

McGurk 効果における時空間パターンとその統合機構

小俣 圭¹(おまた けい), 茂木 健一郎²

東京工業大学総合理工学研究科¹

ソニーコンピューターサイエンス研究所²

(要旨)

1976年に McGurk 効果が発見された (McGurk, 1976)。これは視覚情報によって聴覚情報の知覚が変化するというものである。視覚情報と矛盾する聴覚情報を同時に提示されるとその両者の特徴に整合するように知覚が起こる。これは、脳内における視覚の情報と聴覚の情報統合によって起こると考えられる。本研究では心理物理実験を行いこの情報の統合が脳のどの部位で行われるかを検討する。

Key words: McGurk 効果、視聴覚統合、音素認識

本研究は、McGurk 効果を用いて視聴覚統合について検討した。脳における視覚系は、研究が進んでおりその情報が V1 と呼ばれる部位に投影されていることや、左右の視野がそれぞれ逆の半球の V1 に投影されることなどがわかっている。聴覚系は、その情報処理系にまだ未解明な部分が多い。特に言語機能の遂行という機能に関しては、神経生理学的にはブローカ野やウエルニッケ野と呼ばれる領野において行われていると言われるが実際にそこでどのような神経機構が言語の受動生成に関わっているのかは明らかでない。

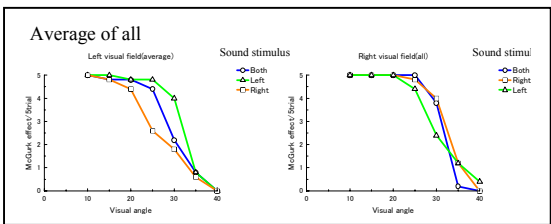
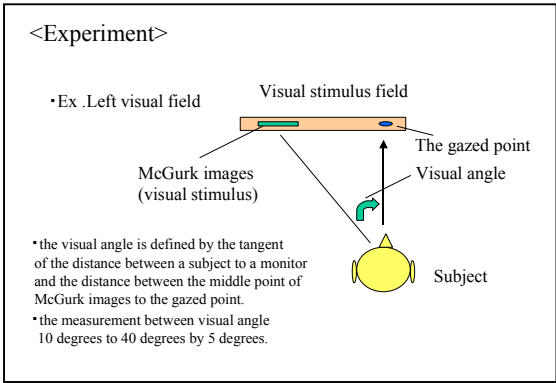
一方でマルチモダリティについて考慮することは、脳内における各感覚情報が同質なものとして解釈される必要があり、その時における情報処理は各感覚に無関係なものとして存在すると考えられる。McGurk 効果とは、視聴覚統合の結果として起こる。具体的には、「ga」という口の動きに対して、「ba」という音を重ねたビデオを見せると、音の知覚が「ba」ではなく「da」として知覚される。これは口の動きによって、「ba」という知覚が「da」にシフトした結果である。McGurk (1976) の説明によれば、視覚刺激「ga」には、[ga]と[da]の特徴、要素が存在し、聴覚刺激「ba」には[da]と[ba]の特徴、要素がある。すると脳はこれらの刺激に対して、矛盾の少ない知覚を選択することになる。つまり、両者の刺激に存在する「da」が合理的な知覚判断となるわけである。では、実際の問題としてこの統合機構は、脳内のどのような場所で、どのように起こるのだろうか。解剖学的事実を踏まえ、刺激の提示を工夫することで、McGurk 効果のラテラリティを

明らかにすることを目的に、心理物理実験を行った。以下にその内容と方法を記す。

左右の各視野における像は左右逆の脳半球に投影されることは既に述べた。聴覚系においては、解剖学より左右の耳からの情報は主に左右逆の脳半球に伝達されることがわかっている。そこで、左右の視野に視覚情報を提示した時に左右の耳に聴覚情報を提示したら、情報の統合がどうなるのかを検討することにした。

まず、McGurk 効果がおこる映像を作成した。口の動きは[ga]または[ka]であり、音は[ba]または[ma]である。作成した映像は画像の中心に顔がうつしだされており背景は白である。その映像をある注視点に対して左右に提示する。提示は注視点から視覚角度として10度から40度まで5度ずつ提示した。このとき被験者は注視点を見るように指示し画像を直接見てはいけないとした。音刺激は、両方の耳に音を提示する場合と左右の各耳に音を提示する場合の3つのパターンを用意した。すると全部で6つのパターンがありそのパターンによって、McGurk 効果の起こり方に変化があるかを検討した。

被験者はモニターの前に25cmのところ座り注視点を注視する。そこで視覚刺激を40度から10度に向かって5度ずつ変化させて提示し、その時、McGurk 効果の有無を判定させた。被験者には McGurk 効果がわずかでも起こった場合それを起こったとし、効果が起きたか、起きてないかの二択として報告させた。その状況とその結果を以下に示す。



上の図は実験の状況を表している。これは左視野における実験を表していて、被験者はヘッドホンにより音を提示される。

そしてその下のグラフはその結果を表している。被験者6人の結果を平均したものである。左側が左視野からの視覚刺激に対するグラフであり、右が右視野からの視覚刺激に対するグラフである。横軸は視覚角度であり、縦軸は McGurk 効果の割合を表している。のラインが両耳への聴覚刺激を表し、のラインが右耳への刺激、のラインが左耳への刺激を表している。視野角度が40度にもなると視覚の刺激がほとんど入らず McGurk 効果はほとんど起こらない。その一方で視覚情報を直視しているわけではないが、10度から25度に提示された視覚情報ではほとんど全ての場合において、McGurk 効果が起こっている。左右の視野に関しては右の視野に示された方が左の視野に提示された場合よりも効果が起こっているといえる。また、左右の耳への刺激であるが提示された視野側の耳への刺激の方がより効果が起こっている。以上の結果と視聴覚の脳のシステムを踏まえて考察を行う。

。脳の言語野は右利きの場合左半球に局在していることがわかっている。上記の結果では、右視野からの視覚刺激が左視野からの刺激よりも効果が起こりやすかった。この結果と言語脳の局在は一致している。また、左右の耳からの刺激は逆の半球に主に刺激が入ってゆくがこれに対する結果は左右対称であり、刺激が同

じ半球に入ってゆく方が起こりやすいのではないかという推測ができる。上記の結果においてラテラルリティという観念に立てば視聴覚の情報統合については、この McGurk 効果に関しては、左半球における優位が考えられる。一方で左右の耳への刺激と左右の視野への刺激の組み合わせとして4パターンによる差は、単純に左半球優位とは言えない。この理由として考えられるのは、左右の耳への刺激は視覚刺激の左右視野への提示ほど、脳のシステムとして分けられておらず、聴覚刺激は同側の大脳半球へも伝達されていることよると考えられる。

現在は、上記の実験の更なる検証と時系列変化による脳内統合処理過程について検討している。「ga」や「da」、「ba」といった音素の物理的な特徴は、その音素を構成するフォルマントの時間変化であることがわかっている。第二フォルマントの発音初期30ms程度の時間変化が上昇するか下降するかということがこれらの音素を捉えるときに重要であることは、範疇的知覚という現象からわかる。

一方で視覚情報としては、口の動きという一連の動作が音素の知覚に影響を与える。実際口の動きによって、一次聴覚野が反応するというfMRIのデータもある。この口の動きという画像の時系列変化と第二フォルマントの時間的な変化というものが統合されることによって、McGurk 効果が起こるとすれば、それはまさに統合には、時間変化が重要であり、その時間変化をコーディングしている神経細胞、もしくは細胞群の存在が必要となると言えよう。

視覚刺激と聴覚刺激の統合のためのタイムウィンドウは、およそ200ms前後と考えられる(Munhall, 1996)。このことから、厳密に視覚と聴覚の刺激が一致しておらずとも統合が起こると言える。すると刺激に対する単純な反射的機構として統合過程があるというよりも刺激の時系列変化をストックしそれを一つの情報として処理する機構が必要となる。

言語機能における脳研究、主に失語症の研究から、第一次聴覚野よりもウェルニッケ野あたりにおいて音素の認識を得ているのではないかと想定されるが、人における聴覚系の研究は視覚系に比べて遅れており、具体的な音素認識の神経細胞やその働きについては今後の研究が待たれる。