

雰囲気アセスメント項目抽出へむけた音響特徴量—印象評価の関係分析

Analysis of the Relationship between Acoustic Features and Impression Rating for the Extraction of Atmosphere Assessment Items

藤澤 充紘†
Mitsuhiro Fujisawa中平 勝子†
Katsuko T. Nakahira

1 はじめに

人間が社会で生活を営むうえで、コミュニケーションは必要不可欠なものである。コミュニケーション中に行われる行動は言語音声、言語非音声、非言語音声、非言語非音声の4つに分類でき、これらが行われている場をコミュニケーション場と後藤ら [3] は定義している。コミュニケーション場とその形成に必要な要素を図1のように示している。

コミュニケーション場では可観測である言語/非言語情報が参加者に明示的/暗示的に伝わり、参加者に知覚された可観測情報をもとに参加者側でどう感じるかなどの認知行動が起こる。そして、認知行動の結果は不可観測量として参加者の内面的な認知情報と共に発話やジェスチャーに含まれる。このような明示的/暗示的情報の相互作用の結果が場の雰囲気を作り出し、参加者の次の言語/非言語情報の外化/内面的な認知行動につながる。

コミュニケーションには雑談や飲み会、会議、報告など様々なシチュエーションが存在する。どのようなシチュエーションにおいても、会話の内容と語気、またジェスチャーや表情に乖離があると、与える印象が会話の主題とは離れたものになってしまう可能性がある。それによりコミュニケーションの行われる場であるコミュニケーション場の雰囲気を捉えられず、コミュニケーションが円滑に行えなくなる可能性がある。そこで、コミュニケーション場の雰囲気アセスメントが必要となる。西田ら [7] は感情の暗黙表現が音響特徴量にも表れるとして笑い、相槌の検出にSVM (サポートベクターマシン)、質問の句末音調を基本周波数 F0 (ピッチ) が上昇するかどうかで識別した。また、識別した情報から雰囲気推移のアセスメントを隠れマルコフモデルで行い、遷移推定の精度は6割程度であった。このようにコミュニケーション場における雰囲気推定は可能になりつつある。

また、Edmondson(1999)[1] が提唱した心理的安全性という概念がある。心理的安全性は、対人関係におけるリスクテイクに対して安全である、という共有された信念のことである。心理的安全性を高めるためにチーム全体でこの信念を共有する必要がある。心理的安全性が高いとチーム内で忌憚のない意見が交換できるため、チームワークが促進され、メンバー同士の信頼が高まり、個々のパフォーマンスが向上する。そのため、アクティブ・ラーニングや、職場環境においては、心理的安全性が重要視される。

しかしながら、心理的安全性が担保されているか確認するために録画録音を行い、プライバシーに配慮できていない環境下だと、却って心理的安全性が低下してしまい、コミュニケーションが円滑に行えないことが考えられる。

以上の懸念を払拭するための一手法として、プライバシーに配慮することを念頭に、音響特徴量や無音抽出を用いて機械的にコミュニケーション場を可視化することが考えられる。本稿では、その実現へ向けた第一段階として、音響特徴量と聞き手の持つ印象との関係分析を行い、アセスメント指標を策定するために、傍参与者視点の被験者に、音響特徴量を調節した2者の会話音声を上で提示した際の印象主観評価を行う。

2 聴取実験

コミュニケーション場は言語/非言語情報のやりとりにより醸し出される雰囲気と構成されるものである。そのため、話者でなくてもコミュニケーション場を認識することは可能であり、そこで行われるコミュニケーションを印象評価することでコミュニケーション場の雰囲気の評価につながると思う。

よく利用される音響特徴量には、ピッチ、パワー、フォルマント周波数 (声道特性)、モーラ数 (発話文字数) などがある。基本周波数 F0 (ピッチ)、パワー、発話速度を利用するものとする。声道特性と声帯特性は個人に依存するものなので取り扱わない。また、モーラ数は主に発話の中に存在する母音の数であり、日本語であれば文字数にほぼ直結する。発話速度はモーラ数/時間で表すことができる。

2.1 音声刺激

ある人物が笑っていれば楽しそう、といった感情を抱く。その場の雰囲気をよくするために笑い返す。このように、コミュニケーションにより感情を変化させ、それに伴い表情や声色など行動も変化し、それがコミュニケーション場の雰囲気を醸し出す。基準となる雰囲気が平常である場合が存在し、雰囲気が変わると人の行動も変化すると考える。これは音響特徴量でもいえることであり、音響特徴量を調節することにより発話の表す感情を表現し、場の雰囲気も調整できると考える。そのため、基準となる音響特徴量からそれぞれ変化させる。

基本周波数 F0: 郡 (1989)[4] は合成母音の「お」を用いてピッチ、高さの変化パターン、長さを段階的に変化させ「喜び」「恐怖」「驚き」「悲しみ」「嫌悪」「怒り」の感情を表しているかを12名の近畿出身の若年女性に評価させた。結果として「喜び」「恐怖」「驚き」にピッチが影響していることが確認できた。また、重野 [5] は、「東京」「河原崎さん」といった単語及び短文に「幸福」「驚き」「怒り」「嫌悪」「恐れ」「悲しみ」を込めて発話したものについて音響特徴量を分析した。結果として、6感情の中で相対的に「幸福」「驚き」では平均 F0 が高く、「嫌悪」では平均 F0 が低くなることが明らかになった。また、武田ら [6] は5名に「平常」「軽い怒り」「怒り」「激怒」の4段階の感情の強さで発声するように指示した。怒りの度合いが大きくなるに従って、最高 F0 が増大する傾向を確認した。

発話速度: 丸島 [8] は自然言語音の「papapapapa」を3種類の

† 長岡技術科学大学

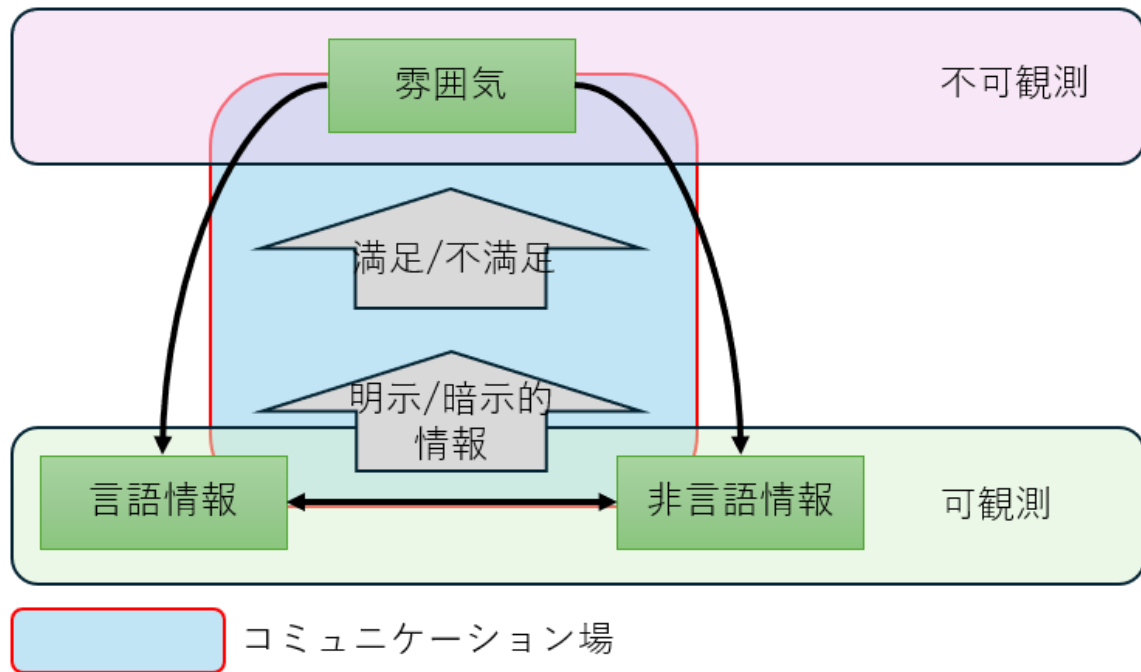


図1 コミュニケーション場

速度で読み上げたものを3名の若年女性に聴取させ、頭の中で繰り返すように指示した。Normalとそれ以外で二項対立的に認知処理されていると考察している。また、武田ら[6]は怒りに関する平均発話速度の解析の結果、アクセント及びモーラ数に関係なく「平常」<「軽い怒り」<「怒り」>「激怒」の傾向になることが得られた。

パワー：武田ら[6]は怒りに関するパワーの解析の結果、怒りの度合いが大きくなるに従って、パワーが上昇することが明らかになった。

2.2 印象評価

原音声データ：実験に用いる原音声は男子大学院生(22歳)2名の雑談を録音したものである。SONY PCM-D10にてサンプリング周波数192kHz、量子化24bitでモノラル録音を行った。

音声刺激の作成：原音声データをもとに、ピッチ、発話速度、パワーをそれぞれ3段階に調整し、 $3 \times 3 \times 3 =$ 計27刺激を作成する。

印象評価：被験者に音声刺激を提示しながら、音声刺激から生じた感情をリアルタイムで回答してもらう。初めに、練習としてBTSJ1000人日本語自然会話コーパス[2]の会話で行い、その後上述した音声刺激にて実験を行う。織田ら[9]が作成した感情・覚醒チェックリストEACLでは持続性がないため驚きが除外されていたが、リアルタイムで回答を求めると、驚きを含めた「喜び」「恐怖」「驚き」「悲しみ」「嫌悪」「怒り」の6つの基本感情から回答してもらう。また、緊張覚醒(TA)/エネルギー覚醒(EA)を各音声刺激の終わりに回答してもらう。

3 おわりに

本稿では、雰囲気アセスメント項目の抽出へむけた音響特徴量と印象評価の関係分析の実験を行うための音声刺激の作成に関して述べた。現在実験中のため、結果は当日発表する。今

後、音声刺激を作成し、印象評価の実験を行う。また、作成した音声刺激が違和感のない人の声を確認するために予備実験を行う必要がある。

謝辞

本研究の一部は科研費JSPS(22K12284, 代表: 岐阜工業高等専門学校・小川信之, 23K11334, 代表: 長岡技術科学大学・中平勝子) および経営改革促進事業の助成を受けたものである。

参考文献

- [1] Edmondson Amy. Psychological safety and learning behavior in work teams. *Administrative Science Quarterly*, Vol. 44, No. 2, pp. 350-383, 06 1999.
- [2] 宇佐美まゆみ監修. Btsj1000人日本語自然会話コーパス(2023年3月ncrb運動完成版), 2023. 科研基盤研究(A)「語用論的分析のための日本語1000人自然会話コーパスの構築とその多角的研究」(研究代表者: 宇佐美まゆみ).
- [3] 後藤義貴, 中平勝子, 北島宗雄. 対話時の音声情報と言語情報に着目した発話意図の符号化. 情報処理学会第77回全国大会講演論文集, 第4分冊, pp. 367-368, 2015.
- [4] 郡史郎. 発話の音調を規定する要因—日本語イントネーション論—. 『吉沢典男教授追悼論文集』吉沢典男教授追悼論文集編集委員会編, pp. 116-127, 1989.
- [5] 重野純. 感情を表現した音声の認知と音響的性質. 心理学研究, Vol. 74, No. 6, pp. 540-546, 2004.
- [6] 武田昌一, 大山玄, 朽谷綾香, 西澤良博. 日本語音声における「怒り」を表現する韻律的特徴の解析. 日本音響学会誌, Vol. 58, No. 9, pp. 561-568, 2002.
- [7] 西田悠. 音響特徴量を用いた討論過程評価のための雰囲気遷移推定手法の開発. 2019年度長岡技術科学大学修士論文, 2020.
- [8] 丸島歩. 自然言語音を用いた発話速度の認知実験. 実験音声学・言語学研究, Vol. 2, pp. 32-43, 2010.
- [9] 織田弥生, ルリ子, 阿部恒之, 菊地賢一. 感情・覚醒チェックリストの作成と信頼性・妥当性の検討. 心理学研究, Vol. 85, No. 6, pp. 579-589, 2015.