

日常会話コーパスのための談話行為タグの設計

居關 友里子* 第十 早織†‡ 伝 康晴§* 小磯 花絵*

* 国立国語研究所 † アルバータ大学 ‡ 東京外国語大学 § 千葉大学

iseki@ninjal.ac.jp, saori.daiju@gmail.com, den@chiba-u.jp, koiso@ninjal.ac.jp

1 はじめに

現在、国立国語研究所では日常場面で生じた多様な会話を扱ったコーパス「日本語日常会話コーパス」(以下 CEJC) の構築が進められており(小磯他 2016; Koiso et al. 2016a, 2016b)、会話場面を収録、文字化したものに対し、研究利用に向けた基本的なアノテーションを施したものの公開が計画されている*1。このアノテーションの一つとして、産出された各発話がどのような行為を担うものなのか、発話同士はどのような関係を結ぶものなのかに関する情報「談話行為タグ」を付与する予定である。本研究では、現在試行中である、アノテーション仕様策定の試みについて報告する。

以下では、まず談話行為のタグ付与に関する先行研究について触れたうえで(2節)、本研究が実験的に試行した談話行為タグ付与の作業とその結果について考察を行う(3節)。これを踏まえた枠組みの修正案を提示し(4節)、今後のアノテーション作業に向けたまとめを述べる(5節)。

2 研究の背景

対話中に産出された発話に対し談話行為の情報を付与する試みは多くある(Carletta et al. 1996; Alexandersson et al. 1997; Allen and Core 1997; Jurafsky et al. 1997; Dhillon et al. 2004; Bunt 2009)。日本においては、荒木 他(1999) やそれを基礎に発展させた人工知能学会「談話・対話研究におけるコーパス利用」研究グループ(2000) が課題指向対話をもとに設定した付与基準を提案している。近年、国際標準化規格 ISO24617-2 が提案され(Bunt et al. 2012; ISO 24617-2 2012)、これを日本語会話に応用する試みもなされている(平岡他 2013)。

これら各談話行為タグの基本的特徴や付与結果などについては、伝(2015)でまとめられている。

付与基準の有効性は、付与対象となる会話データの性質に依存すると考えられる。CEJC は日常生活の中で生じるあらゆる会話場면을対象範囲としており、課題指向対話だけでなく、雑談のようなやり取りを交わすこと自体が目的とされる会話や、作業や食事をしながらのやり取りといったような会話以外の活動が主要に行われる場面も収められる予定である。このような多様な会話への適用を視野に、本研究では既存の枠組みを援用した談話行為アノテーションを試行し、これを修正するという方法で新たな枠組みの設計を検討する。

3 アノテーションの試行

3.1 枠組み

基礎となる枠組みとして ISO24617-2 を用いた。ISO24617-2 は、特定の課題達成に向け活動を進行させる行為に加え、発話順番の取得や談話を構造化するといった、相互行為の調整に関する機能を扱うタグが用意されている点が特徴的である。課題指向型ではない会話における談話行為をアノテーションする際には、このようなタグが有効である可能性が見込まれる。

3.2 方法

■対象データ 対象としたのは「千葉大学3人会話コーパス」*2 (Den and Enomoto 2007) である。友人同士3人1組でなされた、指定されたトピックやそれに関連すること、あるいは関連しないことについてのやり取り10分が12組分、計120分のデータからなる。参加者は用意された時間で自由に会話することが求められており、「雑談」と特徴づけ得るやり取りが観察される。発話は「長い発話単位」(Japanese Discourse Research Initiative 2014) と呼ばれる単位に分割されている。

*1 国立国語研究所共同研究プロジェクト「大規模日常会話コーパスに基づく話し言葉の多角的研究」(プロジェクトリーダー: 小磯花絵) <http://pj.ninjal.ac.jp/conversation/>

*2 <http://research.nii.ac.jp/src/Chiba3Party.html>

表1 次元・談話機能

次元	談話機能
タスク	質問, 情報提供, 申し出, 依頼, 依頼への対処, 提案, 提案への対処, 情報提供 + 質問
自己フィードバック	自己肯定, 自己否定
他者フィードバック	他者肯定, 他者否定, 他者フィードバック誘出
発話順番管理	発話順番取得, 発話順番維持, 発話順番奪取, 発話順番指定, 発話順番開放, 自己肯定 + 発話順番取得
時間管理	滞り, 一時停止
談話構造化	相互行為構造化, 対話開始
自己伝達管理	自己中断, 独り言
他者伝達管理	補完, 訂正
社会的付き合い管理	謝罪, 感謝, 出会いの挨拶, 別れの挨拶, 自己紹介

■**作業者** 2名の作業者(第1・第2著者)が独立に作業を行った。作業期間中に数回の話し合いを行っているが、この内容を受けて既に付与したアノテーションを修正することは行わなかった。

■**作業手続き** 作業には Anvil^{*3}を使用した。一発話単位につき一つずつ、「次元」「談話機能」の各々の集合から最も適当なタグを選択する(表1)^{*4}。各タグの使用例については伝(2015)を参照されたい。

基本的には、同じ行に示された次元と談話機能を組み合わせる。例えば、談話機能「自己肯定」を選択する場合、次元は「自己フィードバック」となる。ただし、次元「タスク」の右枠に挙げた談話機能についてのみ、掲載した全ての次元と組み合わせられる。

ISO24617-2の場合、各タグが適用される発話部分(範囲)を各々指定可能であり、かつ同じ発話箇所にも複数のタグを重複して適用することができるという特徴がある。しかし本研究では、作業の複雑化を避けるため、一つの発話単位に対し、次元と談話機能(表1)を一つずつ選択する方法をとった。複数のタグが付与可能な場合は、最も中心的な機能を担うものを選択することとした。選択可能なタグはISO24617-2を基に設定し、作業を進行させる中で追加や削除、定義の微調整を行った。

^{*3} <http://www.anvil-software.org>

^{*4} 加えて、当該発話が他の発話とどのような関係を結んでいるのか(依存関係)に関する情報についてもアノテーションを行っているが、考察は他の機会に譲る。

表2 談話機能タグの一致率

会話ID	素一致率	κ 値
chiba0132	90.59%	.84
chiba0232	74.28%	.65
chiba0332	76.29%	.65
chiba0432	77.12%	.68
chiba0532	78.46%	.68
chiba0632	78.87%	.70
chiba0732	70.78%	.63
chiba0832	79.69%	.72
chiba0932	81.69%	.73
chiba1032	70.56%	.62
chiba1132	79.82%	.69
chiba1232	75.31%	.68

表3 談話機能タグの使用率

談話機能(次元)	使用率	
	作業者A	作業者B
情報提供(タスク、その他)	38.41%	40.05%
自己肯定(フィードバック)	36.72%	29.86%
質問(タスク、その他)	7.43%	8.88%
発話順番取得(発話順番管理)	4.36%	4.99%
他者肯定(フィードバック)	3.54%	2.84%
自己中断(自己伝達管理)	2.91%	3.92%
独り言(自己伝達管理)	1.65%	2.05%
滞り(時間管理)	1.44%	1.70%
発話順番維持(発話順番管理)	0.72%	1.17%

3.3 結果

「次元」タグについては概ね「談話機能」と連動した付与がなされるため、本節では「談話機能」タグの付与結果を取り上げる。

■**信頼性** 全12データに対して付与した談話機能タグの作業者間の一致率を表2に示す。 κ 値の平均は $\kappa = .69$ ($sd = .06$, $min = .62$, $max = .84$)であった。

■**タグの使用率** 作業者ごとの談話機能タグの使用率(全12データに対する総数)を表3に示す(1%以上のもののみ掲載)。最も多く使用されているのは「情報提供」(38.41%/40.05%)であった。次いで、あいづちや情報提供に対する評価などが含まれる「自己肯定」(36.72%/29.86%)が多く出現していた。

3.4 考察

■**タグ使用率の偏り** タグの使用率を見ると、情報提供と自己肯定が最も多く使用されていた。自己肯定の付与対象の一部には情報提供と類似の機能が含まれ、この二種を合計すると全体のほぼ7割を占めることになる。このような情報提供類に偏るといふ指摘は先行

研究でも複数なされており (荒木 他 1999; 徳久, 寺脇 2007; Fang et al. 2012; 平岡他 2013)、本研究もそれに沿う結果となった。

また、情報提供、自己肯定に偏る他の理由として、一発話単位につき最も中心的な機能を担う談話機能タグを一つ選択するという作業手続きの影響が考えられる。選択可能な談話機能タグには、情報提供や質問など実質的行為を担うもの、自己肯定などそれに反応を返す行為を担うものがある。一方で、それらの行為を行うための調整機能である発話順番開放や一時停止などのタグも同列の選択肢として用意されている。実質的行為と相互行為の調整機能は、しばしばその両者が一発話単位内に生じていた。しかし、前者の方が発話同士の連なりの理解のためにより欠かすことができない情報であるため、中心的機能として捉えやすく、その結果、後者のタグがあまり使用されなかったと予想される。

■**タグ付与の難易** 作業者間の平均一致率は $\kappa = .69$ と比較的高かった。以下では一致率を上下させる要因であるタグ付与の難易に関して、12 データのうち他と比べ著しく一致率の高かったデータ (chiba0132) を詳しく考察する。このデータは付与タグの判断が容易であったという作業者間の共通理解があった。データの特徴として、情報提供の割合が 54.9% と高く (全データの平均 39.2%)、分析対象とした発話数は 255 発話と少ない (全データの平均 499 発話)。発話数の少なさから一発話が長かったことが示唆され、安定した談話機能の判断が可能となったと予想される。加えて、情報提供が多かったことから、会話の進行において「語り型」が他のデータよりも多く生じていた可能性がある。語り型は、同一の参加者による情報提供と、他の参加者からのフィードバックという組み合わせが一定時間繰り返されるため、「話者交替型」と比べると行為連鎖が捉えやすいと考えられる。

全データに対する作業を通して判断が困難であったタグとしては、情報提供と自己肯定の区別、自己肯定と他者肯定の区別といった、連鎖上の位置の判断によって選択されるタグが変わる例が挙げられる。また、情報提供や自己肯定と質問、自己肯定と独り言とといったような、行為判断が発話の産出のされ方に強く影響を受けるものについても、いずれのタグを選択するかに関して、判断の揺れが生じやすかった。これらのタグに関しては、同じ発話に対して作業者間で異なるものを選択していることが多かった (表 4)。

表 4 作業者間で不一致が多い談話機能

作業者 B	作業者 A				
	情報提供	自己肯定	他者肯定	質問	独り言
情報提供	2061	228	37	30	2
自己肯定	30	1621	57	7	12
他者肯定	13	53	94	0	1
質問	76	67	1	368	1
独り言	23	38	1	0	34

4 修正案

3 節のアノテーションの試行を受けて、以下のように枠組みを修正した。

■**手続き** 各発話単位につき一つ以上最大二つの談話機能タグを選択する (ただし、タスク系から二つあるいはフィードバック系から二つを選択することはできない)。

■**タグセット** 選択可能なタグはタスク系、フィードバック系、相互行為管理系の三種に大別する (表 5)。

■**主な変更点** 主な変更点を以下にまとめる。

1. 一発話単位に対し付与できるタグを最大二つまでに増やす: このことによりタスク、フィードバック以外の機能 (新しいタグセットにおける「相互行為管理系」) が反映されやすくなる。
2. 情報提供に関するタグを細分化: 付与対象範囲が広がった情報提供について、「情報提供」「返答」「同意」「不同意」に細分化し、性質の差を反映できるようにする。
3. タグの統合・削除: 区別の困難なもの (自己肯定と他者肯定など)、より汎用的に使える別のタグがあるもの (自己紹介 (< 情報提供) など) などについてタグを整理し、タグ判断の複雑さを下げる。
4. 一つの機能を一つのタグ選択で示す: 次元と談話

表 5 修正案のタグセット

次元グループ	談話機能
タスク系	質問, 情報提供, 返答, 同意, 不同意, 依頼, 依頼への対処, 提案, 提案への対処, 申し出
フィードバック系	肯定, 否定, 質問, 返答
相互行為管理系	発話順番取得, 発話順番維持, 発話順番奪取, 発話順番指定, 発話順番開放, 滞り, 相互行為構造化, 自己中断, 独り言, 補完, 他者訂正

機能の組み合わせは概ね限られていたため、それらの組み合わせを予め組み込んだタグセットを使用し作業手続きを簡略化する。

5 おわりに

本研究では日常会話コーパスに対する談話行為アノテーションの枠組みについて検討した。今後は修正案を用い、他の雑談場面や制度的場面、参加者が主として会話以外の活動に従事している場面など、性格の異なるデータを扱いながら、さらに枠組みの修正を行っていく予定である。付与タグの判断に生じるいくつかの困難については、タグの定義の厳密化や作業者のトレーニングによって解消可能なものであると考えられる。一方で、変則的なタイミングでの発話、産出のされ方が微妙な発話など、そもそも会話に参加している当事者たちすらも、談話機能を特定し兼ねる発話が一定数生じているはずであり、これらの扱いについては検討の必要がある。また、今回報告できなかった、発話間の関係性に関するアノテーションについても、談話機能タグと併せた検討を今後行っていきたい。

■謝辞 本研究は国立国語研究所の共同研究プロジェクト「大規模日常会話コーパスに基づく話し言葉の多角的研究」の研究成果を報告したものである。

参考文献

Alexandersson, J., Buschbeck-Wolf, B., Fujinami, T., Maier, E., Schmitz, N. R. B., and Siegel, M. (1997). "Dialogue acts in VERBMOBIL-2." *Verbmobil report 226*, DFKI Saarbrücken.

Allen, J. and Core, M. (1997). *Draft of DAMSL: Dialogue act markup in several layers*. Department of Computer Science, University of Rochester.

荒木雅弘, 伊藤敏彦, 熊谷智子, 石崎雅人 (1999). 発話単位タグ標準化案の作成. *人工知能学会誌*, **14** (2), pp. 251–260.

Bunt, H. (2009). "The DIT++ taxonomy for functional dialogue markup." In *Proceedings of the AAMAS 2009 Workshop Towards a standard markup language for embodied dialogue acts (EDAML 2009)*, pp. 13–23, Budapest.

Bunt, H., Alexandersson, J., Choe, J.-W., Fang, A. C., Hasida, K., Petukhova, V., Popescu-Belis, A., and Traum, D. (2012). "ISO 24617-2: A semantically-based standard for dialogue act annotation." In *Proceedings of the 8th International Conference on Language Resources and Evaluation (LREC 2012)*, pp. 430–437, Istanbul, Turkey.

Carletta, J., Isard, A., Isard, S., Kowtko, J., Doherty-Sneddon, G., and Anderson, A. (1996). "HCRC dialogue structure

coding manual." Tech. rep. HCRC/TR-82, Human Communication Research Centre, University of Edinburgh.

伝康晴 (2015). 対話への情報付与. 小磯花絵 (編), 講座日本語コーパス 3: 話し言葉コーパス—設計と構築—, pp. 101–130. 朝倉書店.

Den, Y. and Enomoto, M. (2007). "A scientific approach to conversational informatics: Description, analysis, and modeling of human conversation." In Nishida, T. (Ed.), *Conversational Informatics: An Engineering Approach*, pp. 307–330. John Wiley & Sons, Hoboken, NJ.

Dhillon, R., Bhagat, S., Carvey, H., and Shriberg, E. (2004). "Meeting recorder project: Dialogue act labeling guide." Tech. rep. TR-04-002, International Computer Science Institute.

Fang, A. C., Cao, J., Bunt, H., and Liu, X. (2012). "The annotation of the Switchboard corpus with the new ISO standard for dialogue act analysis." In *Proceedings of the 8th Joint ISO-ACL SIGSEM Workshop on Interoperable Semantic Annotation*, pp. 13–18, Pisa, Italy.

平岡拓也, G. Neubig, S. Sakti, 戸田智基, 中村哲 (2013). 説得対話コーパスの構築と分析. *情報処理学会研究報告*, **2013-SLP-99**, pp. 41–46.

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

Japanese Discourse Research Initiative (2014). 発話単位ラベリングマニュアル version 2.0. <http://www.jdri.org/resources/manuals/uu-doc-2.0.pdf>.

人工知能学会「談話・対話研究におけるコーパス利用」研究グループ (2000). 様々な応用研究に向けた談話タグ付き音声対話コーパス. *人工知能学会研究会資料*, **SIG-SLUD-9903**, pp. 19–24.

Jurafsky, D., Shriberg, E., and Biasca, D. (1997). *Switchboard SWBD-DAMSL shallow-discourse-function annotation coders manual*. Institute of Cognitive Science, University of Colorado.

Koiso, H., Tanaka, Y., Watanabe, R., and Den, Y. (2016a). "A large-scale corpus of everyday Japanese conversation: On methodology for recording naturally occurring conversations." In *Proceedings of LREC 2016 Workshop on Casual Talk among Humans and Machines*, pp. 9–12, Portoroz, Slovenia.

Koiso, H., Tsuchiya, T., Watanabe, R., Yokomori, D., Aizawa, M., and Den, Y. (2016b). "Survey of conversational behavior: Towards the design of a balanced corpus of everyday Japanese conversation." In *Proceedings of the 10th International Conference on Language Resources and Evaluation (LREC 2016)*, pp. 4434–4439, Portoroz, Slovenia.

小磯花絵, 土屋智行, 渡部涼子, 横森大輔, 相澤正夫, 伝康晴 (2016). 均衡会話コーパス設計のための一日の会話行動に関する基礎調査. *国立国語研究所論集*, **10**, pp. 85–106.

徳久良子, 寺嶋立太 (2007). 非課題遂行対話における発話の特徴とその分析. *人工知能学会論文誌*, **22** (4), pp. 425–435.