

1. 志望動機

私は早稲田大学人間科学部進学以前より音声に関する研究に興味を抱いてきた。入学当初に菊池研究室を見つけ、指定科目としてあげられていた人間計測・人間情報モデリングや言語情報科学を履修した。これらの授業を聴講していく中で音声の研究に対する興味が深まり、二年次の研究室選択の際に言語情報科学について研究している菊池研究室に配属を希望した。人工知能(AI)の現状と未来[1]において、人工知能の利活用が望ましい分野として、「高度な意味理解や感情認識等によるコンピュータと人間の対話の高度化」が挙げられている。その分野に有識者の 48.1%が期待していることから、人間がコミュニケーションロボット等と、まるで人間同士のようなコミュニケーションをする未来が期待されていることがわかる。一方で研究室での研究を通じて、コミュニケーションロボット等が対話の中で相手の気持ちを読むことが、非常に苦手であることがわかった。たとえ発話者の真の感情がわからなくとも、発話者の声の印象を推定できれば、コミュニケーションロボット等と人間はより深い関係性になれると考えられる。従来の研究では感情表現などの心理学的な内容は文系、音響解析は理系の研究分野であった。しかし文理融合の研究が可能な人間科学学術院ならば、この二つの事柄を統合した、既存の手法では解決ができなかった課題を解決することが期待できる。その中でも私の興味の分野を専門に研究している、菊池英明研究室を進学先として希望する。

[1]総務省(2017)「人工知能(AI)の現状と未来」『平成 28 年版 情報通信白書 第 1 部第 4 章第 2 節』 www.soumu.go.jp/johotsusintokei/whitepaper/ja/h28/pdf/n4200000.pdf
(2018 年 4 月 12 日閲覧)

2. 当研究科での研究課題

近年、表現豊かな音声を対象とした、音響特徴量と印象との関連づけの先行研究は多数ある。たとえば佐藤らの研究[2]によって、表現豊かな音声において基本周波数の値や発話持続時間などの音響特徴量が、発話者に対しての印象を決定付けることが示されている。しかし声質と印象の関係について表現豊かな音声においては未だ明らかになっていないことが多い。仮にこの事柄が明らかになれば、声質を用いた発話者の印象推定が可能となり、対話ロボットや VR コンテンツへの応用が期待できる。またコミュニケーションにおいて意図した印象を与えやすい話し方など、新たな知見を提案できる。したがって、卒業論文及び修士での研究を通じて、この課題を解決していくことを計画し

ている。

まず卒業論文では表現豊かな音声を対象として、声質の一要素である息漏れ声と、発話者印象情報（意図・態度・個人性など、発話者に関する印象全般）の快－不快の関係性について明らかにする。すでに表現豊かな音声を対象として音響特徴量から発話者印象情報を推測できることが[2]によって明らかになっている。特に[2]では音響特徴量から要素感覚（発話者の声の印象）を推定し、要素感覚から発話者印象情報を推定した。要素感覚を用いることで、音響特徴量と発話者印象情報の関連づけを、個人性の影響を受けることなく可能とした。しかし発話者印象情報の中でも音響特徴量が類似している項目において、様々な印象が誤って予測される場合があることが確認されている。このことから[2]において音響特徴量を追加することでさらに推測の精度が向上する可能性が示唆された。

他にも[2]は音響特徴量に含まれている声質の要素として、地声、裏声、緊張した声、ささやき声、きしる声、息漏れ声、ざらざらした声などがあるとまとめている。さらに石井の研究[3]によって、息漏れ声の検出に有効な音響特徴量が提案された。しかし[3]で提案された音響特徴量は[2]において使用されていない。よって[3]の手法で得られた息漏れ声と発話者印象情報の関係は、未だ明らかになっていないことがわかる。また息漏れ声は、発話者印象情報の快－不快の次元と相関が高いことが森らの研究[4]よりわかっている。

以上のことから卒業論文では表現豊かな音声を対象として、声質の一要素である息漏れ声と、発話者印象情報の快－不快の関係性を明らかにする。

表現豊かな音声としては宮島らの研究[5]の一環として収録された「表現豊かな音声（通称「千の声」）コーパス（以下、SEN と呼ぶ）」を使用する。[5]で用いられた収録手法による演技音声は、従来の演技音声と比べて自然性や表現の多様性が向上することが示唆されている。具体的な研究手法としては、MATLAB のパッケージであるVoiceSauce を使用しての音声の解析や、統計学的手法を用いた分析を計画している。また印象評定実験で各音声に快－不快の得点をつけ、息漏れ声との関連づけをすることで、息漏れ声と快－不快の関係性を明らかにしていく。

以上の研究を経て、声質の一要素である息漏れ声と、発話者印象情報の快－不快の関係を卒業論文でまとめた後に、修士以降の研究で声質の他の要素についての同様の研究を行う。

- [2]佐藤安里, 菊池英明, 市川熹, "音声の発話者印象情報の知覚・認知モデル構築", 電子情報通信学会技術報告, 111(471), pp.89-94, 2012.
- [3]石井カルロス寿憲, "母音区間の息漏れに関する音響パラメータの検討", <http://www.irc.atr.jp/~carlos/pdf/carlos-asj0409.pdf> (2018年4月8日閲覧)
- [4] 森大毅, 相田千尋, 粕谷英樹, "活性-評価次元に基づくパラ言語情報ラベルの音響関連量", 音講論, 2005.
- [5]宮島崇浩, 菊池英明, 白井克彦, 大川茂樹, "演技指示の工夫が与える音声表現への影響 ―表現豊かな演技音声表現の獲得を目指して―", 音声研究, Vol.17, No.3, pp.10-23, 2013.

3. その課題について自身で勉強したこと

研究課題で述べた通り, 卒業研究では SEN を使用し声質の一要素である息漏れ声と, 発話者印象情報の快－不快の関係を調べる. したがってその前段階として SEN の特徴及び声質の特徴に関して, 既知の事柄を先行研究調査によって学習した. また解析に必要な技術として, 解析に使用する MATLAB, VoiceSauce の使用法や, 統計学的手法の使い方, 音声コーパスの扱い方などについて学び, 研究基盤を整えた.

3年次の終わりに執筆したプレ卒論では, 解析に必要な技術を用いて印象評価語と音響特徴量の対応関係について実際にまとめた. これは卒業研究で分析をする前に, SEN を用いての音響分析練習として行ったもので, 一連の経験を通じて音響分析に必要な解析ソフトの使用法や表現豊かな音声の重要性について学んだ.

4. 入学後の研究予定

当研究科での研究課題で述べたように, 声質と発話者印象情報の快－不快の関係を明らかにするための研究を行う. 卒業論文では声質の一要素である息漏れ声のみを扱うが, 息漏れ声だけでなく声質全般を解析することで, 声質と発話者印象情報の快－不快の関係についてのさらなる解析を行う.

並行して, さらなる多面的な分析を行うために, 様々な解析技術を習得していくことを予定している.