

無意識な音を意識化させるヒアラブルデバイスの提案

金岡利知^{†1} 長尾正太郎^{†1} 大角耕介^{†1} 山本絵里香^{†1} 宝珠山治^{†1}

概要: 聴覚の知覚・認知処理を補助し認知負荷を軽減することを目指し、“無意識な音を意識化させる”ヒアラブルデバイスを提案する。「ながらスマホ」に代表されるように、複数の認知処理を同時に実施するマルチタスク処理によるトラブルが社会問題化している。人間は常に複数の事柄に集中することはできず、「カクテルパーティ効果」により一方に集中した場合には、集中外の他方の事柄は埋もれてしまう。提案するヒアラブルデバイスは、聴覚の認知処理を必要とするタスクにおいて、聴覚に入力される音を自動的に検知し、蓄積し、必要なタイミングに引き出すことにより、集中外にあった無意識な聞き逃し音を意識化する。本稿では、この提案するヒアラブルデバイスをプロトタイプ化し、日常的に発生しているマルチタスク処理の中から、ビジネスシーンで頻発する「聞き返し」「聞き逃し」を想定した体験デモを被験者 24 名に対して実施し、アンケート・インタビューを通してコンセプトの価値検証を行ったので報告する。

1. はじめに

情報技術の発展により、人は膨大な情報を手にすることができるようになった。もはやインターネットに接続できない世界や、スマートフォンがない日常は考えられないという人も多い。私たちは膨大な情報にアクセスできるようになった一方で、常に情報に触れていたいという欲求は強まり、複数のタスクに注意を分散させるようになってきている。いわゆるマルチタスク処理が日常化している。こうしたマルチタスク処理は、例えば、「ながらスマホ」のように一部では社会問題化している。運転中の「ながらスマホ」は、運転に対する注意がスマートフォンの画面や通話に向かうことで、周囲への注意が散漫になり交通事故を招くことが指摘されている[1][2]。もっと身近なマルチタスク処理の例としては、スマートフォンのゲームや音楽に夢中になるあまり、電車を乗り過ごしてしまったり、病院や役所、銀行などの呼び出し待ちで読書に夢中になり呼び出しのアナウンスを聞き逃してしまったり、講義や会議中にメモを取っているうちに大事なことを聞き逃してしまったり、例を挙げれば切りがないほど多く、人は複数の事柄を同時に処理しようとしている。このように、人間は常に複数の事柄に集中することはできず、一方の集中により集中外の他方の事柄は埋もれてしまう。この現象は「カクテルパーティ効果」[3]としてよく知られている。また、Wickens[4]は、視覚による「見て理解する」と、聴覚による「聴いて理解する」といった同じ知覚資源に集中するマルチタスク処理は、お互いに競合し認知的・精神的負荷が増大することを示している。

情報へのアクセス形態は、近年、スマートフォンだけで

なくワイヤレスイヤホンに代表されるヒアラブルデバイスでも可能である。ワイヤレスイヤホンは音楽鑑賞や会話をするためのデバイスから、スマートフォンと連携することで、音声アシスタント機能[5][6][7]も使えるようになってきている。さらに、声優の朗読を聴ける音声コンテンツ[8]、日常的に人と繋がり会話を共有する音声チャット[9]や音声SNS[10]など、音を中心としたサービスも増加している。このように、視覚だけでなく、聴覚への情報提示量も増加し、人の認知的・精神的負荷、周囲への意識低下が益々増加すると思われる。

増え続ける日常生活の様々なマルチタスク処理の問題に対し、我々は、技術により聴覚に対するタスク処理を補助し認知負荷を軽減することを目指し、“無意識な音を意識化させる”ヒアラブルデバイスを提案する。提案するヒアラブルデバイスでは、聴覚に入力される音を解析、改変することにより、集中外にあった無意識な聞き逃し音から、ユーザに必要な音を検知して意識化させ、音の理解を支援する。

本稿では、まず、「コンセプト」で、解決すべき課題や将来のトレンドを考慮したユーザの体験価値を最大化するヒアラブルデバイスのコンセプトについて述べる。次いで、コンセプトを実現する技術要件について、3章「コンセプトの技術要件」で説明し、4章「プロトタイプ」でプロトタイプ化の詳細について述べる。そして、「コンセプト検証」の章では、日常的に発生しているマルチタスク処理の中から、ビジネスシーンで頻発する「聞き返し」「聞き逃し」を想定し、プロトタイプを用いた体験デモによる被験者検証について述べる。最後に、プロトタイプ化やコンセプト検証を実施した上で得られた理解や気づきについて考察し、

^{†1} 京セラ株式会社 研究開発本部フューチャーデザインラボ
humanaug@gp.kyocera.jp

今後の展望について述べる。

2. コンセプト

マルチタスク処理の増加は、「ながらスマホ」のように社会問題化している。私たちの日常においても、人間は常に複数の事柄に集中できず、選択的集中により処理しているため、集中外の事柄は埋もれてしまう。今後、ヒアラブルデバイスや音を中心としたサービスの普及により周囲に対する意識が低下し、さらに、この問題が顕在化していくと思われる。

我々は、このような日常的に発生しているマルチタスク処理の問題を緩和させることを目的に、提案するヒアラブルデバイスによって、聴覚の認知処理を代替させる。提案するヒアラブルデバイスでは、聴覚の認知処理を必要とするタスクに対し、耳に入ってくる音を自動的に検知し、蓄積し、必要なタイミングで引き出せる。つまり、我々のコンセプトは、このヒアラブルデバイスが耳に入ってくる音を取捨選択し、ユーザにとって必要な音のみ、タイミングよく伝えるメディアータ（媒介者）の役割を担う、いわばヒアラブルデバイスがユーザにとって必要な様々な音や情報をメディアートするロボットとなることを目指している。ユーザにとっては、聴覚の認知処理をヒアラブルデバイスに代替させることで、ユーザは耳に入ってくる音に気を配る必要がなくなり、安心して必要な事柄へ集中できるようになり、心の余裕をもたらされる。

本稿では、この提案するヒアラブルデバイスが、人間の認知処理を補助・拡張することにより、マルチタスク処理の認知負荷を軽減することになり得るのかを研究課題とし、特に、日常的に発生しているマルチタスク処理の中から、ビジネスシーンで頻発する「聞き返し」「聞き逃し」を想定した体験デモを被験者に実施することによりコンセプトの価値を検証する。

3. コンセプトの技術要件

コンセプトで示したユーザにとって必要な音をメディアートするヒアラブルデバイスを実現するためには、人が行っている認知処理と同様に、聞き逃した音を自動的に検知して蓄積し、ユーザに検知したことを通知し、引き出せることが必要である。

3.1 音の検知

音の検知は、自分の名前や番号、案内など重要な言葉や音を検知する。重要な言葉とは、例えば、病院や役所など、読書や音楽の鑑賞に夢中になっている際に、聞き逃してしまうと問題となる言葉である。重要な音とは、家族や上司といった特定の人の声、体調変化に伴う自身の声の変化など、無意識に聞き逃しやすい音である。これらの音の検知は、音声認識や音声識別など既存技術により実現可能である。実用という観点においては、呼び掛けやアナウンスの

音声歪みや周囲雑音に対してロボストに検知する技術が必要となる。次章で示すプロトタイプでは、デモ体験で想定するビジネスシーンに従い、重要な言葉のみ音検知した。

3.2 音の通知

音の通知は、無意識に聞き逃してしまった事象を人に通知し、気付かせ、意識化させる。通知タイミングが悪いと、迷惑・邪魔といった心理を与え、通知を無視することもある。そのため、音の検知の重要性・緊急性を考慮すると共にユーザのコンテキストを理解して、迷惑・邪魔といった心理を与えない工夫が必要である。次章で示すプロトタイプでは、重要な言葉を含む発話の終了タイミングにて通知した。

3.3 音の引き出し

音の引き出しは、ユーザの代わりにヒアラブルデバイスが検知・蓄積していた音を引き出す。音は一瞬にして過ぎ去ってしまうため、音の検知・通知だけでは、人は事象を正確に理解することは難しい。蓄積されていた音を聞き直すことで、発生した事象を理解できる。聞き直しの際には、音の方向や距離といった空間情報、声色や強さなど音の特性を利用するため、人の認知機能を活かせるように、ユーザが耳で聞くのと同様の音を蓄積し、引き出せることが重要である。次章で示すプロトタイプでは、できるだけ耳で聞くのと同様の音を蓄積するため、バイノーラルイヤホンマイクを用いた。

4. プロトタイプ

プロトタイプは、日常的に発生しているマルチタスク処理の中からビジネスシーンで頻発する「聞き返し」「聞き逃し」を対象に、体験デモを被験者に実施することで価値検証することを目的に作成した。そのため、音の検知、通知、引き出しの機能について、デモ体験できる最低限の機能を実装した。

プロトタイプで使用した機器及びアプリケーションサービスを表 1 に示す。音の検知・蓄積を行うための音の入力デバイスには、音の方向や空間情報を耳と同等に収集するため、バイノーラルイヤホンマイクを用いた。音の通知、引き出しのための出力デバイスには、骨伝導ヘッドフォン

表 1 プロトタイプ機器及びアプリケーションサービス

機器/アプリケーション	仕様
バイノーラルイヤホンマイク	Mu6 社製 Scenes LifeLike2
骨伝導ヘッドフォン	AfterShokz 社製 OpenMove
アプリケーションプラットフォーム	Google 社製スマートフォン Pixel 4a
音声認識サービス	Google 社 Google Cloud Speech API

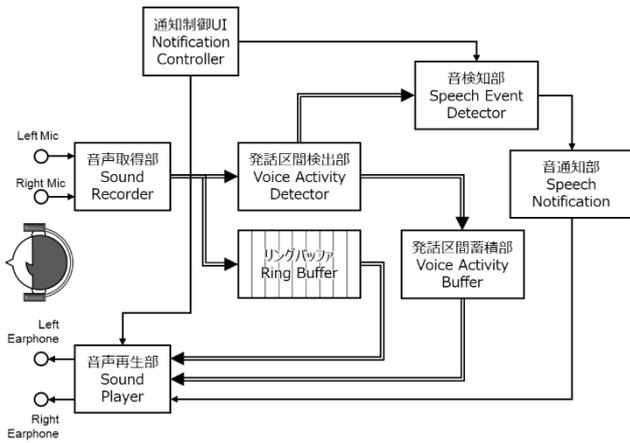


図 1 プロトタイプのアプリケーションブロック図

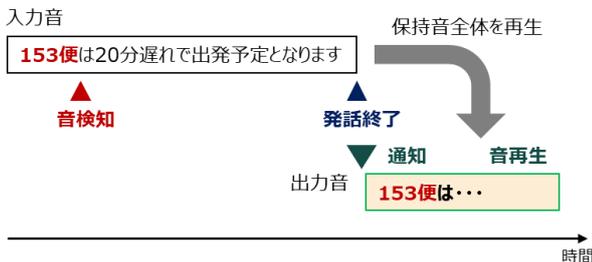


図 2 音検知, 音通知, 音理解の動作イメージ

を用いた。骨伝導ヘッドフォンを用いた理由は、提案するヒアラブルデバイスが人の聴覚の補助を目的としており、耳を塞がないようにするためである。音の検知, 通知, 引き出しの機能を制御するアプリケーションはスマートフォン上に実装した。また、音の検知には、音声認識サービスによる単語検知を用い、重要な言葉の検知のみを実現した。

プロトタイプのアプリケーションブロック図を図 1 に示す。バイノーラルイヤホンマイク (Left Mic/Right Mic) によって集音し、音声取得部においてサンプリングされた音声データは、発話区間検知部により発話部分のみ抽出し、音検知部と発話区間蓄積部に出力される。音検知部は、音声認識サービスを用いて発話音声を変換し、変換されたテキストと、重要な言葉として予め想定した発話検知単語とを部分一致することによりユーザが必要な音を検知する。音通知部は、発話区間終了時に音検知により、発話検知単語と、入力された音声と一致した場合に、ユーザに通知音を出力する。さらに、通知と同時に、発話区間蓄積部から保持されていた発話部分の音声データを、音声再生部を介して骨伝導ヘッドフォンより再生する。また、プロトタイプでは、自動的に音を検知するだけでなく、スマートフォンの通知制御ユーザインタフェース (UI) よりリプレイの指示をすることで、一定時間リングバッファに保持されていた音声を引き出し、聞き直すことができる。

音の検知から音の通知, 音の引き出しの動作イメージを図 2 に示す。価値コンセプト検証のためのビジネスシーンの例として、飛行機の搭乗待ちを想定し、自身が搭乗する

予定の「153 便」を発話検知単語と設定する。図に示すように、音検知は入力される音声データの先頭付近で検知されるが、一連の発話区間を蓄積し、発話区間が終了したときに音通知と共に、蓄積されていた保持音全体を引き出し、再生する。

5. コンセプト検証

本章では、コンセプトの検証方法及び検証結果について説明する。

5.1 検証方法

コンセプトの検証では、プロトタイプしたヒアラブルデバイスを用い、ビジネスシーンで頻発するマルチタスク処理による「聞き返し」「聞き逃し」を想定し、社内従業員 24 名の被験者に対して想定シーンの体験デモ及びアンケート・インタビューを実施した。各被験者は、最初に本研究の背景・課題及びコンセプトについて説明を受け、その後 2 種類のシナリオに沿ってプロトタイプを使ったデモを体験した。一つ目のデモ体験シナリオ「聞き返し」は、講義の聴講中にメモを取るマルチタスクシーンを想定した。被験者には、講義を聴きながら重要な情報・言葉について紙にメモを取り、重要な情報・言葉を聞き逃したと本人が自覚したときに顔を上げるよう指示した。実験者は、被験者に対して講義を行い、被験者が頭を上げたタイミングでスマートフォンの通知制御 UI からリプレイの指示を行った。

二つ目のデモ体験シナリオ「聞き逃し」は、空港の搭乗フロアにおいて、飛行機の搭乗案内アナウンスを待ちながらパソコンを使った仕事をしているマルチタスクシーンを設定した。被験者には、「153 便」の搭乗案内のアナウンスを待ちながら、パソコンを使った仕事の模擬として、画面上に表示されている英文を書き写すよう指示した。実験者は、被験者に対して無関係な搭乗案内アナウンスを行った後、「153 便」の搭乗案内アナウンスを行った。被験者は、自身が搭乗する「153 便」の案内が聞こえると、図 2 に示した動作イメージと同様の動作の体験をした。図 3 にデモ体験のイメージを示す。デモ体験は、実際のシーンとは異



図 3 デモ体験イメージ

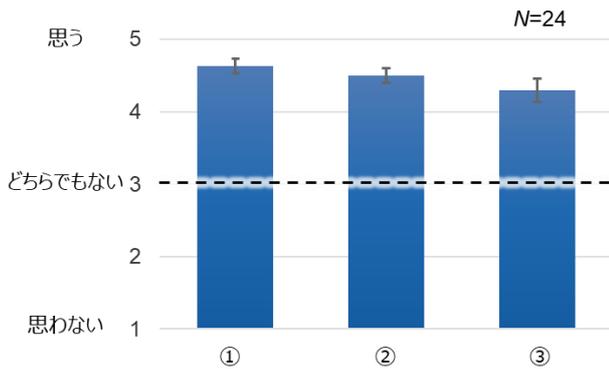


図 4 デモ体験後のアンケート結果

なる事務所の静穏環境下（約 55dBA）において実施した。なお、この環境下においてプロトタイプの音検知性能は、1m 程度の距離から 60dBA 程度の音声入力において、名前の呼び掛けを想定した発話に対して 95%以上の音声検知率であった。

被験者は 2 つのデモを体験した後、以下の 3 つの質問に対して回答した。いずれの質問も 5 段階のリッカート尺度、「1: 思わない」、「2: あまり思わない」、「3: どちらでもない」、「4: やや思う」、「5: 思う」とした。また、各質問項目に対して選択に加えて理由について自由記述を求めた。

- ① あなたの生活の中で「音を聞き返したい」と思うことがあるか？
 - ② あなたの生活の中で「聞き逃した内容を教えて欲しい」と思うことがあるか？
 - ③ 本デモ体験機能は「心の余裕を作り出す」と思うか？
- さらに、実験者は、アンケート終了後にそれぞれのデモ体験について感想・要望を抽出するためインタビューを行った。

5.2 検証結果

被験者から得られたアンケートの結果を図 4 に示す。図より、いずれの回答も、大部分の被験者が「4: やや思う」「5: 思う」の回答を行っており、ウィルコクソンの符号順

表 2 デモ体験後の被験者インタビューコメント

体験デモ 1：聞き返し
<ul style="list-style-type: none"> ・聞き返すことができる安心感が得られる ・緊張状態が和らぐ ・賑やかな場所だとはっきり聞こえないので、聞き返すことも多い ・英語でのミーティングは聞き逃しや間違いをしやすい ・聞き返し時、要約や必要な情報だけでもいいかもしれない
体験デモ2：聞き逃し
<ul style="list-style-type: none"> ・ひとつのことにしか集中できない。空港で焦ることがよくある。 ・聞き逃したかどうか自信が無いときがある ・集中モードを素早く切り替えられるから気が楽になる ・自分に関係があるものかどうか注意しながら聞くのはかなり集中力が必要になる ・話者が重要だと思うことを聞き直したい ・聞き逃す単語の予測ができない

※一部抜粋、コメントの意味が変わらない程度の表現調整あり

位検定により、「3: どちらでもない」に対して有意 ($p = .000$) であることを確認した。このことから、被験者は、デモ体験により筆者らが想定するヒアラブルデバイスを用いたコンセプトに対して強く共感されていることがわかった。

表 2 に被験者のコメントの一部を紹介する。ただし、コメントはアンケートの自由記述及びインタビューから得られた一部を抜粋し、意味が変わらない程度の表現調整を行っている。一つ目のデモ体験シナリオ「聞き返し」については、緊張状態の緩和や聞き返せる安心感が心の余裕に繋がっていることがわかる。また、賑やかな場所では聞き返しが多くなるというコメントや英語でのミーティングは聞き逃しや間違いをしやすいといったコメントから、選択的集中による聞き逃しを補助するというだけでなく、集中して聞いているときにおいても、この聞き返せるという機能がユーザにとって有効であることがわかった。さらに、聞き返し時は途中まで聞いていることもあり、必要な部分だけ聞き返せるようにしたいという欲求もみられた。二つ目のデモ体験シナリオ「聞き逃し」では、デモ体験のシナリオに心当たりを持つ被験者が多く、多くの共感が得られた。そして、複数の事柄に注意を払い続けなければならない事案に対してストレスを感じている被験者が多く、ヒアラブルデバイスが聴覚の認知機能を代替させる機能に期待が見られた。一方で、聞き逃す単語の予測ができない、話者が重要だと思うことを聞き直したい、など検知対象の与え方について現実な課題も見られた。

6. まとめと今後の展望

本研究では、社会問題化しているマルチタスク処理の認知負荷問題に対して、聴覚に入力される音を自動的に検知し、蓄積し、必要なタイミングに引き出すことにより、集中外にあった無意識な聞き逃し音を意識化することで、認知負荷の軽減を目指すヒアラブルデバイスを提案した。そして、この提案するヒアラブルデバイスが、人間の認知処理を補助・拡張することにより、マルチタスク処理の認知負荷を軽減することになり得るのかを研究課題とし、特に、日常的に発生しているマルチタスク処理の中から、ビジネスシーンで頻発する「聞き返し」「聞き逃し」を想定した体験デモを被験者に実施することによりコンセプトの価値を検証した。価値検証の結果、マルチタスク処理問題に対する共感と、人に代わって認知処理をメディエート（代替）するヒアラブルデバイスに対する期待が得られた。

今後の展望としては、コンセプトの技術要件で示したように、ユーザにとって重要な言葉だけでなく、重要な音や、実環境におけるロバスタな音検知、音通知においてユーザに迷惑、邪魔といった感情を抱かせないためのユーザの状態・環境を推定するコンテキスト理解、アンケート・インタビューから得られた音の引き出しにおける要点抽出など、多くの未実装の技術要件が残っている。今後、インタビュ

ーで得られたコメントを基にコンセプト、機能、利用シナリオのブラッシュアップを進め、実用に向けた技術開発・検証を進めていく。また、検証面においては、今回実施したアンケート及びインタビューを中心とした主観的評価だけでなく、開発したヒアラブルデバイスを用いることによる認知負荷の低減効果についての客観的評価を認知実験により明らかにしていく。

参考文献

- [1] “政府広報オンライン:暮らしに役立つ情報:やめよう!運転中の「ながらスマホ」”, <https://www.gov-online.go.jp/useful/article/201707/2.html>, (参照 2021-12-07).
- [2] 篠原一光, 運転中の電子機器使用による注意転導と事故, 交通科学, 2004, Vol.35, 20-26.
- [3] Cherry, E. C., Some experiments on the recognition of speech, with one and with two ears. The Journal of the acoustical society of America, 1953, 25(5), 975-979.
- [4] Wickens, C. D., Multiple Resources and Mental Workload, The Journal of the Human Factors and Ergonomics Society, 2008, 50(3), 449-455.
- [5] “Google:アシスタント”, <https://assistant.google.com/>, (参照 2021-12-07).
- [6] “Amazon:Alexa”, <https://www.amazon.co.jp/meet-alexa/>, (参照 2021-12-07).
- [7] “Apple:Siri”, <https://www.apple.com/jp/siri/>, (参照 2021-12-07).
- [8] “Amazon:オーディオブック Audible”, <https://www.audible.co.jp/>, (参照 2021-12-07).
- [9] “Discord”, <https://discord.com/>, (参照 2021-12-07).
- [10] “clubhouse”, <https://www.clubhouse.com/>, (参照 2021-12-07).