

終助詞と名詞の意味獲得における発話機能の効果の検討

The Effect of Speech Function on Meaning Acquisition of Sentence-Final Particles and Nouns

寺岡 弘貴 服部 侑介 早川博章 深田 智 岡 夏樹
 Koki Teraoka Yusuke Hattori Hirofumi Hayakawa Chie Fukada Natsuki Oka

京都工芸繊維大学
 Kyoto Institute of Technology

We proposed a computational model in which meanings of both sentence-final particles and nouns in Japanese are simultaneously learned, and conducted human-robot interaction experiments. The robot, however, did not talk during the experiments. In this research, we extend the robot so that the robot speaks words being learned. In this paper, we experimentally examine the effect of robot's talk on the learning interaction. As a result of the experiment, it turned out that the robot's utterance tends to change the way people teach. There was no significant improvement due to the introduction of the utterance function with respect to the understandability of the progress of the robot's learning and the accuracy of judgment by the participants to the learning completion of the robot.

1. はじめに

日本語において終助詞は日常会話で頻繁に用いられ、話者の心的態度を表すという重要な働きを持っていることが知られている。日本人同士の会話では終助詞の機能などから発話者の意図を推測し、適切な反応を返すことが求められる。ここでの反応とは、行動や発話による応答や、情報を記憶するなどの脳内の情報処理を含む。岡ら [Oka 13],[岡 15] は将来の家庭用ロボットは、終助詞などから発話者の意図を推測し、適切な反応ができることが望ましいと考え、人とロボットのインタラクションを通して、ロボットに終助詞や名詞の意味を獲得させることを試みた。この終助詞の意味の獲得とは、ロボットが相手の発話に含まれる終助詞に応じて、頷くなどの適切な行動を返したり、情報を記憶するなどの適切な内部処理を行えるようになることである。服部らが報告した実験 [服部 15] では、実験参加者とロボットのインタラクションを通して3種類の果物(リンゴ・ミカン・バナナ)の名前と3種類の終助詞(よ・ね・か)の用法を並行して学習させた。

これまでの服部らの研究では、発話するのは人(実験参加者)だけで、ロボット側は発話しない設定であった。これに対して本研究では、ロボット側も言葉を発する設定とした。本研究の目的は、ロボットが学習中の言葉を発することにより、ロボットの学習がどのように変化するかを調査することである。我々は、ロボットが学習中の言葉を発することの効果について以下の4つの仮説を立てた。

1. ロボットの反応がうなずき/首振りだけである場合と比べて、ロボットが発話することにより、ロボットの学習の進行の様子が人にとって分かりやすくなる。
2. 1.の結果として、人の教え方(どの果物を提示するか、どの終助詞を使うかなど)が、より学習が進みやすい方向に変化する。
3. 1.の結果として、人による、ロボットの学習完了の判断がより正確になる。

4. ロボットが能動学習(自身の学習の進み具合に応じて、意味があやふやな言葉をしゃべってみて相手の反応をうかがうなど)を行った場合は、学習効率が上がる。

本研究では、以上の4つの仮説のうち、まず、1.から3.を検証することを目指す。本論文では、仮説検証のために行った探索的な実験について報告する。

2. 提案手法

2.1 学習部

ロボットに果物を見せた際の画像認識結果とユーザ発話の認識結果であるテキストデータ、これらの入力を受けてロボットは頷く、または首を振るのいずれかの行動を表出しそれを実験参加者が評価する。学習システムはこの評価を報酬として受け取り、学習を行う。学習対象は、「終助詞と内部情報処理の対応付け」、「果物とその名前の対応付け」の2つである。終助詞により、内部情報処理(記憶処理または想起・比較処理。ここで行われる処理の詳細については以降に詳述)を切り替える設計になっており、実験参加者から与えられた報酬によってこの切り替え方を学習する。また、ロボットが行動するたびに毎回必ず実験参加者が評価する実験設定にしたため、強化学習としては遅れない報酬のみを扱うQ学習(式(1))を用いている。また、内部処理の選択はQ値に応じた確率で行動を選択するソフトマックス法(式(2))を採用している。また、下記における s は状態、 a は行動、 $Q(s, a)$ は行動価値、 α は学習率、 r は報酬、 τ は温度定数、 $\pi(s, a)$ は行動選択確率、 A はロボットの選択肢の集合である。本研究でも、服部ら [服部 15] と同様に $\alpha = 0.1$ 、 $\tau = 0.16$ とした。

$$Q'(s, a) \leftarrow Q(s, a) + \alpha(r - Q(s, a)) \quad (1)$$

$$\pi(s, a) = \frac{\exp(Q_t(a)/\tau)}{\sum_{i \in A} \exp(Q_t(i)/\tau)} \quad (2)$$

次に、記憶処理については記憶処理が選択された場合は必ず頷き、実験参加者からの評価を待つ。評価を受けて、認識した物体と名前の対応づけを行うQ値表と、発話された終助詞と内部情報処理選択の対応づけを行うQ値表の2種類の更新を行う。更新を行うと1対話が終了し、再び実験参加者からの発話

を待つ。また、想起・比較処理では、まず目の前の物体からその時点での Q 値表を用いてソフトマックス法 (式 (2)) で名前を想起する。その後ユーザ発話の果物の名前と想起した果物の名前を比較し、比較結果が一致したら頷き、不一致であれば首を横に振る。このロボットの行動に対して実験参加者が報酬を与え、その報酬に基づいて Q 学習を行い、記憶処理の時と同様に 2 種類の Q 値表を更新する。

2.2 発話部

学習部での報酬獲得後に発話処理を行う。ここではまず、想起・比較処理と同様に目の前の物体からその時点での Q 値表を用いてソフトマックス法でその名前を想起する。この時想起される名詞の候補として、正しい果物の名前に加えて、正しくない名詞を複数個採用した。これは、名詞の学習がまだ十分行われていない時に正しくない単語を発話する可能性があるように設定することで学習が不十分である状態を伝えることができると考えて決めた仕様である。また、ここでの正しくない名詞でどのような単語が適切であるかも合わせて探索を行った。以下に本実験で検証した誤り名詞を記載する。

- 「英語風発音」：英語の音声合成で果物の名前を発声する。「lingo」「mican」「banana」
- 「一音誤り」：果物の名前にある三音中 1 つがうまく発声できていない発音。「んごー」「りーご」「みか」「んかん」「んなな」「ばなー」
- 「非果物名称語」：名前の文字を入れ替えて果物の名前では無い単語を作り [佐藤 06]、発声する。「ごんり」「りごん」「かみん」「みんか」「んなば」「ばなば」

またロボットの発話に対する評価は、ロボットの行動に対する評価と同様に発話を「良い」「悪い」で評価する。実験参加者の評価を受けて発話部では、認識した物体と名前の対応づけを行う Q 値表の更新のみを行う。

3. 実験

実験は服部ら [服部 15] と同様のタスクを採用し、ロボットに果物の名前と終助詞の意味を教えることをタスクとして与えた。また、ロボットからの発話が「ない場合」と「ある場合」の二つの実験設定を用いて比較を行う。実験の環境と手順は参考文献 [寺岡 16] と同様のものである。

3.1 実験の流れ

各実験参加者に対して、以下の通り実験を実施した。

1. 実験参加者は 1 つ果物を選んでロボットの目の前の台に置く
2. 実験参加者は 9 通りの台詞 (終助詞×果物名) の中から 1 つ選んで発話する (例:「ミカンだよ」)
3. ロボットは終助詞に応じて内部処理を選択実行し、結果に応じて行動する
4. 実験参加者がロボットの行動を評価する
5. ロボットは目の前にある果物についての発話をする
6. 実験参加者は発話に対する評価を行う
7. ロボットが正しく果物の名前と終助詞を覚えたと判断するまで 1 から 6 の学習を繰り返す
8. 実験が終了した後、実験に関するアンケートに答える

4. 評価手法

仮説 3 の検証のために、果物 x の正解名称想起確率と終助詞 f の正解内部処理選択確率を定義する。果物の正解名称想起確率は、画像から正しい果物の名前が想起される確率である。『ロボットが x という果物を見て “ x ” という名前が想起 (recall) される確率』がある果物 x の正解名称想起 (recall) 確率 $P_r(x)$ (式 (3)) と定義する。 N は想起可能な名詞の集合、 $n(x)$ は x の正しい名前である。この確率の実験参加者内での平均を取り、これを果物の正解名称想起確率の平均 $\bar{P}_r(x)$ と定義する。

終助詞については服部らの研究 [服部 15] で明らかにされた通り、参加者がロボットに対して頷く行動を期待する発話 (Yes 発話) を行なった場合は記憶処理、想起・比較処理どちらの処理を選択しても良く、逆に首を横に振る行動を期待する発話 (No 発話) を行なった場合は想起・比較処理を選択した場合に正解とするのが良い。そこで、Yes 発話の回数 Y と No 発話の回数 N 、学習完了時点での Q 値から想起・比較処理を選択する確率 c として終助詞 f からの内部処理の正解選択 (selection) 確率 $P_s(f)$ を定義する。

$$P_r(x) = \frac{\exp(Q_t(x, n(x))/\tau)}{\sum_{i \in N} \exp(Q_t(x, i)/\tau)} \quad (3)$$

$$P_s(f) = \frac{Y + N \times c}{Y + N} \quad (4)$$

5. 結果

実験の参加者は 18 名 (20 代前半の男性 9 名、女性 9 名) であった。次の図 1 に、特徴的な学習を行なった実験参加者 m4 の学習進度を表すグラフを示す。線のみで記されたグラフは果物の名前を正しく想起できる確率を示す。y 軸の値が大きいほど確率が大きいことを表す。また、線上に点が打たれたグラフは終助詞に対して記憶処理か想起・比較処理、どちらが行われやすいかという確率を示す。y 軸の値が大きいほど記憶処理になりやすく、小さいほど想起・比較処理が行われやすいことを表す。終助詞のグラフの点の位置で実験参加者は各終助詞を用いた発話をした。また、このグラフ中の赤枠線の部分において実験参加者 m4 は、リンゴ、ミカン、バナナの 3 つについて「ね」を用いて発話、ロボットが正しく果物の名前を学習したと確認してから *1「ですか」「だね」「だよ」の順に終助詞を学習させた。

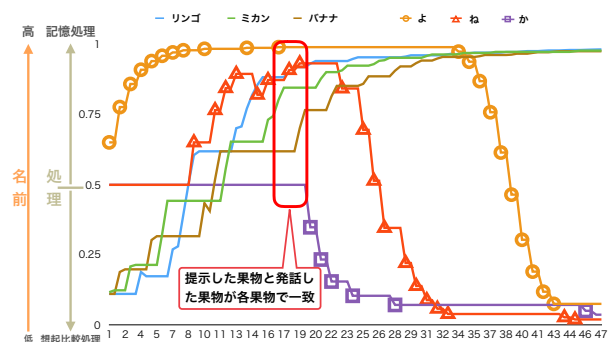


図 1: 実験参加者 m4 の実験結果

*1 事後インタビューで確認した。

図 2 には実験参加者が用いた終助詞の割合の比較結果を示す。ロボットからの発話がない時よりも、発話がある時の方が終助詞「よ」の発話割合が有意に大きくなる傾向 ($p = 0.06818 < 0.1$) が見られ、終助詞「か」では逆に有意に小さくなる傾向 ($p = 0.08497 < 0.1$) が見られた。

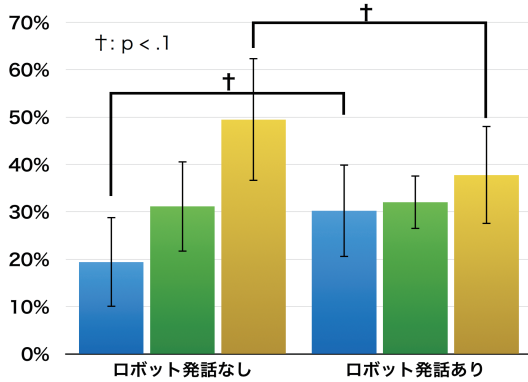


図 2: 終助詞の発話割合 (青: よ, 緑: ね, 黄: か)

図 3 にロボット発話なし条件とロボット発話あり条件の事後アンケート結果の比較を示す。正規分布であるものには t 検定を用い、正規分布でないものには Wilcoxon の順位和検定を用いた。結果、アンケート 5「ロボットの行動に対する評価に迷うことはなかったですか?」にロボット発話なし参加者 ($N = 5$) の方がロボット発話あり参加者 ($N = 12$) よりも有意にスコアが小さくなる傾向 ($p = 0.07675 < 0.1$) が見られた。つまり、ロボット発話あり参加者の方が評価に迷う傾向が見られた。

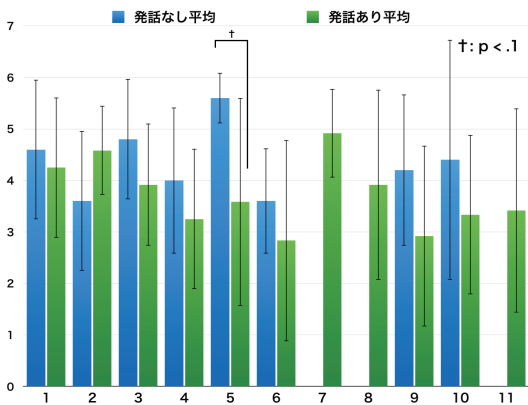


図 3: ロボット発話なし条件とロボット発話あり条件のアンケート回答の比較

表 1 にアンケートと発話回数の検定結果をまとめる。

表 2 に実験終了時の正答率をまとめる。

6. 考察

6.1 3つの仮説に対する考察

1 節で挙げたロボットが発話することによる効果の仮説について考察する。

6.1.1 仮説 1: ロボットが発話することにより、ロボットの学習の進行の様子が人にとって分かりやすくなる

事後アンケート Q2(ロボットの学習がどれくらい進んでいるかどうか、やり取りをしていてわかりましたか?)の結果ではロボットの学習の進度について、ロボットからの発話がない条件よりも発話がある条件の方が分かったと回答した人が多い。今回は実験参加者が少なかったため有意差は出なかったが、効果量を算出すると d [95%CI] = 0.895 [-0.286, 2.076], g [95%CI] = 0.849 [-0.271, 1.97] と大きいため、今後実験参加者を追加した上で分析をすることで差が出る可能性があると考えられる。

6.1.2 仮説 2: 仮説 1 の結果として、人の教え方がより学習が進みやすい方向に変化する

図 2 から、ロボットからの発話がある場合は「よ」を用いてロボットに話しかける割合が多くなる傾向があることがわかる。本実験では「よ」を教示用法で用いると伝えているため、ロボットからの発話がある場合は果物の名前を教える回数が多いことになる。これはロボットの学習の進行の様子がわかることで、ロボットはまだ果物の名前を学習しきれていないということを参加者が把握できたためであると考えられる。

また、参加者が学習内容を切り替える指標としてロボットの発話を用いていると思われた。実験参加者 m4 が最も顕著で、図 1 中の赤枠線の部分において、正しくロボットが果物の名前を発話できたことを確認して「ですか」の学習を初めている。他にも、実験参加者 f8 の事後インタビューで前半は果物の名前を教えて、果物の名前を話せるようになってから「だよ」の学習をどんどん進めると言う順に学習を進めたことを確認した。

これらの結果から「ロボットが発話することにより、人の教え方がより学習が進みやすい方向に変化する」と言うことができると考えられる。

6.1.3 仮説 3: 仮説 1 の結果として、人によるロボットの学習完了の判断がより正確になる

学習判断の正確さを比較するために正解率を式 3 と式 4 と定義して比較を行った。しかしロボット側からの発話の有無による差は見られなかった。

今回の実験ではロボットからの発話を果物の名前のみで設計している。仮説 2 の考察でも述べた通り、果物の名前の発話が正しくできるようになったことを確認してから終助詞の学習を進めている実験参加者もいた。つまり、果物の名前を発話することで、終助詞の意味が理解できているのかの判断もできるようになっていると言える。しかしこのとき、学習初期は Yes 発話を用いて果物と終助詞の学習を進め、後から No 発話を用いて終助詞の学習を進めている。Yes 発話を用いて学習をしている際には記憶処理が選択されるように学習が行われやすいため、後から想起・比較処理が選択されるように学習を変更することになる。今回設定した終助詞による行動選択の正答率は、No 発話で記憶処理が選択されると誤りであると設定しているため、学習方針の変更を行った実験参加者は正答率が下がってしまう。したがって、方針の変更を行ったと思われる実験参加者に対しては異なる正答率の導出式が必要になると考えられる。

6.2 評価の迷い

アンケート 5 番「ロボットの行動に対する評価に迷うことはなかったですか?」ではロボットからの発話がある場合に、評価に迷ったと回答している人が多くなる傾向が見られた。ロボットの行動に対する評価で迷いが見られたのは実験参加者 f7(発

表 1: 2 群間の事後アンケートの回答結果と発話回数と発話割合の平均値の差の検定

	発話なし (n=5)		発話あり (n=12)		t 検定					Wilcoxon の順位和検定		有意差
	平均	標準偏差	平均	標準偏差	p	t	df	信頼区間		p	W	
Q2	3.60	1.356	4.58	0.862						0.265	20.5	n.s.
Q3	4.80	1.166	3.92	1.187						0.2052	42	n.s.
Q4	4.00	1.414	3.25	1.362						0.3548	39	n.s.
Q5	5.60	0.490	3.58	2.019						0.07675	46.5	†
Q6	3.60	1.020	2.83	1.951	0.4463	0.78214	15	-1.322605	2.855938			n.s.
発話回数	37.20	15.471	40.83	20.489	0.7424	-0.33481	15	-26.76345	19.49679			n.s.
発話割合-よ	19.40%	0.095	30.19%	0.098	0.06818	-1.9653	15	-0.22485357	0.00912138			†
発話割合-ね	31.13%	0.096	32.04%	0.057	0.8236	-0.2268	15	-0.09404803	0.07595829			n.s.
発話割合-か	49.47%	0.130	37.77%	0.104	0.08497	1.8444	15	-0.01819812	0.25202006			†

表 2: 実験終了時正答率

		m1	m2	m3	m4	m5	m6	m7	m8	m9	f1	f2	f4	f5	f6	f7	f8	f9
リング	正解率	94.4%	98.8%	73.1%	98.3%	97.9%	84.6%	96.8%	56.1%	98.2%	97.4%	95.7%	99.4%	99.5%	97.3%	79.8%	98.9%	96.7%
ミカン	正解率	88.3%	98.6%	84.8%	97.9%	97.8%	69.0%	97.8%	82.3%	96.8%	98.0%	98.6%	99.0%	99.3%	96.2%	54.8%	98.3%	97.9%
バナナ	正解率	88.4%	98.9%	81.1%	97.6%	98.6%	63.2%	99.7%	92.2%	94.4%	99.3%	98.6%	98.6%	99.0%	91.0%	60.8%	96.6%	94.3%
よ	正解率	71.5%	99.3%	100.0%	96.8%	88.3%	100.0%	100.0%	100.0%	100.0%	100.0%	89.0%	100.0%	81.3%	100.0%	77.6%	98.5%	80.1%
ね	正解率	100.0%	79.1%	100.0%	99.0%	99.3%	100.0%	89.9%	100.0%	92.7%	100.0%	99.6%	85.1%	99.3%	100.0%	100.0%	98.8%	55.4%
か	正解率	97.8%	93.1%	100.0%	98.2%	98.5%	89.9%	99.8%	98.0%	95.9%	99.9%	99.8%	99.9%	99.6%	99.3%	91.1%	98.4%	99.9%
合計	合計発話回数	43	60	11	46	45	13	57	18	31	43	57	44	59	22	12	57	74
	95%超項目	2	4	3	6	5	2	5	3	4	6	5	5	5	5	1	6	3

話あり実験 (一音誤り) で『終助詞の判断が難しかった』と自由記述欄で回答している。実験参加者 m6(発話あり実験 (非果物名称語)) からは『動きは迷うことはなかったが、初め「ミカン」の発音がよくわからず誤った操作をしてしまった。』と自由記述欄での回答を得た。f8(発話あり実験 (非果物名称語)) は『発音が聞き取れないものがあつた』と事後インタビューの時に話していた。m6 と f8 はロボットの発話内容が原因で評価に迷ったと考えられ、これはロボットの発話の内容によっては聞き慣れないものもあり、また実験の説明でロボットの発話の内容については一切話していないために迷ってしまったと考えられる。しかし、各発話内容によつての発話回数や正解率、アンケートの回答に差は見受けられないため個人差が大きく影響すると思われる。

7. 結言

7.1 まとめ

本研究では、人とのインタラクションを通して果物の名前と終助詞の用法を獲得するロボット側から発話を行うことによる効果を検証した。ロボットが発話することによりロボットの学習の進行の様子が人にとって分かりやすくなるという仮説について有意な差は見られなかったが、効果量が大いことを確認した。また、ロボットからの発話で学習状況を把握することで人の教え込み方が『まず果物の名前を教えて、それから終助詞の用法を教え込む』というような戦略的な教育方法を取るようになる可能性があることを確認した。また、『ロボットからの発話で学習状況を把握することで人の学習完了の判断がより正確になる』という仮説については今回は差は見られなかった。

7.2 今後の展望

今回の実験では参加者が少なかつたため、統計的な分析で差が見られなかった可能性がある項目があつた。そのため、今後の展望としては参加者を増やして確実に仮説が成り立つかどうかについての調査を行うことである。また、今回定義した正解率の見直しや人の性格などの個人差を盛り込んでの検証も課題である。

謝辞

本研究は科研費 (25330260) の助成を受けたものである。

参考文献

- [Oka 13] Oka, N., Wu, X., Fukada, C., and Ozeki, M.: Concurrent Acquisition of the Meaning of Sentence-Final Particles and Nouns Through Human-Robot Interaction, in Lee, M., Hirose, A., Hou, Z.-G., and Kil, R. M. eds., *ICONIP 2013, Part I, LNCS 8226*, pp. 387–394 Springer, Berlin, Heidelberg (2013)
- [岡 15] 岡 夏樹, 服部 侑介, 深田 智, 尾関 基行: なぜロボットは「はい-いいえ」質問に対して「はい」と答える傾向があつたか: 人とロボットのインタラクション実験に基づく計算論的考察, 日本認知科学会第 32 回大会・論文集, pp. 187–194 (2015)
- [佐藤 06] 佐藤 久美子, 兼築 清恵: 無意味語反復から分かる、こどもの語彙能力-英語と日本語の比較, 論叢, Vol. 47, pp. 223–236 (2006)
- [寺岡 16] 寺岡 弘貴, 服部 侑介, 深田 智, 岡 夏樹: 終助詞と名刺の意味獲得の計算モデルへの発話機能の導入, Technical Report 436, 信学技報 (2016)
- [中村 15] 中村 友昭, 長井 隆行, 船越 孝太郎, 谷口 忠大, 岩橋 直人, 金子 正秀: マルチモーダル LDA と NPYLM を用いたロボットによる物体概念と言語モデルの相互学習, 人工知能学会論文誌, Vol. 30, No. 3, pp. 498–509 (2015)
- [内田 08] 内田 ゆず, 荒木 健治: 画像に対する発話を対象とした名詞概念獲得システム SINCA, 知能と情報, Vol. 20, No. 5, pp. 685–695 (2008)
- [服部 15] 服部 侑介, 岡 夏樹, 深田 智, 西崎 友規子: 人とロボットのインタラクションを通じた終助詞と名詞の意味獲得実験の分析: 個人差と他者モデルの観点から, HAI シンポジウム 2015, G-12, pp. 69–79 (2015)