

手のひらでわかるリアルタイム認知負荷推定技術 Real-time Monitoring Method for Cognitive Load Using Electrodermal Activity Sensors

森脇 哲人[†]
Akito Moriwaki

金岡 利知[†]
Toshikazu Kanaoka

浦上 ヤクリン[†]
Jacqueline Urakami

1. はじめに

人間が注意できる対象や集中できる情報量には限界があることは、認知負荷理論[1]の中でも提唱・議論されていたが、近年の研究でその具体的な影響が明らかになっている。本理論において、認知負荷とは、人間の認知機能の構造と限界に着目し、学習や作業の際に作業記憶（ワーキングメモリ）にかかる負荷のことを指す。Sweller, J. (1988)では、ワーキングメモリの容量は限られているため、過剰な情報処理が学習効率を低下させることを指摘している[1]。また、タスクの同時処理や切り替えが人間のパフォーマンスに及ぼす影響について、Charron, S. and Koechlin, E. (2010)の研究[2]によれば、脳が同時に処理できるタスクは2つまでであり、それ以上になると注意をスイッチングさせて対処していると報告されている。さらに、Altmann et al. (2013)の研究[3]では、段階的作業の中断は復帰後のエラー率を上昇させることが明らかとなっている。このように、過剰な情報処理やタスクの割込みがヒューマンエラーを増加させ、作業効率の低下を引き起こすことは、多くの研究で指摘されている。

認知負荷に関する議論は、現代社会においても注目されている。マイクロソフトカナダの調査(2015)[4]によれば、約半数弱の人が仕事や勉強中に注意力や集中力の維持が困難であり、関係のないことを考えてしまう傾向があるという結果が報告されている。その要因として、デジタル情報社会の進展、とりわけスマートフォンに代表される情報機器の普及が、人間のマルチタスクやタスクスイッチングを増加させていることが挙げられる。近年の情報機器は多機能化が進み、多数の機能が同時に動作するため、人は様々な情報を同時に扱い、各機能からの割込みによってタスクの切り替えを余儀なくされている。スマートフォンは日常生活においても、メールやSNSの通知によって絶えず人の注意を引き付けており、無意識のうちに注意が散漫になる状況が生じている。これらの影響は日常生活だけでなく運転中にも表れており、近年では運転中に前方を向いて操作しているにもかかわらず、考えごとやぼんやりしながら運転する「漫然運転」状態が増加している。内閣府の調査によれば、死亡事故における安全運転義務違反の原因の第1位が漫然運転であると報告されている[5]。

このように、認知負荷が現代社会の様々な場面でヒューマンエラーや作業効率低下に関連していることが明らかになっている。しかし、日常生活の中でリアルタイムに自身の脳の認知負荷を客観的に把握する方法は、現時点では確立されていない。認知負荷をリアルタイムに可視化する技術が実現すれば、過度な負荷がかかった際に、適切な対策（休憩アラートやタスク調整など）を講じることが可能となり、エラーや効率低下を未然に防ぐことができる。また、個々人が自身の状態を即座に認識することで、セルフマネ

ジメント能力の向上や、より健全な作業環境の構築にも寄与することが期待される。

このような社会的課題を背景として、本論文では、簡便かつリアルタイムに認知負荷を推定し、可視化する技術の実現を目的として、手のひらで測定可能な新しい認知負荷推定技術を開発したので報告する。

2. 関連研究

日常生活や作業中の認知負荷を測定する方法には、大きく三つの方法が存在する。第一に、主観的データ測定方法が挙げられる。これは、アンケートを用いて個人の主観的な認知負荷の状態を把握する方法である。代表的な方法としてはNASA-TLX（National Aeronautics and Space Administration Task Load Index）[6]が広く知られている。NASA-TLXは、精神的要求、身体的要求、時間的圧迫感、作業達成度、努力、不満足度という六つの尺度項目に対する評価点に基づいて、認知負荷を評価している。主観的データ測定方法は、個人の状態を簡便に把握できるという利点がある一方で、特定のタスク終了後の測定に限られること、リアルタイムのフィードバックが得られないこと、さらには個人の主観に左右されバイアスが生じやすいという課題がある。

第二に、行動学的データ測定方法がある。これは、物体への凝視[7]やペンの操作[8]などの行動指標を基に心的状態を推定する方法であり、近年多くの技術が提案されている。これらの方法は、特定の作業下において認知負荷との関係性を示すことが可能であるが、作業条件が限定されている環境でなければ状態を推定できず、日常生活や様々な作業環境への汎用性に課題が残る。

第三に、生理学的データ測定方法が挙げられる。近年では、脳波（EEG: Electroencephalogram）や近赤外分光法（NIRS: Near-Infrared Spectroscopy）など脳に近い活動を測定する方法[9][10]から、呼吸や心拍などの自律神経指標に基づいて推定する方法[11][12]まで、多様な生理学的測定方法が提案されている。脳波は、認知活動に伴う脳の電氣的信号を計測する方法であり、作業中の注意・集中度の変化や認知負荷の評価に広く用いられている。しかし、脳波の計測には頭皮上へ複数の電極装着が必要であり、装着が煩雑であることや、筋電や環境ノイズの影響を受けやすく、測定環境に制約が生じるなど、実用面での課題が多い。また、NIRSは、脳表面の血流変化を非侵襲的に測定できることから、脳波と同様に認知負荷の評価方法として用いられている。NIRSはユーザへの負担が小さい利点がある一方で、測定の際、頭皮や毛髪による光の透過・反射の影響を受けやすく、正確な脳活動の定量が課題となっている。さらに、呼吸や心拍などの自律神経指標を用いる方法は、装着の簡便さや低コスト性から、日常生活環境下でのストレスや認知負荷の評価に適している。しかし、これらの指標は生理的・環境的要因の影響を受けやすく、また一定時間のデータ蓄積を要するためリアルタイム性に欠けるとい

[†] 京セラ株式会社 Kyocera Corporation

う課題がある。上記以外に、精神的ストレスを測定する方法として古くから皮膚電気活動（EDA: Electrodermal Activity）が用いられている。EDA は、ストレスや緊張など自律神経活動や情動活動を反映する生理指標として知られている。実際、山森ら（2023）の研究[13]では、教師の認知負荷と EDA の上昇との関連が検討されている。一方で、EDA は感情や負荷の種類による反応の特異性が低く、個人差や環境要因の影響も大きいので、認知負荷のみに特化した指標としては十分に活用されていないのが現状である。しかし、脳波や NIRS など他の生理学的指標と比較すると、EDA は測定の手軽さやリアルタイム性に優れており、日常生活や多様な作業環境においても認知負荷の評価を行えるという大きな利点がある。

本研究では、「簡便かつリアルタイムに認知負荷を測定する」方法の実現を研究課題とし、EDA に着目する。特に、従来の EDA の課題である認知負荷に対する反応の特異性や個人差の影響に対しては、EDA と認知処理との関係性を実験により調査し、その結果を基に、これらの課題を解決する認知負荷推定アルゴリズムを開発する。さらに、既存の認知負荷測定方法との比較によりアルゴリズムの妥当性を評価し、最後に本技術を活用したケーススタディを通して、作業中における認知負荷可視化の実用性を検証する。

3. 認知処理と皮膚電気活動との関係性調査

本章では、皮膚電気活動（EDA）の概要について説明し、さらに、EDA を用いることで注意や判断、記憶といった認知処理の状態を推定可能かどうかについて、3 つの実験を通して検証する。まず、EDA が従来注目されてきた情動活動だけでなく、注意という認知処理も反映している可能性を示す。次に、情動活動の有無に関わらず、注意という認知処理の有無が EDA に反映されることを示す。最後に、EDA が注意や判断、記憶といった様々な認知処理を行っている状態を反映していることを示す。以上のように、3 つの実験を段階的に実施することで、認知処理の状態と EDA との関係性を明らかにする。

3.1 皮膚電気活動（EDA）の概要

人間の発汗は、その機能の違いから精神性発汗・温熱性発汗・味覚性発汗の3つに分類される[14]。精神性発汗は、感情や情動、精神的ストレスによって主に手のひら、足裏、脇に生じる。味覚性発汗は、甘味や酸味などの味覚刺激によって主に顔面に生じる。一方、温熱性発汗は体温を一定に保つ働きがあり、全身に生じるが、手のひらや足裏などでは精神性発汗が、顔面では味覚性発汗がそれぞれ優位である。EDA とは、精神性発汗に関連した皮膚表面の電気伝導率の変化を指す。

EDA は、皮膚電気伝導レベル（SCL: Skin Conductance



図1 EDA センサ装着形態

Level）と皮膚電気伝導反応（SCR: Skin Conductance Response）の2つに大別される。SCL は、電気伝導率が比較的長い時間をかけて徐々に変化する反応であり、SCR は、短時間で急速に変化する反応を指す。また、SCL は交感神経の活動を、SCR は刺激に対する応答をそれぞれ反映していると言われている[15]。

本論文の実験では、ウェアラブル生体センサ（biosignalsplux, PLUX 社製[16]）を、図1に示すように左手の人差し指、および中指の第2関節付近に装着し、サンプリング周波数 100Hz、16bit で EDA を測定した。得られた EDA データの解析には、神経生理学的信号処理のための Python ツール “Neurokit2” [17][18]を用いた。

3.2 情動を伴う音刺激と EDA との関係性について

これまでの研究においては、犬の鳴き声やリズムカルな音などの情動を伴う音刺激を、特に注意を向けずに聴取した場合でも、SCR が生じることが示されている[19]。そこで、本節では、音刺激に注意を向けて聴取することにより、注意を向けていない場合と比較して、SCR の回数が増加するという仮説を設定し、注意の有無が SCR に反映されるかどうかを検証する。

3.2.1 実験方法

被験者 10 名（男性 8 名、女性 2 名、年齢 26-57 歳）に対し、音刺激への注意の有無が異なる 3 種類のタスクを実施する被験者内実験を行った。表1は実験で使用した音刺激の概要を示す。これらの音刺激は Gomez, P. and Danuser, B. (2007)を参考に、情動を伴う音として作成した[20]。また、表2は実験タスクの概要を示す。各タスクの実施時間は 180 秒間とし、その間に目的音（Plain tone）を 4 回、それ以外の音を 2 回ずつ、計 12 回の音刺激を呈示した。被験者にはタスク実施時間のみ教示し、目的音の呈示回数は知らせなかった。1 回の音刺激の長さは約 2.5 秒間とし、各音刺激に関連する SCR を区別するため、それぞれの音刺激は 15 秒間隔でスピーカ（Creative T15 Wireless, Creative 社製[21]）から呈示した。各タスクの実施順序および各音刺激の呈示順序は、被験者ごとにランダムに設定した。

3.2.2 結果および考察

音刺激呈示後 3 秒以内に発生した SCR を音刺激関連

表1 音刺激の概要

音刺激名	特徴
Plain tone (Target sound)	1つの連続した音（単純音） 本実験の目的音
Consonant	協和的なピアノの音
Dissonant	不協和なピアノの音
Baby cry	赤ちゃんの泣き声
Tempo-Loudness change	短音と長音が交互に変化 音量も変化

表2 実験タスクの概要

タスク名	タスク内容
Listening タスク	音声を聞くだけで、目的音に 注意は向けない
Counting タスク	目的音に注意を向け、目的音が 呈示された回数を数える
Tapping タスク	目的音に注意を向け、目的音が 呈示されるたびに机を軽く叩く

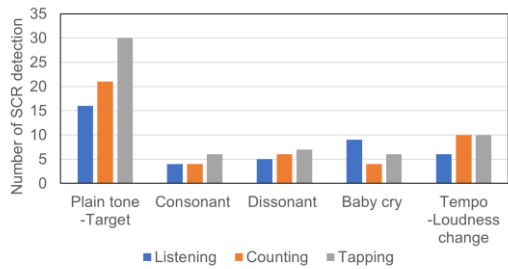


図2 音刺激関連 SCR 検出回数合計

SCR と定義し、それ以外の SCR は自発的 SCR として結果分析には含めない[22].

図2は、各タスクにおける被験者10名分の音刺激関連 SCR の検出回数合計を示す。先行研究[19]と同様に、情動を伴う音刺激に対して SCR が生じることが確認された。さらに、目的音に対する SCR の検出回数は、その音に注意を向けない Listening タスクでは16回であったが、注意を向ける Counting タスクと Tapping タスクではそれぞれ21回、30回と明らかに増加していた。この結果より、音刺激に注意を向けて聴取している場合、より多くの SCR が生じることが示唆される。すなわち、SCR は情動を伴う音刺激に対する反応だけでなく、注意という認知処理も反映している可能性が示された。

3.3 非情動音声と EDA との関係性について

前節の結果から、SCR は情動を伴う音刺激反応だけでなく、注意という認知処理によっても生じる可能性が示唆された。そこで、本節では、非情動音声（情動を伴わない一般的な音声）においても、注意を向けることで SCR が生じることがどうかを検証する。具体的には、非情動音声に注意を向けて聴取した場合には SCR が生じるが、注意を向けない場合には SCR が生じないという仮説を設定し、注意の有無が SCR に反映されるかどうかを明らかにする。

3.3.1 実験方法

被験者6名（男性4名、女性2名、年齢24-30歳）に対し、非情動音声への注意の有無が異なる2種類のタスク（単一記憶タスク、複数記憶タスク）を実施する被験者内実験を行った。実験で利用した非情動音声は、「それでは、〈動物〉は〈色〉の〈数字〉へどうぞ」という構成とし、5種類の動物（イヌ、サル、トラ、ネコ、ブタ）、4種類の色（あお、あか、くろ、しろ）、および8種類の数字（1, 3, 4, 6, 7, 8, 9, 10）を組み合わせで作成した。これらの音声は、古川・小港（2012）の音声コーパスを参考に作成した[23].

単一記憶タスクでは、10個の異なる音声を15秒間隔で再生した（各動物2回ずつ、色と数字はランダムに割り当てた）。被験者には、2回目に再生される「トラ」音声の

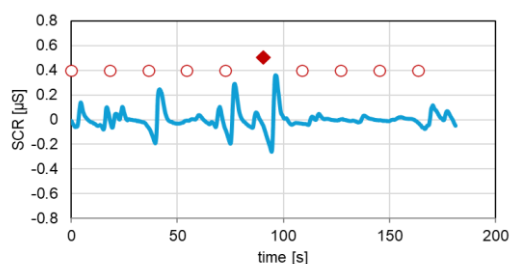


図3 単一記憶タスクにおける SCR 時系列変化

色および数字を記憶し、全音声再生終了後に回答するように指示した。複数記憶タスクでは、10個の音声を15秒間隔で再生した（各動物2回ずつ、2回とも同じ色と数字の組み合わせとした）。被験者には、再生された5種類の動物音声の色および数字を全て記憶し、全音声再生終了後に回答するように指示した。各タスクの実施時間は180秒間とし、音声再生はワイヤレスヘッドホン（ATH-S300BT, オーディオテクニカ社製[24]）を使って行った。また、タスクの実施順序は、被験者の半数で入れ替えて実施した。

3.3.2 結果および考察

図3は、単一記憶タスクにおける被験者1名分の SCR 時系列変化を示す。図中の○は音声再生のタイミングを示し、◆は2回目の「トラ」音声再生のタイミングを示す。2回目の「トラ」音声再生されるまでは、各音声再生直後に SCR が確認されたが、それ以降は SCR が見られなかった。同様の傾向は被験者6名中4名に観察された。これは、2回目の「トラ」音声までは音声に注意を向けて聴取していたが、それ以降は記憶作業に集中し、音声には注意を向けていなかったためと考えられる。実際、タスク終了後、この4名は「2回目の『トラ』音声以降は記憶作業に集中しており、音声には注意を向けていなかった」と述べている。一方、残りの2名は2回目の「トラ」音声以降も SCR が認められた。これは、記憶内容の負荷が比較的低く、記憶作業に集中しなくても記憶できるため、注意が終始音声に向いていた可能性が考えられる。

図4は、複数記憶タスクにおける被験者1名分の SCR 時系列変化を示す。図中の○は音声再生のタイミングを示しており、各音声再生直後に SCR が生じていることが確認された。同様の SCR は、被験者6名中5名で観察された。これは、音声に注意を向けて聴取したことにより SCR が生じたものと考えられる。実際、タスク終了後、全被験者から「動物の色と数字を記憶する必要があったため、音声に注意を向けていた」との報告があった。なお、残り1名については、ノイズのような反応が観察され、EDA の計測不良によるものと考えられる。

以上の結果より、非情動音声に対して注意を向けて聴取した場合のみに SCR が生じ、注意を向けない場合には SCR が認められないことが示唆された。すなわち、情動活動の有無に関わらず、注意という認知処理の有無が SCR に反映されていることが明らかとなった。

3.4 認知処理と EDA との関係性について

本節では、注意や判断、記憶といった認知処理の状態が EDA にどのように反映されるかを検証する。具体的には、音声に注意を向けて聴取し、その音声に対して認知処理を行う際の SCR、および認知処理の負荷レベルによる SCR の強度や SCL の変化を明らかにする。

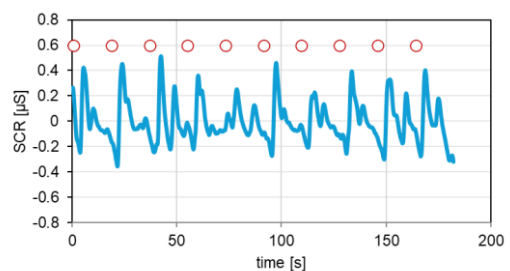


図4 複数記憶タスクにおける SCR 時系列変化

3.4.1 実験方法

被験者 18 名（男性 14 名，女性 4 名，年齢 24-65 歳）に対し，音声への認知処理の内容や負荷レベルが異なる 2 種類のタスク（Tapping タスク，Memory タスク）を実施する被験者内実験を行った．実験音声として，前節と同様の音声コーパスを使用した．

Tapping タスクでは，10 個の異なる音声を 12-15 秒のランダムな間隔で再生した（各動物 2 回ずつ，色と数字はランダムに割り当てた）．被験者には，各音声再生直後に，前のテーブル上に置かれている 5 種類の動物名が記載されたカードを，EDA を測定している左手とは反対の右手で軽くタップするように指示した．この際，各動物の色と数字は記憶する必要はないことも指示した．Memory タスクでは，10 個の音声を 12-15 秒のランダムな間隔で再生した（各動物 2 回ずつ，2 回とも同じ色と数字の組み合わせとした）．被験者には，Tapping タスクと同様にカードのタップと，再生された 5 種類の動物音声の色および数字を全て記憶し，全音声再生終了後に回答するように指示した．各タスクの実施時間は約 150 秒間とし，音声再生は前節と同様のヘッドホンを使って行った．また，タスクの実施順序は，被験者の半数で入れ替えて実施した．

3.4.2 結果および考察

結果の解析は，EDA 測定が不良であった 2 名を除く 16 名のデータを用いて行った．

図 5(a)，(b)は，それぞれ Tapping タスクと Memory タスクにおける被験者 16 名分の SCR 時系列変化を示す．Tapping タスクにおける各音声再生直後の SCR 検出率は 0.91，Memory タスクにおける SCR 検出率は 0.99 であり，いずれのタスクにおいても多くの被験者で音声直後に SCR が認められた．音声への注意や，タップするカードの判断などの認知処理によって SCR が生じたと考えられる．さらに，2 つのタスクにおける SCR の強度（振幅）を比較したところ，有意な差が認められ ($p < .05$)，Memory タス

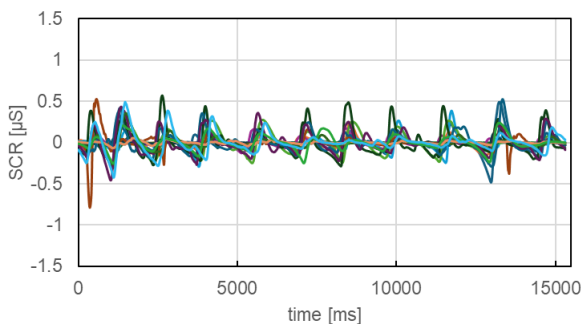


図 5(a) Tapping タスクにおける SCR 時系列変化

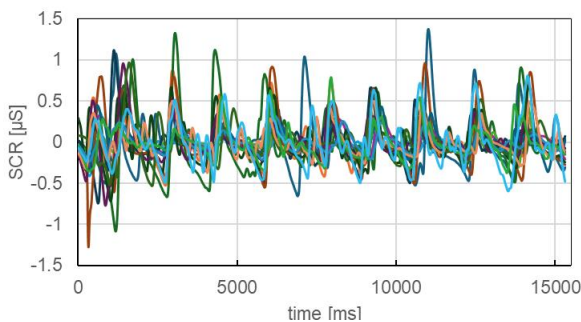


図 5(b) Memory タスクにおける SCR 時系列変化

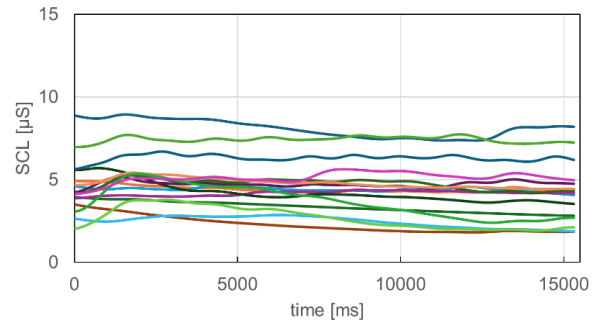


図 6(a) Tapping タスクにおける SCL 時系列変化

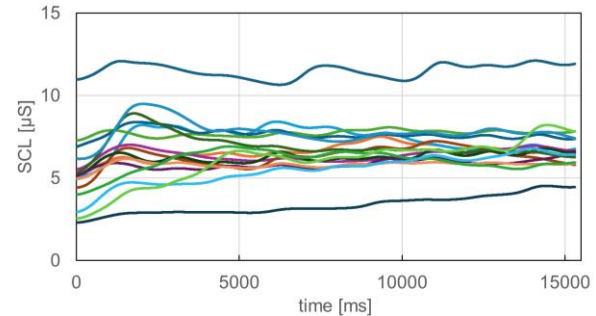


図 6(b) Memory タスクにおける SCL 時系列変化

クの方が SCR の強度が大きかった．Memory タスクでは，カードのタップに加えて音声情報の記憶が求められ，より注意が必要となったため，このような反応の違いが生じたと考えられる．

図 6(a)，(b)は，それぞれ Tapping タスクと Memory タスクにおける被験者 16 名分の SCL 時系列変化を示す．図 6(a)より，Tapping タスクでは SCL がタスク終了にかけて下降傾向を示し，図 6(b)より，Memory タスクでは上昇あるいは均衡傾向を示すことが確認できた．Tapping タスクの SCL について，16 名分それぞれに対して回帰分析および統計検定を行った結果，SCL 時系列データは有意な下降傾向を示した ($p < .01$)．Memory タスクについても同様の方法により，SCL 時系列データが上昇傾向を示した ($p < .10$)．さらに，2 つのタスクにおける傾きには有意な差が認められた ($p < .01$)．Tapping タスクでは，音声聴取後に一時的に情報を記憶し，すぐに記憶内容を解放するため負荷は低いが，Memory タスクでは，各音声情報をタスク終了まで記憶し続ける必要があるため負荷は高い．両タスクにおける SCL の差異は，このような音声情報の記憶に伴う負荷の違いによるものと考えられる．

以上の結果より，音声に対して注意を向けて聴取した場合 SCR が生じ，その音声に対する認知処理の負荷レベルによって SCR 強度や SCL の変位傾向が異なることが明らかとなった．

3.5 認知処理と皮膚電気活動との関係性まとめ

本章では，皮膚電気活動（EDA）を用いて注意や判断，記憶といった認知処理の状態が推定可能かどうかを検証した．3 つの実験の結果，音声に注意を向けた際，SCR が生じることが示され，情報に対する注意反応と SCR との間に強い関係性があることが明らかとなった．さらに，記憶を伴う認知処理の負荷レベルに応じて SCL の変位傾向が異なることから，認知処理の強度と SCL の変位傾向との間に強い関係性があることが明らかとなった．

4. 認知負荷推定アルゴリズム

4.1 認知負荷と皮膚電気活動（EDA）

人の認知処理は、古くから認知心理学の分野においてモデル化されてきた。Atkinson, R.C. および Shiffrin, R. M. により提案された多重記憶モデルと呼ばれる人間の認知処理モデル[25]では、感覚記憶（Sensory Memory）、短期記憶（Short-term Memory）、長期記憶（Long-term Memory）の3つの主要な記憶システムが想定されている。近年では、従来の「短期記憶」に代わり、「作業記憶（Working Memory）」という発展的な概念が用いられるようになり、図7のような認知処理モデルが一般的となっている。これは、短期記憶が単なる情報の一時的な保存にとどまらず、情報の操作や加工、さらに注意の制御といった認知処理を担うシステムであることが Baddeley, A. D. および Hitch, G. J. によって提唱されたためである[26]。図7の認知処理モデルにおいて、感覚記憶は、私たちの周囲で生じている膨大な情報のほとんどをフィルタリングし、一部の情報を追加処理のために作業記憶に渡す働きをする。これは、選択的注意（Attention）と呼ばれ、騒がしい環境下においても聴きたい人の声だけを聞き取ることができる人の能力で、カクテルパーティー効果として良く知られている[27]。作業記憶では、情報を一時的に記憶しながら思考、判断などの認知処理を行うと共に、必要に応じて長期記憶に保存する。また、作業記憶の容量には限界があり、情報は短時間で破棄されるため、リハーサル（Rehearsal）と呼ばれる作業記憶への情報の再入力を行うことで、認知処理の継続が可能となる。さらに、作業記憶への入力には、長期記憶からの思い出し（Retrieval）が利用される場合もある。このようなモデルを基に、Sweller, J. は認知負荷理論を提唱している[1]。本理論では、作業記憶が任意の時点で処理できる情報量や、作業記憶への情報入力、つまり選択的注意、リハーサル、思い出しの量が認知負荷として捉えられる。認知負荷が過剰に増加すると、情報の適切な認知処理が阻害され、ヒューマンエラーや生産効率の低下を引き起こすと考えられる。

3章では、情動的な影響を極力排除した条件下での認知処理と EDA との関係性を調査した。その結果として、人が能動的に情報に注意を向けて取得する際に SCR が検出されることから、SCR は認知処理モデルにおける作業記憶への入力、すなわち選択的注意を反映していると考えられる。また、一時的な記憶を伴う認知処理を行う際には SCL が上昇し、記憶の必要がなくなると SCL が下降することから、SCL は認知処理モデルにおける作業記憶での情報処理を反映していると考えられる。これらの結果と認知心理学における認知処理モデルとの対応付けを仮定して、人の認知負荷を推定、可視化するアルゴリズムを開発した。

4.2 認知負荷推定アルゴリズム

図8は、手のひらから取得した EDA を基に、作業記憶への情報入力レベル、作業記憶での情報処理レベル、そして両方を合わせた認知負荷レベルを算出・推定するアルゴリズムのブロック図を示す。

手のひらから取得した EDA は、0.05-3Hz のバンドパスフィルタ（BPF）を通した後、SCR/SCL 成分分割アルゴリズム（Decomposition algorithm）により4秒のメディアンフ

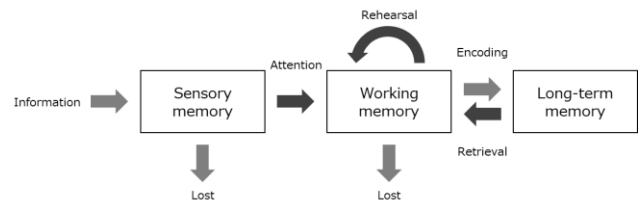


図7 認知処理モデル

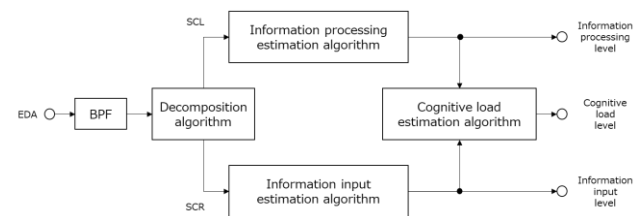


図8 認知負荷推定アルゴリズム

ィルタを用いて、SCL 成分を抽出し、基のデータと SCL との差分を取ることで SCR 成分を抽出する。

抽出した SCL 成分を用いて、情報処理推定アルゴリズム（Information processing estimation algorithm）では、ウィンドウ時間 W にて得られた SCL 成分を基に情報処理レベルを算出する。図9に SCL 成分の時間変化の例を示す。SCL 成分は、時間に対して0に漸近するように減衰する特性を持つため、SCL 成分を常用対数で表す。我々の実験において、SCL は、思考や記憶といった認知処理への負荷が発生する際に、短時間に値が上昇し、負荷がなくなると下降する特性を持つ。そこで、単位時間（ウィンドウ W ）当たりの変位量（ $\Delta \log_{10} \text{SCL}$ ）を情報処理量の指標として、数値化する。SCL 変位量 $\Delta \log_{10} \text{SCL}$ の計算式を（1）式に示す。

$$\Delta \log_{10} \text{SCL}(t) = \frac{1}{W-1} (\log_{10} \text{SCL}(t) - \log_{10} \text{SCL}(t-W+1)) \quad (1)$$

$$\text{if } \log_{10} \text{SCL}(t) < 0 \cdots \log_{10} \text{SCL}(t) = 0$$

W は変位量算出のウィンドウ時間を示し、 t は現在時刻である。情報処理量 P は、（2）式で示すように SCL 変位量 $\Delta \log_{10} \text{SCL}$ を基に、0 から 100 の値で正規化する。

$$P(t) = (\Delta \log_{10} \text{SCL}(t) \cdot W / \alpha + 1.0) / 2 \times 100 \quad (2)$$

このとき、パラメータ α は基準減衰変位量として定義する。基準減衰変位量とは、個人や環境による発汗や蒸発の差を吸収するためのパラメータであり、認知的な情報処理の負荷が掛からないリラックス状態における減衰変位量を基に算出する。具体的には、ウィンドウ W における1秒当たりの減衰量として定義する。本来は、個人や環境において調整した方が良いが、本稿では実験で得られた値を基に標準的な値として $\alpha = 0.1$ としている。

SCR/SCL 成分分割アルゴリズムにより抽出した SCR 成分を用いて、図8の情報入力推定アルゴリズム

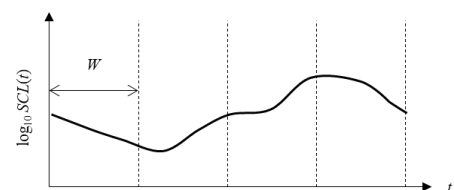


図9 SCL 成分の時間変化

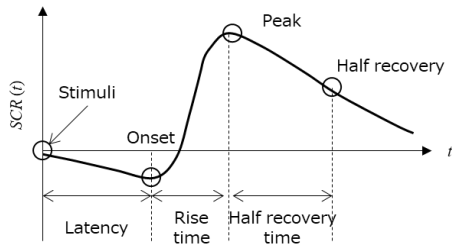


図10 刺激に対するSCR波形

表3 SCRモデルパラメータ

SCRモデルパラメータ	範囲[s]
Rise time	1.0 - 3.0
Half recovery time	2.0 - 10.0
Latency	- 3.0

(Information input estimation algorithm) では、ウィンドウ時間 W にて得られた SCR 成分を基に情報入力レベルを算出する。SCR は、刺激に対して即時的な反応を示す成分であり、刺激入力に対して特徴的な波形を示すことが知られている[22]。図10に刺激に対するSCR波形を示す。図に示すように、刺激(Stimuli)が発生した後、一定の潜伏時間(Latency)を経て、0以下のレベルとなるOnsetが発生する。その後、振幅が急速に上昇(Rise time)することで最大値(Peak)に到達し、その後は緩やかに振幅が低下するという特性を持つ。このとき、振幅がOnsetとPeakの半分の値となる点を半減値(Half recovery)、Peak到達から半減値に至るまでの時間を半減時間(Half recovery time)と呼ぶ。これら各指標は、過去の研究において、概ね表3に示す値の範囲を取ることが知られている[22]。しかしながら、Peak振幅についてはSCLと同様に個人や環境に依存することが、我々の実験で明らかとなっている。また、SCL値と相関がみられることから、ウィンドウ W において、SCR値をSCLの平均値で正規化する。図11に、音刺激を与えたときの正規化SCR値と、何も刺激を与えなかった安静時の正規化SCR値の分布を示す。この分布より、真陽性率と偽陽性率とのバランスが取れる最適点を計算し、正規化SCRのPeakしきい値を0.007とした。ただし、この値は情報入力の感度を示す指標であるため、必要に応じて調整する。このようなSCRモデルを基に、刺激に対するSCRを検出すると、Onsetからバックキャストした刺激入力(Stimuli)からHalf Recoveryまでの区間を情報入力区間 s とし、情報入力レベル I は、以下の(3)式で表す。 $I=0$ は情報入力が全くない場合を、 $I=100$ はウィンドウ内に複数または連続する刺激に対する反応がある場合を示す。

$$I(t) = \frac{1}{W} \sum_{i=1}^{W-1} s(t-i) \times 100 \quad (3)$$

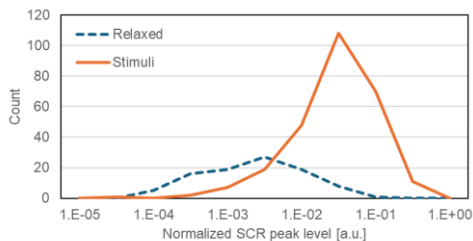


図11 音刺激有無における正規化SCR分布

図8の認知負荷推定アルゴリズム(Cognitive load estimation algorithm)では、ウィンドウ W において算出された情報処理レベル P および情報入力レベル I を基に認知負荷レベルを算出する。完全に負荷が無い状態を $(P, I) = (0, 0)$ 、負荷が極めて大きい状態を $(100, 100)$ と仮定し、認知負荷レベル C は(4)式を用いて算出する。

$$C(t) = \sqrt{(P^2(t) + I^2(t))/2} \quad (4)$$

5. 認知負荷推定アルゴリズムの妥当性検証

本章では、開発した認知負荷推定アルゴリズムによって得られた認知負荷が、実際に測定したい認知負荷を正しく反映しているかどうかを、既存の主観的認知負荷測定方法との比較により、基準関連妥当性の観点から検証する。本アルゴリズムは、人が情報対象に注意を向けた際(すなわち選択的注意が働いた際)の情報入力の検知と、人が思考や記憶などの認知処理を行った際の情報処理の増加を、それぞれ数値化する。さらに、認知負荷理論に基づき、これらの値の合計を認知負荷として定義している。なお、本検証での認知負荷推定アルゴリズムのパラメータは、ウィンドウ W を10[s]、基準減衰変位量 α を0.1とした。

5.1 妥当性検証方法

被験者10名(男性8名、女性2名、年齢25-58歳)に対して、EDAを測定しながら、教示によって音声への認知処理の内容や負荷レベルが異なる3種類のタスクを実施する被験者内実験を行った。各タスクで用いる実験音声としては、3.4節と同様の5種類の動物、4種類の色、8種類の数字を組み合わせた音声コーパスを使用した。

負荷レベルが最も低いタスク(LOW)では、「トラ」音声は何回再生されたかを記憶し、タスク終了後にその回数を回答する。負荷レベルが中程度のタスク(MID)では、「トラ」と4種類の各色が何回再生されたかを記憶し、タスク終了後にそれぞれの回数を回答する。負荷レベルが最も高いタスク(HIGH)では、「トラ」と4種類の各色が何回再生されたかを記憶すると共に、音声に含まれる数字の合計を記憶し、タスク終了後にそれぞれの回数及び合計値を回答する。すべてのタスクにおいて、10個の異なる音声を12秒間隔で再生した(動物、色、数字の組み合わせはすべてランダムに割り当てた)。タスクの流れとしては、最初に3分間の安静状態を取った後、3つのタスクをそれぞれ120秒間実施し、各タスク終了後に2つの主観アンケートを実施する。音声再生は3.4節と同様のワイヤレスヘッドホンを使って行った。また、タスクの順序は、カウンターバランスを取るために被験者間で均等になるように入れ替えている。

主観アンケートには、メンタルワークロードとしてよく知られているNASA-TLX[6]及び精神的努力を測定する尺度であるRSME(Rating Scale Mental Effort)[28]を用いた。NASA-TLXについては、精神的側面だけでなく身体的側面も含めた負荷を測定する指標であるため、表4に示すように本タスクに合わせて日本語訳を行い、各項目、最小・最少を1、最大・最多を7とする7段階評価とした。RSMEについても、「ほぼ努力なし」を1、「極端な努力が必要」を7として、7段階で評価した。

表4 本実験で用いたNASA-TLXの日本語訳

項目	質問内容
1	どの程度の記憶力・読解力を必要としましたか？
2	どの程度の注意力を必要としましたか？
3	どの程度時間的な切迫感がありましたか？
4	どの程度達成できたと思いますか？
5	どの程度一生懸命に取り組む必要がありましたか？
6	どの程度不安感、イライラやストレスを感じましたか？

5.2 妥当性検証結果

結果の解析は、EDA測定が不良であった1名を除く9名分のデータを用いて行った。

図12は、各タスクに対する各主観アンケートの結果を示す。NASA-TLXについては、6指標の平均値と、主に精神的努力を示す1番目の指標も合わせて示す。図に示すように、どちらの主観アンケートもタスク強度に伴って負荷が増加していることがわかる。被験者のタスク従事について、負荷レベルLOWとMIDでは全ての被験者が正答した。一方、負荷レベルHIGHでは8名が一部誤答したものの、いずれの被験者も極端に異なる回答をすることはなく、実験後、被験者全員が「最後まで諦めずにタスクを遂行した」と述べており、完全にタスクを諦めた被験者はいなかった。

図13は、各タスクに対して、開発した認知負荷推定アルゴリズムにより推定した認知負荷レベルを示す。図に示すように、タスク強度に伴って認知負荷レベルが増加していることがわかる。各主観アンケートの値と推定した認知負荷レベルとの相関を表5に示す。表に示すように、精神的な努力・負荷を示すNASA-TLXの1番目の指標（Mental）やRSMEとの間では、相関係数が0.7以上の強い相関を示し、EDAの測定値を基に推定した認知負荷レベルが精神的な努力や負荷を反映した値を示すことが明らかとなった。

6. ケーススタディ

本章では、日常的な事例において、開発した認知負荷推定アルゴリズムがどのような認知処理の状態を可視化でき

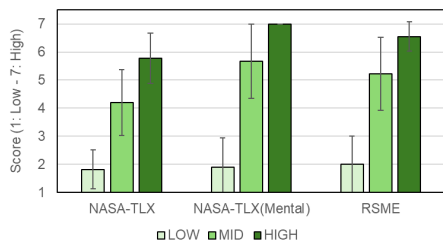


図12 各タスクに対する主観アンケート結果

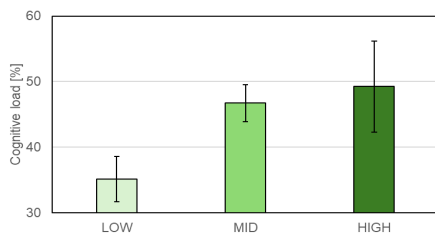


図13 各タスクに対する認知負荷レベル

表5 各主観アンケートの値と認知負荷レベルとの相関

項目	相関係数
NASA-TLX	0.66
NASA-TLX (Mental)	0.79
RSME	0.74

るかを示す。本アルゴリズムは、タスクに対する認知負荷レベルを推定するだけでなく、情報入力の有無や情報処理の負荷レベルを時間推移と共に可視化できる点が特徴である。さらに、5秒程度の遅延でほぼリアルタイムに確認できるため、タスクに対する負荷が低いときの注意喚起や負荷が高いときのタスク調整など、効果的なフィードバックが可能となる。

図14は、1名の日本人被験者を対象に、英語を聞き取る能力を測定するテストを実施した際の情報入力および情報処理レベルの時系列変化を示す。本テストは3つのセクションで構成される。セクション1は、短い質問や発言を聞き、最も自然な応答を選ぶ応答問題である。発言と応答の両方が音声のみで呈示され、紙面の回答選択文はない。セクション2は、2人または3人の会話を聞き、その内容に関する設問を紙面の回答選択文から選択する会話問題である。セクション3は、1人が話すアナウンスやスピーチなどを聴き、関連する設問を紙面の回答選択文から選択する説明文問題である。すべてのセクションにおいて、テストの説明の後に本題を4問実施し、各問題の時間間隔は任意とした。図14の結果から、セクション1では、問題文および回答を常に聴取する必要があるため、情報入力が連続して検出されていることがわかる。一方で、セクション2および3では、回答選択文が呈示されているため、回答選択後は音声を聴取しておらず、聴取の有無が交互になるという特徴的な思考傾向が確認できる。また、情報処理に関しては、音声聴取内容を一時的に作業記憶に保持する際に負荷が上昇し、回答選択後に記憶内容が解放されることで負荷が低下する傾向が見られる。

このように、本指標を用いることで、人がどのタイミングで音に注意し聴取しているか、さらに記憶・思考などの情報処理のタイミングおよび強度を把握することが可能となる。加えて、聞き逃した箇所や、その要因が思考のような情報処理によるものであるかどうかについても推測できる。すなわち、情報入力の有無や情報処理の負荷レベルを時系列で可視化することにより、従来のアンケート調査ではわからなかったタスクの認知負荷要因や注意・思考の特性を把握することが期待できる。

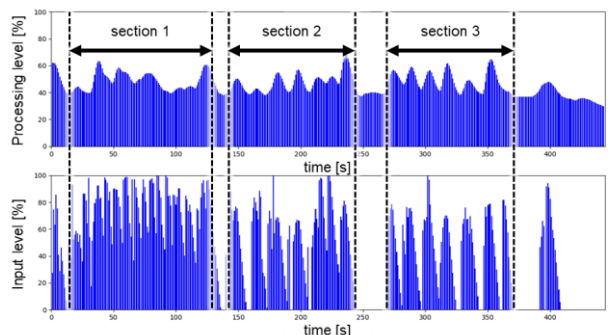


図14 英語聴取テスト実施時の情報入力レベルと情報処理レベル

7. おわりに

本研究では、「簡便かつリアルタイムに認知負荷を測定する」方法の実現を研究課題とし、従来の皮膚電気活動 (EDA) における認知負荷に対する反応の特異性や個人差の影響という課題を解決する認知負荷推定アルゴリズムを開発した。本アルゴリズムでは、音声聴取実験の結果から得られた SCL 成分の変動特性を基に情報処理レベルを、刺激に対する SCR モデルを基に情報入力レベルを推定した。これら 2 つの指標を組み合わせて総合的な認知負荷を推定することにより、EDA 反応の特異性や個人差の課題を解決した認知負荷推定を実現した。さらに、既存の認知負荷測定方法との比較によるアルゴリズム妥当性評価やケーススタディを通して認知負荷可視化の実用性を検証した。その結果、推定した認知負荷が精神的な努力や負荷を反映し、人が注意して聴取しているタイミングや、記憶・思考など情報処理の強度を可視化できることにより、本手法が実用的な認知負荷推定技術として有効であることを示した。今後分析を進めることにより、人の注意傾向や情報処理傾向を把握するだけでなく、脳の効果的な使い方のトレーニングやアシストを行うための指標としても期待される。

本論文では、実験の容易性の観点から、主に音声聴取を対象とした聴覚刺激に対する選択的注意や言語情報処理に基づくモデル構築、妥当性検証を行っており、視覚刺激や触覚刺激、空間情報処理に対する検証は行っていない。また、妥当性検証についても、より多くの被験者数で行った方が、説得性は上がるだろう。さらに、今回は情動的な影響を極力排除した条件下で EDA の評価を行ったが、日常的な状況においては、情動活動と認知処理とが複雑に絡み合うため、「EDA の変化＝認知負荷」とは限らない。様々な刺激での検証や被験者を増やした妥当性評価、情動活動と認知処理との切り分け手法の検討などは、今回の研究における制限であり、今後継続して検討していく必要がある。

謝辞

開発したアルゴリズムの検証実験設計におきまして、ご指導・ご助言を賜りました筑波大学人間系教授 小淵千絵先生に心より感謝申し上げます。

参考文献

- [1] Sweller, J.: Cognitive Load During Problem Solving: Effects on Learning, *Cognitive Science*, Vol.12, No.2, pp.257-285 (1988).
- [2] Charron, S. and Koechlin, E.: Divided Representation of Concurrent Goals in the Human Frontal Lobes, *Science*, Vol.328, No.5976, pp.360-363 (2010).
- [3] Altmann E. M., Trafton J. G. and Hambrick D. Z.: Momentary interruptions can derail the train of thought, *Journal of Experimental Psychology: General*, Vol.143, No.1, pp.215-226 (2013).
- [4] Gausby, A.: Attention Spans, Microsoft Canada, available from: <https://sherpapg.com/wp-content/uploads/2017/12/MAS.pdf> (accessed 2025-06-01).
- [5] 内閣府：令和 6 年交通安全白書（全文），入手先：https://www8.cao.go.jp/koutu/taisaku/r06kou_haku/zenbun/genkyo/h1/h1b1sl_2.html（参照 2025-06-01）。
- [6] Agency for Healthcare Research and Quality: NASA Task Load Index, available from: <https://digital.ahrq.gov/health-it-tools-and-resources/evaluation-resources/workflow-assessment-health-it-toolkit/all-workflow-tools/nasa-task-load-index> (accessed 2025-06-01).
- [7] シルムヨハネス, ペルスキアエルナンデスモニカ, 磯山直也, 内山英昭, 清川清：注視点と視線の交点の深度差を用いた VR 空間におけるユーザの言語負荷の推定, 日本バーチャルリアリティ学会大会論文集, Vol.27, pp.2E1-3 (2022).
- [8] 福林侑也, 永井孝幸：ペンタブレット筆記情報を用いた計算問題遂行中の学習者の認知負荷推定手法の提案, 情報処理学会研究報告, Vol.2022-CLE-36, No.11, pp.1-8 (2022)。
- [9] Keshmiri, S., Sumioka, H., Yamazaki, R. and Ishiguro, H.: Non-parametric Approach to the Overall Estimate of Cognitive Load Using NIRS Time Series, *Frontiers in Human Neuroscience*, Vol.11, No.15, pp.1-14 (2017).
- [10] 鈴木航太, 鈴木達也, 嶋田総太郎, 橘篤導, 小野弓絵：認知負荷を評価するための適切な fNIRS の解析手法の検討, 第 32 回人工知能学会全国大会論文集, Vol.JSAI2018 (2018)。
- [11] 高瀬弘樹：呼吸と認知負荷度の関連, 第 73 回日本心理学会大会発表論文集, pp.712 (2009)。
- [12] 角田啓介, 千葉昭宏, 千明裕, 吉田和広, 渡部智樹, 水野理：心拍変動を用いた認知能力水準の逐次推定, 情報処理学会論文誌, Vol.58, No.2, pp.568-579 (2017)。
- [13] 山森光陽, 長野祐一郎, 徳岡大, 草薙邦広, 大内善広：生理心理学的指標を用いた授業中の教師の認知負荷の把握, 日本教育工学会論文誌, Vol.47, No.1, pp.127-139 (2023)。
- [14] 田村直俊, 中里良彦：味覚性発汗再考, 自律神経, Vol.57, No.4, pp.193-199 (2020)。
- [15] Toomin, M.K. and Toomin, H.: GRS biofeedback in psychotherapy: Some clinical observations., *Psychotherapy: Theory, research and practice*., Vol.12, No.1, pp.33-38 (1975).
- [16] PLUX Biosignals: biosignalsplux, available from: https://www.pluxbiosignals.com/pages/biosignalsplux?srltid=AfmBOor3KSUn5z0ZT_XqgeTkfah7f80o_eLeagW6dlapY5FNsjcJkje (accessed 2025-06-01).
- [17] Makowski, D., Pham, T., Lau, Z.J., Brammer, J.C., Lespinasse, F., Pham, H., Schölzel, C. and Chen SHA: NeuroKit2: A Python toolbox for neurophysiological signal processing, *Behav Res Methods*, Vol.53, No.4, pp.1689-1696 (2021).
- [18] Dominique Makowski and the Team: Neurophysiological Data Analysis with Neurokit2, available from: <https://neuropsychology.github.io/NeuroKit/> (accessed 2025-06-01).
- [19] Greco, A., Valenza, G., Citi, L. and Scilingo, E. P.: Arousal and valence recognition of affective sounds based on electrodermal activity, *IEEE Sensors Journal*, Vol.17, No.3, pp.716-725 (2016).
- [20] Gomez, P. and Danuser, B.: Relationships between musical structure and psychophysiological measures of emotion, *Emotion*, Vol.7, No.2, pp.377-378 (2007).
- [21] クリエイティブメディア株式会社：Creative T15 Wireless, 入手先：<https://jp.creative.com/p/archived-products/creative-t15-wireless>（参照 2025-06-01）。
- [22] Cacioppo, J.T., Tassinary, L.G. and Berntson, G.G. (Eds.): Handbook of psychophysiology, Dawson, M.E., Schell, A.M. and Filion, D.L., *The electrodermal system*, pp.159-181 (2007).
- [23] 古川茂人, 小港省吾：Coordinate Response Measure (CRM) のための日本語版音声・画像コーパスの作成, 信学技報, Vol.112, No.81, pp.387-390 (2012)。
- [24] 株式会社オーディオテクニカ：ワイヤレスヘッドホン ATH-S300BT, 入手先：<https://www.audio-technica.co.jp/product/ATH-S300BT?srltid=AfmBOorX7XlklajPOXQCpmCz6Jvpf2BL2FmBfWLjZydnXCONQGqCXhH>（参照 2025-06-01）。
- [25] Atkinson, R.C. and Shiffrin, R.M.: Human memory: A proposed system and its control processes, *Psychology of learning and motivation*, Vol.2, pp. 89-195 (1968).
- [26] Bower, G. H. (Eds.): Psychology of Learning and Motivation, Baddeley, A. D. and Hitch, G. J., *Working Memory*, Vol.8, pp.47-89, Academic Press (1974).
- [27] Cherry, E.C.: Some Experiments on the Recognition of Speech, with One and with Two Ears, *The Journal of the Acoustical Society of America*, Vol.25, No.5, pp.975-979 (1953).
- [28] 大門樹, 土屋有加：Rating Scale Mental Effort の日本語化に関する研究, シンポジウムモバイル研究論文集, Vol.7(1/2), pp.43-50 (2017)。