

科学研究費助成事業 研究成果報告書

平成 29 年 6 月 21 日現在

機関番号：82404

研究種目：若手研究(B)

研究期間：2014～2016

課題番号：26750215

研究課題名(和文) 音声知覚時における視覚情報と聴覚情報の統合処理過程の解明

研究課題名(英文) Investigate on audio-visual integration of speech sound

研究代表者

遠藤 智美(水落智美)(Tomomi, Mizuochi-Endo)

国立障害者リハビリテーションセンター(研究所)・研究所・脳機能系障害研究部・特別研究員(RPD)

研究者番号：90568859

交付決定額(研究期間全体)：(直接経費) 1,900,000円

研究成果の概要(和文)：音声知覚時において、視覚情報から得る音響要因としてフォルマント周波数に着目し、フォルマント周波数以外の音響要素を揃えた母音を用いて母音弁別課題を実施し、課題遂行時の脳波を計測した。

健常成人を対象とした脳波計測の結果は、有意な波形差は認められなかったが、皮質下に電極を留置した難治性癲癇患者に対して実施した皮質内脳波計測では、音声のみの提示で音声提示から200ms以内に、左外側側頭葉後部で認められた母音ごとの差が、視覚情報と同時に提示すると認められなくなった。これは、口の形という視覚情報があることで聴覚野近傍の神経応答が変化することを示唆する。

研究成果の概要(英文)：Visual speech cue is important for speech perception. In this study, we focus on formant frequency as acoustic information that visual speech cue conveys. To investigate how to use formant information from articulation, we record EEG and ECoG using synthesized vowels which differs only formant frequency.

In healthy adult EEG, there was no significant difference among vowels. On the other hand, ECoG channels on the left posterior superior temporal gyrus showed significant differences among vowels in 200 ms after sound onset in audio-only condition. These significant differences disappeared in audio-visual condition. This result indicated that visual information may change neural response to sound in vicinity of auditory cortex.

研究分野：認知神経科学

キーワード：視聴覚統合 音声 脳波 皮質内脳波

1. 研究開始当初の背景

音声知覚において、口形のような視覚情報も重要な役割を果たす。口形と一緒に音声を提示すると、母語獲得時の乳児や第2外国語学習時の成人の音声弁別能力が向上することが知られている。この特徴より、失語症のリハビリにも、口の形を模した絵や、実際の口形を見せることが行われている。先行研究では、視聴覚統合は後上側頭溝/回、後中側頭溝/回で処理が行われていること、視覚情報による予測が聴覚野の活動に影響を与え、聴覚のみの提示より視聴覚同時提示に対する反応が低下することが知られている。

一方で、音声には様々な音響要因があるが、口形から音声のどの音響要因の情報を得ているのかは未だわかっていない。そこで、本研究では音の周波数スペクトル上の極大点であるフォルマント周波数に着目した。フォルマント周波数は、低い周波数から順に第1フォルマント(F1) 第2フォルマント(F2) ...と呼ばれ、母音の質の違いはF1とF2、子音はF3より高い周波数によって決まることが知られている。先行研究で、フォルマント周波数の情報を視覚情報から得ているというモデルが提唱されているが、実験的には実証されていない。

2. 研究の目的

本研究では、母音弁別時において口形から音響要因、特にフォルマント周波数の情報を得ているという仮説を立て、これを検証することを目的とする。

音響合成の手法を用い、母音/u/のF1を変えることで母音/a/を、F2を変えることで母音/i/を作成する(図1)ことで、フォルマント周波数のみが異なる刺激音を用いて、母音弁別課題を実施し、課題遂行時の脳波及び脳波よりも空間分解能に優れる皮質内脳波を計測する。課題は音声のみの提示(A課題)と発話時の映像との同時提示(AV課題)からなる。母音間、課題間に差が認められた場合も、刺激音は上述の通りフォルマント周波数のみが異なるので、得られた差はフォルマント周波数の違い、もしくは提示方法の違いであると推測できる。

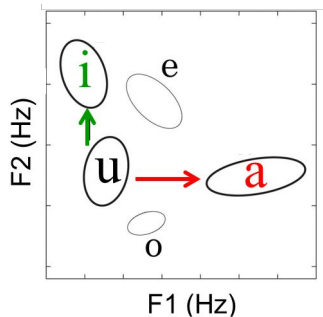


図1 フォルマント周波数

母音/u/の F1 を変えることで母音/a/を、F2 を変えることで母音/i/を作成した。

3. 研究の方法

(1)被験者

脳波は右利きの健常成人19名(男性10名、女性9名 平均年齢24.6歳)に対し、皮質内脳波は右利きの難治性癲癇患者1名(27歳女性)に対し、計測を行った。

(2)刺激

刺激音として、女性日本語話者の発声した母音/u/を採取し、音声分析ソフトを用いて5msごとの第1~第5フォルマント周波数を算出した。フォルマント周波数をパラメータとして音声合成を行うKlatt型音声合成ソフトを用い、算出したフォルマント周波数を持つ長さ180ms(20msのフェードアウト含む)の合成母音/u/を合成した。この合成母音/u/のF1のみ変えて合成母音/a/を、F2のみを変えて合成母音/i/をそれぞれ合成し、これら3種の合成母音/a/、/i/、/u/を聴覚刺激とした。

刺激音を採取した際に撮影した、母音/u/の発話中及び閉口時の顔の静止画を視覚刺激として用いた。撮影は自然光の入らない室内の同一の場所で行い、各画像の明るさやコントラストが著しく変わらないよう配慮した。発話中及び閉口時の計2枚を作成し、対応する画像を順に提示した。本実験では、読唇時に両唇部のみならず顔面全体から読唇情報を得ているという先行研究より、視覚刺激には話者の顔全体を用い、首や髪の毛など顔以外に注意が向かないよう顔以外の背景は同一になるように画像処理ソフトを用いて加工し、各画像の中心に注視点を設けた(図2)。

(3)行動課題

A課題では/u/が70%、/a/、/i/が15%ずつ提示されるようにし、被験者は聞こえた音が/u/か否かをボタン押しで回答した。AV課題ではA課題で用いた音列に、/u/と発話している顔を同時に提示した。被験者は、聞こえた音と口形があっていたか否かをボタン押しで回答した。

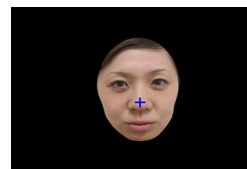


図2 AV課題の一部

画面中央に注視点を置き、音声と同時に口形が閉口から口を開き/u/と発話している画像に、音声提示後に再度閉口の画像に変化する。音声提示から1秒後に注視点の色が赤に変わり、これを合図に被験者はボタンを押して回答する。A条件では注視点以外は背景色とし、被験者は注視点の色の変化を合図にボタンを押して回答する。

(4)解析

脳波は、全頭の 64 チャンネルを対象に事象関連電位の解析、0-50 Hz の活動に対して時間周波数解析を実施した。皮質内脳波は、左右の外側側頭葉後部にある 10 チャンネル、計 20 チャンネルを対象に、事象関連電位の解析、刺激の処理に関連する部位では減衰しそれ以外の部位では増強する 帯域(8-13 Hz)と、刺激の入力に対する活動を反映し、BOLD 信号と相関している high-gamma 帯域(55-140 Hz)の活動に対して時間周波数解析を実施した。

4. 研究成果

健常成人を対象とした脳波計測では、ボタン押しの正答率が 90%を超えたので、正答した試行のみを解析対象とした。しかし、事象関連電位の解析、時間周波数解析の結果から母音間、課題間における有意差は認められなかった。

一方で、皮質内脳波では、被験者である患者に高次脳機能障害があること、実験時に発熱していたことからボタン押しの正答率は 30%程度であったため、全試行を解析対象として解析を実施した。その結果、音の提示から 1 秒以内の左外側側頭葉後部の反応に有意な差が認められた。特に、視覚情報の有無による違いは外側側頭葉後部の中でも後方のチャンネルに、母音の違いは前方のチャンネルに反映された(図 3)。これは、多様な機能を司るといわれる側頭葉後部において、機能ごとに異なる部位が関与している可能性を示唆し、最近報告された側頭回後部が視聴覚情報の統合の重要な役割を担っているという結果を支持する。

また、母音の違いによる差は、A 課題のみで/u/と/a/、/u/と/i/の両方で認められた。差が見られたチャンネルは、左半球の中側頭回にある隣り合うチャンネルであり、これはフォルマント周波数の違いによる音の処理が隣接した異なる部位で行われていることを示唆する。一方で、AV 課題では母音の違いによる差が見られなかった。この理由として、視覚情報により側頭葉の活動が抑制された可能性、視覚情報によって A 課題よりも/u/に近い音として知覚され、母音間における反応の差が有意には至らなかった可能性が考えられる。いずれにせよ、特定のフォルマント周波数ではないが、口形という視覚情報があることで聴覚野近傍の神経応答は変化し、この変化は音声の提示から 200ms 以内に生じること示唆する。

本研究では動画ではなく画像を用いたり、不自然さの残る合成音を用いたため、自然な音声聴取時の脳活動を反映しているとはいえない。また、皮質内脳波では顕著な結果が得られたが、1 例のみであるので、得られた結果を一般化することはできない。しかし、1 例であっても、刺激提示の差、刺激音による差は十分に見られ、これは ECoG の感度の

高さによるものであると考える。今後も、ECoG を用いた知覚・認知の研究が実施され、更なる脳機能の解明が進むことが期待される。

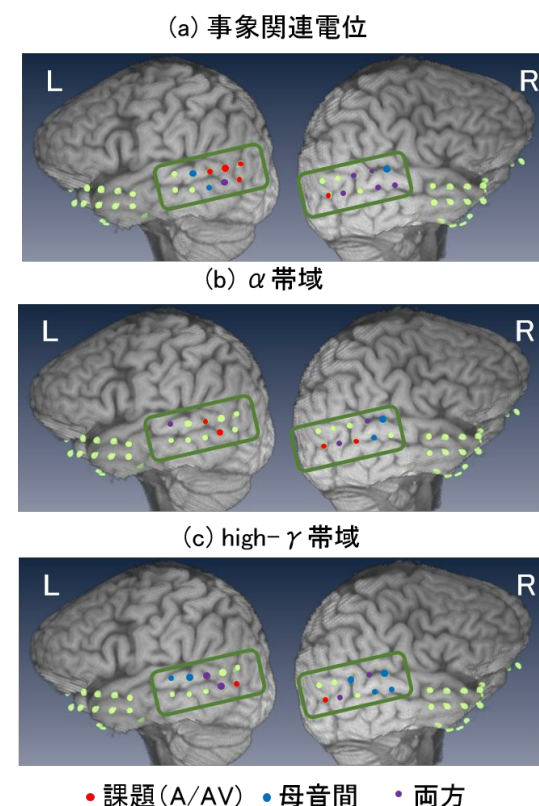


図 3 皮質内脳波で有意差が見られたチャンネル

被験者の 3D MRI 上に皮質内脳波電極を重ねた図を示す(緑の枠内にある左右各 10 チャンネルが解析対象)。図は上から(a)事象関連電位、(b)帯域、(c)high-帯域を対象にした解析の結果、有意差の見られたチャンネルの色を変えて表示する。赤は課題間、青は母音間、紫は課題間、母音間の両方で有意差が認められたチャンネルを示す。

5. 主な発表論文等

(研究代表者、研究分担者及び連携研究者には下線)

[学会発表](計 1 件)

水落(遠藤)智美・岩淵俊樹・幕内充・音声知覚における脳神経ダイナミクス：口形提示による影響。第 1 回坂本勉記念神経科学研究会，国立病院機構沖縄病院(沖縄県宜野湾市)，2016 年 2 月 20 日。

[その他]

ホームページ等

<http://www.rehab.go.jp/ri/noukinou/hibrain/research/index.html>

6 . 研究組織

(1)研究代表者

遠藤(水落)智美 (MIZUOCHI-ENDO,
Tomomi)

国立障害者リハビリテーションセンター
(研究所)・研究所・脳機能系障害研究部・
学術振興会特別研究員 (RPD)

研究者番号 : 90568859