

人はなぜ笑うのか？ 対話における笑いの根拠ラベルの半自動構築

井上 昂治 Mikey Elmers Divesh Lala 河原 達也

京都大学 大学院情報学研究科

inoue.koji.3x@kyoto-u.ac.jp

概要

笑いは人間どうしの対話において多面的なシグナルとして機能するが、これを適切に認識・生成することは、自然な対話の実現を目指す対話システムにとって大きな課題である。本研究では、日本語対話データに対して「笑うことができる文脈 (laughable context)」をアノテーションし、さらにその判断の根拠を分類するためのラベルを、大規模言語モデル (LLM) を活用して半自動的に構築した。まず複数のアノテーターにより、対話データの各発話に対して、対話相手が笑うことができるか否かの二値を判定してもらった。次に、LLM を用いてこの二値の判断の根拠を説明する文を生成し、さらにこれらを分類する根拠ラベルを生成した。その結果、「共感と親近感」「ユーモアと意外性」「リラックスした雰囲気」などを含む 10 種類のラベルが生成された。また、LLM に上記の二値判断を認識させたところ、F1 スコアで 43.1% となり、LLM が自然な対話における笑いを認識することの課題と可能性を明らかにした。

1 はじめに

人間どうしの対話において、笑いはユーモア、共感、驚き、あるいは社会的結束の現れなどのコミュニケーションのシグナルとして機能する [1, 2, 3]。しかし、そのメカニズムは複雑かつ多面的であり、人間らしい応答を目指す対話システムにとって、これを解明することは長期的な課題である [4, 5, 6, 7, 8]。また、笑いやユーモアのモデル化に関する従来のアプローチは、明示的な音声や視覚の刺激を用いた場面に限定されることが多く、自然な対話における機微な文脈を対象としたものは少ない [9, 10, 11, 12, 13]。したがって、自然な対話のデータを対象として、笑いが生じる根拠を明らかにする

表 1: 笑いのアノテーションの例

	発話	笑えるか
B:	素晴らしい心掛けだと思います。わたしはいつも、自分の話ばっかしちゃうので、見習わせてもらいます。	NO
A:	そうなんですねー。でも旦那さんはその話を聞いてくれるんですか？	NO
B:	うんうん、と聞いてくれるんです。負担掛けてないかしら？	NO
A:	かけてないと思いますよー。お優しいですね。わたしの旦那は話を聞いてなさげです。あれ、おかしいな。	YES

ことは、大規模言語モデル (LLM) による対話をより自然かつ共感的にすることに貢献するといえる。しかし、笑いの根拠を何らかの形式でアノテーションすることは時間的・人的コストが大きく、会話分析による質的アプローチに留まってきたのが現状である。

そこで本研究では、この「人はなぜ笑うのか」という問い合わせに対して、情報学の立場から笑いの根拠についての半自動的なアノテーション・ラベル構築を試みる。まず対象箇所を特定するために、対話データの各発話に対して「笑うことができるか否か」という簡単な二値判定を複数のアノテーターに行ってもらった。続いて、その多数決ラベルで「笑える」と判定された文脈について、その判断の根拠を LLM (GPT-4o) を用いて生成し、それらを分類する根拠ラベルも生成した。このような半自動的にラベルを生成するアプローチは一般化することができ、人手によるアノテーションが簡単なラベルに限定される場面（例えば、感情ラベル）において有効である。

2 笑いのアノテーション

対話データセットとして、RealPersonaChat [14] を使用した。このテキスト対話データは、特に指定されたペルソナが与えられるのではなく、参加者自身の設定として、一対一の日本語の対話をを行うもの

表2: 「笑える」と判断したアノテータの割合ごとのサンプル数

割合	サンプル数
1.0 (5/5)	163 (0.64%)
0.8 (4/5)	845 (3.34%)
0.6 (3/5)	2731 (10.80%)
0.4 (2/5)	8143 (32.20%)
0.2 (1/5)	11928 (47.17%)
0.0 (0/5)	1479 (5.85%)

である。1回の対話には約30発話が含まれ、全体で14,000対話が収録されている。今回は、このうち900対話のデータにアノテーションを付与し、残りについては今後の研究でアノテーションを進める予定である。

このデータセットの各発話に対して、その次に対話相手が「笑うことができる文脈 (laughable context)」か否かについて二値の判断を行ってもらった。アノテータは5名で、全員がすべての対話の各発話に対して二値のラベルを付与した。ただし、各対話の最初の2発話は挨拶に相当するためアノテーション対象から除外した。

表2にアノテータ間での一致度の分布を示す。一部のサンプルでは、全員が一致して「笑える」または「笑えない」と判断したが、多くのサンプルでは意見が分かれる結果となり、このタスクの主観性や複雑さを浮き彫りにしている。多数決を適用した場合、3,739発話 (14.8%) が「笑える」とラベル付けされ、21,550発話 (85.2%) が「笑えない」とラベル付けされた。

多数決で「笑える」とラベル付けされた文脈の例を表1に示す。この対話では、話者Aの最後の発話が自己矛盾を含んでおり、対話の文脈を高度に理解する必要がある。このようなアノテーション結果は、笑いの判断において文化的な背景や会話の流れを理解することが重要であることを示している。

3 笑いの根拠ラベルの生成

アノテータがなぜ「笑える」と判断したかを明らかにするために、その文脈の背後にある理由(根拠)についてLLMを活用して半自動的に調査した。本節では、人間アノテーターの過半数が「笑える」と判断したサンプル(3,739サンプル)のみを用いて、その対話の文脈と多数決ラベルの結果(すべて「笑える」)をGPT-4oにプロンプトとして入力し、その根拠を生成した。例えば、表1の文脈に対して生成された根拠は次の通りである。

この会話の最後のAさんの発言に対して、Bさんが笑うと第三者が判断した理由を推測すると、以下のような点が考えられます。

ユーモアの要素: Aさんの「わたしの旦那は話を聞いてなさげです。あれ、おかしいな。」という発言には、少し自嘲的なユーモアが含まれています。自分の状況を軽く笑い飛ばすようなトーンが、聞き手にとって面白く感じられることがあります。

共感と親しみ: Aさんの発言は、日常的な夫婦間のコミュニケーションのズレを軽く表現しており、多くの人が共感できる内容です。共感できる話題は、笑いを誘いややすいです。

(中略)

リラックスした雰囲気: 会話全体がリラックスした雰囲気で進んでおり、Aさんの発言もその流れに沿った軽いジョークのように受け取られた可能性があります。

これらの要素が組み合わさり、Bさんが笑うと第三者が判断したのかもしれません。

続いて、LLMを用いた分類ラベル生成手法[15]を参考に、「笑える」文脈に対して生成された根拠の一覧データから、それらを分類するためのラベルの自動生成を試みた。まず、生成された根拠データをランダムに分割し、それぞれがデータの5%になるようにした。このうちの1つに対して、GPT-4oにその分類ラベルの一覧候補と各ラベルの説明文を生成させ、必要に応じて手動で修正した。その後、この分類情報と、次の分割データに用いて、再度同様のデータを生成することで、分類ラベルおよび説明文を更新した。この処理をすべての部分データに對して続けた。

その結果、表3に示す10個の笑いの根拠ラベルが得られた。これらには、「(1) 共感と親近感」や「(2) ユーモアと意外性」などがラベルとして含まれており、直感的も妥当であるといえる。

根拠ラベルを生成した後に、再びLLMを使用して、各データ(文脈および根拠)にこの根拠ラベルを割り当てた。その際、1つのデータに複数のラベルを割り当てる許可を許可した。表3の右側にラベリングの結果を示す。「(1) 共感と親近感」「(2) ユーモアと意外性」「(3) リラックスした雰囲気」のように多くのデータに割り当たされた典型的なラベルが観察された一方で、その他のラベルも一定数割り当てられていることがわかる。このような分布は、生

表3: 生成された笑いの根拠ラベルおよびそれらの説明の一覧（右側に各ラベルに割り当てられたサンプル数と各ラベルに関する関連文献を示す）

ラベル名	説明	サンプル数	関連文献
(1) 共感と親近感	会話の中で、共通の経験や感情を共有することで、親近感が生まれ、笑いが誘発される状況です。例えば、共通の趣味や日常の「あるある」な状況に対する共感がこれに該当します。	3013 (80.6%)	[16, 17]
(2) ユーモアと意外性	発言に含まれるユーモアや意外性が笑いを引き起こす要因となる場合です。予想外の展開や、言葉遊び、誇張表現などがこれに含まれます。	3233 (86.5%)	[18, 19]
(3) リラックスした雰囲気	会話全体が和やかでリラックスした雰囲気で進んでいるため、自然と笑いが生まれる状況です。軽妙なやり取りや、冗談を交えた会話がこれに該当します。	2955 (79.0%)	[20]
(4) 自己開示と親しみやすさ	自分の失敗談や個人的なエピソードを共有することで、相手に親しみやすさを感じさせ、笑いを誘う状況です。自己開示が相手に安心感を与えることも含まれます。	475 (12.7%)	[21]
(5) 文化的背景と共通認識	特定の文化的背景や共通認識に基づく笑いです。例えば、特定の地域や文化に関連するジョークや、一般的な迷信や言い伝えに基づく発言がこれに該当します。	176 (4.7%)	[22, 23]
(6) ノスタルジーと懐かしさ	過去の思い出や懐かしい話題が笑いを誘う要因となる場合です。共通の過去の経験や、懐かしい出来事に対する共感がこれに該当します。	204 (5.5%)	[24]
(7) 自己卑下のユーモア	自分の欠点や失敗を笑いに変えることで、相手に親しみやすさを感じさせ、笑いを誘う状況です。自分を少し卑下することで、相手に安心感を与えることも含まれます。	404 (10.8%)	[25]
(8) 期待の裏切り	会話の流れや相手の期待を意図的に裏切ることで、意外性を生み出し、笑いを誘う状況です。予想外の返答や展開がこれに該当します。	323 (8.6%)	[26, 27]
(9) ポジティブなエネルギー	会話の中でポジティブな感情やエネルギーが伝わり、相手を笑顔にする状況です。楽しい話題や前向きなコメントがこれに該当します。	338 (9.0%)	[28]
(10) 誇張表現	物事を大げさに表現することで、コミカルな印象を与え、笑いを誘う状況です。誇張された表現が、聞き手にとって面白く感じられることがあります。	478 (12.8%)	[29]

成された根拠ラベルの妥当性を支持しているともいえる。

また、ラベル間での相関係数を図1に示す。例えば、「(1) 共感と親近感」と「(3) リラックスした雰囲気」は弱い正の相関を示している。逆に、「(1) 共感と親近感」は「(10) 誇張表現」と負の相関を示しており、直感的に理解できる。また、「(2) ユーモアと意外性」と「(9) ポジティブなエネルギー」は、どちらもポジティブな感情と関連しているにもかかわらず、負の相関が見られる。これは、前者が暗黙的であり、後者が明示的であるという、異なる表現スタイルのためであると考えられる。

最後に表3の右側に、各根拠ラベルに関する会話分析などの文献を示す。いずれの根拠ラベルにおいても、これまでに何らかの分析や言及がなされており、この結果もまた生成されたラベルの妥当性を支持するものともいえる。

4 LLMによる笑いの文脈の認識

アノテーションしたデータを用いて、LLMがどの程度「笑える」文脈を認識できるのかについても検証を行った。プロンプトでは、ゼロショットでタ

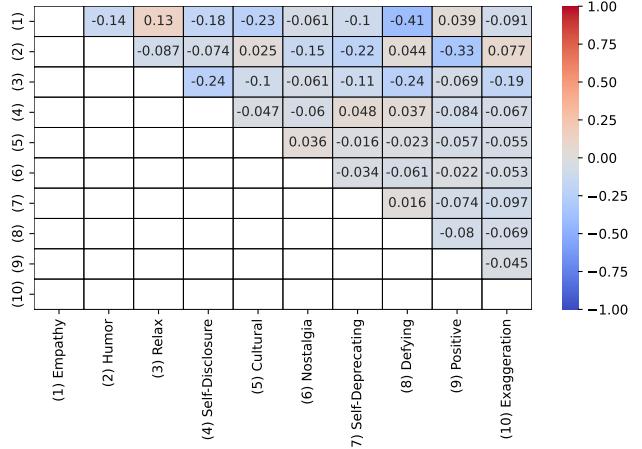


図1: 各文脈に対する割当てにおける笑いの根拠ラベル間の相関係数

スクの説明とChain-of-Thought (CoT) の指示を与えた。CoTでは、まず判断理由を出力したのちに、「笑える」か否かを出力するようにした。多数決ラベルに対するGPT-4oの性能を評価したところ、F1スコアで43.1%、適合率は41.7%、再現率は44.7%であった。このスコアはチャンスレベル(14.8%)は大幅に上回っていたものの、対話における自然な笑いの微妙なニュアンスを捉えることは依然として困難で

表4: 各根拠ラベルにおける GPT-4o の笑いの文脈認識結果

ラベル名	GPT-4o の出力サンプル数	
	笑える(正解)	笑えない(不正解)
(1) 共感と親近感	1226 (40.7%)	1787 (59.3%)
(2) ユーモアと意外性	1571 (48.6%)	1662 (51.4%)
(3) リラックスした雰囲気	1257 (42.5%)	1698 (47.5%)
(4) 自己開示と親しみやすさ	232 (48.8%)	243 (51.2%)
(5) 文化的背景と共通認識	102 (57.9%)	74 (42.1%)
(6) ノスタルジーと懐かしさ	62 (30.4%)	142 (69.6%)
(7) 自己卑下のユーモア	255 (63.1%)	149 (36.9%)
(8) 期待の裏切り	227 (70.3%)	96 (29.7%)
(9) ポジティブなエネルギー	50 (14.8%)	288 (85.2%)
(10) 誇張表現	239 (50.0%)	239 (50.0%)

あることを示唆している。

次に、前節で生成した各分類ラベルにおける LLM の認識性能を個別に調査した。表 4 は、GPT-4o による二値出力の分布と、各ラベルに割り当てられたデータに対する正解率を示している。まず、主要なラベルである (1) から (3) は、40%から 50%の範囲で同様の正解率を示した。さらに、「(5) 文化的背景と共通認識」、「(7) 自己卑下のユーモア」、「(8) 期待の裏切り」では比較的高いスコアが観察され、現在の LLM がこれらの文脈を効果的に捉えられる可能性があることが示唆された。対照的に「(6) スタルジーと懐かしさ」や「(9) ポジティブなエネルギー」のようなカテゴリーでは正解率が低く、LLM の理解に限界がある可能性が示された。

表 5 は、LLM が A の最後の発話に対して「笑えない」と判断したが、多数決ラベルは「(6) スタルジーと懐かしさ」の理由で「笑える」であった例である。なお、この文脈には、「(2) ユーモアと意外性」と「(3) リラックスした雰囲気」のラベルも割り当てられていた。この例では、参加者たちは粉末飲料を牛乳で溶かして飲んだ懐かしい思い出について話している。最後の発話はその郷愁を呼び起こし、暗黙のうちに笑いを誘っている。ここでは、話者 A の感情を捉えることは、現在の LLM にとって難しいようだが、適切な笑いの反応には不可欠である。

表 6 は「(9) ポジティブなエネルギー」のラベルの例を示している。この文脈には「(1) 共感と親近感」と「(2) ユーモアと意外性」のラベルも割り当てられていた。参加者たちは育児における困難な経験について話しているが、最後の発話では、話者 A は時間が経った後、その経験を肯定的に振り返っている。物語自体は困難な時期を語っているが、今では

表5: 「ノスタルジーと懐かしさ」を根拠とする例

発話
A: カルシウムという意味では牛乳とかヨーグルトとかも摂取してますか？
B: 牛乳はミロを入れて飲んでます。ヨーグルトはおやつに食べてます。
A: めちゃくちゃバランス良いですね！
B: そうですね、健康大事なので。
A: ミロって久々に聞きました。

表6: 「ポジティブなエネルギー」を根拠とする例

発話
B: うちの長女も JK ですよ。たらこスパいいですねー！娘も夕飯の支度手伝ってくれますよ。
A: おー大きくなってくるとそういうのが助かりますよね。
B: ほんとほんと！嬉しいですよねー。心強い限りです。
A: 小さい頃はネバーエンディングストーリーばかりに大変だったりしますもんね。
B: ははは、確かに。笑える思い出話ばっかりですよ。
A: 通り過ぎると笑い話だったり、やっててよかったと思えるようになるんですよね。

肯定的に捉えられているため、笑えるものとなっている。この例は、LLM が文脈を「笑える」と正確に解釈するためには、物語の時間的構造と話者の現在の感情を理解する必要があることを示唆している。

5 結論

本研究では、テキスト対話データにおいて、対話相手の立場で「笑える」と判断できる文脈にアノテーションを付与し、この判断の根拠を分類するためのラベルを LLM により半自動的に生成した。その結果、10 個の異なるラベルが生成され、笑いの多面的な性質を理解するための切り口が示された。さらに、GPT-4o がこれらの「笑える」文脈を認識する能力を評価したところ、その精度はチャンスレベルを上回ったものの、会話における笑いの機微な文脈を正確に捉えることの難しさも浮き彫りとなった。

本研究で用いた半自動ラベル生成の手法は、他の類似のタスクにも適用可能と考えられる。具体的には、人間のアノテーターからは簡単な判断のラベル（例えば二値ラベル）のみが得られるものの、より詳細な説明や分類が必要とされる場面への適用が想定される。今後は、データセットの拡張、生成されたラベルの言語学的側面からのさらなる検証、笑いの認識モデルの改良などに取り組む予定である。

謝辞

本研究は、JST さきがけ (JPMJPR24I4) ならびに JSPS 科研費 (JP23K16901) の支援を受けた。

参考文献

- [1] Neal R Norrick. *Conversational joking: Humor in everyday talk*, 1993.
- [2] Phillip Glenn. *Laughter in interaction*. Cambridge University Press, 2003.
- [3] Salvatore Attardo. *Linguistic theories of humor*. Walter de Gruyter, 2009.
- [4] Ye Tian, Chiara Mazzocconi, and Jonathan Ginzburg. When do we laugh? In **Annual Meeting of the Special Interest Group on Discourse and Dialogue (SIGDIAL)**, pp. 360–369, 2016.
- [5] Bekir Berker Türker, Zana Buçinca, Engin Erzin, Yücel Yemez, and T Metin Sezgin. Analysis of engagement and user experience with a laughter responsive social robot. In **INTERSPEECH**, pp. 844–848, 2017.
- [6] Koji Inoue, Divesh Lala, and Tatsuya Kawahara. Can a robot laugh with you?: Shared laughter generation for empathetic spoken dialogue. **Frontiers in Robotics and AI**, Vol. 9, , 2022.
- [7] Bogdan Ludusan and Petra Wagner. The effect of conversation type on entrainment: Evidence from laughter. In **Annual Meeting of the Special Interest Group on Discourse and Dialogue (SIGDIAL)**, pp. 168–174, 2023.
- [8] Nynaeve Perkins Booker, Michelle Cohn, and Georgia Zellou. Linguistic patterning of laughter in human-socialbot interactions. **Frontiers in Communication**, Vol. 9, , 2024.
- [9] Dario Bertero and Pascale Fung. A long short-term memory framework for predicting humor in dialogues. In **Conference of the North American Chapter of the Association for Computational Linguistics: Human Language Technologies (NAACL-HLT)**, pp. 130–135, 2016.
- [10] Akshat Choube and Mohammad Soleymani. Punchline detection using context-aware hierarchical multimodal fusion. In **International Conference on Multimodal Interaction (ICMI)**, pp. 675–679, 2020.
- [11] Sophie Jentzsch and Kristian Kersting. ChatGPT is fun, but it is not funny! humor is still challenging large language models. In **Workshop on Computational Approaches to Subjectivity, Sentiment, & Social Media Analysis (WASSA)**, pp. 325–340, 2023.
- [12] Dayoon Ko, Sangho Lee, and Gunhee Kim. Can language models laugh at YouTube short-form videos? In **Conference on Empirical Methods in Natural Language Processing (EMNLP)**, pp. 2897–2916, 2023.
- [13] Jack Hessel, Ana Marasovic, Jena D. Hwang, Lillian Lee, Jeff Da, Rowan Zellers, Robert Mankoff, and Yejin Choi. Do androids laugh at electric sheep? humor “understanding” benchmarks from the new yorker caption contest. In **Annual Meeting of the Association for Computational Linguistics (ACL)**, pp. 688–714, 2023.
- [14] Sanae Yamashita, Koji Inoue, Ao Guo, Shota Mochizuki, Tatsuya Kawahara, and Ryuichiro Higashinaka. Realpersonachat: A realistic persona chat corpus with interlocutors’ own personalities. In **Pacific Asia Conference on Language, Information and Computation (PACLIC)**, pp. 852–861, 2023.
- [15] Mengting Wan, Tara Safavi, Sujay Kumar Jauhar, Yujin Kim, Scott Counts, Jennifer Neville, Siddharth Suri, Chirag Shah, Ryen W White, Longqi Yang, Reid Andersen, Georg Buscher, Dhruv Joshi, and Nagu Rangan. TnT-LLM: Text mining at scale with large language models. In **SIGKDD Conference on Knowledge Discovery and Data Mining (KDD)**, pp. 5836–5847, 2024.
- [16] Jennifer Hay. *The pragmatics of humor support*. Walter de Gruyter, 2001.
- [17] Dana Garbarski, Nora Cate Schaeffer, and Jennifer Dykema. Interviewing practices, conversational practices, and rapport: Responsiveness and engagement in the standardized survey interview. **Sociological methodology**, Vol. 46, No. 1, pp. 1–38, 2016.
- [18] Marta Dynel. Beyond a joke: Types of conversational humour. **Language and linguistics compass**, Vol. 3, No. 5, pp. 1284–1299, 2009.
- [19] Rod A Martin and Thomas Ford. *The psychology of humor: An integrative approach*. Academic press, 2018.
- [20] Julia Vettin and Dietmar Todt. Laughter in conversation: Features of occurrence and acoustic structure. **Journal of Nonverbal Behavior**, Vol. 28, pp. 93–115, 2004.
- [21] Marc Gelkopf and Shulamith Kreitler. Is humor only fun, an alternative cure or magic? the cognitive therapeutic potential of humor. **Journal of Cognitive Psychotherapy**, Vol. 10, No. 4, 1996.
- [22] Gregory A Bryant and Constance M Bainbridge. Laughter and culture. **Philosophical Transactions of the Royal Society B**, Vol. 377, No. 1863, p. 20200179, 2022.
- [23] Roza G Kamiloğlu, Akihiro Tanaka, Sophie K Scott, and Disa A Sauter. Perception of group membership from spontaneous and volitional laughter. **Philosophical Transactions of the Royal Society B**, Vol. 377, No. 1841, p. 20200404, 2022.
- [24] Doris G Bazzini, Elizabeth R Stack, Penny D Martincin, and Carmen P Davis. The effect of reminiscing about laughter on relationship satisfaction. **Motivation and Emotion**, Vol. 31, No. 1, pp. 25–34, 2007.
- [25] Camille Atkinson. Self-deprecation and the habit of laughter. **Florida Philosophical Review**, Vol. 15, No. 1, 2015.
- [26] Jonathan Ginzburg, Chiara Mazzocconi, and Ye Tian. Laughter as language. **Glossa: a journal of general linguistics**, Vol. 5, No. 1, 2020.
- [27] Ge Xu. An analysis of humor discourse in friends from the perspective of the cooperative principle. **Open Journal of Modern Linguistics**, Vol. 12, No. 4, pp. 460–470, 2022.
- [28] Kexin Wang, Carlos Ishi, and Ryoko Hayashi. A multimodal analysis of different types of laughter expression in conversational dialogues. In **INTERSPEECH**, pp. 4673–4677, 2024.
- [29] Michael McCarthy and Ronald Carter. “there’s millions of them”: hyperbole in everyday conversation. **Journal of pragmatics**, Vol. 36, No. 2, pp. 149–184, 2004.