

研究概要報告書【サウンド技術振興部門】

(/)

研究題目	連続音声の時間情報処理に係る聴覚機能の客観的評価法に関する検討	報告書作成者	田村俊介
研究従事者	田村俊介		
研究目的	<p>我々が連続音声から言語情報を正しく知覚するためには、音声信号に含まれる様々な速さの時間変動成分を正確に抽出する必要がある。特に、10 Hz 以下の時間変動成分(音節の出現頻度と一致)である振幅包絡成分と 50 Hz 以上の時間変動成分である時間微細構造(基本周波数と一致)を捉えることが重要である。振幅包絡成分は音声の聴き取りにおいて重要な手がかりになることが知られており(e.g. Drullman et al., 1994a, b; Elliott et al., 2009; Shannon et al., 1995), 時間微細構造は主に雑音環境下において、雑音の中から音声の情報を抽出する際の手がかりになることが知られている(e.g. Eaves et al., 2011; Hopkins & Moore, 2009; Mahajan et al., 2017)。</p> <p>これらの時間情報を処理する脳内メカニズムについて、律動的に生じる神経活動である神経オシレーションが注目されている(Arnal et al., 2016)。振幅包絡成分に同期して生じる神経オシレーションとして、低周波(主に 4-8 Hz)の θ オシレーションが知られており、聴取する音声が高い明瞭度を持つ際に、θ オシレーションの振幅包絡に対する同期度が高くなることが報告されている(Etard & Reichenbach, 2017; Peelle et al., 2013)。時間微細構造に関しては、音声の持つ基本周波数に同期して生じる高周波(主に 30 Hz 以上)の γ オシレーションが計測されることが知られている(Coffey et al., 2016, 2021; Ross et al., 2020)。そして、音声の持つ基本周波数の神経表現は、雑音下での音声聴取能力と相関することが知られている(Anderson et al., 2011; Song et al., 2010)。</p> <p>上記のように、振幅包絡成分や時間微細構造は音声を聴取する上で重要な手がかりであり、それらに同期して生じる神経オシレーションと音声聴取との関連が示唆されている。しかしながら、時間微細構造に同期する γ オシレーションは主に単音節を用いて評価されることが多いのに対し、振幅包絡成分に同期して生じる θ オシレーションの計測には連続的な音声信号の聴取が必要となるため、それぞれ別々に評価されることがほとんどである。そのため、両者の処理の間の相互作用や音声聴取における役割分担については明らかになっていないことが多い。そこで本研究では、連続的な音声信号に含まれる振幅包絡成分と時間微細構造の処理に係る神経オシレーションを同時に計測し、それらと日常的な音声聴取能力との関係性を検討することで、連続音声の時間情報処理に係る聴覚機能の客観的評価法を確立することを目的としている。</p>		

研究内容	<p>実験方法</p> <p>参加者 26名の健常者(男性13名,女性13名,平均37歳)と26名の統合失調症者(男性14名,女性12名,平均38歳)の合計52名が実験に参加した。統合失調症者については,聴覚刺激に誘発されるγオシレーションに異常が見られることが数多くの研究で報告されているため(e.g. Spencer et al., 2009; Hirano et al., 2015),その異常が音声に含まれる時間情報処理や音声聴取能力にどのような影響を及ぼすのかを調べた。全ての被験者は正常聴力であり(500, 1000, 4000 Hzの聴力レベルが25 dB以下),健常者1名を除き右利きであった。日常的な場面での音声聴取能力はSpeech, Spatial, and Qualities of Hearing Scale(SSQ)(Gatehouse & Noble, 2004)を用いて評価した。</p> <p>刺激 国立情報学研究所が無償で提供している音声データベース(PASL-DSR)内に含まれる,男性話者1名が発した単語音声「いちぶぶん」を聴覚刺激に用いる。ただし,収録された発話音声は基本周波数が複雑に変化する音声であるために,基本周波数に同期して生じるγオシレーションを安定して計測することが難しい。そこで,音声分析合成ソフトウェアPraatを用いて基本周波数を80 Hzに固定した分析合成音声(monotone speech,以下MS)を刺激に用いることで連続音声を用いてのγオシレーション計測を可能とした。また,音声の時間微細構成成分を削除し,振幅包絡成分の情報のみを保持する分析合成音声(Noise-vocoded speech,以下NVS)やモノトーン音声と似た時間構造を持つ非音声刺激(amplitude-modulated click train,以下AMC)も刺激に用いた。具体的な刺激の特徴は説明書内の図1に示す通りで,MS,AMCは80 Hzに大きなパワーを持つ一方で,MSから時間微細構造の取り除いた刺激であるNVSでは80 Hzのパワーは持たない。また,4-8 Hzの振幅包絡特徴について調べたところ,どの刺激も5 Hzの変調周波数にピークを持つ振幅包絡を持つことが確認された。</p> <p>脳波測定 64 chの脳波計(BioSemi, ActiveTwo)を用いて,上記3種類の刺激を聴取している際の脳活動を計測する。各刺激で150回の呈示を行い,約5分間の計測を行なった。実験参加者は,安静閉眼を保ち,分析合成音声に注意を向けるように教示された。</p> <p>データ分析</p> <p>脳波データにウェーブレット変換を用いた時間周波数解析を行うことで,時間微細構造に関連したγオシレーションの解析を行った。さらに,信号源解析を行うことでその発生源となる脳領域を特定した。信号源解析で求めたγオシレーションの振幅包絡の特徴を調べ,刺激の振幅包絡との関連性を相互相関分析で調べた。最終的に,γオシレーションやその振幅包絡に関わる特徴量とSSQの成績の相関を調べることで,音声の時間情報に関わる神経指標と音声聴取能力の関係性について検討した。</p>
------	---

研究のポイント	近年の「隠れ難聴」への高い関心を受け、数多くの研究で聴覚時間情報処理の評価方法が検討されている(古川, 2016)。しかしながら、日常生活で耳にする音とは大きく性質が異なる聴覚刺激を用いられていることが多いこともあり、未だ臨床場面での聴覚機能の評価は「隠れ難聴」を捉えられない純音聴力検査に依存している。本研究課題で行う実験のように、日常的に行う音声コミュニケーションに即した形での刺激設定で聴覚時間情報処理能力を評価することが可能となれば、「隠れ難聴」を検出するための聴覚検査法の臨床応用が促進されるとともに、その根源的な予防法や治療法の開発促進に繋がることが期待される。連続音声聴取時の神経活動については、振幅包絡成分に同期した θ オシレーションの性質が盛んに調べられている。一方、時間微細構造に同期して生じる γ オシレーションについては、連続音声を刺激に用いた場合、基本周波数の変動が大きく、安定した計測が難しいため、主に単音節が刺激に用いられる。連続音声をを用いてそれら両方を計測した研究は数限りあり、さらに、音声の聴取能力の関係性を調べた研究は申請者の知る限りでは存在しない。
研究結果	脳波データにウェーブレット変換を用いた時間周波数解析を適用したところ、主に Fz 電極において MS, AMC 刺激が持つ 80 Hz のパワーに同期した γ オシレーションが計測された(図 2)。健常者(NC)と統合失調症者(SZ)の結果を比較すると、SZ の γ オシレーションの活動が NC よりも低いことが確かめられた。続いて、 γ オシレーションの信号源解析を行ったところ、NC において左半球(LH)の上側頭溝付近(図 3 緑枠)で強い活動が見られた(図 3)。SZ 群では健常者に比べて、LH の上側頭溝における γ オシレーションの活動が有意に低いことが分かった。さらに、上側頭溝における γ オシレーションの振幅包絡を調べたところ、MS, AMC 条件では、主に LH において刺激の振幅包絡と一致する 5 Hz の振幅包絡が見られることが分かった(図 4)。 γ オシレーションの振幅包絡と刺激の振幅包絡との相互相関分析を行なったところ、MS, AMC 条件で高い相関が見られたため、 γ オシレーションの振幅包絡が聴取した刺激の振幅包絡を反映していることが示唆された。パワー(5 Hz)、相互相関ともに、MS 条件において、NC よりも SZ の方が有意に低くなることが分かった。最後に、上記で求めた指標について、SSQ の speech, spatial, quality 及びそれらの合計点数との相関解析を調べたところ、MS 条件における γ オシレーションの活動やその振幅包絡と刺激の振幅包絡との相互相関係数が quality の得点数との有意な正の相関が見られることが分かった。つまり、時間微細構造と振幅包絡を反映した神経オシレーション指標が日常的な音声聴取能力と相関することが分かった。
今後の課題	本研究では、連続音声刺激を用いて、その時間的特徴である振幅包絡、時間微細構造に関連した神経オシレーションの計測を行い、さらにその結果から音声の聴取能力を予測できる可能性があることを示した。しかしながら、当初の計画に含まれていた雑音環境下での音声聴取実験を実施することが出来なかった。さらに、当初予定していた実験参加者(100 名)の半分程度しかデータの収集ができなかったため、被験者の背景情報やその他の認知機能などを考慮した上で、神経オシレーションの分析結果から音声聴取能力を予測するモデルを立てるといった段階まで計画を進めるに至らなかった。今後は、音声の時間情報処理に関わる聴覚機能の客観的評価方法を確立するという目標を達成するために、実験デザインを修正し、被験者数をさらに増やしていく予定である。

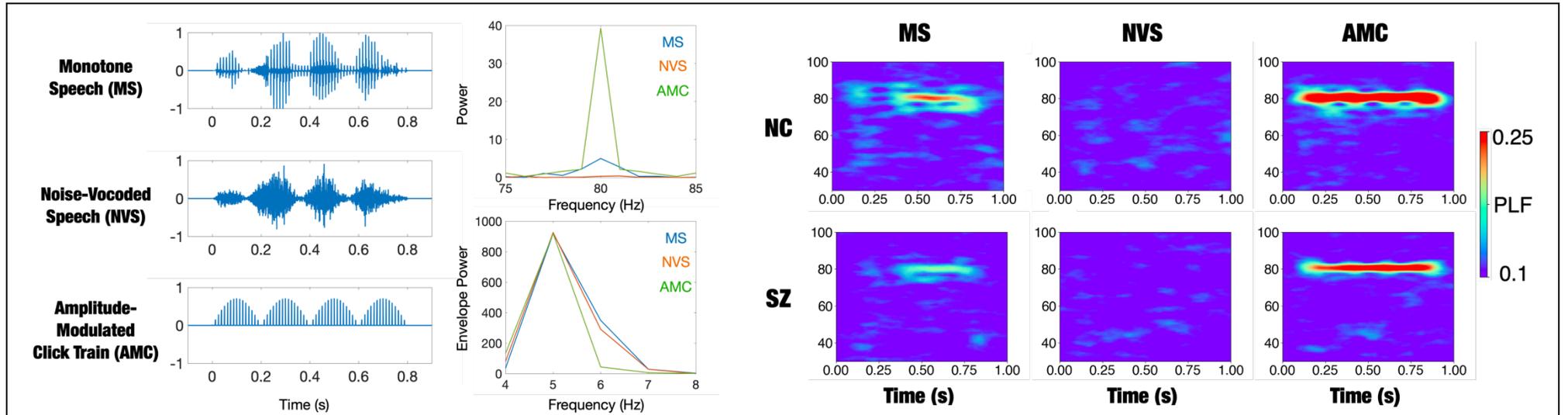


図 1. 刺激の波形と時間特徴

図 2. γ オシレーションの分析結果 (Fz 電極)

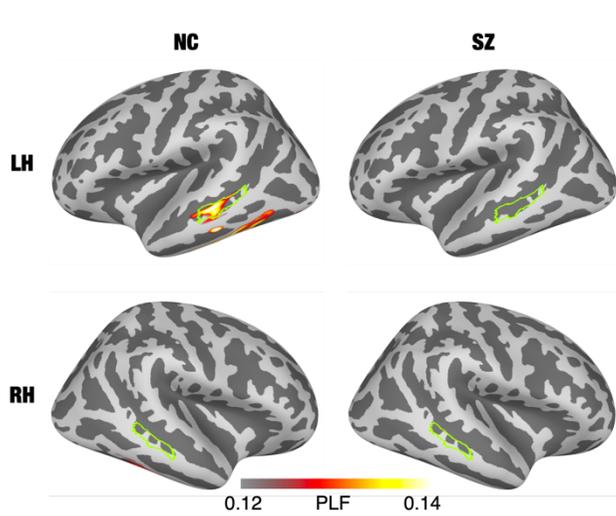


図 3. MS 刺激に対する γ オシレーションの信号源

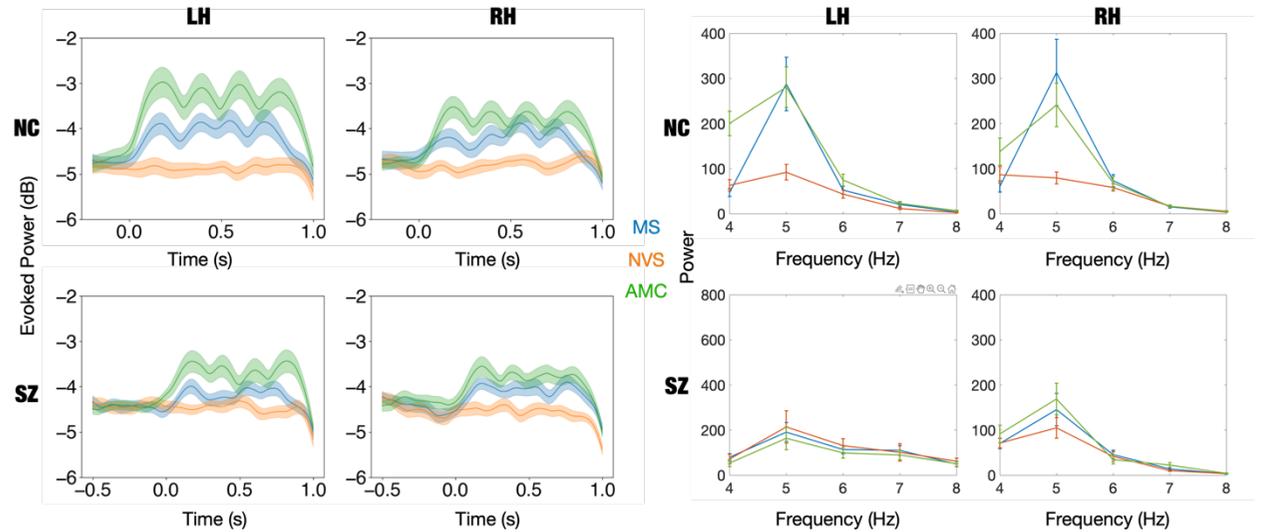


図 4. γ オシレーションが持つ振幅包絡とパワー