

談話行為情報 ユーザーズマニュアル

Version 1.0

1. 概要

付与対象

『日本語日常会話コーパス (CEJC)』コアデータ (20時間) で書き起こされたすべての発話単位*

*Japanese Discourse Research Initiative (2017). 発話単位ラベリングマニュアル version 2.1
<https://www.jdri.org/resources/manuals/uu-doc-2.1.pdf>

付与枠組み

ISO 24617-2* をベースに CEJC 用に整備した枠組みを使用**

- 当該発話を通して話し手が何を行っているかについて、当該発話の産出に関わる特徴 (用いられている言語形式、韻律、身体的振る舞い (視線・身体動作、等)、会話場面の性質、参与者間の関係、先行のやり取り、等) を手がかりに判断
- 当該発話の末尾まで聞いた時点で最も妥当なタグを選択
- 一発話単位につき、以下三種類の情報を付与
 - レベル1タグ: 基本的な談話機能に関わる情報
 - レベル2タグ: 談話の展開や会話の調整に関わる情報 (該当する場合に付与)
 - 依存関係タグ: 発話間で結ばれる関係 (該当する場合に付与、あわせて関係を結ぶ発話番号を指定)

*ISO 24617-2(2012). *Language resource management — Semantic annotation framework (SemAF) — Part 2: Dialogue acts.*

**Yuriko Iseki, Keisuke Kadota, and Yasuharu Den(2019). Characteristics of everyday conversation derived from the analysis of dialog act annotation, *Proceedings of the 22nd Oriental COCOSDA*, pp.1-6.

注意事項

- 本アノテーションは、CEJCの中から特定の談話機能やその連なりが生じている事例を見つけ出すための使用を想定し、やり取りの骨組みとなる特徴が安定的に示されることを目指して設計、付与されたものです。
- 一つの談話機能を担う発話部分が複数の発話単位に分かれていたり、一つの発話単位の中で複数の談話機能が生じている際には、中心的な談話機能以外タグに反映されていないといった箇所が複数含まれます。
- 本アノテーションの使用に際しては、タグの出現数が当該談話行為の出現数に一致するとは限らないという点に留意し、当該会話の実際の音声・映像データとあわせて使用してください。

タグの見方

【表記の例】

①	②	③	④	⑤	⑥	⑦	⑧	⑨	⑩	⑪	⑫
conversation ID	id	start Time	end Time	speaker ID	dialog Act1	relation Type1	relation1	dialog Act2	relation Type2	relation2	text
C001_001	712	513.057	517.22	IC04_美沙	T_Question			Opening			京葉線 え (R 玲)ちゃんちから京葉線 どうやって来るの?。
C001_001	713	517.133	520.26	IC01_玲子	T_Answer	Prospective	712	A-Opening	Retrospective	712	えーっと: (R 斉川)から:蘇我まで出て:。
C001_001	714	517.819	518.02	IC04_美沙	FB_Positive	Retrospective	713				うん。

【各列の意味】

- ① conversationID : 会話 ID(CEJC の仕様を参照)
- ② id : 会話毎に振られた発話の通し番号
- ③ startTime : 発話の開始時間(CEJC の仕様を参照)
- ④ endTime : 発話の終了時間(CEJC の仕様を参照)
- ⑤ speakerID : 話者名(CEJC の仕様を参照)
- ⑥ dialogAct1 : レベル 1 タグ(基本的な談話機能に関わる情報)
- ⑦ relationType1 : レベル 1 タグで示された談話行為の依存関係タグ
- ⑧ relation1 : レベル 1 タグの依存関係を結ぶ発話の id
- ⑨ dialogAct2 : レベル 2 タグ(談話の展開や会話の調整に関わる情報)
- ⑩ relationType2 : レベル 2 タグで示された談話行為の依存関係タグ
- ⑪ relation2 : レベル 2 タグの依存関係を結ぶ発話の id
- ⑫ text : 転記テキスト(CEJC の仕様を参照)

※下線箇所は談話行為情報に独自の情報

※各タグの表記はタグ一覧(2節)の墨付き括弧内を参照

2. タグ一覧

レベル1タグ

1. タスク系 Task

- 情報提供 Inform……………【T_Inform】
 - 情報、考え等を述べる振る舞い
- 情報要求 Question……………【T_Question】
 - 聞き手に特定の情報を要求する振る舞い
- 確認要求 Check Question……………【T_CheckQuestion】
 - 既に共有された情報や同意が強く予想される情報について、正しいかどうか確認する振る舞い
- 返答としての情報・確認提供 Answer……………【T_Answer】
 - 情報要求に応じる振る舞い
- 依頼系 Request……………【T_Request】
 - 依頼・指示・命令・提案・勧誘(受け手に特定の行為をするよう働きかける振る舞い)
- 依頼系への対処 Address-Request……………【T_A-Request】
 - 依頼系に応じる振る舞い(受け入れる・拒否する、等)
- 申し出 Offer……………【T_Offer】
 - 話し手が受け手のために行う・与えるつもりである物事を提示する振る舞い
- 申し出への対処 Address-Offer……………【T_A-Offer】
 - 申し出に応じる振る舞い(受け入れる・拒否する、等)
- 注意獲得 Attention Getting……………【T_Attention】
 - 聞き手の注意を獲得しようとする振る舞い
- 注意獲得への対処 Address-Attention Getting……………【T_A-Attention】
 - 注意獲得に応じる振る舞い(受け入れる・拒否する、等)
- 独り言 SelfSpeech……………【T_SelfSpeech】
 - 自身に向けた発話、あるいは聞き手に向けていると聞かれない発話を行う振る舞い

2. 社会的付き合い管理系 Social Obligations Management

- 挨拶 Greeting……………【S_Greeting】
 - 表現の例: おはよう、こんにちは、さようなら、いただきます、ごちそうさま、等
- 謝罪 Apology……………【S_Apology】
 - 表現の例: ごめん、もうしわけありません、等
- 謝罪への対処 Address-Apology……………【S_A-Apology】
 - 謝罪に応じる振る舞い(受け入れる・拒否する、等)
- 感謝 Thanking……………【S_Thanking】
 - 表現の例: ありがとう、どうも、等
- 感謝への対処 Address-Thanking……………【S_A-Thanking】
 - 感謝の表明に応じる振る舞い(受け入れる・拒否する、等)

3. フィードバック系 Feedback

- フィードバック肯定 Feedback Positive……………【FB_Positive】
 - 先行発話を受け取ったことを表明する振る舞いで、応答系感動詞・感情表出系感動詞・そう系およびその省略形が用いられているもの
- フィードバック了承 Feedback Acceptance……………【FB_Acceptance】
 - 先行発話を受け取ったことを表明する振る舞いで、先行発話の内容を了承したことを明示的に示す形式(オクケー・わかった、等)が用いられているもの
- フィードバック反復 Feedback Repetition……………【FB_Repetition】
 - 先行発話を受け取ったことを表明する振る舞いで、先行発話の一部を繰り返す形式が用いられているもの
- フィードバック語彙的反応 Feedback Lexical……………【FB_Lexical】
 - 先行発話を受け取ったことを表明する振る舞いで、特定のあいづち形式(なるほど・本当、等)が用いられているもの

4. その他 Others

- 非言語行動 Nonverbal……………【O_NV】
 - 発話単位を割り振られている非言語行動(笑い、等)
- 合図 Cue……………【O_Cue】
 - 自身の行う行為の調整や会話の調整に関わるもの(談話の区切れ目を示す、発話の調整を担う、等)
- 不明 Undeterminable……………【O_Undet】
 - どんな行為を担っているか判断困難なもの

レベル2 タグ

1. 回収系 Confirmation
 - 回収 Confirmation……………【Confirmation】
 - 既出の情報や先行発話に含意された情報を再確認する振る舞い
 - 回収への対処 Address-Reconfirm……………【A-Reconfirm】
 - 回収に応じる振る舞い(受け入れる・拒否する、等)
 - 補完 Completion……………【Completion】
 - 他者による先行発話に統語的に続く発話を産出することで、先行発話への理解や予測を表示する振る舞い
 - 補完への対処 Address-Completion……………【A-Completion】
 - 補完に応じる振る舞い(受け入れる・拒否する、等)
2. 修復系 Repair
 - 修復開始 Repair Initiation……………【RepairInitiation】
 - 発話の聞き取りや理解に関する問題に対処する必要があることを表示する振る舞い
 - 修復操作 Repair……………【Repair】
 - 発話の聞き取りや理解に関する問題に対処する振る舞い
3. 談話構造化系 Discourse Structuring
 - 談話開始 Opening……………【Opening】
 - 活動や話題の開始を構造化する振る舞い
 - 談話開始系への対処 Address-Opening……………【A-Opening】
 - 談話開始に応じる振る舞い
 - 談話終了……………Closing【Closing】
 - 活動や話題の終了を構造化する振る舞い
 - 談話終了系への対処 Address-Closing……………【A-Closing】
 - 談話終了に応じる振る舞い
4. 準備系 Preparation
 - 準備 Preparation……………【Pre】
 - 後に行おうとする行為を予告したり、その前提条件を確認したりする振る舞い
 - 準備への対処 Address-Preparation……………【A-Pre】
 - 準備に応じる振る舞い(受け入れる・拒否する、等)
 - 準備が投射する本体 Projected Action……………【ProjectedAction】
 - 準備によって先送りにされていた、本題に相当する振る舞い
5. 順番管理系 Turn Management
 - 順番取得 Turn Take……………【TurnTake】
 - 発話順番を取得する振る舞い
 - 順番開放 Turn Release……………【TurnRelease】
 - 発話順番が完結地点や区切りに達したことを示す振る舞い
 - 順番維持 Turn Keep……………【TurnKeep】
 - 発話順番を維持する振る舞い
 - 順番放棄 Turn Retract……………【TurnRetract】
 - 発話を産出しきらないうちに発話をやめる振る舞い
6. メタ発話情報系 Meta Information
 - 読み上げ Read Out……………【ReadOut】
 - 書いてある・書いている文字を読み上げる振る舞い
 - 冗談 Joking……………【Joking】
 - 行為を冗談として提示する振る舞い

依存関係タグ

関連する談話機能を担う二つの振る舞い(出来事)の関係を示す。うち二つ目の振る舞いに本タグを付与し、依存関係を結ぶ一つ目の振る舞い(出来事)が発話単位を与えられている場合、その発話 id をあわせて指定する。

- 予測的依存関係 Prospective Dependence……………【Prospective】
 - 一つ目の談話機能が特定の談話機能を要求するタイプのつながり(質問とそれに対する返答、等)
- 遡及的依存関係 Retrospective Dependence……………【Retrospective】
 - 二つ目の談話機能が一つ目の談話機能に対する任意の反応を担うタイプのつながり(情報提供とそれに対するフィードバック発話、等)
- 外部予測的依存関係 External Prospective Dependence……………【P-External】
 - 発話以外でなされた一つ目の振る舞いが特定の談話機能を要求するタイプのつながり(物の差し出しに発話で応じる、等)
- 外部遡及的依存関係 External Retrospective Dependence……………【R-External】
 - 二つ目の談話機能が一つ目の発話以外で生じた出来事に対する任意の反応を担うタイプのつながり(周囲で生じた物音に対してフィードバック発話を行う、等)

変更履歴

- Version 1.0 2022年3月31日