

音声対話における親密度と語彙の関係

中里収^{†1} 神崎啓太郎^{†1} 菊池英明^{†2}

あらまし 本研究では、人の話し言葉は親しさによってどのように変化するかを調べた。調査に先立ち、分析に使用するデータとして大学生9名の初対面の時の対話データ（低親密度）と6ヶ月後の対話データ（高親密度）の2種類の対話データを準備した。その後、対話データの中から3ペアの話者に対して形容詞対の質問紙による評価を行った。その結果、23の形容詞対のうち、「礼儀正しい」「力強い」「明るい」「丁寧な」「上手な」「声の大きい」「わかりやすい」の7つの項目に差が見られた。

また、初対面の時に使用される語彙と、知り合ってから半年経過した時に使用される語彙を比較した結果、初対面のデータでは「です」「ます」「はい」が多く、6ヶ月後になると「うん」「おれ」「けど」「まじ」が多くなるという結果が確認された。

The Relationship between the Level of Intimacy and the Vocabulary Used in Dialogues

SHU NAKAZATO^{†1} KEITARO KANZAKI^{†1}
HIDEAKI KIKUCHI^{†2}

Abstract In this study, we analyzed the change in spoken vocabulary with the level of intimacy. For our experiments, we collected two kinds of dialogue data: the initial meeting dialogue data we called "low intimacy" and that after six months we called "high intimacy". In our experiments, subjects listened to the dialogue of three pairs of speakers, and evaluated their impressions of the manner of speech through a questionnaire. Analyzing the results, for 23 factors we found differences for 7 factors: "Polite", "Powerful", "Cheerful", "Considerate", "Skillful", "Good-volume" and "Clear".

Comparing the frequency of words used in low and high intimacy dialogues, we found a high frequency appearance of "desu", "masu" and "hai" forms in the low intimacy dialogues and the use of "un", "kedo", "maji" and "ore" in the high intimacy dialogues.

1. はじめに

1.1 インタフェースとしての音声

最近、ヒューマン-コンピュータのインタフェースとして様々な手法が提案されている。マルチタッチパネルを用いた直観的な画面操作やカメラで身体の動きやジェスチャーを認識する入力装置、脳波などの微弱な電流を検知して入力信号に変換する装置などが実現されてきている。そんな中で、人の音声を用いて人工的なシステムを操作する方法は、利用者が特別な訓練を必要としないため、古くから有効なインタフェースとして注目されてきた。

最近では、スマートフォンやカーナビゲーションシステム、ロボットなど様々な場面で音声対話インタフェースを持つシステムが実現している。コンピュータが今ほど普及していない頃では、システムは不特定多数の話者を想定して設計されることが多く、利用者の年齢や性別に広く対応させることが一つの課題となっていた。

今ではパソコンやスマートフォンが広く普及し、情報機器は一人一台の時代となっている。その結果、システムは特定のユーザーの利用履歴を管理することによって、ユーザーの癖や好み、よく使う語句などに合わせてインタフェースをより使いやすく適応することが可能である。

特に音声を使ったインタフェースでは、擬人化された対象（エージェント）の方がより話しかけやすいとされ[1]、あたかも人格を持った人間と話しているようなインタラクションの実現が期待されている。

人間とエージェントと間のインタラクションにおいて重要な課題は、インタラクションをより効率化するために、エージェントが人間に適応していくように、インタラクションのスタイルに変化を持たせることである。

人間同士の日常の対話でも現実にもみられるように、対話者の属性や、対話者間の関係性によって、ある程度規範的な対話スタイルがあり、それらは話者の関係の変化や状況に応じてダイナミックに変わるものである。

本稿では、このような話者間の関係の変化に注目し、人間同士の2人対話において、同じ対話者のペアの親密度が上がることによって、話し方がどのように変化するかを報告する。

^{†1} 名城大学国際学群
Faculty of International Studies, Meio University

^{†2} 早稲田大学人間科学部
School of Human Sciences, Waseda University

1.2 先行研究および本研究の目的

現在実用化されている音声対話インタフェースでは依然制約が多く、利用者がシステムに適応しているのが現状である。利用者はそのシステムができることを事前に考えて（概念的制約）、単純な文を（言語的制約）、明瞭に話す（音響的制約）必要があると指摘されている[2].

これらの制約は人間同士の対話とはだいぶ異なる。人間同士の自然な対話では、話題も発話文の区切りも明確でないことも多く、聞き手のフィードバックを重要な手がかりとしていたり、話し手と聞き手が同時に発話したりとかなり複雑なことが対話進行中に行われている[3].

このような対話の内容やスタイルを測定する方法として、観察者による主観的な評価と、音声データに基づく語彙や音響的な特徴などを調べる客観的な評価とがある。

人は他人の対話を聞くだけで、話者についての印象を抱くことができる。音声対話研究においても、対話音声の印象を被験者による採点で分析したものは多い。

山住らは「日本語話し言葉コーパス」に収録されている講演形式の音声データに対する印象評価実験を繰り返し行うことで20項目からなる尺度を構築し、潜在的な因子「上手さ」「好悪」「速さ感」「活動性」「スタイル」の5つを抽出している[4].

辻村らは話し方の印象評価で抽出された4つの因子「好ましさ」「速さ感」「活発さ」「雰囲気」と特定の音声の特徴との関係について検討している[5][6].

小川らは大学生が初めて会話する場面で話者の発話量のつりあいに注目し、話者に対する印象の形成過程を分析している[7][8].

水上ら是对話者間の関係性、特に親疎関係をどのように認識しているかを実験によって確認するとともに、親疎関係が、対話者の発話に与える影響を言語的・非言語的な特徴から分析している[9].

筆者らは今までに、人間同士の対話コーパスを収集し、親密度の異なる対話の例として、同じ話者ペアで、初対面の時の対話と、知り合って6か月後の対話データを収集し、対話の印象がどのように変化するか因子分析し、初対面の時の対話が話者に関わらず同じような印象を持つことを報告した[10].

本稿においても、人間関係が「初対面」から「知り合い」になる過程で、話し言葉がどのように変化するのかに注目し、大学生を対象に同じ話者のペアで2回にわたり対話データを収集し、それを分析することで、話し方の印象の変化と実際にその対話で使用された語彙の変化を調べた。

2. 音声対話コーパス

2.1 対話データ収集方法

対話データの収集は早稲田大学人間科学部構内の音声収録用防音室で行った。参加する対話者はお互いに初対面の大学生9名、1回目の収録は1月に、2回目の収録は同年7月に行った。対話収録用防音室は2部屋に仕切られているため、2人の話者の音声を2つのチャンネルに分離して録音ができ、対話収録は相手の顔が見られる対面対話と非対面対話を、時間を区切って収録した。

対話の話題となるテーマは、指定した色（例えば「白」）について連想されるものについて対話してもらった。1つの対話は対面で約6分、非対面で約6分の合計12分とした。

9名の参加者を4名と5名の2つのグループに分け、グループ内の総当たりで対話した。収録された対話データは合計28対話（すべての対話に参加できない話者もいた）、時間にして約340分間であった。このデータに対し発話区間ラベリングと書き起こし作業を行った。発話区間ラベリングでは機械的に200msec以上の無音区間と有声音区間を区切り、その後、図1のように手作業による書き起こしを行った。1単位の発話は「通し番号」、「開始時刻」、「終了時刻」、「左右のチャンネル」、「かな漢字混じり書き起こし」、「カタカナ書き起こし」から成る。

```
0001 00001.539-00002.839 L:  
白色だそうです&シロイロダソーデス  
0002 00004.327-00007.012 L:  
あそう言えばさっき表出たらば白い犬が歩いてまして&アソー  
イエバサッキオモテダトラバシロイイヌガアルイテマシテ  
0003 00009.048-00011.357 L:ん  
そこの外出たらお散歩してて&ンソコノソトダトラオサンボシ  
テテ  
0004 00011.733-00012.332 L:  
3匹&サンビキ  
0005 00012.930-00013.751 L:  
おばちゃんが&オバチャンガ  
0006 00014.179-00015.000 L:  
散歩したんすよ&サンボシタンスヨ  
0007 00016.317-00018.712 L:  
でまその一犬は良いとして&デマソノイーヌワヨイトシテ  
0008 00019.618-00022.611 L:  
急にあの携帯が鳴ったんですねそのおばちゃんの&キューニア  
ノケータイガナツタンデスネソノオバチャンノ
```

図1 書き起こしデータの例

3. 話し方の印象評価

3.1 評価実験方法

収集された対話コーパスを利用して、9名の話者のうち、男性2名女性1名からなる3ペアの対話データを選び、低親密度対話と高親密度対話それぞれの話し方の印象評価実験を行った。

評価実験の被験者は聴力正常な男性11名、女性1名。実験室内で3対話の音声の順に呈示し、話し方の印象を形容詞対の尺度で評価した。

評価者には、対話に親密度の違いがあることは知らされていなかった。また、話の内容や、話し相手の属性（出身地など）を評価するのではなく、あくまでも話し方の印象を評価するように教示した。映像はなく、左右2チャンネルの音声のみで評価した。評価者には1つの評価ごとに「右（左）側の人の話し方を評価する」ように教示した。1対話データは約3分間であり、提示する順序はバランスをとった。

評価には山住らの報告[4]を参考に表1に示す23項目の形容詞対を用いて話し方の印象を項目ごとに5段階（1～5）で評価した。

表1 評価に用いた形容詞対23項目

1	声の大きい	5+++1	声の小さい
2	話し慣れた	5+++1	話し慣れていない
3	礼儀正しい	5+++1	無礼な
4	上手な	5+++1	下手な
5	まとまった	5+++1	ばらばらな
6	スピード感のある	5+++1	スピード感のない
7	丁寧な	5+++1	ぞんざいな
8	せわしげな	5+++1	のんきな
9	力強い	5+++1	弱々しい
10	明るい	5+++1	暗い
11	落ち着きのない	5+++1	落ち着きのある
12	感じの良い	5+++1	感じの悪い
13	元気のある	5+++1	元気のない
14	なめらかな	5+++1	しどろもどろな
15	親しみやすい	5+++1	親しみにくい
16	流暢な	5+++1	たどたどしい
17	まじめな	5+++1	ふまじめな
18	好きな	5+++1	嫌いな
19	速い	5+++1	遅い
20	積極的な	5+++1	消極的な
21	心地よい	5+++1	不快な
22	わかりやすい	5+++1	わかりにくい
23	上品な	5+++1	下品な

3.2 評価実験結果と考察

低親密度の対話と高親密度の対話の評価を比較した結果、形容詞23の項目のうち、「礼儀正しい」「力強い」「明るい」「丁寧な」「上手な」「声の大きい」「わかりやすい」の7つの項目について点数が下がる傾向が見られた（表2）。

すなわち、初対面の時の対話には、「礼儀正しく」、「力強い」などの印象があり、半年後の対話ではそれらの印象が薄れているといった傾向がみられている。

表2 話し方の印象評価結果

	形容詞	低親密度 (初対面)	高親密度 (半年後)	半年間の 変化
1	声の大きい	3.44	2.986	-0.46 *
2	話し慣れた	3.28	3.125	-0.15
3	礼儀正しい	3.35	2.806	-0.54 **
4	上手な	3.18	2.944	-0.24 *
5	まとまった	2.96	3.028	0.07
6	スピード感	3.24	3.097	-0.14
7	丁寧な	3.18	2.903	-0.28 *
8	せわしげな	2.96	2.778	-0.18
9	力強い	3.26	2.875	-0.39 **
10	明るい	3.50	3.125	-0.38 *
11	落ち着き	2.85	2.792	-0.06
12	感じの良い	3.47	3.319	-0.15
13	元気のある	3.35	3.250	-0.10
14	なめらかな	3.06	3.167	0.11
15	親しみやすい	3.40	3.278	-0.13
16	流暢な	3.19	3.125	-0.07
17	まじめな	3.19	3.111	-0.08
18	好きな	3.06	3.181	0.13
19	速い	3.21	3.042	-0.17
20	積極的な	3.40	3.167	-0.24
21	心地よい	3.06	3.028	-0.03
22	わかりやすい	3.24	2.958	-0.28 *
23	上品な	3.00	3.056	0.06

** p < .01 * p < .05

先述の先行研究では、これらの形容詞対を用いて因子分析を行い、4つの因子を抽出しているが[10]、今回の結果からは同じような因子を確認できなかった。これは先行研究では、あえて性差をなくすために男性話者のみのデータで評価したが、今回評価するデータには、男性話者と女性話者の両方を含めたため評価者によって観点が異なったのが原因と考えられる。

4. 使用語彙の調査結果

4.1 親密度の違いによる使用語彙の差

話者ごとの書き起こしデータの総発話回数は、機械的にポーズで区切ったもので26893回であった。その中に現れる特定の語（600語）を抽出し、出現回数を調べた。初対面の時の「低親密度」対話と、知り合って半年後の「高親密度」対話での出現頻度は以下の表の通りであった。対話課題に依存する「アカ」「シロ」「アオ」など一部の単語を除き、出現回数の多いものと、条件により差が大きかった2つの語を載せている（表3および図2）。

表中の数値は単位時間（60秒）に対する語の出現回数であり、（かっこ）内は全対話内の出現回数である。

表3 使用語彙の差（親密度）

	語句	低親密度		高親密度	
		出現割合（出現数）	出現割合（出現数）	出現割合（出現数）	出現割合（出現数）
1	ナイ	2.94	(571)	3.31	(481)
2	デス	4.63	(898)	0.17	(25)
3	ウン	1.51	(292)	3.98	(578)
4	ナンカ	2.19	(425)	2.36	(343)
5	ケド	1.57	(305)	2.29	(332)
6	ハイ	1.91	(371)	0.72	(104)
7	アノ	1.47	(285)	0.99	(144)
8	カナ	1.13	(220)	1.32	(191)
9	ミタイ	1.03	(199)	0.90	(131)
10	マス	1.40	(271)	0.20	(29)
11	アレ	0.93	(181)	0.81	(118)
12	ナンダ	0.69	(133)	1.05	(153)
13	ソレ	0.74	(143)	0.92	(133)
14	アル	0.83	(161)	0.76	(111)
15	ダッタ	0.57	(110)	0.65	(95)
16	グライ	0.49	(95)	0.67	(97)
17	カッタ	0.51	(99)	0.57	(83)
18	ダカラ	0.41	(79)	0.70	(102)
19	ダヨネ	0.36	(69)	0.74	(108)
20	ジャナイ	0.62	(121)	0.37	(54)
26	オレ	0.25	(48)	0.71	(103)
28	マジ	0.18	(35)	0.78	(113)

出現頻度1位の「ナイ」では低親密度（初対面）から高親密度（知人）への違いにあまり変化が見られないが、2位の「デス」、6位の「ハイ」、10位の「マス」の頻度が減少しており、逆に3位の「ウン」5位の「ケド」28位の「マジ」26位の「オレ」の回数は高親密度になるにつれて増えていた。

4.2 他の条件による下位分析

4.2.1 男女話者の条件による差

今回の対話コーパスに収録された話者9名のうち、7名が男性、2名が女性であった。男性話者・女性話者の使用語彙の違いを表4に示す。

表4 使用語彙の差（男女話者）

	語句	低親密度		高親密度	
		男性話者	女性話者	男性話者	女性話者
1	ナイ	2.94	2.95	3.22	3.59
2	デス	3.80	7.59	0.17	0.16
3	ウン	1.62	1.11	4.01	3.89
4	ナンカ	1.83	3.49	1.87	3.84
5	ケド	1.59	1.51	2.18	2.61
6	ハイ	2.11	1.23	0.84	0.36
7	アノ	1.60	1.01	1.06	0.80
8	カナ	1.02	1.56	1.32	1.29
9	ミタイ	0.96	1.25	0.86	1.01
10	マス	1.14	2.31	0.19	0.22
11	アレ	0.98	0.75	0.78	0.90
12	ナンダ	0.65	0.82	0.97	1.32
13	ソレ	0.71	0.82	0.87	1.04
14	アル	0.89	0.61	0.77	0.74
15	ダッタ	0.62	0.38	0.54	0.99
16	グライ	0.52	0.38	0.63	0.77
17	カッタ	0.51	0.49	0.63	0.41
18	ダカラ	0.47	0.19	0.68	0.77
19	ダヨネ	0.42	0.14	0.64	1.04
20	ジャナイ	0.68	0.42	0.41	0.25

先ほど明らかになった表現のうち、低親密度対話において「デス」「マス」を多く使用するのは女性話者、「ハイ」を使うのは男性話者であることが分かった。

4.2.2 対面・非対面による使用語彙の差

また、相手の顔が見える「対面」条件と相手の顔が見えない「非対面」条件によって、使用される語彙がどれほど異なるかを調べるために、各条件による使用語彙の違いを調査した（表5）。条件による差は少ないものの、非対面対話の方が「ウン」と「ハイ」の頻度が高くなる傾向が確認できた。

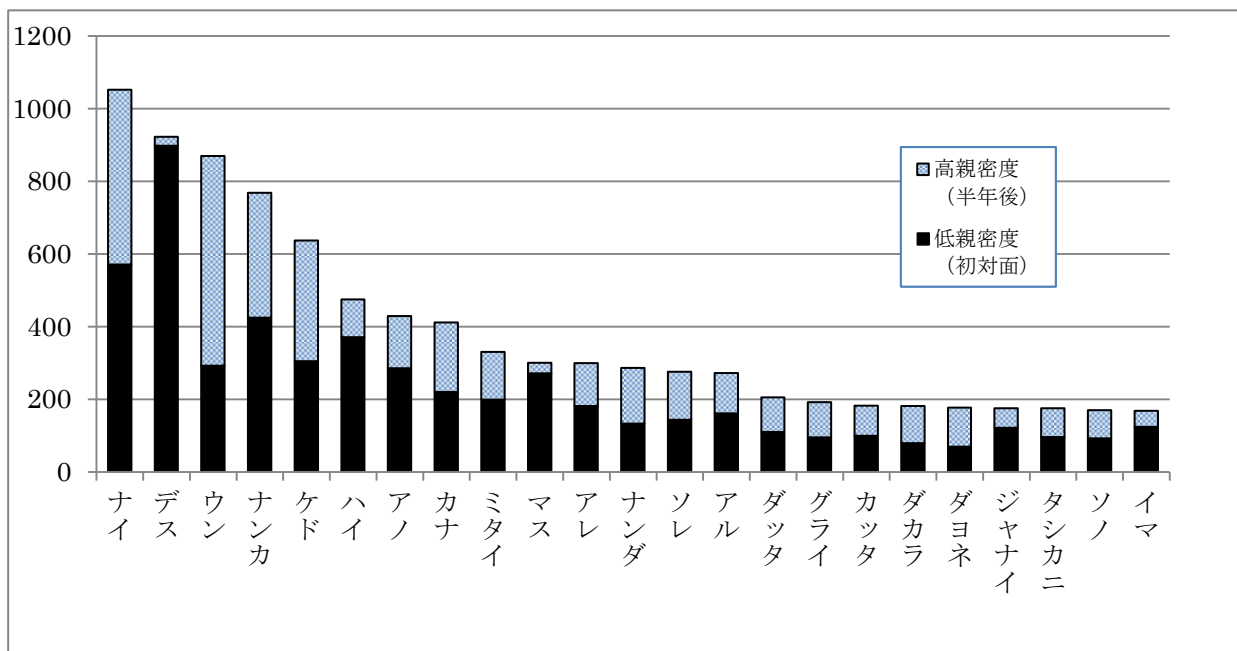


図 2 対話中に出現した語 (回数)

表 5 使用語彙の差 (対面・非対面)

	語句	低親密度		高親密度	
		対面	非対面	対面	非対面
1	ナイ	3.08	2.81	3.63	3.00
2	デス	4.63	4.64	0.21	0.14
3	ウン	1.28	1.73	3.63	4.33
4	ナンカ	2.17	2.22	2.12	2.60
5	ケド	1.67	1.48	2.29	2.28
6	ハイ	1.79	2.03	0.50	0.94
7	アノ	1.31	1.63	0.85	1.13
8	カナ	1.29	0.98	1.46	1.17
9	ミタイ	1.16	0.89	0.95	0.85
10	マス	1.48	1.31	0.19	0.21
11	アレ	0.79	1.08	0.83	0.80
12	ナンダ	0.68	0.69	0.92	1.18
13	ソレ	0.77	0.71	0.80	1.03
14	アル	0.70	0.96	0.77	0.76
15	ダッタ	0.55	0.58	0.73	0.58
16	グライ	0.38	0.59	0.62	0.72
17	カッタ	0.50	0.52	0.63	0.51
18	ダカラ	0.47	0.35	0.73	0.67
19	ダヨネ	0.25	0.46	0.81	0.67
20	ジャナイ	0.68	0.56	0.45	0.29

5. 考察とまとめ

これまでの結果により、対話者間の親密度により使用する語彙に違いがあることが確認できた。特に親密度が高くなるにつれ「デス」「マス」「ハイ」の頻度が減り、逆に「ウン」「オレ」「ケド」「マジ」が増えることが確認された。

また、男女の差による使用語彙の結果では、初対面の時に「デス」「マス」を使うのは女性が多く、「ハイ」を使うのは男性が多いことが分かった。

対面対話と非対面対話の比較では、非対面対話の方が、相槌の機能をもつ「ウン」と「ハイ」の頻度が多かった。これはお互いの表情が見えないために、話を聞いている状態を相槌によって伝えていると考えられる。

6. おわりに

本研究では、初対面の時の対話と、知り合ってから半年後の対話の印象評価と使用語彙の変化を分析した。

2つの対話の印象評価を行い、因子分析では先行研究のような因子を確認できなかったものの、7つの形容詞について有意な差が確認できた。

また、使用語彙の比較では、同じ対話者のペアであっても、初対面の時と知り合ってから半年後の時の対話を比較し使用する語が異なることが確認できた。

コミュニケーションをとっていく中で、同じ相手に対してスタイルを変化させていることを示している。

今後の課題としては、対話の印象評価の際に、対話者の性別や評価者の性別の組み合わせなどについて、評価に影響する要因についてさらなる分析が必要だと考える。

参考文献

- 1) 植田一博, ”相互適応学習 -HAIにおける時間的要因-”, 人工知能学会誌, Vol.21, No.6, pp675-680.(2006)
- 2) 河原達也, “音声対話システムの進化と淘汰-歴史と最近の技術動向-”, 人工知能学会誌, Vol.28, No.1, pp.45-51, (2013).
- 3) 中里収, ”対話中に期待される聞き手の反応について”, 人工知能学会研究会資料, SIG-SLUD-A603-07, pp.33-38,(2007).
- 4) 山住賢司, 籠宮隆之, 槇洋一, 前川喜久雄:”講演音声の印象評価尺度”, 日本音響学会誌 61 卷 6 号, pp.303-311,(2005).
- 5) 辻村壮平, 福田将之, 山田由紀子:”話し方の印象における発話要因 (その1 印象評価に関する検討)”, 日本建築学会大会学術講演梗概集, pp.85-86, (2007).
- 6) 福田将之, 辻村壮平, 山田由紀子:”話し方の印象における発話要因 (その2 物理的特徴と印象評価の関連性)”, 日本建築学会大会学術講演梗概集, pp.87-88,(2007).
- 7) 小川一美, ”二者間発話量の均衡が会話が抱く相手と会話に対する印象に及ぼす効果”, 社団法人 電子情報通信学会, (2003).
- 8) 小川一美, “初対面場面における二者間の発話量のつりあいと会話者および会話に対する印象の関係”, 名古屋大学大学院教育発達科学研究科紀要, 心理発達科学, pp.20-27, (2000).
- 9) 水上悦雄, 善本淳, ”対話者関係性認識における発話の言語的・非言語的特徴の分析”, 人工知能学会研究会資料, SIG-SLUD-B301, pp.1-6.(2013).
- 10) 中里収, 大城裕志, 菊池英明, ”音声対話における親密度と話し方の関係”, 信学技報 HIP2012-96, pp.109-114.(2013).