

## インタビュー対話コーパスの構築と発話意図タグの設計

篠山 学<sup>†</sup> 松本 和幸<sup>‡</sup>香川高等専門学校情報工学科<sup>†</sup> 徳島大学大学院社会産業理工学研究部<sup>‡</sup>

## 1 はじめに

日本の若者は諸外国と比べ、学力が高いにもかかわらず自己肯定感が低いという結果が出ている[1]。これは能力があっても自分でその価値を客観的に認識できていない状態といえる。自身の価値に気づく機会の一つに他者とのコミュニケーションがあるが、近年は若者のコミュニケーション能力が低下していると言われている。

我々は、インタビュー対話システムがユーザ(若者)と対話することで、ユーザの良さを引き出し、ユーザの自己肯定感を高めることを目指している。対話システムがインタビュアーとなり、ユーザがインタビューイとなる。図1にインタビュー対話のイメージを示す。図1中のSは対話システム、Uがユーザである。ユーザが対話開始の時点で部活動の実績がない事実を負い目を感じている状態(自己肯定感が低い状態)を想定しインタビューを開始している。ユーザが忘れていた、あるいは価値に気づいていない事実を引き出すことで、ユーザが自身の価値を認識することができると考えられる。インタビュー対話システムの利点は、対話システムが対話をコントロールしやすいこと、ユーザから引き出した事実の量を使って定量的に評価しやすいことがある。また、インタビュアーが機械のため、人からどう思われるかを全く気にせず済むことから心理的ハードルが下がり、ユーザが利用しやすい。

対話システムの研究は予約対話システムなどのタスク指向型と雑談対話を行う非タスク指向型に大別される。本研究は非タスク指向型に分類される。関連研究では、傾聴対話システムの[2][3]や自動相談システム[4]などがある。また、発話意図タグを提案した研究では、タスクに依存した詳細なタグを設計した研究[5]などがある。本研究では、インタビュアーが対話をコントロールし、ユーザの良さを引き出すことを目的としている。「良さ」とはユーザが特別な事とは感じていない経験や考え、行動のこととする。本稿では、インタビュー対話システムのためのインタビュー対話コーパスの構築と発話意図タグの設計について説明する。

## 2 インタビュー対話コーパス

本研究ではテレビの対談番組[6]に含まれる字幕情報を収集してインタビュー対話コーパスを構築する。具体的には、(a)対談番組を録画、(b)録画した動画ファイルから字幕情報を抽出、(c)抽出した字幕情報から対談者の発話だけを抽出、という手順でインタビュー対話コーパス

Interview dialogue corpus construction and dialogue act tags design

Manabu SASAYAMA<sup>†</sup> and Kazuyuki MATSUMOTO<sup>‡</sup>

<sup>†</sup>Department of Information Engineering, National Institute of Technology, Kagawa College, sasayama@di.kagawa-nct.ac.jp

<sup>‡</sup>Graduate School of Technology, Industrial and Social Sciences, Tokushima University, matumoto@is.tokushima-u.ac.jp

S:今、部活動はなにかしていますか?  
U:いや、なにも。  
S:そうですか。習い事などありませんか?  
U:中学までは吹奏やりました。ホルンです。  
S:すごい。ホルンできる人はなかなかいないですよ。

図1:インタビュー対話のイメージ

表1:インタビュー対話コーパスの特性

収集対話数	388
ゲストが1名の対話数	229
ゲストが2名の対話数	129
ゲストが3名以上の対話数	30
発話意図タグ付与済み対話数	6
総発話数	193046
1対話中の平均発話数	497.54
1対話中の50文字以上の発話数の平均	30.31

を構築する。対談番組を用いる利点としては、字幕の信頼性が高いことや発話者ごとに字幕が色分けされているため発話者を区別できること、第三者が理解できる対話となっていること、量が多いことなどが挙げられる。特に第三者が理解できる対話となっていることは重要であり、発話文やインタビュー技術を分析する場合に十分な情報量が確保されていると言える。例えば、家族間や友人間の対話では暗黙の了解や共通認識により発話内容が省略される場合があり、第三者には対話が分析しにくくなってしまふ。

2017年9月8日から2019年11月29日までの対談番組(1回30分、5回/週)からインタビュー対話コーパスを構築した。インタビュアーは1名のみで、インタビューイ(以下、ゲスト)は1名から3名程度である。再放送の回や録画漏れを除き、全部で388対話を用いた。放送1回分を1対話とする。表1に構築したインタビュー対話コーパスの特性を示す。放送日ごとに1つのファイル(csv形式)に保存し、ファイル名は放送日としている。例えば2017年9月8日分であればファイル名は20170908.csvである。構築したインタビュー対話コーパスの例を表2に示す。発話文IDは発話文を識別するためのIDでファイル名と行番号から成る。発話者は発話した人を表しており、Kがインタビュアーを表し、G1~G3がゲストを表している。発話文は発話者の発話の字幕テキストである。1つの発話は読点または疑問符で終わり、同じ発話者の発話が連続する場合もある。タグは発話文の発話意図を発話の欲求や機能により分類した発話意図タグである。現在、発話意図タグは6対話(ゲストが1名の5対話とゲストが2名の1対話)のみに付与している。

## 3 発話意図タグの設計と付与実験

インタビュー対話コーパスに付与する発話意図タグを設計した。発話意図タグは関連研究[5]を参考に発話の機能で分類した。また、発話は人間の欲求[7]から生起する

表 2:インタビュー対話コーパスの例

発話文 ID	発話者	発話文	タグ
20171019_125	K	でも、あなたお一人の休みの時ってどんな過ごし方してらっしゃるの？	q2k
20171019_126	G1	そうですね。	dk
20171019_127	G1	なかなか、お休みがありがたいことに、今ないんですけど、やっぱり、こうちょっと、仕事の合間にできた2時間とかが、すごく幸せでその2時間の時には例えば、カフェに行ったりとか。	a2m
20171019_128	K	カフェ？	dk
20171019_129	G1	はい。	a1
20171019_130	G1	本屋さんに行ったりとかちょっと、こう日常が見たくてですね。	a2m

表 3:設計した発話意図タグ

タグ	発話の機能	発話の目的 (欲求)
a1	情報の提供 (Y/N)	-
a2s	情報の提供 (Y/N 以外)	共感・称賛・感心 (親和・服従)
a2m	情報の提供 (Y/N 以外)	説明 (承認、顕示、証明)
q1k	情報の要求 (Y/N)	確認・驚き・うながし (認知)
q2k	情報の要求 (Y/N 以外)	確認・驚き・うながし (認知)
ds	注目や理解の提供 (相槌)	共感・称賛・感心 (親和・服従)
dk	注目や理解の提供 (相槌)	確認・驚き・うながし (認知)
w	笑い声	-
z	挨拶・謝罪・感謝	-
f	フィルター	-

と考え、発話の機能をさらに細分化した。試作した発話意図タグを表 3 に示す。情報の提供 (Y/N) は「はい」または「いいえ」で返答した発話に付与され、情報の要求 (Y/N) は「はい」または「いいえ」で返答可能な質問に付与される。

設計した発話意図タグがタグを付与する作業者によってぶれずに付与できるのかを確認するために発話意図タグの付与実験を行った。実験には作成したインタビュー対話コーパスから 6 対話を用いた。発話意図タグ設計者と非設計者の 2 名が 6 対話にタグ付けを行った。2 名によるタグ付与結果の一致率の計算には Cohen の  $\kappa$  を用いた。発話意図タグの一致率を表 4 に示す。0.67 <  $\kappa$  のときに十分な信頼性があるとみなす [8]。1 つの対話を除き 0.67 以上あることから作業者によるぶれが少ないタグが設計できたと言える。

今回は関連研究を参考にして Y/N と Y/N 以外に分類した。しかし、インタビュー対話に必要なかどうかは検討す

表 4:発話意図タグの一致率

対話	$\kappa$
20170926.csv	0.686
20180815.csv	0.700
20180621.csv	0.673
20180413.csv	0.664
20180427.csv	0.683
20171019.csv	0.760

る必要がある。必ずしも Y/N で返答しない場合が多いからである。例えば、「こういう人がいいっていうのはあるんですか？」という発話は Y/N で返答は可能であるが、多くの場合 Y/N 以外で返答されてしまう。

#### 4 おわりに

本稿では、ユーザの良さを引き出すことのできるインタビュー対話システムの構築を目指し、テレビの対談番組の字幕情報を元にインタビュー対話コーパスを構築した。また、発話の機能や発話の欲求の観点から発話意図タグを設計した。設計した発話意図タグを構築したインタビュー対話コーパスの 6 対話に作業者 2 名がそれぞれタグ付けした。発話意図タグの一致率を  $\kappa$  を用いて計算した結果、1 対話を除いて 0.67 以上となり作業者によるぶれが少ないタグを設計できた。

#### 謝辞

本研究は、JSPS 科研費 JP19K12174 の助成を受けたものです。

#### 参考文献

- [1] 高校生の生活と意識に関する調査報告書-日本・米国・中国・韓国の比較-, 国立青少年教育振興機構, <http://www.niye.go.jp/kenkyu-houkoku/contents/detail/i/98/>, (2015).
- [2] 下岡 和也, 徳久 良子, 吉村 貴克, 星野 博之, 渡部 生聖, 音声対話ロボットのための傾聴システムの開発, 自然言語処理, pp. 3-47 (2017).
- [3] 徳久 良子, 寺嶋 立太, 乾 健太郎, 高齢者と家族とのコミュニケーションの質の向上に向けて: 高齢者の Quality of Life 表出發話の分析, 自然言語処理研究会, 研究報告自然言語処理, NL-236 (11), pp. 1-8 (2018).
- [4] 和田 翔, 萩原 将文, 転移学習を用いた階層型潜在変数付きエンコーダ・デコーダによる自動相談システム, 日本感性工学会論文誌, Vol. 18, No. 4, pp. 307-314 (2019).
- [5] 入江 友紀, 松原 茂樹, 河口 信夫, 山口 由紀子, 稲垣 康善, 音声対話コーパスにおける発話意図タグの設計と評価, 電子情報通信学会論文誌, D-II, Vol. J88-D-II, No. 10, pp. 2169-2173 (2005).
- [6] 徹子の部屋, テレビ朝日.
- [7] Henry A. Murray, Explorations in personality, Oxford university press (1938).
- [8] Jean Carletta, Assessing agreement on classification tasks, Computational Linguistics, vol. 22, no. 2, pp. 249-254 (1997).