

ピンクノイズ特性を用いた1対1対話の盛り上げシステムの提案

Proposal of a pink noise-based dialogue enhancement system

小野寺 慶治郎^{1*} 小野寺 佳成² 三宅 芹奈² 今井 倫太¹
Keijiro Onodera¹ Yoshinari Onodera² Serina Miyake² Michita Imai¹

¹ 慶應義塾大学理工学部

¹ Faculty of Science and Technology, Keio University

² 慶應義塾大学大学院理工学研究科

² Graduate School of Science and Technology, Keio University

Abstract: This paper proposes PiNEs (Pink Noise-Based Interactive Dialogue Enhancement System), which improves one-on-one conversations by analyzing pink noise characteristics. Unlike traditional systems that generate simple responses, PiNEs enhances engagement by assessing conversation quality in real time and suggesting relevant utterances. Using PNR (Pink Noise Robustness) to measure adaptability and consistency, PiNEs generates contextually appropriate utterances when PNR declines, sustaining dialogue and fostering trust. Experiments showed that PiNEs increased PNR, revitalizing conversations and improving overall quality. A questionnaire confirmed its effectiveness in enhancing conversational enjoyment and mutual understanding.

1 序論

対話支援システムは、情報提供や応答生成を目的として発展してきた。しかし、既存の手法は対話の流れを維持する機能が限定的であり、対話の質や楽しさを向上させる仕組みが不足している。特に、対話が停滞した際に適切な介入を行う技術の欠如が課題となっている。本論文では、対話の意味的内容のピンクノイズ特性に着目することで、対話支援の改善を行う。ピンクノイズは、パワースペクトル密度が周波数に反比例する特性を持つノイズ (図1) であり、バランスが取れたノイズとして複雑なシステムの適応性や安定性を評価する指標として利用される [1][2]。

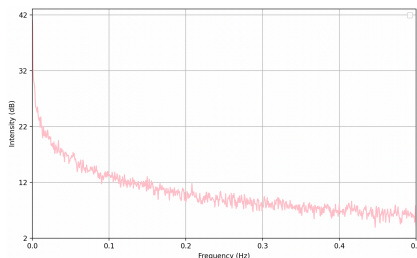


図 1: ピンクノイズのノイズ特性

ピンクノイズが観測される場合、そのシステムが長期間安定しながらも変動可能な状態 (メタ安定性) を持つことが示唆される。本論文では、対話の意味的内容のピンクノイズ特性を考慮した指標 PNR (Pink Noise Robustness)[6] を導入することで、より自然な対話の維持を支援するシステム PiNEs を提案する。

従来の対話支援システムとしては、会話の停滞を検出し、あらかじめ用意された話題を提案する [3] がある。しかし、状況に応じた動的な対応はできず、リアルタイム性の欠如が課題として挙げられる。他にも、AI アシスタントが会話の流れに基づいて応答を提示する手法 [4] もあるが、適切な発話を生成し対話の質を向上させる機能は不十分である。また、現行の研究では、対話の盛り上がりや継続性を測定する明確な指標が確立されておらず、支援の効果が定量的に評価されていない。以上の課題を解決するために、本研究では PNR を活用した新たな手法を導入する。

2 提案

本論文では、従来の対話支援システムの問題点を解決するため、ピンクノイズ特性を活用した1対1対話の盛り上げシステム PiNEs (Pink Noise-Based Interactive Dialogue Enhancement system) を提案する。PiNEs は、対話の適応性や一貫性を示す指標 PNR を導入し、対話の質をリアルタイムで評価することで、適切なタ

*連絡先: 慶應義塾大学理工学部今井研究室
〒 223-0061 神奈川県横浜市港北区日吉 3 丁目 14-1
E-mail: keijiro@ailab.ics.keio.ac.jp

イメージで発話を提案するシステムである。

2.1 PNR の導入と活用

本論文では、対話の適応性を定量的に評価するための指標としてPNR[6]を導入した。PNRは、会話の一貫性や適応性を示す指標であり、対話の同期性を測定する。PNRの計算は、以下の手順で行われる(図2)。

1. 対話のベクトル埋め込み
発話テキストを高次元ベクトル空間に埋め込む。
2. コサイン類似度の算出
各発話間のコサイン類似度を計算し、対話の意味的同期性を表す時系列データを取得する。
3. DFA解析の適用
得られた時系列データをもとに、デトレンドフラクチュエーション解析(DFA解析)[5]を用いてスケール係数を算出する。
4. シャッフルデータとの比較
時系列データを1000回シャッフルして生成されたスケール係数の分布から、元の信号より低い割合を測定することでPNR値を算出する。

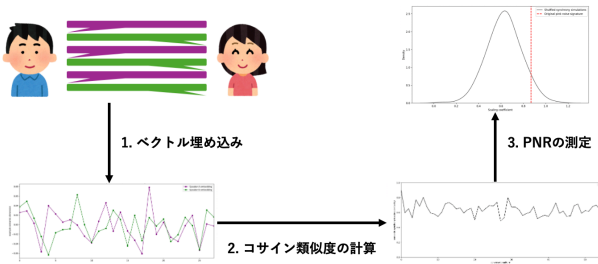


図 2: PNR の計算手順

PNRは、会話の質に対する認識を表す多くの変数と有意に相関があることが示されている[6]。PNRの値が高いほど、対話の適応性が高く一貫性が維持されている一方、PNRが低下すると会話の同期性が失われ、対話の質が低下する可能性が高い。PiNEsでは、PNRをリアルタイムで測定し、対話の流れの停滞を感知して適切な発話を提案する。

2.2 PiNEs のシステム概要

PiNEsは、対話のリアルタイム解析を行い、PNRの変化を基に適切な発話を提案することで、対話の質を向上させる。図3にPiNEsのシステム構成図を示す。

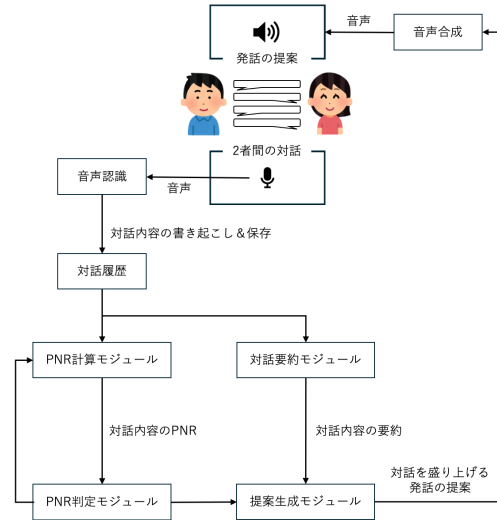


図 3: PiNEs のシステム構成図

PiNEsのシステムは、主に以下の4つのモジュールで構成される。

1. PNR 計算モジュール
対話のピンクノイズ特性を測定し、リアルタイムで毎ターンのPNRを算出する。
2. 対話要約モジュール
これまでの対話履歴を基にして、提案生成に必要な対話情報を要約する。生成された要約は提案生成モジュールで使用される。
3. PNR 判定モジュール
算出されたPNRの値が閾値を上回った場合は対話を継続し、下回った場合は対話が盛り上がっていないと判定し提案生成モジュールに渡される。閾値は[6]におけるPNRの分布の中央値である $PNR = 0.91$ を使用した。
4. 提案生成モジュール
PNRが低下した場合、対話要約をもとに対話の流れに応じた発話提案をLLMを用いて生成する。発話提案においてメタ安定性を実現するためには、二者発話間の過度な同期(両者の発話が似すぎている状態)、および過度な非同期(両者の発話が噛み合っていない状態)は避けるべきである。過度な同期、非同期を回避するため、直近3回分のコサイン類似度の平均をとりACS(Average Cosine Similarity: 平均コサイン類似度)とし、ACSの値に応じて以下の条件分岐を行う。
 - 過度な同期 ($ACS > 0.7$)
会話が停滞し、変化が少なくなっている状態。新しい話題を提案する。

- 適度な同期 ($0.3 \leq ACS \leq 0.7$)
対話のバランスが取れている状態. 対話の流れを維持しつつ, 関連する発話を提案する.
- 過度な非同期 ($ACS < 0.3$)
会話の流れが途切れている状態. 共感を示し, 現在の話題を掘り下げる発話を提案する.

2.3 本研究の貢献

PiNEs により, 従来の対話支援システムの問題点を解決する以下の改良を行なった.

- リアルタイム性の向上
PNR をリアルタイムで測定し, 対話の状態に応じた即時対応を可能にした.
- 文脈適応型の支援
大規模言語モデルを活用し, 対話の流れに適した発話提案を動的に生成する仕組みを導入した.
- 対話の質の定量的評価
PNR を導入することで対話の質を数値化し, 対話支援の効果を定量的に分析できるようにした.

3 実験

提案した PiNEs の有効性を検証するため, PiNEs を使用した場合と使用しない場合の対話における PNR の変化と対話の質に与える影響を評価する実験を行った.

3.1 実験参加者

本実験には, 平均年齢 21.5 歳の慶應義塾大学の学生 6 名 (男性 4 名, 女性 2 名) が参加した. 参加者はすべて日本語を母語とし, 事前に実験手順および対話のルールについて説明を受けた.

3.2 実験条件

本実験では, 2 人 1 組の対面形式で対話を行い, 対話の質および PNR の変化を評価した. 各被験者には「質問者」と「回答者」の役割を与え, 質問者が自由な内容で質問し, 回答者がそれに応答する形式とした. 表 1 に示した 4 つの条件を設定し, PiNEs の有無および役割を入れ替える形で対話実験を行い, それぞれの結果を比較した.

表 1: 実験における 4 つの試行

	PiNEs	被験者 1 の役割	被験者 2 の役割
試行 1	なし	質問者	回答者
試行 2	あり	質問者	回答者
試行 3	なし	回答者	質問者
試行 4	あり	回答者	質問者

3.3 評価項目

本実験では, 各条件における PNR の値を元に行う数値評価と, 対話後の被験者アンケートの結果を元に行うアンケート評価の二軸で検証した.

3.3.1 PNR の変化の評価

各ターンごとに計算された PNR をもとに, PiNEs の使用による PNR の変化を定量的に分析した. 以下の 2 つの視点で評価を行った.

1. PiNEs の使用による対話全体の PNR の変化
PiNEs の使用の有無で, 対話が終了した際の最終的な PNR, および対話開始時から終了時にかけての PNR の変化を分析する.
2. PiNEs の発話提案による PNR の一時的変化
PiNEs が被験者に対して提案を行なった際, 被験者が発話の提案に従った発話をする事によって発話提案前後で PNR がどう変わるかを分析する.

3.3.2 アンケート評価

対話の質を調査するため, 対話終了後に実施したアンケートで以下の質問項目について 5 段階リッカード尺度で回答を得た.

- 対話の楽しさ
- 対話を継続したいと思ったか
- 対話相手との繋がりを感じたか
- 対話相手が自分に興味を持っていたと感じるか
- 対話相手が友好的だったと感じるか

さらに, すべての試行終了後に最終アンケートを実施し, 4 つの試行の中で最も楽しかった試行について選択してもらった.

3.4 実験結果

3.4.1 PNR の変化の評価

- PiNEs の使用による対話全体の PNR の変化
対話終了時の PNR(表 2), および対話開始時から終了時にかけての PNR の変化量 (表 3) をまとめ、代表的な統計値を以下に示す.

表 2: 対話終了時の PNR の代表的な統計値

	PiNEs あり	PiNEs なし
平均	0.483	0.5
中央値	0.49	0.44

表 3: 対話中の PNR 変化の代表的な統計値

	PiNEs あり	PiNEs なし
平均	0.092	0.023
中央値	0.02	-0.035

ウィルコクソン符号順位検定を実施したところ、対話終了時の PNR, 対話における PNR の変化共に p 値は 1.0 となり、優位差は見られなかった。

- PiNEs の発話提案による PNR の一時的変化 PiNEs 使用時における発話提案前後の PNR の一時的変化の結果をまとめ、代表的な統計値を以下の図 4 に示す.

表 4: 提案前後の PNR の差の代表的な統計値

	PNR が上昇	PNR が下降	全提案
度数	15	3	18
平均	0.088	-0.13	0.052
中央値	0.06	-0.14	0.055
標準偏差	0.075	0.075	0.111

片側 t 検定を実施したところ、t 統計量=1.981, p 値=0.032 となり、PNR の一時的変化の平均が 0 よりも有意に大きい、つまり提案が PNR を有意に上昇させたことが示された。また、Cohen'd = 0.467 となり、PNR の一時的変化は中程度の効果を持つと評価される。

3.4.2 アンケート評価

対話後アンケートの結果を以下の図 4 に示す。

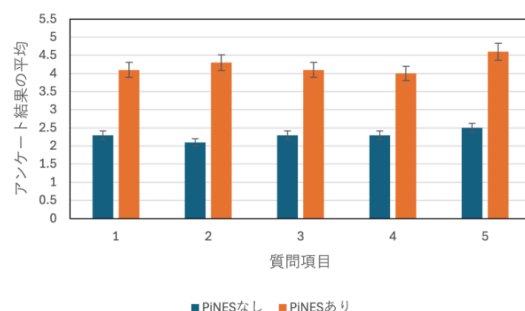


図 4: 対話後アンケート結果

質問項目 1-5 の p 値は、それぞれ 0.0028, 0.0005, 0.0045, 0.0029, 0.0005 となった。したがって、全ての質問項目において「PiNEs あり」の場合と「PiNEs なし」の場合で、アンケート評価に統計的に有意な差があることが確認された。

また、最終アンケートの結果を以下の図 5 に示す。

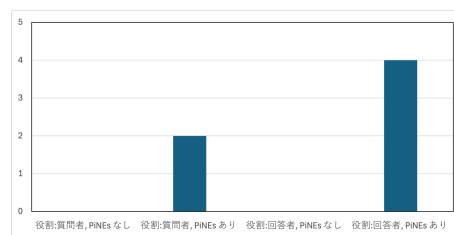


図 5: 最終アンケート結果

したがって、PiNEs あり/質問者に 2 票、PiNEs あり/回答者に 4 票となり、PiNEs を使用した条件に全ての票が入ったことがわかる。しかし、役割が質問者の場合と回答者の場合でアンケート結果に差が出たため、PiNEs の使用の有無で役割間の差を調査するためウィルコクソン符号順位検定を実施した。結果として、PiNEs を用いなかった場合は p 値=0.0072 となり質問者と回答者の間でアンケート評価に有意な差があるのに対し、PiNEs を用いた場合は p 値=0.3381 となり有意な差は見られないという結果が得られた。

4 考察

4.1 PNR の変化の考察

4.1.1 PiNEs の使用による対話全体の PNR の変化

対話全体の PNR, および対話中の PNR の変化に差が見られないのは、PiNEs を使用しなかった場合も対話

参加者が楽しく対話し、システムの介入無しでも PNR が高い値を維持したためだと考えられる。また、PiNEs が提案を行う際には、ある程度の対話の間隔が必要であったため、全体の対話時間に対して提案が相対的に少なかった可能性も考えられる。

4.1.2 PiNEs の発話提案による PNR の一時的変化

表 4 より、18 回の提案のうち 15 回、つまり提案全体の 83% で PNR が上昇したことがわかる。また、3.4.1 における分析では、PiNEs が PNR を上昇させる効果を持つことが統計的に確認された。分析の結果は、PiNEs が対話の同期性を改善し、質の向上に寄与する可能性を示している。したがって、PiNEs の使用による効果は全体には対話全体ではなく局所的なものに留まったことを示しているといえる。

4.2 アンケート評価の考察

4.2.1 対話後アンケート

対話後アンケートの結果 (図 4)、および 3.4.2 における統計分析の結果からも、PiNEs を用いることによって対話の各側面の評価が有意に高くなることが示されている。従って、PiNEs のリアルタイムでの発話提案が対話の活性化に効果的であることが示唆される。

4.2.2 最終アンケート

3.4.2 より、PiNEs を用いない場合は役割間のアンケート評価に差があるのに対し、用いた場合は差はないという結果が得られた。PiNEs を用いない場合、質問者は継続して質問するのに対し、回答者は質問に答えるのみであるため、役割間の差によって対話の難易度に違いが生まれた可能性がある。一方、PiNEs を用いた場合、発話の提案により質問者は質問だけをでなく発話提案をもとに対話を進めることができる。PiNEs の発話提案により、双方の被験者に対話を進めるきっかけが提供され、役割間の認識の違いを軽減する効果があったと考えられる。

5 まとめ

本論文では、ピンクノイズ特性を活用した 1 対 1 対話の盛り上げシステム PiNEs を提案し、その有効性を検証した。PiNEs は、対話の質を測定する指標 PNR を導入し、対話の状態をリアルタイムで評価することで、適切な発話提案を行うシステムである。実験では、PiNEs の発話提案が PNR の局所的な向上に寄与する

ことが示された。また、アンケート結果では PiNEs の使用により対話の質に対する評価が向上し、実験参加者同士の役割間の評価の差が軽減されることが確認された。分析の結果を踏まえ、PiNEs は 1 対 1 対話を支援し、質を向上させる有望なシステムであると結論付けられる。

謝辞

本研究は、JST、CREST、JPMJCR19A1 の支援を受けたものである。

参考文献

- [1] Kerr, W. T., Anderson, A., Lau, E. P., Cho, A. Y., Xia, H., Bramen, J., Douglas, P. K., Braun, E. S., Stern, J. M., and Cohen, M. S.: Automated diagnosis of epilepsy using EEG power spectrum, *Epilepsia*, Vol. 53, No. 11, pp. e189–e192 (2012).
- [2] Raffalt, P. C., Sommerfeld, J. H., Stergiou, N., and Likens, A. D.: Stride-to-stride time intervals are independently affected by the temporal pattern and probability distribution of visual cues, *Neuroscience Letters*, Vol. 792, p. 136909 (2023).
- [3] Isbister, K., Nakanishi, H., Ishida, T., and Nass, C.: Helper agent: Designing an assistant for human-human interaction in a virtual meeting space, *Proceedings of the SIGCHI Conference on Human Factors in Computing Systems*, pp. 57–64 (2000).
- [4] Hohenstein, J., and Jung, M.: AI-supported messaging: An investigation of human-human text conversation with AI support, *Extended Abstracts of the 2018 CHI Conference on Human Factors in Computing Systems*, pp. 1–6 (2018).
- [5] Gorjão, L. R., Hassan, G., Kurths, J., and Witthaut, D.: MF DFA: Efficient multifractal detrended fluctuation analysis in python, *Computer Physics Communications*, Vol. 273, p. 108254 (2022).
- [6] O’Neill, K., and Finn, E. S.: Pink noise in speakers’ semantic synchrony dynamics as a metric of conversation quality, *Proceedings of the Annual Meeting of the Cognitive Science Society*, Vol. 46 (2024).