限らない模倣要素を口癖に注目して調査した。結果として模倣条件と強調模倣条件に有意差は得られなかったが、口癖の強調は好印象につながらない可能性が示唆された。今後の課題として、口癖強調手法を見直すことで常に好印象につながるとは限らない模倣要素についてさらなる調査を進めていく予定である。

謝辞

本研究は JSPS 科研費 17K19993 の支援を受けて行われた。

参考文献

[Chartrand 99] Chartrand, T. L. and Bargh, J. A.: The chameleon effect: the perception-behavior link and social

interaction., Journal of personality and social psychology, Vol. 76, No. 6, p. 893 (1999)

[Kulesza 14] Kulesza, W., Dolinski, D., Huisman, A., and Majewski, R.: The echo effect: The power of verbal mimicry to influence prosocial behavior, Journal of Language and Social Psychology, Vol. 33, No. 2, pp. 183-201 (2014)

[Suzuki 03] Suzuki, N., Takeuchi, Y., Ishii, K., and Okada, M.: Effects of echoic mimicry using hummed sounds on human-computer interaction, Speech Communication, Vol. 40, No. 4, pp. 559-573 (2003)

[Bailenson 08] Bailenson, J. N., Yee, N., Patel, K., and Beall, A. C.: Detecting digital chameleons, Computers in Human Behavior, Vol. 24, No. 1, pp. 66-87 (2008)

[Shimada 08] Shimada, M., Yamauchi, K., Minato, T., Ishiguro, H., and Itakura, S.: Studying the Influence of the Chameleon Effect on Humans using an Android, in Intelligent Robots and Systems, 2008. IROS 2008. IEEE/RSJ International Conference on, pp. 767-772, IEEE (2008)

[大野 10] 大野誠寛, 松原茂樹:音声対話におけるユーザ発話からの口癖検出. 言語処理学会年次大会発表論文集, Vol. 16, pp. 599-602, (2010).

[Bartneck 09] Bartneck, C., Kulić, D., Croft, E., and Zoghbi, S.: Measurement instruments for the anthropomorphism, animacy, likeability, perceived intelligence, and perceived safety of robots, International journal of social robotics, Vol. 1, No. 1, pp. 71-81 (2009)