

対話形式でのストレス管理に向けた チャットボットにおける入力方法の評価

乗瀨 駿平¹ 宮崎 翔¹ 耿 世嫻¹ 下島 銀士¹ 佐藤 安理紗 ジエンジェラ¹ ホシオ シモ²
矢谷 浩司¹

概要：現代社会において多くの人が抱えるとされているストレスは、蓄積してしまうと様々な悪影響をもたらすため、こまめに発散することが重要である。そのためストレスの発散を支援するデジタルアプリケーションが研究・開発されてきているが、ストレスの有効な発散方法とされている感情の表出に着目したものは少ない。そこで本研究では、ストレスから生じる感情を表出し発散することのできるチャットボットシステムの設計を行った。このチャットボットでは、ストレスに感じた出来事とそれにまつわる感情をテキスト入力か音声入力により表出することで、ストレスを管理する事ができる。入力方法による違いを検証するため、2週間にわたり28名の参加者に実際にシステムを利用してもらう実験を実施し、ストレスの発散効果にどのような違いが出るのかについて分析を行った。結果として、ストレス管理のためのチャットボットにおいてはテキスト入力がより好まれ、それぞれの入力方法の利点と欠点が明らかになった。今後の研究においては感情の表出によるストレスの解消を支援する追加の機能の開発と、様々な入力方法が与える影響についてのさらなる検証が必要である。

Examining Input Modality Effects in Interactive Stress Management Chatbots

SHUNPEI NORIHAMA¹ KAKERU MIYAZAKI¹ SHIXIAN GENG¹ GINSHI SHIMOJIMA¹
ARISSA J. SATO¹ SIMO HOSIO² KOJI YATANI¹

1. 序論

ストレスを溜め込まずに発散することは重要である [1] が、実際には十分になされていない。厚生労働省の調査では47.9%の人がストレスを抱えているとされ [2]、世界的にもストレスを感じている人の数は増加している [3, 4]。ストレスは目標達成や成長のためのいい刺激となることもあるが、多くの悪影響をもたらす [5, 6]。抑うつや不安障害といった精神的な影響、頭痛や高血圧といった肉体的な影響、さらには生活習慣の乱れなど社会的な影響も報告されている [5, 7]。これらの悪影響を防ぐため、ストレスチェックの義務化などの取り組みがなされているが、ストレスへの対処については個人に委ねられることが多い [8]。

ストレスの対処は「ストレスコーピング」として心理

学や精神医学の領域において研究されてきた [9]。そして様々なコーピング理論が開発され、医療やカウンセリングなどの場で提供されている。しかし心理カウンセリングを利用したことがある人の割合は6%に留まる [10] など、知見は十分に活用されていない。時間や場所の制約といった構造的な障壁以外にも、羞恥心やスティグマといった心理的・社会的な障壁があり、活用が広まっていない現状がある [11, 12]。

そのような障壁を下げる方法として、スマートフォンやパソコンなどで提供するデジタルコーピングは注目を集めている。好きなタイミングで利用できることや、必要な労力が少ないことなどが利点であり、様々なデジタルコーピング用のアプリやシステムが研究・開発されてきている [13, 14]。しかし既存のアプリは科学的な裏付けが不十分であったり、定着率が低いといった課題が指摘されている [13, 14]。また、単純に手法を紹介してだけのものも多

¹ 東京大学 Interactive Intelligent Systems Laboratory

² オウル大学

く [15,16], 一つの手法についてその効果を高めるような方法の研究はあまりなされていない。

そこで本研究は, ストレスのこまめな発散を支援するシステムを設計する. そして定着率を高め, ストレス発散効果を高めるようなインタフェースについての検討を行う。

2. 関連研究

2.1 ストレスコーピングツール

医学や心理学の領域において研究されてきたストレスコーピングを支援するツールやデバイスについての研究は2010年ごろから盛んになり, スマートフォンの普及とともに加速してきている. 主な研究としては複数のコーピングを紹介したり実行を支援したりするものがある [15,16]. また研究にとどまらず多くのアプリケーションが公開されており, 時間や場所の制約なく日常的に利用することのできる方法として期待されている [13]. しかし利用率や定着率の低さが課題として指摘されている [13].

そこで利用場面に合わせたストレスコーピング支援の方法として, アプリケーションを新規に開発するのではなく, 既存のアプリケーション上で動作するチャットボットを利用する方法が注目されている [14,17]. ビジネスコミュニケーションツール上のチャットボットで仕事におけるストレスを発散できるよう支援する研究では, 「動画による逃避」「落ち着く, 感謝を示す」「ストレスの表出」の3種類のコーピング手法を比較している [14]. 3つ目の「ストレスの表出」が最もストレスを低減させた一方で, ユーザからの主観的な評価は低かった. 書き出したストレスや感情に向き合ってしまう, かえってストレスに感じてしまった参加者がおり, 低減効果を感じにくくなった結果と考察されている.

さらにストレスを抱えている割合が高い女性に注目し, どのようにストレスを感じ, 対処しているのかについて調べた研究では, 既存のデジタルコーピングアプリケーションは「うるさく気が散る (overwhelming and distracting)」ためにあまり好まれていなかった一方で, 音声入力インタフェースについては「すっきりでき, 自然である (refreshing and natural)」とされ, 評価が高かった [18].

そこで本研究ではストレスに伴う感情の表出によるコーピング方法に着目し, より利用しやすくストレス低減効果のあるチャットボットシステムのインタフェースについて検討する。

2.2 感情発散型のコーピング

ストレスコーピングは, 取り組む対象によって「問題焦点型」と「情動焦点型」に大きく分けられる [9]. 本研究においては, こまめなストレスの発散を支援するため, 情動焦点型に注目する. 情動焦点型はさらに「感情発散型」と

「感情抑圧型」に分類されている [19]. ストレスによる怒りや不満, 悲しみなどの感情を, 感情発散型は外に出して気持ちを整理するのに対し, 感情抑圧型は自分の心の中に抑圧してしまう. 前者の感情発散型コーピングの方がより良い精神状態につながりやすく, ストレス関連疾患を予防する効果があるとされている [19]. また2.1節でも述べた通り先行研究 [14]において最もストレス低減効果が高かったため, 本研究においては感情発散型コーピングを基本となるコーピング手法として採用した。

一方で, 「感情への注目や発散」を他のコーピングに対して「あまり有用でない」としている研究もある [20]. このように評価が分かれる原因としては, ストレスフルな体験をめぐる感情の記述は, 感情を客観視でき冷静になれる場合と, 反対に自分の感情を否定的に再評価してしまいストレスを強化してしまう場合があるためである [21].

本研究では感情発散型コーピングにおいて入力方法がストレス発散効果に対してどのような影響を与えるのかについて検証し, 効果に対して影響を与える要因について検討を行う。

2.3 チャットボットにおける入力インタフェース

チャットボットの研究においては入力インタフェースとしてテキスト入力や音声入力が利用されている [22,23]. 画面の操作によるテキスト入力は最もよく使われている [22,23] 一方で, 音声入力の方が有効である状況が存在することも定性的な調査により明らかになってきている [24,25]. 音声はテキストと比べ, 知覚や認知がしやすく, またサービスの満足度を高めやすい特徴があるとされている [24]. また米国とインドにおけるチャットボットの利用経験や期待について調査した研究 [25]では, 母国語での入力において音声入力がテキスト入力よりも高く評価されていた他, 手が塞がっている場合などに特に効果的であると評価されていた。

チャットボットや出来事の記録において, 入力方法が与える影響を定量的に評価した研究もある. 銀行用のチャットボットにおける入力方法の比較研究 [26]では, 音声入力の方がテキスト入力よりも役に立ち, 自己承認的であると感じられていた. またフードログアプリにおける入力方法の比較研究 [27]でも, 音声入力の方が少ない負担で情報量多く記録でき, また内省的な効果もあるとしていた. しかし, ストレス管理のためのチャットボットにおける入力方法について, 定量的に比較し評価した研究はまだない. そこで我々は, 入力方法が利用のしやすさや効果に与える影響についての調査を行う. そしてリサーチクエスチョンとして, 「感情発散型のストレスコーピングチャットボットにおいて, 入力方法がストレス解消効果にどのような影響を与えるのか」を設定した。



図 1: システムの利用画面 (a). ストレスに感じた出来事とその時の感情についての記述を依頼するメッセージ (b). 画面下部のメニュー (c) から好きなタイミングでストレスレベルの報告と、ストレスに感じた出来事についての記述が行える。記述後には感謝のメッセージ (d) と、記述前後でどのようにストレスレベルが変化したかについての質問 (e) が送信される。

3. システム構成

3.1 全体の構成

我々は、ストレスに感じたことについて質問し (3.2 節)、テキスト入力と音声入力による回答を受け付け (3.3 節)、ストレスレベルを計測し (3.4 節)、そして雑談ができる (3.5 節) チャットボットシステムを開発した。

システムは LINE^{*1} 上のチャットボットとして、Google Apps Script (GAS)^{*2} を用いて実装した。デジタルコピーングツールの課題である離脱率の高さ [13, 28] を解消するため、日常的によく使われているメッセージングアプリである LINE 上のチャットボットシステムとした。

利用者のインタフェースを図 1 に示す。

3.2 ストレスの記述

感情発散型コーピングを促すため、ストレスにまつわる記述を支援する機能を持たせた。図 1(b) に示すように、ストレスに感じた出来事やその時の感情についての記述を依頼する。記述後には、図 1(d) に示した感謝のメッセージと、図 1(e) に示したストレスレベルの変化を質問するメッセージを送信する。ストレスレベルの変化は、メッセージ上の「非常に上がった」から「非常に下がった」までの 5

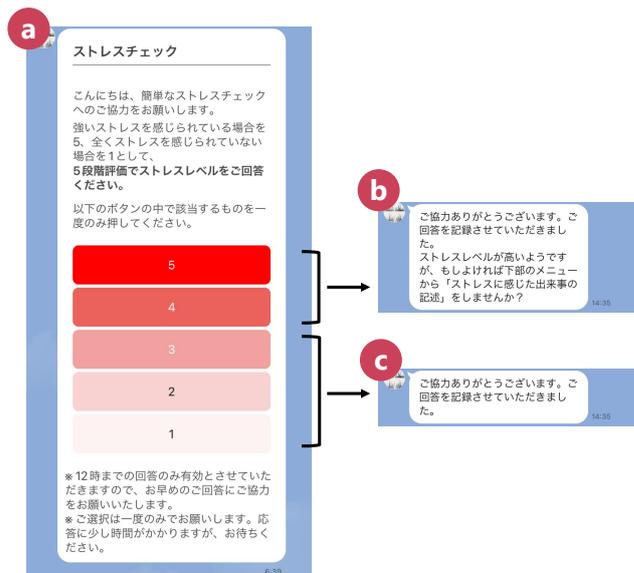


図 2: ESM メッセージの画面。ストレスレベルについて質問するメッセージ (a) 上の 1 から 5 までの 5 段階の選択肢から該当するものを押して回答する。4 か 5 を選んだ場合にはストレス内容の記述を依頼し (b), 3 以下の場合には記録した旨を伝えるメッセージのみ送られる。

つのボタンのうちから当てはまるものを押すことにより回答できる。

また、好きなタイミングで利用できるようにすることが先行研究において推奨されていた [14] ため、画面下部のメニューの右側のボタンから任意のタイミングで記述が行えるように設計した (図 1(c)).

3.3 入力方法

入力方法はテキスト入力と音声入力に対応させた。テキスト入力は、キーボードを使って文字により入力することを指し、キーボード入力かフリック入力などの細かい方法の指定は行わなかった。音声入力は、入力内容を声に出し録音したボイスメッセージを送信する方法を採用した。スマートフォンの音声認識機能により、音声をテキスト変換する音声入力の使用は含めなかった。これはキーボードによるテキスト入力との区別ができず、使用の有無が判別できないためである。

3.4 ストレスレベルの記録

経験サンプリング法 (以下 ESM) に基づいたストレスレベルの記録にも対応させた。ESM は定期的に質問を送り回答を集める手法 [29] であり、想起バイアスを抑え、経時的変化を捉えることができる [30]。利用者の作業を増やしてしまうが、スマートフォンを用いることで有効に測定することができる [31]。

具体的には、図 2(a) に示すようなメッセージを決まった時間に送り、5 段階のストレスレベルを記録することが

*1 <https://line.me>

*2 <https://workspace.google.co.jp/intl/ja/products/apps-script/>

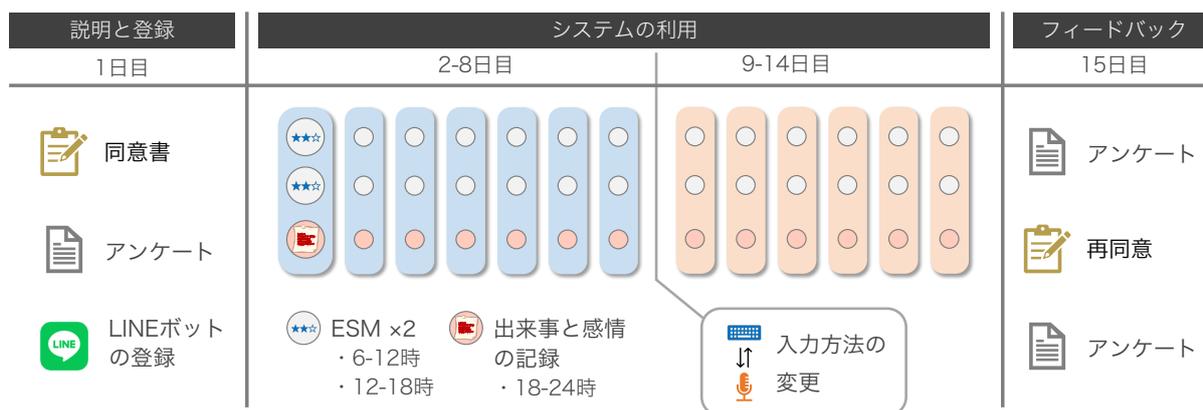


図 3: ユーザ実験は初日と最終日に説明と同意, アンケートなどの作業があり, 間の 13 日間システムを二つの入力方法により利用してもらい, データを集めた。

できる。4 か 5 が記録された場合には, 内容の記述によるコピーを提案する (図 2(b))。

3.5 雑談

記述や記録以外の時間にチャットボットとして機能するよう, 人工知能による自動雑談機能を実装した。リクルートの A3RT が提供している Talk API^{*3}を利用した。Talk API は日本語による日常会話に対応した AI であり [32], いくつかの先行研究において利用されている [33,34]。

4. ユーザ実験

4.1 手順

初日に説明と登録を行った後, 13 日間システムを利用してもらい, 最後にフィードバックをお願いした (図 3)。全ての工程はオンラインで行われた。

初日にはユーザ実験についての説明を行い, 同意書とアンケートへの回答および LINE ボットの登録を参加者をお願いした。実験の目的は, ストレスに感じていることの内容についてのデータを集めるものだと説明した。実際の目的を伝えないことにより, 入力方法によって故意に利用態度を変えられることを防いだ。

アンケートの内容は, 知覚されたストレス尺度の日本語版 [35] と Big Five 尺度の日本語版 [36] を含み, それぞれの質問はランダムな順序で提示した。

2 日目から 14 日目までの 13 日間, LINE ボットを通じて ESM と記述の作業を依頼した。3.4 節で説明した ESM については 6 時台と 12 時台の一日 2 回送信した。3.2 節で説明したストレスに感じた出来事やその時の感情の記述についての依頼のメッセージは 18 時台に送信した。全ての作業は 6 時間以内に行われたもののみ受け付けた。

ただし 9 日目にそれぞれの参加者に割り当てられた入力方法を入れ替えた。G1 の参加者はテキスト入力から音声入力に, G2 の参加者は音声入力からテキスト入力に, 指

定の入力方法を変更した。

最終日である 15 日目には 3 つの作業を依頼した。まず最初のアンケートでは, システムの利用前後での長期的なストレスレベルの変化を計測するため, 知覚されたストレス尺度の日本語版 [35] に再び答えてもらった。

次にユーザ実験の本来の目的について説明した。この説明を受けて, 参加を続行するか終了するかについての質問をした。

参加の続行に同意した参加者については, 最後にフィードバック用のアンケートへの記入をお願いした。アンケートはチャットボット自体と二つの入力方法それぞれについての質問, そして入力方法についての好みと自由記述からなった。

参加者に対しては謝礼として 2500 円を支払い, フィードバックアンケートの回答者には追加で 300 円を支払った。

4.2 参加者

クラウドソーシングサイト^{*4}上で 20 代から 60 代までの 28 名の参加者を集めた (男性, 女性とも 14 名ずつ)。入力方法間の対称性を保つため, 参加者は使用する入力方法の順序によって G1 と G2 の 2 つのグループに分けられた。

G1. テキスト入力 → 音声入力

G2. 音声入力 → テキスト入力

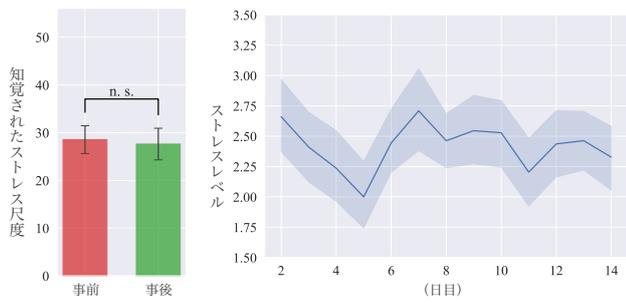
5. 結果

5.1 ストレスの変化

実験参加前後での知覚されたストレス尺度を比較したところ, 0.93 減少したが有意な差は見られなかった (図 4(a))。参加者のうち, 15 名は減少し, 12 名は増加し, 1 名は変わらなかった。ESM で収集されたストレスレベルの日ごとの平均の推移は図 4(b) に示す通りである。平均は 1 から 5 の 5 段階評価の中で 2.41 ($\sigma = 1.06$) であった。6

^{*3} <https://a3rt.recruit.co.jp/product/talkAPI/>

^{*4} クラウドワークス <https://crowdworks.jp/>

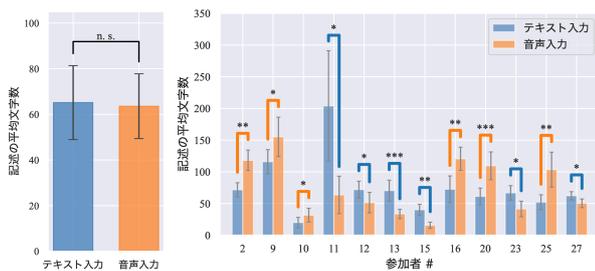


(a) 長期的ストレス (b) 短期的ストレスの変化

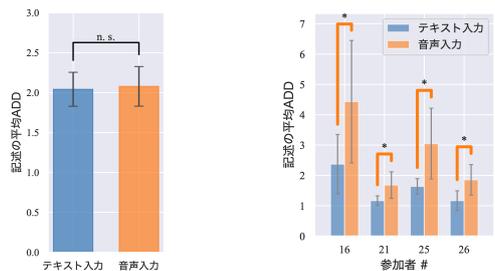
図 4: 参加者によって報告されたストレスレベルの変化. (a) は知覚されたストレス尺度により評価した長期的なストレスレベルを表している. 実験の前後ともに平均値は中庸とされる範囲内であり, 前後での有意な変化はなかった. (b) は ESM により収集された短期的なストレスレベルの日別平均の推移を示している. エラーバーやエラーバンドは 95%信頼区間を示している.

時台の値と 12 時台の値には, 有意な差は見られなかった.

5.2 記述の内容



(a) 入力方法別の平均文字数 (b) 入力方法間で有意差のあった参加者の平均文字数



(c) 入力方法別の平均 ADD (d) 入力方法間で有意差のあった参加者の平均 ADD

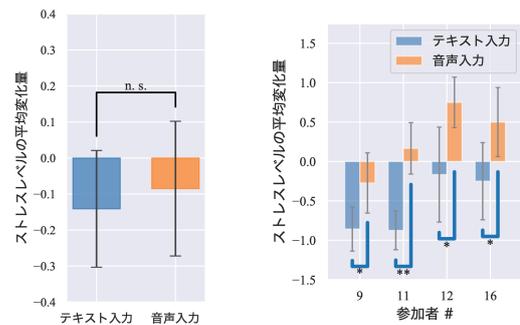
図 5: ストレスについての記述の内容の分析. (a) は全ての参加者の平均文字数であり, 有意差はなかった. (b) は入力方法によって有意差があった参加者 12 名の平均文字数である. (c) は全ての参加者の平均 ADD であり, 有意差はなかった. (d) は入力方法によって有意差があった参加者 4 名の平均 ADD である. エラーバーは 95%信頼区間を示している. (*: $p < .05$, **: $p < .01$, ***: $p < .001$)

記述の内容については, 平均の文字数と平均係り受け距

離 (ADD) を入力方法ごとに分析した. テキスト入力における平均文字数は 65.2 文字 ($\sigma = 43.8$) であり, 音声入力 ($\bar{x} = 63.6, \sigma = 38.4$) に比べてわずかに多かったが, 有意差はなかった (図 5(a)). 平均 ADD については, 音声入力 ($\bar{x} = 2.08, \sigma = 0.67$) がテキスト入力 ($\bar{x} = 2.04, \sigma = 0.57$) に比べ, わずかに長かった (図 5(c)).

全体の平均においては入力方法による有意差は見られなかったが, 参加者ごとの比較においては有意差を確認することができた. 図 5(b) に示した 12 名の参加者は平均文字数に有意差があった. 6 名の参加者 (P2, P9, P10, P16, P20, P25) はテキスト入力の方がより長く, 反対に別の 6 名の参加者 (P11, P12, P13, P15, P23, P27) は音声入力の方がより長く入力していた. 図 5(d) に示した 4 名の参加者 (P16, P21, P25, P26) は, 音声入力の方が平均 ADD が有意に長く, より複雑であったり情報量の多い記述をしていた.

5.3 ストレスの発散効果



(a) 入力方法別のストレスレベルの平均変化量 (b) 入力方法間で有意差のあった参加者のストレスレベルの平均

図 6: 記述前後でのストレスレベルの平均変化量. (a) は全ての参加者の平均であり, 有意差はなかった. (b) は入力方法によって有意差があった参加者 4 名の平均変化量であり, 全員テキスト入力の方がよりストレスレベルが下がっていた. エラーバーは 95%信頼区間を示している. (*: $p < .05$, **: $p < .01$, ***: $p < .001$)

記述後に質問したストレスレベルの変化について, 大きく下がったを-2, 下がったを-1, 変わらないを0, 上がったを1, 大きく上がったを2とした時の5段階評価の平均値で変化量を計算した. 全参加者の平均は, テキスト入力時の方が音声入力時よりわずかに下がっていたが, 有意差は見られなかった (図 6(a)). テキスト入力時については, 記述文字数と負の相関 ($r = -0.38, p < 0.05$) があり, 記述量が多いほどストレスが解消されていた.

また記述内容と同様に, 記述方法によって大きな差が見られた参加者がいた (図 6(b)). 個人ごとに平均を求めたところ, 4 名の参加者 (P9, P11, P12, P16) でテキスト入力の方が音声入力よりもストレスレベルが有意に下がっ

ていた。この4名の参加者は他の参加者に比べテキスト入力時の文字数が多く ($r = 0.48, p < 0.01$)、音声入力のNASA-TLX値が高く ($r = 0.48, p < 0.01$) より負荷を感じていた。

記述直後のストレスレベルの変化量については入力方法による有意差は見られなかったが、アンケートにおけるストレス解消効果の評価については有意差が見られた。図7(a)はそれぞれの入力方法を用いたときのストレス解消効果についての評価の平均値を示しており、テキスト入力 ($\bar{x} = 3.21, \sigma = 0.99$) は音声入力 ($\bar{x} = 2.61, \sigma = 1.10$) に比べ有意に高かった。

ストレス解消効果について、アンケートの自由記述欄ではそれぞれの入力方法の長所と短所が触れられていた。テキスト入力の長所としては、振り返って整理することができる点 (P7, P12, P13, P19, P21) や、客観視できる (P6, P8, P9) などが評価されていた。P14はストレスの原因と対処方法について考えることができたことと答えた。

「他人が読んで分かるような形で文章にするよう意識することで、思考が整理されることがありました。また、記述した文章を後で読み返すことで、自分が感じたよりもショッキングな出来事ではなかったと、認識を新たにすることができた場面もありました。動揺していたり疲れていたりすればするほど、より効果を感じました。」 [P9]

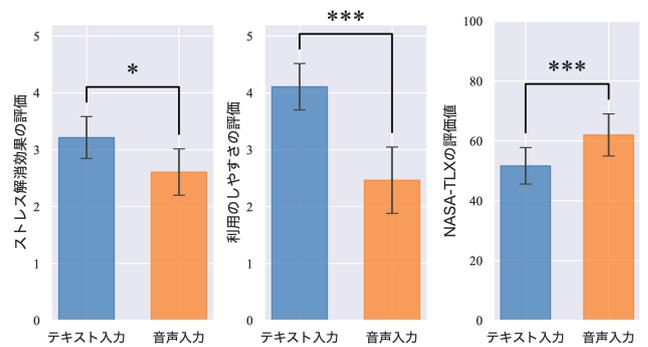
テキスト入力の短所としては、ストレスを感じたことを思い出してしまうこと (P13, P15) や、スマートフォンの小さな画面で入力する負担 (P7, P10, P13, P23) などが挙げられていた。

音声入力については、使いやすさ (P3, P15, P19, P20, P24) や率直な気持ちを表現できること (P14, P15) などが評価されていた。また記述する内容が単純な場合には効果できであると感じた参加者もいた (P5, P13, P15)。P20は実験中に新型コロナウイルス感染症に罹り、声を出すことが健康状態の確認にもつながったという。

「コロナに罹患していた時は家族とも離れ、仕事にも行かなかったので1日にこの実験中にしか言葉を発さなかった日がありました。久しぶりに自分の声を聞きましたし、咽頭痛で喉が荒れている、健康がすぐれていないのだと振り返る事が出来ました。」 [P20]

しかしながら、音声入力は短所も多く挙げられた。音声入力に慣れていないことのほか (P6, P14, P23)、多くの参加者が環境的な制約があることに触れており (P1, P7, P8, P17, P22, P24)、どこでも使うことができることが評価されたテキスト入力 (P22, P26) と対照的な結果となった。

「赤ちゃんがいるので、声を出すのを躊躇する場面が



(a) ストレス解消効果 (b) 利用のしやすさに (c) 利用時の精神的負担についての評価。5段階 評価で、高いほど解 評価で、高いほど利用 の評価値。0-100で、消されたと感じられて 評価値。0-100で、消されたと感じられて 高いほど負担が大きい いる。 れている。 と感じられている。

図7: 最終日のアンケートにおける、それぞれの入力方法の評価。エラーバーは95%信頼区間を示している。(*: $p < .05$, **: $p < .01$, ***: $p < .001$)

多かった (赤ちゃんが眠っているときなど)。」 [P8]

入力内容を読み直すことができず、正しく伝わっているか不安であったという意見 (P5, P9, P18, P27) や、修正できないために緊張したという意見 (P11, P12, P13, P18) も見られた。一度話し始めると、話したい内容を忘れてしまうことがあること (P1, P18, P20) や、忘れないように事前に準備をすることが負担であること (P12, P23, P25) なども短所として挙げられていた。

「やり直しが聞かないので入力前に整理してはいたが上手く出来なかった場合があった。」 [P25]

5.4 利用のしやすさと負担

それぞれの入力方法を使用した際のシステムの利用のしやすさの評価は、テキスト入力 ($\bar{x} = 4.11, \sigma = 1.10$) の方が音声入力 ($\bar{x} = 2.46, \sigma = 1.57$) よりも高かった (図7(b))。5段階中1か2の低評価をつけた人は、テキスト入力では3人だけだったが、音声入力では18人いた。反対に、4か5の高評価をつけた人は、テキスト入力では22人いたが、音声入力では9人だけだった。利用のしやすさはストレスの発散効果と優位な相関がテキスト入力 ($r = 0.52, p < 0.01$) と音声入力 ($r = 0.71, p < 0.001$) とともに見られた。

また、NASA-TLX [37] の平均値はテキスト入力 ($\bar{x} = 51.7, \sigma = 16.5$) の方が音声入力 ($\bar{x} = 62.0, \sigma = 19.0$) よりも低かった (図7(c))。そして利用のしやすさと負担にはテキスト入力 ($r = -0.49, p < 0.01$) と音声入力 ($r = -0.47, p < 0.05$) とともに負の相関が見られた。

5.5 入力方法に対する主観的な評価

最終日のアンケートにおいて、どちらの入力方法をより

使いたいと思うかを尋ねたところ、21名の参加者がテキスト入力と回答し、音声入力よりも有意に多かった。また、テキスト入力と回答した参加者は、音声入力と回答した参加者よりも、音声入力の発散効果と利用のしやすさについての評価が有意に低かった（それぞれ $p < 0.001$, $p < 0.01$ ）。

なぜその入力方法をより使いたいかと思ったのかについての自由記述回答を、選んだ入力方法別に見ていく。テキスト入力を選んだ参加者は、ゆっくりと落ち着いて入力できること（P1, P2, P12, P16, P17, P18, P24）や、見返して修正できること（P5, P6, P11, P12, P13, P14, P17）などを評価した。また、テキストに対してテキストで返ってくること（P11）なども挙げられていた他、音声入力の短所について言及した参加者もいた。環境や状況の制約があること（P1, P3, P11, P20）や緊張してしまうこと（P1, P5, P12）などが、音声入力を選ばなかった理由として挙げられていた。

音声入力を選んだ参加者は、主に利用のしやすさ（P7, P10, P15, P19）を理由として挙げていた。また、P20は音声入力がテキスト入力よりも多い情報量を持つことに触れていた。

「音声でしか伝わらない人間の感触があると思います。」 [P20]

6. 考察

6.1 入力方法の影響

ユーザ実験では、音声入力に比べテキスト入力を使用して記述した場合の方が、参加者はストレス解消効果を強く感じていた。アンケートにおけるストレス解消効果の評価はテキスト入力の方が優位に高かった。また4名の参加者において、記述前後でストレスレベルが優位にテキスト入力の方が下がっていた。そしてテキスト入力が音声入力よりも優位に好まれていた。これらの理由から、テキスト入力が記述によりストレス解消効果をより高めていたと結論づけられる。

6.2 利用しやすさの影響

テキスト入力の方が効果的であるという結果は、音声入力の方が優れているとする2つの先行研究 [26, 27] の結果と矛盾する。この違いにはそれぞれの入力方法の利用しやすさが影響していると考えられる。どちらの入力方法についても、利用のしやすさとストレス解消効果の評価に相関関係が見られた。また、テキスト入力をより好んだ参加者は、音声入力のストレス解消効果だけでなく利用のしやすさについても優位に低く評価した。これらの結果は、利用のしやすさがストレス解消効果や好まれやすさに影響を与えている可能性を示している。

そして参加者は音声入力がストレスの記述において利用

しにくい理由について、主に次の3点を挙げていた。

6.2.1 環境の制約

参加者は環境や状況によっては音声入力の使用が難しかったと述べていた。具体的には、職場にいたり、赤ちゃんが寝ていたりする状況を挙げていた。音声入力のこの短所については2つの先行研究においても触れられていた。銀行のチャットボットの研究 [26] のユーザ実験参加者は、個人情報や公共の場で声に出したくないと述べていた。しかしこのユーザ実験は実験室実験であったため、環境の制約について実感しにくく、大きくは取り上げられていなかった。フードログのアプリの研究 [27] においても、周囲に人がいる場合などには利用が制限されてしまうことに触れていた。

我々の設計したチャットボットはストレスを感じたタイミングで利用されることを期待している。ストレスは周囲の人間の言動が原因となることもあり、そうした場合に相手の目の前で声に出してストレスに感じた内容を記録することは困難である。そのため、音声入力の短所である環境の制約が、他のチャットボットに比べてより影響していたと考えた。無声発話により入力することのできる技術 [38] なども開発されているが、実用化には至っていない。そのためストレス管理チャットボットにおける音声入力の利用にあたり、環境の制約の影響は大きく、音声を使わず指などを使って入力できる方法を搭載する必要がある。

6.2.2 修正不可能性

修正や加筆をしながら入力することができるテキスト入力に対して、音声入力は一度録音が完了し送信されると内容を訂正することができない。この点は、一回で録音を成功させなければならないという緊張感を参加者に与えてしまった。また、一度話し始めると、話そうとしていた内容を忘れてしまうといったことも起きていた。そのため、音声入力によってストレスを感じたことを説明することは難しいと感じられていた。

修正できないことによって生じるこれらの問題については、先行研究では明確には指摘されていなかった。銀行用のチャットボット [26] は、質問や操作の要求など、目的が明確で入力内容が短く簡潔であることが多かった。フードログアプリ [27] は食事内容の記録に音声入力を使っていたが、項目ごとに分けて入力するような設計となっていた。また再録音や一時停止機能があったため、話している途中で内容を忘れてしまった参加者は、これらの機能を利用して内容を整理していた。

我々の設計したチャットボットはストレスに感じた出来事と、その時の感情について、一度に記録するように求めた。そのため記録する内容が長く、複雑になる傾向があった。思い起こし、整理をしながら2つの項目について漏れなく伝わる文章を音声入力で記録することは難しいため、

我々の結果は先行研究と異なる結果になったと考えられる。音声入力で記録する際には、質問や項目を分けることで一度の発話を短く単純なものとする必要がある。また再録音や一時停止などの機能をつけて、修正や整理ができるようにすることで、音声入力の欠点を補い利点を活かすことができるかもしれない。

6.2.3 伝達の不透明性

音声入力の場合、記録したい内容が正しく伝わっているか不安に感じることがある。参加者は送信した音声を再生することができ、録音できているかについて自分で確認することはできたが、記録に自信のない参加者がいた。今回のチャットボットは音声入力の場合もテキスト入力の場合と同様の定型文を返信しており、送信された内容の確認を行っていないことが影響していたと考えられる。銀行のチャットボット [26] は Wizard of Oz 法によるユーザ実験を実施しており、内容に応じて人間が適切な返信をしていた。またフードログアプリ [27] は記録用のアプリでチャットボットではなかったため、他人が認識できるかについて考える必要がなかった。

さらに、ユーザ実験において実験の目的をストレスの原因についての調査であると偽ったことの影響も否定できない。ストレスを感じたことを実験実施者である我々に報告する作業であると捉え、友人に話す際や自分自身のための記録する際とは異なった感覚を与えていた可能性がある。実際に、最終日のアンケートにおいて、「報告」や「作業」といった言葉が使われていた。

そのため、入力方法として音声入力を用いる場合には、音声を即座にテキスト情報に変換し、内容に応じて返信をしたり、正しく伝えることよりも溜め込まずに吐き出すということ自体に集中させるデザインにしたりすることで、この問題点を克服することができると考えられる。

6.3 制約と展望

本研究にはいくつかの制約がある。まず、入力方法についてテキスト入力と音声入力のみを比較した点である。先行研究 [18, 22, 23, 26] において利用や言及されていたためこの2つの入力方法を選んだが、画像入力や選択肢から選ぶといった他の入力方法も検討する必要がある。また今回は音声入力をボイスメッセージの送信によって行ったが、端末の音声認識を利用する方法もある。音声認識を利用した場合、テキストを修正することができるため、異なる結果となる可能性があり、検証が必要である。

次に、ユーザ実験の方法についてである。今回の実験では、目的を偽って説明した。これは入力方法によって故意に利用態度を変えさせないためであったが、報告や作業といった感覚を参加者に与えてしまった可能性がある。また毎日ストレスを感じた出来事について記述するように求め

たため、ストレスを感じた出来事が特にはなかった場合には、記述の作業自体がストレスを感じさせてしまうことがあった。そのため、ストレスを感じた出来事がなかったという報告にも対応し、記述の最低回数を日ごとではなく週ごとなどで規定するなどの工夫が今後は必要である。

7. 結論

本研究では、人々がストレスをこまめに効果的に発散することを支援するため、日常的に使われているメッセージングアプリ上で利用できるチャットボットを開発し、チャットボットの入力方法の影響を検証した。その結果、自分の状況を振り返り客観視できるテキスト入力が音声入力よりも好まれることがわかった。またストレスを感じた経験や感情の共有には多くの内容を入力する必要があり、音声入力にはあまり適していないことが明らかになった。今後は、音声インタフェースや、発散の感覚を与えるような機能により、感情発散型コーピングの長所を強化するようなインタフェースを設計していく必要がある。

謝辞 ユーザ実験にご参加いただいた皆さまならびにパイロットスタディにご協力いただいた皆さまに、厚く御礼申し上げます。本研究は、株式会社メルカリ R4D とインクルーシブ工学連携研究機構との共同研究である価値交換工学の成果の一部です。

参考文献

- [1] Richard S Lazarus. The role of coping in the emotions and how coping changes over the life course. In *Handbook of emotion, adult development, and aging*, pp. 289–306. Elsevier, 1996.
- [2] 厚生労働省. 2019 年 国民生活基礎調査の概況, 2020. <https://www.mhlw.go.jp/toukei/saikin/hw/k-tyosa/k-tyosa19/index.html>.
- [3] Gallup. Americans' stress, worry and anger intensified in 2018, 2019. <https://news.gallup.com/poll/352205/2020-sets-records-negative-emotions.aspx>.
- [4] Gallup. World unhappier, more stressed out than ever, 2022. <https://news.gallup.com/poll/394025/world-unhappier-stressed-ever.aspx>.
- [5] 二木鋭雄. 良いストレスと悪いストレス. *日本薬理学雑誌*, Vol. 129, No. 2, pp. 76–79, 2007.
- [6] Clairice T Veit and John E Ware. The structure of psychological distress and well-being in general populations. *Journal of consulting and clinical psychology*, Vol. 51, No. 5, p. 730, 1983.
- [7] Neil Schneiderman, Gail Ironson, and Scott D Siegel. Stress and health: psychological, behavioral, and biological determinants. *Annual review of clinical psychology*, Vol. 1, p. 607, 2005.
- [8] 島津明人. ストレスチェック制度とセルフケア支援. *医学のあゆみ*, Vol. 263, No. 3, pp. 246–250, 2017.
- [9] Richard S Lazarus and Susan Folkman. *Stress, appraisal, and coping*. Springer publishing company, 1984.
- [10] J-Net21. 市場調査データ 心理カウンセリング, 2020. <https://j-net21.smrj.go.jp/startup/research/service/cons-counseling.html>.

- [11] Sarah Clement, Elaine Brohan, Debra Jeffery, Claire Henderson, Stephani L Hatch, and Graham Thornicroft. Development and psychometric properties the barriers to access to care evaluation scale (bace) related to people with mental ill health. *BMC psychiatry*, Vol. 12, pp. 1–11, 2012.
- [12] David C Mohr, Joyce Ho, Jenna Duffecy, Kelly G Baron, Kenneth A Lehman, Ling Jin, and Douglas Reifler. Perceived barriers to psychological treatments and their relationship to depression. *Journal of clinical psychology*, Vol. 66, No. 4, pp. 394–409, 2010.
- [13] Amit Baumel, Theresa Fleming, and Stephen M Schueller. Digital micro interventions for behavioral and mental health gains: core components and conceptualization of digital micro intervention care. *Journal of medical Internet research*, Vol. 22, No. 10, p. e20631, 2020.
- [14] Esther Howe, Jina Suh, Mehrab Bin Morshed, Daniel McDuff, Kael Rowan, Javier Hernandez, Marah Ihab Abidin, Gonzalo Ramos, Tracy Tran, and Mary P Czerwinski. Design of digital workplace stress-reduction intervention systems: Effects of intervention type and timing. In *CHI Conference on Human Factors in Computing Systems*, pp. 1–16, 2022.
- [15] Pablo Paredes and Matthew Chan. Calmmenow: exploratory research and design of stress mitigating mobile interventions. In *CHI'11 Extended Abstracts on Human Factors in Computing Systems*, pp. 1699–1704, 2011.
- [16] Pablo Paredes, Ran Gilad-Bachrach, Mary Czerwinski, Asta Roseway, Kael Rowan, and Javier Hernandez. Poptherapy: Coping with stress through pop-culture. In *Proceedings of the 8th international conference on pervasive computing technologies for healthcare*, pp. 109–117, 2014.
- [17] Lenin Medeiros, Tibor Bosse, and Charlotte Gerritsen. Can a chatbot comfort humans? studying the impact of a supportive chatbot on users' self-perceived stress. *IEEE Transactions on Human-Machine Systems*, Vol. 52, No. 3, pp. 343–353, 2021.
- [18] Adriana Navarro Sainz, Heekyoung Jung, and Annu Sible Prabhakar. Crafting digital experiences for feminine stress-care: An exploratory approach. In *Extended Abstracts of the 2020 CHI Conference on Human Factors in Computing Systems*, pp. 1–8, 2020.
- [19] 坪井康次. ストレスコーピング—自分でできるストレスマネジメント—. *心身健康科学*, Vol. 6, No. 2, pp. 2.1–2.6, 2010.
- [20] Charles S Carver, Michael F Scheier, and Jagdish K Weintraub. Assessing coping strategies: a theoretically based approach. *Journal of personality and social psychology*, Vol. 56, No. 2, p. 267, 1989.
- [21] Eileen Kennedy-Moore and Jeanne C Watson. How and when does emotional expression help? *Review of general psychology*, Vol. 5, No. 3, pp. 187–212, 2001.
- [22] Alaa A Abd-Alrazaq, Mohammad Alajlani, Ali Abdallah Alalwan, Bridgette M Bewick, Peter Gardner, and Mowafa Househ. An overview of the features of chatbots in mental health: A scoping review. *International Journal of Medical Informatics*, Vol. 132, p. 103978, 2019.
- [23] Lorainne Tudor Car, Dhakshenya Ardhithy Dhinakaran, Bhone Myint Kyaw, Tobias Kowatsch, Shafiq Joty, Yin-Leng Theng, and Rifat Atun. Conversational agents in health care: scoping review and conceptual analysis. *Journal of medical Internet research*, Vol. 22, No. 8, p. e17158, 2020.
- [24] Christine Rzepka, Benedikt Berger, and Thomas Hess. Voice assistant vs. chatbot—examining the fit between conversational agents' interaction modalities and information search tasks. *Information Systems Frontiers*, Vol. 24, No. 3, pp. 839–856, 2022.
- [25] Jennifer Zamora. I'm sorry, dave, i'm afraid i can't do that: Chatbot perception and expectations. In *Proceedings of the 5th international conference on human agent interaction*, pp. 253–260, 2017.
- [26] Songhyun Kim, Junseok Goh, and Soojin Jun. The use of voice input to induce human communication with banking chatbots. In *Companion of the 2018 ACM/IEEE International Conference on Human-Robot Interaction*, pp. 151–152, 2018.
- [27] Yuhan Luo. *Promoting Rich and Low-Burden Self-Tracking With Multimodal Data Input*. PhD thesis, 2022.
- [28] Jake Linardon and Matthew Fuller-Tyszkiewicz. Attrition and adherence in smartphone-delivered interventions for mental health problems: A systematic and meta-analytic review. *Journal of consulting and clinical psychology*, Vol. 88, No. 1, p. 1, 2020.
- [29] Mihaly Csikszentmihalyi and Reed Larson. Validity and reliability of the experience-sampling method. In *Flow and the foundations of positive psychology*, pp. 35–54. Springer, 2014.
- [30] Arthur A Stone and Saul Shiffman. Ecological momentary assessment (ema) in behavioral medicine. *Annals of behavioral medicine*, 1994.
- [31] Yong Sook Yang, Gi Wook Ryu, and Mona Choi. Methodological strategies for ecological momentary assessment to evaluate mood and stress in adult patients using mobile phones: systematic review. *JMIR mHealth and uHealth*, Vol. 7, No. 4, p. e11215, 2019.
- [32] 西村良太, 森雷太, 太田健吾, 北岡教英. 音声対話システムのための自由発話に対応した照応解析による入力発話への話題補充手法. *人工知能学会論文誌*, Vol. 37, No. 3, pp. IDS-F.1, 2022.
- [33] Risako Ono, Yuki Nishizeki, and Masahiro Araki. Virtual dialogue agent for supporting a healthy lifestyle of the elderly. In *9th International Workshop on Spoken Dialogue System Technology*, pp. 253–258. Springer, 2019.
- [34] Yusuke Sawada, Akari Tanabe, and Yohei Nakada. A chatbot to search for similar fashion photos and reference fashion coordinators via body part and fashion item segmentations. In *2022 IEEE Symposium Series on Computational Intelligence (SSCI)*, pp. 1476–1484. IEEE, 2022.
- [35] Chizu Mimura and Peter Griffiths. A japanese version of the perceived stress scale: cross-cultural translation and equivalence assessment. *BMC psychiatry*, Vol. 8, No. 1, pp. 1–7, 2008.
- [36] 和田さゆり. 性格特性用語を用いた big five 尺度の作成. *心理学研究*, Vol. 67, No. 1, pp. 61–67, 1996.
- [37] Sandra G Hart and Lowell E Staveland. Development of nasa-tlx (task load index): Results of empirical and theoretical research. In *Advances in psychology*, Vol. 52, pp. 139–183. Elsevier, 1988.
- [38] Naoki Kimura, Tan Gemiciglu, Jonathan Womack, Richard Li, Yuhui Zhao, Abdelkareem Bedri, Zixiong Su, Alex Olwal, Jun Rekimoto, and Thad Starner. Silentspeller: Towards mobile, hands-free, silent speech text entry using electropalatography. In *Proceedings of the 2022 CHI Conference on Human Factors in Computing Systems*, pp. 1–19, 2022.