

遠隔による第二言語コミュニケーション訓練のための アバター言語学習支援システムの試作

Prototyping of avatar-assisted language learning system for remote training of communication in second language

田中龍弥[†] 山本誠一[‡] 加藤恒夫[†]
Ryuya Tanaka Seiichi Yamamoto Tsuneo Kato

1. はじめに

現代社会では英語が国際的なコミュニケーションの主要言語として広く利用されている。しかし、第二言語を英語とする多くの学習者は、実用的なコミュニケーションスキルを身に付けるには至っていない。この問題に対処するため、対話型 CALL システムが開発されている[1, 2]。第二言語の音声コミュニケーション習得には、文法知識を自動的に運用できるシステムティックな訓練が必要である。しかし、学校の外国語授業では生徒一人ひとりに十分な訓練を提供することが難しい。当研究室では、少人数での会話訓練を模した Joining-in-type Robot Assisted Language Learning (JIT-RALL) システムを提案してきた[3]。同システムは模擬会話を聞いてその応答を繰り返すリピーティング(RP)とは異なり、2体のロボットによる模擬会話で使用される表現形式を借用して自らの内容を表現する質問応答訓練(QA)を採用している。

既存の JIT-RALL システムは対面形式で訓練を行うため、学習者は訓練のために足を運ぶなどの負担がある。さらに、多くの学習者に同時に訓練が困難である。これらの問題に対処するため、ロボットに代わってアバター2体により英会話を行うと共に Windows のリモートデスクトップ機能を利用して遠隔で学習者に英会話訓練を提供するシステムを構築した。本稿では、システム機能の記述と共に小規模な英会話訓練により得られた印象評価について報告する。

2. Joining-in-type ロボット支援言語学習

JIT-RALL システムは、教師を演じるロボット R1 と生徒役を演じるロボット R2 の 2 体のロボットの会話に学習者が参加して英会話を進めることで、表現形式の習得を目指す。会話シナリオは、日本人には運用が難しい表現形式である無生物主語構文や過去形と現在完了形の使い分け等に焦点を当てた表現形式の習得を促進する。具体的にはロボット間の会話の中で R1 が質問を提示し R2 が模範解答を提示した後、R1 が同様の質問を学習者に提示する。学習者は R2 の回答での表現を用いて自分自身の回答を表現するように求められる。日本語では余り使用されない無生物主語構文による表現形式の訓練では従来の RP より QA のほうが訓練後の定着効果が高いことが示されている[4]。

3. アバター言語学習支援システム

3.1 システムのインターフェース

本システムは、アバター2体の対話によって JIT-RALL システムと同様に特定の表現形式に焦点をあてた表現を質問応答で学習者に提示し、学習者が回答することでその習得を促す。図 1 に示すように、教師役アバターと生徒役アバターをバーチャル空間に配置して学習者と 3 人で対話する。

本システムは、予め用意した会話シナリオに沿って、アバター同士が会話を行う。会話シナリオは、対象とする学習者が興味を有するように同年代である英文科の博士後期課程の院生により作成され英語母語話者の校正を経た後、英語母語話者により対面で発話された音声を録音して会話素材として使用される。本システムで使用するアバターと学習者間の対話の一部を表 1 に示す。

身体性を有するロボットと異なり画面に表示されるアバターによる発話が誰に対して向けられているかの明確さが減少する。そのため、聞き手が発話者に向くように発話者は手を動かすなどの話す動作を加えることで動きの制御をおこなった。さらに、アバターの発話内容の冒頭を視覚的に表示し、学習者にどのアバターが話しているかをより明確に示すことを目的として吹き出し機能を追加した。

アバターのモデリングは Blender、動作と音声の制御は Unity を用いた。3種類のアバター対を開発したが今回の実験では図 1 に示すアバター対を用いた[5]。また、将来的には学習者の発話を分析・評価して応答を返す機能を組み込むため、質問に回答した音声を音声認識装置が認識した発話内容等を画面上に表示させる機能も追加した。

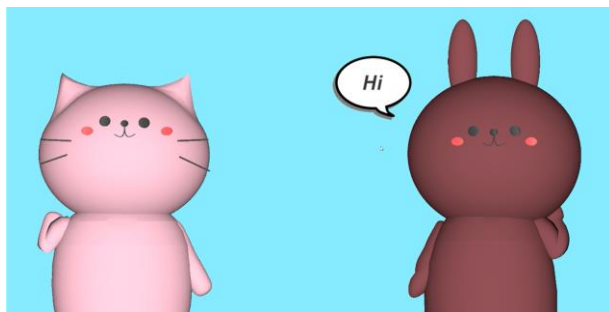


図 1 アバター言語学習支援システムの表示画面

表 1 会話シナリオ例

話者	発話
R1	Recently, many foreigners are interested in Japanese culture.
R1	What attracts foreigners to Japanese culture?
R2	Japan's long history attracts foreigners.
R1	What attracts foreigners to Japanese culture?
学習者	(learner's own answer) (QA)
	Japan's long history attracts foreigners. (RP)

3.2 システム構成

本システムでは、図 2 に示すようにシステムに学習者自身の PC からリモートデスクトップ機能を使用してアクセスする。その後、アバター間の会話を視聴し、続くアバターからの質問に対して音声発話を行うと、アバターから発話に対する応答が返る。

[†]同志社大学院 理工学研究科

Graduate School of Science and Engineering, Doshisha University

[‡]同志社大学研究開発推進機構

Organization for Research Initiatives and Development, Doshisha University

システムの全体制御は Unity で行われている。Unity には、あらかじめ録音された音声ファイルとアバターの 3 次元データが搭載されている。会話の進行に伴い 2 体のアバターが動作する。さらに、吹き出しもアバターの音声が出されるタイミングで表示される。ただし、吹き出しは会話の流れの理解を補助する目的であるため、最初の挨拶以降、アバターの発話はテキストとして表示されない。学習者の音声の認識には、Python プログラムで制御される Whisper が使用されている。学習者が回答する音声データが Unity から Python で記述されたプログラムに送られると、Whisper が呼び出され音声認識が行われる。音声認識が完了すると Unity へ音声認識の結果を文字として返す。学習者への発話の要求から一定時間内に音声入力がない場合、発話要求の過程を飛ばし、次の会話へ自動的に進む。

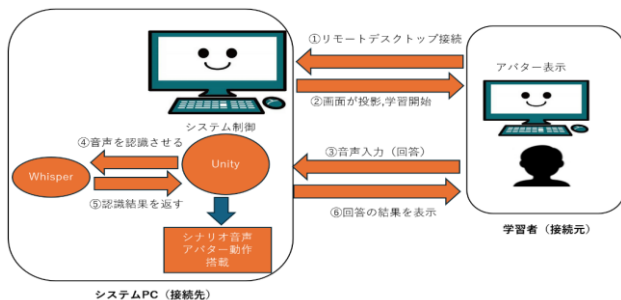


図 2 システム構成

4. 実験

4.1 学習者

日本語の母語話者である 20 歳から 22 歳の同志社大学理工学部生を対象に本システムの使用に関する評価実験を実施した。実験は、通常のアバターシステムと吹き出し機能を追加したシステムを 5 人ずつ、合計 10 人を対象に行った。

4.2 実験方法

実験では、リモートデスクトップを利用して約 20 分間のシナリオで英会話訓練を体験し、その後アンケート調査を行った。アンケートの目的は、アバターを使用した英会話訓練に関する使い勝手等に関する評価を収集することである。評価方法は先に JIT-RALL システムで実施された手法に準じたが、アバターの動作やシステム全体の使いやすさ、吹き出し機能の有無による違いを評価項目として追加している。

4.3 実験結果

図 3 に吹き出しの有無に対する比較評価、図 4 にシステムに対する主観評価を示す。図 3 に示すように、吹き出しありとなしのアバターの動作の印象評価は、どの評価項目でも吹き出しありがアバターの話者交替を明確に区別して認識できていることを示している。アバターが動作していることに好印象を持っている人が多いことも分かった。

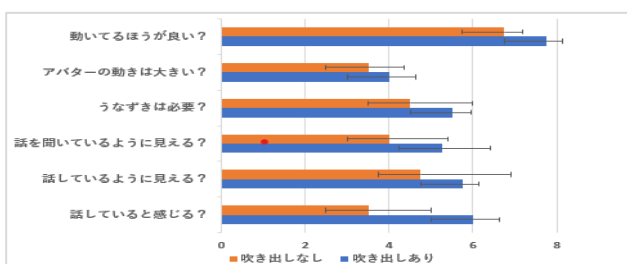


図 3 吹き出しありと無しの比較評価

システムに関する主観評価には比較対象がないので主観的な評価に留まるが、システム全体の動作に対する評価は良好であり、表現習得のシステムとして適切と評価されていると思われる。一方、会話の流れをつかめたという評価に関しては、比較的低い評価に留まっており聴取力の観点から話速の制御等に関する検討が必要である。

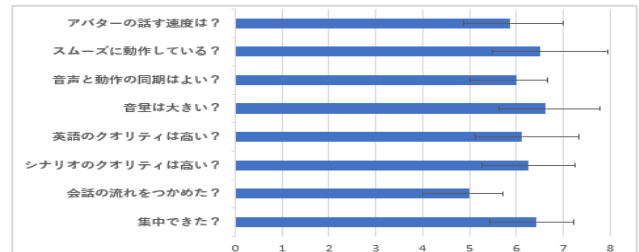


図 4 システムに関する主観評価

5. 考察

吹き出しありの会話は、吹き出しなしの会話と比べて全般に評価が高く、会話交替等の動作の認識に寄与していることが分かった。システムからの音声は男声と女声からなっているが、最初の発話時点で吹き出しの表示によりどちらが発話しているのか理解できるからであると推測される。吹き出しなしではアバターの基本的な動作が発話を明示するには不十分であり、発話時の動作の制御の向上を図る必要があると思われる。システムの総合的な評価に関しては、音声や会話の進行、アバターの動作の同期などが一定の品質を保っており、JIT-RALL システムによる訓練をリモート環境で適用可能と思われる。

6. まとめ

本講ではアバター言語学習支援システムを構築と主観評価実験について報告した。本システムは JIT-RALL システムと比較してロボットの持つ身体性がないことを吹き出し等で補うことにより会話の円滑性に関しては特段の不满はないことが示唆されている。今後大規模な実験により本システムの訓練効果を検証すると共に学習者の発話に対する応答等の機能について検討を進める計画である。

謝辞

本研究は科研費 23K24962 の支援を受けたものである。

参考文献

- [1] Bibauw, S., François, T., & Desmet, P., “Discussing with a computer to practice a foreign language: research synthesis and conceptual framework of dialogue-based CALL” *Computer Assisted Language Learning*, 31(7), 827-877, 2018.
- [2] H. Noh, K. Lee, S. Lee G. G. Lee, “POMY: a conversational virtual environment for language learning in POSTECH”, *Proc. SIGDIAL’11, Association for Computational Linguistics, USA*, 344-346, 2011.
- [3] A. Khalifa, T. Kato, S. Yamamoto, “Joining-in-type Humanoid Robot Assisted Language Learning System,” *Proc. LREC 2016*, pp. 245-249, 2016.
- [4] J.D. Muramoto, T. Kato, A. Tamura, S. Yamamoto, “Retention Effects of Form-focused English Speaking Training Based on Question-Answering Task in Comparison with Repeating Task using Robot-assisted Language Learning System”, *Proc. 9th Workshop on Speech and Language Technology in Education (SLaTE)*,
- [5] <https://booth.pm/ja/items/4634700> (うさぎ VRM アバター)