MCN コーパス：ガイドライン設計とその運用

田中リベカ　（指導教員：戸次大介）

1　序論

自然言語処理研究においては、命題・事象レベルの意味が実テキストから取り出せるようになりつつあるが、テキスト中の命題や事象をただ取り出しても、全てが実世界の事実に対応しているとは限らない。情報の事実性を検知し、命題・事象レベルの情報を有効に活用するには、それらの事実性や確実性などを表すモダリティ表現の認識が必要である。しかし、実テキストにおいてこれらの表現の意味を正しく記述したリソースを作る上では、文脈における対象表現の意味・用法を特定し、曖昧性を解消することが課題となる。

本研究では、モダリティ関連表現にアノテーションを施したコーパス（以下 MCN コーパス）[5] の構築に使用するアノテーションガイドラインの改良を行っている。MCN コーパスのガイドラインにおいては対象表現の意味・用法を特定するために言語学的な分析に基づくテストを用いている。本稿では、MCN コーパスのアノテーションガイドラインで採用している「ネガティブテスト」[3] について、その運用と作成にあたっての問題点を考察する。

2　ネガティブテスト

一貫性のあるアノテーションを行うために、MCN コーパスではアノテータに提示する判断基準としてネガティブテストを採用している。[3, 2, 1]

ネガティブテストとは、「語順の入れ替え」や「対象となる表現の他の語への置き換え」などが可能であるという条件が「不成立」であるという判定が出た場合に、対象の表現が特定の分類に「属さない」とするテスト形式のことである。

> 表現 E' を含む文において条件 c_i が不成立ならば、表現 E' は分類 A_i に属さない。

このような形式を取ることによる利点は主に 2 点挙げられる。1 点目は、ある条件が「成立」するという判定よりも「不成立」であるという判定の方がより基準が明確で、アノテータ間での判断が一致しやすい傾向が見られるということである。これにより、より信頼性の高いアノテーション結果を得ることが可能である。もう 1 点目は、ネガティブテストの形式が「表現 E' が分類 A_i に属するならば、表現 E' を含む文において条件 c_i が成立する」という命題の対偶になっているという点である。これにより、テストの設計者は特定の意味分類の表現が満たす必要条件を見つければ良く、これは十分条件や必要十分条件を見つけるよりも一般に容易である。

ネガティブテストにおいては単独のテストのみでは分類は決定されないが、全ての分類についてそのようなテストを考えることで、消去法によって分類先を特定することができる。このような操作は一見煩雑であるが、信頼性の高いアノテーション結果が期待できるのみならず、見落とされていた新たな分類の発見やテストの問題点の特定がしやすいという利点がある。そのため、明確なフィードバックを元にテストを再構築するという過程を繰り返し、ガイドラインの改良を行うことが可能である。そのようにして作成・改良された最新のガイドラインの例を表 1、表 2 に示す。

3　テストの運用にあたって

ネガティブテストでは主に母語話者の直感に従って個々の判定を行うため、その提示方法などによっては正しく作用しない場合もあり、注意が必要である。

筆者は大学 1 年生〜大学院生を含む日本人学生 40 人程度を対象にネガティブテストにおけるアノテータの判断の一致度を調査した。[2] 被験者には用法の分類などについては一切明らかにせず、表現「（と）いう」が出現する文を提示し、「（と）いう」の箇所を別の表現に置き換えると意味が変化するか、という形式の問いに対して YES/NO/わからない、の回答を得た。その結果から、ネガティブテストの提示方法に関して以下のような傾向が見られた：

- 表現の置き換え操作を行う際、一部の被験者は誤って都合の良い置き換えをしてしまうことがあり、実際に表現を置き換えた文を提示する方が判断の一致度は高い
- 文の表層的な特徴を判定に用いる際、極端に自明な問いは却って一致度を低下させる
- 一文を提示するよりも前後の文脈を併せて提示する方が一致度は高く、判断に要する文脈の情報量は文によってまちまちである

ネガティブテストでは、アノテータがテストを元に判定できなかった場合やアノテータ間での結果が極端に一致しない場合などにテストが再構築される。そのため、判定が設計者の意図と一致しない場合には、その原因がテストの内容にあるのかテストが正しく適用されなかったことにあるのかを見極める必要がある。現時点ではガイドラインを実際に使用する前に、多数の被験者に対してネガティブテストにおける判断の一致度を調査し、被験者の多数が設計者の意図に合わない回答をした場合にテストを修正するという方針をとっている。しかし何人以上の回答が得られたら十分だと言えるのかは明らかではなく、課題が残されている。

4　テストの作成にあたって

ネガティブテストにおいては、条件が「不成立」であるという個々の判定はアノテータ間で判断が一致しやすい簡潔なものを理想としている。しかし実際に作成されたテストが、設計者の意図に反してアノテータにとって判定の困難なものである場合も少なくない。[4, 6]

そのような問題を引き起こす要因の 1 つには、設計者が想定するアノテータの言語知識と、実際にアノテータが有する知識との差がある。たとえば、現時点で作成されているガイドラインにおいては、対象表現の統語環境が判断材料になるケースがある。統語環境は客観的な性質であり表層的に判別できる場合も多いため、

表 1: 「(と) する」関連表現のガイドライン (一部抜粋)

sem	表現	別表記	特徴	例文	テスト	統語環境	備考
1	とする		「宣言する」「主張する示唆する」「述べる」「判断する」「規定する」「結論する」「明記する」に近い意味をもつ。公的な見解など、フォーマルな主張を記述する際に利用されることが多い。	総務省は「状況次第では、今後悪化に転じる可能性もある」としている。 鈴木氏は、増税は根本的な解決にはならないとした。 この契約では、契約者以外の者が事故を起こした場合は、保険金の支払いはできないとしている。 裁判官は、被告の犯行を「計画的できわめて悪質」とした。 検察側は、第一審での判決が不服であるとして、控訴する方針を固めた。 検察側は、第一審での判決を不服であるとして、控訴する方針を固めた。 検察側は、第一審での判決を不服として、控訴する方針を固めた。	「宣言する」「主張する」「述べる」「示唆する」「判断する」「規定する」「結論する」「明記する」(統語環境 2 の場合は「判断する」「規定する」「結論する」) のいずれにも置き換え不可、あるいは置き換えて意味が変化する場合はこのカテゴリではない。	【NP が】【S】とする 【NP が】【NP を】【S】とする	「とする」の直前に単独の NP はこない。(表層上 NP であるものは「だ」や「である」が省略されたものは「だ」や「である」が省略されたもの)
2	とする		仮想的な状況を記述する。「想定する」「仮定する」に近い意味をもつ。	太郎が犯人だったとする。その場合、アリバイはどう説明するんだ? 無人島に、一つだけ物を持っていけるとしよう。君は何を持っていく? 運転中に視界が悪くなったとします。その場合はどうすればよいでしょうか。 来年三月までの収入の合計を 300 万円とする。その場合、税金はいくらになるか。 直線 AB 上の点を Q とする。	「とする」を「想定する」「仮定する」のいずれにも置き換え不可、あるいは置き換えて意味が変化する場合はこのカテゴリではない。	【S】とする 【NP】を【NP】とする	
5	とする		ある行為を試みることを表す。「よう」「まい」のような助動詞を伴う節をとる。	知人から金をだまし取ろうとする。 決してくじけまいとする。 大麻を密輸しようとした疑い。	「とする」の前に「(よ) う」「まい」のような意志を表す助動詞がない場合はこのカテゴリではない。 「試みる」「努める」のいずれにも置き換え不可、あるいは置き換えて意味が変化する場合はこのカテゴリではない。	【NP】が【S-よう/まい】とする	

表 2: 「(と) いう」関連表現のガイドライン (一部抜粋)

sem	表現	別表記	特徴	例文	テスト	統語環境	備考
1	いう	言う	言葉を発するという「意図的な動作」が意味の中心である。また「NP が」という形の項として動作主を要求する (多くの場合節内に動作主が明示される)。 ※動作主が「世の人」「人々」「みんな」「誰か」である場合、「という 2」と意味的に近くなるが、これは「という 1」である。	太郎は「昨日渋谷で花子を見た」と言う。 花子はまだ怒っているようで、太郎を絶対に許さないという。 「太郎が責任をとるべき」と言う人は、どうかしている。 「叶わない夢はない」と人はいう。 花子は、太郎を天才だと言う。 その時警官が通りかかったことは、幸運だったというしかない。 花子が「おいしい」という店には行かない方がいいよ。	「話す」「主張する」「述べる」「表現する」「評価する」「判断する」のいずれにも置き換え不可、あるいは置き換えて意味が変化する場合はこのカテゴリではない。 「わざわざ」「口に出して」「あえて」「しつこく」のいずれを挿入しても意味が不自然になる場合はこのカテゴリではない。	【動作主 (NP)】が【命題 (S)】という 【動作主 (NP)】が【対象 (NP)】を【命題 (S)】という	統語環境が前者のものは「話す」「主張する」「述べる」のいずれかに置き換え可能
2	いう		伝聞の意味をもつ。「言葉を発する動作」よりも、むしろ「言説の存在」あるいは「言説が流布している状態」を表しているもの。 動作主が明示されない。 語用論的には、話者が間接的な言語情報として得たことを表す (直接経験して知っていることについては使わない)。 情報源を表す「〜によると」と共起することが多い。情報源が明示されない場合、「世間一般」「人々」「専門機関あるいは公的機関の公式発表」である。	ニュースによると、インフルエンザが流行っているという。 警察の調べでは、男は以前から現場付近で目撃されていたという。 駅前の焼肉屋は、このあたりで一番おいしいという。 日本人の 9 割が何らかのストレスを抱えているという。 世界には自分と同じ顔の人間が 7 人はいるという。 私たちの普段の生活の中にも、空海が中国からもたらしたというものがあります。それは一体、何でしょう。	「そう (だ)」と置き換えて違和感がある場合はこのカテゴリではない。 「いわれる」「いわれている」に置き換え不可、あるいは置き換えて意味が変化する (尊敬の意味になる等) 場合はこのカテゴリではない。	(〜によると/〜では)【命題 (S)】という	「という 7」との区別が表層的には困難
4	いう		副詞、あるいは「〜と」を伴って擬音語あるいはそれに類する表現をとる。「と」は省略される場合が多い。	人がぶつぶつと言う。 車がガタガタという。	「と」の補部が擬音語あるいは様態の副詞でない場合はこのカテゴリではない。	【動作主 (NP)】が【擬音語 (NP)】(と) いう 【動作主 (NP)】が【副詞】いう	

アノテータ間一致を考える上では有効な判断基準であるといえる。しかし、実際には文と名詞 (句) などの統語範疇の区別がアノテータにとって困難である事例も少なくなく、そのようなテストの使用によってかえって分類の特定が上手くいかないという傾向が見られた。このような高度な言語知識を要するテスト形式についてはフィードバックから気付きを得て、テスト作成時に一般に避けるべきこととしてガイドライン設計に反映させていく必要がある。

また、テスト設計者も想定していなかった、本質的に判定が困難である事例も多い。アノテータに高度な読解力を要求するケースや、非常に長い文脈を把握する必要がある場合、そもそも複数の読みが存在することが原因で判断が定まらないようなものもある。このような問題点は、実際にガイドラインを使用し、多数の文に対してアノテーションを行う中で発見されることも多いため、ガイドラインの修正プロセスをより早期に収束させることが今後の課題である。

5 結論

本稿では、MCN コーパスのアノテーションガイドラインで用られているネガティブテストと、その運用と作成にあたっての問題点を論じた。今後、他の表現についてもテストを構築し、精緻な分類を作成するとともにガイドラインの実用化に向けて取り組む予定である。

参考文献

[1] 戸次大介, 田中リベカ, 川添愛. MCN コーパス:モダリティ関連表現の曖昧性解消のためのアノテーションと言語学的テストの利用. テキストアノテーションワークショップ・コンテスト予稿集. 国立情報学研究所, 2012.

[2] 田中リベカ, 川添愛, 戸次大介. MCN コーパス:言語学的テストに基づくモダリティ・アノテーションの理論と実証. 第 2 回コーパス日本語学ワークショップ予稿集, pp. 135–144. 国立国語学研究所, 2012.

[3] 田中リベカ, 小池恵里子, 戸次大介, 川添愛. 言語学的テストに基づく意味アノテーションのガイドライン設計—確実性判断に関わる表現を中心に. 言語処理学会第 18 回年次大会発表論文集, pp. 401–404, 2012.

[4] 田中リベカ, 戸次大介, 川添愛. MCN コーパス:ガイドライン設計とその運用. 言語処理学会第 19 回年次大会発表論文集 (掲載予定), 2013.

[5] 川添愛, 齊藤学, 片岡喜代子, 崔栄殊, 戸次大介. 言語情報の確実性に影響する表現およびそのスコープのためのアノテーションガイドライン Ver. 2.4. Technical report, Department of Information Science, Ochanomizu University, OCHA-IS 10-4, 2011.

[6] 叢悠悠, 田中リベカ, 中村絢子, 酒向美帆, 佐宗智子, 清水蘭, 劉月晴, 川添愛, 戸次大介. 複合機能表現「という」の分類にみる MCN コーパスの方法論検証. 第 3 回コーパス日本語学ワークショップ予稿集 (掲載予定), 2013.