

Title	目的音と背景音のスペクトル変調・時間変調情報を用いた聴覚探索の研究
Author(s)	堀口, 遼太郎
Citation	
Issue Date	2024-03
Type	Thesis or Dissertation
Text version	author
URL	http://hdl.handle.net/10119/18916
Rights	
Description	Supervisor: 木谷 俊介, 先端科学技術研究科, 修士(情報科学)

修士論文

目的音と背景音のスペクトル変調・時間変調情報を用いた
聴覚探索の研究

堀口 遼太郎

主指導教員 木谷 俊介

北陸先端科学技術大学院大学
先端科学技術研究科
(情報科学)

令和5年3月

Abstract

In our daily lives, listening to sound improves our lives in many ways. Listening to music for healing and enjoying conversation are examples of this. Sound can take in information actively or passively. This information can be used to defend against foreign enemies or to plan actions to take.

Of all the sounds used to improve human life, sound sign are particularly familiar to humans. Among these, alarm sounds are often used. Sound alarms play an important role in comfort because they guide people's behavior. To ensure that the message of the alarm sound is conveyed to all, the alarm sound itself is designed in accordance with guidelines and JIS, which indicate universal design considerations.

These have been studied in noise-free environments. However, noise exists in the real environment. The alarm sound may not convey the message correctly because of interference from the noise. Therefore, it is necessary to ensure that the alarm tone message is conveyed correctly even in the presence of noise. In order for the alarm sound message to be conveyed correctly in the presence of noise, the relationship between noise and alarm sound must be known. The problem of listening to an alarm sound in the presence of noise is defined as an auditory search problem for detecting the target sound in the presence of background noise. The important finding is that the target sound is correctly detected when the message of the alarm sound is correctly conveyed.

Previous research of the auditory search task have focused on the fluctuation of the amplitude envelope of the target and background sounds. Sugita investigated this by mixing time-varying sounds (amplitude-modulated sounds, frequency-modulated sounds, narrow-band noise) and pure tones in a single channel. As a result, the search time increased with the number of interfering sounds when the pure sound was used as the target sound, while the search task for time-varying sounds was independent of the number of pure sounds. They claimed that the temporally fluctuating sound may have popped out. Asemi examined the asymmetry of the auditory search task in the presence or absence of variation in the target and background sounds. The results indicate that it is easier to detect a changing target sound relative to an unchanging background sound than it is to detect an unchanging target sound relative to a changing background sound. Kusaba considered the auditory search problem when both the target and background sounds are variable. The results showed that the target sound is easily detected when the correlation coefficient between the amplitude envelopes of the target and background sounds is low. The results showed that the target sound is more likely to be detected when the correlation coefficient between the amplitude envelope of the target and background sounds is low. Yano defined the movement of the amplitude envelope, proposed a method to control it, and examined the auditory

search problem focusing on differences in movement.

These results indicate that temporal fluctuations in sound are important for the auditory search. On the other hand, humans perceive sounds that fluctuate not only in the temporal direction but also in the frequency direction. Spectro-temporal modulation (STM) information analysis is a method for capturing both temporal and frequency variations of sound. This is a method in which variations in the amplitude of a sound in the time direction are indicated as time modulation (TM) information, and variations in the spectrum in the frequency direction are indicated as spectral modulation (SM) information. If stimuli can be created considering TM and SM information, target and background sounds can be created considering fluctuations in the temporal and frequency directions. These stimuli can be used to present both temporal and frequency perspectives on the auditory search.

This paper aims to clarify the relationship between easily detectable target and background sounds using spectro-temporal Modulation (STM) information. Experimental stimuli generated created based on the STM information. Auditory search tasks were conducted using the generated stimuli. The results of the auditory search task will be used to examine the relationship between target sound and the background sounds that are likely to be detected. Once this relationship is clarified, we can expect to conduct a simulated evaluation of auditory search using STM information. It also leads to the consideration of guidelines for designing sound alarms as an engineering application.

Three experiments were considered in this study; 1. Experiment focusing on SM information; 2. Experiment focusing on TM information; 3. Experiments focusing on the SM and TM information of target and background sounds, respectively.

In the experiment focusing on SM information, a one-way ANOVA with the center frequency of the background noise as a factor showed no main effect. Thus, the relationship between the SM information produced by narrow-band noises with two different center frequencies indicated that the target sound was easily detectable.

In the experiment focusing on TM information, as in the previous experiment, a relationship of ease of detection when the amplitude envelopes of target and background sounds are different. Not only that, but it was also found that the carrier wave affected the ease of detection. However, it was not clear what affects the carrier wave. Therefore, experiments focusing on the bandwidth of the carrier wave and experiments focusing on the center frequency of the carrier wave were conducted. The results showed that auditory search tasks with different amplitude envelopes were more likely to be detected when the bandwidth was wide.

In experiments focusing on the SM and TM information for the target and

background sounds, respectively, it was shown that the bandwidth condition affects the detectability of the target sound, even when the amplitude envelope movement is the same.

These results indicate the following three relationships between target and background sounds. 1. the target sound and the background sound have a relationship with different amplitude envelopes; 2. the center frequencies of the narrow-band noise used for the target sound and the background sound are different; and 3. the relationship is that the width of the target sound is narrower than the spectral width of the spectral modulation information of the background sound.

STM information is known to be related to the spectro-temporal receptive field (STRF). This finding and the experimental results suggest that the auditory search task cued not only information processing in the auditory peripheral system but also features processed in the STRF.

目次

第1章	序論	1
1.1	背景	1
1.2	関連研究	3
1.3	関連研究の問題点	4
1.4	本研究の目的	5
1.5	本論文の構成	5
第2章	STM 情報分析の概要	7
2.1	STM 情報分析の手順	7
2.2	STM 情報でわかっている知見	10
第3章	2種類の音が混合された刺激の STM 情報分析	12
3.1	分析の目的	12
3.2	異なる周波数成分を持つ2音の STM 情報分析	12
3.3	異なる振幅包絡線成分を持つ2音の STM 情報分析	17
3.4	2種類の音が混合された刺激の STM 情報のまとめ	25
第4章	STM 情報をもとに作成した刺激を用いた聴覚探索実験	26
4.1	実験の目的	26
4.2	実験環境	26
4.3	実験参加者	26
4.4	実験方法	29
4.5	評価方法	29
4.6	実験刺激	29
4.7	SM 情報に着目した実験	30
4.8	TM 情報に着目した実験	32
4.8.1	振幅包絡線に着目した実験	32
4.8.2	振幅包絡線と搬送波の帯域幅に着目した実験	36
4.8.3	振幅包絡線と搬送波の中心周波数に着目した実験	40
4.9	目的音と背景音それぞれの TM 情報と SM 情報に着目した実験	42
4.9.1	実験刺激のパラメータ	42
4.9.2	実験結果	42

第5章 総合考察	44
第6章 結論	45
6.1 本研究で明らかにしたこと	45
6.2 残された課題	45
参考文献	47
謝辞	51

目次

1.1	論文の構成	6
2.1	STM 情報分析のブロックダイアグラム. (a) は入力信号, (b) はスペクトログラム, (c) は STM 情報を示す.	9
2.2	STM 情報と音声の特徴の関係を表した図 (文献より引用した). 緑色の領域は言語的意味, 青色は性別, 橙色は領域音の粗さを示す.	11
3.1	中心周波数が 1050 Hz と 300 Hz の狭帯域雑音の 2 音が混合された刺激の STM 情報. 横軸は TM 情報, 縦軸は SM 情報, カラーバーは振幅パワーを示す.	13
3.2	中心周波数が 1050 Hz と 525 Hz の狭帯域雑音の 2 音が混合された刺激の STM 情報. 横軸は TM 情報, 縦軸は SM 情報, カラーバーは振幅パワーを示す.	14
3.3	中心周波数が 1050 Hz と 1600 Hz の狭帯域雑音の 2 音が混合された刺激の STM 情報. 横軸は TM 情報, 縦軸は SM 情報, カラーバーは振幅パワーを示す.	15
3.4	中心周波数が 1050 Hz と 2100 Hz の狭帯域雑音の 2 音が混合された刺激の STM 情報. 横軸は TM 情報, 縦軸は SM 情報, カラーバーは振幅パワーを示す.	16
3.5	搬送波の f_c が 525 Hz, BW が 1/2 oct の時の 2 音を混合した目的音と背景音の振幅包絡線が異なるときの STM 情報. 横軸は TM 情報, 縦軸は SM 情報, カラーバーは振幅パワーを示す.	19
3.6	搬送波の f_c が 525 Hz, BW が 1/2 oct の時の 2 音を混合した目的音と背景音の振幅包絡線が異なるときの STM 情報. 横軸は TM 情報, 縦軸は SM 情報, カラーバーは振幅パワーを示す.	20
3.7	搬送波の f_c が 2100 Hz, BW が 1/2 oct の時の 2 音を混合した目的音と背景音の振幅包絡線が異なるときの STM 情報. 横軸は TM 情報, 縦軸は SM 情報, カラーバーは振幅パワーを示す.	21
3.8	搬送波の f_c が 525 Hz, BW が 132 Hz の時の 2 音を混合した目的音と背景音の振幅包絡線が異なるときの STM 情報. 横軸は TM 情報, 縦軸は SM 情報, カラーバーは振幅パワーを示す.	22

3.9	搬送波の f_c が 1050, BW が 132 Hz の時の 2 音を混合した目的音と背景音の振幅包絡線が異なるときの STM 情報. 横軸は TM 情報, 縦軸は SM 情報, カラーバーは振幅パワーを示す.	23
3.10	搬送波の f_c が 2100, BW が 132 Hz の時の 2 音を混合した目的音と背景音の振幅包絡線が異なるときの STM 情報. 横軸は TM 情報, 縦軸は SM 情報, カラーバーは振幅パワーを示す.	24
4.1	実験間の関連図	27
4.2	実験環境	28
4.3	SM 情報に着目したの実験結果. 横軸は背景音の f_c , 縦軸は d' を示す	31
4.4	振幅包絡線に着目した実験. 目的音の振幅包絡の傾きが 6.0 の場合. 横軸は背景音の振幅包絡線の傾き, 縦軸は d' , 凡例は背景音の f_c を示す.	33
4.5	振幅包絡線に着目した実験. 目的音の振幅包絡の傾きが 8.5 の場合. 横軸は背景音の振幅包絡線の傾き, 縦軸は d' , 凡例は背景音の f_c を示す.	34
4.6	振幅包絡線に着目した実験. 目的音の振幅包絡の傾き 11.0 の場合. 横軸は背景音の振幅包絡線の傾き, 縦軸は d' , 凡例は背景音の f_c を示す.	35
4.7	振幅包絡線と搬送波の帯域幅に着目した実験の結果. 目的音の振幅包絡の傾きが 6.0 の場合. 横軸は背景音の BW, 縦軸は d' , 凡例は背景音の傾きを示す.	37
4.8	振幅包絡線と搬送波の帯域幅に着目した実験の結果. 目的音の振幅包絡の傾きが 8.5 の場合. 横軸は背景音の BW, 縦軸は d' , 凡例は背景音の傾きを示す.	38
4.9	振幅包絡線と搬送波の帯域幅に着目した実験の結果. 目的音の振幅包絡の傾きが 11.0 の場合. 横軸は背景音の BW, 縦軸は d' , 凡例は背景音の傾きを示す.	39
4.10	振幅包絡線と搬送波の中心周波数に着目した実験の結果. 横軸は背景音の BW, 縦軸は d' , 凡例は背景音の傾きを示す.	41
4.11	目的音と背景音の振幅包絡線も搬送波の帯域幅も異なる実験の結果. 横軸は背景音の振幅包絡線の傾き, 縦軸は d' , 凡例は背景音の BW を示す.	43

表 目 次

3.1	バンドパスフィルタのパラメータと振幅包絡線の頂点間の傾き . . .	18
-----	------------------------------------	----

第1章 序論

1.1 背景

日常生活において、音を聞くことは様々な点から生活を豊かにする。例えば、音楽をや会話を楽しむことなどがあげられる。また、音を用いて、情報を得ることにより、周囲に注意を向けたり、身を守ったり、行動の指針を立てたりする。情報を得る過程において、カクテルパーティ効果 [1][2] のように音を選択的に聴取する過程は、特に人が聴覚系で情報を得る過程で代表的なことだと言える。また、救急者のサイレンのような目立ちやすい音を受動的に処理する聴覚的顕著性という現象は、受動的な情報処理の例と言える。このような聴覚による情報処理の過程を経て、ヒトは音を生活の中に取り入れ、用いることで、生活を豊かにする。

ヒトの生活の中で、使われる音の中でも、特にサイン音は、ヒトの身近にある。その中でも特に報知音はよく利用される。報知音は、家電機器の操作の開始・終了といったメッセージを伝える [3]。そのため、快適な生活を行う上で重要な役割を持つ。報知音の持つメッセージが、すべての人に伝わるように、ユニバーサルデザインに配慮した設計を示すガイドライン [4] や JIS [5] によって、報知音自体が工夫されている。報知音の設計指針を再検討する研究もある。倉片らは、家電機器の報知音の発音パターンについて分析を行った [6]。この分析から、報知音の発音パターンが、同じような動作を示すものでも、製品間で異なっていたり、異なるメッセージを伝えるのに似ている音を使っていたりすることなどが指摘された。水浪らは、報知音の非類似度と印象を調べることによって、適切な報知音の ON/OFF パターンについて検討した [7]。

これらは報知音自体の分析や、雑音がないときの聴取による検討であるなど妨害がないときの検討である。実環境では、雑音による妨害が存在する。報知音は雑音による妨害が原因で、報知音が持つ、メッセージの意味が正確に伝わらないことがありうる。そのため、雑音による妨害があっても、報知音の持つメッセージが正しく伝わるようにする必要がある。雑音下で報知音を正しく伝えるための検討として、空間的な報知音の呈示方法に着目した例がある。

自由音場において、信号音と雑音の到来方向が異なるとき信号音が検知されやすくなることが報告されている [8]。この知見を利用して、実際の環境に近い検討である、車載機器の報知音の呈示方法について検討が行われた [9][10]。これらは、報知音が検知されやすいときの報知音と雑音の位置関係を示した。このようなことから、報知音の検知されやすい空間的な呈示方法について、ある程度の解決方

針が示されている。

ここで、雑音と報知音の音響特徴の関係に着目する。雑音と報知音の報知音が検知されやすいときの関係を、音響特徴を用いて示すことによって、音の到来方向によらない検討が行える。呈示方法による検知されやすい報知音と雑音の関係とともに、音響特徴による検知されやすい報知音と雑音の関係を示すことができれば、報知音が正しく伝わる指針となる。そのために雑音と報知音の関係について、雑音がある中から報知音を聴き取る問題を、複数の刺激の中から、目的刺激を検知する視覚探索に倣って、背景音がある中から目的音を検知する聴覚探索として考える。

視覚探索では、塩入のレビュー論文 [11] に紹介されるように、注意に関する研究を中心に様々なことが明らかになった。例えば、目的刺激が他の刺激と単一の特徴のみで違いを持つ場合ではポップアウト現象が起きるが、複数の特徴を持つ刺激の場合は、ポップアウト現象が起きないことが示されている [12]。また、視覚探索は親近性が深くかかわっており、親近性が高い妨害刺激の中から、親近性が低い妨害刺激を探索することが簡単であることが示されている [13]。聴覚でも、注意の研究も同様に行われているが、聴覚探索を用いた注意の研究は、刺激の呈示方法や実験結果の解釈の難しさから視覚探索ほど注意研究の枠組みとして確立されていない [14]。

1.2 関連研究

聴覚探索は、目的音の中から背景音があるかどうか尋ねる問題であり、目的音と背景音が1つのチャンネルに混合された条件（混合条件）と、空間的な手がかりを検討した条件（空間分離条件）がある。

混合条件について、杉田らは時間変動音（振幅変調音、周波数変調音、狭帯域雑音）と純音を1チャンネルに混合することで、聴取実験を行った[15]。その結果、純音を目的音としたときの探索時間は妨害音の数に依存して増加した。一方で、時間変動音の探索時間は純音の数に依存しない結果となった。彼らは、このことについて、時間変動音がポップアウトしたことによって生じたと主張した。

杉田らの研究をもとに、阿瀬見は、目的音と背景音の、変動の有無で聴覚探索の非対称性を検討した[16]。その結果、構成音数に関係なく、変動がある背景音から変動がない目的音を検知するよりも、変動がない背景音から変動がある目的音を検知するほうが検知されやすいことを示した。この実験の結果から、逐次的に構成音が純音であるか判断していると主張した。

実際にある音は、ほとんどが変動がする。そこで草場は、目的音と背景音の両者に変動がある場合の聴覚探索を検討した[17]。この検討では、振幅包絡線の相関係数を類似度として定義した。その結果、目的音と背景音の振幅包絡線の類似度が低いときに検知されやすいことを示した。この傾向は視覚探索でも見られた傾向[18]と同様の傾向である。

草場の研究では、振幅包絡線の系統立った制御はされていなかった。そこで矢野らは、振幅包絡線の動きを定義し、制御する方法の提案を行い、振幅包絡線の動きの違いに着目した聴覚探索を検討した[19][20]。この時、ヒトの聴覚における、音の動き方の知覚[21][22]に着目して刺激が作成された。この検討の結果、目的音と背景音の振幅包絡線の動きに差があるとき目的音が検知されやすいことを示した。

空間的な条件を手がかりとした聴覚探索について、阿瀬見らは空間的な条件でも、目的音と背景音の変動の有無で聴覚探索の非対称性を検討した[16]。その結果、変動がある背景音から変動がない目的音を検知するよりも、変動がない背景音から変動がある目的音を検知するほうが検知されやすいことを示した。ただし、この時の検知の過程は、混合条件と異なり、時間変動音がポップアップしたと主張した。

Eramudugoll はピアノ伴奏、サイレン、目覚まし時計などの様々な実生活のある音の中から、被験者に目的音となる刺激の情報を呈示して検知させた[23]。その結果、同じ場所から呈示するよりも異なる場所から呈示するほうが検知されやすいことが示された。これは空間的な位置関係が目的音の検知されやすさに影響を与えていることを示唆した。

本研究では、1.1節でも述べたように音の到来方向によらない、目的音が検知されやすいときの目的音と背景音の関係を検討するために、混合条件に着目する。

1.3 関連研究の問題点

先行研究から，振幅包絡線の時間変動が重要であることを示された．しかし，実環境下でヒトは周波数の変動についても捉えてる．これらの研究は，周波数の変動について着目していない．実環境下に近い聴覚探索問題を検討するならば，周波数方向について検討する必要がある．音の時間方向と周波数方向の両方の変動を捉える方法として，spectro-temporal modulation (STM) 情報分析 [24] がある．これは音の時間方向の振幅の変動を時間変調 (TM) 情報とし，周波数方向のスペクトルの変動をスペクトル変調 (SM) 情報として示す方法である．TM 情報と SM 情報を考慮して刺激を作成することができれば，時間方向と周波数方向の両方の変動を考えた目的音と背景音が作成できる．これらの刺激を利用し，聴覚探索問題について，検討することによって，背景音中の目的音の検知されやすい条件について，時間方向と周波数方向両方の視点から示すことができる．

1.4 本研究の目的

本研究の目的はSTM情報を用いて、聴覚探索における、目的音が検知されやすいときの、目的音と背景音の関係を明らかにすることである。本研究では、先行研究[20]の刺激を参考に刺激を作成し、STM分析を行う。作成した刺激を用いて聴覚探索実験を行う。聴覚探索実験の結果から、目的音が検知されやすい目的音と背景音の関係を検討する。この関係が明らかになれば、STM情報を用いた、聴覚探索のシミュレーション評価の実装が期待できる。また工学的応用として報知音設計の指針の検討にもつながる。

1.5 本論文の構成

本論文の構成を図1.1に示す。第1章は序論である。研究背景および関連研究を述べる。これを受けて目的を述べる。第2章では、本研究で扱う分析手法であるSTM情報分析について手順を述べ、STM情報で得られた知見を述べる。第3章では、今回検討する目的音と背景音の関係についてSTM情報を用いて示す。第4章では、この関係をもとに作成した刺激を用いて、聴覚探索実験を行い目的音が検知されやすいときの、目的音と背景音の関係を示す。第5章では、実験結果をもとに考察を述べる。第6章では、明らかにされたことと、今後の課題を述べる。

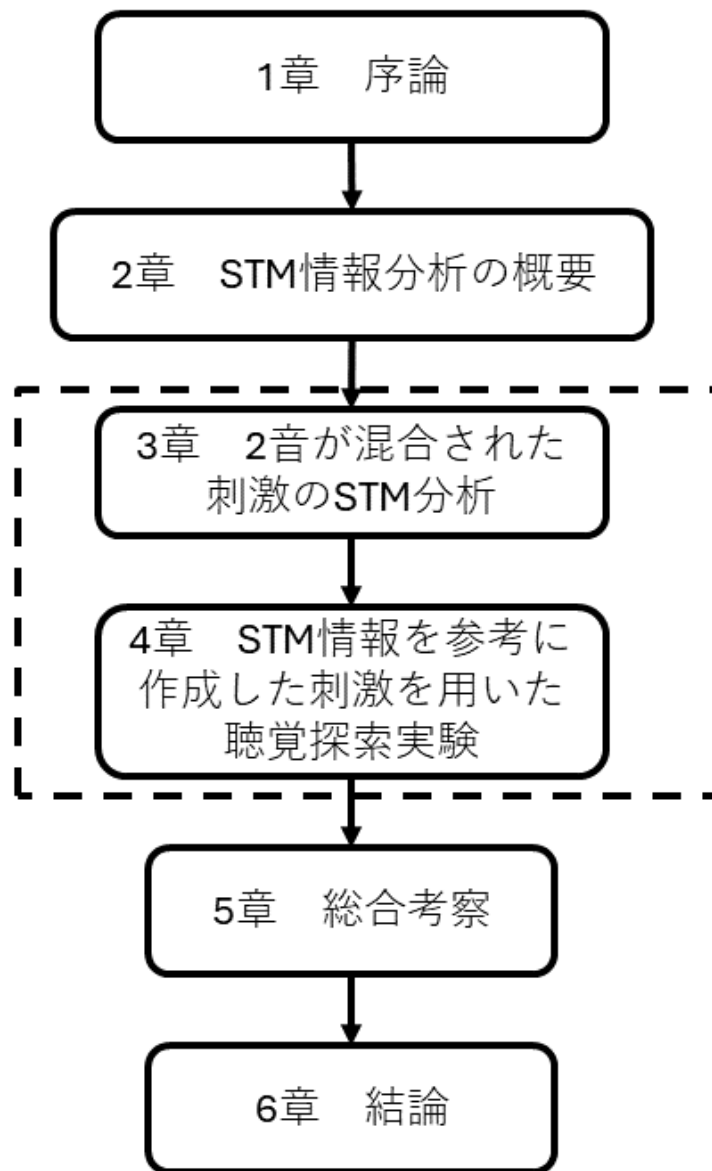


図 1.1: 論文の構成

第2章 STM情報分析の概要

2.1 STM情報分析の手順

本研究では、文献 [24][25] の手法を参考に STM 情報分析を行った。文献 [24][25] では、ガウシアンフィルタを用いて得られたスペクトログラムに対して、2次元フーリエ変換をすることによって STM 情報を求めた。これに対して本稿では、振幅包絡線の動きをより高い分解能で観察したいため、定帯域ガンマトーンフィルタバンクを用いて STM 情報を求めた。STM 情報分析のブロックダイアグラムを、図 2.1 に示す。(a) は入力信号、(b) はスペクトログラム、(c) は STM 情報である。本稿で用いた STM 情報分析では、定帯域ガンマトーンフィルタバンクを利用して、入力信号 $s(n)$ を帯域分割する。 k 番目のチャンネルにおける定帯域ガンマトーンフィルタバンクのインパルス応答 $g_k(n)$ を式 (2.1) に示す。なお、振幅を A 、時間を t 、フィルタの次数を N 、帯域幅を B 、位相を Φ とした。

$$g_k(n) = A(n\Delta t)^{(N-1)} \exp(-2\pi Bn\Delta t) \cos(2\pi f_k n\Delta t + \Phi) \quad (2.1)$$

各フィルタはゲインが -3 dB で交差するように配置された。この時の k 番目のフィルタの出力 $x_k(n)$ は式 (2.2) になる。なお、 $*$ は畳み込み演算子である。

$$x_k(n) = g_k(n) * s(n) \quad (2.2)$$

$x_k(n)$ に対し、式 2.3 のようにヒルベルト変換および 2 乗処理をすることによってパワー包絡線 $e_k^2(n)$ を求める。

$$e_k^2(n) = |x_k(n) + j \cdot \text{Hibert}(x_k(n))|^2 \quad (2.3)$$

ただし、Hibert はヒルベルト変換である。この時、パワー包絡線を、式 (2.4) のように配置することによってスペクトログラムを得る。

$$\mathbf{e}^2(n) = \begin{pmatrix} e_1^2(n) \\ e_1^2(n) \\ \vdots \\ e_k^2(n) \end{pmatrix} \quad (2.4)$$

最後に、パワースペクトログラム $\mathbf{e}^2(n)$ に対して 2 次元フーリエ変換 (2DDFT) することで、STM 情報を得る。

$$\text{STM}(m, l) = 2\text{DDFT}(\mathbf{e}^2(n)) \quad (2.5)$$

ここで、 m は TM 情報のインデックス、 l は SM 情報のインデックスである。これにより、パワースペクトルに対する、周波数方向のスペクトル変化が SM 情報として、時間方向のスペクトル変化が TM 情報として表される。STM 情報は、周波数・時間の両方に対して変化するスペクトル変化が表されたものになる。

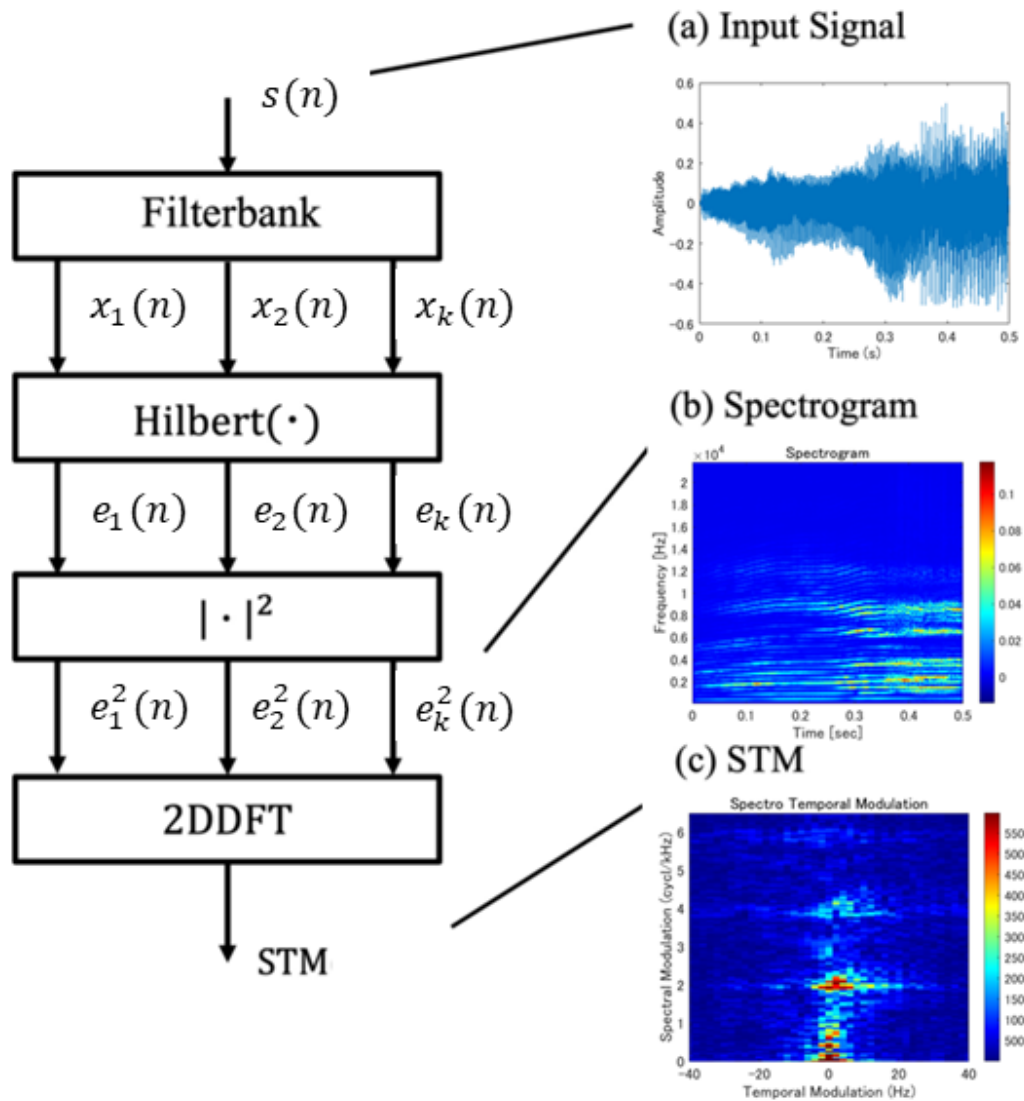


図 2.1: STM 情報分析のブロックダイアグラム. (a) は入力信号, (b) はスペクトログラム, (c) は STM 情報を示す.

2.2 STM情報でわかっている知見

これまでの検討で、音刺激とSTM情報の関係で明らかになっていることを示す。STM情報と音声の関係を図2.2に示す。なおこの図は文献[26]より引用したものである。図2.2の横軸はTM情報、縦軸はSM情報である。TM情報で20 Hz未満の緑色の領域は、語韻などの言語的意味に寄与する領域[27]、SM情報で4 cycle/octaveから6 cycle/octave付近の青色の領域は、話者の性別に寄与する領域[25]、TM情報で30 Hzから150 Hz付近の橙色の領域は、SM情報によらず音の粗さに寄与する領域である[26]。音声にSTM情報分析を利用した研究では、音声の明瞭化を検討した研究や[28] 音声の明瞭度の評価を行った研究[29]、声認識にSTM情報から得られる特徴を用いた研究もある[30]。

これらの他に、本研究の対象である、目的音の検知されやすさに関連する知見として、音の目立ちやすさを表す聴覚的顕著性に関わる知見が報告されている。木所らの報告から、STM情報の部分的な特徴が聴覚的顕著性に関わっていることが示唆された[31]。また田中らの検討では、聴覚的顕著性がSTM情報全体の高次統計量が聴覚的顕著性に関わっていることが示唆された[32]。

STM情報に関して生理学的知見として、高次の変調知覚に係る脳領域であるスペクトル・時間受容野 (spectro-temporal receptive field: STRF) と対応があることが示されている[33][34]。

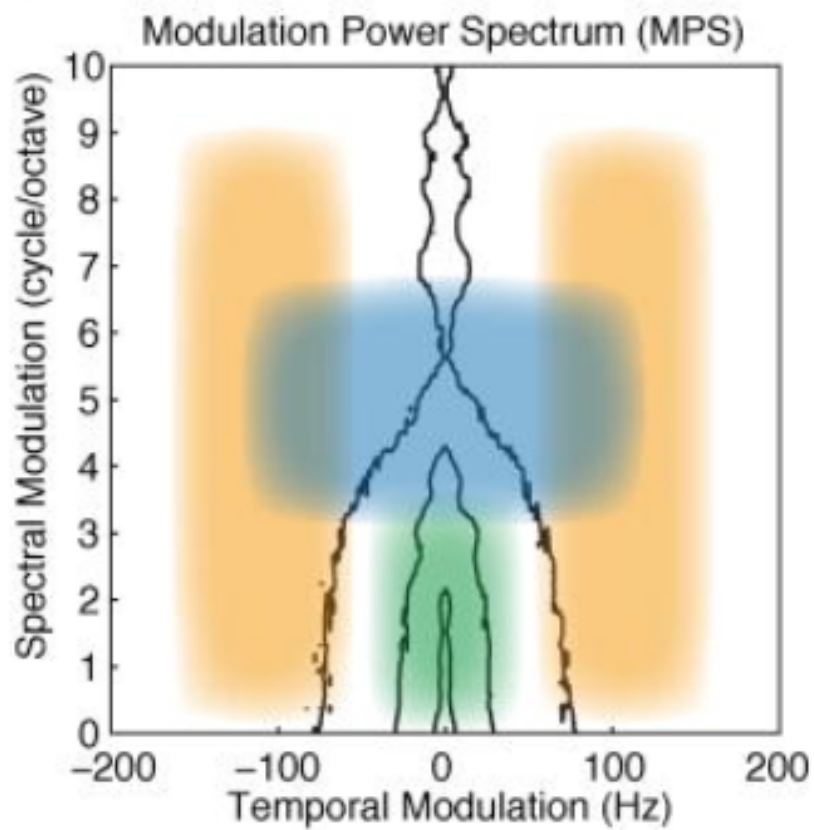


図 2.2: STM 情報と音声の特徴の関係を表した図 (文献 [26] より引用した). 緑色の領域は言語的意味, 青色は性別, 橙色は領域音の粗さに寄与する範囲を示す.

第3章 2種類の音が混合された刺激のSTM情報分析

3.1 分析の目的

分析の目的は聴覚探索実験で使用される目的音と背景音の関係が、STM情報において、どのように表現されるか調べることである。本研究では、先行研究で使用された刺激の作成方法に着目して、SM情報とTM情報にどのような関係が現れるか調べた。

3.2 異なる周波数成分を持つ2音のSTM情報分析

SM情報の検討について、木所の研究では、純音で構成された調波性をもつ音について検討された [35]。また、先行研究 [19][20] では狭帯域雑音を振幅変調した音が使用された。本研究でも、これに倣い、狭帯域雑音を振幅変調した音を実験刺激とする。異なる中心周波数成分を持つ狭帯域雑音を用いた検討を行った。中心周波数が 1050 Hz の狭帯域雑音と中心周波数が、300, 525, 1600 2100 Hz の狭帯域雑音の2音を混合したときのSTM情報を図 3.1 と図 3.2 と図 3.3 と図 3.4 に示す。横軸はTM情報、縦軸はSM情報、カラーバーは振幅パワーである。

これらの結果から、中心周波数の差に対応したSM情報にピークが現れることを確認した。

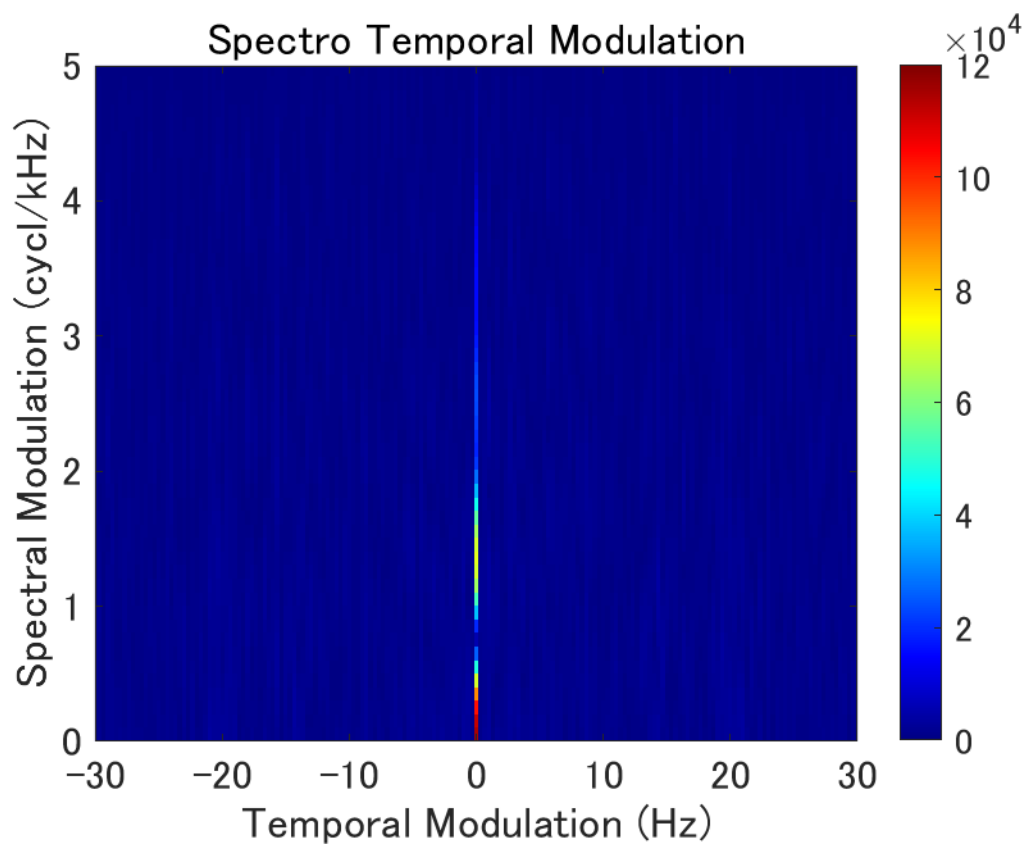


図 3.1: 中心周波数が 1050 Hz と 300 Hz の狭帯域雑音の 2 音が混合された刺激の STM 情報. 横軸は TM 情報, 縦軸は SM 情報, カラーバーは振幅パワーを示す.

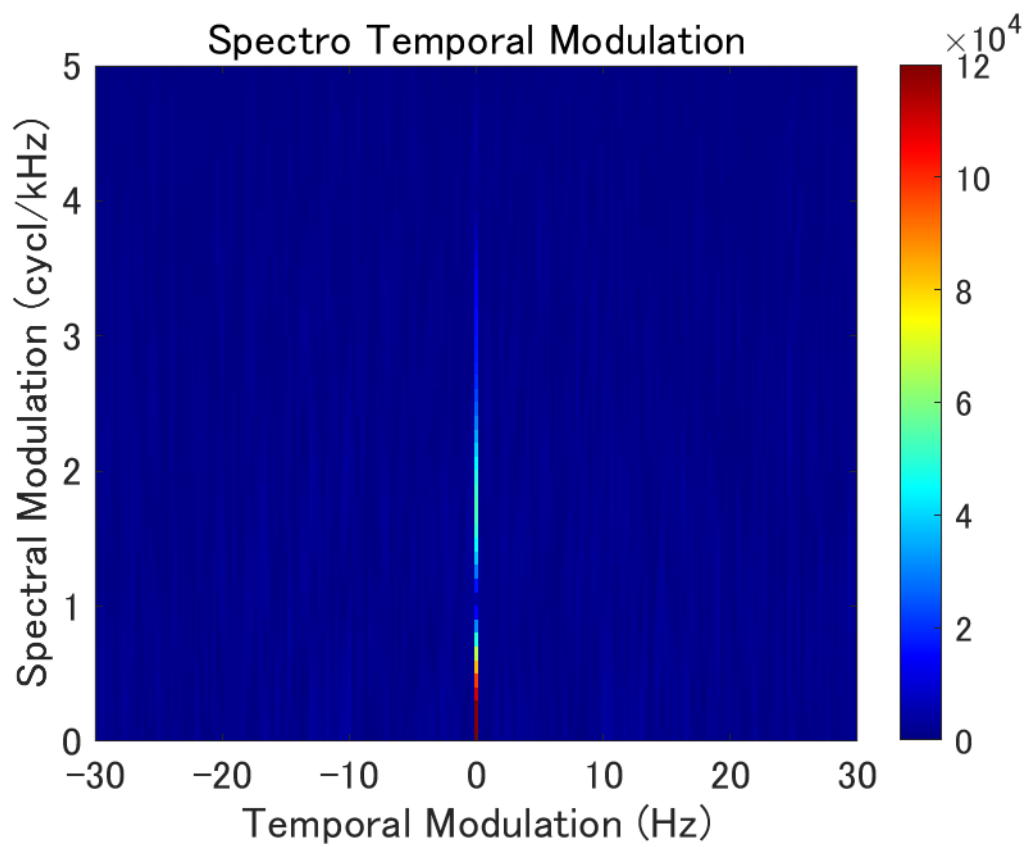


図 3.2: 中心周波数が 1050 Hz と 525 Hz の狭帯域雑音の 2 音が混合された刺激の STM 情報. 横軸は TM 情報, 縦軸は SM 情報, カラーバーは振幅パワーを示す.

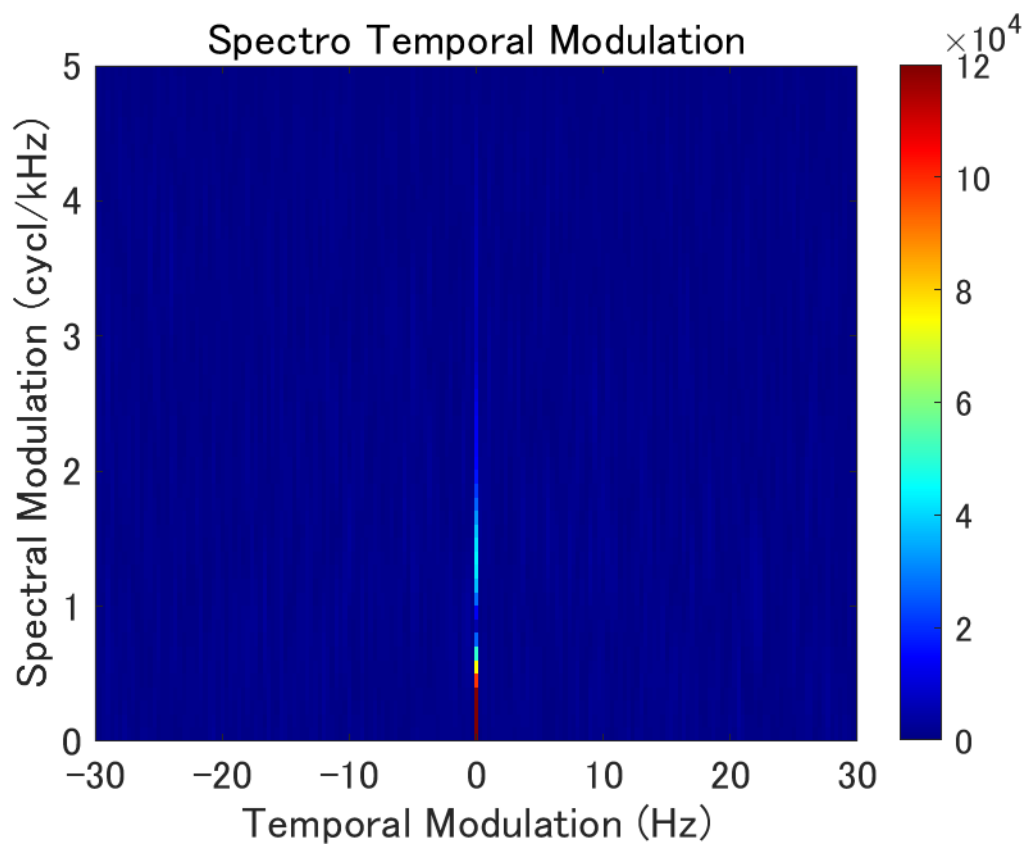


図 3.3: 中心周波数が 1050 Hz と 1600 Hz の狭帯域雑音の 2 音が混合された刺激の STM 情報. 横軸は TM 情報, 縦軸は SM 情報, カラーバーは振幅パワーを示す.

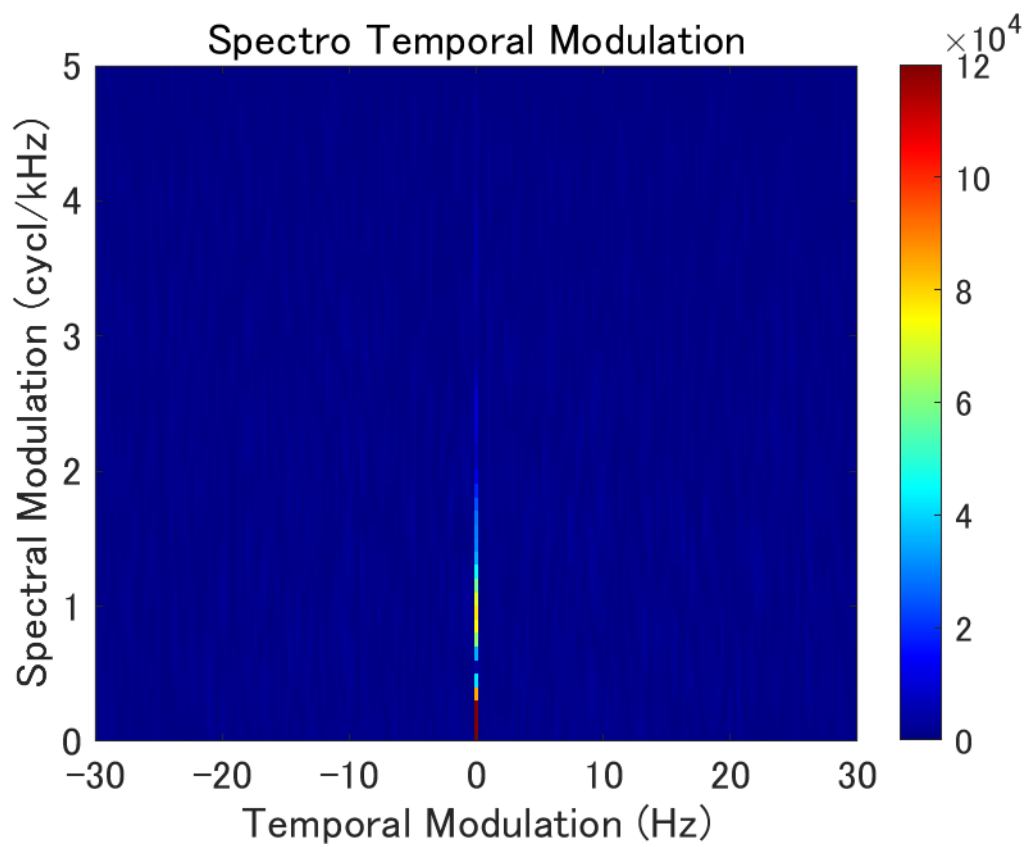


図 3.4: 中心周波数が 1050 Hz と 2100 Hz の狭帯域雑音の 2 音が混合された刺激の STM 情報. 横軸は TM 情報, 縦軸は SM 情報, カラーバーは振幅パワーを示す.

3.3 異なる振幅包絡線成分を持つ2音のSTM情報分析

目的音と背景音の振幅包絡線が、異なるときのSTM情報を調べるために、先行研究[19][20]を参考に刺激を作成した。白色雑音をダウンサンプリングし、FIR型のバンドパスフィルタを通すことで刺激の振幅包絡線を作成した。この時にバンドパスフィルタの帯域幅と中心周波数を変化させることで、振幅包絡線の頂点間の傾きを操作した。その後、得られた振幅包絡線の振幅値の平均を0、変動幅を ± 0.5 に正規化し、元のサンプリング周波数にアップサンプリングした。最後に振幅包絡線と任意の搬送波を掛けることにより刺激を作成した。作成した刺激はすべてサンプリング周波数 20000 Hz、刺激長は3.0秒であった。ダウンサンプリングはサンプリング周波数 100 Hzで行った。搬送波は狭帯域雑音であった。この搬送波の中心周波数を f_c 、帯域幅をBWとした。

FIRフィルタの各種パラメータは、先行研究[19][20]で使用されたものを使用した。各パラメータを表3.1に示す。これらの値は、ヒトが追従できる振幅包絡線の変調周波数が20 Hzまでである[21][22]ということから設定されている。

図3.5に2音の振幅包絡線が異なるときのSTM情報の例を示す。横軸はTM情報、縦軸はSM情報、カラーバーは振幅パワーである。以降のすべての図でも同様である。振幅包絡線の頂点間の傾きが6.0の刺激と、8.5の刺激を混合した。搬送波の f_c が1050 Hz、BWが1/2 octの狭帯域雑音である。この図から、目的音振幅包絡線の頂点間の傾きに対応した、TM情報の2か所からSM情報に向かって、スペクトルが現れるのがわかる。また同様の条件で、BWを変更した例を図3.6と図3.7に示す。なおこの時の2音の振幅包絡線と搬送波の f_c は図3.5と同じである。これらの図からBWが狭まるにしたがって、SM情報のスペクトルが広がるのが分かる。搬送波のBWを132 Hzに固定して f_c を変更させたSTMを図3.8と図3.9と図3.10に示す。なおこの時も、2音の振幅包絡線は図3.5と同じである。これらは f_c が異なるときは、STM情報に系統だった特徴は現れなかった。

表 3.1: バンドパスフィルタのパラメータと振幅包絡線の頂点間の傾き

中心周波数 (Hz)	帯域幅 (Hz)	振幅包絡線の 頂点間の傾き
2.00	1.0	1.0
4.75	1.5	3.5
9.25	1.5	6.0
11.50	1.0	8.5
14.25	1.5	11.0
16.25	1.5	13.5
20.00	1.0	16.0

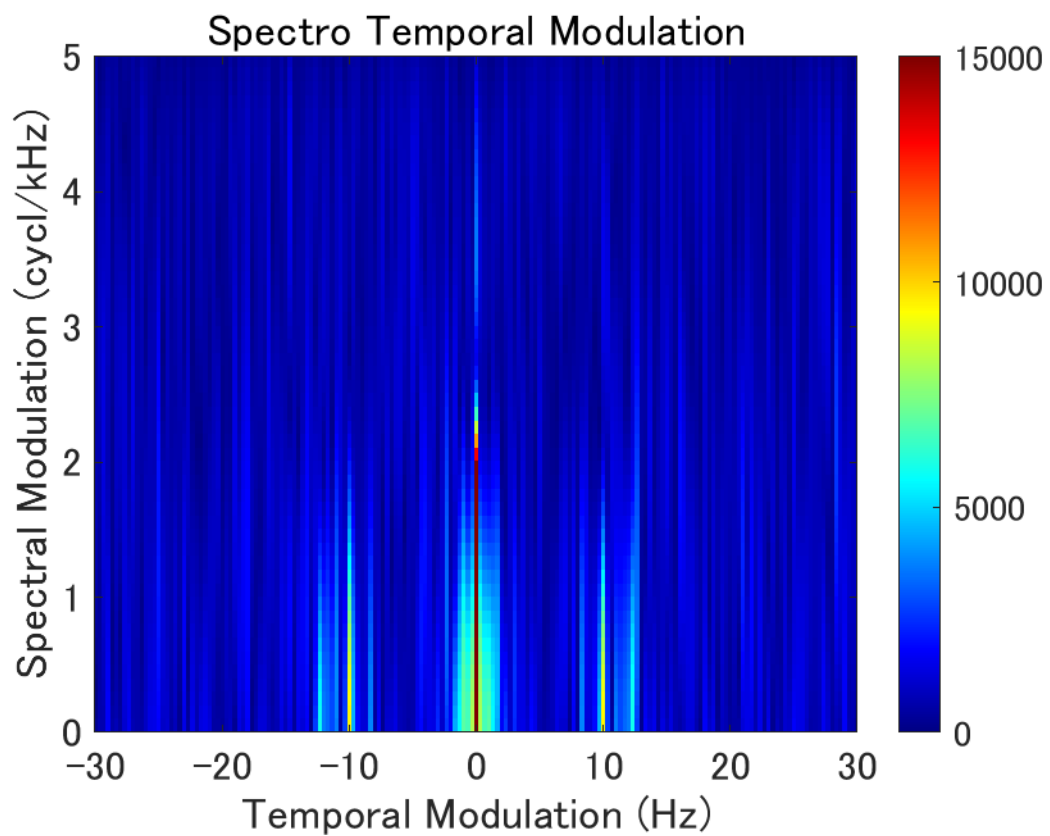


図 3.5: 搬送波の f_c が 525 Hz, BW が 1/2 oct の時の 2 音を混合した目的音と背景音の振幅包絡線が異なるときの STM 情報. 横軸は TM 情報, 縦軸は SM 情報, カラーバーは振幅パワーを示す.

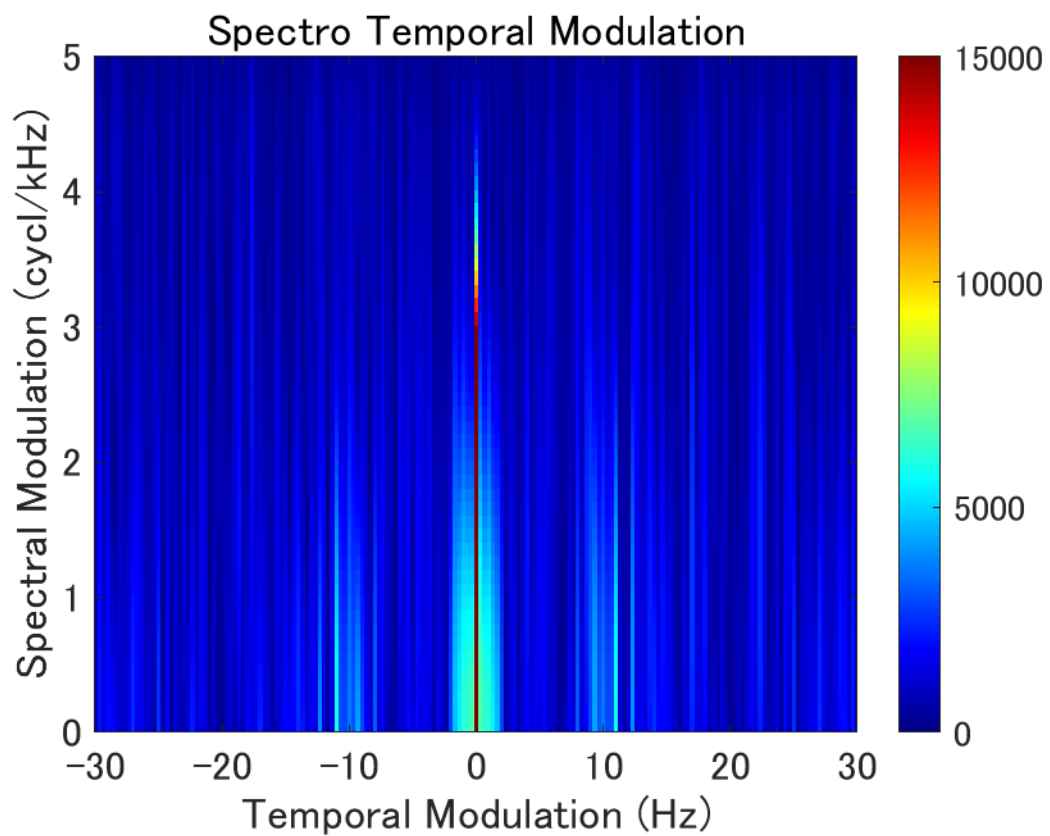


図 3.6: 搬送波の f_c が 525 Hz, BW が 1/2 oct の時の 2 音を混合した目的音と背景音の振幅包絡線が異なるときの STM 情報. 横軸は TM 情報, 縦軸は SM 情報, カラーバーは振幅パワーを示す.

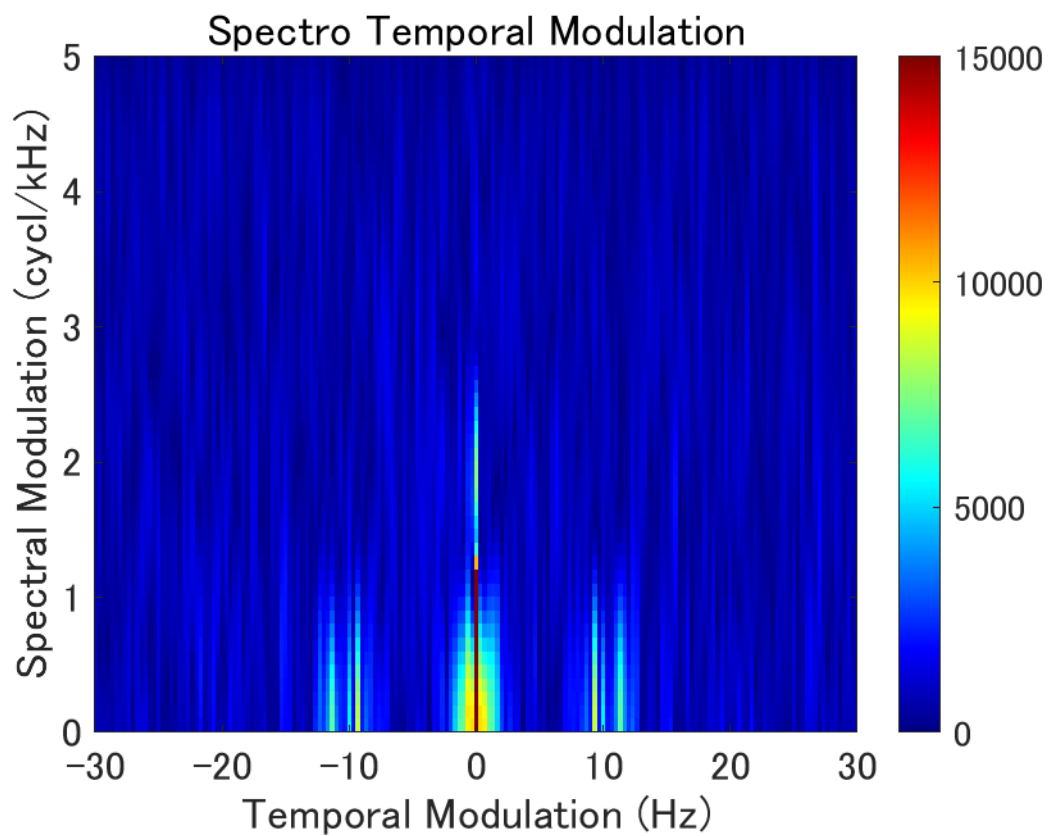


図 3.7: 搬送波の f_c が 2100 Hz, BW が 1/2 oct の時の 2 音を混合した目的音と背景音の振幅包絡線が異なるときの STM 情報. 横軸は TM 情報, 縦軸は SM 情報, カラーバーは振幅パワーを示す.

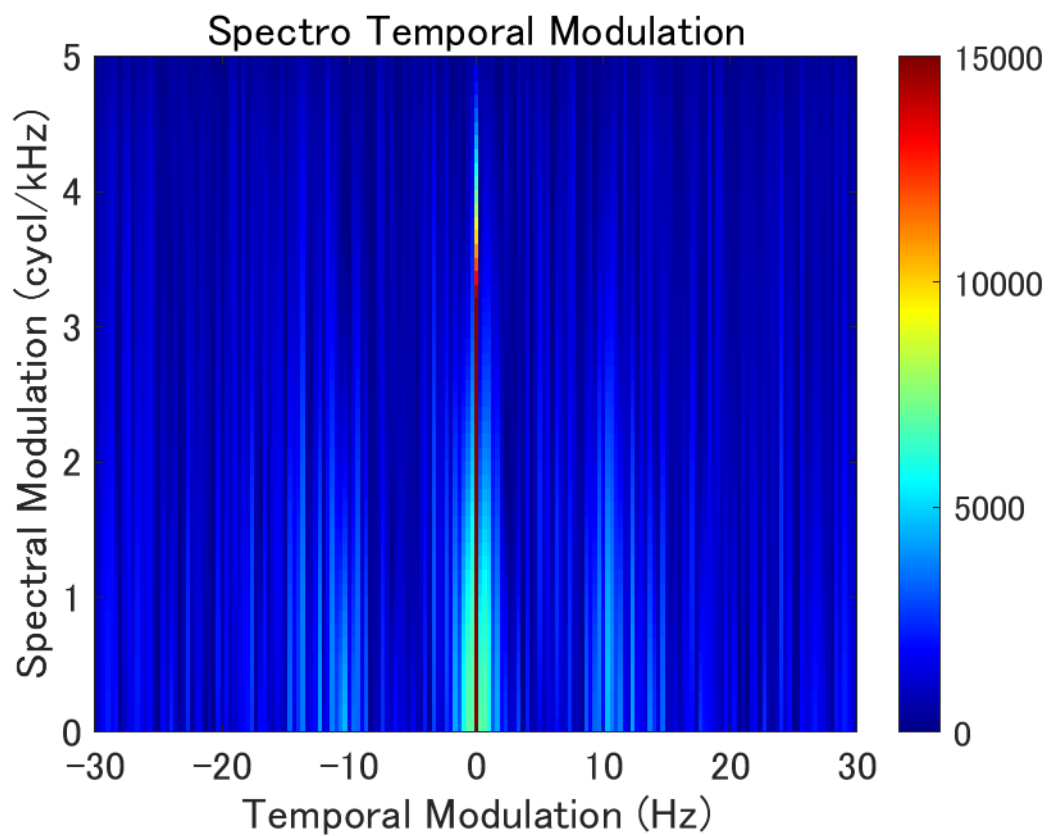


図 3.8: 搬送波の f_c が 525 Hz, BW が 132 Hz の時の 2 音を混合した目的音と背景音の振幅包絡線が異なるときの STM 情報. 横軸は TM 情報, 縦軸は SM 情報, カラーバーは振幅パワーを示す.

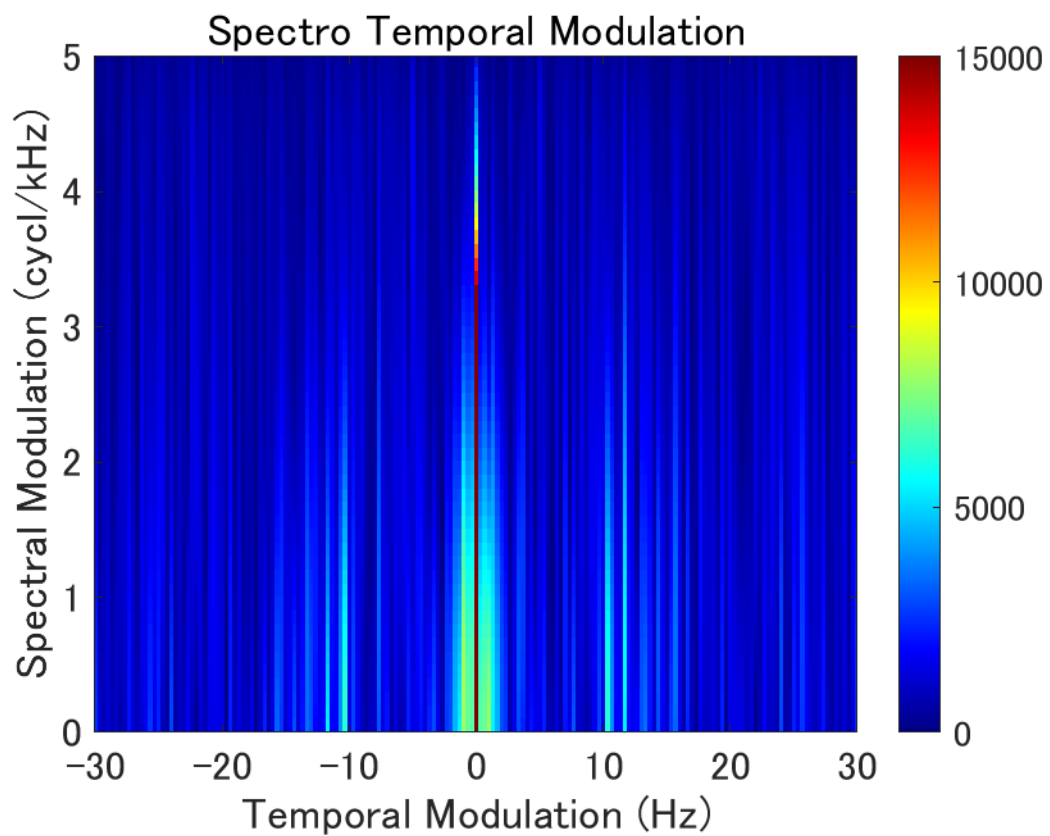


図 3.9: 搬送波の f_c が 1050, BW が 132 Hz の時の 2 音を混合した目的音と背景音の振幅包絡線が異なるときの STM 情報. 横軸は TM 情報, 縦軸は SM 情報, カラーバーは振幅パワーを示す.

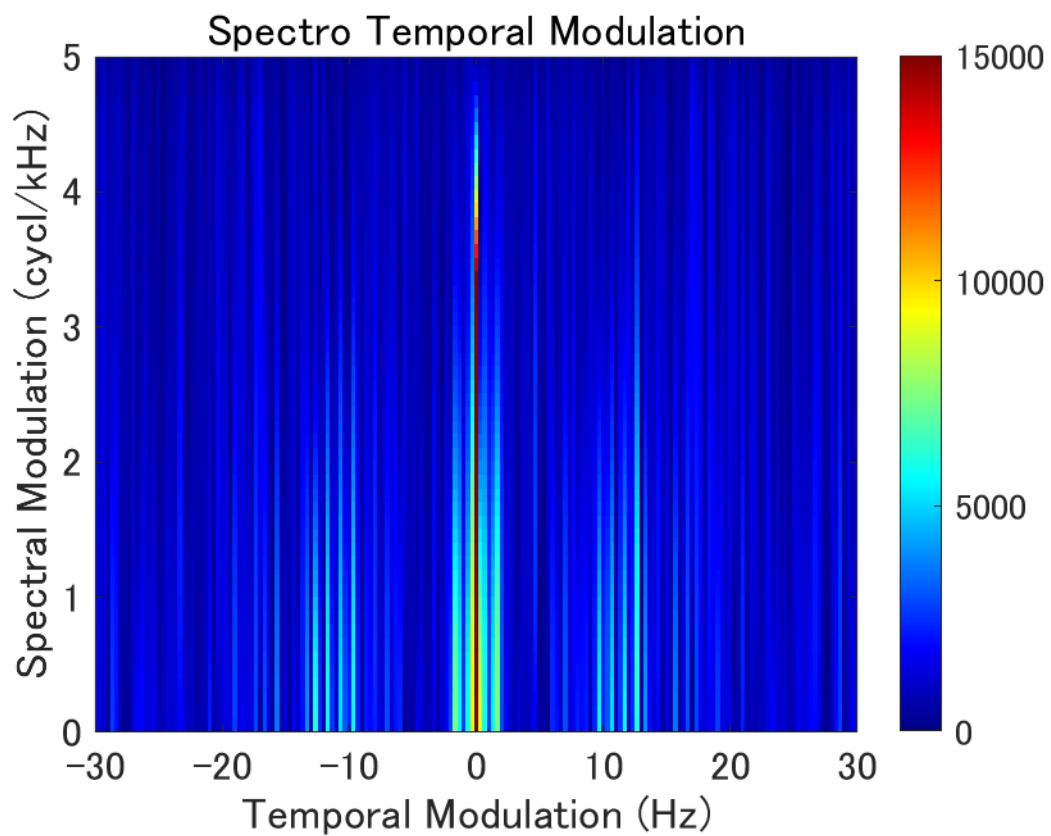


図 3.10: 搬送波の f_c が 2100, BW が 132 Hz の時の 2 音を混合した目的音と背景音の振幅包絡線が異なるときの STM 情報. 横軸は TM 情報, 縦軸は SM 情報, カラーバーは振幅パワーを示す.

3.4 2種類の音が混合された刺激のSTM情報のまとめ

本節では、ここまで目的音と背景音をそれぞれ振幅変調音として考えて分析してきたときの、目的音と背景音のSTM情報についてまとめる。

- 目的音と背景音の中心周波数の差に対応したSM情報にピークが現れる。
- 目的音と背景音のそれぞれの動きに対応したTM情報にピークが現れる。
- 搬送波を狭帯域雑音とした場合、帯域が広くなると、SM情報の広がり狭くなる。
- 搬送波を狭帯域雑音とした場合、搬送波の中心周波数による系統だった関係はSM情報に現れない。

4章でこれらのSTM情報と目的音の検知のされやすさについて聴覚探索実験を行うことにより、どのような関係があるか明らかにする。

第4章 STM情報をもとに作成した刺激を用いた聴覚探索実験

4.1 実験の目的

実験の目的は、STM情報を参考にして作成した目的音と背景音を用いて、検知されやすい目的音と背景音の関係を明らかにすることである。実験間の関連図を図4.1に示す。この図では、3.3節を踏まえて、各実験の刺激の関連をもとに実験の関係を示している。この実験間の関係について、実験の流れとともに説明する。

まず、周波数スペクトルによる検知のされやすさについて明らかにするために、SM情報に着目した実験を行った。次に、先行研究に倣って、時間方向の変動による検知されやすさについて明らかにするために、TM情報に着目した実験を行った。このとき、搬送波に着目した実験を行い、TM情報が目的音の検知されやすさに与える影響を示した。最後に、これらの結果を受けて、目的音と背景音の搬送波が同じ時の検知されやすい関係が、目的音と背景音の搬送波が異なるときでも成り立つかどうかを明らかにするために、目的音と背景音の両方のSM情報とTM情報に着目した実験を行った。

すべての実験に共通する実験条件を、4.2節から4.4節に述べる。各実験の結果について、4.5節から4.7節に述べる。

4.2 実験環境

実験は防音室内で行われた。実験環境を図4.2に示す。実験刺激はコンピュータからオーディオインターフェース (Fireface UCX)、ヘッドホンアンプ (Audio-technica AT-HA 5000) を介し、ヘッドホン (Sennheiser HDA 200) によって呈示された。

4.3 実験参加者

実験参加者は、正常な聴力を持つ大学院生9名 (男性5名、女性4名。平均年齢24.5歳 (標準偏差 ± 2.0)) であった。

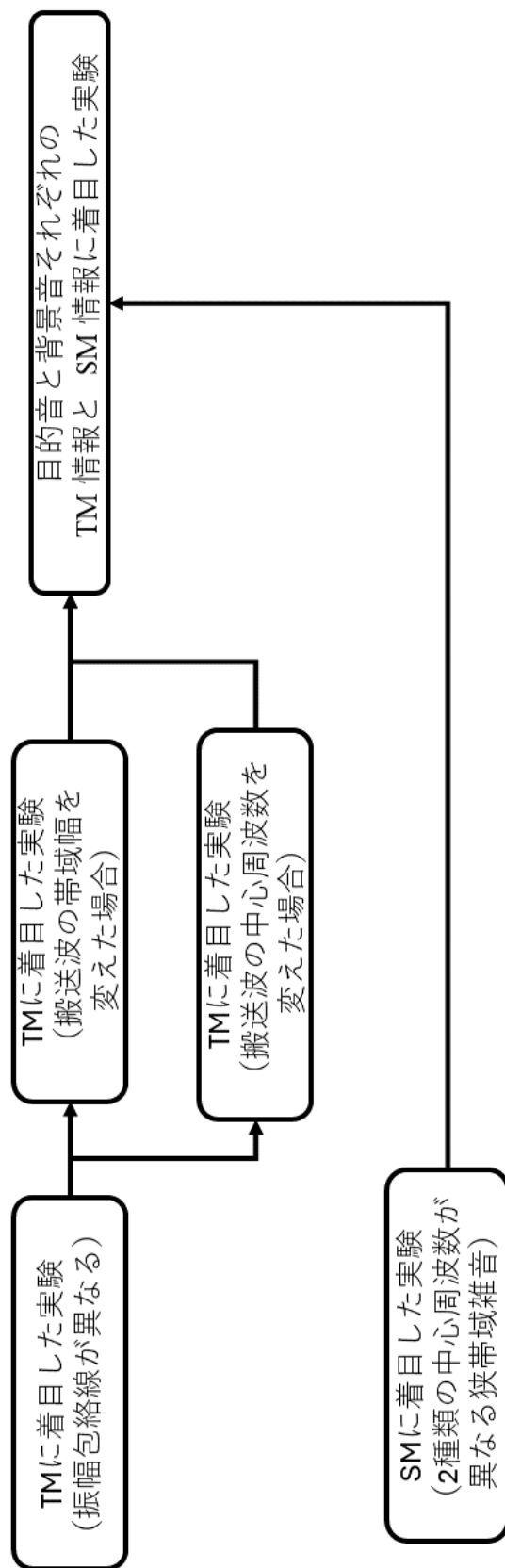


図 4.1: 実験間の関連図

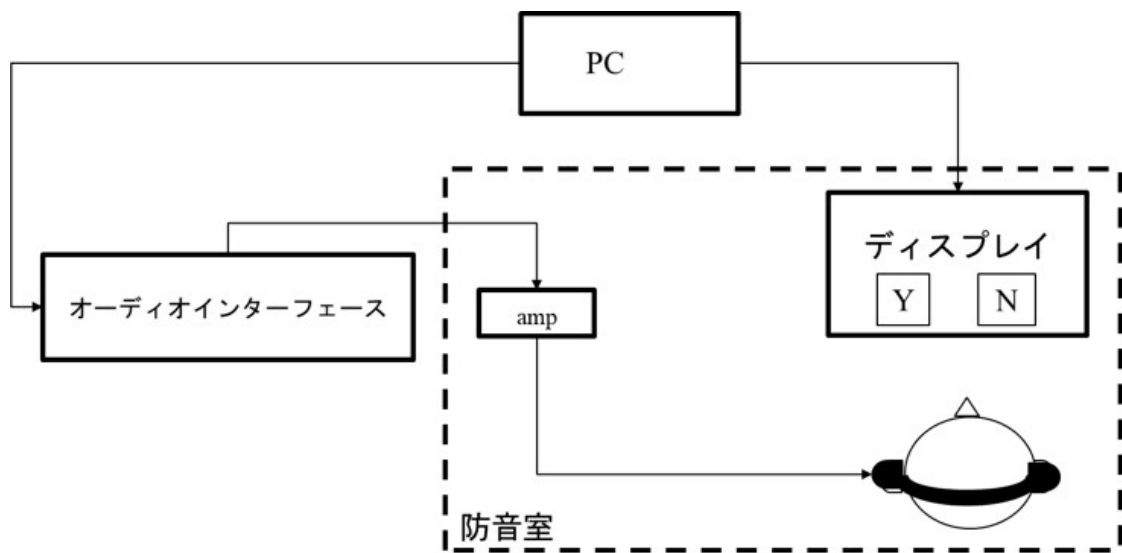


図 4.2: 実験環境

4.4 実験方法

実験参加者に目的音の有無を判断させる Go/No-go 課題をさせた。初めに、実験参加者が目的音を覚えるまで、目的音を繰り返し呈示した。その後、背景音と同時に目的音を含む試行（正試行）と目的音を含まない試行（負試行）をランダムに実施した。負試行は構成音数を正試行と一致させるため、背景音とダミー信号を同時に呈示した。実験途中で目的音を忘れられるのを防ぐために刺激音の1.0秒前に一度だけ目的音を呈示した。目的音が含まれている場合はディスプレイに呈示された Yes ボタンを、含まれていない場合 No ボタンを押すように、実験参加者に教示した。

4.5 評価方法

すべての実験で、実験結果の評価指標として、実験参加者の反応（Hit: 正試行に正解の場合、Miss: 正試行に不正解の場合、Correct rejection: 負試行に正解した場合、False alarm: 負試行に不正解の場合）をもとに検出力 d' を求めた。 d' は信号検出理論 [36] で定義されている。式 (4.1) に定義式を示す。

$$d' = \frac{M_{S+N} - M_N}{\sigma_N} \quad (4.1)$$

ここで、 M_{S+N} は目的音が含まれていると判断する分布の平均、 M_N は目的音が含まれていないと判断する分布の平均、 σ_N はノイズ分布の分散を表す。なおこの時目的音が含まれていると判断する分布と目的音が含まれていないと判断する分布の分散は等しいものとする。実際に実験結果を求めるにあたって文献 [37] を参考にして求めた。

4.6 実験刺激

刺激の作成手順は3.2節と同様について述べた方法と同様である。すべての実験で刺激の呈示レベルはC特性音圧レベル70 dBであった。STM分析を行った刺激と同様に、作成した刺激はすべてサンプリング周波数20000 Hz、刺激長は3.0秒であった。ダウンサンプリングはサンプリング周波数100 Hzで行った。搬送波は狭帯域雑音であった。3章と同様この搬送波の中心周波数を f_c 、帯域幅を BW とした。

4.7 SM情報に着目した実験

実験刺激のパラメータ

目的音は f_c 1050 Hz の狭帯域雑音を，背景音は f_c 300 Hz, 525 Hz, 1600 Hz, 2100 Hz の狭帯域雑音を使用した．目的音と背景音の BW は 1/2 oct であった．負試行では背景音に使用する狭帯域雑音からランダムに 1 音を抽出し，背景音と同時に呈示した．

実験結果

実験結果を図 4.3 に示す．横軸は背景音の f_c ，縦軸は d' を示す．各条件の d' を見るとほとんどの条件で 2.5 を超えている．このことからチャンスレベルを超えた確率で目的音が検知できていることが分かる．背景音の f_c を要因として 1 要因の分散分析を行ったところ，主効果は認められなかった ($F(3, 32) = 0.02, p = 0.996$)．2 つの中心周波数が異なる狭帯域雑音を作る SM 情報に現れる関係は検知がされやすいことが分かった．しかし，BW が原因で検知されやすさに影響を与えているのか， f_c が影響与えているのかはわからなかった．

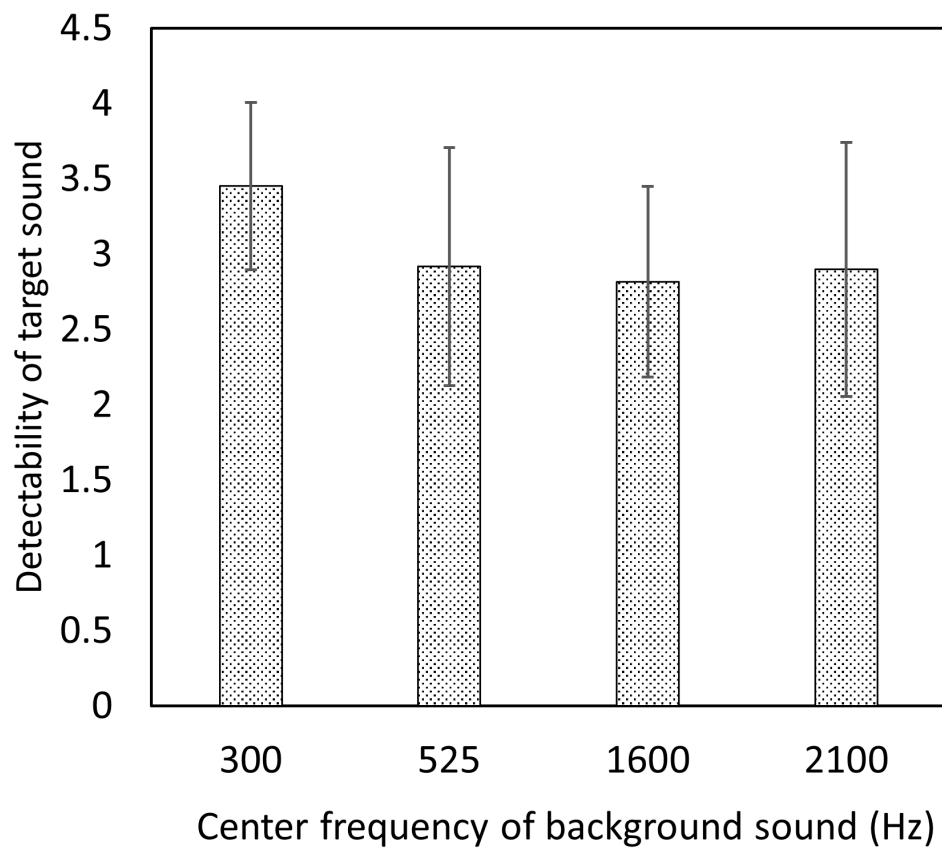


図 4.3: SM 情報に着目したの実験結果. 横軸は背景音の f_c , 縦軸は d' を示す

4.8 TM情報に着目した実験

4.8.1 振幅包絡線に着目した実験

実験刺激のパラメータ

目的音は、振幅包絡線の頂点間の傾きの平均が 6.0, 8.5, 11.0 の 3 条件であった。背景音の振幅包絡線の頂点間の傾きの平均は、目的音の振幅包絡線の頂点間の傾きの平均との差が ± 2.5 , ± 5.0 となるように設定した。また搬送波は狭帯域雑音であり f_c 525 Hz, 1050 Hz, 2100 Hz の 3 条件とし、いずれも BW1/2 oct であった。なお目的音と背景音の搬送波は同一の周波数帯域である。ダミー信号には目的音の搬送波と同じ f_c と周波数帯域幅を持つ狭帯域雑音を使用した。

実験結果

目的音の振幅包絡線の頂点間の傾きの平均が 6.0 の実験結果を図 4.4 に、8.5 の実験結果を図 4.5 に、11.0 の実験結果を図 4.6 に示す。横軸は背景音の頂点間の傾きの平均、縦軸は d' を示す。

各条件の d' の平均値を見るとほとんどの条件で 1.0 を超えている。このことからチャンスレベルである 0.5 を超えた確率で目的音が検知できていることが分かる。ここで目的音の振幅包絡線の頂点間の傾きの平均、振幅包絡線の傾きの差、搬送波の f_c を要因として 3 要因の分散分析を行ったところ、目的音の振幅包絡線の頂点間の傾きの平均 ($F(2, 300) = 10.67, p < 0.01$)、振幅包絡線の傾きの差 ($F(2, 300) = 48.43, p < 0.01$)、搬送波の f_c ($F(3, 32) = 14.36, p < 0.01$) すべての要因で主効果が見られた。また多重比較を行ったところ、目的音と背景音の振幅包絡線の頂点間の傾きが 6.0 で、傾きの差が -5.0 である条件において、搬送波の f_c 525 Hz の場合と搬送波の f_c 2100 Hz の場合に有意差があった ($p < 0.01$)。先行研究 [19] と同様に、目的音と背景音の振幅包絡線が異なる関係が検知されやすいということが示された。これだけにとどまらず、搬送波が検知のされやすさに影響することが分かった。しかし、搬送波の何が影響を与えていたのかわからなかった。

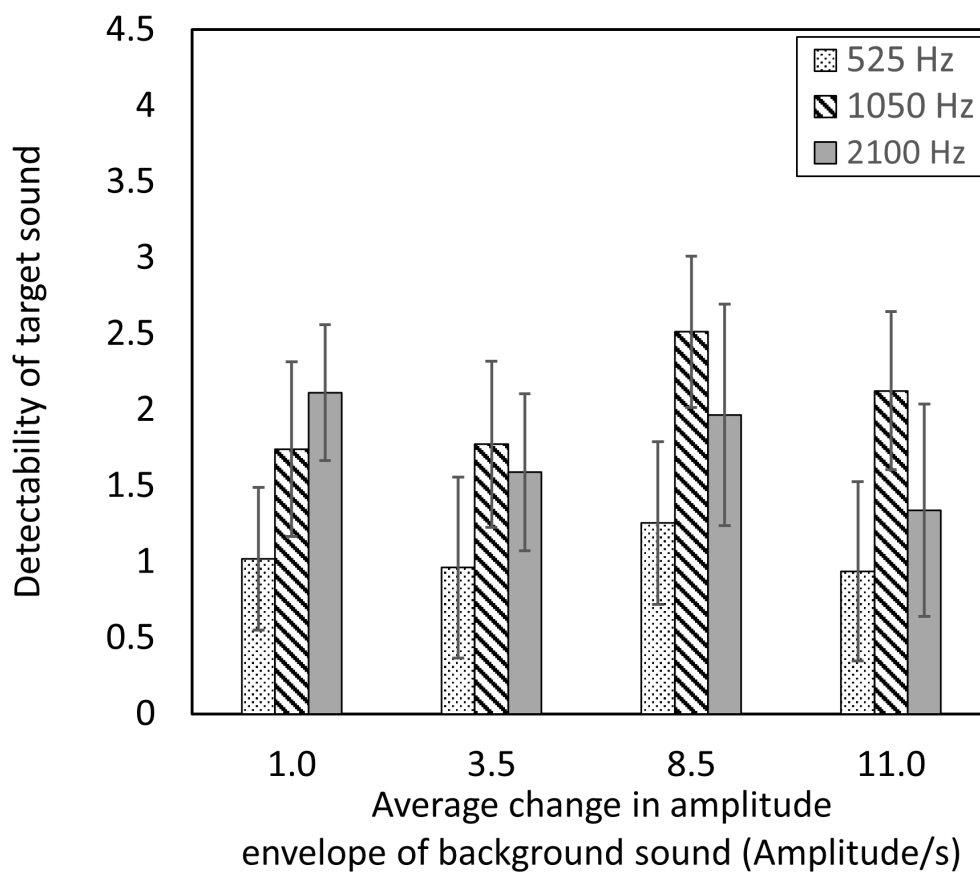


図 4.4: 振幅包絡線に着目した実験. 目的音の振幅包絡の傾きが6.0の場合. 横軸は背景音の振幅包絡線の傾き, 縦軸は d' , 凡例は背景音の f_c を示す.

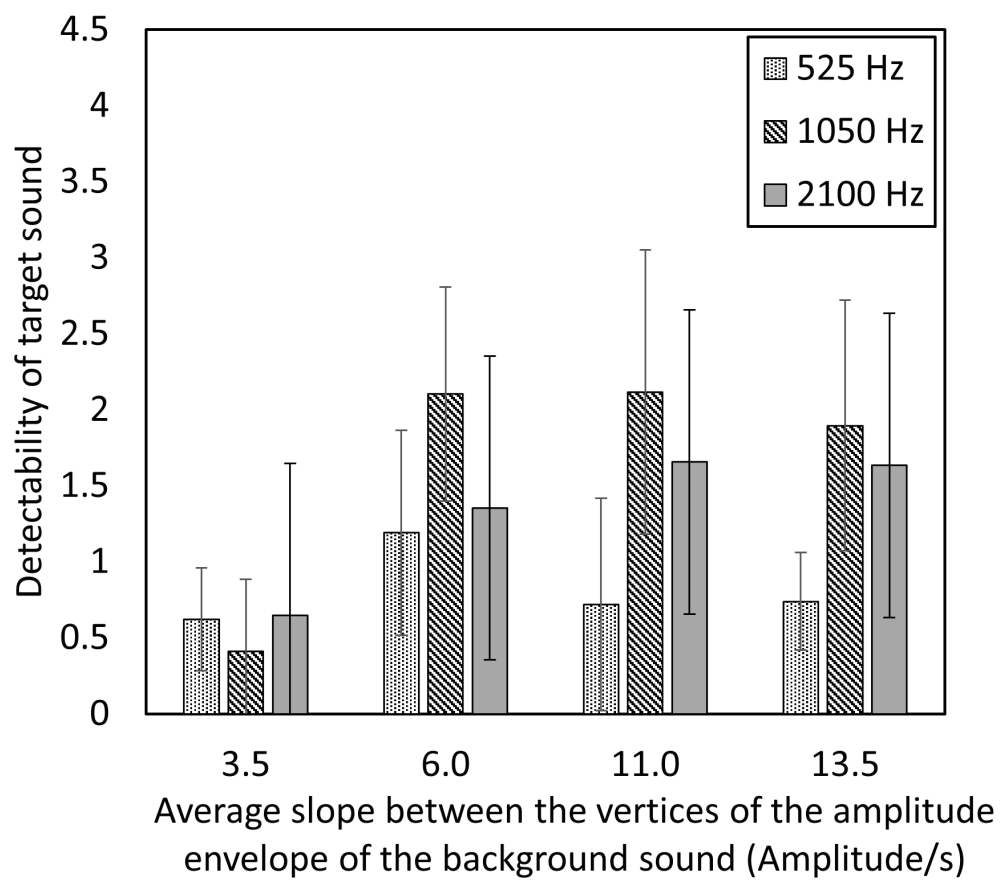


図 4.5: 振幅包絡線に着目した実験. 目的音の振幅包絡の傾きが8.5の場合. 横軸は背景音の振幅包絡線の傾き, 縦軸は d' , 凡例は背景音の f_c を示す.

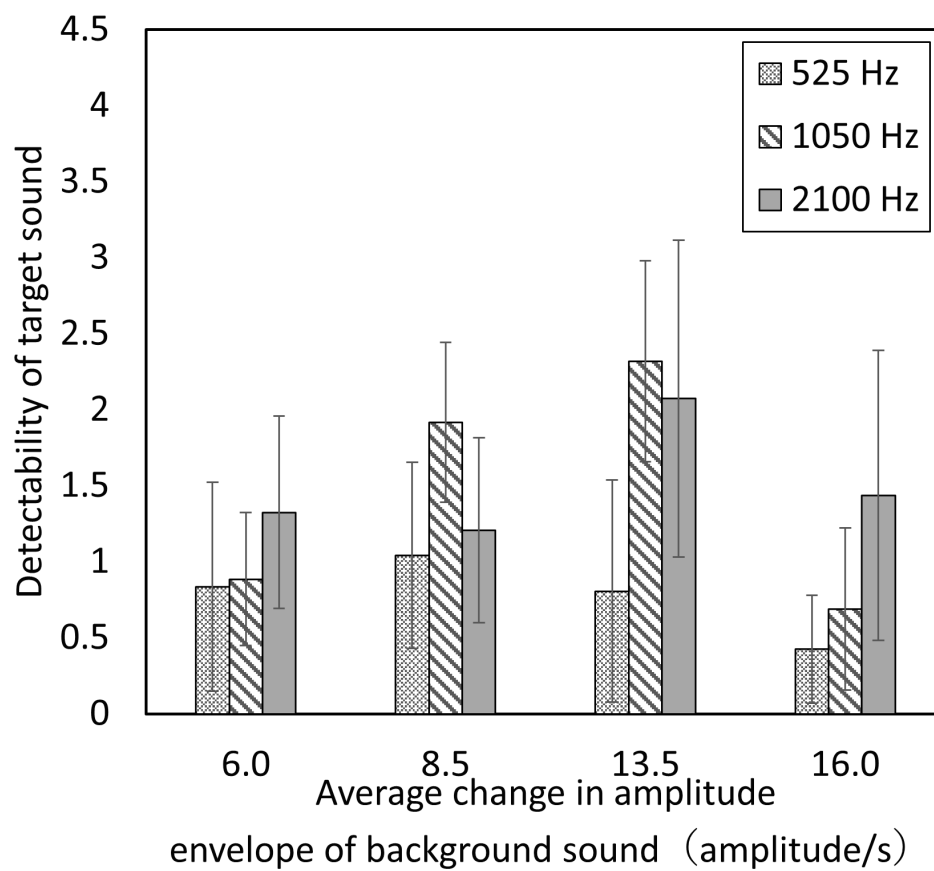


図 4.6: 振幅包絡線に着目した実験. 目的音の振幅包絡の傾き 11.0 の場合. 横軸は背景音の振幅包絡線の傾き, 縦軸は d' , 凡例は背景音の f_c を示す.

4.8.2 振幅包絡線と搬送波の帯域幅に着目した実験

実験刺激のパラメータ

3.6.1 の節の結果と 3.5 節の実験の結果から、搬送波の帯域幅が検知されやすさに影響を与えたかもしれない。そこで搬送波に使用した狭帯域雑音の BW が目的音の検知されやすさに影響を与えているのかどうかを明らかにするために、搬送波の f_c を固定して、BW を変化させた実験を行った。目的音は、振幅包絡線の頂点間の傾きの平均が 6.0, 8.5, 11.0 の 3 条件であった。背景音の振幅包絡線の頂点間の傾きの平均は、目的音の振幅包絡線の頂点間の傾きの平均との差が +2.5, -5.0 となるように設定した。また搬送波は狭帯域雑音であり、 f_c 1050 Hz の 1 条件、BW 1/4 oct, 1/2 oct, 1 oct の 3 条件であった。なお目的音と背景音の搬送波は同一の周波数帯域である。ダミー信号には目的音の搬送波と同じ f_c と周波数帯域幅を持つ狭帯域雑音を使用した。

実験結果

目的音の振幅包絡線の頂点間の傾きの平均が 6.0 の実験結果を図 4.7 に、8.5 の実験結果を図 4.8 に、11.0 の実験結果を図 4.9 に示す。横軸は背景音の頂点間の傾きの平均、縦軸は d' を示す。

各条件の d' を見るとほとんどの条件で 1.0 を超えている。このことからチャンスレベルである 0.5 を超えた確率で目的音が検知できていることが分かる。目的音の振幅包絡線の頂点間の傾きの平均、振幅包絡線の傾きの平均の差、搬送波の BW を要因として 3 要因の分散分析を行ったところ、目的音の振幅包絡線の頂点間の傾きで有意傾向があり ($F(2, 148) = 3.06, p = 0.050$)、振幅包絡線の傾きの平均の差 ($F(1, 148) = 15.26, p < 0.01$)、搬送波の BW ($F(3, 32) = 24.05, p < 0.01$) となり主効果が見られた。また多重比較を行ったところ、目的音の振幅包絡線の頂点間の傾きの平均が 6.0、傾きの差が -5.0、BW 1 oct の条件と、目的音の振幅包絡線の頂点間の傾きの平均が 6.0、傾きの差が -5.0、BW 1/4 oct の条件で有意差があった ($p < 0.01$)。搬送波の BW が検知のされやすさに影響することが分かった。

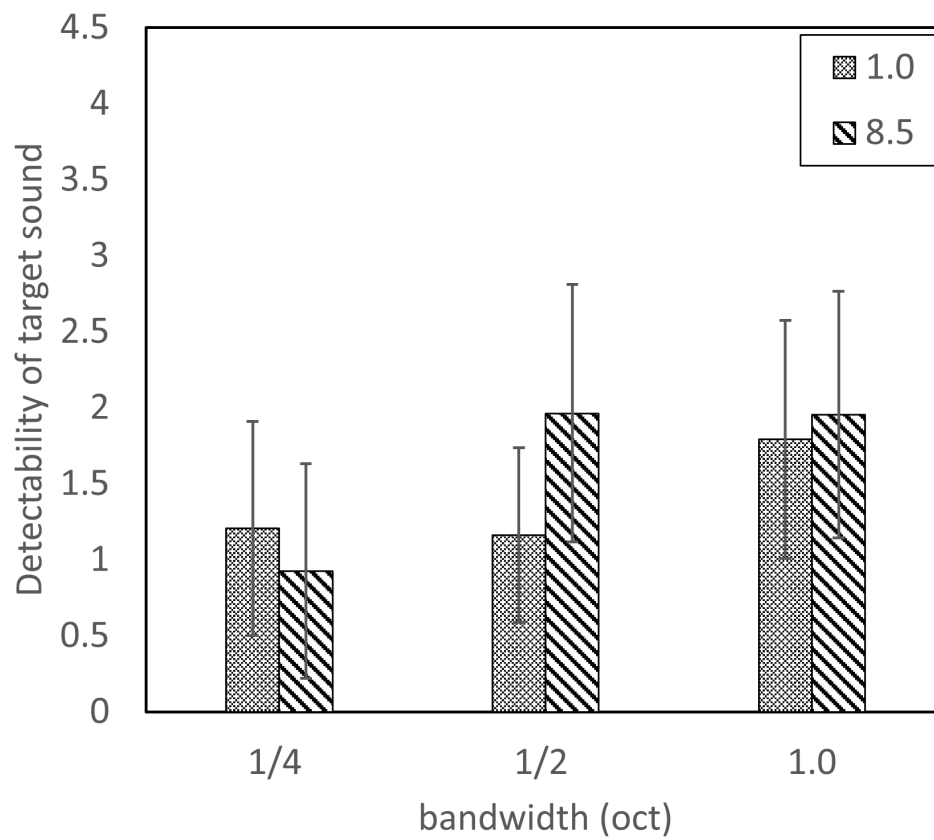


図 4.7: 振幅包絡線と搬送波の帯域幅に着目した実験の結果. 目的音の振幅包絡の傾きが 6.0 の場合. 横軸は背景音の BW, 縦軸は d' , 凡例は背景音の傾きを示す.

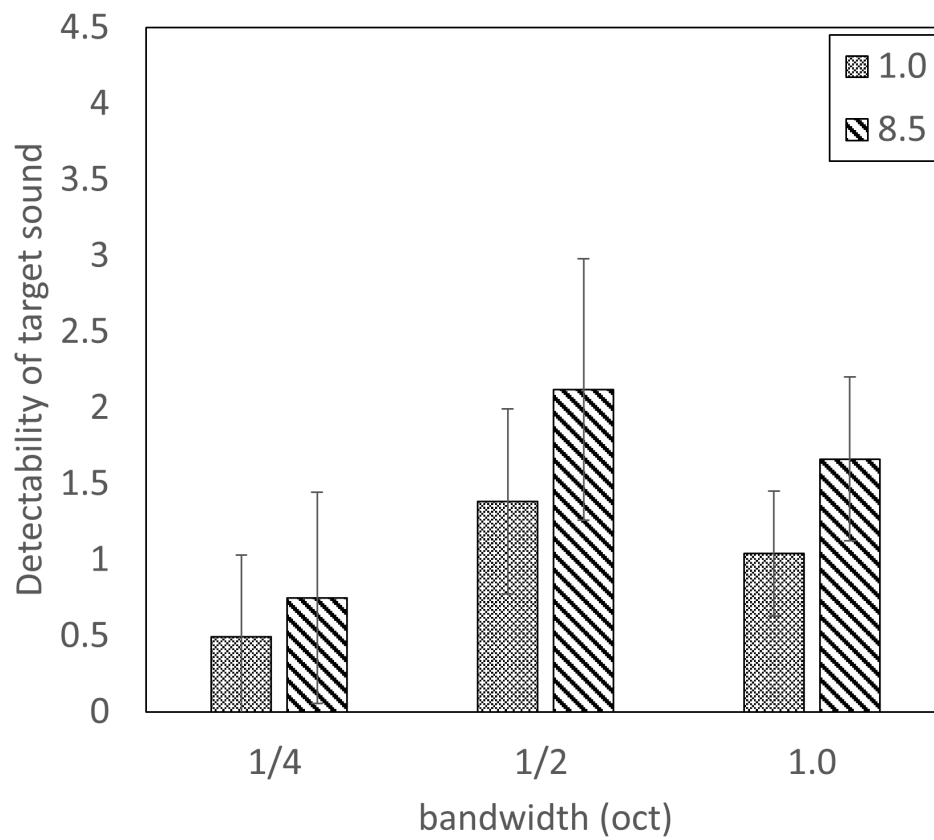


図 4.8: 振幅包絡線と搬送波の帯域幅に着目した実験の結果. 目的音の振幅包絡線の傾きが8.5の場合. 横軸は背景音のBW, 縦軸は d' , 凡例は背景音の傾きを示す.

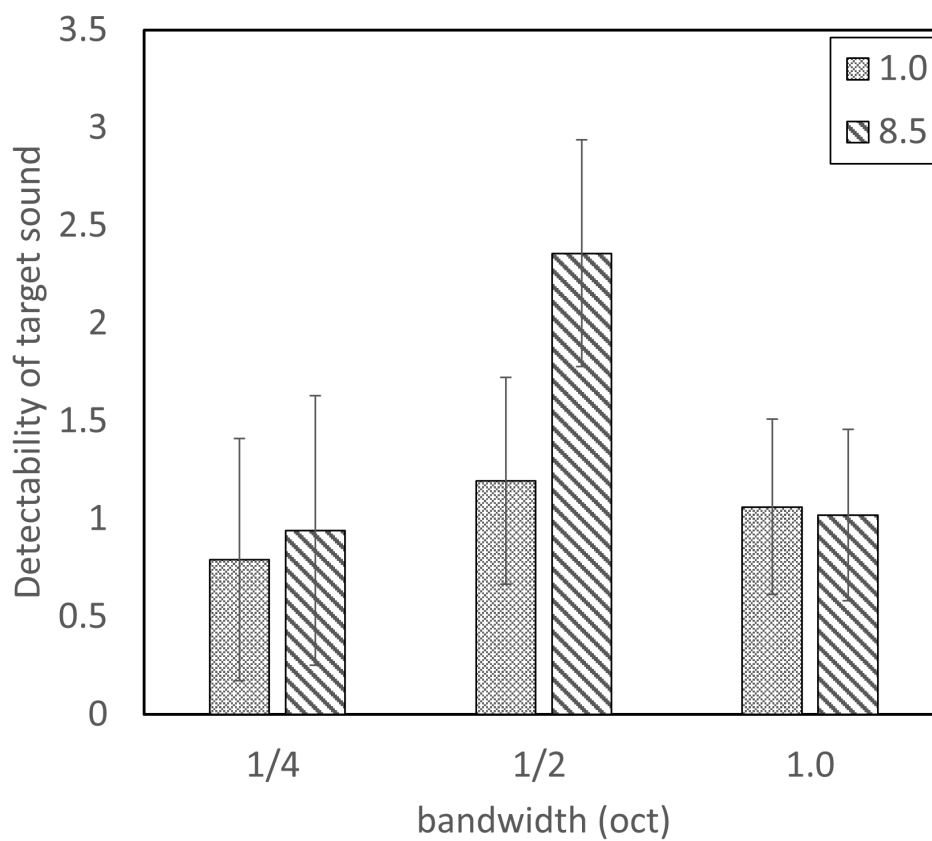


図 4.9: 振幅包絡線と搬送波の帯域幅に着目した実験の結果. 目的音の振幅包絡の傾きが 11.0 の場合. 横軸は背景音の BW, 縦軸は d' , 凡例は背景音の傾きを示す.

4.8.3 振幅包絡線と搬送波の中心周波数に着目した実験

実験刺激のパラメータ

搬送波の f_c が目的音の検知されやすさに影響を与えているのかどうかを明らかにするために、搬送波の BW を固定して、 f_c を変化させた実験を行った。目的音には、振幅包絡線の頂点間の傾きの平均が 6.0 のものを使用した。背景音の振幅包絡線の頂点間の傾きの平均は、目的音の振幅包絡線の頂点間の傾きの平均との差が +2.5, -5.0 の 2 条件であった。また、搬送波は f_c 525 Hz, 1050 Hz, 2100 Hz の 3 条件で、BW が 132 Hz の 1 条件とした。なお、目的音と背景音の搬送波は同一の周波数帯域である。ダミー信号には目的音と同じ f_c と BW を持つ狭帯域雑音を使用した。

実験結果

実験結果を図 4.10 に示す。横軸は背景音の頂点間の傾きの平均、縦軸は d' を示す。振幅包絡線の傾きの平均の差、搬送波の f_c を要因として 2 要因の分散分析を行ったところ、振幅包絡線の傾きの平均の差 ($F(1, 42) = 0, p = 0.983$)、搬送波の f_c ($F(2, 42) = 2.41, p = 0.102$) となり主効果は認められなかった。搬送波の狭帯域雑音の中心周波数が変わっても目的音の検知のされやすさに差がなかった。

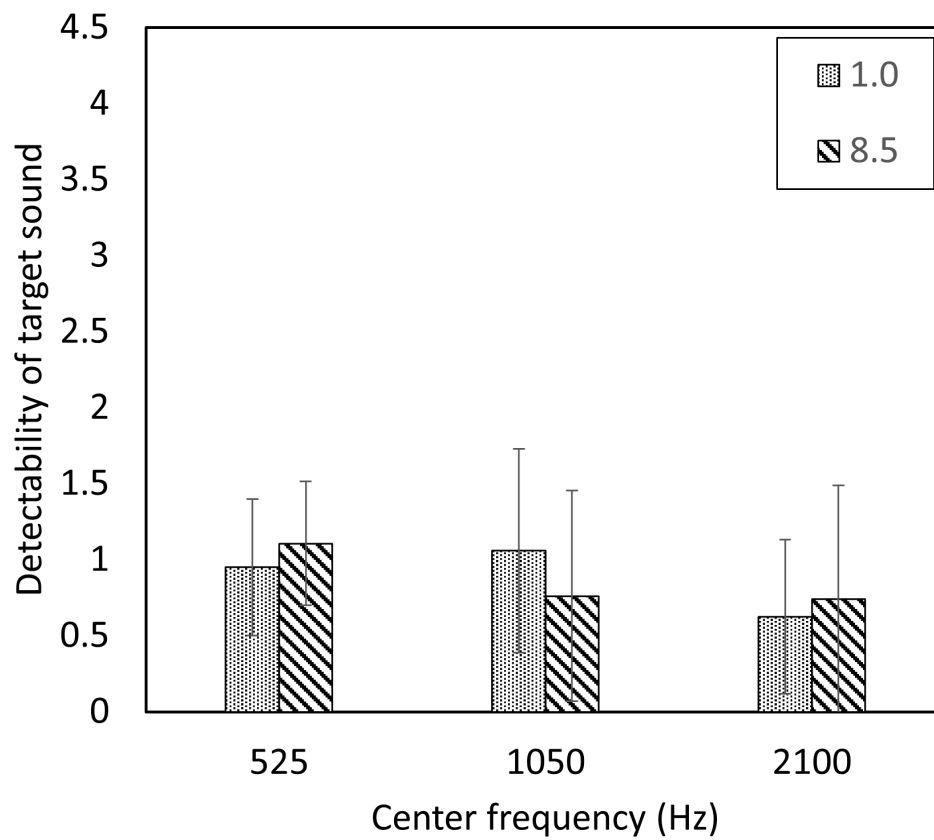


図 4.10: 振幅包絡線と搬送波の中心周波数に着目した実験の結果. 横軸は背景音の BW, 縦軸は d' , 凡例は背景音の傾きを示す.

4.9 目的音と背景音それぞれの TM 情報と SM 情報に着目した実験

4.9.1 実験刺激のパラメータ

目的音には振幅包絡線の頂点間の傾きの平均が 6.0, 搬送波が f_c 1050 Hz, BW 1/2 oct の狭帯域雑音を使用した。背景音の振幅包絡線は, 目的音の振幅包絡線の頂点間の傾きの平均との差が +2.5, 0, -5.0 の 3 条件であった。また, 背景音の搬送波は f_c 1050 Hz の 1 条件, BW 1/4 oct と 1 oct の 2 条件であった。ダミー信号には目的音と同じ f_c と BW を持つ狭帯域雑音を使用した。

4.9.2 実験結果

実験結果を図 4.11 に示す。横軸は背景音の振幅包絡線の頂点間の傾き, 縦軸は d' を示す。背景音の振幅包絡線の頂点間の傾きと背景音の BW を要因として 2 要因の分散分析を行ったところ, 背景音の BW で主効果が認められた ($F(1, 48) = 5.48, p = 0.0234$)。多重比較を行ったところ, 背景音の振幅包絡線が 6.0 で BW が 1/4 oct の条件と, 背景音の振幅包絡線が 6.0 で BW が 1 oct の条件間で有意傾向が見られた ($p = 0.054$)。目的音と背景音の搬送波の帯域幅が異なるときに, 振幅包絡線の動きが同じであったとしても, 帯域幅の条件が目的音の検知されやすさに影響を与えることがわかった。

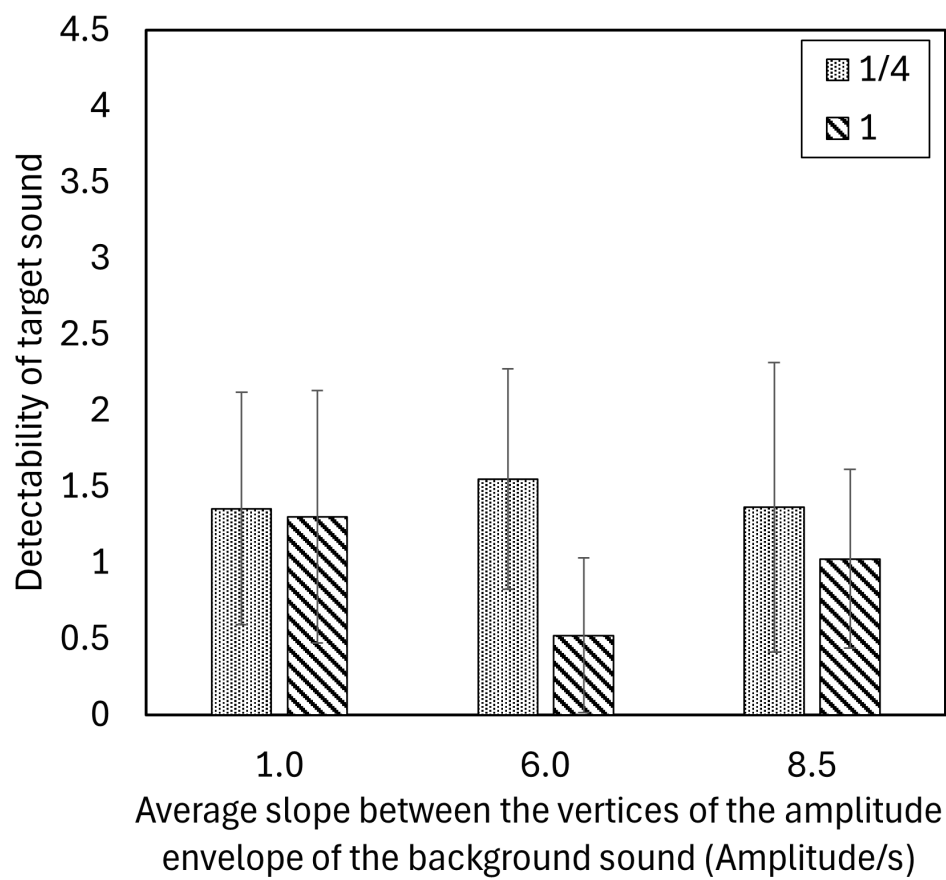


図 4.11: 目的音と背景音の振幅包絡線も搬送波の帯域幅も異なる実験の結果. 横軸は背景音の振幅包絡線の傾き, 縦軸は d' , 凡例は背景音の BW を示す.

第5章 総合考察

SM 情報に着目した実験の結果から、目的音と背景音が異なる聴覚フィルタを通ったことにより周波数選択性 [38] によって、目的音の手がかりがわかりやすかったため、検知されやすかったと考えられる。

TM 情報に着目した実験では、SM 情報に着目した実験とは異なり、目的音と背景音は、同一の聴覚フィルタを通る刺激であった。この場合でも、目的音が検知されることが示された。これは、音の特徴抽出が、聴覚末梢系から脳内に至るまで平行して行われること [39] を考慮すると、振幅包絡線のみが異なる聴覚探索問題では、聴覚末梢系の処理だけでない要因が働いたためと考えられる。

振幅包絡線と搬送波の帯域幅に着目した実験と、振幅包絡線と搬送波の中心周波数に着目した実験の結果から、搬送波が狭帯域雑音の場合、中心周波数ではなく、狭帯域雑音の帯域幅が検知のされやすさに影響を与えていた。また目的音と背景音のそれぞれの TM 情報と SM 情報に着目した実験では、振幅包絡線の動きが同じであっても、検知されやすい条件があることが分かった。聴覚情景分析で Bregman[40] が述べた同時的群化の考え方は振幅包絡線が似たものは群化されるということであったが、今回の結果は単に振幅包絡線が似たものが群化するのではないということを示唆した。このことから、SM 情報が目的音の検知されやすさの手がかりになっていたことを示した。

生理学的な関連を考えると、STM 情報は 3.2 節で述べたように高次の変調知覚に係る脳の領域の STRF と対応があることが分かっている [33][34]。実験刺激は STRF に関わる実験刺激であった。これらことから、今回行った聴覚探索で得られた傾向が振幅変調音以外でも見られる可能性が示唆される。

第6章 結論

6.1 本研究で明らかにしたこと

本研究の目的は、STM 情報を用いて、聴覚探索において、目的音が検知されやすいときの、目的音と背景音の関係を明らかにすることであった。そのために、STM 情報を参考に刺激を作成し、聴覚探索実験を行った。本研究で明らかになったことを以下に示す。

- 2つの中心周波数が異なる狭帯域雑音を作る SM 情報に現れる関係は検知がされやすいことが分かった。
- 先行研究 [19] と同様に、目的音と背景音の振幅包絡線が異なる関係が検知されやすいということが分かった。これだけにとどまらず、搬送波の帯域幅が検知のされやすさに影響する分かった。
- 目的音と背景音の搬送波の帯域幅が異なるときに、振幅包絡線の動きが同じであったとしても、帯域幅の条件が、目的音の検知されやすさに影響を与えることが分かった。

6.2 残された課題

本研究の課題を以下のように列挙する。

- 刺激の条件に関する課題である。現在は、振幅変調音に対して検討を行った。音は調波性、非調波性、周波数変調などといった条件がある。これらの条件を考慮した上で、本研究で、明らかになった目的音と背景音の関係が成り立つのであれば、目的音と背景音が一般的な信号であっても、聴覚探索について STM 情報を用いた議論ができると考えられる。
- STM 情報と目的音の検知されやすさの直接的な関係の結び付けに関する課題である。本研究では、STM 情報を参考にして刺激を作ることにより、目的音が検知されやすいときの、目的音と背景音の関係について議論をした。この関係について、STM 情報と目的音の検知のされやすさが直接結びつくような数理モデルや、アルゴリズムを示せたら、聴取実験を行わずに、様々な条件の目的音と背景音の検知されやすさの議論につながると考えられる。

- 報知音の規格に沿った音同士の聴覚探索が課題である。序論で述べたように報知音設計は、すべてのヒトに伝わるように、工夫されている。これらの指針を考慮した刺激を用いて、聴覚探索について検討することができれば、聴覚探索に沿って、報知音の知覚の議論ができると考えられる。

参考文献

- [1] 赤木正人, “カクテルパーティ効果とそのモデル化,” 電子情報通信学会誌, vol. 78, no. 5, pp. 450–453, 1995.
- [2] 日本音響学会 編, “音響用語辞典,” コロナ社, 1988.4.
- [3] 岩宮眞一郎, “サイン音の科学,” コロナ社, 2012.
- [4] 一般財団法人 家電製品協会 ユニバーサルデザイン技術委員会, “家電製品における操作性向上のための報知音に関するガイドライン,” 2018.
- [5] 日本産業規格 JIS, “アクセシブルデザイン–消費生活用製品の報知音,” 2022.
- [6] 倉片憲治, 松下一馬, 久場康良, 口ノ町康夫, “家電製品の報知音の計測・第3報,” 人間工学, vol. 36, no. 3, pp. 147–153, 2000.
- [7] 水浪田鶴, 倉片憲治, 下迫晴加, 松下一馬, “JIS S 0013 における報知音 (終了音・注意音) の推奨 on/off パターンの追検討,” 人間工学, vol. 40, no. 5, pp. 264–271, 2004.
- [8] K. Saberi, L. Dostal, T. Sadralodabai, V. Bull, and D. R. Perrott, “Free-field release from masking,” *The Journal of the Acoustical Society of America*, vol. 90, no. 3, pp. 1355–1370, 1991.
- [9] 内山英昭, “自動車走行雑御中の報知音知覚に関する研究,” 北陸先端科学技術大学院大学修士論文, 2007.
- [10] 黒田直樹, “方向性の手掛かりが雑音環境下での報知音の検知能力に及ぼす影響,” 北陸先端科学技術大学院大学修士論文, 2009.
- [11] 塩入諭, “注意と視覚探索,” 映像情報メディア学会誌, vol. 58, no. 11, pp. 1619–1625, 2004.
- [12] A. M. Treisman and G. Gelade, “A feature-integration theory of attention,” *Cognitive Psychology*, vol. 12, no. 1, pp. 97–136, 1980.
- [13] H. H. Greene and K. Rayner, “Eye movements and familiarity effects in visual search,” *Vision Research*, vol. 41, no. 27, pp. 3763–3773, 2001.

- [14] 森周司, “聴覚的注意研究の枠組みと動向,” 日本音響学会誌, vol. 75, no. 2, pp. 90–96, 2019.
- [15] Y. Sugita, Y. Suzuki, and T. Sone, “Auditory search asymmetry,” in *Proceedings of the 20th Annual International Conference of the IEEE Engineering in Medicine and Biology Society. Vol. 20 Biomedical Engineering Towards the Year 2000 and Beyond (Cat. No. 98CH36286)*, vol. 6. IEEE, 1998, pp. 3126–3128.
- [16] 阿瀬見典昭, “聴覚の探索非対称性,” 東北大学博士論文, 2003.
- [17] 草場美沙, “聴覚探索における雑音と報知音の時間変動の類似性が報知音知覚に及ぼす影響,” 北陸先端科学技術大学院大学修士論文, 2008.
- [18] 横澤一彦, 熊田孝恒, “視覚探索—現象とプロセス,” 認知科学, vol. 3, no. 4, pp. 119–138, 1996.
- [19] Y. Yano, R. Miyauchi, M. Unoki, and A. Masato, “Study on detectability of signals by utilizing differences in their amplitude modulation,” *Journal of Signal Processing*, vol. 16, pp. 623–627, 2012.
- [20] 矢野雄大, “背景雑音に対する報知音の振幅包絡の動きの違いが報知音検知に与える影響,” 北陸先端科学技術大学院大学修士論文, 2012.
- [21] H. Fastl, “Fluctuation strength and temporal masking patterns of amplitude-modulated broadband noise,” *Hearing Research*, vol. 8, no. 1, pp. 59–69, 1982.
- [22] H. Fastl and E. Zwicker, “*Psychoacoustics:facts and models*”.
- [23] R. Eramudugolla, K. I. McAnally, R. L. Martin, D. R. Irvine, and J. B. Mattingley, “The role of spatial location in auditory search,” *Hearing Research*, vol. 238, no. 1, pp. 139–146, 2008, the Auditory Brain - A Tribute to Dexter R.F. Irvine.
- [24] N. Singh and F. E. Theunissen, “Modulation spectra of natural sounds and ethological theories of auditory processing,” *The Journal of the Acoustical Society of America*, vol. 114, pp. 3394–3411, Dec. 2003.
- [25] T. M. Elliott and F. E. Theunissen, “The modulation transfer function for speech intelligibility,” *PLoS computational biology*, vol. 5, no. 3, p. e1000302, 2009.
- [26] L. H. Arnal, A. Flinker, A. Kleinschmidt, A. L. Giraud, and D. Poeppel, “Human screams occupy a privileged niche in the communication soundscape,” *Current Biology*, vol. 25, pp. 2051–2056, Aug. 2015.

- [27] S. Rosen, “Temporal information in speech: acoustic, auditory and linguistic aspects,” *Philosophical Transactions of the Royal Society of London. Series B: Biological Sciences*, vol. 336, no. 1278, pp. 367–373, 1992.
- [28] T. Chi, Y. Gao, M. C. Guyton, P. Ru, and S. Shamma, “Spectro-temporal modulation transfer functions and speech intelligibility,” 1999.
- [29] M. Elhilali, T. Chi, and S. A. Shamma, “A spectro-temporal modulation index (stmi) for assessment of speech intelligibility,” *Speech Communication*, vol. 41, pp. 331–348, Oct. 2003.
- [30] A. Dutta, G. Ashishkumar, and C. V. Rao, “Improving the performance of asr system by building acoustic models using spectro-temporal and phase-based features,” *Circuits, Systems, and Signal Processing*, vol. 41, pp. 1609–1632, Mar. 2022.
- [31] 木所晃利, 木谷俊介, 鷗木祐史, “Spectro-temporal modulation 分析を利用した聴覚的顕著性の検討,” 聴覚研究会資料, vol. 50, no. 6, pp. 383–388, Nov. 2020.
- [32] 田中聡一郎, 堀口遼太郎, 木谷俊介, 鷗木祐史, “聴覚的顕著性の予測のための spectro-temporal modulation 分析の検討,” 電子情報通信学会技術研究報告, vol. 122, 2022.
- [33] P. Heil, R. Rajan, and D. R. Irvine, “Sensitivity of neurons in cat primary auditory cortex to tones and frequency-modulated stimuli. i: Effects of variation of stimulus parameters,” *Hearing Research*, vol. 63, no. 1, pp. 108–134, 1992.
- [34] D. A. Depireux, J. Z. Simon, D. J. Klein, and S. A. Shamma, “Spectro-temporal response field characterization with dynamic ripples in ferret primary auditory cortex,” *Journal of neurophysiology*, vol. 85, no. 3, pp. 1220–1234, 2001.
- [35] 木所晃利, “聴覚的顕著性とスペクトル・時間変調情報の関係,” 北陸先端科学技術大学院大学修士論文, 2020.
- [36] W. W. Peterson, T. G. Birdsall, and W. C. Fox, “The theory of signal detectability,” *Trans. IRE Prof. Group Inf. Theory*, vol. 4, pp. 171–212, 1954.
- [37] 井上和哉, “信号検出理論の概要と教授法,” 認知心理学会テクニカルレポート, 2019.

- [38] 磯島愿三, “聴覚系の周波数選択性,” *AUDIOLOGY JAPAN*, vol. 34, no. 6, pp. 824–841, 1991.
- [39] 柏野牧夫, “聴覚情景分析の脳内メカニズム,” *計測と制御*, vol. 46, no. 6, pp. 472–478, 2007.
- [40] A. S. Bregman, “Auditory scene analysis,” vol. 198, 1994.

謝辞

主指導教員として、本研究を行う上で多くのことを指導してくださった木谷俊介講師に感謝します。木谷先生の指導無くして本論文の完成はありませんでした。

本学の教授である、鵜木先生には本研究を行う機会や、聴覚分野の最先端の研究を知る機会をいただいたことについて、感謝いたします。

本学の特任助教である上江州先生には公私にわたって研究の面白さや奥深さを教えていただいたこと感謝いたします。

磯山さんをはじめとする博士後期課程の先輩方や、新谷さん、姫子松さんをはじめとする博士前期課程の皆様との日頃の議論は、研究活動にとっても良い影響を与えてくれました。心より感謝いたします。

本学の保健管理センター所属の佐々木先生は、研究室メンバーは相談しづらい研究上の悩み事について、親身に聞いてくださったこと感謝いたします。

本学を昨年度修了された宮家さんは、社会人一年目として忙しい身でもあるにも関わらず公私にわたって様々な助言を頂いたこと感謝いたします。今日まで研究への情熱を保つことができたのは、宮家さんの助言のおかげです。

本研究における実験は、長期間であった。このような長期間の実験に、参加してくださった、鵜木研究室の M1 の学生をはじめとする実験参加者の皆様に深く感謝します。

学内外問わず、関わってくれた友人の皆様に感謝します。皆様が叱咤激励をしてくださったおかげで挫折しそうな時でも頑張ることができました。

最後に学部生活から長く続いた学生生活を経済面と精神面で支えてくれた家族に感謝します。