

AI ナレーションのアクセントが観光地訪問意図に与える影響

— 有能さと擬人性のトレードオフ —

松山大学 経営学部 専任講師

立命館大学 経営学研究科 博士後期課程

立命館大学 経営学部 准教授

崔 琳源

吳 松源

苗 苗

要約

近年、AI によるナレーション技術が発展を遂げているが、観光プロモーションにおける AI ナレーションの効果を実証的に検討した研究はまだ限られている。本研究では、観光プロモーション動画においてナレーションの主体（人間 vs. AI）および音声のアクセント（標準アクセント vs. 地方アクセント）が視聴者の訪問意図に与える影響と、その心理的メカニズムを 2 つの実験研究を通じて検討した。研究 1 では、ナレーション主体と音声のアクセントが訪問意図に与える直接的影響を分析した。その結果、双方の間に有意差は見られなかった。研究 2 では、AI ナレーションに焦点を絞り、知覚された有能さと擬人性を媒介変数とする心理的プロセスを検討した。その結果、AI が標準アクセントを話すとき、有能さを介した正の媒介効果、擬人性を介した負の媒介効果を通じて訪問意図に影響を与えることが確認された。これらの結果に基づき、観光領域における AI ナレーション活用に関して理論的・実践的な示唆が得られた。

キーワード

AI, アクセント, 観光地プロモーション, ステレオタイプ内容モデル, 擬人性

1. はじめに

近年、AI の進化は著しく、動画クリエイターや企業のプロモーション動画制作など、さまざまな場面で AI による音声ナレーションや読み上げ機能の活用が増えている。特に、観光地プロモーション動画においても、AI ナレーションの活用が増えている（Wang et al., 2024）。しかし、AI ナレーションは高品質で自然にテキストを読み上げることが可能である一方で、人間特有の感情表現やイントネーション、さらに育った地域によって異なる地方アクセントなどといった、人間らしさを感じさせる要素には乏しく、冷たく無感情であると指摘されている（Poushneh et al., 2024; Wang et al., 2024）。

観光に関わる従業員が観光地特有の地方アクセントを用いることで、現地の観光客の体験価値または潜在的な観光客の訪問意図を高める役割を果たすことが示されている（Rao &

Tombs, 2022; Li et al., 2025)。したがって、AI がこうした地方アクセントを用いることで、欠如しがちな「人間らしさ」を補完し、説得力を高める可能性があるかどうかについては、いまだ十分に検証されていない。

本稿では、観光プロモーション動画において AI ナレーションのアクセント（標準アクセント vs. 地方アクセント）が視聴者の訪問意図およびその背後にある心理的メカニズムに与える影響を検証することを目的とする。具体的には、以下の 2 つのリサーチクエスチョンを設定した。(1) AI ナレーションは人間ナレーションの代替となり得るのか、(2) AI ナレーションにおける地方アクセントの使用が、訪問意図を高める効果が得られるのか。

II. 理論的背景と仮説の導出

1. AI ナレーションとアクセント

AI は生産コストの低さと効率性の観点から、人間ナレーションを代替する選択肢として広告で多く使われている (Rodero, 2017; Poushneh et al., 2024)。観光地プロモーション動画においても、ナレーションは視聴者の訪問意図の形成に大きく寄与する重要な要素とされており (Gan et al., 2023)、AI によるナレーションの活用が注目されつつある。本稿では、Wang et al. (2024) に基づき、ナレーションを人間ナレーションと AI ナレーションに分類した。人間ナレーションは自然なイントネーションや感情表現に優れ、一方 AI ナレーションは均質で精度の高い読み上げや多様なアクセント対応が可能である。

既存研究では、消費者が AI ナレーションよりも人間ナレーションを好み、また AI ナレーションに人的要素を加えることで消費者の好意的な反応が高まることが論じられていた (Rodero, 2017; Poushneh et al., 2024)。特に、人間の声に含まれるプロソディ (prosody : イントネーション, ストレス, リズムなどの声の要素) が、メッセージに感情や意味を付加させ、信頼性、説得力や自然さの点で AI 音声を上回ることが示されていた (Rodero, 2017)。よって、以下のとおり仮説を設定した。

仮説 1: 観光地プロモーション動画において、人間ナレーションは AI ナレーションに比べて視聴者の訪問意図をより高める。

アクセントは、言語使用における受容されやすさなどの標準性の高さに基づいて分類され、広く理解される標準アクセントと、環境の違いにより地域ごとに多様な形で存在する地方アクセントに分けられる (Xie et al., 2025)。標準アクセントは、言語における標準的な発音として位置づけられており、地方アクセントは特定の地理的地域に固有であるが、完全な方言や他言語とは異なり、音韻やプロソディといった音声上の違いにとどまるため、コミュ

ニケーション上の大きな障壁を生み出すことはほばない (Xie et al., 2025; Li et al., 2025)。

人間の音声には、生まれ育った地域に由来する多様な地方アクセントが含まれており、プロソディを通じて感情やニュアンスを伝達する (Clopper & Smiljanic, 2011; Rodero, 2017)。そして、観光地の地域的文脈と整合する地方アクセントの使用は、知覚された独特性を高め、ポジティブな印象の形成につながることを指摘されている (Xiong et al., 2025)。

初期の AI 音声は標準アクセントで設計されており、人間の声を持つ地方アクセント特有の情緒的ニュアンスが十分に反映されていないと考えられる。しかし、AI に地方アクセントを取り入れることで、より人間らしい感情表現が可能になると期待される。これらを踏まえ、AI ナレーションが紹介する観光地と同じ地方アクセントで話すことで、視聴者の観光地に対する好意的評価を促進する可能性がある。よって、下記の仮説を設定した。

仮説 2: 観光地プロモーション動画において、AI ナレーションが標準アクセントよりも、紹介される観光地と同じ地方アクセントで話す方が、視聴者の訪問意図が高まる。

2. 知覚された能力と知覚された擬人性による媒介効果

ナレーションにおける標準アクセントと地方アクセントが視聴者の訪問意図に与える影響メカニズムを明らかにするために、本稿では、Fiske et al. (2002) によって提唱されたステレオタイプ内容モデル (Stereotype Content Model) に基づき、知覚された有能さ (perceived competence) を媒介変数として位置付けた。

ステレオタイプ内容モデルは、人々が他者を評価する際に、親切さや誠実さなどの対人意図に関わる「暖かさ」と、知性や能力などの目標達成に関わる「有能さ」という 2 つの次元に基づいて判断する傾向を説明するものである (Fiske et al., 2002; Su et al., 2025)。Xie et al. (2025) は、アクセントが話者の発話内容ではなく、声の特徴を通じて社会的印象に影響を与えると指摘していた。そして、標準アクセントは観光などのサービス場面において、有能性や知性のステレオタイプと結びついており、観光客に話者の専門性や信頼性の知覚を高める傾向がある (Su et al., 2025)。これらを踏まえ、次の仮説を導出する。

仮説 3: AI ナレーションにおける標準アクセントの使用は、知覚された有能さを高め、それを介して視聴者の訪問意図に正の影響を与える。

しかし、ステレオタイプ内容モデルは本来、人間に対する認知を説明する枠組みであるため、非人間主体である AI にはそのまま適用できるかという懸念もある。そこで本研究では補完的な視点として、知覚された擬人性 (perceived anthropomorphism) に着目する。

擬人性とは、人間以外の主体や物体に、人間のような身体的特徴、動機、行動、感情など

を付与することを指す (Wang et al., 2025)。Li et al. (2024) は、AI が地方アクセントを用いることで、視聴者は自らの文化的背景が理解されていると感じ、親近感やソーシャルプレゼンスが高まると指摘している。こうした心理的效果により、人々は AI との相互作用に対してより積極的な態度をとるようになる。一方、Wang et al. (2025) は、コンテンツが AI によって生成されたことが明示されると、消費者は当該コンテンツに対して人間らしさ、すなわち擬人性を感じにくくなり、その結果、ソーシャルプレゼンスや信頼感が低下し、最終的に購買意図にも負の影響を及ぼす可能性があるとは指摘している。とりわけ、初期設定となる AI 標準アクセントの場合は、地方アクセントに比べて AI らしさを強調し、擬人性が低く評価されると考えられる。これらを踏まえ、以下の仮説を導出する。

仮説 4: AI ナレーションにおける標準アクセントの使用は、知覚された擬人性を下げ、それを介して視聴者の訪問意図に負の影響を与える。

仮説 5: AI ナレーションにおける標準アクセントの使用は、知覚された有能さを通じて擬人性に影響を及ぼし、視聴者の訪問意図に間接的な影響を与える。

III. 研究方法

本研究では、2 つの実験を通じて仮説の検証を行った。両実験において共通して使用した動画は、中国の東北地方を紹介する観光プロモーション動画である。いずれの動画も視覚的な構成は統一されており、ナレーション以外の差異は存在しない（ただし、動画の左上には「本動画は人間／AI によるナレーションです」と明記した）。また、ナレーターの顔や姿は映っておらず、話者の外見等による先入観の影響を最小限に抑えている。なお、動画のナレーション（人間音声）は標準中国語証書を持つ人と東北地方で生まれ育った人に依頼して録音し、AI 音声および映像編集は動画編集ソフト「CapCut」を用いて行った。研究 1 では、ナレーションの主体（人間 vs. AI）および使用言語（標準アクセント vs. 東北地方アクセント）が、観光地に対する訪問意図にどのような影響を及ぼすかを明らかにすることを目的として、オンライン実験を 2 回に分けて実施した。研究 2 では、AI ナレーションにおける標準語と地方アクセントの使用が視聴者の訪問意図に与える影響の背後にある心理的メカニズムに着目し、その媒介プロセスを検討した。

1. 研究 1

(1) 方法

研究 1 は、2025 年 3 月 27 日から 28 日にかけて、オンライン調査プラットフォーム「Credamo」のサンプリングサービスを通じて実施された。ナレーションの主体を「人間」

と「AI」の2条件で設定し、それぞれ213名と277名の被験者から回答を得た。操作チェックとして、動画視聴後に「最後に紹介された観光地はどこか」という設問を設け、正確に回答できなかった被験者は除外した。最終的に、200名（人間）および250名（AI）の有効回答者数を収集した。

その後、2025年3月29日から30日にかけて、同一の被験者を対象とした追跡調査を実施した。今回はナレーションを東北地方アクセントに変更し、再度同じ動画を提示した。第2回調査では、176名（人間）および229名（AI）の有効回答を回収した。

「訪問意図」はKim et al., (2020)を参照に、「動画で紹介された場所を訪れたいと思う」「近い将来その場所を訪れる予定がある」「できるだけ早くその場所を訪れたいと考えている」「その場所を訪れるために時間とお金を使ってもよいと思う」の4項目（ $\alpha=.726$, $CR=.830$, $AVE=.552$ ）を7ポイントのリッカート尺度で測定し、平均値を求めた。

(2) 結果と考察

本研究では、IBM SPSS Statistics 29を用いて分析を行った。まず仮説1に対して、ナレーション主体（人間 vs. AI）による訪問意図の差異を独立したt検定で比較した。その結果、標準アクセントを用いた条件においては、AIナレーション（ $M_{AI} = 6.104$, $SD_{AI} = .544$ ）の方が人間ナレーション（ $M_{人間} = 6.099$, $SD_{人間} = .446$ ）より平均値がわずかに高かったものの、有意な差は確認されなかった（ $t = -.087$, $p = .931$ ）。

さらに、アクセントの違いによって結果が変化する可能性を考慮し、東北地方アクセントを使用した条件における人間とAIナレーションの比較も行った。その結果、東北地方アクセントを用いた人間ナレーション（ $M_{人間} = 6.092$, $SD_{人間} = .558$ ）とAIナレーション（ $M_{AI} = 6.043$, $SD_{AI} = .708$ ）の間にも有意な差は確認されなかった（ $t = .791$, $p = .430$ ）。これにより、仮説1は支持されなかった。

次に、仮説2を検証するため、AIナレーションにおいて、アクセントの違い（標準アクセント vs. 東北地方アクセント）を対応のあるt検定で比較した。その結果、東北地方アクセントの使用は訪問意図をやや低下させる傾向がみられた（ $MD = .061$, $t = 1.832$, $p = .068$ ）が、5%水準には達しておらず、仮説2も支持されなかった。

そのほか、人間ナレーションに対しても、同じく対応のあるt検定で追加分析を行った。その結果、標準アクセントと東北地方アクセントの間で訪問意図に有意な差はなかった（ $MD = .007$, $t = .184$, $p = .854$ ）。

とりわけ、地方アクセントが標準アクセントと比べて訪問意図に明確な影響を及ぼすとは言えなかった。以上を踏まえ、次章ではAIナレーションに焦点を当て、音声の印象が訪

問意図に与える心理的メカニズムを検討する。

2. 研究 2

(1) 方法

研究 2 では、研究 1 の結果を踏まえ、標準アクセントと東北地方アクセントの間に訪問意図の有意差は確認されなかったが、すべてのアクセントにおいて同様の結果が得られるかどうかを確認するため、地理的に整合性の低い香港アクセントを追加し、比較対象とした。

研究 2 は、2025 年 7 月 13 日から 16 日にかけて、オンライン調査プラットフォーム「Credamo」のサンプリングサービスを通じて実施された。標準アクセント、東北地方アクセント、香港アクセント、それぞれ 250 名、合計 750 名の被験者から回答を得た。操作チェックとして、動画視聴後に「最後に紹介された観光地はどこか」という設問のほか、提示されたナレーションが想定通りのアクセントとして知覚されていたかを「このナレーションは標準的な標準／東北地方／香港アクセントだと思いますか?」という設問(7 段階尺度)でスクリーニングした。その回答が 4 未満であった場合、該当の言語として十分に知覚されていないと判断し、当該回答を除外した。最終的に有効サンプル数は、標準アクセントで 245 件、東北地方アクセントで 239 件、香港アクセントで 203 件だった。

媒介変数となる「知覚された有能さ」は、Fiske et al. (2002) と Kervyn et al. (2012) に参照し、「AI ナレーションは、観光情報を明確に伝えていると感じた」「専門的かつ効率的に情報を伝えていた」「聞いていて専門的な印象を受けた」「コンテンツを的確に伝える能力があるように思えた」「自信にあふれ、流暢であった」「このような観光地紹介の役割を十分に果たせるように感じた」の 6 項目 ($\alpha = .934$, $CR = .948$, $AVE = .752$) を 7 ポイントのリッカート尺度で測定し、平均値を求めた。

「知覚された擬人性」は Schuetzler et al. (2020) に参照し、「人工的／自然的」「機械的／人間的」「不流暢的／流暢的」「不自然／自然」の 4 項目 ($\alpha = .909$, $CR = .937$, $AVE = .787$) を 7 ポイントのリッカート尺度で測定し、平均値を求めた。

そして、従属変数となる「訪問意図」は研究 1 と同様に Kim et al. (2020) を参照した項目を使った ($\alpha = .923$, $CR = .945$, $AVE = .813$)。

(2) 結果と考察

仮説 2 のさらなる検証として、AI ナレーションが用いるアクセントの違いが視聴者の訪問意図に与える影響を検討した。具体的には、三つの音声条件における訪問意図の平均値に差を一元配置分散分析により比較した。その結果、グループ間には有意な差が確認した (F

(2, 684) = 8.089, $p < .001$, $\eta^2 = .023$)。事後比較の結果、香港アクセントは、標準アクセント ($MD = -.310$, $p = .015$) および東北地方アクセント ($MD = -.439$, $p < .001$) と比較して有意に訪問意図が低かった。一方、標準アクセントと東北地方アクセントの間には統計的な有意差は確認されなかった ($MD = -.130$, $p = .440$)。これらの結果は、地理的整合性の低いアクセントは訪問意図を抑制する可能性があることを示唆しており、仮説 2 の知見を補完するものである。

この結果を踏まえ、仮説 3~5 の分析においては、音声条件を「標準アクセント」と「東北地方アクセント」に限定し、訪問意図の影響メカニズムを IBM SPSS Statistics PROCESS (model 6; 5,000 bootstrap samples; Hayes 2013) で分析した。独立変数は「標準アクセント = 1, 東北地方アクセント = 0」にダミーコーディングし、従属変数は訪問意図とした。

まず直接効果として、AI による標準アクセントの使用は、訪問意図に対して有意な負の影響を及ぼすことが確認された ($b = -.243$, $SE = .080$, $t(477) = -3.060$, $p = .002$)。媒介効果として、AI に対する有能さ ($b = .117$, $SE = .057$, 95% CI = [.008, .231])、および擬人性 ($b = -.031$, $SE = .017$, 95% CI = [-.069, -.004]) を測定した結果、仮説 3 と 4 が支持された。総じて、これらの媒介経路は互いに逆方向の影響を及ぼしていたものの、標準アクセントの使用が知覚された有能さを高め、さらにそれが擬人性を通じて訪問意図に影響を与えるという二重媒介効果も有意であることが確認された ($b = .016$, $SE = .011$, 95% CI = [.000, .042])。よって、仮説 5 が支持された。

IV. まとめと今後の課題

1. 本研究のまとめ

本研究では、観光プロモーション動画におけるナレーションの主体 (人間 vs. AI) および音声のアクセント (標準アクセント vs. 地方アクセント) が訪問意図に与える影響を明らかにすることを目的として、2つの実験を実施した。その結果、ナレーション主体やアクセント使用自体には明確な影響が見られなかったものの、AI 音声のアクセントが視聴者の知覚された有能さや擬人性に作用し、それが訪問意図に間接的に影響を及ぼすことが示された。AI ナレーションに標準アクセントを用いる場合、有能さを高める一方で、擬人性感知を低下させるという相反する心理プロセスが同時に生じることが明らかになった。逆に地方アクセントを用いる場合は、有能さの低下という負の影響が生じる可能性があるものの、擬人性感知が高まり、訪問意図を促進する正の影響も生まれることが示唆された。

2. 本研究の示唆

本研究では、これまで主に人間の話者を対象として用いられてきたステレオタイプ内容モデルを、AI ナレーションという非人間的主体に適用し、観光プロモーションの文脈でその有用性を検討した。分析の結果、標準アクセントの使用はAI の知覚された有能さを高め、訪問意図に正の間接効果をもたらす一方、AI の知覚された擬人性を低下させ、訪問意図を抑制する負の媒介効果も示された。これにより、AI ナレーションにおいては、アクセントが視聴者の心理に複数の方向から作用するというメカニズムが明らかになった。

この知見は、今後の非人間的メディア主体の研究に新たな視点を与えるものである。また、観光プロモーションにおけるAI ナレーションの活用に関しても、標準アクセントあるいは地理的整合性の高い地方アクセントを適切に用いることで、視聴者に肯定的な印象を与えることが示唆された。これにより、AI などの非人間主体において知覚された有能さと擬人性がトレードオフの関係にあることを実証的に示した点に、本研究の重要な理論的貢献がある。

また、実務的には、本研究の結果は、従来強調されてきた人間ナレーションの優位性に対して (Rodero, 2017; Wang et al., 2024), AI ナレーションも一定の条件下で人間と同等の説得効果を持つ可能性を示した (研究 1)。観光マーケティング担当者やプロモーション動画制作企業に対して、AI を人間ナレーションの代替手段として積極的に活用し、ナレーションには標準アクセントまたは地理的整合性の高い地方アクセントのいずれを選択しても一定の宣伝効果が期待できることが示された。

3. 今後の課題

本研究には課題も多く残されている。第 1 に、調査対象者が対象地域を訪問した経験の有無を測定していない点が挙げられる。訪問経験の有無により、アクセントやナレーション主体の影響の受け方が異なる可能性があるため、今後の研究ではこれを制御変数として考慮し、より精緻な分析が求められる。

第 2 に、本研究では「知覚された有能さ」および「知覚された擬人性」の媒介効果を検証したが、いずれも部分的な媒介にとどまったことから、他にも影響を及ぼす心理的要因が存在する可能性がある。今後は、信頼性や親近感、AI への態度やテクノロジー受容度といった変数も含めた包括的なモデル構築が求められる。

第 3 に、本研究は中国東北地方を対象としたため、地域的・文化的な一般化可能性に限界がある。今後は、他の地域や異なる文化圏を対象とすることで、地理的整合性やアクセントの効果が文化間で一貫するかを検証し、理論の汎用性を高めることが期待される。

引用文献

- Clopper, C. G., & Smiljanic, R. (2011). Effects of gender and regional dialect on prosodic patterns in American English. *Journal of phonetics*, 39(2), 237-245. DOI: 10.1016/j.wocn.2011.02.006
- Fiske, S. T., Cuddy, A. J. C., Glick, P., & Xu, J. (2002). A model of (often mixed) stereotype content: Competence and warmth respectively follow from perceived status and competition. *Journal of Personality and Social Psychology*, 82(6), 878-902. DOI: 10.1037/0022-3514.82.6.878
- Gan, J., Shi, S., Filieri, R., & Leung, W. K. (2023). Short video marketing and travel intentions: The interplay between visual perspective, visual content, and narration appeal. *Tourism Management*, 99, 104795. DOI: 10.1016/j.tourman.2023.104795
- Hayes, A. F. (2013). *Introduction to mediation, moderation, and conditional process analysis: A regression-based approach*. The Guilford Press.
- Kervyn, N., Fiske, S.T., & Malone, C. (2012). Brands as intentional agents framework: How perceived intentions and ability can map brand perception. *Journal of Consumer Psychology*, 22, 166-176. DOI: 10.1016/j.jcps.2011.09.006
- Kim, M. J., Lee, C.-K., & Jung, T. (2020). Exploring Consumer Behavior in Virtual Reality Tourism Using an Extended Stimulus-Organism-Response Model. *Journal of Travel Research*, 59(1), 69-89. DOI: 10.1177/0047287518818915
- Li, C., Qu, Z., & Yang, Y. (2025). Localizing matters: The effect of AI accent on tourist travel intention. *Journal of Destination Marketing & Management*, 36, 100992. DOI: 10.1016/j.jdmm.2025.100992
- Li, J., Liu, S., Wang, Y., Wang, Q., & Wong, J. W. C. (2024). Should chatbots use dialects? Exploring the influence mechanism of chatbot language form on value co-creation intention. *Current Issues in Tourism*, 1-18. DOI: 10.1080/13683500.2024.2431520
- Poushneh, A., Vasquez-Parraga, A., & Gearhart, R. S. (2024). The effect of empathetic response and consumers' narcissism in voice-based artificial intelligence. *Journal of Retailing and Consumer Services*, 79, 103871. DOI: 10.1016/j.jretconser.2024.103871
- Rao Hill, S., & Tombs, A. G. (2022). When does service employee's accent matter? Examining the moderating effect of service type, service criticality and accent service congruence. *European Journal of Marketing*, 56(7), 1985-2013. DOI: 10.1108/EJM-09-2020-0707
- Rodero, E. (2017). Effectiveness, attention, and recall of human and artificial voices in an advertising story. Prosody influence and functions of voices. *Computers in Human Behavior*, 77, 336-346. DOI: 10.1016/j.chb.2017.08.044
- Schuetzler, R. M., Grimes, G. M., & Scott Giboney, J. (2020). The impact of chatbot conversational skill on

- engagement and perceived humanness. *Journal of Management Information Systems*, 37(3), 875-900.
DOI: 10.1080/07421222.2020.1790204
- Su, L., Gong, L., & Huang, Y. (2025). Humorous or serious? The interaction effect of language style and tourism activity type on tourist well-being. *Journal of Sustainable Tourism*, 33(3), 520-547. DOI: 10.1080/09669582.2024.2354883
- Wang, X., Zhang, Z., & Jiang, Q. (2024). The effectiveness of human vs. AI voice-over in short video advertisements: A cognitive load theory perspective. *Journal of Retailing and Consumer Services*, 81, 104005. DOI: 10.1016/j.jretconser.2024.104005
- Wang, Y., Sauka, K., & Situmeang, F. B. I. (2025). Anthropomorphism and transparency interplay on consumer behaviour in generative AI-driven marketing communication. *Journal of Consumer Marketing*, 42(4), 512-536. DOI: 10.1108/JCM-04-2024-6806
- Xie, P., Long, X., Mu, W., Wang, J., & Wu, Y. (2025). Standard Accent vs. Local Accent: Exploring the Role of Doctors' Accent Characteristics in Online Medical Services. *Health Communication*, 40(7), 1363-1380. DOI: 10.1080/10410236.2025.2499097
- Xiong, X., Sun, D., Wong, I. A., & Lian, Q. L. (2025). Do Tourists Prefer Dialectal Service? The Role of Processing Fluency, Distinctiveness, and Cultural Learning Cues. *Journal of Travel Research*. DOI: 10.1177/00472875251332953