

MCN コーパス : 「ノダ」にみるガイドライン作成の手法

田中 リベカ¹ 川添 愛² 戸次 大介^{1,2,3}

¹ お茶の水女子大学 大学院人間文化創成科学研究科

² 国立情報学研究所

³ 独立行政法人科学技術振興機構, CREST

tanaka.ribeka@is.ocha.ac.jp, zoeai@nii.ac.jp, bekki@is.ocha.ac.jp

1 はじめに

自然言語のテキスト中には、事実性や確実性の異なる命題・事象が混在している。それらの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、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表現の認識が必要である。川添ら [5] はそのような表現を「確実性判断に関わる表現」と呼び、アノテーションのためのスキーマを提供しており、MCN コーパスの構築が進められている。

確実性判断に関わる表現のアノテーションにあたっては、テキスト上での出現についてその意味・用法を特定し適切なラベルを選択するという、曖昧性解消が課題となる¹。筆者らは表現のもつ多義性を明らかにするため、これらの表現に関して意味・用法を分類した辞書を作成している。辞書には曖昧性解消のための言語学的テスト(ネガティブテスト [7])を記載し、作成した辞書をアノテーションガイドラインとしても活用している。分類基準をテストの形式で提示することで信頼性の高いアノテーションを実現すると同時に、時にはアノテーションの過程で分析が反証され、より正確なガイドラインに修正されていくというメリットがある。しかし、ガイドラインの具体的な作成手順が明らかではないという課題もある。

本稿では、ネガティブテストを用いたガイドラインを作成する手順を提案する。まず、ネガティブテストを用いたガイドラインの概要について解説する。第3節では、「ノダ」という表現について、言語学での分析をもとにガイドライン作成を試みた際に明らかになった問題について述べる。第4節では、その反省をもとに現状で効率的だと考えられるガイドラインの作成手順を提案する。

¹本稿では、アノテーション作業を「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。

2 ネガティブテストを導入したガイドライン

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 [6] :

- a 前後の語の形式(接続可能な活用形、補部に取りうる節の形など)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観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、アノテータが言語学の専門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。

上記のようなテストを実際に運用するにあたっては、田中ら [8, 7] 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(不可能)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などの発見がしやすく、またその修正も容易であるという利点がある。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。

一方、田中らの方法論では、ガイドラインを作成する方法が明らかではないという課題が残されていた。アナテータが専門家であることは要求しないが、ガイドラインの作成は言語学者が行うものとしており、アナテーション作業時にも度々言語学者にガイドラインの修正を依頼する必要があった。少数の専門家のみがガイドラインの作成を行っていたため、ガイドラインの構築に時間を要することも問題であった。

3 「ノダ」のガイドライン作成

本研究では、効率的なガイドラインの作成手順を明らかにすることを目標とし、日本語の「ノダ」という表現に関するガイドラインの作成を試みた。

3.1 日本語学における「ノダ」の分析

「ノダ」は状詞性接尾語「の」[3] に活用語尾が後接した形式を指すものとする。たとえば以下のような文に現れる。

1. 悲しくて泣いてる んじゃない。嬉しくて泣いてる んだ。
2. お前に聞いてる んだ。
3. へー、じゃあ君は一人っ子な んだ。
4. なんて間が良い のだ だろう！
5. そうだ、今日は土曜日な んだ わ。
6. あっ、第二土曜日は休みな んだ だった。
7. 咲かないよ。旅行に行った んだ。
8. どうしても知りたい ん です。
9. 大人は働かなきゃいけない んだ よ。
10. 太郎には、そのことだけが心残りに感じられる ので だった。

本研究では、主に野田 [4] の分析にもとづきガイドラインの作成を試みた。

野田によると、「ノダ」はスコープの「の(だ)」とムードの「のだ」に大別され、ムードの「のだ」は更に対事的ムードの「のだ」と対人的ムードの「のだ」とに分かれるという。先に挙げた例文のうち、1,2. はスコープの「の(だ)」、3.-6. は対事的ムードの「のだ」、7.-10. は対人的ムードの「のだ」である。

野田はスコープの「の(だ)」、対事的ムードの「のだ」、対人的ムードの「のだ」を区別するためのテストも提示している。以下に表で示す。

特徴	スコープ	対事的	対人的
名詞に後接する	×		
スコープに「は」 ² が入る	×		
否定にならない ³	×		
聞き手が必要	×	×	

これをもとにネガティブテストを作成すると、以下のようになる。

スコープ

- (1) 名詞に後接している場合はこの分類でない
- (2) スコープ内に取りたて助詞の「は」が出現する場合はこの分類でない
- (補助テスト) 否定形として出現している場合はこの分類である(一部例外がある)

対事的

- (4) 否定形の場合:「後悔」を意味しない場合はこの分類でない(ただし「後悔」を表す場合はこの分類である)

対人的

- (5) 否定形の場合:「命令」を意味しない場合はこの分類でない(ただし「命令」を表す場合はこの分類である)
- (6) 聞き手を必要としない場合はこの分類でない

このガイドラインは分類の特徴から作成したネガティブテストで構成されており、一見十分に思われるが、たとえば以下のような例に対しては機能しない。

教授:「この書類コピー取っておいて。その後こっこのプリントの採点もよろしく。」

学生:「え、あ、はい...」

教授:「あ、採点の前に購買でパン買ってきてね、ついでにこの鍵も事務室に返しておいて。」

²とりたて助詞の「は」。

³一部例外があり、「後悔」を表す場合は対事的ムードの「のだ」、「命令」を表す場合は対人的ムードの「のだ」とであると述べられている。

学生: 「...私は学生なのです! あなたの奴隷なのではありません!」

上の例は野田の分析によるとスコープの「の(だ)」であると考えられるが、名詞に後接していることから、スコープの「の(だ)」ではないと判定されてしまう。よって、スコープの「の(だ)」のテストを修正する必要がある。実際にこのテストを修正するには、スコープの「の(だ)」と認定される表現が満たすべき条件を理解する必要があり、必ずしも容易なことではない。しかし修正すべきという判断が可能なのも、テストが提示されていたからこそであり、このような明確な判断基準が提示されている分析は多くはない。

3.2 複数の分析を統合することの困難

野田は「ノダ」を5つに分類していたが、吉田 [2] は以下の11分類を提示している。

換言、告白、教示、強調、決意、命令、発見、再認識、確認、整調、客体化

野田がいうところのスコープの「の(だ)」にあたる表現は、吉田による分類では明示的に現れてはいない。野田がムードの「のだ」との対応のみを述べており、吉田による分類はムードの「のだ」の下位分類にあたると思われる。また、井島 [1] は野田のスコープの「の(だ)」が更に対期待対比、要素独立対比、要素連動対比の3つに分類されると述べている。

これらを単純にとらえると、「ノダ」は大きく分類するとスコープの「の(だ)」とムードの「のだ」の2つの用法をもち、スコープの「の(だ)」は井島の述べる3つに分類され、ムードの「のだ」は吉田の述べる11の下位分類をもつということになる。しかし、そのように結論づけるためにはいくつか解決しなければならない問題がある。

まず、異なる分析で提示されている分類の完全な対応をとらなければならない。たとえば今回のように、野田のいうスコープの「の(だ)」にあたる用法が吉田の分類には含まれていないということを断言するのは容易ではない。今回は野田の主張を元に、吉田の分類が野田のムードの「のだ」の下位分類に対応していると解釈したが、一般には分析が異なれば表現の分類基準も異なるため、常に「上位分類 - 下位分類」といった単純な関係になっているとも限らない。

また、意味の違いといっても、用法の違いとすべき違いなのか、単なるニュアンスの違いにすぎないのかといった程度の違いがあり、別分類とすべき境界を判

断する必要がある。たとえば野田は、吉田による告白、強調、教示などの分類は「ニュアンスに過ぎない」と述べている。吉田はこれらの用法を区別する明確な根拠(テスト)を提示していないため、ニュアンスの違いにすぎないという野田の主張が適切であるかどうかを、第三者が確認することは困難である。一方で、吉田が主張するように告白、強調、教示などが別の用法として明確に区別できるならば、より細かい分類を採用することが望ましい。アノテーションで「对人的ムードの『のだ』」というラベルが付与されるよりも、「告白」「強調」「教示」などのラベルが付与された方が、アノテーション後のデータとしては情報量が大きく有益であるためである。

「ノダ」の先行研究には他にも、「ノダ」の多様な用法を統一して説明する方法を明らかにすることを目的にしているものもある。そのような分析を元にする、「ノダ」の分類は1つだということになるが、そのような1つの意味ラベルを付与したデータは情報量が全くないと考えられるため、アノテーション作業には複数の分類が必要となる。逆に意味を細かく分類していくと、究極的にはニュアンスの違いや使用される状況の違いにまで分けることになるが、そのような分類は意味の分類だと言い難い。それだけでなく、明らかな違いが見いだせないことからアノテーションの一致度も低くなり、信頼性の高いリソースを構築することも難しくなることが予想される。

まとめると、先行研究の分類からガイドラインを作成する際の困難は2点ある。1点目は、テストとなりうるような特徴が明示されていない場合には分類基準を汲み取る必要があるということ、2点目は複数の分析による分類の対応を理解し統合するのはさらに困難であるということである。

4 ガイドライン作成手順の提案

本節では「ノダ」のガイドライン作成を試みた際の反省をもとに、ガイドライン作成手順として有効かつ効率的であると筆者が考える手順を提案する。

1. ベースとなるガイドラインを用意する
2. 作成したガイドラインを用いて、2名1組でアノテーションを行う
3. アノテーション結果が一致しなかった箇所について、新しい分類が必要か検討する
4. アノテーションの結果から、意味の異なる表現に同じラベルが付与されている場合は、分類の分割を検討する

5. 上記 2.~4. の手順を 2 名の結果が一致するようになるまで繰り返す
6. 他者を交えたアノテーションを実施する
7. Kappa 値が 0.8 を超えたら完成とする

上記の手順は、ネガティブテストを用いたガイドラインの修正・改善のプロセスとほぼ一致している。主な相違点は次の 2 点である。まず、ガイドラインの作成の段階で、設計者を含めた 2 名が試験的なアノテーション作業を行う中で分類を洗練させるため、実際のアノテーション作業にうつる前にガイドラインの信頼性が保証される。次に、ガイドラインの修正・改善のプロセスにおいては最初に用いるガイドラインは言語学者が作成するものとしており、その時点である程度分類が完成されていることを前提としていた。それに対し、ガイドラインの作成手順においては、ベースとなるガイドラインは分類が少なくとも構わない。

1. で用意するガイドラインは、該当表現に関する分類とその特徴、例文、統語環境の情報と言語学的テストを含む。この段階では、分類が網羅的である必要はないが、適切なネガティブテストが用意されていることが必要である。また、ベースとなるガイドラインの作成者は、必ずしも手順 2. 以降に携わる必要はないと考えている。ただし、作成したネガティブテストについて手順 2. 以降に担当する 2 名の設計者の合意がとれていることが必要である。

2.-5. については、2 名 1 組でアノテーションを行うことで、分類とテストの不備が回答の不一致として顕著に現れることが期待される。その場合、ほぼ自動的に 3. のプロセスに突入し、分類は改善される。4. では気づきがなければ修正はされないが、気づきが得られた場合に反映すればよいものとする。先行研究を参考にする場合、この段階で例文に対してアノテーションを行う。

5. までのプロセスを通して、設計者を含む 2 名の間に暗黙の了解が生じてしまうことがある。それによって、テストの表現が客観的なものでないことに気がつかなかったり、テストを用いて判定しているのか、直観的な分類を行っているのかの区別がつかなくなることが危惧される。そこで、6. では他者を交えてアノテーションを行い、作成したガイドラインが客観的に機能するものかどうかを検査する。

7. については、現状で最も普及している評価値であることを考慮して Kappa 値での評価を採用したが、評価値として適切であるかについては検討の余地があると考えている。

5 まとめと今後の課題

本稿では、アノテーション作業を行いながらガイドラインを修正するというサイクルを使うことによって、言語学者でなくてもガイドラインを作成することが可能であるとし、その作成手順を提案した。

用法の違いとニュアンスの違いを適切に区別する方法は現状では明らかではないが、少ない分類からアノテーションを経て分類を増やしていくという手法をとることで、信頼性の高いアノテーションを実現しつつ、直観に沿ってより細かい分類をもつガイドラインへと改良することが可能であると考えている。

参考文献

- [1] 井島正博. ノダ文の機能と構造. 日本語学論集, Vol. 6, , 2010.
- [2] 吉田茂晃. ノダ形式の構造と表現効果. 国文論叢, Vol. 15, pp. 46-55, 1988.
- [3] 戸次大介. 日本語文法の形式理論: 活用体系・統語構造・意味合成, 第 24 巻. くろしお出版, 2010.
- [4] 野田春美. 「の(だ)」の機能. くろしお出版, 1999.
- [5] 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.
- [6] 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, 2012.
- [7] 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.
- [8] 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, 2012.