

AI を活用した次世代型英語スピーキング評価法の開発に向けて －質問紙調査の結果について－

山下 仁司¹、東 淳一²、仁科 恭徳²

要旨

AI による人間との双方向コミュニケーションを英語学習に応用する動きは盛んになりつつあるが、実際にどの程度何に効果があるかを確認するため、英語学習者を対象に AI ロボット Musio を活用した実証実験を行った。その結果、人間ではなくロボットを対象とした方が心理的バリアは低い、英語学習は繰り返しが重要とする信念など、4 因子が抽出され、3 タイプの学習者像が存在する事がわかった。

キーワード：英語教育、AI ロボット、質問紙調査、多変量解析、テキストマイニング

Abstract

In recent years, as learning English using Artificial Intelligence has become popular, it is important to find out what and how such method is good for. We conducted an experiment which some of our students were asked to practice English conversation with AI robot MUSIO and collected opinions of the students using questionnaire. As the result, we found four factors such as “low affective barriers in conversation with robots” or “faith in importance of repeating in English practice”, etc., and classified the students into 3 types.

Key Words: English teaching and learning, AI robot, questionnaire, multivariate analysis, text mining

1. はじめに

本稿では、科学研究費助成事業における研究課題「AI を活用した次世代型英語スピーキング評価法の開発に向けて」（代表者：東淳一 基盤研究（C）18K00809）において、日本人大学生英語学習者を対象に実施した AI ロボット Musio を活用した英語授業に関する質問紙調査（主に 5 件法、3 件法、自由記述）の結果について詳細に報告する¹。本プロジェクトの目的、現在までの流れについての詳細は仁科・東・山下（2021）を参照されたい。

¹ 大阪大学

² 神戸学院大学

2. 質問紙調査について

2.1. 実験協力者について

仁科・東・山下（2021）では、大学生英語学習者を対象に AI ロボット Musio を活用した実証実験の詳細とスピーキング・データの自動取得に関して、現時点での問題点とその改善方法等について報告した。この実証実験の終了後に、同参加者に質問紙調査を実施した。仁科・東・山下（2021）においても既述のとおりであるが、質問紙調査の対象者について整理すると、英語学習に対する意欲が高く、英語圏文化に興味のある大学 2・3 年生計 17 名である。実験協力者の英語力は、イギリス・ピアソン社が開発した VERSANT English Speaking Test（20～80 点の範囲）の平均点が 38.57 点（CEFR 換算で A 2 レベル）、最低点は 32 点（CEFR 換算で A 1 レベル）、最高点は 50 点（CEFR 換算で B 1 レベル）と全体的に英語力は初級レベルであると言える（参照 <https://www.versant.jp/score.html>）。他に、参加者と紐づけることのできる英語力テスト結果や成績データには、大学の授業で実施した英語会話に関するプレゼンテーション試験スコアがある。第 6 章で、これらのスコアを活用した分析も実施する。

2.2. 質問紙の内容

質問紙調査の期間は、2020 年 1 月 7 日から 15 日とし、Google Form を用いて実施した。質問内容は以下の 15 項目である。

表 1. 各質問項目の内容

	質問内容
1	Musio を使って話しているとき、たくさん話そうと思いましたか？また、その理由も書いてください。
2	Musio を使った会話は英語の勉強になったと思いますか？また、その理由も書いてください。
3	Musio を使った会話を体験して、今後もっと英語の勉強をしようと思いましたか？また、その理由も書いてください。
4	Musio を使って会話練習をくり返し行ったら、英語の勉強になると思いますか？また、その理由も書いてください。
5	Musio と人間の先生、どちらの方が会話をしやすいと感じましたか？また、その理由も書いてください。
6	Musio と人間の先生、今後、どちらの方と会話練習をしたいと思いましたか？また、その理由も書いてください。
7	Musio との会話で学習したことを、今後、英語を使う場面で活かすことができると思いますか？また、その理由も書いてください。
8	Musio との会話で良かったことを挙げてください。複数挙げてもかまいません。
9	Musio との会話で困ったことを挙げてください。複数挙げてもかまいません。
10	Musio を使った学習のメリットを挙げてください。複数挙げてもかまいません。
11	Musio を使った学習のデメリットを挙げてください。複数挙げてもかまいません。
12	Musio との会話をしやすかった話題は何でしたか？また、その理由も書いてください。
13	Musio との会話をしにくかった話題は何でしたか？また、その理由も書いてください。

- 14 Musio のシステムは使いやすいものでしたか. 自由に教えてください.
15 これからも機会があれば Musio を使った学習を受けてみたいと思いますか? また, その理由も
答えて下さい.
-

3. 質問紙調査結果の分析手法

5 件法の質問項目 (項目 1~4, 7, 14, 15) については, 5=全面的に賛同する, 4=条件付きで賛同する, 3=中立, 2=条件付きで反対する, 1=全面的に反対するとし, 日を空けて 2 回換算した上で違いの生じた個所を再度検討し整えた。項目 8~13 は, メリット・デメリットを尋ねる項目であることから, 数値変換は行っていない。一方で, 項目 5 と 6 については, 1=人間, 3=どちらともいえない, 5=MUSIO とした。

次に, 質問項目のうち, クローズド・クエスチョンについては回答内容を元に, 1~5 段階で数量化した。数量化にあたっては, 1 回目の結果を見ずに日を変えて変換し直し, 違いが見られた箇所を再検証し客観性を高めた。また, オープン・アンサーの回答内容については, 全て目視で確認し, 肯定的・否定的に分類して同趣旨の回答を集約し, 箇条書きにまとめた。

数量化可能なデータの分析については, 基礎集計 (5 段階項目の平均, 各回答の分布, α 係数, 回答平均値の男女差の検定), 探索的因子分析, クラスタ分析 (因子負荷量を元に, 学生を 3 タイプに分類), クラスタタイプごとの回答平均値, クラスタタイプごとの英語力 (Versant, 授業成績), クラスタタイプごとの男女比の検定を実施した。また, クラスタタイプごとにテキストマイニングを実施した。

なお, 探索的因子分析では 100 名以上を目安にするなど, ある程度のサンプルサイズの確保が必要であるとの指摘があり, (繁樹・柳井・森, 2008; 竹内・水本, 2014), 現時点で確保できたサンプルサイズは 17 名で, 確かに十分な分析対象とは言えない。一方で, 各因子内の因子負荷量が高ければ (具体的には, ある因子につき負荷量が 0.6 を超える変数が 4 つ以上存在している場合) それほど多くのサンプルが必要ではないという議論もある (Guadagnoli & Velicer; 1988)。後述の通り, 抽出された 4 因子の因子負荷量には 0.6 を超える物がそれぞれに 1 設問以上存在し, Guadagnoli らの基準を十分満たしているとは言えないが, ある程度は参考にできると思われる。また, 結果的に被験者の思考パターンのバリエーションとして納得性のある結果となっており, このまま分析を進める。こののち, 機会があればこの結果を裏付けるよりサンプル数を増やした実験を行う予定である。

4. 分析の基礎項目

はじめに各質問項目の数量化の結果を表 2 に, 回答状況分散グラフを図 1 に示す。

表2. 各質問項目の基礎データ

	項目1	項目2	項目3	項目4	項目5	項目6	項目7	項目14	項目15	計	平均	性別	Versant 点数	英会話 授業得点	クラスター
A	4	5	5	1	1	1	5	5	4	31	3.44	m	41	35	Type 1
B	5	3	2	4	1	1	3	1	1	21	2.33	f	50	37	Type 2
C	5	5	5	5	1	5	5	5	5	41	4.56	m	33	19	Type 3
D	5	2	5	5	1	3	4	4	4	33	3.67	m	40	20	Type 3
E	1	5	5	3	1	1	5	1	5	27	3.00	m	37	35	Type 1
F	5	1	5	5	1	1	5	5	3	31	3.44	m	49	26	Type 2
G	1	3	5	1	1	1	1	5	2	20	2.22	f	なし	なし	Type 2
H	5	5	5	5	1	1	5	2	1	30	3.33	f	33	29	Type 2
I	4	4	3	5	1	1	5	5	5	33	3.67	m	なし	なし	Type 1
K	5	5	5	5	5	3	5	5	5	43	4.78	f	34	32	Type 3
L	5	5	5	5	1	1	5	5	3	35	3.89	m	36	20	Type 2
M	5	5	5	5	1	5	5	5	5	41	4.56	f	34	37	Type 3
N	5	5	5	2	1	5	5	3	4	35	3.89	m	39	28	Type 3
O	1	4	5	4	5	1	4	5	4	33	3.67	m	未受験	12	Type 1
P	2	5	5	5	1	5	1	5	4	33	3.67	f	32	27	Type 3
Q	5	5	5	5	1	1	4	2	5	33	3.67	f	40	29	Type 1
R	5	5	5	4	1	1	2	5	2	30	3.33	m	42	30	Type 2

*A~R は回答者を示す。

AI を活用した次世代型英語スピーキング評価法の開発に向けて

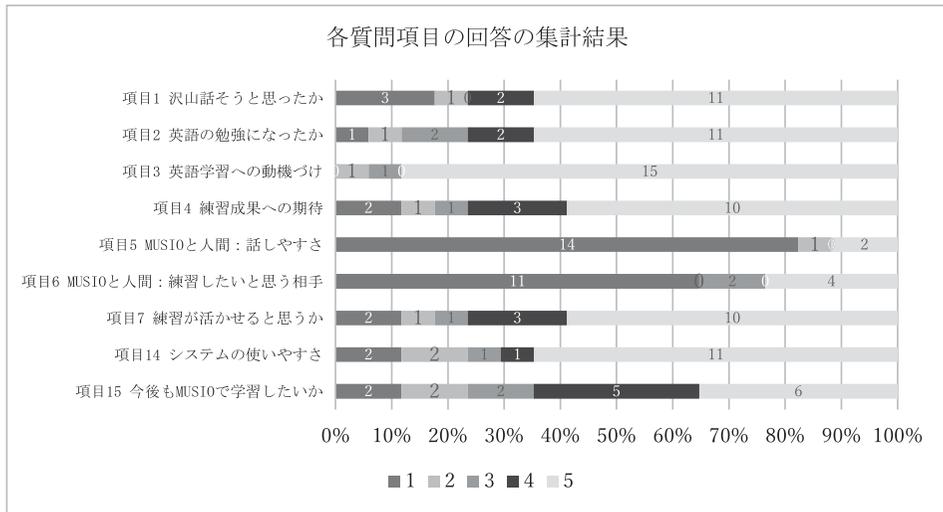


図 1. 各質問項目の回答の集計状況

以下、表 3 では、15 の質問項目における具体的な理由や意見の例を順に示す。

表 3. 質問紙調査における自由記述回答の例

	肯定的意見 (一例)	否定的意見 (一例)
項目 1. Musio を使って話しているとき、たくさん話そうと思いましたか？また、その理由も書いてください。	AI との会話が興味深いから / 自分の英語力を上げるために、たくさん話そうと思った / 英語を話す機会が授業以外にあまりないので、多く話そうと思った / なかなか意思疎通ができなかったので、結果的にたくさん話した	会話にならなかったため、話す気にならなかった。 / MUSIO が勝手に話の腰を折ってしまうので、話す気が失せた
項目 2. Musio を使った会話は英語の勉強になったと思いますか？また、その理由も書いてください。	リスニングの練習になった / 英語を話す機会が日常生活では少ないので、話す訓練にはなった / 会話を成立させるために、色々な事を言おうとしてそれが練習になった / MUSIO に認識させるために、発音 (連結・脱落等) に留意した / ネイティブと話しているときも同じような感じになっているかもしれないと危機感を持った	コミュニケーションが成立しなかった / MUSIO が同じことばかり言う
項目 3. Musio を使った会話を体験して、今後もっと英語の勉強をしようと思いましたか？また、その理由も書いてください。	リスニング力をつけたいと思った / 語彙力をつけたい / スピーキング力をつけたい / 流暢さを身につけたい 様々な話題について話す事 / 会話が成立しない事を、自分の能力不足に帰属させ学習動機・危機感につながった	(MUSIO でなくても) ほかの学習方法もあるので
項目 4. Musio を使って会話練習をくり返し行ったら、英語の勉強になると思いますか？また、その理由も書いてください。	会話・インタラクションをする機会を増やせる、繰り返しすることで身につく / MUSIO から、色々な表現を学べる / 会話をする自信が身につく / スピーキングで、音声認識されることが増加している事が、自分の発音がネイティブに近づいていると思える	会話が成立しないので、練習にならない / 日常会話は良くても、敬語などは身につかないのでは / やや思うが、他にも有効な手段はある

項目 5. Musio と人間の先生、どちらの方が会話をしやすいと感じましたか？また、その理由も書いてください。	人間の先生と話す緊張する / 人間の先生と話す機会が少ない / 気を使わなくていい	人間はこちらの言いたいことをくみ取ってくれたり、インタラクションの調節してくれる / 表情などのノンバーバルコミュニケーションがある / 聞き返せる / 発音が下手でも聞き取ってくれる
項目 6. Musio と人間の先生、今後、どちらの方と会話練習をしたと思いますか？また、その理由も書いてください。	ノンバーバルコミュニケーションがない分、英語力をシビアに上げられる / MUSIO に聞き取ってもらえれば、よりネイティブに近づいたと思える(人間の教師は手加減がある) / MUSIOの方が緊張しなくていい / 今後のAIの発達を考えると、MUSIOなどでの学習も経験しておきたい	人間の方が会話が自然 / 人間との会話の方が止まったりせず、ストレスが少ない / 感情、抑揚などは人間からでないと学べない / 発音が自然 / 人間の先生は、間違いを修正してくれる
項目 7. Musio との会話で学習したことを、今後、英語を使う場面で活かすことができると思いますか？また、その理由も書いてください。	(できると) 思います。話題の振り方などは実際でも使えるからです。 / はい。自らアウトプットするうえで非常に役に立ちます。 / できると思います。AIとはなすことはなかなかないのでいい経験になったからです。 / ある。リスニングを強化できた。 / はい。様々な質問をすることで会話内容が豊富になったから。 / できると思います。この体験がより自分のことを表現しなければならぬ時や学んだことを思い出すよい復習になるからです。 / 思います。なぜなら会話は会話しただけ練習になるとおもうからです。 / できると思う。いろいろなトピックで会話できたから。 / 各トピックからそれぞれの表現を用いたので、活かすことができると思います。	なかなか前に進めないことです。一度詰まったらすれ違いが続いてしまうからです。 / 多少。あまり会話のラリーができなかった。
項目 8. Musio との会話で良かったことを挙げてください。複数挙げてもかまいません。	AI、ロボットと話ができた / (感情がないので) 話しかけやすい / 会話のトピック、テーマが豊富、返しが面白い / 自分の話したことが通じる喜び	
項目 9. Musio との会話で困ったことを挙げてください。複数挙げてもかまいません。		会話に通じない / 何度も同じことを聞かれる / ラリーを続けるのが難しい / 話がそれる、いきなり違う話題を振られる / 難しい表現を使われてわからない事がある
項目 10. Musio を使った学習のメリットを挙げてください。複数挙げてもかまいません。	話す、聞く機会の増加、人間が居なくても会話練習が気軽にできる / 間違っても恥ずかしくないので、気軽に話せる、人目を気にせずに練習ができる / (反面教師的に) 英語教師のすごさが実感できる(できない者に対応するのがストレスがかかることがわかる)	
項目 11. Musio を使った学習のデメリットを挙げてください。複数挙げてもかまいません。		機械的問題(準備が面倒、調子が悪い場合がある) / Musioの話が早くて聞き取れない場合がある / 会話が成立しない場合がある(AIの対話能力がまだ未発達) / Musioが正しく聞き取ってくれない(自分の英語が合っているかわからない)
項目 12. Musio との会話をしやすかった話題は何でしたか？また、その理由も書いてください。	スポーツ、趣味、家族、出身地、旅行、季節の話題などの、比較的身近で単純な話題 理由: 学生が話せる内容を多く持っていること、MUSIO が正確に答えられること	

AI を活用した次世代型英語スピーキング評価法の開発に向けて

<p>項目 13. Musio との会話をしにくかった話題は何でしたか？また、その理由も書いてください。</p>		<p>家族や友達の話。言ったことを繰り返したり、言ったことと違うことを言っていたから。／ 話題はあまり覚えていないが、話を深くしすぎると、Musio が違う話を急に入れるので、しづらかった。／ 世界の問題とか難しい話は聞き取りにくかったです。／ Musio の家族が複雑すぎて、「家族について」が会話しにくかったです。／ 好きな季節や趣味の質問。あいまいな答えや、あるかないかしが答えない。</p>
<p>項目 14. Musio のシステムは使いやすいものでしたか。自由に答えてください。</p>	<p>システムはわかりやすく、スマホレベルであった</p>	<p>反応が鈍い時があった / 面倒くさい等の意見もあり</p>
<p>項目 15. これからも機会があれば Musio を使った学習を受けてみたいと思いますか？また、その理由も答えて下さい。</p>	<p>新しい学習法だから / 留学から帰った時、より上達した時の自分の実力を試したい / かわいいから / より多く話す機会を持ちたい</p>	<p>ほかに効率的な方法がある / より良い機械がありそう / 人間と話す方が好きだから</p>

まず、質問項目 1 への回答内容から、英語を話す機会が（英語会話）の授業以外ではあまりなかったことから、Musio とたくさん話そうと思ったこと、Musio との意思疎通がうまくいかなかった（逆に）結果としてたくさん話したことなどの意見があった。質問項目 2 の回答内容からは、Musio を活用することで、話す訓練になったと実感している学生や、リスニングの練習になったと実感している学生がいることも分かった。特に、自分の言いたいことを Musio に認識させるために、いつも以上に発音に気をつけたと答えた学生もいた。質問項目 3 の回答内容からは、Musio との会話を通して、参加者はスピーキング力、リスニング力、語彙力、そして多様なトピックについて英語で会話できる話題力を身に付けたいと感じたようであった。質問項目 4 の回答内容からは、Musio と会話することで色々な英語表現が学べたこと、会話・インタラクションの機会が増えたことから自信がつき、自分の発音が上達していると実感できたなど、肯定的意見が多かった。質問項目 5 の回答内容から、Musio は機械であるため気を使う必要がなく緊張度が下がったこと、ただし話しかけた内容を汲み取ってくれないことがあること、非言語コミュニケーションが使えないことなどを挙げていた。質問項目 6 の回答内容から、人間の教師には手加減というものがあるが、Musio は機械であることからそのような気遣いが全くないこと、今後の AI の発達を見越すと Musio との学習も経験すべきであると感じること、人間との会話の方が自然でストレスが少ないこと、対人間の方が発音が自然であることなど、Musio と人間を比較した場合に、それぞれに一長一短があることが分かった。

次に、質問項目 7 への回答内容から、Musio との会話を通してリスニング力・アウトプット力が強化でき、話題の振り方が学べて会話内容の充実化が促進されるなど、肯定的な意見が大半を占めた。また、質問項目 8、9、10、11 の回答内容から、AI ロボットと意思疎通が

できる喜びや、AI ロボットならではの話しかけやすさがある一方で、会話の持続が難しく時に話題から話が逸れるなどの否定的な意見も見られた。そして、質問項目 12 の回答内容から、Musio との会話のトピックが予め豊富に揃っていること、また、各トピックについて Musio が正確に会話のやりとりができる能力も有していることに利点を感じている学生もいた。質問項目 13 の回答内容から、Musio の返答内容に問題や誤りがある場合や、Musio による急な話題転換によってコミュニケーションのやり取りが上手くいかないケースがあったことも分かる。項目 14 の回答内容から、学生にとって Musio のシステム自体は分かりやすかった一方で、反応の鈍さを指摘する学生もいた。最後に質問項目 15 の回答内容から、Musio の外観の可愛さや、英語を話す機会を多くもてる点から今後も Musio を使ってみたいという意見があったが、一方では、Musio よりも優れた AI ロボットがあるのではないかという期待感や、やはり人間とのコミュニケーションをより好むことから AI ロボットの活用には消極的であるというコメントも見られた。

次に、数量化が可能な質問項目間の相関係数をまとめた表 4 を参照されたい。質問項目 1 においては質問項目 4 と 7 との間に中程度の相関がみられた。同様に、質問項目 3 と 14、6 と 15、7 と 15 においても中程度の相関が見られたが、他の項目間において顕著な相関は見られなかった。また、今回の回答の平均値（5 段階評価）を算出しウェルチの t 検定を用いて男女差を比較したが、差は見られなかった ($t(7) = 0.40, p = .70$)²。なお、Cronbach α 係数の値が 0.8 を下回っていることから、各質問項目の回答内容には必ずしも一貫性があるとは言えないようである。これは、質問項目の内容自体にばらつきが認められることに起因すると思われる。

表 4. 数量化可能な質問項目間の相関係数と α 係数

	項目 1	項目 2	項目 3	項目 4	項目 5	項目 6	項目 7	項目 14	項目 15
項目 1	1	—	—	—	—	—	—	—	—
項目 2	.000	1	—	—	—	—	—	—	—
項目 3	-.136	.246	1	—	—	—	—	—	—
項目 4	.430	-.043	-.088	1	—	—	—	—	—
項目 5	-.203	.108	.147	.048	1	—	—	—	—
項目 6	.177	.267	.249	.171	.038	1	—	—	—
項目 7	.430	.131	.066	.211	.147	.021	1	—	—
項目 14	-.025	-.065	.334	.057	.214	.233	-.113	1	—
項目 15	-.082	.298	.273	.165	.239	.434	.412	.230	1

Cronbach's $\alpha = .588$

5. 因子分析

次に、因子分析の結果をまとめる。まず、共通性の推定について初期値は最も一般的な重相関係数二乗値 (squared multiple correlation: SMC) を用いた (石川, 2009, p.32)。因子パターンの推定には、近年主流になっている推測統計学手法の最尤法を用いた (竹内・水本, 2014, p.167)。反復回数の上限を 200 回としたが、8 回で収束した。図 2 は固有値のスクリープロットを示す。因子数の決定にはガットマン基準に従い、固有値 1 以上は第 4 因子までであり、第 5 因子以降は再度値が落ち込んでいることが図 2 から分かるので、抽出因子数は 4 とした。なお、因子の回転は、近年教育心理学などの分野で主流となっており (竹内・水本, 2014, p.167)、因子負荷行列のコントラストが際立つプロマックス法を用いた。表 6 の因子構造行列は、各質問項目と回転後の因子との相関係数を示す。表 7 は因子パターン行列である。斜交回転の結果から、これら二つの生成された表に基づき抽出された 4 つの因子は、第 1 因子が「ロボットの方が学習しやすいと感じている」、第 2 因子が「英語発話量・繰り返しの効果に対する信奉」、第 3 因子は「Musio との学習が実践的・役立つと思う」、第 4 因子は「現状限界はあるが、今後の Musio に期待」である。

表 5. 固有値表

因子	初期解			抽出後			回転後
	固有値	寄与率	累積寄与率	固有値	寄与率	累積寄与率	因子構造の平方和
1	2.2309	24.79%	24.79%	1.1687	12.99%	12.99%	1.6253
2	1.7626	19.58%	44.37%	1.4907	16.56%	29.55%	1.2653
3	1.1893	13.21%	57.59%	1.2917	14.35%	43.90%	1.3157
4	1.1305	12.56%	70.15%	0.8026	8.92%	52.82%	1.0763
5	0.8111	9.01%	79.16%				
6	0.7043	7.83%	86.99%				
7	0.5482	6.09%	93.08%				
8	0.4458	4.95%	98.03%				
9	0.1773	1.97%	100.00%				

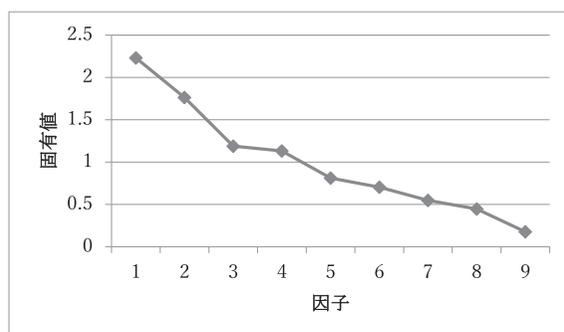


図 2. 固有値のスクリープロット

表 6. 因子構造行列

	変 数	因子 1	因子 2	因子 3	因子 4
項目 1	沢山話そうと思ったか	0.1238	0.5618	0.4852	-0.2497
項目 2	英語の勉強になったか	0.3152	-0.0346	0.1182	0.1681
項目 3	英語学習への動機づけ	0.2940	-0.0923	0.0350	0.1602
項目 4	練習成果への期待	0.1266	0.8708	0.1842	0.0915
項目 5	MUSIO と人間：話しやすさ	0.0144	0.0224	0.0604	0.2772
項目 6	MUSIO と人間：練習したいと思う相手	0.9730	0.3573	0.0061	-0.0795
項目 7	練習が活かせると思うか	0.1404	0.2115	0.9538	0.3359
項目 14	システムの使いやすさ	0.2473	0.0605	-0.1478	0.1502
項目 15	今後も MUSIO で学習したいか	0.6167	0.0715	0.3095	0.8562

表 7. 因子パターン行列

	変 数	因子 1	因子 2	因子 3	因子 4
項目 1	沢山話そうと思ったか	-0.0020	0.4278	0.4603	-0.2964
項目 2	英語の勉強になったか	0.3287	-0.1304	0.0905	0.0890
項目 3	英語学習への動機づけ	0.3261	-0.1711	0.0153	0.0928
項目 4	練習成果への期待	-0.1547	0.9544	-0.0538	0.2441
項目 5	MUSIO と人間：話しやすさ	-0.0423	0.0744	-0.0156	0.2955
項目 6	MUSIO と人間：練習したいと思う相手	0.9810	0.0891	-0.0798	-0.1784
項目 7	練習が活かせると思うか	0.0149	0.0319	0.9154	0.1369
項目 14	システムの使いやすさ	0.2301	0.0693	-0.2266	0.1788
項目 15	今後も MUSIO で学習したいか	0.5022	0.0181	0.0810	0.7754

6. 因子得点に基づくクラスター分析など

次に、因子得点を利用して、分析に用いる変数を抽出した因子 1~4 を指定し、クラスター分析（非階層型クラスタリング手法の k-means 法、クラスターの個数は 3、反復回数の上限は 20）を行った結果、3 つのタイプの学習者が存在している事が分かった。

表 8 と図 3 を参考に図 4 から解釈すると、Type 1 に属する実験協力者は、Musio との会話があまりスムーズに進まなかったため項目 1 では否定的で、項目 6 においても現状は人間を相手にしたいと思っているが、項目 15 で AI ロボットの将来性を買っている。Type 2 に属する実験協力者は、基本的に人間とのコミュニケーションを好み、AI ロボット否定派と言える。Type 3 に属する実験協力者は、英語上達のためには、発話量を増やすことが必要

であると考え、そのための相手として AI ロボットを肯定的に捉えており、その将来性も買っている。

総括すると、Type 1 は AI ロボットとの学習・未来性に期待を感じているが、現状は会話が続きにくいことなどから、やはり人間との練習が良いと感じている集団で「限定的肯定派」と言える。Type 2 は、ジェスチャーなどのノンバーバルコミュニケーションや文脈をくみ取ってくれることから、元々人間との練習の方が良いと思っている集団で「人間重視・ロボット学習否定派」である。Type 3 は、Musio に先進性を感じ、多少不自然なやりとりでも一人で英語の発話機会を増やせることから、Musio の活用を肯定的に見ている集団で「発話量重視・ロボット学習肯定派」と言える。AI ロボットが相手だと緊張感が少なく、間違っても良いとする Affective Barrier の低さなども良い点として挙げている。

表 8. 3 種のクラスターとサンプル数

	Type 1	Type 2	Type 3	計
因子 1	-0.406	-0.877	1.215	
因子 2	-0.634	0.005	0.524	
因子 3	0.271	-0.277	0.051	
因子 4	1.120	-0.908	-0.025	
<i>N</i>	5	6	6	17
%	29.41%	35.29%	35.29%	100.00%

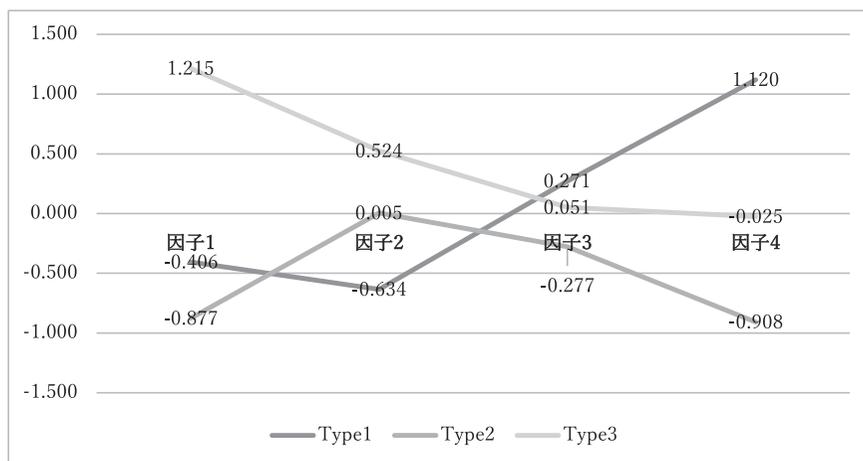


図 3. 各クラスタータイプの中心

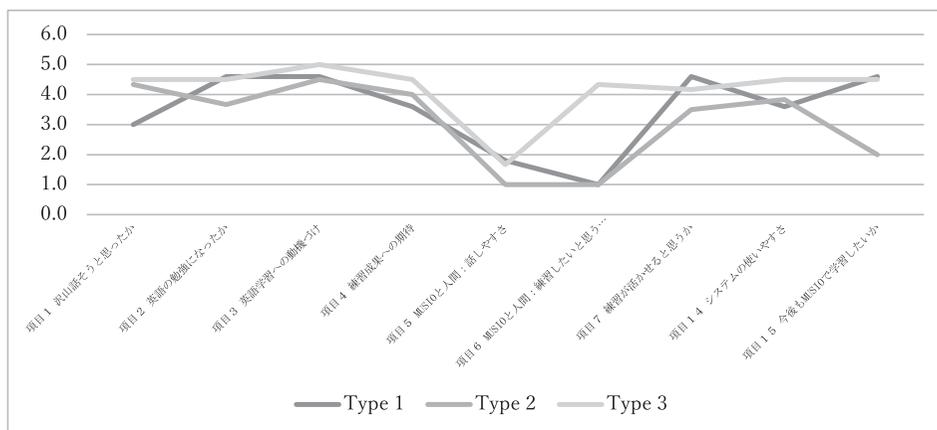


図4. クラスター分析によるタイプ別各項目の平均値

なお、参考までに各タイプに学力差や男女差は認められなかった。具体的には、クラスター分析による男女別タイプ割合の独立性の検定の結果 ($\chi^2=1.311, d=2.00, p=0.519, \text{Cramer's } V=0.278$)、男女で特にどのタイプが多いという偏重性は見られなかった。また、表9、図5、表10からANOVAによる検証によりVersantスコアの学力差はタイプ間で認められなかった。同様に表11、図6、表12からANOVAによる検証により英語会話授業スコアの学力差についてもタイプ間では認められなかった。

表9. 基本統計量 (Versant スコア、範囲は20~80点)

クラスター	n	平均	標準偏差 (SD)	平均-SD	平均+SD	標準誤差 (SE)	平均-SE	平均+SE
Type 1	3	39.333	2.082	37.252	41.415	1.202	38.131	40.535
Type 2	5	42.000	7.583	34.417	49.583	3.391	38.609	45.391
Type 3	6	35.333	3.327	32.007	38.660	1.358	33.975	36.691

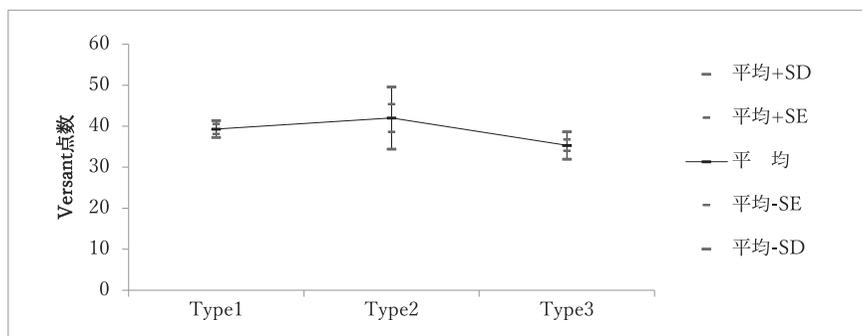


図5. 各水準の平均値 (クラスター)

表 10. VERSANT スコアによる分析

手法	水準 1	水準 2	平均 1	平均 2	差	標準誤差	統計量	P 値
Fisher の 最小有意差法	Type 1	Type 2	39.3333	42.0000	2.6667	3.7755	0.7063	0.4947
	Type 1	Type 3	39.3333	35.3333	4.0000	3.6556	1.0942	0.2972
	Type 2	Type 3	42.0000	35.3333	6.6667	3.1305	2.1296	0.0566
Scheffe	Type 1	Type 2	39.3333	42.0000	2.6667	3.7755	0.2494	0.7835
	Type 1	Type 3	39.3333	35.3333	4.0000	3.6556	0.5986	0.5665
	Type 2	Type 3	42.0000	35.3333	6.6667	3.1305	2.2676	0.1498
Bonferroni	Type 1	Type 2	39.3333	42.0000	2.6667	3.7755	0.7063	1.0000
	Type 1	Type 3	39.3333	35.3333	4.0000	3.6556	1.0942	0.8917
	Type 2	Type 3	42.0000	35.3333	6.6667	3.1305	2.1296	0.1699
Sidak	Type 1	Type 2	39.3333	42.0000	2.6667	3.7755	0.7063	0.8710
	Type 1	Type 3	39.3333	35.3333	4.0000	3.6556	1.0942	0.6529
	Type 2	Type 3	42.0000	35.3333	6.6667	3.1305	2.1296	0.1604

*学力差はどの手法を使っても認められなかった。

表 11. 基本統計量（英語会話授業 output スコア）

クラスター	n	平均	標準偏差 (SD)	平均 - SD	平均 + SD	標準誤差 (SE)	平均 - SE	平均 + SE
Type 1	4	27.750	10.874	16.876	38.624	5.437	22.313	33.187
Type 2	5	28.400	6.189	22.211	34.589	2.768	25.632	31.168
Type 3	6	27.167	6.911	20.255	34.078	2.822	24.345	29.988

*40 点満点

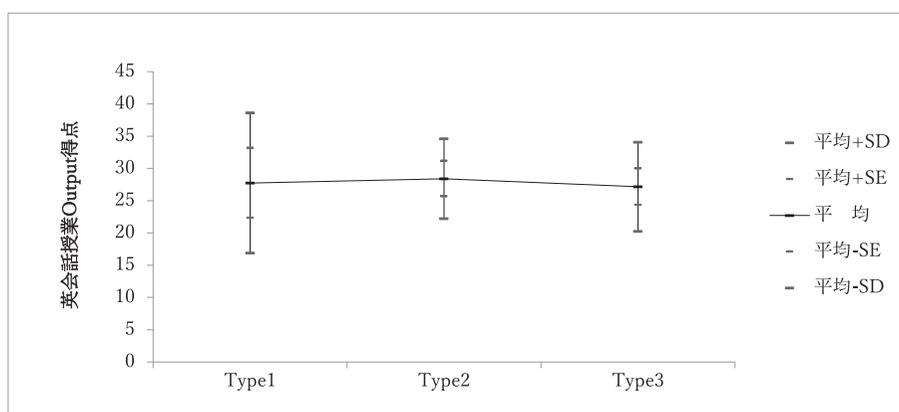


図 6. 各水準の平均値（クラスター）

表 12. 英語会話授業の Output スコアによる分析

手法	水準 1	水準 2	平均 1	平均 2	差	標準誤差	統計量	P 値
Fisher の最小 有意差法	Type 1	Type 2	27.7500	28.4000	0.6500	5.2919	0.1228	0.9043
	Type 1	Type 3	27.7500	27.1667	0.5833	5.0921	0.1146	0.9107
	Type 2	Type 3	28.4000	27.1667	1.2333	4.7769	0.2582	0.8006
Scheffe	Type 1	Type 2	27.7500	28.4000	0.6500	5.2919	0.0075	0.9925
	Type 1	Type 3	27.7500	27.1667	0.5833	5.0921	0.0066	0.9935
	Type 2	Type 3	28.4000	27.1667	1.2333	4.7769	0.0333	0.9673
Bonferroni	Type 1	Type 2	27.7500	28.4000	0.6500	5.2919	0.1228	1.0000
	Type 1	Type 3	27.7500	27.1667	0.5833	5.0921	0.1146	1.0000
	Type 2	Type 3	28.4000	27.1667	1.2333	4.7769	0.2582	1.0000
Sidak	Type 1	Type 2	27.7500	28.4000	0.6500	5.2919	0.1228	0.9991
	Type 1	Type 3	27.7500	27.1667	0.5833	5.0921	0.1146	0.9993
	Type 2	Type 3	28.4000	27.1667	1.2333	4.7769	0.2582	0.9921
Holm	Type 1	Type 2	27.7500	28.4000	0.6500	5.2919	0.1228	1.0000
	Type 1	Type 3	27.7500	27.1667	0.5833	5.0921	0.1146	1.0000
	Type 2	Type 3	28.4000	27.1667	1.2333	4.7769	0.2582	1.0000

*学力差はどの手法を使っても認められなかった。

7. 自由記述部分のテキスト分析

最後に、本節では質問紙調査の自由記述部分をテキストマイニング分析した結果を報告する。以下、図 7~9 は各質問項目において参加者が回答した自由記述の文言を分かち書きし、各単語をコンセプトごとにマッピングした結果である。この表示結果は、関連性の高い単語は近くに、関連度の低い単語はより離れて配置されることから、キーワード同士の関連が把握しやすいというメリットがある³。図 7 は質問項目 5・6 における Musio と人間の教師に関して、図 8 は質問項目 8・9 における Musio を英語学習に活用して良かった点と困った点、そして図 9 は質問項目 15 の Musio を英語学習に継続して活用したいかという意向に関するコメントの分析結果である。

まず、図 7 のコンセプト・マッピングの分析結果から、Type 1 は Musio との会話があまりスムーズにはいかなかったことから、Type 2 と同様に教師は人間である方が文脈を読み取ってくれる点や、表情などの非言語コミュニケーションも用いることができる点が良いと考えているさまが読み取れる。Type 3 は機械が相手だと緊張しないので話しやすいと回答していることが分かる。

AI を活用した次世代型英語スピーキング評価法の開発に向けて

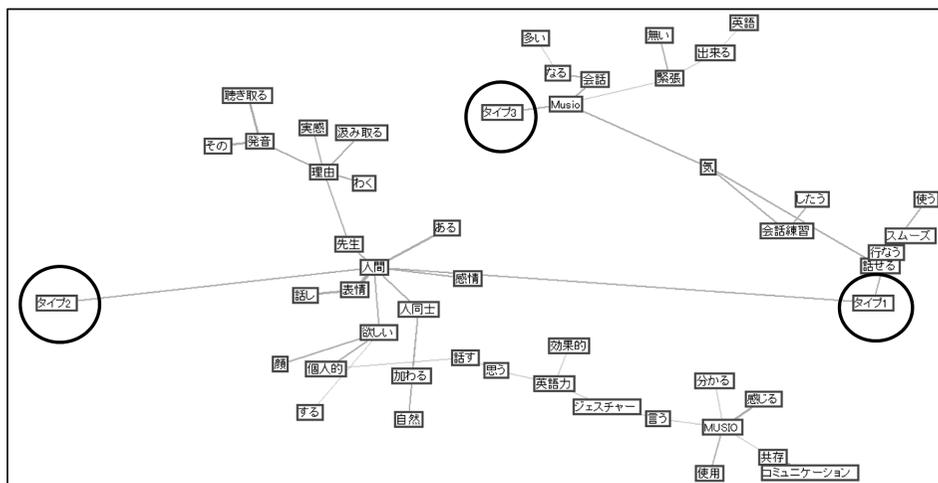


図 7. テキスト分析 1：項目 5・6 Musio と人間の教師について

次に図 8 から、Type 1 は、Musio が話題を何度も変えることから会話が持続しない点を指摘している。Type 2 は機械の不具合などの指摘があるが、話のテーマは多様である点については肯定的な評価である。Type 3 も同様で、Musio の知識の幅には肯定的である。但し、Musio の発話内容や返答内容を何度も聞き取らないといけない点には不満を抱いているようである。

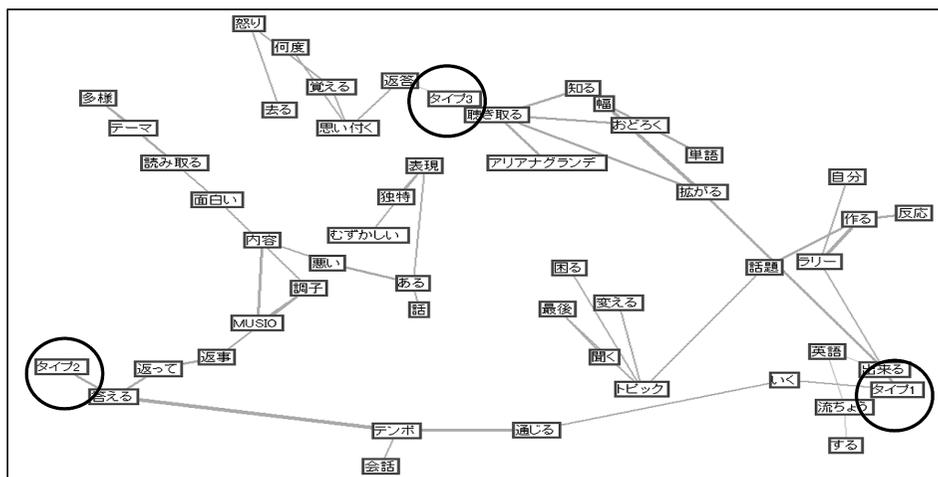


図 8. テキスト分析 2：項目 8・9 良いこと・困ったこと

図 9 からは、Type 1 は、新しい学習法としての AI ロボットとの会話に期待している様子が窺える。Type 2 は人間とのコミュニケーションを好むさまが読み取れる。Type 3 は、Mu-

sio の可愛さ、一人で発話の練習ができる事に肯定的であることが見て取れる。

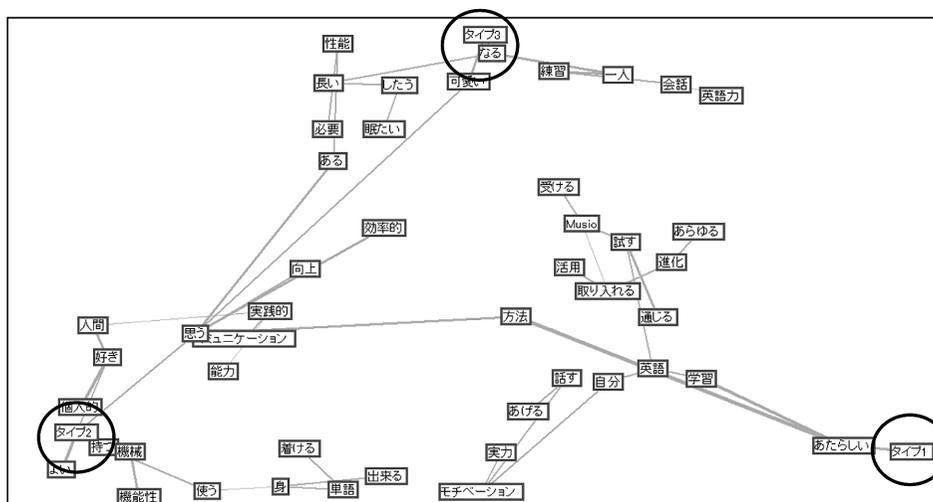


図 9. テキスト分析 3：項目 15 Musio 学習継続意向

8. まとめ

本稿では、AI ロボット Musio を活用した実証実験後に実施した質問紙調査の結果を詳細に報告した。特に因子分析とクラスター分析の結果から、4つの因子が抽出され（第1因子が「ロボットの方が学習しやすいと感じている」、第2因子が「英語発話量・繰り返しの効果に対する信奉」、第3因子は「Musio との学習が実践的・役立つと思う」、第4因子は「現状限界はあるが、今後の Musio に期待」）、3つのタイプに分類することができた（Type 1 は「限定的肯定派」、Type 2 は「人間重視・ロボット学習否定派」、Type 3 は「発話量重視・ロボット学習肯定派」）。AI ロボットが相手だと、緊張感が少なく間違っても良いことから Affective Barrier が低く、一人での英語発話練習にも向いている点など、肯定的な意見が多々見られた。一方で、言語を通じたコミュニケーションでは、表情やジェスチャーなども重視したいという点から人間の教師を好む声も一定数認められた。今後、AI ロボット自体の言語面・感情面の精度をより上げていくことによって、英語学習における AI ロボット活用の肯定派の割合も増えていくであろう。今後の改良に期待したい。

今後の課題として、本調査に先行して報告した仁科・東・山下（2021）で大学生英語学習者から収集した発話データを、テキストマイニングの手法を用いて分析し、AI ロボットと学習者のコミュニケーションがどの程度上手く交わされていたのか、あるいは AI ロボットからの不規則な返答に対して学習者はどのように対応していたのか、そして AI ロボットが的確に認識できなかった学習者の発音や表現にはどのようなものがあつたのかなどを順次解明していく予定である。

謝辞

本研究は JSPS 科研費 JP18K00809 の助成を受けたものである。

注

- 1 本研究で用いた AI (人工知能) ロボットは、英語学習を目的として AKA 社が開発し 2017 年に発売された Musio である (<https://themusio.com/home> を参照)。ディープ・ラーニング技術を採用した会話エンジンを搭載し、人間との英語コミュニケーションを円滑に進めるように設計されている。学校教育においては、2016 年 5 月より同志社中学校・高等学校で導入されており、続いて神戸学院大学においても 2017 年 3 月から全国初の試みとして英語授業で使われている (詳しくは、仁科・東・山下 (2021) を参照されたい)。
- 2 Welch (1938) によれば、Welch の t 検定の使用は非等分散性に左右されない。また、自由度が整数にならないことも多く、四捨五入せず小数点以下を切り捨てることもある。本研究では、一般的な Student の t 検定では自由度は 15 であるが、Welch の t 検定の場合は 7.5592 となる。
- 3 キーワード間の関連性をスプリングに見立てた物理モデルをシミュレーションすることによってマッピングが行なわれる。マッピングの生成過程において、スプリングによってキーワードは他のキーワードと引っ張り合うことでダイナミックに動きながら最終的な適正位置に配置される。

参考文献

- 石川慎一郎 (2009) 「因子分析における因子抽出法間の比較－日本人英語学習者の語彙学習方略データを利用して－」『統計数理研究所共同研究リポート』 232, 25-38.
- 仁科恭徳・東淳一・山下仁司 (2022) 「AI を活用した次世代型英語スピーキング評価法の開発に向けて－実証実験の内容とデータの取得について－」『神戸学院大学グローバル・コミュニケーション学会紀要』 第 7 号, pp.55-62.
- 繁榊算男・柳井晴夫・森敏昭 (編著) (2008) 『Q&A で知る統計データ解析－DOs and DON'Ts (第 2 版)』 東京：サイエンス社.
- 竹内理・水本篤 (編著) (2014) 『外国語教育研究ハンドブック (改訂版)』 東京：松柏社.
- Guadagnoli, E., Velicer, F. W. (1988). Relation of Sample Size to the Stability of Component Patterns, *Psychological Bulletin*, 103, 273-274
- Welch, B. L. (1938). The significance of the difference between two means when the population variances are unequal. *Biometrika*, 29, 350-362.