

対話アンドロイドの主観的意見の不自然さを軽減する対話戦略

Dialogue Strategy to Reduce Unnaturalness of Subjective Opinions of Conversational Androids

内田 貴久 *1*2

Takahisa Uchida

港 隆史 *2

Takashi Minato

石黒 浩 *1*2

Hiroshi Ishiguro

*1 大阪大学大学院基礎工学研究科

Osaka University

*2 株式会社 国際電気通信基礎技術研究所

Advanced Telecommunications Research Institute International

The goal of this research is to construct a conversational android that evokes users' motivation to talk in non-task-oriented dialogue like chatting. It has been said that stating subjective opinions is effective for motivating people to talk; however, users feel it to be unnatural when a conversational android states its subjective opinions. We hypothesized that lacking the background information as to why and how the android has the subjective opinions leads to the sense of unnaturalness because the users cannot accept its subjective opinions without such information. The experimental results showed that stating the background followed by the subjective opinion was significantly more natural than the opposite case; whereas, the naturalness was not influenced by the order if the speaker is a human. These results suggest that sharing background information in advance is an effective strategy for conversational androids to naturally state their subjective opinions.

1. はじめに

1.1 人間の対話意欲をかき立てるアンドロイド

本研究の目的は、雑談のような非タスク指向型対話において、人間と自然な対話を実現するアンドロイド (図1) の構築である。非タスク指向型対話では、対話を継続することが目的となる。そのため、言語的に適切な発話を生成することを目的とする研究が数多く存在する。その中でも、ELIZA [Weizenbaum 66] は非タスク指向型対話システムとして有名なものの一つであり、発話データベースと複数の返答パターンを用いることによって、継続的な対話を実現している。しかし、システムの発話が言語的につながりを持っているだけでは、人間は対話に飽きてしまい対話が継続されないと考えられる。したがって、対話を継続するためには、言語的に適切な発話を生成するだけでなく、人間の対話意欲を保持するように対話を行う必要がある。

ここで、人間同士の対話において、対話が盛り上げる際には主観的な意見 (例:「この絵はとてもきれいですね」) のやり取りが増加し、客観的な事実 (例:「今日は雨が降りました」) の交換は減少する [Tokuhisa 09] ことが報告されている。つまり、アンドロイドが人間と主観的な意見のやり取りを行うことによって、人間の対話意欲をかき立てることができると考えられる。

1.2 アンドロイドの主観的意見に対する人間の認知

しかし、人間はロボットに対して価値 (良し悪し) に関わる主観的な経験を帰属しにくいことが明かにされている [Sytsma 10]。例えば、人間は、食べ物を食べて美味しいと感じる経験をロボットに帰属しにくいと考えられる。したがって、ロボットがある食べ物の味が美味しかったと主張したとしても、それを聞いた人間は、ロボットが本当に美味しいと思っているように感じにくい。このように、ロボット自身の主観的な意見が、話題によってはロボット自身の意見のように感じられず、ユーザの対話意欲を喚起しない、さらには対話意欲を削



図 1: アンドロイド ERICA (左)

ぐことになる可能性がある。また、Uchida et al. [Uchida 16] は、人間の対話意欲に関して、対話相手が人間の場合とアンドロイドの場合の比較を行っている。人間同士の対話では、対話相手がどれだけ同調するかによって対話意欲が変化するのに対し、アンドロイドが対話相手の場合、同調割合が変化しても対話意欲は変化しないことが報告されている。この結果は、アンドロイドの意見に不自然さを感じ、リアリティを損なったため、同調割合の変化に応じてユーザの対話意欲が変化するまでに至らなかったと考えることができる。

1.3 主観的意見の不自然さに関する仮説

対話ロボットに対して価値に関わる主観的意見を帰属しにくいのはなぜだろうか。これまでの研究で、ロボットと自然なインタラクションを実現する上では、ユーザがロボットに想定する性能 (どういことができるのか) 通りにロボットを振る舞わせることが重要であることが示されている [Kiesler 05]。そこで本研究では、ユーザがアンドロイドが主観的意見を述べることに不自然さを感じるのは、その主観的意見を持つに至る理由となる、アンドロイドの能力に関する背景情報が欠落しているためであるという仮説を立てる。例えば、対話しているアンドロイドがどのようなセンサ情報 (例えば、画像情報、音声情報) を利用できるかをユーザが分からない状態で、アンドロイドの主観的な意見 (例えば、美術館はすばらしい) を聞いたと

連絡先: 内田 貴久, 大阪大学大学院基礎工学研究科, 大阪府豊中市待兼山町 1-3, 06-6850-6362, uchida.takahisa@irl.sys.es.osaka-u.ac.jp

しても、ユーザはその主観的な意見に至る背景を想像できないため、不自然な印象を抱くと考えられる。したがって、アンドロイドの主観的な意見に至る背景をユーザに伝達することにより、アンドロイドがその意見を述べることに對してユーザが感じる不自然さを解消できると考えた。一方、対話相手が人間である場合、人間同士であれば互いの能力（例えば、見る、聞く）をある程度共有していると考えられるので、背景情報の欠如という問題は生じないと考えられる。以上をまとめると、本研究の仮説は下記の通りである。

H1（対話相手がアンドロイドの場合）：アンドロイドが主観的な意見を述べる前に、その背景となる情報を共有することによって、主観的意見を述べることに對する自然さが上がる

H2（対話相手が人間の場合）：人間が主観的な意見を述べることに對する自然さは、その背景となる情報を主観的意見の前に共有するかどうかの影響を受けない

2. 実験

2.1 実験方法

上記の仮説を検証するために、下記の4条件を用意した。

- Android-PRE 条件：アンドロイドが主観的意見を述べる前に、主観的意見に至る背景を共有する
- Android-POST 条件：アンドロイドが主観的意見を述べた後に、主観的意見に至る背景を共有する
- Human-PRE 条件：人間が主観的意見を述べる前に、主観的意見に至る背景を共有する
- Human-POST 条件：人間が主観的意見を述べた後に、主観的意見に至る背景を共有する

アンドロイド条件ではアンドロイド ERICA（図1）を使用した。人間条件では、女性の実験者を採用した。本実験は被験者間比較とし、被験者はアンドロイド/女性実験者と男性実験者の間で対話が行われるビデオを観察する。その後、アンドロイド/女性実験者や対話の印象に関して評価を行う。相手と直接対話せずにビデオを観察する方法を採用した理由は、被験者間で対話内容やアンドロイド/女性実験者の動作を統制するためである。ビデオ刺激を作成するため、アンドロイド/女性実験者と男性実験者が予め用意されたスクリプトに従って対話している様子を撮影した。ただし、男性実験者はビデオには登場せず、男性実験者に関しては発話音声のみが聞こえる。実験に使用したビデオの様子をアンドロイド条件に関して図2(a)に、人間条件に関して図2(b)に示す。

本実験では上述した、主観的意見の前に背景について述べるPRE条件と、主観的意見の後に背景について述べるPOST条件を比較する。単に主観的意見の有無で比較しなかった理由は、条件間で発話量及び発話内容を統制するためである。PRE条件とPOST条件を比較することによって、条件間で情報量を統制した形で仮説を検証することができる。

背景情報によってアンドロイドが主観的意見を述べるものの自然さが変化することを確かめる必要があるため、対話の話題としてアンドロイドが意見を述べるのが自然とも不自然とも判断しにくい話題を採用した。具体的には「美術館」を採用し、主観的意見として「美術館の良さ」を採用した。この場合、アンドロイドが美術館を良いと判断するに至る背景を想像するために必要な情報は、美術館の良さの判断に用いられる客



(a) アンドロイド条件

(b) 人間条件

図2: ビデオの様子

観的な情報（例えば美術館の静けさなど）をアンドロイドが観測できることである。ビデオ刺激の中でアンドロイド/女性実験者は男性実験者に、美術館に関する主観的意見を背景情報とともに語る。男性実験者はアンドロイド/女性実験者に対して相槌や同意のみを示し、その他の発話を行わない。このようにした理由は、条件間で主観的意見と背景情報の順序の違いのみが現れるように統制するためである。PRE条件の対話スクリプトを以下に示す。

（導入）

アンドロイド/女性実験者：この前ね、仕事で美術館に行ってきたの。

男性実験者：へー。

（背景）

アンドロイド/女性実験者：大きな白い壁に決まった間隔で1つずつ絵がかけられてて。

男性実験者：ふーん。

アンドロイド/女性実験者：しーんってしてるんだけど、たまに見てる人の足音がコツコツ響いてるの。

男性実験者：なるほどね。

（主観的意見）

アンドロイド/女性実験者：芸術はあんまり詳しくないけど、すごいなって思う絵はたくさんあったな。

男性実験者：あつ、そうなんだ。

アンドロイド/女性実験者：すごく雰囲気よかったんだよね。デートにはぴったりな場所だなんて思ったの。

男性実験者：ほうほう。

（終わり）

アンドロイド/女性実験者：今まであんまり行かなかったけど、美術館も良いもんだね。

POST条件では、上記のスクリプトにおいて「背景」と「主観的意見」を入れ替えた。

アンドロイドの音声は、HOYA株式会社の音声合成ソフトウェア VOICE TEXT ERICA*1 を使用して作成した。ここで、女性実験者、アンドロイド、男性実験者は常に声のトーンが一定になるようにした。アンドロイドの発話に伴う動作は、Ishi et al. [Ishi 12] や Sakai et al. [Sakai 16] によって開発された、発話の韻律情報に従って唇、頭部、胴体が自動的に動く

*1 <http://voicetext.jp/>

システムを使用して生成した。また、ランダムに瞬きをする動作も付加した。

被験者は4つのビデオ (Android-PRE, Android-POST, Human-PRE, and Human-POST conditions) のうち1つを見て、下記の質問に対して7つの選択肢のうち1つを選択することで評価する。

Q1: アンドロイド/女性が美術館を話題にすることの自然さ
かなり自然, 自然, 少し自然, どちらとも言えない, 少し不自然, 不自然, かなり不自然

Q2: 話しの分かりやすさ
かなり簡単, 簡単, 少し簡単, どちらとも言えない, 少し難しい, 難しい,

2.2 結果

本実験で参加した被験者は下記の通りである。

- Android-PRE 条件: 10名 (男性7名, 女性3名, 平均年齢19.8歳, 標準偏差1.79)
- Android-POST 条件: 10名 (男性5名, 女性5名, 平均年齢21.1歳, 標準偏差1.55)
- Human-PRE 条件: 10名 (男性6名, 女性4名, 平均年齢20.9歳, 標準偏差2.13)
- Human-POST 条件: 10名 (男性7名, 女性3名, 平均年齢22.0歳, 標準偏差1.56)

各質問の回答を数値化するために、7つの選択肢に対して-3~3の得点を付した。その平均得点を図3に示す。Android-PRE条件とAndroid-POST条件間、Human-PRE条件とHuman-POST条件間の平均値の差を検定するために、ウィルコクソンの順位和検定を用いて検定を行ったところ、下記の p 値を得た。

表1: 各質問項目に対する p 値

条件/質問項目	Q1	Q2
Android 条件	0.0551	0.538
Human 条件	0.724	0.12

対話相手がアンドロイドの場合、PRE条件とPOST条件間の差に有意傾向が確認された。さらに、Android-POST条件のみ自然さの平均値が0に近い値となった。すなわち被験者はアンドロイドが主観的意見の前に背景情報を述べなかった場合のみ主観的意見に自然さ(不自然さも)感じていないことを示している。また、対話相手が人間の場合、PRE条件とPOST条件の間に有意差は確認されなかった。これらの結果は、対話相手がアンドロイドの場合、主観的な意見を述べる前に背景情報を共有しておくことが美術館の良さを述べることの自然さを向上させる一方、対話相手が人間の場合は主観的意見を述べる前に背景となる情報を共有せずとも自然さが失われなことを示唆している。以上から、本実験で検証する2つの仮説H1とH2が支持された。つまり、アンドロイドが主観的な意見を述べる場合、その前に背景となる情報を共有することでその主観的な意見を述べることにに対して感じる不自然さを軽減できる可能性が示唆された。

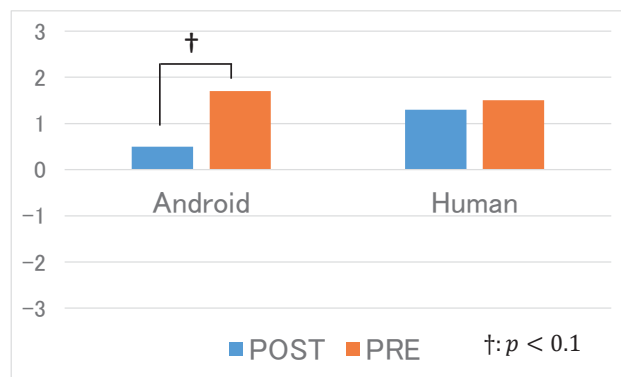


図3: アンドロイド/女性が美術館を話題にすることの自然さ (Q1)

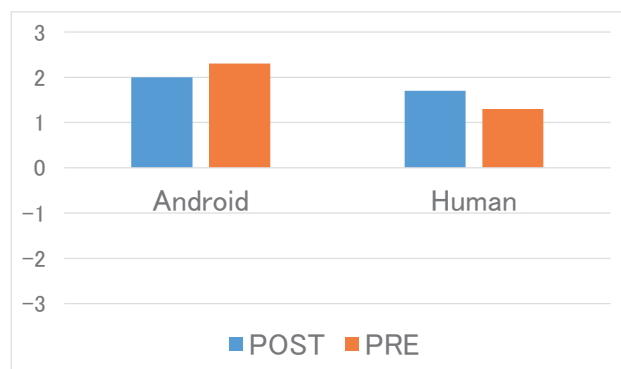


図4: 話しの簡単さ (Q2)

次に、話の分かりやすさに関して、対話相手がアンドロイド・人間の場合ともにPRE条件とPOST条件の間に有意差は確認されなかった(図4)。故に、対話内容自体の分かりやすさは条件間に差はなく、被験者の対話内容に対する理解度は同じであったと考えられる。

3. 考察

対話において、話し手と聞き手は、相互で共有している知識や信念、態度など(共通基盤)を前提としている[Krauss 91]とされている。例えば、人がある食べ物がおいしいという主観的意見を述べた場合、我々はその人が食べ物のおいしさ理解できるだけの知識や信条を持っているがどうか疑問をもつことは少ない。しかし、対話相手がロボットの場合、そのロボットがおいしさを理解できる知識やセンシング能力を持っているかどうかは自明ではないと考えられる。つまり、共通基盤が構築されているかどうかは、対話相手が人間かアンドロイドかどうかによって大きく影響を受けると考えられる。

シミュレーション説[Gallese 98]によると、人間は他者を理解する際に自分を相手の立場に置き換えて理解しようとすると言われている。また、Schank [Schank 90]も、対話において聞き手は話し手の発話に一致する経験を想起しようとすると言っている。したがって、アンドロイドが主観的な意見を述べることに不自然さを感じる理由は、アンドロイドの立場に自身を置くに足る十分な情報が共有されていないためであったと考える

ことができる。

既存の対話システムは、統計的な尤度を計算することで尤もらしい発話を選択して言語的なつながりを保証することを試みているが、その発話に至るプロセスをユーザに想像させてはいない。その結果、ユーザはシステムの主観的な意見に対して不自然さを感じ、引いてはユーザがシステムの発話意図を理解できず、対話破綻が生じるとも考えられる。このことは、ユーザが既存の対話システムに対して対話意欲を減退させている原因とも考えられる。本実験結果が示しているように、主観的な意見の背景を共有することによって、その問題を解決できる。

ここで、不自然さを解消するために、ロボットが対話中に自身の機能に関して逐一説明することで、解決する方法も考えられる。しかし、この方法は説明的なロボットの発話を数多く聞かなければならないことから、ユーザの負荷が大きく対話意欲を損なう可能性がある上に、対話を継続するための高度な自然言語処理機能が前提となる。本章で提案した対話戦略は、対話の滑らかさを保ちつつ、アンドロイドが主観的な意見を述べることを可能にするものであると言える。

本実験結果は、美術館という話題に関して仮説の妥当性を支持するものであった。今後は、美術館以外の話題に関して、この対話戦略が適用可能なかを検証する必要がある。また、本実験では、対話ロボットとして見かけが人間に酷似したアンドロイドを採用したが、同じ話題であっても、他のロボットの場、そのロボットが主観的な意見を述べることに感じる不自然さは異なると考えられる。そこで、提案した対話戦略が適用可能な話題がロボットの種類によってどのように変わるかを明らかにすることが、対話ロボットをデザインする上で必要になる。これらは、今後の課題となる。

4. 結論

本研究では、アンドロイドが主観的な意見を述べることに對してユーザが不自然さを感じる理由は、意見を持った背景知識や状況（共通基盤）がアンドロイドと共有できていないため、どのようにしてその意見を持つに至ったかを想像できないことにあるという仮説を立てて検証を行った。その結果、アンドロイドが話者の場合は共通基盤を形成した上で意見を述べないと、意見を持つことに不自然な印象を与える一方、人間が話者の場合はすでに共通基盤が形成されているため、意見を持つに至る状況等を説明することによる効果は確認されなかった。つまり、アンドロイドが主観的意見を述べる場合、その意見を持つに至る背景である共通基盤の形成が前提として必要であることが示唆された。

5. 謝辞

本研究は JST, ERATO, 石黒共生ヒューマンロボットインタラクティブプロジェクト（グラント番号: JPMJER1401）の助成を受けたものです。

参考文献

- [Gallese 98] Gallese, V. and Goldman, A.: Mirror neurons and the simulation theory of mind-reading, *Trends in cognitive sciences*, Vol. 2, No. 12, pp. 493–501 (1998)
- [Ishi 12] Ishi, C. T., Liu, C., Ishiguro, H., and Hagita, N.: Evaluation of formant-based lip motion generation in tele-operated humanoid robots, in *Proc. of the*

IEEE/RSJ International Conference on Intelligent Robots and Systems, pp. 2377–2382 (2012)

- [Kiesler 05] Kiesler, S.: Fostering common ground in human-robot interaction, in *Robot and Human Interactive Communication, 2005. ROMAN 2005. IEEE International Workshop on*, pp. 729–734 IEEE (2005)
- [Krauss 91] Krauss, R. M. and Fussell, S. R.: Constructing shared communicative environments. (1991)
- [Sakai 16] Sakai, K., Minato, T., Ishi, C. T., and Ishiguro, H.: Speech driven trunk motion generating system based on physical constraint, in *Robot and Human Interactive Communication (RO-MAN), 2016 25th IEEE International Symposium on*, pp. 232–239 IEEE (2016)
- [Schank 90] Schank, R. C.: *Tell me a story: A new look at real and artificial memory*, Scribner New York (1990)
- [Sytsma 10] Sytsma, J. and Machery, E.: Two conceptions of subjective experience, *Philosophical Studies*, Vol. 151, No. 2, pp. 299–327 (2010)
- [Tokuhisa 09] Tokuhisa, R. and Terashima, R.: Relationship between utterances and enthusiasm in non-task-oriented conversational dialogue, in *Proceedings of the 7th SIGdial Workshop on Discourse and Dialogue*, pp. 161–167 Association for Computational Linguistics (2009)
- [Uchida 16] Uchida, T., Minato, T., and Ishiguro, H.: Does a Conversational Robot Need to Have its own Values? A Study of Dialogue Strategy to Enhance People's Motivation to Use Autonomous Conversational Robots, in *the 4th annual International Conference on Human-Agent Interaction (iHAI2016)*, pp. 187–192, Singapore (2016)
- [Weizenbaum 66] Weizenbaum, J.: ELIZA — a computer program for the study of natural language communication between man and machine, *Communications of the ACM*, Vol. 9, No. 1, pp. 36–45 (1966)