

# eラーニングにおけるリスナエージェントを利用した スピーチベーステスト

出店 優一<sup>†</sup> 木村 英雅<sup>†</sup> 林 潤平<sup>†</sup> 長谷川 大<sup>†</sup> 佐久田 博司<sup>†</sup>  
青山学院大学 理工学部 情報テクノロジー学科<sup>†</sup>

## 1 はじめに

従来, eラーニングは主にテキストベースの記述式, 選択式の評価方法を用いている. これは初期段階の知識形成評価には適しているが, 知識利用の習練度や応用力評価には短時間でより多量のアウトプットを課すことのできるスピーチベースのインタラクションによる評価方法が有効であると考えられる. しかし, eラーニングで自学する学習者にとってはPC画面に対して発話するよりも, 話しかける対象が存在する方が自然であると考えられ, またその対象が学習者のスピーチを聞いていると知覚されることにより, テストに対する没入感やそのようなテストを前提とした学習への集中度が向上するだろう.

本稿では, eラーニングにおける知識利用の習練度や応用力評価を想定し, 擬人化エージェントを聞き役として利用したブラウザで動作するスピーチベースのテストシステムを提案する. また, 本テストシステムを利用したeラーニングにおいて, リスナエージェントの聞き手動作の有無による学習態度及びスピーチ態度の変化を調査する.

## 2 聞き手動作

Maatman ら [1] は人間同士の対話において聞き手が頻繁に行う動作を以下に記す3種類に分類した.

- Backchannel:** 聞き手が肯定, 理解していることを示す頷き, あいづち等の話し手の発話中に行う動作
- Disfluency:** 首を傾げる, 顔をしかめる等の話し手の言い淀みや聞き取りが困難な場合に行う動作
- Mimicry:** 発言内容の復唱や同じ方向を見るといった話し手のまねをする動作

また, Maatman らはこれらの聞き手動作を導く話し手の発話内容や顔の動きをトリガーとし, 聞き手動作を



図1 システムUI

生成するリスナエージェントを開発した. Poppe らは Backchannel を詳細に分類し, リスナエージェントに実装した [2]. 発話内容や顔の動きとは別に視線の位置や表情をトリガーとした研究もされている [3].

しかしこれらのような聞き手動作を行うリスナエージェントが eラーニングにおけるテストに利用された例はなく, またブラウザで動作するシステムも存在しない.

本稿で開発するスピーチベーステストシステムは一般的なeラーニングでの利用を考え, HTML5, JavaScript, WebGL で開発され, Google Chrome 上で動作する. また, ここではリスナエージェントの聞き手動作として, Backchannel にあたる「頷き」によるあいづちと Disfluency にあたる「首を傾げる」動作による発話の促しを実装する. 聞き手動作のトリガーには学習者の発話ポーズを用いた. 具体的には, 学習者の発話中の声量が設定した閾値以下になった時間を学習者のスピーチポーズとして検出し, 0.5 秒の短いスピーチポーズを検出した際には頷きを生成し, 6 秒を越える明らかに不自然な長いスピーチポーズが検出された際には首を傾げる動作を生成する.

## 3 実験

本実験ではリスナエージェントの聞き手動作を要因として設定し, 下記の2条件による1要因2水準被験者間実験計画とした.

Speech Based Test by Using a Listener Agent in E-Learning

<sup>†</sup> Yuichi Demise, Hidemasa Kimura, Jumpei Hayashi, Dai Hasegawa, Hiroshi Sakuta

Department of Integrated Information and Technology,  
College of Science and Engineering, Aoyama Gakuin University

条件1 聞き手動作を生成するリスナエージェント

条件2 聞き手動作を生成しないリスナエージェント

参加者は情報系学科の大学生を対象とし、Java プログラミングに関する専門用語の解説をしたテキストを学習コンテンツとする。

各参加者は、以下の順で e ラーニングを行う。

1. 3 問のスピーチベーステストを実施
2. Java プログラミングに関するテキストの学習
3. 1 のテストを再実施

なおテストには 1 問あたり 15 秒の思考時間と 45 秒の発言時間を設定した。

評価には、没入度を示すための指標として、テスト中の学習者の発話時間、学習にかけた時間を計測する。

我々は聞き手動作を行うエージェントが存在することでシステムへの没入度が高まり、より解答に積極的になると仮説を立て、積極的に解答するために解答時間は伸び、集中して学習をするために学習時間は短くなると考えた。

#### 4 実験結果および考察

ボランティアにより参加した情報工学系学科所属大学生 18 名を 9 名ずつ 2 条件に割り当てた。聞き手動作を使用するリスナエージェント条件の参加者 7 名の映像データを用いて本システムの聞き手動作の生成精度を計測したところ、ショートポーズに対しては適合率 79 %、再現率 87 %、F 値 0.82 で頷きを、ロングポーズに対しては適合率 91 %、再現率 91 %、F 値 0.91 で首を傾げる動作をそれぞれ適切なタイミングで生成していた。実験結果を以下に示す。

2 条件のスピーチテストにおける平均発話時間の比較を図 2 に示す。結果、聞き手動作を実装したエージェント条件は平均発話時間が長い傾向にあることがわかった。

解答時の発話時間が長くなる傾向にあるならば、参加者は聞き手動作による発話の促しによってテストに積極的に取り組んでいたと推測でき、スピーチベースの e ラーニングに聞き手動作を利用することで没入度の高いスピーチテストを提供できる可能性があると考えられる。

一方、システム中の学習時間は表 1 のようになった。聞き手動作を生成するリスナエージェントを用いた方が学習時間が短くなっている。このことから聞き手動作によって、リスナエージェントによるスピーチテストを前提とした学習への没入感・集中度が高まったと推測できる。

しかし、2 水準間の平均発話時間及び平均学習時間に有意差は見られなかった。標本数の不足が原因の可能性もあるため、今後実験を継続し、標本数を増やして改めて検証を行う必要がある。

表 1 2 条件による学習時間の比較

条件	学習時間平均 (秒)
エージェント+聞き手動作	230
エージェントのみ	353

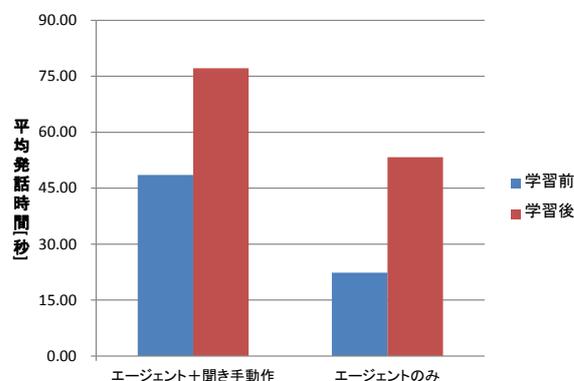


図 2 2 条件による発話時間の比較

#### 5 おわりに

本稿では e ラーニングにリスナエージェントを利用したスピーチベーステストを組み込むことを提案・試作し、聞き手動作が学習に与える影響の調査を行った。

実験の結果、e ラーニングにおけるスピーチベーステストに聞き手動作を行うリスナエージェントを用いた場合、聞き手動作を実装していないエージェントを用いた場合と比較して解答時の発話時間は長く、学習にかかる時間は短くなる傾向にあることがわかった。しかし、被験者の数が不十分であると考えられるため、実験データを増やして再検証する必要がある。

提案システムは、筆記よりも発話が重要とされる英会話等の言語学習に応用することでより効果が発揮されると期待できる。

#### 謝辞

本研究は JSPS 科研費若手研究 (B) 25870698 の助成を受けたものです。

#### 参考文献

- [1] Stacy Marsella R. M. Maatman, Jonathan Gratch. *Natural Behavior of a Listening Agent*. Intelligent Virtual Agents Lecture Notes in Computer Science, 2005.
- [2] Dennis Reidsma Ronald Poppe, Khiet P. Truong and Dirk Heylen. *Backchannel Strategies for Artificial Listeners*. Intelligent Virtual Agents Lecture Notes in Computer Science, 2010.
- [3] E. Cowie R. Eyben F. Gunes H. Heylen D. etal Schroder, M. Bevacqua. *Building Autonomous Sensitive Artificial Listeners*. Linguistics, 2011.