

「ロボットの外見と 発話音声の適合度に関する研究」

大阪工業大学情報科学部情報メディア学科

C05-119 森本 由紀

発表日 2009年2月17日

現状の背景と問題点

- 背景
 - 近年、情報家電などのインタフェースとして、ユーザと円滑なコミュニケーションを構築しようとするエージェントの開発が多く行われている[1]
- 問題点 = 適応ギャップ
 - 同じ音情報であっても外見の異なるエージェントから表出することで、人間は異なった意味として解釈してしまう[2]
 - 大型のロボットを見ると、人間はそのロボットに犬らしい行動や反応を期待する[3]
- 人間は外見に相応しい行動や対応を無意識下で予想し、かつ期待している

[1] 小松 実徳: ビープ音からコンピュータの態度が推測できるのか? - 規律上情報の変動が情報発信者の態度推測に与える影響 - ヒューマンインタフェース学会論文誌 Vol. 7 No. 1, pp19-25(2006)
[2] 小松 実徳: エージェントの外見の違いがユーザの態度解釈に与える影響 - 外見の異なるエージェントからの同一人工音の提示実験: 日本知能情報処理学会論文誌 No.4 pp47-57(2008)
[3] Yamada S., and Yamaguchi T.: training AIBO like a dog: In proceeding of the 13th IEEE International Workshops on robot and Human Interactive Communication pp.431-439(2004)

仮説

『人間らしい外見を持つロボットが表出する音声には、
機械的な音ではなく人間の肉声が適合する』

《実験でのロボットは物理的実体を持つ
エージェントであると解釈している》

実験刺激

【ロボットの種類】

FUJITSU VL-1710SS
カラー液晶ディスプレイ
17インチ

ペット型ロボット
(セガトイズ MIO)

ヒューマノイド型ロボット
(セガトイズ エマ)



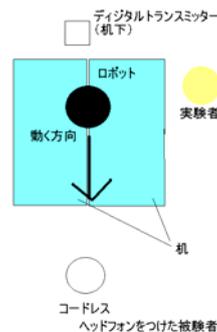
【表出音声】

- ビープ音・合成音声・人間の肉声

実験計画

- 【実験1】 サーストンの一対比較法
 - ロボットの外見に適合する音声を見つけ出すため
- 【実験2】 29の形容詞対におけるSD法
 - 実験1の結果が何の要因で決まるのかを調べるため
 - 5段階で評価
- 【場面設定】
 - 表出音声と行動は帰宅時、玄関扉を開けた後、ユーザが「ただいま」と発話した後のロボットの対応だと想定してもらう

実験の図



被験者(男14人:女16人、計30人)
に位置を指示し、正面のロボットを見
てもらおう

「コードレスヘッドフォンからロボットが
語りかけます」と説明をした後、
ロボットを動かす(ディスプレイは動か
ない)評価を行ってもらおう

■ 実験1と2を通して実験環境は統一

■ コードレスヘッドフォンから聞こえる
音声は実験者がICレコーダを使って
操作する

結果

■ 実験1の目的

ロボットと外見に適合する音声の評価

サーストンの一対比較法の結果、出てきた尺度

値がマイナスに大きいほど、より適合した音声であると解釈する

	ビーブ音	合成音声	人の肉声
ディスプレイ	0.30	-0.39	0.09
ペット型ロボット	-0.1	-0.3	0.4
ヒューマノイド型ロボット	0.62	-0.19	-0.43

ディスプレイには合成音声

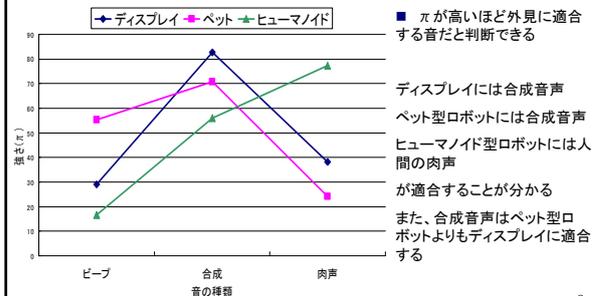
ペット型ロボットには合成音声

ヒューマノイド型ロボットには人の肉声が適合するという結果が得られた

7

■ 実験1

Bradley-terryモデルでの結果



8

■ 実験2の目的

ロボットの外見と音声における印象決定の要因分析

- SPSSで因子分析を行った結果、3つの因子を抽出

因子1から
 感じの良い・近づきやすい・親しみやすい、を「親近性因子」
 (因子負荷量 >0.75)

因子2から
 人間的な・複雑な、を「人間性因子」
 (因子負荷量 >0.7)

因子3から
 速い・すばやい、を「速度因子」
 (因子負荷量 >0.75)

と名づけた

	因子1	因子2	因子3
よい	0.688	0.163	-0.124
やさしい	0.748	0.378	-0.1
可愛らしい	0.735	0.372	-0.019
面白い	0.462	0.552	-0.041
暖かい	0.601	0.644	-0.062
人間的な	0.288	0.742	-0.003
愉快な	0.576	0.531	-0.109
好きな	0.745	0.318	-0.058
好きな	0.505	0.645	0.089
感じの良い	0.791	0.415	-0.045
賢い	0.387	0.547	-0.058
派手な	0.134	0.658	-0.152
複雑な	0.095	0.769	-0.019
数値な	0.073	0.569	0.479
充実した	0.509	0.633	0.017
明るい	0.49	0.639	0.166
積極的な	0.359	0.553	0.418
打ち解けた	0.637	0.453	0.091
速い	0.016	0.273	0.776
すばやい	0.028	0.293	0.797
興味深い	0.482	0.573	0.149
安全な	0.667	0.209	-0.154
激しい	-0.31	-0.23	0.464
落ち着いた	-0.184	-0.131	0.365
面白い	0.687	-0.012	-0.023
近づきやすい	0.868	0.231	-0.009
親しみやすい	0.855	0.279	-0.044
思いやりのある	0.55	0.556	-0.016
分かりやすい	0.545	0.379	0.067

9

多重比較法による有意差

- 因子間の差がLSD値より大きければ有意差がある

- 「親近性因子と速度因子」間の差、「人間性因子と速度因子」間の差に($p < 0.05$)で有意差あり

- 「親近性因子と人間性因子」間の差に有意差なし

	LSD	有意差
親近性因子と人間性因子の差	0.036	なし (n.s)
人間性因子と速度因子の差	0.313	あり (*)
親近性因子と速度因子の差	0.349	あり (*)

- 「人間性因子と速度因子」間の差(0.313)と、「親近性因子と速度因子」間の差(0.349)を比べると「親近性因子と速度因子」間の差の方が大きい

- 「親近性因子」が最も印象決定に貢献している

10

まとめ

- ロボットの外見と音声の適合について
 - ディスプレイには合成音声
 - ペット型ロボットには合成音声
 - ヒューマノイド型ロボットには人間の肉声

「人間らしい外見のロボットが表出する音声には、機械的な音ではなく人間の肉声が適合する」という仮説が支持された

- ロボットの外見と音声の印象決定には
 - 「親近性因子」「人間性因子」「速度因子」が影響し、中でも「親近性因子」が最も貢献している

- ロボットの外見と音声の適合度の印象決定に一番重要なのは親近感である

- 今後、親近感がどのようにしてユーザーに喚起されるのかを調べることで、ロボットはさらにユーザーに受け入れられ、親しみやすい存在として社会の中で活躍するだろう

11