

発話データ解析に基づく 外出意欲増進対話ロボットのシナリオ評価指標

慶金 佑利子¹・松本 浩子²・内田 敬³

¹学生会員 大阪市立大学大学院前期博士課程（〒558-8585 大阪市住吉区杉本 3-3-138）

²正会員 大阪市立大学都市研究プラザ（〒558-8585 大阪市住吉区杉本 3-3-138）

³正会員 大阪市立大学大学院教授（〒558-8585 大阪市住吉区杉本 3-3-138）

E-mail: uchida.ur@eng.osaka-cu.ac.jp

視覚障害者の日常生活向上のための移動支援として、音声案内「ことばの地図」を主とする歩行者ナビシステムの研究を進めてきた。「ことばの地図」は道路等の線の空間では機能するが、ショッピングモール等の広場的空間では進行方向指示が課題となっている。そこで「ことばの地図」と連携し、広場的空間（ショッピングモール等）でも案内する「盲導ロボット犬（対話ロボット）」を構想している。また、対話ロボットの利用により外出頻度が低い視覚障害者及び高齢者の外出意欲増進につながると期待している。

本研究では、「盲導ロボット犬」に付加させる対話機能開発のための1つのステップとして、会話の適切度を定量化するための発話データ解析に基づく客観的指標を確立し、対話実験を行い対話ロボットのシナリオを評価することを目的とする。

Key Words : visually-impaired, elderly people, dialog experiment, utterance analysis, AI

1. 背景と目的

視覚障害者の買い物や通勤など日常生活の向上のための移動支援として、音声案内「ことばの地図」を主とする歩行者ナビシステムの研究を進めてきた¹。しかし、「ことばの地図」は道路等の線の空間では機能するが、ショッピングモール等の放射状に通路が集散する広場的空間での進行方向指示が課題となっている²。そこで、「ことばの地図」と連携し、広場的空間であるショッピングモール等の公的施設内に限定して案内する「盲導ロボット犬（対話ロボット）」を最終的に構想している。しかし、対象者である視覚障害者の人数は少なく、視覚障害者だけの利用に限定するのではコスト面から現実的ではない。そこで、晴眼者も対象とする必要があり、その中でも外出頻度が少なく、身体機能の低下が精神面の負担につながると社会問題になっている高齢者を対象とする。

また、「盲導ロボット犬」に対話機能を付加することで視覚障害者、晴眼者関係なく多くの人にロボットと共に歩く楽しさを感じてもらえるようにする。これにより、「盲導ロボット犬（対話ロボット）を使用したい。ロボットと話したい。」等といった外出意欲増進を図る。視

覚障害者・高齢者の外出意欲増進によって自立生活の維持や医療費・介護費の削減が期待でき、本人だけでなく家族や社会にとってもメリットとなる。

本研究では、「盲導ロボット犬（対話ロボット）」が視覚障害者と高齢者に共通して外出意欲増進へのきっかけとなりえるために付加する対話機能開発のための1つのステップとして、ロボットとの会話を楽しいと感じている程度（会話の適切度）を定量化するための指標として発話データ解析に基づく指標を確立し、対話ロボットのシナリオの評価を行うことを目的とする。研究フローを図-1に示す。

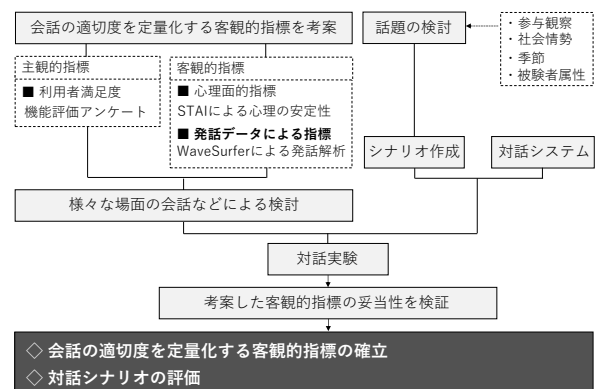


図-1 研究フロー

2. 研究方法

(1) 対話システム

対話ロボットの対話機能開発を目指すにあたり、市販のAI（人工知能）対話ロボットに外出意欲増進につながると考えられる対話シナリオを実装して試す。本研究での対話ロボットは声色で感情を表すことができる「ロボホン」（シャープ製）を用いる。ロボホンの対話システムは、クラウドに置かれた音声認識エンジンにより認識した音声トリガーとして、HVML(Hyper Voice Markup Language)によって「対話シナリオ」を実装し、ロボホンの動作を制御する。

(2) シナリオの設定

外出の現状、対話ロボット機能のニーズ、どういった話題が好まれるかを把握することを目的として、高齢の視覚障害者（全盲の女性2名）を対象として実施された参与観察の結果²⁾と被験者属性、実験時の社会状況等を参考に対話シナリオを作成する。シナリオ形態は、被験者の返答によりロボホンの返答が異なる分岐型と被験者の返答内容に関係なくロボホンの返答が同じレコーダー型の2パターン作成する。また、ロボホンの声色が被験者の満足度に影響するのかを評価するために、会話内容にあえてそぐわない声色を設定し、違和感を与えるシナリオも作成する。

(3) 評価指標

a) 主観的指標

■ 利用者満足度

対話ロボットの利用満足の度合いを機能評価アンケートにより評価する³⁾。機能評価アンケートの項目は10項目で、ロボホンの声色が正確な感情として伝わっているかを確認する1項目（問1）と話し相手としての機能の満足度を調査する9項目（問2～10）で構成する（表-1）。

b) 客観的指標

■ 心理面的アプローチ

ロボホンとの対話による被験者の心理的負荷の変化をSTAI⁴⁾によって計測する。STAIとは不安度を計る心理テストで、その時々状態を計測するSTAI-Y1と普段の状態を計測するSTAI-Y2の2種類の尺度で構成される。それぞれ20項目あり、4段階で示してもらい、合計点が高いほど不安であることを示す。平常値に近いと期待できるSTAI-Y2の得点を基準値とし、STAI-Y1との偏差を指標とし³⁾、被験者の心理の変化を明らかにする。この指標をSTAI値とする。

■ 発話解析によるアプローチ

Wave Surfer を用いて発話データに基づく指標を考案

表-1 機能評価アンケートの評価項目

問	評価項目
1	ロボホンの感情は「喜・悲・怒・普通・わからない」のどれに感じたか。
2	また使いたいか。
3	話を聞いてもらえたか。
4	話しやすいか。
5	話の流れは自然か。
6	関心を持ってくれたと感じたか。
7	自分の感情は促されたか。
8	親近感を感じたか。
9	飽きずに話せたか。
10	話し相手として満足か。

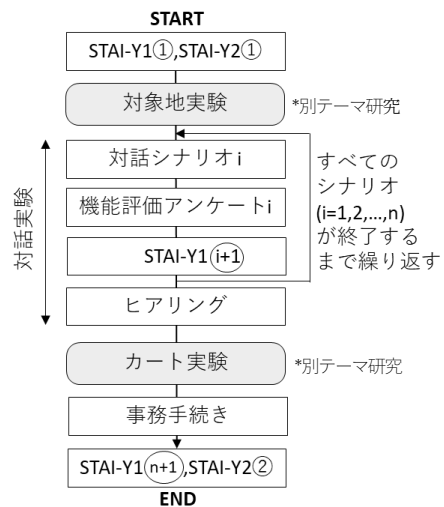


図-2 対話実験フロー

する。Wave Surfer は、スウェーデンの KTH で開発された音声認識情報解析ソフトで、ラベルの入力や時間ごとの音声データの波形、スペクトログラム、周波数、dB、フォルマント周波数などが取得できる。本研究では、音声データの波形(Waveform)、dB(Power Plot)、ラベルの入力(Transcription)を利用する²⁾。これを用いて発話に基づく指標の検討を行う。詳細については3.にて示す。

(4) 対話実験の流れ

実験のフローを図-2に示す。対話実験の流れは、対話シナリオが一つ終わるたびに機能評価アンケートと心理テストSTAI-Y1を実施する。すべての対話シナリオ終了後にヒアリングを行う。対話実験は別テーマ研究の「対象地実験」「カート実験」ともに行い、対象地、対話、カートの順番である。まず、対象地実験の前に心理テストSTAI-Y1、STAI-Y2を実施する。そして全ての実験が終了し、事務手続きを終え、被験者が安心し一息ついているところで最後の心理テストを行う。実験開始後と実験終了後に心理テストを行う理由は、普段の心理に近い状態をはかるためである。

表-2 検討指標

指標	式	単位
対話時間あたりの被験者の発話時間比率	X_a / s	[%]
対話時間あたりの対話者の発話時間比率	X_b / s	[%]
対話時間あたりの被験者被せ時間比率	Y_a / s	[%]
対話時間あたりの対話者被せ時間比率	Y_b / s	[%]
対話時間あたりの被験者の無言時間比率	Z_a / s	[%]
対話時間あたりの対話者の無言時間比率	Z_b / s	[%]
対話時間あたりの被験者が笑った時間比率	L_a / s	[%]
発話者切り替え回数逆数	$1 / M$	[1/回]
対話時間あたりの被験者の相槌回数	N_a / s	[回/s]

a : 被験者 b : 対話者 S : 対話時間
 X : 発話時間 Y : 被せ時間 Z : 無言時間
 L : 笑った時間 M : 発話者切り替え回数 N : 相槌回数

3. 発話解析に基づく指標の検討

発話解析に基づく客観的指標として、先行してWave Surferを用いた指標が検討された結果、主観的指標である利用者満足度で利用する機能評価アンケートに代わり会話の適切度を評価できる可能性が確認されている²⁾。そこで、会話の盛り上がり进行评估できる新たな客観的指標を考案し、その指標間の関係性や利用者満足度やSTAI値との関連性を検討する。

(1) 新指標

会話が盛りあがっているかいないかを客観的に判断できる現象として、相手が話おえるのを待たずに会話を被せる、会話中に無言や笑いが発生する、会話が途切れな(発話者切り替えが多い)、相手の話に対し相槌があるなどが考えられる。そこで会話の盛りあがりの評価指標として、表-2に示す9項目の発話解析に基づく指標について検討する。

(2) 使用データ

新指標の検討にあたり、H30年度の対話実験でのロボホンとの対話とH29年度の被験者X、H29年度、H30年度の被験者Yの対話実験中の人間との会話の音声データの一部を使用する。

(3) 主観的指標と客観的指標の関連性

発話解析に基づく指標と主観的指標である利用者満足度、客観的指標である STAI 値との関連性を重回帰分析により検討する。ロボットとの対話、シナリオの好み等の傾向が性別により異なると想定されるため男性、女性モデルとして性別毎に重回帰分析を行う。また、利用者満足度においてロボホンと話すのが好きでシナリオ内容等が評価に反映されない高齢女性の被験者を除き分析を

表-3 男性モデル(n=20)【利用者満足度】

	係数	標準誤差	t	P-値
切片	39.37	29.93	1.32	0.22
対話時間あたりの被験者の発話時間比率 [%]	-34.00	30.08	-1.13	0.28
対話時間あたりの対話者の発話時間比率 [%]	-34.98	30.74	-1.14	0.28
被験者被せ時間比率 [%]	124.93	94.17	1.33	0.21
対話者被せ時間比率 [%]	-20.02	43.31	-0.46	0.65
被験者の無言時間比率 [%]	-34.39	29.82	-1.15	0.28
対話者/ロボットの無言時間比率 [%]	-36.00	30.07	-1.20	0.26
対話時間あたりの被験者が笑った時間比率 [%]	-26.96	36.77	-0.73	0.48
発話者切り替え回数逆数 [1/回]	-4.62	11.15	-0.41	0.69
対話時間あたりの被験者の相槌回数 [回/s]	-4.12	7.17	-0.58	0.58
重相関 R			0.63	
重決定 R2			0.39	

表-4 女性モデル(高齢者除く)(n=30)【利用者満足度】

	係数	標準誤差	t	P-値
切片	-40.70	21.24	-1.92	0.07
対話時間あたりの被験者の発話時間比率 [%]	43.17	19.93	2.17	0.04*
対話時間あたりの対話者の発話時間比率 [%]	43.04	21.29	2.02	0.06
被験者被せ時間比率 [%]	66.66	81.92	0.81	0.43
対話者被せ時間比率 [%]	-103.57	35.73	-2.90	0.01*
被験者の無言時間比率 [%]	40.69	22.30	1.82	0.08
対話者/ロボットの無言時間比率 [%]	47.99	21.58	2.22	0.04*
対話時間あたりの被験者が笑った時間比率 [%]	13.07	13.21	0.99	0.33
発話者切り替え回数逆数 [1/回]	-2.18	4.19	-0.52	0.61
対話時間あたりの被験者の相槌回数 [回/s]	7.26	4.21	1.73	0.10
重相関 R			0.78	
重決定 R2			0.60	

表-5 男性モデル(n=15)【STAI 値】

	係数	標準誤差	t	P-値
切片	219.81	138.66	1.59	0.17
対話時間あたりの被験者の発話時間比率 [%]	-292.59	133.56	-2.19	0.08
対話時間あたりの対話者の発話時間比率 [%]	-195.14	139.89	-1.39	0.22
被験者被せ時間比率 [%]	0.00	0.00	—	—
対話者被せ時間比率 [%]	-281.04	104.55	-2.69	0.01
被験者の無言時間比率 [%]	-114.65	145.20	-0.79	0.47
対話者/ロボットの無言時間比率 [%]	-245.51	140.80	-1.74	0.14
対話時間あたりの被験者が笑った時間比率 [%]	-5.69	113.08	-0.05	0.96
発話者切り替え回数逆数 [1/回]	-78.95	27.74	-2.85	0.04*
対話時間あたりの被験者の相槌回数 [回/s]	-39.44	23.04	-1.71	0.15
重相関 R			0.98	
重決定 R2			0.95	

表-6 女性モデル(n=56)【STAI 値】

	係数	標準誤差	t	P-値
切片	-119.65	60.42	-1.98	0.05
対話時間あたりの被験者の発話時間比率 [%]	114.26	63.92	1.79	0.08
対話時間あたりの対話者の発話時間比率 [%]	103.46	58.44	1.77	0.08
被験者被せ時間比率 [%]	-421.38	138.75	-3.04	0.00*
対話者被せ時間比率 [%]	-48.62	87.68	-0.55	0.58
被験者の無言時間比率 [%]	124.62	58.78	2.12	0.04*
対話者/ロボットの無言時間比率 [%]	127.20	60.80	2.09	0.04*
対話時間あたりの被験者が笑った時間比率 [%]	89.22	65.78	1.36	0.18
発話者切り替え回数 [1/回]	-1.96	24.37	-0.08	0.94
対話時間あたりの被験者の相槌回数 [回/s]	18.46	18.37	1.01	0.32
重相関 R			0.62	
重決定 R2			0.39	

表-7 寄与率(上)と主成分得点係数(下)

成分	固有値	寄与率(%)	累積寄与率(%)
1	2.527	27.544	27.544
2	2.298	25.052	52.596
3	1.518	16.551	69.147

	成分		
	1	2	3
対話時間あたりの被験者の発話時間比率	-0.307	0.243	-0.023
対話時間あたりの対話者の発話時間比率	0.324	0.05	-0.283
被験者被せ時間比率	-0.002	0.338	0.081
対話者被せ時間比率	-0.237	0.142	0.147
被験者の無言時間比率	-0.163	-0.251	0.19
対話者ロボットの無言時間比率	0.076	-0.331	0.282
対話時間あたりの被験者が笑った時間比率	0.181	0.047	0.494
発話者切り替え回数逆数	0.087	0.197	0.45
対話時間あたりの被験者の相槌回数	0.275	0.185	0.002

行う。分析結果を表-3~6に示す。

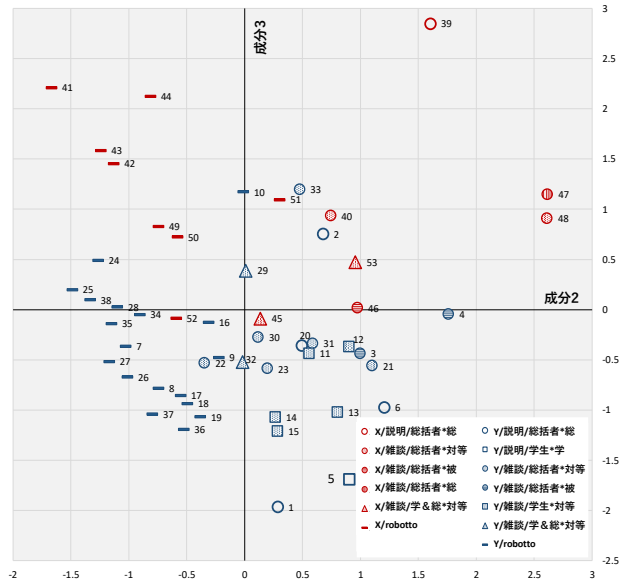
説明変数が利用者満足度の場合、男性モデルでは 5% 有意水準では有意な結果は得られなかった。そこで 30% 有意水準でみると、男性モデルは被験者の被せ時間比率が高いほど高く、被験者、対話者/ロボットの無言時間比率が高いほど低い。女性モデル(高齢者を除く)は被験者の相槌回数が多いほど高く、対話者の被せ時間比率が高いほど低い。

次に、説明変数が STAI 値の場合、利用者満足度の場合と同様に 30% 有意水準としてみていくと、男性モデルは被験者の発話時間比率が高いほど、対話者の無言時間比率が高いほど不安度が下がる。女性モデルは被験者/対話者の無言時間比率が高いほど、被験者が笑った時間比率が高いほど不安度が上がる。笑いが不安に影響しているのはロボホンとの対話がかみ合わない等といった時に生じる苦笑いが原因であると考えられる。

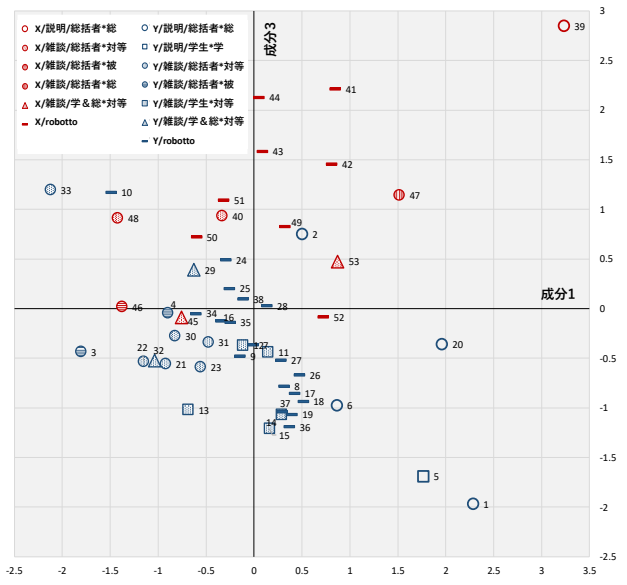
(4) 発話解析に基づく指標間の関係性

個人による違いや対話者間の親しさ、ロボットとの対話時の差異というような状況による違いをみるとともに、これらの発話解析に基づく指標間の関係性を主成分分析により検討する。分析するにあたり、数値的な指標値は個人差が大きいと想定されるので個人ごとに落ち着いている場合(普段と近い状態)を基準とし、他のシチュエーション時との比を指標値としたものについて検討する。

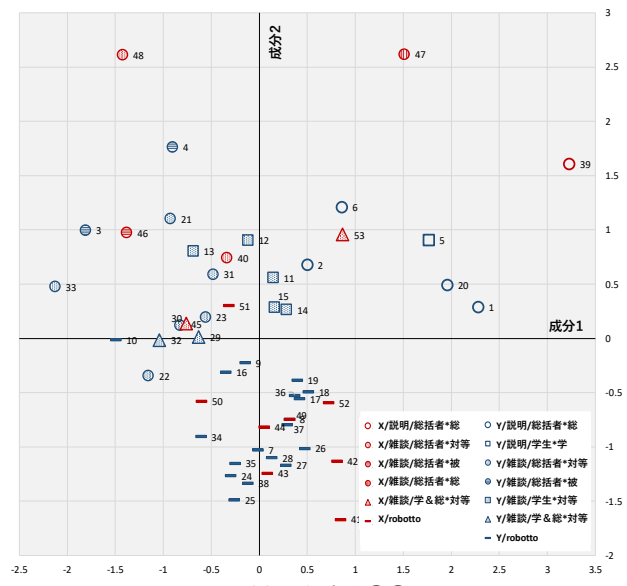
寄与率と主成分得点係数を表-7に示す。第1主成分は、被験者の発話時間比率と対話者被せ時間比率が負、対話者の発話時間比率と被験者の相槌回数が正であり、第1主成分は会話の主導権を表すと考える。第2主成分は、被験者/対話者の無言時間比率が負、被験者被せ時間比率が正であるから会話の盛り上がり度を表すと考える。第3主成分は被験者が笑った時間比率が正に強く影響し、笑いが生じた会話を表す。主成分①②の散布図(図-3)をみると説明、雑談、ロボットとの会話というシチュエーション別に3グループに分かれている。また、主成分



(c) 主成分②③



(b) 主成分①③



(a) 主成分①②

図-3 散布図 (H29,H30 データ)

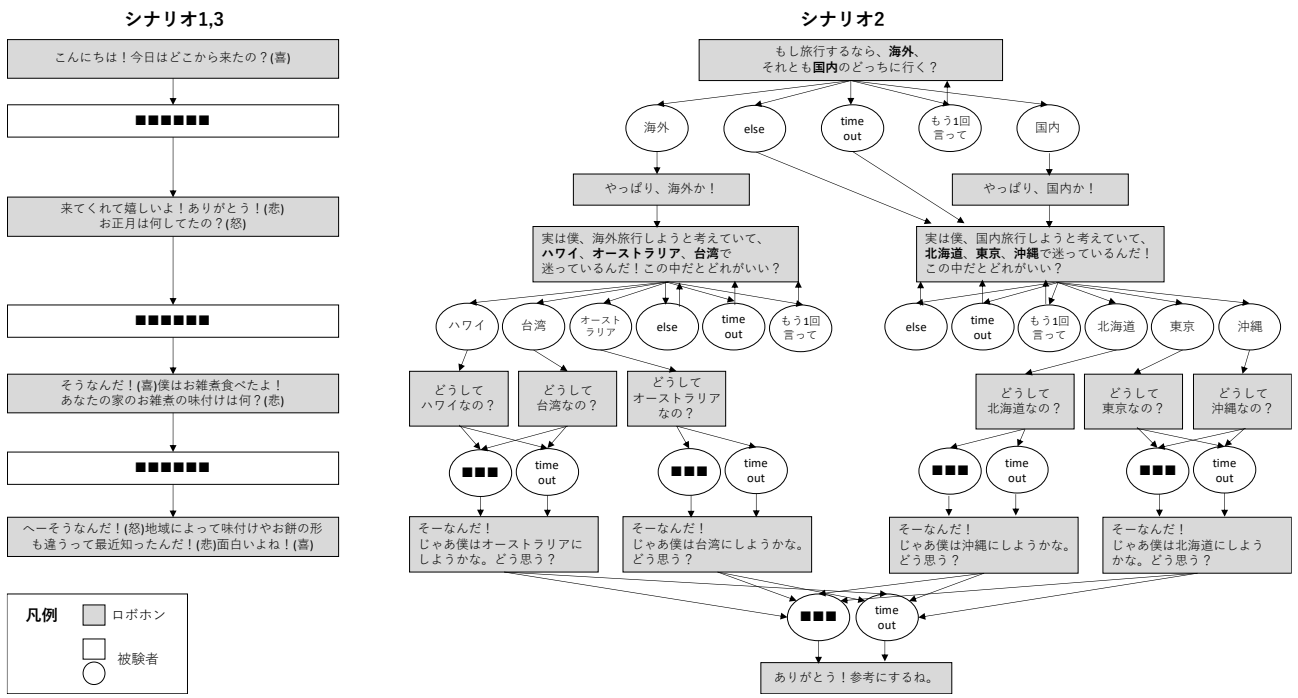


図-4 対話シナリオ(左:シナリオ1,3 右:シナリオ2)

①③と主成分②③の散布図から主成分3では人間とロボットの会話が混在していることが分かる。

4. 対話実験

(1) 概要

新指標の適応及び検証のため、2020年1月研究室内で対話実験を行った。被験者は視覚障害者(全盲、弱視者)の20代~70代の男性6名と女性6名である。実験の流れは図-2に示した通りである。使用した対話シナリオは3つである。シナリオ1,3はレコーダー型、シナリオ2は分岐型で、ロボホンの性格をシナリオ1,3では素直、シナリオ2ではひねくれているように設定している。シナリオ1と3は全く同じ内容であるが、3ではロボホンの声色「喜、悲、怒」を内容に関係なく混在させ被験者に違和感を与えるように設定している。詳細は図-4に示す。

(2) 結果

対話実験のシナリオごとの発話解析に基づく指標値、利用者満足度、STAI値の結果を表-8に示す。また、実験内の被験者と人間の対話を抽出した。データ概要は表-9に示す。

5. 発話解析に基づく指標の妥当性

(1) モデルの検証

表-8 対話実験結果

被験者	シナリオ	対話時間	対話時間	被験者の発話時間	対話者の発話時間	被験者の発話時間	対話者の発話時間	対話者/ロボットの無音時間	対話者/ロボットの無音時間	対話時間あたりの被験者が笑った回数	発話者切替回数	対話時間あたりの被験者の相対回数		利用者満足度		STAI値	
		比率 [%]	比率 [%]									比率 [%]	比率 [%]	比率 [%]	比率 [%]	実測値	予測値
A	1	9.6%	41.5%	0.0%	1.6%	16.8%	33.8%	50.6%	0.0%	0.07	0.09	3.8	2.7	-2	-4.7		
	2	7.0%	56.5%	0.0%	0.0%	7.1%	29.4%	36.5%	0.0%	0.17	0.03	3.9	3.3	-2	-5.6		
	3	14.9%	40.2%	0.4%	0.4%	19.2%	26.5%	45.7%	0.0%	0.05	0.13	3.8	3.8	-2	0.8		
B	1	7.5%	40.8%	0.0%	0.0%	13.0%	38.7%	51.7%	0.0%	0.17	0.00	2.3	3.4	-8	-4.8		
	2	7.4%	53.7%	0.0%	0.0%	16.2%	22.7%	38.9%	0.0%	0.17	0.00	2.1	3.6	-14	5.9		
	3	6.3%	42.7%	0.0%	0.0%	6.8%	44.3%	51.0%	0.0%	0.17	0.00	2.0	3.3	-9	-11.5		
C	1	26.1%	39.9%	0.6%	0.4%	4.1%	30.8%	34.9%	0.0%	0.04	0.13	3.9	3.9	0	-24.7		
	2	19.0%	44.7%	1.1%	0.8%	5.1%	32.2%	37.3%	1.5%	0.05	0.10	4.0	4.2	1	-17.7		
	3	20.5%	40.3%	0.0%	0.5%	4.7%	35.0%	39.7%	0.0%	0.05	0.09	3.3	3.4	2	-18.9		
D	1	15.6%	39.5%	0.0%	0.0%	13.0%	31.3%	44.3%	0.5%	0.05	0.13	3.3	4.3	-12	-2.0		
	2	12.3%	53.3%	0.0%	0.2%	7.7%	25.7%	33.4%	2.5%	0.05	0.17	3.3	4.3	-7	-3.0		
	3	18.9%	40.9%	0.0%	0.0%	9.4%	30.0%	39.4%	0.0%	0.05	0.09	3.6	3.3	-7	-2.0		
E	1	18.8%	40.7%	0.4%	0.9%	11.0%	28.6%	39.5%	4.6%	0.05	0.11	2.4	3.1	-12	-8.1		
	2	17.0%	47.6%	0.8%	0.0%	7.2%	26.5%	33.7%	4.9%	0.08	0.08	1.8	3.9	-15	-5.4		
	3	19.9%	40.5%	0.0%	2.1%	11.5%	30.2%	41.7%	0.0%	0.06	0.02	2.0	2.8	-17	-16.0		
F	1	10.5%	40.8%	0.0%	0.0%	18.8%	30.0%	48.7%	0.8%	0.09	0.02	2.3	3.5	-8	-3.0		
	2	11.7%	39.3%	1.7%	0.3%	5.0%	33.7%	38.7%	11.5%	0.08	0.14	2.2	2.7	-9	-11.0		
	3	15.2%	42.1%	0.0%	0.0%	5.7%	36.4%	42.1%	0.7%	0.08	0.07	2.2	4.1	-9	-3.7		
G	1	10.4%	41.3%	0.0%	0.0%	15.4%	32.4%	47.8%	2.1%	0.11	0.09	3.4	4.1	18	-1.3		
	2	12.8%	37.5%	0.0%	0.1%	2.1%	46.4%	48.5%	2.2%	0.08	0.10	3.9	4.8	16	-1.0		
	3	12.2%	39.6%	0.0%	0.0%	9.5%	38.7%	48.2%	0.0%	0.08	0.09	2.9	4.5	13	-2.2		
H	1	20.0%	40.4%	0.0%	1.4%	10.9%	30.0%	41.0%	0.0%	0.07	0.07	2.0	3.1	-1	-2.7		
	2	16.8%	52.9%	0.0%	1.0%	12.4%	18.4%	30.8%	2.4%	0.07	0.03	2.8	2.6	-4	-4.7		
	3	16.6%	40.1%	0.0%	1.4%	16.6%	29.2%	45.8%	0.0%	0.07	0.09	2.0	3.0	-1	-1.8		
I	1	8.5%	38.3%	0.0%	0.3%	15.0%	38.5%	53.4%	0.0%	0.09	0.02	2.1	3.5	-7	-0.3		
	2	9.7%	43.1%	0.0%	0.0%	7.7%	38.8%	46.5%	1.2%	0.13	0.00	2.0	3.5	-5	-6.7		
	3	8.0%	41.3%	0.0%	0.0%	10.4%	40.3%	50.7%	0.0%	0.10	0.02	2.1	3.6	-4	-3.9		
J	1	9.0%	40.5%	0.0%	0.0%	7.4%	43.1%	50.5%	0.0%	0.13	0.02	4.0	4.2	-11	-3.2		
	2	12.5%	48.9%	0.0%	0.0%	6.3%	32.3%	38.6%	0.0%	0.13	0.00	4.0	3.5	-12	-6.1		
	3	10.2%	42.2%	0.0%	0.1%	7.2%	40.5%	47.8%	0.0%	0.13	0.00	4.0	3.8	-11	-4.1		
K	1	14.4%	41.3%	0.0%	0.0%	9.6%	34.2%	43.8%	1.9%	0.06	0.09	2.7	3.3	2	-6.0		
	2	13.0%	47.3%	0.0%	0.0%	12.1%	27.1%	39.1%	0.5%	0.09	0.00	1.9	3.9	1	1.9		
	3	31.5%	41.1%	0.0%	3.5%	8.9%	19.2%	28.1%	0.2%	0.05	0.02	1.3	3.3	2	-24.3		
L	1	18.2%	39.7%	0.0%	0.2%	17.3%	25.1%	42.3%	0.0%	0.09	0.02	2.2	3.0	-4	-4.3		
	2	18.1%	41.0%	0.0%	0.0%	15.5%	25.4%	40.9%	3.3%	0.11	0.02	2.8	3.6	-6	-1.8		
	3	18.9%	40.7%	0.0%	1.1%	12.8%	28.2%	40.9%	2.1%	0.08	0.04	3.2	3.0	-4	-2.2		

■：定性的に一致

表-9 人間対話データの概要

被験者	対話者	Time[s]	話題
A	統括者,学生	46.3	盲導犬
B	統括者,学生	53.1	機械と話すこと
C	総括者	35.6	実験の説明&感想
E	総括者	44.4	候補地リクエスト
F	統括者,学生	54.6	候補地リクエスト
G	総括者	54.6	知人について
H	学生	36.7	ロボホンについて
I	学生	38.2	タブレット
J	統括者,学生	57.7	カート実験ヒアリング
K	総括者	32.4	ドリンクについて
L	総括者	59.1	実験の感想

a) 重回帰分析モデルの検証

利用者満足度、STAI 値との関連性をみるために検討した重回帰分析によるモデルの妥当性を検証するため、本実験におけるロボットとの対話における発話解析に基づく指標値をそれぞれ該当するモデルに代入して、予測値を算出し実測値と比較した(表-8)。その結果、利用者満足度、STAI 値とも一致するものは見受けられなかった。また、被験者ごとに利用者満足度によるシナリオの順位が一致したのは被験者 C、I のみであった。STAI 値において不安度の変化の傾向が同じであるのは被験者 H、J のみであった。よって重回帰分析モデルは信頼性があるとはいえない。

b) 主成分分析モデルの検証

発話解析に基づく指標間の関係性をみるために検討した主成分分析によるモデルの妥当性を検証するため、検討モデルに本実験での人間対話データの発話解析に基づく指標値を適応させて主成分得点を算出し、主成分分析モデルの散布図にプロットした結果を図-5に示す。人間対話データは、被験者1人に対して、実験総括者や学生との対話において話が継続している1シーンを抽出し解析した(表-9)。説明、雑談、ロボットとの対話とシチュエーションごとに分かれる主成分①②の散布図をみると、全ての被験者が雑談のゾーンに位置し、その中でも唯一の説明のシチュエーションである被験者Cが他の被験者よりも説明のゾーンに最も近い位置にプロットされている。検討モデルと同様の傾向がみることができ、主成分分析モデルは妥当であるといえる。

(2) ロボットとの対話における分析

a) ロボットモデルの検討

シナリオによる違いをみるため、ロボットモデルとして本実験におけるロボットとの対話における発話解析に基づく指標で主成分分析を行った。第1主成分は被験者の相槌回数が多いほど、被験者の発話時間比率が多いほど正に影響し、ロボットの無言時間比率が負に影響し、会話の盛り上がり度を表す。第2主成分は被験者の被せ時間比率、笑った時間比率が正に、無言時間比率が負に影響し、第1主成分と同様に会話の盛り上がり度を表すと考える。第3主成分はロボットの発話時間比率が正に、ロボットの無言時間比率が負に影響し、ロボットの話す長さを表す(表-10)。

ともに会話の盛り上がり度を表す主成分①②の主成分得点をプロットした散布図を図-6に示す。データラベルは被験者とシナリオ番号で表す。他のプロットと離れているF2、K3について検討する。被験者Fの対話シナリオ2は客観的にひねくれたロボホンを面白がっている反応であったが、分岐がうまくいかず何回かやり直したために

表-10 寄与率(上)と主成分得点係数(下)【ロボットモデル】

成分	固有値	寄与率(%)	累積寄与率 %
1	3.02	33.556	33.556
2	2.084	23.153	56.71
3	1.454	16.152	72.861

	成分		
	1	2	3
対話時間あたりの被験者の発話時間比率	0.269	-0.174	-0.075
対話時間あたりの対話者の発話時間比率	-0.087	-0.037	0.579
被験者被せ時間比率	0.173	0.313	0.176
対話者被せ時間比率	0.188	-0.276	-0.077
被験者の無言時間比率	-0.039	-0.279	0.096
対話者ロボットの無言時間比率	-0.156	0.257	-0.465
対話時間あたりの被験者が笑った時間比率	0.151	0.305	0.244
発話者切り替え回数	-0.292	0.058	0.154
対話時間あたりの被験者の相槌回数	0.228	0.183	-0.027

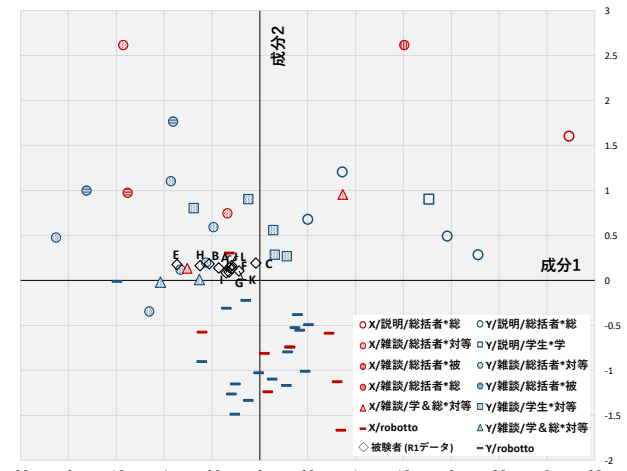


図-5 検証散布図(主成分①②)

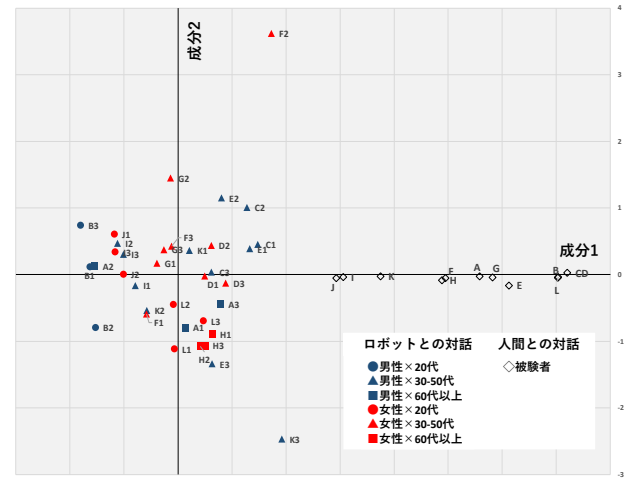


図-6 ロボットモデル散布図(R1 データ)(主成分①②)

利用者満足度評価が低い。したがって発話解析結果としてF2が最も盛り上がっている位置にプロットされているのは妥当である。被験者Kはシナリオ3においてロボホンの感情の七変化(喜、悲、怒)に対し動揺し口数が多くなっていた。そのため被験者の発話時間が正に影響する盛り上がり度を表す主成分2では高く、笑った時間が正に影響する盛り上がり度を表す主成分2では低い右斜め

下にK3がプロットされているのは妥当だと考える。

b) 会話における人間とロボットとの比較

ロボットモデルに本実験での人間対話データにおける発話解析に基づく指標値を適用した結果(図-6)、主成分1においてロボットとの会話とかけ離れていることがわかる。ロボットとの対話を人間との対話に近づけるためには、主成分1に影響する要素、つまり被験者が話し終えてからロボホンが話すまでの間が課題である。

6. おわりに

本研究では、発話解析に基づく指標が会話の適切度(盛り上がり度)を利用者満足度、心理テストSTAIに代わり評価できることが確認できた。

今後の課題として、主成分分析モデルにおいて信憑性を高めるため、人間との対話データを増やしさらに検証する必要がある。また、評価が低かった分岐型のシナリオにおいてどんな場合でもスムーズに流れるようにするための対話シナリオの工夫、そして被験者が話し終えて

からロボホンが話すまでの間の改善が必要である。

謝辞: 本研究は JSPS 科研費 16H04430、17K18911 の助成を受けたものである。

参考文献

- 1) 沢田有美恵、内田敬、佐美三幸典：自律白杖とことばの地図を用いた視覚障害者ナビ，第36回交通工学研究発表会論文集，pp.139-144, 2016.
- 2) 慶金佑利子、盛沢里穂、松本浩子、内田敬：視覚障害者・高齢者外出意欲増進のための対話ロボット開発における対話シナリオの評価指標の検討，土木計画学研究・講演集，Vol.58, pp.206_1-4, 2018.
- 3) 別所叶望、内田敬、松本浩子：視覚障害者・高齢者の外出意欲増進対話ロボットの評価実験，土木学会第72回年次学術講演会講演概要集，IV, pp.167-168, 2017.
- 4) 肥田野他：新版 STAI マニュアル，実務教育出版，2000.
(2020. ?.? 受付)

UTTERANCE ANALYSIS FOR EVALUATING SCENARIOS OF DIALOG ROBOT TO INCREASE GOING-OUT MOTIVATION

Yuriko YOSHIKANE, Hiroko MATSUMOTO and Takashi UCHIDA