

代名詞代用・呼びかけ表現の多言語データセット*

谷口 龍子 大久保 弥 野元 裕樹 南 潤珍
東京外国語大学

要旨：本稿では、代名詞代用・呼びかけ表現研究のための共通調査票（岡野他 2022）に基づいて構築した多言語データセットについて報告する。代名詞代用表現とは、人称代名詞以外の話し手・聞き手を指示する表現のことである。データセットは日本語、朝鮮語、マレー語、インドネシア語、ジャワ語、タイ語、ベトナム語、ビルマ語の 8 言語を対象とし、共通調査票の各概念に対する当該言語の表現とその日英訳、4 つの用法（自称、対称、呼びかけ、敬称）の有無、例文などの情報が含まれる。構築したデータセットから、容認可能な表現の割合に関して、「対称・呼びかけ>自称・敬称」が全言語で成り立つこと、「対称>呼びかけ」の言語ではさらに「自称>敬称」が成り立つことなどが分かった。また、データセットを用いて、田窪（1997）の呼びかけと対称の関係に関する分析、野元他（2021）が指摘する親族名称の代名詞代用における日本語の特殊性の検証も行った。

1. はじめに

日本語をはじめとするいくつかの言語では、人称代名詞ではないものの一人称・二人称人称代名詞のように話し手・聞き手を指示する表現が存在する。本稿では、Sneddon et al. (2010) に従い、そのような表現を代名詞代用表現 (pronoun substitute) と呼ぶ。代名詞代用表現のうち、一人称人称代名詞のように話し手を指示する用法を自称用法、二人称人称代名詞のように聞き手を指示する用法を対称用法と呼ぶ。代名詞代用と関連し、しばしば混同される現象に呼びかけがある。呼びかけ表現 (address term) は、聞き手を同定したり、注意を喚起するために用いられる表現である¹。(1)に3つの用法の例を挙げる。代名詞代用・呼びかけ表現の「お兄ちゃん」、「お母さん」が話し手・聞き手以外の個体（三人称）を指示するのにも用いられていることにも注意されたい。

(1) a. [兄が妹に]

お兄ちゃん_{自称}の言うこと聞かなかったら、お母さんに言いつけるよ。

b. [娘が母親に]

お母さん_{呼びかけ}、お母さん_{対称}のお兄ちゃんの名前、何だっけ？ (野元他 2021: 64)

代名詞代用・呼びかけ表現はさらに敬称 (title) の用法も持つことがある。例えば、「お兄ちゃん」は兄が複数いて区別したい場合に「健お兄ちゃん」のように固有名の直後に生起する。なお、ここでいう敬称は「さん」などの敬称専用表現と同じ形態統語的振舞いをするものをいい、敬意の有無は問わない。

* 本研究は JSPS 科研費 JP20H01255 の助成を受けた。本稿で報告するデータセットのデータ収集は同科科研費のメンバーが共同で行ったものである。

1 「代名詞代用表現」と「呼びかけ表現」のその他の名称については、野元他（2021）を参照されたい。

代名詞代用・呼びかけ表現の研究は、それを持つ言語では比較的好くよくなされてきた。しかし、個別言語内での研究ばかりで、通言語的な研究はほとんどない。代名詞代用・呼びかけの一般言語学的特徴を実証的に明らかにしようとする試みは始まったばかりと言える（野元他 2021、岡野他 2022）。

本研究では、岡野他（2022）の代名詞代用・呼びかけ表現研究のための共通調査票に基づき、多言語データセットを構築した。対象とした言語は、日本語、朝鮮語、マレー語、インドネシア語、ジャワ語、タイ語、ベトナム語、ビルマ語の 8 言語である。

本稿の構成は以下の通りである。まず、2 節で上記調査票と調査方法について説明する。3 節では、調査により構築した多言語データセットの概要を示す。4 節では、多言語データセットを用いて、通言語的視点からいくつかの言語学的な考察を行う。タイ語、ベトナム語、ビルマ語についての個別言語学的な考察は、岡野他（2022）を参照されたい。5 節は、結論である。

2. 調査項目および調査方法

岡野他（2022）の代名詞代用・呼びかけ表現研究のための共通調査票²は、(2)に示す 16 種類の意味カテゴリー（一部は形態統語的カテゴリー）に分けられた 147 項目から成る。「虚構的親族 (fictional family)」とは、血縁関係のない人物に対して仮定される疑似的な親族関係のことである。例えば、日本語では親族以外の中老年の人物に対して、虚構的親族関係に基づき、「おじさん」や「おばさん」という語が用いられる。

(2) 共通調査票で設定されている意味カテゴリー

年齢、親族、虚構的親族、教育、関係、職階、サービス、宗教、王室、照応表現、量化表現、指示詞、固有名、修飾を伴う名詞句、その他、人称代名詞

各項目は日本語・英語により記述された概念として提示されている。表 1 に「親族」の最初の 5 項目を例として示す。

表 1. 代名詞代用・呼びかけ表現研究のための共通調査票からの抜粋

ID	概念（日本語）	概念（英語）	Wordnet synset ID	カテゴリー
8	父	father	10080869-n, 09988063-n	Family
9	母	mother	10332385-n, 10278128-n	Family
10	伯父（父の兄）	uncle (father's elder brother)	10736091-n	Family
11	伯父（母の兄）	uncle (mother's elder brother)	10736091-n	Family
12	叔父（父の弟）	uncle (father's younger brother)	10736091-n	Family

2 https://github.com/matbahasa/ProSub/blob/main/common_questionnaire.tsv

まず、各言語でこれらの概念を表す表現を列挙した。その際、通言語的な定量的分析を見越し、表現の数は最低限に絞った。2つの概念について具体例を表2に示す。この表から分かるように、多義語の異なる意義は、当該の語の後に01、02のような数字を付すことで区別した³。

表2. 各言語における親族の「父」(ID 8) と虚構的親族の「父」(ID 35)

ID	日本語	朝鮮語	マレー語	インドネシア語	ジャワ語	タイ語	ベトナム語	ビルマ語
8	父, お父さん 01, パパ 01, おやじ 01	아버지, 아버지님 01, 아빠 01, 애비 01	ayah, bapa	ayah, bapak01, papi, papa, papah	rama01, bapak01, papah, pak'e	พ่ 01	bố	အဖေ01, ဖေဖေ01, ဒီဖေ, ဖခင်, အဖ, ဖခေအ
35	お父さん 02, お父様 02, おやじ 02	아버님 02	pak	bapak02	rama02, bapak02	พ่ 02, พ่ 03	bố	အဖေ02, ဖေဖေ02

次に、各表現について、自称、対称、呼びかけ、敬称の用法があるかを母語話者としての内省、聞き取り調査、事例調査により調べた。用法が存在する場合には、可能な限りで例文を収集した。

3. データセットの概要

前節で論じた共通調査票に基づく調査により、代名詞代用・呼びかけ表現の多言語データセットを構築した。各調査項目のデータは(3)に示した情報から構成される。「用法」は自称、対称、呼びかけ、敬称の4つである。「容認性」には yes (容認可能、用法あり)、no (容認不可能、用法なし)、* (敬称付きなら yes などの条件付き)、? (不明、要確認) の4種類を設けた。英訳はない場合もある。

(3) 多言語データセットに含まれる情報

言語コード、概念 ID、概念 (日本語)、概念 (英語)、意味カテゴリー、表現、表現和訳、表現英訳、用法、容認性、例文、例文和訳、例文英訳、出典、備考

図1にマレー語の例を示す。この図のようにデータは階層構造を持っている。なお、言語コードは最上位の要素であるため、図の中に含まれていない。

表3は各言語における登録表現・例文数をまとめたものである。表現のうち多義語は語義ごとにカウントしている。例えば、「嫁」は「息子の妻」の意と「自身の妻」の意でそれぞれ「嫁 01」と「嫁 02」のように区別され、2回カウントされる。表現の数は、異なり語数 (タイプ) である。

3 現状、表2中のベトナム語の bố には数字が付いていない。これは単なるミスで、今後修正する。

```

"9": {"concept_jpn": "母",
      "concept_eng": "mother",
      "category": "Family",
      "expression": [
        {"form": "emak",
         "trans_jpn": "お母さん",
         "trans_eng": "mum",
         "function": {"1st": {"judgement": "no", (省略) },
                     "2nd": {"judgement": "yes",
                              "example": [{"form": "Boleh saya ikut emak ke pasar?",
                                           "trans_jpn": "お母さんと一緒に市場に行ってもいいですか?",
                                           "trans_eng": "Can I go to the market with mum (= you)?",
                                           "source": "constructed",
                                           "notes": ""}]},
                     "address": {"judgement": "no", (省略) },
                     "title": {"judgement": "no", (省略) }}}}],

```

図 1. 概念 ID 9 に対するマレー語の情報（一部省略）

表 3. 各言語における登録表現・例文数

言語	日本語	朝鮮語	マレー語	インドネシア語	ジャワ語	タイ語	ベトナム語	ビルマ語	計
表現	191	236	126	123	148	168	109	255	1,356
例文	516	934	340	290	282	378	329	374	3,443

作成したデータセットは、<https://github.com/matbahasa/ProSub> において一般に公開されている。json と tsv の 2 つの形式を用意し、言語学のみならず言語教育や自然言語処理などの応用分野での利用も可能となっている。

4. 言語学的考察

4.1. 容認可能な表現の割合

登録表現のうち、代名詞代用、呼びかけ、敬称の用法が容認されるものの割合をまとめると、表 4 のようになる。最高値と最低値をそれぞれ太字（赤色）と斜体（緑色）で示した。

表 4. 容認可能な表現の割合（単位：％）

言語	日本語	朝鮮語	マレー語	インドネシア語	ジャワ語	タイ語	ベトナム語	ビルマ語
自称	24.6	68.2	19.0	26.0	43.9	34.5	24.8	29.0
対称	57.6	86.4	32.5	45.5	52.7	68.5	55.0	54.5
呼びかけ	58.1	80.1	42.9	54.5	63.5	67.3	63.3	47.8
敬称	28.3	29.2	19.0	23.6	36.5	26.8	36.7	11.8

この表から以下のことが指摘できる。①「対称・呼びかけ>自称・敬称」が8言語すべてで成り立つ。②「対称>呼びかけ」か「呼びかけ>対称」かは言語ごとに異なる。「対称>呼びかけ」の言語は、朝鮮語、タイ語、ビルマ語の3言語である。これらの言語ではさらに、「自称>敬称」が成り立つ。そのため、これらの言語では代名詞代用表現の重要性が相対的に高いと言える。一方、日本語、マレー語、インドネシア語、ジャワ語、ベトナム語の5言語では「呼びかけ>対称」となっている。③日本語とベトナム語以外では、敬称が最も低い値を示す。

4.2. 容認可能な用法の組合せ

各表現ごとに容認可能な用法の組合せを調べ、そのパターンの頻度をまとめると図2のようになる。

	日本語	インドネシア語	ジャワ語	朝鮮語	ビルマ語	タイ語	ベトナム語	マレー語
用法なし	70	48	20	11	91	41	34	73
自	2	5	5	15	9	10	6	6
対	5	4	1	10	19	1	2	3
呼	9	15	13	2	3	0	7	19
敬	7	5	3	1	6	10	4	7
自/対	7	1	0	10	1	1	0	0
自/呼	0	1	1	2	1	0	0	0
自/敬	0	0	0	0	0	1	0	0
対/呼	41	16	20	55	41	55	21	13
対/敬	0	1	1	1	2	0	0	0
呼/敬	2	3	2	0	3	0	6	4
自/対/呼	14	19	15	106	53	30	8	28
自/対/敬	0	0	0	0	0	0	0	0
自/呼/敬	0	0	0	1	0	0	0	0
対/呼/敬	19	12	11	31	8	15	19	7
自/対/呼/敬	60	22	29	83	10	27	27	5

図 2. 容認可能な用法の組合せ

全体として、対称と呼びかけを含む組合せの頻度が高いことが見てとれる。日本語では、4つの用法すべてを持つ表現が最も多いのに対し、他の言語では必ずしもそうになってはいない。

4.3. 対称と呼びかけの関係性

田窪（1997: 22）は日本語に関して「呼びかけ語は、職階等をタイトルとして含んだもので、文内対称詞 [= 代名詞代用対称用法] が職階を残して名前を省略したものである」と述べている。田窪は、(4a)の対称の「課長」は、呼びかけとして既出の「田中課長」から名前「田中」を抜いてできた省略形であると分析している。(4b)で呼びかけの「課長」と対称の「田中課長」が同一人物を指示する解釈を持たないのは、「対称=呼びかけの省略形」という関係が成り立っていないためであるとされる。

- (4) a. 田中課長、課長はこの案件に賛成ですか。
b. 課長、田中課長はこの案件に賛成ですか。 (田窪 1997: 22)

対称と呼びかけの分布の相関を指摘する点で、田窪の一般化は4.1節で指摘した「対称・呼びかけ>自称・敬称」という通言語的特徴や4.2節で指摘した対称と呼びかけを含む組合せの高頻度と整合する。だが、分布の相関の背後にあるとされる「名前の省略」については疑問が残る。もし対称が「名前+敬称」からの名前の省略であるならば、以下の2点が成り立つことが予測される。

- (5) 予測1：対称用法が容認可能ならば敬称用法も容認可能である。
予測2：敬称用法が容認不可能ならば対称用法も容認不可能である。

そこで、本研究のデータセットのうち人称代名詞以外でこれらの予測を検証した。各予測が実際に成り立つ割合を表5に示す。日本語だけ見ても100%からはほど遠く、名前の省略説は支持されない。

表5. (5)の予測が成り立つ割合（単位：%）

言語	日本語	朝鮮語	マレー語	インドネシア語	ジャワ語	タイ語	ベトナム語	ビルマ語
予測1	47.4	33.7	30.8	41.5	48.6	29.7	56.6	16.0
予測2	51.7	14.3	71.3	65.1	42.2	26.8	56.6	46.5

両仮説の反例には、「奥さん」、「ハニー」、「ダーリン」、「そちら」などがある。これらの語は対称として用いられるが、敬称としては用いられない。

ちなみに、名前の省略分析が間違っていたとしても、田窪は(4)のコントラストを正しく説明できる。「同一指示が許されるのは、一番構造的に高い位置にある名詞句が一番情報量が多い場合である。逆の場合は、同一指示が成り立たない」(p. 23)とも述べているためである。これはLasnik (1989) が束縛条件Dと呼ぶ制約で、「対称=名前の省略」という田窪独自の分析とは独立した関係にある。

4.4. 親族名称の非対称性

日本語では、親族名称が代名詞代用表現として用いられるのは目上の場合のみであり、目下の場合にはそれが不可能である（鈴木 1973: 149–150）。例えば、「お父さん」、「お母さん」、「お兄ちゃん」、「お姉ちゃん」は代名詞代用に使えるが、子、弟、妹を表す語で代名詞代用に使えるもの

はない。野元他（2021）は、日本語に見られるこのような目上と目下の間での非対称性は、朝鮮語、マレー語、インドネシア語、ジャワ語、タイ語、ビルマ語には見られないことを指摘している。本稿の多言語データセットにおいてもこのことが追認できた。(6)に朝鮮語の例を挙げる。この例では아이「子」が話し手・聞き手を指示するのに用いられている。

- (6) a. 아가 이렇게 힘든데 엄마는 어떻게 그럴 수 있어?
 「子供（=私）がこんなに大変なのにママはどうしてそうなの。」
- b. 아이들은 조용히 해!
 「子供たち（=あなたたち）は静かにすること！」

5. まとめと今後の課題

本稿では、東アジア・東南アジアの8言語を対象とした代名詞代用・呼びかけ表現の多言語データセットについて、その構築のプロセス、データセット自体の概要、データセットから見えてくる代名詞代用・呼びかけ表現の通言語的特徴を論じた。その過程で代名詞代用・呼びかけ表現に関する日本語の特徴も浮かび上がった。

現在公開中のデータセットには、未入力項目や入力ミスがまだ多く存在している。また、例文がなかったり、例文があっても日本語・英語の訳文がない項目もある。今後、そのような問題点を改善し、段階的にバージョンアップを行う予定である。

本研究で構築したデータセットは一般に公開されている。本稿でも若干の言語学的考察を行ったが今後さらに多くの人達によってこのデータセットが活用され、代名詞代用・呼びかけ表現の研究が個別言語および言語一般で進展することが期待される。その過程でデータに問題が見つかった場合にはご一報いただければ幸いである。

参考文献

- Lasnik, Howard. 1989. On the necessity of binding conditions. In *Essays on Anaphora*, 149–167. Dordrecht: Kluwer Academic Publishers.
- 野元裕樹、スニサー ウィッタヤーパンヤーノン（齋藤）、岡野賢二、トゥザライン、南潤珍、スリ・ブディ・レスタリ. 2021. 「代名詞代用・呼びかけ表現研究の現状：タイ語、ビルマ語、マレー語、インドネシア語、ジャワ語、朝鮮語」 『語学研究所論集』 25: 63–78. 東京外国語大学. doi: 10.15026/100158
- 岡野賢二、野元裕樹、スニサー ウィッタヤーパンヤーノン（齋藤）、トゥザライン、春日淳. 2022. 「アジア三言語における代名詞代用・呼びかけ語の共通項目調査」 『言語処理学会第28回年次大会発表論文集』 69–73.
- Sneddon, James Neil, Alexander K. Adelaar, Dwi N. Djenaar & Michael Ewing. 2010. *Indonesian: A Comprehensive Grammar* (second edition). London: Routledge.
- 鈴木孝夫. 1973. 『ことばと文化』 岩波書店.
- 田窪行則. 1997. 「日本語の人称表現」 田窪行則編 『視点と言語行動』 13–44. くろしお出版.