

# MCN コーパス：ガイドライン設計とその運用

田中 リベカ<sup>†</sup> 戸次 大介<sup>†‡</sup> 川添 愛<sup>‡</sup>

<sup>†</sup> お茶の水女子大学

{g0920522,bekki}@is.ocha.ac.jp

<sup>‡</sup> 国立情報学研究所

zoeai@nii.ac.jp

## 1 はじめに

自然言語処理研究においては、その興味の対象が深い意味処理の研究に移行してきており、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。しかし、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ためには、それらの事実性、確実性、話者の心的態度、情報の出所などの情報を担うモダリティの認識が必要である。

モダリティ関連表現には、「はず(だ)」「よう(だ)」「らしい」などの接尾語や、「(と)いう」「(と)する」などの複合語があり、命題に対する話者の態度を表す。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ことは、実テキストに表れる命題の意味をより有効に活用する上で重要である。しかし、モダリティ関連表現の多くは、文脈によって多様な意味・用法を持つため、テキスト上での曖昧性の解消が必要となる。

筆者らは、主にモダリティ関連表現を対象とし、事実性の認識に関連する表現の意味・用法を特定し、新聞記事等のテキストにアノテーションを施したコーパス(以下MCN コーパス)[4]を構築している。その過程において、表現の意味・用法の特定のために、対象表現の言語学的な分析に基づくテストを構築して用いている。言語学的なテストを用いてアノテーションを行うことにより、言語学的な分析に裏付けられ、かつ信頼性の高い曖昧性解消のためのリソースが得られ、深い意味処理に貢献できると考えている。

本稿では、「(と)いう」と「(と)する」のガイドラインを使用してアノテーションを行った結果をもとに、言語学的テストをアノテーション作業で用いるにあたっての問題点を考察する。まず、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]について解説する。その後、4人の

被験者からなる2つのグループについて、「(と)いう」/「(と)する」のガイドラインを使用したアノテーションを行った結果を示し、問題点に言及する。

## 2 ネガティブテスト

アノテーション作業を行う際、スキーマ設計者はアノテータに対して何らかの形で作業の方針を示す。このガイドラインの説明が適切かどうかアノテーションの質を大きく左右する。しかし、アノテータによる一貫した判断を可能にし、アノテータ間の判断を一致させるような判定基準の提示には工夫を要する。

文中の特定の表現に対して、アノテータはガイドラインの指示を元に、その出現をどの意味・用法としてアノテーションするかを判定し適切なラベルを選択して付与する。したがってここでは、アノテーションを「文中の表現がどの意味・用法で使われているかを判定し分類する作業」として捉えることとする。これにより、質の良いアノテーションは、アノテータによる分類の判定が一貫した判断の元で行われ、アノテータ間での分類が一致するものであると言い換えられる。

### 2.1 言語学的テストの利用

対象表現の実際の出現がガイドライン中のどの意味・用法であるかを判定する基準として、各意味分類について例文を提示する方法が広く用いられているが、MCN コーパスのアノテーションにおいては、特定の文脈の中で出現した表現の曖昧性を解消するために言語学的テストを用いている。ここで言う「言語学的テスト」とは、言語学の理論構築および検証に用いられるテストで、文や文の一部の容認性や適切性を判定するものである。

言語学的テストによる意味分類の判定においては、言語学的な知識をもつアノテータを必ずしも想定して

いないため、明確な語義の違いや特徴を判定材料とする方針をとっている。主に、以下のようなテストを使用している [1] :

- a 前後の語の形式（接続可能な活用形、補部に取りれる節の形など）を見るテスト
- b 語順の入れ替えを利用したテスト
- c 置き換えを利用したテスト
- d 共起可能性を利用したテスト
- e 否定形を利用したテスト

このうち b-e については、アノテータはガイドラインの指示に従って与えられた文をもとに操作後の文（判定文）をつくり、その判定文の容認性・適切性を各自の言語直感に基づいて判定する。このようにテストの設計者が専門知識に基づいて言語学的テストを作成することにより、アノテータが必ずしも言語学的な知識をもたない場合であっても一貫した判定を行うことが可能になると考えている。

## 2.2 ネガティブテストとテストセット設計

上記のようなテストを実際に運用するにあたり、田中ら [2, 3] の手法に基づき、「ネガティブテスト」を採用している。ネガティブテストとは、上の b-e に挙げたような「語順の入れ替え」「対象となる表現の他の語への置き換え」「(共起可能性を見るための)他の語の追加」などが可能であるという条件が「不成立」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。

上記の形式を採用する理由は、主に二点ある。一つは、「置き換え」「語の追加」などが可能であるという条件について、その条件が「成立する」(判定文が容認可能である、意味が通じる)という判断よりも「不成立である」(判定文が容認不可能である、意味不明である)という判断の方が、より基準が明確で、アノテータ間で一致しやすいという傾向が見られるためである。

もう一つは、この形式が「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よく、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。

ネガティブテストでは単独のカテゴリだけに注目すると「分類に属さない」という判定しかできないが、全ての分類についてそのようなテストを考えることに

より、消去法で最終的な分類が決定することが可能となっている。

上記のような方針に基づくガイドライン設計は一見煩雑だが、それまでの分類では見落とされていた新たな分類の発見や、テストの問題点の発見がしやすいという利点がある。ネガティブテストの形式は、「仮説の構築 テストの設計 テストの適用 フィードバック 仮説またはテストの再構築」というサイクルによりガイドラインの改良を行うのに向いているといえる。

## 3 ガイドラインの検証

「(と)いう」関連表現と「(と)する」関連表現の2種類の表現について、ネガティブテストを使用したガイドラインを用いて BCCWJ コーパスに対してアノテーション実験を行った(表1は「(と)する」関連表現のテスト付ガイドライン)。ガイドラインはいずれも改良の途上にあると考えられるが、テスト設計者が作成した意味分類や各々に対するテストを元にアノテータがどのような判定をするかを検証することは、今後のガイドラインの改良に役立つ。A群とB群のいずれのグループも4名のアノテータから構成されており、A群が「(と)いう」のアノテーションを、B群が「(と)する」のアノテーションを行った。50件分のアノテーション結果を得た後、グループ内で4名のアノテーション結果の一致個数を集計するとともに、付与されたラベルからテスト設計者の意図と結果とのギャップを分析した。

各グループ内で付与されたラベルを集計したところ、一致人数別の件数は以下の表のようであった:

	A群(「(と)いう」関連表現)	B群(「(と)する」関連表現)
4人一致	24	4
3人一致	24	26
2人一致	2	20

また、(1)結果が3人一致していた箇所について、3人の付与したラベルと他の1人が付与したラベルの間にどのような関係があるか(2)不一致時の判定の分かれ方に傾向はあるか、の観点から付与されていたラベルを分析した。その結果、(1)2つのラベルの分類間で統語環境が互いに異なっているケースが多く見られる、(2)アノテータの判定が分かれる分類の組として同じものが繰り返し現れている、という傾向が見られた。

## 4 残された問題点

筆者らの考える言語学的テストは、専門的な知識をもたないアノテータが使用することも想定して作成し

表 1: 「(と)する」関連表現のガイドライン

semt	表現	別表記	特徴	例文	テスト	統語環境	備考
1	と		「宣言する」「主張する」の「と」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よく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、被告の行為を「計画的でむね悪質」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と見做し、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見做し、控訴する方針を固めた。	「宣言する」「主張する」「述べた」「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)「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」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		仮定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くとしたら、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」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	
3	と		原則として生物(企業・団体などの組織も含む)による、意図的かつ動的な「位置づけ」や「変更」の行為を表す。	内閣は、通常国会を 5 月 8 日召集とする方向で調整に入った。実行委員会は多数決により、イベントのテーマを「発見」「交流」「ものづくり」の三つとした。その実施は、誰にも関係せずに息子の名前を「太郎」とした。委員らは、学生 100 人を対象として調査を実施した。代官山を拠点として活動する。日本の未来をより輝かせるために、精一杯務めたいと思います。彼は成人すると、幼馴染の女性を妻とした。結婚を前提としてつきあう。我が社は四月より週休二日制とする。(「休暇制度」あるいは「勤務体制」が省略されていると考え)以上をもちまして、臨時議会は閉会とします。大学の四年間は何をもって人生の目標とするかを決定する大切な時期であります。	「と」を「にする」に置き換え不可、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。(統語環境2の場合)特にフォーマルさが失われることがあるが、それは構わない。	【NPが】【NPを】【NP】とする 【NPが】【NPを】もって【NP】とする	「とする4」に分類されるものでも、置き換えてさほど不自然でないものがある。 「にする」の特長: (1)【NPが】の NP の指示対象に主体性がある (2)【NPに】への動的な変化を意味する 「とする1」に分類されるものについて置き換えを行うと、「変化する」の意味になってしまっって音韻性が下がる
4	と		静的な状況や関係を表す。【NPが】が非生物であることが多い(ただし例外もある)	この農産物は、南米を原産とする。この調査は、政策の効果を明らかにすることを目的とする。首都圏とは、東京を中心とする圏域のことである。人間を頂点とする食物連鎖ピラミッド。飲酒運転を原因とする事故が多発している。規制の強化を主な内容とする法案。その機械は、ぜんまいを動力とする。土俵際のかげひきを得意とする力士。(【NPが】が生物)介護を必要とする人(【NPが】が生物)	【NPを】【NP】とする。を「NPがNPである」(名詞句に係る場合は「NPがNPの」で判定しても良い)または「NPがNPにある(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場合はこのカテゴリではない。	【NPが】【NPを】	
5	と		ある行為を試みることを表す。「よう」「まい」のような動詞語を伴う節をとる。	知人から金をたまらなく取るとする。決して受け取らないとする。大麻を密輸しようとした疑い。	「と」の前に「(よ)う」「まい」のような意志を表す動詞語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	
6	と		心情や身体感覚、様態を表す名詞句(の)短縮形をとる。	愕然とする、驚々とする、ドキッとする、イラッとする。背中がぞくぞくとする。さらりとした飲み心地。意識が朦朧とする。彼女のことはしばらくそっとしておいてやれ。	「と」の前に心情や身体感覚、様態を表す表現がない場合はこのカテゴリではない。(補助テスト) 「と」の前の表現を取り去ると意味をなさなくなる場合はこのカテゴリである可能性がある。		
1	として		「という位置づけで」とほぼ同義。「という立場で」「という役割で」「という名目で」などと言いつづると自然な場合がある。	山田氏を議長として採用する予定。大入として制すかしないのか。資金として二十万円が確保されます。働きがいのある会社として注目されている企業。	【NPを】【NPとして】という形で出現している場合: 【NPとして】【NPを】の順序を入れ替えることができない、入れ替えると意味が変化する場合はこのカテゴリではない。【NPとして】を省略することができない、省略すると意味が変化する場合はこのカテゴリではない。【NPとして】単独で出現した場合: 「という位置づけで」「という立場で」「という役割(役職)で」「という名目で」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】として	
2	として		ある出来事についての判断が終了し、他に議論すべき事項へ興味が行くことを示す。	廊りに来るのは当然として、菓子折りの一つも持って来ないのはどういうことだ。廊はそれ以外として、問題は残る。	「(だ)けれど」「(だ)が」に置き換え不可、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】として	
3	として		慣用的な表現	卒別として、部外者の立ち入り厳禁する。結果として、その年の合格者はたったの五人だった。人生には時として、何をしてもうまくいかないことがある。誰一人として理解してくれない。一日として忘れたいことがない。運々として進まない。彼の行方は、密として知れない。			
	としては			彼としては、つらい決断だったに違いない。我が国としても、近隣諸国との関係改善を望んでおるところであります。取引期間としての早期来高値を更新した。何としても増強は阻止しなければなりません。			
	としても						
	としてのものとする						
	こととする						
	よしとする			よしとする 幸いとする 少いとする 少いとしても貢献できれば、我々の最も幸いとするところである。			

ている。また、判定にあたってのアノテータの負担は軽い方が望ましいと考えており、複雑な文脈を読み解かせることは避けたいと考えている。しかしこれらの方針を貫くには、運用上様々な困難・問題が生じることが分析の結果から示唆される。

アノテータの意見が割れた 2 つのラベルの分類間で、それらの統語環境が互いに異なっているケースが多く見られるということは、ガイドラインに統語環境の情報が提示されていても、実際のテキスト上で統語環境を見極めて分類先の候補を絞るという操作がアノテータにとって容易ではないことを示唆している。たとえば、対象表現の前後が名詞句であるか、あるいは文であるかという区別はしばしば言語学的な専門知識を必要とし、必ずしも表層的には判断できない。たとえば、以下は「とする1」に分類されると考えられる文である：

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。

一方、「とする1」の統語環境は「【NPが】【S】する」あるいは「【NPが】【NPを】【S】する」であるとガイドラインに記載されている。上の文は、「不服」の後に「だ」もしくは「である」が省略されていると考えられ、以下の文と同等であるとして「とする1」に分類されると判定されるのである：

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。

また、ガイドラインにおいて統語環境が終止形で記載されていても、実際の出現は必ずしも終止形であるとは限らない。このように実際のアノテーション作業においては、表層的な形式に縛られない、より踏み込んだ考察を求められることがある。統語環境による判

定は意味の判定と比べると客観的な基準となりうる反面、専門的な知識をもたないアノテータにとっては困難なものである。統語環境の判定基準をアノテータに示すことも考えられるが、ガイドラインが必要以上に煩雑なものになってしまう恐れがある。

アノテータの判定が分かれる2分類の組が大抵同じものであるという傾向については、意味や役割、その表現が使われる場面などが似ている2分類間の判定が困難であることを表している。実際、類似度の高い2分類間で表現の振る舞いの差を検知するには、注意深く文脈を読み解くことがしばしば要求される。たとえば以下の文を単独で見た場合、「(と)いう」の表現の分類のうち、言葉を発するという意図的な動作を表す「いう1」であるか、言説の存在や伝聞の意味を表す「いう2」であるかの判定は困難である：

山田社長は、ビルの1階が空いたのを機にカフェ経営に乗り出したという。

一見、伝聞を表す「いう2」に思えるが、この文が山田社長へのインタビュー記事の中に出現した場合、上の文を次のように書き換えて文脈に戻しても違和感がないため、「いう1」に分類される可能性も高い：

山田社長は、ビルの1階が空いたのを機にカフェ経営に乗り出したと話す。

このように、アノテーション時には対象表現を含む文のみに注目するのでは不十分なケースも多い。しかし文脈の読解を伴う判定はアノテータ間で一致しにくいことが予想されるため、ガイドラインにおいてはこのような複雑な判定基準は用いないことが望ましい。

上に見られるテスト運用時の問題は、実際にテストを使用せずしては問題点が明らかになりにくいため、未然に防ぐことが困難である場合が多い。そのためネガティブテストによるガイドラインでは、大多数のアノテータの判定がテスト設計者の意図に合わない場合に、その結果を受けてテストの提示方法や仮説の修正を行うという過程を経て改良されるという方針をとっているが、テストを作成・修正する度に多数のアノテータに対してテストの有効性を再検証することは実際は困難である。何人以上の回答があれば良しと判断するのか、またそのうち何人以上の判定が意図に合わなかった場合にテストの修正を行うのかなど、テストの収束度を評価する指標を作成することが実用上重要であると考えられる。

専門知識を前提とせず母語話者の言語直観で判定できるテストを用いることは、個人の主観によらない言語そのものの客観的な振る舞いを反映した意味リソー

スを作成することにつながる。また、判定に際してのアノテータの負担を軽減することで作業効率が向上するという側面もある。今後、テストを評価する指標を定めるとともに、直観的な判定が可能なテストを作成するためにどのような手順を踏むべきかを明確にし、言語学的テストの修正プロセスをスムーズに収束させることを目指す必要がある。

## 5 おわりに

本稿では、田中ら [2, 3] によるネガティブテストを用いてアノテーションを行い、結果の分析からテスト作成時や運用上の問題点を考察した。

今後、既に作成されたガイドラインのテストに改良を加えるとともに、「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」[4]に含まれるものを中心に、340種類の様相表現、13種類の否定表現、48種類の条件表現の網羅的な分類とネガティブテストの構築を行い、ガイドラインの拡張を行う予定である。

## 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] 川添愛, 齊藤学, 片岡喜代子, 崔榮殊, 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, Technical Report of Department of Information Science, Ochanomizu University, OCHA-IS 10-4, 2011.