

〔短期在外研究〕

研 究 者	東京大学 学際情報学府 博士課程 平城 裕隆	2256103
研 究 課 題 名	複数話者の対面での会話が記録可能な音声入力デバイスの開発	
共同研究者／ 所 属、職 名	Shyam Gollakota/ ワシントン大学 (University of Washington), 教授	
在 留 地 域	アメリカ / シアトル	
渡 航 期 間	2025 年 11 月 30 日～ 2026 年 2 月 26 日 (89 日間)	

〔在外研究の目的・意義〕

以下の 3 つのことを目的・意義としてアメリカシアトルにあるワシントン大学の Mobile Intelligence Lab にて 3 ヶ月滞在し、研究を行なった。

1. 複数話者の会話における音声インタフェース研究の発展

複数話者音声入力システムの研究を出発点として、実世界の自然会話を支援する音声インタフェースの研究開発を行う。特にこれまで自身で取り組んできた音声のデバイス・センシング技術を広げ、対話状態理解を含む会話支援システムへの知見を深める。

2. 国際的研究環境における研究推進方法の習得

インタフェース分野と音声 AI 分野のどちらにも貢献している国際的にもユニークな研究室に滞在することで、両者の垣根を超えた研究を推進する研究者の研究スタイルを学ぶ。

3. 海外の研究環境を支えるエコシステムの理解

日本とは異なる研究環境に身を置くことで最先端の研究が行われるエコシステムに関しても自身の知見を広める。

〔得られた成果・効果〕

本在外研究では、当初の複数話者音声入力システムの研究を出発点として、複数人の自然会話を支援する音声インタフェースの研究に取り組んだ。滞在中は、ワシントン大学 Mobile Intelligence Lab の研究者との議論を通じて、実世界の会話支援においては、単に音声を認識するだけでなく、会話の流れや発話状態を理解することが重要であるという知見を得た。

具体的には、複数話者環境に対応する音声認識、話者識別、ターンテイクング推定、対話状態理解を組み合わせたシステム構成について実装を行った。特に、発話が完了しているのか、まだ話し続ける意図があるのか、あるいは相槌のような短い反応であるのかを判定することに加えて、適切なタイミングで介入することは、AI エージェントが人間の会話を妨げずに支援するために重要である。本研究では、このような会話状態の理解を中心的な課題として設定し、音声認識結果や発話文脈をもとに会話の状態を推定する方法について実装と検討を進めた。滞在期間中に十分な評価まで行うことはできなかった。現在は、在外研究で得られた知見と実装結果をもとに、国際学会への投稿に向けてオンラインで協力しながら準備を進めている段階である。特に、リアルタイムに介入するための設計や AI が「いつ応答するか」だけでなく、「いつ介入しないか」を判断することの重要性は、今回の共同研究を通じて得られた大きな成果であり、今後の研究の中心的な

課題となった。

また、本在外研究では、研究内容だけでなく、世界トップレベルの研究環境における研究の進め方についても非常に多くの学びを得た。

滞在した研究室には世界各国から優秀な学生・研究者が集まり、日常的な議論や実装、研究発表の水準が非常に高かった。研究方針についても週に 2-3 回、多い週はほぼ毎日のペースで受け入れ先の先生と議論する機会があり、仮説を立て、実装し、結果をもとに再度方向性を修正するという短いサイクルで研究が進められていた。現地に着いて初日の議論で、「この研究テーマワクワクする？」とまっすぐ聞かれたのをよく覚えている。

さらに研究室間や教員間の交流も活発であり、ランチやセミナーを通じて新しい研究アイデアや共同研究が生まれる様子を間近で見ることができた。研究室同士の距離や他の研究室の先生と博士学生の距離が近いことも素晴らしいと思った。シェアキッチンがあり度々フードが提供されることや、PhD の学生だけでなく事務の方々や先生も使うことで自然と交流が生まれていた。

またビッグテック企業や OpenAI などスタートアップに所属する卒業生・研究者が頻繁に大学を訪れ、講演や議論を行っていたことも大きな刺激となった。大学と大企業やスタートアップの人材が常に協調的で、人材が循環することで、アカデミックな研究と産業応用が非常に近い距離で結びついていることを実感し、研究成果を社会に届けるためには、技術的な完成度だけでなく、チーム形成、外部連携、ユーザ価値の明確化が重要であることを学んだ。

またこういった循環は、国際的な競争環境からトップ人材を集めて圧倒的な成果を出すという強力なエコシステムに支えられており、一朝一夕では逆転しえないものであることも気付かされた。

世界最先端の人材が集まっていることを学生も先生も事務の方も全員が自然と認識して動いている、そういった雰囲気があった。

〔今後の発展性〕

今後は、本在外研究で得られた知見をもとに、複数人の会話を支援する音声インタフェースの研究を継続する。実験結果や評価をさらに充実させ、論文として発表し、インタフェースや音声分野の研究に貢献することを目指している。

また今回の滞在で学んだ、国際的な研究推進の速度、研究室間・産業界との連携、社会実装を見据えた研究のエコシステムの改善についても、今後の自身の研究者人生に活かしていきたい。