

1. 志望動機

私は、早稲田大学人間科学部入学以前から言語を学ぶことを通じて、言語について興味を抱いてきた。特に、日本語の発音と英語の発音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。そこで、二年次の研究室選択の際に、言語の中でも特に日本語の音声について分析している菊池研究室への配属を希望した。

言語への研究は文系領域の内容を定量的に分析する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。日本語の音声の中でも日常に現れる自然な音声の分析を行いたいと考える。コミュニケーションを通じて現れるパラ言語表現が発話者と聴取者の相互にどのような特徴があるのかについて解明したいと考える。また、日常に現れるパラ言語情報の表出の仕方を分析することで、場面に適したパラ言語表現の教育や表現豊かな音声合成の研究にも寄与できると考えられる。私の研究課題である日常的な音声のデータが収録された『日本語日常対話コーパス』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

音声研究において、パラ言語情報に関する研究は言語情報・非言語情報に比べて遅れていることが述べられている[1]。また、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は難しかった。具体的には、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[2]や講演を行なっている音声のコーパスについての分析である[3]。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うという問題がある。しかし、国立国語研究所では 2016 年から 2022 年までさまざまなタイプの日常会話 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』(CEJC)の構築が進んでいる[4]。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といえる。また、従来までの研究は音声のみに注目したものが多数ある。しかし、話し手から発せられた音声と聴き手が音声を知覚するまでの過程にはつながりがあることが述べられている[5]。そのため、音声のみに注目してしまうとパラ言語情報を認知すること自体や表出方法が曖昧になるという問題がある。音声のみに注目するのではな

く、対話者との関係の違いなどを含めて研究すべきである。これらの先行研究から、日常的かつ自然な音声に対して、話者の関係性を考慮した研究が必要であることが示唆される。また、話者の関係性について行われている研究では、声の高さである Fo を分析することに留まっている[6]。日本語の会話の仕組みやパラ言語情報の発声の仕組みを明らかにするためには Fo のみではなく、会話に含まれる言語情報・パラ言語情報の分析を進めていく必要があると考えられる。

- [1] 前川喜久雄. "韻律とコミュニケーション." *日本音響学会誌* 55.2 (1999): 119-125.
- [2] 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.
- [3] 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.
- [4] 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)
- [5] Denes, Peter B., Peter Denes, and Elliot Pinson, Macmillan, 1993.
- [6] 石本祐一. "日本語日常会話コーパスから見える会話場面と声の高さの関係性." *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, 2020.

3. その課題について自身で研究したこと

音声分析を行うために、音声分析ソフトウェアである Praat の使い方とデータ分析のためのプログラミング言語として Python を習得した。音声に含まれる母音のフォルマントを分析するために Julius で音素セグメンテーションを行い、wav ファイルの母音が発声されている時間から母音のフォルマントを分析する方法を習得した。また、Python を使い TextGrid ファイルから会話の内容、発話時間、発話回数など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。

4. 入学後の研究予定

さまざまなタイプ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 Speaking Style を言語情報から自動推定している手法[6]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プロットすることで発話を分類する。Speaking Style 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 Speaking Style 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分類できている X-JToBI[7]の情報である。Speaking Style を推定した後に、会話ごとを観察する。どのような条件においてパラ言語表現が表出されているのかについて、また表出されたパラ言語表現の音響的特徴はどのようなものかを分析しようと考えている。

[7] 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.

[8] 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.