

Kyutech コーパスにおける発話意図タグの設計と分析

日野 優登[†] 山村 崇[†] 嶋田 和孝[†]

[†]九州工業大学大学院 情報工学府 先端情報工学専攻

〒 820-8502 福岡県飯塚市川津 680-4

E-mail: †{m_hino,t_yamamura,shimada}@pluto.ai.kyutech.ac.jp

あらまし 対話に関する研究では、幅広い分析のためにコーパスへの発話意図のアノテーションが欠かせない。そのため本論文では、我々が現在構築し、公開している意思決定タスクを対象とした対話コーパス (Kyutech コーパス) に対して、発話意図のアノテーションを行った。発話意図のアノテーションでは、AMI コーパスにおいて規定されている発話意図タグと ISO 24617-2 において規定されている発話意図タグの 2 つのタグセットを、それぞれ Kyutech コーパスに対して適用した。そのアノテーション結果を考察、比較すると、現状では、ISO 24617-2 を Kyutech コーパスに適用するほうがアノテーションの一致率やタグの詳細度の面から良いという結論が得られた。

キーワード 複数人対話, コーパス, 発話意図, アノテーション

Design and analysis of dialogue acts for the Kyutech corpus

Masato HINO[†], Takasi YAMAMURA[†], and Kazutaka SHIMADA[†]

[†] Kyushu Institute of Technology, Graduate School of Computer Science and Systems Engineering Kawadu
680-4, Iidukasi, Fukuoka, 820-8502 Japan

E-mail: †{m_hino,t_yamamura,shimada}@pluto.ai.kyutech.ac.jp

Abstract Conversation understanding is one of the most important tasks in natural language processing. Recently, many researchers have studied multi-party conversation. For conversation analysis, corpora about multi-party conversation have an important role. We are developing a freely available Japanese conversation corpus for a decision-making task. We call it the Kyutech corpus. The current version of the Kyutech corpus contains topic tags of each utterance and reference summaries of each conversation. In this paper, we focus on dialogue act tags. We annotate two types of dialogue act tags, the AMI corpus and ISO 24617-2, to the Kyutech corpus, and then discuss the results.

Key words Multi-party conversation, Corpus, Dialogue act, Annotation

1. はじめに

対話は、私たち人間が生活する上で、他人とコミュニケーションを図るための手段である。コミュニケーションは、自らの意思を相手に伝え、相手の意思を受け入れることで成立し、他人との情報共有、双方の感情理解および全体の意思決定を目的として行われる。雑談や会議といった複数人で行うコミュニケーションの形態は多人数インタラクションと呼ばれ、現在対話分析を中心に研究が進められている。

対話分析は、しばしば対話中の発話に発話意図をアノテーションしてから分析を行う。これは発話意図が対話の構造を理解する上で重要な役割を果たすためである。従来の対話研究においても、発話意図は重要視され、多くの発話意図タグ付コーパスが作成された。発話意図タグ付コーパスの作成では、まず

発話意図のタグセットの仕様を設計する必要がある。これまでに発話意図のタグセットの仕様に関して、様々な提案がされている [2] [5] [10]。中でも、著名な発話意図アノテーションスキーマとしては DAMSL [1] などが挙げられる。しかし、こうした既存のアノテーションスキーマの多くが、人と人、もしくは人と機械の二対一対話を想定した設計となっている場合が多い。一方で、今回、我々が研究対象としているのは複数人対話である。したがって、これらを複数人対話に適用するには拡張が必要となる。

複数人対話を対象として設計された発話意図のタグセットとしては、AMI コーパス [4] の発話意図タグセットがある。AMI コーパスは会議の円滑な進行や会議録の効率的な閲覧に関する研究を目的としたコーパスで、そこで用いられている発話意図タグは、情報交換など 15 種類のタグが規定されている。また

近年, Bunt ら [3] によって発話意図アノテーションスキーマの ISO 国際標準化が進められている. ISO 24617-2 として認可されたこのアノテーションスキーマは, ドメイン非依存であり, 発話者の発話目的, 発話の意図を表すことに最適化されていることが特徴である. この ISO 24617-2 では, 既存のアノテーションスキーマでは重要視されなかった発話の非タスク的な機能も含めて発話意図タグを規定しているため, 自然対話などの非タスク指向型対話の構造を網羅的に記述することができる.

我々は, 現在, 意思決定タスクを対象とした対話コーパス (Kyutech コーパス [12]) を構築し, 公開している (注1). 現在の Kyutech コーパスには, 各発話の議論中での意味的な役割を表した 28 個のトピックタグと収録された 9 つの会話に対する要約が含まれている. これは, 議論要約を Kyutech コーパスの最初のタスクとして設定したためである. Kyutech コーパスをそれ以外のタスクで利用するには, トピックタグと要約以外に様々なアノテーションが必要である. その中でも, 発話意図の付与は, 対話分析には欠かせない. そのため, 本論文では, Kyutech コーパスに発話意図のアノテーションを行い, その内容を議論する. 発話意図タグとしては AMI コーパスに基づくものと ISO 24617-2 規格に基づくものの 2 種類を採用し, それぞれとアノテーション結果について比較, 検討する.

2. Kyutech コーパス

まずは今回の処理対象である Kyutech コーパスについて説明する. Kyutech コーパスは複数人対話の対話理解を目的として収集されたものである. 収録されている対話は, 4 名の対話参加者がい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, ショッピングモールに新しく出店するレストランを三つの候補店から一つ選ぶというタスクについて議論を行う形式となっている. 議論の際には, 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. 対話参加者はこの資料を 10 分間黙読した後に, 20 分間議論を行う.

各参加者の発話は人手によって書き起こされている. 具体的には 0.2 秒以上の空白を一つの転記単位の基準とし, 書き起こしている. 対話における発言の自由度は高く, 発話にはフィラーや言いよどみ, 言い誤りや言い直しなどが含まれている. これらに対してルールや主観に基づきタグ付けがされている. また Kyutech コーパスには発話中のトピックを示すトピックタグが付与されている. これは Kyutech コーパスのタスクの一つである議論要約に関連して, 対話中の各発話のトピックを推定する必要性から作成された. トピックタグには, 出店候補店のレストランや閉店するレストラン, またそれらの売り上げやメニューなどを示すタグが用意されており, Kyutech コーパスの書き起こしには, 各転記単位に最大 3 つのトピックタグが

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	台湾モデル	つけ麺 ふじん	中華料理 四川
メニュー例	牛すたーどる: 880円	つけ麺: 700円	麻婆豆腐: 720円
予算	ジャージャー麺: 980円 800円~1,200円	餃子: 200円 700円~1,000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, 知る人ぞ知る台湾すたーどるの老舗. 日本のはラーメンとは違ふ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. パ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) がある	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る有名な, 老若男女, 幅広い客層に人気
別店舗の口コミ	・ここでもか食べられない味がある (20代男性) ・スープのおいさがダメだった (30代男性) ・美味しい, 何度も食べに行きたい (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) ・あともう少し安ければと思う (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. 休日などにカップ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も存在します.

U市内には他に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.

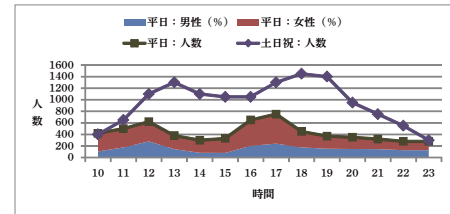


図 1 議論の際に利用した資料の一部抜粋

ID	発話	必須タグ	追加タグ	追加タグ
D	今回の定食はあれですよね, (Dふさいけんだった)不採算だった/	Closed	Sell	
A	そうそう/	Closed	Sell	
A	不採算だから潰されるってことは採算取れないと+	Closed	Sell	
D	<笑>/	Closed	Sell	
A	駄目みたいな空気が凄いあるからさあ/	Closed	Sell	
D	<笑>/	Closed	Sell	
A	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+	Exist4	Sell	
D	(Lんー)/	Exist4	Sell	
A	ファミリープレートが+	Exist4	Sell	
A	なんで良いんだろうね(Q)/	Exist4	Sell	
D	メニュー, 色んな店とか(?かな)/	Exist4	Menu	
C	プラス+	Exist4	Menu	People
C	家族連れが+	Exist4	Menu	People
C	多いって書いてあったんで, 多分色んな種類あるとどこに(?)+	Exist4	Menu	People
A	んー/	Exist4	Menu	People
A	(?)家族があって/	Exist4	Menu	People
C	集まるんじゃやないんすかね/	Exist4	Menu	People
D	あー/	Exist4	Menu	People
D	休日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/	People	Mall	

図 2 書き起こし例

付与されている.

図 2 に書き起こし結果と付与されたタグの例を示す. たとえば, 図中の“(Dふさいけんだった)”は言い直しを意味するタグである.“+”はこの転記単位が次の転記単位と 1 文という意味で同じ発話であるかを意味し, “/”が一つの発話の終わりを意味する.(Q)は音声などから質問という意図を持っていると

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

アノータに判断されたことを意味している。また図中の右側は前述のトピックタグの具体例を示している。たとえば、第1発話から第6発話までは、メインピックは閉店したレストラン (Closed) であり、サブトピックとして売り上げ (Sell) について議論されているということの意味している。詳細については、文献 [12] を参照して頂きたい。

3. タグセット

発話意図は対話の構造理解で重要な役割を持つ。発話意図をアノテーションすることは対話を分析、理解することにつながる。対話中の発話意図を表現する発話意図タグセットは、これまで数多く提案されている。本節では、本論文で比較する2種類の発話意図のタグセットについて紹介する。

3.1 AMI コーパスでの発話意図

まず AMI コーパスの発話意図タグについて記述する。AMI コーパスは会議風景を収録・観察し、その中から会議閲覧システム構築のために利用可能な会話構造を発見することと、収録した会議風景を研究目的のコーパスとして広く公開することを目的として作成されたコーパスである。収録された対話は、4名の対話参加者にそれぞれプロジェクトマネージャーなどの役割が割り振られ、その上で一つの議題について議論を行う形式となっている。対話における発言の自由度は高く、その音声にはフィラーや言いよどみ、言い誤りや言い直しなどが含まれている。AMI コーパスにおいて使用される発話意図タグ^(注2)は、情報交換や前の発話へのコメントなど、他のアノテーションスキーマに見られる基本的なもの以外にも、言いよどみやフィラーなどの自然対話でよく見られる表現もそのタグセットの中で規定している。既存の Kyutech コーパスでは、トピックタグや参照要約を作る際に AMI コーパスのガイドラインを参考にした。よって発話意図タグに関しても、AMI コーパスの発話意図タグ付与マニュアルを参照し、その発話意図タグの適用を試みる。

3.2 ISO 24617-2 の発話意図

次に ISO 国際標準規格である ISO 24617-2 において規定されている発話意図タグセットについて説明する。このアノテーションスキーマは、Bunt らによって既存のアノテーションスキーマの中から著名な 18 のアノテーションスキーマに基づき、構築された。ISO 24617-2 において規定されているタグセットは大きく二つに分けられる。一つ目が一般目的機能タグである。これは既存のアノテーションスキーマにおいて規定されていたような、情報交換や行為に関連するタグのことである。二つ目が次元特有機能タグである。これは既存のアノテーションスキーマでは重視されてこなかった発話の相互作用的な機能を表し、対話中のターン交替を表すターン管理機能やトピックの始まりと終わりを表す談話管理機能、あいづち機能など大きく分けて九つの機能に分けられる。この次元特有機能タグは、その機能ごとにそれぞれ独立しているため、複数のタグを一つの

発話に付与することが可能である。そのため ISO 24617-2 では、これら一般目的機能タグと次元特有機能タグを組み合わせることにより、対話の構造や発話の多機能性をより網羅的に記述できる。

ISO 24617-2 をコーパスに適用した例としては Petukhova ら [11] の研究があげられる。彼らは既存のコーパスに対し、ISO 24617-2 の発話意図タグを適用し、その汎用性を示した。また日本語のコーパスにおける ISO 24617-2 の適用例としては、井岡 [9] らの研究がある。彼らは日本語で収録されたタスク指向型対話コーパスであるホテル予約模擬対話コーパスに対し、ISO 24617-2 による発話意図タグの付与を行い、ISO 24617-2 が日本語においても汎用性をもつことを示した。また非タスク指向型対話かつ自然対話である ESP-C に適用し、その結果、ISO 24617-2 による発話意図タグ付与が網羅的な対話の記述に有効であることを示した。これらの研究から ISO 24617-2 による発話意図タグ付与は Kyutech コーパスにも適用可能であると考えられる。

4. Kyutech コーパスへの発話意図付与

本節では、前節で紹介した二つのアノテーションスキーマを Kyutech コーパスに適用し、アノテーションの安定性やアノテーション結果の分析を行う。

4.1 AMI コーパスの発話意図タグの適用

まずはじめに AMI コーパスの発話意図タグを Kyutech コーパスに適用して、その発話意図タグ付与の安定性とタグを付与した結果の分析をする。以下に、実際に行った発話意図タグ付与の手順と使用したタグについて記述する。

発話意図タグを付与する基本単位については Kyutech コーパスの転記単位を用いる。しかし、この転記単位は 0.2 秒のポーズを目安に発話を区切った単位となっているため、発話内容が継続している位置で言い淀みなどが生じた場合に単位が切れてしまうことや、明らかに文末形の直後で話者が休止を入れなかった場合に、当該位置が境界として認定できないということが考えられる。よって発話意図タグ付与の手順としては、まず転記単位単体をみて、付与できる発話意図タグがないかを判断する。この際、発話が途中で途切れていることなどが原因で発話意図タグが付与できなければ、判断を保留し、次の転記単位を見る。そして、当該の転記単位と保留した転記単位とあわせて考えて、付与できる発話意図タグがないかを判断する。基本的には、以上の手順で発話意図タグの付与を行う。また転記単位は 0.2 秒以上のポーズを単位の区切りとしているため、発話意図が異なる発話が同じ転記単位内に記されることがある。このような場合は転記単位を発話意図の境界で分割しても良いこととする。

次に、今回発話意図タグ付与実験において使用する発話意図タグについて記述する。今回用いる発話意図タグは以下の 15 種類である。

- Backchannel : あいづちなど、ターン (対話中の発言権) を獲得しない短い発話。
- Stall : 言い淀みや、発話内容を考えるための時間を稼ぐ

(注2) : http://groups.inf.ed.ac.uk/ami/corpus/Guidelines/dialogue_acts_manual_1.0.pdf

発話.

- Fragment : 言いかけてやめた発話.
- Inform : 話者が聞き手に情報を与える発話.
- Elicit-Information : 話者が聞き手に情報を要求する発話.
- Suggest : 話題の提案など聞き手に働きかける発話.
- Offer : 話者自身の行動を申し出る発話.
- Elicit-Offer-Suggest : 話題の提案や行動の申し出を求め
る発話.
- Assess : 前の発話に対する評価, 意見.
- Comment-About-Understanding : 話者自身の理解, 不
理解を示す発話.
- Elicit-Assess : 評価や意見を求める発話.
- Elicit-Comment-About-Understanding : 自分の発話を
相手が理解できたか確認する発話.
- Be-Positive : ありがとうなど聞き手との信頼関係に良い
影響を与える発話.
- Be-Negative : 聞き手の気分を害する冗談や攻撃的な発話.
- Other : 上記に該当しない発話意図をもつ発話.

この AMI コーパスの発話意図タグの定義をもとに, これらの
タグセットを Kyutech コーパスに適用するタグ付与マニュアル
を作成した^(注3).

発話意図タグとタグ付与マニュアルを評価するために発話意
図タグ付与実験を行った. 実験には Kyutech コーパスに収録
された 9 対話のうち, 1 対話を用いた. この対話は, 504 転記
単位の発話で構成されている. 実験では 2 人の作業者にタグの
付与を依頼し, その結果からタグ付与が安定的に付与されて
いるかの評価を行なった. 評価の際は, 作業者間でタグ付与結果
がどの程度一致するか, その値を求めた. 本論文では一致率の
尺度として Cohen の κ 値を用いる [6]. Cohen の κ 値は, 観測
された一致率を $P(O)$, 期待された一致率を $P(E)$ とすると,

$$\kappa = \frac{P(O) - P(E)}{1 - P(E)}$$

という式で求められる.

実験の結果, 作業者間の κ 値は, 0.371 という低い値となっ
た. この値から, 今回使用した発話意図タグとタグ付与マニユ
アルでは, 安定的に付与することが現状では難しいことが分
かった.

不一致の傾向を調べるために, 作業者間のタグの不一致が起
こった回数の多い組み合わせを計数した. 表 1 に作業者間でタ
グの不一致の回数が多い上位 5 組を示す. タグの組み合わせ
の後ろの括弧内に記されている数字は, その組み合わせが出
現した回数を表している. 結果を見ると, Assess と Inform,
Suggest と Inform, Assess と Backchannel が混同される場合
が多いことが分かる.

これらを詳しく分析するために, この 4 つのタグの作業者別
の使用回数を求めた. その結果を表 2 に示す. 表 2 を見ると,
作業者ごとに付与するタグにばらつきがあることが分かる. と

(注3) : 音声聞き取れないなどの理由で上記の発話意図タグのどれも付与でき
ない場合, その転記単位には Inaudible を付与することとした.

くに, 作業者 2 は Inform タグを多用しており, Backchannel タ
グや Suggest タグはあまり使用していない. これらの結果から,
今回の発話意図タグ付与において一致率が低いのは, タグの定
義が作業者によって解釈に幅があったこと, タグ同士を区別す
る基準が明確ではなかったことが原因であると考えられる.

表 1 タグの不一致が多い組み合わせ上位 5 組 (AMI)

タグの組み合わせ
Assess - Inform (57)
Inform - Suggest (50)
Assess - Backchannel (45)
Assess - Suggest (29)
Assess - Fragment (8)

表 2 作業者別タグ使用回数 (AMI)

タグ名	作業者 1	作業者 2
Assess	141	148
Backchannel	80	22
Inform	88	196
Suggest	126	76

4.2 ISO 24617-2 の発話意図タグの適用

次に ISO 24617-2 において規定されている発話意図タグを
Kyutech コーパスに適用して, その発話意図タグ付与の安定性
とタグを付与した結果の分析を行った.

アノテーションの手順は AMI コーパスの発話意図タグを用
いた場合と同様のアプローチをとった. タグ付与で用いる発話
意図タグは, 一般目的機能タグについては ISO 24617-2 におい
て規定されている一般目的機能タグのうち 12 種類のタグ, 次
元特有機能タグについては規定されているタグ全てを用いる.
12 種類の一般目的機能タグの選定については AMI コーパスで
使用されている発話意図タグを参考にした.

以下に今回用いた一般目的機能タグ 12 種類を示す^(注4).

- Question : 話者が聞き手に情報を要求する発話.
- Inform : 話者が聞き手に情報を提供する発話.
- Answer : Inform のサブセット. 話者が聞き手の要求す
る情報を提供する発話.
- Agreement : Inform のサブセット. 話者が提供する情報
に対し, それが正しいことを伝える発話.
- Disagreement : Inform のサブセット. 話者が提供する
情報に対し, それ間違いだと伝える発話.
- Confirm : Answer のサブセット. 対話相手の不確実な
情報に対し, それを確認し同意するための発話.
- Disconfirm : Answer のサブセット. 対話相手の不確実
な情報に対し, それを確認し否定するための発話.
- Offer : 話者自身の行動を申し出る発話.
- AcceptSuggest : 対話相手からの提案を受け入れるため
の発話.

(注4) : ただし, 今回用いた 12 種類の一般目的機能タグに当てはまらない発話
には Other というタグを付与することとする. これは AMI コーパスの Other
と等しい. また音声聞き取れないなどの理由で発話意図タグが付与できない場
合, その転記単位には Inaudible を付与することとした.

- DeclineSuggest : 対話相手からの提案を拒否するための発話。
- Suggestion : 聞き手に対し、特定の行為を提案する発話。
- Request : 聞き手に対し、特定の行為を要求する発話。

次元特有機能タグは、発話のもつ機能によって9つのクラスに分けられる。以下にその概要を示す。

- (1) Task : 対話の目的や流れに必要な不可欠なもの。情報提供や問題解決など。
- (2) Auto-Feedback : 聞き手が話者に対する注目や理解などの状態を示すもの。話者に対する相槌などがこれに含まれる。
- (3) Allo-Feedback : 話者の直前の発話に対する聞き手の感じ方や解釈などの状態を示すもの。
- (4) Turn-Management : 発言権の移行に関係するもの。
- (5) Time-management : 言いよどみやポーズなど、対話中の話者の時間を保持するもの。
- (6) Discourse-Structure-management : 対話の開始や話題の開始など、対話の構造を示すもの。
- (7) Own-Communication-management : 話者が自身の発話に対し、補足や訂正を行うようなもの。
- (8) Partner-Communication-management : 聞き手が話者に対し、その発話内容を補足や訂正するもの。
- (9) Social-Obligations-Management : 挨拶や自己紹介など、社交的な対話に用いられるもの。

これらの発話意図タグを用いて、AMI コーパスのときと同じ作業員2名に対話データへの発話意図タグ付与を依頼し、その結果を考察した。対話データにはAMI コーパスの発話意図タグを付与する実験で用いたものと同じ対話データを用いた。ここで、アノテーションでは、一般目的機能タグを必須タグ、次元特有タグを任意タグとした。まず、必須タグである一般機能目的タグのみについて評価する。4.1節同様、Cohenの κ により評価した。

実験の結果、作業員間の κ 値は0.496であった。この値は高いものではないが、文献[2]によれば、 $0.40 < \kappa < 0.60$ はfairであるとFleiss[8]は述べている。また、AMI コーパスの場合の κ 値が0.371であり、それよりも0.1ポイント以上改善していることを考えれば、ISO 24617-2は比較的機能していると考えられる。

次に、AMI コーパスの場合と同様に、不一致の傾向を調べるために、作業員間のタグの不一致が起こった回数の多い組み合わせを計数した。表3に作業員間でタグの不一致の回数が多い上位5組を示す。タグの組み合わせの後ろの括弧内に記されている数字は、その組み合わせが出現した回数を表している。表3を見るとInformタグに関する混同が多いことが分かる。AgreementタグやAnswerタグはInformのサブタグであり、これらとInformタグの混同は、作業員間で当該の発話は情報を提供しているものだという判断に関しては一致していることが分かる。特に表3で問題となるのが、InformタグとQuestionタグの混同である。これらは一方は情報を提供する発話、一方は情報を要求する発話に付与されるタグであり、正反対の働きをもつ。これらが混同された発話を見てみると、“～

て高いですよ”といった、確認質問とも、情報提供とも取れる発話がほとんどであった。これらの区別は難しく、今後考察が必要となる。

表3 タグの不一致が多い組み合わせ上位5組 (ISO)

タグの組み合わせ
Agreement - Inform (30)
Inform - Other (28)
Inform - Question (25)
Answer - Inform (24)
Agreement - Other (24)

一般目的機能タグに関するタグ付与の結果に注目すると、今回使用した一般目的機能タグの定義に当てはまらず、Otherのタグが付与された転記単位が全部で118個存在した。その内、次元特有機能タグも付与されなかった転記単位が22個存在した。これらの発話内容を見てみると、言いかけてやめた発話や独り言のような発話が多かった。前者については、AMI コーパスにおいてFragmentという発話意図タグが用意されており、表現可能な発話となっている。このような発話は人同士の対話に頻出するものであり、これを表現するタグを追加することでより詳細に対話を分析できると考えられる。

次に次元特有タグについての考察を行う。表4は次元特有機能タグのクラスごとの出現数を示す。今回アノテーションの対象としたKyutech コーパスは、対話参加者がある議題について会議を行う様子を収録したコーパスであるためTask機能を持つ発話が多く、全体の8割を占める。またTask機能以外の次元特有機能タグについて見ると、Task機能以外の次元特有機能のタグが付与された平均の数は504転記単位中480個(95%)であり、概ね一つの転記単位にはTask機能以外にも一つの次元特有機能タグが付与されていることが分かる。

表4 次元特有機能タグのクラスごとの出現数

クラス	Times(ALL)
Task	390
auto-Feedback	169
allo-Feedback	0
Turn-Mgt.	185
Time-Mgt.	55
Discourse-Structure-Mgt.	26
Own-Comm-Mgt.	38
Partner-Comm-Mgt.	7
Social-Obligations-Mgt.	0

次元特有機能タグの(4) Turn-Management機能について、このクラスのタグを付与する際、判断が難しいケースが存在した。それは二人以上の話者が同時に話し始めた場合である。この時、一人の話者以外が話すのをやめたならば、発言権を獲得したのはその話者ということになり、その話者の転記単位に発言権の獲得を表すタグの付与を行うことができる。しかし、今回の対話データでは、二人以上の話者が同時に発話を続けるような状況が見られた。このような状況において、誰が発言権を持つのかを正確に特定することは難しい。

次元特有機能タグの(6) Discourse Structure-Management機能には、対話の開始やトピックの始まりを意味するタグが用

意されている。このうち、トピックの始まりを意味するタグの付与が難しい場合が存在することも実験から確認できた。ここで、既存の Kyutech コーパスに含まれているトピックタグも踏まえて考察する。図 1 を例に説明する。図 1 の 7 発話目（「って考えたら...」）から 18 発話目（「あー/」）までは主タグが Exist4（既存店舗 4 の話題）で一貫している。一方で、7 発話目から 10 発話目までは、サブトピックとして Sell（売り上げ）について議論しており、それ以降は Menu（メニュー）について議論している。つまり、特定の対象（ここでは既存店舗 4）の属性（ここでは売り上げ）について話している状態から、その対象の他の属性（ここではメニュー）に議論が移り変わっている。次元特有機能タグ (6) で扱うのは、このような暗黙的にトピックの移行が行われた場面である。従って、トピックの始まりを認識するのは難しいと考えられる^(注5)。一方で、この次元特有機能タグ (6) が扱う内容は、実質的にトピックアノテーションの段階で暗に示されていると考えられる。従って、この次元特有機能タグ (6) を Kyutech コーパスのアノテーションに採用するべきかについては議論の余地が残る。

5. おわりに

本論文では、Kyutech コーパス中の対話の構造を詳細に記述することを目的とし、発話意図タグの付与を行った。発話意図タグ付与では、AMI コーパスの発話意図タグと、ISO 24617-2 において規定されている発話意図タグの 2 種類のタグセットをそれぞれ適用し、そのアノテーション結果を考察した。

AMI コーパスの発話意図タグを Kyutech コーパスに適用した場合、アノテーション結果を見ると、作業員間の κ 値は 0.371 と低く、今回用いたマニュアルでは安定的に発話意図タグを付与することは難しいことがわかった。これはマニュアルにおけるタグの定義やタグ同士の区別の基準が曖昧であったことが原因と考えられるため、マニュアルの改善が必要となる。

一方、ISO 24617-2 に規定された発話意図タグを適用した場合、一般目的機能タグの κ 値は 0.496 と、AMI コーパスの発話意図タグを適用した場合よりも 0.1 ポイント以上改善した。また ISO 24617-2 を適用した場合、次元特有タグにより、対話中の発話の機能をより詳細に記述できることを確認した。

これらの結果から、現状では、ISO 24617-2 を Kyutech コーパスに適用する方がアノテータの一致率の面やタグの詳細度の面からも良いという結論が得られた。しかし、ISO 24617-2 のタグセットをそのまま適用するかどうかは考察の必要がある。ISO 24617-2 はドメイン非依存であり、その汎用性が高いことは確認されているが、Fang ら [7] の研究は ISO 24617-2 にもドメインによって改良の余地があることを示している。こうした研究を念頭に置き、Kyutech コーパスに対する ISO 24617-2 の適用を検討していきたい。

また今回は一つの対話にしかアノテーションをしていない

が、Kyutech コーパスにはまだ 8 つの対話が存在する。今後は、ISO 24617-2 に準拠したアノテーションマニュアルの改訂とあわせて、残りの対話にも発話意図を付与し、一般公開の準備を進める予定である。

謝辞 この研究の一部は科研費 26730176 の助成を受けたものです。

文 献

- [1] J. Allen and M. Core. Draft of DAMSL : Dialog act markup in several layers. <http://www.cs.rochester.edu/research/cisd/resourced/damsl/>, 1997.
- [2] 荒木雅弘, 伊藤敏彦, 熊谷智子, 石崎雅人. 発話単位タグ標準化案の作成. 人工知能学会誌, Vol. 14, No. 2, pp. 251–260, 1999.
- [3] H. Bunt, J. Alexandersson, J. Choe, A. Chengyu Fang, K. Hosida, V. Petukhova, A. Porescu-Belis, and D. Trauim. ISO 24617-2 : A semantically-based standard for dialogue annotation. In *Proceeding of LREC*, pp. 430–437, 2012.
- [4] J. Carletta. Unleashing the killer corpus experience in creating the multi-everything AMI meeting corpus. *Language Resources and Evaluation*, Vol. 41, No. 2, pp. 181–190, 2007.
- [5] J. Carletta, S. Isard, G. Doherty-Sneddon, A. Isard, J. Kowtto, and A. Anderson. The reliability of a dialogue structure coding scheme. *Computational Linguistics*, Vol. 22, No. 2, pp. 249–254.
- [6] J. Cohen. A coefficient of agreement for nominal scales. *Education and Psychological Measurement*, Vol. 20, No. 1, pp. 37–46, 1960.
- [7] A. Fang, J. Cao, H. Bunt, and X. Liu. The annotation of the Switchboard Corpus with new ISO standard for dialogue act analysis. In *Proceeding of Workshop on Interoperable Semantic Annotation*, 2012.
- [8] J. L. Flesis. Measuring nominal scale agreement among many rater. *Psychological Bulletin*, Vol. 76, pp. 378–382, 1971.
- [9] 井岡孝徳. 日本語コーパスにおける ISO / DIS 24617-2 に基づく対話行為情報を用いたアノテーションとその分析. 修士論文, 奈良先端科学技術大学院大学, 2012.
- [10] S. Jekat, A. Klein, E. Maier, I. Maleck, M. Mast, and J. Quantz. Dialogue acts in verbmobil. *Verbmobil Report 65*, 1995.
- [11] V. Petukhova, A. Malchanau, and H. Bunt. Interoperability of dialogue corpora through ISO 24617-2 based querying. In *Proceeding of LREC*, pp. 4407–4414, 2014.
- [12] T. Yamamura, K. Shimada, and S. Kawahara. The Kyutech corpus and topic segmentation using a combined method. In *Proceeding of the 12th Workshop on Asian Language Resources*, 2016.

(注5) : トピックアノテーションでは、既存のトピックタグを意識しながらアノテーションするのに対して、次元特有機能タグのアノテーションではあくまで大ざっぱなトピックの切り替わりだけを認識しようとしている違いがあることに注意。