

音声明瞭度に関連した脳磁界反応の時空間特徴量*

☆嵯峨直樹, 矢野肇 (神戸大/産総研), 滝口哲也 (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

昨今, 脳信号を利用して外部デバイスを制御あるいは操作するような試みが行われている。これらの取り組みはブレイン・コンピュータ・インタフェース (Brain Computer Interface: BCI) と呼ばれ, 新たなインターフェイスとして注目されている。しかしながら, 既存の BCI の多くは, P300 スペラーのような予め用意した選択肢を識別するものであり, ヒトの意思や情動, 思考, 理解などを解読するまでには至っていない。よって, BCI 開発の発展には, このようなヒトの内部状態あるいは認知・知覚状態を脳活動から読み取ることが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ような脳活動の同定を試みた。

一般的にヒトが音を知覚をする際には, まず聴覚野で基本的な音響処理が行われると考えられている。聴覚野の反応の中でも, 代表的な例として N1/N1m 反応が挙げられる。N1/N1m 反応は刺激オンセットからおおよそ 100 ms 後に出現し, ある程度認知過程の影響を受けることが知られているが, 反応潜時を考慮すると高度な言語情報処理の主たる役割を果たしているとは言い難い [1]。

音声に含まれる言語情報に関しては, 聴覚野での処理が行われた後で特定の脳部位で処理されと考えられている。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 [2]。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 [3] 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 [4]。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”が提案されている [5]。

これら領野における活動例として, 例えば雑音環境下における音声知覚脳活動に関しては, 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 [6] 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。

前稿では音声明瞭度の特異的な脳活動特徴を得るために, 明瞭度試験と脳磁界計測を行い, 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 [7]。右半球の知覚性言語野に対応する領野においては, 言語情報を処理する活動が行われている可能性が示唆された。しかし, 明瞭度に関連した有意な差のある活動としては言い難かった。また, ブローカ野も含めた, 音声の認知に関連のあるとされる領野間での活動, 特に明瞭度に関連した活動の相互作用も明らかにされていない。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

本稿は雑音環境下における音声の聞き取りを行うため, 背景雑音としてホワイトノイズ, また音声刺激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 (fhi) による日本語単音節 100 個を利用した。背景雑音のノイズは, 音圧レベルを 50 dB に設定した。音声刺激の音圧レベルは, 背景雑音の音圧レベルに対して, SN 比が特定の値になるように調整された。SN 比は 0, 6, 12, 18, 24 dB の五段階に設定されたが, これらの値は予備実験に基づいて, 明瞭度がおおよそ 25, 30, 55, 70, 85 % になるように決定された。音圧レベルの較正は, 人工耳 (B&K 社製 Type 4157) を用いて行われた。被験者に提示する刺激音としては, 背景雑音を定常的に流し, 一定間隔で音節を挿入したものを用了。

3 明瞭度試験

3.1 方法

聴覚健全者 9 名 (男性 5 名, 女性 4 名, 19-25 歳, 右利き) を被験者とした。明瞭度試験は防音室内で行

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

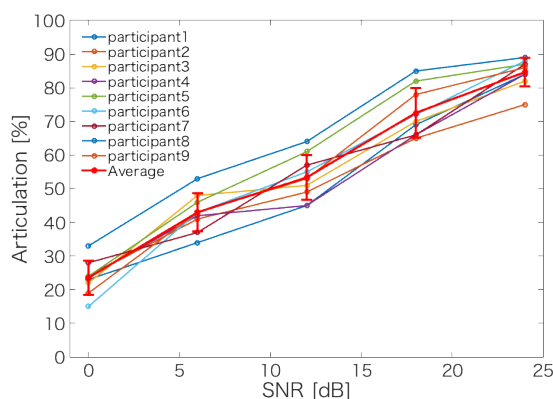


Fig. 1. The relationship between SNR and articulation.

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。音声刺激音の呈示間隔は3 sで、被験者にはこの間に自身が聞こえたと思った通りの音節を回答用紙に記入するように教示した。実験はSN比の条件ごとに行われ、各SN比の刺激音での100音節の再生はランダムな順序で行った。また、SN比の順序もランダムに決定された。

3.2 結果

明瞭度試験の結果を Fig. 1 に示す。SN比が低い条件から順に、明瞭度の平均値は、23.6, 43.0, 53.3, 72.6, 84.7 %であり、標準偏差はそれぞれ5.1, 5.7, 6.6, 7.4, 4.2 %であった。すべての被験者でSN比の向上に伴って、明瞭度も上昇する傾向が観察できた。

4 脳磁界計測

4.1 方法

脳磁界計測の被験者は明瞭度試験と同一にした。計測場所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-122TM, Neuromag Ltd, Finland) を用いて行った。SN比を変化させた音声刺激音に対する脳磁界反応を計測するにあたり、明瞭度に関与する反応と音声刺激音の振幅が変化することによる物理的な原因による反応かを区別するために、コントロール条件として、SN比を音声刺激音と同様に变化させたスピーチノイズを用いた。スピーチノイズは、音声のエンベロープを掛け合わされ、100音節それぞれに対応したものを作成した。音圧レベルは、音声刺激音と同様に背景の白色雑音の音圧レベルに対して、SN比が0, 6, 12, 18, 24 dBとなるように設定された。これら合計1000種類の刺激音を用い、雑音環境下での脳磁界反応を計測した。

実験は、明瞭度試験と同様に、SN比の順序はラン

ダムに設定し、SN比の条件ごとに行った。あるSN比の条件における、100種類の音声刺激音と100種類のスピーチノイズは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/) 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、音声刺激に対する誘発反応、スピーチノイズに対する誘発反応をそれぞれ200回加算し、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。3,000 fT/mmを超える磁界を計測した場合はアーティファクトとして加算から除外した。

4.2 解析

脳磁界計測実験によって得られたデータを用いて、脳内電源推定を行った。電源推定は、(1) 聴覚野が覆われていると思われる18個のチャンネル、(2) 知覚性言語野 (上側頭回後方)、角回、縁上回が覆われていると思われる16個のチャンネル、(3) 下前頭回が覆われていると思われる12個のチャンネルデータをそれぞれ用いて、刺激オンセットから1000 msまで2.5 msごとに行った。推定結果の選別にあたり、Goodness of fit と Confidence volume の2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%3次元信頼領域を示している。パラメータの制約条件は、(i) Goodness of fit $\geq 70\%$ 、(ii) Confidence volume $\leq 16000 \text{ mm}^3$ とした。さらに、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 [8]。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、各領域におけるSN比、明瞭度との関連性を調べた。

4.3 結果

聴覚野、ブローカ野、ウェルニッケ野のそれぞれの領域における活動源強度の時系列変化を Fig. 2 に示す。特に、??は音声刺激を聴取した際の、??はスピーチノイズを聴取した際の活動の変化を示している。??と??は共に音声刺激を聴取した際の活動の変化の例を示している。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、刺激音の種類 (音声刺激聴取時とスピーチノイズ聴取時) とSN比、領域と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、

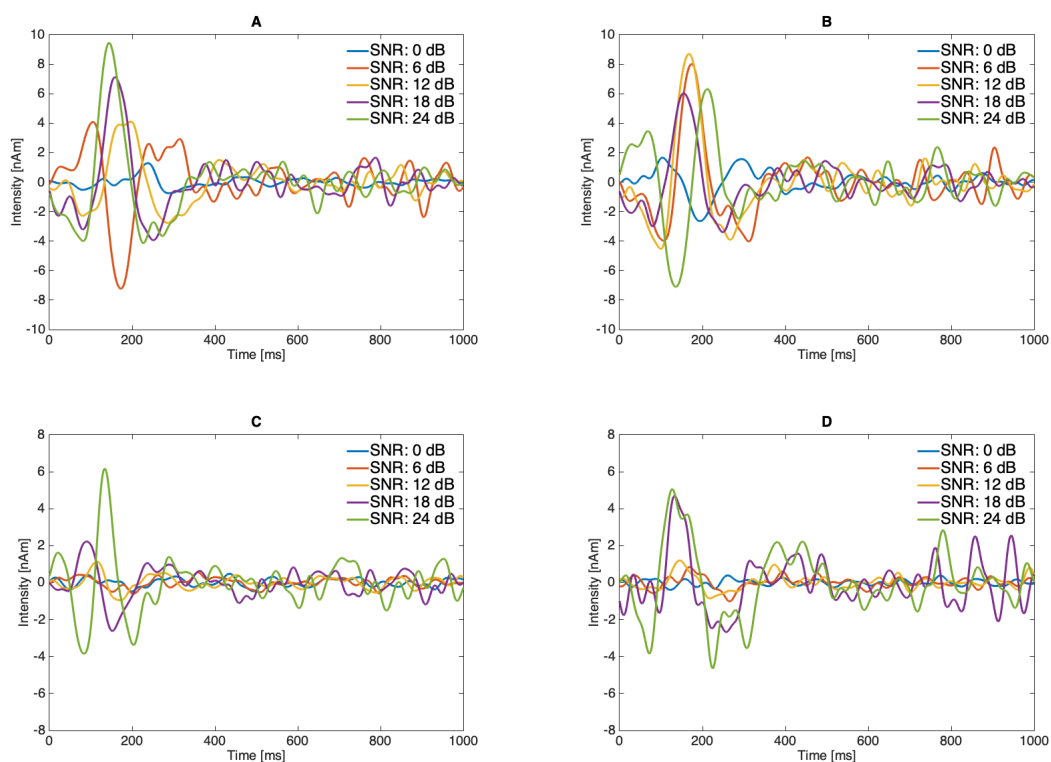


Fig. 2. The time series of the source in the Auditory area in listening to (A) speech sounds and (B) speech noises. The time series of the source in the (C) Broca's area and (D) Wernicke's area in listening to speech sounds.

刺激音の種類に寄らず聴覚野は他の二つの領野に比べ有意にピーク値が大きいことが分かった ($p \leq 0.05$)。また、SN 比の上昇に連れて、強度も上昇していく傾向が確認できたが、音声刺激を聴取時とスピーチノイズ聴取時で同様の傾向であった。ブローカ野での活動は音声刺激を聴取した際の活動の方がスピーチノイズを聴取した際よりも有意に大きいことが確認できた。ウェルニッケ野については、刺激音の種類によって SN 比に対する変化の仕方が異なるように見えるが、有意な差は確認できなかった。潜時については、音声聴取時にはいずれの領野においても SN 比の上昇につれて潜時が早まっていく傾向が確認できた。また、音声聴取時ではブローカ野やウェルニッケ野におけるピーク潜時は聴覚野よりも遅い傾向が見られた。しかし、ブローカ野とウェルニッケ野の間での有意な差は確認できなかった。

5 考察

Fig. 2 から聴覚野周辺における潜時およそ 100～200 ms の間に見られる高いピークを持つ反応は SN 比に関わらず確認できることから N1/N1m 反応であると考えられる。また、Fig. 3 の電源推定の結果から聴覚野での反応は、刺激音に関わらず SN 比の上昇

につれてピーク値が上昇していくことから、刺激の物理特性を反応している傾向が色濃く見られる。潜時について、聴覚野での反応よりも他の二つの領野でのピーク潜時の方が遅いことから、聴覚野からの活動の遷移の傾向が伺えるが、ブローカ野とウェルニッケ野の間での相互作用の可能性は考え難い。ブローカ野では、刺激音の条件によって有意に活動の強さが異なることが確認できた。しかし、活動は SN 比の上昇に伴い上昇していくことから、音声を特定するための活動というよりは、音声として認識できたことにより活動が強まる可能性が考えられる。左半球のウェルニッケ野では、有意的な差は見られなかったものの、SN 比に対するピーク強度の変化の仕方が他とは異なることから、音声理解の過程を反映している可能性が考えられる。特に、SN 比が低いときと高い時では、比較的強度は高いが、潜時は異なる傾向が見られる。SN 比が低いときでは、音声を同定するために活動強度が強まり、また、同定するのに時間がかかると考えられる。SN 比が高い時では、音声理解の結果として活動が強まり、また音声の同定に負荷がかからないため、潜時が早まると考えられる。

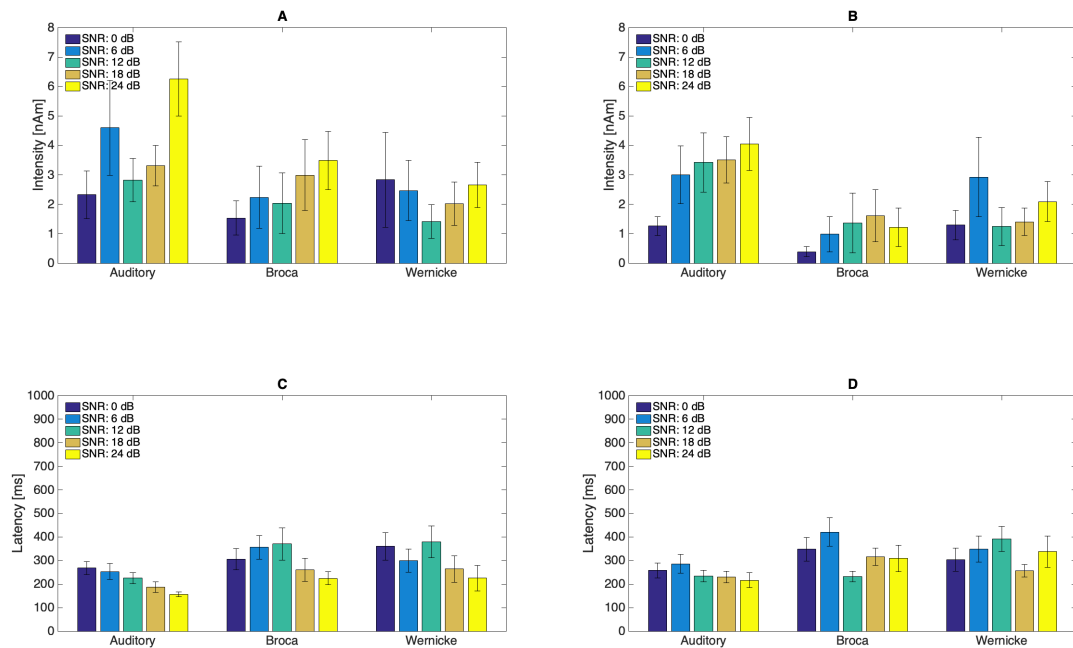


Fig. 3. The peak intensity of brain activity in listening to (A) speech sounds and (B) speech noises. The peak latency of brain activity in listening to (C) speech sounds and (D) speech noises.

6 まとめ

音声情報の処理として、聴覚野からブローカ野やウェルニッケ野に活動が遷移する可能性が考えられた。聴覚野では、刺激音の条件によらずSN比の上昇に伴い、強度も強まっていったため刺激の物理特性を反映している可能性が高い。音声の認知、理解に関しては、ブローカ野やウェルニッケ野が関与する可能性が考えられたが、有意な差は確認できなかった。しかし、左半球のウェルニッケ野では、特に強度について他の二つの領域とは異なる傾向を示しており、ブローカ野とは異なる言語処理をしている可能性が考えられた。

謝辞 本研究の一部は、JSPS 科研費 JP18K19820 の助成を受けたものである。

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, 7865-7870, 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. Rauschecker and S. K. Scott, Maps and Streams in the Auditory Cortex : Nonhuman Primates Illuminate Human Speech Processing, *Nat. Neuroscience*, 12, 6, 718-724, 2009.
- [6] P. C. Wang *et al.*, Cortical Mechanisms of Speech Perception in Noise, *Journal of Speech, Language, and Hearing Research*, 51, 1026-1041, 2008.
- [7] 嵯峨ら, 音声明瞭度に関連した脳磁界反応：聴覚野および知覚性言語野における解析, 日本音響学会 2018 年秋季研究発表会, 2018.
- [8] N. Nishitani and R. Hari, Temporal Dynamics of Cortical Representation for Action, *PANS*, 97, 913-918, 2000.