

学習者の困難度に応じた動的対応型英会話エージェント Study on a Dynamic Adaptive English Conversational Agent Based on Learner Difficulty

下坂 悠真¹ アイエドゥン エマヌエル² 徳丸 正孝²
Yuma SHIMOSAKA¹, Emmanuel AYEDOUN², and Masataka TOKUMARU²
¹関西大学大学院 ²関西大学
¹Graduate School Kansai University ²Kansai University

Abstract: In this paper, we propose an adaptive English conversation agent that dynamically controls support strategies based on learners' conversational difficulty states. The proposed system estimates learners' real-time interaction states (Normal, Silent, UNA, and NUNA) from utterance content and silence duration, and computes activation probabilities for affective backchannels and communication strategies based on recent state frequencies. By probabilistically adjusting intervention timing, the system suppresses unnecessary feedback during smooth interaction while providing emotional or cognitive support when learners experience difficulty. An experiment with 20 university students showed that the adaptive system produced more natural interactions by reducing redundant interventions, while revealing that preferred support frequency varies with learner proficiency and anxiety. These findings highlight the importance of personalized intervention strategies.

1. はじめに

近年のグローバル化の進展に伴い、実践的な英語コミュニケーション能力の重要性は一層高まっている。これに伴い、オンライン英会話レッスンサービスの普及により、英語学習環境は大きく拡充されてきた。しかし、対人での英会話においては、「間違えることへの不安」といった心理的要因が学習者に大きな負担を与え、外国語不安 (Foreign Language Anxiety : FLA) として発話意欲 (Willingness to Communicate : WTC) を低下させることが指摘されている[1][2]。

この心理的負担を軽減する手段として英会話エージェントが注目されているが、多くは学習者の状態や特性を十分に考慮しない画一的な対話設計に基づいており、新奇性による一時的な興味喚起に留まる場合が多い[3]。以上より、英会話エージェントにおいて、学習者の意欲を継続的に支え続けるための対話設計は未だ確立されていない。

事前研究では、対話型進化計算によりエージェントの非言語情報を個別最適化し、親しみやすさおよび英会話意欲の向上可能性を示した。しかし、言語的困難場面では非言語的支援のみでは不十分であり、評価入力に伴う負担という課題も残された[4]。

そこで本研究では、心理的安全性と会話の自律性の両立を目的とし、学習者の困難度状態をリアルタイムに推定し、沈黙時間および発話内容を手がかりとして会話戦略を動的に調整する適応型英会話エージェントを提案する。

2. 提案システム

2.1 システムの概要

図1に、本研究で提案するシステム概要図を示す。本システムは、学習者の音声及び沈黙時間を入力として、対話中に学習者の状態を逐次的に推定し、その結果に基づいてエージェントの会話戦略（支援の有無及び種類）を動的に決定する。

具体的には、まず音声認識モジュールにより学習者の発話をテキスト化する。次に、大規模言語モデルを用いた対話管理モジュールにおいて、対話文脈に基づく応答生成と学習者状態の推定を同時に行う。その結果に基づいて会話戦略を選択し、その戦略に応じた応答を生成して、音声合成モジュールを通じて学習者に提示する。

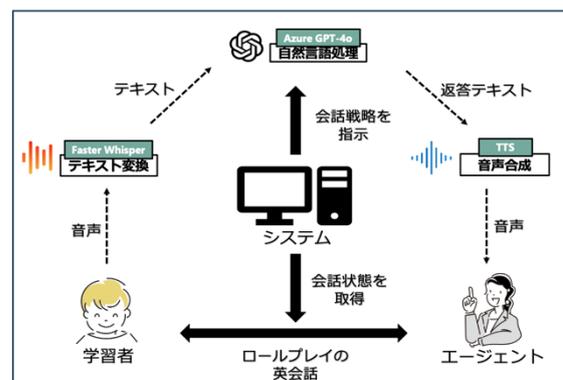


図1 システム概要図

本構成により、学習者の自律的な発話を妨げることなく、支援が必要と判断された場面に限って介入を行う適応的な対話制御を可能とする。

2.2 学習者の状態定義

本研究では、効果的な支援制御を行うため、学習者の会話状態を以下の4種類に分類して定義する。

- **Normal** (通常) : 会話が円滑に進行している状態であり、所定の待機時間内に学習者の発話が検知され、対話が継続している場合を指す。
 - **Silent** (沈黙) : エージェントの問いかけに対して一定時間応答が得られない状態である。この段階では困難の要因が特定できないため、確認質問を行い、次状態の判定に移行する。
 - **UNA** (回答困難) : 質問意図は理解しているものの、語彙不足や表現困難により発話が生成できない状態である。確認質問に対する肯定応答に加え、「"Let me see..."」「"I cannot explain..."」など、思考中または言語化困難を示す発話が検出された場合も本状態と判定する。
 - **NUNA** (理解不能) : 質問自体が理解できていない状態であり、確認質問に対する否定応答や、「"Pardon?"」「"I don't understand..."」などの理解不足を示す発話が検出された場合に該当する。
- 本システムは、これらの4種類の状態を発話内容及び沈黙時間からリアルタイムに推定する。

2.3 会話戦略 (AB・CS)

本研究における会話戦略は、Ayedoun らが提案した DiMaCA モデルを基盤とし、Affective Backchannels (AB) 及び Communication Strategies (CS) の2種類を採用する[5]。以下に2種類の詳細を述べる。本システムでは、この会話戦略を画一的に提示するのではなく、学習者の状態 (Normal, UNA, NUNA) に応じて、その役割及び介入強度を動的に変化させる。

(1) Affective Backchannels (AB)

AB は、学習者の FLA を軽減し、心理的安全性を確保するための共感的反応である。学習者が「間違えたらどうしよう」という不安や、「通じていないのではないか」という孤独感を感じると、WTC は著しく低下する。AB は、こうした心理的障壁を取り除き、学習者が安心して発話できる土台を形成する役割を担う。本システムでは、学習者の状態に応じて以下の4つの機能を使い分ける。

- **承認** : Normal 状態において、会話の流れを維持するための相槌として用いる。
ex) "Okay.", "I see."

- **励まし** : UNA 状態において、焦りを抑制し、思考時間を確保するために提示する。
ex) "Take your time.", "Don't worry."
- **同情・安心** : NUNA 状態において、理解困難の原因をエージェント側に帰属させ、学習者の心理的負担を軽減する。
ex) "I'm sorry, I wasn't clear.", "It's okay."
- **賞賛** : UNA または NUNA 状態から自力で回復し、発話に成功した直後に提示し、自己効力感の強化を図る。
ex) "Great!", "Perfect!"

(2) Communication Strategies (CS)

CS は、学習者が直面している言語的困難を解消するための具体的な足場かけであり、語彙不足や聴解ミスによる会話の断絶を防ぎ、意味交渉の成立を支援する役割を担う。本システムでは、学習者の困難の種類に応じて、支援のアプローチを以下のように切り替える。

- **UNA 時の支援 (ヒント提示)** : 発話意図は明確であるが、適切な英語表現が想起できない状態に対しては、直接的な解答提示を避け、語彙候補や文頭表現などのヒントを提示する。これにより、発話生成に伴う認知負荷を軽減し、学習者の自律的な発話を支援する。
- **NUNA 時の支援 (段階的調整)** : 質問内容自体の理解が困難な状態に対しては、同一内容を繰り返すのではなく、以下の3段階で支援強度を段階的に高める。
 - **Level 1 言い換え** : 別の英語表現で言い換えることで、多角的な理解を促す。
 - **Level 2 単純化** : 語彙や構文を平易なレベル (例 : occupation → job) に落とし、認知負荷を下げる。
 - **Level 3 日本語化** : 最終手段として日本語による補足や翻訳を提示し、確実な意味理解を保証する。

以上のように、本システムは AB による心理的支援と CS による言語的支援を組み合わせることで、学習者の状態に応じた過不足のない介入を実現し、会話の継続を支援する対話環境を構築する。

2.4 動的介入頻度の決定ロジック

本システムは、直近5ターンの会話履歴における各状態の出現頻度に基づき、次ターンで各支援を実行する確率 (Y) を動的に決定する。具体的な算出には、以下の式(1)、式(2)に示す重み付き線形モデルを用いる。なお、Silent 状態は一時的な遷移状態であり、その後の確認質問への応答によって最終的に Normal,

UNA, または NUNA のいずれかに確定される。そのため、介入確率の算出においては、**Silent** 状態自体は独立した変数として扱わず、確定した各状態の出現頻度に基づいて計算を行う。

$$Y_{AB} = Base + (X_{NUNA} \times 0.8) + (X_{UNA} \times 0.4) - (X_{Normal} \times 0.1) \quad (1)$$

$$Y_{CS} = Base + (X_{UNA} \times 0.7) + (X_{NUNA} \times 0.6) - (X_{Normal} \times 0.1) \quad (2)$$

Y_{AB} : ABを出す確率
 Y_{CS} : CSを出す確率
 $Base$: 初期値 0.1
 X_{UNA} : 過去5ターンでのUNAの出現頻度
 X_{NUNA} : 過去5ターンでのNUNAの出現頻度
 X_{Normal} : 過去5ターンでのNormalの出現頻度

重み付けの方針として、第二言語習得 (SLA) の知見に基づき、NUNA 状態では混乱や不安が生じやすいことから、認知的支援に先立って情動的安定化を優先する設計とし、AB に対応する重みを CS より高く設定している。一方、UNA 状態では発話生成の停滞が主因であるため、情動的支援よりも具体的な言語的足場かけを優先し、CS に対応する重みを高く設定している。

2.5 英会話例

提案システムを適用した際の対話生成例を図 2 に示す。本図は、学習者の状態遷移と算出された介入確率Yに基づき、エージェントが会話戦略(静観, AB, CS)を動的に切り替える様子を示している。

Normal 状態や、Yが閾値を下回る軽微な躓きに対しては、過干渉を避けるため静観を選択し、自然な対話進行を維持している。一方、「genre」での躓きに見られる NUNA 状態では、Yが低い段階では CS を提示せず、AB のみを表出する低介入な戦略により自発的回復を促している。その後、明確な聞き返しや発話停滞が生じ、Yが上昇した場面では積極的な介入を行い、NUNA 状態では CS を、UNA 状態では AB と CS を併用することで、心理的負担の軽減と発話支援を同時に実現している。

以上より、本手法は画一的なルールに基づくのではなく、学習者の状態と介入必要度に応じて会話戦略を適応的に生成できることが確認できる。

3. 評価実験

3.1 実験概要

提案システムの有効性を検証するため、日本語を母語とする大学生 20 名を対象とした被験者実験を実施する。比較対象として、支援可能なタイミングで常に (100%の確率で) AB 及び CS を行う「常時支援

エージェント: What do you like to do in your free time?	
学習者: I like watching movies.	(状態:Normal → 介入確率Y: 低 → 静観)
エージェント: So, what is your favorite genre?	
学習者: Gen...? I don't understand.	(状態:NUNA → 介入確率Y: 低 → ABのみ実行)
エージェント: Don't worry, take it easy. [AB: 同情]	
What is your favorite genre?	
学習者: Ah, OK! I like action movies.	(状態:Normal → 介入確率Y: 高 → AB実行)
エージェント: Great! Me too. [AB: 賞賛]	
Do you usually go alone?	
学習者: Pardon?	(状態:NUNA → 介入確率Y: 高 → CSのみ実行)
エージェント: I mean, who do you go with? [CS: 言い換え]	
学習者: I see, I go with my friends.	(状態:Normal → 介入確率Y: 低 → 静観)
エージェント: How often do you go?	
学習者: Umm... I cannot explain.	(状態:UNA → 介入確率Y: 高 → AB/CS実行)
エージェント: Take your time, you are doing fine. [AB: 励まし]	
Do you go "once a month" or "sometimes"? [CS: ヒント提示]	
学習者: Yes, once a month.	(状態:Normal → 介入確率Y: 高 → AB実行)
エージェント: Perfect! [AB: 賞賛]	

図2 提案システムによる会話例



図3 実験時のユーザインタフェース

表1 提案システムによる会話例

Q1	間違えることを恐れず自信を持って話せましたか
Q2	エージェントの助けは適切でしたか
Q3	エージェントの言動に煩わしいと感じませんでしたか
Q4	どちらの方が人間らしさを感じましたか
Q5	どちらの方が今後使用したいと感じましたか

型システム」を用意し、提案手法である「動的支援型システム」との比較を行う。

実験タスクには、海外旅行を想定した「空港」と「レストラン」の2種類のロールプレイシナリオを用いる。これらのシナリオには、意図的に難解な単語や、Yes/No では答えられない開かれた質問 (Open-ended questions) を混入させ、被験者が UNA や NUNA の状態に陥りやすい状況を設計する。実験時のユーザインタフェースを図 3 に示す。

被験者はそれぞれのシステムと対話し、実験後にアンケートによる主観評価を行う。各アンケートを表1に示す。Q1~Q3に対して被験者は、「そう思う/少しそう思う/どちらでもない/あまりそう思わない/そう思わない」の5段階で評価を行う。Q4,Q5は「提案システム/どちらかといえば提案システム/どちらでもない/どちらかといえば比較システム/比較システム」の5段階で評価を行う。

本実験の実施にあたっては、関西大学先端科学技術推進機構の研究倫理審査会の承認(承認番号:25-113)を得ている。

3.2 実験結果

実験後に実施したアンケート結果に基づき、本システムによる英会話エージェントの有効性を検証する。

まず、心理的安全性に関する結果を示す。「会話中に緊張や不安を感じませんでしたか」(Q1, 図4)という問いに対し、肯定的な回答(「そう思う」「少しそう思う」)の割合は、提案システム、比較システムともに85%となった。この結果から、両システムはいずれも高い水準で学習者の心理的安全性を確保できていることが確認された。

次に、エージェントの支援に対する評価について述べる。「エージェントの助けは適切でしたか」(Q2, 図5)という問いでは、比較システムが80%の肯定的回答を得たのに対し、提案システムは50%に留まった。一方で、「エージェントの介入に煩わしさを感じませんでしたか」(Q3, 図6)という問いにおいては、提案システムでは75%の被験者が煩わしさを感じなかったと回答したのに対し、比較システムでは50%に低下した。このQ2, Q3の結果は、常時支援を行う比較システムが「助けの確実性」において優位である一方で、「過干渉」と受け取られやすい傾向を示している。

システム全体の評価として、「今後使用したいと感じたのはどちらか」(Q4, 図7)という問いでは、提案システムを選択した被験者は55%、比較システムは40%となり、提案システムがやや好まれる傾向が見られた。さらに、「人間らしさを感じたのはどちらか」(Q5, 図8)という問いでは、提案システムを選択した被験者が65%に対し、比較システムを選択した被験者は25%であった。この結果は、状態推定に基づいて介入を動的に調整する提案システムの振る舞いが、機械的に支援を行う比較システムよりも、人間的な対話パートナーとして受容されたことを示している。

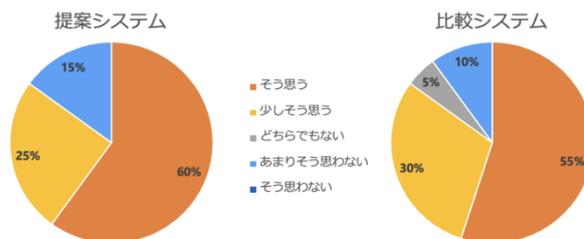


図4: Q1「会話中に緊張や不安を感じませんでしたか」のアンケート結果

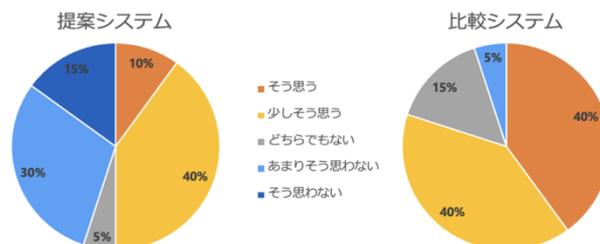


図5: Q2「エージェントの助けは適切でしたか」のアンケート結果

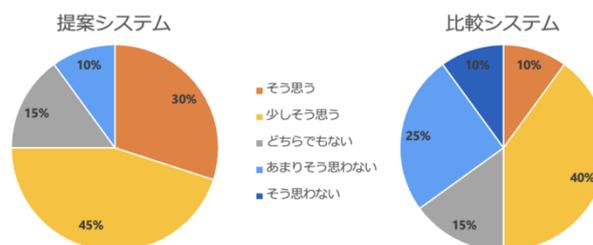


図6: Q3「エージェントの言動に煩わしいと感じませんでしたか」のアンケート結果

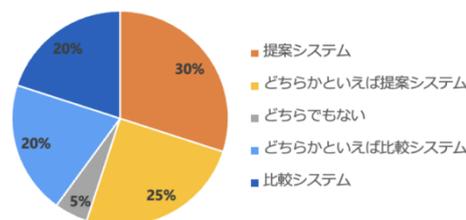


図7: Q4「どちらの方が今後使用したいと感じましたか」のアンケート結果

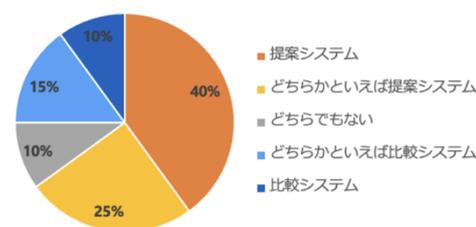


図8: Q5「どちらの方が人間らしさを感じましたか」のアンケート結果

3.3 考察

まず、心理的安全性 (Q1) について述べる。提案システムと比較システムはいずれも高い肯定的評価を獲得し、有意差は認められなかった。これは、本実験が「失敗が許容される練習環境」として認識されていたためと考えられる。比較システムは即時支援による機能的安心感を、提案システムは過干渉の抑制による情緒的安心感を提供しており、異なるアプローチながら双方が心理的障壁の低減に寄与したと解釈できる。

次に、支援の受容性を「支援の適切さ (Q2)」と「煩わしさ (Q3)」の観点から検討する。Q2 では比較システムが高い評価を得た一方で、Q3 では提案システムが煩わしさの低減において優れた評価を得るという対照的な傾向が確認された。これは、英会話支援における「機能的な問題解決」と「対話的自律性」の間にトレードオフが存在することを示唆している。

常時支援型である比較システムは、学習者が困難に直面した瞬間に具体的な解決支援である CS を即座に提示する設計であったため、Q2 において「助けが適切である」と評価されやすかったと考えられる。

一方、煩わしさ (Q3) では提案システムが比較システムを上回った。比較システムは毎回 AB を提示するため、一部の被験者には過剰介入と受け取られ、「急かされる」「考える余裕がない」といった阻害感につながった可能性がある。これに対し提案システムは、沈黙や言い淀みに即時介入せず、状態に応じて AB 頻度を抑制することで、煩わしさの低減に寄与したと考えられる。

ただし、この待機的挙動は、即時的な支援を求める学習者にとっては「支援が遅い」と受け取られ、Q2 における適切さの評価が低くなった可能性がある。

システム全体の評価として、利用意向 (Q4) では提案システムが比較システムを上回った。また、WTC スコアとの関係性を分析した結果、自信の高い被験者ほど提案システムを、不安の高い被験者ほど比較システムを選好する傾向が確認された。自信がある学習者にとっては自律性を尊重する提案システムが好まれ、不安が高い学習者にとっては比較システムの即時的な支援が安心感につながったと解釈できる。

最後に、「人間らしさ (Q5)」においては、提案システムが比較システムに対して顕著に高く評価された。人間同士の対話では、相手が考え込んでいる状況では沈黙を共有し、本当に困難に直面した場合にのみ支援を行うといった「間」の調整が自然に行われる。提案システムにおける状態推定に基づく動的な介入制御は、こうした人間的な対話リズムをより適切に模倣できていたと考えられる。

以上の結果より、万人に最適な単一の支援戦略は存在しないことが示唆された。今後は、学習者の習熟度や性格特性に応じて、介入の基準値自体を動的に適応させる仕組みの導入が必要である。

4. まとめ

本研究では、実ユーザを対象とした評価実験を通じて、学習者の状態に応じて会話戦略の介入確率を動的に制御する英会話エージェントの有用性を検討した。その結果、状態推定に基づく介入制御を導入することで、過干渉を抑えつつ学習者の心理的安全性を維持し、発話継続を支援し得る対話環境を構築できる可能性が示された。また、学習者の特性によって有効な会話戦略が異なる傾向が確認され、個人差を考慮した適応的支援設計の重要性が示唆された。

一方で、本実験は短期間のインタラクションに基づく評価に留まっており、長期的な学習効果や対話行動の変化については十分に検証できていない。今後は、学習者の習熟度や性格特性に応じて介入基準自体を個人適応させる仕組みの導入や、非言語情報を含むマルチモーダルな状態推定を取り入れることで、より精密で自然な適応的英会話学習支援システムの実現を目指す。

参考文献

- [1] Peter M., Richard C., Zoltán D., Kimberly N.: Conceptualizing willingness to communicate in a L2: A situational model of L2 confidence and affiliation, *Modern Language Journal*, Vol.82, No.4, pp.545-562, (1998)
- [2] Elaine H., Michael H., Joann C.: Foreign Language Classroom Anxiety, *The Modern Language Journal*, Vol.70, No.2, pp.125-132, (1986)
- [3] Luke F., Mary A., Andrew T., Aaron G., Zelinda S.: Stimulating and sustaining interest in a language course: An experimental comparison of chatbot and human task partners, *Computers in Human Behavior*, Vol.75, pp.461-468, (2017)
- [4] Yuma S., Emmanuel A., Masataka T.: Optimization of Non-Verbal Information for English Conversation Agents Using Interactive Evolutionary Computation, *The 32nd International Conference on Computers in Education*, (2024)
- [5] Emmanuel A., Yuki H., Kazuhisa S.: Communication strategies and affective backchannels for conversational agents to enhance learners' willingness to communicate in a second language, *International Conference on Artificial Intelligence in Education*, pp.459-462, (2017)