

## 音声明瞭度に関連した大脳皮質活動の時空間的遷移\*

☆嵯峨直樹, 矢野肇 (神戸大/産総研), 滝口哲也, 有木康雄 (神戸大),  
添田喜治 (産総研), 中川誠司 (千葉大/産総研)

### 1 はじめに

近年, 脳信号を機械制御信号とする, いわゆるブレイン・コンピュータ・インターフェイス (Brain Computer Interface : BCI) の開発の試みが盛んに行われている。しかし, その多くは P300 スペラーのように予め用意されている選択肢を識別するものである。従って, ヒトの脳活動から認知・知覚状態や意志, 情動そのものを解読できるようになることが望ましい。本研究では, より汎用的な BCI の開発に有用な知見を得ることを目的として, 音声の知覚状態を反映する脳活動の同定を試みた。

音を聴取した際に, 聴覚野で観測される代表的な反応として N1/N1m 反応が挙げられる。N1/N1m 反応は, ある程度認知過程の影響を受けるが, それ以上に刺激の物理特性によって左右される。また, 刺激のオンセットからおおよそ 100 ms 後に出現することもあり, 言語処理などの高次処理を色濃く反映しているとは言い難い<sup>[1]</sup>。

一方, 音声を聴取した際の言語情報の処理は, 聴覚野で基本的な音響特徴の解析がなされた後で行われると考えられる。言語情報処理に携わる脳部位の代表的な例としては, 聴覚野の後方に位置するウェルニッケ野 (知覚性言語野) や角回, 縁上回が広く知られている<sup>[2]</sup>。また, 近年では音声知覚における, 音声の生成系領野の関与が注目されている。例えば, 運動前野において, ある音素の調音に使われる領野がその音素を知覚する際にも活動するという報告<sup>[3]</sup>や, 発話 (運動性言語処理) に関連する部位とされてきたブローカ野が, 音声知覚にも強く関与するという報告もなされている<sup>[4]</sup>。Wu *et al.* は音声知覚に関する脳内処理の流れとして, 音韻に関する情報を処理する聴覚野から上側頭回後部, 角回, 縁上回, 運動前野へと至る“背側路”と, 意味情報を処理する聴覚野から下前頭回へと至る“腹側路”を提案している<sup>[7]</sup>。

雑音環境下における音声知覚における脳活動に関しては, 明瞭度が低下する条件において, 左右の上側頭回の活動が増加すると報告<sup>[5]</sup>もあるが, この実験は極端な 2 条件の変化を比べたものであり, 明瞭度の段階的な変化による脳活動の空間的特性やその反応潜時については未だ明らかにされていない。

前稿では脳磁界計測を用いて, 知覚する音声の明

瞭度に関連した活動源の推定を行った<sup>[8]</sup>。SN 比を変化させた日本語単音節の刺激オンセットから 1000 ms の間でブローカ野やウェルニッケ野に活動源が推定されたが, SN 比の変化, すなわち明瞭性の変化と活動源強度の有意な差は見出せなかった。本稿では, 単音節の知覚に伴う, 複数の言語知覚関連領野の活動の強さの時間的変化を推定し, 明瞭度の変化に伴う脳内活動源の時間的・空間的変化を調べた。

### 2 刺激音

刺激音には親密度別単語理解度試験用音声データセット (FW03, NTT-AT) に収録された, 日本人女性話者一名 (fto) による日本語単音節 100 個を利用した。また, 長谷らの報告<sup>[9]</sup>に基づき, 単音節ごとのラウドネスが等しくなるように振幅を調整した。次に, これらの音源の SN 比が特定の値になるように白色雑音を付加した。SN 比は -10, -5, 0, 7, 15 dB の 5 段階に設定したが, これらの値は予備実験の結果に基づいて明瞭度がおおよそ 25, 40, 55, 70, 85 % となるように決定された。音声と白色雑音を加算した刺激音の長さは 0.4716 s とした。刺激音の最初の 25 ms 及び, 音声呈示が終了してから刺激音の終端までの区間は白色雑音のみが呈示される。なお, 刺激音量は各被験者について“音がはっきりと聞こえ, なおかつちょうど良いと感じられる大きさ”に調整した。

### 3 明瞭度試験

#### 3.1 方法

聴覚健常者 7 名 (男性 4 名, 女性 3 名, 20-39 歳, 右利き) を被験者とした。実験は防音室内で行われ, 刺激音は挿入型イヤホンを用いて被験者の両耳に呈示された。刺激音の呈示間隔は 5 s で, 被験者にはこの間に聞こえた通りに回答用紙に記入するように教示した。まず, 練習セッションとして, ノイズなしの 100 音節の聴き取りを行わせた。その後, 各 SN 比の刺激音での聴取試験をランダムな順序で行った。

#### 3.2 結果

明瞭度試験の結果を Fig. 1 に示す。SN 比の低下に伴って, 明瞭度も低下する傾向がすべての被験者で観察できた。しかしながら, 音節ごとに SN 比と明瞭

\*Spatiotemporal characteristics of cortical activities associated with articulation of speech perception. by SAGA, Naoki, YANO, Hajime(Kobe Univ/AIST), TAKIGUCHI, Tetsuya, ARIKI, Yasuo(Kobe Univ), SOETA, Yoshiharu(AIST), NAKAGAWA, Seiji(Chiba Univ/AIST)

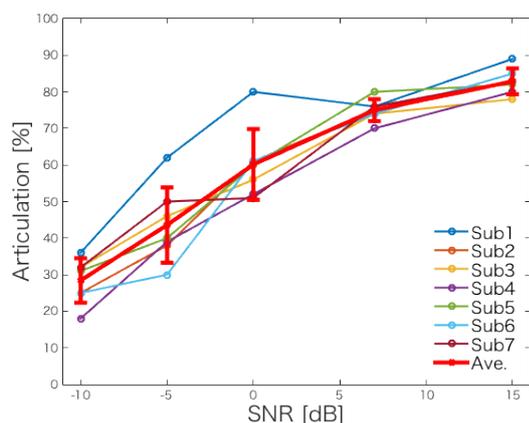


Fig. 1: The relationship between SNR and articulation.

度の相関を見ると、音節によってはSN比の大小に関わらず、明瞭度が常に低い音節もあった。

## 4 脳磁界計測

### 4.1 方法

明瞭度試験で用いた単音節のうち、SN比と明瞭度との相関係数が最も大きかった音節(/ぺ/, /きょ/, /しゅ/)に対する脳磁界を計測した。被験者は明瞭度試験と同じである。脳磁界計測は磁気シールドルーム内で122ch全頭型脳磁界計測システム(Neuromag-122™, Neuromag Ltd, Finland)を用いて行った。刺激音は明瞭度試験の時と同様に、挿入型イヤホンによって被験者の両耳に呈示した。刺激音の呈示間隔は2sとした。明瞭度試験と同様に、まず練習セッションとしてノイズなしの100音節の聴き取りを行わせた。その後、明瞭度試験で決定した高相関音節を、SN比を-10, -5, 0, 7, 15 dBに調整したうえでランダムな順序で呈示した。なお、被験者に聴き取りへの集中を維持させるために、一定頻度で挿入したターゲット音(/あ/, /い/, /う/)を聴取した際にはボタン押しにて応答するように教示した。

脳磁界データは0.03 - 100 Hzのアナログフィルタを通した後、サンプリング周波数400 HzでA/D変換をした。各SN比条件において、音節ごとの誘発反応を70回以上加算し、加算平均から得られたデータに2 - 30 Hzのデジタル帯域通過フィルタを適用した。

### 4.2 解析

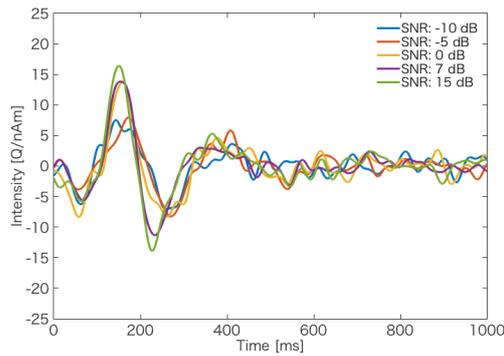
まず、音声知覚に関与するとされる、(1)聴覚野、(2)下前頭回の弁蓋部と三角部(ブローカ野)、(3)上側頭回後方(ウェルニッケ野)、角回、縁上回、(4)運動野/運動前野において、それぞれ単一電流双極子モデルも用いた脳内電源推定を行った。チャンネル選択にあたっては、それぞれの領野が覆われていると思われる(1)18個、(2)14個、(3)16個、(4)16個の

チャンネルを用いた。これらチャンネルのデータを用いて、それぞれの領野で刺激オンセットから1000 msまで2.5 msごとに推定を行った。推定結果の選別にあたり、Goodness of fitとConfidence volumeの二つの指標を用いた。Goodness of fitは観測値と推定値の適合性を、Confidence volumeは、活動源位置における95%三次元信頼領域を示している。Confidence volume  $\leq 10000 \text{ mm}^3$ の制約を満たす活動源の中からGoodness of fitの値が最も優れたものを代表の活動源とし、位置と向きを固定し、活動源強度の振幅のみを推定するAmplitude fittingを行った<sup>[10]</sup>。活動源強度の時間変化が得られた後、平均値と標準誤差範囲よりも大きなピーク値とその潜時について、各領野におけるSN比、明瞭度との関連性を調べた。

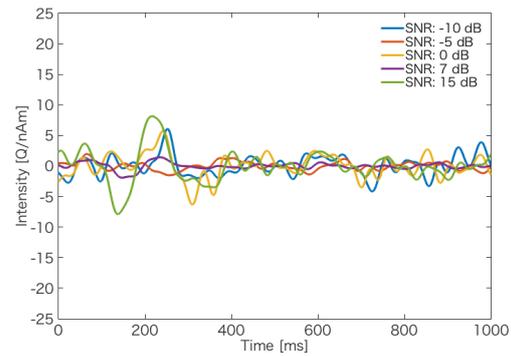
### 4.3 結果

ある被験者一名における(1)~(4)のそれぞれの領野における活動源強度の時間的変化をFig. 2に示す。また、活動源強度のピーク値とその潜時についてSN比との関係をFig. 3に示す。統計的に差があるかを調べるため、領野とSN比を主効果とする二元配置分散分析を行った。左半球における活動源強度のピーク値について、領野を主効果とする有意差( $p \leq 0.05$ )は、聴覚野とブローカ野、聴覚野と運動/運動前野であり、どちらも聴覚野における活動の方が大きかった。SN比を主効果とする有意差は確認できなかった(Fig. 3(a))。また、潜時について、領野を主効果とする有意差は、聴覚野とブローカ野、聴覚野とウェルニッケ野、聴覚野と運動/運動前野で確認でき、聴覚野は他の領野よりも活動のピーク早かった。SN比を主効果とする有意差は確認できなかった(Fig. 3(b))。右半球においては、活動源強度のピーク値について、領野を主効果とする有意差は、聴覚野と下前頭回、聴覚野と運動/運動前野、下前頭回と運動/運動前野、上側頭回後方と運動/運動前野で確認できた。こちらもSN比を主効果とする有意差は確認できなかった(Fig. 3(c))。潜時について、領野を主効果とする有意差は、聴覚野と上側頭回後方、聴覚野と運動/運動前野で確認でき、聴覚野と下前頭回で有意傾向( $p = 0.054$ )が見られた。SN比を主効果とする有意差は、上側頭回後方の領野と他の3つの領野の間で確認された(Fig. 3(d))。

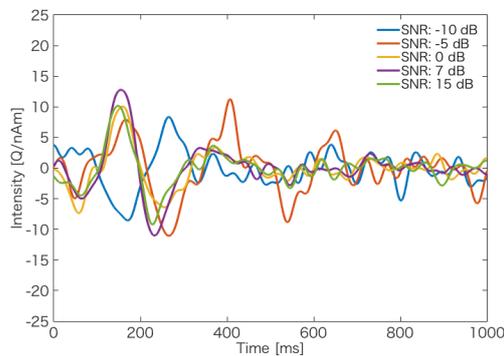
さらに、領野ごとにSN比を主効果とする一元配置分散分析を行ったところ、右半球の上側頭回後方では活動源のピーク値の潜時において、SN比の有意差が存在した。この反応潜時と明瞭度とにおける関係をFig. 4に示す。



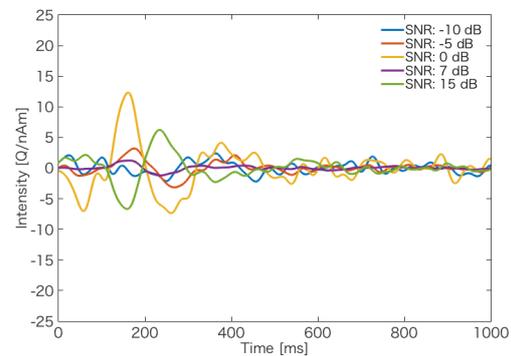
(a) The intensity of source localization in/around the auditory cortex in the left hemisphere.



(b) The intensity of source localization in/around Broca's area in the left hemisphere.



(c) The intensity of source localization in/around Wernicke's area in the left hemisphere.



(d) The intensity of source localization in/around the motor and the premotor cortex in the left hemisphere.

Fig. 2: Time series of the source intensity.

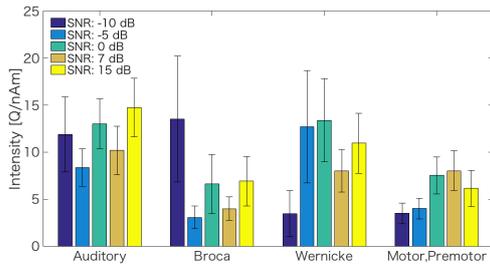
## 5 考察

聴覚野においては、潜時 100 ms~200 ms の間に大きなピークを持つ反応が見られた。聴覚野において潜時 100 ms で見られる成分は、SN 比の違いによる振幅の差異が見られないことから、N1/N1m 反応を捉えていると思われる。ウェルニッケ野における活動源強度は、聴覚野における反応と大きな差はないものの、潜時は聴覚野よりも遅い。このことは、音声知覚の処理が、聴覚野からまずウェルニッケ野へと移行していると判断できる。ブローカ野や運動/運動前野は聴覚野に比べピーク値の潜時は遅いが、ウェルニッケ野から活動が移行しているとは言い難い。右半球でも、左半球と同様に聴覚野から他の領域への活動の遷移が考えられるが、それら領域間における活動の遷移の傾向は見られない。運動/運動前野は他の領域のよりも活動が弱く、音声知覚、理解へ関与は比較的弱いと考えられる。SN 比、ひいては音声の明瞭性によるピーク値の変化は確認できなかったが、潜時については右半球の上側頭回後方において、SN 比の変化に伴う有意な変化が確認できた。音声の明瞭度と

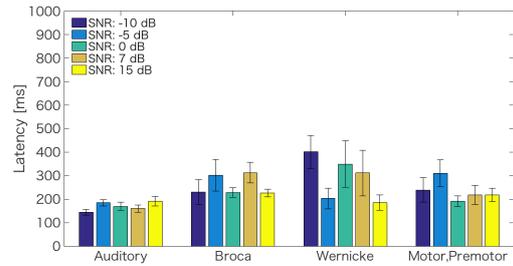
その活動潜時の変化の仕方を見ると、SN 比が-5 dB のときに最も遅く、それ以降 SN 比が向上すると潜時は早くなっている傾向が確認できる (Fig. 4)。このことから、雑音がある程度大きく、音声を同定しようとするために負荷がかかる時に、反応が遅れる可能性が考えられる。この時、SN 比が-10 dB の条件の時に反応潜時が早まっているが、これは雑音が大き過ぎることにより、ほとんど聴き取ることができず、音声を同定するための負荷が逆に少なくなってしまうと考えられる。

## 6 まとめ

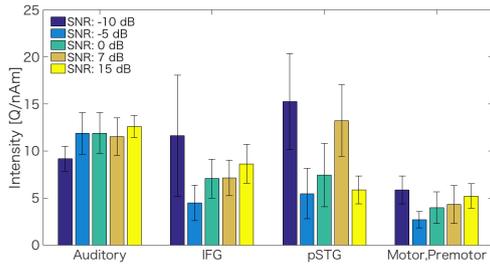
音声聴取時には、聴覚野から音声知覚に関連のあるとされる、上側頭回後方のウェルニッケ野や運動/運動前野、下側頭回などの領域に活動が遷移していることが推測された。しかし、今回の結果からは、音声知覚関連領域間での活動の遷移の傾向や明瞭度に関連した活動源強度の変化について、システムティックな傾向は見出せなかった。一方、右半球の上側頭回後方での反応潜時には、SN 比の有意な影響が観察さ



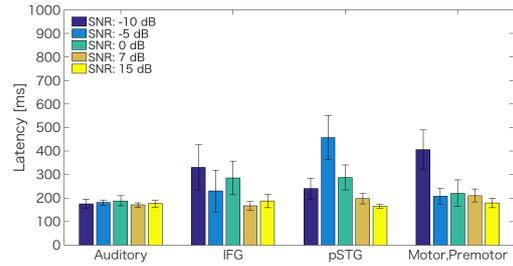
(a) The peak intensity of brain activity in the left hemisphere.



(b) The peak latency of brain activity in the left hemisphere.



(c) The peak intensity of brain activity in the right hemisphere.



(d) The peak latency of brain activity in the right hemisphere.

Fig. 3: The relationship between SNR and the features of the source localization in both hemisphere. IFG means inferior frontal gyrus and pSTG means posterior superior temporal gyrus.

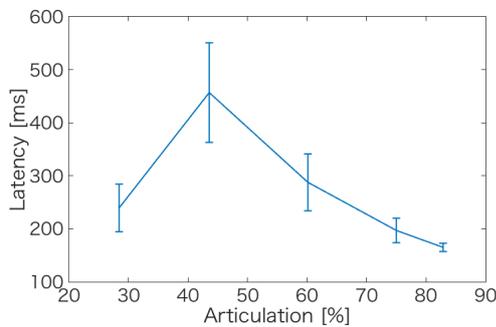


Fig. 4: The relationship between articulation and latency in the posterior superior temporal gyrus in the right hemisphere.

れたことから、音声知覚における明瞭性が反映されている可能性が示唆された。

### 参考文献

[1] R. Hari *et al.*, Auditory Evoked Transient and Sustained Magnetic Fields of the Human Brain, *Experimental Brain Research*, 40, 237-240, 1980.  
 [2] M. S. Gazzaniga *et al.*, "Cognitive Neuroscience: The Biology of the mind (4th Ed)," W. W. Norton & Company, 2014.  
 [3] F. Pulvermüller *et al.*, Motor Cortex Maps Articulatory Features of Speech Sounds, *PNAS*,

103, 7865-8770, 2006.

[4] K. Watkins *et al.*, Modulation of Motor Excitability During Speech Perception: The Role of Broca's Area, *Journal of Cognitive Neuroscience*, 16, 978-987, 2004.  
 [5] P. C. Wang *et al.*, Cortical Mechanisms of Speech Perception in Noise, *Journal of Speech, Language, and Hearing Research*, 51, 1026-1041, 2008.  
 [6] A. D' Ausilio *et al.*, The Role of Motor System in Discriminating Normal and Degraded Speech Sounds, *Cortex*, 48, 882-887, 2012.  
 [7] Z.M.Wu *et al.*, Interaction between auditory and motor system in speech perception, *Neurosci Bull*, 30(3), 490-496, 2014.  
 [8] 嵯峨ら, 音声明瞭度に関連した脳磁界計測一聴覚野および運動野における活動源解析一, 日本音響学会 2017 年秋季研究発表会, 2017.  
 [9] 長谷ら, 親密度別単語了解度試験用音声データセット (FW03) 単音節音声のラウドネス校正, 日本音響学会誌-, 64, 647-649, 2008.  
 [10] N. Nishitani and R. Hari, Temporal Dynamics of Cortical Representation for Action, *PANS*, 97, 913-918, 2000.