

ロボットによる終助詞「よ」「ね」の意味獲得：

「終助詞の意味」を「発話に対する適切な行動」と捉える立場から

Meaning Acquisition of Sentence-Final Particles *yo* and *ne*:

Viewing the Meaning of Sentence-Final Particles as Appropriate Actions for Utterances

大上 涼麻 松岡 啓 荒木 修 柴田 諒子 高岡 勇紀

土坂 恭斗 呉 霞 深田 智 尾関 基行 岡 夏樹

Ogami Ryoma Matsuoka Hiromu Araki Osamu Shibata Ryoko Takaoka Yuki

Tsuchisaka Yasuto Wu Xia Fukada Chie Ozeki Motoyuki Oka Natsuki

1 はじめに

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つ。また、子どもによる獲得時期も早いことが知られている。しかし、その意味獲得の計算モデルはあまり研究されていない。ロボットが日本人と日常会話をするためには終助詞の機能を知っていることが不可欠であると考えられる。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、適切な反応をすることが求められる。この反応とは、行動や発話を返したり、情報を記憶するなどの内部情報処理をしたりすることを意味する。ロボットに言語を獲得させようとする研究は近年盛んである（たとえば、[1, 2]）が、助詞の獲得は、ほとんど試みられていない。

そうした中、野々口ら [3, 4] は、終助詞の意味を、終助詞を含む発話に対する適切な行動であると捉え、終助詞を含むユーザの発話に対する適切な反応、つまり、適切な行動や発話、内部情報処理をロボットができるようになることを最終目標と定め、手始めに、ユーザの発話に対する適切な反応のうち、目に見える行動として現れる部分の学習を目指した。

彼らは、終助詞「よ」の教示用法（聞き手が知らないと思われる情報を聞き手に告げ知らせる用）[5] と終助詞「ね」の同意要求用法（話し手・聞き手が共有していると目される情報について、聞き手に同意を求める用法）[5] を取り上げ、それらに対してロボットが適切に行動できるようになることを目指した。野々口らは、ロボットが対話相手から与えられる報酬に基いて適切な行動を学習するという枠組みで、学習実験用のシステムを構築し、予備実験を行ったが、そこでいくつかの不備が見つかり、本実験には至らなかった。

本研究では、野々口らの研究を引き継ぎ、終助詞を含む発話に対する適切な行動をロボットが学習できるようになることを実験的に示す。ここで獲得された行動を手がかりにして、内部情報処理を獲得させる（たとえば、

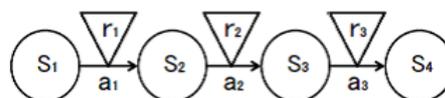


図 1: 行動系列の強化学習

対話相手が注目しているものの方向に顔を向けることを獲得し、それを手がかりに、そこで見たものの名前を覚える) ことを、今後試みる計画である。

2 学習方法

野々口らの研究では強化学習を用いて、ロボットが発話に対する適切な行動を獲得することを目指した。

強化学習とは、ある環境内におけるエージェントが、環境の状態を観測し、その状態に対して行動を選択し、その結果環境から報酬を得ることで、任意の状態においてどの行動を選択すべきかを学習する手法である。強化学習の目的は、この報酬の期待値を最大にする方策 π を学習することにある。

野々口らの研究では、話し手の発話に対する適切な行動を学習することを目的としているが、その行動は1つだけで完了するとは限らず、複数の行動からなる行動系列を扱う必要がある。このような場合、図1のような状態遷移図で行動系列を扱い、行動価値 $Q(s_t, a_t)$ を Q 学習により得ることができる。

しかし、図1のような状態遷移図で行動系列を扱う場合、以下の問題が生じる。

例えば話し手が、「リンゴ」の名前を知らないロボットに対して、リンゴを見せながら「リンゴだよ」と発話したとき、図1の状態 s_1 が、

- 話し手の発話：「リンゴだよ」
- 話し手の持っているもの：「赤い球体」
- 話し手の持っているものの名前：「知らない」

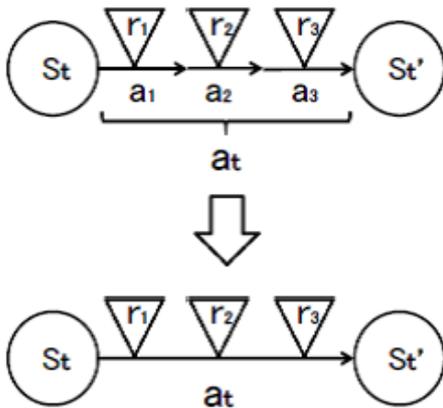


図 2: 本実験での定式化

- 今見ているもの: 「話し手が持っているもの」

というものであったとする. このとき行動 a_1 として「話し手の顔を見る」という行動を選んで状態 s_2 に遷移した場合, 状態 s_2 は,

- 話し手の発話: なし
- 話し手の持っているもの: 「赤い球体」
- 話し手の持っているものの名前: 「知らない」
- 今見ているもの: 「話し手の顔」

というものである. このとき行動 s_2 として, 「うなずく」という行動を選んで状態 s_3 に遷移したとき, 状態 s_2 と状態 s_3 では状態の持つ情報に変化がなく, 状態 s_2 と s_3 は等しくなり, 現在の状態と同じ状態に遷移したと言える. このように, 現在の状態と同じ状態に遷移する行動が存在する状態遷移図では, 行動の順番を学習することはできない. これを学習可能とするためには, 各状態において, 「現時刻までに何回行動したか」や「今までどのような行動をしたか」といった過去の行動の履歴に関する情報を持つ必要がある. このように状態が今までの行動の履歴に関する情報を必要とする場合, マルコフ性が成立しないと言う.

野々口らの研究では, 非マルコフ過程を扱う事を避けるため, 行動系列の長さを 1~3 回のいずれかと決めてしまい, 図 2 のように取りうる行動系列をそれぞれ 1 行動 (以下, 行動パターン) とし, 状態 s で行動パターン a を取り報酬 r を得る即時報酬課題として定式化した. 本研究でもこれを踏襲する.

3 評価実験

本実験の目的は, 終助詞「よ」の教示用法や終助詞「ね」の同意要求用法を用いた発話に対し, 2 節で述べた学習法を用いて適切な行動が学習できることを確認することである.

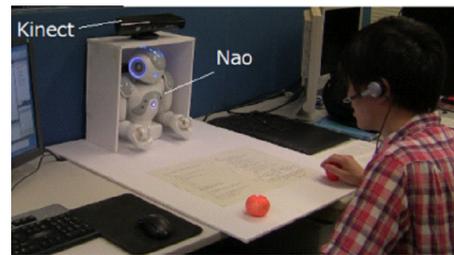


図 3: Nao がリンゴを見て, 被験者が「リンゴだよ」と発話している場面

実験は話し手 (被験者), 聞き手 (ロボット) がいる状態で, 「リンゴだよ」のように, 話し手が聞き手に対して「[word] だ [final]」と発話し, 聞き手は発話に対する反応として行動を返し, 話し手がその行動を評価するというやりとりを通して, 終助詞を用いた発話に対する適切な行動をロボットが獲得する. [word] には話し手と聞き手の間に置く物体名である普通名詞が入り, [final] には終助詞「よ」か「ね」が入る.

3.1 実験環境

図 3 に示す環境で実験を行う.

使用するロボット機体は ALDEBARAN Robotics 社の Nao T14 (以下, Nao) である. また, Nao の上部に Microsoft 社の XBOX360 Kinect センサー (以下, Kinect) を設置する.

被験者には Nao の正面に座ってもらい, テーブル上には「リンゴ」「ミカン」を模した 2 つ物体を, それぞれ被験者の右側手前と左側手前の決められた位置に置く.

Kinect は非接触型制御装置であり, RGB カメラ, 深度センサーなどをもつ. 深度センサーにより被験者の骨格情報を取得し, 被験者の顔や手の位置を得ることができる.

3.2 実験手順

実験は以下の手順を 1 エピソードとして行う.

1. Nao がテーブル上の物体のどちらか一方にランダムに顔を向け, 開始の合図として「Talk to me」と話す.
2. 被験者はタスクに応じて発話対象とする物体を選び, それを右手で把持, もしくは右手で指さして発話する.
3. Nao があらかじめ組み込まれた行動パターンからランダムに 1 つ選び行動する.

4. 被験者は、Nao の行動が適切であるかを評価する。これは Nao が行動中であっても評価可能である。行動終了後、Nao が「How was my action?」と話すので、被験者は途中で評価した場合であっても、このタイミングで必ず評価する。

評価内容は「良い」か「悪い」のどちらかとし、Visual C # の Windows フォームアプリケーションのボタンツールを用いてマウスを左手で扱い評価してもらう。この評価はあくまでも被験者の発話に対する Nao の行動が適切か否かの評価である。

学習は全ての実験が終了した後、実験データを基に 2 節で示した方法により行う。

3.3 実験設定

3.3.1 発話内容

被験者の発話は「リンゴだよ」「ミカンだよ」「リンゴだね」「ミカンだね」のいずれかとし、音声認識を用いて入力する。

また、被験者は発話する際、必ず右手で発話対象物を把持する、もしくは指さすとする。ただし、物体を移動させることはできない。

3.3.2 行動パターン

Nao は、以下の 4 種類の行動をすることができる。

- 被験者の顔（以下、顔と称する）を見る。
- 被験者が発話対象に選んだ物体（以下、発話対象物と称する）を見る。
- 被験者が発話対象に選ばなかった物体（以下、その他の物体と称する）を見る。
- うなづく。

Nao は、これらを 1~3 個組み合わせる行動する。上に示した 4 種類の行動を上からそれぞれ A, B, C, D とし、例えば 3 個組み合わせる $A \rightarrow B \rightarrow C$ と行動する行動パターンを ABC のように表すとすると、Nao は次のような行動パターンで行動することができる。

1 つの行動からなる行動パターン

A, B, C, D

2 つの行動からなる行動パターン

$AA, AB, AC, AD, BA, BB, \dots, DB, DC, DD$

3 つの行動からなる行動パターン

$AAA, AAB, AAC, AAD, ABA, \dots, ADD, BAA, BAB, \dots, DDB, DDC, DDD$

表 1: Nao の行動パターン

1 番目	2 番目	3 番目	1 番目	2 番目	3 番目
顔	顔	顔	その他	顔	顔
顔	顔	対象物	その他	顔	対象物
顔	顔	その他	その他	顔	その他
顔	顔	うなづく	その他	顔	うなづく
顔	対象物	顔	その他	対象物	顔
顔	対象物	対象物	その他	対象物	対象物
顔	対象物	その他	その他	対象物	その他
顔	対象物	うなづく	その他	対象物	うなづく
顔	その他	顔	その他	その他	顔
顔	その他	対象物	その他	その他	対象物
顔	その他	その他	その他	その他	その他
顔	その他	うなづく	その他	その他	うなづく
顔	うなづく	顔	その他	うなづく	顔
顔	うなづく	対象物	その他	うなづく	対象物
顔	うなづく	その他	その他	うなづく	その他
顔	うなづく	うなづく	その他	うなづく	うなづく
顔	顔	-	その他	顔	-
顔	対象物	-	その他	対象物	-
顔	その他	-	その他	その他	-
顔	うなづく	-	その他	うなづく	-
顔	-	-	その他	-	-
対象物	顔	顔	うなづく	顔	顔
対象物	顔	対象物	うなづく	顔	対象物
対象物	顔	その他	うなづく	顔	その他
対象物	顔	うなづく	うなづく	顔	うなづく
対象物	対象物	顔	うなづく	対象物	顔
対象物	対象物	対象物	うなづく	対象物	対象物
対象物	対象物	その他	うなづく	対象物	その他
対象物	対象物	うなづく	うなづく	対象物	うなづく
対象物	その他	顔	うなづく	その他	顔
対象物	その他	対象物	うなづく	その他	対象物
対象物	その他	その他	うなづく	その他	その他
対象物	その他	うなづく	うなづく	その他	うなづく
対象物	うなづく	顔	うなづく	うなづく	顔
対象物	うなづく	対象物	うなづく	うなづく	対象物
対象物	うなづく	その他	うなづく	うなづく	その他
対象物	うなづく	うなづく	うなづく	うなづく	うなづく
対象物	顔	-	うなづく	顔	-
対象物	対象物	-	うなづく	対象物	-
対象物	その他	-	うなづく	その他	-
対象物	うなづく	-	うなづく	うなづく	-
対象物	-	-	うなづく	-	-

顔 … 顔を見る、対象物 … 発話対象物を見る、

その他 … その他の物体を見る

1 つの行動からなる行動パターンは 4 通り、2 つの行動からなる行動パターンは $4^2 = 16$ 通り、3 つの行動からなる行動パターンは $4^3 = 64$ 通りあり、Nao はこれら計 84 通りの行動パターンの中からランダムに 1 つ選び、行動するものとする。行動中は Nao の LED が黄色に点灯する。

全ての行動パターンを表 1 に示す。

3.3.3 物体の把持判定

3.3.1 節で説明したように、被験者は発話する際、必ず右手で発話対象物を把持する、もしくは指さす。この節では物体の把持判定の仕方について説明する。

被験者の右手の位置は、Kinect の深度センサーにより被験者の骨格情報を取得することで求めることができる。実験では、実験に使う 2 つの物体はテーブル上の決められた位置に置き、被験者はこれらを動かすことはできないので、物体の場所はあらかじめプログラムに入力

しておくことができる。

被験者の右手の位置を点 $B(x_b, y_b, z_b)$ 、任意の物体の位置を点 $C(x_c, y_c, z_c)$ とおくと、この2点間の距離 d は、

$$d = \sqrt{(x_b - x_c)^2 + (y_b - y_c)^2 + (z_b - z_c)^2} \quad (1)$$

という式で求めることができる。この d の値が 50mm 未満、つまり被験者の右手と任意の物体の距離が 50mm 未満のとき、その物体を把持していると判定する。

3.3.4 物体の指さし判定

被験者の右肘の位置から被験者の右手の位置へのベクトルを、被験者が指さしている方向と捉えて指さしている物体を判定する。

被験者の右肘の位置も右手の位置と同様、骨格情報を取得することで求めることができる。被験者の右肘の位置を点 $A(x_a, y_a, z_a)$ 、被験者の右手の位置を点 $B(x_b, y_b, z_b)$ 、任意の物体の位置を点 $C(x_c, y_c, z_c)$ とおくと、右肘から右手へのベクトル \vec{AB} は、

$$\vec{AB} = (x_b - x_a, y_b - y_a, z_b - z_a) \quad (2)$$

となる。ここで、ベクトル \vec{AB} を

$$\vec{AB} = (x_v, y_v, z_v) \quad (3)$$

とおくと、直線 AB 上の点 $P(x, y, z)$ は、任意の定数 t を用いて

$$(x, y, z) = \vec{OA} + t\vec{AB} = (x_a, y_a, z_a) + t(x_v, y_v, z_v) \quad (4)$$

と表すことができる。ここで、点 C から直線 AB に降ろした垂線の足が点 P のとき、ベクトル \vec{CP} とベクトル \vec{AB} の内積は 0 であるから、

$$\begin{aligned} \vec{CP} \cdot \vec{AB} &= (x_a + tx_v - x_c)(x_v) \\ &\quad + (y_a + ty_v - y_c)(y_v) \\ &\quad + (z_a + tz_v - z_c)(z_v) \\ &= 0 \end{aligned} \quad (5)$$

という式が成り立つ。上式を t について解くと、

$$t = \frac{x_v(x_a - x_c) + y_v(y_a - y_c) + z_v(z_a - z_c)}{x_v^2 + y_v^2 + z_v^2} \quad (6)$$

となり、 t が求まる。これを式 4 に代入することで、直線 AB 上にある点のうち、点 C に最も近い点 P の座標が求まる。この点 P と、点 C 間の距離 d を式 1 により求めることで、直線 AB と点 C 、つまり被験者の右肘と右手を通る直線と任意の物体の位置との距離 d が求まる。この d の値が 50mm 未満のとき、その物体を指さしていると判定する。

3.3.5 4通りの実験

実験は以下に示す 4 通りの実験を行う。

実験 1 「Nao はテーブル上の物体の名前を知らない」という設定で行う。被験者には Nao が物体の名前を知らないことを伝え、Nao がリンゴを見ていたら「リンゴだよ」と見ている物体の名前を教えるというタスクを与える。また、Nao は実験 1 では物体の名前を教えても覚えられない設定で行うことも被験者に伝えておく。実験 1 の例を図 4(a) に示す。

実験 2 「Nao はテーブル上の物体の名前を知っている」という設定で行う。被験者には Nao が物体の名前を知っていることを伝え、Nao がリンゴを見ていたら「リンゴだね」と見ている物体の名前を Nao と共有できていることを確認するというタスクを与える。実験 2 の例を図 4(b) に示す。

実験 3 実験 2 と同様「Nao はテーブル上の物体の名前を知っている」という設定で行う。被験者には Nao が物体の名前を知っていることを伝え、Nao がリンゴを見ていたら「ミカンだよ」と Nao の視界外に見ていない物体があることを教えるというタスクを与える。実験 3 の例を図 4(c) に示す。

実験 4 「Nao はリアルタイムで学習する」という設定で行う。被験者には、実験 4 開始時点では Nao は物体の名前を知らないことを伝え、Nao に物体の名前を教えたり、物体の名前が Nao と共有できていることを確認したり、自由に発話するというタスクを与える。

これらの実験を 1 人の被験者に対し、10 回ずつ計 40 回行う。

今回の実験では、実験中のやりとりを記録しておくだけで学習は行わず、実験終了後に記録に基づき学習を行う方式(オフライン方式)を採用しているため、実際には Nao はリアルタイムで学習は行わない。実験 4 で「自由に発話する」というタスクを与えることで、被験者がどのような場面でもどのように終助詞を用いた発話をするかが観察できる。

3.4 学習

学習は 2 節で示した方法により行う。実験中は学習に必要な情報を時刻と共にデータとして保存し、実験終了後に学習を行う。

環境の状態 s 、Nao の取る行動 a 、得られる報酬 r について、それぞれ 3.4.1 節、3.4.2 節、3.4.3 節に示す。

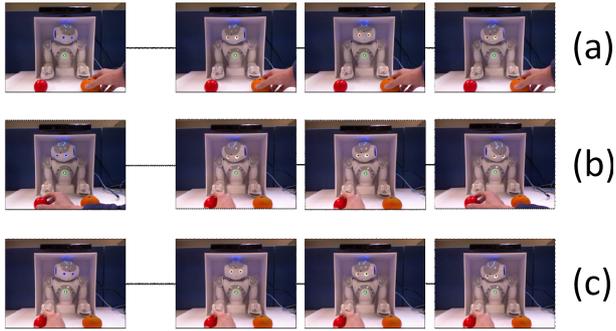


図 4: (a)「ミカン」を見ている Nao に対し、ミカンを把持して「ミカンだよ」と発話したときの Nao の不適切な行動例 (b)「リンゴ」を見ている Nao に対し、リンゴを把持して「リンゴだね」と発話したときの Nao の適切な行動例 (c)「ミカン」を見ている Nao に対し、リンゴを把持して「リンゴだよ」と発話したときの Nao の適切な行動例

3.4.1 状態

学習エージェントは環境から観測した情報に基づいた状態をもつ。環境から観測できる情報には以下のものがある。

被験者の発話

被験者の発話は、3.3.1 節に示したように、「[word]だ [final]」の形式である。この 4 通りの発話を音声認識の辞書に登録し、どの発話がなされたのかを判別する。

被験者が発話対象に選んだ物体 (= 発話対象物)

被験者が発話対象に選んだ物体が何であるかという判定は、3.3.1 節に示したように、発話時に発話対象物を右手で把持、もしくは指さしてもらうことで判定する。把持の判定、指さしの判定についてはそれぞれ 3.3.3 節と 3.3.4 節で説明した。

Nao は発話対象物の名前を知っているか

Nao が発話対象物の名前を知っているかについては、「知っている」か「知らない」の 2 通りのみとする。

被験者の発話時に Nao は発話対象物を見ているか

被験者の発話時に Nao が発話対象物を見ているかについては、3.2 節の手順 1. において Nao が顔を向けた物体と、手順 2. で被験者が把持・指さしていると判定した物体が同じであれば「見ている」、違っていれば「見ていない」とする。

今回行った実験では、「Nao は発話対象物の名前を知っているか」については、実験 1 では「知らない」、実験 2,3 では「知っている」となり、「被験者の発話時に Nao は発話対象物を見ているか」については、実験 1,2 では「見ている」、実験 3 では「見ていない」となる。

3.4.2 行動

状態 s で取る行動 a は、3.3.2 節で示した 84 通りの行動パターンからランダムに選ぶ。

3.4.3 報酬の与え方

状態 s で行動 a を取った結果得られる報酬 r について説明する。

報酬は被験者が Nao の行動を評価することにより得ることができる。評価は 3.2 節に示したように、「良い」「悪い」の 2 種類である。評価「良い」に対しては +1、評価「悪い」に対しては -1 の報酬を与える。

また、被験者は Nao が行動し始めてからエピソードが終わるまでの間、任意のタイミングで、何度でも評価できるので、ある行動パターンにおいて、1 番目の行動を始めた時刻から 2 番目の行動を始める直前までの評価を 1 番目の行動に対する評価、2 番目の行動を始めた時刻から 3 番目の行動を始める直前までの評価を 2 番目までの行動に対する評価、3 番目の行動を始めた時刻からエピソードが終わるまでの評価を 3 番目までの行動に対する評価とした。

本研究では報酬を、評価があった時点までの行動パターンのみを与える。

例として、Nao が 1 番目に A、2 番目に B、3 番目に C と行動する行動パターン ABC を選択し、それぞれの行動において「良い」「悪い」「悪い」と評価された場合について説明する。

まず、1 番目の行動 A に対する評価は「良い」であるので、1 つの行動だけからなる行動パターン A へのみ報酬 +1 を与える。

次に、2 番目の行動 B に対する評価は「悪い」であるので、2 つの行動だけからなる行動パターン AB へのみ報酬 -1 を与える。

最後に、3 番目の行動 C に対する評価は「悪い」であるので、3 つの行動からなる行動パターン ABC へのみ報酬 -1 を与える。

3.4.4 行動価値の更新

行動価値の更新には以下の式を用いる。

$$Q(s, a) \leftarrow Q(s, a) + \alpha(r - Q(s, a)) \quad (7)$$

4 実験結果と考察

4.1 実験結果

今回は工学系の男子大学生 5 名、女子大学生 3 名に対して実験を行った。実験データを基に学習した結果と考察

を以下に記述する。学習の更新係数 $\alpha = 0.5$ とした。データを基に学習を行い、行動価値が高い順に行動パターンを上位 20 個、下位 20 個抜粋した結果を、表 2, 表 3, 表 4 に示す。表 2, 表 3, 表 4 はそれぞれ実験 1, 実験 2, 実験 3 の結果である。ただし実験 1, 3 では下位 20 個のなかに評価が与えられていない (行動価値が初期値の) 行動パターンが含まれていたため、これは除外した。

実験 1 では、Nao が対象物を見ている状態で「うなずく」という行動を取ったとき、発話に含まれる普通名詞が対象物の名前であることを覚えたという意味と解釈され、評価「良い」が与えられると考える。また実験 2 でも同様に、Nao が対象物を見ている状態で「うなずく」という行動を取ったとき、Nao と対象物の名前を共有できていると解釈され、評価「良い」が与えられると考える。

表 2, 3 を見ると、実験 1, 2 では「対象物を見る」の直後に「うなずく」、もしくは「顔を見る」の直後に「うなずく」という行動パターン (表 2, 3 中の a) が上位に多く見られる。「対象物を見る」の直後に「うなずく」という行動パターンは、先に述べたとおり被験者の発話を理解したと解釈したためであると考えられる。対象物を見ている状態から「顔を見る」の後に「うなずく」という行動は、被験者の発話を理解したと解釈され、評価「良い」が与えられたと考えられる。Nao が見ている物体を発話対象物として発話するため、Nao が最初に「うなずく」(表 2, 3 中の b) という行動をとった場合も、実験 1 の場合は発話対象物の名前を覚え、実験 2 の場合は物体の名前を共有できていると解釈され評価「良い」が与えられたと考える。

下位では、「その他の物体を見る」(表 2, 3 中の c) という行動が多く見られる。これは、Nao が見ている物体を発話対象物とするのに「その他の物体をみる」という行動が不自然であると解釈されたためであると考えられる。

また実験 1, 2 に関しては被験者の発話時に Nao がすでに対象物を見ている状態であるため、「対象物を見る」(表 2 中の d) という行動から始まる行動パターンは、Nao が何も行動していないと被験者が錯覚した可能性があり、次に「うなずく」行動をとった場合には対象物を見ている状態であったため評価「良い」が与えられたが、連続して「対象物を見る」行動をとった場合には反応がないと解釈されたり、「その他の物体を見る」という行動を次にとった場合には反応が遅い上に適切でない行動をとったため評価「悪い」が与えられたと考えられる。

実験 3 では、Nao はエピソード開始時に発話対象でない物体を見ているため、被験者が発話対象に選んだ物体に視線を移すことが求められるため「対象物を見る」(表 4 中の e) という行動に対し評価「良い」が与えられると考えられる。

また、被験者が発話対象に選んでいない物体を見ている状態で「うなずく」行動をとった場合には評価「悪い」が与えられると考えられる。表 4 では、下位に見られる

表 2: 実験 1 での各行動パターンの行動価値

1 番目	2 番目	3 番目	行動価値
顔	対象物	対象物	0.875
顔	うなずく ^a	-	0.875
顔	うなずく ^a	対象物	0.875
うなずく ^b	対象物	対象物	0.875
うなずく ^b	その他	対象物	0.875
顔	顔	その他	0.750
対象物 ^d	その他	うなずく	0.750
対象物 ^d	うなずく ^a	対象物	0.750
その他	うなずく	対象物	0.750
うなずく ^b	その他	-	0.750
うなずく ^b	その他	その他	0.750
顔	顔	うなずく ^a	0.500
顔	対象物	うなずく ^a	0.500
顔	その他	うなずく	0.500
顔	うなずく ^a	顔	0.500
顔	うなずく ^a	その他	0.500
対象物 ^d	顔	-	0.500
対象物 ^d	顔	うなずく ^a	0.500
対象物 ^d	うなずく ^a	顔	0.500
対象物 ^d	うなずく ^a	その他	0.500
対象物 ^d	顔	その他 ^c	-0.250
対象物 ^d	その他 ^c	顔	-0.250
その他 ^c	うなずく	その他 ^c	-0.250
対象物 ^d	対象物	うなずく	-0.375
顔	顔	顔	-0.500
対象物 ^d	顔	顔	-0.500
対象物 ^d	その他 ^c	-	-0.500
その他 ^c	顔	-	-0.500
その他 ^c	対象物	-	-0.500
その他 ^c	その他 ^c	その他 ^c	-0.500
うなずく	対象物	その他 ^c	-0.500
うなずく	うなずく	その他 ^c	-0.500
うなずく	うなずく	うなずく	-0.500
その他 ^c	顔	うなずく	-0.625
対象物 ^d	対象物	その他 ^c	-0.750
その他 ^c	顔	顔	-0.750
その他 ^c	その他 ^c	-	-0.750

顔 … 顔を見る, 対象物 … 対象物を見る,
その他 … その他の物体を見る

最初の行動が「うなずく」である行動パターンがこれに該当する (表 4 中の f)。

実験 4 は、「自由に発話する」というタスクで、被験者がどのように終助詞を用いるか観察する目的であった。初めに名前を教え、次に教えた名前を確認し覚えていないようなら再度教え、覚えているようなら同意要求や存在を教える、といったように実験 1, 2, 3 で行った発話を順次していくことを想定していた。実験後、各被験者に実験 4 でどのように発話していったかをインタビューにより確認すると、それぞれほぼ想定通りの発話を行っていたことが分かった。表 5 に一被験者の例を示す。想定と違った発話では、以下の 2 つがあった。Nao が見えない物体に対し終助詞「ね」を用いて、同意を要求するという発話と、Nao が見ている物体を発話対象物とし、発話対象物と異なる名前と終助詞「ね」を組み合わせ同意を要求する発話とがあった。これらの発話もインタビューにより、Nao が本当に物体の名前を覚えたかを試すという意図があったことを確認した。

表 3: 実験 2 での各行動パターンの行動価値

1 番目	2 番目	3 番目	行動価値
うなずく ^b	顔	-	0.938
うなずく ^b	うなずく	うなずく	0.875
顔	顔	うなずく ^a	0.750
顔	うなずく ^a	うなずく	0.750
その他	顔	うなずく ^a	0.750
うなずく ^b	その他	対象物	0.750
顔	顔	その他	0.500
顔	対象物	対象物	0.500
顔	その他	その他	0.500
顔	うなずく ^a	その他	0.500
対象物 ^d	顔	対象物	0.500
対象物 ^d	対象物	うなずく ^a	0.500
対象物 ^d	うなずく ^a	-	0.500
その他	うなずく	その他	0.500
うなずく ^b	-	-	0.500
うなずく ^b	顔	顔	0.500
うなずく ^b	顔	対象物	0.500
うなずく ^b	顔	うなずく ^a	0.500
うなずく ^b	対象物	-	0.500
うなずく ^b	対象物	顔	0.500
顔	顔	-	-0.500
顔	対象物	顔	-0.500
顔	その他 ^c	-	-0.500
顔	その他 ^c	対象物	-0.500
対象物 ^d	-	-	-0.500
対象物 ^d	対象物	顔	-0.500
対象物 ^d	対象物	対象物	-0.500
対象物 ^d	その他 ^c	顔	-0.500
対象物 ^d	その他 ^c	対象物	-0.500
その他 ^c	-	-	-0.500
その他 ^c	顔	その他 ^c	-0.500
その他 ^c	対象物	-	-0.500
その他 ^c	対象物	対象物	-0.500
その他 ^c	対象物	その他 ^c	-0.500
その他 ^c	その他 ^c	対象物	-0.500
その他 ^c	その他 ^c	うなずく	-0.500
その他 ^c	うなずく	対象物	-0.500
その他 ^c	うなずく	うなずく	-0.500
対象物 ^d	顔	顔	-0.750
対象物 ^d	その他 ^c	その他 ^c	-0.875

顔 … 顔を見る, 対象物 … 対象物を見る,

その他 … その他の物体を見る

4.2 考察

今回は 8 名の被験者に実験を行い、各行動パターンの行動価値を学習した。実験 1,2,3 において、それぞれ筆者が想定していた通りの行動が上位または下位に概ね見られた。

今回の実験では、実験中には学習を行わず、実験後に適切な行動を学習したが、インタラクション中に学習するようにすれば、実験中にロボットのふるまいが変化し、被験者のふるまいや被験者が与える評価もそれに連れて変化することが予想され、ロボットの学習結果に影響を与える可能性がある。今後はこの影響を明らかにする実験も行っていきたい。

5 まとめと今後の展望

本研究では、発話に対する適切な行動の学習について、終助詞「よ」の教示用法と終助詞「ね」の同意要求用法を中心に、即時報酬課題での強化学習の評価実験を行っ

表 4: 実験 3 での各行動パターンの行動価値

1 番目	2 番目	3 番目	行動価値
顔	うなずく	対象物 ^e	0.875
顔	その他	-	0.750
対象物 ^e	うなずく	その他	0.750
その他	対象物 ^e	うなずく	0.750
顔	その他	うなずく	0.625
顔	顔	顔	0.500
顔	顔	対象物 ^e	0.500
顔	対象物 ^e	-	0.500
顔	対象物 ^e	顔	0.500
顔	うなずく	-	0.500
対象物 ^e	-	-	0.500
対象物 ^e	顔	うなずく	0.500
対象物 ^e	対象物 ^e	-	0.500
対象物 ^e	対象物 ^e	顔	0.500
対象物 ^e	その他	-	0.500
対象物 ^e	その他	対象物 ^e	0.500
対象物 ^e	その他	その他	0.500
対象物 ^e	うなずく	-	0.500
その他	対象物 ^e	顔	0.500
その他	対象物 ^e	対象物 ^e	0.500
対象物	その他	顔	-0.313
顔	対象物	その他	-0.500
顔	うなずく	その他	-0.500
その他	-	-	-0.500
その他	顔	その他	-0.500
その他	うなずく	その他	-0.500
うなずく ^f	-	-	-0.500
うなずく ^f	顔	うなずく	-0.500
うなずく ^f	対象物	-	-0.500
うなずく ^f	うなずく	-	-0.500
うなずく ^f	うなずく	うなずく	-0.500
その他	顔	顔	-0.750
その他	その他	顔	-0.750
うなずく ^f	顔	顔	-0.750
うなずく ^f	うなずく	その他	-0.750

顔 … 顔を見る, 対象物 … 発話対象物を見る,

その他 … その他の物体を見る

た。8 名の被験者で評価実験を行ったところ、概ね適切な行動を学習することができることが分かった。

今後の展望としては、インタラクション中に学習することでロボットの反応が変わり、その結果被験者の反応も変わることが予想されるため、インタラクション中に学習する実験を行い、そこで観察されるであろう動的な相互適応過程を分析したい。

また、現状では、終助詞の意味のうち、行動に関わる部分だけの獲得を試みたに過ぎないが、今後は、行動を獲得するだけでなく、発話や内部情報処理も含めた獲得を可能にし、終助詞の意味の全体を扱うことを目指したい。目に見える行動や耳に聞こえる発話の獲得と比べて、外部から直接観測できない内部情報処理を獲得することは難しいことが予想されるが、獲得した行動を手がかりに内部情報処理を獲得する（例えば発話対象物を見ることでそこに注意が向き、名前を覚えるといった内部情報処理を獲得する）ことでこの壁を乗り越えることができるのではないかと考えている。

謝辞

本研究は科研費 (21500137) の助成を受けた。

表 5: 実験 4 結果例

	Nao が		発話意図	評価
	見ている物体	被験者の発話		
1 回目	ミカン	ミカンだよ	名前を教える	悪い
2 回目	ミカン	ミカンだよ	名前を教える	良い
3 回目	ミカン	ミカンだね	名前を覚えたか確認	良い
4 回目	リンゴ	リンゴだよ	名前を教える	悪い
5 回目	ミカン	ミカンだね	名前を覚えたか確認	良い
6 回目	ミカン	ミカンだね	名前を覚えたか確認	良い
7 回目	ミカン	ミカンだよ	名前を教える	良い
8 回目	リンゴ	リンゴだよ	名前を教える	良い
9 回目	ミカン	ミカンだね	名前を覚えたか確認	良い
10 回目	リンゴ	リンゴだよ	名前を教える	悪い

参考文献

- [1] 岩橋直人. “ロボットによる言語獲得—言語処理の新しいパラダイムを目指して”. 人工知能学会誌, Vol. 18, No. 1, pp. 49–58, 2003.
- [2] 田口亮, 岩崎直人, 船越孝太郎, 中野幹生, 能勢隆, 新田恒雄. “統計的モデル選択に基づいた連続音声からの語彙学習”. 人工知能学会論文誌, Vol. 25, No. 4, pp. 549–559, 2010.
- [3] 野々口直宏. “発話に対する適切な行動としての意味獲得：終助詞「よ」「ね」を中心として”. 京都工芸繊維大学 卒業論文, 2012.
- [4] 野々口直宏, 深田智, 尾関基行, 岡夏樹. “発話に対する適切な行動としての意味獲得：終助詞「よ」「ね」を中心として”. 人工知能学会第 2 種研究会 ことば工学会資料, SIG-LSE-C003, pp. 45–55, 2012.
- [5] 金水敏. “<言語学の最新情報>—日本語学：終助詞ヨ・ネ”. 月刊言語, Vol. 22, No. 4, pp. 118–121, 1993.