

機能語の意味を表現する推論テストセット ——JSeM とりたて助詞テストの構築——

川添 愛¹ 田中 リベカ² 峯島 宏次² 戸次 大介²

国立情報学研究所¹ お茶の水女子大学大学院²

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

1 目的

日本語を母語とする人々の中で、「さえ」や「も」の意味を説明せよ、あるいは「だけ」と「しか」の違いを説明せよと言われて、すぐに適切な答えができる人はどれくらいいるだろうか。さらに、外国語におけるこれらの表現の対応物を用いずに説明するよう強いられた場合、何の難しさも感じずに対処できる人は、さらに少数ではないだろうか。しかし日本語母語話者は普段、これらの助詞的表現を難なく使っており、間違った使い方にも気がつくことができる。これは助詞だけに限らず、機能語一般についてもあてはまる。

我々は母語について膨大な知識を持っている。もしそのすべてが我々の意識にとって明らかであれば、言語学——少なくとも母語についての分析はほぼ不要になり、機械の上でそれを実現することにもさほど苦労しないはずである。しかし実際は、我々の言語知識の大部分は無意識のものであり、それを分析し体系化しようとする試みは継続的に行われている。そして、機械による意味理解の研究には、現在のところ、言語学によるそのような分析・体系化の「現時点での」成果を利用するものとそうでないものが混在している。

このような状況においては、何を基準にして、ある意味処理システムが「意味を理解できている（／できていない）」と判断すべきだろうか。現実的な方針は、人間が言葉の意味に基づいて行う観察可能な行動に焦点を当て、それらをシステムが近似できるかどうかを見ることであろう。文間の推論関係を判断する含意関係認識は、そのようなタスクの一つである。そしてその基盤となるのは、人間による推論を集積したテストセットである。

本研究では、日本語の機能語、中でも特にとりたて助詞に焦点を当て、その意味を推論テストの束として表現し、機械による機能語の意味理解の指標とするこ

とを提案する。以下では、「日本語意味論テストセット (Japanese SeMantics test suite; 以下 JSeM)」のとりたて助詞テスト群を紹介する。

2 日本語意味論テストセット JSeM 概要

JSeM[5][6] は、日本語の意味理論および含意関係認識システムの評価に資することを目的として筆者らが構築している推論テストセットである。既存の多くの推論テストセットと同様、JSeM の個々の推論テストは (1) のように、1) 正しいと仮定される「前提 (群)」(P)、2) 前提から推論されるかどうかが問われる「仮説」(H)、3) 前提から仮説が推論できるかどうかについての判断 (yes, no, unknown) の三つの部分によって構成される。

(1) answer: yes inference_type: entailment

P1 太郎もレポートを出した。

H 太郎がレポートを出した。

JSeM では FraCaS test suite (Cooper et al. 1996) の方針にならい、言語現象ごとにテストをまとめ、ターゲットとなる現象以外の要因を極力排除している。これまでに公開しているバージョンでは、FraCaS が扱っている量化、複数性、照応、テンス、比較、命題的態度をカバーし、FraCaS のテストと対応付けている。

また JSeM では、上記の例に示されるように、inference_type タグによって前提 P と仮説 H のペアに関わる推論過程のタイプを明示している。取りうる属性値には、形式意味論・語用論の文脈でよく知られている推論の分類 ([2][8] 等) に基づき、含意 (entailment)、前提 (presupposition)、慣習的含み (conventional implicature)、一般的な会話の含み (GCI)、個別的

な会話の含み (PCI) 等がある。また、前提 (presupposition) に前提調節 (presupposition accommodation) [1][7] が関わる場合は、それも属性値として明示する。以下の例示では、便宜上 inference_type は省略する。

3 JSeM とりたて助詞テスト

日本語のとりたて助詞は、名詞 (句) を始めとして文中の様々な要素に後続する機能語であり、その文の意味解釈に多大な影響を及ぼす。とりたて助詞の範囲には諸説あるが、沼田 (2003) において考察の対象されているものを例として以下に挙げる。

- (2) 「だけ」「のみ」「ばかり」「しか」「も」「さえ」「すら」「まで」「は (対比)」「だって」「くらい (／ぐらい)」「など (なぞ/なんぞ/なんか)」「なんて」「こそ」

とりたて助詞に関しては、日本語学において、各表現の用法の分類と一般化や、「とりたて」という文法的カテゴリの定義について詳細な議論がなされている。また、理論言語学 (生成文法、形式意味論等) においても、主にとりたて助詞の統語的特徴や量子子としての性質を中心とした研究の蓄積がある。

とりたて助詞の意味、および各表現の特徴を記述するにあたっては、いくつかの観点を考慮する必要がある。以下では、JSeM におけるとりたて助詞の意味を表現する推論テストの例を、その観点別に紹介する。なお、以下で挙げる具体例は現時点で JSeM に含まれている「だけ」「のみ」「ばかり」「しか」「も」「さえ」「すら」「まで」「は (対比)」に限定し、今後追加する予定の「だって」「くらい (／ぐらい)」「など (なぞ/なんぞ/なんか)」「なんて」「こそ」は含まない。

3.1 取り立てられる事物とそれ以外との関係

とりたて助詞は助詞的な表現であり、名詞句あるいは名詞句 + 格助詞を修飾することが多いが、副詞や動詞、節を修飾する場合もある。多くの場合、修飾先の表現が表す事物を取り立て、それと他の事物との関係を表す。具体的には、当該の文において主張されている内容が、他の事物についても成り立つ、あるいは成り立たないことを付随的に表現する。たとえばとりたて助詞「だけ」を含む「太郎だけがレポートを出した」という文は、「だけ」の修飾する「太郎」に関して、彼がレポートを出したということを主張すると同時に、太郎以外の人についてはレポートを出していないこと

を意味する。また「も」を含む文「太郎もレポートを出した」では、「太郎がレポートを出した」と同時に、「太郎以外の人 (誰か) レポートを出した」ことも意味する。この点は過去の多くの研究にとっての分析対象となっており、たとえば沼田 (2003) はこの点を「自者と他者」「肯定と否定」(および「主張と含み」「断定と想定」といった基本的特徴によって捉えることを提案している。

この点を推論テストとして表現する際に難点となるのが、「とりたて助詞の取り立てる事物 (沼田 2003 いう『自者』) とそれ以外の事物 (『他者』) からなる集合をどう設定するか」である。日常的な会話においては、そのような集合は話し手・聞き手に暗黙に共有されていることが多い。しかし推論テストにおいては、その言語的な自然さを保ちつつ、自者と他者からなる集合を明確に表現する必要がある。

JSeM では現在のところ、1) 前提群における集合の明示と、2) 適当な上位概念の導入の二通りの方法で対応している。1) を利用して作成したテストには (3) (4) のようなものがある。これらは、前提群 (P1、P2) において「ジョン」「ビル」の二人からなる集合 {ジョン、ビル} を設定し、P2 の「だけ」の修飾先である「ジョン」について「太郎に会う」が成り立つこと、および「ビル」についてはそれが成り立たないことを推論の形で表現する。

- (3) answer: yes
P1 太郎がジョンとビルに会いに来た。
P2 ジョンだけが太郎に会った。
H ジョンは太郎に会った。
- (4) answer: no
P1 太郎がジョンとビルに会いに来た。
P2 ジョンだけが太郎に会った。
H ビルは太郎に会った。

2) の「適当な上位概念の導入」とは、「ジョン」に対する「人」、「ケーキ」に対する「食べ物」、「ヘリコプター」に対する「乗り物」など、とりたて助詞の取り立てる事物をインスタンスあるいは下位概念として持つ上位概念のうち、最も顕著なものを「自者と他者を含む集合」として導入するものである。たとえば、「しか」「も」の例では次のようなものがある。

- (5) answer: yes
P1 太郎しかレポートを出さなかった。
H 太郎以外の人にはレポートを出さなかった。
- (6) answer: yes
P1 太郎もレポートを出した。
H 太郎の他に、レポートを出した人がいる。

ただし、実際にどのような上位概念が想定されるかは文脈によって変わるため、2) の方略は集合の全要素を前提群にて明示する 1) に対して補助的なものにすぎない。また上の (5) (6) に関しては、仮説 H の中に前提 P1 にはない情報(「太郎は人である」)が入っているという点で、純粹に含意(entailment)のみによって P1 から H が導かれているわけではないことに注意が必要である。むしろ、仮説を解釈する際に「太郎は人である」という前提の調節(accommodation)が関わっている。そのことを明示するために、inference_type 属性値に presupposition accommodation を加える。

3.2 話者の想定

沼田(2003) 他が指摘するように、とりたて助詞の中には、それが取り立てる事物あるいはそれ以外の事物に対する話者の想定を表すものがある。「さえ」「すら」「まで」などといったいわゆる「極限のとりたて」にはそれが顕著である。これらは、文中でそれらが取り立てる事物について、話し手が当該の命題が成立しないと想定していたことを表す。たとえば話し手が「太郎さえが試験に受かった」と言う場合、話し手はこの文によって「太郎が試験に受かった」ことを主張するのみならず、それ以前に「太郎は試験に受からない」と想定していたことを表現している。

このような想定を推論テストによって表現するために、JSeM ではいくつかのテストに「話し手」を明示的に導入している。(7) はその一例である。

- (7) answer: no

P1	花子は「太郎さえが試験に受かった」と言った。
H	花子は太郎が試験に受かると思っていた。

3.3 単調性

前述の通り、とりたて助詞は日本語学のみならず、理論言語学(統語論、形式意味論)においても盛んに研究されている。これらの文脈では、とりたて助詞の量子化としての側面が取り上げられることが多い。そのような側面の一つに、単調性(monotonicity)がある。以下に「のみ」の単調性の例の一部を挙げる。

- (8) answer: yes

P1	この授業には、文学部の女子学生のみが出席している。
H	この授業には、女子学生のみが出席している。

- (9) answer: unknown

P1	この授業には、女子学生のみが出席している。
H	この授業には、文学部の女子学生のみが出席している。
- (10) answer: unknown

P1	女子学生のみがレポートを期日までに提出した。
H	女子学生のみがレポートを提出した。
- (11) answer: unknown

P1	女子学生のみがレポートを提出した。
H	女子学生のみがレポートを期日までに提出した。

3.4 否定、数量表現との関係

とりたて助詞の量子化的側面として、否定表現や数量表現と共にした場合の解釈、特にそのスコープに関する特性も数多く分析されており([4][11][13]等)、JSeM でもそれを反映したテストを作成している。以下に、中村(2008)等で観察されている「遊離数量詞+も」と否定とのスコープ関係を示す推論テストの例を挙げる。(12) は前提 P1 において「数量詞+も」が否定より広いスコープを取ると解釈した場合の推論、他方(13) は狭いスコープを取ると解釈した場合の推論の例である¹。

- (12) answer: yes*

P1	引越しの荷物が五つも届いていない。
H	届いていない引越しの荷物が五つある。
- (13) answer: yes*

P1	今日のライブには、観客が十人もいなかった。
H	今日のライブにいた観客は十人未満だ。

3.5 他の表現との言い換え、表現の多義性

限定のとりたて助詞「しか」「だけ」「のみ」「ばかり」の間の言い換え、極限のとりたて助詞「さえ」「すら」「まで」の間の言い換えなど、似た用法を持つ表現間の言い換えも推論テストとして反映している。(14) は「さえ」と「まで」の例である。

- (14) answer: yes

P1	花子は太郎にさえチョコレートあげた。
H	花子は太郎にまでチョコレートあげた。

ただし、このような言い換えは常に可能であるとは限らない。言い換えが不可能になる場合としては、一

¹いずれの例においても、「数量詞+も」と否定のスコープ関係が逆になる解釈が可能である。このように、可能な複数の解釈のうちの一つを問題にしている場合は、answer 属性値に“yes*”のようにアスタリスクを付ける。

方の表現が多義的であり他方ない意味で使われている場合と、各々の表現が適切に使用される状況にずれがある場合などがある。前者の例 (15) では、前提 P1 の「さえ」は菊池 (1999) で指摘されている「十分条件性を強調する『さえ』」であるが、「まで」にはそのような意味がないために、推論が成り立たなくなる²。

(15) answer: unknown*

P1 金さえ出せば、好きなものが手に入る。

H 金まで出せば、好きなものが手に入る。

後者に関しては、定延 (2003) に挙げられている次のような例がある。

(16) うどんしか食べなかったから、ビタミンが不足している。(定延 2003, p.146, (7))

(17) ??うどんだけ食べたから、ビタミンが不足している。(同上, (6))

上のような、言い換えの結果が不自然あるいは容認不可能になるという観察は、意味の理論と処理の双方にとって重要であるものの、推論テストの形では反映しにくい。こういった事実をどのように適切にデータ化するかについては、現在検討中である。

4 おわりに：現状と課題

本論文では、JSeM とりたて助詞テスト群の概要を述べ、機能語の意味を推論テストの形で表現する試みを紹介した。2016年1月13日現在、JSeM とりたて助詞テストは385個のテストを含んでおり、近くJSeM HP (<https://researchmap.jp/community-inf/JSeM/>) にて公開する予定である。

今後は、「だって」「くらい」等のとりたて助詞やその他の機能語についてもセットを拡張する。重要な課題として、1) 上に述べた容認不可能性をどう反映するか、2) とりたて助詞に関わる現象を言語学的に体系立てて整理しつつ、意味処理の評価のために適切にテスト間のバランスを取る (yes/no/unknown のバランス等) にはどうすればよいか等について検討中である。

参考文献

[1] D. Beaver and H. Zeevat. 2007. "Accommodation." In G. Ramchand and C. Reiss (eds.) *The Oxford*

Handbook of Linguistic Interfaces. Oxford: Oxford University.

[2] G. Chierchia and S. McConnell-Ginet. 2000. *Meaning and Grammar: An Introduction to Semantics*. MIT Press.

[3] R. Cooper, D. Crouch, J. van Eijck, C. Fox, J. van Genabith, J. Jan, H. Kamp, D. Milward, M. Pinkal, M. Poesio, S. Pulman, T. Briscoe, H. Maier, and K. Konrad. 1996. "Using the framework." Technical report, FraCaS: A Framework for Computational Semantics. FraCaS deliverable D16.

[4] Y. Kato. 1985. Negative sentences in Japanese. *Sophia Linguistica* 19 [Monograph]: 1-229. Sophia University.

[5] A. Kawazoe, R. Tanaka, K. Mineshima, and D. Bekki. 2015a. "A Framework for Constructing Multilingual Inference Problem." In *Proceedings of MLKRep2015*, Vienna, Austria, July 2015.

[6] A. Kawazoe, R. Tanaka, K. Mineshima, and D. Bekki. 2015b. "An Inference Problem Set for Evaluating Semantic Theories and Semantic Processing Systems for Japanese." In *Proceedings of LENLS12*, Tokyo-Yokohama, Japan.

[7] D. Lewis. 1979. "Scorekeeping in a Language Game." *Journal of Philosophical Logic* 8: 339-359.

[8] C. Potts. 2005. *The Logic of Conventional Implications*. Oxford University Press.

[9] 菊池康人. 1999. サエとデサエ. 『日本語科学』6:7-31. 国立国語研究所.

[10] 定延利之. 2003. 現代語の限定とりたて. 沼田善子、野田尚史 (編) 『日本語のとりたて 現代語と歴史的变化・地理的変異』145-158、くろしお出版.

[11] 中村ちどり. 2008. 日本語の取り立て助詞と限定詞・名詞句フォーカス. 『言語と文化・文学の諸相』263-274. 岩手大学人文社会科学部.

[12] 沼田善子. 2003. 現代語のとりたての体系. 沼田善子・野田尚史 (編) 『日本語のとりたて 現代語と歴史的变化・地理的変異』225-241. くろしお出版.

[13] 茂木俊伸. 1999. とりたて詞「まで」「さえ」について - 否定との関わりから -. 『日本語と日本文学』28: 27-36. 筑波大学国語国文学会.

²ただし、かなり marginal ながら、推論が可能になる解釈も存在する。