

博士論文

論文題目 視聴覚不一致を伴う聴覚情報処理の神経基盤
に関する皮質脳波研究

氏名 宇野 健志

目次

I. 要旨	1
II. 序文	2
はじめに.....	2
本論文の構成	2
A. 音素認知	4
B. 視聴覚統合	5
C. マガーク効果（融合認知）	6
D. マガーク効果（融合認知）の脳活動検証（研究①）	9
E. 視聴覚不一致の検出.....	11
F. 視聴覚不一致処理に関わる脳活動の部位	11
G. 単一種感覚入力の認知処理	12
H. 不一致を伴う視聴覚情報の2つの処理機構（研究②）	13
I. 皮質脳波（electrocorticography; ECoG）	15
III. 目的	19
A. 本研究の目的	19
B. 研究デザイン	19
IV. 対象と方法	21

A. 対象.....	22
B. 方法.....	24
研究① マガーク効果（融合認知）の脳活動.....	24
1. 課題.....	24
2. 認知行動成績の統計解析.....	28
3. 応答時間.....	28
4. 皮質脳波（ECoG）記録.....	29
5. 皮質脳波（ECoG）解析.....	32
6. 電極の選択.....	38
7. HGA と認知行動成績の関連解析.....	39
研究② 視聴覚不一致を伴う聴覚情報処理.....	41
1. 課題.....	41
1-1. 視聴覚課題.....	41
1-2. 聴覚課題.....	43
2. 認知行動成績の統計解析.....	44
2-1. 視聴覚課題.....	44

2-2. 聴覚課題	44
3. 応答時間	45
4. 皮質脳波 (ECoG) 記録	46
5. 皮質脳波 (ECoG) 解析	46
6. 電極の選択	46
7. HGA と認知行動成績の関連解析	47
8. 皮質の活動時間の解析	48
V. 結果	49
研究① マガーク効果 (融合認知) の脳活動	49
A. 認知行動成績	49
B. 応答時間	51
C. Event related spectral perturbation (ERSP)	51
D. 事象関連 HGA	56
E. HGA と認知行動成績	59
研究② 視聴覚不一致を伴う聴覚情報処理	66
A. 認知行動成績	66
B. 応答時間	67
C. 事象関連 HGA	68

D. HGA と認知行動成績.....	70
E. 皮質の活動時間.....	75
VI. 考察	77
A. 研究① マガーク効果（融合認知）の脳活動.....	77
1. 予期と不一致検出.....	79
2. IFG/STS と視聴覚統合.....	80
3. 本研究の新奇性.....	82
B. 研究② 視聴覚不一致を伴う聴覚情報処理基盤の動的変化.....	85
1. トップダウン処理・ボトムアップ処理.....	86
2. 本研究の新奇性.....	91
3. 課題の有効性.....	92
C. 総合考察.....	94
1. 研究の限界.....	94
2. 展望.....	95
VII. 結論	97
VIII. 謝辞	98
IX. 引用文献	99

I. 要旨

異なる感覚器官への入力がある脳内でどのように処理されヒトに認知されるのか、その情報処理機構は明らかとされていない。本研究では、言語に関連した不一致を伴う視聴覚情報に対する聴覚情報処理の神経基盤を検討した。信号対雑音比がよく時空間解像度に優れた皮質脳波を用いて、high gamma activity (HGA) の変化を調べたところ、正しい聴覚認知と関連する脳活動を捉えた。また、不一致の小さい視聴覚入力には下前頭回 (IFG) が、不一致の大きい視聴覚入力には上側頭溝 (STS) が関与する可能性が示された。入力情報の視聴覚不一致の程度に応じて処理過程が変化することが示唆された。

II. 序文

はじめに

当科では脳神経外科手術における言語機能温存の重要性から、言語機能マッピングを中心とした脳機能の研究を行ってきた。これまで、言語の側性化、局在、ダイナミクスなどについて多くの成果を得てきた¹⁻⁴。言語機能に対する理解をさらに深めるには言語機能ネットワークを対象とした研究が必要であることは自明であるが、従来の言語課題では、たとえ単語の処理であっても賦活される脳領域の統制が難しく、解析が困難であった。例えば、「りんご」という単語を視覚的に提示しただけで、りんごの映像が脳内で認識され、りんごの香りや味、食感などが想起される。更に、りんごを食べるための運動野の賦活が起こり、場合によっては最後にりんごを食べたときの誰かとの会話が思い出されることもあるかもしれない。ただでさえ複雑な多領域間のネットワークに関する検討を行う上で、認知課題を極めて要素的にする必要があったと考えた。そこで今回、言語を形成する最小単位である音素に焦点を絞り、言語処理の最も基礎的な視聴覚認知に関する検討を行った。

本論文の構成

本研究は、視覚情報を伴った聴覚情報の処理機構に関する 2 つの皮質脳波

研究からなっている。研究①では、後述するマガーク課題によって起こる脳活動がどのような認知と関連するかを検討した。そして研究②では、この結果から得られた視聴覚不一致検出機構について、視聴覚不一致の程度に着目し、より詳細に検討した。

A. 音素認知

言語の構造は、文 (sentence)、単語 (word)、音節 (syllable)、音素 (phoneme) などの階層に分けられる。一般に、音素とは意味の相違をもたらす最小単位の音であり、音として他と区別可能なものを指す。音節とはひとまとまりに発音される最小単位のものであり、日本語ではふつう子音と母音のペアからなる。音韻論、音声学などでは古くから音節、モーラ、拍などの概念を扱うが、それぞれについての定義や解釈、お互いの関係など、いまだに定まった見解はなく、諸説あるモーラや拍の概念についてはここでは触れない。

このように言語の構成要素の中で基礎となる音素であるが、スムーズな言語理解のためにはそのひとつひとつの正しい認知が必要である。一方、音声における音の区別は周波数の違いという物性の違いとして捉えられるが、単純にその物理的な違いのみで音素認知がなされるわけではない。日本語を母語とするヒトにおいては、異なる物性を持つにもかかわらず英語の r と l を区別することが難しいことや、個人差や声の高さの違いは周波数の違いであるのに苦も無く音素認知が可能であることは、音素認知が単純な入力情報の物性の違いに基づくものではないことを意味する。そのような音素認知の神経基盤を解明することは、ヒトの認知情報処理機構を明らかにすることに通じる重要なものである⁵。

B. 視聴覚統合

聴覚認知とは音声情報のみで決定されるものではない。対面での会話では、相手の声（聴覚情報）だけでなく口唇の動きをはじめとした視覚情報も同時に入力されるが、実際にこのような状況での言語処理では両情報が統合的に処理されることが知られている。本研究では視覚情報と聴覚情報を矛盾なく処理されて生じる認知を統合認知として扱い、後述するマガーク効果での視覚『か』と聴覚『ぱ』の入力情報からこれらと異なる第3の『た』という認知を得る場合を融合認知として区別する。従来の知覚・認知研究では、それぞれの感覚器に入力された情報は独立して処理されるものとして扱われ、各知覚モダリティについて個別に検討されてきた^{6,7}。しかしながら現実世界では、外部環境からの情報が同時に複数の感覚器官へ入力される。このように複数の感覚器官への入力情報を統合して処理することで、我々は外部環境をより正確に認知している。実際、複数種類の感覚入力があることで単一感覚入力よりも認知精度が向上するという現象は、古くからよく知られている⁸⁻¹⁰。例えば Sumby らは、話者の音声聞き取りにくい環境（雑音や他者の話声中）での聴音課題において、話者の顔を提示することで課題成績が向上することを示した⁸。これは、口唇の動きから発声内容を無意識に予測することで聴覚認知精度が向上するためと考えられている¹¹。このような視覚と聴覚の統合処理は、言語獲得時に無意識に身

に付く処理機構である。言語は獲得の過程で音素認識が構築されるが、これには視覚的要素も大きく関わり視覚情報と聴覚情報のペアリングが起こる^{12,13}。そして、正常な音素認識学習により構築されたニューラルネットワークとしての視聴覚統合機構は、音素認識精度を上げることに寄与している⁷。音素は物性として異なっていた方が区別して認識しやすいが、音響的に似た音素は音声のみでの区別が難しいこともあり、視覚情報がこれを補う情報として有用である。本研究では、視覚情報を伴う聴覚情報の処理について検討を行う。

C. マガーク効果（融合認知）

一方で、複数の感覚入力に情報の不一致がある場合、本来の入力情報とは異なる認知をしてしまうことも知られている。この代表的な現象として視聴覚分野におけるマガーク効果（＝融合認知）が挙げられる。マガーク効果とは、視覚情報によって正しい聴覚認知が歪められる現象である。例えば、音声『ぱ』と同時に『か』と発声しているときの顔の動画を提示することで、『た』と認知するものである^{14,15}（図1）。

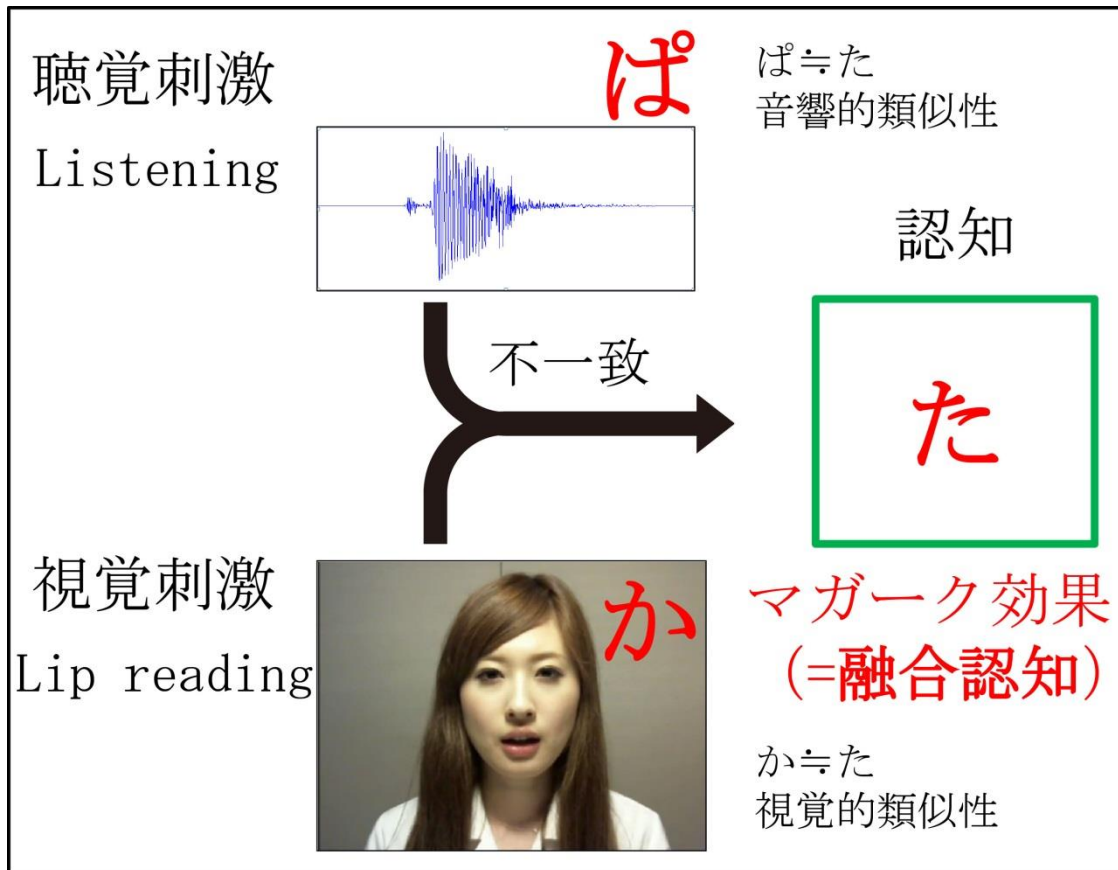


図 1. マガーク効果の模式図

マガーク効果とは、『か』と発声している動画の音声を『ぱ』と入れ替えると、本来提示されていない『た』という音声として融合認知する現象である。視聴覚情報に不一致があることがわかっているにもかかわらずこの現象は起こる。

McGurkらは視覚『ka』、聴覚『pa』、融合認知『ta』の他に視覚『ba』、聴覚『ga』、融合認知『da』の組み合わせを融合認知の起こる組み合わせとして報告しているが¹⁴、本研究では前者の視覚『ka』、聴覚『pa』、融合認知『ta』の組み合わせを検討対象とし、視覚『ka』、聴覚『pa』の組み合わせの動画をマガーク条件とした。

融合認知は幼少からでも（3～5 歳）起こる現象であるが、とりわけ成人においてよく起こることから、言語発達に伴い起こりやすくなると考えられている¹⁴。さらに、融合認知は、顔の提示を小さくしても口唇の動きを知覚できる範囲であれば起こることや¹⁶、音声と口唇の動きの時間的なズレを生じさせても知覚上同時に感じられる程度のズレ（100～200 ms 程度）であれば起こる¹⁷ 強固なものであることが知られている。

さらに融合認知は、たとえ音声と異なる発声動画が提示されていることを知らされていたとしても、自動的に起こる現象である¹⁸。ただし、融合認知は100%の確立で起こる現象ではなく、施行ごとに融合認知が起こることも起こらないこともある。個人間での融合認知生起率もばらつきが大きい^{14,15,19}。また、言語の種類に寄らず起こる現象であるが、日本人では起こりにくいことも知られる¹⁵。

この現象を利用して、異なる感覚器官への入力を統合する情報処理の神経基盤を明らかにすべく多くの研究報告がなされている。頭皮脳波、functional Magnetic Resonance Imaging (fMRI)、magnetoencephalography (MEG) など低侵襲的画像検査の発達により、その機能と活動部位の研究は盛んになってきている。これまでの報告では、マガーク条件で賦活される部位として下前頭回 (IFG)、上側頭溝 (STS) などがあげられている^{20,21}。Sekiyama らは、fMRI や positron

emission tomography (PET) を用いて、同じ視聴覚課題でも視聴覚情報に不一致のない課題と比べ、マガーク条件を課した際に下前頭領域および上側頭領域が活動することを観察した²¹。しかしながら、これらの部位がマガーク条件を課すことによって賦活されることは示されているものの、融合認知との直接的な関連は示されていない。

D. マガーク効果（融合認知）の脳活動検証（研究①）

従来、融合認知が起こるときには何らかの脳活動があると考えられ、これを捉えようと様々な研究が行われてきた。しかし、マガーク条件によって賦活される脳活動は捉えているものの、本当に融合認知と関連した脳活動を検出していると言えない。

これまで挙げた知見の通り、ヒトは幼少の言語獲得段階から視覚と聴覚のニューラルネットワークが構築され、自動的に、不可避的に視聴覚統合認知が起こる。これとは別に、視聴覚情報が乖離しているときにだけ働く融合処理機構があるのだろうか。

融合認知とは、統合認知の結果として起こるエラーのようなものである。本来、視覚情報と聴覚情報が一致しない場合、両者を融合させるのは不適切である。視聴覚情報が一致しない場合は不一致を検出し、聴覚情報をそのまま（＝

視覚情報と分離して) 処理する必要がある。マガーク条件でみられる脳活動は、むしろ不適切な統合を避けるために視聴覚不一致を検出する機構を反映しているのではないだろうか。

このような考察に基づき、本研究ではマガーク効果（融合認知）の脳活動について検証を行う。まず視聴覚一致条件の動画を提示した際の脳活動と比較してマガーク条件でより強く賦活される脳活動を捉え、その上でその活動が何を反映するものであるかを検討する。ここでは、信号対雑音比のよい皮質脳波の利点を生かしてトライアルベースで認知結果（認知行動成績）を分類し、それぞれの脳活動の比較を行う。具体的には下記の 2 つの認知結果で分類したモデルを比較検証する。

- モデル 1：融合認知の脳活動

マガーク条件で誘発される脳活動は融合認知と関連したものである。

- モデル 2：不一致検出の脳活動

マガーク条件で誘発される脳活動は、不適切な融合認知を避け、正しい聴覚認知をするための脳活動である。

E. 視聴覚不一致の検出

研究①は、マガーク条件で得られる脳活動は、融合認知ではなく正しい聴覚認知と関連した。これまで、視覚情報と聴覚情報の統合的認知については多くの研究がなされてきたが、研究①で示された視聴覚情報の不一致を検出する機構についてはほとんど検証されていない。視聴覚の統合は意識せずとも、自動的に、不可避におこるほど強固なものである¹⁸。このため、視覚と聴覚の情報に不一致があると誤った聴覚処理が行われることになってしまう。そこで、不一致な視聴覚入力となされた場合には、この不一致を検出することが必要となる。研究②では、視聴覚不一致入力から不一致を検出する神経機構について調べる。視聴覚情報の不一致検出の神経基盤を検討することは、正確な聴覚処理の機序を明らかにする上で極めて重要である。具体的には、局在と不一致の程度（レベル）に関する検討を行う。

F. 視聴覚不一致処理に関わる脳活動の部位

前述の通り、マガーク条件を代表とする、不一致のある視聴覚入力となされた際には、IFG と STS が関与することが知られている^{20,22-24}。視聴覚の不一致入力となされると、IFG（特にブロードマン 44 と 45 野）と STS 後方が活動する^{19,23,25,26}。両部位は視覚単独刺激や聴覚単独刺激と比べて視聴覚刺激でより強く

活動し、視聴覚情報処理に関与する部位とされている²⁷⁻³³。また、視聴覚刺激において、視覚情報と聴覚情報の一致しているときよりも一致していないときにより強く活動し、視聴覚不一致入力の処理により強く関与していることが知られている^{25,31,34-37}。一方で、不一致を伴う視聴覚情報処理における IFG と STS の機能的差異については明らかとされていない。

G. 単一種感覚入力の認知処理

単一種感覚入力の認知処理においては目標刺激と妨害刺激の違いの大きさにより情報処理の神経基盤が変化することが知られている。聴覚単独入力の研究においては、目標刺激である聴覚刺激（標的信号；Target Signal）と妨害刺激である聴覚ノイズ刺激（雑音；Noise）との差が顕著である場合には上側頭領域、両者の差が顕著でない場合には下前頭領域で情報処理がなされると報告されている^{38,39}。また、同じく単一種感覚刺激である視覚探索課題では、色や形などの点で妨害刺激との違いが顕著な目標刺激は感覚野において処理され、一方で妨害刺激との違いが不明瞭な目標刺激は、より高次の皮質である前頭葉で処理されることがわかっている⁴⁰⁻⁴³（図 2）。つまり、目標刺激と妨害刺激の違いの程度が、目標刺激の処理に関わる部位の一次決定因子と考えられている。

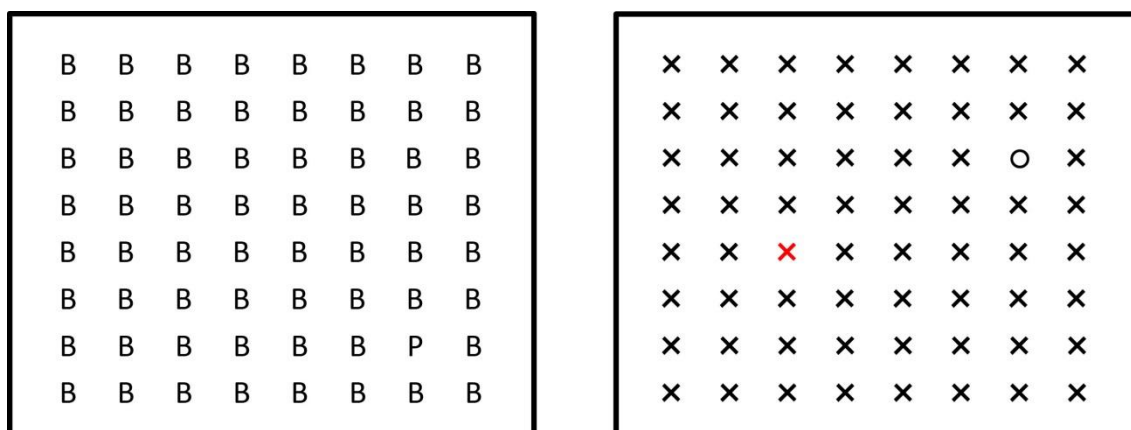


図 2. 視覚探索課題における 2 つの処理

(左) 前頭葉で処理がなされる視覚探索課題の例。P の記号が B の記号の中では顕著でなく、主に前頭葉での処理で検知される。(右) 感覚野で処理がなされる視覚探索課題の例。記号や色の違いが顕著であると主に感覚野での処理で検知される。

H. 不一致を伴う視聴覚情報の 2 つの処理機構 (研究②)

入力情報の不一致の大きさに応じて脳の担当領野が動的に切り替わるという単一種感覚入力の知見を踏まえると、視聴覚情報の不一致処理機構もまた、入力情報の乖離度に応じて情報処理の主座が前頭葉と感覚野とに動的に変化する可能性があると考えられる。異種感覚間入力である視聴覚入力時の聴覚情報の処理においては、実際に提示される音声 (Target signal) と口唇の動きから予測される音声 (Noise) との違いが大きければ、より容易に目標刺激である音声を正しく認知できる。すなわち、視覚情報から予測される音声情報が目標刺激である実際の音声情報と顕著に異なっている場合には、これらの情報は感覚野で処理され、これらの情報の違いが顕著でない場合には、前頭葉で処理される

ということである。

前述のとおり、これまでの研究により不一致を伴う視聴覚情報処理には IFG と STS が関与する。視聴覚入力における聴覚情報の選択的処理には、口唇の動きから予測される音声と実際の音声との乖離が小さいときには IFG が、乖離が大きいときには STS が処理にあたると考えられる（図3）。

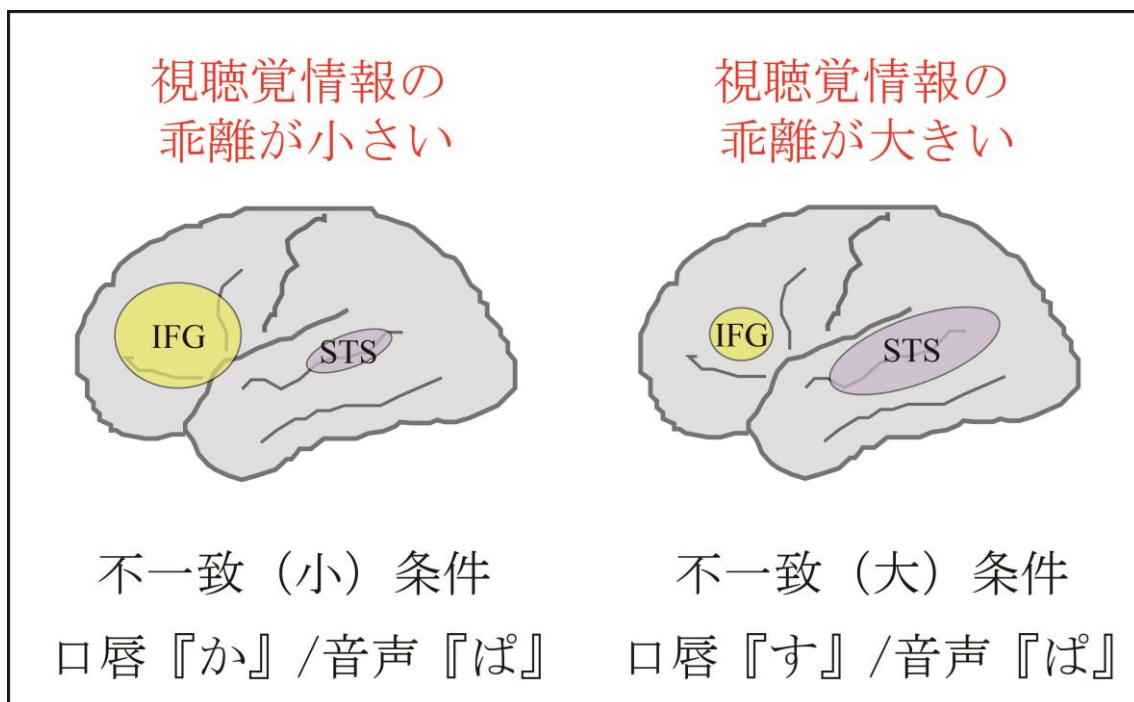


図3. 不一致を伴う視聴覚情報の2つの処理機構の概念図

（左）視聴覚情報の乖離が小さいときには前頭葉で処理がなされる。特に下前頭回（IFG）の関与が想定される。（右）視聴覚情報の乖離が大きいときには感覚野で処理がなされる。特に上側頭溝（STS）の関与が想定される。

これまでなされてきた視聴覚情報処理研究においては、視覚情報と聴覚情報がどのようにして一つの知覚として統合されるのかに焦点が当てられ、入力

情報の乖離の大きさと脳内の処理領域の関係については調べられていない。加えて、入力情報の性質による処理基盤の動的変化の知見も視覚単独や聴覚単独といった単一感覚入力の研究が主である。そのため、視聴覚情報処理の神経基盤は視聴覚情報の乖離の大きさに応じて変化するのか、この情報処理において IFG と STS の機能的差異はあるのか、については明らかとなっていない。そこで本研究では、視聴覚情報の不一致の程度を操作することで、被験者が視聴覚情報から聴覚情報を選択的に処理する際の脳活動について検討する。これによって視聴覚情報処理で賦活する IFG と STS という 2 領域の機能的差異を明らかにする。

I. 皮質脳波 (electrocorticography; ECoG)

この問題を検討するに当たり、本研究では皮質脳波を用いる。難治性てんかんの患者に対し、てんかんの焦点を同定するために、開頭手術によって硬膜下に電極を留置することがある。この硬膜下電極より計測できる皮質脳波 (electrocorticography) は、通常用いられる頭皮脳波が髄液、頭蓋骨、頭皮で減弱・拡散した信号であるのに対し、軟膜越しに大脳皮質直上で電位変化を記録できるため、極めて高い信号対雑音比で脳活動を捉えることが可能である (図 4)。

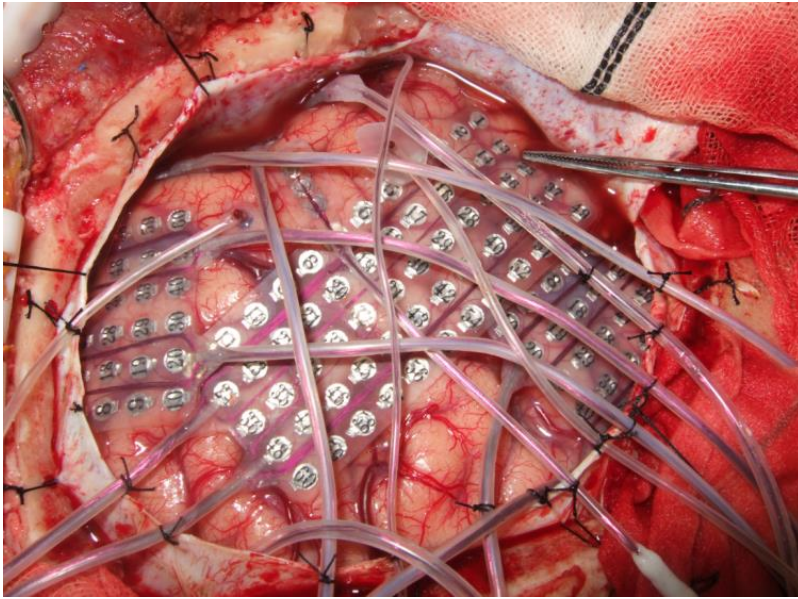


図 4. 硬膜下電極

硬膜下電極留置時の術中写真。硬膜下電極により計測できる皮質脳波は、軟膜越しに大脳皮質から直接電氣的活動を記録できる。頭皮電極に比べ、髄液、頭蓋骨、頭皮による電位変化の減弱・拡散がない為、信号雑音比・空間分解能が高い。

一般に、低い周波数で律動する信号ほど波の振幅は大きい。頭皮脳波では 30 Hz を超えるような周波数信号は減弱し雑音に埋もれ有意な信号として観察することは難しい。しかし、皮質脳波はこれを超える周波数信号を捉えることができ、高周波脳律動を観察できるところに大きな利点がある。高周波脳律動は近年、大脳皮質における局所的な脳機能と強い関連を示す脳活動の指標として注目されている。特に、High gamma 帯域 (70~150 Hz) における事象関連同期 (High gamma activity; HGA) は神経発火と強く相関し⁴⁴、運動、感覚、言語、注意など脳局所活動を反映すると知られている^{3,45-47}。HGA は脳波解析により得

られる電気信号であることから、ミリ秒単位の高い時間分解能を有している。一方で、fMRIは脳活動に伴う脳血流変化（酸化ヘモグロビンと脱酸化ヘモグロビンの比）から局所脳活動を捉えているため、時間分解能は高くない⁴⁸。脳波やMEGはHGA同様に高い時間分解能をもつが、HGAに対応する高周波成分を捉えることは困難である。

空間解像度に関しては、HGAは電極直下にある神経集団の活動を反映するため、空間解像度は電極密度に依存することになる。このため通常使用される硬膜下電極の場合、解像度は5~10 mmとなる。本研究で用いた硬膜下電極は中心間距離5mmの電極を用いており、5mmの分解能を得ている。fMRIは3~4mm程度の空間分解能をもつためこれよりは少し劣るものの、頭皮脳波の3~5 cm、MEGの1~2 cmと比較すると比較的優れている。また、頭皮脳波やMEGでは逆問題という信号源推定における局在精度の不確定性が問題となる。

そして何よりも皮質脳波は信号対雑音比がよく、event relatedの解析において最も威力を発揮する。fMRIでもevent relatedの解析は可能であるが、信号対雑音比が劣るため、今回皮質脳波で行ったような認知行動成績に基づくトライアルベースの解析を行うことは通常難しい。

このように、HGAは高い空間解像度と時間分解能を有し、ヒトを対象とした研究では計測の困難な直接的な神経活動を捉えられることから、ヒトの複雑

な脳機能を解き明かす神経科学の手法としてだけでなく、臨床における脳機能マッピングの手法としても、高い有用性をもつと考えられる。皮質脳波を用いることで、視聴覚入力からの選択的聴覚情報処理において下前頭回と上側頭溝がどのような役割を担っているのかを明らかにすることができ、正しい聴覚情報処理を行う神経基盤の理解を深めることに繋がると考えられる。

Ⅲ. 目的

A. 本研究の目的

視覚情報を伴う聴覚情報処理の神経基盤について検討する。研究①ではマ
ガーク効果（融合認知）の脳活動を、研究②では視覚入力を妨害刺激とした際
に目標刺激である聴覚入力を正しく認知するための情報処理機構を、認知行動
成績と皮質脳波を用いて詳細に検討することを目的とした。

B. 研究デザイン

本研究は研究①と研究②からなる。頭蓋内電極留置を行った被験者に対し
て自作の視聴覚動画を提示し、認知行動成績を評価した。神経生理学的評価を
加えるために、認知行動成績とともに皮質脳波を計測した。マガーク条件（研
究①）および視聴覚不一致な入力（研究②）に対する情報処理機構を調べるた
め、それぞれの条件と視聴覚一致入力時の HGA の差を比較し、下記の項目につ
いて調べた。

1. 認知行動成績

＝マガーク効果（融合認知）が起こるのか、視聴覚不一致が聴覚認知を変
容させるのかを確認する（研究①）。不一致の程度が正しい聴覚認知と
関連するのかを確認する（研究②）。

2. 応答時間

＝情報処理の多寡があるかどうか調べるため、課題条件間および認知行動成績間での応答時間を比較する。

3. 課題条件と脳活動分布

＝マガーク条件（研究①）および視聴覚不一致のある聴覚情報（研究②）を処理する神経活動、関与する部位を明らかにする。マガーク条件と一致条件（研究①）、視聴覚不一致の有無や大きさ（研究②）と脳活動分布を比較する。

4. 認知行動成績と脳活動パターン

＝神経活動と認知行動成績との関連から情報処理の神経基盤を明らかにする。

5. 脳活動の時間（研究②）

＝視聴覚不一致の大きさによる脳活動潜時の違いを明らかにする。

また、聴覚課題により、視覚情報を有さない聴覚課題によっても十分に課題が成立する（認知成績が保たれる）ことを確かめた。

IV. 対象と方法

本研究は、視聴覚不一致を伴う聴覚情報処理の神経基盤を解明するため、認知行動成績と皮質脳波を記録した。対象は、東京大学医学部附属病院において難治性てんかんの治療を目的として頭蓋内電極を留置した患者を対象とした。頭蓋内電極留置は、MRI 上器質的病変が明らかでなく、術前の非侵襲的検査によっててんかん焦点の同定に至らなかった例に対して、てんかん焦点の同定を目的として行われた。電極を留置する部位や数に関しては、臨床症候や術前の非侵襲的検査、言語領域など臨床上の必要性に基づいて決定した。頭蓋内電極は焦点検索、脳機能マッピングに必要な期間（2～4 週間）留置し、これらの検査を行わない時間を利用して、本研究を目的とした皮質脳波計測を行った。頭蓋内電極を用いた脳電位記録による機能的神経回路の研究に関しては、東京大学医学系研究科倫理委員会審査で承認を受けており（#1797）、本研究に参加するすべての患者本人から事前に書面での同意を得た。本研究の皮質脳波計測は、電極留置術から 7 日以上経過し、最終のてんかん発作から 24 時間以上経過した後に行った。集中力が保てる状態のときに課題を行い、一回の計測には 30 分から 1 時間程度を要した。皮質脳波計測時には抗てんかん薬を服用していた。本研究の計測はいずれも東京大学医学部附属病院内の生理検査室（シールドルーム）にて行った。

A. 対象

2012年11月から2014年6月に、東京大学医学部附属病院において難治性てんかんの治療を目的として左大脳半球を広範に覆う頭蓋内電極を留置した連続9名の患者を対象とした。このうち1名は、てんかん焦点を同定するための脳波計測中にてんかん重積状態となり、発作後意識障害の遷延がみられたために研究を目的とした脳波計測は行わず、本研究から除外した。結果として8名（男性4名、女性4名）についての解析を行った。この8名の年齢は20歳から43歳（平均28.8歳）であり、Wechsler Adult Intelligence Scale-III (WAIS-III)(Wechsler, 1981)における verbal intelligence quotient score は全員60以上であった。8名全員が右利きであり、てんかんを除いて神経学的な疾患の既往はなく、視覚および聴覚機能にも異常を認めなかった。また、臨床上の目的で行ったfMRIおよびWada testによって、言語優位半球は8人全員が左半球と同定された。表1に解析対象となった8名の患者の詳細を示した。

表 1. 対象患者情報詳細

被験者	年齢 /性別	てんかん焦点	言語優位半球	言語性 IQ	電極数
1	30/M	左上側頭回	左	104	179
2	24/F	左内側側頭葉	左	65	179
3	20/M	左後頭葉	左	71	179
4	25/M	左内側側頭葉	左	99	164
5	43/M	左内側側頭葉・上側頭回	左	90	174
6	25/F	両側多焦点	左	64	180
7	26/F	左内側側頭葉	左	92	180
8	37/F	両側側頭葉	左	73	173

B. 方法

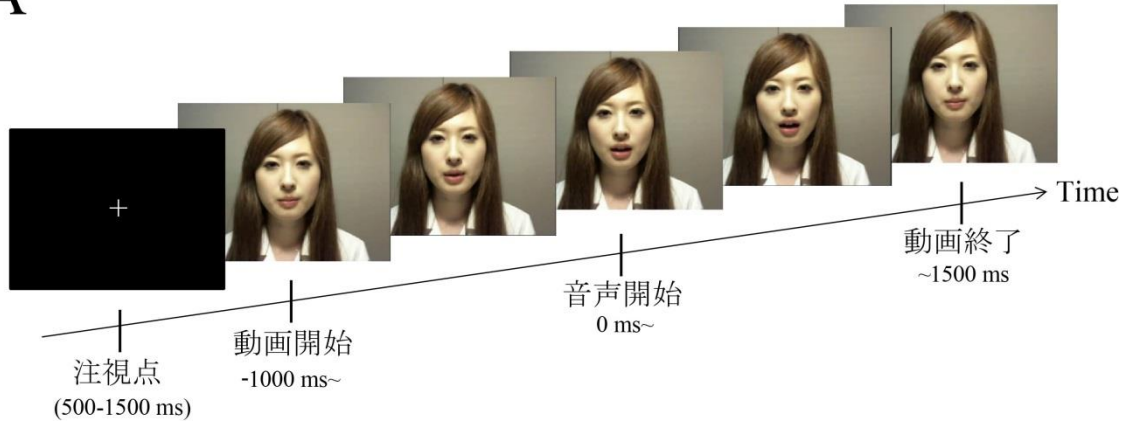
研究① マガーク効果（融合認知）の脳活動

1. 課題

言語に関連した視聴覚動画を提示した際に、融合認知が起こる時と正しい聴覚認知ができる時の脳活動を比較検討するため、視聴覚課題を提示した。提示課題はプログラミング言語である MATLAB (Mathworks, Inc., アメリカ) を用いて作成した。課題では、それぞれが 2500 ms からなる 6 種類の動画クリップがランダムな順番で提示される構成とした。動画クリップは、日本人女性が発声する際の、口唇の動き (lip-movement) と音声との組み合わせで作られた。2500 ms の動画クリップはすべて、音声開始前 1000 ms から後 1500 ms までを切り出したものであり、動画と音声を組み替えたクリップにおいても音声開始のタイミングは同様とした。具体的な方法として、例えば話者が『か』の口唇動画に『ぱ』の音声を重ね合わせる場合であれば、最初に音声データをもとに『か』の音が立ち上がるタイミングを分析した後、別途録音しておいた『ぱ』の音声をそのタイミングに合わせることで、『か』の音声が立ち上がるタイミングで『ぱ』の音声が立ち上がるようにした。そのため、発声のための口唇の動き始めは、発声内容のタイプによってそれぞれ音声開始の 270 ms から 570 ms 前と異なっていた。動画クリップ開始前には 1000 ms \pm 500 ms (ランダムジッタ) の注視点を

持つ画面を提示した（図 5A）。動画クリップは具体的には、口唇の動きと音声の一致している動画（一致条件）が 3 種類、マガーク条件が 1 種類、口を閉じたまま口唇が動かない動画（閉口条件）が 1 種類の合計 5 種類であった。一致条件は、『ぱ』『か』『た』のいずれかを発声している動画であり、口唇の動きと音声は一致していた。マガーク条件では、口唇の動きは『か』、音声は『ぱ』が提示された。閉口条件では同じく音声『ぱ』が提示されたが、口唇は閉じたままで動かないものであった（図 5B）。閉口条件の動画クリップは、被験者が口唇の動きを注視し続けるようにするためのものであったため、この条件の結果はすべての解析から除外した。

A



B

視聴覚課題

	Lip movement	Voice
一致 条件	ば か た	ば か た
マガーク 条件	か	ば
閉口 条件	閉口	ば

図 5. 研究①の視聴覚課題

A. 被験者に、日本人女性の話者が発声する動画クリップを提示した。動画クリップは、発声時の映像（『ば』、『か』、『た』、口を動かさない）と音声（『ば』、『か』、『た』）との組み合わせからなる。動画クリップは、500 ms から 1500 ms 提示される注視点の後にランダムに提示される。いずれの動画クリップも 2500 ms であり、1000 ms の時点で音声が始まる。B. 各条件の動画と音声の組み合わせを示す。マガーク条件は口唇の動き『か』、音声『ば』が提示された。

それぞれの動画クリップを視聴した直後に、被験者は聴取した音を回答した。回答は押しボタンを用いて、『ば』、『か』、『た』、『閉口』の中から選択した。閉口条件が提示された際には音声を回答するのではなく、『閉口』を選択することとした。提示した動画クリップに、口唇の動きと音声異なる組み合わせが含まれることは事前に被験者には教示しなかった。

一致条件、マガーク条件、閉口条件は、をそれぞれ約 53%、35%、12%の頻度で提示された。課題は、それぞれの被験者の体調、疲労に合わせて 56 から 224 回の動画クリップ提示を行った。動画クリップの音声開始を 0 秒として、回答までの時間を計測した。

動画クリップは 640×480 ピクセル、フレームレートは 30 で提示した。音声 (44.1 kHz、16 ビット) は、約 75dB の音圧でカナル式イヤホンを通じて提示された。2500 ms の動画クリップの中で、音声はすべての動画クリップにおいて動画開始後 1000 ms で提示されるが、発声のための口唇の動き始めは、発声内容のタイプによってそれぞれ音声開始の 270 ms から 570 ms 前と異なっていた。動画クリップ開始前には 1000 ms ± 500 ms (ランダムジッタ) の注視点を持つ画面を提示した。被験者はシールドルーム内の椅子に座わり、約 70 cm 前方のノート型 PC 画面に提示される動画クリップを視聴した。

2. 認知行動成績の統計解析

マガーク条件においては、『た』と回答したものを融合認知、『ぱ』と回答したものを正しい聴覚認知（正答）とした。融合認知の生起率、正答率を求め比較した（特に記載のない場合、両側検定で有意水準は $\alpha=0.05$ とした）。

3. 応答時間

課題に対する応答時間を、動画クリップの音声開始から回答までの時間で比較した。被験者ごとに応答時間を算出し、一致条件内での情報処理の多寡に違いがあるか否かを見るために、一致条件内における『ぱ』、『か』、『た』での応答時間を対応のある一元配置分散分析で比較した。次に条件間の情報処理の多寡に違いがあるか否かを見るために、一致条件とマガーク条件の間での応答時間を対応のある t 検定で比較した。また、認知行動成績の違いによって情報処理の多寡に違いがあるか否かを明らかにするため、マガーク条件における融合認知時と非融合認知時、正答時（正しい聴覚認知時）と誤答時の間においても同様の手順で比較を行った。応答時間は、被験者ごとに平均値から 2 標準偏差以上を逸脱した値は外れ値として応答時間の解析からは除外した。

4. 皮質脳波（ECoG）記録

視聴覚課題遂行時の脳活動を皮質脳波にて捉えた。言語に関連する視聴覚情報処理には主に左大脳半球が関与すると知られていることから^{21,23,31}、左大脳半球外側に留置した頭蓋内電極に着目して評価を行った。対象とした電極は、シリコンシートに格子状に配置された60極のプラチナ電極であり、電極の大きさは直径1.5mm、電極中心間距離は5mmの高密度電極であった（ユニークメディカル、日本）（図6）。

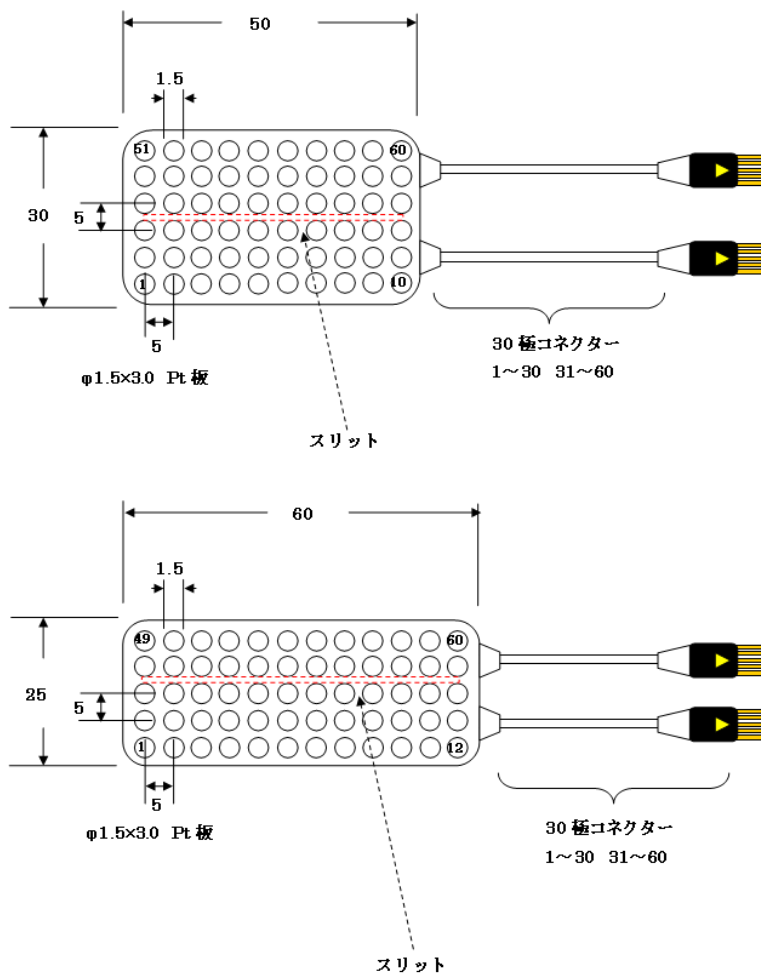


図 6. 留置電極シート

電極シートは、直径 1.5mm の電極が中心間距離 5mm で 60 極並ぶものを使用。本研究では、6×10（上）または 5×12（下）の電極シートが合計 3 枚留置されている患者を対象とした。

電極位置は、頭蓋内電極留置術前に撮影した MRI 画像に、術後撮影した頭蓋内電極の位置情報を持つ CT 画像を Dr. view/Linux の mutual information 法を用いて合成することで調整した（旭化成、日本）¹。3 次元脳表画像の再構築にはサイバーネット社の Real INTAGE を用いた（サイバーネット社、日本）。皮質脳波は

多チャンネル脳波計（EEG1200、日本光電、日本）を用いて、サンプリング周波数 2000 Hz、アナログフィルター0.08～600 Hz で記録した。皮質脳波には、動画クリップの開始と同時にトリガーとして 1 ミリ秒の矩形波も記録された。また、提示する音声もアナログ情報として直接脳波データとともに記録した。基準電極は、右頭頂葉近傍の硬膜に内側から設置した電極を用いた。

5. 皮質脳波 (ECoG) 解析

皮質脳波データの解析はMATLAB上で作成したプログラムを用いて行った。視認によって、顕著なノイズやてんかん性異常波を含むデータは除外した。結果として対象とした電極数は患者一人当たり平均で176(標準偏差;SD=5.2)極であった(表1)。

マガーク条件提示によって賦活される部位を検出するため、視聴覚動画の一致条件提示とマガーク条件提示における皮質脳波の変化を比較した。初めに、課題によって変化する各周波数成分の概要を捉えるため、MATLAB上で動作するEEGLAB toolbox⁴⁹のWavelet解析を用いて時間周波数パワー(event related spectral perturbation; ERSP)を評価した。

Wavelet解析とは周波数解析手法のひとつであり、広い周波数領域に対して時間周波数解析を行う上で適した解析である。周波数解析の代表的手法であるフーリエ変換では、一般に時間領域の情報が失われる。時間領域の情報を得るためにはフーリエ解析に窓関数を用いるが、窓フーリエ変換では窓幅は周波数に関わらず固定される。すると、高い周波数帯域ではある窓に対して何周期も観測データが得られるのに対して、低い周波数帯域では十分に長い時間の窓でないと1周期分のデータが得られない。例えば時間窓を500msとした場合、100Hzの活動は50周期の波が含まれるが、1Hzの活動は半周期の波しか含まれ

ないことになる。これに対して Wavelet 解析では、定義した Wavelet 関数 (Mother wavelet、本研究では Morlet Wavelet を使用、式 1) を拡大縮小、平行移動することで解析対象の波形に相似な波形を抽出して時間情報を残したままスケール情報を出力する。これにより、周波数に合わせて窓を変えることができる。つまり、高周波では周波数分解能を下げ、時間分解能を上げ、低周波では周波数分解能を向上させて時間分解能を低下させることで適切な解析結果が得られる (式 2)。

$$\Psi_{a,b}(x) = \frac{1}{\sqrt{a}} \Psi\left(\frac{x-b}{a}\right) \quad (1)$$

$\Psi_{a,b}(x)$: wavelet 関数

a : スケールパラメータ

b : シフトパラメータ

$$W(a,b) = \int_{-\infty}^{\infty} \frac{1}{\sqrt{a}} \overline{\Psi\left(\frac{x-b}{a}\right)} f(t) dt \quad (2)$$

$W(a,b)$: wavelet 係数

以上の Wavelet 解析を用いて、視聴覚動画提示に伴う脳活動の変化を計測した。音声の提示は動画クリップ開始後 1000 ms であったが、本研究は視聴覚課題の情報処理を研究の対象としているため、音声開始時を基準 (0 ms) とした。音声開始前 1350~1050 ms (画面中央に注視点のみ提示されている時間帯) を背

景脳活動として baseline に定め、音声開始前後 1500 ms の ERSP (5~200 Hz) を描出し、一致条件、マガーク条件を比較した。

次に、神経発火との深い関連が示されている High gamma 帯域の活動を評価した。High gamma 帯域の活動を抽出するため、まず、電極ごとに 70~150 Hz の範囲で帯域通過フィルター (有限インパルス応答フィルター finite impulse response filter; FIR filter) をかけた。つぎに、このフィルタリング後のデータにヒルベルト変換を行い絶対値を取ることで、包絡線 (Envelope) の計算即ちパワーの推定を行った⁵⁰ (図 7)。

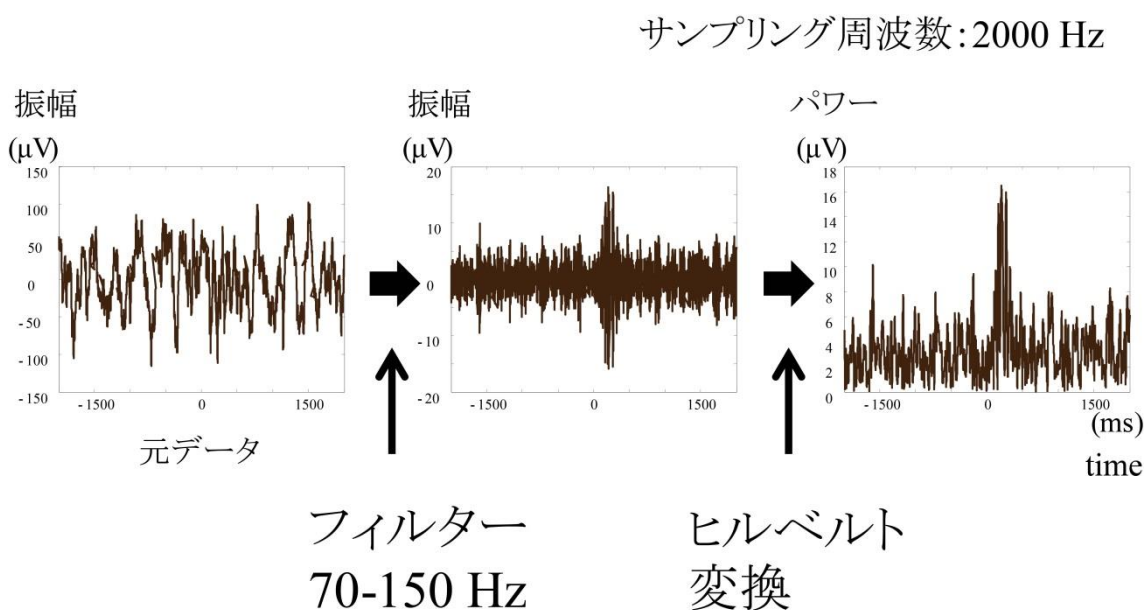


図 7. High gamma activity の定量化

本研究では、70~150 Hz の High gamma activity を定量化することで脳活動の評価を行った。70~150 Hz の帯域通過フィルターを用いた後、ヒルベルト変換を行って事象関連パワーを得た。

これにより、元のデータと同じ時間点をもつ行列として帯域パワーが得られ、事象関連パワーの解析が可能となる⁵¹。High gamma 帯域の事象関連パワーを図示するために、100 ms の boxcar kernel スムージングを行った⁵²。ここで用いたスムージングの手法は、ある観測点の値を、前後 50 ms の範囲での平均値に置き換える操作を全観測点（両端 50 ms を除く）に対して行うものである。活動時間解析を除き、図示の目的以外の統計解析ではスムージングを行っていないデータを用いた。baseline の平均パワーを基準として、そこからの変化率を算出した。

音声開始後 0～500 ms を early period、500～1000 ms を later period として解析対象期間を定義した。マガーク条件提示による神経活動を明らかにするため、視聴覚不一致の無い一致条件とマガーク条件について各 period における baseline からの HGA 平均上昇率の比較を行った。各被験者のそれぞれの電極について、一致条件よりもマガーク条件で HGA の平均上昇率が高くなったか否かを t 検定によって調べた（片側検定、 $p < 0.05$ ）。得られた p 値に対しては、電極数で false discovery rate 法（FDR 法）による多重比較の補正を行った。

多重比較の補正は、帰無仮説が正しいのにもかかわらず棄却してしまう“第一種の過誤”を許容しないために行う補正であるが、一方で帰無仮説が誤っているのにもかかわらず採択してしまう“第二種の過誤”が生じることが問題となる。

つまり、有意水準が厳しくなりすぎるため統計学的有意差が出にくい場合がでてくる。ここで用いた FDR 法は、棄却された仮説のうち誤って棄却された真の帰無仮説の割合 (false discovery rate) を制御する方法である。他の多重比較法よりも第二種の過誤が少なくなる。今回の検定では、N 個の検定対象電極で得られた p 値を小さい順に並べ $p_1 \leq p_2 \leq p_3 \leq \dots \leq p_N$ とし、

$$p_i \leq q \times i/N \quad (i=1,2, \dots, N)$$

を満たす最大の i を r とすると、 $p_1, p_2, p_3, \dots, p_r$ にあたる電極を有意差ありと判定した。q は本来期待する有意水準であり、本研究では 0.05 とした。

この統計の結果、音声開始後 0~500 ms の early period においては一致条件よりもマガーク条件で HGA 平均上昇率が高かった電極は、全被験者の全電極 (1408 極) のうち 4 極しかみられなかった。一方で音声開始後 500~1000 ms の later period では、全被験者で合計し、51 極認めた (詳細は『V. 結果』に示す)。このことから、この後の解析においては 500~1000 ms の later period のみを対象とした。

また、一致条件では 3 つの異なった音声を用いたが、マガーク条件との比較をするにあたり、音声の種類によらず一致条件内が一様であることを確かめるため、一致条件の中での 3 つの音声に対して HGA の変化を評価した。3 つの音声での baseline からの HGA 変化率を各電極で一元配置分散分析 (有意水準 $\alpha <$

0.05) を行った。電極数による FDR 法での多重比較補正を行ったところ⁵²、全被験者の全電極中で 7 つの電極に HGA 変化率に違いが見られた。本研究では、視聴覚の一致条件と比べてマガーク条件で特異的に起こる脳活動を検討の対象としているため、視聴覚一致条件内において異なる脳活動を示すこれらの電極は解析結果から除外した。また、これらの除外電極は IFG 上にも STS 上にも認めなかった。

6. 電極の選択

関心領域である IFG と STS の電極は次のように選択した。

IFG 電極については、電極面積の半分以上が IFG 上に位置している電極を IFG 電極とした。IFG は弁蓋部（ブロードマン 44 野）および三角部（ブロードマン 45 野）とした⁵³。

STS 電極については、STS 後方部位を STS として扱った³⁴。シルビウス裂と中心溝の交点から STS へ垂線を下した点から後方を STS とした。実際には電極は脳溝内には留置できないため、STS の外側表面において、STS から 1.5mm 以内の距離に中心が存在する電極を STS 電極とした。

この結果、1 人の被験者あたり、IFG 電極は平均 27.4 極 (SD=4.7)、STS 電極は平均 10.1 極 (SD=3.1) であった (表 2)。IFG や STS において有意な HGA 上昇を示した電極の割合が、非 IFG/STS 領域で有意な HGA 上昇を示した電極の占める割合よりも高いかどうかを、対応のある t 検定にて比較した (片側検定)。

7. HGA と認知行動成績の関連解析

マガーク条件提示における HGA と融合認知、正しい音声認知の関連を調べるため、認知行動成績によって分類した際の HGA を比較した。認知行動成績は融合認知（回答『た』）と非融合認知（回答『ば』、『か』）、正答（回答『ば』）と誤答（回答『か』、『た』）に分類した。本研究で対象とした被験者 8 例は、全体としてはほぼ先行研究と同様の認知傾向であった（詳細は『V. 結果』に示す）。しかしながら、融合認知の生起率は範囲が 10%～95%、標準偏差が 26.8 であり、8 名の中でも融合認知の起こりやすさに大きな個人差があった。この要因としては、本研究が頭蓋内電極留置の被験者を対象としており、多くの被験者を望むことが困難であったため、個人差が誤差として相殺されにくかったことによると考えられる。

この点については外れ値となるデータを除外する方法もあり得るが、本研究では皮質脳波のデータが希少であることと、このような個人差まで含めた上でヒトの認知および脳活動の全体的な傾向を捉えるという目的を考え、8 例全体を対象として解析を行った。融合認知に関連した HGA 上昇や正答に関連した HGA 上昇を調べるため、各被験者における IFG 電極の合計パワーを、融合認知時と非融合認知時（または正答時と誤答時）で対応のある t 検定にて比較した（片側検定）。STS においても同様の操作を行った。

次に、融合認知時または正答時に HGA が上昇した電極がどこに分布するの
かを調べるため、各電極について融合認知時（または正答時）に非融合認知時
（または誤答時）よりも HGA が有意に上昇する電極を t 検定（片側検定）にて
算出し、脳模式図にプロットした。ただし、認知行動成績に極端な傾向がある
場合にはどちらかの条件の試行数が少なくなりデータの信頼性が損なわれる可
能性があるため、認知に関連した電極の分布を求める本解析においてのみ、対
象を融合認知率が 90%以上の 2 名と 10%以下の 1 名を除いた 5 名として行った。
融合認知時に HGA が上昇する電極と正答時に HGA が上昇する電極を全被験者
から集め、IFG や STS への分布が異なるかどうかカイ 2 乗検定にて比較した⁵⁴。

研究② 視聴覚不一致を伴う聴覚情報処理

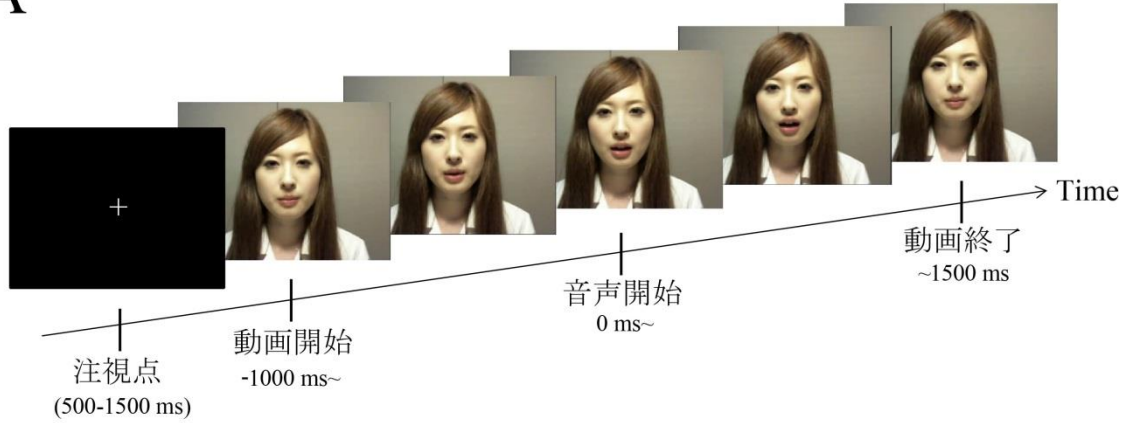
1. 課題

1-1. 視聴覚課題

視覚情報と聴覚情報に不一致があるときの脳活動を調べるため、視聴覚課題を提示した。提示課題は研究①と同様プログラミング言語である MATLAB (Mathworks, Inc., アメリカ) を用いて作成した。課題は、研究①の条件に非マガーク条件の視聴覚不一致条件（口唇の動き『す』、音声『ぱ』）を1種類加え、6種類の動画クリップがランダムな順番で提示される構成とした。本研究は、視聴覚不一致における正しい聴覚認知に関連する脳活動に着目したものであり、この観点では、マガーク条件も視聴覚不一致のある動画（不一致条件）に分類される。視聴覚不一致条件2種での脳活動の比較を行った。不一致条件においては、音声は常に『ぱ』が提示されたが、音声との不一致が小さい口唇の動き『か』が提示される不一致（小）条件（マガーク条件）と、不一致が大きい口唇の動き『す』が提示される不一致（大）条件からなっていた。『ぱ』と『か』はともに破裂子音と非円唇母音からなっているが、『す』は摩擦子音と円唇母音からなる^{55,56}。音声上の違いも、不一致（小）条件は視聴覚入力の差異が子音だけ（『ka』と『pa』）、不一致（大）条件は視聴覚入力の差異が子音と母音両方（『su』と『pa』）である。したがって、口唇の動きから予期される音節と実際に呈示さ

れる音声『ば』との差は不一致（大）条件の方が大きいと考えられた。（図 8）。

A



B

視聴覚課題

	Lip movement	Voice
一致 条件	ば か た	ば か た
不一致（小） 条件	か	ば
不一致（大） 条件	す	ば
閉口 条件	閉口	ば

聴覚課題

Auditory - only 条件	-	ば、か、た
--------------------	---	-------

図 8. 研究②の課題

各課題の動画と音声の組み合わせを示す。不一致な組み合わせは、不一致（小）条件は口の動き『か』音声『ば』であり、不一致（大）条件は口の動き『す』、音声『ば』とした。聴覚課題では口唇の動きは隠して提示された。

一致条件、不一致（小）条件、不一致（大）条件、閉口条件は、をそれぞれ約 45%、30%、15%、10%の頻度で提示された。課題は、それぞれの被験者の体調、疲労に合わせて 66 から 264 回の動画クリップ提示を行った。その他の課

題、計測環境は研究①と同様であった。

1-2. 聴覚課題

聴覚課題では本実験で用いた音声は視覚情報（口唇の動き）の無い状態でも正しく認知できるか否かを調べるため、視聴覚課題と同じ 2500 ms の動画クリップから、話者の口元を隠して口唇の動きによる情報を削除したものを提示した（図 9）。動画クリップの種類は音声『ば』、『か』、『た』の 3 種からなった。これらをランダムに提示し、視聴覚課題と同様の視聴環境、回答方法で行った。それぞれの被験者で 30 回または 60 回の動画クリップ提示を行った。提示課題は視聴覚課題と同様にプログラミング言語である MATLAB を用いて作成した。



図 9. 聴覚課題の動画

聴覚課題では、被験者に話者の口元を隠した動画を提示した。

2. 認知行動成績の統計解析

2-1. 視聴覚課題

視聴覚課題においては閉口条件を除き、口唇の動きの内容にかかわらず、提示された音声と一致する音節（音素）を回答した場合に正答とした。つまり、本研究では不一致（小）条件、不一致（大）条件とも『ぱ』の回答を正答とした。話者の口唇の動きが被験者の音声認知に与える影響を調べるために、視聴覚動画の一致条件提示と不一致（小）条件提示における正答率を比較した。また、不一致条件の中でも不一致（小）条件提示と不一致（大）条件提示における正答率も比較した。被験者ごとの正答率を対応のある t 検定で比較した。

2-2. 聴覚課題

視聴覚課題で用いた聴覚情報が、聴覚情報単独でも正しく認知される音声であることを確かめるため、聴覚課題の正答率と一致条件の正答率を被験者ごとに算出し、正答率を対応のある t 検定で比較した。また、視覚情報によって聴覚認知が妨害されることを確かめるため、聴覚課題の正答率と不一致条件の正答率を比較した。

3. 応答時間

被験者ごとに応答時間を算出し、条件間の情報処理の多寡に違いがあるか否かを見るために、一致条件と不一致条件の間、不一致（小）条件と不一致（大）条件の間について、対応のある t 検定で比較した。また、認知行動成績の違いによって情報処理の多寡に違いがあるか否かを明らかにするため、不一致（小）条件における正答時と誤答時の間、および不一致（大）条件における正答時と誤答時の間についても同様の手順で比較を行った。応答時間は、被験者ごとに平均値から 2 標準偏差以上を逸脱した値は外れ値として除外した。

4. 皮質脳波 (ECoG) 記録

研究①と同様の ECoG 記録環境にてデータを取得した。

5. 皮質脳波 (ECoG) 解析

皮質脳波データの解析は研究①同様 MATLAB 上で作成したプログラムを用いて行った。視聴覚入力の不一致を検出するための神経基盤を明らかにするため、各被験者のそれぞれの電極について、一致条件よりも不一致条件で HGA の平均上昇率が高くなった電極を t 検定にて検出した (片側検定、 $p < 0.05$ 、FDR にて補正)。検定、除外電極などの選定はすべて研究①と同様とした。

6. 電極の選択

研究①と同様の方法を用いた。

7. HGA と認知行動成績の関連解析

不一致条件提示における HGA と正しい音声認知の関連を調べるため、認知行動成績によって分類した際の HGA を比較した。認知行動成績は正答と誤答に分類した。正答に関連した HGA 上昇を調べるため、各被験者における IFG 電極の合計パワーおよび STS 電極の合計パワーを、正答時に誤答時よりも活動が大きいかどうか、対応のある t 検定にて比較した（片側検定）。

さらに、全被験者から IFG 電極と STS 電極を集め、電極ごとに正答時と誤答時の HGA 平均変化率（baseline から 500～1000 ms）を算出した。正答時と誤答時の HGA を、全 IFG 電極と全 STS 電極のそれぞれで、対応のある t 検定を用いて比較した⁵⁴。

8. 皮質の活動時間の解析

視聴覚不一致の提示によって起こる脳活動の活動潜時を明らかにするため、提示課題の条件による前頭葉と側頭葉の活動時間を比較した。5. 皮質脳波 (ECoG) 解析の結果に基づき、一致条件と不一致条件で HGA に有意差のあった電極で評価した。各電極において、音声提示後に、条件間で baseline からの HGA 変化率の差が最大となる時間 (ピーク潜時) を求めた。各条件における前頭葉と側頭葉の電極を全被験者から集め、前頭葉のピーク潜時と側頭葉のピーク潜時を t 検定で比較した。

V. 結果

研究① マガーク効果（融合認知）の脳活動

A. 認知行動成績

マガーク条件の課題提示により融合認知が起こることを確認するため、認知行動成績を計測した。融合認知の生起率は 50.3% (SE=9.5%) であった (表 2) (図 10)。

表 2. 被験者結果詳細

被験者	IFG 電極数	STS 電極数	融合認知 (%)	正答率 不一致(小)条件 (%)	正答率 不一致(大)条件 (%)
1	34	14	10	68	90
2	20	9	35	40	70
3	33	15	36.3	64	80
4	29	10	40	3.3	20
5	26	10	46	50	60
6	21	11	50	3	5
7	29	6	90	10	10
8	27	6	95	5	10
平均	27.4 (SD=4.7)	10.1 (SD=3.1)	50.3 (SE=9.5)	30.3 (SE=9.3)	43.1 (SE=11.7)

IFG = inferior frontal gyrus, STS = superior temporal sulcus, SD = standard deviation, SE=standard error

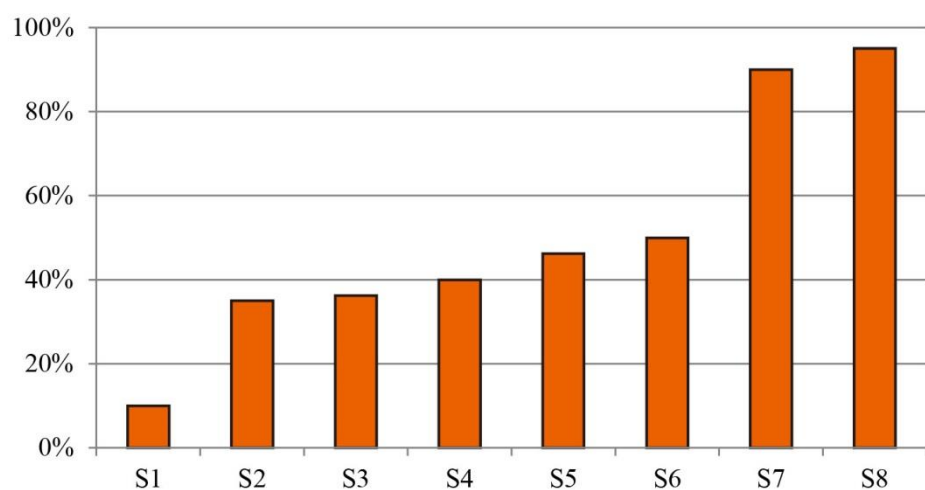
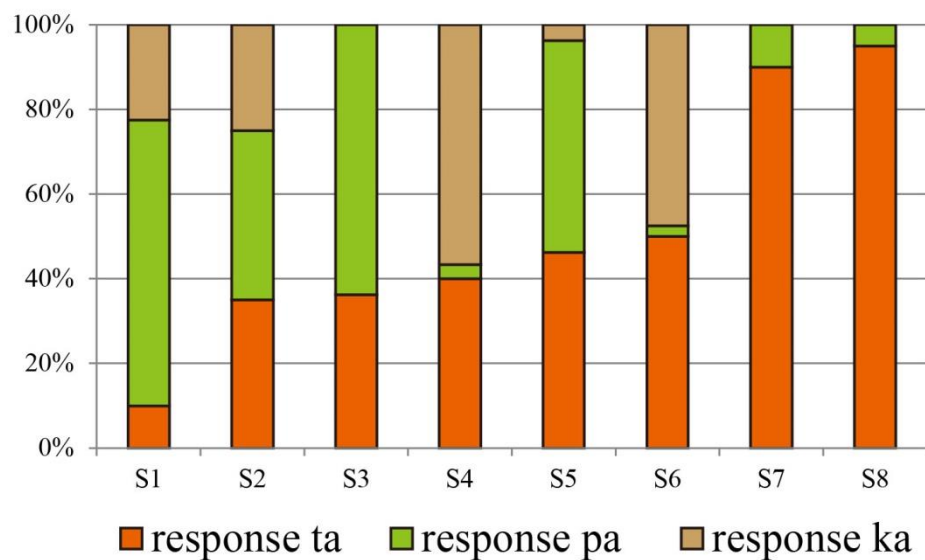


図 10. 認知行動成績

頭蓋内電極を留置した被験者 8 人の、マガーク条件提示における認知行動成績を示す。融合認知の生起率は 50.3% (SE=9.5%) であった。

B. 応答時間

心的情報処理の多寡を測る指標の一つとして、音声開始から回答までの応答時間を計測した。課題の条件間で差があるか否か、認知行動成績の違いで（融合認知と非融合認知、正答と誤答）差があるか否かを比較した。まず、一致条件において、『ば』、『か』、『た』の提示の間に応答時間の差を認めなかった（【 $F(2,21)=3.47, p=0.88$ 】、一元配置分散分析）。条件間の比較では、一致条件（1.94秒）と不一致条件（2.18秒）の間に有意な差を認めた（ $P=0.01$ 、両側検定）。一方で、マガーク条件においては認知行動成績間に有意差はみられなかった（融合認知 $p=0.15$ 、正答 $p=0.28$ 、両側検定）。

C. Event related spectral perturbation (ERSP)

課題提示による皮質脳波の変化を明らかにするため、動画クリップ提示時の ERSP を描画した。マガーク条件提示時、一致条件提示時どちらにおいても、主として High gamma 帯域の活動が上昇した。また、特にマガーク条件によって強く活動する電極がみられた（図 11-13）。High gamma 帯域の活動（70～150 Hz）を対象に、課題による時系列変化を図示した（図 14）。

動画と音声一致

Patient 4
IFGの1電極

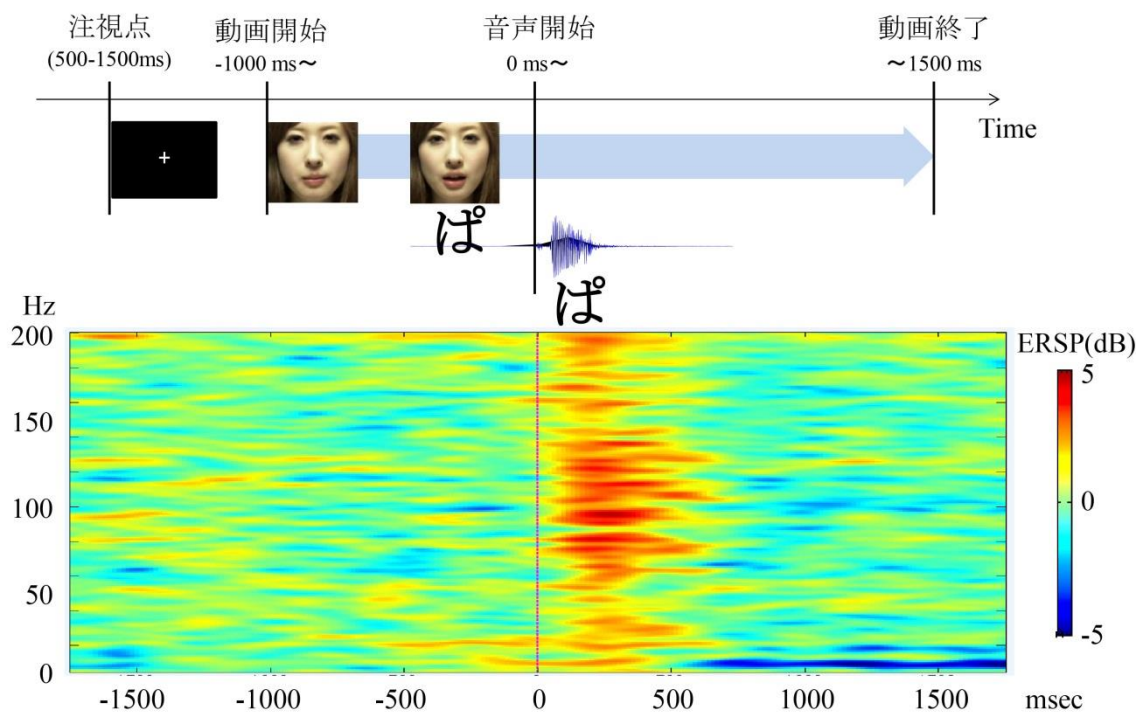


図 11. 課題と ERSP (被験者 4 の IFG 電極)

一致条件の提示をした時の IFG 電極 1 つの活動を示す。音声提示前 1350~1050 ms を baseline としたパワーの加算平均を示す。

動画と音声不一致

Patient 4
IFGの1電極

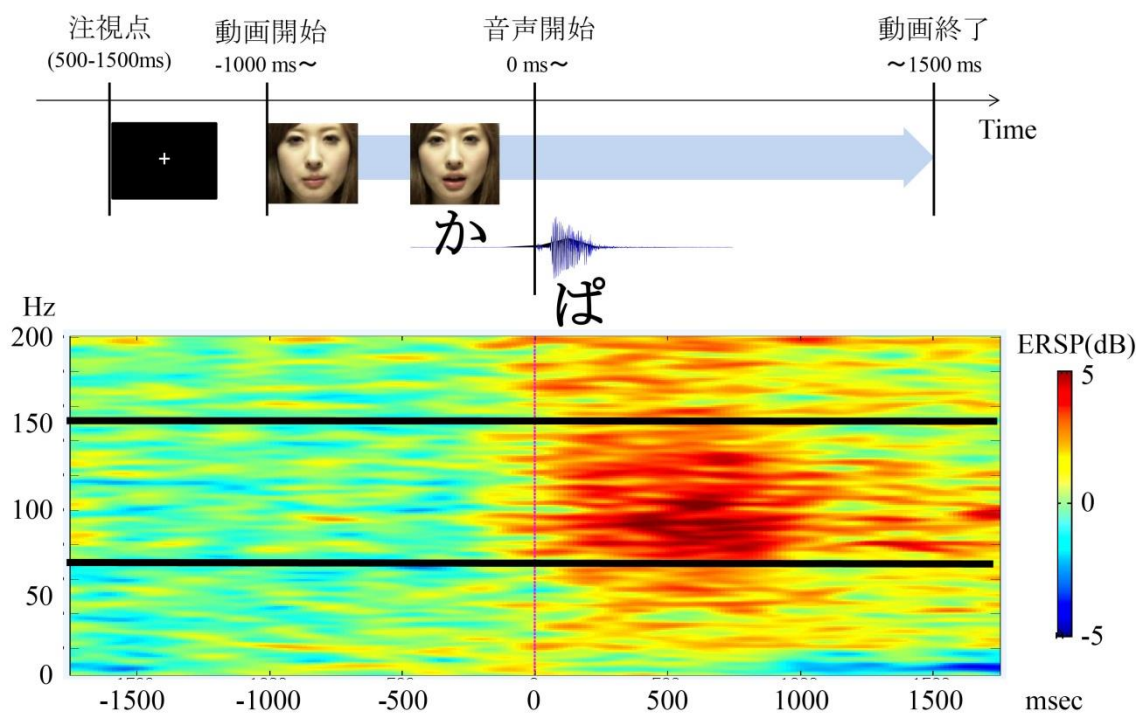


図 12. 課題と ERSP (被験者 4 の IFG 電極)

図 11 と同一の電極。マーカー条件の提示をした時の活動を示す。音声提示前 1350 ~1050 ms を baseline としたパワーの加算平均を示す。

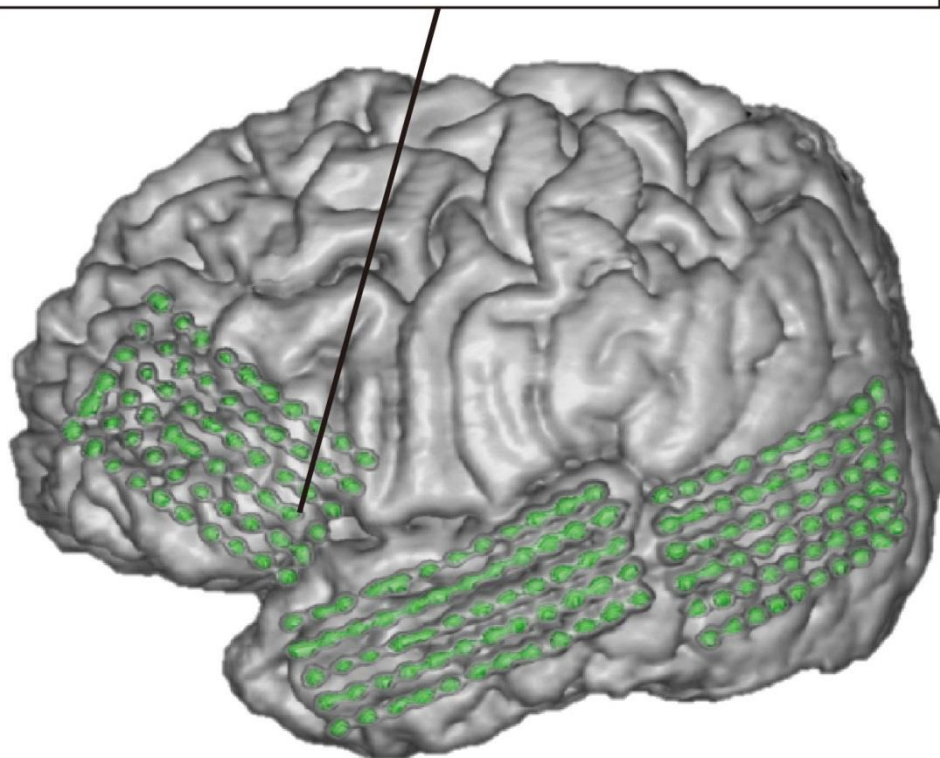
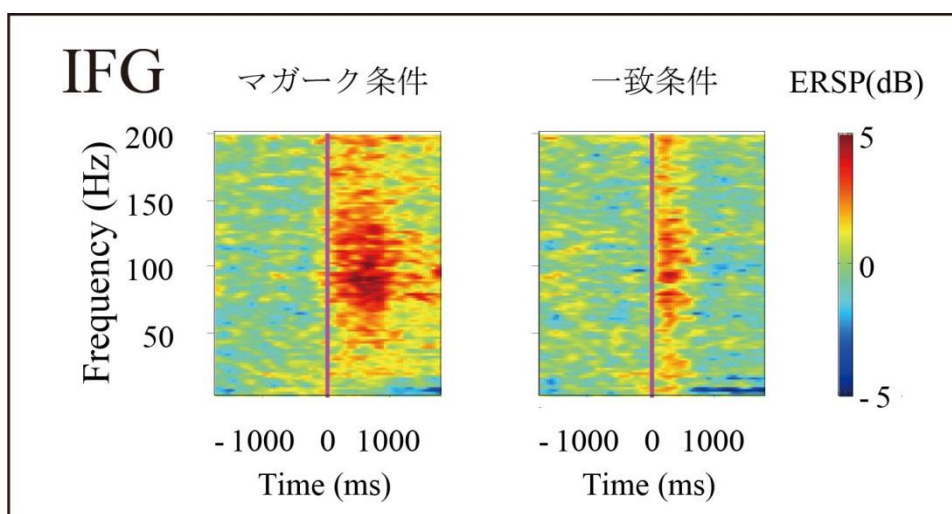


図 13. 時間周波数マップ (被験者 4 の IFG 電極)

IFG 電極における ERSP を示す。課題提示により、特に High gamma 帯域 (70~150 Hz) の脳活動が上昇する。また、視聴覚一致課題に比べて、不一致な課題でより強い活動上昇がみられる。

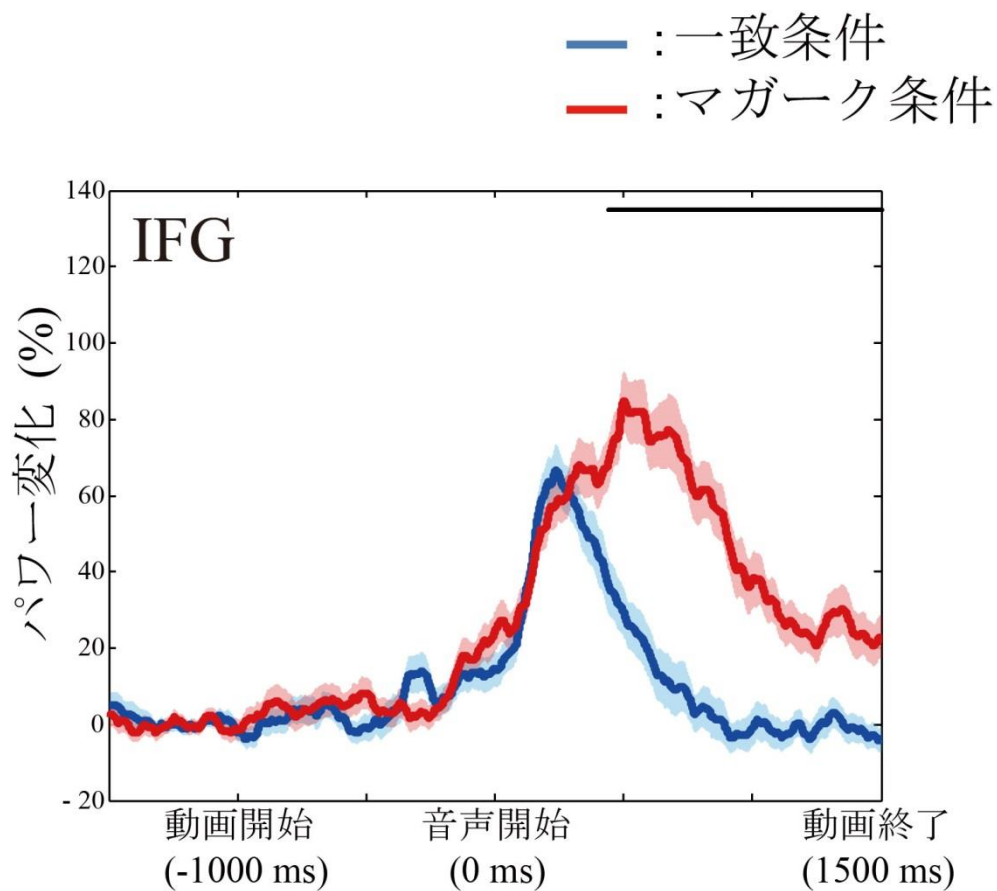


図 14. High gamma activity (HGA) の時系列変化 (被験者 4 の IFG 電極)

baseline 期間 (音声開始前 1350~1050 ms) からの HGA 変化率 (%) を示す。青の線は一致条件、赤の線はマガーク条件の平均変化率であり、色付きの部分は標準誤差 (一致条件 $n=90$ 、マガーク条件 $n=60$) を示す。上部にある黒い線は、マガーク条件が一致条件に比べて有意に高い HGA 上昇を示した期間を示す ($p < 0.001$)。

D. 事象関連 HGA

マガーク条件に対する脳内の処理機構を調べるため、マガーク条件と視聴覚不一致のない一致条件における HGA の違いを比較した。マガーク条件において HGA が上昇した電極位置を脳模式図に描画すると、IFG および STS 周辺に集積する傾向がみられた。また、一致条件に比べてマガーク条件で上昇した HGA は、音声提示後 500～1000 ms において主に観察された (図 15)。

マガーク条件 > 一致条件

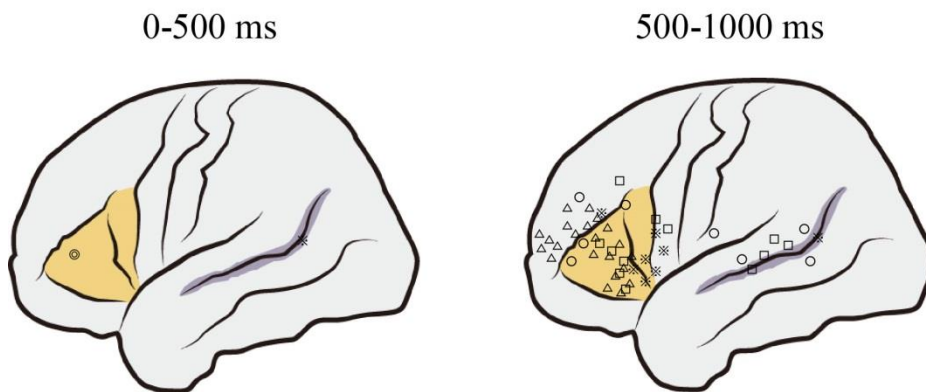


図 15. High gamma activity (HGA) 上昇電極の分布図

一致条件に比べてマガーク条件で有意な HGA 上昇を示した電極の分布を脳模式図に示した。baseline (音声提示前 1350～1050 ms) から 0～500 ms の early period に HGA 変化を比較したものを左、baseline から 500～1000 ms の later period に HGA 変化を比較したものを右に示した。解析対象電極全 1408 極のうち、early period では 4 極のみ、later period では 51 極の HGA 上昇電極を認めた。多重比較の補正は FDR 法による。電極の記号の違いは被験者の違いを示す。

音声提示後 0～500 ms の early period においては、一致条件よりもマガーク条件で HGA 上昇の大きい電極はほとんどみられなかった。HGA の有意な上昇

がみられた電極の具体的な割合は、IFG 電極の中で 0.43% (SE=0.40%)、STS 電極の中で 1.39% (SE=1.30%)、その他の部位（非 IFG、非 STS）で 0.18% (SE=0.16%) であった。

音声提示後 500~1000 ms の later period においては、一致条件よりもマガーク条件で HGA が上昇する電極の割合は、IFG 電極の中で 9.59% (SE=4.38%)、STS 電極の中で 3.17% (SE=1.96%)、その他の部位（非 IFG、非 STS）で 2.33% (SE=0.96%) であった。しかし、HGA 上昇電極は 8 例全例にみられたわけではなく、4 例においてのみ観察された。そしてこの 4 例は、融合認知の起こりにくい順に被験者から順に並べた 4 例と一致した。

HGA 上昇電極の分布は、その他の部位の電極分布と比較して IFG において有意に集積した ($p=0.046$ 、片側検定) (図 18)。

N=8
* : p < 0.05

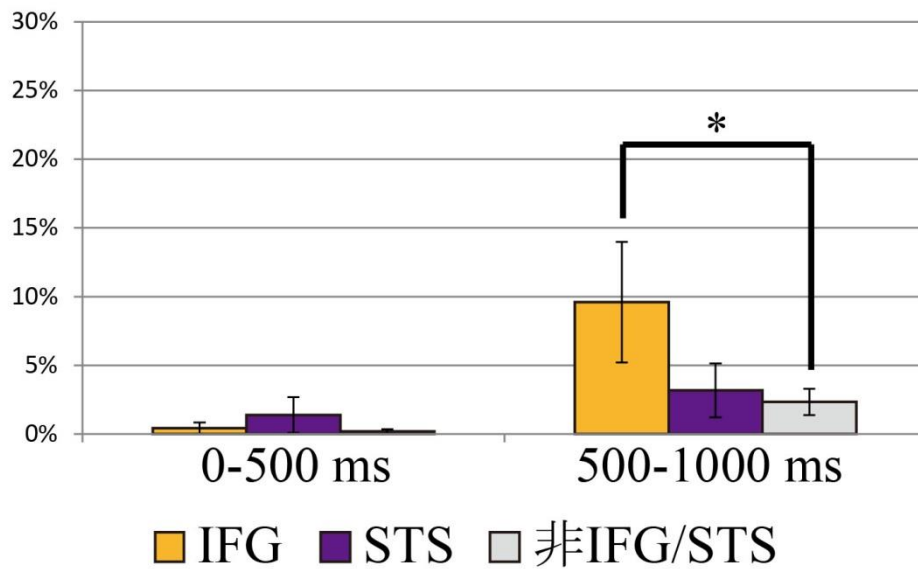


図 16. 領域に占める High gamma activity (HGA) 上昇電極の割合

一致条件に比べてマガーク条件で有意な HGA 上昇を示した電極の下前頭回 (inferior frontal gyrus; IFG ; 橙)、上側頭溝 (superior temporal sulcus; STS ; 紫)、その他の領域 (The other regions ; 灰) に占める割合を示した。later period (音声提示後 500~1000 ms)において、IFG に HGA 上昇電極が集積する (p=0.046)。エラーバーは標準誤差 (n=8) を示す。

E. HGA と認知行動成績

ここまではマガーク条件と一致条件の比較（提示条件の違いによる比較）による活動部位の検索であった。次に、マガーク条件提示時の脳活動を認知行動成績に基づいて分類、解析した。

● モデル1に基づいた検討

融合認知の際の脳活動を捉えるため、融合認知時と非融合認知時の試行で分類し、各被験者における IFG、STS の HGA を対応のある t 検定で比較した。IFG は融合時 1.08 (SE=0.02)、非融合時 1.11 (SE=0.03)、STS は融合時 1.05 (SE=0.02)、非融合時 1.02 (SE=0.04) であった。IFG、STS どちらにおいても融合認知時の有意なパワーの上昇は認めなかった (IFG: $p=0.41$ 、STS: $p=0.29$) (図 17)。

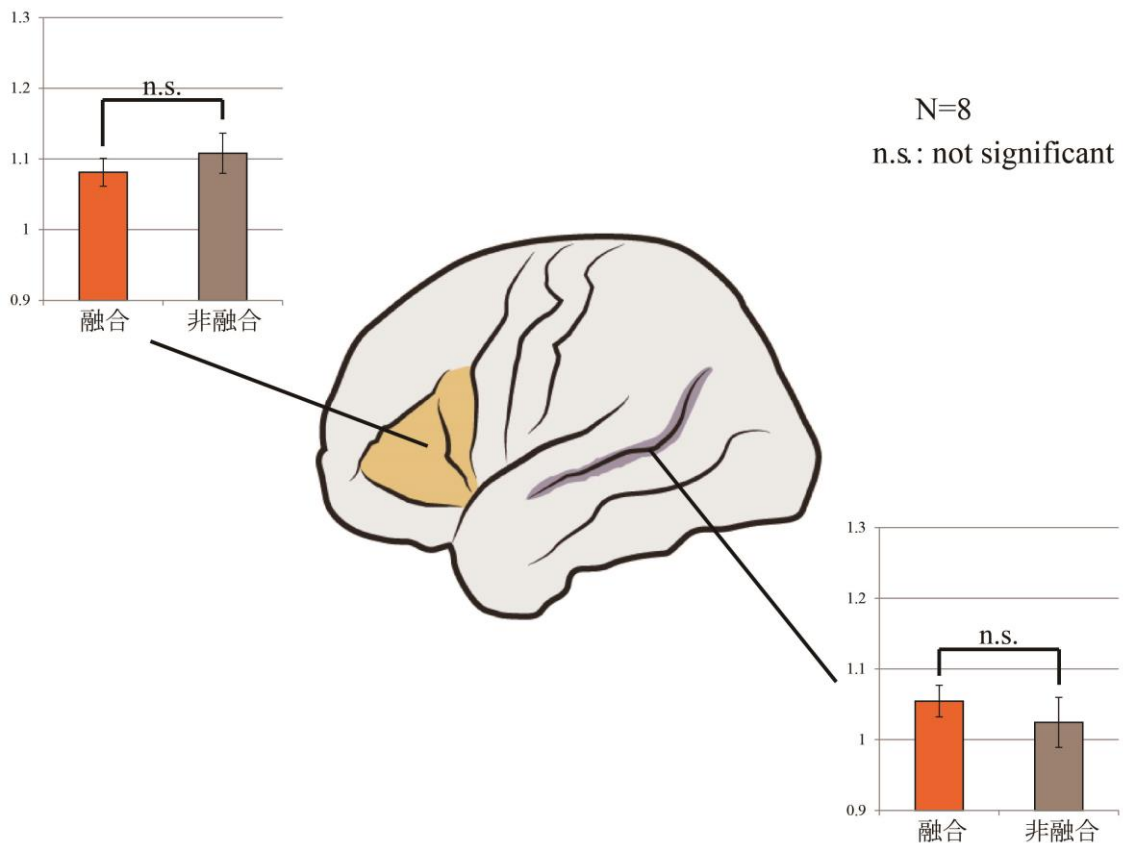


図 17. 融合認知時と非融合認知時の HGA

ママーク条件を提示した際の脳活動を融合認知の時と非融合認知の時に分け、8人の被験者の平均を比較した。IFG は融合時 1.08 (SE=0.02)、非融合時 1.11 (SE=0.03)、STS は融合時 1.05 (SE=0.02)、非融合時 1.02 (SE=0.04) であった。融合認知時の有意な活動はみられなかった (IFG: $p=0.41$ 、STS: $p=0.29$)。エラーバーは標準誤差を示す。

次に、融合時に HGA が上昇する電極が IFG や STS に集積するかどうか調べるため、融合認知で HGA が上昇する電極を脳模式図にプロットした(図 18)。

融合 > 非融合

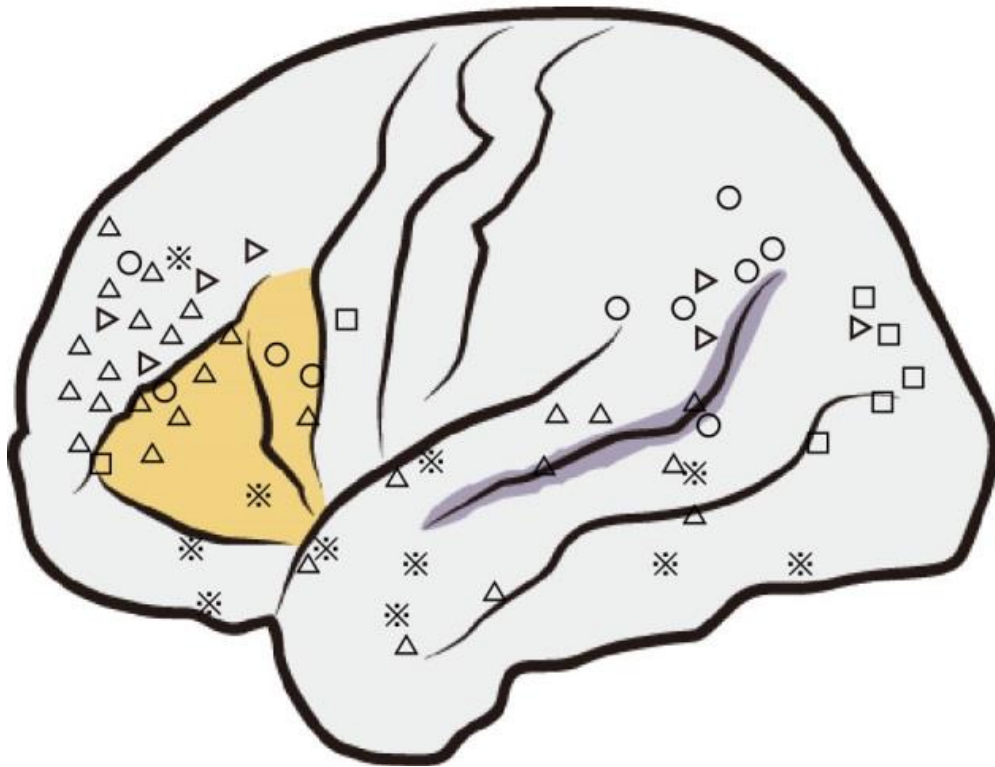


図 18. 融合認知時に HGA が上昇した電極

融合認知時に HGA が上昇した電極を示す。このような電極は、IFG の 7.75% (11/142)、STS の 5.2% (3/58)、非 IFG/STS の 7.1% (48/675) にみられた。IFG や STS への有意な集積はみられなかった (IFG: $p=0.79$ 、STS: $p=0.58$)。

融合認知時に HGA の上昇した電極は、IFG の 7.75% (11/142)、STS の 5.2% (3/58)、非 IFG/STS の 7.1% (48/675) にみられた。このような電極の占める割合を、IFG および STS と非 IFG/STS とでカイ 2 乗検定で比較したところ、融合時の HGA 上昇電極は IFG や STS へ有意な集積を示さなかった (IFG: $p=0.79$ 、STS: $p=0.58$)。

- モデル 2 に基づいた検討

正しい聴覚認知（正答）の際の脳活動を捉えるため、正答時と誤答時の HGA を、各被験者における IFG、STS の各部位のパワーを対応のある t 検定で比較した。IFG は正答時 1.20 (SE=0.05)、誤答時 1.07 (SE=0.02)、STS は正答時 1.04 (SE=0.04)、誤答時 1.05 (SE=0.02) であった。IFG においては、正答時の HGA 上昇が誤答時に比べて有意に大きかった (IFG: $p=0.036$)。一方で、STS においては正答時と誤答時で HGA に差を認めなかった ($p=0.63$) (図 19)。

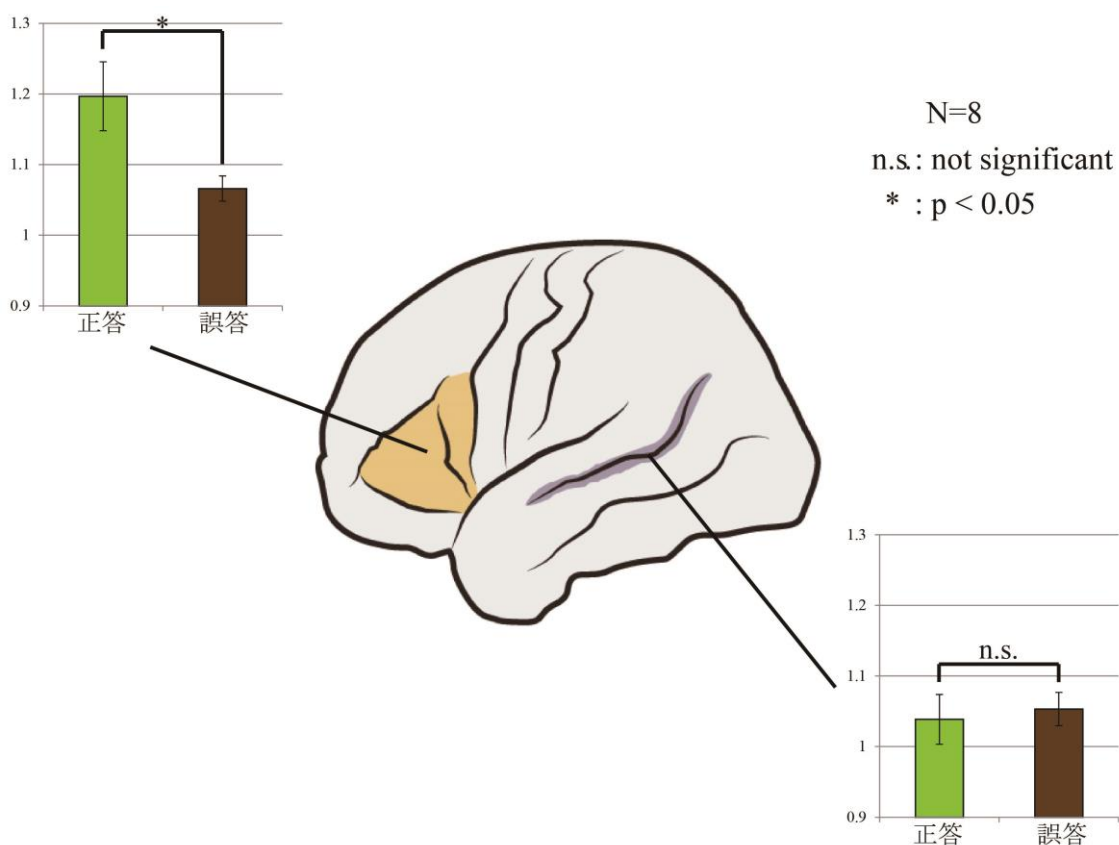


図 19. 正答時と誤答時の HGA

ママーク条件を提示した際の脳活動を正答時と誤答時に分け、8人の被験者の平均を比較した。IFGは正答時1.20 (SE=0.05)、誤答時1.07 (SE=0.02)、STSは正答時1.04 (SE=0.04)、誤答時1.05 (SE=0.02)であった。IFGにおいては正答時のHGAが誤答時に比べて有意に大きかった ($p=0.036$)。STSでは差を認めなかった ($p=0.63$) エラーバーは標準誤差を示す。

次に、正答時にHGAが上昇する電極がIFGやSTSに集積するかどうか調べる

ため、正答時にHGAが上昇する電極を脳模式図にプロットした (図 20)。正答

時にHGAが上昇した電極は、IFGの19.0% (27/142)、STSの5.2% (3/58)、非

IFG/STSの10.5% (71/675)にみられた。IFGやSTSに占めるこれらの電極の割

合を、非IFG/STSの電極割合とカイ2乗検定にて比較すると、IFGで有意な集

積がみられた ($p=0.0046$)。一方、STS へは有意な集積はみられなかった ($p=0.20$)。

正答 > 誤答

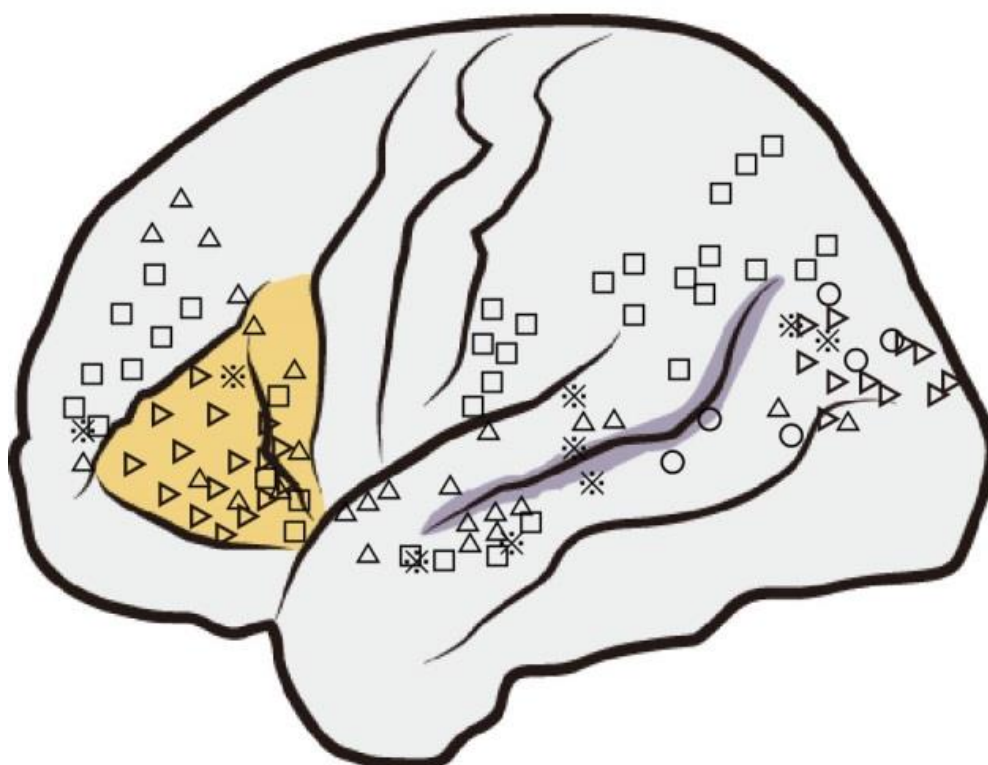


図 20. 正答時に HGA が上昇した電極

正答時に HGA が上昇した電極を示す。このような電極は、IFG の 19.0% (27/142)、STS の 5.2% (3/58)、非 IFG/STS の 10.5% (71/675) にみられた。領域に占める電極の割合を比べると、正答時に HGA 上昇を示す電極は、IFG に有意な電極の集積がみられた ($p=0.0046$)。一方、STS へは有意な集積はみられなかった ($p=0.20$)。

IFG、STS など各領域の活動がどのような認知行動成績と関連するものか検討するため、同一領域において融合認知の際に HGA が上昇する電極の割合と正

しい聴覚認知の際に HGA が上昇する電極の割合をカイ 2 乗検定にて比較した。

IFG および非 IFG/STS 領域において、正答時の HGA 上昇電極が有意に多かった

(IFG: $p=0.0053$ 、非 IFG/STS: $p=0.027$) (図 21)。

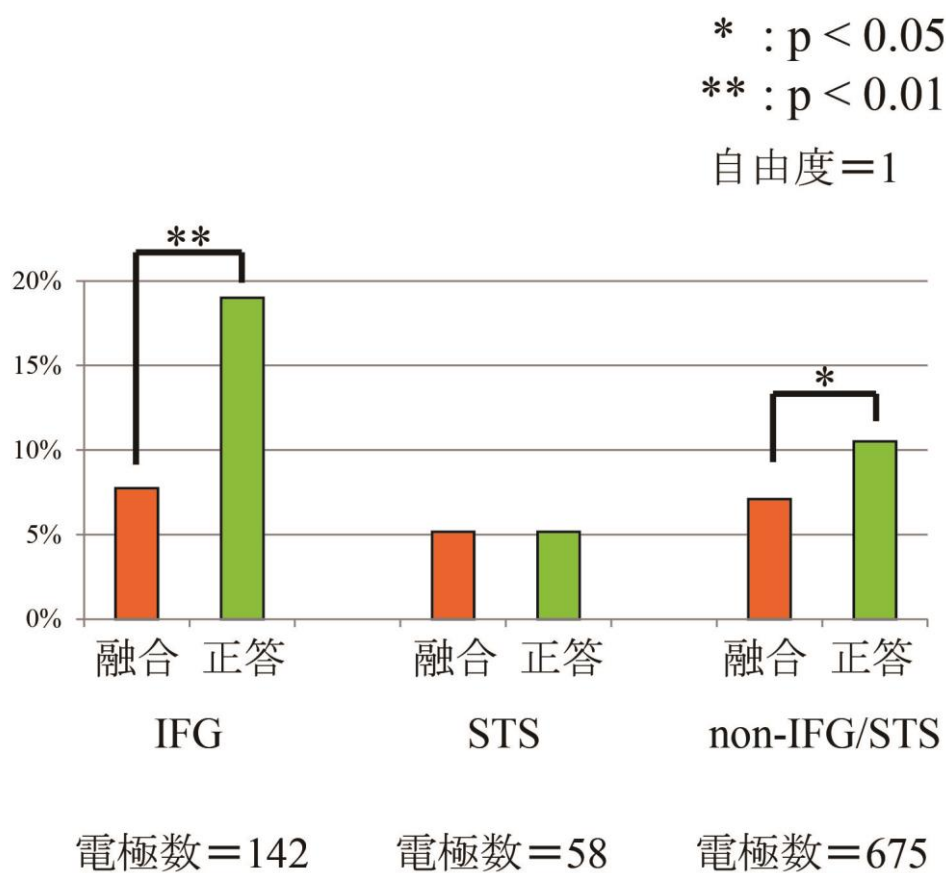


図 21. 融合認知および正答時に HGA が上昇した電極

融合認知および正答時に HGA が上昇した電極が各領域内において占める割合を示す。IFG と非 IFG/STS 領域では、融合時よりも正答時に HGA が上昇する電極が多く見られた (IFG: $p=0.0053$ 、非 IFG/STS: $p=0.027$)。

研究② 視聴覚不一致を伴う聴覚情報処理

A. 認知行動成績

聴覚情報処理に対する視覚刺激の影響を調べるため、各種条件間での話者の音声聞き取り成績の結果を比較した（表 2）。音声を正しく認知した割合は、視聴覚情報一致条件では 96.7%（SE=2.36%）であったのに対し、不一致条件（不一致（小）条件および不一致（大）条件）では 34.5%（SE=10.0%）であった。この違いは対応のある t 検定を用いて比較すると有意差を認めた（ $p=0.00049$ 、両側検定）。また、不一致条件の中でも不一致（小）条件と不一致（大）条件を比較すると、正答率はそれぞれ 30.3%（SE=9.3%）と 43.1%（SE=11.7%）であり、視聴覚不一致の程度が大きい方が正答率は高かった（ $p=0.0099$ ）（図 22）。また、聴覚課題の正答率は 95.8%（SE=3.89%）であり、視聴覚課題（一致条件）との間に、有意差は認めなかった（ $p=0.35$ ）。

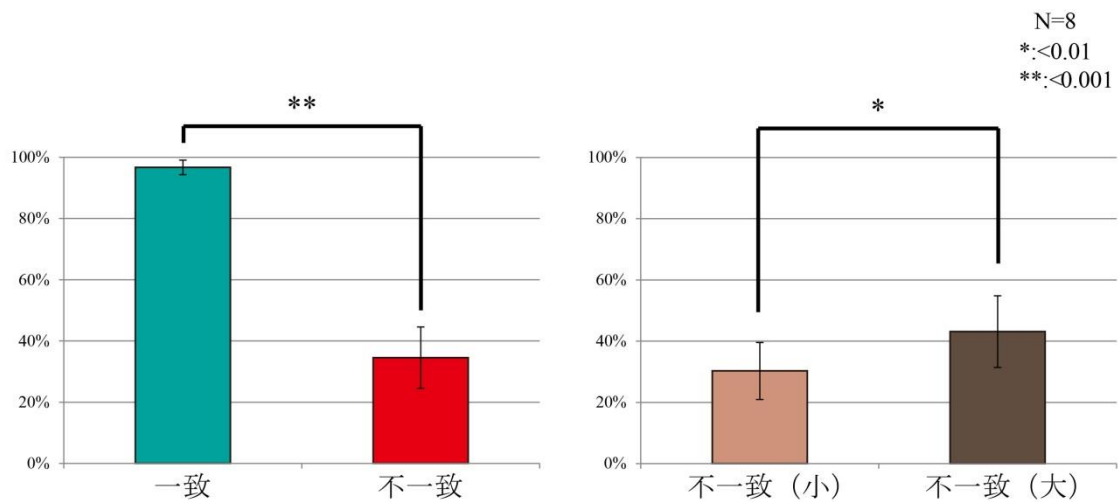


図 22. 認知行動成績

頭蓋内電極を留置した被験者 8 人での、聴覚情報を正しく認知できた率（正答率）を条件間で比較した。一致条件の正答率と不一致条件の正答率の比較を左に、不一致（小）条件の正答率と不一致（大）条件の正答率の比較を右に示した。両側 t 検定にて比較を行ったところ、それぞれ不一致条件、不一致（小）条件での有意な正答率の低下を認めた。エラーバーは標準誤差（N=8）を示す。

B. 応答時間

心的情報処理の多寡を測る指標の一つとして、音声開始から回答までの応答時間を計測した。課題の条件間で差があるか否か、回答の正誤で差があるか否かを比較した。条件間の比較では、一致条件（1.94 秒）と不一致条件（2.18 秒）の間に有意な差を認めた（ $P=0.01$ 、両側検定）。一方で、不一致（小）条件（2.09 秒）と不一致（大）条件（2.43 秒）の間には有意差はみられなかった（ $p=0.28$ 、両側検定）。不一致（小）条件における回答の正誤結果は、応答時間に有意差を

認めなかった ($p=0.28$ 、両側検定)。また不一致 (大) 条件においても同様に、回答の正誤で応答時間に違いはみられなかった ($p=0.99$ 、両側検定)。

C. 事象関連 HGA

不一致な視聴覚入力に対する処理機構を調べるため、視聴覚情報の不一致あり (不一致条件) または不一致なし (一致条件) における HGA の違いを比較した。視聴覚入力の不一致において HGA が上昇した電極位置を脳模式図に描画すると、IFG および STS に集積する傾向がみられた (図 23)。

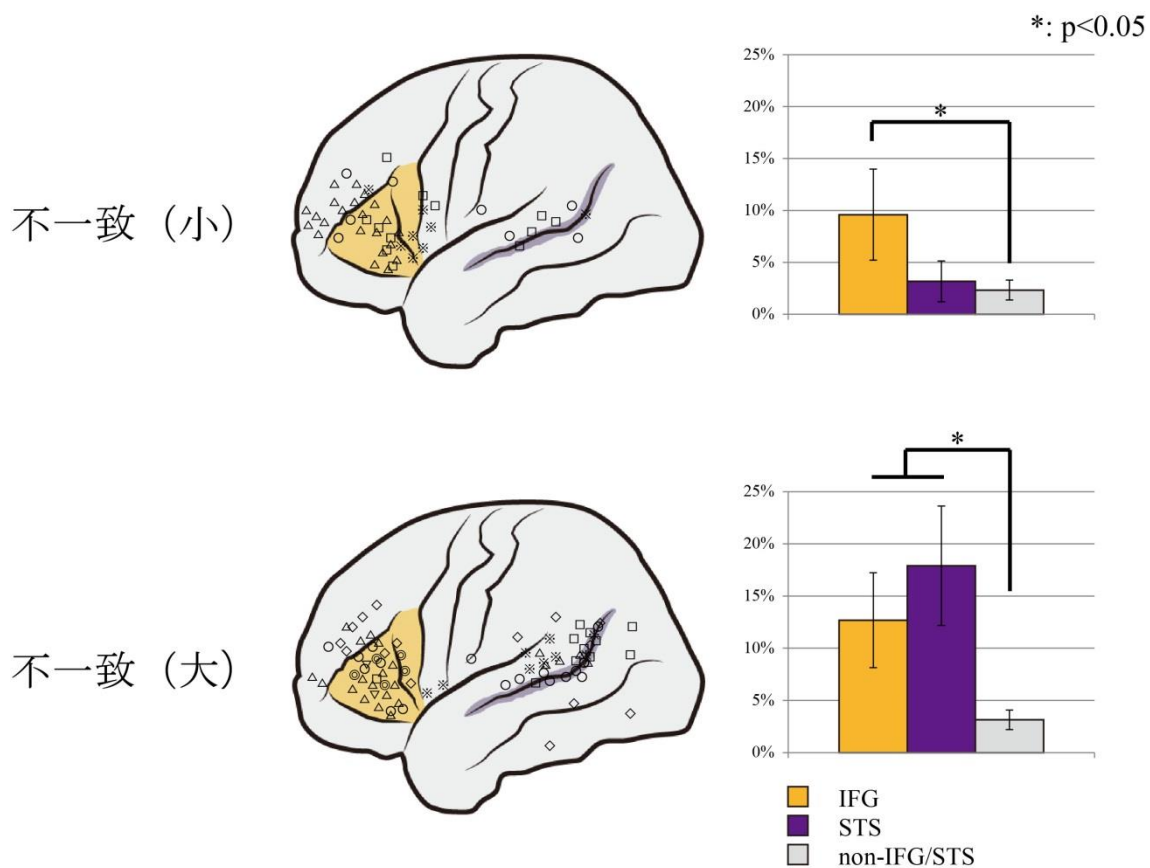


図 23. High gamma activity (HGA)上昇電極の分布図

左に一致条件に比べて不一致条件で有意なHGA上昇を示した電極の分布を脳模式図に示した。上段は不一致(小)条件、下段は不一致(大)条件と一致条件の比較である。Baselineは音声提示前1350~1050msとし、解析対象期間は500~1000msとした。解析対象電極全1408極のうち、不一致(小)条件で51極、不一致(大)条件で82極のHGA上昇電極を認めた。多重比較の補正はFDR法による。電極の形の違いは被験者の違いを示す。右はHGA上昇電極の下前頭回(inferior frontal gyrus; IFG; 橙)、上側頭溝(superior temporal sulcus; STS; 紫)、その他の領域(the other regions; 灰)に占める割合を示した。不一致(小)条件ではIFGに($p=0.046$; 上段)、不一致(大)条件ではIFG($p=0.038$; 下段)とSTS($p=0.016$; 下段)にHGA上昇電極が集積する。エラーバーは標準誤差($n=8$)を示す。

一致条件よりも不一致条件でHGAの大きな電極の割合は、不一致(小)条件ではIFG電極の中で9.59%(SE=4.38%)、STS電極の中で3.17%(SE=1.96%)、

その他の部位(非 IFG、非 STS)で 2.33% (SE=0.96%)であった。一方の不一致(大)条件では、IFG 電極の中で 12.7% (SE=4.54%)、STS 電極の中で 17.9% (SE=5.72%)、その他の部位 (非 IFG、非 STS) で 3.15% (SE=0.93%)の電極で、HGA 上昇がみられた。

IFG、STS の電極分布をその他の部位の電極分布と比較した結果、不一致(小)条件では IFG において有意に HGA が上昇する電極の集積がみられ ($p=0.046$ 、片側 t 検定)、不一致(大)条件では IFG ($p=0.037$ 、片側 t 検定と STS ($p=0.016$ 、片側 t 検定)において有意に HGA が上昇する電極が集積した (図 23)。

D. HGA と認知行動成績

ここまで、視聴覚不一致条件と一致条件の比較(提示条件の違いによる比較)により不一致処理に関わる活動部位の特定を行った。IFG や STS における HGA 上昇と視聴覚不一致の程度の詳細に調べるため、認知行動成績に基づいた解析(認知行動成績の違いによる比較)を行った。視聴覚不一致な動画クリップの提示(不一致(小)条件と不一致(大)条件)を行った際の、正答と誤答という認知行動成績の違いで分類した際の HGA を比較した。

正しい聴覚認知に関連する脳活動が不一致の大きさによって異なるのかを調べるため、正答の際の HGA が誤答の際の HGA よりも大きいかどうか、8 人

の被験者における IFG、STS それぞれについて対応のある t 検定で比較した。不一致（小）条件においては IFG において正答時の HGA 上昇がみられ（ $p=0.018$ 、片側検定）、不一致（大）条件においては、STS において正答時の HGA 上昇がみられた（ $p=0.038$ 、片側検定）（図 24）。

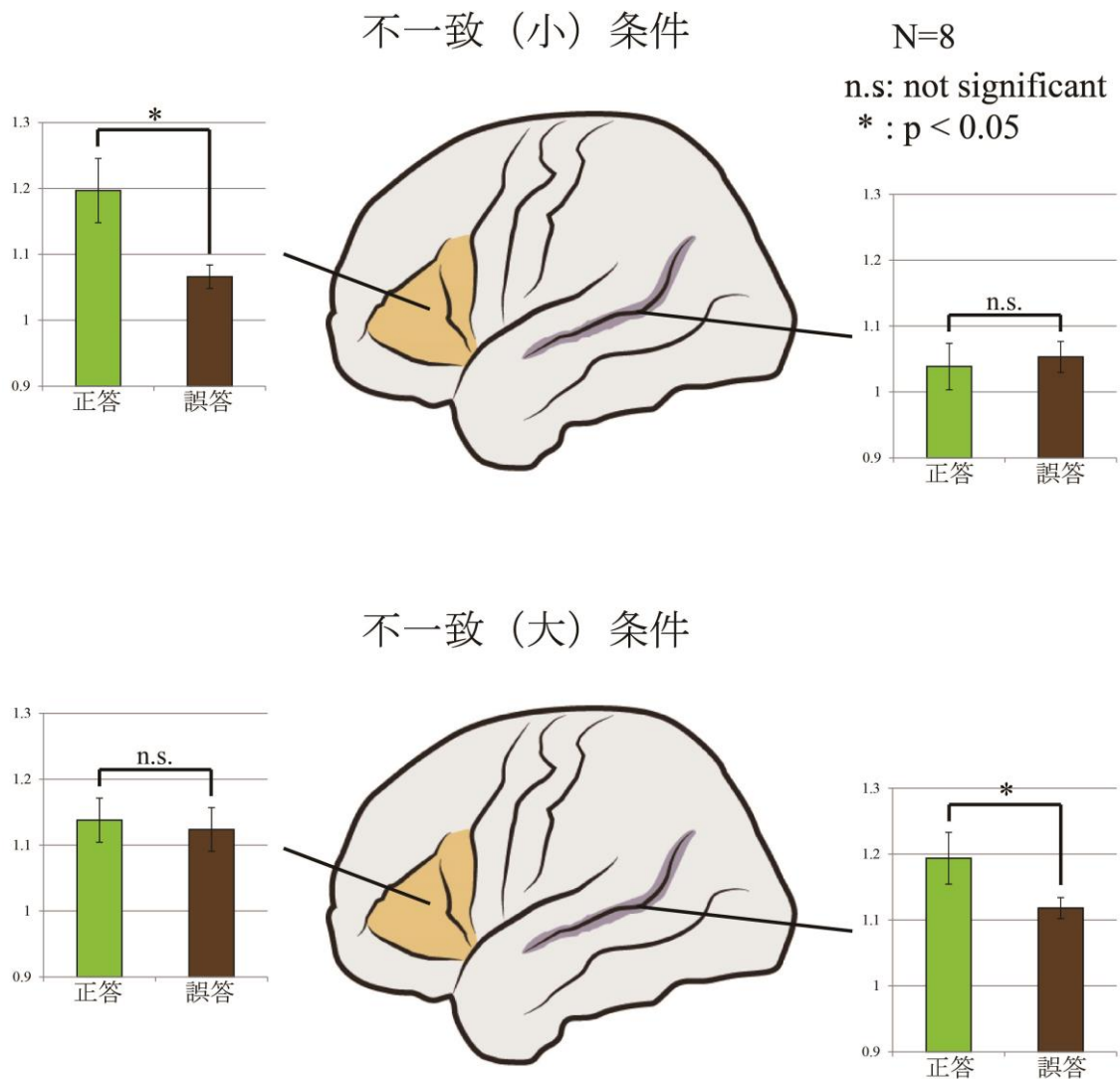


図 24. 正答時と誤答時の HGA : グループ解析 (N=被験者数)

不一致（小）および（大）条件を提示した際の脳活動を正答の時と誤答の時に分け、正答時の HGA が誤答時よりも大きいかどうか 8 人の被験者で対応のある t 検定にて比較した。不一致（小）条件では IFG において正答時の有意な活動がみられ ($p=0.018$)、不一致（大）においては STS において正答時の有意な活動が観察された ($p=0.038$)。エラーバーは標準誤差を示す。

さらに、全被験者から IFG 電極 219 極と STS 電極 81 極を集め、正答時の HGA が誤答時の HGA よりも大きいかどうか、対応のある t 検定で比較した。不

一致（小）条件においては、IFG の HGA は正答時に誤答時よりも有意に上昇した（ $p=1.36E-12$ 、片側検定）が、STS の HGA は正答時と誤答時で差はみられなかった（ $p=0.121$ 、片側検定）。逆に、不一致（大）条件においては、IFG の HGA は正答時と誤答時で違いはみられなかった（ $p=1$ 、片側検定）が、STS の HGA は正答時に有意な上昇がみられた（ $p=3.26 E-6$ 、片側検定）（図 25）。

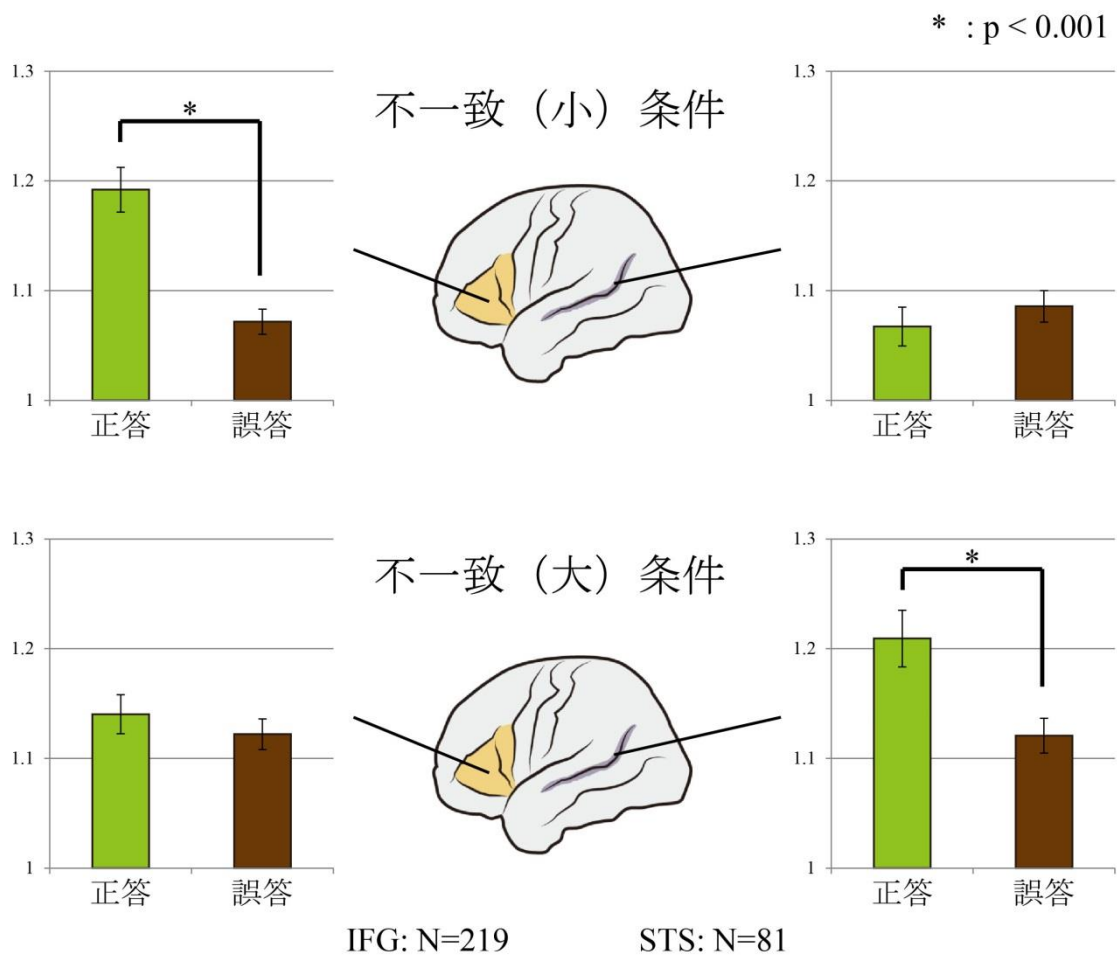


図 25. 正答時と誤答時の HGA : グループ解析 (N=電極数)

視聴覚不一致の提示をした際、音声認知の正答時と誤答時での HGA を比較した。全電極に対して正誤時の HGA を求め、全被験者から下前頭回 (inferior frontal gyrus; IFG) 219 極、上側頭溝 (superior temporal sulcus; STS) 81 極の電極を集めて比較した。不一致 (小) 条件では、IFG で正答時の HGA が高く ($p=1.36E-12$ 、片側検定)、不一致 (大) 条件では、STS で正答時の HGA が高かった ($p=3.26 E-6$ 、片側検定)。N は IFG と STS の電極数を示す。エラーバーは標準誤差を示す。

E. 皮質の活動時間

前頭葉と側頭葉の活動の時間的評価のため、事象関連 HGA の解析で有意差のあった電極において、課題条件による前頭葉と側頭葉の活動時間を比較した。不一致（小）条件におけるピーク潜時は、前頭葉で 775 ms (SE=21.3)、側頭葉で 775 ms (SE=50.0)であったが有意差は認めなかった ($p=0.996$)。不一致（大）条件におけるピーク潜時は、前頭葉で 804 ms (SE=17.7)、側頭葉で 725 ms (SE=31.1)であり、側頭葉の活動が有意に速かった ($p=0.034$ 、両側検定) (図 26)。

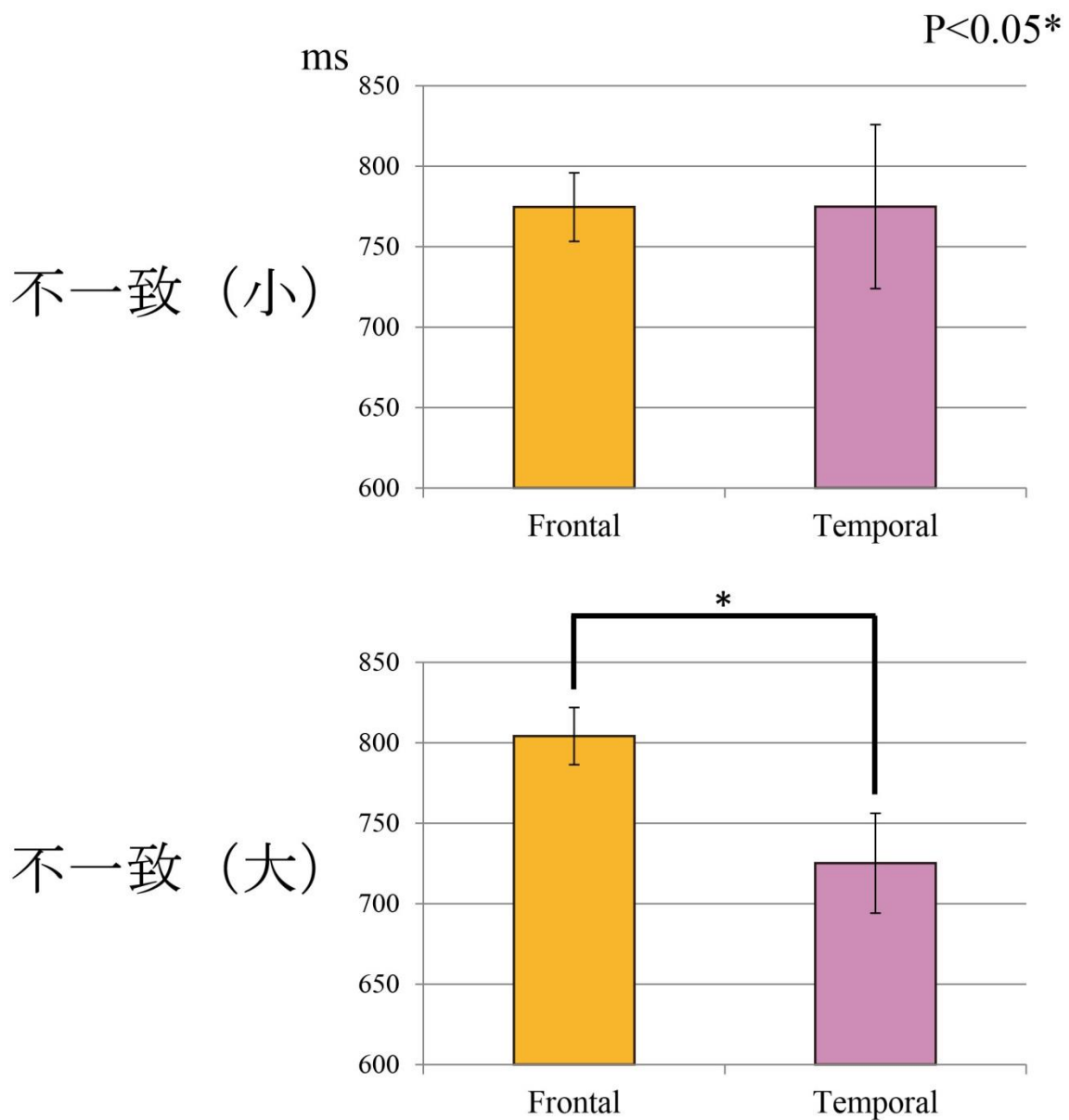


図 26. ピーク潜時の比較

8人の被験者において、不一致条件での事象関連 High gamma activity に有意差のある電極を集め、前頭葉と側頭葉での課題に対する活動時間の違いを検討した。不一致 (小) 条件では、前頭葉 42 極 (775 ms; SE=21.3) と側頭葉 6 極 (775 ms; SE=50.0) で有意な差は認めなかった ($p=0.56$)。一方、不一致 (大) 条件では前頭葉 43 極 (804 ms; SE=17.7) と側頭葉 35 極 (725 ms; SE=31.1) で活動開始時間に有意な差を認めた ($p=0.034$)。エラーバーは標準誤差 (不一致 (小) : 前頭葉 $n=42$ 、側頭葉 $n=6$ 、不一致 (大) : 前頭葉 $n=43$ 、側頭葉 $n=35$) を示す。

VI. 考察

A. 研究① マガーク効果（融合認知）の脳活動

研究①では、マガーク効果（融合認知）の脳活動について、マガーク条件で誘発される脳活動は、融合認知によるものなのか（モデル 1）、融合認知に特異的な脳活動よりもむしろ不適切な聴覚認知を避けて正しい聴覚認知を得る処理機構と関連しているのか（モデル 2）を検証した。

本研究では、マガーク条件を提示した際に起こる脳活動を、認知行動成績によって分類し、融合認知に関わる脳活動（モデル 1）と、正しい聴覚認知に関わる脳活動（モデル 2）を比較した。

マガーク条件を課した際に起こる脳活動を視聴覚一致条件の動画視聴時と比較し、マガーク条件提示（視聴覚情報の違い）によって起こる脳活動を、HGAを用いて評価した。結果、マガーク条件提示では IFG と STS 周辺が賦活された（有意な HGA 上昇電極の分布は IFG のみ）。また、HGA 上昇電極が観察される被験者は融合認知が起こりにくい被験者である傾向がみられた。これは、本研究で捉えている HGA が融合認知とは異なる認知に関連した活動である可能性を示唆する。マガーク条件で起こる脳活動がどのような認知と関連するのか調べるため、認知行動成績で分類して IFG と STS の脳活動を比較した。その結果、融合認知の時の HGA は、IFG、STS とも有意な上昇は見られなかったが（IFG:

p=0.41、STS: p=0.29) (図 17)、正しい聴覚認知 (正答) の時の HGA は IFG において有意に上昇した (IFG: p=0.036、STS: p=0.63) (図 19)。さらに、融合認知時に HGA が上昇する電極の分布と正しい聴覚認知時に HGA が上昇する電極の分布を検討したところ、融合認知時には (非 IFG/STS と比較して) IFG にも STS にも電極の集積は見られなかったが (IFG: p=0.79、STS: p=0.58) (図 18)、正しい聴覚認知時には (非 IFG/STS と比較して) IFG において有意な電極の集積がみられた (IFG: p=0.0046、STS: p=0.20) (図 20)。また、IFG、STS、非 IFG/STS の各同一領域内において、融合認知に関連した HGA 上昇と正しい聴覚認知に関連して HGA 上昇した電極の占有率を比較すると、IFG と非 IFG/STS 領域において正しい聴覚認知時の電極密度が高く、これらの領域においては融合認知に関与する電極よりも正しい聴覚認知に関与する電極が有意に多かった (IFG: p=0.0053、非 IFG/STS: p=0.027) (図 21)。これらの結果は、マガーク条件提示時に賦活される脳活動は融合認知よりも、不適切な聴覚認知を避け正しい聴覚認知を得ることと関連した脳活動であること (モデル 2) を支持するものである。つまり、マガーク条件の視聴覚不一致を検出し視覚情報を排除して不適切な融合認知を避ける (正しい聴覚認知を得る) 時に働く情報処理が存在することを示唆している。

1. 予期と不一致検出

ヒトは入力される情報から常に無意識に外界の環境を認知し次の状況を予測している。これは変化する外界に適応するためにも必要な能力であり、次の入力を予期することで情報処理の負荷を軽減することにもつながっている。一方で、予期せぬ入力に対してはこれを検出して処理を行うという神経基盤 (predictive coding) が必要であるが⁵⁷、このような予測に反する事象の検出に関わる反応として、ミスマッチネガティビティ (MMN) と呼ばれる事象関連電位が知られている^{58,59}。MMN とは、連続した同一の音刺激を提示しているなかで、周波数や音圧、刺激間隔が他とは異なる音刺激を提示した時に発生する反応である。つまり、本来期待される入力と異なる入力がなされた場合（または期待される入力がなされなかった場合にも）に観察される応答であり、本人が刺激に注意を向けていなくても、また気が付いていなくても発生する。つまりヒトの脳機能として、予期とそれに見合わない入力を検出する機構が存在するということである。しかし、MMN は聴覚または視覚等の単一モダリティに対する反応であるため、ここから得られる知見はあくまで単一モダリティにおける不一致検出の神経活動・神経基盤に関するものである。言い換えると、“視覚と聴覚”という異なる感覚器官からの入力に対する不一致検出の神経基盤については、MMN の実験手法では検討することが難しく、現在まで十分な検証がなされてい

ない。

マガーク条件は融合認知を起こす特殊な条件としての位置づけがあるものの、視覚と聴覚に不一致のある条件としても捉えられる。このことから、視覚情報（口唇の動き）から予期される聴覚情報と実際に提示される聴覚情報に乖離があった場合には、この不一致を検出する脳活動が惹起される可能性が考えられる。本研究で得た結果は、視聴覚情報から不一致な視覚情報を検出して排し、正しい聴覚認知を得るための情報処理機構の活動を反映したものと考えられる。

2. IFG/STS と視聴覚統合

IFG は視聴覚の統合に関与するという報告⁶⁰も関与しないという報告⁶¹もあり、いまだ結論は出ていない。そもそも IFG は言語を統一的に処理する部位であり⁶²、音素認知課題においては賦活されるのは妥当である。しかし、本研究では音声に対する視覚的要素の違いが IFG をより強く賦活させており、視聴覚の統合的利用を反映していることが示唆された。また、IFG は視聴覚不一致の処理に関与するとされるが²³、その役割は本研究で示した不一致検出機構であるかもしれない。マガーク条件は、融合認知が起こることからわかるように視聴覚の不一致を検出することが困難と考えられる。マガーク条件は STS で不一致を

検知できず、IFG でのより高次の処理によってのみ検知できた可能性がある。IFG と STS はそれぞれ優位半球においては言語領野であるブローカ野、ウェルニッケ野として知られる。特に、IFG は言語関連領野の中でも中枢的役割として、語義 semantic、構文 syntactic、音声 phonological レベルでの処理を統一して担っていると考えられている⁶²。さらに、言語表出や高度な言語的選択⁶³、記憶関連の選択⁶⁴などにも関与しているとされ、多くの高次脳活動に携わる部位である。このようなことから、IFG の活動は、不一致の小さい視聴覚入力(マガーク条件)の不一致をも検出でき、正しい聴覚認知を可能にしているのかもしれない。

サルにおいては STS に視覚刺激と聴覚刺激の両方に反応する神経細胞が存在することが知られている⁶⁵⁻⁶⁸。そして一次視覚野と一次聴覚野の間に位置し、両者からの投射を受ける視聴覚の連合野である⁶⁹⁻⁷⁴。過去の報告ではマガーク条件に関連した活動がみられる部位であるが^{20,21}、本研究ではマガーク条件の視聴覚入力がなされた際、STS の活動は少なくとも脳表電極では乏しかった。これには、日本人はマガーク効果が起こりにくいということ¹⁵が関与しているのだろうか。しかし、今回計測した被験者における融合認知の生起率は 50.3%に達しており過去の報告と同様であった¹⁵。加えて、過去の日本人を対象とした fMRI などの研究でも STS の活動は捉えられていることや、本研究における融合認知の起こりやすかった被験者においてもマガーク条件提示で STS の活動が強いと

いう結果は得られなかったことから、この可能性は考えにくい。また、今回は神経発火と関係が深いとされる HGA⁴⁴に着目しているが、違う周波数帯で STS が活動しているのであろうか。しかし、HGA ともっともよく相関するとされる fMRI を用いた研究では STS の活動は捉えられており²¹、主な要因とは考えにくい。本研究では皮質脳波を用いて脳活動を捉えたが、外側に留置した頭蓋内電極を用いているため脳溝内の活動を過小評価したために活動の差を捉えられなかった可能性がある。この評価には、深部電極や脳溝内電極⁷⁵が必要となる。

また、本研究は事象関連データであることから、MMN などのように刺激の提示頻度によって HGA の変化が観察されにくくなる (i.e., マガーク効果の馴化が起きる) 可能性も指摘されうる⁷⁶。しかし、融合認知は被験者に視聴覚情報が一致していないことを教示していても生じるうえ¹⁸、通常の認知行動研究では同一の被験者に 400 回ほど課題を提示しても融合認知は生じる⁷⁷。このように融合認知は被験者の予測や短期的な訓練に左右されにくいため、本研究における課題の提示頻度がマガーク効果の生起や HGA の変化に大きな問題となる可能性は低いと考えられる。しかし、本問題に関する更なる検討を行うことで、視聴覚情報処理に関してさらに頑健な知見が得られることが期待される。

3. 本研究の新奇性

本研究の結果は、マガーク条件で賦活される IFG や STS の活動が、これまで考えられていたような融合認知と関連して起こる活動でなく、不一致な視聴覚入力となされた際の正しい聴覚認知に関連するものである可能性を支持する。提示する課題の違いで分類し脳活動の比較を行った研究は多いが、本研究で行ったように融合・非融合や、正答・誤答という認知結果（認知行動成績）で分類し、定量化した IFG や STS の活動量との関係を明らかにした報告はない。これは、皮質脳波の高い信号対雑音比から、トライアルベースでの認知結果を用いて脳活動を分類することができたことによる。本研究により、マガーク条件のように視覚情報と聴覚情報に不一致があった時、視聴覚不一致の検出に関連した情報処理が働き、その活動が正しい聴覚認知に関連することが示唆された。

しかしながら、本研究において融合認知に関連した HGA が観察されなかったからといって、融合認知に関連した脳活動が存在しないとは言いきれない。今回、時間周波数解析によって活動の変化が大きい周波数は high gamma 帯域であったと確認したが、他の帯域における解析を baseline の取り方を変更するなどしながら検討することで異なる結果を得るかもしれない。実際に Keil らは、音声提示前の beta 帯域の活動が融合認知の生起率に影響すると報告している⁷⁷。このように音声提示前の脳活動の関与も指摘されており、今後、本研究で対象としていない解析時間帯や周波数帯域と認知の関連も評価していくことが必要

と思われる。

B. 研究② 視聴覚不一致を伴う聴覚情報処理基盤の動的変化

研究②では、言語に関連する視聴覚入力の中から聴覚情報を目標刺激として処理する神経基盤が、視聴覚情報の不一致の程度によって動的に変化するかどうかを検証した。不一致のある視聴覚情報処理には IFG と STS が関与することが広く知られていたが、これまでこれらの機能的差異については不明であった。視覚や聴覚の単一種感覚入力の先行研究では、目標刺激 (Target signal) と妨害刺激 (Noise) の違いが顕著である場合には感覚野で処理され、顕著でない場合には前頭葉で処理されることが知られている。この知見を踏まえると、視聴覚情報処理においても目標刺激と妨害刺激の乖離の程度に応じて情報処理基盤が変化すること、特に視聴覚情報の乖離が大きい場合は感覚連合野である STS、乖離が小さい場合は前頭葉言語野である IFG が関与することが予想された。そこで本研究では、視聴覚情報の不一致の程度を操作した課題を提示することで、不一致を伴う視聴覚情報処理における IFG と STS の機能的差異を検討した。

実験では、被験者に口唇の動きが見える話者を前に音声の聞き取りを課すとともに、IFG と STS について視聴覚情報が一致している場合と一致していない場合における HGA 上昇の差を評価した。その結果、IFG では、不一致の程度によらず HGA 上昇の差が顕著であった (図 23)。しかし、正しい音声認知と HGA 上昇は、不一致 (小) 条件では有意な関連を示したが、不一致 (大) 条件では

関連を示さなかった（図 24、25）。一方で STS では、不一致（大）条件の時のみ HGA 上昇の差が顕著で（図 23）、正しい音声認知と HGA 上昇が関連した（図 24、25）。これらの結果は、視聴覚情報処理における神経基盤が、視聴覚情報の不一致の程度に応じて動的に変化するという仮説を支持するものである。

1. トップダウン処理・ボトムアップ処理

以上の知見は、視聴覚情報処理における STS と IFG の機能差を理解するうえで有益である。これまでの単一種感覚入力の研究において明らかにされた、入力情報の乖離の程度に応じて変化する情報処理機構の背景には、トップダウン処理とボトムアップ処理というヒトの 2 種類の情報処理過程が想定されている^{41,42,78}。トップダウン処理は概念駆動型処理とも言われる処理方法である。初めに高次水準にある概念や理論などから駆動され、予期、期待などをもとに外界からの刺激を処理する過程である。そのため、高次機能処理を要し、感覚野での処理に先行して主として前頭葉で処理が行われると考えられている。例えば、騒音の中で特定の相手の声を聞き分けるときの脳活動は、オーケストラの演奏中に特定の楽器の音を聞き分けるなどはトップダウン処理とされる。また、図 2 として先に挙げた視覚探索課題のように、似た形状や色などの記号の中から検出目標とする記号を検知するにはトップダウン処理を要する（図 2 左）。こ

のように、目標刺激と妨害刺激の乖離の程度が小さいときにはトップダウン処理を行っていると考えられている。これに対してボトムアップ処理はデータ駆動型処理とも言われる。感覚刺激によって駆動され、入力データに基づいた部分的処理が先行して行われた後に高次のレベルへ情報処理が進む処理過程である。例えば自分へ向かって飛んでくるボールが視野に入った場合、ただちにボールへ注意が向くのは背景に対して入力刺激の顕著さが明らかであるためであり、このような刺激はボトムアップ処理の対象となる。図 2 の視覚探索課題で言えば、目標刺激と妨害刺激に顕著な形状の違いや色の違いがある図 2 右は、ボトムアップ処理による検知が行われる(図 2 右)。ヒトの脳は、入力情報によって、これらの処理を使い分けると考えられている。視聴覚不一致処理に関与する部位が IFG と STS であるという知見もあわせると、トップダウン処理には IFG が、ボトムアップ処理には STS が関与することが想定される(図 27)。IFG は言語中枢として高次処理を担うなど⁶²、高次皮質としてトップの処理に関わり、STS は視聴覚処理で一次聴覚野と機能的結合を持ち⁶⁹、感覚野(低次皮質)としてボトムの処理に当たる領域として矛盾しない。

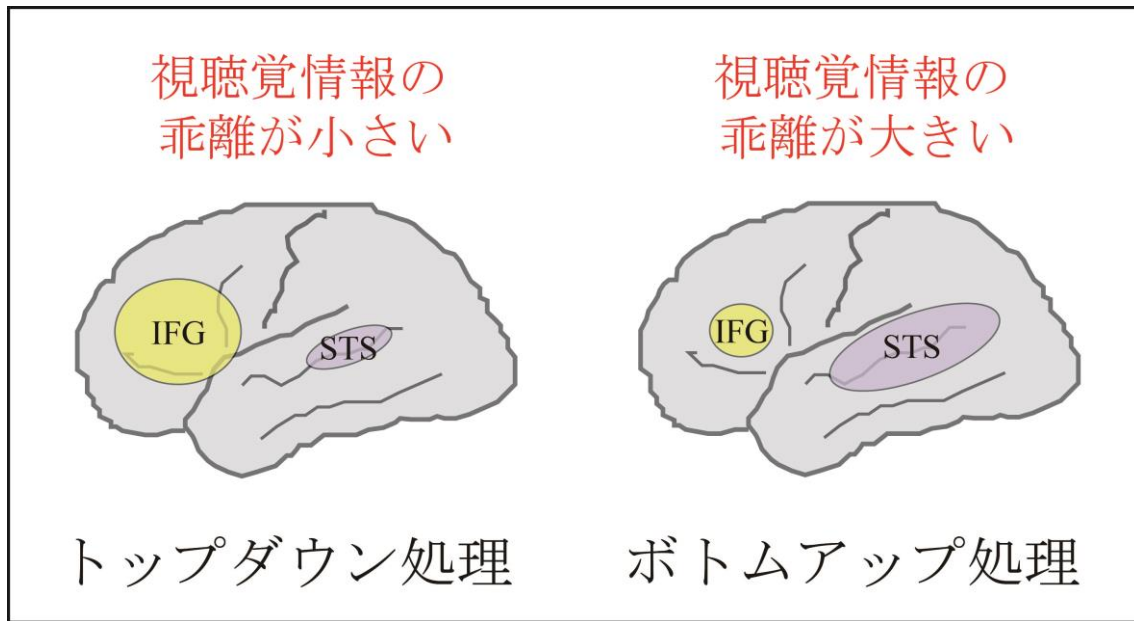


図 27. 視聴覚情報のトップダウン処理とボトムアップ処理の概念図

(左) 視聴覚情報の乖離が小さいときには前頭葉での高次処理として IFG がトップダウン処理にあたる。(右) 視聴覚情報の乖離が大きいときには感覚野である STS においてボトムアップ処理がなされる。

不一致(小)では IFG において HGA 上昇が観察されたのに対して不一致(大)では IFG と STS 両部位において HGA 上昇が観察された。IFG はなぜ不一致(大)条件でも活動したのであろうか。単一感覚入力の研究では、IFG はトップダウンだけでなくボトムアップ処理にも関与することが示されており、ボトムアップ処理におけるこの IFG の活動は、処理の主たる感覚野近傍での活動の後、情報が高次皮質に送られることで起こると考えられている⁷⁹⁻⁸¹。本研究での認知行動成績を加味しない場合の IFG における HGA 上昇は、このボトムアップ処理の結果を反映していると推察される。

また、これらの処理の時間的評価は動物実験で報告されている。例えば Buschman らはサルに対して視覚課題を行い、前頭葉と頭頂葉の神経活動を記録した。その結果トップダウン処理を要する場合には前頭葉が、ボトムアップ処理を要する場合には頭頂葉が、目標刺激の位置情報をより早く表現することが示された⁴⁰。このように、トップダウン処理では前頭葉の活動が先行し、ボトムアップ処理では感覚野の活動が先行するという活動時間特性も示されている。

本研究結果がトップダウン処理やボトムアップ処理という処理過程を反映したものであるのか検討するため、前頭葉と側頭葉の活動時間の比較を行った。その結果不一致（大）条件時に側頭葉のピーク潜時が前頭葉に先行しており（図26）、前述の動物実験の知見に合致した。一方で、不一致（小）条件でのピーク潜時は前頭葉と側頭葉で差はみられなかった。これには、視不一致（小）条件時には側頭葉の賦活が乏しく、解析対象の電極が少なかったことや、トップダウン処理の遅さ故に前頭葉と側頭葉の活動の差が埋没した可能性が考えられる。本研究の活動時間からはトップダウン処理・ボトムアップ処理を支持する結果は得られなかった。

ただし、トップダウン処理やボトムアップ処理という概念はひとつの推察であり、本研究で得られた結果のみで断定するには不十分である。特に今回得られた脳活動パターンの違いについては、注意の程度が脳活動の違いとして結

果に影響を及ぼしたという解釈もありうる。本研究では、不一致（小）条件と不一致（大）条件を設けることで、視聴覚情報の乖離の大小を操作するとともに脳活動パターンの違いを明らかにした。しかし課題正答率が示すように、不一致（大）条件に比べて、不一致（小）条件では視聴覚情報の乖離が小さく聴覚情報を正しく回答することが難しかった。そのため、視聴覚不一致の大小に伴う脳活動パターンの違いは、処理様式の違いではなく、課題遂行に要した注意の量の違いを単に反映しているだけという可能性も考えられる。すなわち、視聴覚不一致が大きく課題を容易に遂行できる場合は STS が、視聴覚不一致が小さく課題を遂行するには多くの注意が必要であった場合は IFG が活動するということである。

確かに、本研究では視聴覚情報の不一致の程度を操作しただけであるため、このような注意の配分量が不一致（小）条件と不一致（大）条件で異なっていた可能性は否定できない。だが本研究からは、“不一致（小）の正誤は IFG と、不一致（大）の正誤は STS とそれぞれ関連する”という結果も同時に得られている。もし IFG/STS の活動パターンが単に注意の配分量を反映しているだけであれば、誤答した場合（i.e., 注意が十分に配分されなかった場合）は、不一致（小）条件であれ不一致（大）条件であれ、同様の脳活動が観察されることになり、課題の正誤と関連する脳活動が条件間で異なるという結果は得られない

であろう。これは、本研究の結果 (i.e., IFG/STS という脳活動パターンの違い) が注意の配分量ではなく処理様式の違いを反映している可能性を示唆する。今後、注意の配分量を統制した実験を行い、本問題について更なる検討を加えることで、IFG/STS の機能分担についてより確実性の高い知見が得られると期待される。

2. 本研究の新奇性

視聴覚不一致の情報処理において、IFG と STS 両部位の活動が観察されていたが、これらの部位の機能的差異はこれまで明らかでなかった。IFG と STS はどちらも視覚や聴覚の単独入力よりも視聴覚同時入力²⁷⁻³³、さらに視聴覚一致入力よりも不一致入力で強く賦活される^{25,31,34-37}点は共通している。これまで視聴覚情報処理における IFG と STS の機能的差異を評価するうえで、不一致の程度の異なる課題を用いた研究はほとんどみられない。また、提示する課題の違いで分類し脳活動の比較を行った研究は多いが、研究①と同様に正誤という認知行動成績で分類し、定量化した IFG や STS の活動量との関係を明らかにした報告はない。本研究は、不一致の程度の異なる課題を提示し、認知行動成績と HGA の定量値から IFG、STS の機能を検討した点において、これまでにはない新しいものである。このような観点で分析したことにより、不一致の程度に

応じた処理過程の違いを捉えることが可能になったと思われる。

3. 課題の有効性

今回の課題は自作の動画を用いたが、視覚情報のない聴覚課題や不一致のない視聴覚課題では十分に高い正答率が得られ、不一致条件では有意に正答率が低下することが確かめられた。また、視覚情報によって聴覚情報が妨害されることを確認するため課題提示から回答までの応答時間を計測した。応答時間は課題遂行成績の指標とされ、応答時間の長短は処理の多寡を反映すると考えられている^{82,83}。本研究では被験者にできるだけ早く回答するという教示をしていないので解釈に注意を要するが、視聴覚不一致条件で一致条件よりも応答時間が遅延した ($p=0.03$) という結果は、視覚情報が聴覚情報を妨害したことを反映していると考えられる。

本研究において、不一致（小）条件は実際の音声と口唇の動きが子音のみ異なり、不一致（大）条件は実際の音声と口唇の動きが子音と母音の両方が異なる。しかし、実際に被験者にとって前者は視聴覚の乖離が小さく、後者は乖離が大きいという条件となっていたのであろうか。通常、目標刺激と妨害刺激の乖離が小さい条件では、乖離が大きい条件に比べて目標刺激を検出する正答率は低下する。本研究結果は、不一致（小）条件は不一致（大）条件に比べて

有意に正答率が低かった ($p=0.005$ 、 $p=0.01$)。この結果は、被験者にとって、不一致（小）条件は、不一致（大）条件よりも視聴覚の乖離が小さい課題として認知されたことを意味する。以上より、本研究で用いた課題は、対象とした被験者に対して適切であったと考えられ、結果の妥当性を支持する。

C. 総合考察

本研究（研究①、研究②）から、視聴覚入力に対して不一致検出機構が働き、その不一致の程度によって IFG と STS の関わりが変化するという結果を得た。これまで視聴覚不一致提示と不一致検出機構との関連を調べた報告はなく、新たな知見を提示することができた。現時点での研究の限界と展望について述べる。

1. 研究の限界

本研究で用いた皮質脳波は高い時空間解像度を持ち脳機能の解明に有力であるが、電極が覆っていない大脳皮質に関しては計測を行えないことから、注目する脳領域を十分にカバーした症例を選ぶ必要があり、対象となる被験者を多数望めない。そのため、他の低侵襲的診断機器の利点を理解し、これらの検査と並行して研究を行っていくことで、より皮質脳波研究の結果を生かすことが可能となると考えられる。

また、本研究で対象とした脳溝内など電極の留置が困難な部分に関しては、外側の皮質に比べて活動が過小評価される可能性がある⁷⁵。今後、深部電極や脳溝内電極などを用いることで、詳しく評価することが可能になると考えられる。

さらに研究②で課した課題は、不一致条件の課題の種類が不一致の程度の大小ひとつずつと少なく、十分とは言えないかもしれない。しかしながら、本

研究のように入院中の患者に協力を得て行う研究は倫理的に十分な配慮が必要である。患者負担も考え、長時間の課題を課すことは望ましくない⁸⁴。したがって限られた時間で可能な、より統制された課題を工夫していく必要がある。

また、本研究の対象は難治性てんかんを持つ患者群である。皮質脳波研究や単一細胞発火の研究など、難治性てんかん患者を対象とした研究は多数行われているが、脳実質の器質的・機能的異常が存在していることに加えて、てんかんの長期罹患により機能部位が移動し得ることや抗てんかん薬の服用などの複雑な因子が関与するため、健常被験者を対象とした研究と同様に解釈してよいのか、これまで議論されてきてきた⁸⁴。特に、本研究は高次の脳機能を対象としたものである。健常被験者と比べIQが低いてんかん患者群が正しく課題を遂行できたのか、注意が必要である。ただし、本研究における被験者の融合認知（マガーク効果）の生起率は、先に述べたようにこれまでの健常被験者を対象とした報告と類似していることから、本課題を十分遂行でき、結果は信頼できるものと考えられた。

2. 展望

我々は不一致のある視聴覚入力を用いて情報処理基盤についての研究を行ったが、その中でいくつかの課題が見つかった。我々は本研究の結果の解釈に

においてトップダウン・ボトムアップ処理を想定した考察を行った。残念ながら、本研究の結果は、これを裏付けるだけの十分な内容を備えていない。今後は、頭蓋内脳波の利点を生かした時間軸を重視した解析を進めていく必要がある。

ここには、領域間の同期性に関する検討や各処理過程における皮質電気刺激による影響の検討などが含まれるであろう。本研究で用いた認知課題は音素処理に関するものであったが、複雑なネットワークの関与が想定されるトップダウン・ボトムアップ処理の機構を明らかにする上では、このような要素的な応答に対して詳細な電気生理学的検討を加えることが極めて有効と思われる。一方で、このような要素的な課題から得られる情報は、臨床や現実世界への即時応用性という観点では、あまり有用とは言えない。本研究の結果を踏まえて、単語レベル、更には文・文章レベルでの認知課題を用いた複雑なネットワーク機構の解明を行っていく必要がある。

VII. 結論

本研究は、言語に関連する視聴覚情報の脳内処理機構を検討したものである。脳活動の解析は、左大脳半球外側に広く留置した信号対雑音比と時空間解像度に優れた皮質脳波を用いて行い、神経発火と関連する HGA を評価した。我々が捉えた脳活動は、不一致な視聴覚情報が提示された際、不適切な視聴覚の統合を避け正しい聴覚認知を得ることと関連していた。この脳活動は、目標刺激と妨害刺激の不一致の程度に基づいて、動的に変化することが示唆された。すなわち、視聴覚不一致を伴う聴覚情報処理において不一致の小さい視聴覚入力には IFG が、不一致の大きい視聴覚入力には STS が関与する可能性がある。

VIII. 謝辞

本研究の機会を与えて下さり、ご指導を賜りました東京大学大学院医学系研究科脳神経外科学教授・斉藤延人先生に深く感謝いたします。本研究の遂行にあたり臨床・研究両面で直接的なご指導・ご助言を賜りましたNTT 東日本関東病院脳神経外科部長（元東京大学大学院医学系研究科脳神経外科学准教授）・川合謙介先生に深く感謝いたします。研究のアイデアから論文執筆に至るまで、研究についての具体的なご指導を賜りました玉川大学脳科学研究所（元東京大学大学院認知言語神経科学分野准教授）・坂井克之先生に深く感謝いたします。研究内容、計測、解析から統計に至るまで、あらゆる面での具体的な技術および論文執筆のご指導を賜りました東京大学大学院医学系研究科脳神経外科学助教・國井尚人先生、NTT 東日本関東病院脳神経外科・松尾健先生に深く感謝いたします。実際の実験プロトコル作成から計測、分析、解釈、論文執筆までご指導・ご協力を賜りました元東京大学大学院医学系研究科・茨木拓也氏、東京大学大学院教育学研究科日本学術振興会（元東京大学大学院医学系研究科日本学術振興会）・分部利紘氏に心より感謝いたします。

IX. 引用文献

1. Kunii N, Kamada K, Ota T, Kawai K, Saito N. A detailed analysis of functional magnetic resonance imaging in the frontal language area: a comparative study with extraoperative electrocortical stimulation. *Neurosurgery* 2011;69:590-6; discussion 6-7.
2. Ota T, Kamada K, Kawai K, Yumoto M, Aoki S, Saito N. Refined analysis of complex language representations by non-invasive neuroimaging techniques. *British journal of neurosurgery* 2011;25:197-202.
3. Kunii N, Kamada K, Ota T, Kawai K, Saito N. Characteristic profiles of high gamma activity and blood oxygenation level-dependent responses in various language areas. *NeuroImage* 2013;65:242-9.
4. Kunii N, Kamada K, Ota T, Greenblatt RE, Kawai K, Saito N. The dynamics of language-related high-gamma activity assessed on a spatially-normalized brain. *Clinical neurophysiology : official journal of the International Federation of Clinical Neurophysiology* 2013;124:91-100.
5. 小俣圭. 視聴覚統合における McGurk 効果の研究. 博士論文 東京工業大学. 2007.
6. Zion Golumbic EM, Ding N, Bickel S, et al. Mechanisms underlying selective

- neuronal tracking of attended speech at a "cocktail party". *Neuron* 2013;77:980-91.
7. Mesgarani N, Chang EF. Selective cortical representation of attended speaker in multi-talker speech perception. *Nature* 2012;485:233-6.
 8. Sumbly WH, Pollack I. Visual contribution to speech intelligibility in noise. *The Journal of the Acoustical Society of America* 1954;26:212-5.
 9. von Kriegstein K. A Multisensory Perspective on Human Auditory Communication. In: Murray MM, Wallace MT, eds. *The Neural Bases of Multisensory Processes*. Boca Raton (FL); 2012.
 10. Soto-Faraco S, Kingstone A. Multisensory integration of dynamic information. *The handbook of multisensory processes* 2004:49-67.
 11. van Wassenhove V, Grant KW, Poeppel D. Visual speech speeds up the neural processing of auditory speech. *Proceedings of the National Academy of Sciences of the United States of America* 2005;102:1181-6.
 12. Kuhl PK, Meltzoff AN. The bimodal perception of speech in infancy. In; 1982: *American Association for the Advancement of Science*; 1982.
 13. Kuhl PK. A new view of language acquisition. *Proceedings of the National Academy of Sciences* 2000;97:11850-7.

14. McGurk H, MacDonald J. Hearing lips and seeing voices. *Nature* 1976;264:746-8.
15. Sekiyama K, Tohkura Y. McGurk effect in non-English listeners: few visual effects for Japanese subjects hearing Japanese syllables of high auditory intelligibility. *The Journal of the Acoustical Society of America* 1991;90:1797-805.
16. Jordan TR, Sergeant PC. Effects of facial image size on visual and audio-visual speech recognition. *Hearing by eye II: Advances in the psychology of speechreading and auditory–visual speech* 1998:155-76.
17. Munhall KG, Gribble P, Sacco L, Ward M. Temporal constraints on the McGurk effect. *Perception & psychophysics* 1996;58:351-62.
18. Lifshitz M, Aubert Bonn N, Fischer A, Kashem IF, Raz A. Using suggestion to modulate automatic processes: from Stroop to McGurk and beyond. *Cortex; a journal devoted to the study of the nervous system and behavior* 2013;49:463-73.
19. Nath AR, Beauchamp MS. A neural basis for interindividual differences in the McGurk effect, a multisensory speech illusion. *NeuroImage* 2012;59:781-7.
20. Arnal LH, Wyart V, Giraud A-L. Transitions in neural oscillations reflect

- prediction errors generated in audiovisual speech. *Nature neuroscience* 2011;14:797-801.
21. Sekiyama K, Kanno I, Miura S, Sugita Y. Auditory-visual speech perception examined by fMRI and PET. *Neuroscience research* 2003;47:277-87.
 22. Campbell R. The processing of audio-visual speech: empirical and neural bases. *Philosophical transactions of the Royal Society of London Series B, Biological sciences* 2008;363:1001-10.
 23. Ojanen V, Mottonen R, Pekkola J, et al. Processing of audiovisual speech in Broca's area. *NeuroImage* 2005;25:333-8.
 24. Blank H, von Kriegstein K. Mechanisms of enhancing visual-speech recognition by prior auditory information. *NeuroImage* 2013;65:109-18.
 25. Hein G, Doehrmann O, Muller NG, Kaiser J, Muckli L, Naumer MJ. Object familiarity and semantic congruency modulate responses in cortical audiovisual integration areas. *The Journal of neuroscience : the official journal of the Society for Neuroscience* 2007;27:7881-7.
 26. McGettigan C, Evans S, Rosen S, Agnew ZK, Shah P, Scott SK. An application of univariate and multivariate approaches in FMRI to quantifying the hemispheric lateralization of acoustic and linguistic processes. *Journal of*

cognitive neuroscience 2012;24:636-52.

27. Calvert GA, Campbell R, Brammer MJ. Evidence from functional magnetic resonance imaging of crossmodal binding in the human heteromodal cortex. *Current biology : CB* 2000;10:649-57.
28. Stevenson RA, James TW. Audiovisual integration in human superior temporal sulcus: Inverse effectiveness and the neural processing of speech and object recognition. *NeuroImage* 2009;44:1210-23.
29. Van Atteveldt N, Formisano E, Goebel R, Blomert L. Integration of letters and speech sounds in the human brain. *Neuron* 2004;43:271-82.
30. Wright TM, Pelphrey KA, Allison T, McKeown MJ, McCarthy G. Polysensory interactions along lateral temporal regions evoked by audiovisual speech. *Cereb Cortex* 2003;13:1034-43.
31. Olson IR, Gatenby JC, Gore JC. A comparison of bound and unbound audio-visual information processing in the human cerebral cortex. *Brain research Cognitive brain research* 2002;14:129-38.
32. Skipper JI, Nusbaum HC, Small SL. Listening to talking faces: motor cortical activation during speech perception. *NeuroImage* 2005;25:76-89.
33. Sohoglu E, Peelle JE, Carlyon RP, Davis MH. Predictive top-down integration

- of prior knowledge during speech perception. *The Journal of neuroscience : the official journal of the Society for Neuroscience* 2012;32:8443-53.
34. Szycik GR, Jansma H, Munte TF. Audiovisual integration during speech comprehension: an fMRI study comparing ROI-based and whole brain analyses. *Human brain mapping* 2009;30:1990-9.
 35. Irwin JR, Frost SJ, Mencl WE, Chen H, Fowler CA. Functional activation for imitation of seen and heard speech. *Journal of neurolinguistics* 2011;24:611-8.
 36. Kaiser J, Hertrich I, Ackermann H, Mathiak K, Lutzenberger W. Hearing lips: gamma-band activity during audiovisual speech perception. *Cereb Cortex* 2005;15:646-53.
 37. Hocking J, Price CJ. The role of the posterior superior temporal sulcus in audiovisual processing. *Cereb Cortex* 2008;18:2439-49.
 38. Binder JR, Liebenthal E, Possing ET, Medler DA, Ward BD. Neural correlates of sensory and decision processes in auditory object identification. *Nature neuroscience* 2004;7:295-301.
 39. Zekveld AA, Heslenfeld DJ, Festen JM, Schoonhoven R. Top-down and bottom-up processes in speech comprehension. *NeuroImage* 2006;32:1826-36.
 40. Buschman TJ, Miller EK. Top-down versus bottom-up control of attention in

- the prefrontal and posterior parietal cortices. *Science* 2007;315:1860-2.
41. Corbetta M, Shulman GL. Control of goal-directed and stimulus-driven attention in the brain. *Nature reviews Neuroscience* 2002;3:201-15.
 42. Desimone R, Duncan J. Neural mechanisms of selective visual attention. *Annual review of neuroscience* 1995;18:193-222.
 43. Kastner S, Ungerleider LG. Mechanisms of visual attention in the human cortex. *Annual review of neuroscience* 2000;23:315-41.
 44. Ray S, Crone NE, Niebur E, Franaszczuk PJ, Hsiao SS. Neural correlates of high-gamma oscillations (60-200 Hz) in macaque local field potentials and their potential implications in electrocorticography. *The Journal of neuroscience : the official journal of the Society for Neuroscience* 2008;28:11526-36.
 45. Axmacher N, Schmitz DP, Wagner T, Elger CE, Fell J. Interactions between medial temporal lobe, prefrontal cortex, and inferior temporal regions during visual working memory: a combined intracranial EEG and functional magnetic resonance imaging study. *The Journal of neuroscience : the official journal of the Society for Neuroscience* 2008;28:7304-12.
 46. Crone NE, Hao L, Hart J, Jr., et al. Electrocorticographic gamma activity

- during word production in spoken and sign language. *Neurology* 2001;57:2045-53.
47. Tallon-Baudry C, Bertrand O. Oscillatory gamma activity in humans and its role in object representation. *Trends in cognitive sciences* 1999;3:151-62.
 48. Ogawa S, Lee T, Kay A, Tank D. Brain magnetic resonance imaging with contrast dependent on blood oxygenation. *Proceedings of the National Academy of Sciences* 1990;87:9868-72.
 49. Delorme A, Makeig S. EEGLAB: an open source toolbox for analysis of single-trial EEG dynamics including independent component analysis. *Journal of neuroscience methods* 2004;134:9-21.
 50. Chandrasekaran C, Ghazanfar AA. Different neural frequency bands integrate faces and voices differently in the superior temporal sulcus. *Journal of neurophysiology* 2009;101:773-88.
 51. Flinker A, Chang EF, Kirsch HE, Barbaro NM, Crone NE, Knight RT. Single-trial speech suppression of auditory cortex activity in humans. *The Journal of neuroscience : the official journal of the Society for Neuroscience* 2010;30:16643-50.
 52. Chang EF, Niziolek CA, Knight RT, Nagarajan SS, Houde JF. Human cortical

- sensorimotor network underlying feedback control of vocal pitch. *Proceedings of the National Academy of Sciences of the United States of America* 2013;110:2653-8.
53. Fischl B, van der Kouwe A, Destrieux C, et al. Automatically parcellating the human cerebral cortex. *Cerebral cortex* 2004;14:11-22.
54. Fisch L, Privman E, Ramot M, et al. Neural "ignition": enhanced activation linked to perceptual awareness in human ventral stream visual cortex. *Neuron* 2009;64:562-74.
55. Association IP. *Handbook of the International Phonetic Association: A guide to the use of the International Phonetic Alphabet*: Cambridge University Press; 1999.
56. Ogata S, Murai K, Nakamura S, Morishima S. Model-Based Lip-Based Lip-Based Lip Synchroniza Synchroniza Synchronization With Automatically Transla Y Transla Y Translated Synthetic Voice Ted Synthetic Voice Toward A Multi-Modal Translation System Tion System. 2001.
57. Huang Y, Rao RP. Predictive coding. *Wiley Interdisciplinary Reviews: Cognitive Science* 2011;2:580-93.
58. Naatanen R, Paavilainen P, Tiitinen H, Jiang D, Alho K. Attention and

- mismatch negativity. *Psychophysiology* 1993;30:436-50.
59. Garrido MI, Kilner JM, Stephan KE, Friston KJ. The mismatch negativity: a review of underlying mechanisms. *Clinical Neurophysiology* 2009;120:453-63.
 60. Romanski LM. Integration of faces and vocalizations in ventral prefrontal cortex: implications for the evolution of audiovisual speech. *Proceedings of the National Academy of Sciences of the United States of America* 2012;109 Suppl 1:10717-24.
 61. Matchin W, Groulx K, Hickok G. Audiovisual speech integration does not rely on the motor system: evidence from articulatory suppression, the McGurk effect, and fMRI. *Journal of cognitive neuroscience* 2014;26:606-20.
 62. Hagoort P. On Broca, brain, and binding: a new framework. *Trends in cognitive sciences* 2005;9:416-23.
 63. Fan J, Flombaum JI, McCandliss BD, Thomas KM, Posner MI. Cognitive and brain consequences of conflict. *NeuroImage* 2003;18:42-57.
 64. Badre D, Poldrack RA, Pare-Blagoev EJ, Insler RZ, Wagner AD. Dissociable controlled retrieval and generalized selection mechanisms in ventrolateral prefrontal cortex. *Neuron* 2005;47:907-18.
 65. Barraclough NE, Xiao D, Baker CI, Oram MW, Perrett DI. Integration of visual

- and auditory information by superior temporal sulcus neurons responsive to the sight of actions. *Journal of cognitive neuroscience* 2005;17:377-91.
66. Benevento LA, Fallon J, Davis BJ, Rezak M. Auditory--visual interaction in single cells in the cortex of the superior temporal sulcus and the orbital frontal cortex of the macaque monkey. *Experimental neurology* 1977;57:849-72.
67. Bruce C, Desimone R, Gross CG. Visual properties of neurons in a polysensory area in superior temporal sulcus of the macaque. *Journal of neurophysiology* 1981;46:369-84.
68. Hikosaka K, Iwai E, Saito H, Tanaka K. Polysensory properties of neurons in the anterior bank of the caudal superior temporal sulcus of the macaque monkey. *Journal of neurophysiology* 1988;60:1615-37.
69. Noesselt T, Rieger JW, Schoenfeld MA, et al. Audiovisual temporal correspondence modulates human multisensory superior temporal sulcus plus primary sensory cortices. *The Journal of neuroscience : the official journal of the Society for Neuroscience* 2007;27:11431-41.
70. Beauchamp MS, Lee KE, Argall BD, Martin A. Integration of auditory and visual information about objects in superior temporal sulcus. *Neuron* 2004;41:809-23.

71. Cusick CG. The superior temporal polysensory region in monkeys. In: *Extrastriate cortex in primates*: Springer; 1997:435-68.
72. Ghazanfar AA, Chandrasekaran C, Logothetis NK. Interactions between the superior temporal sulcus and auditory cortex mediate dynamic face/voice integration in rhesus monkeys. *The Journal of neuroscience : the official journal of the Society for Neuroscience* 2008;28:4457-69.
73. Kayser C, Logothetis NK. Directed Interactions Between Auditory and Superior Temporal Cortices and their Role in Sensory Integration. *Frontiers in integrative neuroscience* 2009;3:7.
74. Szycik GR, Stadler J, Tempelmann C, Munte TF. Examining the McGurk illusion using high-field 7 Tesla functional MRI. *Frontiers in human neuroscience* 2012;6:95.
75. Matsuo T, Kawasaki K, Osada T, et al. Intrasulcal electrocorticography in macaque monkeys with minimally invasive neurosurgical protocols. *Frontiers in systems neuroscience* 2011;5:34.
76. Naatanen R, Paavilainen P, Rinne T, Alho K. The mismatch negativity (MMN) in basic research of central auditory processing: A review. *Clinical Neurophysiology* 2007;118:2544-90.

77. Keil J, Muller N, Ihssen N, Weisz N. On the variability of the McGurk effect: audiovisual integration depends on prestimulus brain states. *Cereb Cortex* 2012;22:221-31.
78. Norman DA, Bobrow DG. On data-limited and resource-limited processes. *Cognitive psychology* 1975;7:44-64.
79. Friederici AD. The cortical language circuit: from auditory perception to sentence comprehension. *Trends in cognitive sciences* 2012;16:262-8.
80. Rinne T, Degerman A, Alho K. Superior temporal and inferior frontal cortices are activated by infrequent sound duration decrements: an fMRI study. *NeuroImage* 2005;26:66-72.
81. Salmi J, Rinne T, Koistinen S, Salonen O, Alho K. Brain networks of bottom-up triggered and top-down controlled shifting of auditory attention. *Brain research* 2009;1286:155-64.
82. Sternberg S. Memory-scanning: mental processes revealed by reaction-time experiments. *American scientist* 1969;57:421-57.
83. Donders FC. On the speed of mental processes. *Acta psychologica* 1969;30:412-31.
84. Mukamel R, Fried I. Human intracranial recordings and cognitive neuroscience.

Annual review of psychology 2012;63:511-37.