

漫画における表情に着目した情緒注釈付き対話コーパスの作成

Construction of dialogue corpus with emotion focused on facial expression in comics

徳久雅人 松浦大樹 村上仁一 池原悟

Masato Tokuhisa Taiki Matsura Jin'ichi Murakami Satoru Ikehara

鳥取大学 工学部 知能情報工学科

Department of Information and Knowledge Engineering, Tottori University

1 はじめに

テキストからの情緒推定について分析を行うために、テキストから読み取れる情緒を注釈として付与したコーパスが必要である。情緒注釈の付与の問題点は、注釈付け作業者の主観に強く依存する点である。複数の作業者に注釈付与を依頼した際、情緒注釈を付与する箇所や、推定する情緒の深さに違いが生じやすい。表情や音声などを手がかりとすると、こうした不安定さの問題が幾分解消されると予想される。一方、漫画には、登場人物の表情が豊かに表現され、対話場面が大量に描写されている。

そこで、本稿では、漫画「ちびまる子ちゃん」[1]を対象として、情緒注釈付き対話コーパスを安定して作成することを目的とする。まず、1つの話につき2人の作業者がコーパスを試作する。ここで、漫画の登場人物の表情で注釈付与のタイミングを計り、文脈を理解した上で登場人物の情緒の注釈を付与する。次に、両者の協議により注釈の正解を定める。安定性は注釈の一致率や表情と情緒の関係から評価する。最後に、表情を参照しない情緒注釈付与の正解率について考察する。

2 対話テキストの入力

2.1 漫画テキストの入力条件

漫画には、登場人物が直接発話した台詞が「吹き出し」で記述される。その他に、想像している事柄、独り言、効果音が幾つかの形式で言語記述される。本稿では、話者の明確な部分について、これらの種類を区別しながら入力する。なお、登場人物以外に、ナレータの部分も入力する。

漫画「ちびまる子ちゃん」には、句読点が無い。ひらがな書きの部分が多い。これらは適宜補いながら入力する。

2.2 入力結果

漫画「ちびまる子ちゃん」の第1巻から第10巻までの104話を入力した結果、規模は10,213コマ、29,538文(388,809文字)であった。

3 表情と情緒の注釈付与

3.1 表情・情緒の分類

表情タグは、〈幸福〉、〈嫌悪〉、〈悲しみ〉、〈驚き〉、〈恐れ〉、〈怒り〉、および、〈背後〉の7種類を設ける。〈背後〉以外の6種類は、エクマンらの分類に従った[2]。漫画に特有な「青ざめ」、「冷や汗」、「震え」等の描写は登場人物の内心の情緒を表すと予想されるので、予備情報として付与する。登場人物の「後ろ

姿」に対しても、予備情報の描写がある際、表情と同様に情緒推定のきっかけになりうるので、〈背後〉という表情タグを付与する。

情緒タグは、8分類系と4分類系を用意する。8分類系は、《喜び》、《悲しみ》、《好ましい》、《嫌だ》、《驚き》、《期待》、《恐れ》、《怒り》、《なし》の9種類を設ける。複雑な情緒を注釈付けする際、複数の情緒タグで表現することが適切であると予想して、プルチックの分類に従った[3]。表情タグが付与されつつ情緒がみられない場合は《なし》を付与する。

情緒の8分類系は、区別が容易ではないことがある。そこで、8分類系を4分類系に圧縮して、P(《喜び》、《好ましい》、《期待》)、N(《悲しみ》、《恐れ》)、A(《嫌だ》、《怒り》)、S(《驚き》)、および、Z(《なし》)の5種類のタグを設ける。これら4分類系のタグは8分類系のタグより自動生成する。

表情と情緒の関係は、〈幸福〉は《喜び》、《好ましい》、《期待》の素直な表出とみなす。その他の表情は、それぞれ残りの情緒の素直な表出といえる。

3.2 注釈の付与条件

表情・情緒の注釈は、次の手順で付与する。

手順1: コマ単位の登場人物の表情を見て、表情タグを付与
 手順2: 表情タグの付与されたコマの前後の文脈より登場人物の情緒を作業者が推定し、コマ単位の情緒タグを付与
 手順3: 手順1および2は、2人の作業者が独立に付与を行うが、手順3では、2人の作業者が両者のタグを比較して、正解の表情・情緒タグを決定

手順1を踏むことで、作業者間で情緒推定を行う同期がとられるメリットがある。その他に、作業者にとつて、直接受け取れる情緒と、登場人物に内心存在すると思われる情緒とが、区別しやすくなるため、作業の負担が軽減されるメリットが期待される。手順2において両作業者の付与した情緒が大きく異なるならば、手順3において正解が作成しにくくなり、第4章で示す一致率が低くなると予想される。

3.3 付与結果

6人の作業者(当研究室の学生)が104話に情緒注釈を付与した。1話につき2人の作業者を割り当てた。手順1から3にかかった時間は、1話あたり約2時間であった。

コーパスの一部を表1に示す。表情・情緒のタグは、正解のタグである。タグは、台詞ごとではなく、コマごとに付与されているが、1コマ内で表情の変化がある場合には複数のタグが付与された。

正解のタグ数は、表情タグが12,606カ所(15,502個)、情緒タグが12,431カ所(16,861個)であった(1カ所に複数個のタグが付与)。正解の情緒タグの内訳は、《喜び》が4,588個、《悲しみ》が1,458個、《好ましい》が192個、《嫌だ》が3,010個、《驚き》が2,026個、《期待》が2,255個、《恐れ》が1,772個、《怒り》

が1,350個、《その他》が210個であった。

表情と情緒の関係について、(背後)の表出は192カ所であった。6種類の表情は12,401カ所あり、89.6%は素直な表出であった。たとえば表1の#4は素直な表出ではない(顔で笑って、心で泣いて…)。漫画のストーリーが単純ではあったが、表情の表す情緒を超えて深読みをしない傾向があるともいえる。

表1 コーパスの一部(話者・台詞は[1]第5巻より)

#	頁	コマ	話者	台詞	表情	情緒
3	22	3	まる子	うちのもみの木は小さいね。	悲しみ	悲しみ
4	22	3	お姉ちゃん	仕方ないじゃん。	幸福	悲しみ
5	22	5	お姉ちゃん	ぎゃあっ!	驚き、 恐れ	驚き、 恐れ
6	22	5	まる子		驚き	驚き
7	23	1	お姉ちゃん	まる子、あんたもみの木に、金魚の死骸埋めたでしょ。		
8	23	2	まる子	そうだよ。		
9	23	2		だって肥料になると思っ。		
10	23	2	お姉ちゃん	やめてよ。	嫌悪	嫌だ
11	23	2	ちゃん	気持ち悪い。	(汗)	

4 情緒注釈付与の安定性の評価

手順3で定めた正解の情緒タグと、手順2で付与した各作業者の情緒タグとを比較して、一致率により情緒タグ付与の安定性を調べる。一致率は次の式で評価する。

$$\text{適合率} = (\text{一致タグ数}) / (\text{付与タグ数}) * 100 (\%)$$

$$\text{再現率} = (\text{一致タグ数}) / (\text{正解タグ数}) * 100 (\%)$$

作業者毎の一致率を表2にまとめる。注釈付与を行った二者に共通する正解が定められたことがわかる。したがって、安定した注釈付けが行われたことが確認できた。また、8分類系と4分類系の差異は8ポイント以内であったことから、8分類系の各情緒の使い分けができていたことが伺える。

表2 表情に着目した付与の場合の一致率

作業者	適合率(%)	再現率(%)	(一致タグ数)
A者	90.3 / 90.9	86.7 / 88.8	(9,993 / 9,758)
B者	84.7 / 86.9	80.5 / 82.6	(4,176 / 3,884)
C者	70.3 / 76.1	68.3 / 76.0	(7,367 / 7,402)
D者	84.7 / 86.0	78.6 / 82.6	(7,990 / 7,839)
E者	94.1 / 94.5	89.4 / 92.1	(5,492 / 5,405)
F者	90.6 / 92.9	83.0 / 88.9	(8,623 / 8,354)
合計	85.1 / 87.4	80.5 / 84.9	(43,641 / 42,642)

※ xx / yy は、xx=情緒8分類系、yy=情緒4分類系
※ 情緒タグの付与されていない部分はタグ《なし》として扱った

5 考察

5.1 台詞のみからの付与との比較

本コーパスを用いてテキストからの情緒推定を行う際の推定精度の目安を得るために、表情を参照せずに台詞のみから情緒を推定して注釈を付与した場合の一致率を調べる。

対象は、[1]の第2巻1話・2話(計569文)、および、第6巻1話・2話(計648文)である。作業者は新たにG者が行う。第2話の作業者であるA者・E者、および、第6話の作業者であるD者・F者の当該範囲の一致率と比較する。

表3によると、表情の参照が「無」の方が「有」よりも、適合率および再現率がともに低い。

表3 表情の有無による一致率の比較

巻	表情	作業者	適合率(%)	再現率(%)	(一致タグ数)
2	無	G者	61.7 / 66.7	61.7 / 67.1	(274 / 284)
2	有	A者	86.2 / 87.2	84.2 / 85.1	(374 / 360)
2	有	E者	93.6 / 94.8	86.0 / 90.5	(382 / 383)
6	無	G者	59.7 / 64.4	51.9 / 61.5	(315 / 328)
6	有	D者	83.5 / 84.4	74.5 / 82.2	(452 / 438)
6	有	F者	89.5 / 92.4	81.5 / 88.6	(495 / 472)

※ xx / yy は、xx=情緒8分類系、yy=情緒4分類系

5.2 不一致の原因分析

不一致箇所から約200カ所をサンプリングして、不一致の原因を分析した。結果を以下に示す。

- (i) 表情をみれば正解の情緒が判明(28%)
- (ii) 顔の描かれていない場面/無表情の場面(26%)
- (iii) 「喜びと好ましい」など解釈ずれ(15%)
- (iv) 正解と異なるが別の作業者と一致(11%)
- (v) 普段の台詞なので表情で情緒を決定(9%)
- (vi) その他(11%):空想的なシーン(3%), …

不一致原因は、約37%が表情に依存した情緒推定であったこと[(i), (v)]、約26%が表情の有無による推定タイミングにズレが生じたこと、すなわち、言語に依存した情緒推定であったこと[(ii)]、約26%が情緒推定の正解が非決定的であること[(iii), (iv)]であった。

6 おわりに

対話テキストへの安定した情緒注釈付与のために、漫画の登場人物の表情に着目し、漫画104話(10,213コマ、29,538文)を対象に1話につき2人の作業者が付与を行った。正解と各者の一致率は適合率85.1%、再現率80.5%であったこと、情緒を素直に表現する割合が89%であったこと、情緒の8分類系と4分類系における差が8ポイントであったことより、付与の種類とタイミング、情緒推定の深さ、情緒の解釈の揺れについてそれぞれ安定性が評価できた。

また、表情を参照せずに注釈を付与する方法と比べた結果、情緒推定が表情に依存する側面、言語に依存する側面、および、正解決定の難しさについて定量的な感触を得た。本コーパスは、表情に着目して作成したものであるため、言語のみの情報から情緒を推定すると結果に差異が生じた。本考察では、言語情報から情緒が推定できる割合は50~60(%)程度であると推定する。

謝辞 本研究は科研費(若手(B):17700151)の下で行いました。台詞入力に協力頂きました田中勝弘様・東弘之様、そして、注釈付与に協力頂きました研究室メンバーに感謝します。

参考文献

- [1] さくらももこ:「ちびまる子ちゃん」, 第1~10巻, 集英社, 1987~1993.
- [2] P.エクマン, W.V.フリーゼン:「表情分析入門」, (訳編:工藤力), 誠心書房, 1990.
- [3] R. Plutchik: "The Multifactor-Analytic Theory of Emotion," *Psychology*, 50, pp.153-171, 1960.