

音声による心の健康度測定の妥当性について

篠原 修二[†] 中村 光晃[†] 大宮 康宏^{††} 高木 宏昌^{†††} 牛渡 智^{††††}
光吉 俊二[†] 徳野 慎一[†]

†† PST 株式会社 〒231-0023 横浜市中区山下町 2-905

† 東京大学 〒113-8655 東京都文京区本郷 7-3-1

††† 株式会社トライプロ 〒155-0033 東京都世田谷区代田 6-6-8 旭マンション 201

†††† C I I ソフト開発株式会社 〒335-0023 埼玉県戸田市本町 1-24-18-12

E-mail: †{shinokan99,m-nakamura,mitsuyoshi,tokuno}@m.u-tokyo.ac.jp, ††omiya@medical-pst.com, †††h __
takagi@trypro.co.jp, ††††ushiwatari@ciisoft.co.jp

あらまし これまで我々は音声から心の元気を測定する手法を開発してきた。本研究の目的は、この音声指標を精神健康調査に使用される心理テスト General Health Questionnaire 30 (GHQ30) と実験によって比較し、その妥当性を評価することである。実験は、企業の協力を得て 2015 年 1 月上旬から 4 月下旬にわたり実施された。実験期間中、毎日始業前と終業後に管理者とのインタビュー音声を録音した。GHQ30 は、実験期間の初期と終盤の二度、約 80 日の期間をあけて実施した。音声指標は、GHQ30 実施の前後 1 週間ずつ計 2 週間分の音声データから算出し、GHQ30 実施時の音声指標とした。これら二指標の比較検証の結果、各被験者の GHQ30 得点差と音声指標の差の間には、負の相関が見られた。GHQ30 の項目別に見ると、特に「一般的疾患傾向」「身体的症状」「希死念慮とうつ傾向」と音声指標の間に強い負の相関が見られた。

キーワード 音声, General Health Questionnaire 30, 元気度

Validity of mental health check by voice compared to psychological test (General Health Questionnaire 30)

Shuji SHINOHARA[†], Mitsuteru NAKAMURA[†], Yasuhiro OMIYA^{††}, Hiroaki TAKAGI^{†††}, Akira
USHIWATARI^{††††}, Shunji MITSUYOSHI[†], and Shinichi TOKUNO[†]

†† PST Corporation, Inc. 2 Yamashita-cho, Naka-ku, Yokohama-shi, 231-0023 Japan

† Faculty of Engineering, The University of Tokyo 7-3-1 Hongo, Bunkyo-ku, Tokyo, 113-8655 Japan

††† Trypro Inc. Asahi mansion 201, 6-6-8 Daita, Setagaya-ku, Tokyo, 155-0033 Japan

†††† Cii software Development Co.,Ltd. 1-6-10 Kawagishi, Toda, Saitama, 335-0023 Japan

E-mail: †{shinokan99,m-nakamura,mitsuyoshi,tokuno}@m.u-tokyo.ac.jp, ††omiya@medical-pst.com, †††h __
takagi@trypro.co.jp, ††††ushiwatari@ciisoft.co.jp

Abstract It is well known that people's way of expressing emotion alters when they have a mental disorder. We investigated the relationship between scores on a self-administered psychological test, namely the General Health Questionnaire 30 Items Version (GHQ30), and vitality derived from emotion as recognized by voice. We asked participants to undertake the GHQ30 test twice with an approximately 80 day period in between, and we recorded their voices for one week before and after each test in order to obtain their vitality scores. Regarding the results, the vitality score did not show a correlation with the GHQ30 scores. A comparison of the amount of change between the first and second administration of the test, on the other hand, revealed a correlation between the amount of change in the vitality score and the amount of change in GHQ30 total scores. Especially, it revealed a correlation between the amount of change in the vitality score and the amount of change in the GHQ30 subscales "general disorder trends," "physical conditions," and "suicidal ideation and depression".

Key words voice, General Health Questionnaire 30, vitality score,

1. はじめに

メンタルヘルスの不調による損失は全世界的に無視できないほど大きく、適切な対処が必要とされている課題である [1] [2]. この問題に対処するには、低コストで日常的に利用可能なメンタルヘルスのスクリーニング手法が必要である。

現在、メンタルヘルスを評価する主な手段は、医師等の専門家による問診や General Health Questionnaire 30 items version (GHQ30) [3], Beck Depression Inventory) [4] などの自記式アンケートである。専門家による問診は実施可能な数に制限があり、自記式アンケートにはレポーティングバイアス^(注1)の問題がある。また、唾液や血液などバイオマーカーによる評価も研究されているが [5] [6], コストや被験者の負担に問題が残る。

一方、近年のスマートフォンの普及に伴って、音声データを用いた病態分析が注目されている [7]. スマートフォンを利用した音声分析は、非侵襲である上、専用装置を必要とせず、手軽にかつ遠隔的に行えるという利点がある。

このような観点から、筆者らはこれまで、音声データからストレス状態や抑うつ状態を推定する手法の開発を進めてきた [8] [9] [10]. ストレスは感情に影響を及ぼすことが知られており [11], これらの研究では音声データから直接ストレス状態を分析するのではなく、音声から抽出された感情成分の変遷からストレス状態を推定している。ここで用いられている音声感情認識技術 (Sensibility Technology, 以下 ST) [12] [13] [14] は、多数の人を対象に音声サンプルを集めることにより、個人差のバイアスを解消しており、また発話の言語的内容に依存しないため、広い範囲での応用が期待できる。

本研究の目的は、先行研究で提案されている元気度と呼ばれる音声指標 [9] と GHQ30 の得点を実験によって比較し、音声指標の妥当性を評価することである。音声指標とストレスの関係性を扱った研究としては、人が強いストレスにさらされていると考えられる状況での感情をモニタした研究 [15] がある。この研究では、参加者の状況の違い以外、ストレスやメンタルヘルス状態の評価が明示的に行われていないため、参加者に加わったストレスについては推測に基づいている。また、より大規模な実験 [16] では音声録音と GHQ30 による評価、さらに一部の参加者に対して専門家による面接も実施されている。しかし、大規模な実験であるため、参加者一人あたりの音声録音は 1 回のみにとどまり、音声から認識された感情と GHQ30 の評価や専門家の診断との有意な相関を見出すには至っていない。

本研究では約 80 日の期間において GHQ30 によるメンタルヘルス状態評価を 2 回実施し、継続的に録音された音声から得られる元気度と比較することで、継時変化を考慮した比較を行った。以降、2 章では実験および分析方法について述べる。3 章では結果について述べる。最後に 4 章で本論文をまとめる。

(注1) : 病歴や喫煙歴など特定の情報が、被験者によって選択的に隠されたり表面化させられたりすること。

2. 実験および分析

実験は、企業の協力を得て 2015 年 1 月上旬から 4 月下旬にわたり実施した。実験では、トラックドライバー約 50 名を対象にして、始業直前と終業直後の音声録音および GHQ30 によるメンタルヘルス状態の評価を実施した。これらのデータは、被験者の同意を得て収集した。

2.1 GHQ30 データの取得

GHQ30 による評価は二度実施した。一度目の調査は実験期間の初期に、二度目の調査は一度目の調査の約 80 日後に行った。GHQ30 の質問票を参加者に配布し、各自が記入した後に随時回収を行ったため、回答日には個人間で数日のずれがある。

回収した質問票は直ちに採点した。GHQ30 は

- 一般的疾患傾向
- 身体的症状
- 睡眠障害
- 社会的活動障害
- 不安と気分変動
- 希死念慮とうつ傾向

の 6 項目の下位尺度を含んでおり、合計点とともに各項目に対する点数も個別に算出した。

2.2 音声指標の取得

音声は、業務に関する被験者と管理者との対話を録音した。この録音は、この実験のために作成されたアプリケーションがインストールされたノートブック PC に、USB オーディオインタフェース (X2u, SHURE, Niles, IL, USA) を介して接続されたハンドマイク (PM-240, TOA, Kobe, Japan) を用いて行われた。録音フォーマットは Linear pulse codemodulation で、サンプリング周波数は 11025Hz, 量子化ビット数は 16bit とした。

音声収集後、まず音声データを、呼吸から次の呼吸までを単位とする発話に分割し、ST 解析によって発話ごとに感情指標を算出した。ST で分析される感情は興奮、平静、怒り、喜び、哀しみの 5 種類であり、各成分が音声に含まれる度合いが 10 段階で算出される [12] [13] [14]. 元気度は、これら感情指標の時間的変遷、および各成分のバランスから録音音声毎に [0.0, 1.0] の範囲の値で出力される [9]. 本実験では、GHQ30 実施日の前後 1 週間ずつ計 2 週間分の各音声データに対して元気度算出し、それらの平均値を GHQ30 実施時の音声指標とした。

3. 結果

被験者のうち、音声と GHQ30 両方のデータが得られた被験者は 26 名であった。その内、男性は 20 名、女性は 6 名であった。以下では、この 26 名のみを分析対象とする。

表 1 に、元気度平均値と、GHQ30 の項目別得点および合計点との相関係数を示す。GHQ30 の評価は 2 回行ったため、表では実施回ごとのデータ、および全てのデータでの相関係数を表示している。

表 1 元気度平均値と GHQ30 項目別得点の間の相関係数.

GHQ30 項目	相関係数		
	1 回目 (N = 26)	2 回目 (N = 26)	全体 (N = 52)
一般的疾患傾向	0.056	0.069	0.057
身体的症状	0.051	-0.202	-0.092
睡眠障害	0.199	0.232	0.199
社会的活動障害	-0.044	0.223	0.073
不安と気分変調	0.080	0.026	0.046
希死念慮, うつ傾向	-0.133	0.090	-0.020
合計	0.049	0.098	0.065

1 回目, 2 回目, および全体で元気度平均値との相関係数絶対値が最大であったのは, 全て睡眠障害であり, 相関係数は各々, $r = 0.199$ ($t(24) = 0.995, p = 0.330$), $r = 0.232$ ($t(24) = 1.169, p = 0.254$), $r = 0.199$ ($t(24) = 1.435, p = 0.158$) であった^(注2). このように, 音声指標である元気度と自記式アンケートである GHQ30 得点の間には直接的な相関関係は見られなかった.

次に, 各指標の 1 回目と 2 回目の間の変化量について分析する. 元気度平均値, GHQ30 ともに 2 回目の評価値から 1 回目の評価値を引きその差分を求める. 表 2 に, GHQ30 項目別得点の差分と元気度平均値差分との間の相関係数を示す.

「一般的疾患傾向」「身体的症状」「希死念慮, うつ傾向」「合計」の各得点差分と元気度平均値差分の間に負の相関が見られた. 相関係数は各々, $r = -0.547$ ($t(24) = 3.205, p = 0.004$), $r = -0.367$ ($t(24) = 1.934, p = 0.065$), $r = -0.416$ ($t(24) = 2.239, p = 0.035$), $r = -0.468$ ($t(24) = 2.593, p = 0.016$) であった.

一方他の 3 項目「睡眠障害」「社会的活動障害」「不安と気分変調」については, 有意な相関は見られなかった.

表 2 GHQ30 項目別得点差分と元気度平均値差分との間の相関係数.

表中 * は 10%, ** は 5%, *** は 1% の有意水準を表す.

GHQ30 項目	相関係数 (N = 26)
一般的疾患傾向	-0.547***
身体的症状	-0.367*
睡眠障害	0.016
社会的活動障害	-0.281
不安と気分変調	-0.004
希死念慮, うつ傾向	-0.416**
合計	-0.468**

4. 議論とまとめ

GHQ30 の得点と元気度平均値で表される音声指標について, 元気度平均値そのものと GHQ30 のスコアによる比較では相関が見られなかった. 一方個人内の変化量に着目すると, 負の相関が見られるようになった. メンタルヘルス状態の音声への影響には個人差があると予想されるが, 個人内の変化を見ること

で影響の個人差が相殺され, 個人間で共通する傾向が抽出できるようになったと考えられる. GHQ30 は, 得点が高いほどメンタルヘルス状態が悪いことを表す一方, 元気度は得点が高いほど健康であることを意味する. したがって, 両者の変化量に負の相関が見られたことは妥当な結果と考えられる.

ただし, GHQ30 の項目別に相関を見ると, 全ての項目に関して負の相関が見られたわけではない. 「一般的疾患傾向」「身体的症状」「希死念慮, うつ傾向」の 3 項目については相関が見られたが, 残りの 3 項目「睡眠障害」「社会的活動障害」「不安と気分変調」に関しては相関が見られなかった. 他の母集団についてもこのような傾向が見られるのか, 追加実験を行って検証したい. また, なぜこのような違いが見られるのかについての分析は, 今後の課題としたい.

本研究では, メンタルヘルスを低コストかつ日常的にスクリーニングするための手法を構築するために, 音声から抽出される心の元気度と自記式アンケートで得られるメンタルヘルス状態の関係について調べた. その結果, GHQ30 得点の変化量と音声指標の変化量の間には負の相関が見られた. これらの結果は, 音声から認識された元気度を元に発話者のメンタルヘルス状態を推測可能であることを示している. ただし, 正確な推定を行うためには, ある程度の期間モニタリングを実施し, 音声の個人差を除去するようなキャリブレーションを行う必要がある. この技術の開発についても今後の課題としたい.

本研究では元気度抽出のために音声による感情認識を利用しているため, データの収集にかかるコストは極めて低い. また, 電話をかけている場面などを利用することで, 利用者が特別に意識せずともメンタルヘルス状態をモニタできるようになると考えられ, 日常的なスクリーニングに適した技術になると期待できる.

文 献

- [1] R. C. Kessler, H. S. Akiskal, M. Ames, H. Birnbaum, P. Greenberg, R. M. A. Hirschfeld, R. Jin, K. R. Merikangas, G. E. Simon, P. S. Wang, Prevalence and effects of mood disorders on work performance in a nationally representative sample of U.S. workers, *Am. J. Psychiatry* 163 (2006) 1561-1568.
- [2] World Health Organization, *The Global Burden of Disease: 2004 update*, WHO Press, Geneva, Switzerland (2004), 46-49.
- [3] D. P. Goldberg, *Manual of the General Health Questionnaire*, NFER Publishing, Windsor, England, 1978.
- [4] A. T. Beck, C. H. Ward, M. Mendelson, J. Mock, J. Erbaugh, *An Inventory for Measuring Depression*, *Arch. Gen. Psychiatry*, 4 (1961) 561-571.
- [5] Shuhei Izawaa, Nagisa Sugayab, Kentaro Shirotsub, Kosuke Chris Yamadac, Namiko Ogawab, Yuko Ouchib, Yuichiro Naganod, Katsuhiko Suzukie, Shinobu Nomura, Salivary dehydroepiandrosterone secretion in response to acute psychosocial stress and its correlations with biological and psychological changes, *Biological Psychology* Volume 79, Issue 3, (2008) 294-298.
- [6] Atsuo Sekiyama, Interleukin-18 is involved in Alteration of Hypothalamic-pituitary-adrenal axis activity by Stress, *Society of Biological Psychiatry Annual Meeting.*, San Diego, USA, 2007.
- [7] Shunji Mitsuyoshi, "Emotion recognizing method, sensibility creating method, device, and software," U.S. Patent 7

(注2) : 検定には Microsoft Excel 2010 の検定関数を利用した.

340 393, Sep. 25, 2003.

- [8] Mitsuteru Nakamura, Shuji Shinohara, Shunji Mitsuyoshi, Yasuhiro Omiya, Gentaro Tsumatori and Shinichi Tokuno, PsychoAnalyzer Paper here.
- [9] Shuji Shinohara, Shunji Mitsuyoshi, Mitsuteru Nakamura, Yasuhiro Omiya and Shinichi Tokuno, A mental health assessment method based on emotional level derived from voice, 準備中
- [10] Shuji Shinohara, Shunji Mitsuyoshi, Mitsuteru Nakamura, Yasuhiro Omiya, Gentaro Tsumatori and Shinichi Tokuno, Validity of a voice-based evaluation method for effectiveness of behavioural therapy, MINDCARE2015, Milan, Italy, 2015
- [11] R. S. Lazarus, From psychological stress to the emotions: A history of changing outlooks, *Annu. Rev. Psychol.*, vol. 44, Jan. (1993) 1-21.
- [12] S. Mitsuyoshi, F. Ren, Y. Tanaka, S. Kuroiwa, Non-verbal voice emotion analysis system, *Intl. J. Innovative Comp., Info. and Control* 2 (2006) 819–830.
- [13] S. Mitsuyoshi, Y. Tanaka, F. Ren, K. Shibasaki, M. Kato, T. Murata, T. Minami, and H. Yagura, Emotion voice analysis system connected to the human brain, *IEEE NLP-KE 2007*, Beijing, China, (2007) 479–484.
- [14] S. Mitsuyoshi, Development of Verbal Analysis Pathophysiology, *Econophysics, Sociophysics & Other Multidisciplinary Sciences Journal* 5 (2015) 11–16.
- [15] S. Tokuno, G. Tsumatori, S. Shono, E. Takei, T. Yamamoto, G. Suzuki, S. Mitsuyoshi, and M. Shimura, Usage of emotion recognition in military health care, *DSR 2011*, Singapore, Singapore, (2011) 1-5.
- [16] S. Tokuno, Stress Evaluation by Voice: From Prevention to Treatment in Mental Health Care, *Econophysics, Sociophysics & Other Multidisciplinary Sciences Journal* 5 (2015) 30-35.