

インタビュー回答のあいまいさの定量的評価と分析

中川航輝^{†1}
富山県立大学

1. はじめに

現代社会における意味の探求において、データの収集は不可欠である。その中でも、アンケート、インタビュー、エスノメソドロロジーなどの手法が一般的に用いられている。特にインタビューは、人々の意味世界を深く理解するための有力な手段として、量的な研究と質的な研究の両方で広く採用されており、情報収集や研究において非常に有用な手法である。アンケートと異なり、生の声を通じて情報を収集することにより、回答者の意見をより深く捉えることが可能である。また、質問に対する回答後にさらなる深掘りを行ったり、異なる角度からの質問を投げかけることができるため、拡張性にも優れている。

しかし、Nederhof[1]は、回答者は社会的に望ましいとされる回答をする傾向があるという。木戸[2]によるとインタビューであいまいな表現を使用すると、回答者の意図や情報が十分に伝わらず、誤解を招く可能性がある。また、調査結果や意見の分析が困難になり、正確な判断が難しくなることがある。研究でインタビューを行い、分析に使用する場合、研究者がインタビューのデータを使用すべきかを判断する必要がある。

そこで、本研究ではインタビューにおけるあいまいな表現の回答を定量的に分析し、その影響を明らかにすることを目的とする。インタビューであいまいな回答が多用されることは、回答者の意図する答えを正確に捉える上で障害となり得るため、あいまいな回答を効果的にフィルタリングする手法を提案する。そして、インタビューにおけるあいまいな表現の定量的分析を通じて、社会調査の精度向上に寄与することを目指す。

2. インタビュー回答のあいまいさ

2.1. あいまいさの特徴

Lyons[3]によると、あいまいさは、言語的な不明確さであり、複数の解釈が可能な表現を指す。この特徴は、特に言語学の分野で研究されており、言語の多義性や文脈依存性

が原因とされる。インタビューにおけるあいまいさは、質問の仕方、回答者の主観、文化的背景などによっても引き起こされる。例えば、回答者が自分の意見を明確に述べることを避けたり、一般的な表現を使用することで、意図的または無意識的にあいまいな回答をすることがある。

2.2. あいまいさの影響

あいまいな回答は、データの解釈と信頼性に大きな影響を与える。調査研究において、あいまいさはデータの質を低下させ、分析の信頼性を損なう[4]。また、あいまいな回答は、回答者の真意を隠したり、社会的望ましさを反映することで、誤った結論に導く可能性がある。このような影響は、特に政策決定や社会科学の研究において重要である。

3. 実験方法

実験1では構造化インタビューの形式でインタビューを行い、回答を収集する。実験2では、インタビューの回答(文章)を読んでもらい、それがあいまいであるかどうかを判定してもらうことを目的としている。収集したデータは、後の機械学習にて分析を行う。

3.1. 実験1 インタビュー回答のあいまいさの測定

本実験では、1問1答形式のインタビューを実施した。質問は合計20問あり、それぞれ「理解」、「考え」、「経験」に関する内容である。インタビューのテーマはIT関連と自身の人生やキャリアに関するものである。インタビューの対象者は富山県立大学工学部の学生6名で、インタビューは録音した後、文字起こしをした。「理解」、「考え」、「経験」について、以下のような質問を行った。

1. ITとデジタルの違いについて説明してください。
2. これからどういったことに挑戦していきたいですか。
3. 大学では、これまでにどんなことを学んできましたか。

3.2. 実験2 あいまいな回答の影響評価

本実験では、インタビューでの回答を「あいまいである」「あいまいではない」のどちらかに分類する作業を行った。使用されたデータは合計94個であり、非常に短い回答や

Quantitative Evaluation and Analysis of Ambiguity in Interview Responses

^{†1} KOKI NAKAGAWA, Toyama Prefectural University

「分からない」といった回答はデータセットから除外した。評価は Google Forms を用いたアンケート形式で実施した。アンケートの対象者は富山県立大学工学部の学生 6 名である。あいまいであるかどうかの判断は、「表現の目的に対して、意味するものが広すぎる。すなわち、ぼやけた表現になっている」という定義 [5] を提示したうえ、アンケート回答者の主観によるものとした。

4. あいまいさの定量的評価モデルの実装と検証

4.1. モデルの設計

回帰モデルを用いて各インタビュー回答のあいまい度を数値で予測する。説明変数は 1 件の回答 (テキスト) であり、目的変数はその回答のあいまい度 (数値) である。使用するデータは、3 章で収集した 94 件のインタビュー回答である。あいまい度については、「あいまいである」を 1, 「あいまいではない」を 0 として、それらを平均したものをあいまい度として使用する。

伊藤 [6] によると、「回答」が文脈依存的なものだとすれば、その分析・解釈を「文脈」に即して行わなければならない。そのため、モデルの開発には、文脈の理解に優れている BERT という NLP (自然言語処理) モデルを選定した。BERT は Google によって開発されたモデルで、事前学習済みであるため、文章全体の意味を捉え、その中のあいまいさを評価するのに適していると考えた。

4.2. モデルの実装

この節では、BERT モデルの実装と学習プロセス、さらにその評価方法について詳細に論じる。まず、データセットは訓練、検証、テストの 3 部分に分割された。訓練データセットの 90% を訓練用に、残りの 10% を検証用に割り当てた。トレーニングは 4 エポックにわたって行い、各エポックでの損失計算とパラメータ更新を行った。検証プロセスでは、検証データセットにおけるモデルの一般化能力が評価され、検証損失が計算された。

4.3. モデルの評価

モデルの精度評価では、テストデータセットでの平均二乗誤差 (MSE) で行った。テストデータセットでのモデルの実行結果として、MSE は 0.022133011799471913 と計算された。これはモデルがテストデータ上で高い精度で回答のあいまいさを予測していることを示している。MSE は予測値と実際の値の差を二乗した値の平均であり、この数値が低いほどモデルの予測精度が高いことを意味する。

5. 考察

第 3 章で行われた実験では、時間的制約のため、インタビューのデータセットは 6 名のみの協力に基づいており、データ量が限定的であった。この小規模なサンプルサイズは、回答の「あいまい度」の平均値を計算する際の統計的信頼性に影響を及ぼし、その結果、平均値は必ずしも正確ではない可能性がある。また、同様の理由で、第 4 章で述べられた NLP モデルの機械学習においても、データセットの不足はモデルの性能評価における信頼性の限界を示唆している。実際に高い予測精度が得られたものの、小さいデータセットがバイアスや過学習を引き起こす可能性があるため、これらの結果には慎重な解釈が必要である。

教師あり学習と NLP モデル開発においては、理想的にはインタビュー回答データは少なくとも 1,000 件から 10,000 件が必要であり、各回答に対しては 100 人以上の評価者からのフィードバックを得ることが望ましい。このような大規模なデータセットは、モデルの精度と信頼性を大幅に向上させるであろう。これらの数値は、業界標準や以前の研究に基づいて提案されており、今後の研究でこれらの基準を満たすことが重要である。

6. おわりに

本研究ではインタビュー回答のあいまいさを定量的に評価することに焦点を当てた。今後の研究では、このあいまいさが回答者の年齢、性別、職業といった属性とどのような関係にあるかを探求することが有益だろうと考える。また、インタビュー時の録音データから、話者の間のポーズや声のトーンといった非言語情報を分析し、これがあいまいさの評価にどのように影響するかを検討することも考えられる。

さらに、インタビューの質問項目を「理解」、「考え」、「経験」の 3 つに分けて分析することで、これらのカテゴリーが回答のあいまいさにどのように影響を及ぼすかを明らかにすることができたかもしれない。各項目に基づいて回答をラベル付けし、機械学習を適用することで、項目ごとの回答の特徴や傾向を把握することが可能となるだろう。これらの分析は、インタビュー回答のあいまいさをより深く理解するための切り口になるのではないだろうか。

参考文献

- [1] Nederhof, A. J.: Methods of coping with social desirability bias: A review, *European Journal of Social Psychology* (1985).
- [2] 木戸彩恵調査者の先行知識及び概念認識がインタビュー調査に与える影響, 共同対人援助モデル研究 (2011).
- [3] Lyons, J.: *textsSemantics*, Cambridge ; New York : Cambridge University Press (1977).

- [4] Krosnick, J. A. and Presser, S.: Question and Questionnaire Design, *Journal of Personality and Social Psychology* (2009).
- [5] 阿部圭一情報伝達型の日本語文章に表れるあいまい表現の類型化とその改善例, 情報処理学会デジタルプラクティス, Vol. 5, No. 1 (2014).
- [6] 伊藤勇質的インタビュー調査の再概念化, 福井大学教育地域科学部紀要 (2009).