

7M-01

自律型アンドロイドのキャラクタ表現のための対話の振る舞い制御

山本 賢太[†] 井上 昂治[‡] 中村 静[‡] 高梨 克也[‡] 河原 達也[‡][†] 京都大学 工学部情報学科[‡] 京都大学 大学院情報学研究科 知能情報学専攻

1. はじめに

近年、対話システムの研究開発が盛んに行われている。その中で、我々は人と自然にインタラクションを行う自律型アンドロイドの研究開発を進めている [1]。自律型アンドロイドは外見、振る舞い共に非常に人間に酷似させている。このためユーザはアンドロイドに対して何らかのキャラクタを感じることが期待される。より自然な対話を実現するためには、違和感のないキャラクタ表現が求められる。自律型アンドロイドは、利用されるタスクの種類に応じて社会的役割が与えられる。そのような社会的役割などの要因によって特定のキャラクタを持つことが想定される。例えば、飲食店の店員には「外交的で丁寧」といったキャラクタが期待される。

対話システムにおけるキャラクタの表現方法としては、言語表現の生成確率の違いによるキャラクタ表現 [2] などがある。本研究では対話におけるアンドロイドの発話において、発話量、相槌、フィラー、交替潜時といった対話中の振る舞いを制御することでキャラクタ表現方法の構築を目指す。与えられたキャラクタに対して適切な特微量調整を行うことにより、キャラクタに合わせた印象を与える振る舞いの実現を目指す (図 1)。本稿では、対話中の振る舞い制御による外向性、情緒不安定性、丁寧さに関する印象への影響について実験により検証を行った。

2. キャラクタ表現

2.1 キャラクタの定義

本研究ではキャラクタを、外交性、情緒不安定性、丁寧さという 3 特性を用いて表現されるものとして定義した。外交性、情緒不安定性は性格を表現する因子として広く用いられている [3]。丁寧さはアンドロイドを社会の中で利用するという場合を想定して追加した。これら 3 特性の組み合わせで任意のキャラクタを表現する。この 3 特性をそれぞれ直交する「外交的 - 内向的」「情緒安定的 - 情緒不安定的」「丁寧 - カジュアル」という 3 軸に対応させ、キャラクタはこの 3 軸により張られる 3 次元空間上の 1 点に対応させる。

2.2 対話における振る舞い

先行研究では、対話における印象に対して発話量 [4]、相槌 [5]、交替潜時 [6] といった要因の影響が確認されている。本研究ではこれらにフィラーを加えたものをキャラクタを表現するための対話中の振る舞いとした。制御する特微量としては、発話量 (発話時間の長さ)、相槌の頻度 (対話中の相槌回数)、相槌の種類 (対話中の相槌の語彙数)、フィラーの頻度 (対話中のフィラー回数)、交替潜時 (話者交替時の時間間隔) とした。本稿では、音声合成でこれらの特微量を制御することでキャラクタ特性に関する印象へ与える影響について調査した。

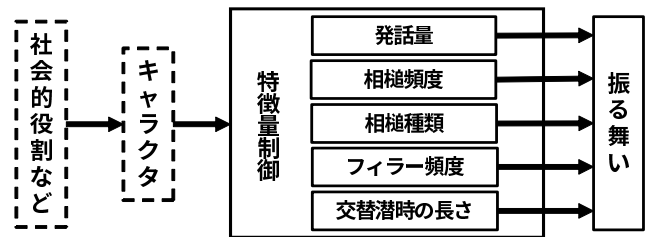


図 1: システム概念図

3. 実験

3.1 実験内容

実験に用いる音声は、アンドロイドを遠隔操作して収録した対話を参考にシナリオを 2 つ用意した。日常的な内容の特定の場面に限らない話題で、相槌やフィラーが適度に含まれ、1 分ほどで切り出せるものを選定した。このシナリオに対してアンドロイドの発話を音声合成ソフト Voice Text ERICA を用いて合成した。ユーザの声として実際の被験者の代わりに実験実施者 (第一著者) の声を用いて対話音声を作成した。

2 つのシナリオに対して、参考対話から 1 分ほどの基準対話を作成した。相槌、フィラーは参考とした対話に現れたものを使用し、交替潜時はユーザの発話終了からアンドロイドの発話開始までの時間間隔を 0.5 秒とした。実験条件ごとに使用する比較対話は、基準対話に対して対応する 1 つの特微量のみを調整し、残りの特微量は基準対話のままとした。相槌の高頻度条件はユーザの発話内の全ての節境界 [7] に相槌を挿入し、低頻度条件は相槌を全て取り除いた。相槌の多種類条件はアンドロイドの発話内の相槌を置き換えて種類を増やし、少種類条件は相槌を全て「はい」に置き換えた。フィラーの高頻度条件はアンドロイドの発話内の全ての節境界及び文頭にフィラーを挿入し、低頻度条件はフィラーを全て取り除いた。交替潜時の長時間条件は 3 秒間とし、短時間条件はアンドロイドの発話開始をユーザの発話終了に 0.5 秒オーバーラップさせた。発話量比較用音声のみ 1 分より長い対話内容を使用し、発話量が多い音声はユーザの発話時間が 25.3 秒、アンドロイドが 49.2 秒、発話量が少ない音声はユーザの発話時間が 38.8 秒、アンドロイドが 25.5 秒となった。

キャラクタ特性に関するアンケート項目を、表 1 に示す。外交性、情緒不安定性の評定には音声からの人物像評価にも用いられている性格特性 5 因子尺度 (Big Five 尺度) の短縮版 [8] を使用し、これに丁寧さに関する項目を追加した。各特性に対応する複数の形容詞に対して 7 段階で評定をしてもらい、評定結果は対応する項目の平均値とした。また、対話の自然さについての項目も 7 段階評定で用意した。発話量が多く相槌が多いほど外交的、相槌の種類が少なくフィラーが多いと情緒不安定的、交替潜時が長いと丁寧だとみなされるという仮説を立てた。

表 1: アンケート項目

キャラクター特性	項目			
外交性	話し好き	無口な*	陽気な	外交的
情緒不安定性	悩みがち	不安になりやすい	心配性	気苦労の多い
丁寧さ	丁寧な	礼儀正しい		

(* 逆転項目)

表 2: 対話の種類

	相槌頻度	相槌種類	フィラー頻度	交替潜時	発話量比較用
シナリオ 1	多	多	多	長	発話量少
	基準				
シナリオ 2	なし	「はい」のみ	なし	短	発話量多
	多	多	多	長	
	基準				
	なし	「はい」のみ	なし	短	

3.2 実験手続き

実験には本学の学生で男性 28 名, 女性 18 名の計 46 名 (18-23 歳) に参加してもらった。各実験参加者はヘッドフォンで対話音声聞き, 各対話終了毎にアンドロイドに対する印象について, アンケートの各項目に対して“あてはまらない”から“あてはまる”まで 7 段階での評定をもらった。全被験者にこの評定を全条件について実施してもらった。

対話の提示順序は, まず対話シナリオの基準対話を最初に提示した後, 各条件について提示順序はランダムに配置し, 発話量比較用の対話は最後に提示した。全体で 20 個の対話を提示し, 参加者 1 人につき実験時間は約 40 分であった。

3.3 実験結果

相槌の頻度, フィラーの頻度, 交替潜時は, 印象に対して大きく影響を与えた。発話量も印象に対して影響があったが, 相槌の種類は印象への影響が小さかった。また, 交替潜時の長いもの, 短いものに対する自然さの評価は低めとなっていた。

表 2 に示すように, 各特徴量毎に基準対話を含めた 3 群間での分散分析を行った。発話量に関してはシナリオ間での比較となるため t 検定を用いた。多重比較に関しては, Tukey の HSD 法を用いた。

2 つのシナリオで多重比較により高頻度条件と低頻度条件の間に有意差の見られたものについて以下に述べる。外交性について表 3 に示す。相槌が多く, フィラーが少なく, 交替潜時が短いほど外交的とみなされた。多重比較により, 相槌が少ない場合と基準音声では有意差が見られなかったが, フィラーの頻度に関しては低頻度条件 > 基準 > 高頻度条件という順序の有意差が見られた。情緒不安定性について表 4 に示す。フィラーが多く, 交替潜時が長いほど情緒不安定的とみなされた。フィラーの頻度に関しては高頻度条件 > 基準 > 低頻度条件という順序の有意差が見られた。丁寧さについて表 5 に示す。相槌が少なく, 交替潜時が長いほど丁寧だとみなされた。フィラーの頻度に関しても有意差が見られたが, シナリオ 2 での多重比較による高頻度条件と低頻度条件の間に有意差が見られなかった。発話量について表 6 に示す。発話時間が長いほど外交的とみなされ, 発話時間が短いほど丁寧だとみなされた。相槌の種類に関しては, 2 つのシナリオの両方に印象への影響が見られた特性はなかった。

仮説のうち, 相槌の種類以外に関しては概ね仮説通りの結果となった。相槌の種類に関しては, 対話が 1 分であり相槌の数が多い中での種類の数の比較のため, 実験参加者が差異に気づきにくかった可能性が考

表 3: 外交性に関する分散分析結果

	多い(長い)		基準		少ない(短い)		F 値	
	平均	標準偏差	平均	標準偏差	平均	標準偏差		
シナリオ 1	相槌頻度	5.44	1.08	3.99	1.16	4.08	1.07	27.174*
	相槌種類	4.64	1.22	3.99	1.16	4.76	0.90	9.342*
	フィラー頻度	3.51	1.02	3.99	1.16	4.84	1.07	18.963*
	交替潜時	2.56	0.91	3.99	1.16	5.20	1.06	70.495*
シナリオ 2	相槌頻度	5.35	0.89	4.76	1.02	4.73	0.83	10.757*
	相槌種類	4.88	0.94	4.76	1.02	4.80	0.88	0.289
	フィラー頻度	3.53	1.06	4.76	1.02	5.17	0.89	41.892*
	交替潜時	2.70	0.97	4.76	1.02	4.64	1.19	83.329*

(* $p < 0.001$)

表 4: 情緒不安定性に関する分散分析結果

	多い(長い)		基準		少ない(短い)		F 値	
	平均	標準偏差	平均	標準偏差	平均	標準偏差		
シナリオ 1	相槌頻度	2.55	1.05	3.38	1.17	3.10	1.10	9.149*
	相槌種類	3.08	1.15	3.38	1.17	2.73	0.94	7.060*
	フィラー頻度	4.71	1.42	3.38	1.17	2.42	1.09	56.378*
	交替潜時	4.76	1.16	3.38	1.17	2.38	1.03	62.988*
シナリオ 2	相槌頻度	2.91	1.21	3.29	1.45	3.22	1.16	1.921
	相槌種類	3.05	1.24	3.29	1.45	3.16	1.11	1.921
	フィラー頻度	5.07	1.50	3.29	1.45	2.37	1.03	70.268*
	交替潜時	4.84	1.26	3.29	1.45	3.09	1.31	48.321*

(* $p < 0.001$)

表 5: 丁寧さに関する分散分析結果

	多い(長い)		基準		少ない(短い)		F 値	
	平均	標準偏差	平均	標準偏差	平均	標準偏差		
シナリオ 1	相槌頻度	4.37	1.62	5.38	1.19	5.21	1.05	9.512*
	相槌種類	5.10	1.17	5.38	1.19	5.46	1.05	1.508
	フィラー頻度	4.77	1.33	5.38	1.19	5.21	1.00	4.078**
	交替潜時	4.60	1.18	5.38	1.19	2.84	1.09	63.181*
シナリオ 2	相槌頻度	4.12	1.62	3.48	1.59	4.85	0.96	13.939*
	相槌種類	5.01	0.96	3.48	1.59	4.88	0.94	31.274*
	フィラー頻度	4.54	1.02	3.48	1.59	4.75	1.21	17.613*
	交替潜時	4.41	1.12	3.48	1.59	3.10	1.59	18.896*

(* $p < 0.001$, ** $p < 0.05$)

表 6: 発話量に関する t 検定結果

	発話時間長い(シナリオ 2)		発話時間短い(シナリオ 1)		t 値
	平均	標準偏差	平均	標準偏差	
外交性	5.74	0.82	5.03	0.84	4.991*
情緒不安定性	2.74	0.93	2.83	1.04	0.545
丁寧さ	4.76	1.22	5.91	1.04	5.688*

(* $p < 0.001$)

えられる。

4. おわりに

本稿では, 音声対話における振る舞いが与える印象について分析した。発話量, 相槌の頻度, フィラーの頻度, 交替潜時の違いが外交性, 情緒不安定性, 丁寧さといった印象に影響することが確認できた。

本稿では自然さについては分析から除いたが, 自然な音声として許容される範囲内での特徴量制御についても検討する必要がある。この結果をもとにシステムにキャラクターを入力として与えた際に, そのキャラクターに応じた振る舞いの特徴量を自動的に決定する手法について検討する。

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の支援を受けて実施された。

参考文献

- [1] 井上昂治ほか。自律型アンドロイド ERICA のための音声対話システム。人工知能学会研究会資料 SIG-SLUD B502(02), 21-24, 2015。
- [2] 宮崎千明ほか。文節機能部の確率的書き換えによるキャラクター性変換。言語処理学会 第 21 回年次大会 発表論文集, 277-280, 2015。
- [3] 小塩真司。Progress & Application パーソナリティ心理学。サイエンス社, 2014。
- [4] 小川一美。二者間の発話量の均衡が観察者が抱く会話者と会話者に対する印象に及ぼす効果。実験社会心理学, 43(1):63-74, 2003。
- [5] E.de Sevin *et al.* Influence of personality traits on backchannel selection. IVA 2010 LNAI 6356, 187-193, 2010。
- [6] 長岡千賀ほか。音声対話における交替潜時が対人認知に及ぼす影響。ヒューマンインタフェースシンポジウム 2002 論文集, 171-174, 2002。
- [7] 小磯花絵ほか。話し言葉コーパス 設計と構築。朝倉出版, 2015。
- [8] 内田照久。音声の発話速度が話者の性格印象に与える影響。心理学研究, 73(2):131-139, 2002。