

対話形式が時間経過に伴って敬体から常体へと変化する 音声アシスタントがユーザに与える印象の調査

How changing the dialog format affects the user experience of a voice assistant

服部 健心^{1*} 松村 耕平¹ 岡藤 勇希^{1,2}
Kenshin Hattori¹ Kohei Matsumura¹ Yuki Okafuji^{1,2}

¹ 立命館大学

¹ Ritsumeikan University

² 株式会社サイバーエージェント

² CyberAgent, Inc.

Abstract: 音声アシスタントに代表される音声によるユーザインタフェースが普及してきている。本論文では、日本語対話の親密度に応じて対話形式を敬体から常体へと変化させる特徴を導入した音声アシスタントを提案する。これにより、音声アシスタントの印象を向上できる可能性がある。ユーザが音声アシスタントを用いてタスクを実施する際、音声アシスタントの対話形式を敬体から常体へと変化させたときのユーザの印象を実験から調査した。ユーザ評価からは従来法と提案法の印象に違いは見られなかったが、インタビューからは提案法が選好されることが確認された。

1 はじめに

スマートスピーカーなどの音声アシスタントは音声ユーザインタフェース (Voice User Interface, VUI) の形態の一つとして普及してきている。しかし、音声アシスタントの使用率は依然高いものではない。スマートスピーカーについては、MMD 研究所 (Mobile Marketing Data Labo.) が 2021 年 10 月にスマートフォンを所有している 20 歳から 69 歳の男女に対して行ったスマートスピーカーに関する調査 [1] で、スマートスピーカーの所有率は 21.6 % と報告されている。音声アシスタントがよりユーザと自然に対話を行うことは、ユーザの満足度の向上や音声インタフェースの更なる普及に貢献すると考えられる。

新たな音声アシスタントの提案として、スマートスピーカーの話し方 (対話形式) に注目した研究が行われている。現在、一般に普及しているスマートスピーカーの対話形式は、時間や場所、ユーザの違い、時間経過によって異なることはなくほぼ一定である。しかし、日本語は「デス・マス体」や「ダ体」等のスピーチスタイルが文末形式に組み込まれている数少ない言語であり、相手との関係や場面によってこれらを使い分け、また、同じ相手でもスピーチスタイルの切り替

えを行う [2]。例えば、聞き手が目上の人や初対面の人に対しては「デス・マス体」を使用した敬語などの丁寧な言葉で話し、この話し方により相手に信頼性があるという印象を与える。逆に、対等な関係の人や目下の人、何度も会話をしたことがある人に対しては「ダ体」を使用した普通語などの砕けた言葉で話し、相手に親しみやすいや楽しいといった印象を与える。これはスピーチスタイルシフト [2] やスピーチレベルシフト [3] と呼ばれるものである (以下スピーチスタイルシフトと呼ぶ)。日本語での対話ではこのスピーチスタイルシフトがよく起こり、対話の特徴といえる [2]。

このスピーチスタイルシフトを考慮した対話形式を導入したスマートスピーカーは、これまでの敬語のみの対話形式ではユーザに与えられなかった親しみやすさや楽しさの印象を与え、よりユーザと自然に対話することできる可能性がある。これは音声アシスタントの機能として検討の余地がある。ここで、スピーチスタイルシフトを考慮した対話形式をスマートスピーカーに導入した場合に、ユーザにどのような印象を与えるのかを検証する必要がある。2018 年に大内らは、スピーチスタイルシフトの、「デス・マス体」や「ダ体」などの文末形式を相手との関係や場面によって使い分けるところに注目した研究 [4] を行っている。しかし、スピーチスタイルシフトの同じ相手でも文末形式の切り替えを行う部分に注目した研究はまだされておらず、ユーザへの影響を検証する必要がある。

*連絡先: 立命館大学情報理工学部情報理工学科
〒525-0058 滋賀県草津市野路東1丁目1-1
E-mail: is0519si@ed.ritsumeik.ac.jp

これらの背景を踏まえ本論文の研究目的は、スマートスピーカーが同じユーザに対する発話の文末形式を変化させた場合にユーザにどのような印象を与えるのかを検証することで、スマートスピーカーが日本でより普及し長い期間使用されるために実装するべきである新たな対話形式を提案するというものである。

2 関連研究

関連研究として、大内らの敬体と常体の使い分けに注目した研究 [4] と、申の時間軸での敬体使用率についての研究 [2] についてまとめる。

大内らの研究 [4] では、スピーチスタイルシフトの、「デス・マス体」や「ダ体」などの文末形式を相手との関係や場面によって使い分けるところに注目し、スマートスピーカーの対話形式によるユーザの影響を調査する実験を行った。この実験では円滑な人間関係を確立・維持するための言語行動を理論化したポライトネス理論 [5] に基づき、「デス・マス体」を用いる文末形式を敬意を示す敬体、「ダ体」を用いる文末形式を仲間内を示す常体とした。また、終助詞である「よ」、「ね」、「よね」、「さ」は敬体と常体のどちらの発話にも違和感を与えず、相性が良いため用いた。これら 2 条件の対話形式をとるスマートスピーカーを用いて、ユーザがスマートスピーカーに対してどのような印象を受けるのかを比較した。結果として、敬体は親切な印象をユーザに与え、「常体に比べて情報に信頼性がある」とユーザに感じさせていたが、「機械的で感情がなさそう」という印象も与えることがわかった。一方で、常体は人間的、親しみやすさ、楽しいといった印象をユーザに与え、「敬体に比べて友好的である」とユーザに感じさせていたが、「初対面では馴れ馴れしく、言葉が軽く感じる」という印象も与えることが示された。

申の研究 [2] では、日本語におけるスピーチスタイルシフトの生起要因について、韓国語との比較を通し、日本語のスピーチスタイルシフトが「どのように」起こるのかを考察した。この実験では 2 つの言語の比較を行うために、初対面の 2 人に 10 分程度の会話を数日間の日時を隔て 3 回の対話を行わせていた。話す内容は「夏休みの計画」などの話題を提示はしたが、それに縛られず自由に話すように指示をしていた。そして、1 回の対話内での文末形式使用率の変化の比較と、3 回の対話を課されるの連れて文末形式使用率の変化の比較を行なった。結果として、日本語対話の場合、3 回の対話間では文末形式が敬体から常体に変化することが示された。また、1 回の対話内でも時間経過で敬体から常体への変化が存在し、さらに相手のスタイル変化に合わせる「同調」が存在することが示された。

3 調査実験

本論文では、スマートスピーカーとユーザの 1 回の対話内において、スマートスピーカーの対話形式が時間経過で敬体から常体に変化した場合、ユーザはどのような印象を受けるのかについて調査を行う。敬体と常体の設計は、大内らの研究 [4] で使用されていた「デス・マス体」を用いる文末形式を敬意を示す敬体、「ダ体」を用いる文末形式を仲間内を示す常体とし、1 回の対話時間は申の研究 [2] から 10 分程度とした。また、大内らの研究 [4] で示された、敬体と常体のそれぞれのユーザに与える印象から以下のように仮説を立てた。

仮説 1 回の対話内でスマートスピーカーが、スピーチスタイルシフトを考慮した敬体から常体に変化させる人間らしい対話形式をとることでユーザに与える印象は、敬体の親切な印象と情報への信頼性、常体の人間的、親しみやすい、楽しいという印象、その両方を与えることができる。

スマートスピーカーの対話形式の違いによって人々がどのような印象を受けるのかを調査するために、実験参加者とスマートスピーカーが対話をする被験者間実験を行う。スマートスピーカーの対話形式は、敬語で話す敬体条件、普通体で話す常体条件、敬語から普通体に時間経過で遷移する変化条件の 3 条件で比較する。

3.1 実験手順

本実験では、まず参加者に実験の内容説明と実験でのデータの記録及び実験参加に関する同意確認を行う。その後、参加者はスマートスピーカーの対話形式が敬体条件、常体条件、変化条件のうち一つでタスクを行う。各条件の参加者への割当ては、敬体条件、常体条件、変化条件を一つのセットとして順番をランダムに入れ替えたセットを数セット用意し、それらを参加者の実験参加順に割当てた。タスク終了後、参加者はアンケートに回答し、全て回答することで実験終了となる。

本実験の参加者は大学生・大学院生 19 名 (男性 16 人、女性 3 人) で平均年齢は 21.74 歳である。参加者の選定条件は、参加者とスマートスピーカーの対話は日本語で行うため、日本語での流暢な対話が可能であることとした。参加者 19 名のうち、過去にスマートスピーカーを使用した経験があると回答したのは 8 名であった。参加者それぞれの年齢、性別、実験条件、スマートスピーカーの使用経験を参加者情報として表 1 に示す。スマートスピーカーの使用経験については、使用経験がある人に＊を付けるものとする。実験での拘束

時間は 30 分程度である。参加者には謝礼として 1000 円の謝金を支払った。

表 1: 参加者情報

参加者	年齢 [歳]	性別	実験条件	使用経験
P1	21	男性	敬体条件	
P2	22	男性	敬体条件	
P3	24	男性	敬体条件	
P4	21	男性	敬体条件	
P5	23	男性	敬体条件	*
P6	22	男性	敬体条件	*
P7	21	男性	敬体条件	
P8	20	男性	変化条件	
P9	24	男性	変化条件	
P10	22	女性	変化条件	*
P11	21	女性	変化条件	*
P12	22	男性	変化条件	
P13	22	男性	変化条件	*
P14	22	男性	常体条件	
P15	22	女性	常体条件	*
P16	21	女性	常体条件	*
P17	22	男性	常体条件	
P18	21	男性	常体条件	
P19	20	男性	常体条件	*

3.2 実験方法

実験で使用するスマートスピーカーを図 1 に示す。これは Amazon Echo を分解し Bluetooth 接続ができる機器をコードで取り付け、遠隔で音声を出力することができるように改造したものである。Amazon Echo に本来搭載されている音声アシスタントの Alexa は使用しなかった。



図 1: 実験で使したスマートスピーカー

音声は Voice Peak を用いて女性の声で作成した。Voice Peak は株式会社 AHS が提供している、AI 音声合成技術を用いて打ち込んだ文字から音声を作成することができる読み上げソフトである。男性や女性、音の速さやピッチ、感情などの項目を調整できる。

本実験でのタスクは、大内らの研究 [4] や Chaves らの研究 [6] を参考に「スマートスピーカーにおすすめの旅行先を尋ね、興味のある旅行先の情報について聞いていき、行きたい旅行先を決める」ために対話をするというものである。タスクは 10 分程度行い、3 条件全てにおいてタスク開始から 5 分程度のタイミングで「面白い情報はありましたか?」という対話を挟ませた。変化条件ではこのタイミングで敬体から常体へと変化させることにした。

参加者とスマートスピーカーの対話は、参加者の話す内容に対するスマートスピーカーの回答を実験者が選択することで対話を実現させる Wizard of Oz 法 [7] を用いて行った。実験システムを図 2 として示す。実験者は Amazon Echo に取り付けられた Bluetooth 接続ができる機器と実験で使用する音声保存されたパソコンを接続し、参加者から操作しているところが見えない位置であり、参加者の発した言葉が聞き取れる位置からスマートスピーカーを操作した。また、Bluetooth 接続ができる機器は参加者から見えない位置に設置した。そして、スマートスピーカーからの「どこの都道府県にご興味がありますか?」などの問いかけに対して参加者が「京都」や「京都でお願いします」と話した場合、「京都ですね、観光地と料理のどちらにご興味がありますか?」と参加者の話した内容の復唱と次の問いかけを同時に回答するようにした。また、参加者が「別の情報を知りたい」と話した場合は、「わかりました。では同じ都道府県の料理の情報にご興味がありますか?」や「別の都道府県の情報にはご興味がありますか?」などと回答し、その情報の音声を再生した。参加者の話す内容に対する回答として適切なものがない場合は、「すみません、選択肢にないためお答えできません。」という内容を選択し回答した。

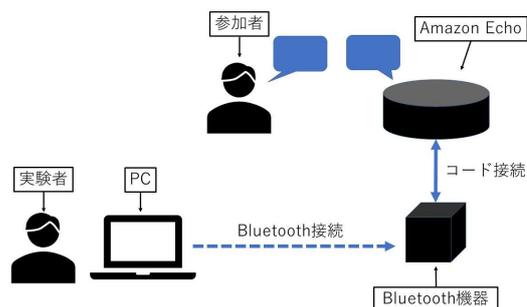


図 2: 実験システム

タスクの具体的な内容としては、まず旅行先の選択肢として「北海道」、「愛知」、「京都」を提示する。次にカテゴリーとして「観光地」と「料理」を提示する。

その後さらにカテゴリとして、観光地は「文化(寺院, 博物館など)」と「自然(自然, 水族館など)」, 料理は「ご当地の定番料理」と「スマートスピーカーがおすすめする料理」を提示する. 最後に, それぞれのカテゴリに割当てた観光地名や料理名を提示し, 選択された観光地名や料理名の情報を読み上げるように設定した. また, タスクを繰り返し行えるようにするため, 一つ前の選択肢に戻る事ができる設定にした. 参加者は, 提示された選択肢から興味のあるものを選択してスマートスピーカーに話しかけ, 最後の観光地や料理の情報を聞くことを繰り返し行う. これらの音声を敬体条件と常体条件の両方作成することで, 3条件を全て行えるようにした. 実際に行われた敬体条件の対話例を表2, 常体条件を表3, 変化条件を表4にそれぞれ示す.

10分程度のタスク終了後, すぐにアンケート用紙を参加者に渡し, スマートスピーカーの印象について回答させた.

表 2: 敬体条件の対話例

話者	内容
スマートスピーカー	今のおすすめの国内の旅行先は, 北海道, 愛知, 京都, 福岡です. どの県の情報を知りたいですか?
参加者	京都で.
スマートスピーカー	京都ですね. 観光地と料理のどちらにご興味がありますか?
参加者	観光地をお願いします.
スマートスピーカー	観光地ですね. 観光地のジャンルは, 自然, 文化のどちらにご興味がありますか?
参加者	自然をお願いします.
スマートスピーカー	自然ですね. 自然でのおすすめは, 天橋立, 渡月橋, 竹林の小径, 花見小路通, 哲学の道となっております. どちらにご興味がありますか?
参加者	竹林の小径で
スマートスピーカー	竹林の小径ですね. 竹林の小径は, 京都嵐山にある手入れされた竹林が約 400m も続く道です. 平安時代は貴族の別荘地だったと言われており, 木漏れ日に映える緑が美しい竹林の小径は, 優雅な雰囲気漂っています. 歩いて楽しむほか, 人力車に乗ってゆっくりと楽しむこともできます.

表 3: 常体条件の対話例

話者	内容
スマートスピーカー	今のおすすめの国内の旅行先は, 北海道, 愛知, 京都, 福岡だね. どの県の情報を知りたい?
参加者	愛知で.
スマートスピーカー	愛知だね. 観光地と料理のどちらに興味がある?
参加者	料理で.
スマートスピーカー	料理だね. 料理のジャンルは, ご当地の定番料理と私のおすすめ料理のどちらに興味がある?
参加者	おすすめ料理でお願い.
スマートスピーカー	おすすめ料理だね. おすすめ料理は, ひつまぶし, 小倉トースト, きしめんとなっているよ. どれに興味がある?.
参加者	やっぱり興味ない.
スマートスピーカー	わかった. ならご当地の定番料理に興味はある?
参加者	ないかな.
スマートスピーカー	わかった. なら同じ都道府県の観光地の情報には興味ある?
参加者	ある, お願い.
スマートスピーカー	わかった観光地だね. 観光地のジャンルは, 自然, 文化のどちらに興味がある?

表 4: 変化条件の対話例

話者	内容
スマートスピーカー	観光地ですね. 観光地のジャンルは, 自然, 文化のどちらにご興味がありますか?
参加者	文化をお願いします.
スマートスピーカー	文化ですね. 文化でのおすすめは, 北海道庁旧本庁舎札幌時計台, 小樽運河, 白い恋人パーク, 北海道神宮となっております. どちらにご興味がありますか?
参加者	小樽運河をお願いします.
スマートスピーカー	小樽運河ですね. 小樽運河は, 北海道開拓の玄関口として, 大正 12 年に完成された運河です. 海岸の沖合を埋め立てて造られた運河で, 緩やかに湾曲した運河となっている特徴があります. 夕暮れ時には, 運河沿いの石造倉庫群がガス灯でライトアップされて, とてもきれいな風景を見ることができます!!
(5 分後)	
スマートスピーカー	面白そうな情報はありましたか?
参加者	あったよ.
スマートスピーカー	それは良かった! もっとたくさん探そう! 文化の別の情報には興味はある?
参加者	ある.
スマートスピーカー	文化だね. 文化でのおすすめは, 北海道庁旧本庁舎, 札幌時計台, 小樽運河, 白い恋人パーク, 北海道神宮となっているよ. どこに興味がある?

3.3 評価方法

アンケート評価は, Q1-Q24 は人の主観評価として使用される Godspeed Questionnaire [8] と, Q25-Q31 は米津の研究のロボットと対面しての会話印象評価 [9] を用いて行う. 各項目は 5 段階で評価を行い, Q1 「偽物のような - 自然な」であれば 1 であるほど参加者はスマートスピーカーを「偽物のような」と感じる. アンケートは全て紙で作成し, アンケート用紙に書かれた数字に丸を記入する形で回答してもらった.

本研究では, 対応のない 3 つ以上のグループ間の差の有無を調べる時に使われる, データごとに順位を付けを行うことで比較を行うノンパラメトリック検定の一つの Kruskal-Wallis 検定を用いて, Godspeed Questionnaire は 5 つの下位尺度 (擬人化, 生命性, 好ましさ, 知性の知覚, 安全性の知覚) ごとに, 米津のロボットの会話印象評価は質問項目ごとに, 敬体条件, 常体条件, 変化条件の比較をする. 質問項目を表 5 で示す. また, 「スマートスピーカーと話してみよう感じたか」と「スマートスピーカーを今後も使用したいと思ったか」という 2 つの自由記述の質問を行う.

3.4 結果

敬体条件, 変化条件, 常体条件のそれぞれで Godspeed Questionnaire の下位尺度ごとの平均値と標準誤差の結果を図 3, 米津のロボットの会話印象評価の質問項目ごとの平均値と標準誤差の結果を図 4 として示す. 図の縦の値はアンケートの印象評価の数字であり,

表 5: 質問項目

質問番号	質問項目
Q1	偽物のような - 自然な
Q2	機械的 - 人間的
Q3	意識を持たない - 意識を持っている
Q4	人工的 - 生物的
Q5	ぎこちない動き - 洗練された動き
Q6	死んでいる - 生きている
Q7	活気のない - 生き生きとした
Q8	機械的な - 有機的な
Q9	人工的な - 生物的な
Q10	不活発な - 対話的な
Q11	無関心な - 反応のある
Q12	嫌い - 好き
Q13	親しみにくい - 親しみやすい
Q14	不親切な - 親切な
Q15	不愉快な - 愉快的な
Q16	ひどい - 良い
Q17	無能な - 有能な
Q18	無知な - 物知りな
Q19	無責任な - 責任のある
Q20	知的でない - 知的な
Q21	愚かな - 賢明な
Q22	不安な - 落ち着いた
Q23	冷静な - 動揺している
Q24	平穏な - 驚いた
Q25	楽しくない - 楽しい
Q26	発言をためらう - ためらわない
Q27	相手の様子が浮かばない - 浮かぶ
Q28	内容が伝わらない - 伝わる
Q29	難しい - 簡単
Q30	相手が自分に同意していないように感じる - 同意しているように感じる
Q31	気分が悪い - 気分が良い

青色の棒グラフは敬体条件，オレンジ色の棒グラフは変化条件，灰色の棒グラフは常体条件の平均値である。

米津のロボットの会話印象評価の Q26 では敬体条件，変化条件，常体条件の平均値は (3.14, 3.17, 4.00) であり，常体条件のみ高い値となっていた。また逆に，Q27 では (2.71, 2.50, 1.83) となっており，常体条件のみ低い値となっていた。

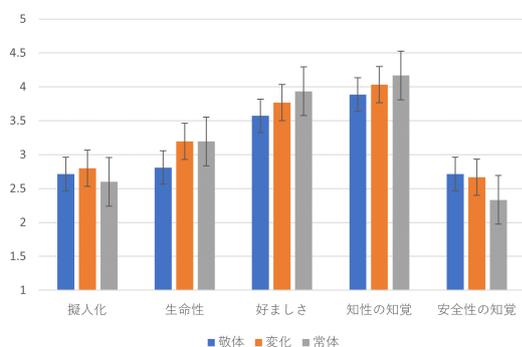


図 3: Godspeed Questionnaire の下位尺度ごとの平均値と標準誤差の結果

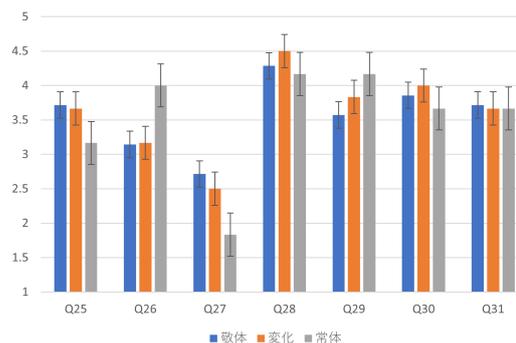


図 4: 米津のロボットの会話印象評価の項目ごとの平均値と標準誤差の結果

そこで，Kruskal-Wallis 検定で Godspeed Questionnaire の下位尺度ごとに比較した結果を表 6，米津のロボットの会話印象評価の項目ごとに比較した結果を表 7 として示す。表の平均ランクとは Kruskal-Wallis 検定で順位付けを行い，それらの順位をグループごとに平均値化した値のことであり，漸近有意確率とは Kruskal-Wallis 検定の結果 (p 値) のことである。Kruskal-Wallis 検定を行う際，有意水準は 5% ($p < .05$) とした。

結果としては，Godspeed Questionnaire の 5 つの下位尺度，また米津のロボットの会話印象評価 Q25-Q31 の全てにおいて有意差 ($p < .05$) は認められなかった。

表 6: Kruskal-Wallis 検定 (Godspeed Questionnaire の下位尺度)

下位尺度	検定統計量	自由度	漸近有意確率 (p)
擬人化 (Q1-Q5)	0.566	2	0.753
生命性 (Q6-Q11)	1.751	2	0.417
好ましさ (Q12-Q16)	0.385	2	0.825
知性の知覚 (Q17-Q21)	1.40 ⁸	2	0.495
安全性の知覚 (Q22-Q24)	1.471	2	0.479

表 7: Kruskal-Wallis 検定 (米津のロボットの会話印象評価の項目)

質問項目	検定統計量	自由度	漸近有意確率 (p)
楽しい Q25	2.152	2	0.341
発言をためらわない Q26	2.415	2	0.299
相手の様子が浮かぶ Q27	3.448	2	0.178
内容が伝わる Q28	0.204	2	0.903
簡単 Q29	1.047	2	0.593
自分に同意 Q30	0.668	2	0.716
気分が良い Q31	0.038	2	0.981

アンケートの最後に行った自由記述の質問に対する回答を以下に示す。

敬体条件では、「スマートスピーカーと話してみてもう感じたか」に関して、P3, P6 から「しっかり会話になっていると感じた」やP4, P5 から「回答がしっかりとしており理解しやすかった」という回答があった。一方で、P3 から「人間と対話している感覚は少なかった」やP2, P5 から「決められた言葉だけを話していると感じた」、P2, P7 から「機械と話している感が強かった」という回答が得られた。「スマートスピーカーを今後も使用したいと思ったか」に関して、使用したいと回答したのは無回答1人を除く6人中2人であった。

変化条件では、「スマートスピーカーと話してみてもう感じたか」に関して、P12, P13 から「途中の話し方の変化に驚いた」やP9 から「仲良くしたいということが伝わり、好感が持てた」、P11, P12 から「日常会話で使いたい」という回答があった。一方で、P9 から「距離を縮めてくるのが早い」やP10 から「もう少し敬語の時間が欲しかった」という回答が得られた。「スマートスピーカーを今後も使用したいと思ったか」に関して、使用したいと回答したのは6人中5人であった。

常体条件では、「スマートスピーカーと話してみてもう感じたか」に関して、P14, P18 から「とても話しやすい印象を持った」やP15 から「スマートスピーカーのイントネーションが敬語にはない可愛さがあった」という回答があった。一方で、P14 から「情報は詳細だがため口のためイメージと異なった」という回答が得られた。「スマートスピーカーを今後も使用したいと思ったか」に関して、使用したいと回答したのは6人中4人であった。

また全ての条件において、P1, P8, P17, P19 から「対話の速度が遅く、対話のずれが生じていたのが気になった」という回答と、P5, P11, P15, P16 から「会話が一方通行に感じた」という回答があった。

4 考察

仮説として「1回の対話内でスマートスピーカーが、スピーチスタイルシフトを考慮した敬体から常体に変化させる人間らしい対話形式をとることでユーザに与える印象は、敬体の親切的な印象と情報への信頼性、常体の人間的、親しみやすい、楽しいという印象、その両方を与えることができる。」と考えていた。また、図4の米津のロボットの会話印象評価の項目ごとの平均値と標準誤差の結果から、Q26(発言をためらわない)とQ27(相手の様子が浮かぶ)に関しては敬体条件・変化条件と常体条件の間に差があるように見える。

しかし、アンケートに対する Kruskal-Wallis 検定での分析では、全てにおいて有意な差は認められなかった。これは、敬体条件と常体条件の違いを文末形式の「デス・マス体」と「ダ体」で表現したが、実験でのタ

スクの参考にした大内らの研究 [4] や Chaves らの研究 [6] では2,3回程度会話回数で終わる設計になっていたのに対して、本実験では10分程度まで繰り返し行う設計にしたために、敬体と常体の違いが生じない単語の対話時間が多くなったためであると考えられる。また、今回の実験では参加者が敬体条件7人、変化条件6人、常体条件6人と少なかったために、標準誤差が大きくなってしまっていたためだと考えられる。さらに、全ての条件の自由記述の質問において回答されていた、対話のずれが10分程度のタスク時間にしたことによって生じやすくなってしまい、参加者のスマートスピーカーに対する印象を悪くさせていたためだと考えられる。

変化条件ではp9の「距離を縮めてくるのが早い」という悪い印象を与えてしまっていた。これは、敬体と常体が同時に存在する時間を考慮せず、敬体から常体に突然変化させたために、参加者を驚かせてしまっていたためだと考えられる。また、全ての条件の自由記述の質問において回答されていたP5, P11, P15, P16の「会話が一方通行に感じた」から、本実験のタスクは大内らの研究を参考に「旅行先を決める」という内容で行ったが、この内容では申の研究で行っていた話題自由な対話に比べて、スマートスピーカーの話す機会が多く、参加者が話す機会が少ないという一方的な対話になってしまっていた可能性が考えられる。そのためにスマートスピーカーの変化条件を行ったとしても良い印象を得られなかったと考えられる。

一方で、自由記述の質問においては、敬体条件ではP3, P6の「しっかり会話になっていると感じた」やP2, P7の「機械と話している感が強かった」、常体条件ではP15の「スマートスピーカーのイントネーションが敬語にはない可愛さがあった」やP14の「情報は詳細だがため口のためイメージと異なった」と、大内らの研究でのインタビュー結果と同様の回答となっており、変化条件ではP9の「仲良くしたいということが伝わり、好感が持てた」やP11, P12の「日常会話で使いたい」という意見もあった。また、「スマートスピーカーを今後も使用したいと思ったか」という質問に対して、敬体条件では2人、変化条件では5人、常体条件では4人であったことから、変化条件がほかの2条件と比べて良い印象を与えていた可能性がある。

5 おわりに

本論文では、スマートスピーカーの新しい対話形式として、スピーチスタイルシフトの同じ相手でも文末形式の切り替えを行ったりするという部分に注目した、時間経過で敬体から常体に変化する条件を提案し、人にどのような印象を与えるのかを調査した。敬体のみの条件、常体のみの条件、敬体から常体に変化する条

件で比較を行った結果, Godspeed Questionnaire と米津のロボットの会話印象評価を用いた評価では有意な差は認められなかった. しかし, 自由記述の質問では, 「スマートスピーカーを今後も使用したいと思ったか」という質問に対して, 敬体条件では 2 人, 変化条件では 5 人, 常体条件では 4 人であったことから, 変化条件がほかの 2 条件と比べて良い印象を与えていた可能性が示唆された.

本実験で有意な差が認められなかった原因として, 本実験で行ったタスクは参加者へのインタビューから一方通行な対話になっており, 変化条件を加えた比較を行うのに適していなかった可能性がある. 参加者とスマートスピーカーの話す時間が同程度であり, スマートスピーカーの対話形式の変化が起こることが自然に感じるようなタスクでの調査が必要であると考えられる. また, 実験参加者が少ないことにより分散が大きくなり, 第二種の過誤を引き起こしてしまった可能性もある. 大内らの研究 [4] では 14 名の参加者による実験を実施していた. 今後, これらの実験デザインの問題を改善し, 変化条件がほかの条件より良い印象を与えていた可能性と, より自然なスマートスピーカーと人での対話を実現する方法を探る.

参考文献

- [1] 伊藤 南美. スマートスピーカーに関する調査. MMD 研究所, November 2021. https://mmdlabo.jp/investigation/detail_2007.html (2023/2/15).
- [2] 申 媛善. 日本語と韓国語における文末スタイル変化の仕組み: 時間軸に沿った敬体使用率の変化に着目して. *日本語科学*, 22:173–195, oct 2007.
- [3] 宇佐美 まゆみ. 談話レベルから見た敬語使用-スピーチレベルシフト生起の条件と機能. *学苑*, (662):p27–42, 02 1995.
- [4] Subaru Ouchi, Kazuki Mizumaru, Daisuke Sakamoto, and Tetsuo Ono. Should speech dialogue system use honorific expression? In *Proceedings of the 7th International Conference on Human-Agent Interaction*, HAI '19, page 232–233, New York, NY, USA, 2019. Association for Computing Machinery.
- [5] Richard J Watts. *Politeness*. Cambridge University Press, 2003.
- [6] Ana Paula Chaves and Marco Aurelio Gerosa. Single or multiple conversational agents? an interactional coherence comparison. In *Proceedings of the 2018 CHI Conference on Human Factors in Computing Systems*, CHI '18, page 1–13, New York, NY, USA, 2018. Association for Computing Machinery.
- [7] Nils Dahlbäck, Arne Jönsson, and Lars Ahrenberg. Wizard of oz studies — why and how. *Knowledge-Based Systems*, 6(4):258–266, December 1993.
- [8] Christoph Bartneck, Dana Kulić, Elizabeth Croft, and Susana Zoghbi. Measurement instruments for the anthropomorphism, animacy, likeability, perceived intelligence, and perceived safety of robots. *International Journal of Social Robotics*, 1(1):71–81, November 2008.
- [9] 米津 壮二 and 大澤 博隆. テレプレゼンスロボットにおける同調動作の効果. *人工知能学会全国大会論文集*, JSAI2017:3Q11in1–3Q11in1, 2017.