

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

木下 裕一朗^{1,a)} 中村 聡史¹

概要: 企業の利益となるようにユーザの行動を妨害したり誘導したりする Deceptive Pattern (Dark Pattern) は、グラフィカルユーザインタフェースだけでなく、音声ユーザインタフェースにも潜んでおり、今後音声ユーザインタフェースがさらに普及するにつれて、大きな問題となる可能性がある。本稿は、新たな Deceptive Pattern の一つとして言語の壁を悪用したデザインに着目し、特に、非母語話者にとっての発音容易性による選択行動の誘導に焦点を当てた。本稿では、選択肢の発話選択時に、ユーザの母語とは異なる言語の使用を強制したとき、ユーザの発話選択行動は選択肢の発音容易性によって誘導されるか実験により検証した。実験の結果、ユーザが非母語で発話選択するとき、ユーザは発音しやすい選択肢を選びやすく、誘導できることが明らかになった。

1. はじめに

ウェブサイトやモバイルアプリケーションなどで、企業の利益となるようにユーザの行動を妨害したり誘導したりする Deceptive Pattern^{*1} (DP) が多く使用されている。例えば、ユーザに選ばせたい選択肢を目立つように表示してクリックするよう誘導したり、サブスクリプションの解約手続きを煩雑にしてユーザに解約を諦めさせようとしたりするものは DP に該当する。DP は、ユーザに経済的な損失や心理的なストレスを与えるため問題視されている。

これまでの DP に関する研究の多くはグラフィカルユーザインタフェース (GUI) を対象としていたが、DP は GUI だけでなく、音声ユーザインタフェース (VUI) にも潜んでいる可能性が指摘されている [2-4]。GUI で DP が次々と出現したように、VUI の普及が進み様々な場面で利用可能になると、VUI にも DP が蔓延する可能性がある。そのため、VUI における潜在的な DP について早期に検討することが、ユーザを保護するうえで重要となる。

DP の文脈で VUI に着目した研究として、VUI における現在および将来の潜在的な DP について検討した研究 [2] や、VUI で用いる合成音声の特徴によってユーザの意思決定が操作される [5,6] ことを明らかにした研究などがある。

VUI は、音声入力と音声出力を用いるユーザインタフェース [7] であるが、先行研究 [5,6] では音声出力における DP のみが対象となっている。一方で、我々の知る限り、音声入力に関する DP については、これまでほとんど検討されていない。そこで本稿では、音声入力における潜在的な DP に着目し、その誘導効果を明らかにする。

ここで、音声入力では、選択肢の発話/発音容易性の違いを利用してユーザの選択を誘導できてしまう可能性がある。実際に、日本語話者を対象とした日本語選択肢の音声入力とクリック入力の選択傾向を比較した研究 [8] で、選択肢の長さや表現の違いによって、ユーザの音声選択が誘導されることが明らかになっている。[8] は母語を用いた音声入力という状況であるが、ユーザが非母語を用いて VUI を使用するとき、操作に困難が生じる [9] ことや、精神的作業負荷が大きい [10] ことが知られているため、発話/発音容易性の違いを利用した誘導は、ユーザが非母語で発話選択する場合に、より効果を発揮する可能性がある。

自動翻訳技術の進歩により言語の壁はなくなりつつあるものの、GUI ではコンテンツの一部をユーザの使用言語に翻訳せず、ユーザの理解を妨げたり誤認させたりする DP (Untranslation [11]) が存在する。ユーザが非母語を用いて VUI を使用する状況は、VUI が多言語対応していない場合だけでなく、このような DP によって意図的に生み出される可能性もあるため、非母語を用いた発話選択において、発話/発音容易性による選択誘導可能性を明らかにする必要がある。

本稿は、これまで十分に検討されていない音声入力にお

¹ 明治大学
Meiji University

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

^{*1} Harry Brignull [1] は、2010 年に Dark Pattern と名付けたが、人種の問題を考慮して呼称を Deceptive Pattern (または Deceptive Design) に変えている。そのため本稿では、Deceptive Pattern という用語を使用する。

ける潜在的な DP の一つとして、選択肢の発音容易性の違いによる誘導可能性に着目し、ユーザが非母語で発話選択を行う場合に、選択肢の発音容易性の違いによってユーザの選択が誘導されるか検証する。

2. 関連研究

2.1 Deceptive Pattern

DP は、E コマースサイト [12, 13] や動画配信プラットフォーム [14] などに存在し、SNS [15, 16] やゲーム [17, 18] には特有の DP が存在することが明らかになっている。また、プライバシーに関する DP [19] も存在し、同意インタフェースにおける DP [20–22] についても盛んに研究が行われている。オンラインサービスの人気度と DP の使用には弱い相関があり [23, 24]、240 の人気モバイルアプリのうち約 95% に DP が含まれていたことが明らかになっている [25]。

従来の DP 研究は主に GUI を対象として行われてきたが、DP は GUI だけでなく、物理空間のデザイン [26] や IoT デバイス [27]、ソーシャルロボット [28, 29]、VUI [2, 4, 5] にも存在し、これから新たな DP が出現する可能性が指摘されている。しかし、Digital Services Act [30]（デジタルサービス法）をはじめとした DP に関する現在の規制の枠組みでは、GUI 以外における DP が十分に考慮されていない。そのため、GUI 以外における DP についても焦点を当て、ユーザの行動への影響を明らかにする必要がある。

本稿では VUI に焦点を当て、音声入力における潜在的な DP の一つとして発音容易性の違いに着目し、その誘導効果を実験により明らかにする。

2.2 音声ユーザインタフェースにおける選択の誘導

これまで、VUI に用いる合成音声の発話速度や音の高さなどといった音声特徴を操作することで、ユーザの選択が誘導されることが明らかになっている [5, 6]。また、ユーザは GUI における DP に遭遇したときに苛立ちを感じる [31] ことが示されている一方で、ユーザは音声特徴の意図的な操作によって誘導されても、懐疑心を抱かない可能性が示されている [6]。

ユーザが発話して選択を行う場合、選択肢の表現 [8] や長さ [32] の違いによって選択が誘導されることが明らかになっている。また、GUI と VUI における有料アイテムの選択率を実験により比較した研究 [33] では、GUI に比べ VUI を用いた方が有料アイテムがより多く選ばれる傾向があったことが報告されている。これらの研究から、GUI においては選択の誘導に寄与しない要因であっても、VUI では選択を誘導する可能性がある。

音声入力における DP については、これまで十分に検討が行われていない。本稿では、ユーザが非母語を用いて発話選択を行う場合に、選択肢の発音容易性の違いによって

選択が誘導されるか検証し、DP として機能する可能性があるか明らかにする。

3. 発話選択実験に用いる質問と選択肢の選定

非母語を用いた発話選択において、選択肢の発音容易性の違いによる選択誘導可能性を検証するために、発音容易性以外の要因である選択肢の人気度や知名度の差などによって選択の偏りが生じない選択肢を選定する必要がある。そこで本章では、選択肢をクリックで選ぶ場合に選択の偏りが生じない選択肢の抽出を目的とし、質問に対して 2 択の選択肢から 1 つをクリックで選び回答してもらう実験を行う。発音容易性による発話選択の誘導可能性を検証する実験（4 章）では、本章のクリック選択実験において選択の偏りが無い選択肢を使用する。

3.1 実験設計

実験は、実験用に実装したシステム（図 1）を使用して行う。実験では、参加者に対して「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 17 個の出来事が起こり、それぞれの出来事で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。質問は日本語で、選択肢は英語で表示されます。意味がわからない英語選択肢が出てきた場合は、調べていただいて構いません。」と指示し、用意した 17 の各質問^{*2}に対して、2 択の選択肢から 1 つをクリック選択により回答してもらう。

17 問の質問のうち、発音容易性による選択誘導可能性を検証する実験（4 章）で検証に使用するための質問は 10 問であり、残りの 7 問のうち 5 問は、実験の意図を気づかれないようにするためのダミー質問で、他の 2 問は参加者が注意して実験を行っていることを確かめるための質問である。質問は 1 問ずつ表示し、質問文は日本語で、選択肢は英語で提示する。

発音容易性による誘導可能性を検証するための 10 問における 2 択の選択肢は、それぞれ発音が容易な選択肢と難しい選択肢で構成する。発音が難しい選択肢は、先行研究 [34] により明らかになっている、日本人にとって発音するのが難しい音（/r/や/th/など）を含む単語とし、反対に発音が容易な選択肢は、そのような音を含まない単語とした。また、選択肢の長さによって発話選択が誘導される可能性がある [32] ため、発音容易性による選択誘導可能性を検証する 10 問については、同一質問の選択肢間で出来るだけ文字数の差が生じないようにした。2 択の選択肢は横並びに表示し、各選択肢を左右のどちらに表示するかは質問ごとにランダムに決定した。なお、実験に用いた 17 の質問は時系列で起こる出来事に対する質問としたため、質問の提示順序は参加者間で同一とした。

^{*2} クリック選択実験に使用した質問と選択肢は、<https://github.com/yuichiroooo/exp202509> で公開している。



図 1 クリック選択実験 (3 章) に用いた実験システム

参加者が 17 の質問すべてに対する選択が終わった後、実験に使用した英語選択肢の単語難易度が適切であるか、また 2 択の選択肢間で発音容易性が異なっているかを確かめるためにアンケートを実施する。アンケートでは最初に、実験で提示された英語選択肢のうち知っている単語がどの程度あったかを 7 段階リッカート尺度を用いて回答してもらう。次に、発音容易性による選択誘導可能性を検証するための 10 問で提示された 2 つの英語選択肢について、どちらの選択肢の発音がより難しいと感じるか回答してもらう。アンケートの途中では、実験参加者がアンケート項目をよく読んで回答していることを確かめるために、注意力確認の質問を実施し、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う。

実験参加者の募集には、Yahoo!クラウドソーシング^{*3} (YCS) を利用した。YCS は、国内で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 [35]。YCS では、対象を男性もしくは女性に限定して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。パイロットテストにより実験時間は 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。本稿では、日本人にとって発音が苦手な音を含む単語の選択肢を発音が難しい選択肢としているため、実験参加者の対象は、英語の非母語話者である日本人の成人とした。

3.2 実験手順

実験参加者は、PC を用いて Google Chrome を利用し、YCS のページから実験用のシステムにアクセスする。実験用システムは最初に実験の概要や、実験参加の条件、実験を行う際の注意点についての説明が記載されたページを表示する。実験参加者が説明を読み、参加に同意して設置されたボタンをクリックすると、実験のシナリオに関する指示 (3.1 節) が記載されたページに遷移する。参加者がシナリオについての説明を読み、ボタンをクリックすると、

実験が開始となり 1 つ目の質問と選択肢が表示される。参加者が質問への回答をクリック選択で行い、「次へ」のボタンを押すと 2 つ目の質問と選択肢が表示される。このようにして、参加者が計 17 の質問に対して選択を終えると、システムはアンケートのページへと遷移する。そして、参加者がすべてのアンケート項目に対して回答を終えて設置されたボタンをクリックすると、システムは実験の終了を知らせるページに遷移し、ユニークな実験参加者 ID と実験コードを表示する。参加者がその ID とコードを YCS のページに戻って入力すると、参加者は報酬を受け取ることができる。

3.3 結果

実験参加者は 200 名であり、実験時間の平均は約 3 分 57 秒であった。注意力確認のタスクをクリアしていない参加者や、実験時間が他の参加者と比べて極端に短い参加者は分析対象として適切でないと考えて除外した。本稿では、Rousseuw ら [36] の基準に従い、全実験参加者の実験時間の第一四分位数から四分位範囲の 1.5 倍の値を引いた時間よりも、実験時間が短かった参加者を分析対象から除外した。除外の結果、本稿では 181 名の参加者によるデータを分析対象とした。181 名のうち、男性は 89 名、女性は 88 名、トランスジェンダーは 1 名、ノンバイナリーや他のジェンダーはともに 0 名、未回答は 3 名であった。また、181 名の参加者のうち、20 代が 4 名、30 代が 26 名、40 代が 54 名、50 代が 64 名、60 代が 21 名、70 代以上が 9 名、未回答が 3 名であった。

発音容易性による選択誘導可能性を検証する実験 (4 章) で使用するための 10 の質問において、それぞれ 2 択の各選択肢の選ばれた回数を求め、選択に偏りがあるかカイ二乗検定を行った (表 1)。多重比較を行ったため、 p 値は Bonferroni 法により補正した。また、有意水準は 0.05 とした。カイ二乗検定の結果、10 問のうち 5 問 (Q1, Q4, Q5, Q10, Q11) は選択が有意に偏っており ($p < .05$)、残りの 5 問 (Q2, Q7, Q8, Q14, Q15) については選択の偏りに有意差は認められなかった ($p > .05$)。実験後に実施したアンケートで、実験に使用された英語選択肢のうち、知っている単語がどの程度あったかを参加者に 7 段階 (1: すべて知らなかった, 7: すべて知っていた) で評価してもらった結果、平均値は 5.72 ($SD = 0.96$) であった。このことから、参加者は多くの英語選択肢の意味を理解していたと考えられるため、英単語の知識による選択への影響は小さいと言える。

上記と同様の 10 問における 2 択の英語選択肢に対して、実験後のアンケートで「2 つのうち、どちらの発音がより難しいと感じますか?」と尋ねた。参加者が回答した結果を表 2 に示す。発音がより難しい選択肢として選ばれた

^{*3} <https://crowdsourcing.yahoo.co.jp/>

表 1 発音容易性による選択誘導可能性を検証する 10 問において、クリック選択で発音の容易な選択肢と発音の難しい選択肢が選ばれた回数。選択回数に偏りがあるかはカイ二乗検定を行い、多重比較のため p 値の補正には Bonferroni 法を用いた。

質問文	発音が容易な選択肢 (選択数)	発音が難しい選択肢 (選択数)	p 値
Q1. これから飛行機に乗ります。チェックイン時に、どちら側の席が良いか係員に聞かれました。どちらを選びますか？	Window ($N = 148$)	Aisle ($N = 33$)	$p < .001$
Q2. 飛行機内では軽食が提供されます。客室乗務員にどちらを食べたいか聞かれました。どちらを選びますか？	Cookie ($N = 105$)	Crisps ($N = 76$)	$p = .26$
Q4. 次に、以下のどちらかを見に行きます。どちらを選びますか？	Palace ($N = 39$)	Castle ($N = 142$)	$p < .001$
Q5. 宿泊しているホテルで夕食をとります。食事形式を選ぶことができ、ウェイターにどちらが良いか聞かれました。どちらを選びますか？	Set menu ($N = 64$)	Buffet ($N = 117$)	$p < .001$
Q7. 次の日の朝に朝食をとります。ウェイターにどちらのパンを食べたいか聞かれました。どちらを選びますか？	Bread ($N = 96$)	Croissant ($N = 85$)	$p = 1.00$
Q8. 今日は世界遺産を見に行きます。どちらの遺産を見たいですか？	Natural heritage ($N = 103$)	Cultural heritage ($N = 78$)	$p = .53$
Q10. 昼食の注文をします。どちらを食べたいですか？	Hot dog ($N = 45$)	Sandwich ($N = 136$)	$p < .001$
Q11. ショッピングをするために、アパレルショップに入りました。どちらを買いたいですか？	Bag ($N = 69$)	Clothes ($N = 112$)	$p = .01$
Q14. 旅行の最終日になりました。まずは以下のどちらかに行きます。どちらを選びますか？	Garden ($N = 104$)	Harbor ($N = 77$)	$p = .37$
Q15. 次に、以下のどちらかを観に行きます。どちらを選びますか？	Concert ($N = 96$)	Theater ($N = 85$)	$p = 1.00$

表 2 発音容易性による選択誘導可能性を検証する 10 問において、2 択の選択肢のうち、どちらの発音がより難しいと感じるかをアンケートで尋ねた結果。選択数に偏りがあるかはカイ二乗検定を行い、多重比較のため p 値の補正には Bonferroni 法を用いた。

質問番号	選択肢内容と選択数	p 値
Q1	Window ($N = 10$) / Aisle ($N = 171$)	$p < .001$
Q2	Cookie ($N = 24$) / Crisps ($N = 157$)	$p < .001$
Q4	Palace ($N = 89$) / Castle ($N = 92$)	$p = 1.00$
Q5	Set menu ($N = 17$) / Buffet ($N = 164$)	$p < .001$
Q7	Bread ($N = 7$) / Croissant ($N = 174$)	$p < .001$
Q8	Natural heritage ($N = 25$) / Cultural heritage ($N = 156$)	$p < .001$
Q10	Hot dog ($N = 26$) / Sandwich ($N = 155$)	$p < .001$
Q11	Bag ($N = 6$) / Clothes ($N = 175$)	$p < .001$
Q14	Garden ($N = 30$) / Harbor ($N = 151$)	$p < .001$
Q15	Concert ($N = 32$) / Theater ($N = 149$)	$p < .001$

回数に偏りがあるかカイ二乗検定を行い、多重比較のため Bonferroni 法により p 値の補正を行った結果、10 問中 9 問 (Q1, Q2, Q5, Q7, Q8, Q10, Q11, Q14, Q15) において選択の偏りに有意差が認められた ($p < .001$)。このことから、有意差があった 9 問については、選択肢間で発音容易性が異なっていたと言える。

本章の実験によって、発音容易性が異なっており、クリック選択では選択に偏りが生じない 2 択の選択肢ペアと質問 (Q2, Q7, Q8, Q14, Q15) を抽出することができた。発音容易性の違いによる誘導可能性を検証する発話選択実験 (4 章) には、抽出した 5 つの質問とそれらの選択肢を用いる。

4. 発話選択時における選択の偏りの検証

選択肢の発音容易性の違いによる発話選択の誘導可能性を検証するため、3 章のクリック選択実験により選定した質問を用いて、発話選択実験を行う。

4.1 実験設計

実験は、実験用に実装したシステム (図 2) を用いて行う。実験用システムは 3 章で使用したシステムをベースとし、発話によって選択が行えるように変更した。発話内容の認識には Web Speech API を用いた。

実験では、参加者に「あなたは、これから 1 人で海外旅行に行くとして、その中で 12 個の出来事が起こり、それぞれの出来事で質問が提示されるため、2 つの選択肢か



図 2 発話選択実験 (4 章) に用いた実験システム

ら 1 つを音声で選び、質問に答えてください。質問は日本語で、選択肢は英語で表示されます。英語選択肢の意味を知らなかったり、発音がわからなかったりする場合は、調べていただいても構いません。」と指示し、12 の質問^{*4}に対して 2 択の選択肢から 1 つを発話選択により回答してもらう。クリック選択実験 (3 章) と同様、質問は 1 問ずつ表示し、質問文は日本語で、選択肢は英語で提示する。

12 問のうち 5 問は発音容易性による選択誘導可能性を検証する質問であり、3 章の実験結果を踏まえて選定した 5 問を用いる。残りの 7 問のうち 5 問は実験の意図を気づかれないようにするためのダミー質問で、他の 2 問は参加者が注意して実験を行っていることを確かめるための質問である。本章の実験では、時系列で起こる出来事の順序として適切となるように、質問の一部の表現を、選択への影響が変わらないように配慮した上で、3 章の実験時と変更した。2 択の各選択肢の表示位置は質問ごとにランダムに決定し、質問の提示順序は参加者間で同一とした。

発話選択の際、システムの認識した発話内容が選択肢と一致せず、参加者が実験を終えることができない可能性がある。そこで本実験では、必ず 2 回目の発話試行時に選択肢が選ばれるようにした。1 回目の発話時は、認識した発話内容と選択肢内容の完全一致により選択の判定を行い、2 回目の発話時は、認識した発話内容と 2 択の各選択肢内容のレーベンシュタイン距離を求め、より類似度が高い選択肢を発話された選択肢として決定した。レーベンシュタイン距離とは、2 つの文字列がどの程度異なっているかを示

^{*4} 発話選択実験に使用した質問と選択肢は、<https://github.com/uyuchiroooo/exp202509> で公開している。

す距離であり、値が小さいほど類似度が高いことを表す。なお、認識された選択肢が参加者の選ぼうとした選択肢とは異なる可能性もあるため、一度認識された場合でも参加者が再度選択を行えるようにした。また、実験参加者が正確な発音を試みなくなることを防ぐために、2回目の発話時にいずれかの選択肢が選ばれるようにしていることは、参加者に知らせなかった。

参加者が12の質問に対して選択を終えた後に、選択肢に用いた英単語の意味を参加者がどの程度知っていたか、また発話選択の難しさについて取得するためにアンケートを実施する。アンケートでは最初に、表示された英語選択肢の中に知っている単語がどの程度あったか、また発話選択がどの程度難しかったかを、それぞれ7段階リッカート尺度を用いて回答してもらう。次に、参加者が注意してアンケート項目を読んでいることを確かめるための質問を行い、最後に性別と年齢を答えてもらう。

実験参加者の募集にはYCSを利用した。実験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。パイロットテストにより実験時間は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、英語の非母語話者である日本人の成人とし、3章の実験に参加した者は、本実験に参加できないようにした。

4.2 実験手順

実験参加者は、PCを用いてGoogle Chromeを使用し、YCSのページから実験用のシステムにアクセスする。実験用システムは最初に実験の概要や参加条件、注意事項を表示し、参加者はマイクを使用できることを確かめるために、システムで用意された音声テストを行う。参加者がマイクを使用できることを確認し、実験参加に同意して設置されたボタンをクリックすると、実験のシナリオに関する説明と、発話選択の練習タスクが表示される。参加者がシナリオについての説明を読み、練習タスクを完了して実験開始ボタンを押すと、1つ目の質問と選択肢が表示される。参加者は質問への回答を発話選択により行う。参加者が12問の質問に対する回答を終えると、システムはアンケートページへと遷移する。参加者がすべてのアンケート項目に対して回答を終え、設置されたボタンを押すと、システムは実験の終了を知らせるページに遷移し、ユニークな実験参加者IDと実験コードを表示する。参加者がそのIDとコードをYCSのページに戻って入力すると、参加者は報酬を受け取ることができる。

4.3 結果

実験参加者は200名であり、実験時間の平均は約3分44秒であった。注意力確認のタスクをクリアしていない参加者は、分析対象として適切でないと考えて除外し、本稿では106名の参加者によるデータを分析対象とした。106名のうち、男性は51名、女性は55名で、トランスジェンダーやノンバイナリー、他のジェンダーと回答したのはいずれも0名、未回答も0名であった。また、106名の参加者のうち、20代が11名、30代が12名、40代が33名、50代が31名、60代が16名、70代以上が3名、未回答が0名であった。

発音容易性による選択誘導可能性を検証した5つの質問について、発音の容易な選択肢が選ばれた数の合計は361、発音の難しい選択肢が選ばれた数の合計は169であった。これらの選択に偏りがあるかカイ二乗検定を行った結果、統計的有意差が認められ($\chi^2 = 69.55, p < .001$)、選択肢の発音容易性によって発話選択が誘導されることが示された。選択誘導可能性を検証した5つの質問について、2択の各選択肢が選ばれた数を質問ごとに求めた結果を表3に示す。各質問において選択の偏りがあるかカイ二乗検定を行い、多重比較のためBonferroni法を用いて p 値の補正を行った結果、5問中4問で選択の偏りに有意差が認められた($p < .05$)。実験後に実施したアンケートで、実験に使用した英語選択肢の中に知っている単語がどの程度あったかという項目に対して7段階リッカート尺度(1: すべて知らなかった, 7: すべて知っていた)を用いて参加者に回答してもらった結果、回答の平均値は6.06 ($SD = 0.93$)であった。このことから、英単語の知識による選択の偏りへの影響は小さかったと考えられる。また、音声選択がどの程度難しかっ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: とても簡単だった, 7: とても難しかった)を用いた回答の平均値は4.50 ($SD = 1.78$)であった。

各実験参加者が、選択誘導を検証した5つの質問のうち、何問において発音が容易な選択肢を選んだかを求めた結果を図3に示す。実験参加者は5問のうち、平均3.41問($SD = 1.02$)において発音の容易な選択肢を選んでいったことがわかった。また、実験参加者の80%以上が、5問中3問以上で発音の容易な選択肢を選んでいったことが明らかになった。

本章の実験では、参加者の1回目の発話時に、認識された発話内容と選択肢内容の完全一致により選択の判定を行ったが、参加者が実験を終了できない可能性を考慮し、2回目の発話時は、認識された発話内容と2つの選択肢内容のレーベンシュタイン距離を計算し、より類似度が高い選択肢を参加者によって発話された選択肢とみなした。参加者の1回目の発話時に、認識内容と選択肢内容が一致したことによって選択できた割合は55.1% ($N = 292$)であ

表 3 発音容易性による選択誘導可能性を検証した 5 問において、発話選択で発音の容易な選択肢と発音の難しい選択肢が選ばれた回数。選択回数に偏りがあるかはカイ二乗検定を行い、多重比較のため p 値の補正には Bonferroni 法を用いた。

質問文	発音が容易な選択肢 (選択数)	発音が難しい選択肢 (選択数)	p 値
Q1. 飛行機に乗りました。機内では軽食が提供され、客室乗務員にどちらを食べたいか聞かれました。どちらを選びますか？	Cookie ($N = 90$)	Crisps ($N = 16$)	$p < .001$
Q4. 次の日の朝になりました。ホテルで朝食をとります。ウェ이터にどちらのパンを食べたいか聞かれました。どちらを選びますか？	Bread ($N = 68$)	Croissant ($N = 38$)	$p = .02$
Q5. 今日は世界遺産を見に行きます。どちらの遺産を見たいですか？	Natural heritage ($N = 80$)	Cultural heritage ($N = 26$)	$p < .001$
Q9. 旅行の最終日になりました。まずは 2 つのうちどちらかに行きます。どちらを選びますか？	Garden ($N = 70$)	Harbor ($N = 36$)	$p < .01$
Q10. 次に、2 つのうちどちらかを観に行きます。どちらを選びますか？	Concert ($N = 53$)	Theater ($N = 53$)	$p = 1.00$

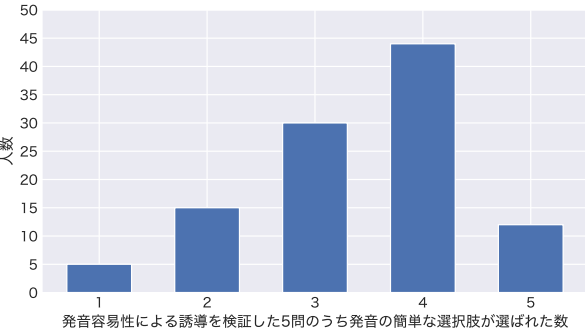


図 3 発音容易性による選択誘導を検証した 5 問において、実験参加者ごとに発音が簡単な選択肢を選んだ質問数を求めた結果

り、残りの 44.9% ($N = 238$) については 1 回目で認識されなかった。2 回目以降の発話時 ($N = 238$) において、参加者が 1 回目に選ぼうとした選択肢と最終的に選んだ選択肢が異なっていた割合を求めることで、発話選択行動の変化を分析した。参加者が 1 回目の発話時に選ぼうとした選択肢は、レーベンシュタイン距離を用いて推測した。複数回発話したとき ($N = 238$) に、最初に選ぼうとした選択肢と最終的に選んだ選択肢が異なっていた割合は、15.1% ($N = 36$) であった。この割合について、発音が容易な選択肢が選ばれたとき ($N = 133$) と、発音が難しい選択肢が選ばれたとき ($N = 105$) に分けてそれぞれ求めたところ、発音が容易な選択肢が選ばれたときは 11.3% ($N = 15$) であった一方で、発音が難しい選択肢が選ばれたときは 20.0% ($N = 21$) であった。つまり、発音の容易な選択肢を最初に選ぼうとし、認識されなかったために最終的に発音の難しい選択肢を選んだ割合の方が、その逆よりも高い傾向にあったと言える。

次に、発音が容易な選択肢が選ばれたときに発話内容と選択肢内容が完全に一致していた割合と、発音が難しい選択肢が選ばれたときに発話内容と選択肢内容が完全に一致していた割合をそれぞれ求めた。ここでは、参加者の 2 回目以降の発話において、認識された発話内容と選択肢内容が完全に一致していたデータも含めて分析した。発音が容易な選択肢が選ばれた回数は 361 回であり、そのうち発話内容と選択肢内容が完全に一致していた割合は 77.01% ($N = 278$) であった一方で、発音が難しい選択肢が選ばれた 169 回において、発話内容と選択肢内容が完全に一致していた割合は 49.70% ($N = 84$) と低いことがわかった。

5. 考察

5.1 発音容易性による選択の誘導

4 章の発話選択実験により、発音容易性の異なる選択肢間で発話選択の偏りに統計的有意差が認められたため、非母語で発話選択する場合には選択肢の発音容易性によって選択が誘導されることが明らかとなり、DP として成立する可能性が示された。本稿では、非母語の発話選択に焦点を当てたが、母語であっても方言を話すユーザはスマートスピーカを操作する際に言葉や発音を変える必要があり、認知的負担が大きい [37] ことが知られている。そのため、母語を用いた発話選択においても、発音容易性の違いによって選択を誘導できてしまう可能性がある。

本稿では、発話選択実験において、参加者が実験を確実に終わらせるようにするために、参加者の 2 回目の発話時に 2 択のいずれかの選択肢が選ばれるようにした。発音が容易な選択肢が選ばれたときに発話内容と選択肢内容が完全に一致していた割合が 77.01% であったのに対し、発音が難しい選択肢が選ばれたときに完全一致していた割合は 49.70% であったため、発話内容と選択肢内容が完全に一致しないと選択することができないような VUI デザインが使用された場合では、発音が難しい選択肢を選ぼうとしたユーザは選択に成功しない可能性が考えられる。本稿の実験では、最初に発音の容易な選択肢を選ぼうとし、認識されなかったために最終的に発音の難しい選択肢を選んだ割合が、その逆（最初に発音の難しい選択肢を選ぼうとし、最終的に発音の容易な選択肢を選ぶ）よりも高い傾向がみられた。しかし、発話内容と選択肢内容が完全一致するまで選択肢を選べないようなデザインでは、ユーザは発音の難しい選択肢を選ぶことを諦め、発音の容易な選択肢を選ぶように変更する傾向が大きくなると考えられる。

GUI における選択の誘導では、選ばせたい選択肢に色をつけて目立たせたり、ボタンサイズを大きくしたりするデザイン (False hierarchy [38]) が使用されている。このような誘導戦略と比較して、発音容易性の違いを悪用した誘導は、ユーザにとって誘導の意図がわかりづらいと考えられる。ユーザは強引な DP に対して否定的な感情を抱く一方で、目立たない DP に対してはそのような感情を抱かない [39] ため、発音容易性の違いを悪用した選択の誘導に対して、ユーザは不信感を抱かない可能性があり、知らず識

らずのうちに選択が誘導されてしまう可能性が考えられる。

5.2 言語を考慮した音声インタラクションの設計

VUIの設計には利用可能なガイドラインが少なく [4, 40], デザイナーが悪意なく DP を使用してしまうことも考えられる。特に、我々が着目した発音容易性の違いを利用してユーザの選択を誘導するデザインは、デザイナーが無意識のうちに使用してしまう可能性がある。本稿は、選択を誘導する言語要素の一つとして発音容易性に着目したが、選択肢表現の難易度の違いなどによっても、ユーザの選択が誘導される可能性が考えられる。このような言語の操作によるユーザの選択行動の誘導可能性についてさらに研究を行い、デザイナーが意図せず DP を使用しないために、言語表現に関する VUI の設計ガイドラインの確立を実現したいと考えている。

5.3 制約と展望

本稿にはいくつかの制約がある。まず、本稿では実験参加者の対象を、英語を非母語とする日本人とし、英語を用いた発話選択において発音容易性の違いを利用した選択誘導可能性について検証した。そのため、実験で得られた結果は、異なる言語や異なる国籍の参加者を対象としたときに変わる可能性がある。また、実験では海外旅行時における選択というシナリオで検証を行ったため、異なる状況における選択に対して結果が一般化できるとは限らない。

本稿は非母語を用いた発話選択に着目し検証を行ったが、今後は母語による発話選択においても発音容易性によって同様に選択が誘導されるか検証し、誘導効果の大きさを比較したいと考えている。また、今後はショッピングなど異なるシナリオを設定して実験を行うことにより、発音容易性の違いを利用した発話選択の誘導可能性について、さらなる検証を行う予定である。

6. おわりに

本稿は、音声ユーザインタフェースを用いたユーザの音声入力において、Deceptive Pattern となりうる問題が存在する可能性に着目し、ユーザが非母語を用いて音声操作する場合に選択肢の発音容易性の違いによって、ユーザの発話選択が誘導されるか検証した。実験の結果、非母語による発話選択では発音の容易な選択肢が選ばれやすいことが明らかとなり、発音容易性の違いを利用した誘導が Deceptive Pattern として成立する可能性が示された。このような誘導はユーザにとって気づきづらく、無意識のうちに行動を操作されてしまう恐れがある。倫理的な音声ユーザインタフェースの設計には、言語の問題を考慮する必要がある、デザイナーは意図せずユーザの行動を操作しないようにするために慎重にデザインする必要がある。

参考文献

- [1] Brignull, H.: Deceptive Patterns. <https://www.deceptive.design/> (参照: 2025/07/04) .
- [2] 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).
- [3] Conca, S. D.: The Present Looks Nothing Like the Jetsons: Deceptive Design in Virtual Assistants and the Protection of the Rights of Users, *Computer Law Security Review*, Vol. 51, p. 105866 (2023).
- [4] Mildner, T., Cooney, O., Meck, A.-M., Bartl, M., Savino, G.-L., Doyle, P. R., Garaialde, D., Clark, L., Sloan, J., Wenig, N., Malaka, R. and Niess, J.: Listening to the Voices: Describing Ethical Caveats of Conversational User Interfaces According to Experts and Frequent Users, *Proceedings of the 2024 CHI Conference on Human Factors in Computing Systems, CHI '24*, pp. 1–18 (2024).
- [5] 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).
- [6] 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).
- [7] Schnelle-Walka, D.: I Tell You Something, *Proceedings of the 16th European Conference on Pattern Languages of Programs, EuroPLoP '11*, pp. 1–26 (2011).
- [8] 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).
- [9] 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).
- [10] 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).
- [11] 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).
- [12] Sin, R., Harris, T., Nilsson, S. and Beck, T.: Dark Patterns in Online Shopping: Do They Work and Can Nudges Help Mitigate Impulse Buying?, *Behavioural Public Policy*, Vol. 9, No. 1, pp. 61–87 (2025).
- [13] Moser, C., Schoenebeck, S. Y. and Resnick, P.: Impulse Buying: Design Practices and Consumer Needs, *Proceedings of the 2019 CHI Conference on Human Factors in Computing Systems, CHI '19*, pp. 1–15 (2019).
- [14] 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).
- [15] 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).
- [16] Seyson, S. and Willett, W.: Exploring the Evolution of Dark Patterns and Manipulative Design on Instagram, *Proceedings of the Extended Abstracts of the CHI Conference on Human Factors in Computing Systems*, CHI EA '25, pp. 1–8 (2025).
- [17] Zagal, J. P., Björk, S. and Lewis, C.: Dark Patterns in the Design of Games, *International Conference on Foundations of Digital Games* (2013).
- [18] Niknejad, S., Mildner, T., Zargham, N., Putze, S. and Malaka, R.: Level Up or Game Over: Exploring How Dark Patterns Shape Mobile Games, *Proceedings of the International Conference on Mobile and Ubiquitous Multimedia*, MUM '24, pp. 148–156 (2024).
- [19] Bösch, C., Erb, B., Kargl, F., Kopp, H. and Pfattheicher, S.: Tales from the Dark Side: Privacy Dark Strategies and Privacy Dark Patterns, *Proceedings on Privacy Enhancing Technologies*, Vol. 2016, pp. 237–254 (2016).
- [20] Nouwens, M., Liccardi, I., Veale, M., Karger, D. and Kagal, L.: Dark Patterns after the GDPR: Scraping Consent Pop-ups and Demonstrating their Influence, *Proceedings of the 2020 CHI Conference on Human Factors in Computing Systems*, CHI '20, pp. 1–13 (2020).
- [21] Gray, C. M., Santos, C., Bielova, N., Toth, M. and Clifford, D.: Dark Patterns and the Legal Requirements of Consent Banners: An Interaction Criticism Perspective, *Proceedings of the 2021 CHI Conference on Human Factors in Computing Systems*, CHI '21, pp. 1–18 (2021).
- [22] Habib, H., Li, M., Young, E. and Cranor, L.: “Okay, whatever”: An Evaluation of Cookie Consent Interfaces, *Proceedings of the 2022 CHI Conference on Human Factors in Computing Systems*, CHI '22, pp. 1–27 (2022).
- [23] 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).
- [24] Gunawan, J., Pradeep, A., Choffnes, D., Hartzog, W. and Wilson, C.: A Comparative Study of Dark Patterns Across Web and Mobile Modalities, *Proc. ACM Hum.-Comput. Interact.*, Vol. 5, No. CSCW2, pp. 1–29 (2021).
- [25] 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).
- [26] 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).
- [27] 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).
- [28] 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).
- [29] Dula, E., Rosero, A. and Phillips, E.: Identifying Dark Patterns in Social Robot Behavior, *2023 Systems and Information Engineering Design Symposium (SIEDS)*, pp. 7–12 (2023).
- [30] European Parliament: Digital Servicess Act***I (2022). https://www.europarl.europa.eu/doceo/document/TA-9-2022-0014_EN.html (参照: 2025/10/13) .
- [31] 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).
- [32] 村上楓夏, 重松龍之介, 大石琉翔, 中川由貴, 中村聡史, 柴崎礼士郎, 鳥居武史, 高尾英行, 清水紗英里, 水原悠貴: 英文選択肢の長さが発話選択による選択に及ぼす影響, 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-211, No. 11, pp. 1–8 (2025).
- [33] Buchta, K., Wójcik, P., Nakonieczny, K., Janicka, J., Gahuszka, D., Sterna, R. and Igras-Cybulska, M.: Microtransactions in VR. A Qualitative Comparison Between Voice User Interface and Graphical User Interface, *2022 15th International Conference on Human System Interaction (HSI)*, pp. 1–5 (2022).
- [34] Shudong, W. and Higgins, M.: An Online Pronunciation Training Support System Designed for Japanese Learners of English, *Fifth IEEE International Conference on Advanced Learning Technologies*, ICALT '05, pp. 171–173 (2005).
- [35] Seaborn, K. and Nakamura, S.: Quality and Representativeness of Research Online with Yahoo! Crowdsourcing, *Frontiers in Psychology*, Vol. 16 (2025).
- [36] Rousseeuw, P. J. and Hubert, M.: Robust Statistics for Outlier Detection, *Wiley Interdisciplinary Reviews: Data Mining and Knowledge Discovery*, Vol. 1, No. 1, pp. 73–79 (2011).
- [37] Harrington, C. N., Garg, R., Woodward, A. and Williams, D.: “It’s Kind of Like Code-Switching”: Black Older Adults’ Experiences with a Voice Assistant for Health Information Seeking, *Proceedings of the 2022 CHI Conference on Human Factors in Computing Systems*, CHI '22, pp. 1–15 (2022).
- [38] Gray, C. M., Kou, Y., Battless, 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).
- [39] Luguri, J. and Strahilevitz, L. J.: Shining a Light on Dark Patterns, *Journal of Legal Analysis*, Vol. 13, No. 1, pp. 43–109 (2021).
- [40] Murad, C., Tasnim, H. and Munteanu, C.: “Voice-First Interfaces in a GUI-First Design World” : Barriers and Opportunities to Supporting VUI Designers On-the-Job, *Proceedings of the 4th Conference on Conversational User Interfaces*, CUI '22, pp. 1–10 (2022).