

中空型音声コミュニケーション・メディアのデザインのための 声質と対話時間に関する基礎的調査

江波亮^{†1} 高島健太郎^{†1} 西本一志^{†1}

概要: 近年スマートフォンの普及により、インターネットを用いたコミュニケーションはより一般的なものになった。Skype 等のオンラインでのコミュニケーションツール上ではしばしば匿名の見ず知らずの他者と、伝えたい意図や内容を意識せず、会話を行っていること自体を楽しむためのコミュニケーションが行われている。本研究ではこのようなコミュニケーションを「中空型コミュニケーション」と呼び、これを支援するメディアを設計するための基礎的研究として、会話を継続させるための要因について調査を行う。筆者の実体験から、その要因の一つが話者の声質ではないかと考え、声質と会話時間の関係について分析する。

1. はじめに

近年 PC・スマートフォンの普及によりインターネットに触れる機会が増えている。特にスマートフォンの普及はめざましく、インターネットに接続するデバイスは1人1台の時代となった。総務省の調べ[1]ではスマートフォンの世帯保有率はパソコンを上回っており、2017年の世帯保有率は75.1%と、大きな伸びを見せている。

このような社会の到来により、インターネットを介したオンラインでのコミュニケーションはより一般的なものになった。インターネットを用いた代表的なコミュニケーションツールのひとつとして、2000年代から利用されている Skype[2]が挙げられる。Skype は世界中のユーザの間でオンラインでの会話を可能にするソフトウェアであり、多くのユーザが個人またはビジネスで Skype を使用し、無料の音声通話、ビデオ通話、1対1の通話、グループ通話、インスタントメッセージ (IM) の送信、ファイルの共有を行っている。Skype はスマートフォン、タブレット、PC など、自分にとって最適な端末にインストールし手軽に利用することができる。

本稿第1筆者は、インターネットを用いたコミュニケーションの普及とともに、ある私的な娯楽目的でこの Skype を利用し始めた。匿名の見ず知らずの他者とのオンラインでの会話である。インターネットの匿名発言が可能な掲示板にて、会話をする相手を探し、自身が望むトピックとマッチした、匿名の相手と会話を行うという一連の流れを、娯楽として楽しみながら利用していた。

しかしながら、通話中にも関わらず、相手に一方的に電話を切られてしまうことが度々起きる。このように、通話の相手がまだ話しているにもかかわらず、がちゃんという音を立てて受話器を置き、乱暴に電話を切ってしまう行為は俗にガチャ切りと呼ばれる[3]。このような行為が起きる

のは、筆者が行っている匿名のオンラインでの会話が、話のコンテンツよりもその入れ物、すなわち好みの相手の声を聞いていることやインタラクションを行っていること自体を目的とした、非シグナル的側面を重視するコミュニケーションであることが理由だと考える。このようなコミュニケーションでは、会話の内容はあまり意識されず、印象の良い話し相手が欲しい、その会話を行っていること自体を楽しみたいという気持ちが、会話の行為を継続させている。本研究ではこのようなコミュニケーションを「中空型コミュニケーション」と呼ぶ。Skype には筆者のように、特定の意図の伝達というより、他者と気楽に会話をしたい、他愛のない話をしたい、人の声を聞いていたい、通話をつなげていたい、という目的で、掲示板を利用し会話の相手を探しているユーザが一定数存在している。

中空型コミュニケーションを行うためには、今のところ Skype などの既存のコミュニケーションツール用いられることが一般的であろう。しかし、これらのツールは従来の意図伝達を目的とした「コンテンツ重視型コミュニケーション」のためのメディアであり、本研究で述べる「中空型コミュニケーション」を達成することは考慮に入られていない。そのため、明瞭に音声伝えるための技術は適用されているものの、相手の印象が期待と異なる、といった理由から前述のガチャ切りが行われてしまう。ユーザの嗜好を考慮した上で、望ましい話者のマッチングやコミュニケーション体験の質の向上ための介入を行い、ガチャ切りを未然に防止する、中空型コミュニケーションのための理想的なメディアが必要であると考える。

本研究では、中空型コミュニケーションのためのメディアを設計するための基礎的検討として、このような匿名の見ず知らずの他者とオンラインで行う、特段伝えたい意図のないコミュニケーションにおいて会話を継続させる要因の研究を行う。特に、第1筆者の実体験から、その要因のひとつがユーザの声質ではないかと考え、これに着目し、会話の継続時間の関係性について分析する。

^{†1} 北陸先端科学技術大学院大学 先端科学技術研究科
Graduate School of Advanced Science and Technology,
Japan Advanced Institute of Science and Technology

2. 関連研究

声質に着目した印象評価に関する関連研究は多数存在する[4][5]. 発話を直接聴取してもらい評価してもらう方法が主であり, 対象となるシーンは日常生活から講演まで幅広い. 声質は心理状態に左右され, 無意識でもその声質から自分の感情や状況が相手に伝わってしまう. 2 者間のコミュニケーションでは, こうしたお互いの感情や状況を声質から読み取って相手の印象形成をしており, 話しているうちに互いの感情が伝染しあうということが起きる[6]. この仕組みを実験的に調べるには, 会話中に声を特定の感情方向にリアルタイムで変化させる必要があり, そのための実験装置の開発も行われている. 例えば渡邊らは DAVID (Da Amazing Voice Inflection Device) という装置を開発し[7], 話し手の声を音量や話し方はそのままに, 感情的な側面だけを変化させることを可能にした. これにより, 話者の声の感情的な側面が, どのように聴者の感情や印象形成に影響しているかを調査することが可能になるとしている.

一方, 音声対話における声質が会話時間に及ぼす影響を直接的に調査した関連研究は見受けられない. しかしながら, 本研究で述べる中空型コミュニケーションの場合, 会話を行っていること自体を楽しむものであるから, 会話時間は会話の満足度を反映する重要な指標であると言えるだろう. 会話時間には話者の体調や会話の行い方など, 様々な要素が影響を与えると考えられるが, 本研究では会話の相手の声質に着目する. 声質は相手の好感度や信頼性を含む印象の形成に影響を及ぼすことが指摘されており[4], 表情や身体動作が十分に伝わらないオンラインコミュニケーションではなおさら, 非言語コミュニケーションの重要な要素であるといえるだろう. さらに, 印象そのものと比べ, 声質はソフトウェアによる操作が容易であり, メディア設計のための具体的な示唆が得られやすいと考えられる.

3. 予備実験

3.1 概要

声質の対話時間に対する影響を調査するために実験を行う. 3 種の音声を用いた Q&A 形式での, 会話を模した聴取実験を行う. 被験者は実際に他者と会話を行うわけではなく, 1 人でオーディオプレイヤーを用いて録音された質問の音声を順に再生し, 口頭で回答していく. 質問の音声は第 1 筆者の地声, 高音化の加工あり, 低音化の加工ありの 3 パターンでそれぞれ事前に 50 問程度を録音して用意する.

- 被験者: 筆者らが所属する大学院の男子学生 10 名 (予定)
- 実験手順:
 - (1) 被験者に, 匿名の見ず知らずの他者とオンラインの会話であるという想定を伝える.

- (2) 地声の質問音声での Q&A を行う. 被験者は 1 問ずつ質問を再生し, 一時停止を行い, 口頭で返答することを繰り返す. 用意したすべての質問に答えてもらう必要はなく, 飽きた, もしくは何らかの理由で続けなくなった任意のタイミングで止めてもらう.

- (3) 音声を高音化の加工あり, 低音化の加工ありに変え, (2)を繰り返す.

- (4) アンケート用紙を配布し, 記入してもらう. さらに, 記入内容に基づいたインタビューを行う.

被験者は性別が会話時間に影響を与えることを配慮し, 第 1 筆者と同じ男性に限定する. 同様に, 被験者の興味による会話時間の変化を可能な限り防ぐため, 自由会話ではなく, 事前に準備した質問を用いた Q&A 形式を用いる. 質問内容は「好きな食べ物は?」といった他愛ないものであり, 雑談のための質問集から抜粋したものである. 声質によってそれぞれ異なった質問リストを準備する. 3 種の声質の音声のうち, 高音と低音については, ボイス・トランスフォーマー (Roland 社製 VT-4) を使用し, 第 1 筆者の地声を加工したものをそれぞれ用意する. 第 1 筆者の主観に基づき, 内容が聞き取り可能である限界の程度までピッチとフォルマントを変化させたものを録音する. 質問の返答は数秒間を目安に, 一言二言程度で行ってもらうよう指示する.

3.2 評価

実験後, 質問への返答状況とアンケートの回答に基づいた評価を行う. 前者については, 3 つの声質の条件により返答される質問数に違いがあるかを比較する. 回答の長さによって影響を受ける可能性があるものの, Q&A の継続時間についても比較する. アンケートは Q&A 中の心境と声の印象について自由記述形式で尋ねるものであり, これを用いて質問数と時間に差が生じた(あるいは生じなかった)理由について考察する. アンケートの具体的な質問項目を表 1 に示す.

表 1. 質問項目

設問番号	設問内容
1	(各音声の条件ごとに)なぜ聴取・回答をやめようと思ったのか
2	(各音声の条件ごとに)Q A 中どのような心境であったか. どのようなタイミングで飽きを感じたか.
3	三種類の声の印象はそれぞれどうだったか.
4	三種類の声のうち, 最も話したいと感じたのはどれか. またその理由は何か.

4. まとめ

本研究では、「中空型コミュニケーション」に適したメディアの要件を明らかにすることを背景に、会話相手の声質と会話の継続時間の関係性について分析する。予備実験では、通常の音声、高音化の加工あり、低音化の加工ありの3パターンで、会話を模したQAの継続時間の違いについて調査する。予備実験は、現在実施中である。さらに今後予備実験の結果を踏まえ、声質の効果の検証実験、実際のオンラインコミュニケーションでの効果の有無の調査実験を検討する予定である。これらについては、インタラクシオン 2020 において発表する予定である。

参考文献

- [1] “平成 30 年通信利用動向調査の結果”
https://www.soumu.go.jp/menu_news/s-news/02tsushin02_04000062.html, (参照 2019-12-23)
- [2] “Skype”.
<https://www.skype.com/>, (参照 2019-12-23).
- [3] “goo 国語辞書”.
<https://dictionary.goo.ne.jp>, (参照 2019-12-23).
- [4] 岡田陽介. 政治家の印象形成における声の高低の影響: 音声合成ソフトを用いた女声による実験研究. 応用社会学研究. 2016, vol. 58, p. 53-66.
- [5] 内田照久, 中畝菜穂子. 声の高さと発話速度が 話者の性格印象に与える影響. 心理学研究. 2004, vol. 75, no. 5, p. 397-406.
- [6] “人によい印象を与えるには～印象形成における姿勢と声のプロソディの効果”
<https://www.waseda.jp/inst/wias/news/2018/09/19/5469/>, (参照 2019-12-23).
- [7] “DAVID; Da Amazing Voice Inflection Device”
<http://cream.ircam.fr/?p=44>, (参照 2019-12-23).