

やりとりの力を鍛える 会話AIエージェント

松山 洋一, Ph.D

株式会社エキュメノポリス 代表取締役

早稲田大学 知覚情報システム研究所 客員准教授





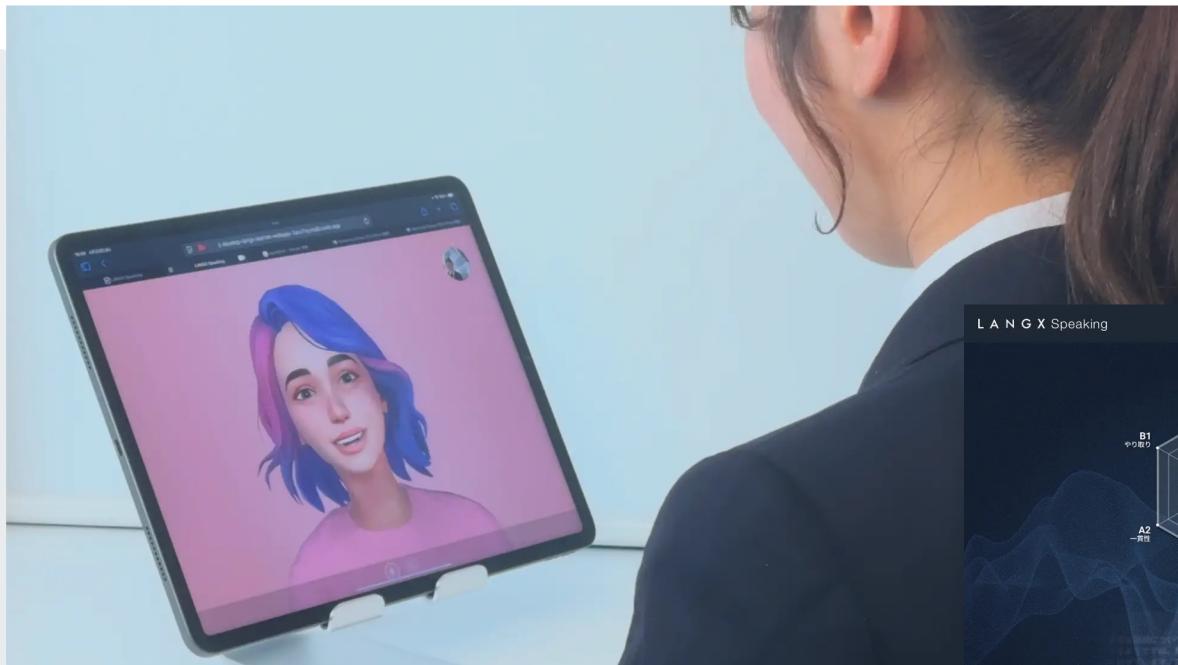
松山 洋一, Ph.D

株式会社エキュメノポリス 代表取締役
早稲田大学 知覚情報システム研究所 客員准教授

対話システムの基礎研究から 社会イノベーションへ

早稲田大学 基幹理工学研究科 情報理工学専攻で博士（工学）を取得後、
米国カーネギーメロン大学にてダボス会議公式バーチャルアシスタントの研究
開発プロジェクトや米Google、Microsoft、Yahoo!などとの各種会話AI产学連
携プロジェクトを主導。帰国後の2019年度に早稲田大学 知覚情報システム研
究所に准教授として着任。

2022年に株式会社エキュメノポリスを創業。代表取締役に就任。
対話型診断AIエージェントプラットフォーム「EQU AI」を開発。



オンラインAI英会話レッスン

Realistic conversational tasks

Local Governmental Customers



スピーキング能力判定

Your strength and weakness

University Customers



東京大学
THE UNIVERSITY OF TOKYO



九州大学
KYUSHU UNIVERSITY



早稲田大学
WASEDA University



中央大学



立教大学



関西大学
KANSAI UNIVERSITY

A screenshot of the LANGX learning feedback interface. It shows a conversation transcript between a user and an AI, with a green box highlighting a response. Below the transcript is a section titled 'アドバイス' (Advice) with a tip about using 'and' to connect activities.

学習フィードバック

For skill improvement

Governmental Supporters



文部科学省



OECD



NEDO



JST



NICT

<https://www.youtube.com/watch?v=xYITmyQIW14>



Yes.
はい。



言語学習・指導・評価のためのヨーロッパ言語共通参照枠 [欧州評議会, 2001, 2020]



Conversation between **social agents**

Social Agent

人間を社会的環境に状況づけられた存在として捉え、
特定のコミュニケーションタスクを達成するために
どのように**基礎能力 (Competence) を巧みに活用できるか**
について広範かつ包括的に整理した能力判定基準

Level: A2

Can-do: Can understand and use frequently used expressions related to areas of immediate relevance.

Role-play: Ordering food at a restaurant



Level: B1

Can-do: Explore solutions to address the declining customer satisfaction.

Role-play: A discussion between restaurant managers



Level: B2

Can-do: Can interact with a degree of fluency and spontaneity that makes regular interaction with native speakers.

Role-play: A meeting where the learner gives a progress report



あなたは、明日締切の最終レポートに取り組んでいる学生です。

しかし、友人があなたのPCを借りたいと言っています。

あなたは、お互いにとって良い解決策を見つけるために交渉しなければなりません。

- 自分の状況を説明し、まずはリクエストを断るようにしてください。
- その後、お互いにとって都合の良い解決策を交渉してください。
- 友人にPCを貸しても良いですが、自分のレポートを完成させるための十分な時間を確保してください。
- 友人に対して、苛立ちや迷惑そうな態度を見せないでください。



Our CEFR-based IC scale based on (Galaczi & Taylor 2018)

Turn Management

The ability to keep the conversation cooperative and collaborative through **appropriate turn-taking**.

Topic Management

The ability to **develop ideas collaboratively** towards the expected outcome of the interaction.

Interactive Listening

The ability to encourage the current speaker to keep their turn and continue talking by using **reactive actions** (e.g., backchanneling, acknowledgement).

Breakdown Repair

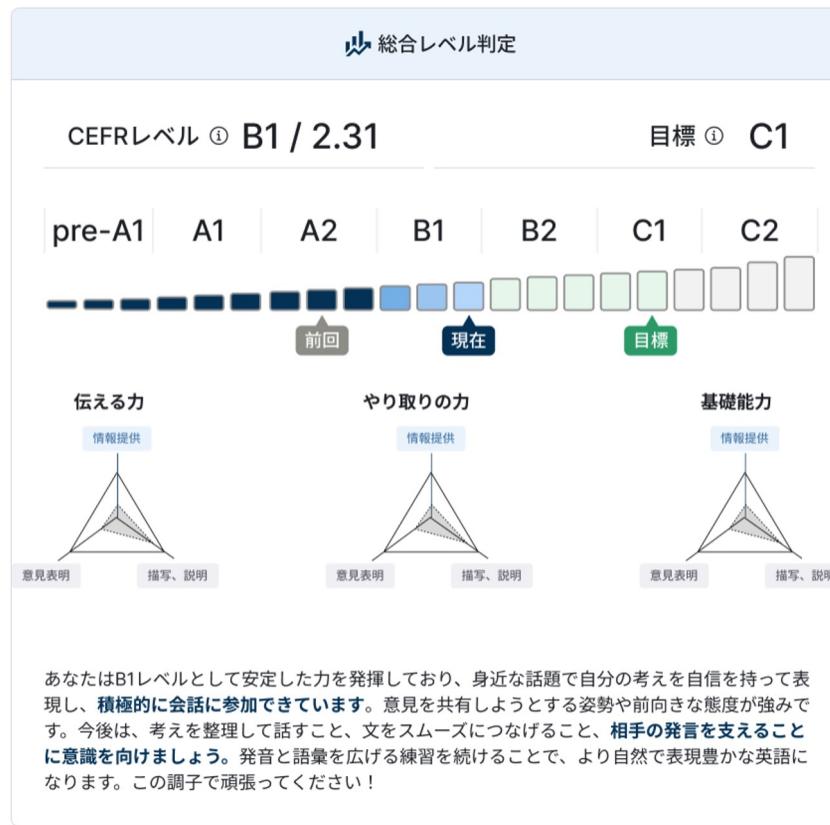
The ability to **repair breakdowns** in interaction proactively and preemptively.

(Galaczi & Taylor 2018) Evelina Galaczi, and Lynda Taylor, **Interactional competence: Conceptualisations, operationalisations, and outstanding questions**, Language Assessment Quarterly 15, no. 3 (2018): 219-236.

(Eguchi, 2025) Eguchi, M., Takizawa, K., Saeki, M., Kurata, F., Suzuki, S., Matsuyama, Y., & Sawaki, Y.(2025) **Human-versus artificial-intelligence-delivered roleplay tasks for assessing interactional competence: An applied conversation analytic study**. TESOL Quarterly (to be published).

やりとりの力の判定を含む「学習志向アセスメント」を提供

1. 全体スコア表示



2. Can-Do チェックリスト

※ トピック毎の達成状況

Part 1: 伝える力 ① Good So so Bad

天気について話す B2

- ✓ How is the weather today? 情報提供 >
- ⚠ What is your favorite type of weather? 意見表明 >

家族と兄弟姉妹について話す C1

ペットについて話す B2

スポーツについて話す B2

Part 2: やりとりの力 ①

勉強会の計画 B2

- ⚠ 適切なタイミングで話の順番を取り、維持し、終える 話す・聞くの順番のやりとり >
- ✓ 自発的に話の順番を始める >
- ✗ 相手に話す機会を与え、話の場を共有する >
- ⚠ 同意・不同意の際に相手の立場を理解して示す >
- ✓ 文脈に関連した新しい話題を始める >
- ✓ 明瞭な発音で伝えましょう >

出題や例示を通りで音目を發展させます >

3. 改善アドバイス

⚡ 次のレベルへのアドバイス

理由を添えて好みを明確に伝えましょう 意見表明 ^

判断基準
自分の好みを述べ、その理由を説明する力は、メッセージをより明確で説得力のあるものにします。

あなたの回答の傾向
晴れや寒い天気が好きなど、好みをよく表現できていますが、説明が繰り返しになったり不明瞭な部分があります（例：「I prefer cold weather... because...」）。理由を述べようとする意識はありますが、考えをもう少し整理すると良いでしょう。① センテンス

練習 - 次のレベルに向けて
「because」や「so」を使って理由を述べる練習をしましょう。
例：
• I like cold weather because...
• I prefer studying in the library because... \n自分の声を録音して、理由が明確に伝わっているか確認してみましょう。

ためらいを減らして流暢に話しましょう 流暢性 ^

会話の中で考えをつなげましょう 一貫性 ^

相手を促し、積極的に聞く姿勢を示しましょう やりとりの力 ^

スムーズに話す順番を管理しましょう やりとりの力 ^

明瞭な発音で伝えましょう 発音 ^

アウトプット仮説

会話のアウトプットを通して インプットの質を上げる

「覚えてから使う」ではなく「使いながら覚える」

1. Long, M. (1991). **Focus on form: A design feature in language teaching methodology.** In K. de Bot, R. Gisberg, & C. Kramsch (Eds.), Foreign language research in cross-cultural perspective. John Benjamins.
2. Gass, S. M., & Mackey, A. (2006). **Input, Interaction and Output: An Overview.** AILA Review, 19, 3–17.
<https://doi.org/10.1075/aila.19.03gas>
3. Swain, M. 2005. **The output hypothesis: Theory and research.** In Handbook on Research in Second Language Learning and Teaching, E. Hinkel (ed.), 471–483. Mahwah NJ: Lawrence Erlbaum.

早稲田大学で実施したスピーキングテストの対話データを用いて、ユーザーエミュレータを構築 [Obi 2026]



インタビュー会話（3年分）で学習

A レベル：8,171対話（722時間）

B レベル：13,031対話（1,804時間）

C レベル：3,118対話（645時間）

Full-duplex 音声対話モデル

発話の重なりや相槌といった対話中の相互行為を再現 [Défossez 2024]

[Défossez 2024] Défossez, A., Mazaré, L., Orsini, M., Royer, A., Pérez, P., Jégou, H., Grave, E., & Zeghidour, N. (2024). **Moshi: a speech-text foundation model for real-time dialogue.** arXiv preprint arXiv:2410.00037.

[Obi 2026] Obi, K., Yoshikawa, S., Saeki, M., Eguchi, M., & Matsuyama, Y. (2025). **Reproducing proficiency-conditioned dialogue features with full-duplex spoken dialogue models**, International Workshop on Spoken Dialogue Systems Technology (IWSDS 2026), 2026.



Aレベル
(初級者)



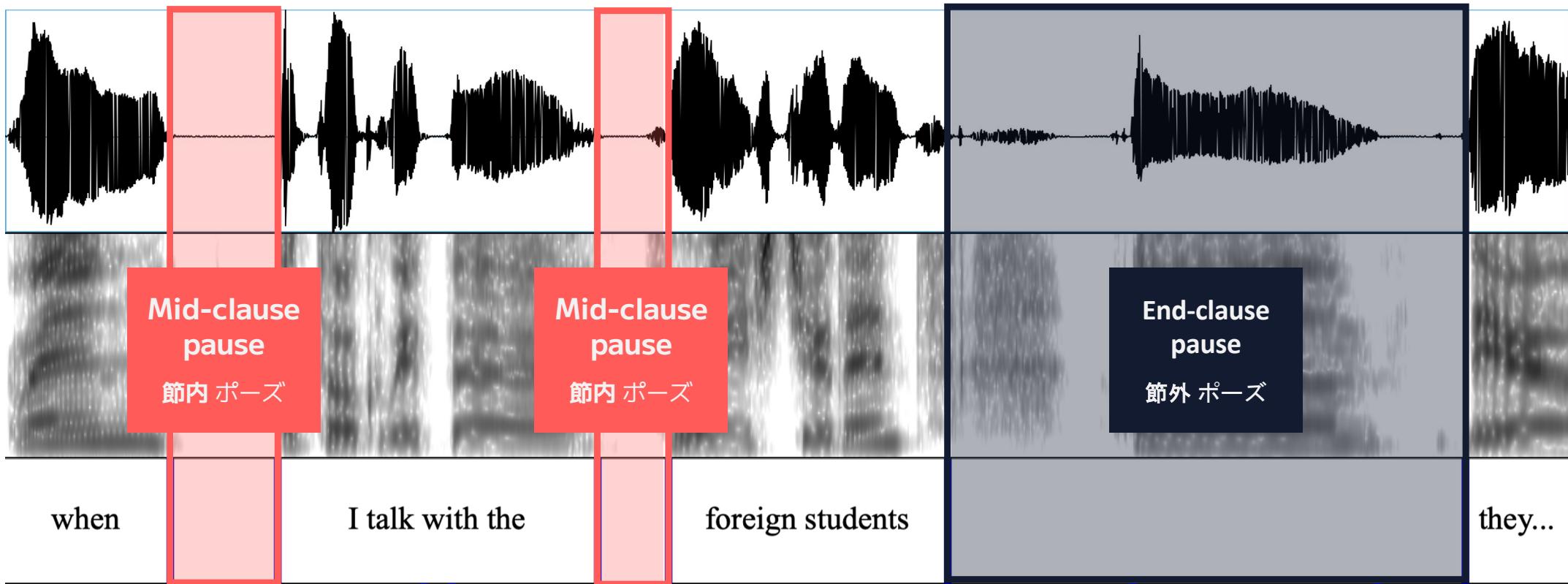
Bレベル
(中級者)



Cレベル
(上級者)

[Obi 2026] Obi, K., Yoshikawa, S., Saeki, M., Eguchi, M., & Matsuyama, Y. (2025). **Reproducing proficiency-conditioned dialogue features with full-duplex spoken dialogue models**, International Workshop on Spoken Dialogue Systems Technology (IWSDS 2026), 2026.

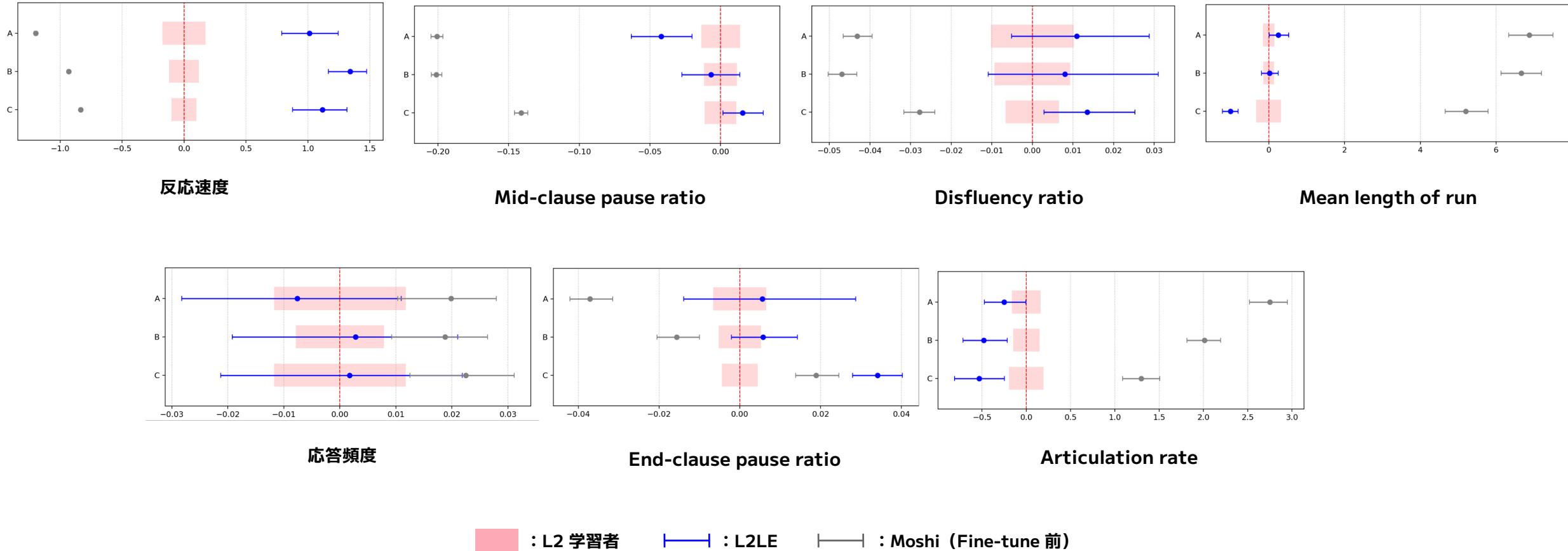
初級者ほど文章の中（節内）でポーズが発生しやすい



Ryuki Matsuura, Shungo Suzuki, Mao Saeki, Tetsuji Ogawa, and Yoichi Matsuyama. "Refinement of utterance fluency feature extraction and automated scoring of L2 oral fluency with dialogic features." In 2022 Asia-Pacific Signal and Information Processing Association Annual Summit and Conference (APSIPA ASC), pp. 1312-1320. IEEE, 2022.

Emulation Result | 流暢性に関する比較

実データ/模倣データの平均差の変動範囲比較により、CEFRレベルごとの再現度を比較



[Obi 2026] Obi, K., Yoshikawa, S., Saeki, M., Eguchi, M., & Matsuyama, Y. (2025). **Reproducing proficiency-conditioned dialogue features with full-duplex spoken dialogue models**, International Workshop on Spoken Dialogue Systems Technology (IWSDS 2026), 2026.

LANGX® Japanese



LANGX

はじめまして
Hajimemashite
Halo / Senang bertemu denganmu

あなたのなまえはなんですか?
Anata no namea wa nan desu ka?
Siapa namamu?

わたしのなまえはインテラです
Watashi no namea wa InteLLA desu
Namaku InteLLA

よろしくおねがいします
Yoroshiku onegaishimasu
Senang bertemu denganmu / Mohon kerjasamanya

あなたのしゅっしんはどこですか?
Anata no shusshin wa doko desu ka?
Dari mana asalmu?

わたしはにほんからきました
Watashi wa Nihon kara kimashita
Aku dari Jepang



✓ 自己紹介と患者の名前の確認

✓ 主訴の聴取

✓ 現病歴の確認

- ✓ Onset (発症)
- ✓ Provocation/Palliation (増悪・軽減因子)
- ✓ Quality (性状)
- ✓ Region/Radiation (部位・放散)
- ✓ Severity (重症度)
- ✓ Additional Symptoms (随伴症状)
- ✓ Timing (経過)

✓ 既往歴 (Past Medical History : PMH)

✓ 服薬歴 (Medications)

✓ アレルギー (Allergies)

✓ 家族歴 (Family History)

✓ 社会歴 (喫煙・飲酒・薬物使用・職業・等)

✓ 要約

✓ ICE (Ideas, Concerns, Expectations)

痛みはどの辺りでしょうか

人的資本は「効率化」から「信頼」の時代へ

人口構造的危機

日本の労働人口の減少による質的転換

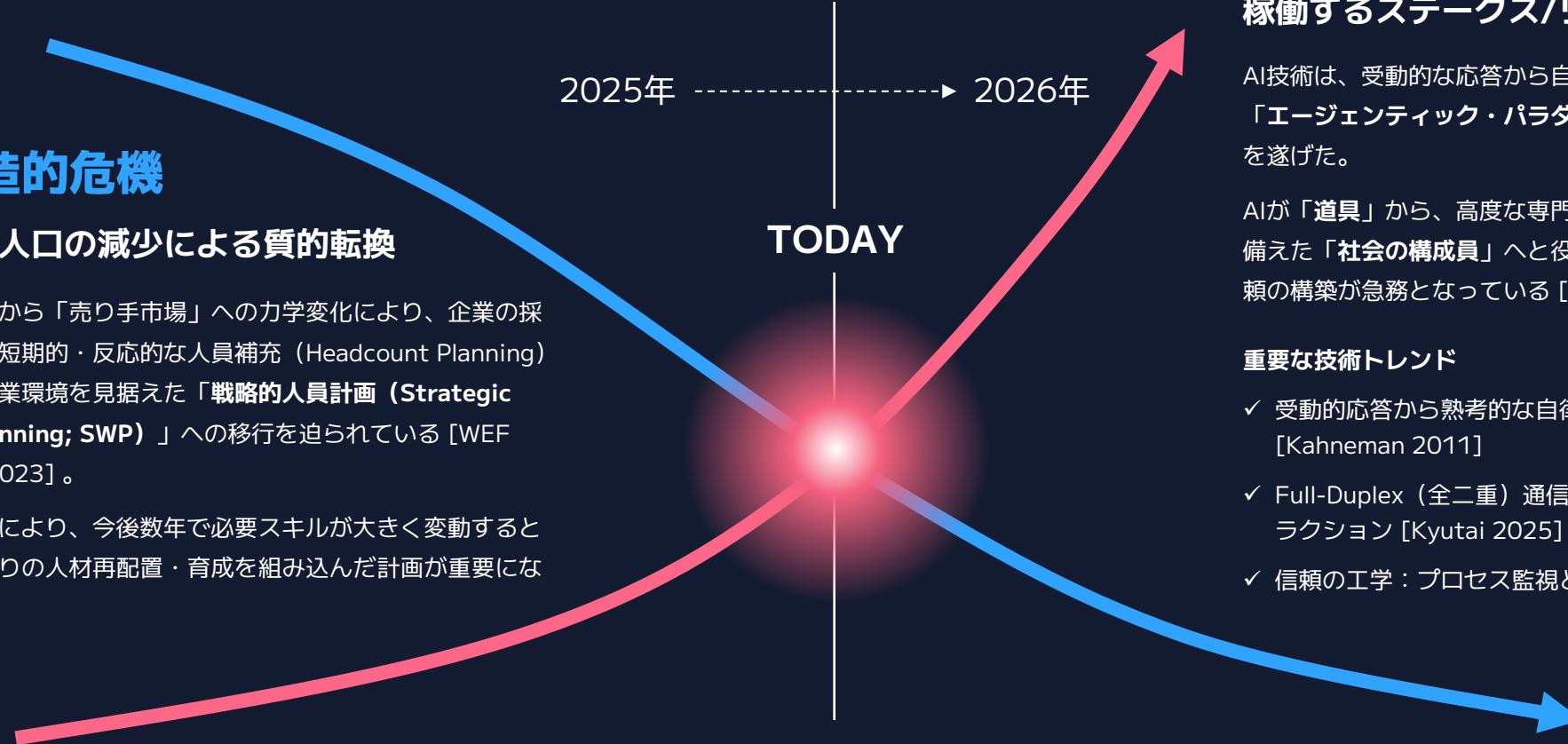
「買い手市場」から「売り手市場」への力学変化により、企業の採用活動は従来の短期的・反応的な人員補充（Headcount Planning）から、未来の事業環境を見据えた「戦略的人員計画（Strategic Workforce Planning; SWP）」への移行を迫られている [WEF 2025, KPMG 2023]。

AI等の技術変化により、今後数年で必要スキルが大きく変動すると見込まれ、先回りの人材再配置・育成を組み込んだ計画が重要になる。

2025年

TODAY

2026年



[WEF 2025] World Economic Forum, **Future of Jobs Report 2025**.

[KPMG 2023] KPMG International, **Creating Path to Continuous SWP**, 2023.

[a16z 2025] Andreessen Horowitz, **Big Ideas 2026 Podcast Series: Voice Agents and High-Stakes Trust**, 2025.

[Kahneman 2011] Kahneman, D, **Thinking, fast and slow**. Farrar, Straus and Giroux, 2011.

[Kyutai 2025] Kyutai, **Moshi: a speech-text foundation model for real-time dialogue**, 2025.

[JST 2024] 科学技術振興機構, さまざまな分野に広がるトラスト研究、総合知による取り組みへ, 2024.

Agentic AIの台頭

稼働するステークス/リスクの高まり

AI技術は、受動的な応答から自律的に計画・行動する「エージェンティック・パラダイム」へと非連続な進化を遂げた。

AIが「道具」から、高度な専門性と倫理的ガバナンスを備えた「社会の構成員」へと役割を変える中で、深い信頼の構築が急務となっている [a16z 2025]。

重要な技術トレンド

- ✓ 受動的応答から熟考的な自律的認知主体への転換 [Kahneman 2011]
- ✓ Full-Duplex（全二重）通信とマルチモーダル・インターラクション [Kyutai 2025]
- ✓ 信頼の工学：プロセス監視とガバナンス [JST 2024]

教育・雇用・医療 —— 「対話の質」を評価する全産業を対象に

EQU AI Platform

対話を通して真の能力を評価し科学的エビデンスへと変換。

個人と組織の最重要局面における意思決定を支える、

Agenticな診断プラットフォーム。

Gnowsis 診断フィードバック

グノウシス

- ✓ **マルチモーダル評価**：言語・音声・映像を統合しAIエージェントが「合議制」で多角的に診断。
- ✓ **客観的根拠**：監査可能性を高める客観的エビデンス（ログ・根拠・妥当性レポート）。



InteLLA 対話エージェント

インテラ

- ✓ **能動的対話戦略**：高度なインタラクション能力を見るとともに、思考の質を深掘りする。



Awards | 最近の主な受賞

Forbes Japan 日本発スタートアップ100選 (2026)



JEITAベンチャー賞 (2025)



大学発ベンチャー表彰 (2024)



SIGDIAL Best Paper Award (2024)



SXSW EDU Finalist (2023)



東洋経済すごいベンチャー100 (2023)



- Obi, K., Yoshikawa, S., Saeki, M., Eguchi, M., & Matsuyama, Y. (2025). **Reproducing proficiency-conditioned dialogue features with full-duplex spoken dialogue models**, International Workshop on Spoken Dialogue Systems Technology (IWSDS 2026), 2026.
- Eguchi, M., Takizawa, K., Saeki, M., Kurata, F., Suzuki, S., Matsuyama, Y., & Sawaki, Y. (2025). **Human- versus artificial intelligence-delivered roleplay tasks for assessing interactional competence: An applied conversation analytic study**. TESOL Quarterly. Advance online publication. <https://doi.org/10.1002/tesq.70028>
- Suzuki, S., Takatsu, H., Matsuura, R., Koyama, M., Saeki, M., & Matsuyama, Y. (2025). **Feedforwarding diagnostic language assessment: Artificial intelligence-driven weakness identification and contextualised feedback for second language speaking**. Language Testing, 42(4). Advance online publication. <https://doi.org/10.1177/02655322251348725>
- Nakano, M., Takeuchi, H., Yoshikawa, S., Matsuyama, Y., & Komatani, K. (2025). **Dialogue systems engineering: A survey and future directions**. arXiv preprint arXiv:2508.02279. <https://doi.org/10.48550/arXiv.2508.02279>
- Takatsu, H., Suzuki, S., Eguchi, M., Matsuura, R., Saeki, M., & Matsuyama, Y. (2025). **Gnowsis: Multimodal multitask learning for oral proficiency assessments**. Computer Speech & Language, 95, Article 101860. <https://doi.org/10.1016/j.csl.2025.101860>
- Saeki, M., Takatsu, H., Kurata, F., Suzuki, S., Eguchi, M., Matsuura, R., Takizawa, K., Yoshikawa, S., & Matsuyama, Y. (2024). **InteLLA: Intelligent Language Learning Assistant for assessing language proficiency through interviews and roleplays**. In Proceedings of the 25th Annual Meeting of the Special Interest Group on Discourse and Dialogue (SIGDIAL 2024) (pp. 385–399). <https://doi.org/10.18653/v1/2024.sigdial-1.34> [Best Paper Award]
- Kurata, F., Saeki, M., Fujie, S., & Matsuyama, Y. (2023). **Multimodal turn-taking model using visual cues for end-of-utterance prediction in spoken dialogue systems**. In Proceedings of the 24th Annual Conference of the International Speech Communication Association (Interspeech 2023) (pp. 2658–2662). <https://doi.org/10.21437/Interspeech.2023-1111> [Best Student Paper Award]
- Saeki, M., Miyagi, K., Fujie, S., Suzuki, S., Ogawa, T., Kobayashi, T., & Matsuyama, Y. (2022). **Confusion detection for adaptive conversational strategies of an oral proficiency assessment interview agent**. In Proceedings of the 23rd Annual Conference of the International Speech Communication Association (Interspeech 2022) (pp. 3988–3992). <https://doi.org/10.21437/Interspeech.2022-10884>