

Title	方向性マスク解除を説明する両耳聴モデルに関する研究
Author(s)	水川, 慎也
Citation	
Issue Date	2011-03
Type	Thesis or Dissertation
Text version	author
URL	http://hdl.handle.net/10119/9610
Rights	
Description	Supervisor:赤木正人, 情報科学研究科, 修士

修 士 論 文

方向性マスキング解除を説明する
両耳聴モデルに関する研究

北陸先端科学技術大学院大学
情報科学研究科情報科学専攻

水川 慎也

2011 年 3 月

修 士 論 文

方向性マスキング解除を説明する 両耳聴モデルに関する研究

指導教官 赤木正人 教授

審査委員主査 赤木正人 教授
審査委員 党建武 教授
審査委員 徳田功 准教授

北陸先端科学技術大学院大学
情報科学研究科情報科学専攻

0810060 水川 慎也

提出年月: 2011 年 2 月

概要

目的音と雑音の到来方向が異なっている場合、両耳受聴時に我々の目的音検知能力が著しく向上することが知られている。この現象は、方向性マスキング解除 (Spatial release from masking:SRM) と呼ばれる。これまで、多くの心理物理測定が行われ、両耳聴取時の目的音検知についての聴覚特性が徐々に明らかとされてきた。しかし、いまだ SRM のモデル化は行われておらず、SRM がどのように生じるのか明らかとされていない。両耳聴取時のマスキング解除についてのモデルは、両耳間時間差 (Interaural Time Different:ITD) を手掛かりとして生じるマスキング解除を説明するモデルが存在する。両耳間レベル差 (Interaural Level Different:ILD) を手掛かりとしても SRM が生じることが報告されているが、現在提案されている両耳聴モデルで ILD により生じるマスキング解除を説明可能なものはない。そこで、本論文では ITD と ILD を手掛かりとして生じる SRM のモデルの構築を行うことを目的とする。本モデルは、ヒトの聴覚中枢でマスキングを一括して除去するメカニズムが存在するという、方向性を利用したマスキング除去の考え方を採用している。ITD と ILD を手掛かりとしてどれだけマスキングの除去が行え、目的音対マスキングのパワー比がどれだけ改善をするかにより、方向性マスキング解除を示している。本モデルの特徴として ITD と ILD の影響を独立して考えることが可能で、最終的に 2 つの出力を結合する構成をしている。それにより、ITD か ILD だけが強く影響する場合も、両方が影響をあたえる場合も議論できるモデルとなっている。提案したモデルでいくつか SRM のシミュレーションを行った。その結果、これまで知られている SRM について精度良く模擬することが可能なことを示す出力が得られた。本研究で、ITD と ILD を手掛かりとして生じるマスキング解除がどのようなメカニズムで生起するのかを説明する両耳聴モデルの提案が行えた。

目次

第1章 序論	1
1.1 はじめに	1
1.2 本研究の背景	1
1.3 本研究の目的	3
1.4 本論文の構成	3
第2章 代表的な両耳聴モデルについて	5
2.1 はじめに	5
2.2 代表的な両耳聴モデル	5
2.3 Equalization - Cancellation モデル	7
2.4 おわりに	8
第3章 方向性マスキング解除の両耳聴モデルの構築	13
3.1 はじめに	13
3.2 提案モデルの概要	13
3.2.1 本モデルの構成についての検討	13
3.2.2 提案モデルの処理の流れ	14
3.3 入力音の分割フィルタの構成	15
3.4 両耳間時間差, 強度差の算出	16
3.5 ITD と ILD を利用したマスキング解除量の算出手法	17
3.5.1 ITD と ILD を利用した両耳のマスキングの除去	17
3.5.2 ITD を手掛かりとして生じるマスキング解除量の算出手法	18
3.5.3 ILD を手掛かりとして生じる方向性マスキング解除量の算出手法	18
3.6 両耳間時間差, 強度差を用いた算出手法の連結	19
3.7 おわりに	20
第4章 モデルを用いたシミュレーション	29
4.1 はじめに	29
4.2 目的音が純音で周波数が 1.0 kHz の場合	29
4.3 目的音が純音で周波数が 2.5 kHz の場合	29
4.4 目的音がパルス列信号の場合	30
4.5 複数の雑音源が存在する場合	30

4.5.1	提案モデルの問題点	31
4.5.2	解決策	31
4.6	考察	32
4.7	おわりに	32
第 5 章	モデルの評価	37
5.1	はじめに	37
5.2	モデルのパラメータに関する検討	37
5.3	モデルの評価実験	38
5.3.1	目的	38
5.3.2	実験刺激	38
5.3.3	実験方法	38
5.3.4	結果と考察	39
5.4	おわりに	39
第 6 章	結論	41
6.1	本論文で達成されたこと	41
6.2	今後の展望	42

第1章 序論

1.1 はじめに

我々は、日常生活を送る上で、聴覚から多くの情報を受け取っている。身の回りでは、人の会話の音声、スピーカーから流れる音楽、木々のざわめきなど常に多くの音情報が発せられている。我々はそのような音から、今までの経験や知識などを頼りに様々な情報を取り出すことができる。人間生活を送る上で、最も重要なコミュニケーションのひとつである会話も聴覚機能により成り立っている。

見方を変えて音から受け取ることができるもう一つの情報としてあるのが、その音が発せられている方向や距離である。その音が本来持つ情報とは別に、音源の存在する位置情報から間接的に情報を受け取ることができる。例えば、動いている救急車のサイレンが鳴っているとする。音源の方向、距離は常に変化し続ける、それにより救急車が移動している、または近くに存在するといった付随的な情報を受け取ることができる。

しかし、様々な音に囲まれているからこそその問題もある。友人の話し声が電車の走行音に妨害され聞こえなかったり、家電から発せられる報知音が雷の音で聞こえなかったりした経験があると思う。日常生活において、聞こえなければならない音が、何か別の妨害音の存在によって聞こえないことは多々ある。しかし、ヒトの聴覚機能は優れた機能を持っており、特定の方向の音を聴こうと考え、注意を向けることで音源の位置の違いを手掛かりに妨害音の中から欲しい音を拾う事ができる。音源の方向や距離の情報は音の知覚の手助けをする役割もある。

では、ヒトはどのように音源の存在する方向、距離を知覚するのだろうか。多くの人は人間が2つの耳で音を聴くことによって、方向、距離を知覚していると答えるだろう。しかし、ヒトが2つの耳で受け取った音響信号からどのようにして音源を知覚しているかについては、聴覚研究者の間で日々研究が進められているが、未だどのような処理が行われて方向定位が生じるか全貌は示されていない。ヒトの聴覚系でどのように両耳で受け取った音響信号が処理され方向定位が生じるか明らかとすることは、工学的な応用へと繋がる重要な課題である。

1.2 本研究の背景

ヒトは両耳で受け取った音響信号から、音源がどの方向、どの程度の距離に存在するのか知覚できる。音源方向を知覚できることにより、音の種類の違いとは別に、音の存在す

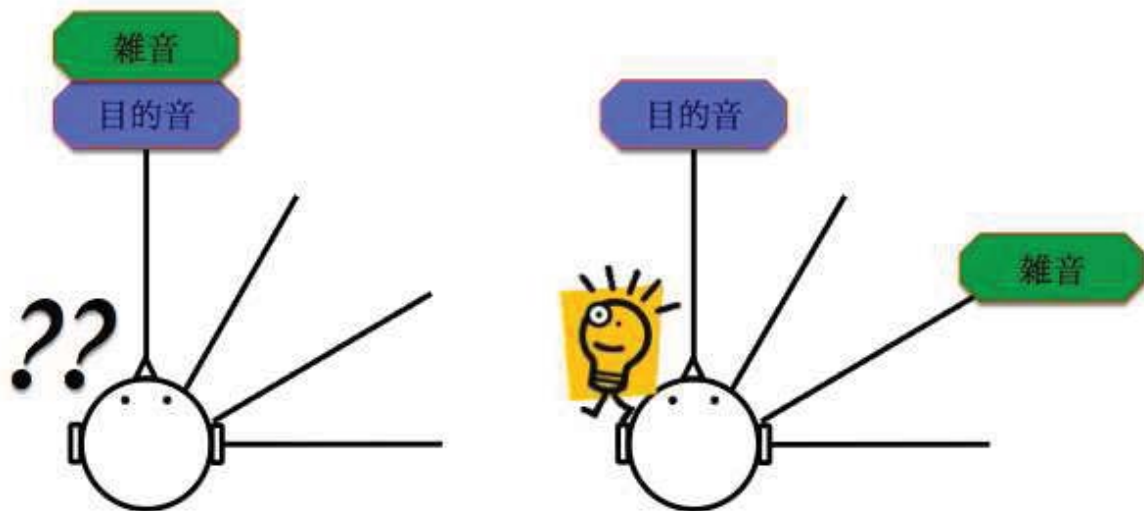


図 1.1: SRM が生じる状況. 右の図のように雑音と目的音の音源が異なった方向に存在しているときに目的音の検知能力が向上する

る方向的な違いを手掛かりに複数存在する音を聞き分けることが出来る. では, 身の回りの音について考えてみたとき, 複数の音源が同一の方向, 距離に存在していることは稀であり, 特定の方向に注意をはらい音を聴くことで, その方向に存在する音が聴こえやすくなることがある. この現象に注目した聴覚心理実験が Ebata らによって行われ, 信号音の方向性の手掛かりを利用することで, 信号音の検知能力が向上することが報告された [1]. 両耳受聴を行い, 音の音源方向を知覚することにより信号音の検知能力が著しく向上するのである. 関連して, Saberi らは, 自由音場において信号音と雑音が異なる方向から呈示されるとき, 信号音と雑音の音源が同じ方向に存在している場合に比べ, 2つの音源の存在する方向が異なっているほうが信号音の知覚が容易になると報告している [2]. この現象は, 方向性マスキング解除 (Spatial Release from Masking:SRM) と呼ばれる [3]. 図 1.1 に SRM が生じる状況を示す.

ヒトが両耳により音を知覚するとき, 方向定位の手掛かりとして両耳間時間差 (Interaural Time Difference:ITD) や両耳間レベル差 (Interaural Level Difference:ILD) を利用していると考えられている. ITD とは両耳に入力される信号に生じる時間差であり, ILD は両耳に入力される信号の音圧差である. 方向定位と同様に, SRM にも ITD と ILD が大きな影響を与えていることが知られている. これまでに SRM の知見に基づいて行われた調査について紹介する. 中西らは方向性の重要な手がかりである ITD が SRM に与える影響について調査している [4]. 白色雑音環境下において, 雑音, 信号音の呈示方向を ITD の制御によりコントロールし, 信号音の検知できるマスキング閾値を調べた. 実験の結果 ITD

のみを方向性の手掛かりとして SRM が生起することが明らかとなった。また、ITD のみにより生じる SRM は両耳間位相差 (Interaural Phase Difference:IPD) により生じる両耳マスキングレベル差 (Binaural Masking Level Difference:BMLD) と同様の傾向を示すことが報告された。また、内山らにより、雑音に自動車走行雑音を用い、報知音のマスキング閾値が測定された [5]。その結果、白色雑音下の実験と同様に ITD による SRM が生じることが明らかとされた。そして、黒田らにより、頭部伝達関数 (Head-Related Transfer Function:HRTF) を畳み込んだ雑音、信号音を用いて SRM の調査が行われた [6]。その結果、信号音が正中面に存在し、かつ信号音の高周波数成分を含む信号の場合 ITD だけではなく ILD によっても SRM が生起することが明らかとされた。以上より、ヒトの聴覚特性を知るための心理物理測定が行われ、ITD と ILD を手掛かりとして SRM が生じることが明らかとされたが、この現象のモデル化はこれまでに行われていない。もし、この現象を説明可能なモデルを構築できれば、ヒトの優れた聴取能力をコンピュータ上で実現し、利用することに繋がる。これをふまえて、本研究の目的を述べる。

1.3 本研究の目的

実環境における音の多くは方向性を持つ音であるため、SRM はヒトの目的音知覚において非常に有用であり、工学的に利用できれば、知覚されやすい目的音の呈示の確立に貢献できる。しかし、方向性情報を手掛かりとして生じるマスキング解除がどのようなメカニズムで生じるのか、その全貌は未だ明らかではない。これまで様々な実験により両耳聴取時の目的音検知についての聴覚特性が徐々に明らかとされてきた。しかし、この現象を説明可能でどのように SRM が生じるか示している両耳聴モデルは提案されていない。現在提案されている両耳聴モデルの多くが、ITD の検出やそれに基づく方向定位を示すモデルである。最近の両耳聴モデルの研究においては ILD についての議論が研究の中心となっている。また、両耳聴モデルの目的として ITD や ILD の検出、方向定位、特定方向の信号抽出などが中心であり、両耳聴取時に生じるマスキング解除現象を示すモデルは少ない。そのようなモデルの一つとして Equalization-cancellation (EC) モデル [7] があるが、このモデルは BMLD のモデルであり、SRM を説明可能かどうか明らかとなっていない。SRM を示すようなモデルが提案できれば、この現象の工学的な応用に繋がり、ヒトの聴覚特性をいかした目的音の呈示方向の検討や、補聴器などの両耳機能を模すアプリケーションの機能向上への利用が期待できる。そこで、本研究では ITD、ILD を手掛かりとして、ヒトの聴覚中枢においてどのようなメカニズムで SRM が生起するかについて説明する両耳聴モデルの提案を目的とする。

1.4 本論文の構成

本論文の構成を以下に示す。

第1章 本研究の背景と本研究に関わりのある先行研究について紹介する。また、未解決の問題について触れ、本研究で解決すべき目的を示す。

第2章 これまでに提案されてきた代表的な両耳聴モデルについて紹介する。そのなかで、自分の研究に取り入れることができるようなモデルの考え方などを見つけ出し、そのモデルについて検討を行う。

第3章 本研究で提案する方向性マス킹解除を説明する両耳聴モデルがどのように構成されており、どのようにしてマス킹解除現象について示しているのかについて記す。

第4章 本研究で提案した方向性マス킹解除のモデルを用いて、先行研究で行われた聴覚心理実験と同条件でSRMのシミュレーションを行う。それにより、本モデルがSRMのモデルと見なせるかの確認を行う。

第5章 本研究で提案したモデルがSRMを示すモデルとして適切なのか、モデルの持つパラメータの検討と今まで行われていないような実験刺激を用いてモデルの評価実験を行う。

第6章 本研究で明らかとなったことと今後の展望について述べる。

第2章 代表的な両耳聴モデルについて

2.1 はじめに

これまでに、音源方向推定や雑音除去を目的とした様々な両耳聴モデルが提案されている。多くの両耳聴モデルの基礎となるものが1948年にJeffressにより提案された両耳聴モデルである[8]。Jeffressの提案したモデルは聴覚中枢において左右両耳の入力に対応する遅延線が存在し、遅延を与えた左右の入力が最も良く一致する場所で音源方向が知覚されるというモデルである。Jeffressの提案したモデルは、ITDの検出のための基本的な考え方として広く知られ、時間差に基づく両耳聴モデルの多くはJeffressのモデルに基づいて提案されている。

2.2 代表的な両耳聴モデル

両耳間時間差検出モデルはJeffressのモデルに基づくものがほとんどである。Jeffressのモデルを拡張したモデルとしてよく知られるものにJeffress - Colburnモデルがある[9][10][11][12][13]。このモデルはCross - Correlationに基づく処理を行うモデルとして知られている。このモデルはITDの検出に基づいた、音源方向推定を行う事を目的としたモデルであり、左右両耳の入力に対して等間隔に複数の遅延回路が存在し、各場所で左右の入力に対して一定の遅延が与えられ、遅延が与えられた左右両耳信号について相互相関を行い、各場所での出力が決定される。もし、ITDが特定の値を持つならば左右からの入力に与える遅延がITDと対応する場所で最大の値が出力されることとなりITDが算出される仕組みとなっている。このモデルでは等間隔に配置された遅延回路によって場所情報が時間情報に対応することでITDが算出され、音源方向の推定が行われる。Jeffress - ColburnモデルにILDを考慮した処理は存在せずITDについてのみ考慮されている。もし左右の入力がILDを持つような音であったとしても、このモデルにおいてはITDに対応する場所で出力は最大となり、ILDを利用しての音源方向推定は行えないということになる。

さらに、Jeffress - Colburnモデルを拡張したモデルがいくつか提案されているのでそれについて紹介する。Blauertのグループは両耳聴取により生じるカクテルパーティー効果のモデル化を目的とする両耳聴モデルを提案している[14]。このモデルは音源方向推定と特定方向の信号音抽出機能を持つモデルである。まず、BlauertとCobbenによりDuifhuisにより提案された聴覚末梢系のモデルをJeffress - Colburnモデルに結合するこ

とが行われた。彼らはその後、両耳信号が持つ ILD の影響をモデルで表現するため複数のメカニズムをモデルの処理の中に取り入れた。そのなかで最もよく知られるメカニズムが Lindemann によって提案されたものである [15][16]。Lindemann は Jeffress - Colburn モデルに両耳の入力に対して対側抑制を行う機能を追加し、さらに、モデルの両端に片耳からの入力に対応する機能を実装した。Gaik がさらに Lindemann のモデルを拡張し、実測された頭部伝達関数の ITD と ILD の組み合わせから、ILD を ITD に変換可能な重み付け関数を用いて ITD と ILD の影響を、より自然に表現することを行った。

また、Breebaart らは生理学的な面から、Jeffress - Colburn モデルが聴覚中枢系に存在するニューロンの内の両耳の入力から興奮を受けるニューロンである EE (Excitatory - Excitatory) ニューロンの働きを漫然と示すものであると述べ、聴覚中枢系に存在するニューロンの内、片側の入力から興奮を受け、反対側の入力から抑制を受けるニューロンである EI (Excitatory - Inhibitory) ニューロンの働きを示す処理ブロックをモデルへ組み込むことを行った [18][19][20]。それにより、Breebaart のモデルは ITD と ILD の推定を同時に行うことを可能としている。

他にも、信号分離を目的とするモデルとしてよく知られる両耳聴モデルとして周波数領域両耳聴モデルというものが提案されている [21]。前述した両耳聴モデルは、帯域分割を行った信号に対して、時間領域での処理を行っているのに対して、周波数領域両耳聴モデルでは周波数領域で処理を行っていることが特徴としてあげられる。このモデルは音源の方向推定を行うとともに、方向性に基づいた信号分離を行うモデルである。このモデルでは、まず時間信号を周波数領域に変換し、クロススペクトルを利用し、IPD と ILD を算出する。その算出した IPD、ILD とあらかじめ保持している HRTF データベースより算出できる IPD と ILD の値を対応させることで音源方向推定を行ない、その推定方向を利用し、信号分離を行う。このモデルは IPD と ILD のどちらも利用する両耳聴モデルとなっている。信号音の周波数が分かっているならば IPD は ITD に変換可能であり、このモデルは ITD と ILD の両方を利用する両耳聴モデルである。

これらのモデルは音源方向推定や雑音除去に基づく音源分離などを目的としており本研究で問題としたいマスキング解除について説明するモデルとは目的が異なる。両耳聴取時のマスキング解除を説明するモデルとしては、IPD を手掛かりとして生じる BMLD を説明する Durlack の提案した Equalization - Cancellation (EC) モデル [7] が知られている。EC モデルは、ヒトの聴覚中枢系において、片耳に入力されたマスキング音ともう一方の耳に入力されるマスキング音を等価化する Equalization 処理、そして、等価化したマスキング音を打ち消す Cancellation 処理を行うメカニズムが存在するという仮定に基づいたモデルである。この一連の処理を EC 処理と呼ぶ。EC 処理後の出力が、方向性情報を利用してマスキング音を打ち消したあとに残る信号音を検知するための指標となり、これを用いて、ヒトが方向性を持つ音を両耳聴取する際の、目的音の検知能力の向上を説明している。

2.3 Equalization - Cancellation モデル

ここでは、本研究に関連した両耳聴取により生じるマスキング解除のモデルである EC モデルについての検討を行う。Equalization 処理で行っていることは次のような処理になる。特定の方向に音源の存在するマスキャーがあると考え、両耳に入力されるマスキャーは特定の ITD を持つ信号である。両耳の入力信号から算出される ITD を用いて、方側に入力されたマスキャーをその ITD と同じ時間だけ遅らせる処理を行う。両耳のマスキャーの時間的なズレを無くすことを行えば、両耳のマスキャーは ITD を持たないこととなり時間的な面だけで方向定位が生じるとするならば、音源が正中面に存在する状態に等しくなる。その後、Cancellation 処理で一致させた両耳のマスキャーについて、片側のマスキャーからもう一方の側のマスキャーを引くことでマスキャーの打ち消しを行う。EC 処理によってどの程度マスキャーが除去され、どの程度知覚したい目的音が残っているかを示す指標として次式が与えられている

$$f_m = \frac{(1 + a^2)(1 + \sigma_\epsilon^2) - 2a \cos(\omega\tau) \exp(-\omega^2\sigma_\delta^2)}{2[1 + \sigma_\epsilon^2 - \exp(-\omega^2\sigma_\delta^2)]} \quad (2.1)$$

ここで、 a は両耳信号の強さの比であり、 τ は ITD に対応するものであり、 ω は検知したい目的音の周波数になる。 σ_ϵ と σ_δ はそれぞれ強度差と時間差の分散を表すモデルのパラメータとなっている。出力である f は EC factor と呼ばれ、EC 処理後の信号音とマスキャーのパワー比に相当するものである。添字の m は目的音の音源の存在している方向を基準として、マスキャーの音源の存在している方向が相対的な角度で何度異なっているを示している。この EC factor を利用して BMLD を算出するためには、信号音、マスキャーがともに正面にある S_0N_0 条件のときの EC factor f_0 と信号音、マスキャーのどちらかが方向性を持つ N_0S_m もしくは N_mS_0 条件のときの EC factor f_m を求めることにより

$$\text{BMLD} = 10 \log_{10} \frac{f_m}{f_0} \quad (2.2)$$

式 2.2 により S_0N_0 条件のときに比べ N_0S_m もしくは N_mS_0 条件のときどれだけ目的音を知覚可能なマスキング域値が低下するかを算出することが可能である。

EC モデルは簡単な仮定により ITD を手掛かりとして生じるマスキング解除現象を説明しており、方向性マスキング解除のモデルを構築する上で利用出来る考え方であるが、EC モデルが先行研究のどの程度まで説明可能であり、推定精度はどの程度なのかの検討が行われていないため、まず、EC モデルについての検討を行う必要がある。検討には中西らが行った聴取実験のデータを利用する [4]。EC モデルを用いて中西らが行った聴覚心理実験と同じ実験刺激を用いてシミュレーションを行う。この聴覚心理実験は ITD のみ持つ両耳信号を用いて行われたため EC モデルでこの現象をシミュレートすることが可能であると考えられる。このとき用いた ITD は Kuhn によって提案された ITD の近似式を用いて算出した [22]。

$$\text{ITD} = \frac{d}{c} = \frac{r(\theta + \sin \theta)}{c} \quad (2.3)$$

r [m] は平均的な人間の頭部半径, θ [rad] は音源の到来方向, c [m/s] は音速, d [m] は両耳間の経路差を表す. 本シミュレーションでは, 頭部半径 r を 0.09 m, 音速 c を 343.5 m/s とした. 目的信号の周波数は 0.1 kHz ~ 2.5 kHz まで 0.1 kHz 刻みで変化させる. 刺激音の提示方向 θ は, 聴取者の正面を 0° として, 右前方水平面上を 1° 間隔で 90° まで変化させた. 時間差とレベル差のゆらぎを表す EC モデルのパラメータ σ_ϵ と σ_δ は Durlack の論文にて用いられている $\sigma_\epsilon = 0.25$, $\sigma_\delta = 105 \times 10^{-6}$ を利用しマスキング解除量を算出した.

その結果を, 図 2.1 に示す. Azimuth の軸が目的音の呈示角度を示し, Frequency の軸が目的音の周波数である. 高さの軸がマスキング解除量となる. 色が青に近い場所ほどマスキング解除量が多い条件になる. また, 中西らが明らかとした, 目的音が純音でその周波数が 1.0 kHz, 1.5 kHz, 2.0 kHz, 2.5 kHz のときに ITD を手掛かりとして生じる SRM とそれと同条件の EC モデルの出力を並べて表示したものを図 2.2 に示す. EC モデルによる出力の模擬精度を確認するため, 聴覚心理実験の結果とシミュレーション結果の平均二乗誤差を求めると, EC モデルにより算出されたマスキング解除量は 0.95 dB であり, かなり精度良く模擬可能なことが確認された. また, ITD を手掛かりとして生じるマスキング解除量は, 入力される目的音信号の周波数によって特定の位置にピークが現れるが, その特徴も示している. この結果から EC モデルの出力はマスキング解除量, 特徴とも中西らの聴取実験の結果と一致し, EC モデルは ITD を手掛かりとして生じるマスキング解除のモデルであると考えられる. EC モデルを利用することで ITD により生じるマスキング解除量は良い精度で算出可能であることが示された.

次に, EC モデルに黒田らが行った聴取実験 [6] と同様の, ITD だけではなく ILD も同時に存在する音を入力した場合, どのような結果を出力するのか検討を行う. 結果を図 2.3 に示す. 結果を見てわかるように EC モデルの出力は聴取実験の結果と大きくかけ離れた結果を示すことが確認された. EC モデルでは ILD を手掛かりとして生じるマスキング解除については表現できない問題があるため, ILD の存在する音を入力した場合に模擬精度は大きく低下する.

2.4 おわりに

ここでは, 既存の両耳聴モデルについて述べてきた. 既存の両耳聴モデルは Jeffress の提案した両耳間時間差算出を目的とする古典的なモデルに対して様々な研究者が拡張を行い, 研究が進められてきた. しかし, そのモデルの多くは時間差情報を利用するようなモデルであり, レベル差情報をどのように扱うかが, 最近の両耳聴モデルの研究においてひとつの課題となっていることがわかる. またモデルのコンセプトとしても, 音源方向推定や聴覚生理学的な面から ITD や ILD をどのように検出しているかといったことを説明するようなモデルが研究の中心となっており, 方向性情報を手掛かりとして生じるマスキング解除現象についてのモデルは EC モデルの他に報告されていない.

両耳聴モデルのなかでも本研究に関連して重要なモデルであると考えられる EC モデ

ルについての検討を行い、その結果、EC モデルでは入力される両耳信号が ITD のみ持っているような場合は、精度良くシミュレートが可能であったが ILD の存在している音を入力とした場合、模擬精度は著しく低下することが確認された。この原因は EC モデルにおいて両耳信号の amplitude の情報にあたる入力部分が ILD を想定して設計されたものではなく、片耳信号か両耳信号かを表すためだけに存在しているからである。式 2.1 において a は本来ならば両耳信号の比から求めるもので ILD に対応するものと考えることができるが EC モデルでは a には 1 か 0 の入力時しか正しくシミュレートすることができない。 $a = 1$ のときというのは両耳信号にレベル差が存在していない時であり、 $a = 0$ というときは雑音信号は両耳信号にもかかわらず目的信号は単耳信号であるときを表している。この制約条件のもと、EC モデルは精度良くマスキング解除量の算出が可能となる。EC モデルは音の方向性情報のうちの時間差情報だけを利用してのマスキング除去理論を提案し、それを用いることで ITD を手掛かりとして生じるマスキング解除量の算出を可能としたモデルである。

EC モデルの最も重要な考え方はマスキアーに時間的なズレを与えて一致させ、両耳のマスキアーを打ち消し打ち消し前と後で、信号とマスキアーの比がどれだけ改善されたかである。両耳の信号が単一の周波数で構成されたもので時間差がわずかしかない場合、EC 処理後の出力は小さくなるが、両耳の信号がちょうど半周期時間差があるような両耳信号であった場合、EC 処理後の出力は元の信号より大きくなる。このように EC モデルはシンプルな仮定でニューロンの働きである興奮と抑制のどちらも表現することができる。ITD と ILD は興奮と抑制によって検出されるならば、EC モデルの考え方を利用することで ITD と ILD の両方に対応したモデルを構築することが可能であり、本研究の目的とする方向性マスキング解除を説明する両耳聴モデルを構築する上で非常に重要なモデルである。

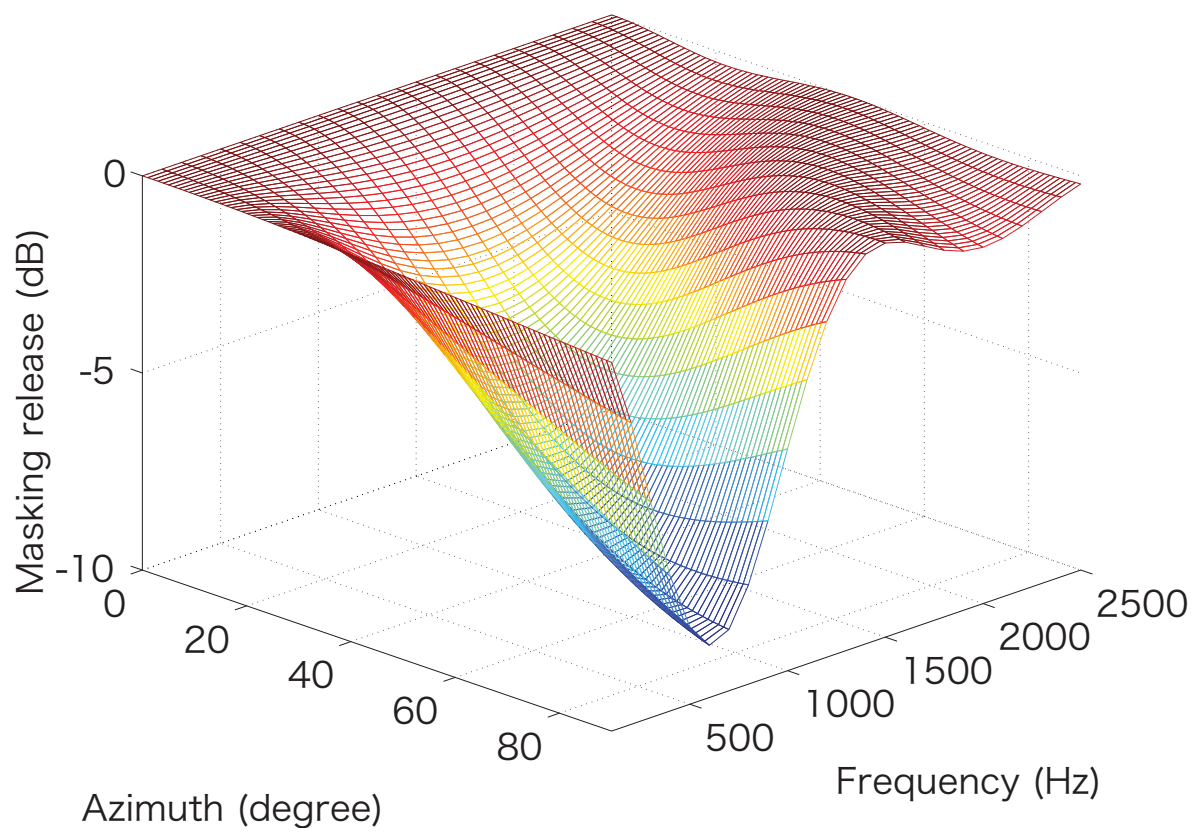


図 2.1: EC モデルで計算したマスキング解除量. 入力信号の周波数, 音源の存在する角度とマスキング解除量の関係を示した図

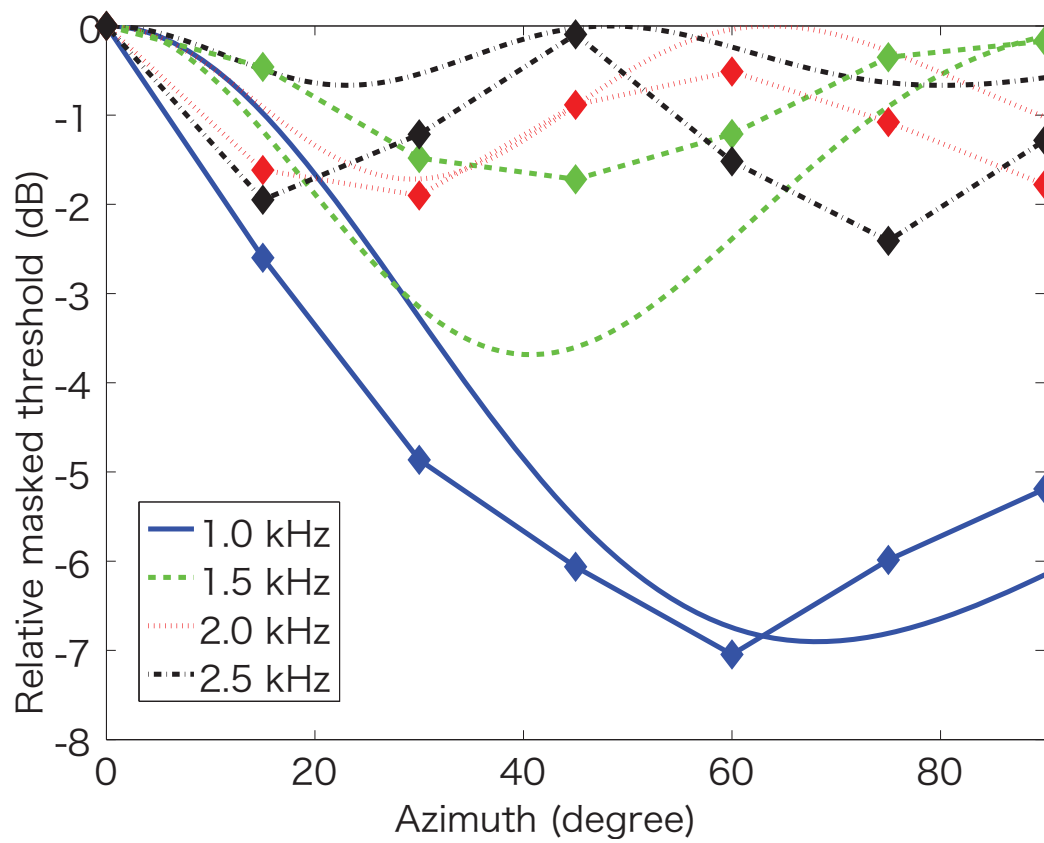


図 2.2: マーカーなしのラインが EC モデルを用いて算出したマスキング解除量, マーカー付きのラインは中西らの聴取実験の結果

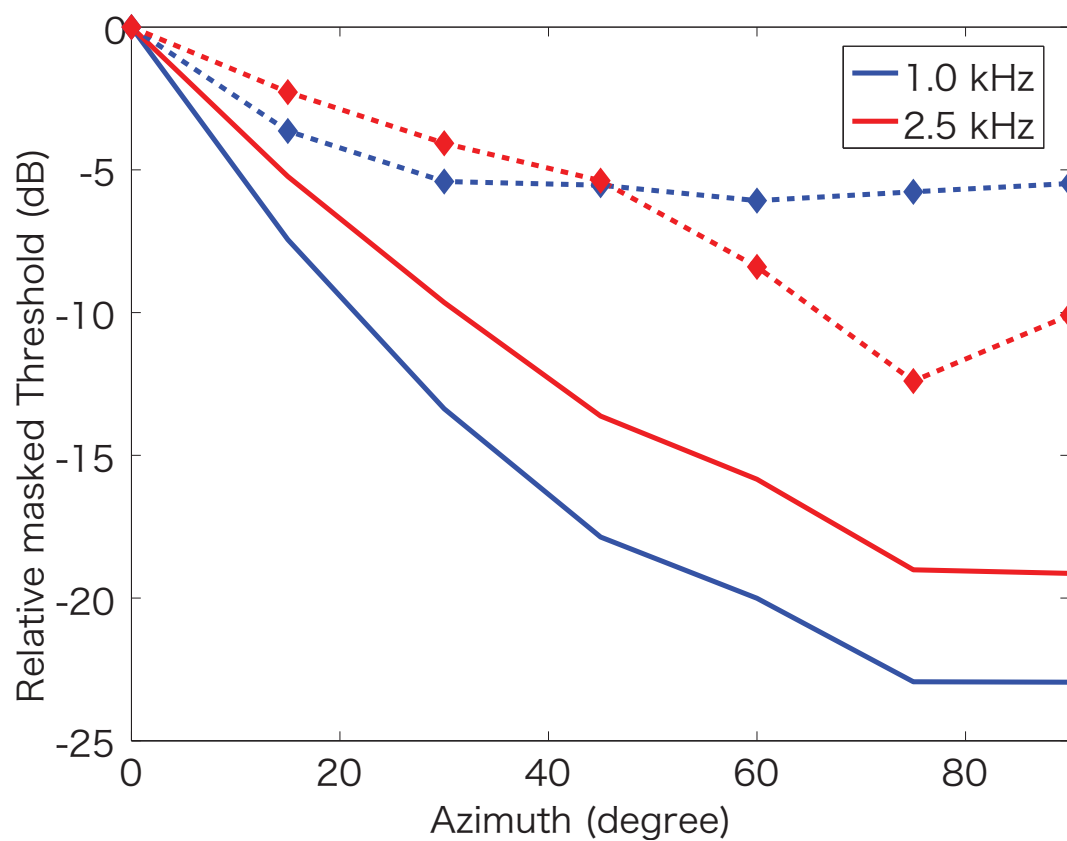


図 2.3: EC モデルの入力に黒田らの聴取実験と同様の ITD と ILD をともに持つような音を用いたときの出力. 青のラインが目的音の周波数が 1.0 kHz の場合, 赤のラインが目的音の周波数が 2.5 kHz の場合, とともにマーカー付きのラインが黒田らの聴取実験の結果で, マーカー無しのラインがモデルの出力

第3章 方向性マスキング解除の両耳聴モデルの構築

3.1 はじめに

前章では、音の方向性に基づくマスキング解除モデルの一つとして EC モデルについて検討した。その結果、EC モデルでは ITD を手掛かりとして生じるマスキング解除について精度良く模擬する結果を示したが、入力信号が大きな ILD を持っている場合、模擬精度は大きく低下することが確認された。しかし、方向性情報を利用したマスキング解除モデルは EC モデル以外に報告されておらず、ITD と ILD 両方を手掛かりとして生じるマスキング解除現象を示すようなモデルは提案されていない。実環境において特定方向に音源の存在する音を両耳聴取する場合、ITD のみが生じる場合は殆ど存在せず、ITD と ILD の両方を手掛かりに方向定位が行われ、マスキング解除が生じる。そのため、ITD だけでは不十分であると考え、ITD と ILD 両方を考慮したモデルの構築を行う。

3.2 提案モデルの概要

本研究で提案する方向性マスキング解除モデルでは、ITD と ILD の 2 つの両耳間差により生じるマスキング解除量の算出を可能にすることを目的とする。構築の方法としては Durlach の提案した EC メカニズムの考え方を基本とし、EC モデルでは考慮されていない ILD を手掛かりとして生じるマスキング解除についても説明できるようにモデルの拡張を行う。しかし、提案モデルでは出来る限りヒトの聴覚中枢系の機構を考慮し、信頼性のあるモデルとしていきたいと考える。そのため、まず提案モデルをどのように構成するかを決定するため聴覚中枢系の機構について検討する。

3.2.1 本モデルの構成についての検討

音の方向性により生じるマスキング解除は両耳で音を聞くことにより生じ、両耳からの入力が交差する場所である聴覚中枢系と深く関係している。そのため、モデルの構成を考える前にヒトの聴覚中枢系の機構について整理する。ヒトの聴覚中枢系で左右の耳からの入力初めて交差するのが上オリーブ複合体 (Superior Olivary Complex:SOC) と知られている。これは複数の核からなり、主要なものとして上オリーブ内側核 (MSO:Medial

Superior Olive), 上オリーブ外側核 (Lateral superior olive:LSO), 大形体内側核 (Medial Nucleus of Trapezoid Body:MNTB) がある.

MSO は左右からの入力に対する時間差検出を行う器官であると考えられる. MSO に存在するニューロンの多くは同側の入力及び, 反対側の入力から興奮を受ける EE ニューロンである. これらのニューロンは反応を引き起こす最適な時間差で並んでいる. 両耳間時間差に基づき音源方向推定を示しているような Jeffress のモデルや相互相関モデルは MSO についてのモデルといえる. EC モデルも ITD を利用しての時間的な処理を行う点から, MSO より出力されたものを利用したモデルであると考えられる.

もうひとつの器官である LSO は左右からの入力に対するレベル差検出を行う. LSO に存在するニューロンの多くは同側の入力から興奮を受け, 反対側の入力から抑制を受ける EI ニューロンである. 両耳の発火の強さの差から対側抑制後に残るものが強度差として検出される. MSO と LSO は並列して存在し, 左右の入力がそれぞれに対して入力される.

本モデルではヒトの聴覚中枢系において ITD と ILD が別々の器官で検出されていることを考慮し, ITD と ILD をあらかじめ混合して考えるのではなく, 時間差とレベル差を処理するブロックを並列に定義することを考える. それにより, ITD と ILD を手掛かりとして生じるマスキング解除量をそれぞれ独立に求める方法を提案する. また, MSO, LSO はそれぞれ低周波に良く反応するニューロン, 高周波に良く反応するニューロンを多く持つ. 関連して, 比較的低周波の純音の検知には ITD が重要な役割を果たし, 高周波の純音では ILD が重要な役割を示すことが知られている [23]. 方向性マスキング解除も同様の傾向を示すことが明らかとされている. ITD を利用したモデルと ILD を利用したモデルの出力に, 入力信号の周波数成分に応じた重み付けを行い最終的に結合する構成とする. モデルの入力は目的音とマスキングの両耳信号であり, それを S_l , S_r , N_l , N_r と表記する. モデルの最終的な出力は方向性マスキング解除量とする.

3.2.2 提案モデルの処理の流れ

提案モデルの処理の流れを図 3.1 に示す. 本モデルの処理の流れは次のようになる

1. 聴覚末梢系を考慮したフィルタバンクを用いて, 両耳の入力信号 S_l , S_r , N_l , N_r をそれぞれ帯域分割する.
2. 帯域分割された両耳信号から, 各帯域ごとに ITD と ILD の算出を行う.
3. 算出された ITD と ILD を用いて, 各帯域ごとに ITD を手掛かりとして生じるマスキング解除量と ILD を手掛かりとして生じるマスキング解除量の算出を行う.
4. ITD と ILD それぞれ独立に算出されたマスキング解除量に対して, ITD と ILD のどちらを有効に利用出来るか決定するためそれぞれの出力に重み付けを行い, 各帯域のマスキング解除量を決定する.

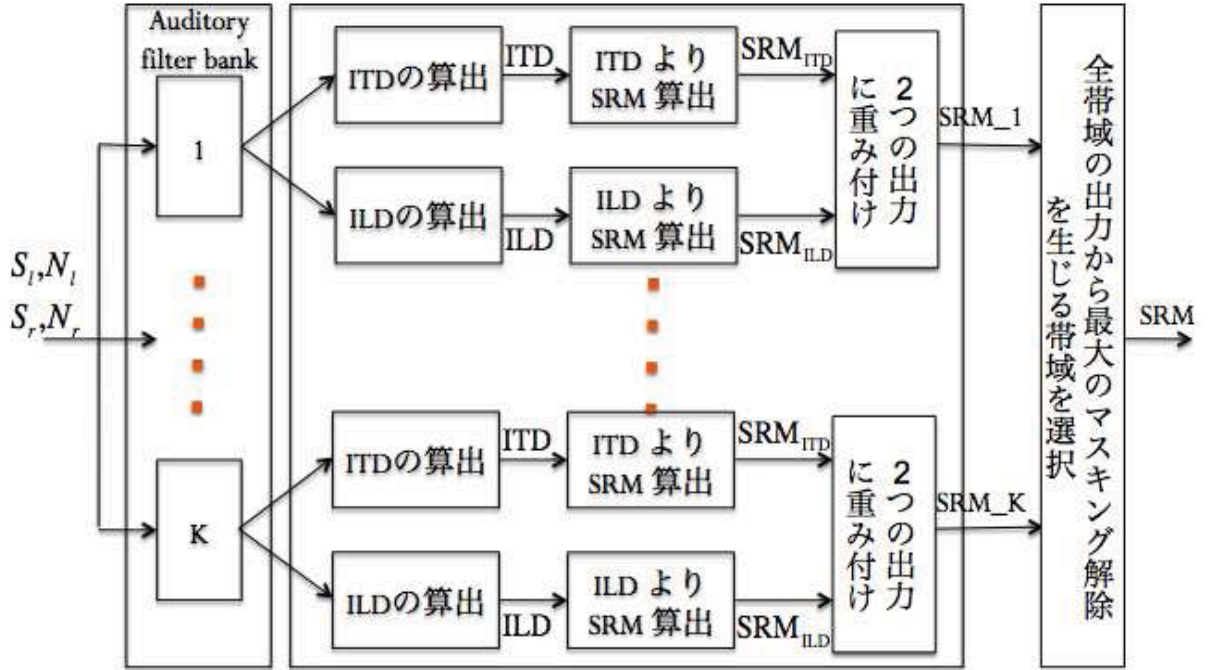


図 3.1: 新しく提案する方向性マスキング解除のモデルの処理の流れ

5. 各帯域ごとに算出されたマスキング解除量の中から最大のマスキング解除を生じる帯域を知覚に用いられる帯域として選択，最終的なマスキング解除量を決定する。

今後，各ブロックで行われている事について詳細に記していく。

3.3 入力音の分割フィルタの構成

ITD は基本的に入力される両耳信号の周波数に依存せず，大きな違いは生じないが，ILD は入力される両耳信号の成分周波数によって大きく変化する．純音で考えるならば入力信号の周波数が低いほど ILD は小さくなり，高いほど ILD は大きくなる傾向がある．入力信号が純音の場合は問題ないが，白色雑音や，パルス列信号などの複合音を考える場合，帯域分割を行い各帯域ごとに処理を行わなければならない．そこで，両耳入力信号である S_l, S_r, N_l, N_r を周波数分割し，各帯域ごとに今後の処理を行う．周波数分割に用いるフィルタバンクは聴覚末梢系の基底膜で行われる周波数分析機能を模擬するガムマトーンフィルタバンクを用いる [24][25]．分割後の信号は $S_{l|n}, S_{r|n}, N_{l|n}, N_{r|n}$ と表記する． n は帯域分割した後の n チャンネル目の出力であることを示す．

図 3.2 のガンマトーンフィルタバンクは最低中心周波数を 500 Hz 最高中心周波数を 3000 Hz とし 20 帯域に分割するフィルタバンクである。図 3.3 にガンマトーンフィルタバンクの中心周波数と帯域幅の関係について示す。本モデルではこの条件で設計したフィルタバンクを利用しているが、最低中心周波数を 500 Hz 最高中心周波数を 3000 Hz としている理由は、ヒトが水平方向の定位に用いる主要な周波数成分をカバーしていることが挙げられる。垂直方向についても考える場合は、それに対応するフィルタバンクが必要になる可能性がある。

帯域分割した信号とマスキングについて各帯域ごとの SNR (Signal-to-Noise Ratio:SNR) を算出し、最も大きな SNR の帯域と他の全帯域を比較し、マスキングに比べて信号音の成分が著しく小さいと考えられる帯域に関しては、マスキングに埋もれた信号音を知覚するのが困難な帯域であると考えてこの段階で取り除く。この目的は本モデルが出力するマスキング解除量は正面にマスキングと信号音が存在するときに比べて、側面にマスキングが移動する場合にどれだけマスキング閾値が低下するかであり、どれだけ SNR が小さい帯域であっても SNR に関係なく相対的な低下量であるマスキング解除量は算出される。しかし、その帯域で信号音が全く知覚することができないような弱いパワーしか持っていない場合、その帯域を利用して信号音の知覚を行っていると考えすることは難しい。そのため、あらかじめ取り除くことを行っている、それにより、今後の計算量を減らす目的もある。これ以降の全ての処理は分割後の各帯域の信号音、マスキングについて行う。

3.4 両耳間時間差，強度差の算出

図 3.4 に帯域分割後の両耳信号の一例を示す。図から見て取れるように、両耳信号は強い相関を持つような形状をしている。まず、周波数分割を行った信号音とマスキングから、各周波数帯毎の ITD と ILD を算出する。

ITD の算出は両耳入力信号について、右耳の入力と左耳の入力で相互相関処理を行い算出する。

$$R_{lr|n}(\tau_n) = \int_{-\infty}^{\infty} S_{l|n}(t)S_{r|n}(t - \tau)dt \quad (3.1)$$

このとき $R_{lr|n}$ が最大の値となる τ_n が求める ITD である。本来、相互相関モデルは入力信号そのものではなく、入力信号より生成される神経インパルスの発火列である。今回は、ITD の推定モデルを用いての検出ではなく、工学的な処理で入力信号そのものについて相互相関処理を行い、ITD を算出している。HRTF をたたみ込むことで作った方向性を持つ音では問題なく ITD の算出が行えるが、実環境で録音された音などから上式で ITD を算出しようとする、 $R_{lr|n}$ が最大の値となる τ_n が非常に大きな値を示したりすることがあり、ITD が上手く求められないこともある。明らかに異常な値が算出された場合は、Kuhn の近似式で補うことを行う。方向性を持つパルス列信号を入力とし、算出した ITD を図 3.5 に示す。縦軸が ITD [s] を示し、横軸が呈示角度である。各帯域の ITD は図から見て取れるように多少の違いは存在するが、どの帯域でもほぼ一定の値を

示すことがわかる。この多少の違いは算出誤差とも考えられるが、Kuhn の近似式で算出する値と大きな違いは存在せず許容範囲と考える。

ILD は各帯域の両耳入力信号について、右耳の入力と左耳の入力のパワー比を求めることで算出する。

$$a_n = \frac{\int_0^T S_{r|n}^2(t)dt}{\int_0^T S_{l|n}^2(t)dt} \quad (3.2)$$

a_n を dB 表記したものが ILD であるが、今後の処理では a_n の値そのものを利用する。方向性を持つホワイトノイズを入力とし、算出した ILD を図 3.6 に示す。縦軸が ILD [dB] を示し、横軸が呈示角度である。ILD は各帯域である程度のばらつきが生じている。1.0 kHz 以下の低い周波数の帯域では ILD は 10 dB 以下の値を示すことが多いが、2.5 kHz の高い周波数を含む帯域では ILD が 20 dB 以上の値を示すこともある。また、音源の位置が側頭部に移動する程、各帯域で ILD の値は大きく異なってくることがわかる。

3.5 ITD と ILD を利用したマスキング解除量の算出手法

ここでは、各帯域ごとに算出した、ITD と ILD を用いてマスキング解除量を算出する方法を提案する。

3.5.1 ITD と ILD を利用した両耳のマスキングの除去

算出方法は EC 理論に基づくものであり、両耳間差を用いてマスキングを一致させ除去し、その後残る信号音によって、その音の知覚が行われているとする。

はじめに、両耳に入力されるマスキングを一致させる方法を提案する。図 3.4 で示したように、両耳信号は高い相関を持つ信号であり、両耳に入力される信号音の違いとして ITD と ILD の 2 つが重要な要素で他を無視できるものと仮定すると、両耳信号について次式が成り立つ。

$$N_{l|n} = a_n N_{r|n}(t - \tau_n) \quad (3.3)$$

ここで τ は ITD であり、 a_n は ILD に相当するパワー比である。実際に ITD と ILD を用いて両耳信号を可能なかぎり一致させたものを図 3.7 に示す。実際の信号は ITD と ILD 意外にも微量な違いを有しており、ITD と ILD だけで完全に一致させることができるものではない。しかし、図 3.7 に見て取れるように ITD と ILD という 2 つの重要な両耳パラメータによってかなりの精度で両耳のマスキングを一致させることができる。次に両耳のマスキングを打ち消すことを考える。マスキングを除去するためには、一致させたマスキングの片側からもう一方を引けばよい。図 3.7 の両耳信号について打ち消し処理を行った結果を図 3.8 に示す。マスキング成分の多くがこの打ち消し処理によって除去されることが見て取

れる。この例の信号の打ち消し後の信号と元の信号のパワー比は -22 dB となっていた。以降、この両耳のマスキングを一致させ、除去するという考え方にに基づきマスキングが除去された後に残る信号音の成分によって方向性を手掛かりとして生じるマスキング解除現象のモデルの構築を行う。

3.5.2 ITD を手掛かりとして生じるマスキング解除量の算出手法

まず、ITD を手掛かりとして生じるマスキング解除量の算出法を示す。両耳のマスキングを一致させるために ITD を利用して片側のマスキングに τ だけ時間差補正を行っている。このとき、同時に補正を行った側の信号音に対しても同様の時間差補正を行う。その後、マスキングと信号音それぞれについて片側の入力ともう一方の入力で打ち消し処理を行う。マスキングを打ち消すために行った時間差補正分だけ、信号音にズレが生じ、打ち消し処理をしても信号音の成分が残る。それがマスキング中に存在する信号音の知覚に用いられる指標となりそれを次のように求める

$$f_{\tau m} = \frac{\int_0^T (S_{l|n}(t) - S_{r|n}(t - \tau_n - \sigma_\delta))^2}{\int_0^T (N_{l|n}(t) - (a_n + \sigma_\epsilon)N_{r|n}(t - \tau_n - \sigma_\delta))^2} \quad (3.4)$$

σ_ϵ と σ_δ はマスキングを一致する際にゆらぎを持たせるためのモデルのパラメータである。ITD を手掛かりとして生じる SRM を算出するためには、 S_0N_0 条件のときの $f_{\tau 0}$ に比べて S_0N_m 条件のときの $f_{\tau m}$ を求めて、 $f_{\tau 0}$ に比べて $f_{\tau m}$ が相対的にどれだけ大きいのか

$$\text{SRM}_{\tau_n} = 10 \log_{10} \frac{f_{\tau m}}{f_{\tau 0}} \quad (3.5)$$

で算出する。

3.5.3 ILD を手掛かりとして生じる方向性マスキング解除量の算出手法

次に、ITD を手掛かりとして生じるマスキング解除量の算出と同様の考え方にに基づき ILD を手掛かりとして生じるマスキング解除量の算出法を提案する。両耳のマスキングを一致させるときに、ILD を利用してパワーの小さい側のマスキングを a だけ増幅する事を行っている。それと同時に補正を行った側の信号音に対して信号音の増幅処理を行うこととする。その後、マスキングと信号音それぞれについて片側の入力ともう一方の入力で打ち消し処理を行う。信号音の増幅を行ったため、反対側の信号では打ち消しきれず、信号音の成分が残る。それを次のように求める

$$f_{am} = \frac{\int_0^T (S_{l|n}(t) - (a_n + \sigma_\epsilon)S_{r|n}(t))^2}{\int_0^T (N_{l|n}(t) - (a_n + \sigma_\epsilon)N_{r|n}(t - \tau_n - \sigma_\delta))^2} \quad (3.6)$$

SRM を算出するためには ITD の処理と同様に f_{a0} に比べて f_{am} が相対的にどれだけ大きいかが、

$$\text{SRM}_{a_n} = 10 \log_{10} \frac{f_{am}}{f_{a0}} \quad (3.7)$$

で算出する。

以上より、ITD と ILD それぞれを手掛かりとして生じる SRM を独立に求めることが可能となる。

3.6 両耳間時間差，強度差を用いた算出手法の連結

一般的に ITD は 1.5 kHz 以下の低周波の信号音の知覚に有効に働き、ILD は 2.0 kHz の高周波の信号音の知覚に有効に働くと言われている。入力された信号音の周波数成分を基に ITD と ILD がそれぞれどの程度知覚に影響を与えるかを決定する必要がある。そこで、最終的な SRM の決定方法として ITD を手掛かりとして生じる SRM_{τ_n} と ILD を手掛かりとし生じる SRM_{a_n} にそれぞれ重み付けを行う。重み付けの方法として人の聴覚神経の位相同期指数を用いることにした。ヒトの聴覚系では、刺激に対して位相同期した神経インパルスの発火が生じ、その発火が聴覚中枢で相互相関処理され、時間差情報を知覚する。位相同期指数は入力される刺激が低周波の場合比較的高い値を示すが、高周波では非常に低くなる。Jhonson の研究によれば特徴周波数 (Characteristic Frequency: CF) 1000 Hz 以下の位相同期指数は 0.7~0.9 程度と報告されている [26]。CF が高くなるにつれて位相同期指数は大きく低下し、ITD による音源定位能力も低下する。位相同期指数の低下が、高周波音での ITD を手掛かりとして生じるマスキング解除量が減少することの原因であると考えられる。また、時間情報と強度差情報が相補的な関係にあると考え、時間情報の有効性の低下に基づき強度差情報が有効になるようにした。重み付けには Blackburn と Sachs によって与えられた聴神経の位相同期指数の近似曲線を用いている [27]。

$$\alpha = 0.97 - 0.16cf - 0.01cf^2 \quad (3.8)$$

図 3.9 に ITD の出力に与える重みを示す。最終的な SRM を ITD から得られる SRM_{τ_n} と ILD から得られる SRM_{a_n} により次式のように決定している。

$$\text{SRM}_n = \alpha \text{SRM}_{\tau_n} + (1 - \alpha) \text{SRM}_{a_n} \quad (3.9)$$

また、ここまでの処理はすべて帯域分割した各帯域ごとに行われ、SRM は各帯域ごとに算出される。ヒトが複数の周波数成分を持つ音を知覚するとき、最も知覚しやすい周波数帯の音を聞くと仮定し、最大のマスキング解除を生じる周波数帯の出力を最終的な出力とする。

$$\text{SRM} = \max(\text{SRM}_n) (n = 1, 2, 3 \dots) \quad (3.10)$$

3.7 おわりに

本章では、新しい方向性マスキング解除のモデルの提案を行った。このモデルは聴覚中枢系の上オリーブ複合体において ITD と ILD を検出する器官が並列して存在することにヒントを得て、ITD と ILD を手掛かりとして生じるマスキング解除をそれぞれ独立に議論できるモデルとなっている。その利点として ITD と ILD 両方存在するような入力であろうとも、ITD と ILD それぞれどちらかを手掛かりとして生じるマスキング解除量を算出することもできるし、ITD と ILD のどちらも影響するような状況でもマスキング解除量の算出が行える。ITD と ILD が相補的に働くような状況においては、最終的に ITD と ILD のどちらがどの程度手掛かりとして利用されるか、入力信号の成分周波数をもとに聴神経インパルスのゆらぎである位相同期指数を用いて決定している。そして、最終的に分割した全帯域から最大のマスキング解除を生じる帯域を選択し、最終的な出力を決定した。以上より、ITD と ILD を手掛かりとして生じるマスキング解除量を独立して求めることが可能なモデルの構築を行った。

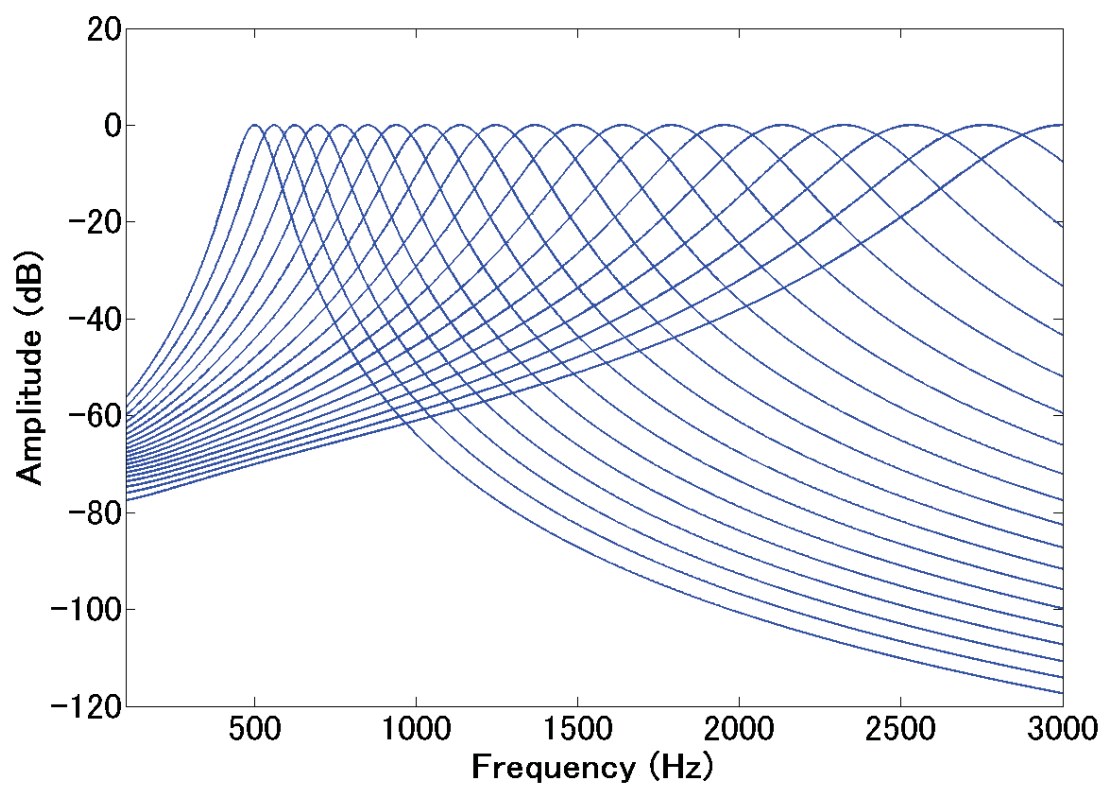


図 3.2: 帯域分割に用いるガンマトーンフィルタバンクの周波数特性, 最低中心周波数 0.5 kHz 最高中心周波数 3.0 kHz でその間を 20 分割するようなフィルタバンク

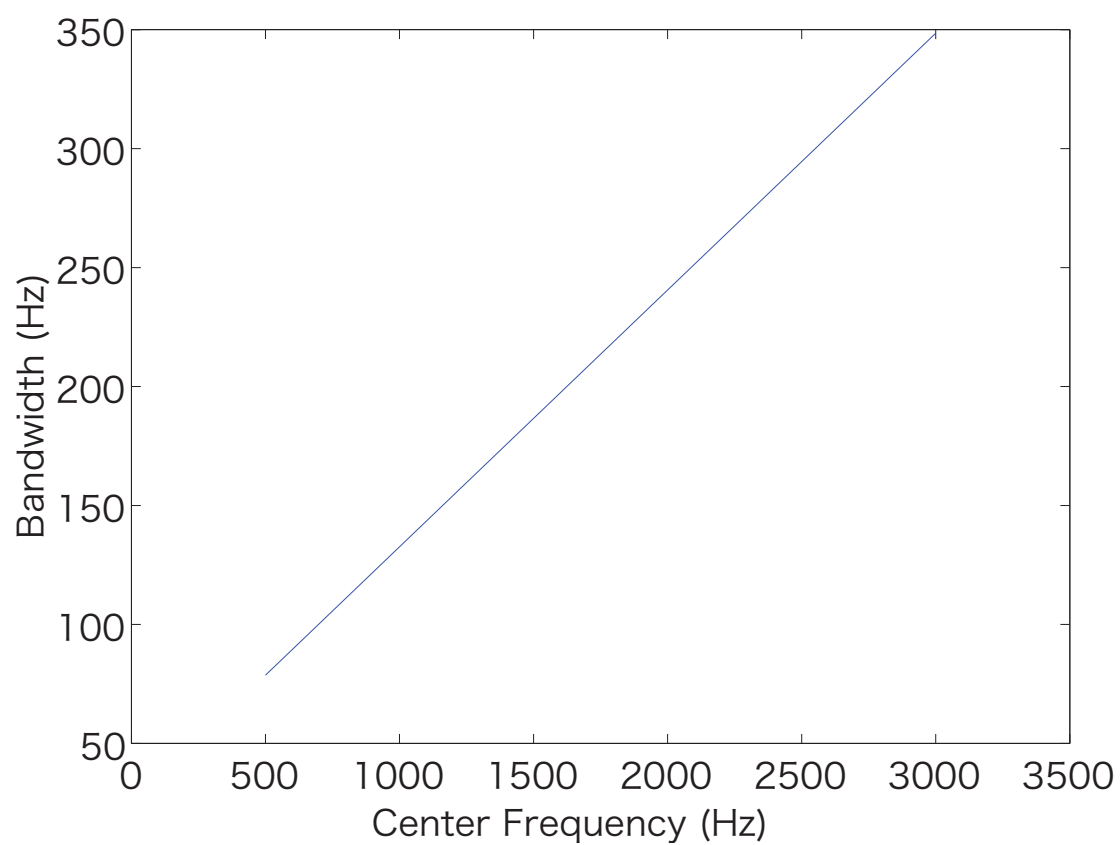


図 3.3: 帯域分割に用いるガンマトーンフィルタバンクの中心周波数と帯域幅の関係

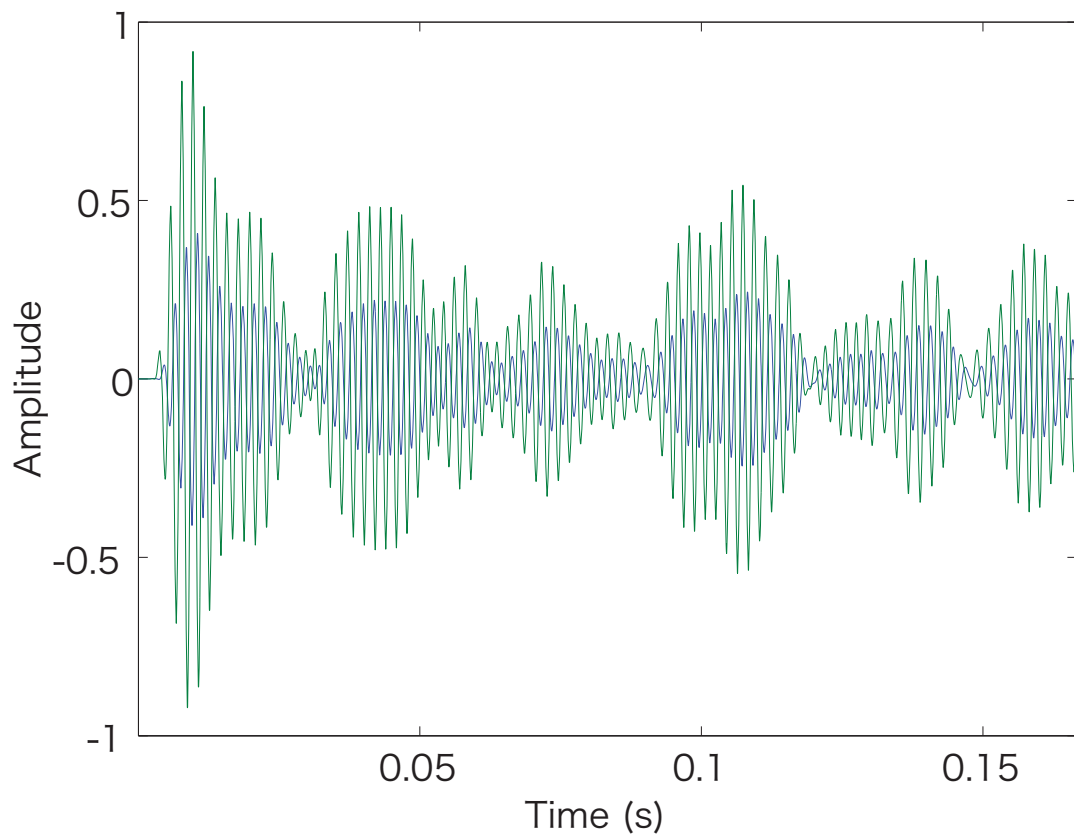


図 3.4: ITD と ILD を持つような両耳信号の一例. これは, 右側面 90° に音源が存在する走行雑音を, ガンマトーンフィルタバンクで帯域分割した 1 チャンネル目の信号の一部分を切り出したもの. 青で表される信号が左耳に入力される信号で, 緑で表される信号が右耳に入力される信号.

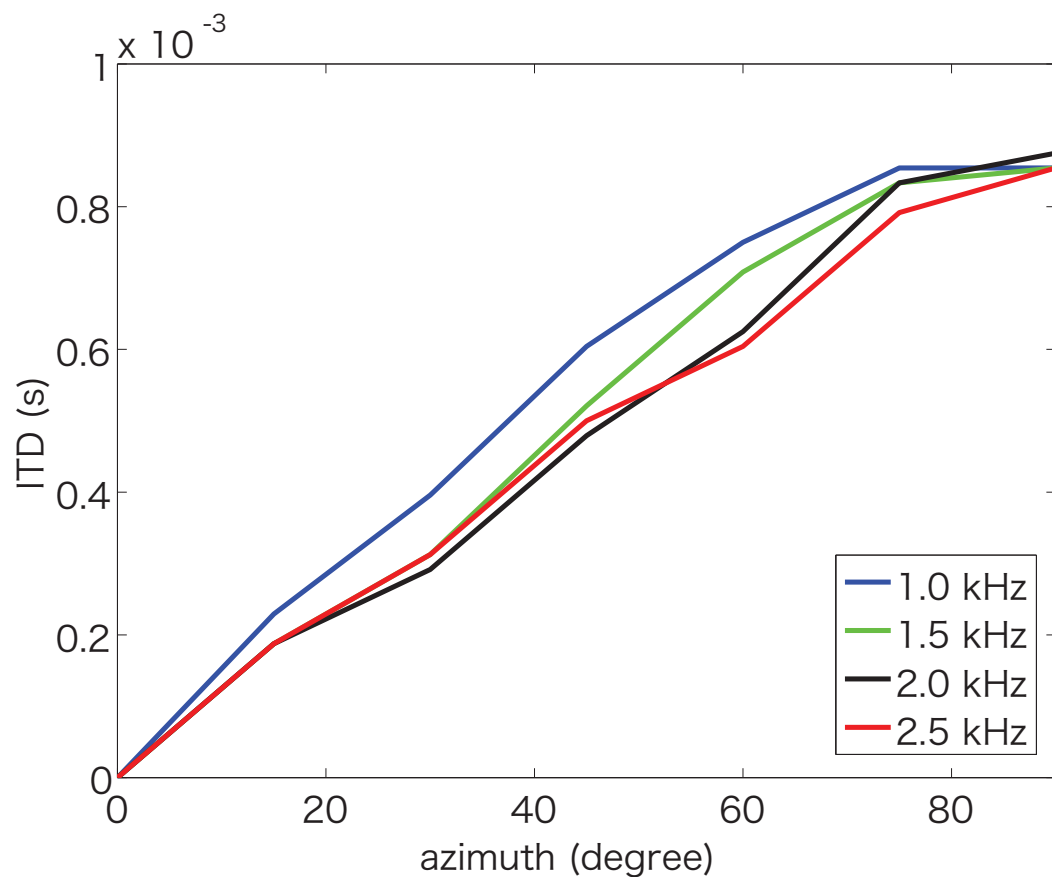


図 3.5: 算出された ITD のなかから 1.0 kHz, 1.5 kHz, 2.0 kHz, 2.5 kHz に対応する帯域の結果を表示している

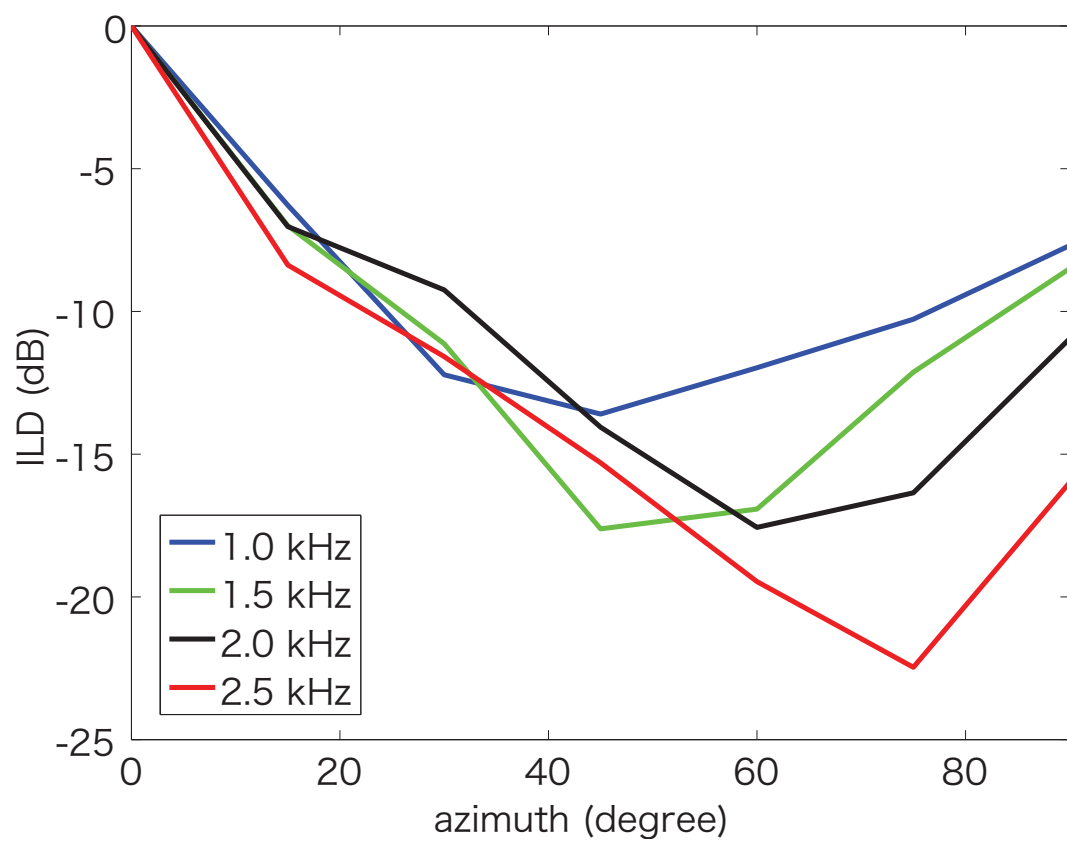


図 3.6: 算出された ILD のなかから 1.0 kHz, 1.5 kHz, 2.0 kHz, 2.5 kHz に対応する帯域の結果を表示している

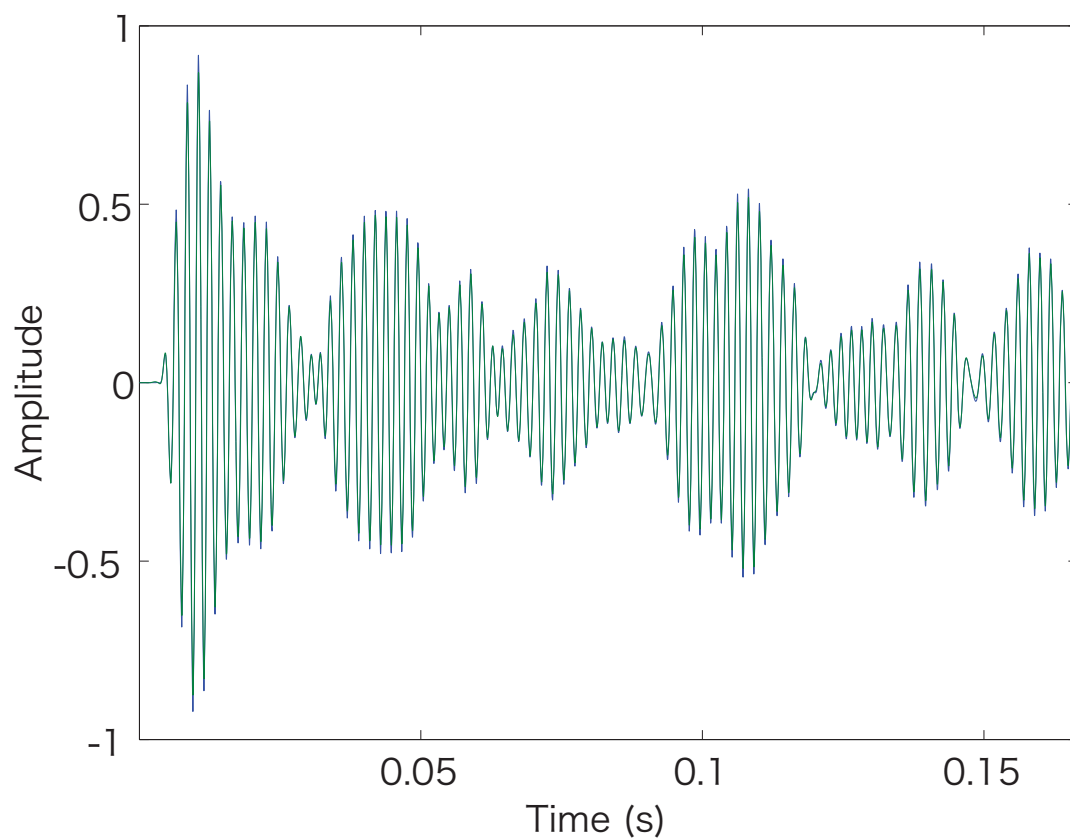


図 3.7: ITD と ILD を用いて両耳のマスキングを一致させた後の両耳信号の一例。これは、図 3.4 を ITD と ILD によって補正したもの。青で表される信号が左耳の信号で、緑で表される信号が右耳の信号。

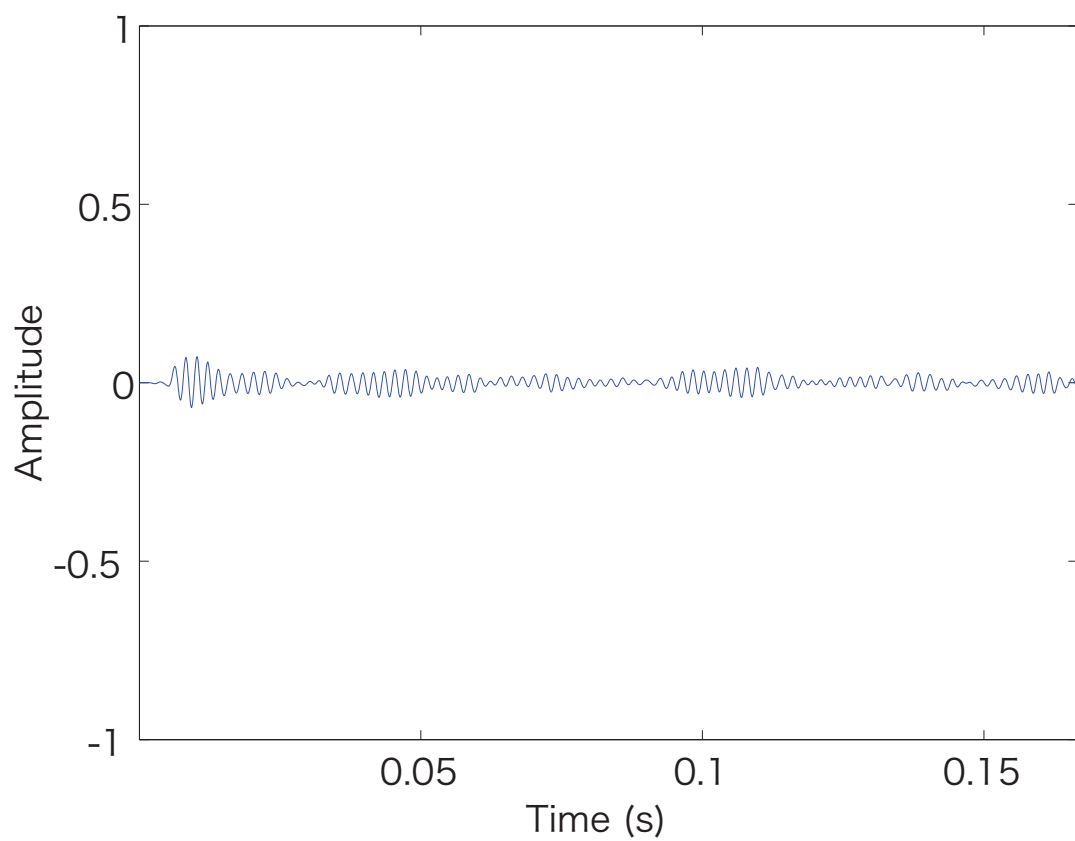


図 3.8: 一致させた両耳信号を打ち消した後の信号, ITD と ILD で一致できなかったものが打ち消し後に残る

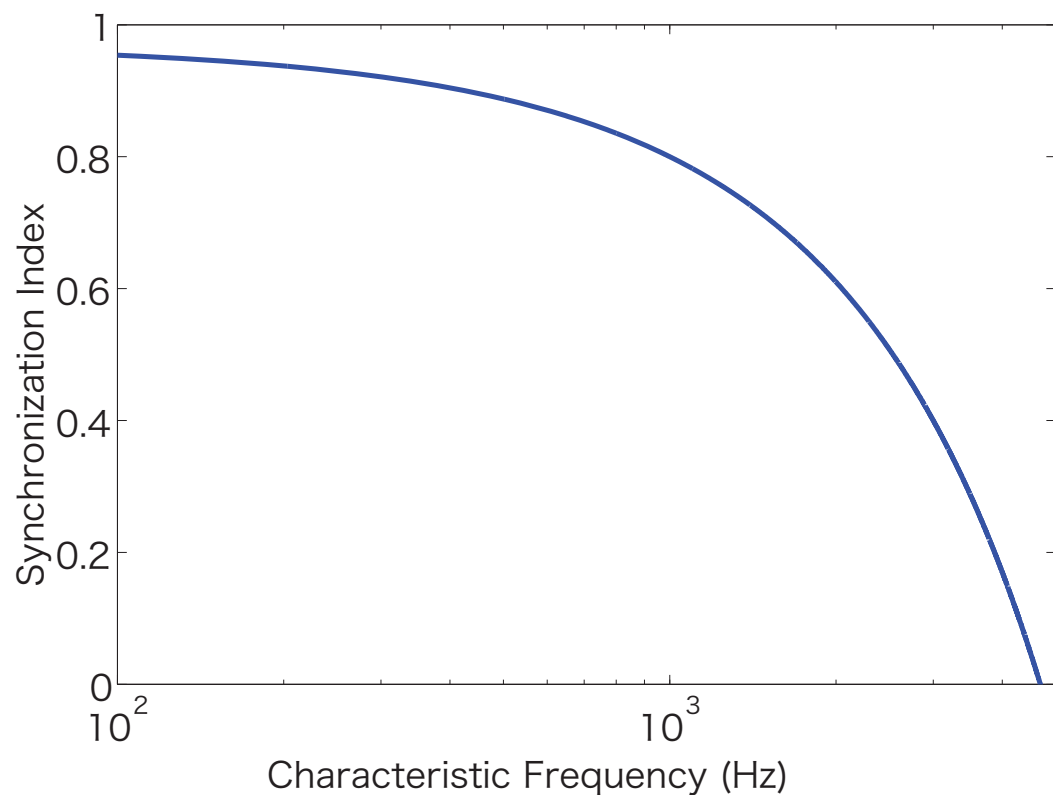


図 3.9: Blackburn と Sachs によって与えられた聴神経の位相同期指数の近似曲線

第4章 モデルを用いたシミュレーション

4.1 はじめに

前章で、本研究で目的とする ITD と ILD を手掛かりとして生じるマスキング解除量の算出を可能とする SRM のモデルの提案を行った。しかし、そのモデルが SRM を正しく模擬できるモデルであるということの確認を行う必要がある。そこで、黒田らによって行われた聴取実験 [6] と同様の実験刺激をモデルへの入力信号として用いて SRM のシミュレーションを行う。提案したモデルの出力と黒田らの聴取実験の結果を比較して、本モデルが SRM のモデルと見なせるか確認する。それにより、本提案モデルが先行研究で行われた聴覚心理実験で明らかにされている SRM について、模擬可能なモデルであることを証明する。

4.2 目的音が純音で周波数が 1.0 kHz の場合

目的信号が純音で、周波数が 1.0 kHz の場合について、SRM のシミュレーションを行う。マスキャーとして用いる雑音は自動車走行雑音を用いる。走行雑音は走行速度 60 km/h の窓を開けた環境下において、自動車内でデジタル録音された走行雑音を利用した。モデルのパラメータである時間誤差とレベル誤差の値は $\sigma_e = 0.25$, $\sigma_\delta = 105 \times 10^{-6}$ を用いて SRM を算出した。シミュレーション結果と黒田らの聴取実験の結果をあわせて図 4.1 に示す。縦軸がマスキング解除量を示し、横軸が呈示角度である。マーカーのついてない青のラインがモデルにより得られたマスキング解除量で、マーカー付きの赤のラインが黒田らの聴取実験により測定された結果である。推定誤差を求めると 1.30 dB となり、SRM においても個人差のバラつきなどで 2 ～ 3 dB の測定誤差は生じることから、本モデルはそれなりに良い精度で SRM の推定が行えていると考えられる。

4.3 目的音が純音で周波数が 2.5 kHz の場合

目的信号の純音の周波数が 2.5 kHz の場合の SRM のシミュレーションを行う。マスキャーとして用いる雑音は、1.0 kHz の時と同様に走行雑音を用いる。モデルのパラメータである時間誤差とレベル誤差の値は $\sigma_e = 0.25$, $\sigma_\delta = 105 \times 10^{-6}$ を用いて SRM を算出した。シミュレーション結果と黒田らの聴取実験の結果をあわせて図 4.2 に示す。縦軸がマ

スキング解除量を示し、横軸が呈示角度である。マーカーのついてない青のラインがモデルにより得られたマスキング解除量で、マーカー付きの赤のラインが黒田らの聴取実験により測定された結果である。推定誤差を求めると 1.27 dB となり、1.0 kHz とほぼ同じ精度でマスキング解除量の推定が行えていることが確認できた。

4.4 目的音がパルス列信号の場合

ここでは、目的信号にパルス列信号を用いた場合の SRM のシミュレーションを行う。マスキングとして用いる雑音は、純音の時と同様に走行雑音を用いる。モデルのパラメータである時間誤差とレベル誤差の値は $\sigma_e = 0.25$, $\sigma_\delta = 105 \times 10^{-6}$ を用いて SRM を算出した。シミュレーション結果と黒田らの聴取実験の結果をあわせて図 4.3 に示す。縦軸がマスキング解除量を示し、横軸が呈示角度である。マーカーのついてない青のラインがモデルにより得られたマスキング解除量で、マーカー付きの赤のラインが黒田らの聴取実験により測定された結果である。推定誤差を求めると 0.99 dB となり、純音の時よりも多少ながら模擬精度が良い結果を出力した。

4.5 複数の雑音源が存在する場合

提案モデルでは、信号音源と雑音源ともに 1 つの場合に利用出来るモデルとなっている。しかし、雑音源が 2 つ以上存在するような場合についても利用可能なモデルであればモデルの適応範囲が広がる。黒田らは雑音源が 2 つ存在するときの方向性マスキング解除についての聴取実験も行っており、それは次のような 3 条件で行われた。

1. $S_0 N_{45} N_{315}$
2. $S_0 N_{90} N_{315}$
3. $S_0 N_{90} N_{270}$

それぞれの条件でマスキング解除量は次のようになると報告された

1. 約 5 dB
2. 約 8 ～ 9 dB
3. 約 10 dB

4.5.1 提案モデルの問題点

提案モデルでは両耳に入力されるマスキングの両耳間差を用いて EC 処理を行うが、1. , 3. の条件においては 2 つの雑音がまったく同じものである場合、両耳から入力されるマスキングが同じになってしまい、本モデルでは 2 つの雑音が合わさったものは正面に存在するマスキングであるとみなされてしまう。このとき、ITD も ILD も存在せず、方向性情報を利用することができない。そのため、本モデルではマスキング解除量を算出することができないといった問題点がある。しかし、黒田らが行った聴取実験においてはかなり大きなマスキング解除が生じることが報告されている。そのため、2 つ以上の雑音源がある場合についても何らかの方法でマスキング解除量を算出する方法を考える。

本モデルで扱えない条件は図 4.4 のように複数の雑音源が左右対称となるような位置に存在する場合である。この問題をモデルで用いる処理の見直しなどで解決することはモデルのコンセプトから変更を加える必要があり、モデルで用いる処理に変更を加えて解決することは困難であると考えている。

4.5.2 解決策

複数の雑音源が存在するときに例外的な処理を行い、マスキング解除量の算出を可能とする。黒田らの聴取実験の結果を、雑音源が一つしか存在していない場合と比較すると、1. の場合は S_0N_{45} 条件のマスキング解除量とほぼ等しくなり、2. の場合は 45° と 90° の間に雑音源が存在するような場合と等しいマスキング解除量となっている。3. の場合は S_0N_{90} 条件のマスキング解除量とほぼ等しい結果になっている。このことから、複数の雑音源が左右対称となるような位置に存在する場合のマスキング解除量は片側だけ存在するときと近い値を示し、片側だけに存在する条件に置き換えて考えることができる。

柏野と平原は、同時に提示された単語に対する話者の人数判断は、話者数が 2 名までの場合は正確であるが、3 名以上の話者から提示された場合人数判断能力は著しく低下することを報告している [28]。そのことから、雑音源が複数存在する場合のマスキング解除量が、片側だけに存在しているときと同じ傾向を示す原因は、複数の雑音源すべてを利用してマスキング解除が生じるのではなく、一つの目的音ともう一つの音源のマスキングの 2 つまでしか有効に働かないからだと考えられる。そのため 2 つの雑音源が存在する場合は、片側だけに雑音源が 1 つ存在するような状況に次のように置き換える。2 つのマスキングから入力される信号を $N_{1|l}$ $N_{1|r}$ $N_{2|l}$ $N_{2|r}$ とすると、

$$N_l = (N_{1|l} + N_{2|r})/2 \quad (4.1)$$

$$N_r = (N_{1|r} + N_{2|l})/2 \quad (4.2)$$

複数の雑音源が左右対称となるような位置に存在する場合は片側に存在する状況に置き換えられ、 N_{90} N_{315} のような場合は N_{45} と N_{90} のちょうど中心の方向にマスキングが存在するような状況に置き換えられる。以上の例外的な処理を行うことで、2 つの雑音源が存

在する場合でもマスキング解除量の算出が可能になる。しかし、黒田らの結果と近い出力は得られるが、比較する聴取実験の結果がそれほど多くあるわけではないため、検討の余地を残す。

4.6 考察

提案モデルにより算出されたマスキング解除量はどのシミュレーションにおいても、ピークの出方、解除量とも、黒田らの聴取実験の結果と同様の特徴、近い値を示した。この結果から、提案モデルが ITD だけではなく ILD が存在している音により生じるマスキング解除量も算出可能なモデルとなっていることが確認された。また、モデルに雑音源が 2 つ存在している場合に対応するための例外的な処理を追加した、それにより、雑音源が 2 つ存在している場合についてのマスキング解除量の算出も行えるようになった。

4.7 おわりに

新しく提案したマスキング解除のモデルが SRM を正しく模擬できるのかを確認した。提案したモデルを用いて、黒田らの聴取実験を参考とし SRM のシミュレーションを行った結果、本モデルの出力は ITD と ILD の両方が存在している信号を入力とした場合でも、聴取実験の結果を精度よく模擬する結果を得た。この結果から、本モデルは ITD と ILD を手掛かりとして生じる SRM についてのモデルと解釈できることが確認された。

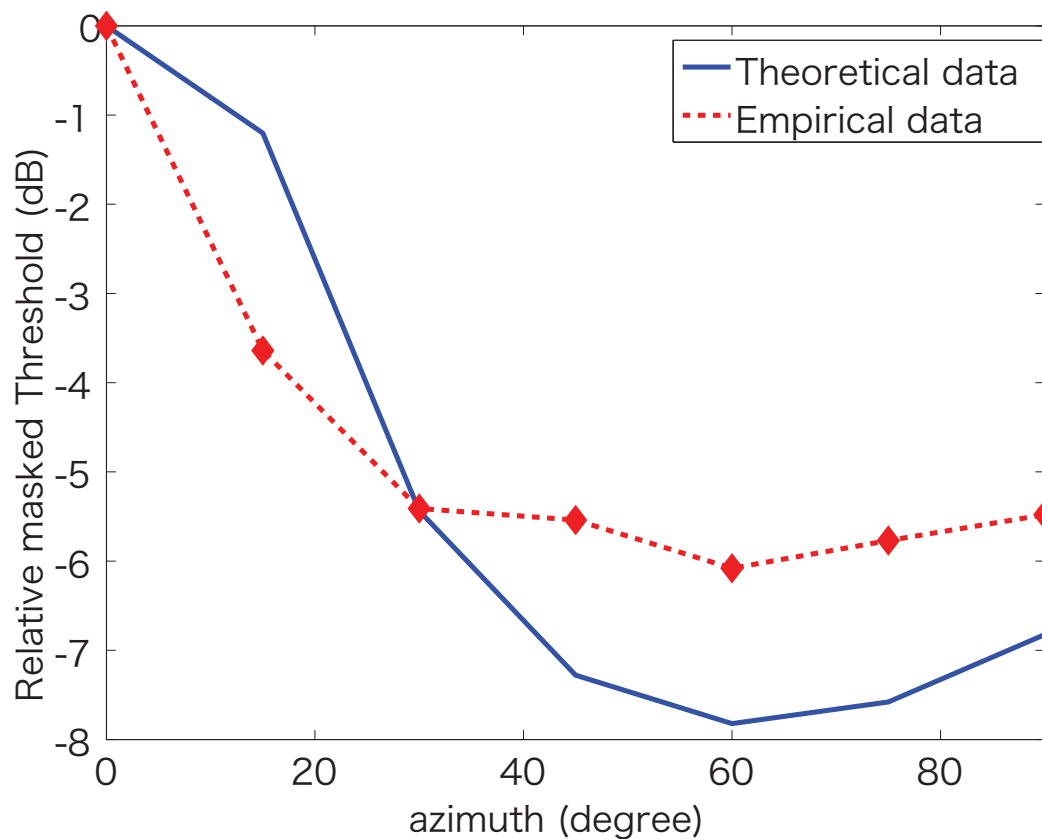


図 4.1: 目的音に 1.0 kHz の純音を用いてシミュレーションを行った結果. マーカーなしの青のラインが提案モデルを用いて算出したマスキング解除量. マーカー付きの赤のラインは黒田らが行った聴取実験結果

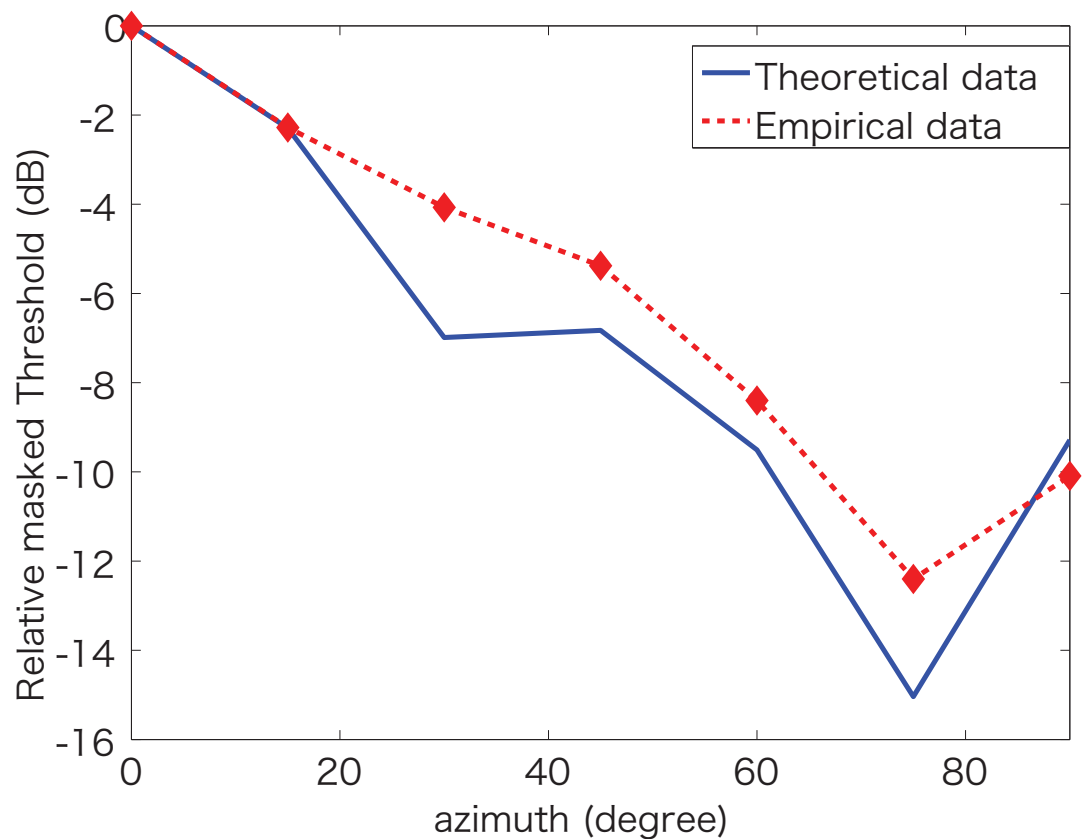


図 4.2: 目的音に 2.5 kHz の純音を用いてシミュレーションを行った結果. マーカーなしの青のラインが提案モデルを用いて算出したマスキング解除量. マーカー付きの赤のラインは黒田らが行った聴取実験結果

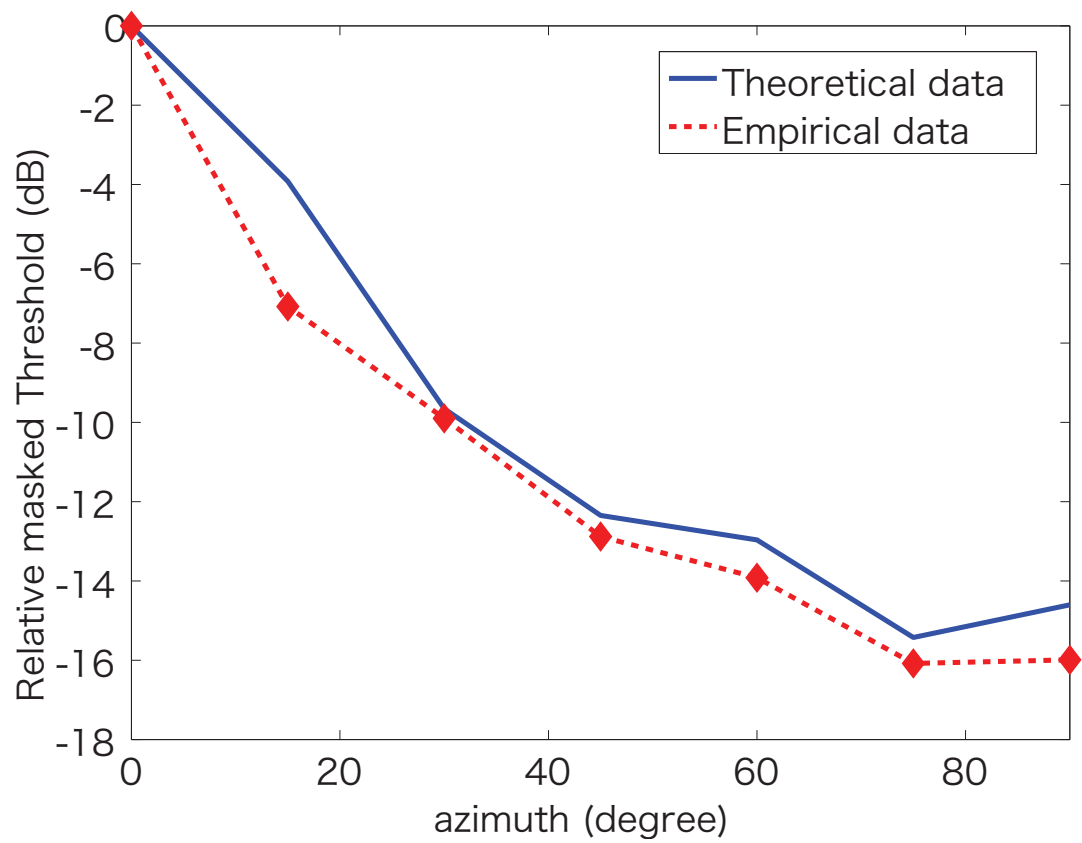


図 4.3: 目的音にパルス列信号を用いてシミュレーションを行った結果. マーカーなしの青のラインが提案モデルを用いて算出したマスキング解除量. マーカー付きの赤のラインは黒田らが行った聴取実験結果

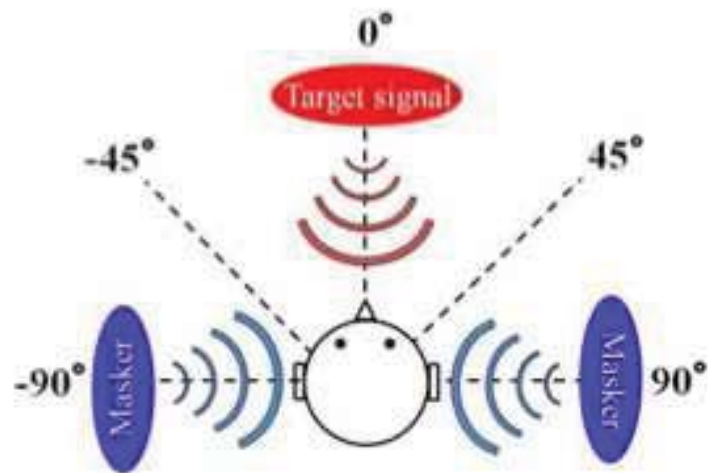


図 4.4: 複数のマスキャーが聴取者を中心として左右対称となるような位置に存在する場合

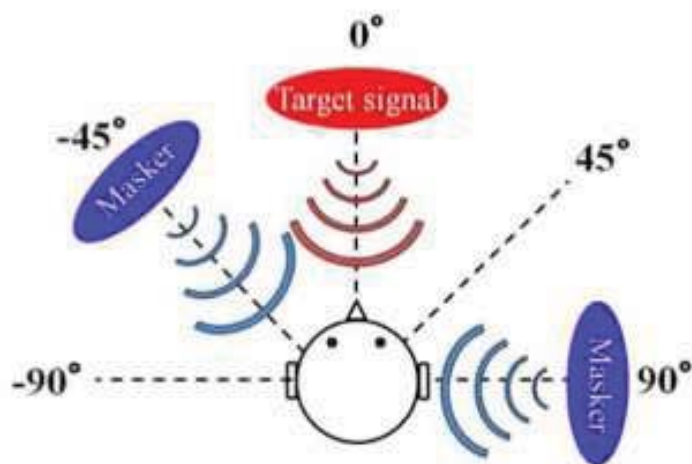


図 4.5: 複数のマスキャーが聴取者を中心として左右非対称となるような位置に存在する場合

第5章 モデルの評価

5.1 はじめに

前章で、提案したモデルを利用して、既に聴覚心理実験が行われて明らかとされている SRM についてのシミュレーションを行った。その結果、本提案モデルは SRM のモデルであると考えられる結果を出力した。ここでは、提案したモデルについての検討、評価を行いたいと考える。まず、本モデルのパラメータについての検討からモデルの評価を行う。次に、先行研究で行われたことが SRM の全てを網羅するわけではないということから、これまでに聴取実験が行われていない刺激を用いて本モデルで SRM のシミュレーションを行い、シミュレーション結果が現象を模擬するような結果を示すのか確認するため、モデルの評価実験を行う。それにより、本モデルが実験が行われていない未知の事柄についても示すことができるモデルであることを確認する。

5.2 モデルのパラメータに関する検討

モデルのなかにある自由度のあるパラメータは位相同期指数を表す α と時間誤差とレベル誤差の値を示す σ_e と σ_δ である。まず、位相同期指数についてであるが、Blackburn と Sachs らの近似曲線はネコの位相同期指数を元に提案されたものであり、CF 2.5 kHz で位相同期指数 0.5 と高い値を示し、ヒトの聴覚とは多少の開きがあると考えられる。伊藤らが行った、特徴周波数 - 位相同期指数の関係についての検討では 2500 Hz 付近で位相同期指数 0.1 程度を示すとされている [29]。現状、ヒトの位相同期指数についての生理学的なデータを手に入れることは困難でありネコの位相同期指数を参考として、それをもう少しヒトの値に近くなるように 1.5 kHz ~ 2.0 kHz あたりの傾きを急峻にし、モデルの調整を行っている。今後ヒトの聴覚系の生理学的な研究が進みヒトの聴神経の位相同期指数が明らかとされれば、1.5 kHz ~ 2.0 kHz の出力をよりヒトの実験結果に近い値に改善することができると期待できる。

次に、時間誤差とレベル誤差を示すパラメータについてであるが、このパラメータが存在する目的はヒトが生じるマスキング解除量にある程度の限界を持たせるためである。 σ_e と σ_δ が存在しなければ S_0N_0 条件のとき、両耳に入力されるマスカーは完全に一致しているため、両耳のマスキングを打ち消したときの出力は 0 となる。このとき、モデルの出力するマスキング解除量は無限の値を示す。もちろん、ヒトのマスキング閾値が無限に改善されるわけではない。そのため、ヒトが生じるマスキング解除量にある程度の限界を持た

せる必要がある。そこで、 σ_e と σ_δ というパラメータで両耳のマスカーの一致能力の不完全さを表している。すなわち、両耳のマスカーの一致能力にある程度の限界を持たせておくことでヒトが生じ得る範囲のマスクング解除量をモデルが出力するようにしている。

5.3 モデルの評価実験

5.3.1 目的

方向性を手掛かりとした方向定位能力は目的音の周波数が 1.5 kHz 以下の場合 ITD に強く影響され、周波数が 2.0 kHz 以上の場合 ILD に強く影響される。先行研究によりマスクング解除現象も同様の傾向が見られることが明らかとされた。しかし、そのどちらも上手く手掛かりとして利用できないとされる 1.5 kHz ～ 2.0 kHz の周波数の信号で、かつ ITD も ILD も存在する音を用いて行われた聴取実験は報告されていない。モデルはどんな音の入力に対しても出力を返すが、モデルの出力と比較するデータを持たず有効性の確認が行えない。そこで、モデルの出力と比較できるようなもので、今まで実験が行われていないような条件の信号音を用いてモデルの評価実験を行うことにする。その聴取実験の結果が、本モデルの出力と近い値、同様の特徴を示すならば、方向性を手掛かりとして生じるマスクング解除現象を模擬するモデルとして有用性の高いモデルであると考えることができる。

5.3.2 実験刺激

目的信号は 1.8 kHz の周波数の純音とし、雑音信号には走行雑音を用いた。また、報知音の断続パターンは 1 秒あたり、ON = 0.1 s、OFF = 0.05 s とする、ON、OFF を交互に繰り返すパターンとした。雑音信号である走行雑音は走行速度 60 km/h の窓を開けた環境下において、自動車内でデジタル録音された走行雑音を利用した。雑音は全部で 4 パターンの刺激を用意し、雑音はそこからランダムに選択される。提示される刺激は 2 s の雑音信号中に 1 s の目的信号が存在する音で、目的信号が呈示開始位置は、聴取者に予測されないように雑音信号が提示してから 1 s までの間のランダムな時間に目的信号の呈示が開始する構成とした。雑音信号の音圧は実験開始前に 65 dB にキャリブレーションを行う。これらの実験刺激のサンプリング周波数は 48 kHz とした。

5.3.3 実験方法

実験は防音室内で行い、実験刺激はスピーカーを利用して呈示を行った。スピーカーの配置は、聴取者を中心として半径 1.2 m の円周上に正面 0° から右側面 90° まで 30° 刻みで配置した。マスクング閾値の測定法には極限法を利用した。この方法は聴取者が目的信号を確実に検知できる高い音圧から低い音圧へ変化させる下降系列と、目的信号を確

実に検知できない低い音圧から高い音圧へ変化させる上昇系列の二つの信号呈示順序を繰り返す。聴取者には呈示された実験刺激に対して、目的信号が「検知できる」か「検知できない」を判断してもらう。下降系列ならば「検知できない」が3回連続して選択された場合、上昇系列に移行し、上昇系列では「検知できる」が3回連続して選択された場合、2回目の下降系列に移行する。下降系列、下降系列を交互に3回繰り返す。二つの系列で測定されたそれぞれの閾値の差が2 dB以内のときのデータを有効なものとし、測定値の平均値をマスキング閾値とした。本実験には男性大学院生5名が参加した。全ての聴取者は正常な聴力(0.125 kHz ~ 8 kHz のオクターブ周波数内で 15dB HL 以下)を有し、他の聴取実験に参加した経験がある。

5.3.4 結果と考察

実験を行った結果を図5.1に示す。結果を見るとモデルの出力に比べ、聴取実験のほうがマスキング解除量が小さくなっているが、ピークの位置についてはモデルと似た傾向を示している。この結果を見ると、モデルの出力するマスキング解除量が全角度においても少し小さな出力であれば実験結果と近い値を示すと考えられる。モデルと、実験結果で多少開きが出てしまった原因について考察する。モデル側の問題として考えられることとして、位相同期指数についての問題が大きいと考える。モデルで利用している位相同期指数がどうしても 1.5 kHz ~ 2.0 kHz あたりでヒトの位相同期指数と開きが生じてしまうことが原因の一つとして考えられる。また、黒田らの実験が HRTF を畳み込んだ音を用いて、ヘッドフォン呈示で行われたのに対し、本実験ではスピーカーを用いて呈示を行った。この実験は防音室で行ったが、無響室で行ったわけではなく、壁からの反射音が存在し、それが方向定位能力を低下させる原因となったことが考えられる。

5.4 おわりに

ここでは、提案モデルについての検討、評価実験を行った。その結果、モデルのもつパラメータから見た結果と、聴取実験を行った結果を合わせて 1.5 kHz ~ 2.0 kHz あたりで多少ながら模擬精度が低下することが確認された。この原因はモデルの持つ位相同期指数が 1.5 kHz ~ 2.0 kHz の刺激が入力されたときに、ヒトの位相同期指数と開きがあること、また、位相同期指数から ITD と ILD それぞれの影響を α と $1 - \alpha$ という ITD の有効性に伴い、ILD の有効性が低下するという単純な重み付けで表現していることが原因と考えられる。1.5 kHz ~ 2.0 kHz では音源定位に ITD も ILD も有効には働かないということが Mills らの MAA の結果から読み取れる [30]。SRM も同様に 1.5 kHz ~ 2.0 kHz では ITD も ILD も手掛かりとして有効に利用できずマスキング解除量が低下する可以考虑することができる。しかし、模擬精度は低下するながらも参考となるようなマスキング解除量の算出が行えること、モデルの出力が示す特徴は聴取実験の結果と類似することから、本モデルは方向性マスキング解除のモデルであると解釈できる。

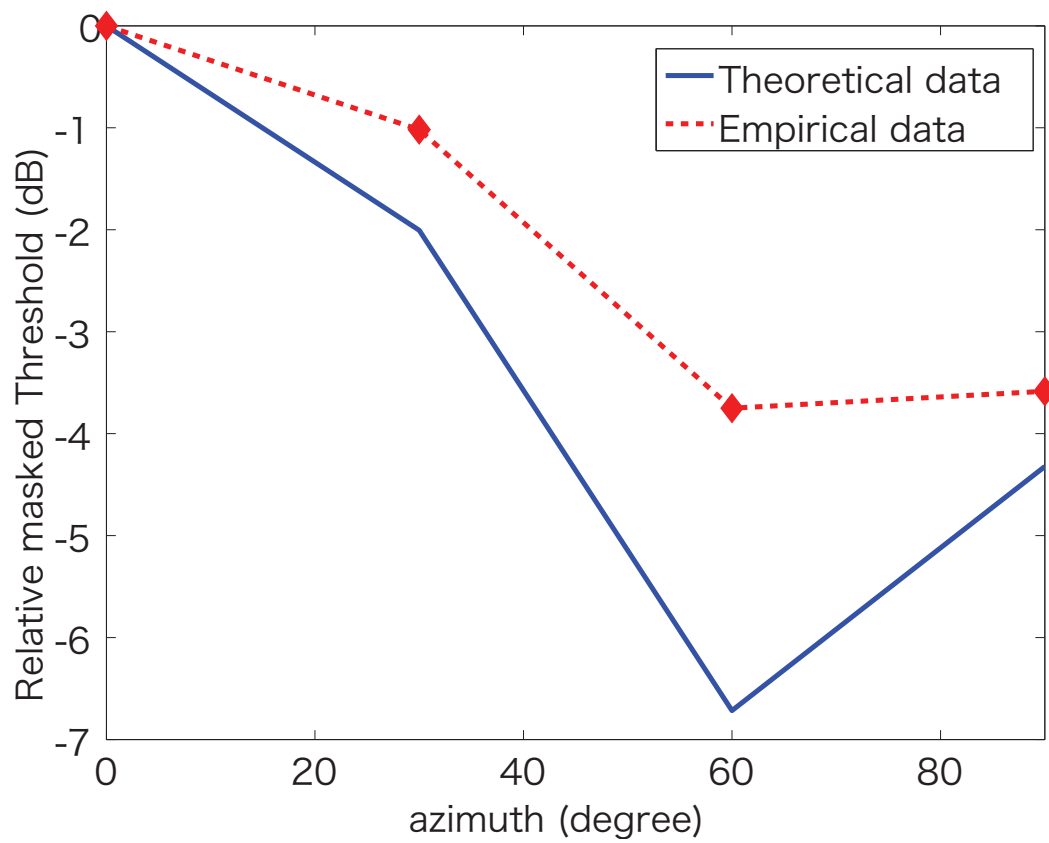


図 5.1: 目的音に 1.8 kHz の純音を用いてシミュレーションを行った結果. マーカーなしの青のラインが提案モデルを用いて算出したマスキング解除量. マーカー付きの赤のラインが聴取実験の結果

第6章 結論

6.1 本論文で達成されたこと

本研究では、ヒトの聴覚中枢においてどのような処理が行われて SRM が生じるかを 2 つの両耳間時間差 ITD と ILD を用いて説明する両耳聴モデルの提案を目的とした。

まず、既存の両耳聴モデルについて第 2 章で調査と検討を行った。近年の両耳聴モデルに関する研究では、ILD の検出がどのように行われており、ILD を手掛かりにどのように方向低位が生じるのかといったことを説明するようなモデルを構築することに焦点が当てられている。Jeffress の提案した時間差検出モデルを拡張する形で様々なモデルが提案されてきたが、ILD の影響を完全に示すモデルは現れていない。また、両耳聴モデルの中で、本研究において特に重要であると考えられる EC モデルについての検討を行った。このモデルは音の方向性を利用してマスキング除去が行われるという仮定によりマスキング解除を説明するモデルである。EC モデルを用いて ITD のみ持つ音、ITD と ILD の両方を持つ音を入力として用い、マスキング解除量算出を行った結果、EC モデルでは ITD を手掛かりとして生じるマスキング解除については説明可能であったが、ILD を手掛かりとして生じるマスキング解除については SRM を模擬する結果を示さないことが明らかとなった。

続く第 3 章では、EC モデルの考え方を取り入れ新しい方向性マスキング解除のモデルを提案した。本モデルはヒトの聴覚中枢系に ITD と ILD を手掛かりとして用いて、両耳のマスキングを一致させ打ち消すメカニズムが存在するという、方向性を利用したマスキング除去の考え方に基づき SRM についてのモデルを構築した。本モデルの特徴として ITD と ILD の影響を独立して考え、最終的に 2 つを結合するという処理の流れをしており、ITD と ILD それぞれだけが強く影響する場合も、両方が寄与する場合も議論できるモデルとなっている。

第 4 章では、新しく提案したマスキング解除のモデルが既存の知られている SRM 現象を正しく模擬できるのかを確認した。提案したモデルを用いて黒田らの聴取実験を参考に、SRM のシミュレーションを行った。その結果、本モデルの出力は ITD と ILD の両方が存在している信号を入力とした場合でも、聴取実験の結果を精度よく模擬する結果となった。それにより、本モデルは ITD と ILD により生じる SRM についてのモデルと解釈することができる。

最後に、第 5 章で、モデルの評価を行った。モデルの評価はモデルが持つパラメータについての検討と、今まで行われていないような SRM を正しく模擬できるかを評価実験

で確認した。その結果、本モデルでは入力信号の周波数が 1.5 kHz ～ 2.0 kHz において模擬精度が多少低下することが確認された。しかし、模擬精度は低下するながらも、その結果が示す特徴は聴取実験の結果と類似することから、本モデルを SRM のモデルとみなせる結果が得られたと考える。

以上より、ITD と ILD を手掛かりとして生じるマスキング解除がどのようなメカニズムで生じるか説明する両耳聴モデルの提案を行った。本研究が、今後の両耳聴モデルの発展や、知覚されやすい目的音の呈示法の確立に役立つことを願って已まない。

6.2 今後の展望

本研究では方向性マスキング解除のモデルを提案した。しかし、本モデルをより精巧で適応範囲の広い使いやすいものへと改善するために、いくつか今後の展望を示す

1. 位相同期指数の改善
2. ITD と ILD のデータベースの利用

の 2 点について考える。

1. これは第 5 章も述べたが、本モデルで利用している位相同期指数はネコの位相同期指数を元に提案されたものでありヒトのものとは多少開きがある。生理学的な面から研究が進み、ヒトの聴神経の位相同期指数が明らかとされたらより正確な出力の得られるモデルに改善することができると考える。
2. 各周波数を入力としたと各呈示角度に対応する ITD と ILD のデータベースをモデルに組み込むことを考える。データベースとして持つ ITD と ILD を利用することで、例えば、実環境で録音した音を入力としたときでもマスキング解除量の算出を可能にするモデルへ改善することができると考えている。このモデルにおける、実環境で録音した音を用いる場合の問題点は ITD と ILD が雑音、残響などにより正しく算出されないことである。ITD と ILD のデータベースを持つことで、あらかじめ持っているデータベースから方向性情報を利用でき、それにより生じるマスキング解除量の算出が可能となる。

付録

目的

本研究でモデル化の対象としたのは音源の水平方向の違いにより生じるマスキング解除現象であり、水平方向の方向定位に用いられるとされる ITD と ILD に着目して SRM を説明している。しかし、上下方向の違いによっても何らかのマスキング解除が生じることが明らかとされ、それがどの程度生じるのか、何を要因として生じるのかが明らかとなれば、モデルを水平方向だけではなく空間的な入力に対しても利用出来るモデルへと拡張することが期待できる。しかし、上下方向に雑音源を配置したときにマスキング解除が生じるかの検討は行われていない。そこで、音源の垂直方向の違いにより生じるマスキング解除についての研究の基礎的な調査として、雑音源を上下方向に配置した場合にマスキング解除が生じるかどうかを調べるため、聴取実験を行う。この調査が、水平方向についてのみではなく音源の垂直方向の違いがマスキング解除に与える影響についての議論に役立つことを期待する。

実験刺激

目的信号はパルス列信号を用い、雑音信号には走行雑音を用いた。目的信号であるパルス列信号は Saberi ら [2] の実験と同様に 1s あたり 100 個の $65 \mu\text{s}$ 幅の矩形パルスで構成した。雑音信号である走行雑音は走行速度 60 km/h の窓を開けた環境下において、自動車内でデジタル録音された走行雑音を利用した。雑音は全部で 4 パターンの刺激を用意し、呈示される雑音はそこからランダムに選択される。提示される刺激は 2s の雑音信号中に 1s の目的信号が存在する音で、目的信号が呈示開始位置は、聴取者に予測されないように雑音信号が提示してから 1s までの間のランダムな時間に目的信号の呈示が開始する構成とした。雑音信号の音圧は実験開始前に 65 dB にキャリブレーションを行う。これらの実験刺激のサンプリング周波数は 48 kHz とした。

実験方法

実験は防音室内で行い、実験刺激はスピーカーを利用して呈示を行った。スピーカーの配置は、聴取者を中心として半径 1.2 m の円周上に正面 0° - 垂直方向 0° 、水平方向 45° - 垂直方向 0° とし、水平方向 45° の上下にも、水平方向 45° - 垂直方向 40° 、水平

方向 45° - 垂直方向 -40° の位置に雑音源を配置した。マスキング閾値の測定法には、モデルの評価実験の時と同様に極限法を利用した。本実験には男性大学院生 3 名が参加した。全ての聴取者は正常な聴力 ($0.125\text{ kHz} \sim 8\text{ kHz}$ のオクターブ周波数内で 15 dB HL 以下) を有し、他の聴取実験に参加した経験がある。

結果と考察

実験結果を図 6.1 に示す。結果から、音源が上下方向に存在している場合に $1, 2\text{ dB}$ 程度大きなマスキング解除が生じていることがわかる。これは、垂直方向の違いによってもマスキング解除が生じることを示唆する結果である。しかし、今回の実験は水平方向 45° - 垂直方向 40° 、水平方向 45° - 垂直方向 -40° という単純に水平方向 45° の上下に配置しての実験であり、球面上に配置したものではないためスピーカーと聴取者までの距離が異なっており、それにより違いが生じた可能性が排除できない。また、上下で 1 dB 程度の差が生じていることも興味深い結果に思える。聴取者 3 人の結果がすべて水平方向 45° - 垂直方向 0° 、水平方向 45° - 垂直方向 -40° 、水平方向 45° - 垂直方向 40° の順に大きくなる傾向を示しており上下で知覚に差が生じる可能性も示唆された。しかし、これに関しても、水平方向 45° - 垂直方向 -40° では非常に床に近い位置にスピーカーがあるため、床からの反射音の影響を強く受け、マスキング解除量が上方向に比べて小さいと解釈することもできる。今回の実験で垂直方向の違いによってもマスキング解除が生じる可能性があるということが示唆されたが、より詳細に現象を調べていくためには、なるべく実環境下で生じる要因を排除しての検討を行う必要がある。無響室を用いての聴取実験や HRTF を畳み込んだ音を用いてヘッドフォン呈示する方法での実験などを行うことで、より詳しく現象を調べることが求められる。また、音の上下弁別に利用されるとされるスペクトラルキューを利用するような両耳聴モデルの提案も今後の課題としてあるのではないかと考える。

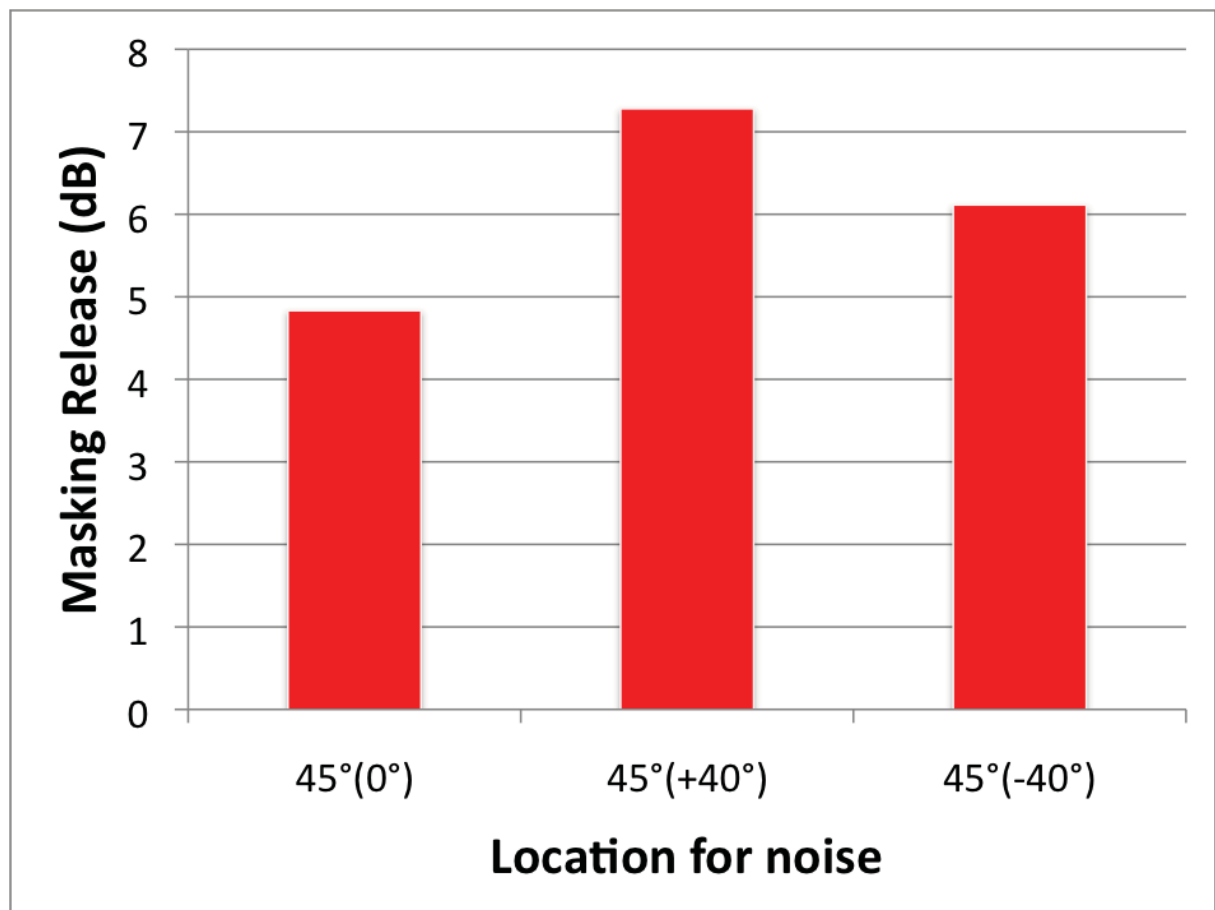


図 6.1: 水平方向 45° とその上下 +40° と -40° に雑音源を配置した場合に生じたマスキング解除量

謝辞

本研究を行うにあたり，終始ご指導を賜りました北陸先端科学技術大学院大学 情報科学研究科 赤木正人 教授に深く感謝いたします。また，折に触れてご指導いただきました，北陸先端科学技術大学院大学 情報科学研究科 鵜木祐史 准教授，宮内良太 助教に深く感謝いたします。加えて，本研究を始めるに当たり貴重な助言を賜りました，中国科学院声学研究所 李軍鋒 教授に深く感謝いたします。さらに本研究を遂行していく中で，熱心な議論と多面に渡る協力を賜った北陸先端科学技術大学院大学 情報科学研究科 党建武 教授，徳田功 准教授，末光厚夫 助教，川本真一 助教に深く感謝いたします。

その他，日頃の研生活において，議論ならびに助言をくださった赤木，鵜木研究室の皆様および諸先輩方に厚くお礼申し上げます。

最後に，本研究を行う機会を与えてくださり，また研究活動を暖かく見守ってくれた多くの皆様に心より感謝いたします。

参考文献

- [1] M. Ebata, T. Sone, and T. Nimura, "Improvement of hearing ability by directional information," J. Acoust. Soc. Am, Vol. 43, pp. 289-297, 1968.
- [2] K. Saberi, L. Dostal, T. Sadralodabai, V. Bull, D.R. Perrott "Free-field release from masking," J. Acoust. Soc. Am, Vol. 90, no. 3, pp. 1355-1370, 1991.
- [3] C. Lane, N. Kopco, B. Delgutte, B. G. Shinn-Cunningham, and H. S. Colburn "A cat's cocktail party: Psychophysical, neurophysiological, and computational studies of spatial release from masking," Proc. ISH 2003, pp 341-347, 2003 J. Acoust. Soc. Am, Vol. 90, no. 3, pp. 1355-1370, 1991.
- [4] J. Nakanishi, M. Unoki, M. Akagi "Effect of ITD and Component Frequencies on Perception of Alarm Signals in Noisy Environments," Jounal of Signal Processing, Vol.10-4, pp.231-234, 2006
- [5] H. Uchiyama, M. Unoki, M. Akagi "Improvement in detectability of alarm signal in noisy environments by utilizing spatial cues," Proc.WASPAA2007, New Paltz, NY, pp.74-77, 2007
- [6] N. Kuroda, L. Junfeng, Y. Iwaya, M. Unoki, M. Akagi "Effects from spatial cues on detectability of alarm signals in noisy environments," IWPASH, pp. 1-4, Nov.11-13, 2009
- [7] N. I. Durlach, "Equalization and cancellation theory of binaural masking level differences," J. Acoust. Soc. Am, Vol. 35, pp. 1206-1218, 1963.
- [8] L. A. Jeffress, "A place theory of sound localization," J.Camp.Psych., Vol. 41, pp. 35-39, 1948.
- [9] H. S. Colburn, "Theory of binaural interaction based on auditory - nerve data. I. general strategy and preliminary results on interaural discrimination," J. Acoust. Soc. Am., Vol. 54, pp 1458-1470, 1973.
- [10] H. S. Colburn, "Theory of binaural interaction based on auditory - nerve data. II. detection of tones in noise," J. Acoust. Soc. Am., Vol. 61, pp 525-533, 1977.

- [11] H. S. Colburn and J. S. Latimer, "Theory of binaural interaction based on auditory-nerve data. III. Joint dependence on interaural time and amplitude differences in discrimination and detection," J. Acoust. Soc. Am., Vol. 64, pp 95-106, 1978.
- [12] R. M. Stern and H. S. Colburn, "Theory of binaural interaction based in auditory-nerve data. IV. A model for subjective lateral position," J. Acoust. Soc. Am., Vol. 64, pp 127-140, 1978.
- [13] R. M. Stern, H. S. Colburn. "Lateral-position based models of interaural discrimination," J.Acoust.Soc.Am., Vol. 77, pp 753-755, 1985.
- [14] J. Blauert and W. Cobben, "Some considerations of binaural cross-correlation analysis," Acustica, 39:96-103, 1978.
- [15] W. Lindemann, "Extension of a binaural cross-correlation model by contralateral inhibition. I. simulation of lateralization for stationary signals," J. Acoust. Soc. Am., Vol. 80, pp 1608-1622, 1986.
- [16] W. Lindemann, "Extension of a binaural cross-correlation model by contralateral inhibition. II. the law of the first wavefront," J. Acoust. Soc. Am., Vol. 80, pp 1623-1630, 1986.
- [17] W. Gaik, "Combined evaluation of interaural time and intensity differences: Psychoacoustic results and computer modeling," J. Acoust. Soc. Am., Vol. 94, pp 98-110, 1993.
- [18] J. Breebaart, S. van de Par, and A. Kohlrausch. "Binaural Processing model based on contralateral inhibition. I. Model strcture," J.Acoust.Soc.Am., Vol. 110:1074-1088, 2001.
- [19] J. Breebaart, S. van de Par, and A. Kohlrausch, "Binaural processing model based on contralateral inhibition. II. Dependence on spectral parameters," J.Acoust.Soc.Am., Vol. 110, pp 1089-1103, 2001.
- [20] J. Breebaart, S. van de Par, and A. Kohlrausch, "Binaural processing model based on contralateral inhibition. III. Dependence on temporal parameters," J.Acoust.Soc.Am., Vol. 110, pp 1125-1117, 2001.
- [21] N. Nakashima, Y. Chisaki, T. Usagawa and M. Ebata, "Frequency domain binaural model based on interaural phase and level differences," Acoustical Science and Technology, 24, 4, pp. 172-178, 2003

- [22] Kuhn, G. F., “Physical Acoustics and Measurements Pertaining to Directional Hearing,” in Yost, W. A. and Gourevitch, G. (eds), *Directional Hearing*, Springer-Verlag, pp. 3-25, 1987.
- [23] Brian C. J. Moore, *An Introduction to the Psychology of Hearing* Fifth Edition, Chap. 7, Academic Press, London, 2003
- [24] 赤木 正人, “聴覚フィルタとそのモデル,” 電子情報通信学会誌, Vol. 77, No. 9, pp 948-956, 1994-9.
- [25] B. R. Glasberg, and B. C. J. Moore, “Derivation of auditory filter shapes from notched-noise data,” *Hearing Research*, 47, pp 103-138, 1990.
- [26] Don H. Johnson, “The relationship between spike rate and synchrony in responses of auditory-nerve fibers to single tones,” *J. Acoust. Soc. Am.* Vol. 68, Issue 4, pp. 1115-1122, 1980
- [27] C. C. Blackburn and M. B. Sachs, “Classification of unit types in the anteroventral cochlear nucleus: PST histograms and regularity analysis,” *Journal of Neurophysiology*, Vol 62, Issue 6 1303-1329
- [28] 柏野 牧夫, 平原 達也, “文音声における同時複数話者の人数判断,” 日本音響学会講演論文集, 2-8-12, 369-370, 1997
- [29] 伊藤 一仁, “初期聴覚中枢系の神経細胞レベルでの時間情報処理機構に関する研究,” 北陸先端科学技術大学院大学博士論文, 2003
- [30] A. W. Mills, “On the minimum audible angle,” *J. Acoust. Soc. Am.* Vol. 30, pp237-246, 1958

本研究に関する研究業績

学会発表リスト

Shinya Mizukawa and Masato Akagi, “A binaural model accounting for spatial masking release,” Proc. NCSP11, March 2011

水川慎也，赤木正人，“方向性マスキング解除を説明する両耳聴モデルの提案,” 日本音響学会 2010 年 春季研究発表会 講演論文集, 2-P-56(b), March 2011