

# fMRIを用いた聴覚野からの 音声デコード周波数解像度に関する研究

佐藤 公信<sup>1,a)</sup> 金子 竜也<sup>2</sup> 芝田 京子<sup>2,b)</sup>

**概要:** 現在, ヒト脳から非侵襲により, 聴覚野における情報を再構築する技術は, 開発されていない。本研究は, 非侵襲の fMRI (functional magnetic resonance imaging) により聴覚情報のデコードを目的とする。fMRI を用い周波数に依存する活動部位を測定した先行研究における周波数解像度は, 西洋音楽の音階で表すと 2 オクターブで, 西洋音階を基準とする音声デコードが実現可能か不定である。そこで, デコード技術開発の初段階として, 西洋音で半音毎の脳活動マッピングを行う。一般的な fMRI の空間解像度は, 3mm ボクセルであり, 一つの神経細胞と比較し, 大きい。ピッチが半音異なる聴覚刺激で, 神経細胞と比較し, 大きな空間解像度である fMRI を用いて, 脳活動の優位差が計測可能は疑問である。本論文では, fMRI によって西洋音階で半音ピッチが異なる純音を聴覚刺激として得られる脳活動に有意差が計測可能か検証を行う。C7, C#7, D7 および D#7 を聴覚刺激とした際, 半音差のピッチの各組み合わせで, 脳活動で有意差を得ることができた。2kHz 周辺では, 西洋音階で半音差の周波数解像度で, ヒト脳からのデコード実現の可能性が示唆された。

## 1. はじめに

現在, 脳科学分野において, 脳機能の解析が進んでいる。聴覚野においても, その機能解明が進んでいる [1], [2], [3], [4]。これら先行研究の結果を踏まえ, ヒト脳から情報を抽出するデコード技術に関する研究がなされている。しかし, 聴覚において未だ, 音声デコード技術は実現されていない。

そこで, 本研究は, 非侵襲の fMRI (functional magnetic resonance imaging) により計測されたデータを対象とし, 識別器とする機械学習により, 聴覚野において周波数に依存した活動領域を特定し, 刺激提示された周波数のデコード技術実現を目指す。先行研究では, 一次聴覚野および二次聴覚野において, 特定周波数に反応する神経細胞が存在し [5], 電極により脳活動を測定することで聴覚刺激のピッチを抽出可能である。fMRI を用いた周波数に依存する活動部位測定研究として, Dave R.M. Langers らによる研究が挙げられる。Dave R.M. Langers らの研究 [6] により明らかとなった周波数解像度は, 125Hz, 500Hz, 2kHz, 8kHz と

西洋音楽の音階で表すと 2 オクターブである。先行研究では, fMRI を用い西洋音階を基準とする音声デコードが実現可能か不定であった。一般的な fMRI の空間解像度は, 3mm ボクセルであり, 一つの神経細胞と比較し, 大きい。神経細胞と比較し大きな空間解像度である fMRI を用いて, ピッチが半音異なる聴覚刺激で, 脳活動の優位差が計測可能か未知である。

本論文では, fMRI による聴覚情報デコード技術開発の初段階として, fMRI を用い脳活動を計測する際, 異なるピッチの音から脳活動に有意差を得ることができるか, つまり, デコード時の周波数解像度を明らかにする。

## 2. fMRI

本研究に用いる fMRI について説明する。脳内では, ニューロンの発火により, 酸素消費量は増加し, 脳血流が増加する。発火するニューロン付近で, 磁性体である還元型ヘモグロビンが減少する [7], [8]。この変化を可視化する装置が fMRI である。

fMRI は, 高磁場を発生させる装置ではあるものの, 非侵襲であり, 高磁場による長期的な影響は確認されていない。

## 3. デコード

本研究の最終目標であるデコードとは, 予め脳活動を学習させた識別器, 例えばニューラルネットワークによって,

<sup>1</sup> 高知工業高等専門学校  
Kochi National College of Technology, Monobeetsu 200-1,  
Nangoku, Kochi 783-8508, Japan

<sup>2</sup> 高知工科大学  
Kochi University of Technology, Miyanokuchi185, Kami,  
Kochi 782-8502, Japan

a) satoh@ee.kochi-ct.ac.jp

b) shibata.kyoko@kochi-tech.ac.jp

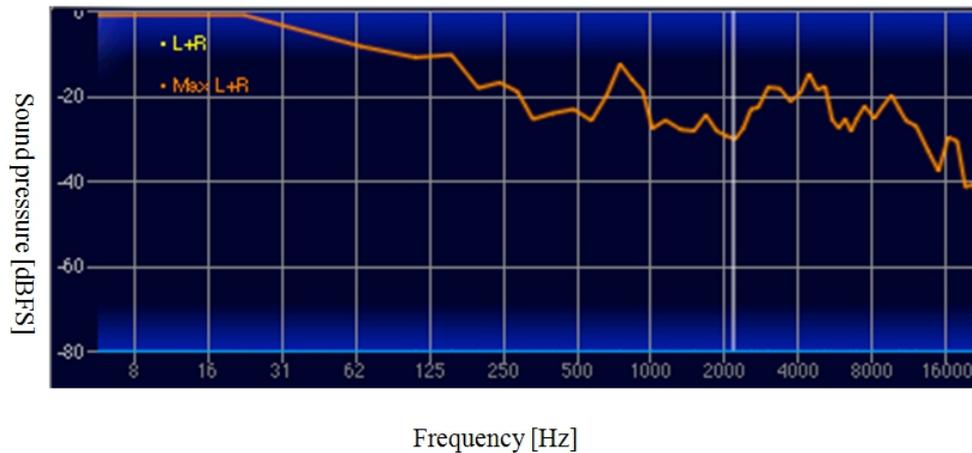


図 1 Frequency analysis of driving sound of fMRI

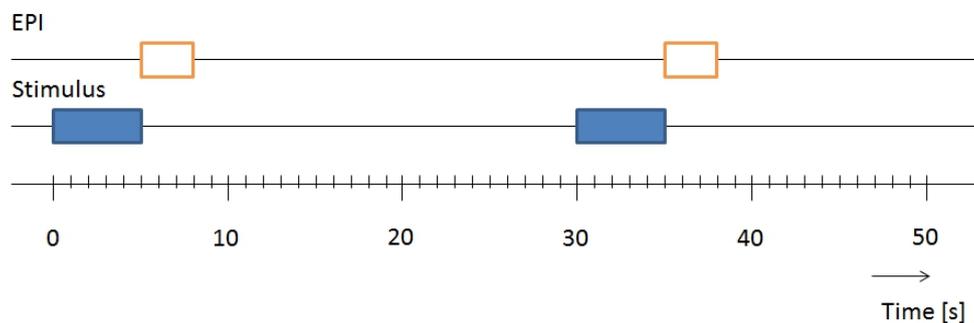


図 2 Double column figure

新たに得られた脳活動を対象としてパターン認識を行い、脳活動から情報を取り出す技術である。脳からのデコード技術は、視覚野において先行している [9]。

機械学習は、あるタスクにより引き起こされる脳活動と、異なるタスクにより引き起こされる脳活動の間には、差が必要である。なぜなら、異なる二つのタスク間に差がなければ、機械学習により、教師あり学習を行っても学習収束することは無く、正しい識別結果を得ることができない。本研究にて提案する音声デコードでは、聴覚刺激の周波数に依存し、活動する脳領域を、識別器とする機械学習により学習する。デコード時は、学習済みの識別器によりパターン認識し、聴覚野から情報を抽出する。

音声デコードの初段階として、本論文では、fMRI の 3mm ボクセル空間解像度において異なる二つの聴覚刺激間に差が観測されるか確認する。差が確認された場合、機械学習を用いた識別器により、正しい脳活動のパターン認識が可能、つまり、音声デコード実現可能であることが示唆される。

#### 4. 実験

本実験では、fMRI により、計測されるデータから西洋音階で表現される半音ピッチ異なる音間で脳活動有意差が観測可能な検証する。

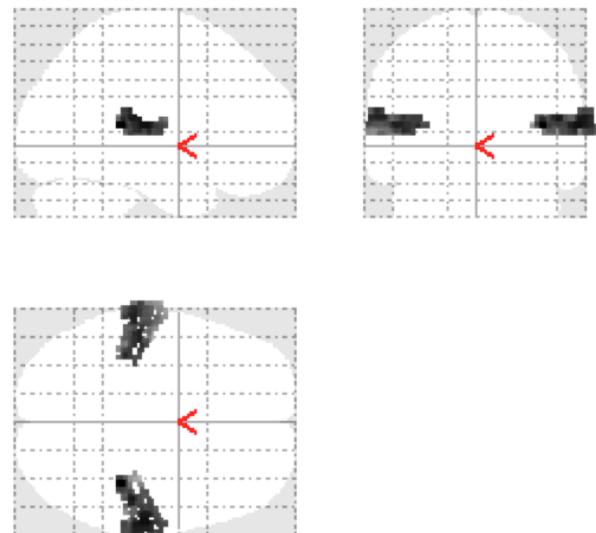


図 3 Brodmann 41-42 area

##### 4.1 被験者

被験者は、20代男性9人とする。各被験者は、オージオメーターによる聴力検査 [10] で異常が確認されていない。

##### 4.2 実験使用機材

fMRI (SIEMENS 社 MAGNETOM Verio3T) であり、磁場強度は、3T である。空間解像度は、3mm とする。ま

た、EPI(echo planer imaging) の設定は、TR 3000 ms, TE 30 ms, flip angle 90°, field of view 192 × 192 mm とする。聴覚刺激提示には、Steinberg 社 Nendo6 を用いる。量子化ビット数は、16bit, サンプリングレートは、44.1KHz とする。また、量子化エラーによる歪みを低減するため、Apogee 社 UV22 プラグインを Nuendo 上で使用する。デジタル・アナログ・コンバータは、RME 社 FireFaceUC を使用し、USB(universal serial bus) によりパーソナルコンピュータと接続する。実験は、磁場の影響を受けない光マイクロフォンを被験者が装着し実施する。

### 4.3 聴覚刺激

fMRI は、スキャン時以外も常時駆動音を発している。ヒトの可聴域である 20Hz-20kHz において、fMRI のスキャン実行時以外の駆動音を解析したところ、2KHz 周辺がもっとも低かった。図 1 は、60 秒間 fMRI の動作音が録音された音声のピークホールドを示す。解析結果を踏まえ、駆動音の影響を最低限とするため、聴覚刺激は C7(2093.00Hz), C#7(2217.46Hz), D7(2349.32Hz) および D#7(2489.02Hz) とする。聴覚刺激は、純音とすとし、提示する刺激の定位は、センターとする。

### 4.4 実験デザイン

実験デザインは、タスクとレストを繰り返すブロックデザインする。タスクとして、5 秒間の刺激提示を行い、EPI スキャンを行う。刺激による脳活動は、刺激提示後 4~6 秒後にピークに到達し、20~30 秒後にベースラインとなる。EPI の動作音による脳活動への影響を最小限に抑えるため、EPI から 22 秒間のレスト後、タスクである聴覚刺激を実施する。22 秒は、本研究に使用する fMRI が設定可能な最大値である。

### 4.5 解析手法

解析は、MathWorks 社 MatLab r2007b 上で SPM8(statistical parametric mapping8) を使用し実施する。ROI(region of interest) は、ブロードマン 41 および 42 とする (図 3)。

解析は、まず、実験中の頭の動きによる雑音除去を目的とした位置補正を行う。さらに、各個人毎に異なる頭の大きさを標準脳に標準化する。ガウスフィルターを適応し雑音除去し、一般線形モデルを用いる。これら前処理がなされたデータに対して、t 検定 [11] による統計的推論を行う。

### 4.6 結果と考察

各聴覚刺激と無音とのグループモデル解析による解析結果を図 4, 図 5, 図 6 および図 7 に示す。t 検定の閾値は、0.1 とする。聴覚刺激に応じた、脳活動が観測されることが分かる。

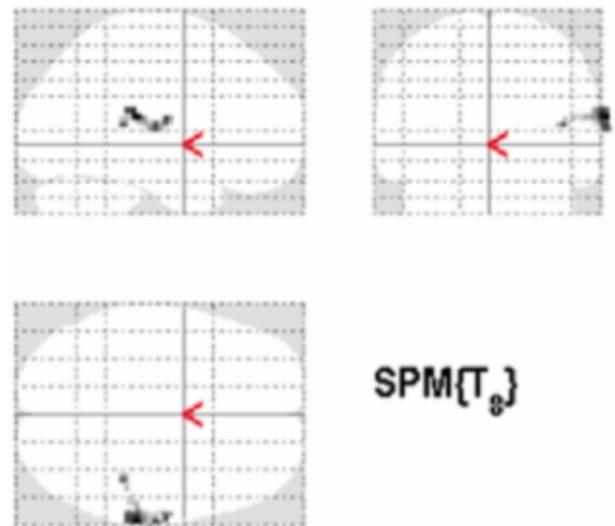


図 4 Group model analysis of C7

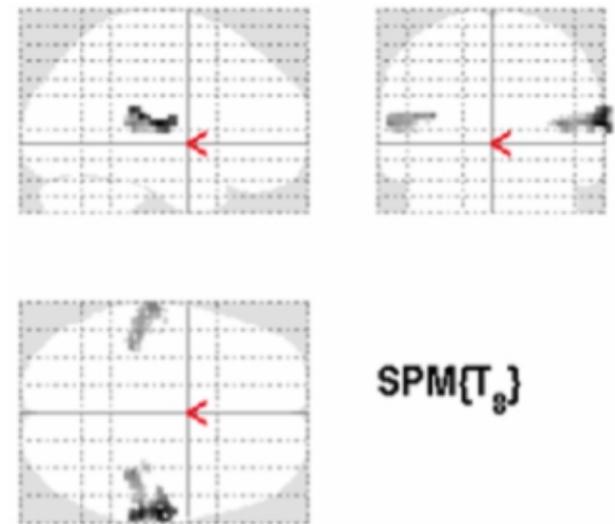


図 5 Group model analysis of C#7

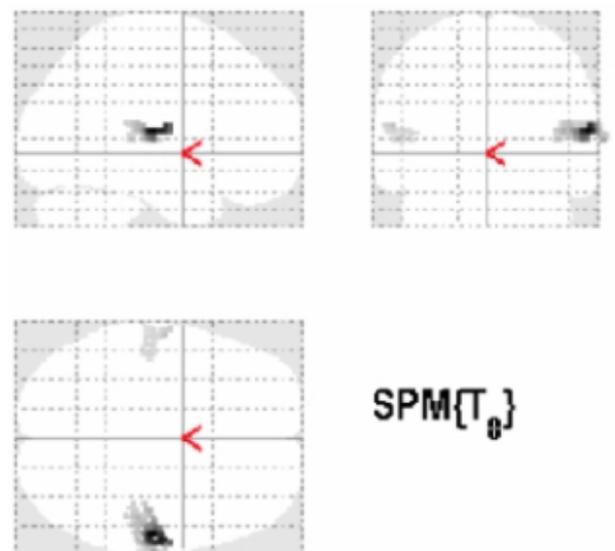


図 6 Group model analysis of D7

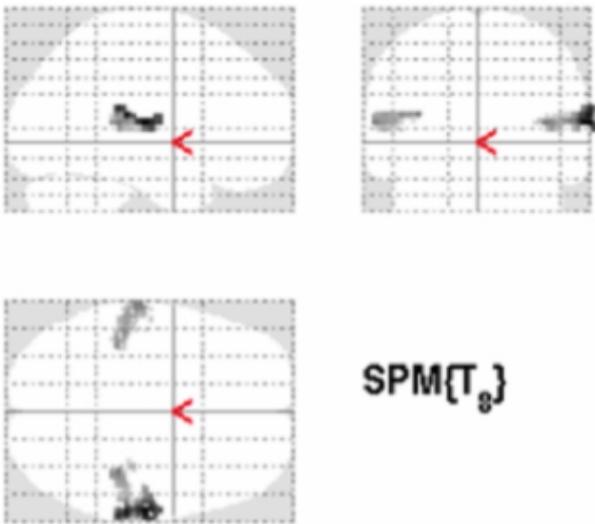


図 7 Group model analysis of D#7

次に、聴覚刺激半音ピッチ差による脳活動有意差についてグループモデル解析を行う。C7とC#7のt検定の閾値は0.08で行う。また、C#7とD7およびD7とD#7の解析は、t検定の閾値として0.05を用いる。結果を図8、図9および図10に示す。

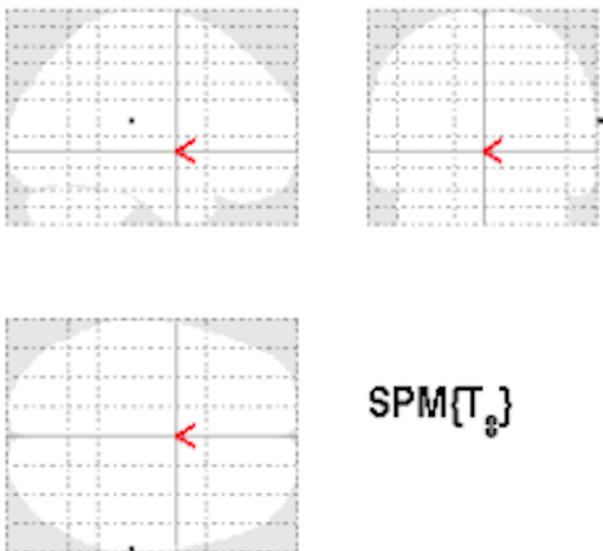


図 8 Significant differences of C7 vs C#7

解析の結果、聴覚刺激半音ピッチの違いにより、脳活動に有意差を得ることができた。半音差の聴覚刺激に対して脳活動に有意差が確認できたことにより、機械学習により、音声デコードの実現可能性が示唆された。

## 5. おわりに

本研究では、非侵襲であるfMRIによる聴覚情報デコード技術実現を目指す。研究の初段階として、本論文では、fMRIを用いた際、脳活動に有意差を得ることができるか、つまり、デコード時の周波数解像度を明らかにする実験を

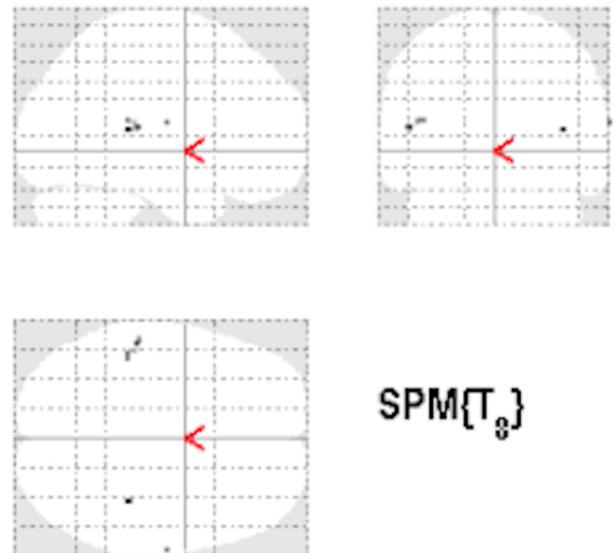


図 9 Significant differences of C#7 vs D7

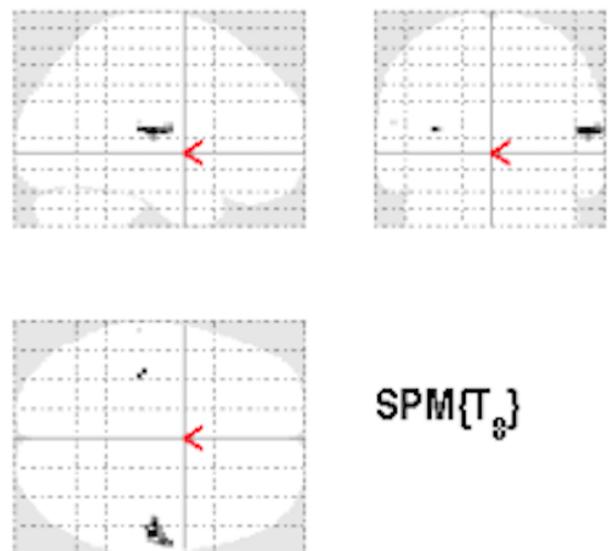


図 10 Significant differences of D7 vs D#7

行った。ヒトの可聴域の周波数の中でもfMRIの駆動音が最も小さい周波数帯を聴覚刺激として選択した。実験の結果、C7、C#7、D7およびD#7の半音差において脳活動に優位差が確認された。

脳活動に有意差が得られることが確認できたことにより、機械学習により、2kHz周辺では、西洋音階で半音差の周波数解像度で、ヒト脳からの音声デコード実現の可能性が示唆された。今後の研究として、機械学習により、脳活動からの聴覚刺激周波数特定を目指す。

## 参考文献

- [1] John C. Strainer, John L. Ulmer, F. Zerrin Yetkin, Victor M. Haughton, David L. Daniels, Steven J. Millen, "Functional MR of the Primary Auditory Cortex: An Analysis of Pure Tone Activation and Tone Discrimination", *Neuroradiol* 18(1997), pp.601-610

- [2] Brodmann, K., "Vergleichende Lokalisationslehre der Grosshirnrinde in ihren Prinzipien dargestellt auf Grund des Zellenbaues", Barth, Leipzig(1909)
- [3] Rademacher, J., Morosan, P., Schormann, T., Schleicher, A., Werner, C., Freund, H.J., Zilles, K., "Probabilistic mapping and volume measurement of human primary auditory cortex" *NeuroImage* 13(2001), pp.669-683
- [4] Belin, P., Zatorre, R.J., Hoge, R., Evans, A.C., Pike, B., "Event-related fMRI of the auditory cortex", *NeuroImage* 10(1999), pp.417-429.
- [5] Shamma, S.A., "Topographic organization is essential for pitch perception," *Proceedings of the National Academy of Science* 101, pp.1114-1115.
- [6] Dave R.M. Langers, Qalter H. Backes, Pim van Dijk, "Representation of lateralization and tonotopy in primary versus secondary human auditory cortex", *NeuroImage*, Vol. 34, (2007), pp. 264-273.
- [7] Bilecen, D., Scheffler, K., Schmid, N., Tschopp, K., Seelig, J., "Tonotopic organization of the human auditory cortex as detected by BOLD-fMRI", *Hear. Res.* 126(1998), pp.19-27.
- [8] Fujioka, T., Ross, B., Okamoto, H., Takeshima, Y., Kakigi, R., Pantev, C., "Tonotopic representation of missing fundamental complex sounds in the human auditory cortex" *Eur. J. Neurosci.* 18 (2003), pp.432-440.
- [9] Y.Miyawaki, H.Uchida, O. Yamashita, M. Sato, Y. Morito, H.Tanabe, N.Sadato, Y. Kamitani, "Visual Image Reconstruction from Human Brain Activity using a Combination Multiscale Local Image Decoders" *Volume 60, Issue 5, Neuron* (2008), pp.915-929
- [10] Katz, J., "Handbook of Clinical Audiology", Lippincott Williams and Wilkins, Philadelphia (2001)
- [11] Friston, K.J., Holmes, A.P., Price, C.J., Buchel, C., Worsley, K.J., "Multisubject fMRI studies and conjunction analyses", *NeuroImage* 10(1999), pp.385-396.