

1. 志望動機

私が大学院進学を志す理由は、音声コミュニケーションに現れる表現の違いに感心を持っているからである。日常的なコミュニケーションを通じて、発話者と聴取者の相互にどのような音声の特徴があるのかについて解明したいと考える。この研究を達成することで言語学習者には学習の難しい様々なコミュニケーション場面の表現の教育に貢献したいと考えている。私は、早稲田大学人間科学部入学以前から英語学習を通じて、言語について興味を抱いてきた。特に、日本語と英語のコミュニケーションの仕方の違いに興味を持ち、日本語の音声の表現の仕方に関心を持った。大学入学後、言語情報科学や人間計測・人間情報モデリングの授業を受講した。これらの授業を履修することで言語を定量的に分析する手法への関心が強くなった。さらに、言語の研究には文系領域の分析手法のみでは不十分であり、理系領域の分析手法が不可欠であることを知った。そこで、二年次の研究室選択の際に、日本語の音声を理系領域の分析手法を用いて研究を行っている菊池研究室への配属を希望した。

言語の研究では、文系領域の分析手法だけでは問題を解決することは難しい。文系領域の分析手法だけでなく理系領域の分析手法を用いて分析を行うという文理融合の研究が可能な人間科学学術院ならば、既存の手法では解決できなかった課題の解決が期待できる。そして、私の研究に用いる日常的な音声のデータが収録された『日本語日常会話コーパス』は菊池英明研究室で扱うことができる。さらに、音声の分析のための専門的な知識を学ぶことができるため、人間科学研究科の中でも菊池英明研究室を進学先として希望する。

2. 当研究科での研究課題

国立国語研究所では 2016 年から 2022 年まで様々な対話場面の日常会話 200 時間をバランス良く収録した大規模な『日本語日常会話コーパス』が構築され公開された [1]。そのため、日常的かつ自然な音声のデータを分析することが可能になった。一方で、コーパスの完成が 2022 年であるため、コーパスの音声の分析は進んでいないという課題がある。

音声研究において、自発音声は朗読音声に比べて複雑であるため、自発音声におけるパラ言語情報の研究は難しいことが述べられている [2]。パラ言語情報とは、話し手が意図的に生成した、文字情報に書き起こすことのできない情報である [3]。自発音声の複雑さゆえに自発音声の研究は避けられてきた。日々行っている対話の多くは自発音声であり、自発音声の研究は重要である。

従来までは、対話場面が限定されたコーパスしかなかったため、日常的かつ自然な音声の分析を行うことはできなかった。例えば、オンラインゲームを行いながら話してい

るコーパスについての分析[4]や講演を行なっている音声のコーパスについての分析[5]である。これらの研究には、パラ言語情報の表出としての表現が限定的になってしまいうという問題がある。また、従来までの研究は、異性と同性・親密さの違いなどの限られた関係の分析にとどまっている。そのため、多様な関係性を考慮した包括的な分析を行うことができていない。そこで、日常的かつ自然な音声に対して多様な話者の関係性を考慮した研究が必要であると考えられる。日本語の会話の仕組みやパラ言語情報の生成の仕組みを明らかにするためには、会話者の関係性や状況を考慮して分析を進めていく必要がある。

[1] 国立国語研究所, 大規模日常会話コーパスに基づく話し言葉の多角的研究,

<https://www2.ninjal.ac.jp/conversation/index.html> (2022年3月31日参照)

[2] 前川喜久雄. "コーパスを利用した自発音声の研究." 東京工業大学リサーチレポジトリ

[3] Fujisaki, Hiroya. "Prosody, models, and spontaneous speech." Computing prosody. Springer, New York, NY, 1997. 27-42.

[4] 有本泰子, et al. "感情音声のコーパス構築と音響的特徴の分析-MMORPGにおける音声チャットを利用した対話中に表れた感情の識別." 情報処理学会研究報告音楽情報科学 (MUS) 2008.12 (2008-MUS-074) (2008): 133-138.

[5] 前川喜久雄. "『日本語話し言葉コーパス』を用いた自発音声の分析." 研究報告人文科学とコンピュータ (CH) 2011.4 (2011): 1-4.

3. その課題について自身で勉強したこと

当研究の予備分析として菊池研究室に収録された新入生対話コーパスを使って、親密度の変化によって会話の非流暢性がどのように変化するか分析した。収録されている音声に対して発話時の言い間違いや言い淀みのアノテーションを行なった。分析を通して、会話の内容、発話時間、発話回数などのデータをどのように扱うかを学んだ。分析の結果から、親密さが高くなることによって、言い淀みや言い間違いが増加するという傾向は得られなかった。しかし、このことから会話全体を観察するのではなく、話者ごとに観察すべきであるという結果が得られた。また、音声ファイルから母音のフォルマントを分析する方法を習得した。さらに、分析を行う上で必要になる統計的な手法について学んできた。具体的には、多数の変数からどのようなデータの傾向があるかを調べる主成分分析や多数の変数を分類するクラスター分析などである。また、自身の研究の先行研究調査として、国立国語研究所が主催した日常会話コーパスのシンポジウムに参加した。話者の関係性について行われている研究について以下に述べる。会話の相手・場所・状況の違いによって声の高さである基本周波数が多様に変化することが述べられていた [6]。しかし、分析対象は基本周波数のみに留まっ

ており、パラ言語情報についての分析を行うことはできていない。また、対話形式と参加者の関係が対話に影響を与えていることがわかっている[7]。このことから、話者の関係性がパラ言語表現に与える影響があることが考えられる。

[6] 石本祐一. "『日本語日常会話コーパス』見る日常会話音声の基本周波数の多様性." 言語資源活用ワークショップ発表論文集 = Proceedings of Language Resources Workshop. No. 5. 国立国語研究所, 2022.

[7] 千葉 祐弥, 東中 竜一郎. "日常会話エージェントの構築を目標とした対話状況認識と日常会話の構成要素の特定." 言語資源活用ワークショップ発表論文集 = Proceedings of Language Resources Workshop. No. 5. 国立国語研究所, 2022.

4. 入学後の研究予定

入学後の研究予定は、対話者間の関係性がパラ言語情報の生成にどのような影響を与えるのか。そのために様々な対話場面の日常会話を定量的に分類する。分類することで、異なる場面同士の会話を比較することが可能になる。分類する方法としては、さまざまなコーパスの発話様式を言語情報から自動推定している手法[8]を参考にする。明瞭さ、親しさ、社会階層という3つの軸上に各発話を付置することで発話を分類する。発話様式を分類し、微細なパラ言語情報の変化を捉えることで、従来までの分析手法である快-不快の関係や基本8感情などに捉われることなく、パラ言語情報を認識・識別できると考えている。分類したこれらの発話様式を推定する変数として音響的特徴を使用する。使用する音響的特徴は、明瞭さに影響を与えているとされる母音のフォルマント情報、自発発話と講演発話を分類できている X-JToBI[9]の情報である。対話者間の関係性がパラ言語表現の生成にどのような影響があるのかを調べるために、発話様式を推定した後に、会話ごとに音声を比較・分析する。

[8] 沈睿, and 菊池英明. "音声言語コーパスにおける speaking style の自動推定—転記テキストに着目して—." 自然言語処理 21.3 (2014): 445-464.

[9] 前川喜久雄, 菊池英明, and 五十嵐陽介. "X-JToBI: 自発音声の韻律ラベリングスキーム." 情報処理学会研究報告音声言語情報処理 (SLP) 2001.123 (2001-SLP-039) (2001): 135-140.