

# 自律型アンドロイドの対話の振る舞い制御モデルによる キャラクタ表現法の検討

## Investigation of a Character Representation Method by Dialogue Behavior Control Model of an Autonomous Android

山本 賢太\* 井上 昂治 中村 静 高梨 克也 河原 達也

Kenta Yamamoto, Koji Inoue, Shizuka Nakamura, Katsuya Takanashi, Tatsuya Kawahara

京都大学 大学院情報学研究科

Graduate School of Informatics, Kyoto University

**Abstract:** This paper addresses character expression for humanoid robots that play a social role via spoken dialogue, so that the character matches to the given social role such as a lab guide or a counselor. While conventional methods of character expression are mostly based on changing the style of utterance texts, this study focuses on speech-based features that may affect the impression of spoken dialogue. Specifically, we use five features: utterance amount, backchannel frequency, backchannel variety, filler frequency, and switching pause length. We adopt the three-dimensional character space of extroversion, emotional instability and politeness, and investigated the relationship with the speech-based features. A statistical analysis of subjective evaluations shows that the speech features except for the backchannel variety are related to either of the dimensions. Moreover, by using the subjective evaluation scores on the relevant dimensions, we can train classifiers to control the speech-based features and thus control dialogue behaviors according to the desired character. An experimental evaluation demonstrates the feasibility of character expression with regard to the dimensions of extroversion and politeness.

## 1 はじめに

人と自然なインタラクションを行う人間型ロボットの研究開発が行われている [1, 2]. このようなロボットの一つに、自律型アンドロイドがある [3]. 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たものである。そのため、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。自律型アンドロイドには、利用されるタスクの種類に応じてカウンセラーや研究室ガイドなどの社会的役割が与えられる。このような社会的役割に応じたキャラクタを付与することで、落ち着いているカウンセラーの方が話しやすいといったように、ユーザに良い印象を与えるだけでなく、対話によるタスクの遂行にもよい影響があると考えられる。

対話システムにおけるキャラクタの表現方法として、沈ら [4] は、言語パターンによるキャラクタの表現法を提案している。Mairesse ら [5] は、ユーザに与える印象をコントロールしながら対話するシステムとして

PERSONAGE を開発した。これらの研究では、文章の書き換えによりキャラクタが表現されている。

一方で、音声対話では話し方などの要因も印象への影響があると考えられる。Levitan ら [6] は、ユーザの話速に同調する音声対話システムを構築した。話速以外の要因についてもたとえば、よく話す人は外向的で、フィラーが多いと落ち着きがないといった印象を与えることが予想される。そこで、本研究では、アンドロイドの対話の振る舞いを制御することでキャラクタを表現するモデルを構築する。

## 2 キャラクタの表現

本研究におけるキャラクタの定義と対話の振る舞い制御に用いる音声特徴量について述べる。

### 2.1 キャラクタの定義

Eysenck は、外向性、神経症傾向（情緒不安定性）という特性を用いて性格を 2次元で表現している [7]. 性

\*連絡先：京都大学 大学院情報学研究科 知能情報学専攻  
〒 606-8501 京都府京都市左京区吉田本町  
E-mail: yamamoto@sap.ist.i.kyoto-u.ac.jp

格 5 因子尺度 (Big Five 尺度) は、外向性、情緒不安定性、誠実性、調和性、開放性の 5 つの特性を用いて、性格を表現する [8]。Sevin ら [9] は、エージェントの性格表現に外向性、情緒不安定性を用いている。また、Wang ら [10] は、エージェントに丁寧さをもたせた場合の影響について実験を行っている。外向性、情緒不安定性は心理学で性格を表現する際に広く用いられている指標であり、先行研究においてもエージェントの性格表現に用いられている。丁寧さも、コミュニケーションにおける重要な要因である。

そのため本研究では、この 3 特性をそれぞれ「外向的 - 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 振る舞い制御のための音声特徴量

対話における発話者の印象に影響が考えられる音声特徴量を検討する。発話量は、たくさん話す人ほど外向的と感じられるなど、対話における印象への影響があると考えられる。相槌は、対話において感情の表出や理解を示すなど様々な役割をもつため、印象にも影響すると考えられる。Sevin ら [9] は、外向性と情緒不安定性について、相槌の頻度と種類が、外向性、情緒不安定性の印象に対して影響することを示している。フィラーは、音声対話特有の表現であり印象にも大きく影響すると考えられる。Shiwa ら [11] によると、「えっと」というフィラーが反応の遅延に対する印象の改善に効果があることが示されている。音声対話においてシステムの反応時間である交替潜時は、ユーザが認識しやすい特徴であると考えられる。長岡ら [12] は、交替潜時の長さが話者についての印象評定に影響があることを確認している。

以上をふまえて、音声の特徴量 (以降、音声特徴量) としては、発話量、相槌の頻度、相槌の種類、フィラーの頻度、交替潜時の長さを用いる。相槌と異なりフィラーの種類に関しては、印象との関係性が明確に示唆されていないため、今回は採用しない。

## 3 キャラクタの印象への振る舞いの影響の調査

振る舞い制御モデルを構築するに際して、音声特徴量のキャラクタに対する影響を調べる。音声特徴量のキャラクタ特性に対する印象への影響について以下の仮説を立てる。

1. 発話量が多く相槌が多いほど、対話に対して積極的と感じられ、外向的と受けとられやすい。

表 1: 音声特徴量の制御内容

音声特徴量	条件	制御内容
発話量	多	A: 49.2 秒, U: 25.3 秒 *
	少	A: 25.5 秒, U: 38.8 秒 *
相槌の頻度	高	ユーザ発話中の節境界及び文末
	低	相槌を削除
相槌の種類	多	「えー」、「ふん」、「あー」、「はい」
	少	「はい」のみ
フィラーの頻度	高	ロボット発話中の節境界及び文頭
	低	フィラーを削除
交替潜時の長さ	長	3 秒
	短	0.5 秒のオーバーラップ

\* A: アンドロイド, U: ユーザ  
基準対話に使用した対話の後に内容を追加した。

2. 相槌の種類が少なくフィラーが多いと、発話に自信がないと感じられ、情緒不安定的と受けとられやすい。

3. 交替潜時の長さが長いと、相手の話をよく聞き応答していると感じられ、丁寧と受けとられやすい。

### 3.1 音声試料

音声試料は人間がアンドロイドを遠隔操作する WOZ 法により収録した 2 者による対話を参考に、シナリオを 2 つ用意した。収録した対話の中から、特定の場面に限られない日常的な内容の話題で、相槌やフィラーが適度に表出し、1 分程度で内容がまとまっているものを選定している。対話内容は、任意のキャラクタで不自然にならない対話内容であること、また音声特徴量 (交替潜時を除く) を調整しても不自然にならないことを考慮した。このシナリオ内のアンドロイドの発話部分を、音声合成ソフト VoiceText ERICA を用いて合成している。ユーザの発話部分は、実際の話者の代わりに実験実施者 (著者) の声を用いている。

2 つのシナリオに対して、1 分程度の基準対話を作成した。相槌とフィラーは、参考とした対話に現れたものを使用し、交替潜時の長さを 0.5 秒としている。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 つの音声特徴量のみを調整し、残りの音声特徴量は基準対話のままとしている。音声特徴量の制御内容は表 1 にまとめる。相槌、フィラーを追加する際には、不自然にならないよう考慮して選択した。

キャラクタ特性に関するアンケート項目は、パーソナリティ心理学や対人印象研究で広く用いられ、音声からの人物像評価にも用いられている心理尺度である性格特性 5 因子尺度 (Big Five 尺度) の短縮版 [14] から、外向性、情緒不安定性に関する項目を採用した。これに丁寧さに関する項目を独自に追加したものを使用した (表 2)。ただし、丁寧さの 2 項目の選定に際し、里

表 2: アンケート項目

キャラクター特性	項目			
外向性	話し好き	無口な*	陽気な	外向的
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)

見 [15] の研究を参考にした。各特性に対応する複数の形容詞に対して「あてはまらない:1」から「あてはまる:7」までの7段階で評定してもらった。各キャラクター特性の評定点是对应する項目の平均値とする。また、対話の自然さについても「不自然:1」から「自然:7」まで7段階での評定をしてもらった。全被験者はこの評定を全条件について行う。

### 3.2 実験手続き

実験には男性 28 名、女性 18 名の計 46 名の大学生 (18 ~ 23 歳) が参加した。各実験参加者はヘッドフォンで対話音声聞き、アンケートフォームにて回答を行う。対話音声を聞いた後、女性アンドロイドに対しての印象を回答してもらった。女性の発話が音声合成であることは明示していない。最初に基準対話、最後に発話量比較用対話を提示するよう順序を固定し、残りの条件については実験ごとに順序をランダムに並べ替えた。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 分析結果

音声特微量ごとに基準対話を含めた 3 群間での分散分析を行った。各キャラクター特性で、多重比較により高頻度条件と低頻度条件の間に有意差 ( $p < 0.05$ ) の見られたものについて以下に述べる。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。

外向性に関する分析結果を表 3 に示す。相槌の頻度が高く、フィラーの頻度が低く、交替潜時の長さが短いほど、外向的とみなされた。情緒不安定性に関する分析結果を表 4 に示す。フィラーの頻度が高く、交替潜時の長さ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。丁寧さに関する分析結果を表 5 に示す。相槌の頻度が低く、交替潜時が長いほど、丁寧とみなされた。発話量に関する分析結果を表 6 に示す。発話時間が長いほど外向的とみなされ、発話時間が短いほど丁寧とみなされた。情緒不安定性に関しては、有意差が見られなかった。

## 3.4 考察

音声資料に対する自然さの評定の平均点については、2 シナリオともに交替潜時の 2 条件のみ基準対話よりも低く、残りのすべての条件に関しては基準対話と同等以上となっていた。そのため交替潜時以外の音声特微量においては本実験の調整において、問題となる不自然さは生じていないと考えられる。

相槌の頻度は、分散分析による有意差が見られた場合でも、多重比較による低頻度条件と基準対話との間の有意差が見られない場合も多い。これは基準対話の実際の対話コーパスから取り出した相槌の回数 (シナリオ 1:4 回, シナリオ 2:3 回) が高頻度条件 (シナリオ 1:14 回, シナリオ 2:8 回) と低頻度条件 (0 回) の低頻度条件側によっていたためと考えられる。相槌の種類は、2 シナリオともに印象評定の有意差が見られた特性はほとんどなかった。1 分間の対話では種類の差異に気づきにくかったことが考えられる。フィラーの頻度は、交替潜時の長さについて印象評定への影響が大きかった。多重比較においても、低頻度条件から基準音声、高頻度条件という順序のついた傾向が見られるものが多かった。交替潜時の長さは、印象への影響が特に大きくすべての印象特性に影響していた。交替潜時の長さの違いは、差異が知覚しやすかった可能性が高い。また、印象への影響が大きいが自然さの評価が低かった。本実験では外向性、丁寧さへの影響が見られたが、内容依存である可能性も考慮しなければならない。本実験結果においてキャラクター特性の丁寧さと他の 2 つの評定点の相関係数は小さく、キャラクター特性は独立であると考えている。これらの結果から、発話量、相槌の頻度、フィラーの頻度、交替潜時の長さを制御することによるキャラクター表現の可能性が示された。

## 4 振る舞い制御モデル

本節では、音声特微量を利用して、与えられたキャラクターに応じた振る舞いの制御モデルを構築する。さらに、そのモデルの評価実験について述べる。

### 4.1 ロジスティック回帰

ロジスティック回帰を用いて、与えられたキャラクターから音声特微量の制御量を決定するモデルを構築する。ロジスティック回帰では、入力には各キャラクター特性を 1 から 7 の間で数値化したものとし、与えられたキャラクターのための各音声特微量の制御量 [0 ~ 1] を出力する。3 章の実験で用いた音声特微量の条件をそれぞれ 0, 1 とラベル付けし、各キャラクター特性の評定点

表 3: 外向性に関する分散分析結果

	音声特徴量	高/多/長		基準		低/少/短		F 値	多重比較
		平均	SD	平均	SD	平均	SD		
シナリオ 1	相槌の頻度	5.44	1.08	3.99	1.16	4.08	1.07	27.174*	高 > 低, 高 > 基準
	相槌の種類	4.64	1.22	3.99	1.16	4.76	0.90	9.342**	多 > 基準, 少 > 基準
	フィラーの頻度	3.51	1.02	3.99	1.16	4.84	1.07	18.963**	低 > 基準 > 高
	交替潜時の長さ	2.56	0.91	3.99	1.16	5.20	1.06	70.495**	短 > 基準 > 長
シナリオ 2	相槌の頻度	5.35	0.89	4.67	1.17	4.73	0.83	12.020**	高 > 低, 高 > 基準
	相槌の種類	4.88	0.94	4.67	1.17	4.80	0.88	0.813	
	フィラーの頻度	3.53	1.06	4.67	1.17	5.17	0.89	35.103**	低 > 基準 > 高
	交替潜時の長さ	2.70	0.97	4.67	1.17	4.64	1.19	64.468**	短 > 長, 基準 > 長

(\*  $p < 0.05$ , \*\* $p < 0.01$ )

表 4: 情緒不安定性に関する分散分析結果

	音声特徴量	高/多/長		基準		低/少/短		F 値	多重比較
		平均	SD	平均	SD	平均	SD		
シナリオ 1	相槌の頻度	2.55	1.05	3.38	1.17	3.10	1.10	9.149**	低 > 高, 基準 > 高
	相槌の種類	3.08	1.15	3.38	1.17	2.73	0.94	7.060**	基準 > 少
	フィラーの頻度	4.71	1.42	3.38	1.17	2.42	1.09	56.378**	高 > 基準 > 低
	交替潜時の長さ	4.76	1.16	3.38	1.17	2.38	1.03	62.988**	長 > 基準 > 短
シナリオ 2	相槌の頻度	2.91	1.21	3.92	1.45	3.22	1.16	15.119**	基準 > 高, 基準 > 低
	相槌の種類	3.05	1.24	3.92	1.45	3.16	1.11	13.799**	基準 > 多, 基準 > 少
	フィラーの頻度	5.07	1.50	3.92	1.45	2.37	1.03	67.347**	高 > 基準 > 低
	交替潜時の長さ	4.84	1.26	3.92	1.45	3.09	1.31	32.082**	長 > 基準 > 短

(\*  $p < 0.05$ , \*\* $p < 0.01$ )

表 5: 丁寧さに関する分散分析結果

	音声特徴量	高/多/長		基準		低/少/短		F 値	多重比較
		平均	SD	平均	SD	平均	SD		
シナリオ 1	相槌の頻度	4.37	1.62	5.38	1.19	5.21	1.05	9.512**	低 > 高, 基準 > 高
	相槌の種類	5.10	1.17	5.38	1.19	5.46	1.05	1.508	
	フィラーの頻度	4.77	1.33	5.38	1.19	5.21	1.00	4.078*	基準 > 高
	交替潜時の長さ	4.60	1.18	5.38	1.19	2.84	1.09	63.181**	基準 > 長 > 短
シナリオ 2	相槌の頻度	4.12	1.62	4.05	1.41	4.85	0.96	7.382**	低 > 高, 低 > 基準
	相槌の種類	5.01	0.96	4.05	1.41	4.88	0.94	15.347**	多 > 基準, 少 > 基準
	フィラーの頻度	4.54	1.02	4.05	1.41	4.75	1.21	5.887**	高 > 基準, 低 > 基準
	交替潜時の長さ	4.41	1.12	4.05	1.41	3.10	1.59	18.001**	長 > 短, 基準 > 短

(\*  $p < 0.05$ , \*\* $p < 0.01$ )

多重比較において 2 シナリオ共に両極端の条件間 (太字) に有意差の見られた音声特徴量 (太字) をモデルに採用した

表 6: 発話量に関する t 検定結果

キャラクター特性	発話時間長い		発話時間短い		t 値
	平均	SD	平均	SD	
外向性	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**

(\*  $p < 0.05$ , \*\* $p < 0.01$ )

太字はモデルに採用した特性

[1 ~ 7] から音声特徴量の各条件 (2 値) へのマッピングを学習する。学習には 3 章の実験結果で有意差の見られた特性のみを使用する (表 7)。

各音声特徴量に対して実験参加者による評定点を 92 サンプル用いて学習を行った。各データセットのうち

4 分の 3 を学習データ, 4 分の 1 をテストデータとして交差検定を行った。閾値を 0.5 に定めた時の 2 値分類結果を表 7 に示す。F 値は相槌の頻度を除いて 0.8 以上であった。

## 4.2 音声特徴量の制御方法

振る舞い制御モデルに各キャラクター特性を 1 から 7 の整数値で入力し, 前節のロジスティック回帰の出力値を用いて, 音声特徴量を制御する。制御内容の値の範囲は, 3 章の実験にて設定した範囲に合わせた。

表 7: ロジスティック回帰に用いた学習データと各音声特徴量の分類の予測精度

音声特徴量	使用した特性			クラス	適合率	再現率	F 値
	外	情	丁				
発話量	○		○	多	0.778	0.778	0.778
				少	0.857	0.857	0.857
相槌の頻度	○		○	高	0.680	0.810	0.739
				低	0.810	0.680	0.739
フィルターの頻度	○	○		高	0.750	1.000	0.857
				低	1.000	0.720	0.837
交替潜時の長さ	○	○	○	長	0.903	0.900	0.900
				短	0.901	0.902	0.900

使用した特性 外：外向性, 情：情緒不安定性, 丁：丁寧さ

#### 4.2.1 発話量の制御

発話量を制御するために、あらかじめ発話内容の多いパターンと少ないパターンの応答文を用意する。音声特徴量制御モデルにおける発話量のモデルの出力結果に応じて応答文を選択する。出力値が 0.5 以上の場合は発話量が多く、0.5 未満の場合は発話量が少なくなるよう応答文を選択する。

#### 4.2.2 相槌の制御

相槌は統計モデルによって制御可能である [16, 17, 18]。今回の実験では、モデルを単純化するため、ユーザ発話内のすべての節境界において相槌が生起する確率が等しいと仮定し、振る舞い制御モデルの出力を閾値とする。ユーザ発話内の節境界において 0 から 1 の範囲の一様分布からサンプリングした値を相槌の生起確率とし、この確率が閾値を超えた場合に相槌を生成する。

#### 4.2.3 フィルターの制御

フィルターの頻度も相槌と同様に統計モデルにより制御する。今回の実験では、モデルを単純化するため、アンドロイド発話内のすべての節境界においてフィルターが生起する確率が等しいと仮定し、振る舞い制御モデルの出力を閾値とする。アンドロイド発話内の節境界において 0 から 1 の範囲の一様分布からサンプリングした値をフィルターの生起確率とし、この確率が閾値を超えた場合にフィルターを生成する。

#### 4.2.4 交替潜時の長さの制御

応答文を発話するタイミングは、キャラクター特性  $c$  を与えたときの振る舞い制御モデルの交替潜時の制御量  $f_{pause} \in [0, 1]$  を 3 章の実験条件の範囲  $[-0.5, 3]$  に

表 8: 与えたキャラクタと正規化された印象評定点のピアソンの積率相関係数

キャラクター特性	相関係数	t 値
<b>外向性</b>	0.570	9.163**
情緒不安定性	-0.004	-0.056
<b>丁寧さ</b>	0.235	3.185**

(\*  $p < 0.05$ , \*\*  $p < 0.01$ )

太字は相関の見られた特性

合わせて正規化した値 (1) を交替潜時の長さとする。負の値の場合はユーザの発話末尾にオーバーラップする。

$$length = f_{pause}(c) \times 3.5 - 0.5 \quad (1)$$

### 4.3 評価実験

1 つの対話シナリオに対してアンドロイドの発話量が多い対話と少ない対話を用意した。この対話に対して、振る舞い制御モデルへの入力として、外向性、情緒不安定性、丁寧さの 1 から 7 の組み合わせのうち、16 通りの制御内容とした。

この評価実験には大学生及び大学院生の男性 11 名が参加した。実験参加者はこの 16 対話を聞き印象評定した。印象評定に用いたアンケートは 3 章の実験と同一のものを用いた。対話の提示順序は実験参加者ごとにランダムに並べ替えた。

実験参加者ごとに評定点を平均と分散で正規化した Z スコアを求めた上で、あらかじめ与えたキャラクター特性の 7 段階の点数とのピアソンの積率相関係数を求めた (表 8)。外向性と丁寧さについて p 値が有意水準 1% で有意な相関が確認された。

外向性はすべての音声特徴量の制御に関与しているため (表 7)、印象に与える影響が最も大きいと考えられる。丁寧さに関しては外向性に比べて相関が小さく、情緒不安定性に関しては有意な相関が見られなかった。

また、各条件において自然さの平均値は 3.5 を超えており、本実験における音声特徴量の調整において極端な不自然さは生じていないと考えられる。

## 5 おわりに

本研究では、自律型アンドロイドのキャラクター表現法として、対話における音声特徴量の制御によるキャラクターに応じた振る舞いを生成するモデルを構築した。

そのため、音声特徴量を制御することによるキャラクターの印象への影響について調査した。この実験結果から、発話量、相槌の頻度、フィルターの頻度、交替潜時

の長さの違いが外向性、情緒不安定性、丁寧さといった印象に影響することを明らかにした。

この実験結果をもとにキャラクタに応じた振る舞い制御モデルを提案した。このモデルを実装し実験を行い、振る舞いの制御により、外向性と丁寧さのキャラクタ特性については表現可能であることが確認された。

今回は相槌やフィラーの生起確率を一定と仮定したが、相槌やフィラーの位置によっても印象が変化すると考えられるため、今後の展望として、相槌やフィラーの統計的生成モデルと組み合わせたモデルを検討する。

## 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ショナルプロジェクト（課題番号：JPM-JER1401）の一環で行われた。

## 参考文献

- [1] Breazeal, C., Scassellati, B.: A context-dependent attention system for a social robot, *Proceedings of the Sixteenth International Joint Conference on Artificial Intelligence*, pp. 1146–1153 (1999)
- [2] Ishiguro, H., Ono, T., Imai, M., Kanda, T.: Development of an Interactive Humanoid Robot “Robovie” — An interdisciplinary approach, *Robotics Research*, Springer, pp. 179–191 (2003)
- [3] Inoue, I., Milhorat, I., Lala, D., Zhao, T., Kawahara, T.: Talking with ERICA, an autonomous android, *SIGdial Meeting on Discourse and Dialogue (SIG-DIAL)*, pp. 212–215 (2016)
- [4] 沈睿, 菊池英明, 太田克己, 三田村健: 音声生成を前提としたテキストレベルでのキャラクタ付与, *情報処理学会論文誌*, Vol. 53, No. 4, pp. 1269–1276 (2012)
- [5] Mairesse, F. and Walker, M. A.: Controlling User Perceptions of linguistic style: Trainable generation of personality traits, *Computational Linguistics*, Vol. 37, No. 3, pp. 455–488 (2011)
- [6] Levitan, R., Benus, S., Galvez, R. H., Gravano, A., Savoretti, F., Trnka, M., Weise, A., Hirschberg, J.: Implementing Acoustic-Prosodic Entrainment in a Conversational Avatar, *In Proc. INTERSPEECH*, pp. 1166–1170 (2016)
- [7] Eysenck, H. J.: *The biological basis of personality*, Springfield, IL: Charles C. Thomas Publisher, (1976)
- [8] 和田さゆり: 性格特性用語を用いた Big Five 尺度の作成, *心理学研究*, Vol. 67, No. 1, pp. 61–67 (1996)
- [9] Sevin, E. de, Hyniewska, S. j., Pelachaud, C.: Influence of Personality Traits on Backchannel Selection, *IVA 2010 LNAI 6356*, pp. 187–193 (2010)
- [10] Wang, N., Johnson, W. L., Mayer, R. E., Rizzo, P., Shaw, E., Collins, H.: The politeness effect: Pedagogical agents and learning outcomes, *Interaction Journal of Humam-Computer Studies*, Vol. 66, No. 2, pp. 98–112 (2008)
- [11] Shiwa, T., Kanda, T., Imai, M., Ishiguro, H., Hagita, N.: How quickly should communication robots respond?, *International Journal of Social Robotics*, Vol. 1, pp. 153–160 (2009)
- [12] 長岡千賀, Draguna, M., 小森政嗣, 中村敏枝: 音声対話における交替潜在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*, pp. 171–174 (2002)
- [13] 小磯花絵, 前川喜久雄, 五十嵐陽介, 丸山岳彦, 伝康晴, 籠宮隆之, 西川賢哉, 菊地浩平: *話し言葉コーパス 設計と構築*, 朝倉出版, (2015)
- [14] 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*, Vol. 73, No. 2, pp. 131–139 (2002)
- [15] 里見文: 「丁寧な」と「cortés」の連想語の比較 -日本語母語話者とメキシコ人とスペイン語母語話者の質問紙調査から-, *日本語教育方法研究会誌*, Vol. 16, No. 2, pp. 50–51 (2009)
- [16] Fujie, S., Fukushima, K., Kobayashi, T.: A conversation robot with back-channel feedback function based on linguistic and nonlinguistic information, *Proceeding of International Conference on Autonomous Robots and Agent*, pp. 379–384 (2004)
- [17] Truong, K. P., Poppe, R., Heylen, D.: A rule-based backchannel prediction model using pitch and pause information, *Interspeech 2010. International Speech Communication Association (ISCA)*, pp. 3058–3061 (2010)
- [18] Kawahara, T. and Yamaguchi, T. and Inoue, K. and Takanashi, T. and Ward, N.: Prediction and generation of backchannel form for attentive listening systems, *In Proc. INTERSPEECH*, pp. 2890–2894 (2016)