

音声聴取における Deceptive Pattern: 発話速度の操作による非母語話者の選択誘導可能性の検証

木下 裕一朗^{1,a)} 中村 聡史¹

概要: ユーザの行動を誘導する Deceptive Pattern (Dark Pattern) は、音声ユーザインタフェースにも存在する。これまで、音声特徴の操作によりユーザの選択が誘導される可能性が示されているが、先行研究はユーザの母語による音声提示という状況で検証を行っている。しかし、非母語を用いて音声ユーザインタフェースを使用する状況下では、誘導効果に変化する可能性がある。そこで本研究では、非母語による音声インタラクションに焦点を当て、音声の発話速度の意図的な操作による非母語話者の選択誘導可能性を検証した。また、母語と非母語において発話速度操作による誘導効果の比較を行った。実験の結果、母語、非母語の両方の場合において、発話速度の操作による選択の誘導はみられなかった。しかし、言語の違いはユーザの選択傾向を変化させる可能性や、ユーザの意図する選択を妨げる可能性が示唆された。

1. はじめに

サービス提供者の利益となるように、ユーザを騙し行動を誘導するデザインは Dark Pattern または Deceptive Pattern^{*1} (DP) と呼ばれており、多くのウェブサイトやモバイルアプリケーションで使用されている [2–4]。例えば、ショッピングサイトで「定期購入」のボタンをユーザに選ばせるように目立たせたり、サブスクリプションサービスの解約を困難にしたりするデザインは DP に該当する。

これまで、主に GUI を対象として DP の発見 [5–7] や分類 [2,8,9]、自動検出 [10,11] などについて研究が行われてきた。しかし、DP は GUI だけでなく音声ユーザインタフェース (VUI) にも存在することが明らかになっており、有料機能利用のため課金するよう VUI が勧めてきたり、VUI によりアプリケーションの終了を妨害されたりした事例が報告されている [12]。これらはスマートフォンやスマートスピーカといったユーザが個人で使用する VUI における DP であるが、VUI が今後ますます普及し、様々な場所でサービス特化したものとして利用されるようになると、DP の使用はさらに増え、これまで見られなかった新たな DP が出現する可能性も考えられる。今後ユーザを保護するために、VUI における潜在的な DP について早期

に検討することが重要である。

VUI は音声入力と音声出力を備えたユーザインタフェースであり、音声入力する場合にユーザの選択が誘導される可能性がある [13] ことから、我々はこれまで音声入力における言語の壁を悪用した DP に着目してきた [14]。GUI では、ユーザの使用言語と異なる言語を用いて理解を困難にし、ユーザの意図した行動を妨げる DP が存在する [4]。このような DP は VUI にも適用される可能性があり、非母語による音声インタラクションは、VUI が多言語対応していない場合だけでなく、意図的に生み出される可能性がある。言語の壁を悪用した DP は音声入力だけでなく、VUI による音声出力においても使用される可能性があり、ユーザの望む行動が妨げられ、誘導されてしまう恐れがある。

そこで本研究では、VUI の音声出力における言語の壁を悪用した DP に着目し、VUI がユーザの非母語を用いて音声提示を行う状況で意図的に異なる発話速度によって選択肢を提示することにより、ユーザの選択が誘導されるか検証する。また、ユーザの母語においても発話速度の操作による選択誘導可能性を検証することで、言語の違いによる選択誘導効果の差異を明らかにする。

本研究における問いを以下に示す。

RQ1 VUI によって選択肢を音声提示する際に、発話速度を意図的に操作することでユーザの選択は誘導されるか？

RQ2 VUI が選択肢を音声提示するときの使用言語が、ユーザの母語であった場合と、非母語であった場合でユーザの選択に及ぼす影響は異なるか？

¹ 明治大学
Meiji University

^{a)} kinoshita@nkmr-lab.org

^{*1} 2010 年に Harry Brignull [1] はユーザを騙し誘導するデザインを Dark Pattern と名付けたが、後に人種の問題を考慮して呼称を Deceptive Pattern (または Deceptive Design) に変えた。本研究では、Deceptive Pattern という用語を使用する。

2. 関連研究

2.1 Deceptive Pattern

DP はアメリカにおける 240 の人気モバイルアプリケーションのうち約 95% に使用されており [3], 日本においても 200 のモバイルアプリケーションのうち 93.5% が 1 つ以上の DP を含んでいたことが報告されている [4]. ユーザは DP に対して苛立ち [15] や嫌悪 [16] を抱く一方で, DP を見つけられなかったり [3], DP の存在を意識していたとしても行動が誘導されたりする [17] ことが明らかになっている. また, DP の特性を理解することで, ユーザは DP を含むものとそうでないもののある程度識別可能であるが, 識別は容易でないことが示されている [18].

DP は GUI だけでなく VUI にも存在する [12, 19, 20] ことが明らかになっている. また, GUI で操作するスクリーンベースのインタフェースだけでなく, ホームロボット [21] や IoT デバイス [22], 物理空間のデザイン [23] にも潜んでいることが明らかになっている. しかし, これまで GUI 以外における DP について十分に検討されておらず, 異なるインタフェースでどのような DP が出現するか, またユーザにどのような影響を及ぼす可能性があるのか, さらに明らかにする必要がある. 本研究では VUI における潜在的な DP に焦点を当て, 音声の発話速度を意図的に操作することによって非母語話者の選択行動が誘導されるかを検証する.

2.2 音声ユーザインタフェースにおける選択の誘導

VUI では, 音量やピッチ, 速度などの音声特徴を操作することによりユーザの選択を誘導できる可能性がある [12]. Dubiel ら [19] は, 合成音声の質による選択誘導可能性に着目し, 発話速度が速く, ピッチ変化の大きい音声で提示した選択肢が選ばれやすいことを明らかにしている. また, 音声のピッチだけでなく, トーンやイントネーションなどの操作が, ユーザの選択や情報の信頼性評価に強く影響することも明らかになっている [20]. さらに, アプリケーション内のユーザへの許可リクエストを音声提示する際に, 同意と拒否の 2 つの選択肢の発話速度を変化させると, ユーザは同意を選びやすくなる傾向が示されている [24]. このような音声特徴の操作による誘導は, ユーザにとって気づきづらい可能性が示唆されている [19, 20, 24].

プライバシーに関する同意選択において発話速度の操作による影響を検証した先行研究 [24] に対して, 本研究は, ユーザの嗜好がより強くあらわれる状況での選択に焦点を当てている点や, 発話速度の操作による選択誘導効果がユーザの母語と非母語の場合で異なるかを比較し, 言語の違いによる影響を検証する点で異なる.

2.3 非母語を用いた音声ユーザインタフェースの操作

VUI は多文化背景をもつユーザに十分に対応できておらず [25], ユーザが非母語を用いて VUI を操作する場合には困難が伴い [26], 精神的作業負荷が大きい [27] ことが知られている. Amazon の Alexa を用いて, 英語の母語話者と, 韓国語を母語とする英語の非母語話者の音声認識精度の違いを検証した研究 [28] では, 英語の母語話者の音声認識精度が 98% であったのに対し, 非母語話者の音声認識精度は 55% であったことが示されている.

言語の違いを悪用した DP は, ユーザの音声入力だけでなく, VUI による音声出力においても使用される可能性がある [12, 19, 20, 24]. 非母語のリスニングでは, 知っている単語であっても聞き取ることができなかったり, 単語を聞き取れたとしても意図されたメッセージを理解できなかったりするといった問題があるため [29, 30], 非母語により音声提示される場合に, 音声特徴の操作によってユーザはより簡単に誘導されてしまう可能性が考えられる. 本研究では, 発話速度の操作による選択誘導可能性をユーザの母語の場合と非母語の場合において検証し, 結果を比較することで, 言語の違いによる選択への影響を明らかにする.

3. 選択の偏りが生じない質問と選択肢の選定

発話速度の操作による選択誘導可能性を検証するうえで, 発話速度以外の要因によって選択に偏りが生じないようにする必要がある. そこで本章では, 同一の発話速度で選択肢を音声提示する実験を行うことで, 発話速度以外の要因によって選択の偏りが起こらない選択肢ペアを選定する.

本研究では, ユーザの母語と非母語において発話速度操作による選択誘導可能性を検証するため, 本章の実験においても母語条件と非母語条件の 2 条件を設けて実験を実施した. 実験では, 英語を非母語とする日本人の参加者を対象とし, 母語条件を日本語, 非母語条件を英語とした.

3.1 実験設計

実験用に実装したシステム (図 1) を用いて, 15 問の各質問に対して 2 択の選択肢から 1 つを選び回答してもらう実験を行った.

3.1.1 実験シナリオ

本研究では, 選択が求められる多様な状況を設定可能で, かつ実験参加者が非母語による対話に直面する状況を想像しやすくなるを考え, 過去に実施した選択実験 [14] と同様に, 海外旅行というシナリオを採用した. 具体的には, 「あなたは, これから 1 人で海外旅行に行くとします. その中で 15 個の出来事があり, それぞれの出来事に対して質問が提示されるため, 2 つの選択肢から 1 つを選んで質問に答えてください.」と参加者に指示し, 海外旅行時における選択というシナリオで各質問にクリック選択で回答しても

らった。

3.1.2 選択肢内容とアンケート

選択肢は、英語表記の場合に5～7語となるように決定し、同一質問における選択肢間で単語数の違いが生じないようにした。また、使用する英単語のほとんどは、日本の高校卒業時まで学習する英単語とした。

実験後に、音声提示された選択肢の内容を実験参加者が理解できたかを確認するためにアンケートを実施した。また、アンケートでは、実験参加者が注意深く文章を読んでいるかを確かめるために注意力確認の質問を行い、アンケートの最後に実験参加者の性別および年齢を取得した。

3.1.3 使用する音声

選択肢提示に用いる音声の作成には、Google Cloud Text-to-Speech を利用した*2。Google Cloud Text-to-Speech で利用可能な音声のうち、母語条件ではja-JP-Wavenet-Bを、非母語条件ではen-US-Wavenet-Hを使用した。商用の音声アシスタントには女性音声が多く採用されている[31]ため、本研究でも女性の合成音声を用いた。

本実験では、2択の選択肢の両方を人間の平均的な発話速度で提示した。日本語の平均的な発話速度は8.0 モーラ/秒であり[32]、英語の平均的な発話速度は152～170WPMである[33]ことが知られているが、英語の母語話者の平均的な発話速度は、非母語話者にとって速く、発話内容を正しく認識できない可能性がある。そこで本研究では、著者が所属する研究室内で、英語を非母語とする日本人学生13名を対象にパイロットテストを実施し、非母語話者でも発話内容を理解できると考えられる発話速度に設定した。各条件に使用した音声の平均発話速度と、基本周波数の平均および標準偏差について表1に示す。

3.2 実験手順

質問は1問ずつ画面にテキストで表示した。母語条件では「選択肢1」という言葉の後に1つ目の選択肢を音声提示し、続けて「選択肢2」という言葉の後に2つ目の選択肢を音声提示した。非母語条件では、「Option 1」という言葉の後に1つ目の選択肢を音声提示し、続けて「Option 2」という言葉の後に2つ目の選択肢を音声提示した。2つの選択肢を音声で提示する順序は、試行ごとにランダムに決定した。選択肢の内容は画面上に表示せず、「選択肢1/Option 1」「選択肢2/Option 2」と表記した。

実験参加者は、15問の質問に対して選択を行った後、アンケートに回答した。アンケートへの回答が終わると、実験用システムは参加者に対して実験の終了を通知した。

3.3 実験参加者の募集

実験参加者の募集には、Yahoo!クラウドソーシング(YCS)

表1 実験に用いた合成音声の平均発話速度と、基本周波数の平均および標準偏差。

	平均発話速度	基本周波数の平均 (標準偏差)
母語条件 (日本語)	8.1 モーラ/秒	255.7Hz ($SD = 48.0$)
非母語条件 (英語)	139WPM	213.3Hz ($SD = 37.1$)

を利用した。YCSは、広く利用されているクラウドソーシングサービスの一つであり、高いタスク完了率とアカウントの信頼性から、その有用性が示されている[34]。YCSでは、対象を男性もしくは女性に限定して参加者を募集することができるため、参加者の男女比が同程度となるように、男性限定と女性限定の2回に分けてそれぞれ同数の参加者を募集した。実験に要する時間は5分程度と予想されたため、地域の賃金基準に基づき、実験参加に対する報酬として約100円分のPayPayポイントを参加者のアカウントに付与した。母語条件、非母語条件ともに、実験参加者の対象は英語を非母語とし、日本語を母語とする成人とした。また、先に実施する母語条件の実験に参加した者は、後に実施する非母語条件の実験に参加できないようにした。

3.4 結果

母語条件の実験は2025年11月12、13日にかけて実施し、非母語条件の実験は2025年11月19日に実施した。各条件でそれぞれ200名の実験参加者を募集した。注意力確認の質問に対する回答が適切でなかった参加者は分析対象として適切でないと考え除外し、母語条件では191名(男性96名、女性93名、未回答2名)、非母語条件では192名(男性95名、女性96名、未回答1名)を分析対象とした。

各条件における、質問ごとの選択結果を表2に示す。本研究では有意水準を0.05とし、母語、非母語条件それぞれで質問ごとに選択が偏っているか二項検定を実施した。多重比較のため、 p 値はBonferroni法により補正した。二項検定の結果、母語条件では7問(Q4, Q5, Q7, Q8, Q9, Q10, Q11)、非母語条件では10問(Q3, Q5, Q6, Q7, Q8, Q10, Q11, Q12, Q13, Q15)において選択の偏りが生じなかった($p > .05$)。

次に、音声再生回数と、音声提示された選択肢の内容をどの程度理解できたかというアンケート項目に対する7段階リッカート尺度(1:まったく理解できなかった, 7:すべて理解できた)による回答結果を表3に示す。音声の再生回数については、母語条件と非母語条件の間で大きな差はみられなかった。選択肢内容の理解度については、母語条件と比べて非母語条件の方が低いものの、7段階のうち平均4.33($SD = 1.44$)であったため、合成音声の発話速度が過度に速かったり、選択肢の英単語が過度に難解であったりした可能性は低いと考えられる。

4章の実験では、本章の実験で母語、非母語条件ともに選択の偏りが生じなかった5つ(Q5, Q7, Q8, Q10, Q11)

*2 実験に使用した音声は、<https://github.com/yuichiroooo/exp202511> で公開している。

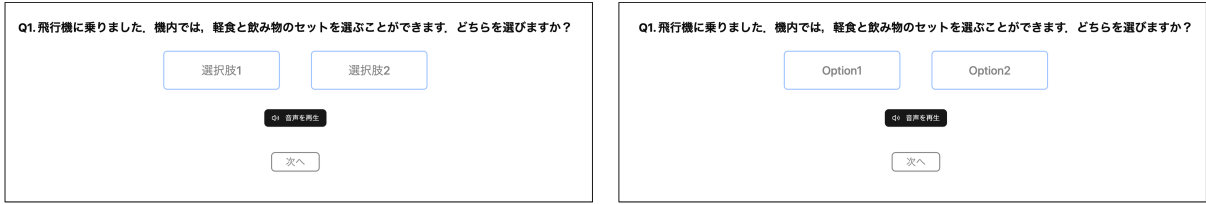


図 1 実験に用いたシステムのスクリーンショット。左は母語条件におけるシステム、右は非母語条件におけるシステムとなっており、選択肢ボタンの表記が異なっている。また、再生ボタンを押すと、母語条件では日本語音声で選択肢が提示され、非母語条件では英語音声で選択肢が提示される。

表 2 母語条件と非母語条件における各質問の選択結果。表中で選択肢を日本語で示しているものは母語条件における結果であり、英語で示しているものは非母語条件における結果である。選択回数に偏りがあるか二項検定を行い、多重比較のため p 値の補正には Bonferroni 法を用いた。

質問文（一部省略）	選択肢（選択数）	p 値
1. 機内食として、どちらのセットを選びますか？	チョコクッキーとコーヒー ($N = 119$) / チョコクッキーと紅茶 ($N = 72$) Chocolate cookies with a cup of coffee ($N = 123$) / Chocolate cookies with a cup of tea ($N = 69$)	.012 .002
2. レストランでは、どちらの料理を食べますか？	グリルチキンとご飯とサラダ ($N = 140$) / グリルフィッシュとご飯とサラダ ($N = 51$) Grilled chicken with rice and salad ($N = 148$) / Grilled fish with rice and salad ($N = 44$)	< .001 < .001
3. どちらの自然遺産を見に行きたいですか？	滝のある美しい自然遺産 ($N = 147$) / 森のある美しい自然遺産 ($N = 44$) Beautiful natural heritage with waterfalls ($N = 102$) / Beautiful natural heritage with forests ($N = 90$)	< .001 1.000
4. ホテルの部屋はどちらを選びますか？	庭が見える部屋 ($N = 104$) / 街が見える部屋 ($N = 87$) Room with a garden view ($N = 121$) / Room with a city view ($N = 71$)	1.000 .006
5. どちらのショートコンサートに参加しますか？	ジャズのショートコンサート ($N = 99$) / クラシックのショートコンサート ($N = 92$) Short concert with jazz music ($N = 105$) / Short concert with classical music ($N = 87$)	1.000 1.000
6. 朝食では、どちらのパンを食べたいですか？	バターとジャムを添えたクロワッサン ($N = 144$) / バターとジャムを添えた食パン ($N = 47$) Croissant with butter and jam ($N = 92$) / Bread with butter and jam ($N = 100$)	.001 1.000
7. どちらの博物館に行きたいですか？	歴史を学べる地元の博物館 ($N = 112$) / 文化を学べる地元の博物館 ($N = 79$) Local museum to learn about history ($N = 99$) / Local museum to learn about culture ($N = 93$)	.305 1.000
8. 昼ご飯はどちらのパスタを食べたいですか？	トマトソースとベーコンのパスタ ($N = 115$) / クリームソースとベーコンのパスタ ($N = 76$) Pasta with tomato sauce and bacon ($N = 110$) / Pasta with cream sauce and bacon ($N = 82$)	.087 .766
9. どちらの庭園に行きたいですか？	バラと木のある庭園 ($N = 102$) / ラベンダーと木のある庭園 ($N = 89$) Garden with roses and trees ($N = 118$) / Garden with lavender and trees ($N = 74$)	1.000 .028
10. どちらのレストランに行きたいですか？	洋食を出す伝統的レストラン ($N = 115$) / 洋食を出すモダンレストラン ($N = 76$) Traditional restaurant serving Western foods ($N = 109$) / Modern restaurant serving Western dishes ($N = 83$)	.087 1.000
11. どちらの夜景を見たいですか？	高いタワーからの街の景色 ($N = 107$) / 静かな海辺からの街の景色 ($N = 84$) City view from a tall tower ($N = 110$) / City view from a quiet seaside ($N = 82$)	1.000 .766
12. 朝食のメニューはどちらを選びますか？	卵とトーストの朝食 ($N = 118$) / フルーツとヨーグルトの朝食 ($N = 73$) Breakfast with eggs and toast ($N = 102$) / Breakfast with fruits and yogurt ($N = 90$)	.021 1.000
13. どちらのオペラを観に行きたいですか？	古代を舞台にしたオペラ ($N = 121$) / 現代を舞台にしたオペラ ($N = 70$) Opera set in ancient times ($N = 80$) / Opera set in modern times ($N = 112$)	.004 .375
14. どちらのカフェに行きますか？	ケーキが美味しい人気のカフェ ($N = 135$) / パフェが美味しい人気のカフェ ($N = 56$) Popular cafe with delicious cakes ($N = 126$) / Popular cafe with delicious parfaits ($N = 66$)	< .001 < .001
15. 機内では、どちらの映画を見たいですか？	心温まるコメディ映画 ($N = 136$) / ハラハラするコメディ映画 ($N = 55$) Comedy movie with a touching story ($N = 84$) / Comedy movie with a thrilling story ($N = 108$)	< .001 1.000

表 3 音声再生回数と、音声提示された選択肢内容の理解度。理解度は、7段階リッカート尺度（1：まったく理解できなかった，7：すべて理解できた）を用いた実験参加者の主観評価により取得した。

	音声再生回数の平均（標準偏差）	理解度の平均（標準偏差）
母語条件（日本語）	1.07 ($SD = 0.92$)	6.81 ($SD = 0.92$)
非母語条件（英語）	1.27 ($SD = 0.51$)	4.33 ($SD = 1.44$)

の質問を発話速度操作による選択誘導可能性を検証するための質問として使用し、他の 10 問は実験意図を気付かれないようにするためのダミーの質問として実験に使用する。

4. 発話速度の操作による選択誘導実験

本章では、発話速度の操作による選択誘導可能性を、母語の場合と非母語の場合でそれぞれ検証し、結果を比較する。3 章の実験と同様、英語を非母語とし、日本語を母語とする成人を実験参加者の対象とし、母語条件を日本語、

非母語条件を英語とした。

4.1 実験設計

実験用のシステム（図 1）を用いて、質問に対して 2 択の選択肢から 1 つを選び回答してもらう実験を行った。3 章と同様に、実験参加者には海外旅行時における選択というシナリオで実験に取り組んでもらった。質問は合計 15 問あり、3 章の実験に使用した質問と同じ質問を用いた。15 問のうち、3 章の実験で母語条件、非母語条件の両方において選択に偏りがみられなかった 5 つの質問を検証用の質問とし、他の 10 問を実験の意図に気付かれないようにするためのダミー質問とした。

検証用の 5 問では、選択肢を音声提示する際に、一方を 3 章の実験時と同じ発話速度（以降、標準発話速度と呼ぶ）で、もう一方をそれよりも遅い発話速度で提示した。ダ

ミー質問では、2 択の選択肢のどちらも標準発話速度で提示した。発話速度の操作によるプライバシー同意選択への影響を検証した先行研究 [24] では、発話速度の差を 20%程度としていたため、本研究でも母語条件、非母語条件ともに、標準発話速度の約 0.8 倍を遅い発話速度として設定した (表 4)。実験に使用する音声は 3 章の実験と同じ音声を使用し、検証用の質問に使用する音声のみ、速度を操作した。実験では、全ての質問においてどちらの選択肢を先に音声提示するか試行ごとにランダムに決定し、検証用の質問では、どちらの選択肢を遅い速度で提示するかも試行ごとにランダムに決定した。

4.2 実験手順

3 章と同様に、質問は 1 問ずつ日本語のテキストで表示し、実験参加者がシステムの音声再生ボタンをクリックしたときに選択肢を音声提示した。母語条件では日本語音声、非母語条件では英語音声再生され、実験参加者は音声を聞いた後に「選択肢 1/Option 1」「選択肢 2/Option 2」と表記された選択肢のいずれかをクリックした。

15 問の質問に対する選択が終了した後、実験参加者はアンケートに回答した。アンケートでは、音声提示された選択肢の内容を参加者がどの程度理解できたかを尋ねた。また、ユーザは VUI に潜む DP によって選択が誘導されても、懐疑心を抱かない可能性がある [20] ため、実験参加者がどの程度自身の選択をコントロールできたと感じているかを尋ねた。アンケートの途中では、実験参加者が注意深く質問を読んでいることを確かめるために注意力確認の質問を行い、アンケートの最後に実験参加者の性別と年齢を回答してもらった。

4.3 実験参加者の募集

実験参加者の募集には YCS を利用し、母語条件、非母語条件の各実験において男性限定と女性限定の 2 回に分けてそれぞれ参加者の募集を行うことで、参加者の男女比が同程度となるようにした。本章の実験に要する時間は 5 分程度と予想されたため、地域の賃金基準に基づき、実験参加に対する報酬として約 100 円分の PayPay ポイントを参加者のアカウントに付与した。実験参加者の対象は、英語を非母語とし、日本語を母語とする成人であり、3 章の実験に参加した者は本章の実験に参加できないようにした。また、先に実施する母語条件の実験に参加した者は、後に実施する非母語条件の実験に参加できないようにした。

4.4 結果

母語条件の実験は 2025 年 11 月 21 日に実施し、非母語条件の実験は 2025 年 11 月 25 日に実施した。母語条件、非母語条件の各実験でそれぞれ 200 名の参加者を募集し、

表 4 検証用の質問における選択肢提示時の発話速度の違い。3 章の実験時の発話速度を標準発話速度とし、その約 0.8 倍の速度を遅い発話速度とした。

	標準発話速度	遅い発話速度
母語条件 (日本語)	8.1 モーラ/秒	6.6 モーラ/秒
非母語条件 (英語)	139WPM	113WPM

注意力確認の質問に対する回答が適切でなかった参加者は分析対象から除外した。本章では、母語条件 175 名 (男性 93 名, 女性 81 名, 未回答 1 名)、非母語条件 177 名 (男性 91 名, 女性 85 名, 未回答 1 名) によるデータを分析対象とした。

各条件において、検証用の 5 つの質問でそれぞれ 2 択の各選択肢が選ばれた合計回数を求め、二項検定を行ったところ、母語条件ではどの質問においても選択の偏りがみられなかった ($p > .05$)。一方、非母語条件では 2 つの質問 (Q5, Q10) で選択に偏りがみられた ($p < .05$)。なお、多重比較のため、Bonferroni 法を用いて p 値を補正した。非母語条件で選択に偏りが生じた 2 つの質問は、発話速度の違いによる選択誘導可能性を検証するうえで適切ではないと考え除外し、選択に偏りがなかった 3 問のデータについて分析する。

条件ごとに、標準発話速度で音声提示された選択肢が選ばれた合計回数と、遅い発話速度で音声提示された選択肢が選ばれた合計回数を表 5 に示す。選択に偏りがあるか二項検定を行った結果、どちらの条件においても有意差は認められず ($p > .05$)、発話速度の操作はユーザの選択を誘導しない可能性が示唆された。

母語条件で発話速度を操作した検証用の 5 問について、実験参加者ごとに標準発話速度で提示された選択肢を選んだ質問数と、遅い発話速度で提示された選択肢を選んだ質問数を求めた結果を図 2 に示す。また、非母語条件についても、分析対象とした 3 問の検証用質問において標準発話速度で提示された選択肢が選ばれた質問数と、遅い発話速度で提示された選択肢が選ばれた質問数を参加者ごとに求めた結果を図 3 に示す。母語条件は、標準発話速度の場合と遅い発話速度の場合の選択質問数の分布が類似していた。一方、非母語条件では、標準発話速度の場合と遅い発話速度の場合で最頻値が異なっており、分布の違いがみられた。

音声の再生回数と、アンケートに対する回答結果を表 6 に示す。音声の再生回数については、全ての質問におけるデータを分析に使用した。音声の再生回数と選択肢内容の理解度については、3 章の実験と同様の結果が得られた。自分の選択をコントロールできていると感じたというアンケート項目に対する 7 段階リッカート尺度 (1: 強く反対する, 7: 強く同意する) を用いた評価値について、母語条件と非母語条件の間に差があるか Brunner-Munzel 検定を

表 5 標準発話速度で音声提示された選択肢が選ばれた回数と、遅い発話速度で音声提示された選択肢が選ばれた回数。選択回数に偏りがあるか二項検定を実施した。

	標準発話速度で提示	遅い発話速度で提示	p 値
母語条件 (日本語)	434 回	441 回	.839
非母語条件 (英語)	279 回	252 回	.259

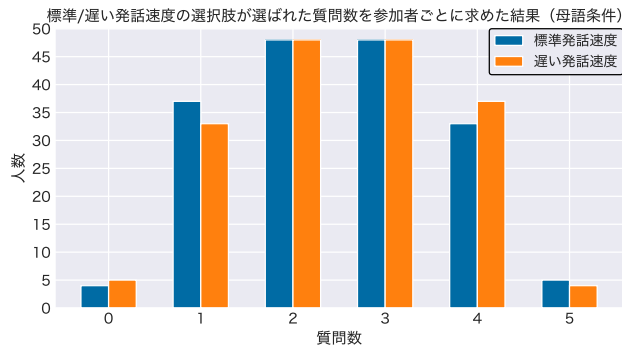


図 2 母語条件において、発話速度の操作による選択誘導可能性を検証した 5 問のうち、標準発話速度で提示した選択肢が選ばれた質問数と、遅い発話速度で提示した選択肢が選ばれた質問数を実験参加者ごとに求めた結果。

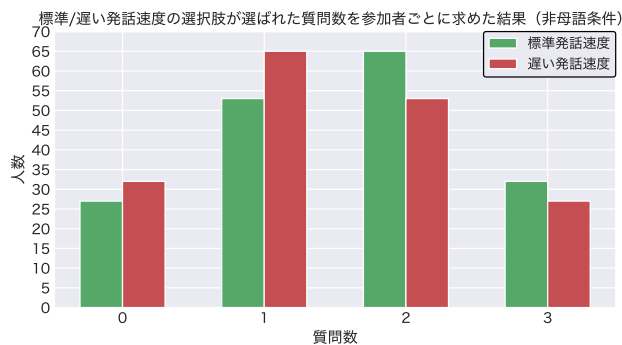


図 3 非母語条件において、発話速度の操作による選択誘導可能性を検証した 5 問のうち、標準発話速度で提示した選択肢が選ばれた質問数と、遅い発話速度で提示した選択肢が選ばれた質問数を実験参加者ごとに求めた結果。

行った結果、有意差が認められ ($W = -12.272, p < .001$), 非母語で音声提示された場合は、母語の場合よりも自分の選択をコントロールすることが難しい可能性が示唆された。

5. 考察

5.1 発話速度の操作による誘導 (RQ1)

母語条件、非母語条件ともに、発話速度の操作によって選択に偏りが生じなかったため、本研究の結果からは発話速度の操作が DP として機能しない可能性が示唆された。Dubiel [19] らの研究では、発話速度が速くピッチの変化が大きい音声で選択肢を提示した場合に、発話速度が遅くピッチの変化が小さい音声で提示する場合よりも選ばれやすい可能性が示されている。本研究では音声のピッチを操作せず、発話速度のみを操作した場合に誘導効果がない可

能性が示されたため、VUI により選択肢が音声提示される状況下では、音声のピッチがユーザの選択により強く影響する可能性が考えられる。また、Dubiel [19] らは、一方の音声の発話速度を 180WPM、もう一方の音声を 100WPM に設定しており、発話速度の差が大きかったが、本研究では非母語条件における標準発話速度を 139WPM、遅い発話速度を 113WPM と設定し、Leschanowsky ら [24] と同様に速度の違いは 20% 程度であった。この発話速度の差が異なると、選択誘導効果は変化する可能性がある。

全体として発話速度の操作による誘導効果はみられなかった一方で、母語条件、非母語条件ともに分析対象とした誘導検証用質問のうち、多くの質問で標準あるいは遅い発話速度によって提示された選択肢を選んだ参加者が一定数いることが確認された。GUI における DP によって行動を誘導されやすいユーザ層が存在する [35] ように、発話速度の操作によって選択を誘導されやすいユーザが存在する可能性も考えられる。今後は、誘導されやすいユーザとそうでないユーザの特徴について明らかにする予定である。

5.2 言語の違いによる選択への影響 (RQ2)

実験参加者ごとに標準/遅い発話速度で提示された選択肢を選んだ質問数を求めた結果 (図 2, 3), 母語条件では分布形状に違いがみられなかったが、非母語条件では最頻値が異なり分布に違いがあった。また、3 章の実験と 4 章の実験の両方で、意味は同じであるにも関わらず、母語条件で選択に偏りが生じず、非母語条件では選択が偏った質問があり、その逆も存在した。これらのことから、言語が異なることによって、ユーザの選択傾向が変化する可能性が考えられる。本研究では発話速度に着目したが、選択肢自体の聞き取りやすさがユーザの選択に影響する可能性がある。例えば、聞き取りやすい選択肢を遅い発話速度で提示し、聞き取りづらい選択肢を標準あるいは速い発話速度で音声提示した場合に、非母語話者は聞き取りやすく遅い発話速度で提示された選択肢を選びやすくなる可能性が考えられる。

4 章の実験において、非母語条件の参加者の選択肢内容の理解度は 7 段階中で平均 4.63 ($SD = 1.34$) と中程度であったため、正確にすべての選択肢内容を把握できていた参加者は少なかったと考えられる。そのため、選択肢内容を理解せずにランダムに選択を行った参加者が存在した可能性も考えられる。また、非母語条件における自身の選択のコントロールについての評価値は、母語条件よりも有意に低かった ($p < .001$)。これらのことから、言語の違いは、ユーザが情報を正しく理解した上で意思決定を行うことを阻害する可能性が高いと考えられる。言語障壁の悪用とユーザへの影響に関する知見は不足している [4, 14, 36] ため、今後さらなる研究が必要である。

表 6 音声再生回数と、音声提示された選択肢内容の理解度および自身の選択のコントロールに関する参加者の主観評価。理解度と選択コントロールは、7段階リッカート尺度による回答結果となっており、値が大きいほど理解度が高いことや自身の選択をコントロールできたことを示している。

	音声再生回数の平均 (標準偏差)	理解度の平均 (標準偏差)	選択コントロールの平均 (標準偏差)
母語条件 (日本語)	1.04 ($SD = 0.22$)	6.69 ($SD = 0.54$)	6.34 ($SD = 0.97$)
非母語条件 (英語)	1.20 ($SD = 0.46$)	4.63 ($SD = 1.34$)	5.12 ($SD = 1.31$)

5.3 制約と展望

本研究にはいくつかの制約がある。まず、英語を非母語とし、日本語を母語とする成人を実験参加者の対象とし、母語条件を日本語、非母語条件を英語と設定したため、本研究で得られた結果は、異なる文化背景のユーザや異なる言語を対象とした場合に変わる可能性がある。また、実験では海外旅行時の選択というシナリオを参加者に指示したため、異なる状況下での選択に対して、本研究の結果が適用できるとは限らない。さらに、発話速度の操作による選択誘導可能性を検証した実験では、先行研究 [24, 37] と同様に発話速度の差を 20% 程度にしたが、差の大きさが異なる場合は、誘導効果が変化すると考えられる。

今後は、発話速度の差を大きくした場合の選択誘導効果を検証するとともに、速度差がどの程度大きくなった場合にユーザが不信任を抱くのかも明らかにしたいと考えている。また、言語の違いによって選択傾向が変化する可能性が示唆されたため、同一の参加者を対象に、非母語で選択肢を音声提示した場合と、母語で選択肢を音声提示した場合で、選ぶ選択肢が一致するかを検証したいと考えている。さらに、本研究では選択肢を音声により提示し、参加者はクリックにより選択を行ったが、音声のみで操作を行う VUI も存在するため、VUI による音声提示に対して、ユーザが発話により操作を行う場合に潜む DP についても検討する予定である。

6. おわりに

本研究は、選択肢を音声提示する際に、発話速度を操作することによってユーザの選択が誘導されるか検証し、その誘導効果がユーザの母語で音声提示する場合と非母語で音声提示する場合で異なるか比較を行った。実験の結果、母語、非母語のどちらの場合においても発話速度の操作による選択の偏りはみられず、誘導効果は確認されなかった。一方、言語の違いは、ユーザの選択傾向を変化させる可能性があり、非母語によって音声提示された場合は、母語の場合よりも自身の選択のコントロールが困難になる可能性が示唆された。

本研究は音声の発話速度に着目したが、音声には他にも様々な要素があり、それらがユーザの行動に影響を及ぼす可能性がある。音声のどのような操作がユーザの行動を誘導するか、また、言語の違いによってその誘導効果は増幅する可能性があるか、今後さらに研究を行う必要がある。

参考文献

- [1] Brignull, H.: Deceptive Patterns. <https://www.deceptive.design/> (参照: 2025/07/04) .
- [2] Mathur, A., Acar, G., Friedman, M. J., Lucherini, E., Mayer, J., Chetty, M. and Narayanan, A.: Dark Patterns at Scale: Findings from a Crawl of 11K Shopping Websites, *Proc. ACM Hum.-Comput. Interact.*, Vol. 3, No. CSCW, pp. 1–32 (2019).
- [3] Di Geronimo, L., Braz, L., Fregnan, E., Palomba, F. and Bacchelli, A.: UI Dark Patterns and Where to Find Them: A Study on Mobile Applications and User Perception, *Proceedings of the 2020 CHI Conference on Human Factors in Computing Systems*, CHI '20, pp. 1–14 (2020).
- [4] Hidaka, S., Kobuki, S., Watanabe, M. and Seaborn, K.: Linguistic Dead-Ends and Alphabet Soup: Finding Dark Patterns in Japanese Apps, *Proceedings of the 2023 CHI Conference on Human Factors in Computing Systems*, CHI '23, pp. 1–13 (2023).
- [5] Soe, T. H., Nordberg, O. E., Guribye, F. and Slavkovik, M.: Circumvention by Design - Dark Patterns in Cookie Consent for Online News Outlets, *Proceedings of the 11th Nordic Conference on Human-Computer Interaction: Shaping Experiences, Shaping Society*, NordiCHI '20, pp. 1–12 (2020).
- [6] Chaudhary, A., Saroha, J., Monteiro, K., Forbes, A. G. and Parnami, A.: “Are You Still Watching?”: Exploring Unintended User Behaviors and Dark Patterns on Video Streaming Platforms, *Proceedings of the 2022 ACM Designing Interactive Systems Conference*, DIS '22, p. 776–791 (2022).
- [7] Mildner, T., Savino, G.-L., Doyle, P. R., Cowan, B. R. and Malaka, R.: About Engaging and Governing Strategies: A Thematic Analysis of Dark Patterns in Social Networking Services, *Proceedings of the 2023 CHI Conference on Human Factors in Computing Systems*, CHI '23, pp. 1–15 (2023).
- [8] Gray, C. M., Kou, Y., Battles, B., Hoggatt, J. and Toombs, A. L.: The Dark (Patterns) Side of UX Design, *Proceedings of the 2018 CHI Conference on Human Factors in Computing Systems*, CHI '18, pp. 1–14 (2018).
- [9] Gray, C. M., Santos, C. T., Bielova, N. and Mildner, T.: An Ontology of Dark Patterns Knowledge: Foundations, Definitions, and a Pathway for Shared Knowledge-Building, *Proceedings of the 2024 CHI Conference on Human Factors in Computing Systems*, CHI '24, pp. 1–22 (2024).
- [10] Mansur, S. M. H., Salma, S., Awofisayo, D. and Moran, K.: AidUI: Toward Automated Recognition of Dark Patterns in User Interfaces, *Proceedings of the 45th International Conference on Software Engineering*, ICSE '23, pp. 1958–1970 (2023).
- [11] Chen, J., Sun, J., Feng, S., Xing, Z., Lu, Q., Xu, X. and Chen, C.: Unveiling the Tricks: Automated Detection of Dark Patterns in Mobile Applications, *Proceedings of*

- the 36th Annual ACM Symposium on User Interface Software and Technology*, UIST '23, pp. 1–20 (2023).
- [12] Owens, K., Gunawan, J., Choffnes, D., Emami-Naeini, P., Kohno, T. and Roesner, F.: Exploring Deceptive Design Patterns in Voice Interfaces, *Proceedings of the 2022 European Symposium on Usable Security*, EuroUSEC '22, pp. 64–78 (2022).
- [13] Shigematsu, R., Oishi, R., Nakagawa, Y., Nakamura, S., Torii, T. and Takao, H.: Guiding Task Choice in Japanese Voice Interfaces through Vocalization Cost: Click-based vs. Voice-based Selection, *Proceedings of the 7th ACM International Conference on Multimedia in Asia*, MMAsia '25, pp. 1–7 (2025).
- [14] 木下裕一朗, 中村聡史: 非母語話者の発話選択における Deceptive Pattern: 英語選択肢の発音容易性による選択誘導可能性の検証, 情報処理学会 研究報告ヒューマンコンピュータインタラクション (HCI), Vol. 2025-HCI-215, No. 12, pp. 1–8 (2025).
- [15] M. Bhoot, A., A. Shinde, M. and P. Mishra, W.: Towards the Identification of Dark Patterns: An Analysis Based on End-User Reactions, *Proceedings of the 11th Indian Conference on Human-Computer Interaction*, IndiaHCI '20, pp. 24–33 (2021).
- [16] Kojima, M., Katie, S., Kinoshita, Y. and Nakamura, S.: Reactions of Japanese Consumers on Social Media to Dark Patterns and Deceptive Designs, 情報処理学会 研究報告ヒューマンコンピュータインタラクション (HCI), Vol. 2025-HCI-213, No. 4, pp. 1–8 (2025).
- [17] Bongard-Blanchy, K., Rossi, A., Rivas, S., Doublet, S., Koenig, V. and Lenzini, G.: “I am Definitely Manipulated, Even When I am Aware of it. It’s Ridiculous!” - Dark Patterns from the End-User Perspective, *Proceedings of the 2021 ACM Designing Interactive Systems Conference*, DIS '21, pp. 763–776 (2021).
- [18] Mildner, T., Freye, M., Savino, G.-L., Doyle, P. R., Cowan, B. R. and Malaka, R.: Defending Against the Dark Arts: Recognising Dark Patterns in Social Media, *Proceedings of the 2023 ACM Designing Interactive Systems Conference*, DIS '23, pp. 2362–2374 (2023).
- [19] Dubiel, M., Sergeeva, A. and Leiva, L. A.: Impact of Voice Fidelity on Decision Making: A Potential Dark Pattern?, *Proceedings of the 29th International Conference on Intelligent User Interfaces*, IUI '24, pp. 181–194 (2024).
- [20] Zhang, S., Chen, H., Wang, Y., Xu, Y., Bai, J., Wu, Y., Li, S., Yi, X., Wang, C. and Li, H.: The Manipulative Power of Voice Characteristics: Investigating Deceptive Patterns in Mandarin Chinese Female Synthetic Speech, *Proc. ACM Interact. Mob. Wearable Ubiquitous Technol.*, Vol. 9, No. 3, pp. 1–32 (2025).
- [21] Lacey, C. and Caudwell, C.: Cuteness as a ‘Dark Pattern’ in Home Robots, *Proceedings of the 14th ACM/IEEE International Conference on Human-Robot Interaction*, HRI '19, pp. 374–381 (2020).
- [22] Kowalczyk, M., Gunawan, J. T., Choffnes, D., Dubois, D. J., Hartzog, W. and Wilson, C.: Understanding Dark Patterns in Home IoT Devices, *Proceedings of the 2023 CHI Conference on Human Factors in Computing Systems*, CHI '23, pp. 1–27 (2023).
- [23] Greenberg, S., Boring, S., Vermeulen, J. and Dostal, J.: Dark Patterns in Proxemic Interactions: A Critical Perspective, *Proceedings of the 2014 Conference on Designing Interactive Systems*, DIS '14, pp. 523–532 (2014).
- [24] Leschanowsky, A., Sergeeva, A., Bauer, J., Vijapurapu, S. and Dubiel, M.: Exploring the Impact of Modality and Speech Rate Manipulation in Voice Permission Requests—Limits of Applicability and Potential for Influencing Decision-Making, *International Journal of Human-Computer Studies*, Vol. 205, p. 103590 (2025).
- [25] Wenzel, K. and Kaufman, G.: Designing for Harm Reduction: Communication Repair for Multicultural Users’ Voice Interactions, *Proceedings of the 2024 CHI Conference on Human Factors in Computing Systems*, CHI '24, pp. 1–17 (2024).
- [26] Pyae, A. and Scifleet, P.: Investigating Differences Between Native English and Non-native English Speakers in Interacting with a Voice User Interface: A Case of Google Home, *Proceedings of the 30th Australian Conference on Computer-Human Interaction*, OzCHI '18, pp. 548–553 (2018).
- [27] Wu, Y., Edwards, J., Cooney, O., Bleakley, A., Doyle, P. R., Clark, L., Rough, D. and Cowan, B. R.: Mental Workload and Language Production in Non-Native Speaker IPA Interaction, *Proceedings of the 2nd Conference on Conversational User Interfaces*, CUI '20 (2020).
- [28] Song, J. Y., Pycha, A. and Culleton, T.: Interactions Between Voice-Activated AI Assistants and Human Speakers and Their Implications for Second-Language Acquisition, *Frontiers in Communication*, Vol. 7 (2022).
- [29] Goh, C. C.: A Cognitive Perspective on Language Learners’ Listening Comprehension Problems, *System*, Vol. 28, No. 1, pp. 55–75 (2000).
- [30] Cao, X., Yamashita, N. and Ishida, T.: How Non-native Speakers Perceive Listening Comprehension Problems: Implications for Adaptive Support Technologies, *Collaboration Technologies and Social Computing*, CollabTech '16, pp. 89–104 (2016).
- [31] Cambre, J. and Kulkarni, C.: One Voice Fits All? Social Implications and Research Challenges of Designing Voices for Smart Devices, *Proc. ACM Hum.-Comput. Interact.*, Vol. 3, No. CSCW, pp. 1–19 (2019).
- [32] Maekawa, K.: Corpus of Spontaneous Japanese: Its Design and Evaluation, *Proceedings of the ISCA & IEEE Workshop on Spontaneous Speech Processing and Recognition*, SSPR '03, pp. 7–12 (2003).
- [33] Yuan, J., Liberman, M. Y. and Cieri, C.: Towards an Integrated Understanding of Speaking Rate in Conversation, *Proceedings of the Interspeech 2006* (2006).
- [34] Seaborn, K. and Nakamura, S.: Quality and Representativeness of Research Online with Yahoo! Crowdsourcing, *Frontiers in Psychology*, Vol. 16 (2025).
- [35] Rossi, A., Carli, R., Botes, M. W., Fernandez, A., Sergeeva, A. and Sánchez Chamorro, L.: Who is Vulnerable to Deceptive Design Patterns? A Transdisciplinary Perspective on the Multi-Dimensional Nature of Digital Vulnerability, *Computer Law & Security Review*, Vol. 55, p. 106031 (2024).
- [36] Seaborn, K., Itagaki, T., Watanabe, M., Wang, Y., Geng, P., Fujii, T., Mandai, Y., Kojima, M. and Yoshida, S.: Deceptive, Disruptive, No Big Deal: Japanese People React to Simulated Dark Commercial Patterns, *Extended Abstracts of the CHI Conference on Human Factors in Computing Systems*, CHI EA '24, pp. 1–8 (2024).
- [37] Sánchez Chamorro, L., Toebosch, R. and Lallemand, C.: Manipulative Design and Older Adults: Co-Creating Magic Machines to Understand Experiences of Online Manipulation, *Proceedings of the 2024 ACM Designing Interactive Systems Conference*, DIS '24, pp. 668–684 (2024).