

# Kyutech コーパス：意思決定タスクを対象とした複数人対話コーパス\*

嶋田和孝<sup>†</sup> 山村 崇<sup>†</sup> 河原真太郎<sup>†</sup> Giuseppe Carenini<sup>‡</sup> Raymond Ng<sup>‡</sup>  
<sup>†</sup>九州工業大学 <sup>‡</sup>The University of British Columbia

shimada@pluto.ai.kyutech.ac.jp

## 1 はじめに

対話理解は、自然言語処理の研究の重要なタスクの一つである。従来の対話研究は、何らかの案内システムや予約システムなどの人とシステム間の対話に関するものが多かった。一方で、近年、雑談や複数人による対話に関する研究が盛んに行われるようになった。それに応じて様々な対話に関するコーパスが作成されたり、公開されている [2, 6]。

対話理解の研究のタスクの一つとして、対話要約がある。英語を対象とした研究では、AMI コーパス [1] や ICSI コーパス [3] などが存在する。これらのコーパスでは、各発話のトピックタグや正解要約など様々なアノテーションがされている。一方で、日本語においてはこのような対話要約に自由に利用できるコーパスは存在しない。

本論文では、複数人による意思決定タスクを対象としたコーパス (Kyutech コーパス) について述べる。対話コーパスのタスク設定や書き起こし方針、要約研究のためのトピックアノテーションおよび正解要約の作成方法について説明する。作成されたコーパスは Web ページ上に順次公開する<sup>1</sup>。

## 2 Kyutech コーパス

本節では、Kyutech コーパスのタスク設定と書き起こしテキストの作成方針について述べる。

### 2.1 対話環境とタスク

本コーパスでは、4名の参加者による複数人対話を対象としている。4名の参加者は、ある都市のショッピングモールの経営者であるという設定のもと、そのショッピングモールのレストラン街にあるレストランの閉店後に来店させるレストランを3つの候補の中から1つ選ぶというタスクについて議論する。議論のた

この度あなた方には、商業施設「UBCモール」内のレストラン街において、「定食和屋(わや)」閉店後、次に来店させる店舗を、UBCモールの経営者の立場で、決定していただきたいと思ひます。

店名	台湾スーパードル	つけ麺 ふうじん	中華料理 四川
メニュー例	牛ステーキ: 880円 ジャージャー麺: 980円 予算: 800円~1,200円	つけ麺: 700円 餃子: 200円 700円~1,000円	麻婆豆腐: 720円 中華丼: 900円 900円~1,500円
座席数	25	30	25
営業時間	11:00~23:00	11:00~23:00	11:00~23:00
概要	20年前から県内にある、知る人ぞ知る台湾スーパードルの老舗。日本のラーメンとは違う独特なスープは癖があるが好きな人にはたまらない	全国で有名なつけ麺。県内のショッピングモールには初出店。バラエティのあるメニューと麺を食べた後のシメの雑炊(+100円)が有名。	全国で300店舗以上を有する中華料理チェーン店。手頃な値段で本格的な中華を味わえと有名。老若男女、幅広い客層に人気。
閉店後の口コミ	・ここでしか食べられない味がある (20代男性) ・スープのおいさが良かった (30代男性) ・美味しい、何度でも食べに行きたい (30代女性)	・シメの雑炊まで食べればお腹いっぱいになる (40代男性) ・つけ麺だが、普通のラーメンも美味しい (30代男性) ・あっさりしていて美味しい (20代女性)	・手頃な予算で食べられる (20代女性) ・あともう少し食べたいと思う (30代男性) ・種類も豊富で家族連れも多い (30代男性)

#### 【UBCモールに関する情報】

UBCモールは、A県U市の中心部付近にあるショッピングモール (複合商業施設) で、スーパーマーケットと60の専門店、ゲームセンター、映画館、7つのレストランなどから成ります。

#### 主なターゲット:

所在地のU市と隣接するX市の住民です。休日などにカッパルや家族連れが長時間滞在する時間消費型の施設としてのみならず、U市内に古くからある商店街の衰退に伴い、日用品の買い物を含む多くの役割を担っています。

#### 立地・交通機関など:

UBCモールには十分な無料駐車場の他に、U市内の各所からUBCモール向けのバスもあります。郊外型のショッピングモールではないため、UBCモールの周りには規模は大きくありませんが、いくつかのオフィスビルなども存在します。

U市内には他にもスーパーマーケットやディスカウントストアなどの単体の商業施設がありますが、ショッピングモールはUBCモールのみです。

#### 来客者に関する情報:

営業時間は店舗によって異なります。最も大きな施設であるスーパーマーケットは10時から23時まで開いています。以下のグラフは、UBCモール全体への時間別平均来客数 (平日と土日祝) および平日の時間別の男女比率を表しています。土・日・祝日の男女比は、男性:女性=4:6で、時間帯での男女比の大きなばらつきはありません。

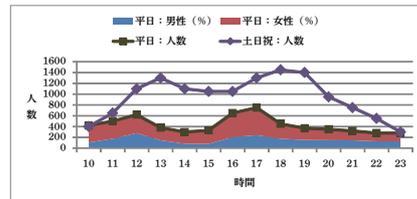


図 1: 議論の際に利用した資料の一部抜粋。

めに、出店候補店のレストラン情報、閉店するレストランの情報と閉店理由、レストラン街の既存店の情報、ショッピングモールの立地や来客者の時間帯・性別分布、ショッピングモールのある都市の人口などの統計情報、さらに隣接する市町村の情報などが書かれた10ページほどの資料が準備されている。資料の一部を図1に示す。参加者は、この資料を10分間黙読し、次に20分間の議論ののち、出店レストランを1つ決定する。現在4つのシナリオを用意しており、16名の学生から4名を選び、9つの対話データを収録している<sup>2</sup>。

\*The Kyutech corpus: a multi-party conversation corpus on a decision-making task

<sup>1</sup><http://www.pluto.ai.kyutech.ac.jp/~shimada/resources.html>

<sup>2</sup>あるタスクに対する4名を選ぶ場合は、過去の組み合わせと2名以上が被らないように配慮し、さらに同じシナリオを議論しないようにした。



図 2: 対話環境.

タグ	説明
(F) タグ	フィラー.
(D) タグ	言いよどみ・言い直し.
(Q) タグ	疑問文: 文末の表記やイントネーションから疑問文であると判断される場合.
(%) タグ	上記では判断が難しい疑問文.
(?) タグ	聞き取りに自信がない箇所.
(L) タグ	ささやき声や独り言のような小さな声.
< 笑 >	発話中もしくは単独の笑い.

表 1: 書き起こしの際に利用されるタグ.

参加者はテーブルを囲む形で座り、議論をする。議論の様子は、4人それぞれを KINGJIM MR360<sup>3</sup> で撮影し、さらに通常のビデオカメラでも撮影する。図 2 にその様子を示す。対話収録後には、最終的な選定結果への満足度や誰が決定に貢献したと思うかなどのアンケートも取っている。

## 2.2 書き起こしのルール

KINGJIM MR360 で録画された画像と音声から ELAN<sup>4</sup> を利用して書き起こしテキストを作成する。書き起こしの際には、日本語話し言葉コーパス (CSJ) の書き起こし基準 [4] を参考にする。具体的には 0.2 秒以上のポーズを転記単位の区切りとし、フィラーなどのタグ付けを行う。表 1 に現在使用しているタグの一覧を示す。

CSJ コーパスの転記単位で書き起こすと 1 つの発話が多数に分割されることが多く生じる。そこで、各書き起こしデータの行 (1 つの転記単位) に対して、以

<sup>3</sup><http://www.kingjim.co.jp/sp/mr360/>

<sup>4</sup><https://tla.mpi.nl/tools/tla-tools/elan/>

ID	Begin	End	Utterance
D	00:00:24.490	00:00:25.530	現時点で+
D	00:00:26.585	00:00:27.615	どこがいいと思う(Q)/
C	00:00:29.985	00:00:31.195	私はつけ麺/
A	00:00:31.815	00:00:33.965	自分も同じくつけ麺ふうじんがいいかな/
B	00:00:35.045	00:00:38.185	僕は中華料理かなと考えてました/
D	00:00:38.405	00:00:40.665	自分も中華料理かなっていう/
C	00:01:02.650	00:01:06.880	なんで中華料理が良いと(D 思ったか)思ったのか/
D	00:01:08.300	00:01:08.950	んー/

図 3: 書き起こしの例.

下のルールに基づいて、発話単位のアノテーションを行った<sup>5</sup>。+ はその転記単位が次の転記単位と繋がる (1 つの発話内) であることを意味し、/ は 1 つの発話の終わりを意味する。

**Step1:** まず、すべての転記単位の末尾を + とする。

**Step2:** 次の転記単位まで 1 秒以上の空白があり、以下の条件のどちらかにマッチした場合は + を / に変更する。

**条件 1:** 「動詞 (もしくは助動詞) + 接続詞」の形で終わる

**条件 2:** 次の発話が「接続詞」、「フィラー」、「副詞」のいずれかで始まる

**Step3:** Step2 で発話の終わりと認定されないが、作業者が 1 つの発話の終わりであると感じる部分は + を \* に変更する。

図 3 に実際のコーパスの例を示す。各行のアルファベットは話者 ID を意味する。例えば、「現時点で」と「どこが良いと思う」は発話認定ルールによって + と判断され、音声でのイントネーションから疑問文である (Q) と認定されていることを表している。9 つの対話データに書き起こしを行った結果、転記単位での発話数は 4511 となった。

## 3 アノテーション

本研究の最終的な目的の一つは対話の要約生成である。対話の要約には様々な手法が考えられるが、第 1 段階として、各発話のトピックを推定するということが考えられる。そこで、Kyutech コーパスでは転記単位に対してトピックセグメンテーション (トピックのアノテーション) を行う。また、最終的な要約システムの評価のために正解要約を作成する。

### 3.1 トピックセグメンテーション

本節では議論をトピック毎に分割 (トピックセグメンテーション) するための各転記単位へのアノテーション作業について述べる。まず、トピックアノテーショ

<sup>5</sup>ただし、転記単位の末尾のみを対象とする。

タグ	説明
CandX	候補店 1 に関するトピック
CandY	候補店 2 に関するトピック
CandZ	候補店 3 に関するトピック
Cands	候補店 1~3 の複数に関するトピック
Closed	閉店したレストランについてのトピック
Exist1	既存店 1 に関するトピック
Exist2	既存店 2 に関するトピック
Exist3	既存店 3 に関するトピック
Exist4	既存店 4 に関するトピック
Exist5	既存店 5 に関するトピック
Exist6	既存店 6 に関するトピック
Exists	既存店 1~6 の複数に関するトピック
CIEx	既存店および閉店したレストランの両方に関するトピック
Mall	UBC モール全体についてのトピック
OtherMall	他の場所にあるショッピングモールに関するトピック
Location	モール内でのレストラン同士の位置関係に関するトピック
Area	地域や都市についてのトピック
People	ターゲットにする顧客についてのトピック
Price	価格に関するトピック
Menu	メニューに関するトピック
Atomos	雰囲気に関するトピック
Time	営業時間などに関するトピック
Seat	座席数や回転率などに関するトピック
Sell	売り上げに関するトピック
Access	交通手段などモールへのアクセスに関するトピック
Meeting	話を進めるための議事提案や最終決定部に関連するトピック
Chat	レストラン選択には直接関係ない雑談
Vague	前後の文脈からも何を言っているかわからない場合

表 2: トピックセグメンテーションに利用されるタグ。

ンの前段階として、考えられるトピックタグの洗い出しを行う必要がある。今回は著者らを含む 4 名がそれぞれ Kyutech コーパスの中の別々の対話データに対して、その対話でどのようなトピックが存在するかを調査した。この作業を 2 回繰り返して、表 2 に示す 28 個のトピックを選定した。

次にこのトピックタグに基づき、この研究に関係していない者も含め 6 名で実際の書き起こしデータにトピックのアノテーションをした。各転記単位について、必ず 1 つのタグ（必須タグ）を付けることを条件とし、1 つでは足りないと思う場合には、追加でもう 1 つまでタグ（追加タグ）を付けることができたようにした。アノテータは 2.1 節で説明した議論のための資料にも目を通しつつ、発話の前後の流れなどを踏まえてタグ付けを行った。アノテーションは 1 対話に対して 2 名で独立に行った。事前に用意したトピック

タグ以外に新たにタグを作成することも許可したが、このアノテータによる作業では新しいタグは生まれなかった。アノテーション終了後、2 名のアノテータの付けたタグについて、その一致率を計算した。具体的には、それぞれのアノテータが付けた 2 つ以内のタグのうち、少なくとも 1 つは一致している率をここでは一致率（すなわち部分一致率）とした。その結果、2 名のアノテータ間の部分一致率は 0.879 という高い値であった。

2 名のアノテータによるアノテーション処理の後、著者のうち 3 名<sup>6</sup>で、アノテータによるタグをチェックした。ここでは必須タグ以外の追加タグを 2 つまで許すこととし、著者ら 3 名の合議によって、各転記単位のタグを決定した。この際に、アノテータの勘違いや誤記なども議論を通して修正した。この処理によって、全 4511 個のデータについて、3 つ以内のトピックタグを付与した。この 3 つの最終的なトピックタグが少なくとも 2 名のアノテータの付けたタグのうち 1 つ以上を含んでいる率は 0.965 となり、アノテータの意向を十分に考慮したトピックラベルが得られた。実際のアノテーション結果を図 4 に示す。現在の Kyutech コーパスでは、必須タグを 1 つのトピックだとみなした場合、およそ 10 転記単位で 1 つのトピックが構成されている。

### 3.2 正解要約

本節では、正解要約を作成する手順について述べる。正解要約は句読点などを含め 250 字以上・500 文字以内という制限を設けた。

正解要約作成は AMI コーパスのガイドライン（Abstractive Hand Summaries Guidelines (Scenario)）に準拠した<sup>7</sup>。このガイドラインに従って、要約生成は 3.1 節の処理の直後に、同じアノテータによって行われる。アノテータには「対話の内容を知らない人が読んだ場合でも、議論の内容を理解できるような要約を作成する」ことを依頼した。3.1 節での結果（トピックセグメンテーション）を意識はしても良いが、正解要約がそれと完全に一致する必要はないとして、要約を作成した。

ある対話に対して、あるアノテータが作成した正解要約を図 5 に示す。これらのアノテータが作成した要約とは別に、3.1 節のトピックアノテーションと同様に、2 名のアノテータの要約の総意を取った要約を著者らが作成することで、1 つの対話に対して合計 3 つ

<sup>6</sup>この 3 名中 2 名は、タグ選定には関わったが、前段の 2 名によるアノテーション作業には関わっていない。

<sup>7</sup><http://groups.inf.ed.ac.uk/ami/corpus/guidelines.shtml>

	アノテータ1		アノテータ2		最終タグ			発話
	必須タグ	追加タグ	必須タグ	追加タグ	必須タグ	追加タグ	追加タグ	
D	Closed	Sell	Closed		Closed	Sell		今回の定食はあれですよ、(D ぶさいけんだった)不採算だった/
A	Closed	Sell	Closed		Closed	Sell		そうそう/
A	Closed	Sell	Sell		Closed	Sell		不採算だったから潰されるってことは採算取れないと+
D	Closed	Sell	Sell		Closed	Sell		<笑>/
A	Closed	Sell	Sell		Closed	Sell		駄目みたいな空気が凄いあるからさあ/
D	Closed	Sell	Sell		Closed	Sell		<笑>/
A	Exist4	Sell	Exist4	Sell	Exist4	Sell		って考えたら、採算取れる店って何なんだろうつつつたら、ファミリープレートが一番売り上げが良い訳で+
D	Exist4	Sell	Exist4		Exist4	Sell		(L んー)/
A	Exist4	Sell	Meeting		Exist4	Sell		ファミリープレートが+
A	Exist4	Sell	Meeting		Exist4	Sell		なんで良いんだろうね(Q)/
D	Exist4	Menu	People		Exist4	Menu		メニュー、色んな店とか(? かな)/
C	Exist4	People	People		Exist4	Menu	People	プラス+
C	Exist4	People	People		Exist4	Menu	People	家族連れが+
C	Exist4	People	People		Exist4	Menu	People	多いて書いてあったんで、多分色々な種類あるとこに(?)+
A	Exist4	People	People		Exist4	Menu	People	んー/
A	Exist4	People	People		Exist4	Menu	People	(?)家族があつて/
C	Exist4	People	People		Exist4	Menu	People	集まるんじゃないんすかね/
D	Exist4	People	People		Exist4	Menu	People	あー/
D	People		People		People	Mall		休日はなんか消費が少なく長時間滞在、消費型って書いてありましたね、カップルや家族連れ/

図 4: 2 名のアノテータのアノテーション結果と最終タグの例。

既存店の中で最も売り上げが良いファミリープレートを例とし、主に 3 つの候補店の値段、客層、メニューの多さについて議論がなされた。値段については既存店の中に値段が安くも売り上げが少ない店舗があることや値段が安くても売り上げを出している店舗があるため、優先順位は低いとされた。また客層については、ショッピングモールの主な客層が家族連れや女性客を多くことから、それらの客層に受けの良い店舗が優先するとされた。その結果、ラーメンかいぶつは 3 つの候補店のうち値段はもっとも安い、メニューが少なくショッピングモールの主な客層である女性客に受けが悪いと判断され、棄却された。最終的には既存店に女性向けの店舗がないことを考慮し、メニューの多さよりも女性客に受けが良いかという点を優先した。その結果、つけ麺ふうじんとポノバスタを比べたとき、つけ麺ふうじ人はメニューの豊富さではポノバスタに優っているが、ポノバスタはつけ麺ふうじんよりも女性受けが良いと判断され、ポノバスタが採択された。(433字)

図 5: 正解要約の例。

の正解要約を用意する。

## 4 おわりに

本論文では、複数人による意思決定タスクを対象とした Kyutech コーパスについて説明した。現在、9 対話について書き起こしおよびアノテーション作業が終了している。1 つの対話に対するアノテーションに掛かる時間は 2 時間半程度<sup>8</sup>であった。

今後はこの Kyutech コーパスを対象に対話要約の研究を進めていく。トピックセグメンテーションの結果を用い [8]、生成型要約 [5, 7] により議論内容の要約を生成する予定である。また、抜粋型要約の正解データ (転記単位での重要度付与) の作成なども今後の課題となる。さらに対話データを収集し、コーパス自体の規模を大きくすることも今後の課題である。

謝辞：本研究は科研費 26730176 の助成を受けたものです。

<sup>8</sup>アノテータによる作業の時間。内訳はトピックセグメンテーションに 2 時間弱で、残りが要約作成。そのあと、1 対話に 2 時間程度の時間をかけて著者 3 名の合議が行われた。

## 参考文献

- [1] Jean Carletta. Unleashing the killer corpus: experiences in creating the multi-everything AMI meeting corpus. *Language Resources and Evaluation*, Vol. 41, No. 2, pp. 181–190, 2007.
- [2] 東中竜一郎, 船越孝太郎, 荒木雅弘, 塚原裕史, 小林優佳, 水上雅博. Project Next NLP 対話タスク: 雑談対話データの収集と対話破綻アノテーションおよびその類型化. 言語処理学会 第 21 回年次大会ワークショップ「自然言語処理におけるエラー分析」, 2015.
- [3] Adam Janin, Don Baron, Jane Edwards, Dan Ellis, David Gelbart, Nelson Morgan, Barbara Peskin, Thilo Pfau, Elizabeth Shriberg, Andreas Stolcke, and Chuck Wooters. The ICSI meeting corpus. In *Proceedings of the IEEE ICASSP*, 2003.
- [4] 国立国語研究所. 日本語話し言葉コーパスの構築法. No. 124. 2006.
- [5] Yashar Mehdad, Giuseppe Carenini, and Raymond T. Ng. Abstractive summarization of spoken and written conversations based on phrasal queries. In *Proceedings of the 52nd Annual Meeting of the Association for Computational Linguistics*, pp. 1220–1230, 2014.
- [6] Fumio Nihei, Yukiko I. Nakano, Yuki Hayashi, Hung-Hsuan Hung, and Shogo Okada. Predicting influential statements in group discussions using speech and head motion information. In *Proceedings of the 16th International Conference on Multimodal Interaction, ICMI '14*, pp. 136–143, 2014.
- [7] Tatsuro Oya, Yashar Mehdad, Giuseppe Carenini, and Raymond Ng. A template-based abstractive meeting summarization: Leveraging summary and source text relationships. In *Proceedings of INLG 2014*, 2014.
- [8] 山村崇, 嶋田和孝. 語彙的連鎖とトピックモデルに基づくテキストセグメンテーション. 言語処理学会 第 22 回年次大会 (NLP2016), 2016.