

遠隔による第二言語コミュニケーション訓練のための ロボットアバター言語学習支援システムにおける視線情報の分析

Analysis of Gaze Information in Robot-Avatar-Assisted Language Learning system for Remote Learning

名島大生^{1*} 寺本龍生¹ 加藤恒夫¹ 田村晃裕¹ 山本誠一¹
Taisei Najima¹ Ryusei Teramoto¹ Tsuneo Kato¹ Akihiro Tamura¹ Seiichi Yamamoto¹

¹ 同志社大学大学院理工学研究科

¹ Graduate School of Science and Engineering, Doshisha University

Abstract: We have been developing a robot-assisted language learning system that uses two humanoid robots to let a second language learner join in English conversation between them and practice his/her communicative skill by him/herself in an immersive environment. However, all kinds of learning are expected to be given remotely due to the spread of COVID-19. We developed a robot-avatar-assisted language learning system that simulates conversations with the robots appearing on a PC through the Internet. We conducted a 6-day straight experiment with eight Japanese university students to measure the effect of learning and eye-gaze as an index of the learners' concentration on the system. The experimental results showed the effectiveness of learning with the system, and that the learners' eye-gaze activities changed between two tasks: repeating and question answering.

1 はじめに

近年のグローバル化により英語でのコミュニケーションの機会は増え、その重要性は増している [1]. 第二言語の学習はインストラクタとの 1 対 1 形式が最も効果が高いと言われる [2] が、実際には経済的、時間的な面で難しいことが多く、英語の授業や英会話教室での会話訓練が一般的である。そこでも生徒一人ひとりに十分な訓練の時間を与えることは簡単ではない。第二言語 (L2) でのコミュニケーション能力の向上を支援するツールとして、コンピュータと音声対話を行うコンピュータ支援言語学習 (Dialogue-based Computer-Assisted Language Learning, DB-CALL) システムの研究開発が行われてきた [3]. しかし、従来の CALL システムは必ずしもリアルなエージェントが用いられず、身振りや視線動作などの非言語情報が欠落していることがあった。そうすると、実際の英会話のような臨場感や緊張感を感じさせることが難しい。

当研究室では臨場感のある英会話訓練を実現するために、人型ロボット 2 体を用いた Joining-in-type Robot-Assisted Language Learning (JIT RALL) システムを提案してきた。2 体のロボットが学習者の前で英会話を

展開し、学習者の参加を促す。ロボット間の英会話は日常的な話題を取りあげながらも、例えば無生物主語の使い方や時制の使い分けなど特定の表現形式に焦点を当てる。学習者はロボットの中のやりとりを参考にして質問に回答することで、表現形式の習得が促進されることを狙いとしている。人型ロボットを用いることで、臨場感を与えながらも、人間ではない相手に発話することで誤りをおかすことに対する過剰な恐れを軽減する効果も期待できる [4]. これまでの研究で、様々な表現形式の習得効果を測定してきた。その中で、質問への回答を繰り返し発声させることで焦点を当てた無生物主語の正しい使用が増加するとともに、発声中のフィールドポーズが削減される効果があることを確認した。また、学習者個人の能力に応じた訓練内容を設定するため、訓練の途中で聞き取り能力を測定し、英会話の難度を変更する取り組みも行ってきた。

しかしながら、2020 年は新型コロナウイルス感染拡大のため実験室での実験が困難になった。また今後、様々な学習がリモート化される社会的な動きもある。本研究では、遠隔でも英会話訓練ができるようにバーチャル空間に 2 体のロボットアバターを作成した。学習者の集中度を確認するために視線追跡システムを追加して小規模な英会話訓練実験を行い、表現形式の習得効果と視線の分布を分析したので報告する。

*連絡先: 同志社大学大学院理工学研究科
〒 610-0394 京都府京田辺市多々羅都谷 1-3
E-mail: ctwf0126@mail4.doshisha.ac.jp

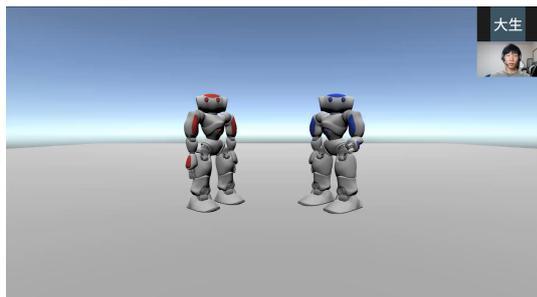


図 1: ロボットアバターによる言語学習支援システム

2 ロボットアバターによる言語学習支援システム

本システムは、ロボットアバター 2 体の対話によって焦点を当てた表現形式を学習者に提示し、その習得を促す。図 1 に示すように、学習者 (Learner) 1 名に対して教師役アバター (A1) と生徒役アバター (A2) を配置し 3 人対話を模擬する。ロボットアバターの発話が誰に向けたものであるかを明確にするため、発話中ロボットアバターの顔や手を聞き手に向けてるように制御している。ロボットアバターのモデリングは Blender を、動作と音声の制御は Unity を用いた。

このシステムではあらかじめ用意したシナリオに沿って、ロボットアバター同士が対話を行う。英語教育に適したシナリオを作成するため英語を母国語とするネイティブ英語教師と共同でシナリオを作成した。また、ロボットアバターの音声には、ネイティブ英語教師の音声を使用した。

このシステムでは主に質問・回答タスクを通して学習を進める。質問・回答タスクでは、ロボットアバター同士の対話の中で A1 による質問と A2 による模範的な回答例を学習者に提示する。続いて A1 が同様の質問を学習者に行うことで学習者に適切な発話を促す。この A1 による質問と A2 による模範的な回答例は 1 つのシナリオで 5 つずつ提示され、それにともない学習者には回答の機会が 5 回与えられる。このシステムで用いるロボットアバターと学習者間の対話の一部を表 1 に示す。

焦点を当てた表現形式の定着化を図るため、学習者は質問に対して同一の回答を 2 回繰り返す。学習者が聞き返しを要求する発話をした場合、A1 は最大 3 回まで同じ質問内容を繰り返し発話する。学習者が質問に対する回答が思いつかなかった場合には "I don't know" と発話してもらうことで次の会話へと進む。

第二言語の訓練では、学習者の言語運用能力に応じた訓練内容の設定が必要とされるが、その判断基準を今回高難度グループで焦点を当てる無生物主語を用いた文を正しく聞き取れるか否かとし、高難度グループ、

表 1: 高難度シナリオ例 (一部)

話者	発話
A1	There are many attractive places in Japan. What places make you happy in Japan?
A2	Kinkakuji temple makes me happy in Japan.
A1	I see. What places make you happy in Japan?
Learner	...

低難度グループの 2 つのグループに振り分ける。学習者の聞き取り能力を評価するために A2 の模範的な回答例を復唱するリピーティングタスクを設けた。リピーティングタスクでは、2 つのシナリオを用意しており、各シナリオで模範的な回答例が 5 つずつ提示される。模範的な回答例の文章に含まれる単語を 60% 以上復唱できれば聞き取れていると判断し、模範的な回答例を 6 つ以上聞き取れている場合に無生物主語の使い方に焦点を当てる高難度グループ、それ以外を現在完了形と時制の使い方に焦点を当てる低難度グループへと振り分ける。

現在は、Zoom の画面共有機能を用いて学習者の PC ディスプレイに映像を投影し、音声認識の失敗による対話破綻を防ぐために Wizard-of-Oz 法により制御している。しかし、今後音声認識精度が向上し、対話破綻の恐れがなくなれば学習者はこのシステムを用いて一人での言語学習が可能である。

A1, A2, 学習者の発話、身体動作は Zoom のレコーディング機能を用いて収録している。

3 視線データ収集

学習者の視線データは Tobii 社製の Tobii Pro ナノを用いて収集した。同製品は PC に装着型のスクリーンベースアイトラッカーでグラスを装着する必要がないため、学習者への負担が少ない。このアイトラッカーを用いてサンプリング周波数 60Hz で画面上の視点の位置を取得した。

視線データは各シナリオの開始時刻から終了時刻まで記録した。

学習者の集中度を調査するためにロボットアバターへの注視率を測定した。測定方法として、取得した視線の各サンプル点についてロボットアバターと重なる場合は注視していると判定し、重ならない場合は注視していないと判定した。そして、学習者の集中度の指標として各シナリオにおけるロボットアバターの注視時間の割合である注視率を求めた。

表 2: 実験スケジュール

グループ	day1	day2	day3	day4	day5	day6
高難度	T1	N1s	N2s	N3	N4	N5
	N1	N1	N2	N3d	N4d	N1
低難度	T1	N1s	N2s	T1	T2	T3
	N1	N1	N2	T1d	T2d	T1

表 3: 回答レベルとその基準

Level	概要
A	A-1 適切な表現かつ文法としても正しい
	A-2 適切な表現ながら文法誤りが含まれる
B	質問に応じた回答だが R2 の回答に類似しない
C	C-1 質問に対して見当違いな回答
	C-2 無回答または”I don't know” 等の回答

4 実験

4.1 実験方法

20 歳から 23 歳の日本語を母語とする日本人大学生 8 名を対象に実験を行った。実験協力者に対して、1 日に 10 分程度のシナリオを 2 つ行い、6 日間 Wizard-of-Oz 法によりシステムを制御する。6 日間の実験スケジュールを表 2 に示す。

学習者はロボットアバター間の対話内容を正確に聞き取る必要があるため、学習者の聞き取り能力に合わせてグループ分けを行う。学習者の聞き取り能力を評価するために、期間中 2 日目及び 3 日目に A2 による模範的な回答例を復唱するリピーティングタスク (N1s, N2s) を設けた。リピーティングタスクでは、2 つのシナリオを用意しており、各シナリオで模範的な回答例が 5 つずつ提示される。模範的な回答例の文章に含まれる単語を 60% 以上復唱できれば聞き取れていると判断し、模範的な回答例を 6 つ以上聞き取れている場合に高難度グループ、それ以外を低難度グループへと振り分ける。高難度グループでは無生物主語の使い方に焦点を当てる N(non-animate subject) シナリオ、低難度グループでは過去形と現在完了形の使い方に焦点を当てる T(tense) シナリオを用いて実験を進める。N および T の後に続く数字ならびに記号 s と d はそれぞれシナリオの種類、リピーティングタスク、部分的に変更したシナリオを表す。学習効果を定量的に評価するため、それぞれのグループで期間中 1 日目と 6 日目にプレテスト、ポストテストとして同一のシナリオ (T1, N1) を実行した。

今回の試行実験の実験協力者 8 名は、高難度グループに 5 名、低難度グループに 3 名振り分けられた。

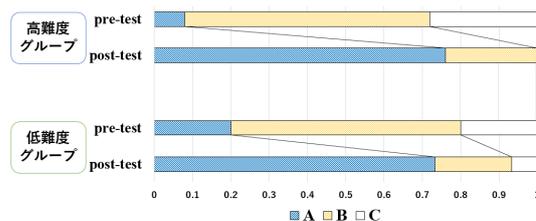


図 2: 各グループにおける 3 段階の回答レベルの割合

表 4: 各グループにおける A1, A2 への注視率

	A1 (%)	A2 (%)	合計 (%)
高難度	34.9	20.1	55.0
低難度	26.2	17.4	43.6

4.2 学習効果の評価

プレテスト、ポストテストにおける実験協力者の 2 回目の回答に対して表 3 の基準に基づき、主観評価を行った。表 3 より、適切な表現が用いられているときの学習者の回答はレベル A に振り分けられる。そして、質問を理解しているが、適切な表現が用いられていないときの学習者の回答はレベル B に、質問を理解できていないときの学習者の回答はレベル C に振り分けられる。評価値の相対度数分布を図 2 に示す。レベル A の発話割合は高難度グループでプレテストの 8.0% からポストテストの 76.0% に 68.0% 増加し、低難度グループでプレテストの 20.0% からポストテストの 73.3% に 53.3% 増加した。また、レベル C の発話割合は高難度グループでプレテストの 28.0% からポストテストの 0.0% に 28.0% 減少し、低難度グループでプレテストの 20.0% からポストテストの 6.7% に 13.3% 減少した。

4.3 視線データの分析

収集した視線データを用いて学習者ごとに各シナリオにおける A1 と A2 への集中度を計測した。まずは、グループ間における集中度の違いを測るために、各グループにおける A1 と A2 への注視率の平均値を算出した。その結果を表 4 に示す。低難度グループに比べて高難度グループの方が A1 を注視している割合は 8.7% 高く、A2 を注視している割合は 2.7% 高かった。

次に、シナリオごとに学習者の視線の変化を調査した。高難度グループ、低難度グループそれぞれにおける 12 種類のシナリオ別の A1 と A2 への注視率の平均の推移を図 3, 図 4 に示す。各図において赤字で示した N1s と N2s はリピーティングタスクであり、その 2 点だけ A1, A2 それぞれに対する注視率が逆転している。タスク別に集計した A1 と A2 への注視率の平均を表 5 に示す。高難度グループにおいて、質問・回答タスク

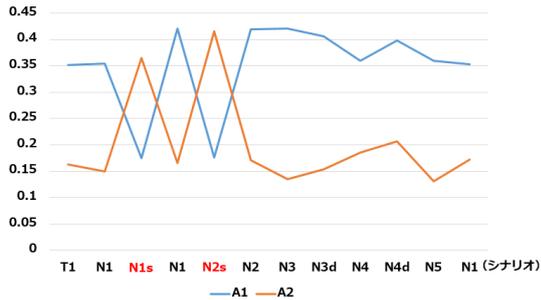


図 3: 高難度グループの各シナリオにおける 2 体のロボットアバターに対する注視率の推移

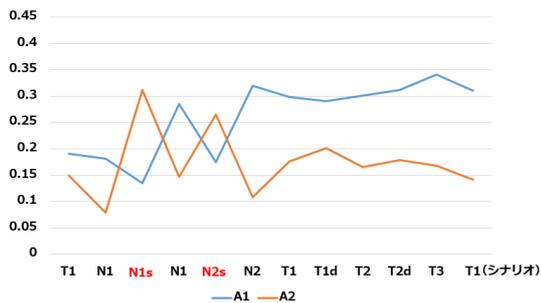


図 4: 低難度グループの各シナリオにおける 2 体のロボットアバターに対する注視率の推移

では A2 より A1 を注視している割合は 22.1%高いのに対してリピーティングタスクでは A1 より A2 を注視している割合は 21.4%高かった。低難度グループでも同様に質問・回答タスクでは A2 より A1 を注視している割合は 13.2%高いのに対してリピーティングタスクでは A1 より A2 を注視している割合は 13.4%高かった。

5 考察

低難度・高難度の両グループにおいてレベル A の割合が大幅に増加したことから、システムを使い続けることで学習者に適切な表現が定着化したことが考えられる。また、両グループでレベル C の割合が大幅に減少したことから、システムを使い続けることによって英語の聞き取り能力も向上したと考えられる。

質問・回答タスクでは A1 への注視率が高く、リピーティングタスクでは A2 への注視率が高いことから、注視率は学習者の集中度の指標となりえ、学習者はロボット間の対話よりも自分の回答に直接影響のある A1 からの質問や A2 からの模範的な回答例に集中していると考えられる。一方で、質問・回答タスクで A2 への注視率が低いことから、質問・回答タスクでは模範的な回答例への集中が必ずしも高くないことが考えられる。

表 5: 各タスクにおける A1, A2 への注視率

	タスク	平均 (%)		SD (%)	
		A1	A2	A1	A2
高難度: 5 名 (全 12 シナリオ)	質問・回答	38.4	16.3	13.2	4.6
	リピーティング	17.6	39.0	9.4	11.7
低難度: 3 名 (全 12 シナリオ)	質問・回答	28.3	15.1	10.3	5.4
	リピーティング	15.5	28.9	3.1	9.3

6 まとめ

バーチャル空間にロボットアバターを作成し Zoom の画面共有機能を用いることでリモートでの第二言語学習を可能にした。8 名の実験協力者に 6 日間試用してもらった結果、高難度、低難度の両グループで適切な表現 (レベル A) の割合を大幅に増加させた。

視線分析の結果、高難度グループの実験協力者の方がアバターの注視率が高かった。また、質問・回答タスクでは教師役を演じる A1 を注視しているのに対して、リピーティングタスクでは生徒役を演じる A2 を注視しており、タスクによって集中する発話が変化する様子が分かった。

今後は、実験協力者を増やすことで学習効果を統計的に検証するとともに、収集した視線データを細かい時間区切りで分析し、A1 の質問時や A2 の模範的な回答例の提示時、学習者の回答時の視線動作の変化を調査したい。

謝辞

本研究の実験環境構築に協力していただいたネイティブ英語教師の皆様へ感謝いたします。本研究は科研費 19K00927 の支援を受けた。

参考文献

- [1] 文部科学省: 今後の英語教育の改善充実方策について, https://www.mext.go.jp/b_menu/shingi/chousa/shotou/102/houkoku/attach/1352464.htm.
- [2] B. S. Bloom: The 2 sigma problem: The search for methods of group instruction as effective as one-to-one tutoring, *Educational researcher*, pp.4-16, (1984).
- [3] S. Bibauw et al: Discussing with a computer to practice a foreign language: research synthesis and conceptual framework of dialogue-based CALL, *Computer Assisted Language Learning* 32(8), pp.1-51, (2019)
- [4] C. Chang, J. Lee, P. Chao, C. Wang, G. Chen, et al: Exploring the possibility of using humanoid robots as instructional tools for teaching a second language in primary school, *Ed. Technol. Soc.*13, pp.13-24, (2010).
- [5] A. Khalifa et al: Joining-in-type Humanoid Robot Assisted Language Learning System, *Proc.LREC*, pp.245-249, (2016)