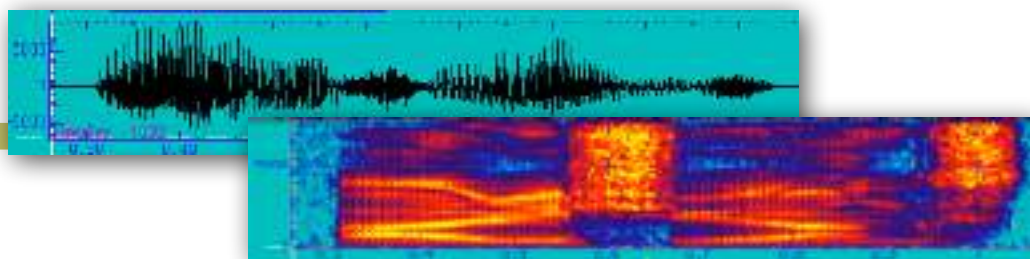


音響音声学

(Topics in Acoustic Phonetics)



峯松 信明

工学系研究科電気系工学専攻

様々な情報＝四要素の組み合わせ

音の高さ

- 高い音, 低い音

音の大きさ

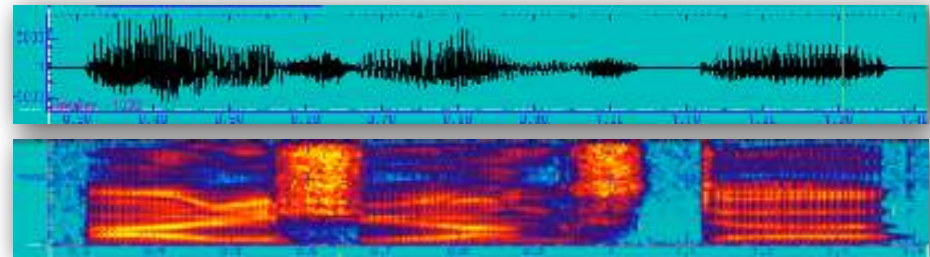
- 大きい音, 小さい音

音の長さ

- 長い音, 短い音

音の音色 (声色)

- 太い声, 細い声, 黄色い声, 甘い声, 渋い声, 色っぽい声 . . .
- 高さ・大きさ・長さが同じ2音を「違う音」と認識した場合, その2音は**音色**が異なる。
- 「あ」と「い」 「あ」と「あ」



- 1) 高さ, 2) 大きさ
3) 長さ, 4) 音色

ある情報

峯松研から文系学生への提案

テキスト読み上げ技術 → 日本語韻律教育支援

The screenshot shows the OJAD (Online Japanese Accent Dictionary) website. The header features the OJAD logo and navigation buttons for Twitter, YouTube, and Facebook. A sidebar on the left lists 'OJAD 4機能' (4 features) and '教科書別' (by textbook). The main content area includes a 'OJADとは？' (What is OJAD?) section, a search bar, and a '教科書別データ' (Textbook data) section listing 50 textbooks.

OJADとは？

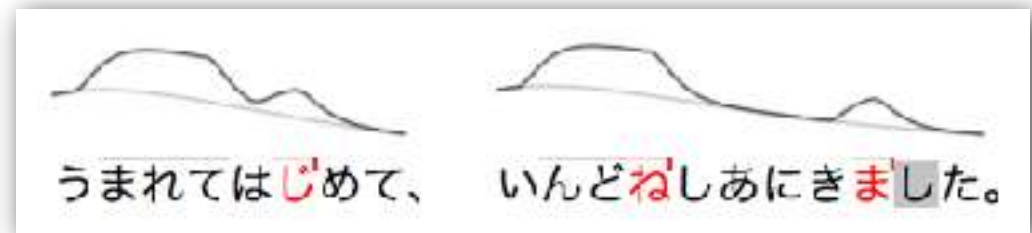
OJADは日本語教師・学習者のためのオンライン日本語アクセント辞書です。約9,000の名詞の東京方言アクセントを選べることができます。約3,500の動詞（動詞、い形辞、な形辞）について、基本12活用、約42,300のアクセントを選べることができます（男女1名ずつの音声サンプルも無料）。それ以外の約300の使役語表現に対しても、アクセントを選べることができます。更に任意の肯定文・疑問文に対して、特定の語に強いフォーカスを入れずに読み上げた際に予想されるピッチパターンを表示することもできます。学習者の自主学習、日本語教室や日本語教師養成講座での韻律学習にご使下さい。日本語母語話者の東京方言参照にもお使いいただけます。

利用される場合は、まず「ご利用にあたって」をご覧ください。

教科書別データ

みんなの日本語 初級【スリーエーネットワーク】

第01課	第02課	第03課	第04課	第05課	第06課	第07課	第08課	第09課	第10課
第11課	第12課	第13課	第14課	第15課	第16課	第17課	第18課	第19課	第20課
第21課	第22課	第23課	第24課	第25課	第26課	第27課	第28課	第29課	第30課
第31課	第32課	第33課	第34課	第35課	第36課	第37課	第38課	第39課	第40課
第41課	第42課	第43課	第44課	第45課	第46課	第47課	第48課	第49課	第50課



峯松研から文系学生への提案

“読み上げる”支援から“聞く”支援へ

- リスニング教材の多くは無雑音のクリーン音声
- 現実世界の音声は、雑音や歪みが混入したいい音声
- 聞き取り能力の頑健性向上に対する技術支援は可能か？

様々な外的要因による音声の歪み

goo.gl/jJdtuK

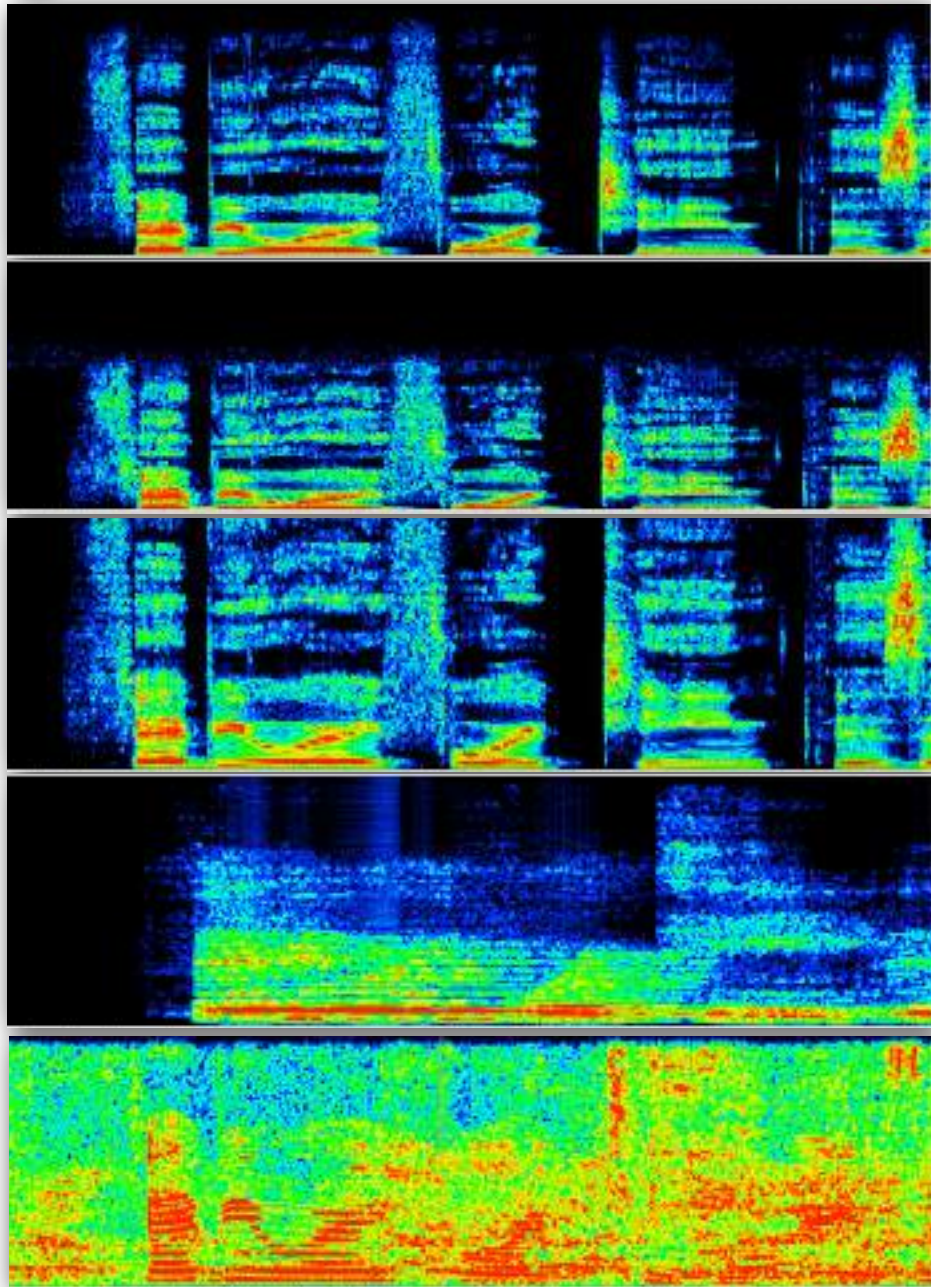
- エフェクタをかけたような巨人・小人の音声 [体格]
- ホールの反響（リバーブ）がかかった音声 [反響]
- 集会・駅・空港などの様々な背景雑音 [雑音]
- 電話、タクシー／警察無線、などのチャネル歪み [伝送]

なぜ母語話者は聞き取れて、学習者には難しいのか？

- 単なる慣れの問題？適切な acoustic cue が分かってない？
- 外的要因の変化の一方で、**共通項**は存在するのか？それは何？
 - 体格、反響、雑音、伝送の違いに不変な共通項なんてあるのか？

峯松研から文系学生への提案

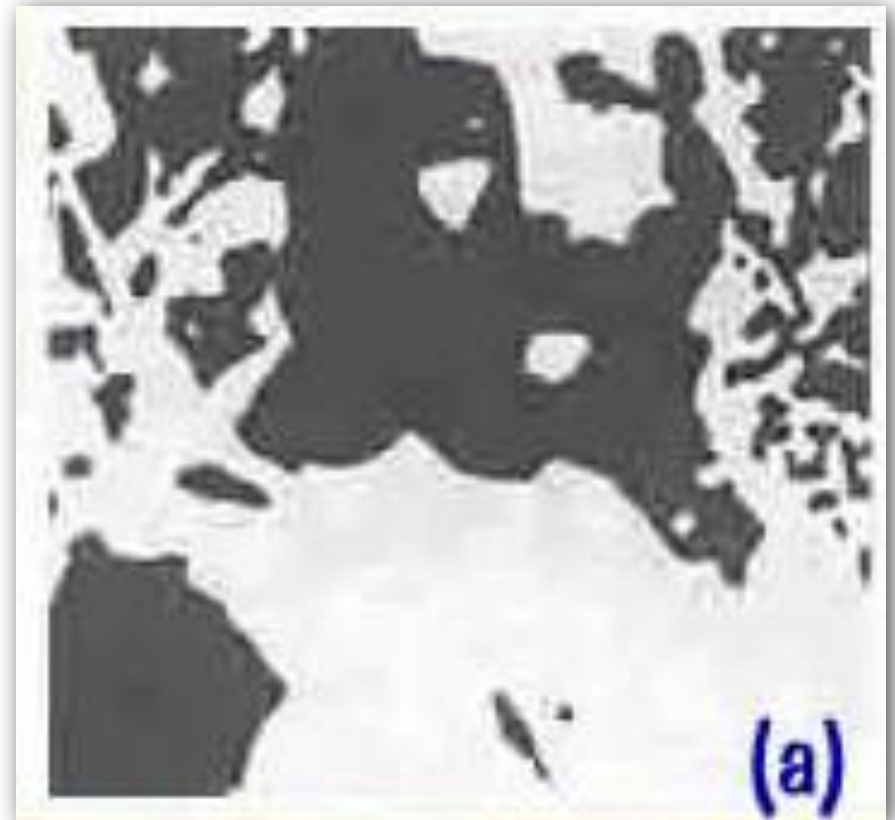
何が共通しているのか？共通項なんてあるのか？



峯松研から文系学生への提案

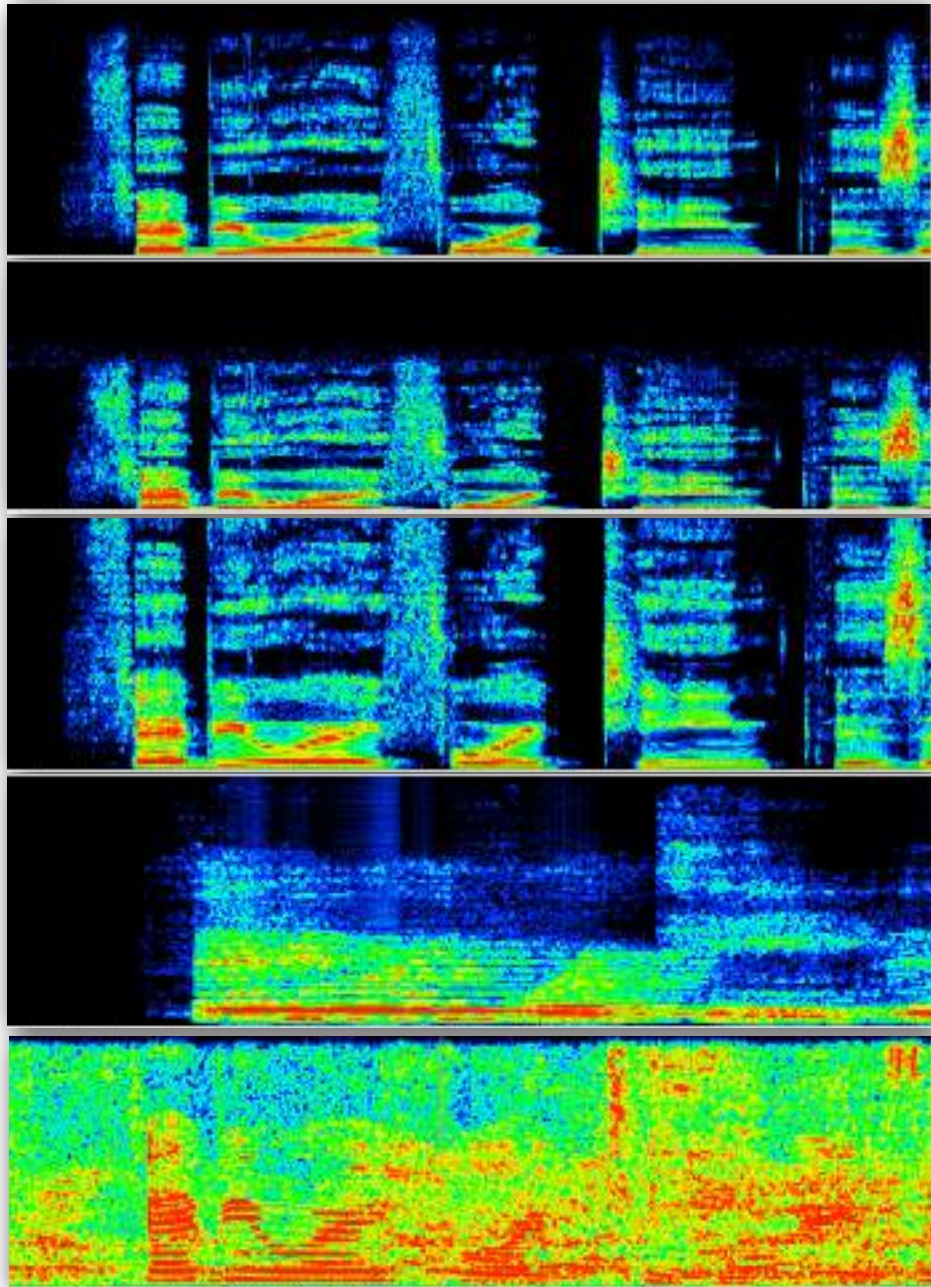
スペクトルダイナミクスの中にあるパターンを見つける

- 白い斑点パターンの中にXXがいる, YYがいる!
- ゲシュタルト知覚の例
 - 背景のパターンが変化すれば, 対象物の知覚は極めて容易に
 - 共通項は動かさず, 外的要因を動かすと, パターンを見つけやすい。



峯松研から文系学生への提案

何が共通しているのか？共通項なんてあるのか？



共通項は fix して
外的要因のみを
あれこれ変えた
音声刺激は
簡単に作れます。

峯松研から文系学生への提案

雑音や歪みに強い耳も作れるかも？

- 色の違い：歪ませ方の違い
- 太さの違い：歪み量の違い
- どこまで歪みの量を落としていくと、聞き取れるのか？
 - Listening Challenge !!



様々な音響歪みのシミュレータは既に構築済み

- あとは、こういう技術の教育的な効果、価値を検証が必要
- 誰か、やりませんか？

マスキング (p.119)

マスキング量と周波数特性

● マスキング

- 信号音（目的音）＋雑音を呈示。当然，信号音が聞こえ難くなる。
- その信号音の聴覚閾値（それ以下だと聞こえない）が上る。
- 雑音：マスカー，信号音：マスキー

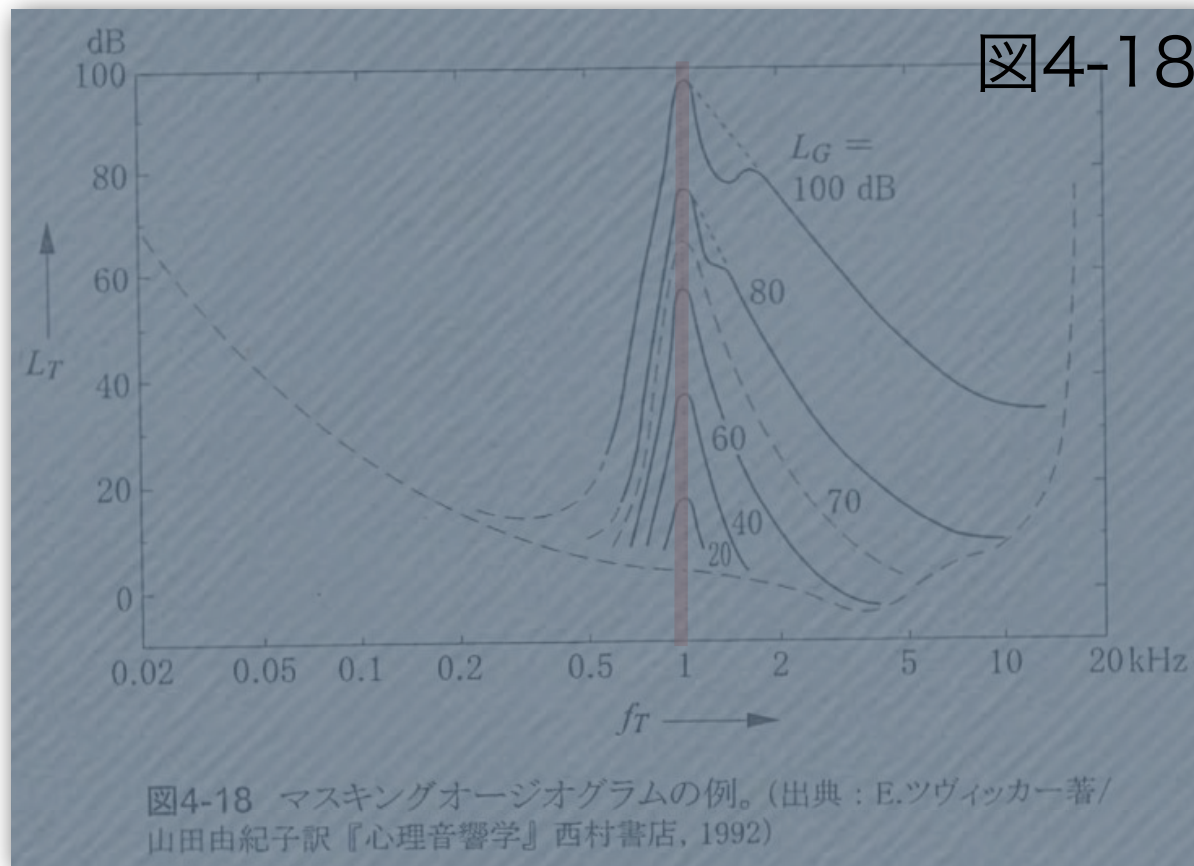
● 聴覚閾値の上昇量＝マスキング量

- あるマスカーのマスキング量が大きい → 聞こえ難くする効果大
- マスキング量は一般にマスキー（信号）の周波数特性により異なる。
- マスカーを固定して，マスキーの周波数と強さを色々変え，そのマスカーのマスキング量を測定する。
- 当然，マスキーと似た周波数特性（スペクトル）を持つマスカーで妨害されると，マスキーを知覚することが難しくなる。

マスキング (p.119)

マスキングオーゾグラム

- マスカー固定, マスキの周波数色々変えて測定
- マスカー無の閾値とマスカー有の閾値との差=マスキング量



- マスカー：中心1kHz, 帯域幅160 Hz, 6種類のレベルのいずれか
- 波線：マスカー無の閾値, そこからの差がマスキング量

代表的なマスキング

3種類のマスキング

- 純音, バンドノイズ, ホワイトノイズ (全て80dB SPL)
- バンドノイズ: ある帯域のみにエネルギーが (等しく) 分布する。
- ホワイトノイズ: 全帯域に等しくエネルギーが分布する。
- 純音: 400Hz, バンド: 中心410Hz, 幅90Hz,

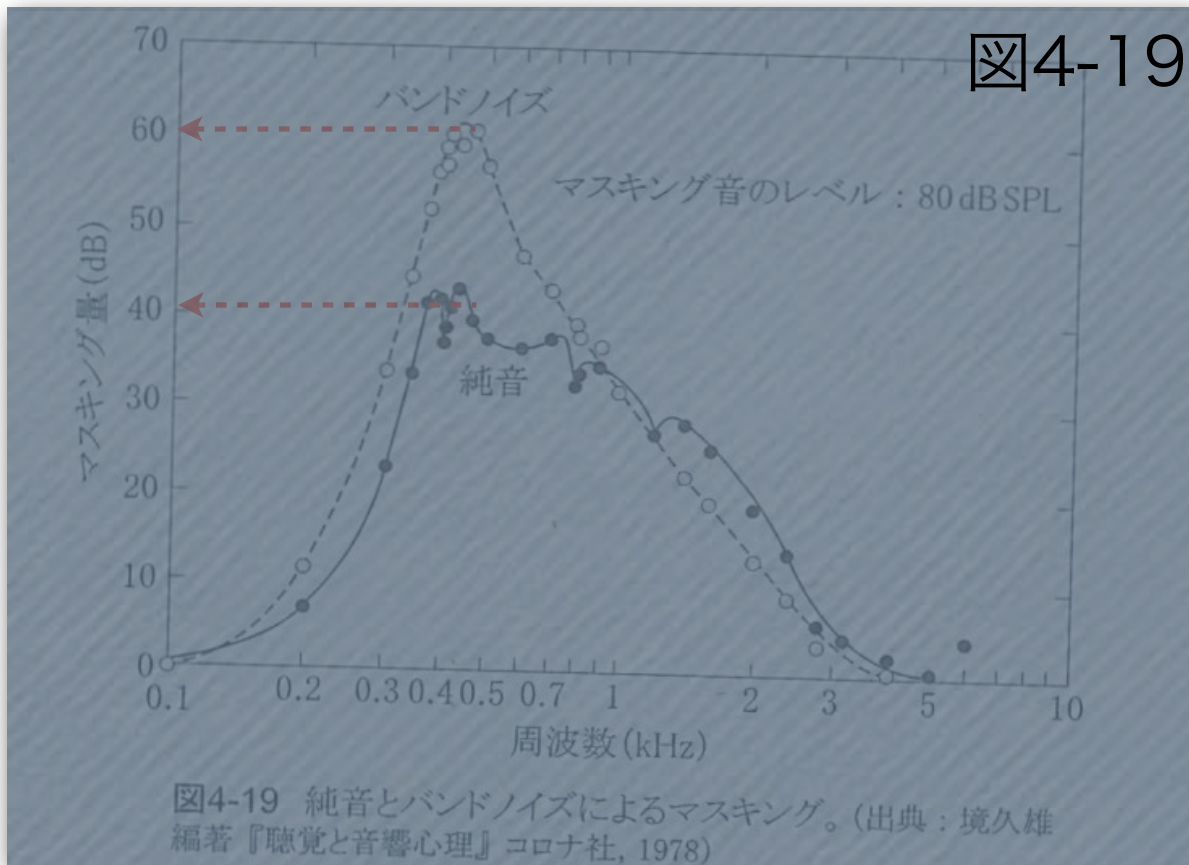


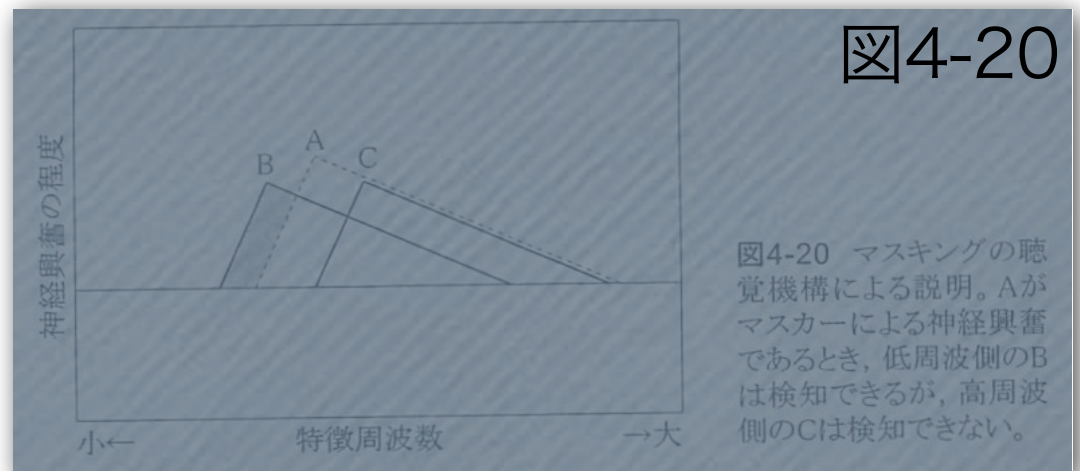
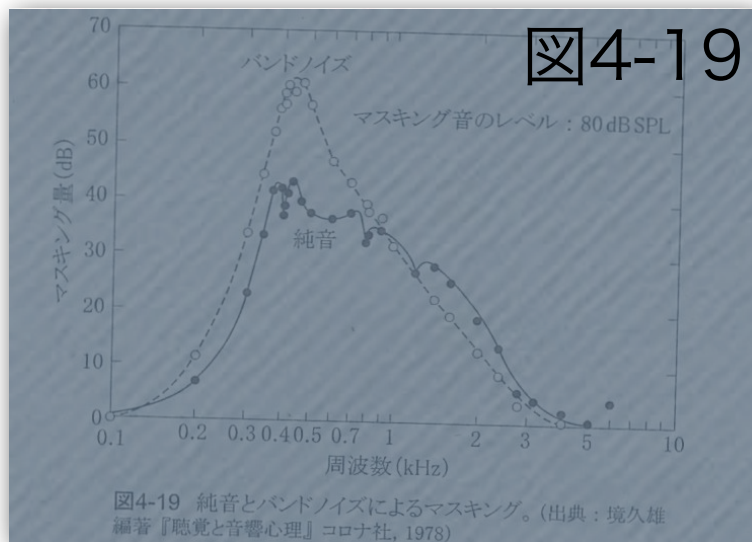
図4-19 純音とバンドノイズによるマスキング。(出典: 境久雄 編著『聴覚と音響心理』コロナ社, 1978)

- 純音マスキングの場合, 純音信号音との間でうなりが発生
- 純音マスキングしか聞こえて無い場合でも, うなりによって信号音の存在を認識可能
- 20dB = 音の強さ 100 倍

代表的なマスキャー

3種類のマスキャー

- 純音, バンドノイズ, ホホワイトノイズ (全て80dB SPL)
 - ホホワイトノイズの場合も, 純音ほどの効果となってしまう。
 - 純音とホホワイトノイズはスペクトル的には全く異なる。
 - あるマスキャーに対して, マスキング効果があるのは限られた帯域
 - ホホワイトノイズはマスキング効果が薄い成分 (帯域) を多く持つ
- マスキング量の非対称性
 - あるマスキャーに対するマスキング量は高周波数側の方が大きい



マスキングされた音を聞いてみる

マスキャーの違い

- マスキャー：純音を弱い音から順々に強くしていく。

- ケース 1 (65)

 - マスキャー：バンドノイズ

- ケース 2 (66)

 - マスキャー：ホワイトノイズ

マスキャーに対するマスキャーの周波数軸上の位置の違い

- マスキャー：バンドノイズ

- ケース 1 (67)

 - マスキャー：純音がマスキャーの位置から次第に高周波側にずれる

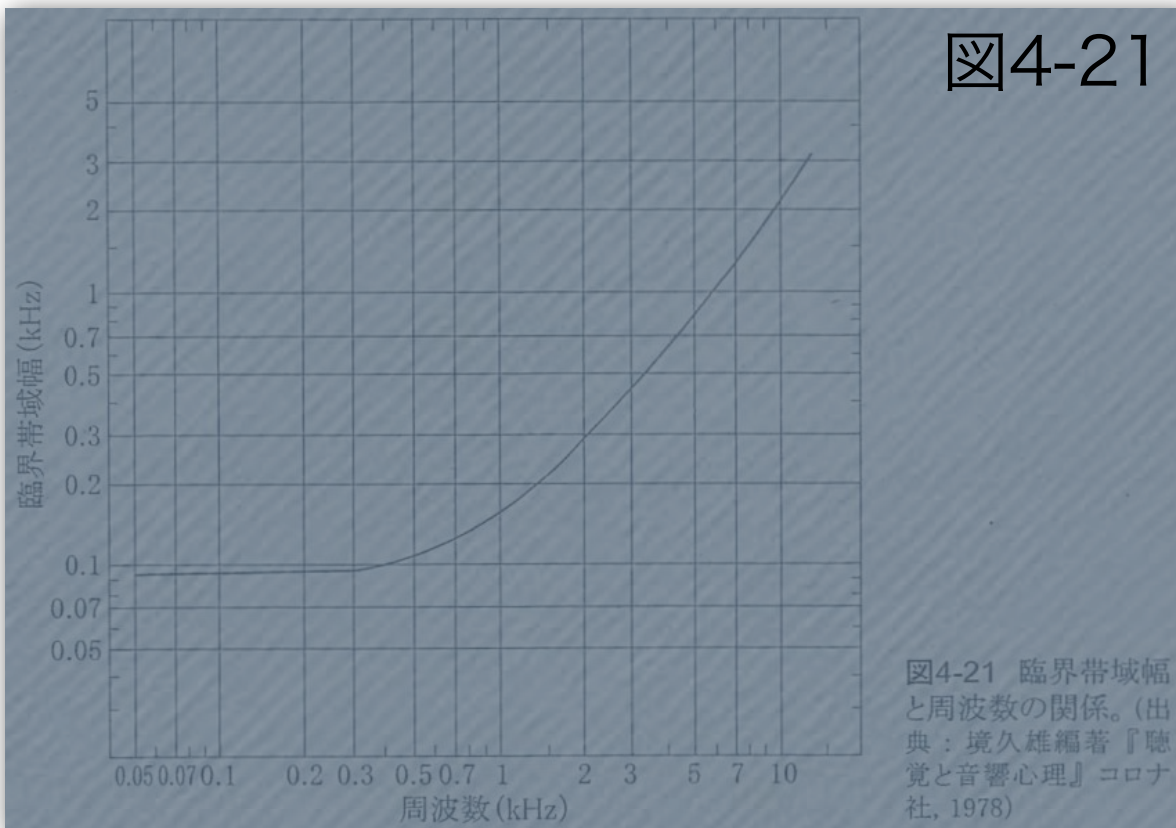
- ケース 2 (68)

 - マスキャー：純音がマスキャーの位置から次第に低周波側にずれる。

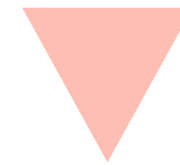
臨界帯域 (p.123)

基本的理解と一つの問題

- マスカーとマスキーが類似した周波数帯域にエネルギーを有する場合に、マスキング量は大きくなる。
- マスカー (バンドノイズ) の帯域幅とマスキング量の関係は？



- 帯域幅が狭いうちは、幅の増加 → マスキング量増加
- ある幅に到達すると、それ以上幅を増やしても効果は無し
- マスキング量が増えなくなる
帯域幅 = 臨界帯域

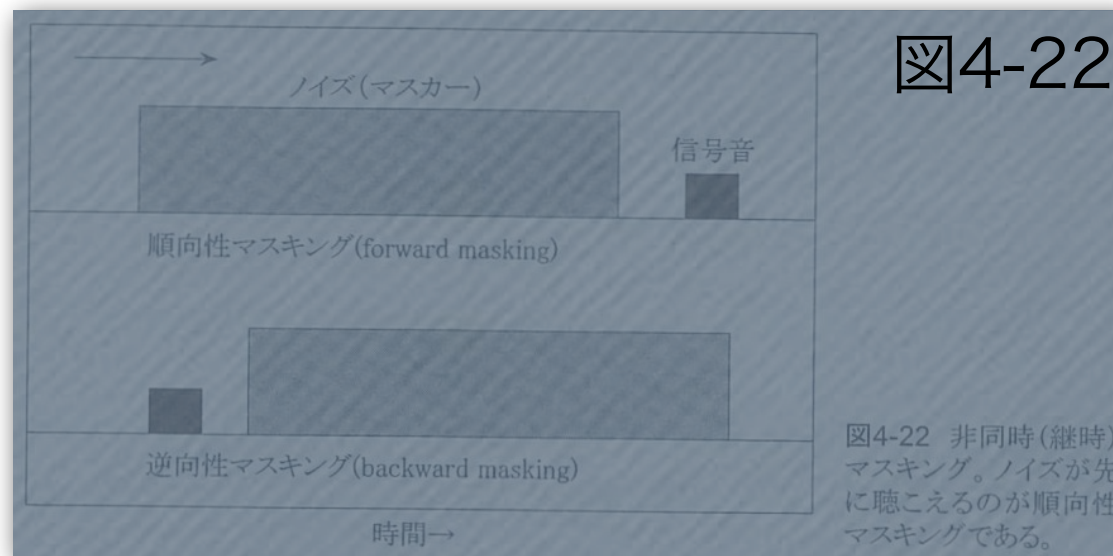


基底膜 = 有限個の多チャンネル

非同時マスキング

masker と maskee は時間的に同時である必要はない？

- 非同時（継時）マスキング（temporal masking）
 - 前向性マスキング（forward masking）
 - 大きなノイズ聴取直後は、神経興奮が収まるのに時間がかかる
 - 後向性マスキング（backward masking）
 - 強い刺激ほど神経を伝わる速度が速くなる（？）
 - ノイズによる神経発火の伝搬が信号のそれを追い抜く（？）
 - それよりも、心理的な要因？



非同時マスキングを聞いてみる

順向性マスキング

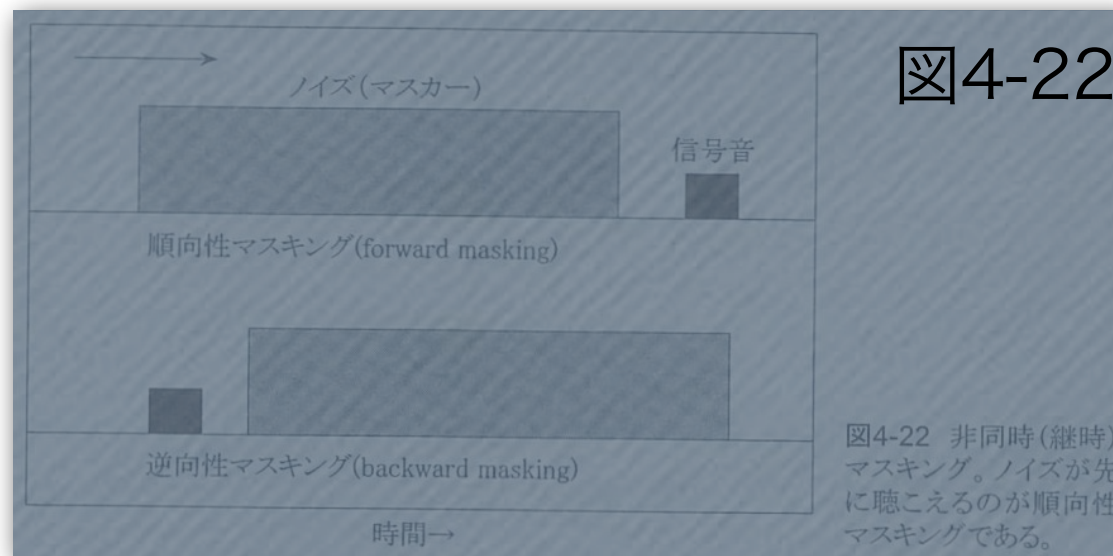
● ノイズ + 信号音 (69)

● ノイズ+信号音 (70)

逆向性マスキング

● 信号音 + ノイズ (71)

● 信号音+ノイズ (72)



両耳聴

両耳加算効果

- 聴覚閾値は片耳聴と両耳聴で異なるのか？
 - 両方の片耳に、聴覚閾値より小さな音を呈示しても聞こえないはず。
- 両耳聴では、片耳聴に比べて聴覚閾値が約 3dB 下る。
 - 片耳聴では、聞こえ、という意識体験は無い、場合でも・・・
 - 両耳からの刺激が中枢に届き、加算された結果、知覚される。

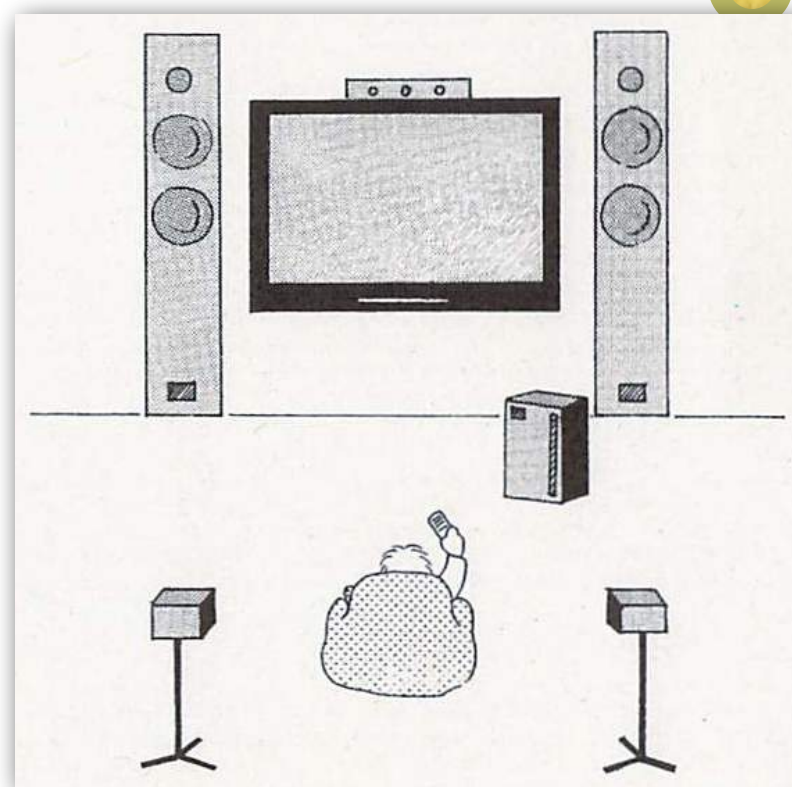
両耳融合効果

- 耳は左右にあるにも拘わらず、左右以外の方向性を感じる。
 - サラウンド効果、音（音像）の定位
 - 左右から全く同じ音を呈示すれば、音像は頭内の中央に定位。
- 5.1 chシステムを例にとって
 - 本来耳は二つなので、録音時に耳の距離だけはなれた2本のマイクでヘッドホン聴取で完全なサラウンドは実現できるはず。
 - なぜ 5.1 本のスピーカが必要??

両耳聴

5.1chサラウンドシステム

- 何故、5.1本ものスピーカが必要？
- 通常はヘッドホンではなく、スピーカ呈示になる。
 - で、聴取者の音響体験が映画館のそれと類似するように計算して再生。
- 録音がそもそも「完全なステレオ」録音ではない。
 - ホールで両耳の位置にマイクを立てて録音している訳ではない。
 - 実際には、ボーカルの音、伴奏の音、XXXの音、など多チャンネルで録音を、それを編集した結果が音楽 CD として販売される。
 - 5.1chの環境を想定した編集がされることも。



両耳ビート

うなり (ビート)

- 僅かに異なる周波数の純音と重畳して生じる低周波のうなり
- 両耳にこれらの純音を聞かせるとどうなるか？
- ある種の「うなり」のような独特の感覚が生じる。

とにかく聞いてみよう (73)

- 1000 Hz 以上では知覚が困難
- 440 Hz + 441 Hz
- 1200 Hz + 1201 Hz

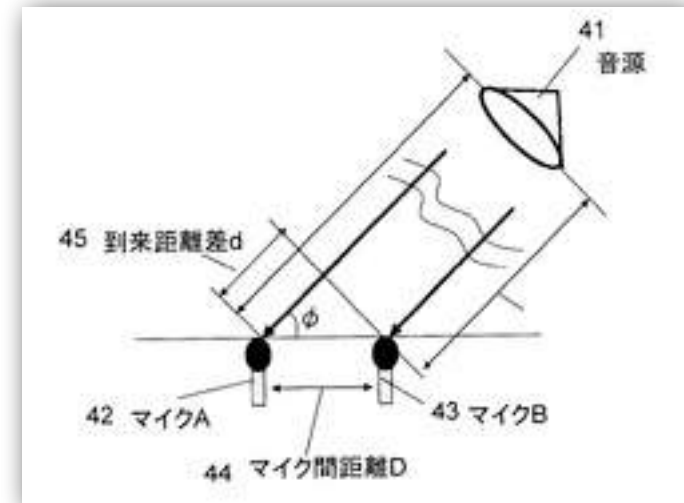


方向知覚

ある音源から両耳に至る経路の違い

$$\frac{d}{D} = \cos \phi$$

- D は事前に分る。
- d は（波形の）測定により分る。
- その結果， ϕ が推定できる。
- 遠い方が強度も弱まる。
 - 左右の強度差から ϕ が推定される。
- 二つの情報源の使い分け
 - 1500 Hz よりも低周波：時間差による方向推定
 - 低周波だと，回折現象（回り込み）により強度差が不明瞭に
 - 1500 Hz よりも高周波：強度差による方向推定
 - ここまで高周波になると，時間ピッチ的な処理が困難に



方向知覚における二つの情報源

時間差による知覚

- 500Hz, 2000Hzに対して等しい位相差を実現 (74)

強度差による知覚

- 440Hz, 4000Hzに対して等しい強度差を実現 (75)
- ヘッドホンだと回析しない



非定常音（クリック音）の方向知覚

パルス的な波形の到着時刻差

- 時間差知覚の時間分解能は極めて高くなる。
- 1度の違いまで区別できるとの報告例もある。
 - 約9マイクロ秒=10万分の1秒の違いを弁別

聞いて確認してみる

- 1万分の1秒=100マイクロ秒のパルス時刻差（99）



Masking Level Difference (MLD)

マスキングと到来方向

- マスカーとマスキーが同方向から到来：マスキング量大
- マスカーとマスキーが異方向から到来：マスキング量小
 - 後者の場合は，方向のずれによって両者を区別しやすくなる

同相・異相と音像定位

- 両耳に同じ音（同相）が与えられる：頭内中央で定位
 - つまり，音源はある一つの箇所にあるように聞こえる
- 両耳に反転音（異相）が与えられる：どちらが先か後かが不明
 - 音源方向の定位は位相情報からは困難に
 - 頭内全体から聞こえてくるように感じる

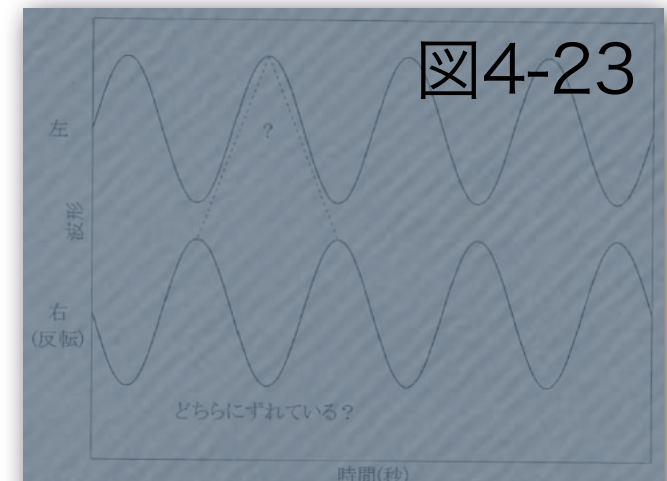
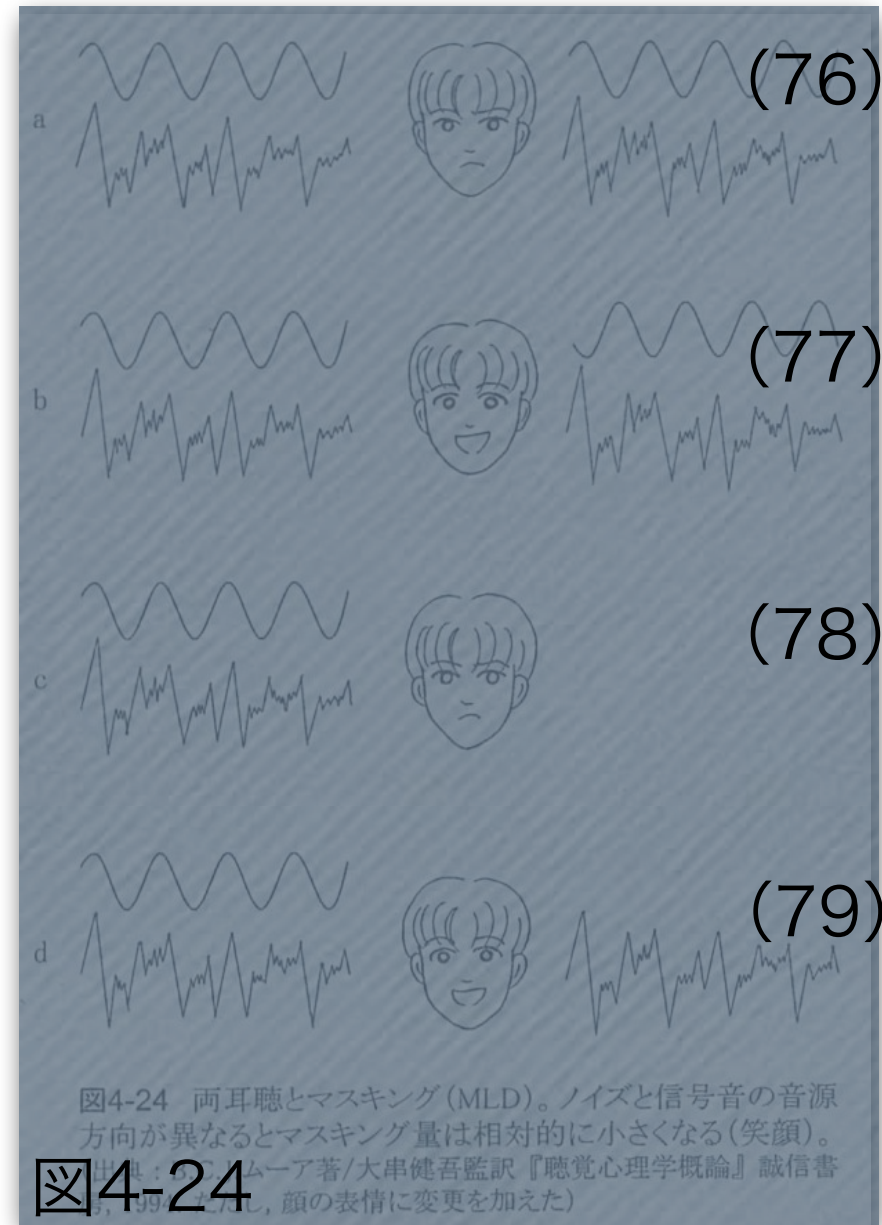


図4-23 波形の反転と位相差。左右耳の音を上下反転すると、ちょうど半波長分だけずれたのと同じことになり、どちらの音が遅れているのかわからない。

Masking Level Difference (MLD)

同相・異相と両耳聴

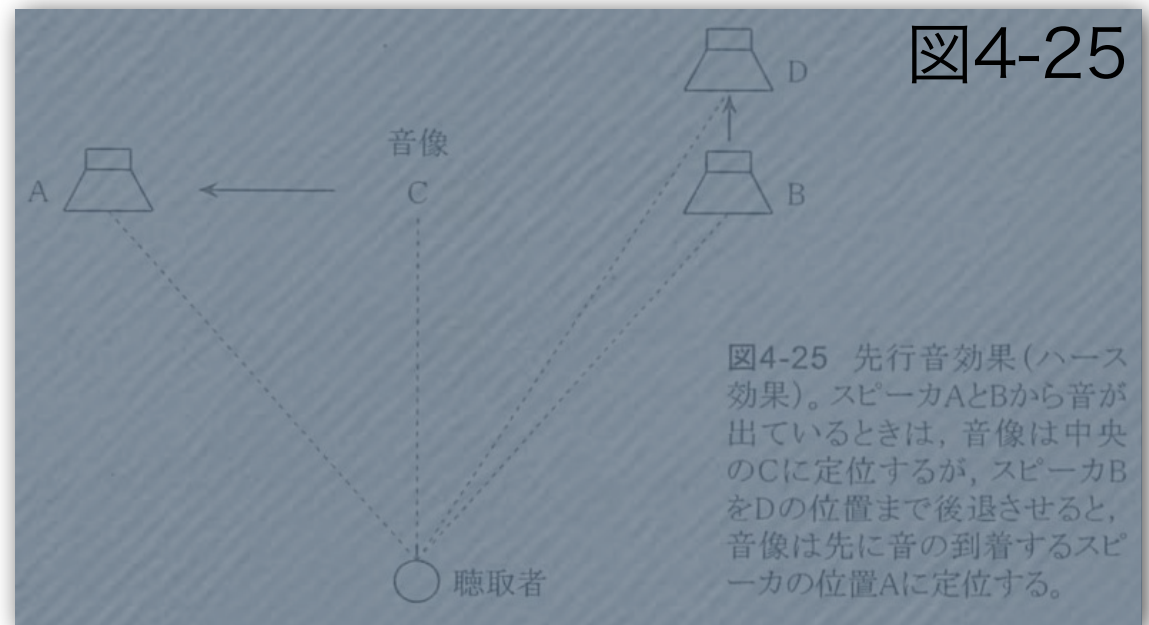
- 同相のものが両耳(a), あるいは片耳に呈示されると(c), ×
- 両耳に異相の信号が呈示されれば, マスキング効果が薄れる(b)
- 信号は頭全体に定位。よってマスクーとマスクーが同位相ではなくなる。
- 両耳にマスクー, 片耳のみにマスクーが呈示されれば両者の方向性が異なるので(d), ○
- 信号は右のみ。この場合も, マスクーとは同位相ではなくなる。



先行音効果

複数音源の音に対してどの方向に音像を感じるか？

- 先に音が聞こえてくる（先に音が到着する）方向に音像を感じる
- 先行音効果（precedence effect）
- ハース効果（Hass effect）
 - $A + B \rightarrow C$
 - $A + D \rightarrow ?$

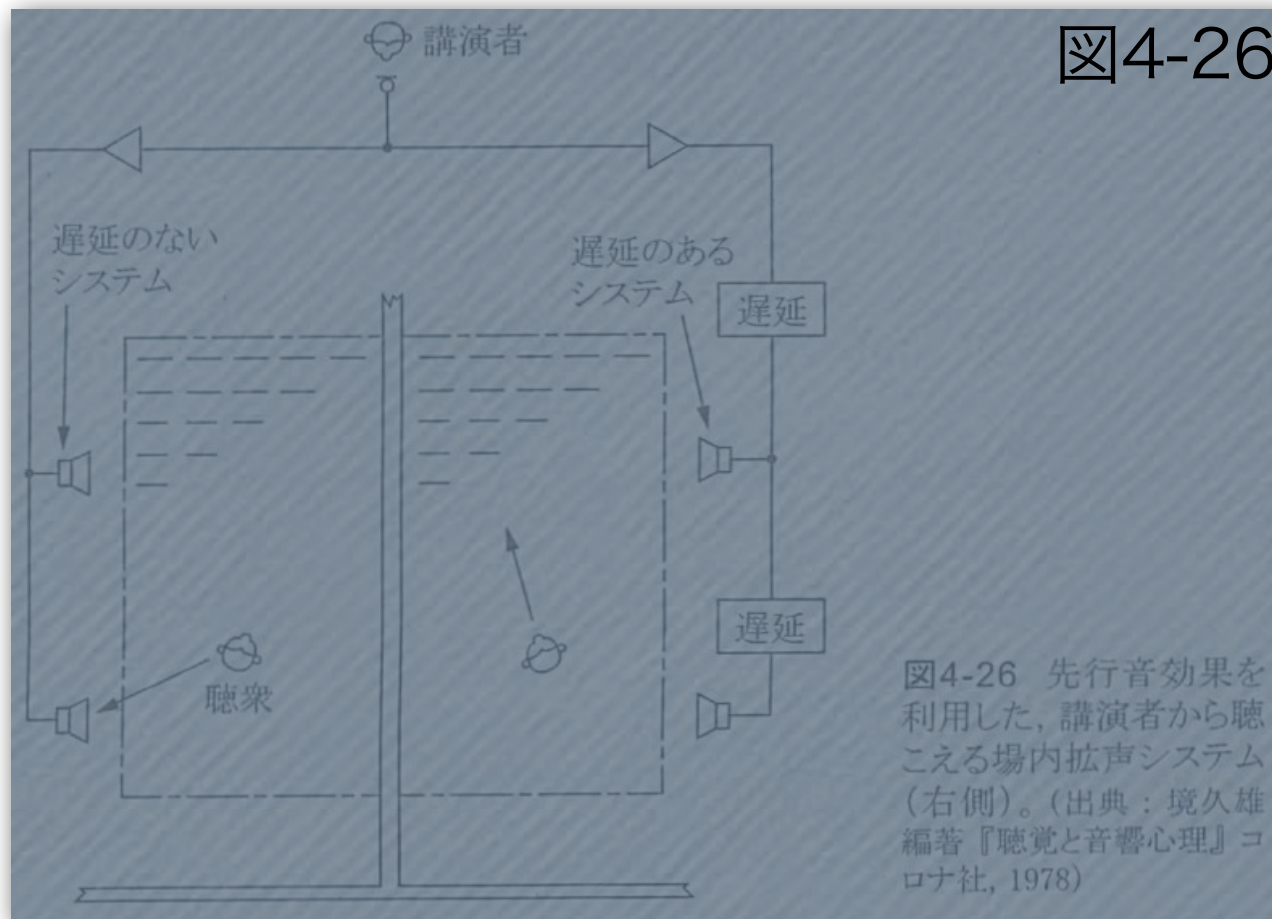


- 2本のスピーカを真正面に置いて聴取する。音像は中心に。
- 首をひねって、僅かに右を見る。左からの距離が短くなる。
 - 音像は左にあるように感じる。

先行音効果の応用例

講演者の方から音が聞こえてくるように見せかける

- 右スピーカは遅延を入れる → 右聴衆は講演者から聞こえる
- 左スピーカは遅延なし → 左聴衆はスピーカから聞こえる



先行音効果の応用例

聴覚で環境を捉える視覚障害者

● 足音の知覚

● 足→耳， と来る直接音

● 足→障害物（壁など）→耳， と来る間接音（反射音）

● 直接音と反射音

● 障害物が十分遠く， 直接， 反射の時間差大 → 別々に知覚

● 障害物が近づき， 時間差小 → 反射音は直接音に吸収

（先行音効果）

→ 障害物が近いことを知る

