

対話型擬人化エージェントの言語的配慮に対する 受容性の6か国異文化比較に関する研究 - クラウドソーシングによる大規模印象調査 3 -

松尾篤¹ 宮本友樹² 片上大輔¹

¹ 東京工芸大学 ² 電気通信大学

e1913122@st.t-kougei.ac.jp miyamoto@uec.ac.jp

katagami@t-kougei.ac.jp

概要

本研究では、対話型擬人化エージェントとの対話において、言語的配慮による受容性の文化差について考察することを目的とする。Brownら[1]によれば、ある発話行為によるフェイス侵害度は、話者同士の社会的距離 (Distance)、相対的な力の差 (Power)、当該文化における負担度 (Rank of impositions) によって求まるとされている。つまり、AIや擬人化システムに対する考え方の違いなども含めて、ある文化圏では適切に選択された発話方略も、別の文化圏では親和性の低い発話になる可能性がある。従って、本研究では擬人化エージェントとの対話において文化による受容性の違いを調査するために、文化圏や言語が異なる日本、アメリカ、中国、フランス、イギリス、オーストラリアの6か国を対象とした大規模な調査を行う。調査の結果、国間による受容性の違いが認められ、国によって性別や年齢による影響の度合いに大きな違いが見られた。

1 はじめに

昨今の人工知能技術の発展により、様々な知的対話システムが開発されている。身近なところではスマートフォンの音声応答アプリケーションである米Apple社のSiriをはじめ、Google社のGoogleアシスタント、NTTドコモ社のしゃべってコンシェルなど、様々な場面において人間とシステム間の対話が行われている。それに伴い、対話システムに関する研究も盛んになっている[2][3][4]。心理的距離や立場を考慮した対話システムの研究として、Brownらによるポライトネス理論[1]の発話方略を用いた研究がある。

心理的距離や立場を考慮した対話システムとして、ポライトネス理論[1]の発話方略を対話システムの発話設計に応用し、その受容性を実験的に検証する取り組みとして、例えばSrinivasanらの研究[2]においては、ロボットがユーザに対して助けを求めるシチュエーションで、対話相手との心理的な距離を縮めるポジティブ・ポライトネス・ストラテジー (PPS) を用いる実験条件と対話相手との心理的な距離を維持するためのネガティブ・ポライトネス・ストラテジー (NPS) を用いる実験条件を設定し印象評価を行った。実験の結果、PPSが有意に高評価であった。このように、従来研究ではエージェントによる言語的配慮が受容性に影響を及ぼすことが示されている。

また、これらの言語的配慮に、文化による受容性の違いを調査した研究がある。先行研究として片上らの研究[3]では、運転支援、家庭内、旅行代理店の3シチュエーションにおいて、PPSとNPSの日米間の印象評価を行った。実験の結果、日米間のエージェントを利用しようと思う意向を示すIntention to Use (ITU) における評価の違いが大きく、発話条件の違いが異文化間の受容性の違いに影響していることが示された。また、片上らの研究[3]を6か国について分析した研究[4]では擬人化エージェントとの対話に対して日本の印象評価が低く、中国の印象評価が高かった。

本稿では、片上らの研究[3]を発展させ、6か国について分析した研究[4]についてより深い分析と考察を行っていく。6か国の異なる文化的背景を持つ実験参加者を対象として以下2つの調査を実施した。

- 調査1：社会に進出するロボットとの対話に関する受容性の質問紙調査

- 調査 2：マルチモーダル情報を備えた対話型擬人化エージェントとの対話を想定した印象評価。

調査の方法として、クラウドソーシングサービスを用いて実験参加者を募り、オンライン上で質問紙調査を実施する。本研究では、エージェントによる言語的配慮に対する受容性の異文化比較研究として、先行研究 [3] の日本語母語話者とアメリカ英語母語話者に加え、イギリス英語母語話者、オーストラリア英語母語話者、フランス語母語話者、中国語母語話者の 6 か国を対象とする。それぞれ募集には SurveyMonkey を用いる。

2 異文化比較調査実験

本研究で行った異文化比較調査実験 [4] の概要について説明する。

2.1 実験目的と概要

本実験の目的は、1 章で述べた調査 1 及び調査 2 を通して対話型擬人化エージェントにおける受容性の文化差について考察することである。以下に各調査の概要を述べる。なお本調査は、東京工芸大学研究倫理委員会の承認（倫理承認番号：2020-06）を得て行った。

2.1.1 調査 1：社会に進出するロボットとの対話に関する受容性の質問紙調査

固有の擬人化エージェントを指定せずに、社会に参画するロボット全体に対して抱えている不安印象を実験参加者に評価してもらう。評価尺度としてロボット不安尺度（Robot Anxiety Scale：RAS）[5] を使用する。RAS はロボット会話能力不安（3 項目）、ロボット行動特性不安（4 項目）、ロボット対話不安（4 項目）の 3 つの下位尺度から構成される。各項目は 6 件法（1：全く不安に思わない、2：ほとんど不安に思わない、3：あまり不安に思わない、4：少し不安に思う、5：かなり不安に思う、6：非常に不安に思う）で評価される。

2.1.2 調査 2：マルチモーダル情報を備えた対話型擬人化エージェントとの対話を想定した印象評価

対話型擬人化エージェントと対話する想定した動画を参加者に視聴してもらい、その印象を評価しても

らう。評価には Technology Acceptance Model (TAM) [6][7] を使用する。TAM はエージェントが人々に受け入れられる・拒否される要因を予測することを目的として提案された 5 段階のリッカート法（1：全くそう思わない、2：あまりそう思わない、3：どちらとも言えない、4：ややそう思う、5：とてもそう思う）で評価する心理尺度である。

2.1.3 調査に関する自由記述

調査 1 と調査 2 が終わった後に、人型エージェントとの対話について感じたこと、本調査について感じたことについて自由記述してもらった。

- 設問 1：「人型エージェントとの対話について感じたことがあれば自由にお書きください」
- 設問 2：「本調査について感じたことがあれば記入してください」

2.2 実験参加者

調査 1 及び調査 2 には、1 か国あたり 300 人、計 1800 人を募集した。その中から、実験に同意した参加者の中で、回答に不備があった場合や募集した国籍と異なっていた場合は分析の対象から除いた結果、アメリカ人参加者 179 人、イギリス人参加者 230 人、オーストラリア人参加者 225 人、フランス人参加者 229 人、日本人参加者 267 人、中国人参加者 242 人、計 1372 人となった。

調査に関する自由記述では、空欄や特に意見がないものを分析の対象から除いた結果、アメリカ人参加者 124 人、イギリス人参加者 156 人、オーストラリア人参加者 145 人、フランス人参加者 151 人、日本人参加者 149 人、中国人参加者 208 人、計 933 人となった。

2.3 実験動画について

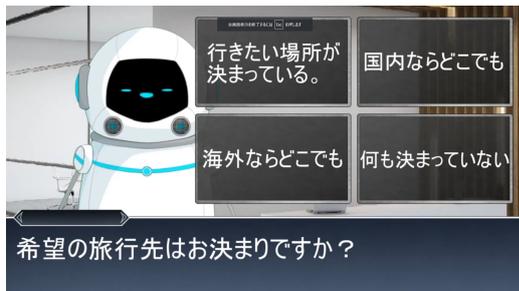
調査 2 で使用する実験動画について、先行研究 [3] の動画をもとに、6 か国の調査向けに作り直した。また、発話内容の英語、フランス語、中国語への翻訳は、翻訳業者に発話意図が変わらないように翻訳を依頼した。動画の内容はマルチモーダル情報を備えた対話型擬人化エージェントとのやり取りの再現動画となる。シチュエーションによる印象の違いに対応するために異なるシチュエーションを用意している。シチュエーションの違いとして挙げられるのは、「タスク指向か非タスク指向か」と「エージェン



(a) 運転支援エージェント



(b) 家庭内エージェント



(c) 旅行代理店エージェント

図 1: 各シチュエーション動画におけるエージェントの様子

トとユーザの関係」の違いである。また、それぞれのシチュエーションに合わせたシナリオを用意し、発話内容に合わせて擬人化エージェントの表情を動かす。シチュエーションごとのエージェントの様子を図 1 に示す。シチュエーションによるタスクの違いだけでなく、エージェントの用いる言葉遣いで印象は変化する。それぞれのシチュエーションに適した言葉遣いを比較するために、エージェントの発話シナリオも NPS 的敬体条件、PPS 条件の 2 通り用意した。エージェントの発話設計のために、社会言語学、語用論の知見であるポライトネス理論 [1] を活用する。

2.4 実験結果

本稿では、調査 2 についての結果を述べる。また、Technology Acceptance Model との相関関係を探る手段として数量化理論 1 類を行った。この理由と

して、国以外に性別や年齢による違いの影響を考慮する必要があるためである。また、説明要因とする項目には質的データが多く含まれており、各々の説明変数をカテゴリー化することによって、カテゴリー別の目的変数に対する影響の方向と大きさを考察することが可能である。

2.4.1 調査 2 について

Technology Acceptance Model についての結果を表 1～表 4 に示す。本研究では、各カテゴリー内容のスコアの最小値と最大値の差をレンジとする。従って国のレンジが最も大きいので、国による違いは 1 番影響度の高い要素である。全体を通して中国のスコアが高いことが示された。つまり、中国は擬人化エージェントとの対話への印象評価が高かったと言える。ITU (Intention to Use: 利用しようとする意向) (表 1) では、日本の評価が低く、次点でアメリカやイギリスの評価が低くなっている。PENJ (Perceived Enjoyment: 使用する際の楽しさの知覚) (表 2) では、日本の評価が低く、次点でアメリカ、イギリス、フランスの評価が低くなっている。SP (Social Presence: 対話の際の社会的存在の有無) (表 3) と Trust (技術の信頼性) (表 4) では、日本、アメリカの評価が低く、次点でイギリス、フランスの評価が低くなっている。また、性別に比べて年齢による影響が大きいことが分かる。

2.5 考察

表 1～表 4 の結果より、調査 2 の結果には国間による違いが大きく影響している。つまり、国の文化によって擬人化エージェントとの対話への印象に違いが示されている。ITU の項目においてはレンジが 1.445 と他の項目より高いことから、利用しようとする意向に関してより国の文化差が出やすいと言える。さらに、ITU の項目ではシチュエーションと発話条件の影響度が逆転している。このことより、他の印象評価と比べて利用しようとする意向においては発話条件による違いがより影響していることが分かる。ただし、発話条件の違いは、全体で見ると国間による違いや性別による違い、年齢層による違いに比べて、影響力は小さいと言える。従って、発話方略による違いよりも、概要や 1 章で述べた Brown らの発話行為によるフェイスの侵害度の要因の 1 つである「相対的な力の差 (Power)」による影響が大きいことが示された。つまり、Brown らのポライト

表 1: ITU (Intention to Use) の分析結果

アイテム	カテゴリー内容	分布数	スコア	レンジ	影響度
国	AU	450	0.020	1.445	1
	CN	484	0.917		
	FR	458	-0.035		
	JP	534	-0.528		
	UK	460	-0.174		
	USA	358	-0.209		
性別	女性	1754	-0.093	0.258	3
	男性	990	0.165		
年齢層	20代	634	-0.139	0.428	2
	30代	1036	0.077		
	40代	466	0.225		
	50代	358	-0.203		
	60代以上	250	-0.092		
シチュエーション	driving	962	0.004	0.014	5
	home	894	-0.009		
	travel	888	0.005		
発話条件	NPS	1372	-0.009	0.018	4
	PPS	1372	0.009		

注：重相関係数は 0.405

表 2: PENJ (Perceived Enjoyment) の分析結果

アイテム	カテゴリー内容	分布数	スコア	レンジ	影響度
国	AU	450	0.032	0.879	1
	CN	484	0.525		
	FR	458	-0.046		
	JP	534	-0.355		
	UK	460	-0.068		
	USA	358	-0.075		
性別	女性	1754	-0.044	0.123	3
	男性	990	0.079		
年齢層	20代	634	-0.077	0.235	2
	30代	1036	0.058		
	40代	466	0.111		
	50代	358	-0.124		
	60代以上	250	-0.072		
シチュエーション	driving	962	-0.000	0.075	4
	home	894	-0.037		
	travel	888	0.038		
発話条件	NPS	1372	-0.002	0.003	5
	PPS	1372	0.002		

注：重相関係数は 0.350

ネス理論における言語的配慮の発話方略である NPS と PPS については、発話方略の選択ミスによる相手への印象の変化は比較的少なく、相手との関係性による違いが大きいことから、相手との関係性によって相手への印象への影響がある程度決定していることが示唆される。

レンジの値に着目すると、国に対するレンジの大きさに対し、それ以外の要因のレンジが圧倒的に少ない。つまり、文化による受容性の違いを生む要因として、性別と年齢だけではなく国ごとの違いの影響が大きいと推測する。

表 3: SP (Social Presence) の分析結果

アイテム	カテゴリー内容	分布数	スコア	レンジ	影響度
国	AU	450	-0.013	0.805	1
	CN	484	0.591		
	FR	458	-0.107		
	JP	534	-0.214		
	UK	460	-0.107		
	USA	358	-0.189		
性別	女性	1754	-0.051	0.141	3
	男性	990	0.090		
年齢層	20代	634	-0.033	0.282	2
	30代	1036	0.070		
	40代	466	0.105		
	50代	358	-0.177		
	60代以上	250	-0.150		
シチュエーション	driving	962	-0.054	0.120	4
	home	894	0.067		
	travel	888	-0.009		
発話条件	NPS	1372	-0.006	0.013	5
	PPS	1372	0.006		

注：重相関係数は 0.354

表 4: Trust の分析結果

アイテム	カテゴリー内容	分布数	スコア	レンジ	影響度
国	AU	450	-0.011	0.920	1
	CN	484	0.656		
	FR	458	-0.095		
	JP	534	-0.264		
	UK	460	-0.103		
	USA	358	-0.226		
性別	女性	1754	-0.069	0.192	3
	男性	990	0.122		
年齢層	20代	634	-0.118	0.425	2
	30代	1036	0.124		
	40代	466	0.138		
	50代	358	-0.287		
	60代以上	250	-0.060		
シチュエーション	driving	962	0.068	0.162	4
	home	894	-0.094		
	travel	888	0.021		
発話条件	NPS	1372	0.016	0.033	5
	PPS	1372	-0.016		

注：重相関係数は 0.340

3 おわりに

本研究では擬人化エージェントの対話において、国ごとの文化による受容性の違いを調査し、明確な違いが示された。また、その違いを生む要因として性差、年齢、発話条件、シチュエーションなどについて検討した。その結果、国ごとに性別や年齢による受容性の違いが見られた。擬人化エージェントとの対話の受容性を高めるにあたって、性別や年齢など、国ごとの文化による受容性の違いになる要因を探ることが今後の研究の課題になるだろう。

謝辞

本研究は科研費（課題番号：22H04869）の支援を受けました。記して感謝いたします。

参考文献

- [1] Penelope Brown, Stephen C Levinson, and Stephen C Levinson. **Politeness: Some universals in language usage**, Vol. 4. Cambridge university press, , 1987.
- [2] Vasant Srinivasan and Leila Takayama. Help me please: Robot politeness strategies for soliciting help from humans. In **Proceedings of the 2016 CHI conference on human factors in computing systems**, pp. 4945–4955, , 2016.
- [3] 片上大輔, 山本隆太郎, 宮本友樹, 宇佐美まゆみ. 対話型擬人化エージェントの言語的配慮に対する受容性の6か国異文化比較に関する研究 -クラウドソーシングによる大規模印象調査-. HAI シンポジウム 2021, 03 2021.
- [4] 松尾篤, 宮本友樹, 片上大輔, 宇佐美まゆみ. 対話型擬人化エージェントの言語的配慮に対する受容性の6か国異文化比較に関する研究 -クラウドソーシングによる大規模印象調査 2 -. ファジィシステムシンポジウム講演論文集, Vol. 38, pp. 244–249, 09 2022.
- [5] Tatsuya Nomura, Tomohiro Suzuki, Takayuki Kanda, and Kensuke Kato. Measurement of anxiety toward robots. In **ROMAN 2006-The 15th IEEE International Symposium on Robot and Human Interactive Communication**, pp. 372–377. IEEE, , 2006.
- [6] Fred D Davis. Perceived usefulness, perceived ease of use, and user acceptance of information technology. **MIS quarterly**, pp. 319–340, 1989.
- [7] Marcel Heerink, Ben Kröse, Vanessa Evers, and Bob Wielinga. Assessing acceptance of assistive social agent technology by older adults: the almere model. **International journal of social robotics**, Vol. 2, No. 4, pp. 361–375, 2010.

4 付録

表 5: RAS（ロボット不安尺度）[5] より抜粋した下位尺度と質問項目

RAS-S1：ロボット会話能力不安
ロボットが会話中に的外れなことを話すのではないか。 ロボットとの会話は融通が利かないのではないか。 ロボットは難しい話が理解できないのではないか。
RAS-S2：ロボット行動特性不安
ロボットがどのような動きをするのか。 ロボットが何をしてくるのか。 ロボットがどれくらいの力を持っているのか。 ロボットがどのくらいの速さで動くのか。
RAS-S3：ロボット対話不安
ロボットにどう話しかけたらいいのか。 ロボットから話しかけられた時にどう答えたらいいのか。 ロボットに自分のした話の内容が理解されているのか。 ロボットから話しかけられた内容が自分には理解できないのではないか。

表 6: TAM[6][7] より抜粋した下位尺度と質問項目

利用しようとする意向（Intention to Use：ITU）
近い将来、私はこのロボットを使うだろうと思う。 近い将来、私はこのロボットを使うことになると思う。 近い将来、私はこのロボットを使うつもりである。
使用する際の楽しさの知覚 （Perceived Enjoyment：PENJ）
このロボットに話しかけられるのは楽しい。 このロボットと一緒に何かをするのは楽しい。 このロボットは楽しいと思う。 このロボットは魅力的だと思う。 このロボットはつまらないと思う。
対話の際の社会的存在の有無 （Social Presence：SP）
このロボットと対話していた時、本当の人間と話しているかのように感じた。 ときどき、このロボットが本当に私の方を見ているかのように感じた。 このロボットが生き物であると想像できる。 このロボットが実在の人間ではないと思う。 ときどき、このロボットが本音を出しているように見える。
技術の信頼性（Trust）
このロボットがアドバイスをくれたら、私はこのロボットを信頼するだろう。 私はこのロボットがくれたアドバイスに従うだろう。