

未知の人との電話対話における印象評価分析 —話題や対話相手ごとの印象評価と音響特徴—*

○高椋琴美, 谷田泰郎 (シナジーマーケティング株式会社)

1 はじめに

人のコミュニケーションはいい加減であるが、そこに人間らしさや個性がある。我々は音声対話から得られる人の普遍性や個性など「心のモデル」に通じる知見を音声対話システムに応用することで、言語理解や人間理解に役立てたいと考えている。

先行研究[1]では小説を題材に、内容を「聞いた」場合と「見た」(黙読した)場合の記憶情報の違いについて、伝達内容の言語特徴や朗読者の音響特徴、受信側の性別や個性性などの観点から調査を行った。また音響特徴と言語特徴を用いた重要箇所抽出(=記憶に残る箇所)を試み、人間と同程度の結果を得ることが出来た。しかし、[1]の実験は受信方向のみの実験であり、双方向のコミュニケーションではなかった。また、比較的解析しやすいプロの朗読者による恣意的な音声データであり、自然対話の音声データの質とは程遠いものであった。本稿では、自然対話の実験を行うことで取得した音声対話データのいくつかを取り上げ、記憶や印象に残った情報の音響特徴や言語特徴と話者の特性についての分析結果を報告する。

2 対話実験の内容

2.1 実験の目的

実験は、1対1の自由対話における記憶や印象と、話者の個性や属性、個性×個性の関係、音響特徴や言語特徴などについて被験者ごとに観察し、これらの関係性を比較することを目的としており、なるべく被験者の先入観を排除し、音声言語情報のみから評価を行ってもらった必要がある。

2.2 実験概要

初対面の被験者どうしを集め、被験者にそれぞれ別の会議室に入ってもらい、30分間電話で自由に対話してもらった。全員が初対面であることを考慮し、話をしやす

いように、入社3年目までの社会人(22-25歳)6名(男女各3名ずつ)を被験者として集めた。また、各自が記入した自己紹介シート(項目:出身地・職業・あだ名・好きなこと・休日の過ごし方・最近気になっていること)を全員に配布し、対話中の参照を可とした。音声は被験者ごとにIC Recorderで録音した(LPCM22.05kHz 16bit)。対話終了後、被験者が記憶している話題と対話内容の書き起こし、その話題に関する印象(いい・悪い・どちらでもない)、対話相手の総合的な印象評価を記述してもらった。

全対話終了後、被験者本人に関する価値観アンケートと実験に関するアンケート、及び個別インタビューを実施した。実験の流れを以下に示す。(実験詳細は[2]を参照頂きたい)。

- 1) 自己紹介シート記入
- 2) 1対1の自由対話(30分)
- 3) 記憶している対話内容の書き起こし(20分)
- 4) 相手の印象についてのアンケート(10分)
- 5) 価値観アンケートと感想記述(10分)
- 6) 個別インタビュー(10分)

各被験者は2)~4)を5回繰り返すことで全被験者と対話してもらい、15対話セット、30音声データ(雑音が入ったため音響処理に使用可能なのは24音声データ)を得た。

3 対話実験の分析と考察

最初に対話セットごとの特徴と相手の印象について俯瞰し、その中から特徴的な対話セットを選択した。選択した対話セットについて、話題に関する印象と、話題ごとの音響特徴と言語特徴(以降あわせて対話特徴と呼ぶ)の両方を比較分析した。

3.1 対話セットごとの特徴分析

対話セットごとに対話相手の印象と互いの対話特徴を比較した。

3.1.1 対話相手の印象評価

対話終了後に行った対話相手の印象評価結果をTable 1に示す。被験者を①~⑥、評価を

* Analysis of impression on telephone conversation between strangers, by TAKAMUKU, Kotomi and TANIDA, Yasuo (Synergy Marketing, Inc.)

5段階（5かなり気が合いそうだ／4.気が合いそう／3 どちらともいえない・まだわからない／2.あまり気が合いそうにない／1.気が合わない）で表記している。

Table 1 対話相手の印象評価

被評価者	評価者	評価者						被評価平均
		男性			女性			
		①	②	③	④	⑤	⑥	
男性	①		4	4	1	3	2	2.8
	②	4		5	3	4	4	4.0
	③	4	4		4	3	5	4.0
女性	④	2	3	4		3	3	3.0
	⑤	3	5	4	4		5	4.2
	⑥	4	5	5	3	5		4.4
評価平均		3.4	4.2	4.4	3.0	3.6	3.8	

女性は評価にばらつきがある。男性では、被験者③は全員に対し4以上の評価を付けておりばらつきは小さいが、他2名の被験者は女性に対してのみ評価にばらつきがみられる。

被験者①と被験者④は互いに最低の評価をつけており、他者への評価も他者からの評価も全体的に低めでばらつきがある（低めの評価をするのは心理的抵抗が大きいと思われる）。このような理由から、被験者①と被験者④を中心に印象と対話の特徴を調査した。

3.1.2. 対話ごとの音響特徴

音声分析ソフト Praat version 5.4.09 を用い、それぞれの音声を発話区間と無音区間とに分類し（無音区間条件＝最小ピッチ：10Hz/無音閾値(dB)：-25/無音時間：0.75 以上/音声最小時間：0.1）、SoundEngine Free ver.5.02 で無音区間を削除した発話区間のみの音声ファイルを作成し、Praat を用いて発話区間全体の音響特徴量（Pitch と Intensity の平均・標準偏差・最大値・最小値）を取得した。音声には雑音も含まれており（音声データから雑音を除去する際に一部音声と雑音が重なり除去できないものがあつた）、今回は平均と標準偏差のみを採用した。被験者①と被験者④の対話ごとの音響特徴を Fig.1 と Fig.2 に示す。

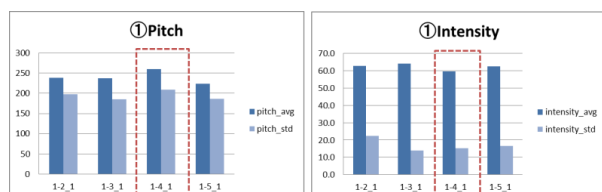


Fig. 1 被験者①の音響特徴

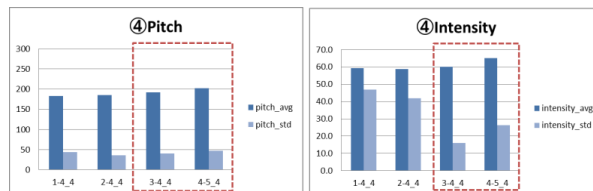


Fig. 2 被験者④の音響特徴

被験者①は、「あまり気が合いそうにない」と回答した被験者④との対話では他と比べ Pitch 平均が高く、Intensity 平均が小さい。被験者①は緊張したりストレスを感じている場面で、声が高くなったり小さくなっていると考えられる。被験者④は女性にしては声が高く抑揚も少ないが、「気が合いそう」と回答した被験者③と被験者⑤との対話では Pitch 平均と Intensity 平均が高めで Intensity 標準偏差が小さい。ネガティブな感情時には低く小さくなるなどの特徴が声に表れると考えられる。

次に、30 分間の対話時間中の各被験者の発話時間を発話割合として Table2 に示す。

Table 2 対話時間中の発話割合

対話	話者	相手の印象	発話割合
①②	①	4	48.4%
	②	4	52.7%
①③	①	4	59.6%
	③	4	39.9%
①④	①	2	49.9%
	④	1	32.6%
①⑤	①	3	47.2%
	⑤	3	70.2%
①④	④	1	32.6%
	①	2	49.9%
②④	④	3	42.1%
	②	3	56.8%
③④	④	4	37.6%
	③	4	37.9%
④⑤	④	4	47.9%
	⑤	3	79.5%

被験者①と被験者④の対話（以降対話①④とする）では、被験者④の発話割合が 32.6%、被験者①の発話割合が 49.9%とバランスが悪い。同様に、対話④⑤でも被験者④の発話割合 47.9%に対し被験者⑤は 79.5%でバランスが悪いが、こちらは被験者④からの評価が良い。発話時間には相槌なども含まれており、発話割合だけでは印象評価との関係は判断できない。そこで、対話①④と、その比較対象として対話③④（被験者③は被験者④からの評価が高く被験者①と同じ男性である）を採用して言語特徴を調査した。

3.1.3. 対話ごとの言語特徴

「無意味ではないが、無意味に思える発言、すなわち相槌や同調、相手の発話の繰り返し、内容語を含まない発話、を除いたもの」を『有意義発話』と定義し抽出した (Fig.3)。

①印象	④印象	①書き起こし原文	④書き起こし原文	
話題探索	0	1	そうですね、あの、ま、好きっちゃ好きなんすけど、これまた、人数が集まらへんと出来へんものが好きなんでやっばり、あの～、ま、ボーリングは別ですけどね	スポ、スポーツ、好きなんですか？
			テニスと野球っていうのは、ちょっと何か、集まらへんと厳しいかなあ～と集まりませぬ(w)	はい
話題探索	0	1	声かけても、みんなして、仕事の予定があって、言いながら、入れ違うんでね、はい	あ、そうですね(w)え、集まってるんですか～、今んとこ(w)
			うん	そうですね～、野球なんか難しいっすね(w)

Fig. 3 有意味発話の抽出

Fig.3 で色をつけた単位で話者の主導権が入れ替わっているものと考え（本稿ではこれを『ターン』と呼ぶ）、それぞれのターン数を数えた。また、抽出した書き起こし文から「(W)」や「…」などの不要な記号を取り除き、肯定発話は”。”、疑問発話は“？”で終わるように修正した（③修正後、④修正後の列）。また、対話相手が殆ど反応していない場合は、発話文としての意味内容が成立する単位で発話を接続する場合もあった。こうして作成したデータから話題ごとに以下の情報を取得した。

Table 3 言語特徴に用いた指標

指標	意味
ターン数	話者の主導権が入れ替わっている数
発話文字数	1発話者の発話文字数
発話数	1発話者の発話数
肯定発話数	1発話者の肯定発話数
疑問発話数	1発話者の疑問発話数
ターン平均発話文字数	1発話者の1ターンにおける発話文字数の平均
ターン平均発話数	1発話者の1ターンにおける発話数の平均
疑問発話割合	疑問発話数÷発話数
発話文字数倍率	1発話者の発話文字数÷対話の相手の発話文字数
発話数倍率	1発話者の発話数÷対話の相手の発話数

Table 4 対話①④の言語特徴

①の総合評価	④の総合評価	①ターン	④ターン	①発話文字数	④発話文字数	①発話文字数倍率	④発話文字数倍率
2	1	110	103	8513	2976	2.9	0.3
		213		11489			
①肯定発話数	④肯定発話数	①疑問発話数	④疑問発話数	①発話数	④発話数	①発話数倍率	④発話数倍率
354	137	13	29	367	166	2.2	0.5
491		42		533			

Table 5 対話③④の言語特徴

③の総合評価	④の総合評価	③ターン	④ターン	③発話文字数	④発話文字数	③発話文字数倍率	④発話文字数倍率
4	4	165	158	4387	5392	0.8	1.2
		323		9779			
③肯定発話数	④肯定発話数	③疑問発話数	④疑問発話数	③発話数	④発話数	③発話数倍率	④発話数倍率
168	242	78	47	246	289	0.9	1.2
410		125		535			

対話①④と対話③④を比較すると、発話数は同程度だが、対話①④は発話文字数が多い。特に被験者①の発話文字数が多く、発話文字

数倍率は被験者④の 2.9 倍である。被験者④は被験者①との対話では聞き手であったことが分かる。対話③④では発話量のバランスもよく、疑問発話数が多いことから互いに興味を示していたと言える。対話①④の総合評価が低いのは、話し手と聞き手の発話量の差が大きいこと、発話量の割に疑問発話の数が少ないことなど、被験者①が一方的に話していたことに起因していると考えられる。

3.2 話題ごとの特徴分析

対話を複数の話題に分割し、話題ごとの特徴と印象との関係について比較した。

3.2.1. 話題定義と記憶と印象の評価

話題定義では、対話セットごとに別の作業員 2 人が階層構造で定義した話題区間を調整した。話題の粒度や内容は定義者の個人差がみられた。被験者の記憶の書き起こしがかなり要約されたものであったこともあり、個人差が出にくく、かつ分析がしやすい比較的大きめの定義を採用した。定義した話題数は対話①④では 14、対話③④では 27 となった。被験者が話題を記憶していたかどうかの判断は、被験者の書き起こしやアンケートに、その話題に関する記載が少しでもあれば、該当の話題を記憶していたとした。被験者は記憶を書き起こした際、その話題に関する印象評価（良い／悪い／どちらでもない）も行っている。そこで話題ごとの記憶と印象を印象評価スコアとして水準化した(Table 6)。

Table 6 話題ごとの印象評価スコア

スコア	変数の水準
2	記憶に残っていて良い評価
1	記憶に残っていて普通の評価
0	記憶に残っていない
-1	記憶に残っていて悪い評価

3.2.2. 話題ごとの音響特徴と言語特徴

3.2.1 で定義した話題の単位で音声データを分割し、対話ごとの場合と同じ方法(3.1.2, 3.1.3)で対話特徴を取得し、Table 6 の印象評価スコアとの相関を調べた。結果を Table 7 に示す。記憶や印象は、自分の対話特徴だけでなく相手の対話特徴とも相関があると考え、それぞれとの相関を調べた。表頭が対話特徴、表側が被験者である。例えば先頭行は対話①④の被験者①の印象評価と、自分や相手の対話特徴との相関となっている。

対話①④では双方とも印象評価で 2 をつけていない。また対話③④では双方とも -1 をつ

けておらず更に被験者③は全対話の全話題の印象評価を2としており相関が出にくかった。

Table7 印象評価と特徴量の相関

印象評価との相関		Pitch_avg	Pitch_std	Pitch_avg	Pitch_std	intensity_avg	intensity_std	intensity_avg	intensity_std	話題継続時間
自分	相手	自分		相手		自分		相手		
①	④	-0.31	-0.20	0.53	0.25	0.21	-0.18	0.54	0.20	0.26
③	④	0.07	0.03	-0.12	0.20	0.35	-0.11	-0.08	-0.09	0.34
①	③	0.54	0.28	-0.05	0.20	0.85	0.04	0.58	0.49	0.36
③	④	-0.20	-0.05	-0.06	0.21	0.01	-0.08	0.27	0.13	0.50

印象評価との相関		ターン数		発話文字数		発話数		肯定発話数		疑問発話数	
自分	相手	自分	相手	自分	相手	自分	相手	自分	相手	自分	相手
①	④	0.44	0.47	-0.01	0.58	0.05	0.58	0.01	0.59	0.60	0.21
③	④	0.36	0.37	0.35	0.32	0.32	0.35	0.31	0.31	0.20	0.26
①	③	0.39	0.35	0.29	0.35	0.34	0.39	0.28	0.39	0.59	0.20
③	④	0.56	0.58	0.48	0.55	0.54	0.58	0.45	0.56	0.54	0.35

印象評価との相関		ターン平均発話文字数		ターン平均発話数		疑問発話割合		発話文字数倍率		発話数倍率	
自分	相手	自分	相手	自分	相手	自分	相手	自分	相手	自分	相手
①	④	-0.46	0.60	-0.45	0.53	0.51	-0.35	-0.55	0.62	-0.48	0.60
③	④	0.01	0.26	-0.06	0.35	0.06	-0.02	-0.13	0.25	-0.21	0.34
①	③	0.38	-0.13	0.43	-0.07	-0.06	0.13	0.11	-0.48	0.15	-0.41
③	④	0.11	0.01	0.26	0.14	0.20	-0.07	0.20	0.08	0.21	0.06

話題継続時間に関しては、長いと評価が低い傾向にある。

対話①④では、被験者④の Pitch 平均・Intensity 平均と双方の印象に相関が見られる。被験者④は声に感情が出やすくそれを被験者①も感じて印象に影響が出ていると考えられる。被験者①の Intensity 平均が大きいと被験者④の印象はいい。被験者①は自分の Pitch 平均と印象に負相関が見られ自分の Intensity 平均とは正の相関が見られる。つまり印象が悪い話題では声が高く小さくなる。これは対話ごとの場合と同じ傾向である。

対話③④では、被験者③の Intensity 平均と双方の印象に相関が見られる。被験者④の Pitch 平均が高いと双方評価は低めである。話題探索中や被験者④が聞き手に回っている話題で被験者④の声が高い傾向がみられた。

また話題中のターン数が多ければ印象評価スコアが高い傾向にある。話が弾んでキャッチボールが起こる方が評価は高いようだ。自分が主に話している（自分の発話文字数が多い、発話数が多い、自分の肯定発話が多く疑問発話が少ない、相手の肯定発話が少なく疑問発話が多い）ときと、自分が主に聞いているときの印象評価スコアを見ると、全体的に相手が話している時にスコアが高い。被験者③と被験者④は自分が話している時もスコアが高めなのに対し、被験者①は自分が話しているときのスコアが低い。

ターン平均発話文字数、ターン平均発話数、疑問発話割合、発話文字数倍率、発話数倍率を見ると、被験者①のその傾向がより顕著に見られる。被験者④は自分が話しているとき

の方が印象スコアが高く、被験者①との対話では相手が話しているときの印象スコアが低い傾向があるが、被験者③との対話では相手が話しているときの印象スコアは悪くない。

4 おわりに

本稿では、自然な対話を題材とした実験データの一部で印象評価の異なる対話①④と対話③④を主に調査した。その結果、互いの反応が相手の印象にも影響を与えていることが音響特徴や言語特徴から示された。

- 互いに気が合わないと回答した対話①④では、発話量は多かったものの、被験者①が一方向的に話すことが多かった。被験者①は被験者④の声からネガティブな感情を感じ取り、ストレスや緊張からか自らの声が高くなった可能性がある。
- 対話③④では発話バランスもよく、対話が弾んでいたようである。互いにどちらかが話し手の場合でも印象評価がよく、聞き手と話し手が気持ちよく話をしていたことがうかがえる。双方とも被験者③の声が大きいと印象がよいことから、被験者③がよい反応を示していたのではないだろうか。

自然対話では、被験者の動作による雑音の発生、不明瞭な発音、早口、文にならない発言、言い淀み、言い間違い、話題定義など、音響特徴、言語特徴ともに分析の前処理が難しく、笑い声や息づかいなど取扱いが難しいものがあつたが、その中に話者の無意識の気持ちが表れているはずである。

今回は対話や話題の単位で調査したが、さらに小さい話題や発話単位の特徴なども調べていく必要がある。加えて、まだ調査できていない他の被験者や、他の対話の特徴などとの比較を通じて人の普遍性や個性性などについて考えてきたい。

参考文献

- [1] 谷田, 高椋, 「見る」と「聞く」の言語理解の観察, 第223回自然言語処理研究発表会
- [2] 谷田, 高椋, コミュニケーションにおける記憶と個性の評価 - 初対面の人どうしの対話における記憶内容の観察実験 -, 人工知能学会, 2016年度人工知能学会全国大会