

報告番号	※甲	第	号
------	----	---	---

## 主 論 文 の 要 旨

論文題目 語りに傾聴を示す応答の生成に関する研究

氏 名 伊藤 滉一朗

## 論 文 内 容 の 要 旨

人間には、自らの考えや気持ち、状況や出来事などを伝えたいという欲求が存在する。この欲求を満たす方法の1つは、聴き手に語ることである。語りによって聴き手の理解や共感が得られれば、語り手の自己肯定感を高める効果や、孤独感やストレスを和らげる効果が期待できる。語ることは人間にとって重要な活動であり、あらゆる場面で行われている。語るためには、聴き手の存在が不可欠である。しかし、現代社会では、孤独化の進行に伴う聴き手不在の生活シーンが増加し、日常生活において聴き手がいることは、当たり前ではなくなりつつある。このような状況においては、人が語れる機会を増やすことは重要な課題であるといえる。

上述の課題に対して、コミュニケーションロボットやスマートスピーカーなどの会話エージェントが語りを聴く役割を担うことが考えられる。これらが聴き手として認められるには、語りを熱心に聞いていること、すなわち、傾聴していることを語り手に伝達する機能を備える必要がある。このための明示的な手段は語りに応答することであり、頷きなどのジェスチャが有効な手段の1つである。また、語りを傾聴していることを伝える手段としては、発話の表出も有力である。本論文では、傾聴を示す目的で語りに応答する発話を傾聴応答と呼ぶ。傾聴応答の代表例は、「はい」や「ええ」などの相槌であるが、感心や評価、繰り返すなど、相槌以外の応答も存在する。傾聴応答を適切に生成できれば、語り手の語る意欲を促進するなどの効果が期待できる。

傾聴応答は、任意のタイミングで生成すればよいわけではなく、適切なタイミングで生成する必要がある。また、傾聴応答には、「はい」「そうですね」「すごいですね」など、様々な応答表現が存在し、語りの内容に合わせた適切な応答表現を選定する必要がある。傾聴応答の生成タイミングや応答表現が不適切である場合には、かえって話し手の語る意欲を減退させかねない。そのため、語りの聴き手を担う会話エージェントによる傾聴応答の生成を実現するには、傾聴応答タイミングの検出と傾聴応答表現の選定を解決できる必要がある。そこで、本論文では、以下を目的とする。

- (1) 傾聴応答の生成に利用可能な傾聴応答コーパスの構築
- (2) 傾聴応答の表出タイミングとしての適切さの推定手法の提案
- (3) 傾聴応答の一種である不同意応答の生成に利用可能なコーパスの作成、並びに、その生成の実現性の検証

本研究では、主に自然言語処理の側面から、上述したそれぞれの目的の達成を目指す。

本論文は全5章から構成される。第1章は本論文の序論であり、傾聴応答の生成に関する課題及び研究動向を示すとともに、本論文の位置づけとアプローチを述べる。

第2章では、本研究が目指す傾聴応答の生成に利用可能な傾聴応答コーパスの構築について述べる。傾聴応答に関しては、その代表である相槌を中心に、データの収集と分析、生成タイミングの検出手法の提案が行われている。ただし、傾聴応答は相槌以外にも存在している。相槌に限定されない多様な傾聴応答の生成方式については、実際の傾聴応答を幅広く収集し、データの観察と分析を通じた今後の検討が待たれる状況にある。そこで、本研究では、語りに対する傾聴応答を収集するために、事前に収録された語りの音声に対し、表出するに相応しい傾聴応答の表現と表出タイミングを作業者が付与するという収集方式を採用する。この収集方式には、作業者は聴き手として傾聴応答の付与に集中できるため、本研究の想定に沿った傾聴応答を収集できるといった利点がある。さらに、傾聴応答が語りに影響しないため、単一の語りに対して複数の傾聴応答を収集できる。この収集方式によって、単一の語りに対する聴き手11名分の傾聴応答を148,962個収集した。多頻度性、多様性、網羅性、自然さの観点から収集した応答を評価することで、本方式により自然で多様な応答を大量かつ網羅的に収集できることを確認した。最後に、収集した応答を利用して、傾聴応答の生成タイミングの検出実験を実施した。

第3章では、適切なタイミングでの傾聴応答の自動生成に向けて、傾聴応答の表出タイミングとしての適切さの度合いを推定する手法を提案する。傾聴応答は、任意のタイミングで表出すればよいというわけではなく、適切なタイミングで表出する必要がある。これまでに、相槌を中心に、傾聴応答タイミングの検出に関する研究が行われてきた。人間同士の対話における傾聴応答を抽出することで、傾聴応答タイミング検出のためのデータを作成し、作成したデータに基づき検出手法を開発することが主流であった。これらの検出手法は、与えられたタイミングを傾聴応答の生成に適するか否かの二値に分類するものであり、その結果に従って傾聴応答の生成を決定することを想定している。一方、本論文では、適切なタイミングでの傾聴応答の自動生成に向けて、傾聴応答の表出タイミングとしての適切さの度合いを推定する手法を提案する。この適切さを表す指標として、所与のタイミングにおいて傾聴応答を表出する聴き手の割合である表出率を導入し、その推定を行う。表出率は、語りを傾聴する会話エージェントが傾聴応答を表出するか否かを判断するための材料として利用できる。例えば、推定された表出率をもとに、語り手の嗜好や会話エージェントの個性などを加味して、表出するか否かを柔軟に決定することが考えられる。本手法では、Transformerベースの手法でエンコードされた語りの音響情報と言語情報を用いて、表出率を推定する。本手法の推定性能を評価するために、本研究で構築した傾聴応答コーパスを用いて実験を行い、本手法の有効性を確認した。また、表出率を利用して応答タイミングを検出する実験を行い、その有用

性を確認した。

第4章では、語りの傾聴における不同意応答の生成について述べる。傾聴応答の生成に関する多くの研究では、典型的な応答である相槌が研究対象とされてきた。語りを傾聴する聴き手の基本的な応答方略は、語り手の発話を受容することである。例えば、傾聴応答の生成に関する主な研究対象である相槌は、「語り続けて」というシグナルや内容理解を示す機能を持っており、これも、語りを受容していることを伝えるものといえる。一方で、語りでは、時として自虐や謙遜などの発話が行われることがある。この場合、その発話内容を否定することなくそのまま受容することは必ずしも適切ではなく、語り手の発話に同意しないことを示す応答、すなわち、不同意応答を積極的に表出することが求められる。このように、語りの傾聴を担う会話エージェントが不同意を示すべき発話を検出し応答できることは不可欠な機能であるものの、傾聴応答生成に関する従来研究において、不同意応答の生成に関する試みは行われていない。そこで本研究では、語りを傾聴する会話エージェントによる不同意応答生成の実現性を示すために、不同意応答の生成に利用できる応答コーパスを作成できること、並びに、応答コーパスを用いて不同意応答を適切に生成できることを検証する。まず、不同意応答の生成に利用できる応答コーパスを作成するために、時間制約のない環境で語りデータに不同意応答のタイミングと表現をタグ付けし、応答コーパスを作成した。評価の結果、不同意応答タイミングを網羅的に、不同意応答表現を安定的にタグ付けできることを確認した。続いて、事前学習済みの Transformer ベースのモデルに基づく、不同意応答タイミングの検出手法、及び、不同意応答表現への分類手法を実装し、実験により応答コーパスを用いた不同意応答生成の実現性を確認した。

最後に、第5章で本論文をまとめ、今後の研究課題、及び、将来の展望について述べる。