

秋田県立大学大学院博士学位論文

2音源により生じる音像と
先行音効果の関係に関する研究

播 摩 敏 雄

2015 年 3 月 20 日

緒 言

近年，DAT，CD，MD，DVD 等のメディアの急速な進歩により，オーディオ環境はアナログからデジタルへと劇的な変化を遂げた．また，サラウンド音の付加などによる音空間の創出も行われている．更に，コンサートホールなどで交響曲を聴取している音場の臨場感を，他の場所で正確に再現する技術の研究や音と映像を組み合わせたバーチャルリアリティシステムにおいても，デジタル信号処理技術が駆使され，積極的に音を加工したり，音と映像を組み合わせる試みなどが行われている．このようにオーディオ技術やデジタル信号処理技術が進歩しても，実空間で音の方向を聞き分けたり，音を通じた認知を行ったりするのは人間であり，実空間における人間の知覚過程の仕組みについて探究することが不可欠である．

人間は，聴取した音から，音そのものが持つ性質，例えば，音の大きさ，高さ，音色等を知覚することができる．その他に，音の発生した方向や広がり感のような空間的情報を得ることもできる．音から音源の方向や距離を感じる能力は音像定位と呼ばれ，知覚された音の像は音像と呼ばれる．音像定位は，距離知覚と方向知覚に大きく分けられる．距離知覚とは，音像の遠近に関する知覚である．通常の距離知覚に関する手掛かりの一つとして，直接音と反射音の時間差がよく知られている．さらに，聞きなれた信号であれば，異なる位置にある音源から到来するその音の音量を比較することによって，音源間の相対的距離を知覚することもできる．その他に，距離が遠くなるにつれて高い周波数成分が低い周波数成分より減衰しやすい，周波数スペクトルの変化も音像定位の手掛かりと言える．

方向知覚では，両耳受聴の場合，両耳間位相差（左右耳に到達する音波に生じる位相の差），両耳間時間差（左右耳に到達する音波に生じる時間の差），両耳間強度差（左右に到達する音波に生じる強度の差）などの手掛かりから音像の方向を判断しているものと言われる．また，単耳聴の場合でも，頭あるいは耳介による回折現象のために生じるスペクトル構造の変化を手掛かりとして方向が知覚されることが知られている．

単一音源の場合は，音源方向や距離が知覚に上述の要因が手掛かりとなる．しかし，音源が二つ以上になった場合には，複数の音源から出た音がそれぞれに到達時間差や音圧レベル差を持ち，音像の知覚に影響を与えると考えられるため，単一音源より複雑と

なる．音源が複数存在する状況下で最も単純な構成，すなわち音源が二つ存在し，類似した音が発生した場合を考えると，二つの音源間の時間差によって音像の変化が起こることが知られている．この状況は，二つの音源がほぼ同時に提示されると，二つの音源の位置の中間から合成された音像が聞こえるように感じる．これは加法定位 (summing localization) と呼ばれている．加法定位から先行音側に一つの音像として知覚される．この現象を先行音効果 (precedence effect) と呼ばれる．2 音源の時間差の境界を SLT (summing localization threshold) と本研究で名付けた．2 音源間に十分な時間差が存在する場合には，それぞれの音源から音が別々に聞こえる．これはエコー (echo) と呼ばれ，先行音効果は，二つの音源にわずかな時間差が存在する場合には，先行する音源側からのみ音が聞こえるように知覚される現象で，過去に先行音効果について調べられた代表的な実験には，Wallach 及び Haasが行ったものがある．Haasの実験では，二つのスピーカを聴取者から等距離に配置して，両スピーカから出力したパルス音の到着時間差や音圧レベル差を変化させて，どのスピーカから聞こえてくるのかを調べた．実験の結果，先に出る音の到着時間を 1～30ms遅らせた場合，2 音源からの音を聞いているにもかかわらず，先行したスピーカ側からしか音が聞こえないという現象が起こった．その現象は 2 音のレベル差を 10～15dB程度大きくなるまで見られた．類似した音が複数の音源からわずかな時間差で放射された場合，最初に到達した音の方向に音源を知覚する現象であると定義し，これを先行音効果と呼んだ．Wallach, Haasらの実験以降に行われた研究によって得られた先行音効果に関する知見を以下に示す．

- 聴取者の耳に連続して到達する二つの短音は，その時間差が十分に短ければ，一つの音として聞こえる．融合の生じる時間差は音の種類によって異なる．この時間差の上限は，パルス音については約 5ms程度であるが，音声や音楽などのような複雑な音については，30ms程度となることもある．
- 二つの短音が融合して一つの音に聞こえるならば，全体の音の聞こえてくる位置は，第 1 音の位置によってほとんど決定される．
- 第 1 音の位置が第 2 音の方向の差が最大約 10° までの間のときには，第 2 音の方に音全体を引っ張る．そして，それ以上では第 2 音の影響が小さくなっていく．二つの音の到着時間差が 1ms以下ならば，第 1 音と第 2 音の間に音像が生じる．
- 第 2 音が十分に第 1 音より (10～15dB) 強ければ，第 2 音は先行音効果に打ち勝つ．
- 先行音効果は質的に似た音に対して生じやすい．第 2 音が第 1 音と大きく異なれば

融合は生じない．特に，第 1 音と第 2 音のスペクトルは似ていなくてははいけない．

従来の先行音効果の研究では，2 音源の配置方向が限定された前方の狭い範囲に設定され，刺激音としては，主にパルス音が用いられてきた．そのなかで，先行音効果からエコーへと変化する時間差の境界 ET については，実験の容易さからか（音が二つに聞こえたかの判断の容易さ）比較的多くの実験結果が残されている．古くは Cherry and Taylor (1954) らの研究をはじめとして，ヘッドホン受聴を含め 20 件以上の研究結果がある．一方，加法定位から先行音効果の時間差の境界 SLT に関する研究は ET の研究より，はるかに少ないように見える，その理由は先行音効果の評価方法によるところが大きいと考えられる．つまり，加法定位から先行音源側に変移して聞こえる一つ音像がどの方向から聞こえてくるのかを回答させるものがほとんどで，音像がどこにあるのかを定量的に示す基準が定められていないため音像の方向を明確に示すことが困難であったと考えられた．このようなことから，先行音効果を利用した拡声装置を設計するとき，ある条件で音源から音を放射した場合，聴取者の知覚する音像の方向を，従来の研究から推定することは依然として困難であると考えられる．本研究では，上述した問題点に重点を置き，2 音源によって生じる音像が，音源の位置，音源信号の種類および 2 音源間の時間差やレベル差などの条件によってどのように加法定位から先行音効果へと移動し，その時間差の境界がどの方向にあるのかを聴取実験を通して検討した．

このような研究では，刺激音として，日常的に存在する音を対象とするが当然であり，音声や楽器音などについても調べる必要がある．また，実生活では，音の方向を確認するとき，両耳で音の方向を決めて，視覚で音の鳴っている方向を確定している．このような，聴覚と視覚に刺激が同時に存在する場合は，ほとんどである．音の定位に対する視覚刺激の影響に関して，多くの研究がある．例えば，視覚刺激の付加により，聴覚刺激の知覚方向が視覚刺激の方向に引き寄せられることが知られている．このような音の方向定位に伴う，聴覚と視覚の要因が顕著に現れる例として，腹話術効果が挙げられる．また，言葉の認識では，視覚に影響される現象として，マガーグ効果 MacGurk, MacDonald が知られている．このような，聴覚刺激と視覚刺激が同時に提示されたとき，先行音効果は聴覚刺激の物理的特性と知覚との対応が異なる現象であるものの，基本的には，聴覚刺激に基づいて起こる現象であるが，そこに視覚刺激が付加されたとき，先行音効果の生じ方に影響が現れることは十分に考えられる．

上述したように，本研究では，2 音源によって生じる音像が，音源の位置，音源の種

類および 2 音源間の時間差やレベル差などの条件によってどのように変移するのかを明らかにすることを第 1 の目的とした。更に、視覚刺激が付加された場合、2 音源間の諸条件に対し、先行音効果の振る舞いに、どのような影響を与えるか明らかにすることを第 2 の目的とした。

上述した目的を達成するため、本論文を以下のように構成した。

第 1 章は、2 音源によって生じる音像に関する過去の研究を、特に先行音効果に着目して概観しながら本研究の背景を述べ、従来の研究の問題点を指摘したうえで本研究の目的を述べる。

第 2 章は、2 音源から時間差・レベル差を持って提示される同一な音刺激によって生じる音像の方向が、一般的に定義されている加法定位、先行音効果とどのような対応関係にあるかを調べるために、聴取実験による検討を行う。刺激音は、バースト音(10ms)および長い継続時間のピンクノイズを用いた。実験では、共通の被験者に対して、音像の方向を判断させる実験(絶対方向判断)と視認できる音源のどこに音像が最も近いかを判断させる実験(3 肢強制選択)を行い、その結果を比較する。本章では、特に、時間差の境界 SLT を詳細に調べた。また 2 音源の配置は、被験者の正面を 0° として、 $\pm 30^\circ$ および $\pm 60^\circ$ の 2 種類、音源間のレベル差として、先行する音源のレベルを基準として後続する音源のレベルを 0, +3, +6dB の 3 種類に設定し、先行音効果の下限の境界 SLT と上限 ET にどのような関係があるのかを調べた。

第 3 章では、刺激音の種類とレベル差が音像の知覚に及ぼす影響について、加法定位から先行音効果への移動だけではなく、エコーへの知覚の変化も含めて実験的に調べた。刺激音は、スピーチ(男声・女声)、ギター、サキソフォンの 4 種類とレベル差は 0, +3, +6, +9dB と設定し、先行音効果の下限の境界 SLT と上限 ET にどのような関係があるのかを調べた。

第 4 章では、実空間を考えて、音源となる話者や演奏者、拡声用のスピーカなどが視覚的に確認できることが多い。そのことを考慮すると、視覚から得られる情報が 2 音源によって生じる音像にどのような影響を与えるかについて、SLT と ET に重点をおいて実験的に検討した。刺激音は、第 3 章で用いたと同じものを用いた。また、音声の発声時や楽器の演奏時の映像を同時に収録し、被験者の前方正面に設置したスクリーンにプロジェクターを用い等身大の大きさで投影した。聴覚刺激のみを提示したときと聴覚刺激と視覚刺激を同時に提示したときの比較により、音像の知覚方向の対応関係について調べた。第 5 章は結論である。

目 次

緒言

第1章	1
-----	---

序 論	1
-----	---

1.1 研究の背景	1
1.1.1 1 音源による音像定位	2
1.1.2 2 音源による音像定位および先行音効果	3
1.1.3 音像定位における視聴覚の相互作用	7
1.2 2 音源による音像に関する過去の研究の問題点	9
1.3 研究の目的	11

第2章 先行音効果の生起と知覚される音像の関係	13
-------------------------	----

2.1 はじめに	13
2.2 先行音効果と方向定位の実験方法	14
2.2.1 音源配置と実験系	15
2.2.2 聴取実験に用いた刺激音	15
2.2.3 絶対方向判断の実験（実験Ⅰ）	20
2.2.4 3肢強制選択の実験（実験Ⅱ）	20
2.3 実験結果の評価	22
2.3.1 絶対方向判断の実験結果	22
2.3.2 考察	36
2.3.3 3肢強制選択の実験の結果	37
2.3.4 考察	38
2.4 実験Ⅰ、Ⅱの結果の比較	45
2.5 まとめ	48

第3章 刺激音の種類とレベル差が先行音効果の境界に与える影響	51
--------------------------------	----

3.1 はじめに	51
3.2 刺激音の継続時間の知覚と認知	51

3.3	刺激音の種類が先行音効果に与える影響	54
3.3.1	音源配置と実験システム	54
3.3.2	口頭教示と回答方法	57
3.3.3	実験結果	60
3.3.4	考察	63
3.4	レベル差の違いが先行音効果に与える影響	73
3.4.1	実験系, 実験条件及び回答方法	73
3.4.2	実験結果	76
3.4.3	考察	77
3.5	まとめ	87
第4章	先行音効果に及ぼす視覚刺激の影響	89
4.1	はじめに	89
4.2	聴覚と視覚の相互作用	90
4.3	実験概要	93
4.3.1	視聴覚刺激	93
4.3.2	実験系	99
4.3.3	刺激音の再生と回答データ収集装置	101
4.4	SLT 実験	103
4.4.1	実験条件と回答方法	103
4.4.2	SLT 実験の結果	104
4.4.3	考察	107
4.5	ET 実験	141
4.5.1	実験条件と回答方法	141
4.4.2	ET 実験の結果と考察	142
4.6	まとめ	166
第5章	結論	167
	謝辞	175
	参考文献	177

著者発表論文	-----	183
--------	-------	-----

第 1 章

序 論

本章では、本研究の背景として、人間が音の方向を知覚する要因を概観し、その上で、室内空間における先行音効果に関する聴覚事象について述べ、先行音効果の発生する聴覚の時間的な要因の効果を考察する。さらに、音の方向定位と先行音効果に関する従来の研究について言及しながら、先行音効果の基礎的な振る舞いを概観し、先行音効果が室内空間での音の聴取にどのように関わっているかを考察する。また、それらの知見を基に、先行研究者がまだ検討していない事象について述べ、先行音効果の関わりと問題点を整理し、本研究の目的を述べる。

1.1 研究の背景

近年、音環境に対する関心が高まっている。音を物理的存在として扱うだけではなく、人間が生活の中でどのような音を聞き、いかに意味付け、利用しているかについてである。この生活環境の中で、音響伝送系のエンドユーザーの関心は、単に音が聞こえればよいというレベルから、あたかもその場にいるような、よりリアルな音を求めるレベルへと変化している。そのような要求に対応していくために、現在の AV 技術の延長線上に位置する「あたかもその場にいるような」視聴覚環境を作るバーチャルリアリティ技術に対し、各分野からのアプローチが行われている。一方、音響機器を利用するオーディオ環境では、デジタル信号処理技術の進歩により、アナログからデジタルへと著しい変革が遂げられ、デジタル録音の進歩や CD, DAT, MD, DVD などのデジタルオ

オーディオ機器の普及は、高品質で忠実な音をもたらした。しかし、このような機器を利用する主体となるのは人間であり、それらの機器の性能の評価や人間にとっての高品質を追求するためには、音を聴く、感じるなどを知覚する人間の聴覚機構の特性を知ることが不可欠である。

人間の聴覚は、音の3属性といわれる 大きさ、高さ、音色を知覚できる能力に加えて、左右の耳に入った音波の差やスペクトルから得られる情報によって、その音が聞こえてくる方向や距離、あるいはその音のもつ響きや広がりなどの空間的情報を得ることができる。そのうち、音の聞こえる位置を知覚できる能力は、音像定位と呼ばれる。では人間は、音の含まれるどのような情報から音像定位をしているのだろうか。その物理的な要因としては、耳入力信号の両耳間時間差、音圧レベル差、それらのスペクトル、反射音と直接音の関係、頭部の移動による耳入力信号の変化などが挙げられる。さらに、人間は、音がどこで鳴っているかよく分からない場合には、頭を動かして周辺を見回すなどの動作により、音源を探す。このように、頭部の移動や視覚情報を利用して音の方向を精度良く認識できる聴覚と視覚の相互作用によるところも大きいと考えられる^{1,2)}。

1.1.1 1 音源による音像定位

1 音源の場合、音像定位の基本的な要因は、音源の位置によって三つの立場から考察されている。すなわち、両耳を含む水平面上に音源がある場合の水平面定位、頭部の中心を通り、両耳を結ぶ直線に垂直な面上に音源がある場合の正中面定位、及び、距離の違いを知覚する距離定位である。これらに関する従来の研究結果によると、水平面定位については、音源の方向が異なれば、音源から両耳までの距離の差が異なり、両耳間差が変化することから、主に両耳間差の情報を利用していると考えられている^{1,2)}。一方、正中面定位では、両耳間差は小さいかほとんどないと考えてよい。この場合、音源の仰角あるいは前後の位置が異なることにより、音源から鼓膜までの音響伝達系の情報(頭

部伝達関数)を用いていると考えられる^{1,2)}。また、距離定位は、ラウドネス、スペクトル、聴取空間に存在する反射音などに依存しているといわれ、さまざまな要因を総合的に用いて知覚されることが考えられる。従来の研究では、水平面定位、正中面定位、横断面定位、距離定位、頭部移動の効果、反射音効果、音の臨場感、音像の広がり感、音色感などについて、心理音響、電気音響、室内音響等の多種多様にわたり研究が行われて来た。音像定位に関する初期研究が遂行されていた頃は、現在のような安価で高速なコンピュータを用いたデジタル信号処理等の音源作製(音声・楽器音・自然音)の時間差・レベル差などの制御が困難であったため、純音や帯域ノイズなどに限られていた。そのため、音像定位の要因として、耳入力信号の両耳間差(時間差と音圧レベル差)の物理的なパラメータを正確に提示するのが難しく、聴覚系システムの知覚過程を詳細に知るとは困難であると思われた。最近では、パーソナルコンピュータのオーディオインタフェースを用いて、刺激音の作製・加工も容易に提供できるようになり、音像定位の手掛かりを明確にするため、刺激音の物理量だけを変化させる方法によって、個々の手掛かりの特徴が明らかにされている^{1,2)}。

1.1.2 2 音源による音像定位および先行音効果

音源が二つ以上になった場合には、複数の音源から出た音がそれぞれに到達時間差や音圧レベル差を持ち、音像の知覚に影響を与えると考えられるため、1音源より複雑となる。音源が複数存在する状況下で最も単純な構成、すなわち音源が二つ存在し、類似した音が発生した場合を考えると、二つの音源間の時間差によって音像の変化が起こる。この状況は、二つの音源がほぼ同時に提示されると、2音源間の中間から合成された音像が聞こえるように感じる。これは加法定位(summing localization³⁾)と呼ばれている。加法定位から先行音側に変化する時間差の境界を SLT (summing-localization threshold) と本研究で定義した。この二つの音源に十分な時間差が存在する場合には、

それぞれの音源から音が別々に聞こえる．これはエコー(echo)と呼ばれている．音像が二つに分離する境界をエコー検知限 ET と呼ぶ．この加法定位とエコーの中間の時間差で生じる現象が先行音効果(precedence effect)である．図 1.1 に、2 音源間の時間差による知覚現象の流れの概要を示した．

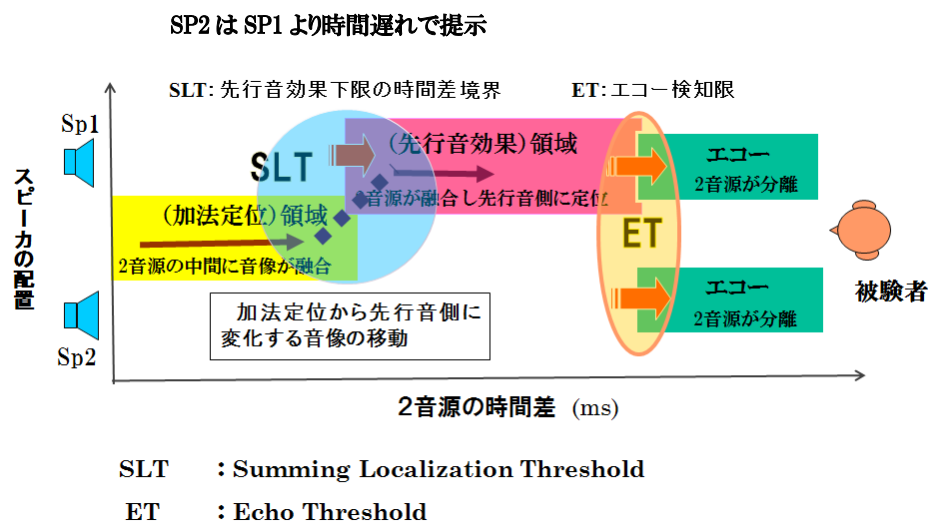


図 1.1 2 音源の時間差による知覚現象の流れ

SLT : Summing localization threshold と本研究で定義した．

過去に先行音効果について調べられた代表的な実験には、Wallach ⁴⁾及び Haas ⁵⁾が行ったものがある．Haas の実験では、二つの音源を聴取者から等距離に配置して、一つの音源から出る音の到達時間差や音圧レベル差を変化させて、どの音源から聞こえてくるのかを調べた．実験の結果、先に出る音の到達時間差を 1～30ms 遅らせた場合、2 音源を聞いているにもかかわらず、先行した音源側からのみ音が聞こえるという現象が起こった．その現象は二つのレベル差が 10dB 程度大きくなるまで見られた．このような類似した音が複数の音源からわずかな時間差で放射された場合に、最初に到達した音の方向に音源を知覚する現象が、先に述べたように先行音効果と定義されており、ハース

効果(Haas effect), 第一波面の法則 (law of the first wave front⁶⁾)とも呼ばれる.

その他の先行音効果に関する検討としては, 古くは 永井らが, 2つのスピーカから放射される信号の時間差およびレベル差と音像の方向の関係を調べている⁷⁾. また, de Boer⁸⁾によると, 方位角が約 $\pm 20^\circ$ の範囲では, 音像方向が音源間の時間差やレベル差と直線関係にあることを示している. また, 1/3 オクターブフィルタを通したトーンバーストに対する Wendt⁹⁾の結果では, レベル差が大きい範囲のインパルスの場合には de Boer による結果とほとんど等しくなるが, レベル差が小さい範囲のインパルスの場合には, その傾きが周波数によって異なっていることが示された. さらに, 1/3 オクターブバンドノイズの場合も同様の傾向を持つことを示している. 時間差と先行音効果による音像の関係については, 中心周波数に強く依存し, 純音を連続して提示した場合の合成音像の時間差に対する音像推移の軌跡は, 周波数に強く依存することを示している. また, 1030 Hz 以上の周波数では有効な結果が得られなかった. おそらく, 被験者が回答した方向のばらつきが大きかったためと推測される.

これらの現象を概観してまとめると, 2音源が同レベルの場合, 一方の音源の遅れ時間を増加させると先行音側に音像が変移するようになる. また, 2音源間の時間差がない場合に比べてラウドネスの変化も増し, 音像空間の中で先行音側の近くにまで変移し, 約 1ms 以上の遅れ時間になると音像は先行音側に片寄ってしまい, 空間的な広がりを変化して, 音像の重心が先行音の方向に移動する. 遅れ時間がある限度を越えると, 最終的には音像は分離して二つになり, それぞれ異なった方向に音像が定位される. 分離前に音像が一つ知覚されている場合には, 音像方向は先行音源方向によって定まる. 音像の方向を回答させる手法として, Queen¹⁰⁾は, 被験者位置を中心とした円周上に複数個並べた音源の番号を回答させ, それらの方向の平均として回答方向を求めている. また, Leakey¹¹⁾は, 2音源による合成音と, 2音源間に等間隔に配置された単一音源による音を継時的に提示して, 合成音の音像が単一音源の音像に対して左・右のどちらに知覚さ

れたかを回答させている。他にも、Braasch¹²⁾らをはじめとした幾つかの研究では、ヘッドホンを通じて提示された両チャンネル間に遅延時間をもつ刺激による合成音像と同等と感ずるように、両耳に提示される単一の刺激の ITD を調整させた。この方法では、知覚される音像として同等の方向となる ITD が得られるが、知覚される音像の方向を直接知することはできないという欠点がある。そこで、本章では、過去行われてきた先行音効果に対する音像の回答方法を見直し、被験者が知覚した音像の方向を直接得ることを重視し、音像の定位方向を絶対方向判断で回答させる方法と知覚した音像の位置がどの音源に最も近いのか 3 肢強制選択による二つの実験を対応することで加法定位および先行音効果といった知覚上の現象の関係をより詳細に議論できるよう留意している。

1.1.3 音像定位における視聴覚の相互作用

人間の周囲の環境に対する認知情報処理は、得られた大量の感覚情報を極めて効率よく処理することで実現されている。その中でも特に、聴覚と視覚の二つの感覚に依存するところが大きい。この二つの感覚の感覚器において、聴覚情報は主に時間構造で表現され、視覚情報は主に空間構造で表現されることが大きく異なる。一方、注意力や記憶と言った認知情報処理の段階では共通していることも多く、両方の感覚からの情報を相補的に利用した情報処理が行われているといえる。視覚刺激が音像定位に影響を与える例としては、腹話術効果 (Ventriloquism effect) がある。腹話術効果は、口の動きである視覚刺激と音声である聴覚刺激が聴取者から見て 10° 程度しか離れていないときに視覚刺激の方向に引き寄せられる現象である。その他の現象も含め、一般に定位に関する実験では、視覚情報が知覚上優位となる傾向が見られる。また、言葉の認識では、視覚に影響される現象として、マガーグ効果¹³⁾が知られている。

先行音効果は聴覚刺激の物理的特性と知覚した結果が矛盾した現象であるものの、基本的には、聴覚刺激に基づいて起こる現象である。しかし、前述したように定位判断

においては視覚情報が優位であるため、視覚刺激が聴覚刺激と同時に提示された場合、先行音効果の振る舞いに影響が現れることは十分に考えられる。つまり、コンサートホールなどで音楽を「生」で聴く場合、聴衆者は演奏者(視覚刺激)を視認している。このような場合に、先行音効果の生じ方に影響があることなどが考えられる。視覚刺激の存在が聴覚的知覚に影響を及ぼす例として、森本¹⁴⁾らのエコー検知限に対する視覚刺激の影響に関する研究がある。この研究では、刺激音として女性によるスピーチと二種類のリコーダ演奏を用いている。直接音の方向は正面、反射音の方向は正面と左側方の二種類で実験を実施している。その結果、映像の方向が直接音の方向と一致し、反射音の方向と異なっている場合には、エコー検知限は映像がない場合に比べて刺激音の種類に関係なく5ms程度長くなった。即ち、エコーが検知され難くなる傾向が示された。このように視覚は、可視光を物理的入力とする感覚器によって、外界にある物体の色、形、運動、テクスチャ、奥行きなどについての物体のカテゴリ情報を得て、物体の位置関係のような外界の空間的な情報や時間的な情報を知覚している。別の観点から見ると、映像と音の時間継続と映像の動きが同期している場合の調和感に対して、時間的なずれによる注意力の割り当て方がどのように影響を受けるのかについても明らかにすることが不可欠であると言える。Bolivarら¹⁵⁾と Lipscombら¹⁶⁾は、時間的に同期していれば聴覚と視覚に対する注意はより強められ、両者が同期していなければ注意は弱められるという仮説を提案している。また、視覚的な注意に対してGrimes¹⁷⁾は、音と映像の意味的な一致の程度によって、聴覚に対する注意と視覚に対する注意がどのように変化するかを調べた。その結果、反応時間のデータから、映像と音声が意味的に強く一致している場合は、不意に提示される映像に対して反応が遅くなるということが示された。視覚刺激に対してより注意が向けられたために不意に提示されると反応が遅れたのであると解釈できる。つまり、視聴覚刺激の意味的な高い一致は、視覚刺激に対する注意力を促進する。チャールズ スペンス¹⁸⁾は、聴覚と視覚のモダリティ間相互作用を扱う研究

について焦点をあてて、視聴覚統合について考察している．岩宮^{19, 20, 21)}らによる視聴覚の相互作用，また，「音楽と映像のマルチモーダル・コミュニケーション」(財)九州大学出版会では，オーディオ・ヴィジュアル・メディアを通しての音楽聴取行動における視覚と聴覚の相互作用，液晶プロジェクタによる大画面映像が音と映像の相互作業に及ぼす影響，音楽のリズムと映像の動きの同期が音楽と映像の調和に及ぼす効果などの研究が示されている．また，大串ら²²⁾は，人間が情報を受容する際に受ける心理的な反応と，これと関連する生理的な反応を手掛かりにして評価しようとする客観評価の方法について，心理的な測定技術を考察し検討している．このように従来の研究でも示されているように，聴覚の方向定位に視覚の影響が関与していることから，本研究では，聴覚の現象である先行音効果に着目して，時間的知覚と視覚での空間的知覚において視聴覚の相互作用としてどのような「振る舞い」を行うのか，また，聴覚刺激の時間的なずれに対して視覚刺激がどのような関係にあるのかを検討する．

1.2 2 音源による音像に関する過去の研究の問題点

既に，1.1.2 節で述べたように，先行研究者の研究によって得られた先行音効果の知見から，先行音効果からエコーへと変化する時間差の境界 ET に関しては，実験の容易さからも，比較的多くの実験結果が残されている．古くは Cherry and Taylor²³⁾らの研究をはじめとして，ヘッドホン受聴を含め 20 件以上の研究結果がある．しかし，加法定位から先行音効果の時間差の境界については，定量的に評価されておらず，その時間的境界において知覚した音像がどの方向にあるかも明確ではない．このため，先行音効果を利用した拡声装置を設計する場合，配置条件や遅延時間の設定が決定され音源から音を放射したときに，聴取者の知覚する音像の方向を従来の研究から推定することは依然として困難である．日常的な空間で，先行音効果が果たしている役割について考えて見る．通常，音が発生した場合，聴取者は音源からの直接音だけではなく，周辺の壁や

物体からの反射音を聞いている。通常、反射音の到来する方向は、音源のある方向とは異なっており、あたかも複数の音源から音が到来しているように解釈できる。しかし、聴取者は、大概の場合、反射音の存在に惑わされることなく音源の位置を特定することが出来る。それは、反射音が直接音に比べて聴取者へ到達するのがわずかに遅く、エネルギーもより小さいためである。わずかな時間差しかない場合、音像は分離せず一つとなって聞こえ、またそこに先行音効果が働くため音源のある方向が知覚されと考えられる。コンサートホールや講演会場のように大多数に情報を伝達する必要のある場所では、音源信号の伝達を補助するために拡声装置がよく使われている。このような場所がよく見られるように天井や壁面などに拡声装置を設置した場合、前述した直接音と反射音の関係が成り立たなく場所が生じることは容易に想像ができる。このように拡声装置が使われた場合、近い拡声装置からの音が最初に到達する聴取者がいる可能性が考えられる。また、各々の拡声装置からの到来音に到達時間差が生じてしまうことも考えられる。もし、到達時間差に大きな差が生じてしまった場合、エコーとして聞こえてしまうことになる。仮に、その到達時間差がエコーになるほど大きくなかったとしても、先行音効果により最初に耳に到達するスピーカの方向へ音が定位してしまい、本来定位してほしい音源方向とは異なる定位が生じることとなる。こういった多くの問題を解決するためには、拡声器の配置や各々に設ける到達時間差などを制御し、先行音効果を適切に発生させることが出来れば、違和感なく音源方向から音を聞くことができると考えられる。室内における音源定位の改善を目的に作られたシステムで、先行音効果を利用した代表的なものに、竹中工務店の音方向制御拡声システム(OZ-T10)が挙げられる^{24, 25)}。そのシステムは、ステージ上にいる話者から発せられた音をマイクで集音して、話者に近いスピーカからの音を聴取者に早く到達するように、時間差を付けることによって音像定位を改善するというものである。実際にこのシステムは、教会の礼拝堂で導入された。この礼拝堂は、教育施設に設置されているなかで、パイプオルガンや讃美歌の響きの美

しいホールとして定評がある。しかし、音響設計において響きの豊かさを最優先したため残響時間が長く、礼拝の説教や講演の声が聞き取りにくいという問題が見られた。そこで、パイプオルガンや讃美歌の美しい響きを保持したまま、建物の内装にほとんど手を加えることなく、その問題を解決するために、音方向制御拡声システムを導入した。このような例から分かるように、残響時間が長く、講演の演説が聞き取りにくいコンサートホールなどには、再度音響改善を施すのはコスト面からも望ましくない。そのため、既存の拡声装置を用いて音響改善を行うことができれば、望ましいといえる。また、実際のコンサートホールや講演会場では、演奏者や講演者が視認され、聴覚と視覚の情報も存在する。1.1.2 節で述べたように、視覚刺激が音像定位に影響を与える例として、腹話術効果がある。言葉の認識では、視覚に影響される現象として、マガーグ効果が知られている。このような、聴覚刺激と視覚刺激が同時に提示されたとき、先行音効果は聴覚刺激の物理的特性と知覚との対応が異なる現象であるものの、基本的には、聴覚刺激に基づいて起こる現象であるが、そこに視覚刺激が付加されたとき、先行音効果の生じ方に影響が現れる。また、先行音効果の ET に関して視覚刺激の影響について、森本らのエコー検知限に対する視覚刺激の影響に関する研究がある。それは、刺激音として女性によるスピーチと二種類のリコーダ演奏を用いていた。その結果、エコー検知限 ET は映像がない場合に比べて刺激音の種類に関係なく 5ms 程度、エコー検知限が長くなった。即ち、エコーが検知され難くなる傾向があることを示している。しかし、この研究では、被験者 3 名、刺激音が 2 種類、音源配置 2 方向のデータであり、加法定位から先行音効果に至るまでの時間差については視覚刺激の存在の影響を検討していない。現在、視覚刺激の影響を考慮して先行音効果の発生する条件について調べた研究は極めて少ない。コンサートホールなどの臨場感を考えるためには、広範な時間差に対して視覚刺激の影響を検討することが不可欠であると考えられる。

上述したように、本研究では、2音源によって生じる音像が、加法定位から先行音側に

変移した先行音効果下限の時間差の境界を SLT と定義し、先行音効果の上限 ET に至るまでの音像の変化について、音源の位置、音源の種類、2音源の時間差、レベル差などの条件の変化が及ぼす影響を検討する。更に、視覚刺激が付加された場合、2音源間の諸条件に対し、先行音効果の振る舞いがどのように変化するのかについても検討する。

1.3 研究の目的

1.1.2 節と1.2 節で、既に述べたように、これまで、先行音効果現象の従来の研究では、2音源によって生じる先行音効果の知覚現象が広く研究されてきた。しかしながらその結果を概観しても、2音源によって生じる音像が、時間差・レベル差・音源信号の種類・音源の位置などの条件により、知覚される音像の方向がどのように変化するのかについて十分に明らかとなっているとはいえない。それは、従来研究における先行音効果の評価方法によるところが大きい。つまり、2音源間の時間差の経過に伴って生じる加法定位から先行音側へ変移される時間差の境界については定量的に評価されておらず、その時間的境界において知覚した音像がどの方向にどの位置にあるのか明確ではない。このことは、先行音効果を積極的に利用して、拡声装置を設計するとき、ある条件で音源から音を放射したときに、聴取者の知覚する音像の方向を、従来研究から推定することは困難であることを意味している。そのため、本研究では、2音源間の時間差の境界をSLT (summing-localization threshold)と定義し(図1.1参照)、2音源によって生じる音像が、2音源間の時間差に応じてどのように推移し、どの程度の時間差(SLT)そして音像方向で、加法定位から先行音効果へ変化するのか、そしてそれが音源の位置、音源信号の種類および2音源間のレベル差等の条件によって影響を受けるかを明らかにする。また、先行音効果の上限のエコー検知限 ETにおいても諸条件とどのような関係があるのかを明確することを第1の目的にした。更に、1.1.3 節でも述べたように、コンサートホールや講演会場で講演・音楽などを聴いている場合には、聴覚

と視覚の情報が混在していることがほとんどである．先行音効果は聴覚に基づいて起こる現象であるが，空間知覚における視覚情報の優位性を考えると，視覚刺激が付加されると先行音効果の振る舞いに影響が現れると考えられる．そこで，視覚刺激が付加されたとき，先行音効果の振る舞いにどのような影響があるかを明らかにすることを第2の目的とした．

第2章

先行音効果の生起と知覚される 音像の関係

2.1 はじめに

第1章で述べたように，室内で知覚される音像に關与するものとして，直接音と反射音の時間的融合と分離といった聴覚現象が挙げられる．融合(合成音像)と分離(エコー)の境界付近に關する音像の定位についてはあまり詳しく知られていない部分が存在している．特に，2音源の時間差によって，時間差が無い場合には2音源の中間位置に音像が定位し，ある程度時間があれば先行音源方向に音像が定位する．その中間の時間差の場合には，音像もその二つの状態の中間にあると思われるが，具体的に2音源間の時間差に対して知覚される音像の方向がどのように変化するのは，これまでの研究では，明確に説明することができていない．また，従来の先行音効果を積極的に応用したPAシステムにおいても，コンサートホールなどの大規模な空間で，拡声を行いながら，実音源の方向に音像を向けるために配置された拡声用音源に対し，先行音効果を利用して，実音源方向に音像を向けることを主張しているものの，その効果が定量的に評価されているとは言えない．

本研究では，加法定位から先行音側に変移する時間差の境界をSLT (summing

localization threshold) と定義して、次の二つの実験方法で行った。一つ目は、知覚した音像の方向を回答させる（絶対方向判断）実験Ⅰ、二つ目は、実際に音を発する2つの音源(スピーカ)とその中間にもう一つスピーカを配置し、知覚した音像が三つの音源のどれに最も音像が近いかを回答させる（3肢強制選択）実験Ⅱである。2音源の中間においたスピーカは時間差がない場合に発生する加法定位の時の音像方向として配置した実際には音を発しない視覚的なダミーである。これらの実験を、共通な系と被験者で、可能な限り共通な条件のもとで行った。そのことにより、両実験結果の対応を直接とることが可能となり、加法定位および先行音効果といった現象について有益な知見が得られると考えた。

2.2 先行音効果と方向定位の実験方法

本章では、2音源間の時間差やレベル差などが音像定位に及ぼす影響と、音像の定位方向と加法定位及び先行音効果の関係を明らかにするため、共通の実験系、刺激音そして被験者で、回答方法の異なる2種類の実験（実験Ⅰ、Ⅱ）を行った。実験Ⅰは被験者に2種類の刺激音と二通りの音源配置、19段階の音源間の時間差、3段階のレベル差の組み合わせで提示し、知覚された音像の方向を直接回答させる方法で回答させた。以降、この回答判断を絶対方向判断と呼ぶこととする。2音源間の時間差がごく短いときには2音源の中央付近に音像が知覚される。実験Ⅱでは、この位置にダミースピーカを置き三つの音源となりうるスピーカを視覚的に見えるように配置して、知覚した音像が、どの音源に最も近いかを、3肢強制選択で回答させた。以降、この回答判断を3肢強制選択と呼ぶ。なお、実験Ⅱの基本的な実験条件は、実験Ⅰと同様としている。実験ⅠとⅡの主な違いは、被験者に求める回答の内容である。実験は視覚情報の問題から実験Ⅰ、実験Ⅱの順番で実施した。被験者には、それぞれの実験を行う前にどのような回答を行うかを教示した。被験者は、両実験に通して参加している。また被験者は、正常な聴力を有

する20歳代の男子学生10名である。

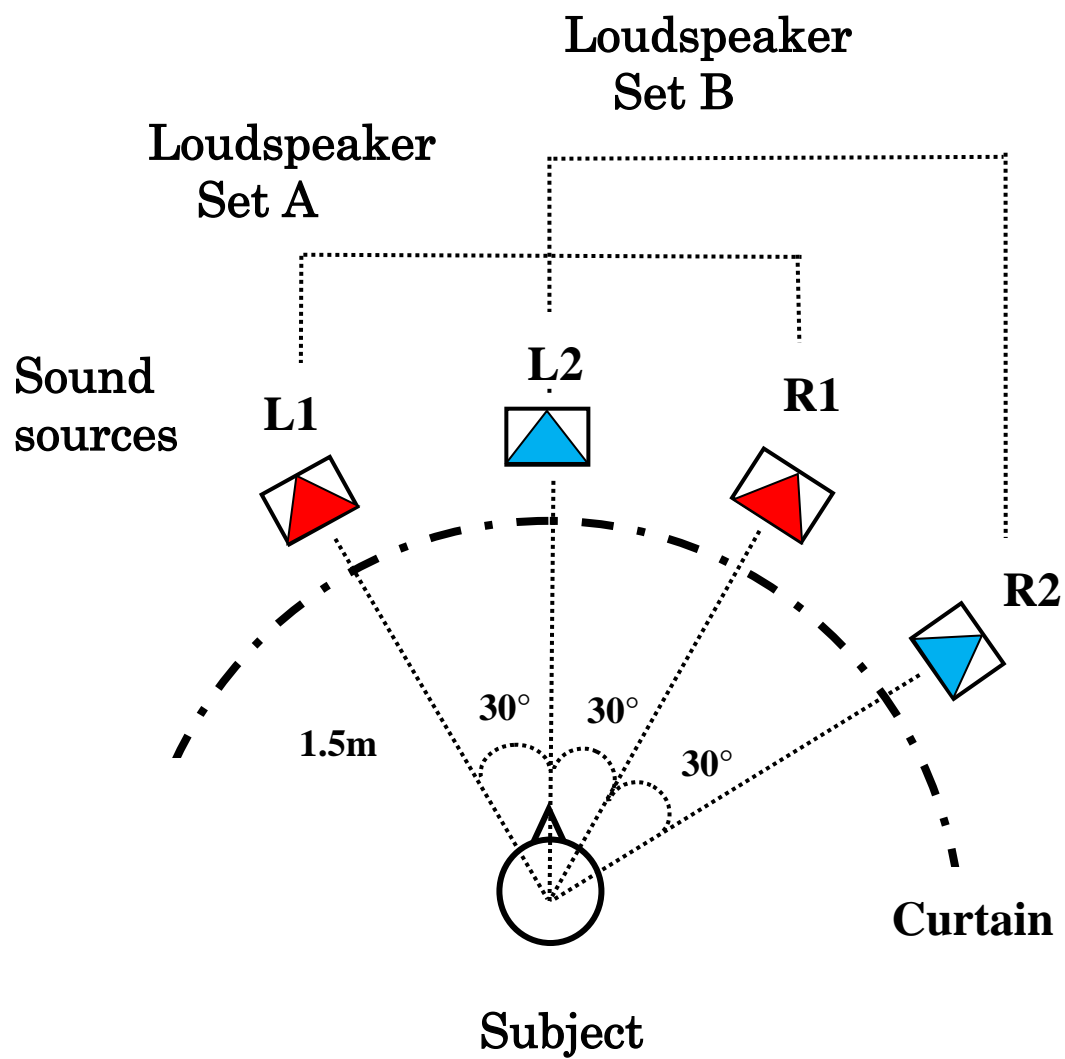
2.2.1 音源配置と実験系

実験Ⅰ，Ⅱの両実験は，ともに無響室内で実施した．実験系を図2.1(a)～(c)に示す．音源の配置方向の基準及び正負は，被験者正面を 0° ，左方向を負，右方向を正とする．実験は，2音源をそれぞれ $\pm 30^\circ$ に配置する条件と， 0° と 60° 方向に配置する条件の2通りで実施した．それぞれを以降，セットA，及びセットBと呼ぶこととする．図2.1に示すように，被験者とスピーカの距離は1.5 mとし，スピーカの高さは被験者が椅子に着座した際の頭部中心位置で1.4 mとした．被験者の頭部は軽く固定し，被験者には頭を動かさないよう教示した．刺激音は，パソコンを用いて出力し，アンプを経由して，それぞれの音源に供給された．刺激音は，クリック（ピンクノイズバースト）とピンクノイズの2種類である．遅延時間は，音源Rから音源Lの時間差として定義し， -2 から 2 msの範囲で19段階とした．音源の音圧レベルは，被験者の頭部中心位置で60 dBとし，2音源のレベル差は，レベル差なしの条件 0 dBに加え，音刺激の提示が時間的に遅れている音源に対して $+3$ ， $+6$ dBとなる3段階で行った．以上の条件のもと，実験Ⅰ，Ⅱを共通の被験者に対し，実施した．

2.2.2 聴取実験に用いた刺激音

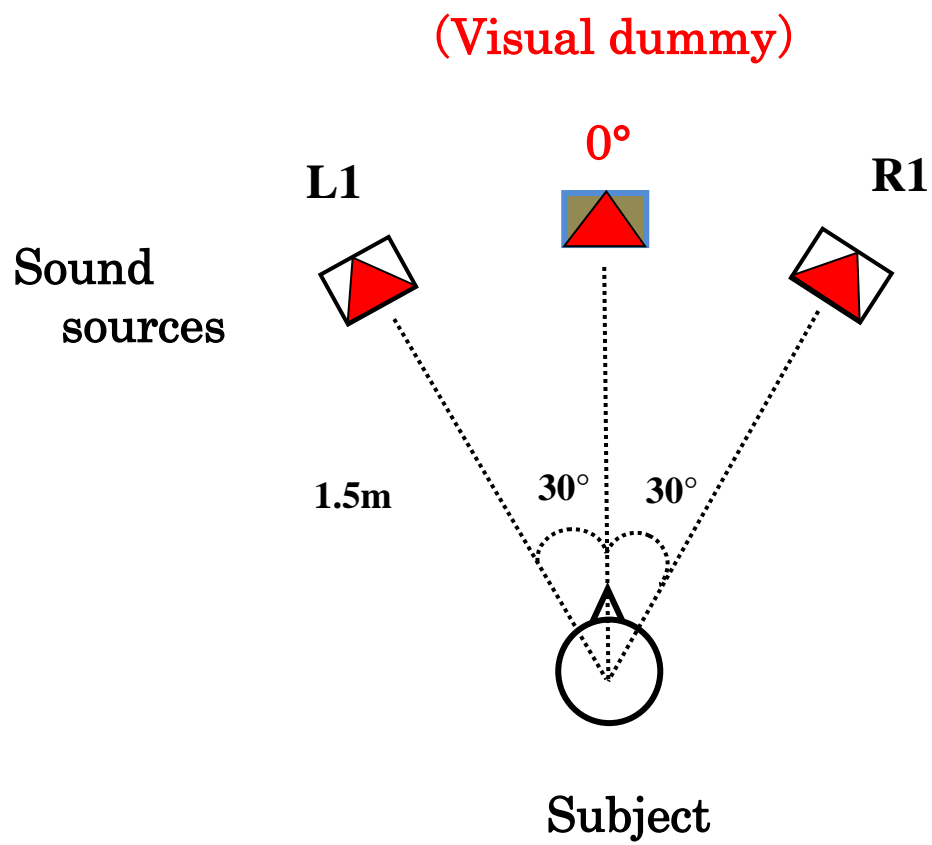
音に対する人間の知覚は提示される刺激音の継続時間に依存することが良く知られている．たとえば，ラウドネスについては，継続時間が約 100 ms までは継続時間が長くなるにつれて，同じ振幅であってもラウドネスが増加していき，それ以上の継続時間ではラウドネスが飽和する．同様に音の方向定位においても，継続時間の短い衝撃的な

立ち上がりや立ち下がりを持つ音源や継続時間の長い音源に対して、加法定位から先行音効果の音像変移の生起に継続時間の影響があるものことが考えられる。そこで、本研究では、刺激音を比較的短い継続時間のピンクノイズのバースト（継続時間：10 ms, 立ち上がり：2 ms, 立ち下がり：2 ms を含む、以降この刺激音をクリックと呼ぶ）と比較的長い継続時間のピンクノイズ（継続時間：600 ms）とした。2 音源の時間差は、図 2.1 の音源 R から見た音源 L の時間差として、 $-2.0, -1.75, -1.5, -1.25, -1.0, -0.8, -0.6, -0.4, -0.2, 0, 0.2, 0.4, 0.6, 0.8, 1.0, 1.25, 1.5, 1.75, 2.0$ [ms] の 19 段階である。この時間差の範囲において、被験者の回答する音像の方向は、2 音源の中間から先行して音が発せられる音源の方向へ変移することが想定される。各音源からの出力は、単独で被験者の頭部中心位置で 60dB となるように調整した。また、先行音効果を拡声へ応用することを考えるとき、拡声用音源に設ける利得が先行音効果に与える影響を考察することが重要であることから、両実験では、2 音源間にレベル差を設けた。先行して出る音が本来の音源、後続して出る音が拡声用の音源に対応すると考え、本実験では、後者の方が前者に比べ 0, +3, +6dB となる 3 段階で行った。



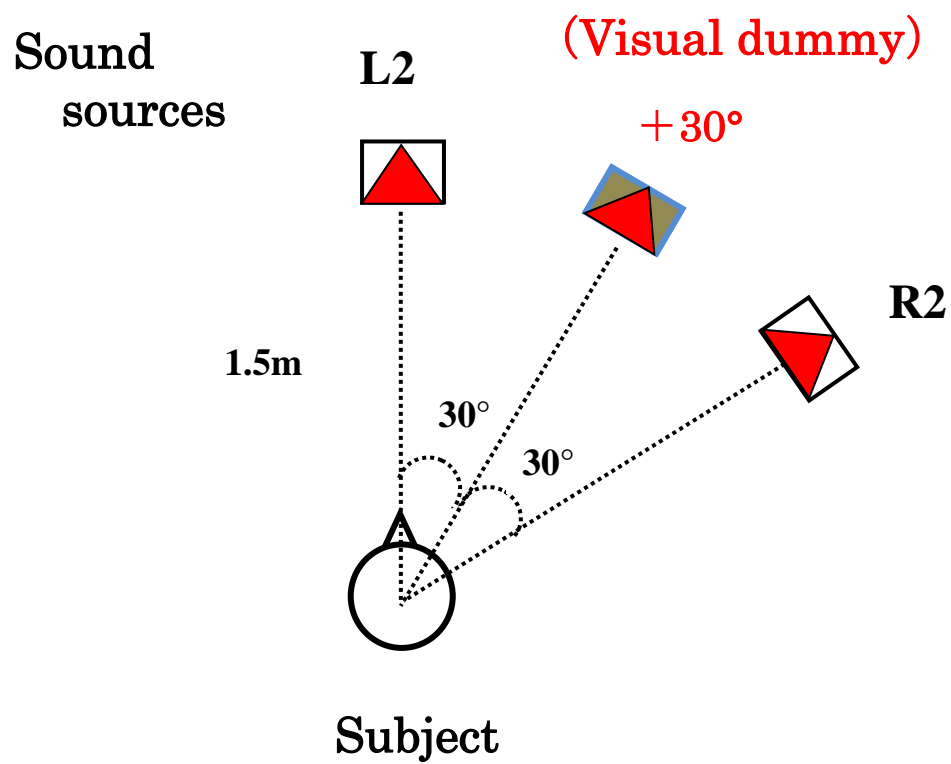
(a) 実験 I (音源配置, Set A, Set B)

図 2.1 実験の音源配置



(b) 実験Ⅱ（音源配置, Set A）

図 2.1 実験の音源配置



(c) 実験Ⅱ（音源配置，Set B）

図 2.1 実験の音源配置

2.2.3 絶対方向判断の実験（実験Ⅰ）

1.1.1節でも述べたように、定位方向の回答方法の例として、Queen¹⁰⁾が行ったように、被験者位置を中心とした円周上に実際に番号を付けた音源を複数並べて、知覚した音像を番号で回答させ、その回答方向を平均して音像方向として示す方法や、Leakey^{11, 26)}のように、2音源による合成音と、2音源間に等間隔に配置された単一音源による音を継時的に提示して、合成音が単一音源に対して左・右のどちらに知覚されたかを回答するものなどがある。また、Braasch¹²⁾らをはじめとした幾つかの研究では、ヘッドホンを通じて提示された、両チャンネル間にある遅延時間をもつ刺激による合成音像と、あるITDで両耳に提示される単一の刺激の定位位置が一致するようにITDを被験者に調整させる方法が知られている。この方法では、知覚される音像として同等の方向となるときのITDが得られるが、知覚される音像の方向を直接知ることにはできないなどの問題点がある。

そこで、本実験Ⅰでは、上述の回答方法を参考にしながら、知覚した音像の方向を被験者に直接回答させる方法をとった。この際、図2.1(a)に示すように、音響的に透明かつ視覚的には不透明なカーテンを音源と被験者の間に下ろし、視覚的な手掛りを被験者に与えることを極力避けた。被験者は、頭部を中心とした円が印刷された回答用紙に、音像が定位した絶対方向を「×」で記入するよう教示した。刺激音の提示の後、回答のための時間を4秒とった。刺激音の提示順序はランダムとした。ただし、同条件の刺激音が繰り返し提示されることはないようにした。

2.2.4 3肢強制選択の実験（実験Ⅱ）

1.1.1節でも述べたように、実験Ⅱでは、2音源間の時間差がごく短いときに知覚される音像は2音源の中央付近に定位される。この位置に視覚的なダミーとしてのスピーカ

を置き、三つの音源を視覚的に見える状態とし、その後提示した音から知覚した音像が3音源のどの音源に最も近いかを、3肢強制選択で回答させた。なお、実験条件は実験Ⅰと同様である。これは、Litovsky^{27,28)}らの実験で用いられた判断基準の1つである。この方法では、図2.1 (b), (c)に示されるとおり、回答のために音源位置の情報を被験者に視覚的に提示する必要があるため、実験Ⅰで使われていたカーテンはなくし、被験者に音源を見せた。ただし、刺激音の提示に用いられた音源は両端の2つだけであり、その間に置かれた音源からは音は発せられない。また、そのことを被験者に知らせていない。また、この実験を実施するには、音源を視認できる状態にする必要がある。この経験が実験Ⅰにおける絶対方向判断に影響しないように、全ての被験者において実験Ⅰを先に実施し、その後で実験Ⅱを行った。ただし、このような実験順序の問題とは別に、先行音効果の振舞いが同時に提示される視覚刺激の影響を受けることが報告されている²⁹⁾。このことは、実験Ⅱの結果を、音源を見せずに行った実験Ⅰの結果と比較する上で問題となる可能性もあると考えられる。しかし、文献では、音声の聴覚刺激に対する話者の映像を提示するといった、音源の性質と関連の深い視覚刺激の提示がなされたときの検討結果であり、本実験と同じような条件とは必ずしも対応しないと考えた。先行音効果の拡声への応用を考えたとき、拡声用の音源は視覚的に確認できる場合は多いと思われる。上記の回答方法は、過去に行われてきた先行音効果に関する研究において用いられてきたものと必ずしも共通なものではない。特に、Wallachらの行った実験^{30,31)}とは、ダミー音源の存在についても、回答の選択肢として設定したものも異なる。前述のWallachらの回答方法では、被験者の知覚する音像が左右のどちらにあるかはわかるが、その定位方向が加法定位および先行音効果のどちらに対応しているのかが不明確になってしまう。それに対して、Litovskyらの実験に沿った回答方法においては、2音源の中央付近にダミー音源を置くことで被験者の強制選択が容易となり、定位方向と、加法定位および先行音効果といった知覚上の現象との対応がとりやすくなると考えた。ま

た，そのような回答方法で得られた実験Ⅱの結果から，先行して出力された音源を選択した確率に基づいて，先行音効果が生じたとみなせる音源間の遅延時間を定量的に定めることができる．さらに，実験Ⅰとの比較を行うことで，先行音効果が生じているとみなせる2音源間の時間差において，音像はどの方向にあるか，そこに至るまでに音像の方向は時間差によってどのように変化するかが議論できる．実験Ⅱにおける被験者の回答内容は，実験Ⅰに比べて単純であるため，刺激音提示後にとられた回答時間は実験Ⅰよりも短くし2秒とした．

2.3 実験結果の評価

本論文では，先行音効果の現象を拡声システムへの応用を目指す見地から，音像定位の結果(実験Ⅰ)と，先行して音を出力する音源の方向を選択した確率(実験Ⅱ)との関係を，被験者ごとではなく，全ての被験者の実験結果を平均したものを通じて評価することとした．

2.3.1 絶対方向判断の実験結果

図2.2 (A)～(F)～図2.3 (A)～(F)は，2音源間の遅延時間と音像の関係を示したものである．横軸は2音源間の遅延時間を示し，縦軸は被験者の回答方向を示している．プロットした円の大きさは，その方向に回答した被験者の数に対応した大きさとなっている．図中の点線ラインは提示した2音源の方向を示している．例えば，図2.2 (A)を見ると，スピーカの開き角度が $\pm 30^\circ$ ，2音源の音圧レベル差が0 dB，2音源のレベル差が0 dBの場合，遅延時間が約0～0.3ms近辺では，おおよそ2音源の中央付近に音像が知覚されていることが分かる．2音源間の時間差が大きくなると音像が，先行音側によっていく傾向が見られる．時間差が十分に大きくなると音像の方向は，被験者の約 20° 方向に収

束している。時間差が無い場合に中央付近にあった音像が、時間差が大きっていくと約 20° 方向へ推移していくが、この変化は急激に生ずるわけではなく、時間差の増大に伴って徐々にシフトしていくことをこの図は示している。先行する音源方向への音像の移動状況は、左右どちらの音源が先行しているかによって、若干その傾向は異なっており、被験者から見て右側面の絶対判断のばらつきがやや大きい。しかし、その差異はそれほど大きくなく基本的な傾向は、ほぼ同様であるとみなせる。図2.3 (A)を見て分かるように、音源配置が 0° と 60° 方向となるセットBにおいても、同様な傾向が見て取れる。しかし、セットBの場合には、被験者の前方である 0° の音源が先行するとき、短い時間差で先行音側に変移する傾向が見られた。また、レベル差では、レベル差を増すと特にセットBにおいてばらつきが大きくなる傾向が見られた。

図2.2 (B)から(F)、図2.3 (B)から(F)についても、レベル差による音像のシフトが見られるものの、基本的な傾向は図2.2 (A)、図2.3 (A)と似ているといえる。まとめると、2音源間の時間差の設定が増加し、加法定位から先行音効果へと移り変わる過程において、音像は2音源の中央から先行音源の方向へと徐々に推移することが確認された。また、推移の際には、回答のばらつきがある程度見られるものの、ほぼ平均的な一塊の音像を為していることが分かった。そこで、続いて、被験者の回答結果を集計し、その平均値に基づいて議論を進める。

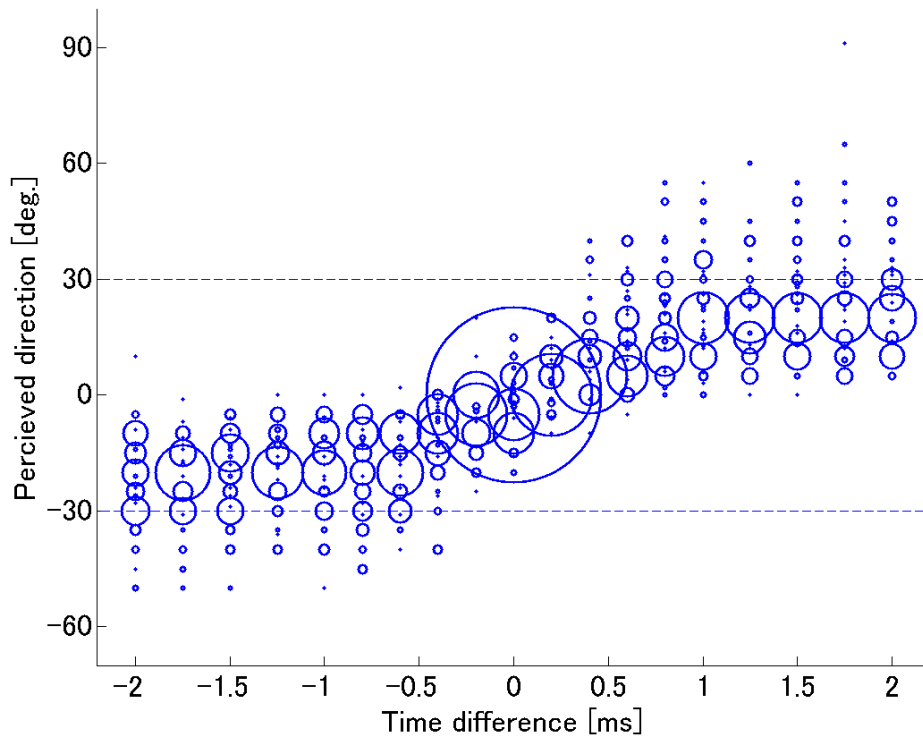
図2.4～2.7に、各被験者が回答した定位方向の平均値と2音源間の時間差についてプロットした図を、音源間に設定したレベル差ごとに示す。図2.4及び2.5は音源配置がセットAの結果を、図2.6及び2.7は音源配置がセットBの結果を表している。各図の軸の表記は図2.2及び2.3と同様であり、横軸の時間差は値が負であれば被験者からみて左側の音源からの音が先行し、正であれば右側の音源からの音が先行することを意味する。縦軸は、回答した知覚音像方向を示しており、その平均値とばらつきを、マーカと土標準偏差の誤差棒として示した。また、図2.4 (a)はレベル差が0 dB、図2.4 (b)はレベル差

が+3 dB, 図2.4 (c)はレベル差が+6 dBのものを表し, 各レベル差による, 時間差が0 ms以下と以上の回答が知覚音像方向に線で引かれている. レベル差を設けた実験では, 後続する音の方のレベルを高くしているため, 時間差が0msのときは, 左側の音源のレベルが大きい場合と, 右側の音源のレベルが大きい場合が存在する. そのため, 0 msの実験結果だけは, 値が二つ存在している.

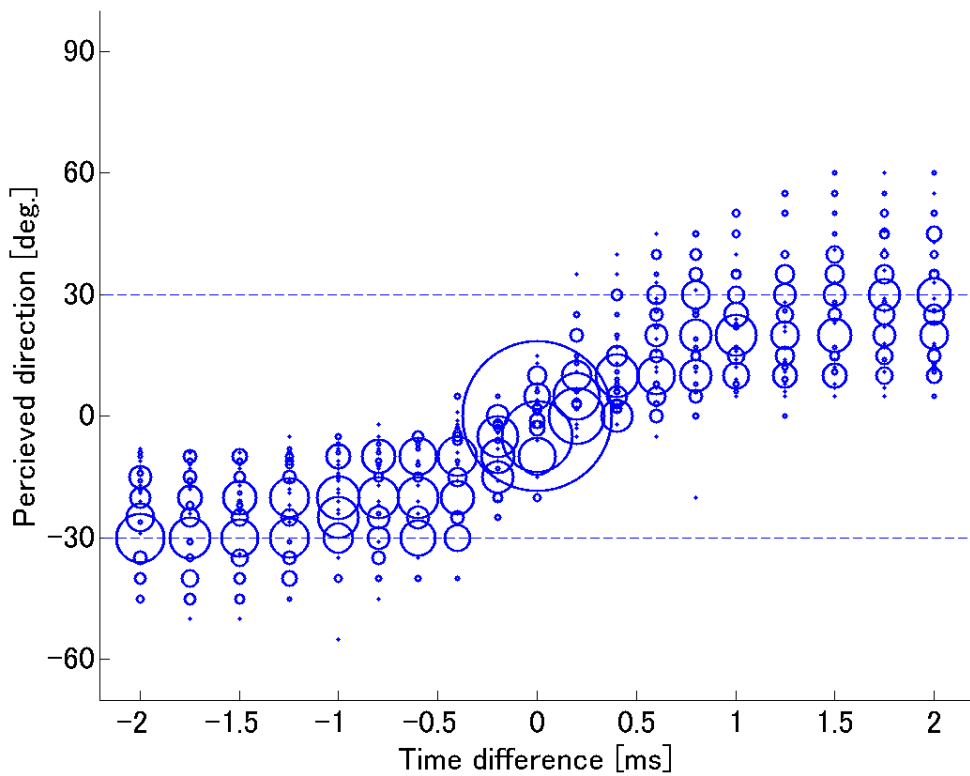
図2.4 (a), 2.5 (a)は音源配置がセットAで2音源のレベル差が0 dBの実験結果を表している. それぞれの図を見ると, 時間差が0 msのときの定位方向がほぼ0°となっており, 音源間の時間差の絶対値が大きくなるにつれて, 先行して音の出力される音源の方向に音像の定位方向が徐々に移動しているのが分かる. 刺激音がピンクノイズ (図2.4(a)) の場合, 時間差が ± 1.5 ms程度で, 設置した音源方向と同じ $\pm 30^\circ$ にほぼ収束している. それに対して, 刺激音がクリック (図2.5(a)) のときは, 定位方向は $\pm 30^\circ$ に到達せず, ± 1.25 ms程度以上の時間差で $\pm 20^\circ$ 強の値に収束している. この差については, 2.3.2 節で考察する. 次に, レベル差が+3 dBの図2.4 (b), 2.5 (b)を見ると, 時間差が0msのとき, 左側の音源が先行している場合には約 $+5^\circ$, 右側の音源が先行している場合には約 -5° 付近となっている. これは, 本実験では後続して出力される音源のレベルが高いため, 左側の音源が先行し, 右側のレベルが大きくなっている場合は定位方向が正の方向, つまり右側へ, 右側の音源が先行し, 左側のレベルが大きくなっている場合は負の方向, つまり左側へ移動する傾向があることを示している. 同様の傾向は, 図2.4 (c), 6 (c)でも見られ, 時間差による音像の移動は, +3 dBの場合よりも若干小さくなり, 時間差の絶対値が0 msから大きくなったときの音像の移動も小さくなっており, 回答のばらつきも比較的大きい.

音源配置がセットBの場合である図2.6, 2.7の結果を見ると, レベル差が0 dBの図2.6 (a), 2.7 (a)では, 時間差が0 msのときの定位方向が約 20° となっている. これは, 音像が2音源間の中間である 30° ではなく, やや正面寄りに定位されていることを示しており,

中央の方向を回答しているセットAの結果とは異なっている。一方、音源間の時間差の絶対値が大きくなるにつれて、先行して音出力される音源の方向に音像の定位方向が徐々に移動する傾向はセットAの場合と同様といえる。しかし、ピンクノイズ、クリックのいずれの刺激音についても、左側つまり 0° の音源が先行する時間差が負のときは、十分な時間差が設けられた場合、その定位方向は 0° 付近の近づくものの、右側つまり 60° の音源が先行する正の時間差のときには、定位方向は音源方向である 60° に収束せず、平均で見ても若干正面寄りとなっている。ピンクノイズの場合は、図2.6 (a)から、その方向は 50° 付近、クリックの場合は、図2.7 (a)から、 40° 付近となっている。また、セットAで、ピンクノイズ、レベル差が0 dB、時間差が0.8 msの場合、平均値を中心として $\pm 10^\circ$ 程度の回答のばらつきがあるが、セットBのピンクノイズ、レベル差が0 dB、時間差が0.8 msでは、平均値を中心として $\pm 20^\circ$ 程度の回答のばらつきが見られ、これを比べても、明らかにセットBで提示された刺激音の方の定位が困難であることが分かる。

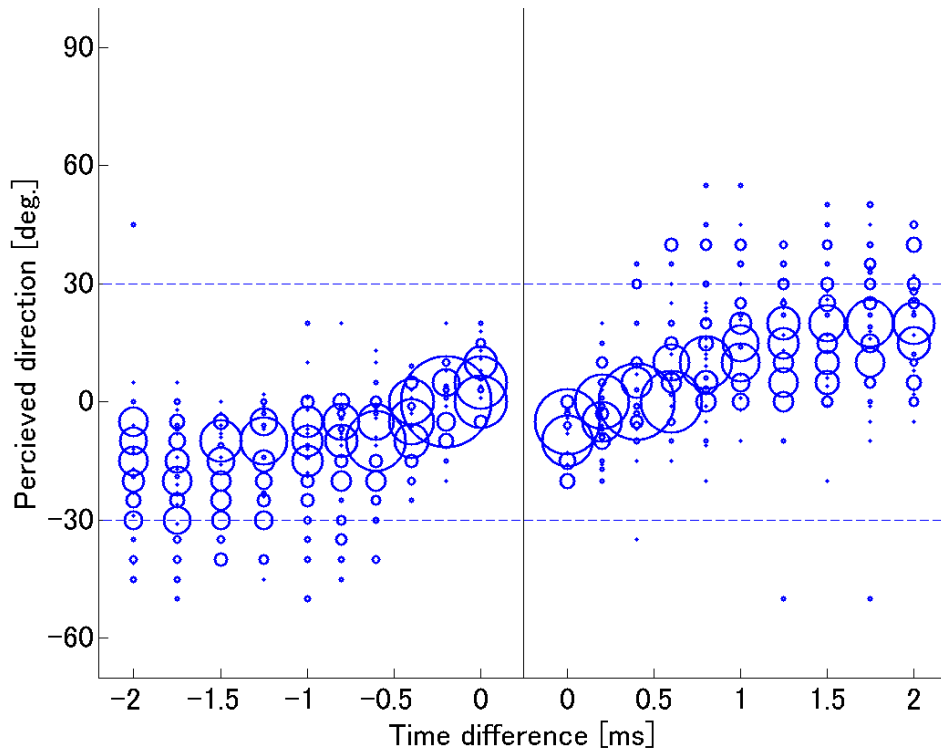


(A) 刺激音：クリック，2音源間のレベル差：0dB，音源方向： $\pm 30^\circ$ セット A

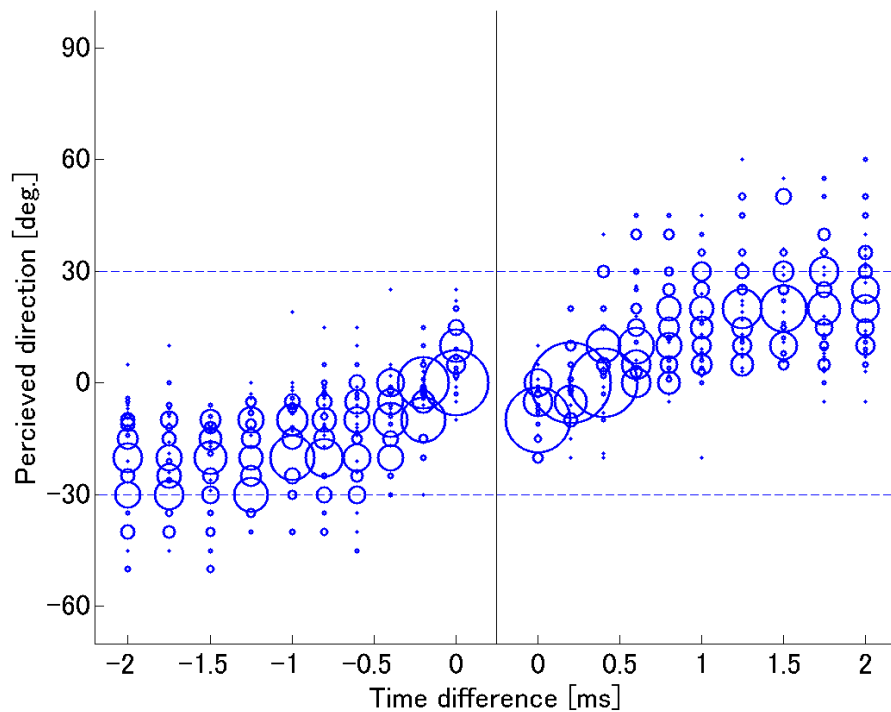


(B) 刺激音：ピンクノイズ 2音源間のレベル差：0dB 音源方向： $\pm 30^\circ$ セット A

図 2.2 2音源間の遅延時間に対する定位方向の変化

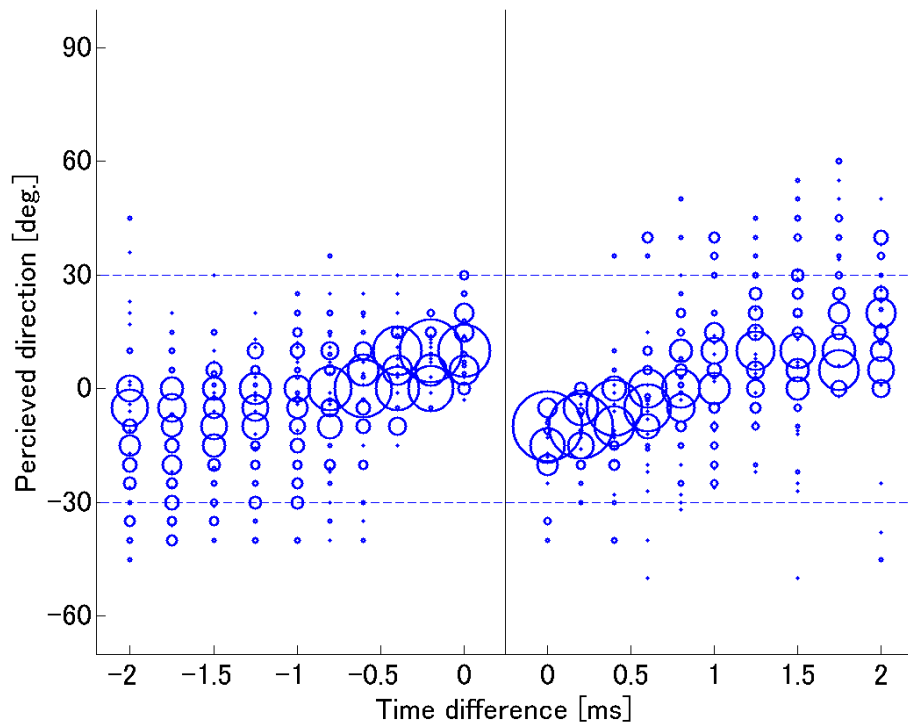


(C) 音源方向： $\pm 30^\circ$ セット A, 刺激音：クリック 2 音源のレベル差：+3dB

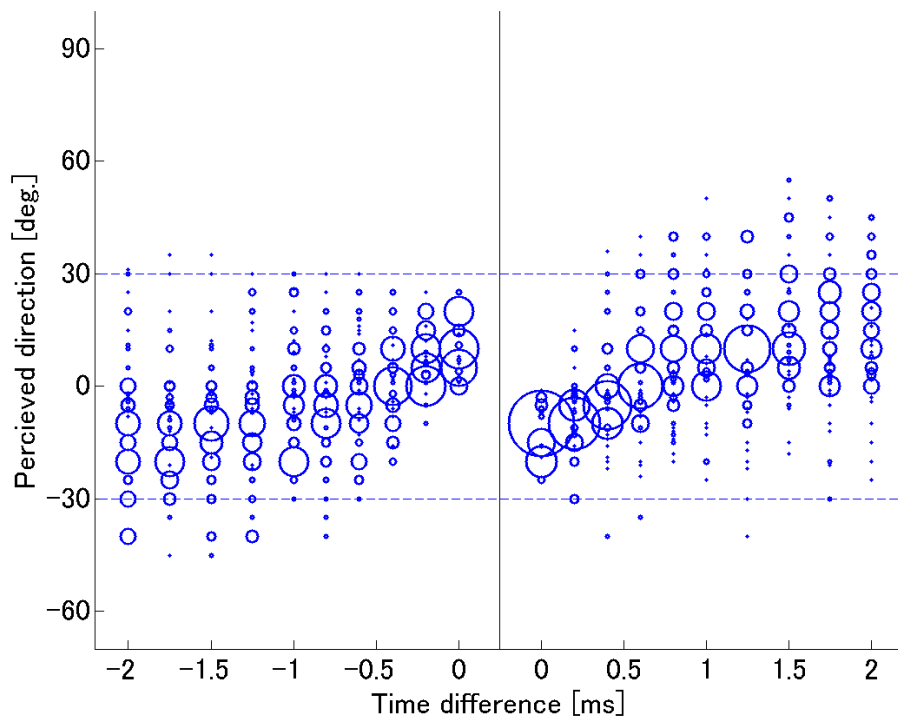


(D) 音源方向： $\pm 30^\circ$ セット A, 刺激音：ピンクノイズ, 2 音源間のレベル差：+3dB

図 2.2 (続き) 2 音源間の遅延時間に対する定位方向の変化

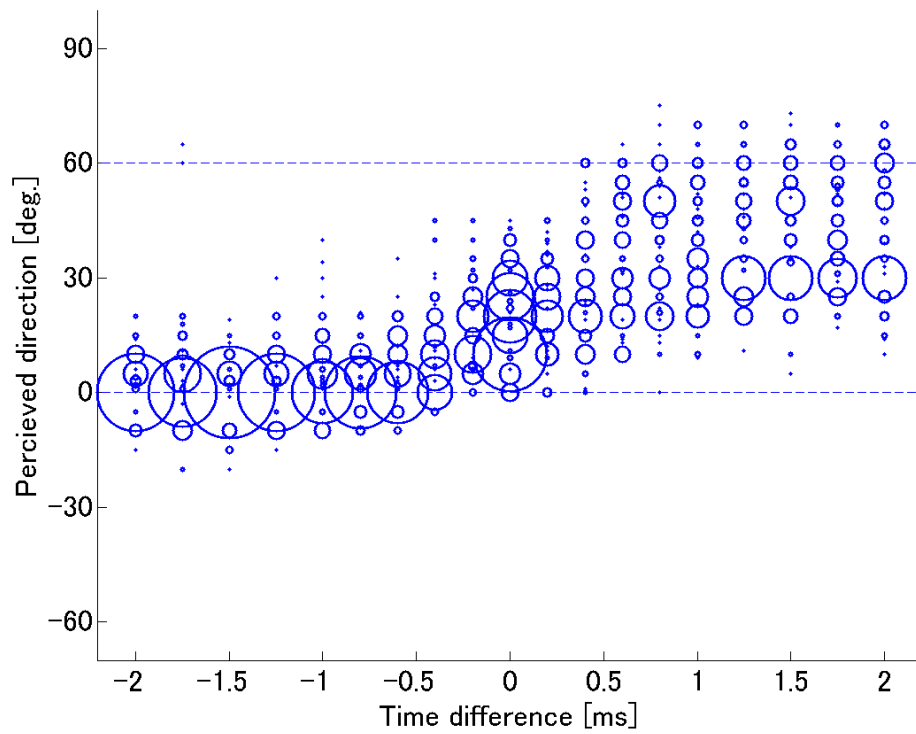


(E) 音源方向： $\pm 30^\circ$ セット A、刺激音：クリック、2 音源間のレベル差：+6dB

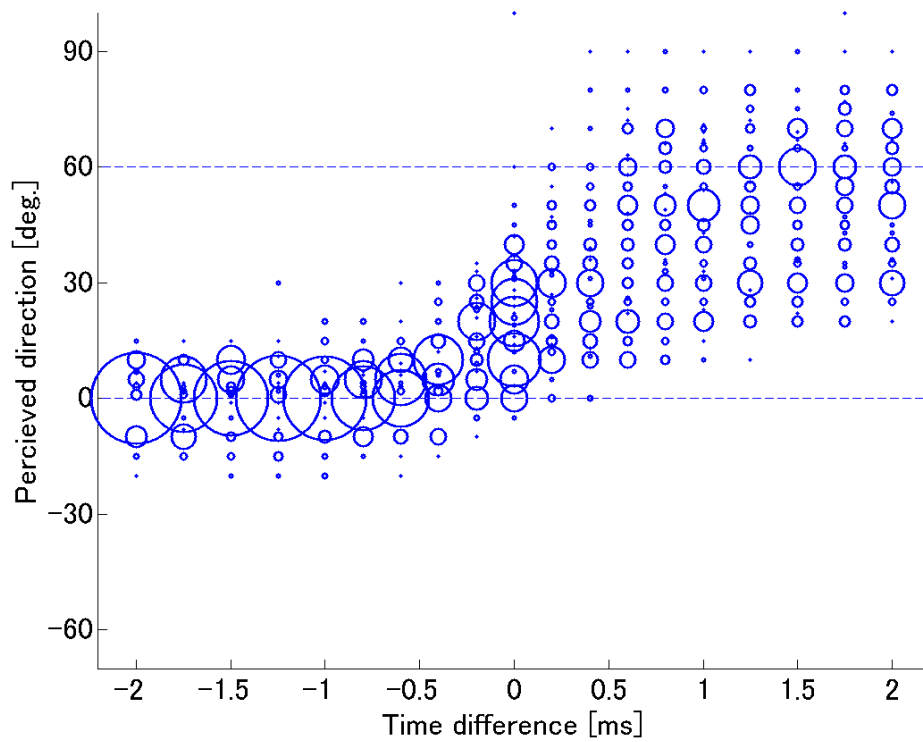


(F) 音源方向： $\pm 30^\circ$ セット A、刺激音：ピンクノイズ、2 音源間のレベル差：+6dB

図 2.2 (続き) 2 音源間の遅延時間に対する定位方向の変化

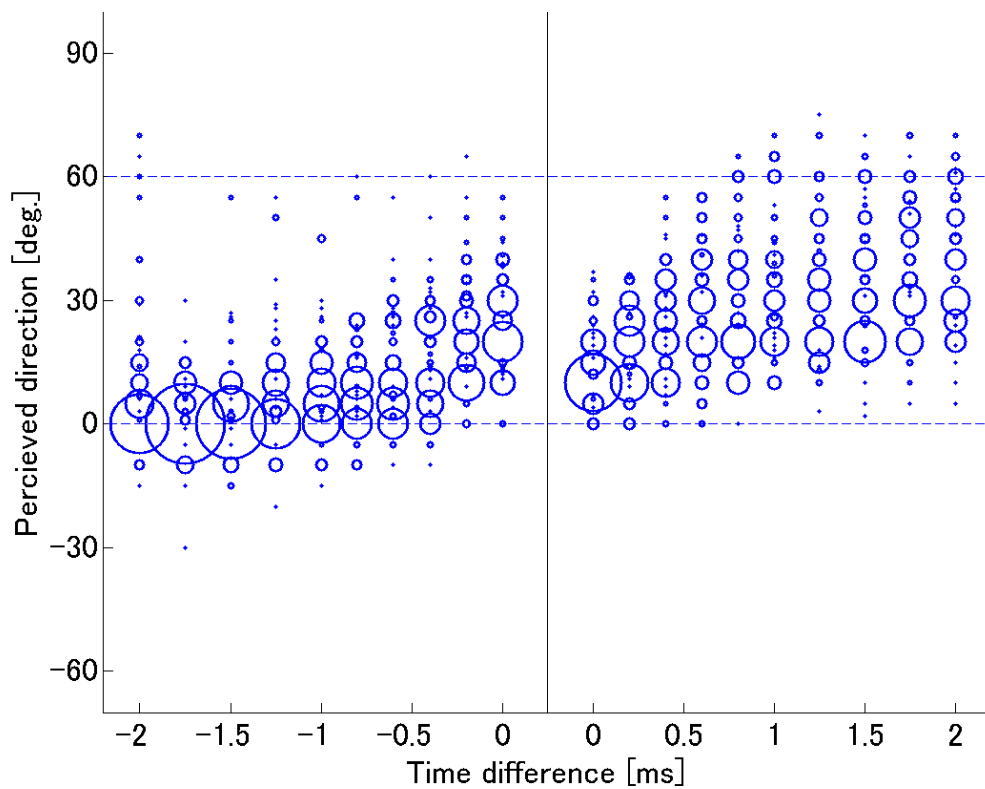


(A) 音源方向： $\pm 60^\circ$ セット B, 刺激音：クリック, 2 音源間のレベル差：0dB

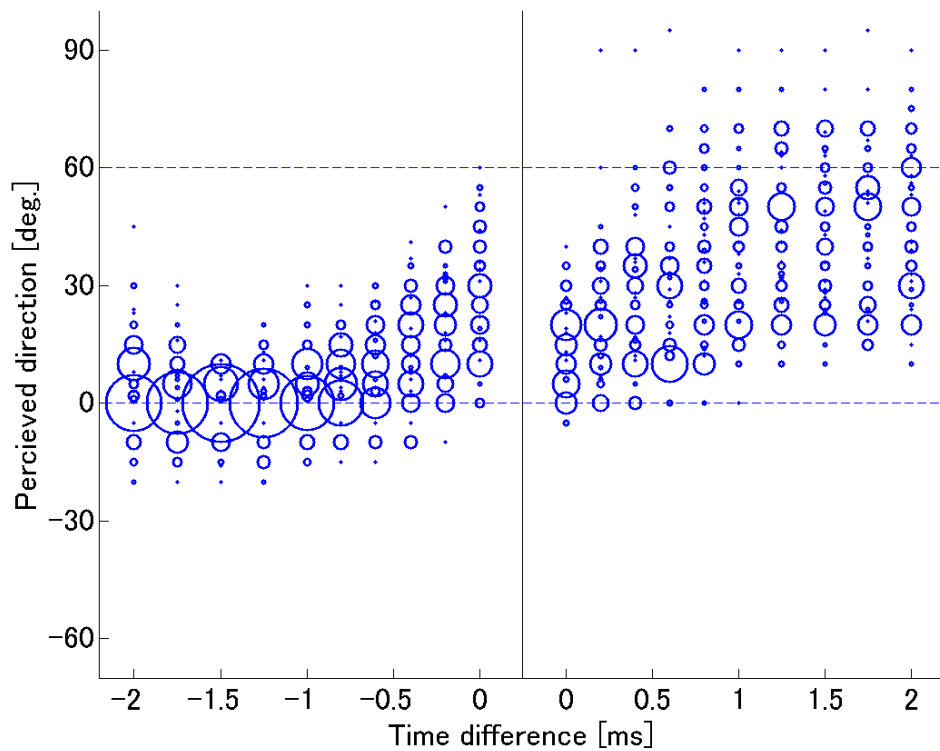


(B) 音源方向： $\pm 60^\circ$ セット B, 刺激音：ピンクノイズ, 2 音源間のレベル差：0dB

図 2.3 2 音源間の遅延時間に対する定位方向の変化

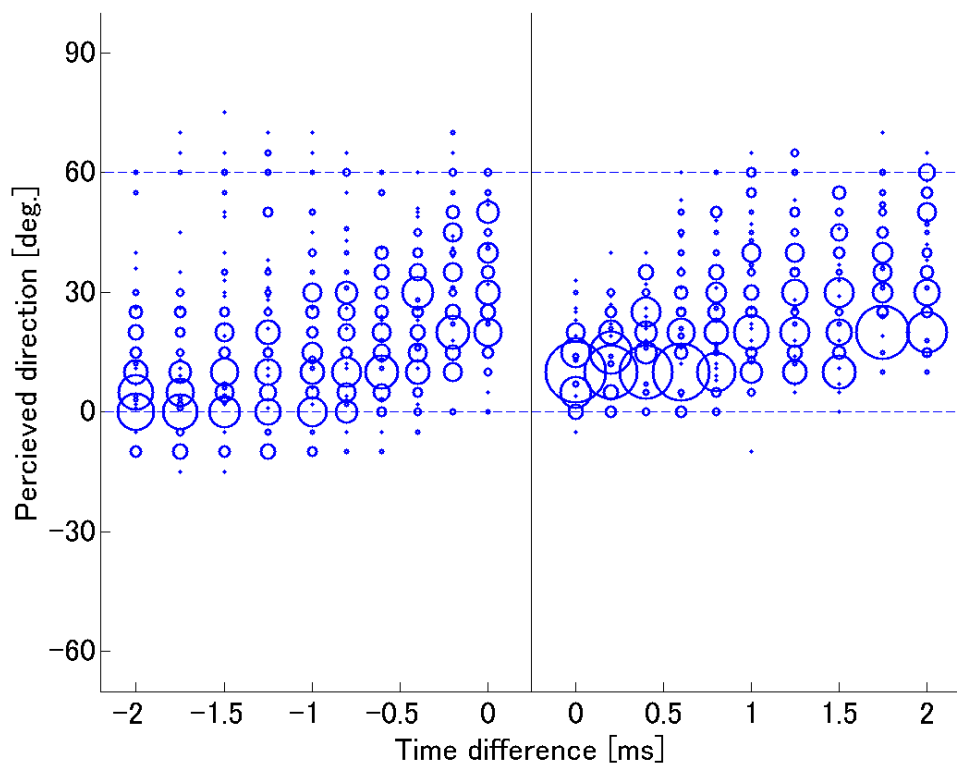


(C) 音源方向： $\pm 60^\circ$ セット B, 刺激音：クリック, 2 音源間のレベル差：+3dB

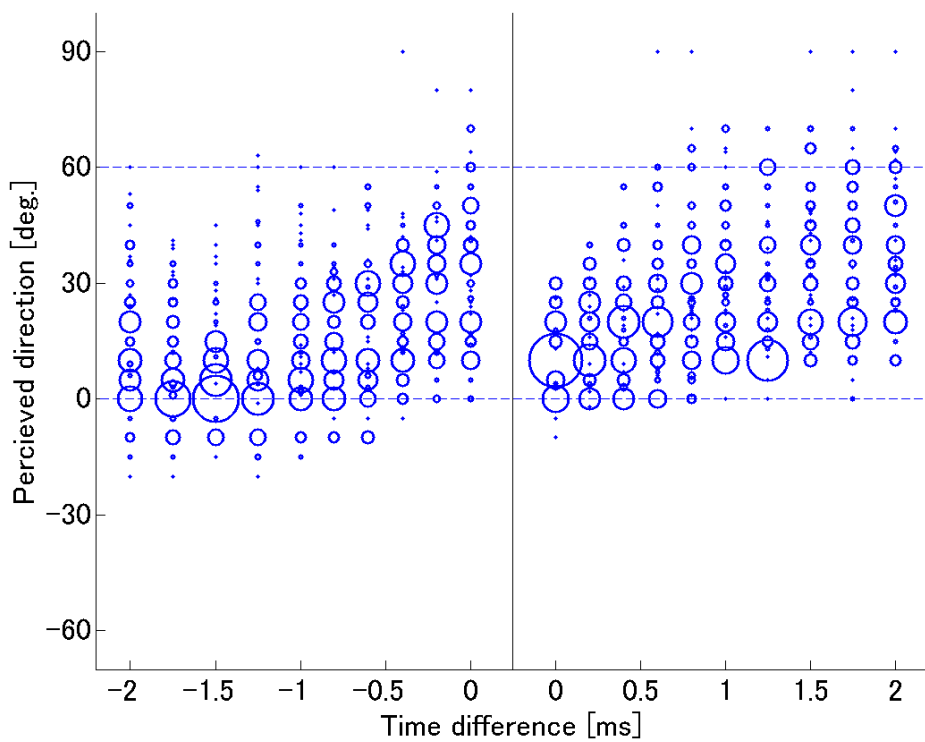


(D) 音源方向： $\pm 60^\circ$ セット B, 刺激音：ピンクノイズ, 2 音源間のレベル差：+3dB

図 2.3 (続き) 2 音源間の遅延時間に対する定位方向の変化

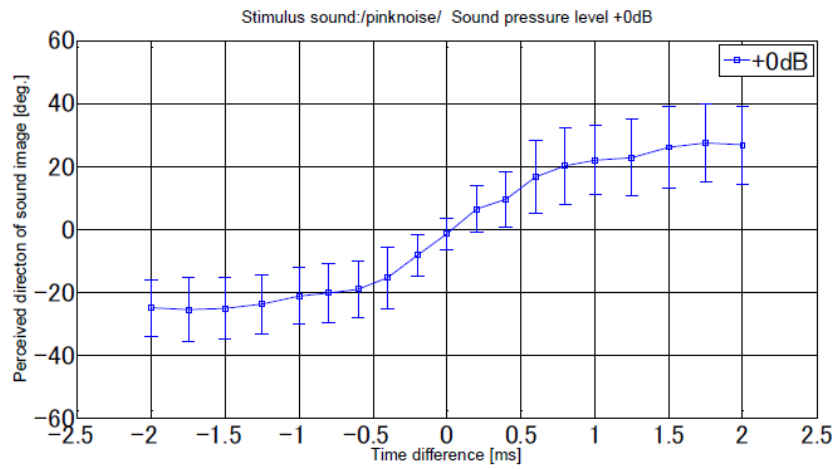


(E) 音源方向： $\pm 60^\circ$ セットB，刺激音：クリック，2音源間のレベル差：+6dB

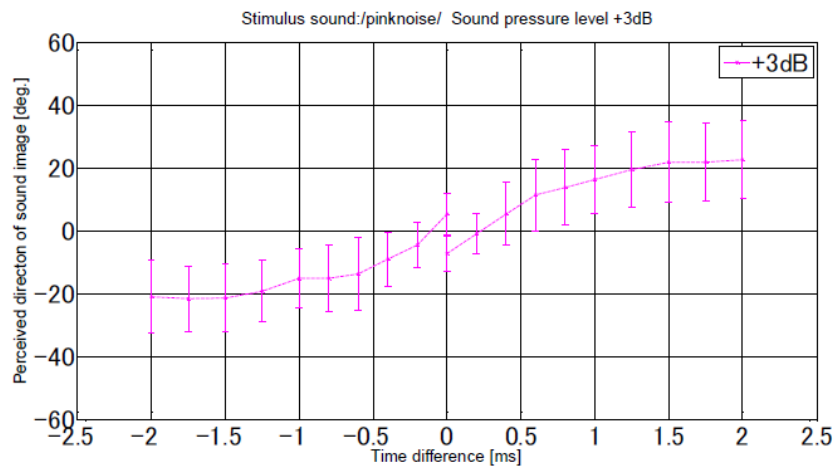


(F) 音源方向： $\pm 60^\circ$ セットB，刺激音：ピンクノイズ，2音源間のレベル差：+6dB

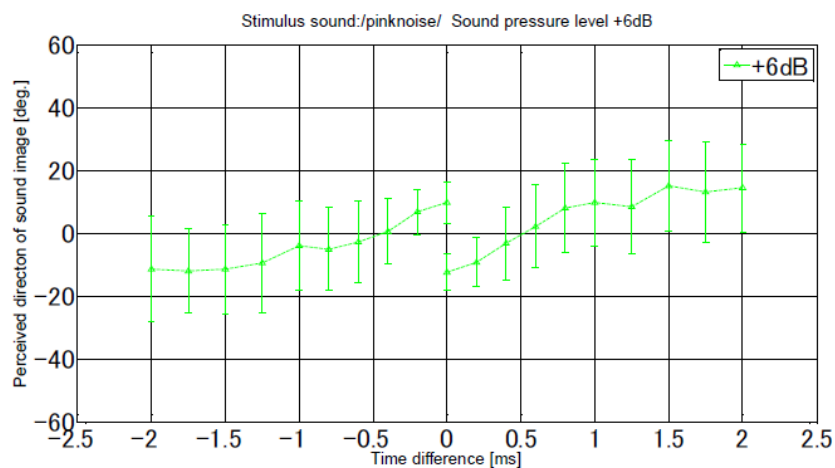
図 2.3（続き） 2音源間の遅延時間に対する定位方向の変化



(a) Level difference: 0 dB



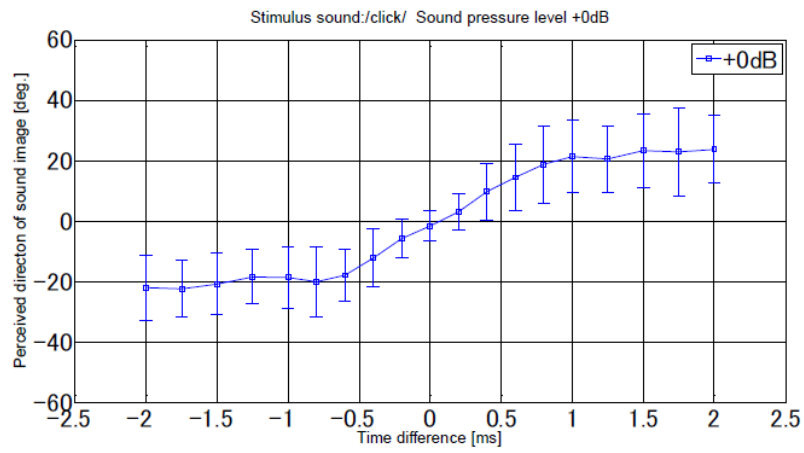
(b) Level difference: +3 dB



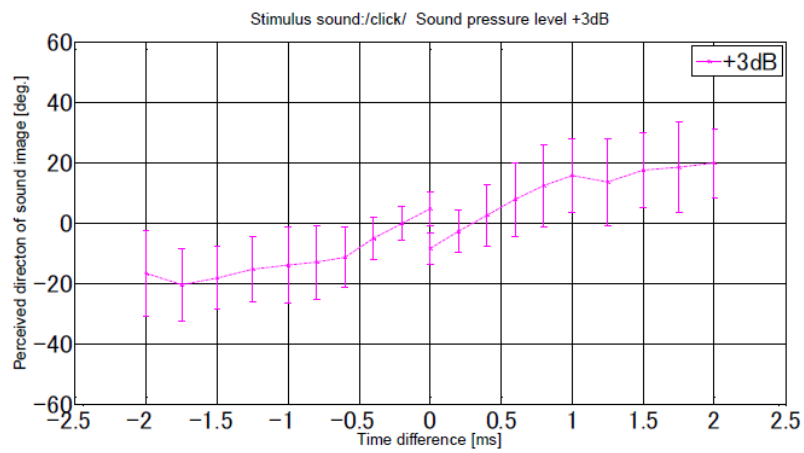
(c) Level difference: +6 dB

図 2.4 時間差による定位方向の変化

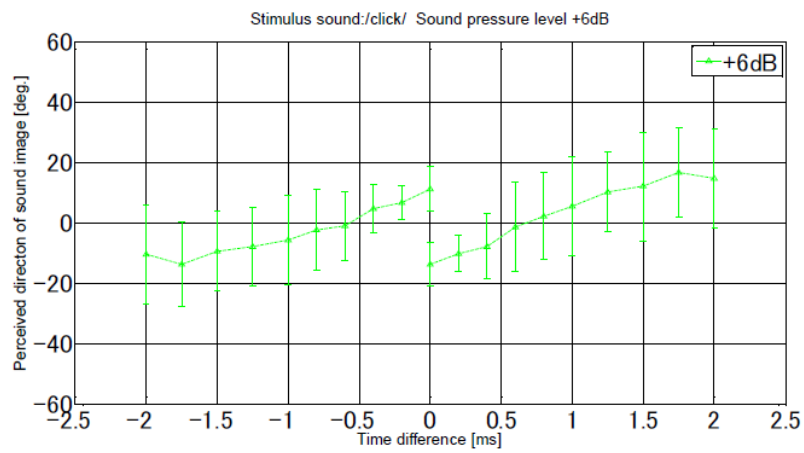
(刺激音：ピンクノイズ，音源配置：セット A)



(a) Level difference: 0 dB



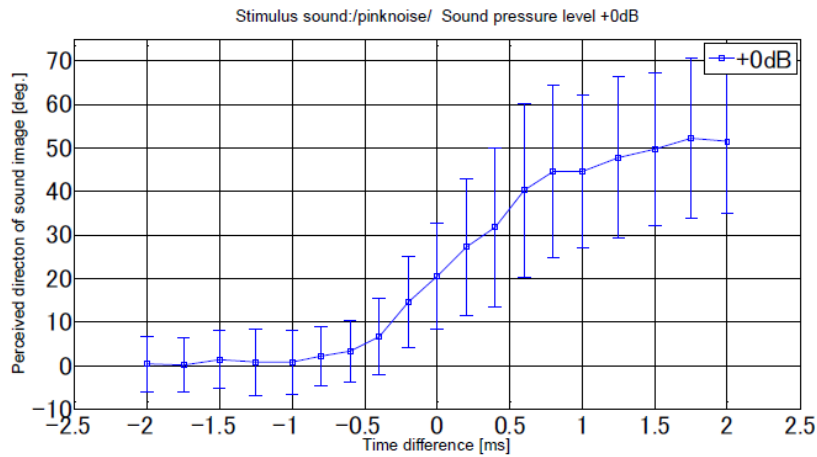
(b) Level difference: +3 dB



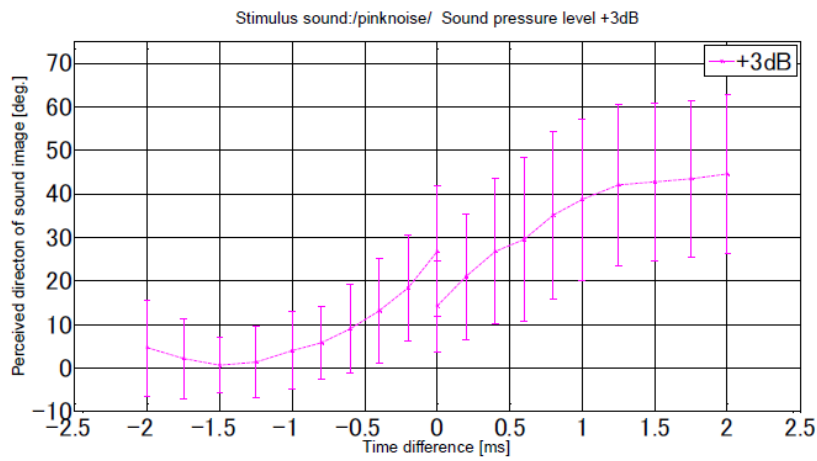
(c) Level difference: +6 dB

図 2.5 時間差による定位方向の変化

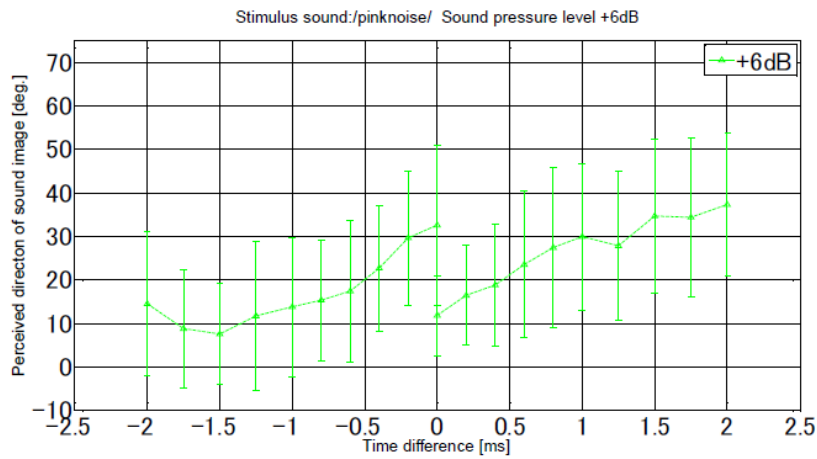
(刺激音：クリック，音源配置：セット A)



(a) Level difference: 0 dB



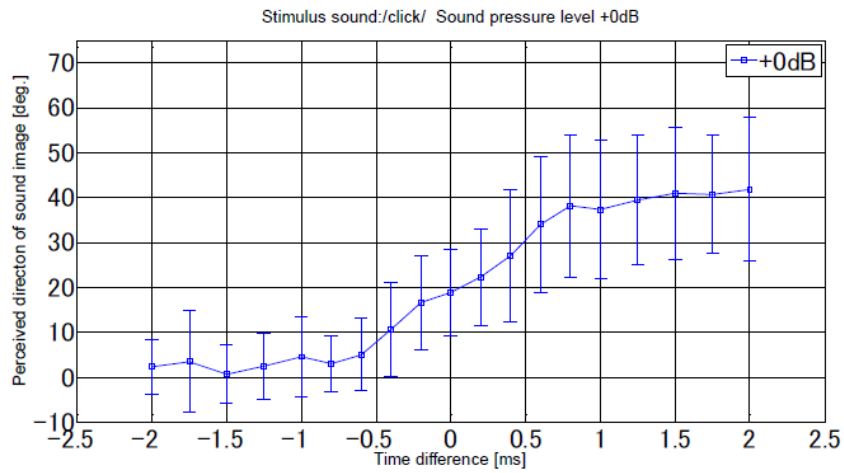
(b) Level difference: +3 dB



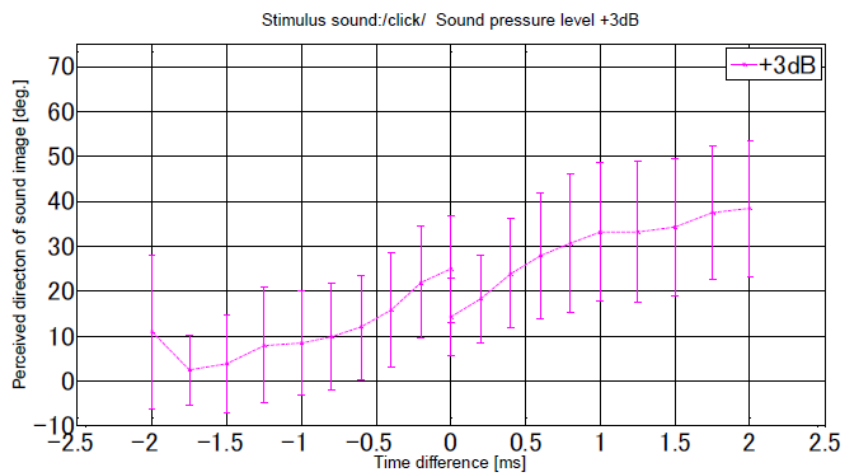
(c) Level difference: +6 dB

図 2.6 時間差による定位方向の変化

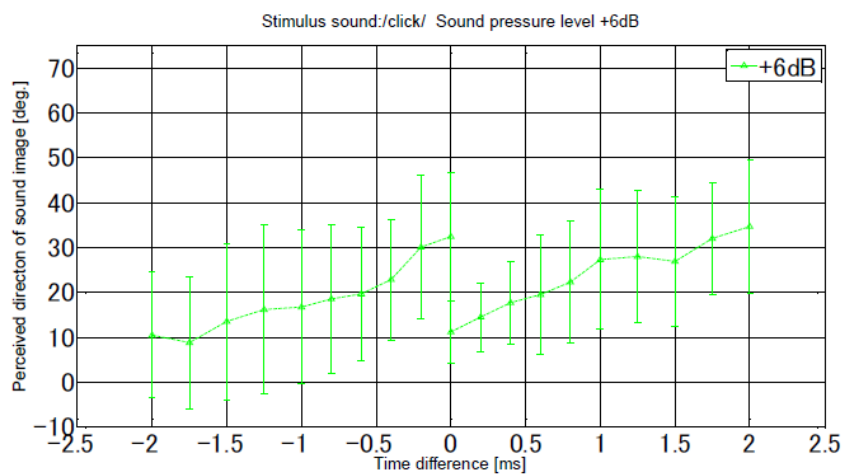
(刺激音：ピンクノイズ，音源配置：セット B)



(a) Level difference: 0 dB



(b) Level difference: +3 dB



(c) Level difference: +6 dB

図 2.7 時間差による定位方向の変化

(刺激音：クリック，音源配置：セット B)

2.3.2 考察

音源配置がセットAで、レベル差が0 dB、時間差が0 msの場合、定位方向がほぼ0°となっており、音源間の時間差の絶対値が大きくなるにつれて、先行音側に音像の定位方向が徐々に移動している。刺激音がピンクノイズの場合、時間差が±1.5 ms程度で、設置した音源方向と同じ±30°にほぼ収束する。それに対して、刺激音がクリックの場合、定位方向は±30°に到達せず、±1.25 ms程度以上の時間差で±20° 強の値に収束している。この差について、考察する。

Leakey ³²⁾らは、2 音源間における時間差やレベル差と音像方向の関係を 2 つの音源信号について調べている。刺激音がインパルス、スピーチの場合、被験者前方の約±20°の範囲では、2 音源間の時間差やレベル差による音像方向の変化との関係は直線的（比例関係）であることを示している。本実験 I の結果は、Leakey らの結果とその点ではほぼ同様の傾向といえる。ただし、本実験 I の結果を見ると、各図に書かれている誤差棒の幅は、時間差やレベル差による定位方向の変化と比べて決して小さくないため、2 音源間の時間差による定位方向の有意な変化を定量的に議論するには、限界がある。2 音源間の時間差、レベル差を変化させたときの合成音像の定位方向のばらつきがある程度大きいことは、音像に広がりがあることや、回答として広がった音像のどこを定位方向として答えるかが影響していると推定される ^{33, 34)}。セット A, B の音源配置両方の場合で、時間差が正のときに、クリックの方がピンクノイズよりも定位方向の平均値が内側に寄っているように見えるのは、合成音像の定位方向が安定するのにある程度の継続時間が必要となる可能性を示唆している。また、セット A とセット B の音源配置の比較では、セット B の場合、角度の弁別精度が悪くなることが知られているように、単純に側方になるほど合成音像の定位判断が難しくなり、側方のばらつきが大きくなったものと考えられる。

2.3.3 3肢強制選択の実験結果

実験Ⅰと同条件で提示された刺激に対して、仮想音源を中間に置いた三つの音源に対して、知覚した音像の定位方向が最も近い音源を選択させた。このことで、先行して出力された音源の方向を選択した確率が、それ以外の音源方向に対して有意に高いとみなせる時間差を得ることができる。

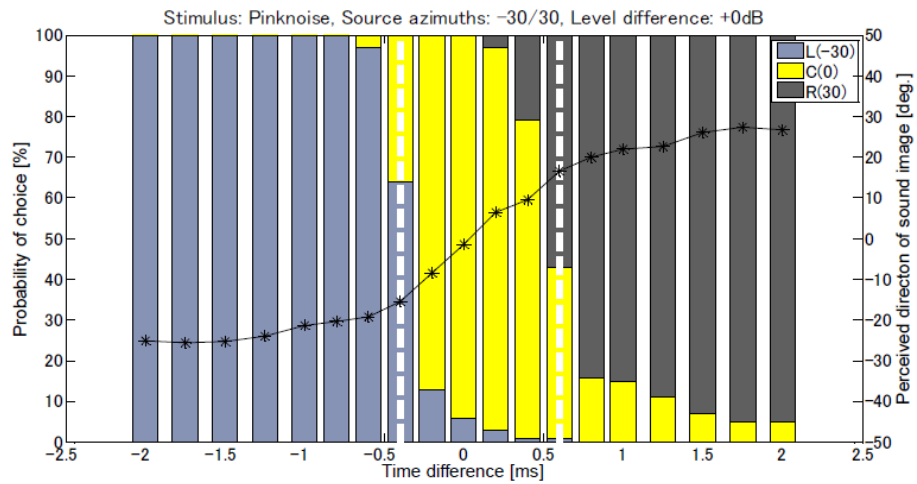
実験Ⅱでは、それぞれの音源が選択された割合に基づき、 χ^2 検定を用いて危険率5%でその音源が有意に多く選択される時間的境界を求める。本論文では、10名の被験者全員の回答結果をまとめて取り扱う。その結果、先行して出力される音源を選択する確率が、他の音源を選択する確率に比べて有意に大きいとみなせる2音源の時間差の条件を決定し、その後、その絶対値が最小となるの時間差を、本論文では先行音効果が生じた時間差（SLT）とみなすこととする。この基準で決定される時間差は、もし2音源からその時間差を持った音が提示されれば、知覚される音像が先行して音が出力される音源に最も近いと有意に判断される時間差であり、チャンスレベルよりも厳しい基準と言える。2音源間の時間差による判断確率の変化を、実験Ⅰにおける定位方向の平均値の変化とともに、図2.8～2.11に示す。2.3.1節で述べられたこと同様の理由で、各図の(b), (c)では時間差0 msのときの実験結果が二つ存在するため、時間差がマイナスの場合とプラスの場合でプロットが分けられている。実験Ⅰにおける音像方向の回答結果から考えると、時間差がマイナスのときは、被験者からみて左側の音源を選択する確率が高く、0 ms付近に近づくにしたがって、2音源の中間にある音源を選択する確率が高くなると思われる。同様に時間差がプラスの方向へ変化すると、被験者からみて右側の音源を選択する確率が高くなると推測される。図2.8～2.11を見ると、上記のような傾向が、全ての図において共通して見られる。また同じ音源配置であっても、両音源間の提示音圧レベル差が大きくなるにつれて、2音源の中間の音源が選択される時間差の範囲が広がっている。図2.8と図2.10、図2.9と図2.11を比べると、どちらの刺激音についても、

時間差による回答確率の変化の傾向については、音源配置による目立った違いは見られない。その一方で、両方の音源配置に共通して、どのレベル差においても、被験者から見て左側の音源を選択する確率が右側を選択する確率よりも高いように見える。各図から、前述の基準で有意に先行音源側に寄ったとみなせるSLTを決定した。また、そのSLTにおける実験Ⅰの音像方向を読み取ることによって、SLTにおける音像の定位方向の平均値を求めた。それらをまとめたものが表2.1, 2.2となる。SLTはそれぞれの音源が先行する場合に対して定まるので、SLTとなる時間差は、正と負の二つ求められる。実験では、刺激音が2種類、レベル差の条件数が3なので、ある音源配置につき、2種類×3条件×2つ=12の時間差が求められる。音源配置はセットA, Bの2種類なので、総数は24となる。これらの表を見ると、音源配置がセットAのときは、レベル差が+3 dB, +6 dBでいずれの刺激音でもプラス側の境界は求められた。しかし、音源配置がセットBのときは、レベル差が+6 dBの場合、刺激音がピンクノイズにおいては、短い時間差においても右側を選択する確率が高く境界となる時間差は求められなかった。

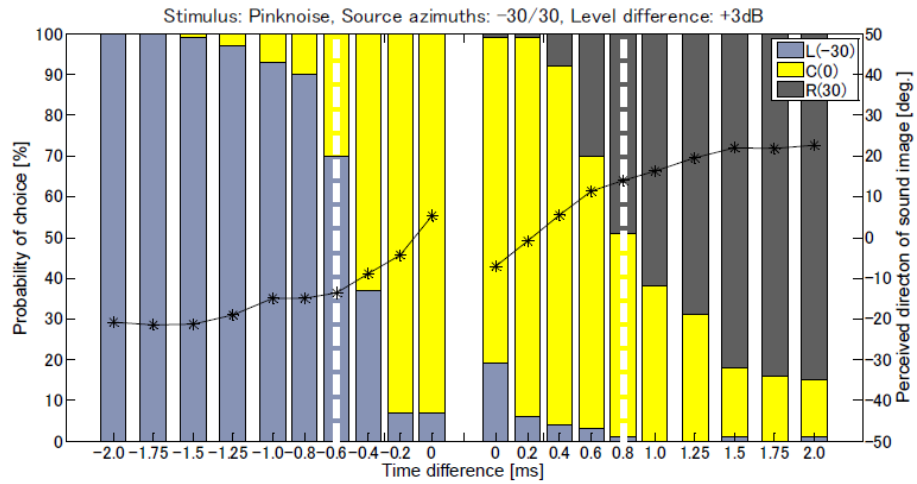
2.3.4 考察

Blauert²⁾は、加法定位から先行音効果に変化する2音源間の時間差について、標準的な刺激音でレベル差が0 dBであれば0.63~1 msであると述べている。表2.1及び2.2を見ると、これらの研究で求められた時間差に比べて同等か小さな値となった。また、レベル差が+3 dB, +6 dBのとき、SLTの時間差が求められた場合には、上記の研究と同程度となっている。求められた条件のみにおいて、SLTの2音源間レベル差による変化を見ると、レベル差の増加にともなって、SLTの絶対値は大きくなるように見える。このことから、本実験においてSLTを求められなかった音源配置、刺激音およびレベル差の条件については、SLTの絶対値は、実験で設定された2音源の時間差の範囲を超えて大きくなっていると考えられる。音源配置による違いを見ると、セットAの結果の表2.1

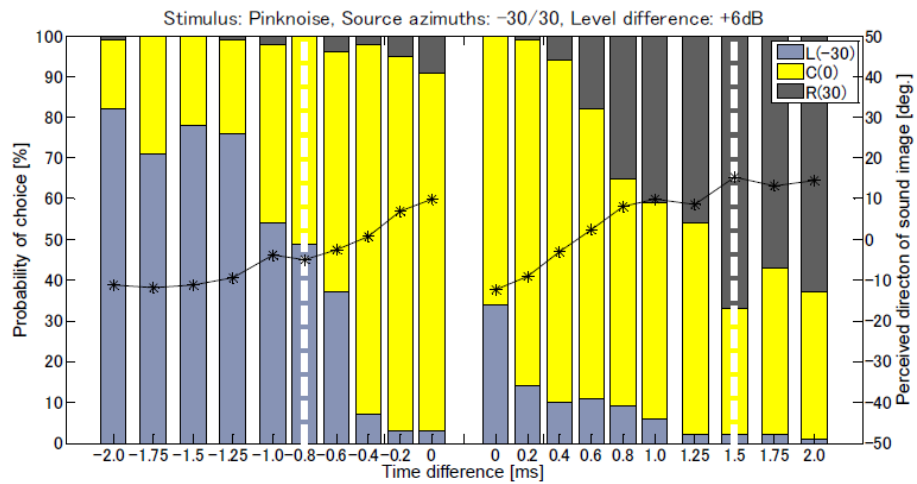
では、レベル差が0 dBのとき、SLTの絶対値は、マイナス側とプラス側で大きな差はないといえる。また、刺激音の種類によって、その値は大きく変化しない。しかし、レベル差が+3 dB, +6 dBのときには、SLTの絶対値は、マイナス側とプラス側で1ms以上の差がある。これは、実験Ⅱの結果において、被験者からみて左側の音源を選択する確率が右側を選択する確率よりも高いように見えることと関係があると考えられる。しかし、セットAにおいて2音源間のレベル差が+3 dB, +6 dBのときに、2音源間の時間差がマイナス側、プラス側の両方でSLTが求められたが、SLTの絶対値のもつ差の範囲が広くなるようにレベル差の影響を受けたものと考えられるが、このことについての議論は困難である。セットBの結果の表2.2も同様に、SLTの時間差の大きさは、マイナス側とプラス側で大きく異なっている。これは、2音源のうち一つが正面にある場合、短い時間差でその正面の音源に寄ったと判断しやすい傾向があることを表しているものと考えられる。逆に、音源が正面から方向的に離れている場合には、合成音像はそちらに寄りづらいと見ることもできる。これは、実験Ⅰの結果との整合性を考えると合理的と考えられる。



(a) Level difference: 0 dB



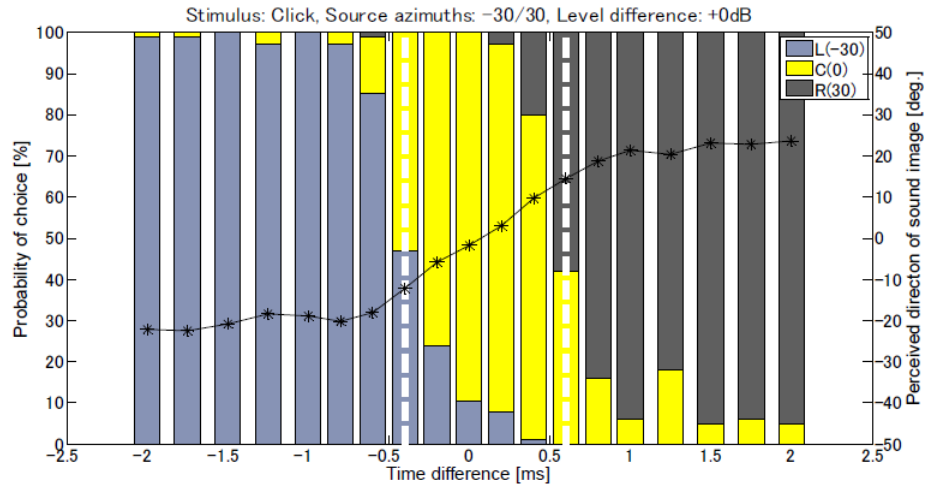
(b) Level difference: +3 dB



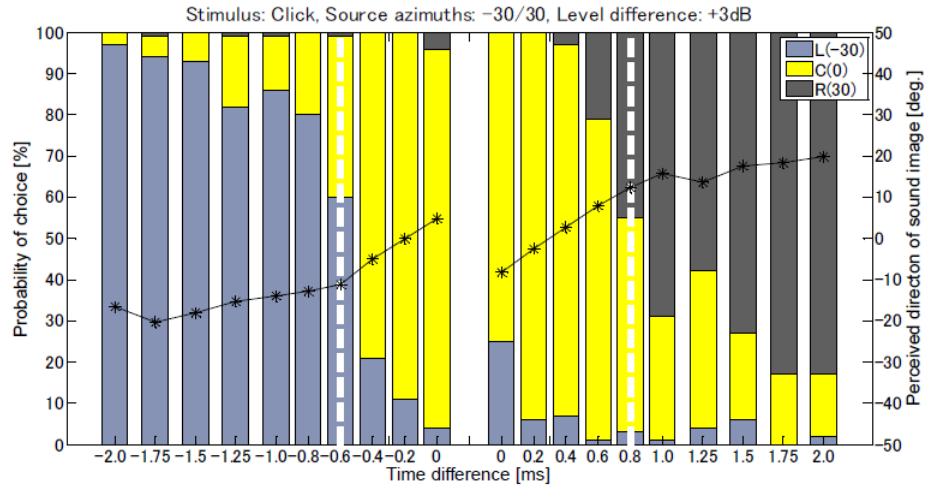
(c) Level difference: +6dB

図 2.8 時間差による回答確率の変化

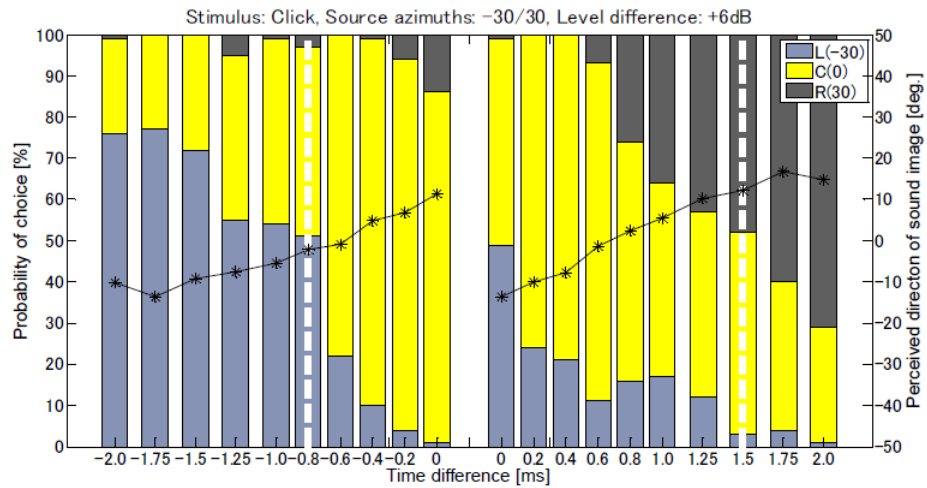
(刺激音：ピンクノイズ，音源配置：セット A)



(a) Level difference: 0 dB



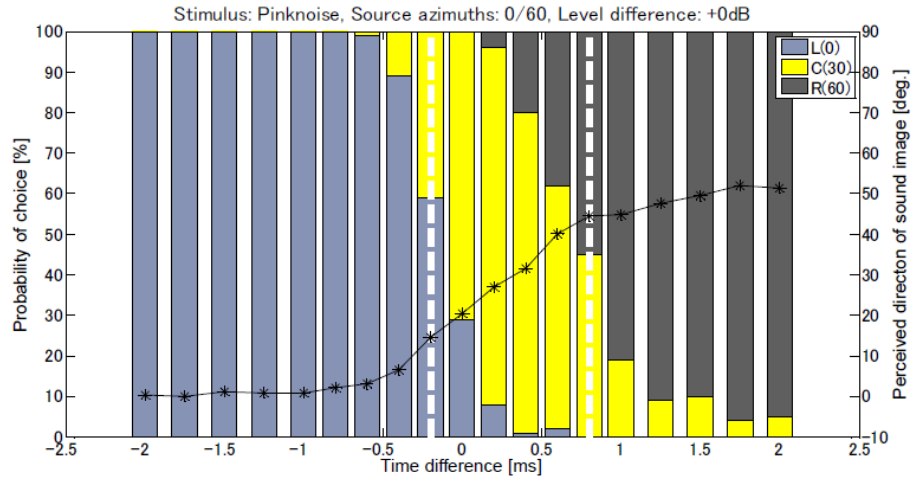
(b) Level difference: $+3\text{dB}$



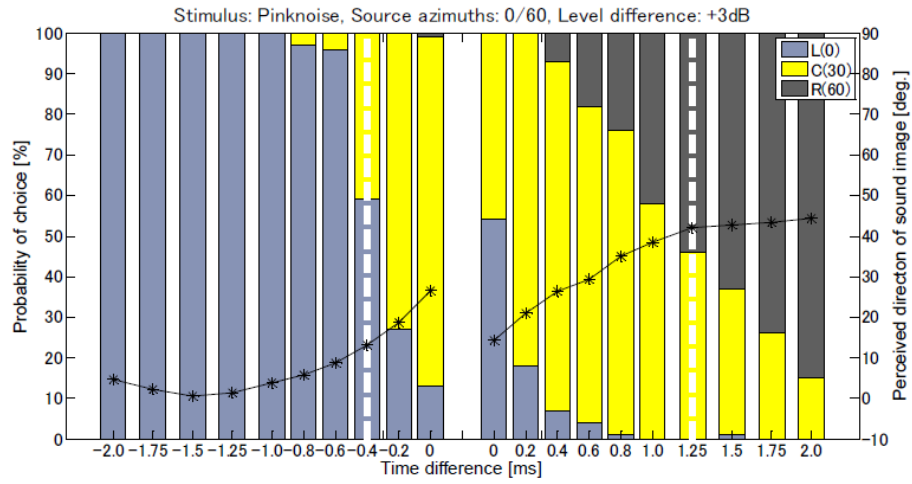
(c) Level difference: $+6\text{ dB}$

図 2.9 時間差による回答確率の変化

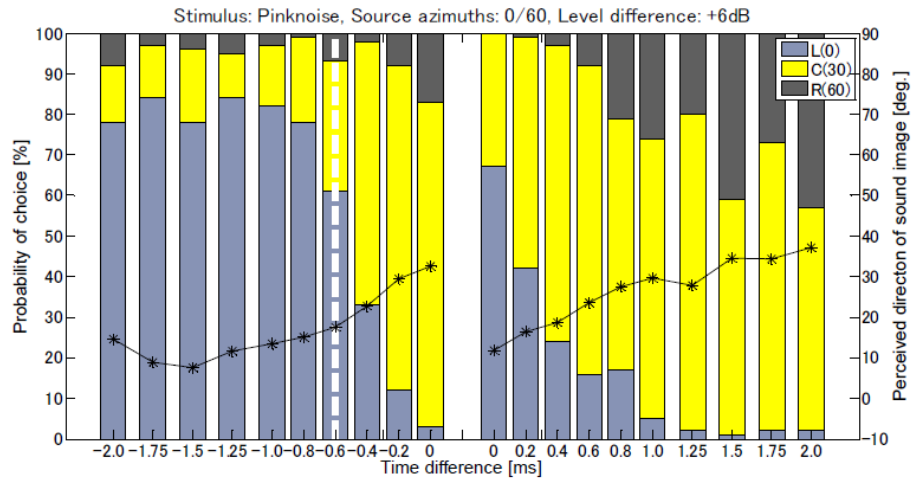
(刺激音：クリック，音源配置：セット A)



(a) Level difference: 0 dB



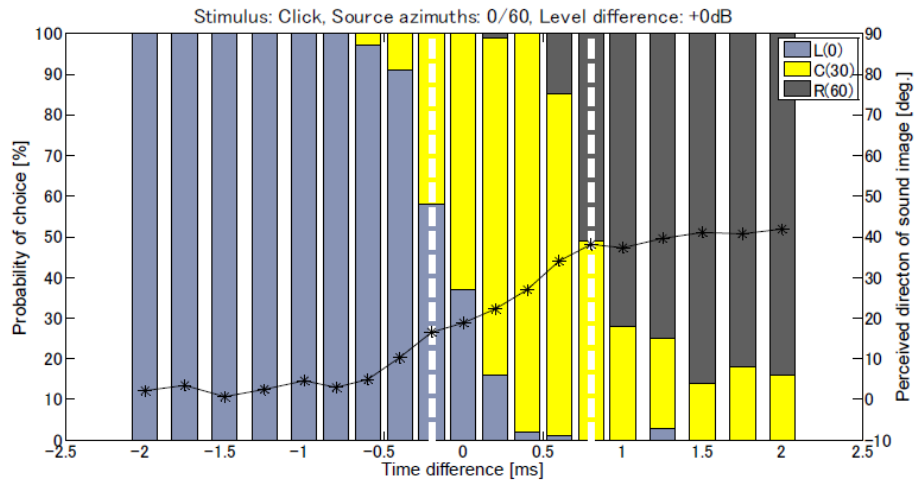
(b) Level difference: +3 dB



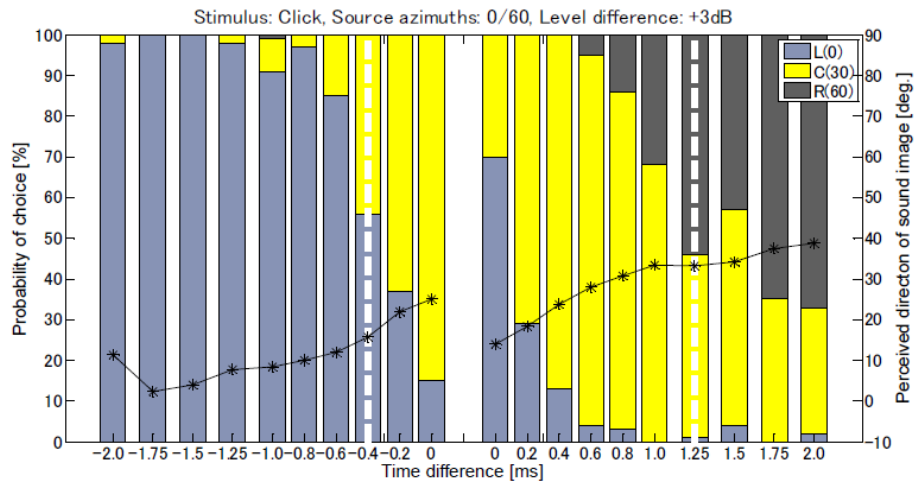
(c) Level difference: +6dB

図 2.10 時間差による回答確率の変化

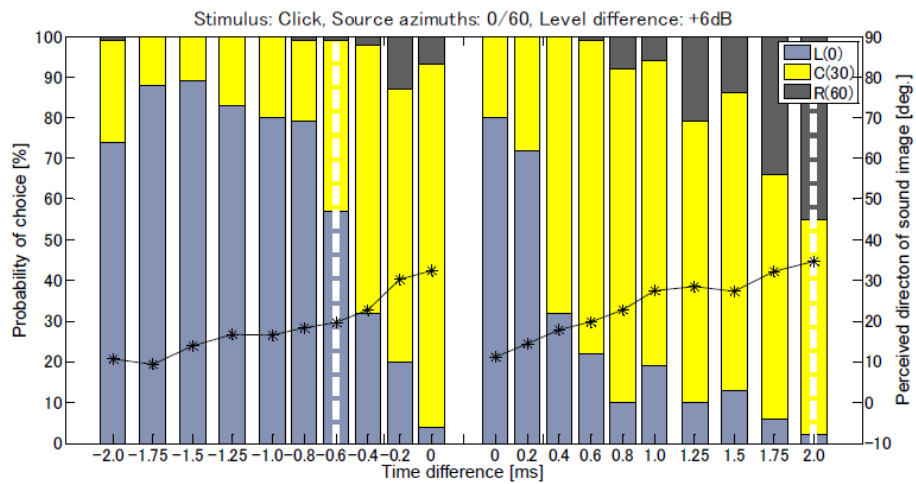
(刺激音：ピンクノイズ，音源配置：セット B)



(a) Level difference: 0 dB



(b) Level difference: +3dB



(c) Level difference: +6 dB

図 2.11 時間差による回答確率の変化

(刺激音：クリック，音源配置：セット B)

表2.1 先行音効果の生じる境界(SLT)の時間差と定位方向の平均値
(音源配置：セットA)

Table 1: Boundary of time difference at which the perceived sound image is significantly moved to the leading sound source (calculated from the results of Exp. II) and the average perceived direction (calculated from the results of Exp. I), sound source location: Set A.

Level difference	Stimulus			
	Pink noise		Click	
0dB	Time difference[ms]	−0.4, +0.6	Time difference[ms]	−0.4, +0.6
	Perceived direction[°]	−15, +17	Perceived direction[°]	−12, +14
+3dB	Time difference[ms]	−0.6, +0.8	Time difference[ms]	−0.6, +0.8
	Perceived direction[°]	−14, +14	Perceived direction[°]	−11, +12
+6dB	Time difference[ms]	−0.8, +1.5	Time difference[ms]	−0.8, +1.5
	Perceived direction[°]	−5, +15	Perceived direction[°]	−2, +12

表2.2 先行音効果の生じる境界(SLT)の時間差と定位方向の平均値
(音源配置：セットB)

Table 2: Boundary of time difference at which the perceived sound image significantly is moved to the leading sound source (calculated from the results of Exp. II) and the average perceived direction (calculated from the results of Exp. I), sound source location: Set B.

Level difference	Stimulus			
	Pink noise		Click	
0dB	Time difference[ms]	−0.2, +0.8	Time difference[ms]	−0.2, +0.8
	Perceived direction[°]	+15, +45	Perceived direction[°]	+17, +38
+3dB	Time difference[ms]	−0.4, +1.25	Time difference[ms]	−0.4, +1.25
	Perceived direction[°]	+13, +42	Perceived direction[°]	+16, +33
+6dB	Time difference[ms]	−0.6, —	Time difference[ms]	−0.6, +2.0
	Perceived direction[°]	+18, —	Perceived direction[°]	+20, +35

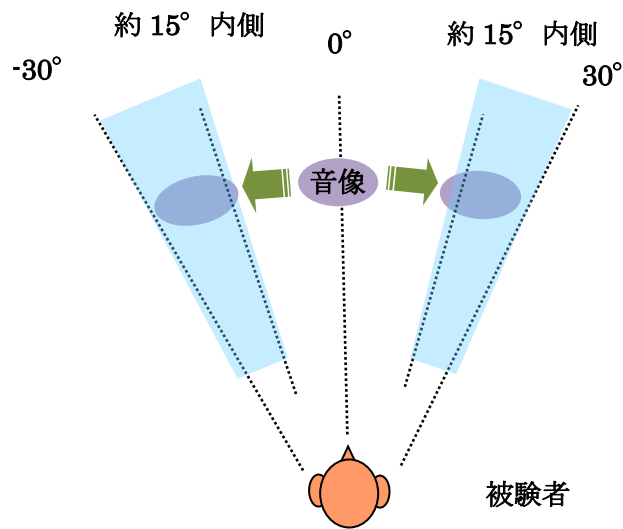
2.4 実験Ⅰ,Ⅱの結果の比較

図2.8～2.11の各パネルには、実験Ⅰにおいて対応する条件で定位された方向の平均値と一緒にプロットされている。また、表2.1, 2.2では、SLTの時間差と、その時間差において実験Ⅰで得られた定位方向の平均値と一緒に載せている。これらから、SLTである時間差において音像の定位方向がどこにあるか、すなわち、先行して信号が出力される音源の方向に合成音像が定位したと見なせる合成音像の定位方向の範囲を決定することができる。表2.1, 2.2には、その範囲をSLTとなる時間差とともに載せている。その範囲について、音源配置ごとに考察する。表2.1を見ると、実験ⅠのセットAでピンクノイズ、レベル差が0 dBの条件で得られた加法定位と先行音効果が切り替わる定位方向の境界は、 -15° , $+17^{\circ}$ で、クリックで -12° , $+14^{\circ}$ の間となっている。レベル差が+3 dBのピンクノイズのきは -14° , $+14^{\circ}$, クリックで -11° , $+12^{\circ}$ 付近となり、若干正面に寄っているが、定位方向の絶対値も小さく、レベル差のないときと同程度で、先行音方向に音像が定位したものと見なせる。レベル差が+6 dBのピンクノイズでは -5° , $+15^{\circ}$ で、クリックは -2° , $+12^{\circ}$ となり、マイナス側は中央寄り、プラス側ではレベル差がない場合と同程度であるが、一方のレベル差が+6 dBを超えると知覚する音像の境界値が極端に変化してしまうものと考えられる。セットAの音源配置では、レベル差が6dBを除き、ピンクノイズとクリックの刺激音で、若干正面寄りではあるが、ほぼ左右対称の定位方向にあり、実験Ⅰ,Ⅱともに、その傾向はほぼ同等である。

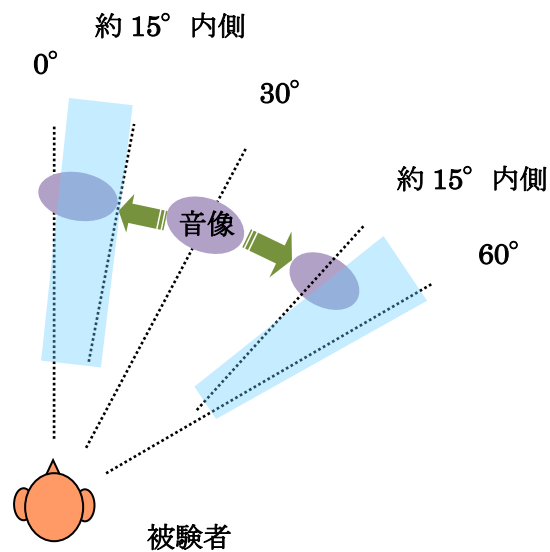
一方、セットBの結果である表2.2では、実験Ⅱにおいて境界が算出された条件で、境界となる時間差における定位方向を同様に算出すると、レベル差が0 dBの場合、ピンクノイズで $+15^{\circ}$, $+45^{\circ}$ となる。セットBの場合、2音源の間は 30° となるため、この条件ではほぼ 30° を中央としてほぼ対称といえる。一方、クリックでは $+17^{\circ}$, $+38^{\circ}$ となっており、 60° が先行している場合に、やや正面方向によった状態で先行音効果が発生している。レベル差が+3 dBの場合、ピンクノイズで $+13^{\circ}$, $+42^{\circ}$ となり、クリ

ックは $+16^\circ$ 、 $+33^\circ$ でレベル差0 dBの場合とほぼ同様の傾向と言える。レベル差が $+6$ dB、刺激音がピンクノイズではマイナス側で $+18^\circ$ 、プラス側は算出不能であった。一方、クリックでは $+20^\circ$ 、 $+35^\circ$ と算出されたが、図2.11を見ても分かるように、2音源の時間差に対する判断の割合の変化が緩やかになっている。これはおそらく定位があいまいになったことが原因で判断の難易度が上がり、境界を決定するのが困難であったためと考えられる。

まとめると、少なくともレベル差が0 dBにおいて、セットAである音源配置が $\pm 30^\circ$ の場合、SLTにおける音像の方向は、ピンクノイズでは -15° 及び $+17^\circ$ 、クリックでは -12° 及び $+14^\circ$ となっている。一方、セットBである音源配置 0° 及び 60° の場合、SLTにおける音像の方向は、ピンクノイズでは 15° 及び 45° 、クリックでは 17° 及び $+38^\circ$ となっている。多少乱暴ではあるが、総括すると $\pm 30^\circ$ の場合には、 -15° と 15° , 0° と 60° の場合には、 15° と 45° が加法定位と先行音効果の判断が分かれる音像の方向と言える。この方向は、2音源の中央方向とそれぞれの音源側のちょうど中間となっており、この範囲を3分割する境界と考えることができる。つまり、定位した音像の方向が単純にどの音源に近いかでどの音源に近いかを判断していると言える。逆に言えば、音源の配置によってSLTが変化するのは、2音源の時間差に対して音像がどこに定位するかが異なっているということになる。この関係を図示したものを図2.12に示す。



(a) 刺激音：ピンクノイズ，音源配置：セット A，レベル差：0dB



(b) 刺激音：ピンクノイズ，音源配置：セット B，レベル差：0dB

図 2.12 時間差による先行音効果が生じる領域

それぞれの音源方向から開き角度の4分の1内側を先行音効果と加法定位が分かれる境界の方向とする．セット A とセット B の場合では，1/4 で 15° 内側から先行音効果が生じはじめる．

2.5 まとめ

2音源によって生じる合成音像の定位方向の変化が、加法定位及び先行音効果とどのような関係にあるかを定量的に明らかにすることを目的として、絶対方向判断と3肢強制選択の二つの定位実験を行った。検討条件とした時間差は、音像が2つの音源の中間付近に定位する加法定位が生じると考えられる極短い時間差から、先行して音出力される音源側に移るまでの比較的短い時間差（2ms）とした。絶対方向判断の実験Ⅰでは直接知覚した音像の方向を回答させ、3肢強制選択の実験では視認できる三つの音源のうちのどれに最も知覚した音像が近いかを強制的に選択させた。この2つの実験を、共通の被験者に対し行い、比較することで目的の達成を試みた。

実験Ⅰの結果として、音源間のレベル差が0, +3, +6 dBのいずれの場合においても、2音源の時間差の増加にともなって、定位方向の平均値は2音源間を徐々に直線的に移動することが示された。また、この傾向は、音源配置によらず共通であった。ただし、刺激音がクリックの場合は、この実験の中で絶対値の大きな時間差となる場合でも、音源方向には寄りきらず、細部においては刺激音の違いが見られた。

実験Ⅱの結果から、先行して音出力された音源方向に定位していると思わせる2音源間の時間差であるSLTを、各音源配置について算出した。その時間差は、算出された条件で比べた限りでは、刺激音による違いはなかった。また、レベル差が大きくなると、その時間差の絶対値も大きくなることが示された。音源配置については、左右非対称となる音源配置において、得られるSLTについても非対称性が見られた。

両実験の結果の比較から、実験Ⅱで算出された境界の時間差における音源の定位方向は、2音源が配置された音源方向から15°程度内側に寄ったところであることが示された。今回の実験で用いた2音源の開き角度が60°であることから、加法定位と先行音効果を分ける空間的な境界は、2音源の方向から2音源の開き角の四分の一内側、つまり、この範囲を単純に三等分する境界であると言える。この結果は、拡声において先行音効果を

利用した定位方向の制御を考えたときに、受聴位置からみた音源の方向を意識することなく、そこで生じる音像の方向のみを考慮すればよいということを示唆するものであり、有益な知見であると考ええる。ただし、本論文の実験結果は、限られた音源配置における広帯域音による合成音像の知覚を調べたものである。実験Ⅰにおける定位方向のばらつきは大きく、本論文では示さないが、個人差もうかがえる。

第3章

刺激音の種類とレベル差の違いが 先行音効果の境界に与える影響

3.1 はじめに

2章では、2種類の回答方法による実験を通して、加法定位と先行音効果を定量的に区別することを試みた。その結果、3肢強制判断による実験によって時間的な境界である SLT を明らかにできた。また、絶対方向判断の実験結果と比較することで、方向としての境界が2音源の方向から2音源の開き角度の四分の一内側であることを示した。

本章では、2章の知見を元に、絶対方向判断の実験を基本として、2音源の音像定位における二つの境界である SLT と ET が、刺激音の違いによってどのような影響を受けるのかを検討する。

3.2 刺激音の継続時間の知覚と認知

人間が日常生活で、聞く音はその大半が変化する音で非定常音（例えば音声や音楽など、時間の経過と共に周波数成分が変化する音）である。知覚は元来時間経過と共に変化するもので、持続的に発せられている音をある時間範囲にわたって統合し、それらの情報を抽出して知覚され認知が行われている。言うまでもなく音は空気の物理的な振動にすぎないが、それを人が認知することによって、音声、物音、音楽としての意味のあるものとなる。どちらかといえば人間は音の物理的変化そのものに着目し

て聴くことはむしろまれで、音の持続する形(パターン)がもたらす情報、例えば言葉の意味、音源(物音)の種類や音楽の流れを知覚し認知しているのだろう。研究において聴取実験に用いられてきた刺激音は、日常生活において耳にする音とは著しく異なり、正弦波・ホワイトノイズ・ピンクノイズ・バースト音などの人工音で、かつ定常的な音が使われ、物理的に見れば極めてよく統制された刺激音で実験が行なわれてきた。

反対に、研究における実験において、用いる刺激を日常耳にするものにする、被験者に働きかける要因が複雑になり、どんな要因が結果に影響を与えたのか明確に言えないことが多くなるという問題が出てくる。このように、日常生活で耳にする音に近い非定常音より、人工的な定常音の方が、刺激の測定、制御は、はるかに容易である。

しかし、非定常音を用いずに、音の知覚におけるパタンの役割について知ることはできない。そこで、実験精度を確保した上で、非定常音の特性について明らかにし種々のパターンについて、知覚の影響を調べる必要がでてくる。人間が普通聞く非定常音は継続時間が約 50～500ms 程度で、レベルは約 50～80dB 程度である。これまでの先行研究では、ラウドネスの観点から言うと、定常状態に達する音の継続時間は約 100～300ms とされており、これは臨界継続時間 (critical duration) と呼ばれる³⁵⁾。表 3.1 に、先行研究者が求めた刺激音の種類による臨界継続時間の値を示す。関連する研究の例として、E. Port³⁶⁾の実験結果を図 3.1 に示す。この図は、刺激音の継続時間による等感曲線の例である。刺激音が 60dB の狭帯域持続雑音を基準音とし、それと等しい大きさに聞こえる短音の音圧レベルを求めたものである。横軸は、狭帯域雑音(ノイズバースト)の継続時間であり、縦軸は、提示した刺激音の大きさが基準音の大きさと等しく感じられる音圧レベルを示している。これによれば、大きさ(ラウドネス)は 70ms 程度で定常状態に達すると考えられ、継続時間が 100～300ms あれば音の大きさは十分定常であると言ってよい。しかし、ラウドネスに関する聴覚の時定数については、他にもさまざまな研究があり、必ずしも一定していない³⁷⁾。以上のような背景から、本章

の先行音効果の実験を行うにあたって、短い継続時間を持つ衝撃音から長い継続時間を持つパタンの異なる音声・音楽などの刺激音を用いるのが適当と考えられる。

表 3.1 信号音の種類による臨界継続時間

著 者 年 代	信号音の種類	臨界継続時間(実験結果)
Miller (1948)	白色雑音	60～150ms
Pollack (1978)	白色雑音	100ms
Stevens & Hall (1966)	白色雑音	150ms
Zwicker (1966)	白色雑音	100ms
Békésy (1929)	800Hz	120～180ms
Pedersen 他 (1977)	1000Hz	160～320ms
Reichardt & Niese (1970)	1000Hz	100ms
Port (1963)	0.3, 2, 10kHz 狭帯域雑音	150～300ms

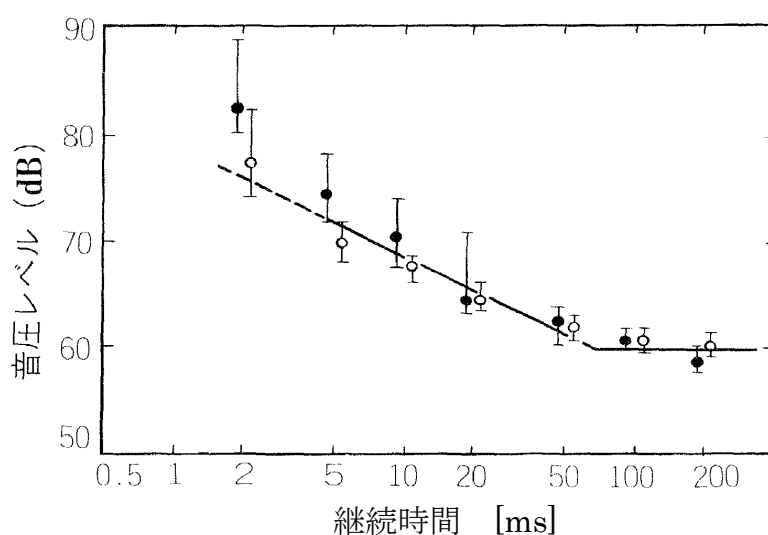


図 3.1 継続時間の異なる狭帯域雑音の等感曲線

(2.5kHz を中心とする狭帯域雑音．直線はエネルギーが等しくなる線)

●：ノイズバーストのレベルを変えた場合，○：基準となる持続雑音のレベルを変えた場合，

3.3 刺激音の種類が先行音効果に与える影響

実験は半無響室で実施した。刺激音は、日常耳にするような音を選び、2音源の時間差が0～50 msの範囲とした場合の先行音効果の生じ方(振る舞い)について検討した。音源は、被験者の真正面から左右各60°間隔で5個を配置した。音源配置を図3.2に示す。真正面の音源を常に先行させる音源とし、正面以外のその他の一つの音源を後続音源とした。各音源から提示する音圧レベルは受聴者の頭部中心で60 dBとした。2つの音源の時間差は、0～50 msの範囲において40段階で変化させた。実験を行った。設定した時間差の詳細は表3.2に示す。回答は、あらかじめ被験者に回答用紙を配布し、音像を知覚した位置を記入する方法で行った。回答用紙の例を図3.3に示す。被験者は男子学生5名の健聴者である。

3.3.1 音源配置と実験システム

実験は、床面に吸音材料を敷いた半無響室(縦7.5 m 横6.0 m 高さ4.6 m)で行なった。被験者とMain-speaker (SP1)、Sub-speaker (SP2, 3, 4, 5, 6)の位置関係は図3.2に示している。SP1は、被験者の真正面に配置し、その他のスピーカは、被験者を中心とする半径3 mの円周上に、いずれも床面上1.4 mの高さに配置した。また、SP2～6は、SP1を基準にして、60°置きに5個のスピーカを配置した。被験者は、頭部が半径3 mの円の中心になるように椅子に腰掛け、頭はヘッドレストで軽く固定した。SP1からは、特殊な条件を除き全試行で必ず音が提示され、常に先行する。SP1の提示音の音圧レベルは、被験者の頭部中心で60 dBと一定にし、SP2～6の音圧レベルも、頭部中心で60 dBとなるようにした。時間差はSP1を基準とし、SP2～6は実験条件に応じて随時刺激音がSP1より0～50msの範囲で40段階の時間差遅れで提示した。設定した時間差の詳細は表3.2に示したとおりである。実験に用いる刺激音は

全部で 4 種類とした。一つ目は、長い継続時間の音声（スピーチ）で、ある日の天気概況を朗読したものである。その継続時間は約 15 秒間である。二つ目は、楽器音であり、サキソフォンの演奏音とした。継続時間は約 12 秒間となる。三つ目は、同じく楽器音で、継続時間約 9 秒のギター演奏音である。最後に、衝撃的な音として、ピンクノイズのトーンバースト音を用いた。そのタイムパターンは、継続時間 200 ms のバースト提示後インターバルが 500 ms でそれを 10 回繰返すこととした。全体の継続時間は約 6.5 秒となる。図 3.4 に、実験システムの概要を示す。また、図 3.5 に、使用した刺激音の時間波形を、図 3.6 に、その信号の周波数スペクトルを示す。信号(刺激音)は、Personal Computer（以下 PC）上で、量子化密度 16 bit、標本化周波数 50 kHz で合成し、Tucker Davis Technology System II（以下 TDT）上に転送することで、時間差を制御した。なお、SP2～6 の切り替えは、TDT により制御し、刺激音はチャンネルセレクトを介してランダムに自動選択できるようにプログラミングした。

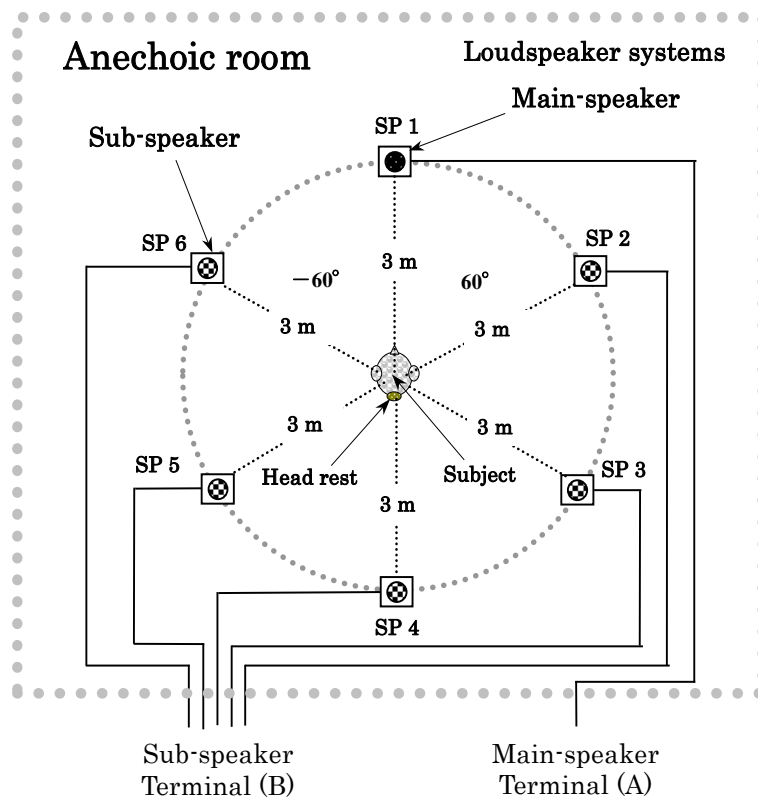


図 3.2 被験者とスピーカの配置図

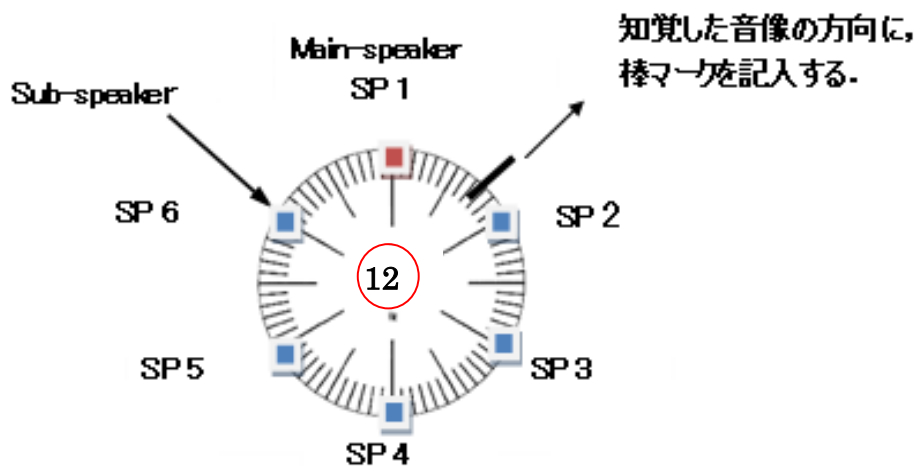


図 3.3 回答用紙

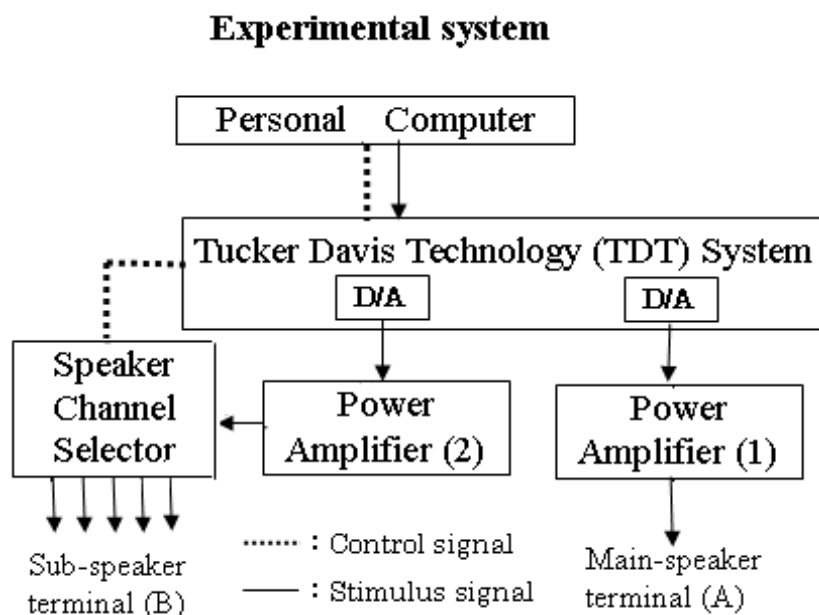


図 3.4 実験システム

3.3.2 口頭教示と回答方法

被験者は、無響室中央にある測定椅子に腰掛けて、顔を正面の音源に向けるようにしてから、ヘッドレストで頭部を軽く固定した。信号音が出されたら、どの方向から音が聞えて来たかを回答用紙に記入させた。図 3.3 に示したように回答用紙には、円周上に 5° 間隔に目盛の付いた円グラフが描かれており、その上には 60° 置きに設置されたスピーカの位置も示されている。被験者には、次の回答の教示を行った。

- 1) 被験者が音像の位置をはっきり知覚した場合、その方向に棒マークで記入する。
- 2) 音像が 2 箇所に知覚された場合、知覚できた 2 箇所に棒マークを記入する。
- 3) 音がどこから聞こえてくるか判断できなかった場合には、回答用紙の円中の番号に (×) 印を付ける。
- 4) 音像が動いて聞こえたらその方向を矢印で表示する。

以上の回答方法で、記入用紙に知覚した音像の方向を記入させた。

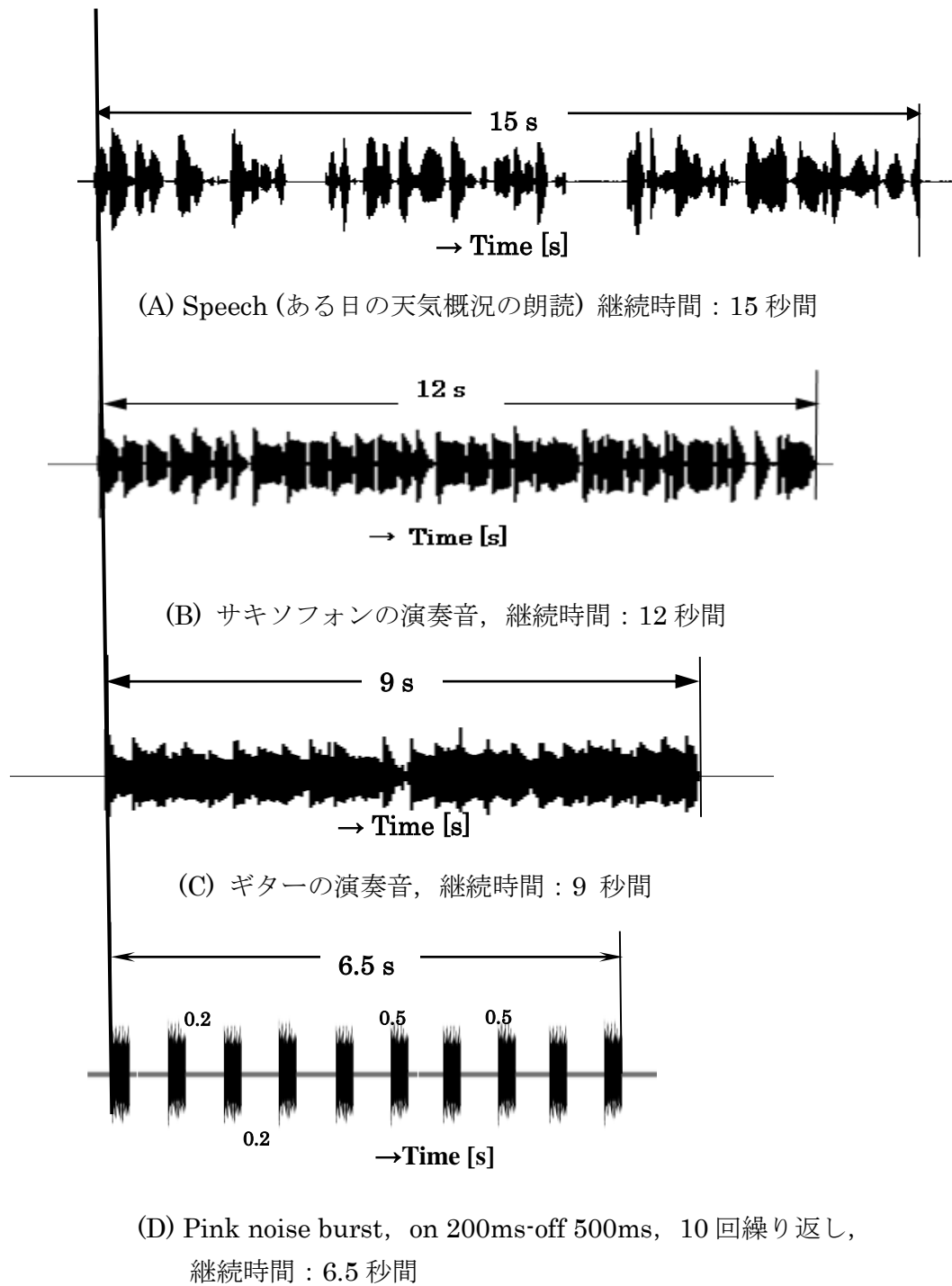


図 3.5 刺激音の波形

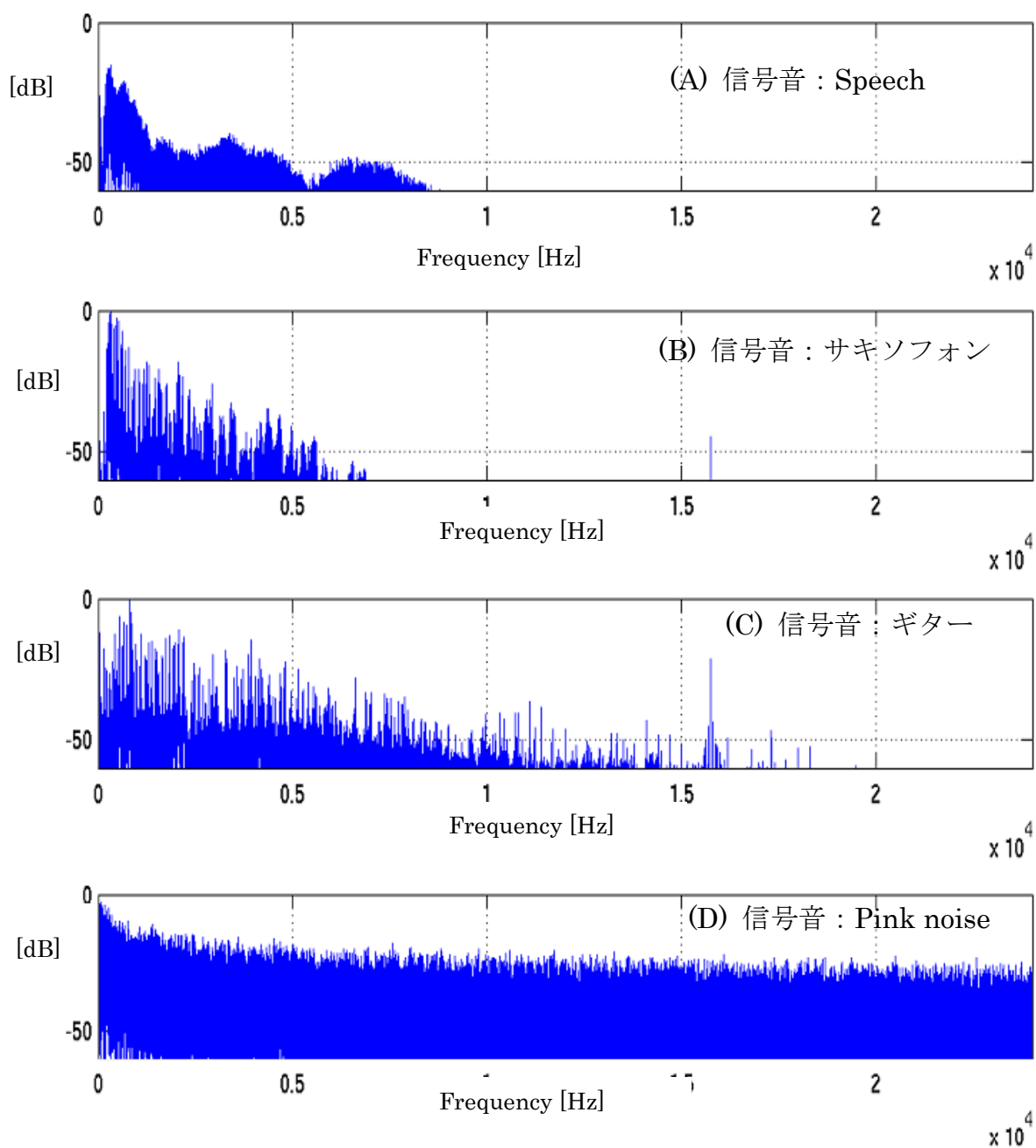


図 3.6 刺激音のスペクトル

表 3.2 被験者に提示した遅延時間

Main-speaker に対する Sub-speaker からの音の遅れ時間 (提示方向 60° , -60° , 120° , -120° , 180°)									
0.0	0.1	0.2	0.3	0.4	0.5	0.6	0.7	0.8	0.9
1.0	1.2	1.4	1.6	1.8	2.0	2.2	2.4	2.6	2.8
3.0	3.5	4.0	5.0	6.0	7.0	8.0	9.0	10.0	12.0
15.0	17.0	20.0	22.0	25.0	27.0	30.0	35.0	40.0	50.0

信号音の遅延時間の条件番号：40 段階，数字は遅延時間[ms].

3.3.3 実験結果

2章で述べたように，2音源から類似した音に時間差を設けて提示した場合，それぞれの2音源の方向から，2音源の開き角度の4分の1内側に加法定位と先行音効果が分かれる空間的な境界があることが明らかとなった．そこで，本章の実験結果は，左右の定位判断は対称であるとみなし，左右対称な条件の回答結果を右側面に折り返した上で，提示音源が 0° と $\pm 60^\circ$ の場合は，音源の開き角度の四分の一にあたる 15° 以内，提示音源が 0° と $\pm 120^\circ$ の場合は，同じく開き角度の四分の一にあたる 30° 以内であれば先行音効果が発生したとみなすこととする．その後，各時間差に対する総試行回数（本実験の場合，10回）から，片側の二項検定に基づき，先行音効果発生割合が8/10以上となった場合を先行音効果が発生したとし，その時間差の境界をSLTとした．

図3.7(A)～(D)，図3.8(A)～(D)に各音源配置での聴取実験の結果をまとめたものを示す．各図の横軸は，2音源間に設けられた時間差を示している．各図の上のパネルは各時間差に対して被験者が回答した絶対方向判断による定位方向をバブルプロットで描いたものであり，縦軸が定位方向を示している．一方，下のパネルは，前述したように2章での知見を基に，2音源間を4等分し， 0° と $\pm 60^\circ$ の音源配置においては， 15° 以内，

0° と ±120° の音源配置においては30° 以内であれば先行音効果が発生したとみなし、先行音効果の発生の有無を割合として描いたものである。グラフの色が黒の場合には、定位方向が15° もしくは30° より大きくなった割合を示しており、グラフの色が灰色の場合は、定位方向が15° もしくは30° 以内となった割合を示している。各時間差に対する試行回数が10回となっているため、片側の二項検定に基づき、図中に示した線のように灰色で示された先行音効果発生の割合が0.8以上となった場合に先行音効果が発生するとみなすこととする。この考えに基づき、時間差が増加していき、先行音効果の発生割合が0.8以上となる時間差を本実験で得られるSLTとした。厳密な意味では0.8以上となる時間差とその直前の時間差との間にSLTが存在すると考えられるが、刺激音の違いがSLTに及ぼす影響を検討するという観点では本研究における求め方でも問題はないと考えられる。一方、本実験ではエコーを知覚した場合には、それぞれの音像を知覚した方向を答えるよう教示しているので、確実にエコーを知覚できるようになると、黒と灰色の回答割合が0.5/0.5となる。そこでETの求め方としては、回答割合が半々となる遅延時間をETとすることとする。これも同様に厳密な意味では今回の求め方よりも若干短い時間差がETであると考えられるが、比較という意味では問題ないといえる。

以上のような手順で図 3.7, 3.8 から求めた先行音効果の下限の境界 SLT と上限の境界 ET を、表 3.3～3.7 に示す。表をみて分かるように、音源配置が 0° と ±60° 方向のピンクノイズバーストの場合、SLT は 0.6 ms, ET は 5 ms となり、これらから先行音効果の生じる範囲は 0.6～5 ms であると言える。スピーチでは、SLT は 0.6 ms, ET は 10 ms, 範囲は 0.6～10 ms となる。同様に、ギター演奏音では、SLT は 0.8 ms, ET は 7ms, 範囲は 0.8～7 ms となり、サキソフォン演奏音の場合では、SLT は 0.6 ms, ET は 10 ms, 先行音効果の範囲は 0.6～10ms となる。

音源配置が 0° と ±120° 方向のピンクノイズバーストの場合も同様に読み取ることができ、SLT が 1ms, ET は 7ms, 先行音効果の生じる範囲は 1～7ms となる。ス

ピーチでは、SLT は 0.8 ms, ET は 10 ms, その範囲は 0.8~10 ms となる。ギター演奏音では、SLT は 1.2 ms, ET は 15 ms, 先行音効果の範囲は 1.2~15 ms となっており、今回用いた刺激音の中では ET が最も大きくなっている。サキソフォン演奏音の場合には、SLT は 0.8 ms, ET は 8 ms, 先行音効果の範囲は 0.8~8 ms であった。

この結果を、2 音源の到来方向による違いと刺激音の種類による違いという観点で見てみる。SLT, ET とともに、後続音源が $\pm 60^\circ$ の時よりも、 $\pm 120^\circ$ の時のほうが、値が大きくなっている。このことは開き角が大きいほど先行音源側に音像を持っていくには時間差を大きくとる必要がある可能性を示唆している。一方、エコーについては、検知限の値が大きくなっていることになるため、開き角が大きいほど分離しにくい可能性を示唆している。

一方、刺激音の種類に関する違いについて見てみると、後続音源が $\pm 60^\circ$ の場合 SLT の平均が約 0.65 ms で、 $\pm 120^\circ$ の場合は、0.95 ms となっている。それぞれの標準偏差が、0.1 ms と 0.19 ms であることを考えると、SLT については刺激音の種類による値の変動は比較的小さいといえる。ET については、後続音源が $\pm 60^\circ$ の場合 ET の平均が 8.0 ms で、 $\pm 120^\circ$ の場合は、10.3 ms となっている。それぞれの標準偏差は、2.4 ms と 3.3 ms となっており、SLT に比べて刺激音の種類による変動が大きくなっている。傾向としては、刺激音がスピーチ・ギター・サキソフォンのように日常的な音の場合では約 10ms 程度と長めになったが、ピンクノイズバーストのように非日常的で衝撃性の強い、短区間の音では、約 6ms 程度と他の刺激音よりも ET が短く知覚されることが分かった。

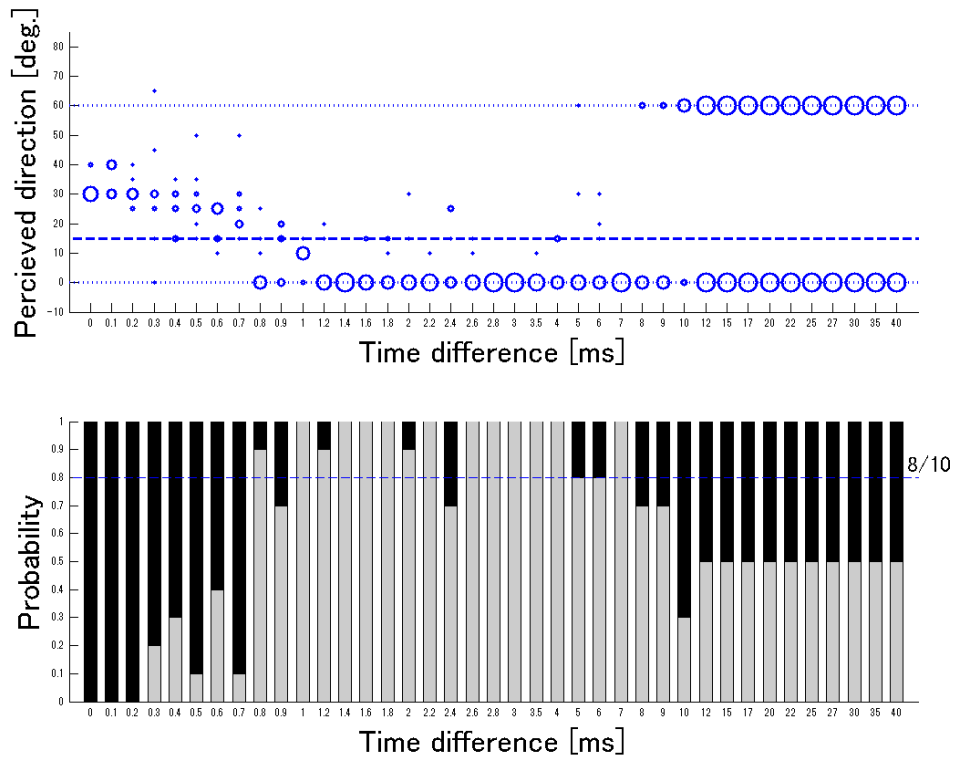
3.3.4 考察

3.3.3 節で述べたように、刺激音の種類が SLT の値に及ぼす影響はほとんどないという結果になった。これは本研究において SLT を定義し、明確に加法定位と先行音効果の境界を定義したことによって得られた新たな知見と言える。ET については、刺激音の違いによって値の変動が見られた。刺激音の種類が ET に及ぼす影響に関する研究には過去にいくつか参考になるものが存在する。例えば、Schubert ら³⁹⁾の実験では三角包絡線の継続時間が 50 ms の信号音を用いてエコー検知限 ET を調べており、その結果 ET は約 7~12 ms であるという結論を示している。また、Cherry ら²³⁾は、スピーチのような長い継続時間を持つ連続した刺激音では、エコー検知限 ET が約 20ms 程度となることを報告している。岩宮¹⁹⁾の研究では、刺激音の立ち上がりが ET に影響することが示されている。立ち上がりが急峻な場合には ET は短くなり、刺激音の立ち上がりが緩慢な場合はエコー検知限 ET が長くなる。刺激音の種類による ET に関する知見をまとめたものを表 3.8 に示す。こういった過去の知見から、刺激音の包絡のパターンと継続時間の長さが ET に影響し、その結果先行音効果が生じる時間差の範囲が変動するのではないかと想定した。本実験で用いた各信号音の周波数スペクトルを確認すると(図 3.6 を参照)。まず、ピンクノイズバースト音の場合、他の信号に比べるかなり広い帯域の周波数成分を有しており、周波数分布もなだらか (-3dB/octave) となっている。バースト音の継続時間は 15 秒で、10 回の繰り返し音であり、信号の立ち上がり・立ち下がり急峻である。スピーチの周波数特性としては、0.2, 1, 3 kHz 付近に緩やかな山谷が分布していて、8kHz 付近まで周波数成分が存在する。信号の時間波形を示した図 3.5 を見ると、バースト音やギター音ほど、立ち上がり・立ち下がり急峻ではない。また、サキソフォンの演奏音の場合には、周波数の分布は 250Hz 付近に激しい凹凸が見られ、6kHz 付近まで特徴的なスペクトル分布が存在する。時間波形をみると信号の立ち上がりは、スピーチほど緩やかではないが、バースト音やギター音ほど急峻ではない。続い

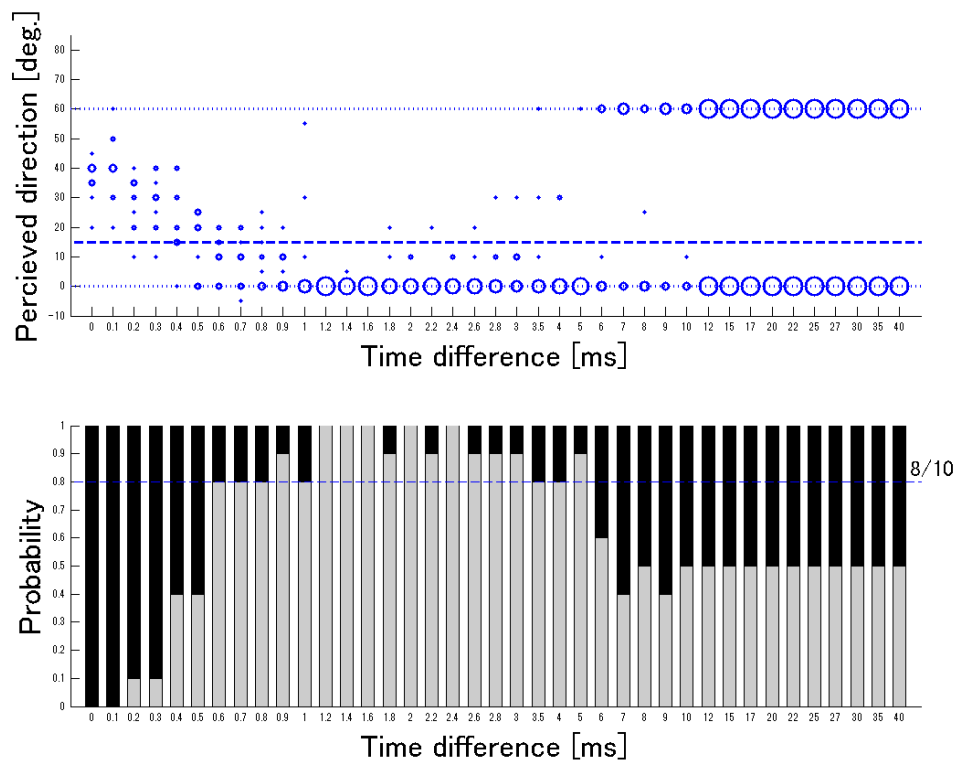
て、ギター音の周波数分布を見ると、スペクトルは 100Hz から 15kHz と比較的広い範囲に分布し、さらに高い周波数のスペクトルが不規則な間隔で現れている。一方、時間波形としては、バースト音と同様に立ち上がりが急峻である。このような各刺激音の立ち上がりや減衰の形と周波数帯域の広さが、先行音効果に關与している可能性がある。そこで、これらの刺激音の周波数スペクトルの分布帯域などを比べてみると、①ピンクノイズバースト、②ギター音、③スピーチ、④サキソフォンの順で、含んでいる周波数スペクトルの帯域が広い。実験結果から先行音効果が働く領域を見ると、先行音効果の下限 SLT は、各刺激音で約 0.8ms 程度で、ほぼ同じとなっている。このことから、先行音効果の下限の時間差では、刺激音の違いの影響はほとんどないものと言って良い。一方、先行音効果が発生する時間帯の上限のエコー検知限 ET に対応する時間差は、①ピンクノイズバーストは 5.0 ms、②ギター音は 7.0 ms、③サキソフォン 8 ms、④スピーチは 10 ms と、この順序で先行音効果の上限 SLT の時間差が大きくなっており、周波数帯域が広い刺激音ほど ET の値が小さくなっている。このように、刺激音の種類の違いで、先行音効果領域の上限 ET で、先行音効果の生じる範囲が影響されていることが分かり、若干短いが、過去の研究でも示された値とほぼ同等といえる。

また、各刺激音に対して先行音効果領域の判定精度を見るため先行音効果の生起頻度の割合を見ると、ギター、ピンクノイズバースト、スピーチ、サキソフォンの順番で被験者は、先行音効果の変移過程を適確に判断していることが分かった。

以上のことから、刺激音の種類の違いで、加法定位から先行音効果までは、各刺激音を形成する継続時間と刺激音の持つ周波数スペクトルのパターンが下限 SLT には、ほとんど影響はなく、先行音効果からエコーに至るエコー検知限 ET に対して大きい影響を与えることが分かった。今回の実験で得られた刺激音の種類に対する SLT 及び ET の平均をグラフとして表したものを図 3.9 に示す。この傾向は過去の研究より若干小さいがほぼ同様な傾向であると言える。

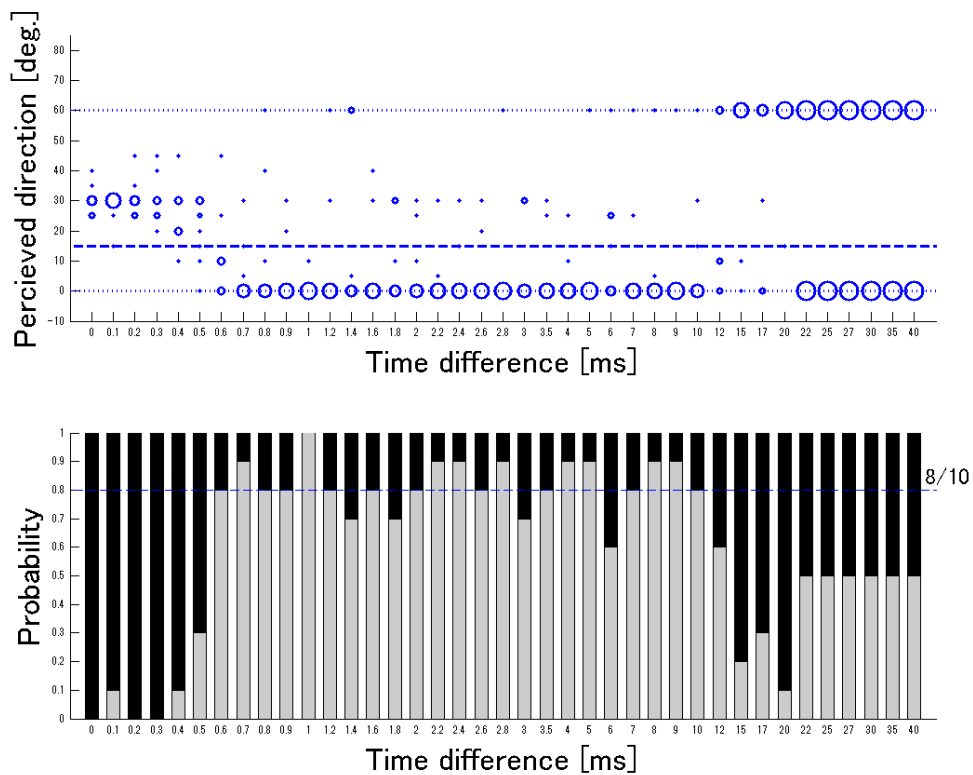


(A) 刺激音：ギター，音源配置： $0^{\circ} \sim \pm 60^{\circ}$ ，レベル差：0dB

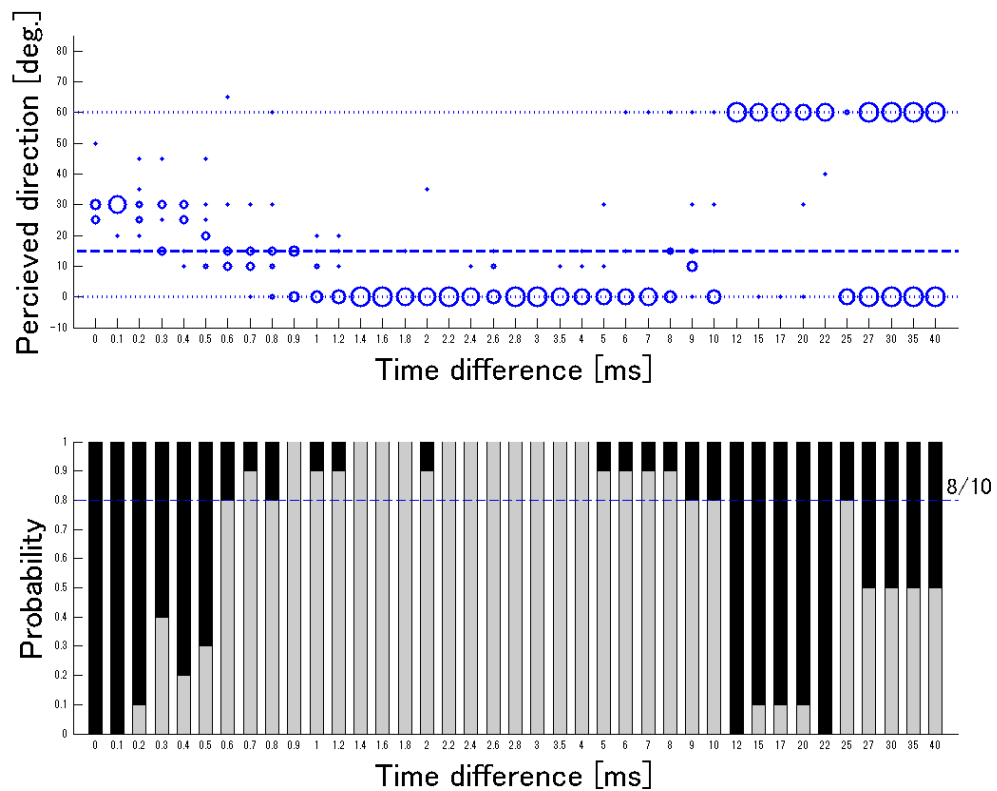


(B) 刺激音：ピンクノイズバースト，音源配置： $0^{\circ} \sim \pm 60^{\circ}$ ，レベル差：0dB

図 3.7 各時間差条件に対する定位方向及び先行音効果が発生した割合

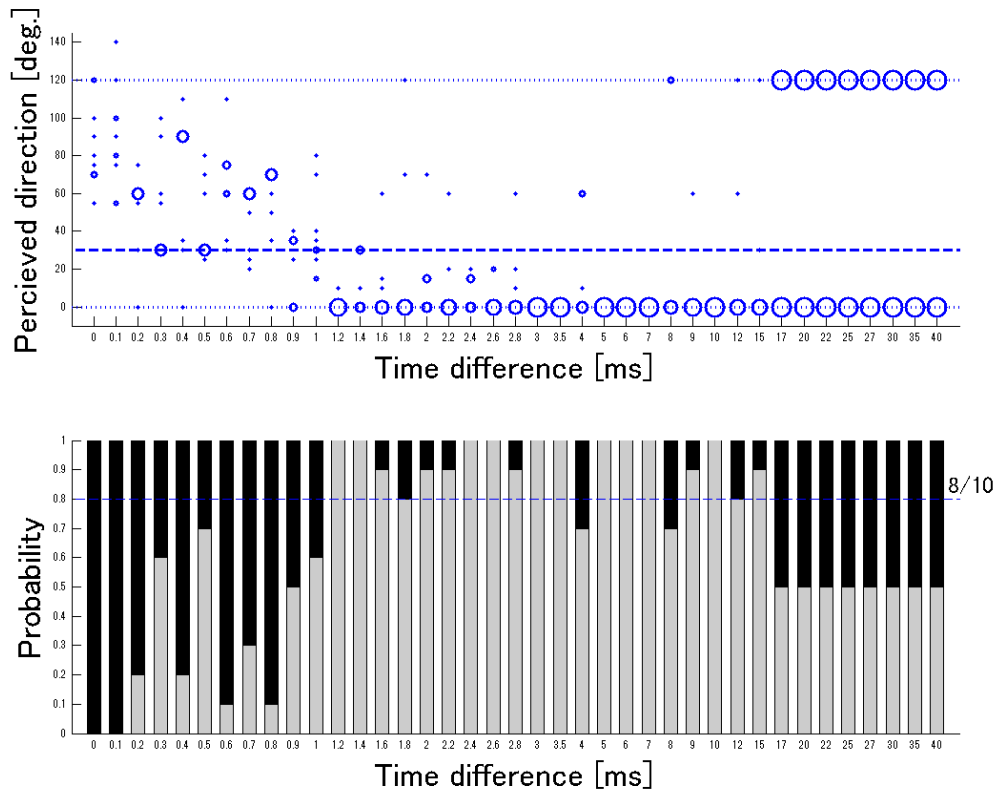


(C)刺激音：サキソフオン，音源配置： 0° ～ $\pm 60^{\circ}$ ，レベル差：0dB

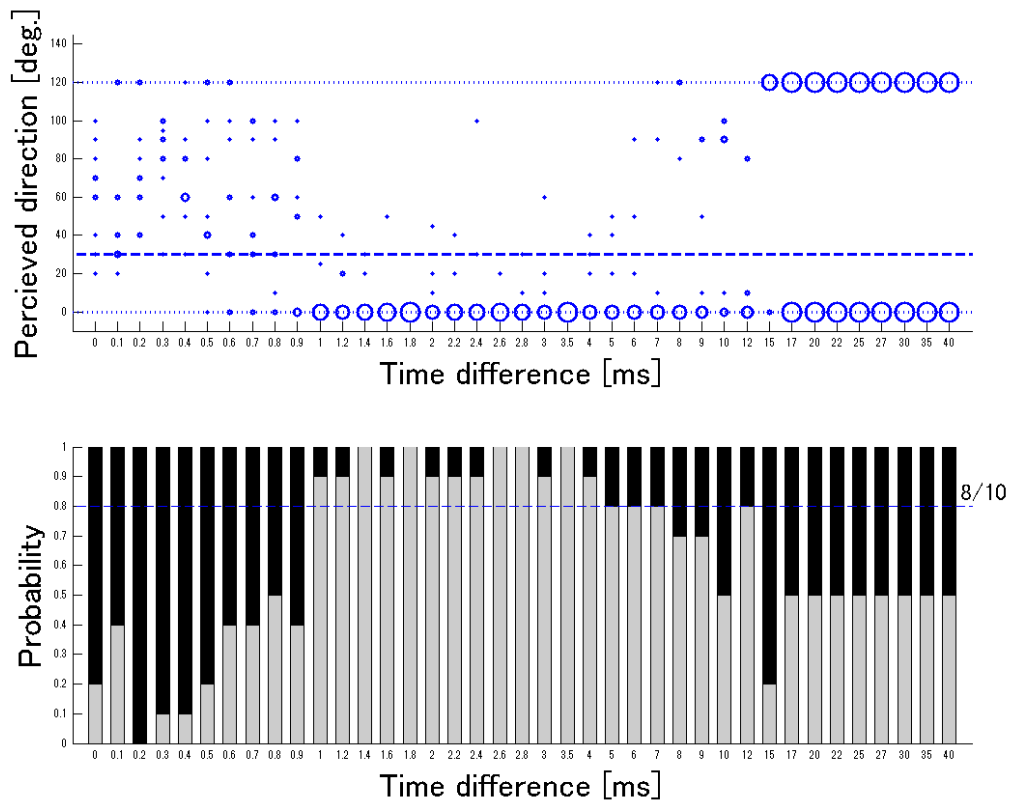


(D)刺激音：スピーチ，音源配置： 0° ～ $\pm 60^{\circ}$ ，レベル差：0dB

図 3.7（続き）各時間差条件に対する定位方向及び先行音効果が発生した割合

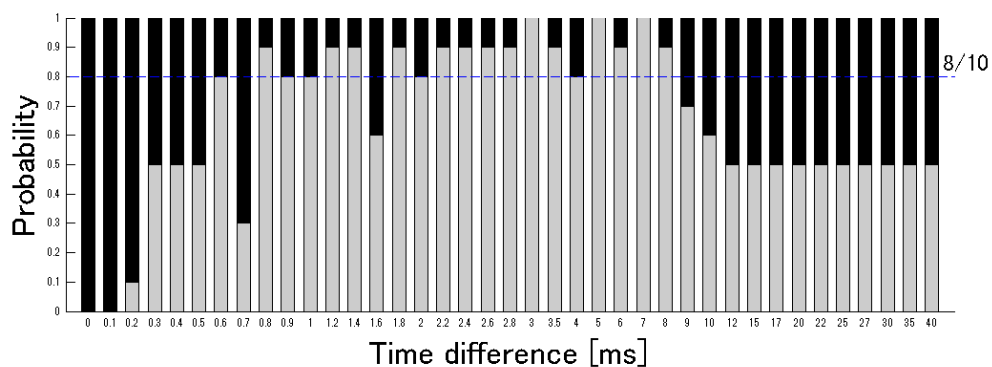
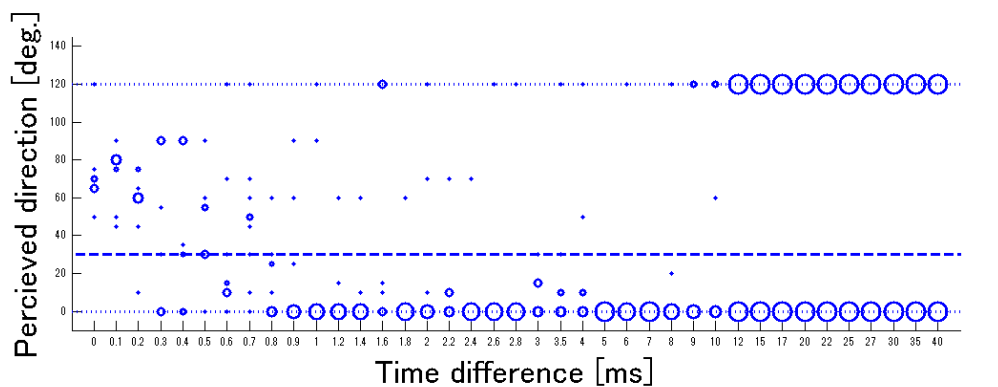


(A) 刺激音：ギター，音源配置： 0° ～ $\pm 120^{\circ}$ ，レベル差：0dB

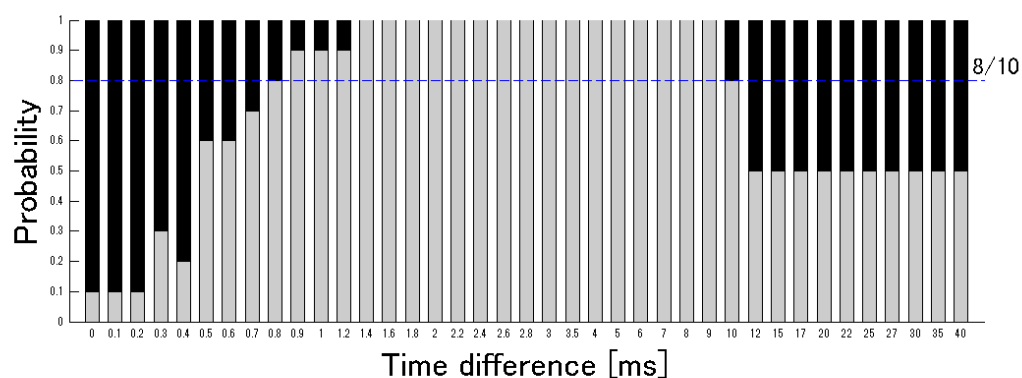
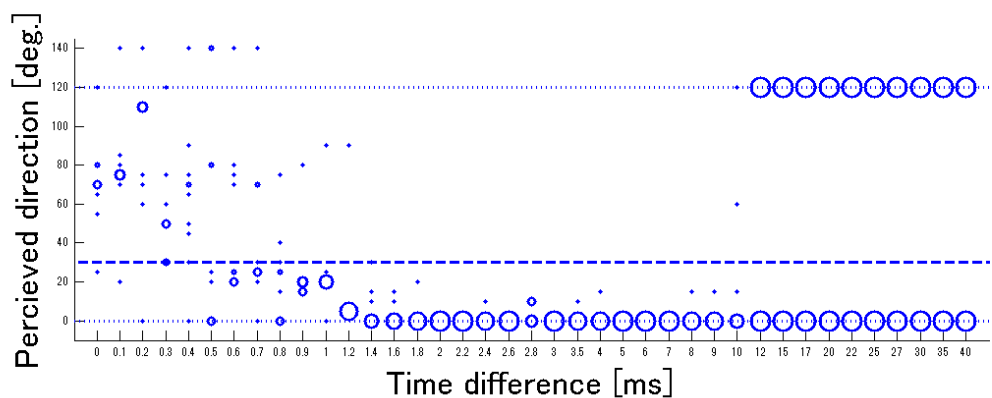


(B) 刺激音：ピンクノイズバースト，音源配置： 0° ～ $\pm 120^{\circ}$ ，レベル差：0dB

図 3.8 各時間差条件に対する定位方向及び先行音効果が発生した割合



(C) 刺激音：サキソフォン，音源配置： 0° ～ $\pm 120^{\circ}$ ，レベル差：0dB



(D) 刺激音：スピーチ，音源配置： 0° ～ $\pm 120^{\circ}$ ，レベル差：0dB

図 3.8（続き）各時間差条件に対する定位方向及び先行音効果が発生した割合

表 3.3 2 音源間の時間差により発生した先行音効果の範囲

2 音のスピーカの開き角度	SLT	先行音効果の範囲	ET
±60°	0.6	0.6～5.0	5.0
±120°	1.0	1.0～8.0	8.0
平 均	0.8	0.8～6.5	6.5

刺激音：Pink noise on 200ms-off 500ms 10 回 繰り返し，継続時間：6.5 秒間，
音源の位置：±60°，音圧レベル：Main-speaker と Sub-speaker 共に 60dB

表 3.4 2 音源間の時間差で発生した先行音効果の範囲

2 音のスピーカの開き角度	SLT	先行音効果の範囲	ET
±60°	0.6	0.6～10.0	10
±120°	0.8	0.8～10.0	10
平 均	0.7	0.7～10.0	10

刺激音：スピーチ，継続時間：15 秒間，音源の位置：±60° ±120°
音圧レベル：Main-speaker と Sub-speaker 共に 60dB

表 3.5 2 音源間の時間差で発生した先行音効果の範囲

2 音のスピーカの開き角度	SLT	先行音効果の範囲	ET
±60°	0.8	0.8～7.0	7.0
±120°	1.2	1.2～15.0	15.0
平 均	1.0	1.0～11.0	11.0

刺激音：ギター音，継続時間：9 秒間，音源の位置：±60° ±120°
音圧レベル：Main-speaker と Sub-speaker 共に 60dB

表 3.6 2 音源間の時間差で発生した先行音効果の範囲

2 音のスピーカの開き角度	SLT	先行音効果の範囲	ET
±60°	0.6	0.6～10.0	10.0
±120°	0.8	0.8～8.0	8.0
平 均	0.7	0.7～9.0	9.0

刺激音：サキソフォン，継続時間：12 秒間，音源の位置：±60° ±120°
音圧レベル：Main-speaker と Sub-speaker 共に 60dB

表 3.7 刺激音の種類により発生した先行音効果の範囲

刺激音の種類	SLT	先行音効果の範囲	ET
ピンクノイズバースト	0.8	0.8～6.0	6.0
スピーチ	0.7	0.7～10.0	10.0
ギター	1.0	1.0～11.0	11.0
サキソフォン	0.7	0.7～9.0	9.0
平均	0.8	0.8～9.0	9.0

刺激音：4 種類，音源の位置：±60° ±120° ， 音圧レベル：Main-speaker
と Sub-speaker 共に 60dB

表 3.8 先行研究者によるエコー検知限(ET 値)

文 献	信 号	音圧レベル (dB SL)	閾の種類	閾 値
宮坂栄一, 境 久雄 (1979)	2kHz 立上クリック音	60	弁別限	10 ms
	2kHz 立下クリック音	60	弁別限	7.5 ms
江崎正直, 曾根敏夫 二村忠元(1974)	パルス幅 30 μ sec 継続時間 100ms	---	弁別限	8 ms
Klemm(1920)	クリック音	---	エコー検知限	2 ms
Rosenzweig and Rosenblith(1950)	クリック音	---	エコー検知限	2 ms
Cherry and Taylor(1954)	スピーチ	---	エコー検知限	20 ms
Blodgett, et al.(1956)	425Hz を中心とする 狭帯域ノイズ	---	エコー検知限	7.6 ms
	広帯域ノイズ 0.1-4.8kHz	---	エコー検知限	9.5-10 ms
Guttman(1962)	クリック音	36	エコー検知限	3 ms
		16		4.6 ms
Babkoff and Sutton(1966)	クリック音	50	エコー検知限	2.4 ms
Schubert and Wernick(1969)	純音 0.25,0.5,2kHz 三角包絡線 継続時間 20ms	40	エコー検知限	5-7 ms
	継続時間 50ms			7-12 ms
	継続時間 100ms			up to 100 ms
	ハイパスノイズ, ローパスノイズ (遮断周波数 1kHz) 三角包絡線 継続時間 20ms	40	エコー検知限	5.5(HP),6(LP)
	継続時間 50ms			8(HP),12(LP)
	継続時間 100ms			16(HP),22(LP)

ダイコティックにヘッドホン受聴した場合に, 生じたエコー検知限の値を示している.

直接音: 左耳, 反射音: 右耳, あるいは, その逆, 直接音と反射音は同レベルである.

Jens Blauert 著: 空間音響 鹿島出版会 (1986)²⁾.

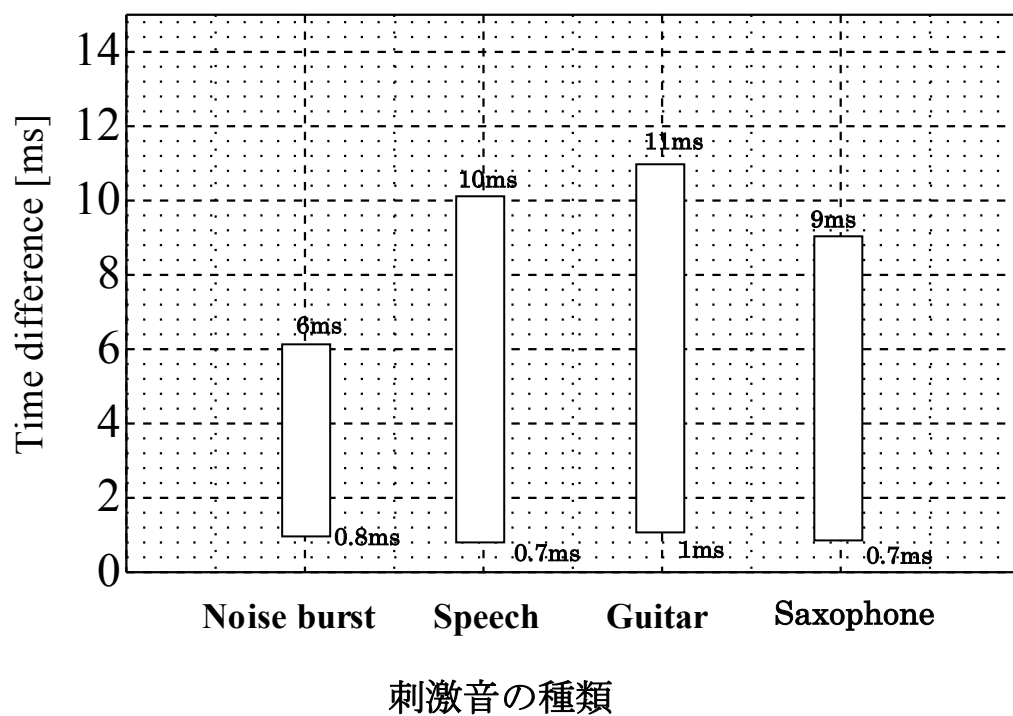


図 3.9 刺激音の種類により発生とした先行音効果の範囲

図 3.7(A~H)のデータから先行音効果の生じた範囲を示した。

音源位置は、 $\pm 60^\circ$ ， $\pm 120^\circ$ の方向であり，平均値として求めた値である。

3.4 レベル差の違いが先行音効果に与える影響

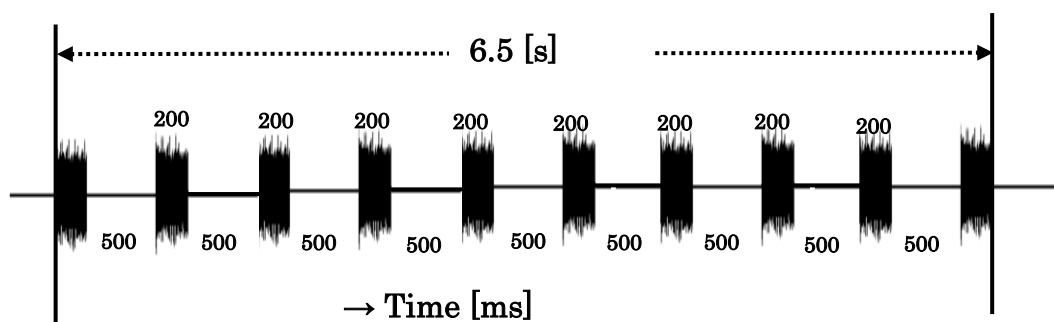
3.1 節でも述べたが、室内空間で拡声装置を使用する場合、聴取者は、講演者(主音源)より客席に近いスピーカからの音が定位において優位に働くことがあり、主音源の方向とは異なる方向へ音像を知覚して、違和感を生じることがある。本実験では、遅れて到来する刺激音の音圧レベルが、先行して到来する刺激音のレベルより 3 dB, 6 dB,あるいは 9 dB と高くなった場合に、2 音源の時間差による先行音効果の生起がどのような影響を受けるのかを検討する。2 音源のレベル差が先行音効果に及ぼす影響に関する先行研究としては、Kietz ら³⁹⁾は、どんな周波数でも両耳間レベル差が 10~15dB になれば、音像が片一方の耳に完全に移動し、音像は強い刺激を受けている耳の方へ生じると報告している。また一般的に、定位判断において、両耳間差において、レベル差のほうが、時間差よりも支配的であると言われている。基本的に自然な空間において、レベル差及び時間差のどちらかしか存在しないということは稀で、常に種々の両耳間信号差が生じている。そういったことから、音像を制御する応用を踏まえると、どの両耳間差が比較的重要であるか、そして、それらがどのように互いに関係しているのかを明確にする必要がある。本実験では、刺激音として、ピンクノイズバースト音を用い、2 音源の時間差をランダムに提示し、レベル差を 0, 3, 6, 9 dB と変化させて、先行音効果の生起にどのような影響があるのかを検討した。

3.4.1 実験系、実験条件及び回答方法

本実験には、3.3 節で実施したものと同一実験系を用いた。被験者の回答方法も共通である。前述したように、刺激音としては 3.3 節でも用いたピンクノイズバーストを用いた。そのタイムパターンとスペクトルを図 3.10 に示す。Main-speaker(SP1)からの信号音のレベルは被験者の頭部の中心位置で 60 dB 一定となるように固定し、

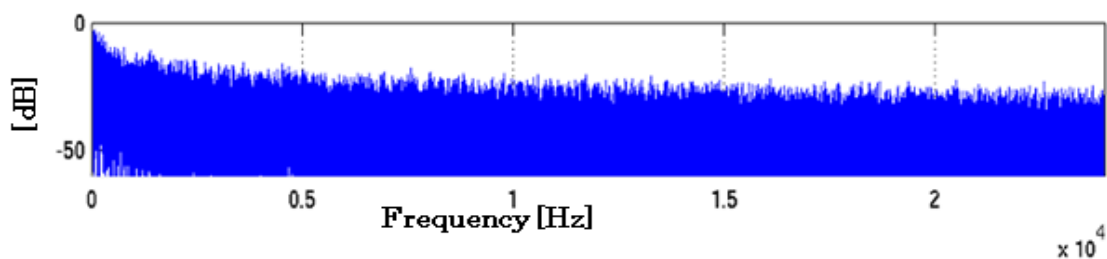
Sub-speaker (SP 2, 3, 4, 5, 6)からのレベルを 60, 63, 66, 69 dB のいずれかになるように変化させた。2 音源間に設けた時間差, 実験に参加した被験者は 3.3 節の実験と全く同じである。被験者 1 人当たりの回答数は, Sub-speaker の数が 5 つ(スピーカ), 時間差条件が 40 段階で, 計 200 施行, それにレベル差条件が 4 条件加わり, 4×200 で 800 試行となる。したがって, 被験者 5 名を含めると実験全体の判断数は $5(\text{被験者}) \times 800 = 4000$ 回となる。

被験者に課したタスクについても, 3.3 節で実施した刺激音の種類の影響を調べた実験と全く同様であり, その集計方法の考え方も同じとなる。



(A) ピンクノイズバーストのタイムパターン

信号音：ピンクノイズバースト音， 継続時間：6.5 秒間， on200ms-off500ms， 10 回繰り返し波形， 音圧レベル：ピークホールドで 60dB.



(B) 周波数スペクトル

図 3.10 (A),(B) ピンクノイズのタイムパターンと周波数スペクトル

3.4.2 実験結果

図 3.11 (A)～(D), 図 3.12 (A)～(D), 図 3.13 (A)～(D)に, 各音源配置及びレベル差での聴取実験の結果をまとめたものを示す. 図の作成の仕方は 3.3 節の実験と同様であり, 左右の方向定位の特徴は対称であるとみなし, 図 3.11 は 0° と $\pm 60^{\circ}$ の音源配置の結果を, 図 3.12 は 0° と $\pm 120^{\circ}$ の音源配置の結果を示している. 図 3.13 は 0° と 180° の音源配置での実験結果を表す. 各図の横軸は, 2 音源間に設けられた時間差を示している. 各図の上のパネルは各時間差に対して被験者が回答した定位方向をバブルプロットで描いたものである. その下のパネルは繰り返しとなるが 3.3 節でも示したように 2 章での知見をもとに決定した加法定位と先行音効果の方向的境界を用いて, 先行音効果の発生割合を求め, それを図示したグラフである. 各図から刺激音の種類の影響を検討したときと同様に SLT と ET を決定し, 先行音効果の生じる範囲を求めた. その結果を表 3.9 に示す.

先行音効果の発生する範囲から, レベル差の違いが先行音効果に及ぼす影響について見てみる. 2 音源の開き角度が $\pm 60^{\circ}$ 方向で 2 音源が共に 60dB の場合, SLT が 1.2 ms で ET が 7 ms となっている. レベル差が 3 dB の場合は, 先行音効果が発生する範囲が 0.9～3.5 ms となっており, 6 dB では, 1.8～3.0 ms となる. このように, 後続音のレベルが徐々に大きくなるにつれ, 先行音効果の発生する範囲は狭くなっていく傾向があることが分かる. レベル差が 9 dB となると, 音像が先行音源側へ移動しないためと思われるが, 今回の手法では SLT も ET も算出することが出来なかった. したがって, レベル差が 9 dB あると, 先行音効果は発生せず, 常に高いレベルを持つ後続音源側に音像が定位すると言える.

音源配置が $\pm 120^{\circ}$ では, レベル差が 0dB, 2 音源が共に 60dB の場合, SLT が 1.0 ms, ET が 15 ms となり, 先行音効果の発生した範囲は 1.0～15.0 ms と分かる. 2 音源のレベル差が 3dB では, 先行音効果の範囲は 1.0～8 ms, 6dB のレベル差では, 1.0～3.0 ms

となる．このようにレベル差が増すと徐々に先行音効果の範囲が狭くなる傾向は音源の配置が 0° と $\pm 60^\circ$ と同じである．しかし，SLT はレベルが増加しても 1.0 ms のままで変動しておらず，ET のみがレベルの増加に従い，徐々に減少している．また，3.3. 節でも述べたように，本研究の音源配置では先行音源と後続音源の開き角度が大きくなると ET が増加する傾向にあった．レベル差が増加するとその ET の増加傾向は，0 dB の場合は 15.0-7.0 で 8.0 ms の増加が見られていたが，レベル差が 3 dB になる 4.5 ms の増加となり，レベル差が 6 dB となると ET の増加が見られなくなった．

また，音源配置が 0° と 180° ，レベル差が 0dB，2 音源が共に 60dB でレベル差 0dB の場合，SLT が 0.5 ms，ET は 4 ms で，先行音効果が発生した範囲は 0.5～4 ms となる．レベル差が 3 dB では，先行音効果の範囲が 0.6～9 ms で，6 dB のレベル差では，0.5～2.8 ms となる．レベル差が大きくなると徐々に先行音効果の範囲が狭くなる傾向や，SLT はレベルの影響をほとんど受けない傾向は 0° と $\pm 120^\circ$ の結果と同じと言える．しかし，音源配置が 180° の真後ろにある場合は，始めから先行音側にやや優勢な音像定位が見られたものの，前後誤りが起きやすいためか，音像の定位方向が前後し曖昧で不安定であることが分かった．

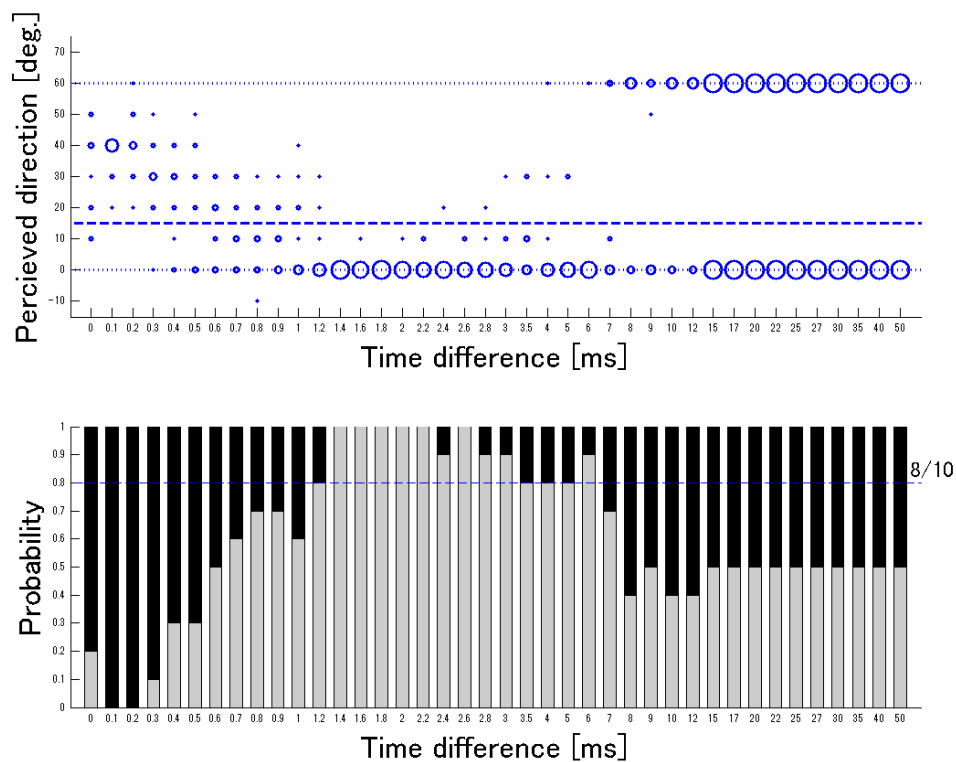
3.4.3 考察

先行音と後続音の間にレベル差がある場合の先行音効果に関する実験の結果から，先行音側に音像を定位した頻度の割合を求めた．それによると，後続音のレベルが先行音より高い場合，先行音とのレベル差が 6 dB までは，先行音効果の生じる範囲の影響はそれほど大きくはないことが分かる．しかし，先行音と後続音のレベル差が 3～6 dB と大きくなるにつれて，先行音効果の成立する確率が減少する傾向にあることが示されている．先行音効果の生じる範囲内の総試行回数に対し，先行音効果が成立した

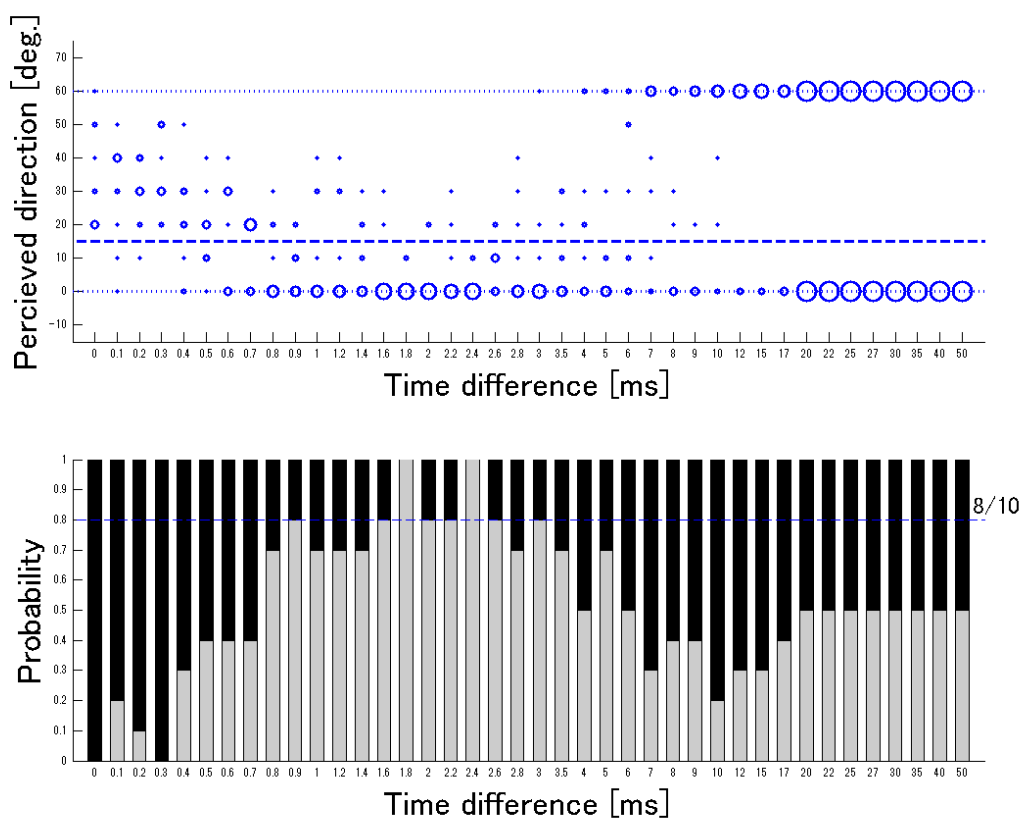
試行の割合を求めた。その結果を、図 3.14 に示す。後続音のレベルが先行音と等しい場合には、先行音効果の生じる割合は 88 %、レベル差が 3 dB 高ければ 76 %、6 dB 高ければ 57 % となり、後続音のレベルが増すと、先行音効果の生じる割合は徐々に減少する傾向を示した。この傾向は、先行音と後続音の音圧レベルが等しい場合の割合を基準として、3 dB 高くなるごとに 15 % 程度、先行音効果の生じる割合が減少することを示している。また、後続音のレベルが先行音より 9 dB 高くなれば、遅延音の提示方向にかかわらずレベルの高い側、すなわち、後続音に音像を定位する傾向が見られ、先行音効果は、ほとんど現れなかった。さらに、後続音の提示方向が 180° の場合では、遅延時間と音のレベル差の効果は曖昧で、加法定位および先行音効果の成り立つ範囲は特定することは困難であった。この結果は、後続音のレベルが先行音のレベルより 6 dB を越えない範囲で高くなっても、先行音効果は生じるが、その差が 9 dB になると、時間差よりレベル差の効果が優位なることを示している。この傾向は 田原ら⁴⁰⁾の結果とよく似ている。また Damaske⁴¹⁾は、先行音効果が生じる 2 音源のレベル差が約 10 dB 程度が限界としており、それと比較すれば、本実験の結果は、それよりも低い値となっている。この違いの原因は、Damaske がエコー検知限を先行音効果領域の上限とした評価基準設定の違いや提示した刺激音の方向の違いなどによるものと推定される。しかしながら、ここで得られた結果は、先行音効果領域を直接求めたものであり、より確かな値が得られたものと考えられる。この結果から、先行音と後続音の音圧レベル差が 6 dB 付近あるいは、その近傍に、音像定位の変化を引き起こす限界があり、この条件を考慮に入れて拡声装置等の設計を行う必要があるのではないかと考えられる。

表 3.9 2 音源のレベル差による先行音効果領域と先行音側に音像を定位した頻度の割合

2 音源 Main/Sub スピーカの時間差と音圧レベル差による 先行音側に判断した頻度の割合				
Main/Sub の 音圧レベル差	0 dB	3 dB	6 dB	9 dB
±60°の先行音 効果の範囲	1.2～7.0ms	0.9～3.5ms	1.8～3.0ms	生じない
±120°の先行音 効果の範囲	1.0～15.0ms	1.0～8.0ms	1.0～3.0ms	生じない
180°の先行音 効果の範囲	0.5～4.0ms	0.6～9.0ms	0.5～2.8ms	生じない

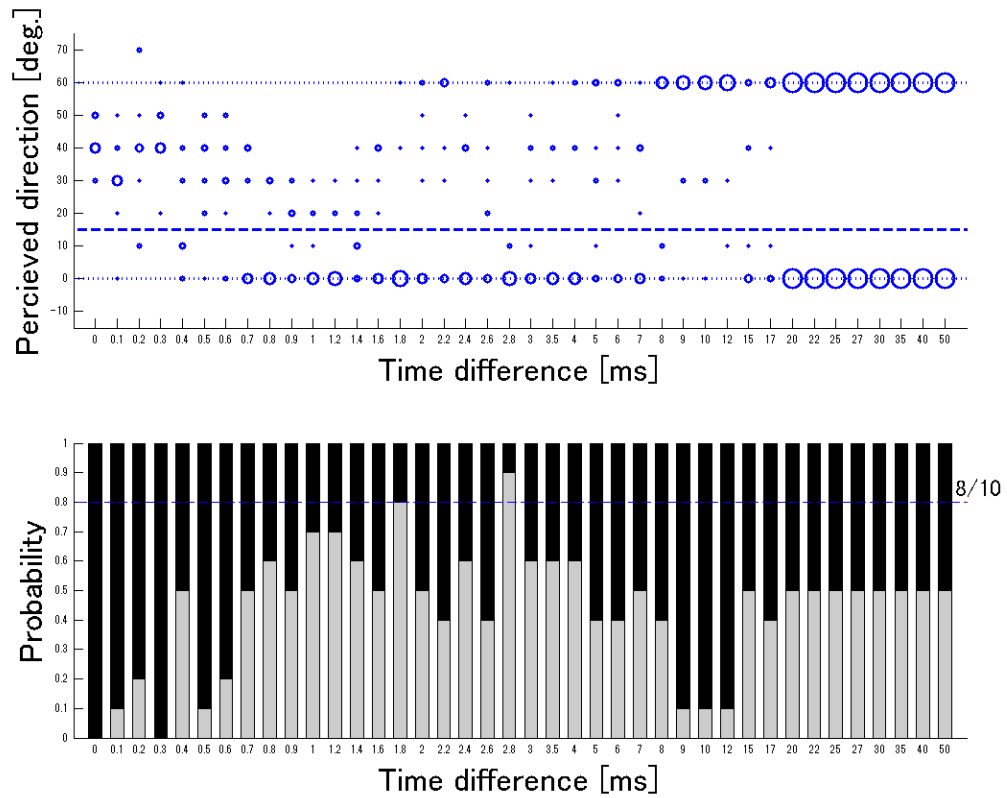


(A)刺激音：ピンクノイズバースト，音源配置： 0° ～ $\pm 60^{\circ}$ ，レベル差：0dB

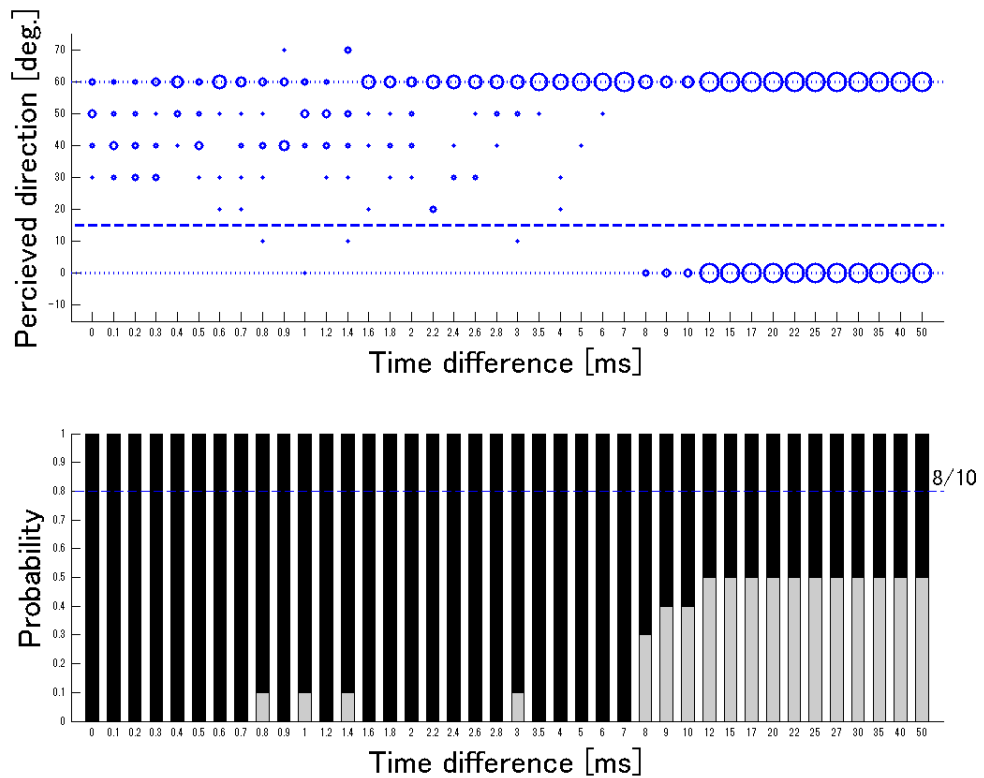


(B)刺激音：ピンクノイズバースト，音源配置： 0° ～ $\pm 60^{\circ}$ ，レベル差：3dB

図 3.11 各時間差条件に対する定位方向及び先行音効果が発生した割合

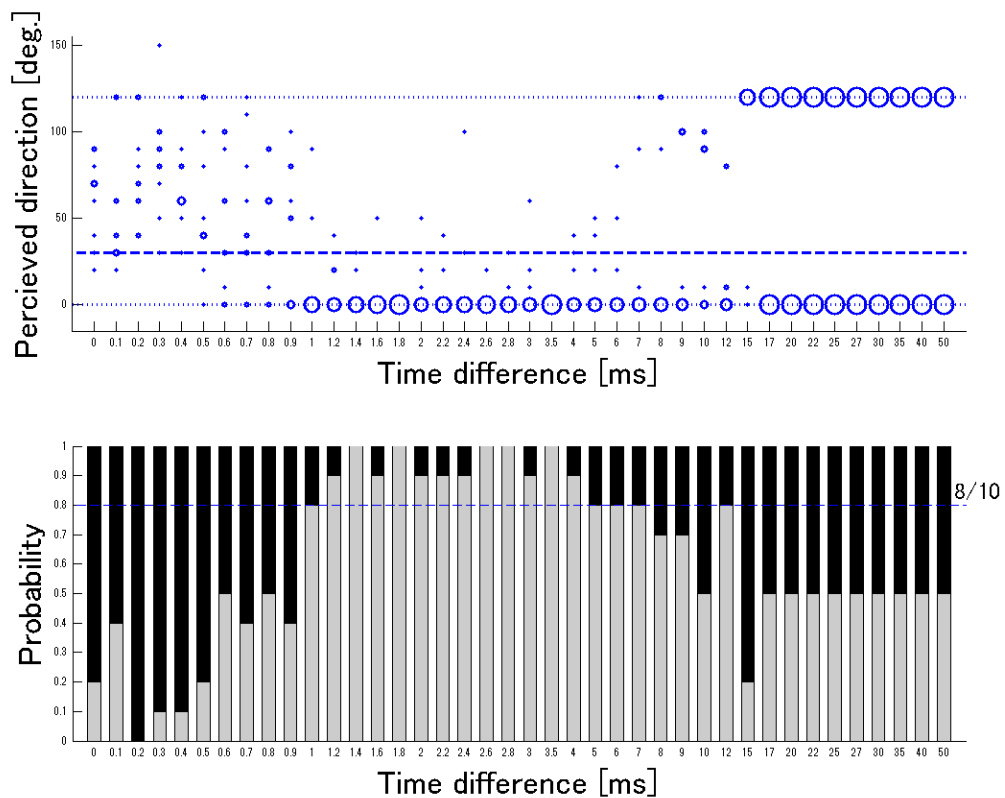


(C)刺激音：ピンクノイズバースト，音源配置： 0° ～ $\pm 60^{\circ}$ ，レベル差：6dB

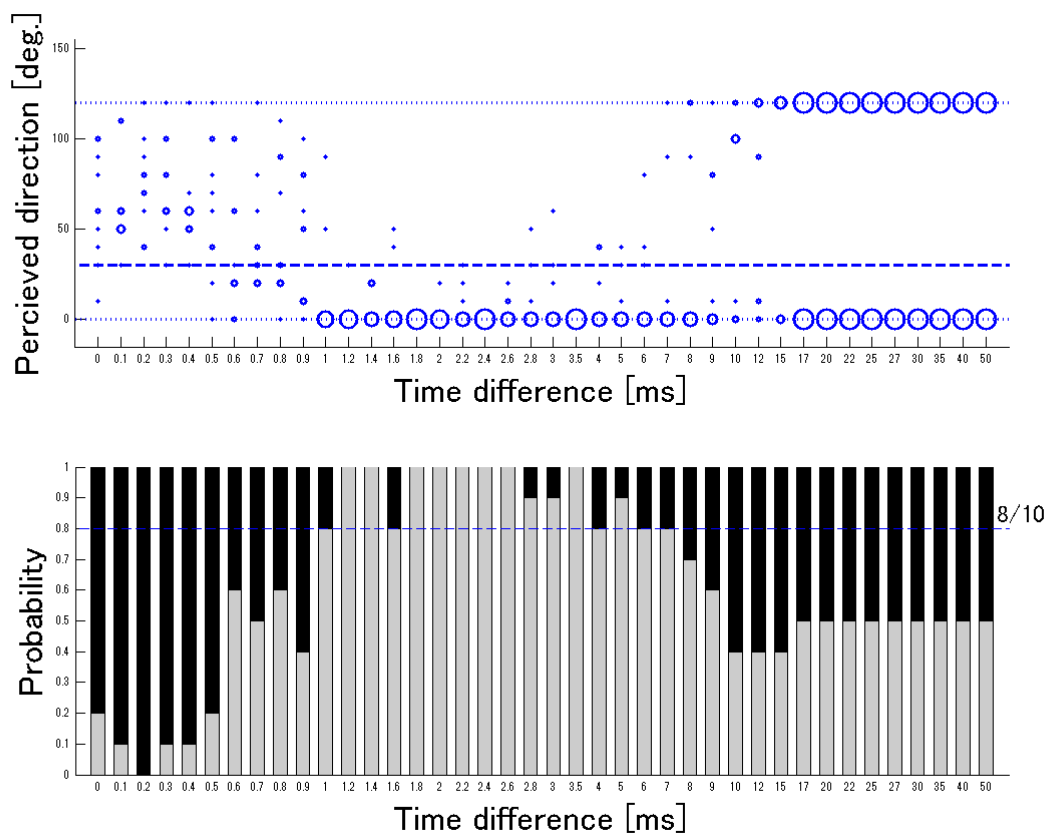


(D)刺激音：ピンクノイズバースト，音源配置： 0° ～ $\pm 60^{\circ}$ ，レベル差：9dB

図 3.11 (続き) 各時間差条件に対する定位方向及び先行音効果が発生した割合

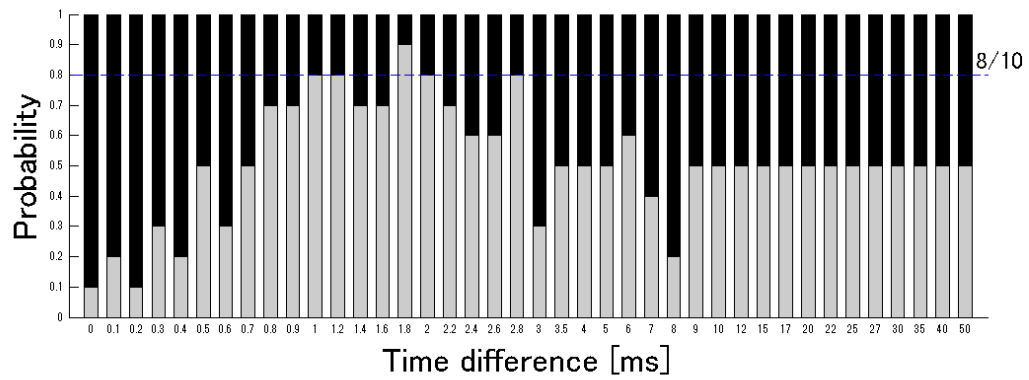
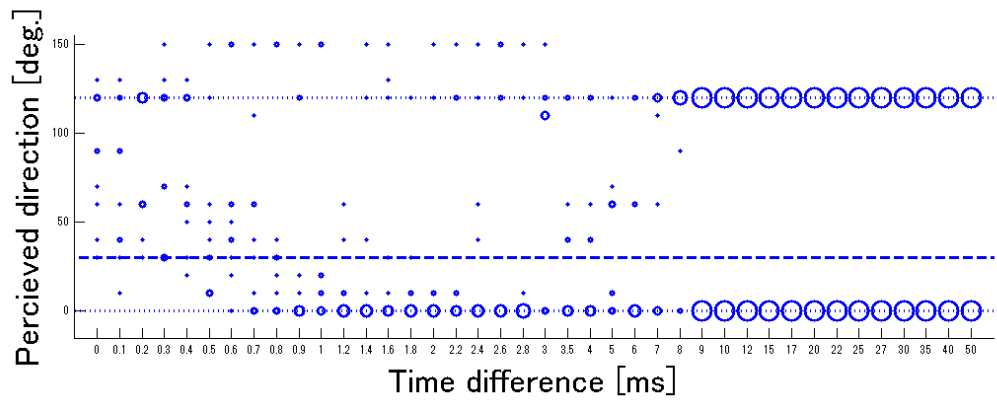


(A)刺激音：ピンクノイズバースト，音源配置： 0° ～ $\pm 120^{\circ}$ ，レベル差：0dB

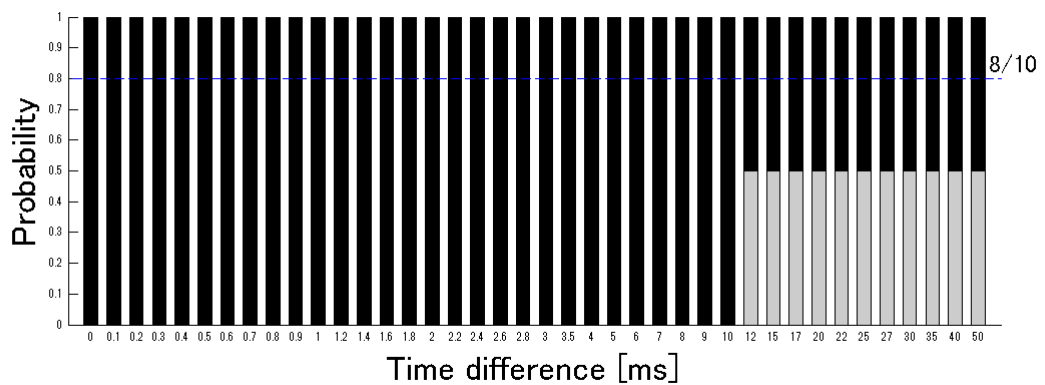
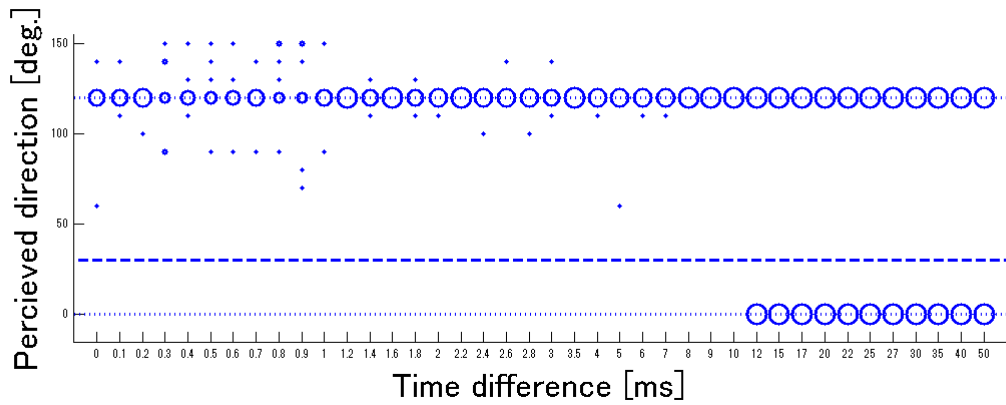


(B)刺激音：ピンクノイズバースト，音源配置： 0° ～ $\pm 120^{\circ}$ ，レベル差：3dB

図 3.12 各時間差条件に対する定位方向及び先行音効果が発生した割合

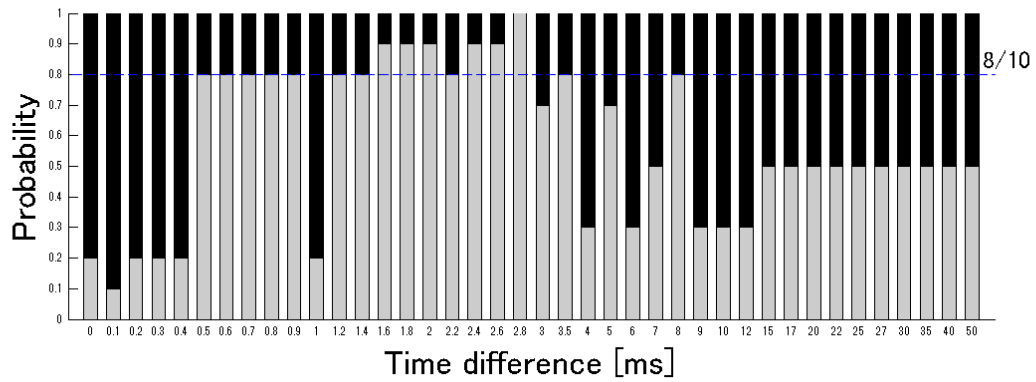
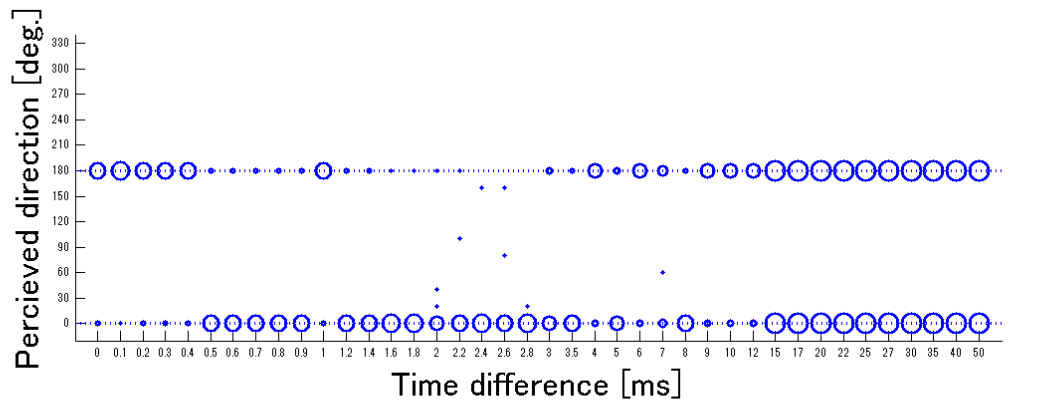


(C)刺激音：ピンクノイズバースト，音源配置： 0° ～ $\pm 120^{\circ}$ ，レベル差：6dB

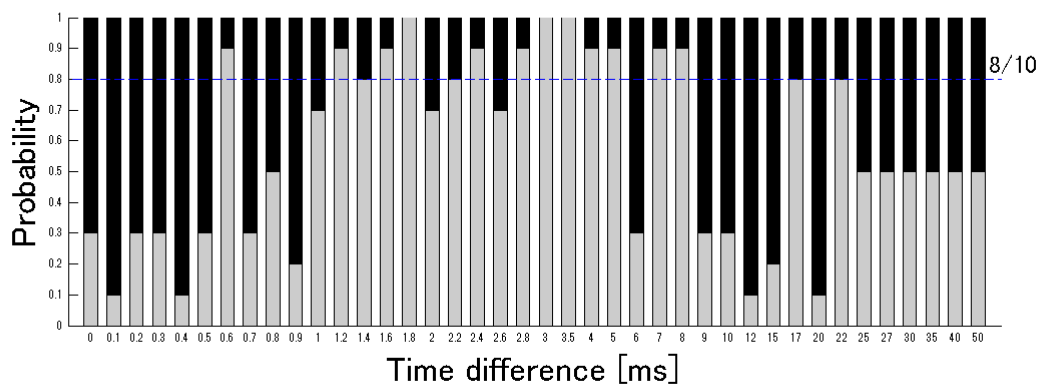
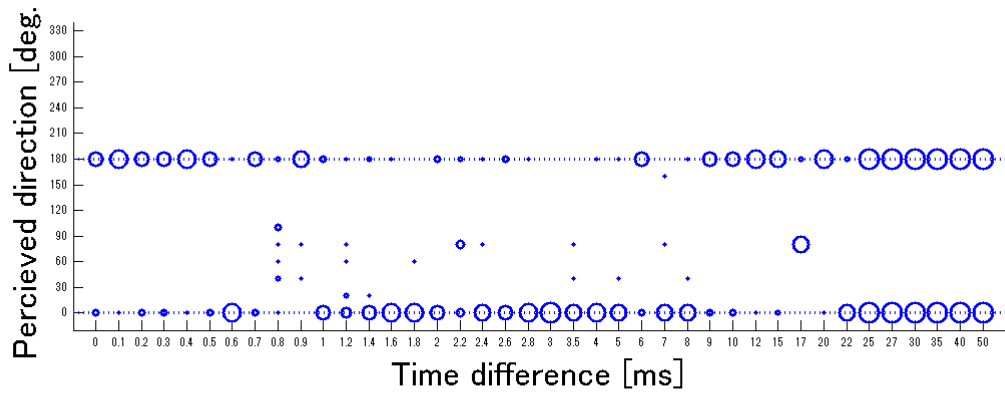


(D)刺激音：ピンクノイズバースト，音源配置： 0° ～ $\pm 120^{\circ}$ ，レベル差：9dB

図 3.12（続き） 各時間差条件に対する定位方向及び先行音効果が発生した割合

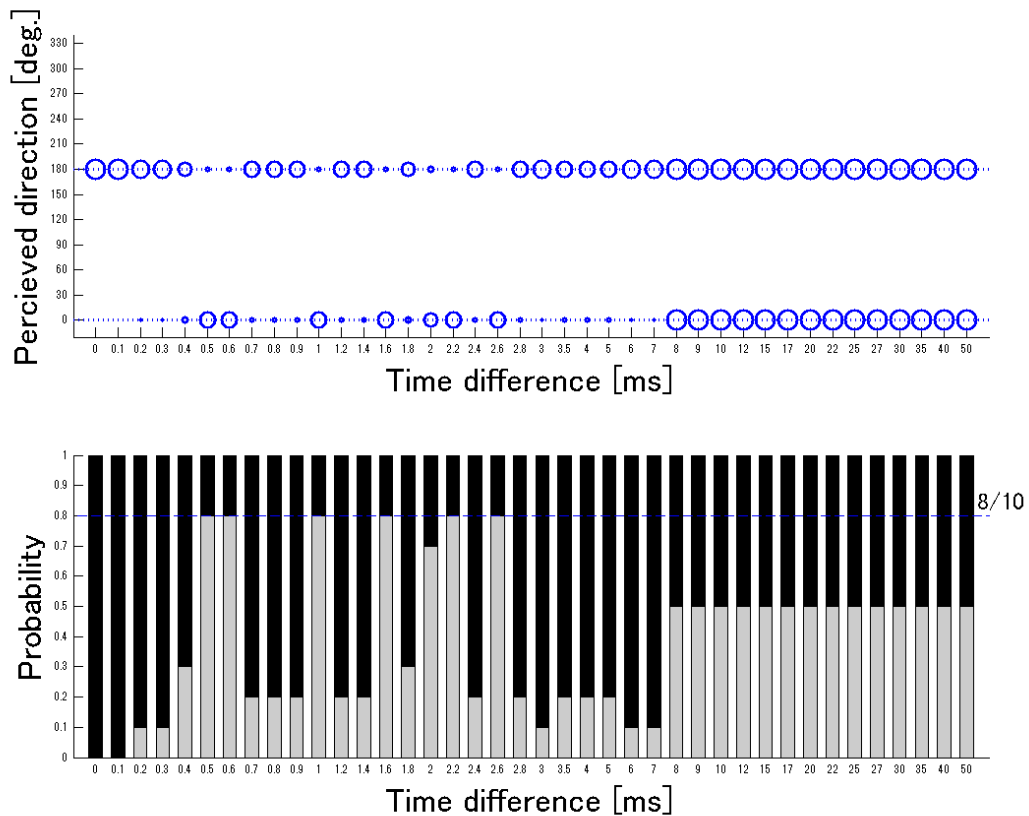


(A)刺激音：ピンクノイズバースト，音源配置： 0° ～ $\pm 180^{\circ}$ ，レベル差：0dB

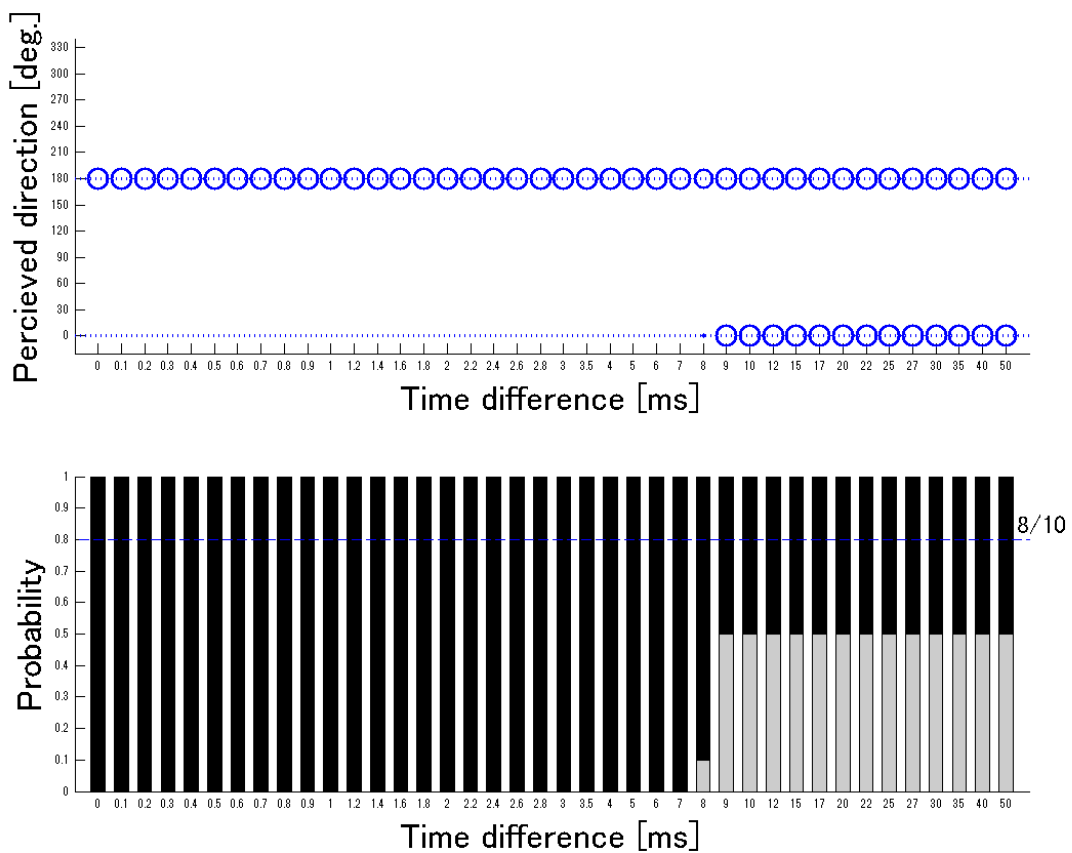


(B)刺激音：ピンクノイズバースト，音源配置： 0° ～ $\pm 180^{\circ}$ ，レベル差：3dB

図 3.13 各時間差条件に対する定位方向及び先行音効果が発生した割合



(C)刺激音：ピンクノイズバースト，音源配置： 0° ～ $\pm 180^{\circ}$ ，レベル差：6dB



(D)刺激音：ピンクノイズバースト，音源配置： 0° ～ $\pm 180^{\circ}$ ，レベル差：9dB

図 3.13（続き） 各時間差条件に対する定位方向及び先行音効果が発生した割合

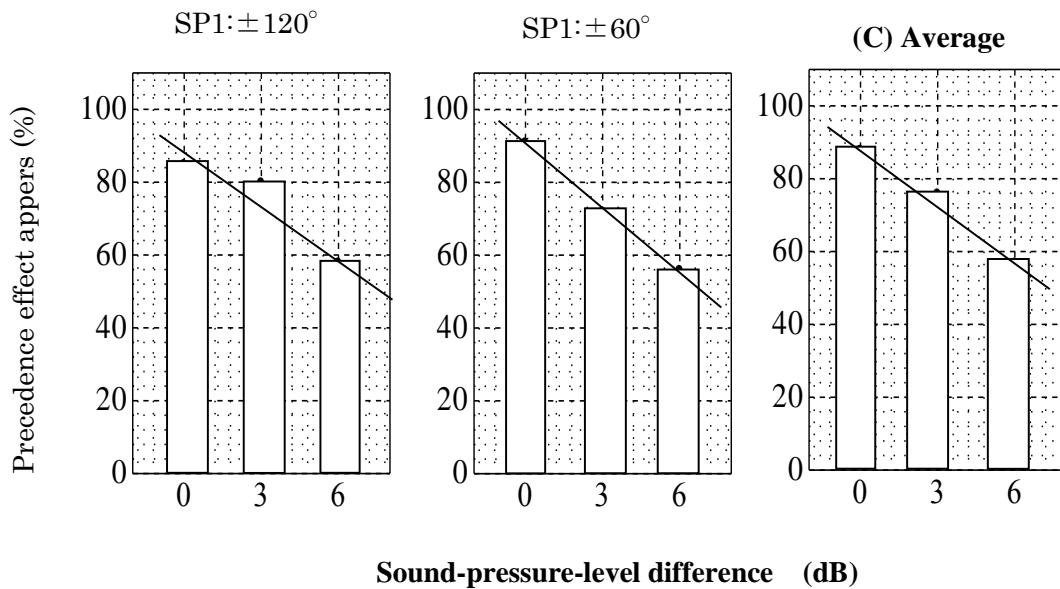


Figure 3.14 Rate of response frequency at which the precedence effect took place for sound level difference between the second sound and the first one. 0, 3 and 6 along the horizontal axis mean 60/60dB SPL, 60/63dB SPL and 60/66dB SPL in the level of sound from the main-speaker/sub-speaker. SP1: Separation angle between two loudspeakers is $\pm 60^\circ$. SP2: Separation angle between two loudspeakers is $\pm 120^\circ$. Regression lines are drawn for the average rate of response frequency in SP1 and SP2.

図 3.14 先行音と遅延音のレベル差および遅延音の提示方向による 先行音効果が生ずる割合

レベル差が 6dB 以上では，先行音効果は生じ難い，レベルが優位となる．レベル約 6dB 未満で，遅延時間差が優勢で，先行音効果が生じる．レベル差 6dB 以上で，レベル差が優勢で，先行音効果が生じ難い．レベル差 9dB の場合，先行音効果は生じない．過去の研究では，レベル差が 10dB までが遅延時間が優勢であると報告されている．本実験では，6dB 程度までという結果となり，先行音効果が生じる範囲が異なる^{62, 63, 64)}．

3.5 まとめ

本章では、刺激音の種類とレベル差の違いが、先行音効果の生起にどのような影響があるかを聴取実験を通して、SLT と ET を求めることで検討した。その結果、次の知見を得た。

実験に用いた刺激音で、先行音効果の発生する範囲は、ピンクノイズバースト音で 0.75～6 ms、スピーチで 0.7～10 ms、ギターで 1～11 ms、サキソフォンで 0.7～9 ms となった。この結果から、SLT の値は、刺激音の種類にあまり影響を受けず、平均で約 0.8 ms 程度と安定していることが明らかとなった。この知見は本研究で SLT を定義し、定量的に SLT を求めることができるようになったため得られた知見といえる。一方、SLT については、刺激音の種類による値の変動があることが示された。この知見は過去の研究でも示された結果であり、同様の知見を確認できたといえる。この変動をもたらす要因としては、刺激音が含む周波数成分や、時間波形としての立ち上がりの急峻さなどが考えられる。

刺激音のレベル差の違いが先行音効果の発生する時間領域に与える影響としては、レベル差が増加するにつれ、その時間幅が減少していく傾向があることが示された。音源配置による影響が若干見られるが、基本的な傾向としてレベルの増加に伴い、SLT は増加し、ET は減少することが示された。ただし、その変動は SLT については比較的小さいものであり、ET の減少がそれに比べ大きいものであった。このことから刺激音の種類の違いは、主に ET に影響するということが出来ることが明らかとなった。また、レベル差の増加は、SLT と ET のみに影響するだけでなく、全体的な先行音効果の発生割合を減少させることも実験的に明らかとなった。

第 4 章

先行音効果に及ぼす視覚情報の影響

4.1 はじめに

第 1 章でも述べたが、人間にとって、音の方向を知る機能は、重要な能力の 1 つである。特に、音の方向を知りたいときは、音を放射している物体に視覚的注意を向けたりして、音源の方向を視覚的にも探る行動をとる。実際のコンサートホールや講演会場においても、演奏者や講演者の姿といった視覚的な音源に関する情報が存在している。視覚刺激が音像定位に影響を与える例として、腹話術効果(ventriloquism effect)がよく知られている。視覚刺激が聴覚刺激に影響を与えるならば、聴覚現象の先行音効果が発生する条件に対しても影響を与える可能性があると考えられる。もしも、視覚刺激が先行音効果に影響を及ぼすとしたら、PA 機器への応用の際にも考慮すべきことだと考えられる。しかし、視覚刺激の影響を考慮した先行音効果の発生条件についての研究はあまりない。その中で、視覚刺激が存在する場合の先行音効果が発生する条件について研究したものに、砂田⁴²⁾らの研究がある。この研究では、同期した映像がある場合とない場合での先行音効果が生じる条件について比較し検討している。刺激音の種類は 4 種類について行われた。実験の結果、視覚刺激を付加した場合、2 音源の遅延時間差が 1ms 以下において、先行音効果への影響が見られ、先行音効果がより短い時間差で発生したと報告している。しかし、加法定位と先行音効果の境界、すなわち SLT に相当する部分の知覚判断はかなり曖昧であったため、加法定位か先行音効果のどちらが起っているのか、また視覚刺激との関係がどのようにあるのか明らかにされていない。そこで、本研究で

は、本論文で定義した加法定位と先行音効果の境界である SLT とエコー検知限 ET を視覚刺激を付加した条件のもとで求める 2 種類の実験を行い、視覚刺激が先行音効果に及ぼす影響について検討した。

4.2 聴覚と視覚の相互作用

4.1 節で視覚情報が先行音効果に影響を与える可能性があることを述べた。本節では、聴覚と視覚の知覚上の関係について更に詳しく触れる。古くから、聴覚と視覚の間で生じる相互作用については多くの研究が行われている。それらの研究から、音像定位に関わる視覚と聴覚の相互作用について概観する。

視覚と聴覚のように、情報の媒体の異なる感覚が、外界の認識を行う上で、互いにどのような作用を及ぼし合うかという問題は、心理学その他の分野で広く取り扱われている。一言に、音像定位に関する、視覚と聴覚の相互作用といっても様々な側面が存在する。丸山^{43,44)}、盛永⁴⁵⁾らは、腹話術効果とは反対に、光に対する感度が音によって影響を受けることを実験的に示している。それによると、音が視感覚の感受性に及ぼす影響は、促進、抑制、無影響の3種類に分類され、どのような影響を受けるかは、音の継続時間、音の周波数成分、音色などに依存しているという。促進は、閾値が下がること、つまりよく見えるようになること、あるいはよく聞こえるようになることで、抑制は、逆に、閾値が上がって、見えにくくなること、あるいは聞こえにくくなることを意味する。高音では促進作用があり、1k~3kHz程度が最も効果が大きいとしている。周波数が1kHz以下になると促進作用は弱くなって、音の影響が現れにくくなる。

また、直接音と反射音についてのエコー検知限に対する視覚刺激の影響については、ホールなどの空間で、音を聴く場合に妨害要素として働くエコーが、視覚刺激によって、聞こえにくくなるという森本ら¹⁴⁾の研究がある。その実験では、直接音と同一レベルの

反射音を用いて、反射音が直接音と分離し始めるエコー検知限を求めており、視聴覚刺激として、リコーダ演奏及びスピーチの映像とその演奏音を用いている。その結果、映像と直接音の方向が同方向で、反射音の方向が異なる場合、映像の付加により、エコーの検知限が約5 ms程度長くなることが示されている。これは、視覚刺激により、聴覚のエコーに対する感度が鈍くなって、エコーの存在が分りにくくなることを示している。

また、4.1節でも述べたが、音像定位における視覚の優位性を示す現象として、腹話術効果がある。腹話術師は、人形の役の時は、自分の唇を動かさず声色を変えて喋っているのであるが、人形の口をパクパク動かすことにより、聴衆は、あたかも人形が喋っているように感じる。腹話術効果は、テレビ、コマーシャルをはじめマルチメディアに用いられる音と映像などでも検討されている¹⁹⁻²¹⁾。視覚と聴覚の相互作用の問題は、従来、心理学の基礎的な研究対象であったが、テレビ音声のデジタル化、ビデオやレーザー・ディスクの普及、デジタル・テレビ画面の大型化、高品質化を反映して、視聴空間における問題に関心が持たれるようになった。特に、音像の定位や、音と映像のズレの問題などに積極的な取り組みがなされている。音像の定位に現れる視覚と聴覚の相互作用に関しては、アナウンサーが喋っている映像と音の左右方向の食い違いが、どう知覚されているのかを調べたものなどがある²²⁾。それによると、映像に対して音源の方向が0° から10° 程度ずれた範囲は、映像の支配領域で、音は映像に引き寄せられてしまう。10° から20° にかけては、映像と音の競合領域で、やや映像に引き付けられるが、20° 以上では、音は映像の影響から開放されて、実際に音が出てきた方向から聞こえる。また、映像の影響が大きい領域では、音が映像に引き付けられる力は、画面が大きくなるほど強くなる。音が映像に引き付けられる影響の程度は、映像への注意力によっても異なる。喋っている人が同性の場合は影響が小さいが、異性の場合には影響力が強いという実験結果もある。音が何処から聞こえるかという、音像定位に関する実験では、このような視覚優位の傾向が見られている。つまり、聴覚と視覚が矛盾した情報を受け取

ったとき、中枢系は、対象を一つにまとめようとする。このとき、視覚が、基準系となりがちである。

腹話術効果は、聴覚情報による知覚より視覚情報による知覚が優先される現象であるが、視聴覚の相互作用には、視覚の知覚でもなく聴覚による知覚でもなく、両方の情報が存在することで異質な知覚が生じるような現象も存在する¹⁹⁾。例えば、言葉の認識では、視覚に影響される現象として、マガーグ効果¹³⁾がよく知られている。「/ba/」という音声とあわせて「/ga/」と発音している人の映像を見せられると、「/da/」と聴こえるという現象であり、音の情報と唇の動きの情報が融合して別な音に聞こえるものである。この効果は、言葉を聞き取るときに、相手の唇の動きが知覚に重要な働きをしていることを示している。特に、まわりが騒がしい場合のような、言葉が聞き取りにくい状態では、視覚の影響はかなり大きいと考えられる。

このように、異なる感覚の矛盾する情報を統合しようとする能力は、かなり高次の部位での処理であると考えられている。その特徴的な傾向として、聴覚と視覚に同期した刺激によって感化する共鳴現象がある。これは、暗い音によって映像の印象が暗くなり、明るい音によって映像が明るく感じられるという現象である。つまり、視覚情報を単独で提示した場合と、視聴覚両情報を与えた場合の印象の変化で、明るい音は、映像の印象をより明るく、暗い音は、映像の印象をより暗くするという傾向が見られる。たるんだ音が、映像の印象をたるんだ方向に、引き締まった音が映像を引き締まった方向に変化させる。こういう聴覚から視覚への共鳴現象は広く見られている^{20, 46)}。音と映像の調和とは関係なく、明るい音が映像を明るくし、明るい映像が音を明るくする。視覚的明るさは、網膜に達するエネルギーと対応し、聴覚的明るさは1kHz～3kHzの周波数帯域と対応するとされている。これに対して、引き締まりや美的因子においては、ある程度、調和のとれた音と映像の組合せでないと、共鳴現象が生じにくい。聴覚から視覚の方向への共鳴現象として、例えば、「きれいさ」に関係する美的因子だと、ノイズとか歪と

かの要素が関与してくる⁴⁷⁾。こういった、妨害成分と信号成分を分離して認識する機能は、ある程度上位のレベルで処理されているといわれる。引き締まり因子も、時間的な処理に関わってくると考えられる。共鳴現象は、いずれも、聴覚から視覚の方向への方が顕著であると考えられている。定位の場合においても、情緒的な側面の相互作用には、聴覚優位の傾向が強いことが示され、感覚の感受性の変化に現れる相互作用では、映像の付加が音の汚さを目立たなくさせるような感覚の感受性の変化にも、相互作用は現れる。これは、例えば、音を刺激として与えることによって、視覚の感度が増す現象（よく見えるようになる）の存在が、心理学の分野で認められている^{48, 49)}。AVメディアの情報伝達に、この現象をあてはめて見ると、映像を音と同時に提示することによって、音響再生系の音質に対する聴覚系の感受性(感度)が低下し、音質の悪さに対して寛容になることが考えられる^{50, 51)}。あるいは、逆に、聴覚系の感受性が上昇し、音質に対して厳しい判断をするようになるという場合もあるかも知れない。

以上の述べたことを踏まえて、先行音効果の現象に対して視覚刺激を付加したとき、どのような影響があるのかを明らかにしたい。

4.3 実験概要

本章では、3 章とは違い、SLT と ET を別々の実験で検討する。以降、SLT に関する実験を SLT 実験とし、ET に関する実験を ET 実験とする。それぞれの実験に対し、視覚刺激の「有／無」によって閾値の値が変化するかを検討する。

4.3.1 視聴覚刺激

SLT 実験及び ET 実験に用いる視聴覚刺激は共通のものとなる。視聴覚刺激としては、スピーチ（男声および女声）、太鼓を叩く音及びギター演奏音の 4 種類を用いた。

刺激音の選択理由は、次のとおりである。まず、スピーチは、日常、われわれが接する機会の最も多い視聴覚刺激であり、音と同期した視覚刺激があると、腹話実効果と同様な現象が起ることを想定して選択した。小太鼓の音は、視覚的に音の出るタイミングが捉えやすく、聴覚的には、衝撃的な音でスペクトルが広帯域で時間的な立ち上がりな明快なものとして選択した。ギターの演奏音は、3章の実験結果から、時間的変動はゆるやかで、かつ周波数特性が比較的広帯域であるが、合成音像定位が不安定な様相を示したことから、視覚刺激の影響が大きくなる可能性があると考え、選択した。いずれの音も比較的身近にあるもので、過去耳にした経験があると思われる刺激音となる。

各々の刺激音に対する映像（視覚刺激）は、スピーチの場合、それぞれ成人男性及び女性が椅子に腰掛けて喋っている情景を、小太鼓の場合、立った状態で太鼓を叩いている情景を、ギターの場合、椅子に座ってギターを演奏している情景を撮影した。実験の音と映像は、半無響室内でデジタルビデオカメラレコーダ（SONY DCR-VX9000）を用いて同時に収録した。スピーチの内容は、ある日の天気概況の朗読で、内容は「天気概況です、日本海付近を前線や低気圧が通過し、再び冬型の気圧配置に戻るでしょう」となっている。読みあげにかかった、刺激の継続時間は約7秒である。太鼓は9回連続に叩いた情景を、ギターは適当なワンフレーズを演奏した情景を撮影した。それぞれ継続時間は約5秒となる。音圧レベルは、被験者の頭部の中心位置で60 dBとなるように設定し、視覚刺激は、被験者が着座した上後方から、音と同期したほぼ原寸大となる動画像を被験者正面のスクリーンに投映した。撮影した動画のキャプチャ画像及び録音した刺激音の波形を図4.1～図4.8に示す。



図4.1 Speech 話者(男性)の静止画像

天気概況(男性の声)と動画映像を提示

- Speechは、男声の天気概況の朗読で「天気概況です。日本付近を前線や低気圧が通過し、再び冬型の気圧配置に戻るでしょう」という文章を朗読した。
- Speechは、被験者から3mの位置にあるスクリーン(H:2.6m、W:2.5m)に映された動画映像である。
- 天気概況の朗読は、継続時間7秒間であった。
- SLT実験の場合、後続音の時間差は、12ポイントを用意した。
- ET実験の場合、後続音の時間差は、30ポイントを用意した。
- 音と映像は、同期させた。

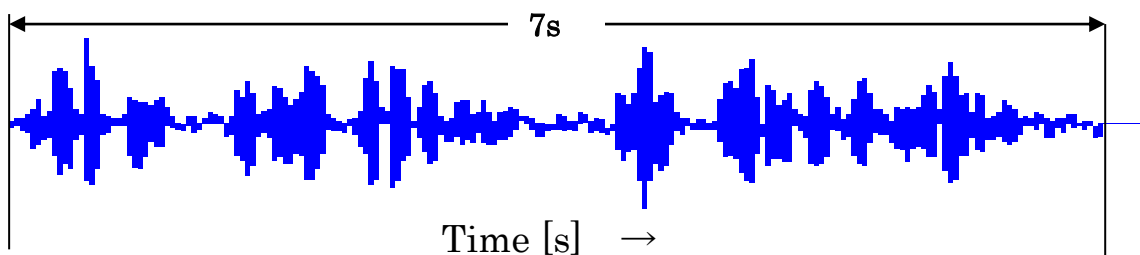


図 4.2 男声のタイムパターン



図4.3 Speech 話者(女性)の静止画像

天気概況(女性の声)と動画映像を提示

- Speechは、女声の天気概況の朗読で「天気概況です．本付近を前線や低気圧が通過し，再び冬型の気圧配置に戻るでしょう」という文章を朗読した．
- Speechは、被験者から3mの位置にあるスクリーン(H:2.6m、W:2.5m)に映された動画映像である．
- 天気概況の朗読は、継続時間7秒間であった．
- SLT実験の場合、後続音の時間差は、12ポイントを用意した．
- ET実験の場合、後続音の時間差は、30ポイントを用意した．
- 音と映像は、同期させた．

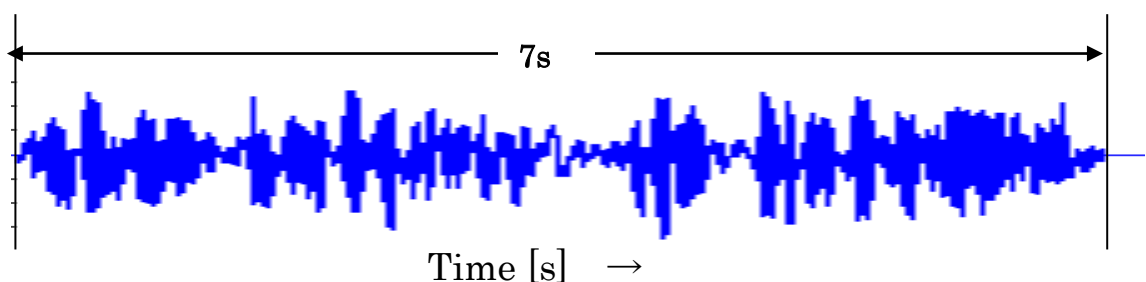


図 4.4 女声のタイムパターン



図4.5 小太鼓演奏者の静止画像

小太鼓を叩く演奏者の動画像を提示

- 小太鼓を演奏音は，被験者から3mの位置にあるスクリーン(H: 2.6m, W:2.5m)に映された動画像である.
- 小太鼓を叩く音は，継続時間 5秒間であった.
- SLT実験の場合，後続音の時間差は，12ポイントを用意した.
- ET実験の場合，後続音の時間差は，30ポイントを用意した.
- 音と映像は，同期させた.

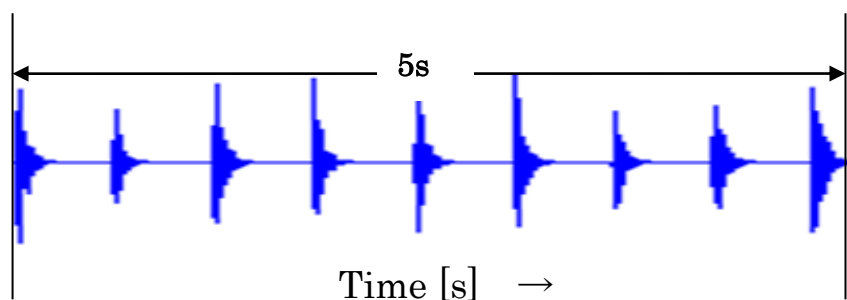


図 4.6 小太鼓の音のタイムパターン



図4.7 ギター演奏者の静止画像

ギター演奏音と動画像を提示

- ギター演奏音は、被験者から3mの位置にあるスクリーン (H:2.6m, W:2.5m)に映された動画像である.
- ギター演奏音は、継続時間 5秒間であった.
- SLT実験の場合、後続音の時間差は、12ポイントを用意した.
- ET実験の場合、後続音の時間差は、30ポイントを用意した.
- 音と映像は、同期させた.

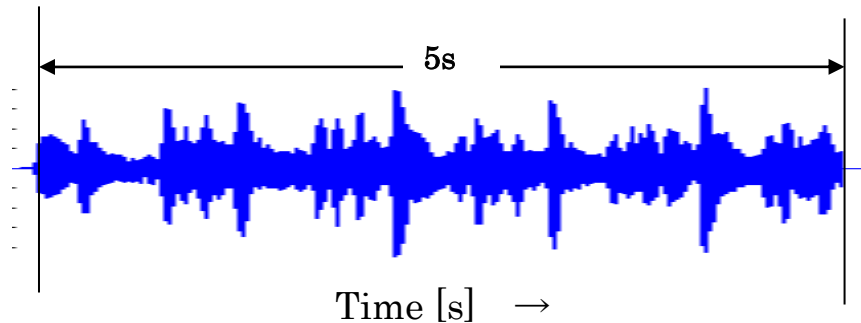


図 4.8 ギター音のタイムパターン

4.3.2 実験系

視聴覚刺激の提示に用いる実験系も SLT 実験及び ET 実験で共通となる．実験系の概要を図 4.9, 4.10 に示す．実験は、縦 7.5m、横 6.0m、高さ 4.6m の半無響室で、床面に吸音マットを敷いた状態で行った．被験者は椅子に腰掛け、その頭部をヘッドレストで軽く固定している．スピーカは、図 4.9 に示すように、被験者を中心とする半径 3 m の円周上に 30° 間隔で配置した．その高さは床上 1.3 m である．音源からの直接音に相当する音を放射するためのスピーカを正面に、後続音を放射するスピーカを左右 30° 間隔で 6 個配置した．以後、被験者正面のスピーカをメインスピーカ (M_{Sp0}) と呼び、その他をサブスピーカ (S_{Sp1-6}) と呼ぶ．サブスピーカは、正面を 0° として、右側 30° のスピーカを S_{Sp1} , 60° を S_{Sp2} , 90° を S_{Sp3} , また、左側 -30° を S_{Sp4} , -60° を S_{Sp5} , -90° を S_{Sp6} と表示する（図 4.9 参照）．被験者が視覚的にスピーカの位置を知ることができないように、配置したスピーカと被験者の着座する椅子の間にスピーカが見えないように音響的にほぼ透明とみなせる薄い布を垂らした．視覚刺激は、プロジェクタ (EPSON ELP-810) を用いて被験者正面の M_{Sp0} の前に設置したスクリーン (高さ 2.6m, 幅 2.5m) に音と同期したほぼ原寸大となる動画像を投射することによって与えた．この実験系で実験を行い、被験者には、データ収集装置を用いて、知覚した音像方向をレーザポインタで直接回答させるタスクを課した．ポインタで音像の方向を指示し、回答が確定したら、決定ボタンを押すことにより、回答をコンピュータに転送している．本章では SLT 実験及び ET 実験ともに、現実的なホールや講演会場の拡声装置を想定し、常に被験者正面の M_{Sp0} から先行音が放射されるように実験条件を設定した．したがって、遅延する音刺激は、常にサブスピーカ S_{Sp1-6} の中のどれか 1 個から提示される．図 4.10 に示す提示系を用いて、 M_{Sp0} の信号はアンプから直接、 S_{Sp1-6} への信号は、遅延を与えた後、チャンネルセレクトを介し、出力した．

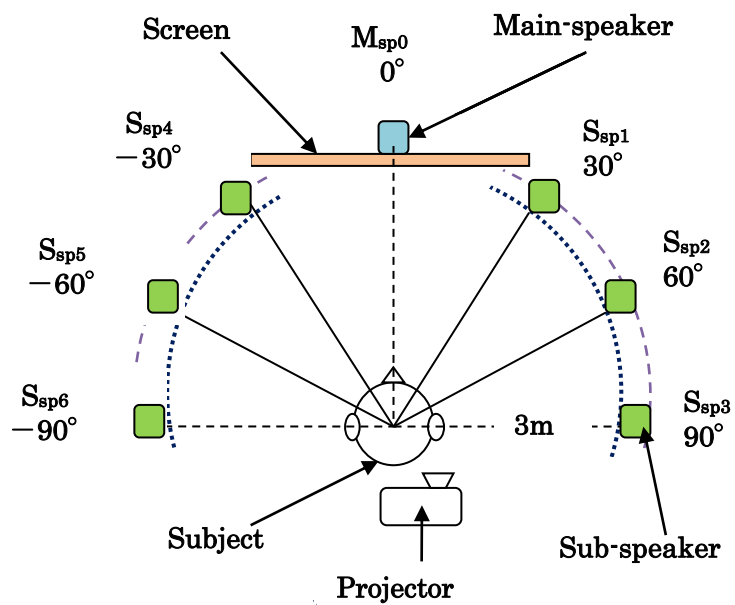


図 4.9 被験者とスピーカ 配置

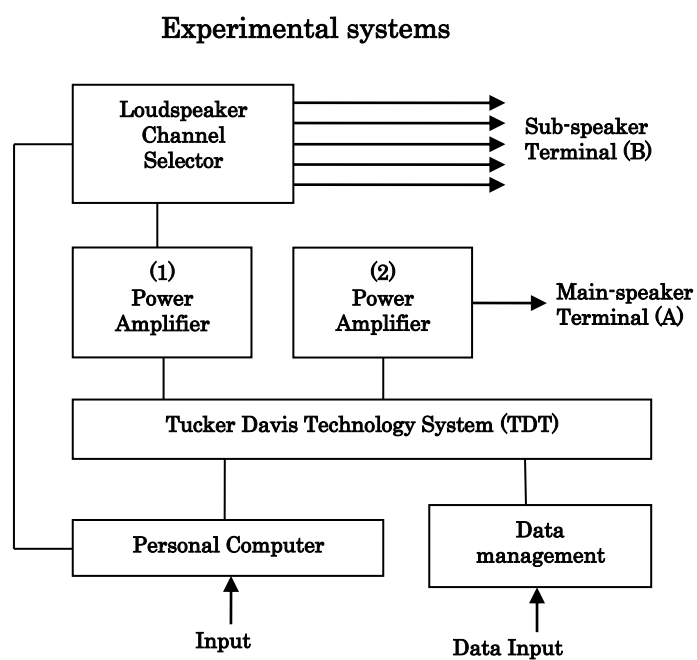


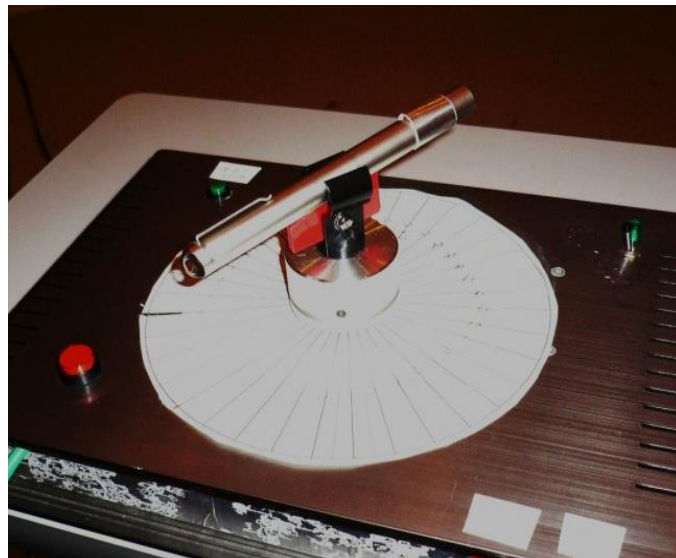
図 4.10 実験装置の構成

4.3.3 刺激音の再生と回答データ収集装置

図4. 10に示したように、実験を実施する無響室の外部制御室から、スピーカケーブルとデータ収集装置のケーブルが無響室内部に引き込まれている。2音源間の時間差はあらかじめ所望の時間差を持つように刺激音データの前に必要な分だけ振幅0のデータを付加したPC上の2つの刺激音データをTDT上のメモリに転送・再生することで制御した。このように生成した音の出力を2台のパワーアンプを通してメインスピーカとサブスピーカから音を提示している。ここで、メインスピーカへの出力はパワーアンプから直接接続したが、サブスピーカへの出力は所望のサブスピーカから音を出力するために、パワーアンプからチャンネルセクタを介しサブスピーカへ接続されている。また、精度の良い方向の回答を収集するために、レーザポインタで音が聞こえてきた方向に向け、決定ボタンを押すことにより被験者の回答した方向を記録できるデータ収集装置を作成した。データ収集装置の外観を図4. 2に示す。レーザポインタを設置した回転軸には、可変抵抗が接続されている。可変抵抗に対して一定電圧を付加しておくことで、レーザポインタの回転に合わせて変化する電圧から指示されている方向を得る構造となっている。具体的に被験者が回答した方向は、PCに電圧値として与えられる。したがってこの電圧値から実際の角度に変換する必要がある。まず、出力される電圧値はTDTのAD変換によって値が変換されてPC内部に取り込まれる。実験時にデータ収集装置を使用している様子とデータ収集装置の外観を図4. 11に示す。被験者の頭部中心から60 cm前方にデータ収集装置のレーザポインタの中心が配置されている。そのため、レーザポインタが指示する角度は被験者の知覚した角度とは異なっているので、その差異は実験後に補正している。



(a) 実験中の被験者 風景



(b) データ収集装置の外観

図 4.11 実験中の被験者とデータ収集装置の外観

各スピーカは，視覚的効果を排除するため，被験者から見えないようにカーテンで蔽っている．被験者の手元にあるデータ収集装置で，知覚した音像方向をレーザポインタで直接回答する．被験者の正面には，スクリーン(高さ 2.6m，幅 2.5m)が設置されている．データ収集装置は，レーザポインタの指示が $\pm 1^\circ$ 間隔の精度で得られるようにした．

4.4 SLT 実験

4.4.1 実験条件と回答方法

SLT 実験では、2 音源の遅延時間を 0.1~1 ms の間は 0.1 ms おきに設定し、その上に 2 ms, 3 ms の時間遅れをとって、全 12 条件とした。2 音源間の時間差をまとめたものを、表 4.1 に示す。

被験者には、提示された映像を注視しながら、スピーカから提示される刺激音を聞き、聞こえた方向を回答することを求めた。まず、実験内容を口頭で説明し、知覚した音像の方向を指示するデータ収集装置の取り扱いを説明した後、回答の練習を 10 回程度行った。全刺激に対し、視聴覚刺激を提示後、3 秒間の回答時間を設けた。回答は、音が聞こえた方向にレーザポインタを向けて、音の位置が確定したら、第一決定ボタンを押すことで得られるが、もしも、音像が 1 点に集約せず広がって知覚された場合、広がった音像の中心を回答するように教示した。また、被験者が回答時間内に回答できなかった場合、もう一度同じ条件の視聴覚刺激を繰り返し提示することによって、必ず回答を得るようにした。実験は、1 セット 60 試行で行われた。後続音の提示方向は 6 方向、刺激音 4 種類、視覚刺激の有無、時間差 12 種類、繰り返し各 5 回で、実験を通して、被験者一人当たりの試行は、 $6 \times 4 \times 2 \times 12 \times 5 = 2880$ 回である。被験者は、大学生で健聴者 5 名、その内 3 名が男性、2 名が女性である。

表 4.1 SLT の実験に用いた時間差

実験で用いた 2 音源の時間差					
0.1	0.2	0.3	0.4	0.5	0.6
0.7	0.8	0.9	1.0	2.0	3.0

被験者に提示した時間差：12 point, 単位[ms], 表の時間差を持つ刺激を被験者にランダムに提示し, 知覚した音像方向をデータ収集装置(レーザポインタ)により収録した.

4.4.2 SLT 実験の結果

各音源配置での聴取実験の結果を図 4. 12~4. 14 に示す. これらの図は, 左右対称な音源配置における音像定位は対称であるとみなし, 0° と $\pm 30^\circ$, 0° と $\pm 60^\circ$, 0° と $\pm 90^\circ$ の音源配置に対して, マイナス側つまり左側の定位を, プラス側つまり右側へ回答データを折り返して集計した結果である. 各図の横軸は, 2 音源間に設けられた時間差を示している. 3 章の結果と同様に, 各図の上のパネルは各時間差条件に対して被験者が回答した音像の定位方向をバブルプロットで描いたものである. 一方, 縦軸は音像の定位方向を示している. 同様に, 図の下のパネルは, 2 章での知見を基に, 先行音源方向から 2 音源間の開き角度の四分の一内側に音像が含まれた場合に, 先行音効果が生じたと考え, 先行音効果の割合を算出したもの示している. したがって, 0° と $\pm 30^\circ$ の場合はその境界が 7.5° , 0° と $\pm 60^\circ$ の場合は 15° , 0° と $\pm 90^\circ$ の場合は 22.5° となる. 各時間差に対する試行回数が 50 回となっているため, 片側の 2 項検定に基づき, 図中に示した線のように先行音効果の発生した割合が 0.6 以上となった場合に先行音効果が発生するとみなす. 3 章と同様に, この考えに基づき, 先行音効果の発生割合が 0.6 以上となる最小の時間差を本実験で得られる SLT とする. 各音

源配置の SLT を算出し、その差をとることで視覚刺激の影響をまとめた結果を表 4.2 に示す。表 4.2 は全被験者 5 名に対して得られた SLT の値を平均したものとなっている。視覚刺激の影響は、映像無し (OFF) の SLT から映像有り (ON) の SLT を引いた値で評価する。

図 4.12～図 4.14 の結果を集計した表 4.2 から分かるように、音源配置が 0° と $\pm 30^{\circ}$ で、各刺激音がスピーチ（男声・女声）・小太鼓・ギターの場合、この 4 種類の刺激音から視覚刺激が無い場合の SLT である SLT_{off} の値が約 0.5ms 程度と算出された。一方、視覚刺激を付加した場合の SLT である SLT_{on} の値は、各刺激音において約 0.3ms 程度となることが算出された。したがって、視覚刺激の存在によって、約 0.2ms 程度先行する音源側に映像が引き寄せられ、短い時間差で先行音効果が生じることが示された。しかし、音源配置が 0° と $\pm 60^{\circ}$ では、 SLT_{off} が約 0.5ms 程度で、 SLT_{on} で約 0.45ms 程度なり、視覚刺激の影響は約 0.05ms と小さい。音源配置が 0° と $\pm 90^{\circ}$ においても、この傾向は同様に見られた。

個人による差異が無いかどうかを考慮して、視覚の有無の影響を検討するために、各個人に対して視覚刺激が加わることによって、各刺激における SLT が大きくなるのか、小さくなるのかをまとめた結果を、表 4.3 に示す。表中の各刺激の下に示した①が「視覚刺激を加えた結果、SLT が短くなった被験者の割合」、②が「SLT が長くなった被験者の割合」を示している。したがって、映像が ON 及び OFF の条件間で SLT に差が生じなかった被験者は、この表ではカウントされていないため、①と②の和が 100 %にならないこともある。図中で示した網掛けの部分は、視覚刺激によって 6 割以上の被験者、つまり 5 人中 3 名以上が同じ傾向を示したことを表している。この結果から、全刺激音において、大半の被験者で、先行音効果への移行が短い時間差で起こっていることが分かる。また、SLT が長くなる傾向を示す被験者は同じ人物であり、先行音効果に対する視覚刺激の影響が、他の被験者とは異なっていた。この現象については、以

下のような理由が考えられる．聴覚と視覚には，相互の感覚に対する注意を促進したり抑制したりする働きがある．すなわち，能動的(意図的に注意を向ける)あるいは受動的(目立った刺激に対して付随的に注意を向ける)に働く注意力によって，視覚刺激の付加の影響が変化するとされている．この被験者の場合，音刺激に注意を集中して音の方向を判断しているものと考えられ，音優位の制御が強く働いたものと想定される．視覚刺激が無い場合の SLT_{off} は，Haas⁵⁾や Wallach⁴⁾によって，1ms 程度であることが示されている．本実験の結果では，スピーチ，小太鼓，ギターで 約 0.5 ms 程度(SLT_{off})であり，4 種類の刺激音と 2.3 章でも求めた SLT_{off} も若干短いように見えてはいるが，その傾向は同様であると言える．また，刺激音に対する視覚刺激の影響では，音が視覚刺激に引き寄せられる視覚優位の現象で，古くから知られている腹話術効果に類似した現象であると捉える．本実験では，音源配置が $\pm 30^\circ$ で視覚刺激の影響が顕著に現れ，先行音効果の下限 SLT_{on} に対応する時間差に約 0.2 ms 程度短く知覚される．特に，音源配置が映像の近くにある場合，視覚刺激の影響が大きい．本実験においても，音源配置が $\pm 60^\circ$ ， $\pm 90^\circ$ よりも $\pm 30^\circ$ 方向で視覚刺激に強く引き寄せられる傾向がある．SLT に対する影響は，映像と音源の距離が約 30° 程度離れた距離が限界と思われ，それ以上離れている場合には，視覚刺激の影響はそれ程起こらないものと言える．

図 4.15～図 4.26 に各被験者の視覚刺激無しの SLT と図 4.27～図 4.38 に視覚刺激有りの SLT の結果を示した．これらの図から SLT を求めた結果を，表 4.4 に各被験者の視覚刺激有無に対する SLT と視覚刺激有りから視覚刺激無しの SLT を引いた値を示した．表中では，視覚刺激がある場合に SLT が短くなった被験者の文字を赤文字で，視覚刺激がある場合に SLT が長くなった被験者の文字を青文字で，変化が見られなかった被験者の文字は黒文字で示している．その結果，ほぼ全ての被験者，ほぼ全ての刺激で視覚刺激が有る場合に，SLT は短くなっていることが分かる．わずかな組み合

わせで SLT が長くなる条件も見られるが、それは男性スピーチの結果でしかみられず、被験者も限定されることが分かる。おおむね SLT が短くなっていることから、視覚刺激が有る場合には、腹話術効果に類似した現象が生じ音像が正面方向に引き付けられていると考えられる。

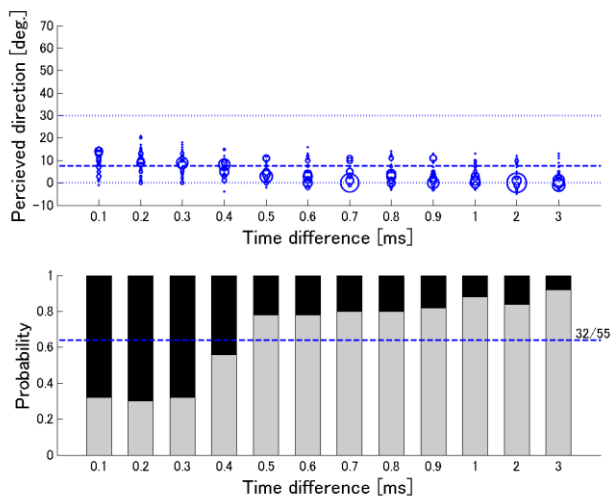
4.4.3 考察

音像の方向が視覚刺激の影響を受けて変化する現象を腹話術効果と呼ばれている。従来の研究で Klemm⁵¹⁾による電話機を用いた実験がある。その検討では、被験者の前に電話機を置き、音を別の隠した電話機で鳴らした場合、2 台の電話機を 12 度以上離さなければ、音と電話機が分離しないことが示された。その研究では、音源と映像が 11 度ずらした映像を用いた場合には、その判断割合(音と電話機は同じ場所か分離したか)が 48%程度となり、その範囲は 11 度から 20 度程度であることも示されている。また、この値は音響の専門家と非専門家ではかなり差があると報告している。

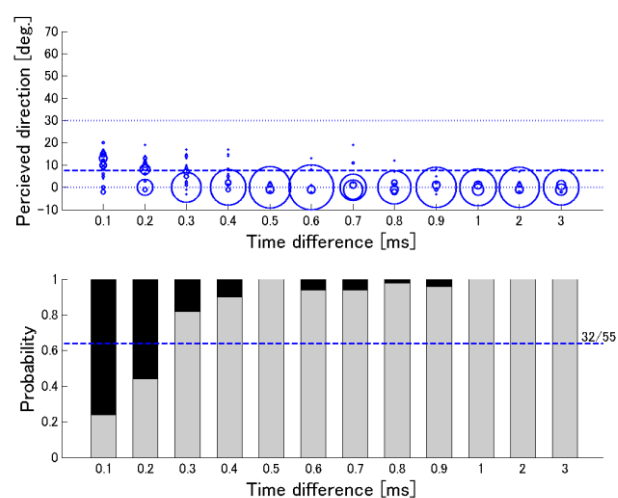
Jack&Thurlow⁵²⁾は腹話術の人形の人間的特徴を取り除いた場合の影響を調べた。目鼻を取り除いた場合にも音像をひきつける度合いが減少するという。また、顔はそのまま音声や口の動きに同期した純音に変えても視覚の影響は減じるという。これらの要因は視覚と聴覚の一体感に関係あり、腹話術効果に経験的要因があることを示している。提示した視聴覚刺激が同じであっても、被験者が映像に注目している度合いが高いほど腹話術効果が現れ易いことも過去の研究によって示されている⁵³⁾。

SLT 実験の結果から、聴覚刺激と視覚刺激が同時に提示されとき、音像の方向が視覚刺激の影響を受けて変移し移動することが分かった。それは、被験者 5 名の平均で、先行音効果の時間差の境界 SLT が約 0.2 ms 程度 短くなるように知覚されることが分かった。しかし、各被験者においては SLT の値に若干の差異が見られた。また、相対的に映像と音像の空間的な距離を増すにつれ、その効果は減少し、ある範囲(2 音源

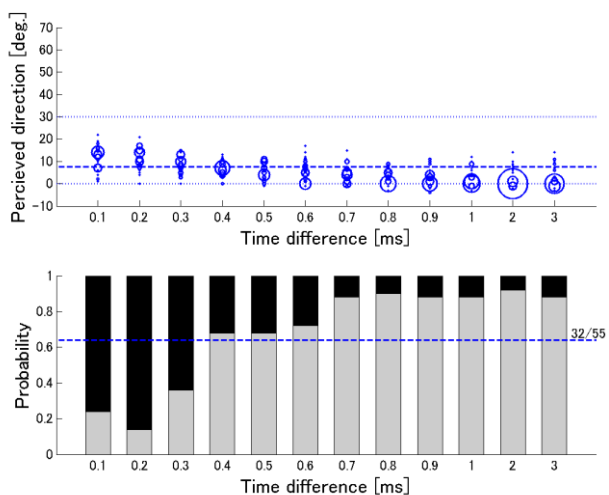
間の開き角度が $\pm 60^\circ$) を超えると急に消滅した。また、本研究の結果においては、被験者個人に対する視覚刺激の影響では時間差の境界 SLT を短くしたり、影響が無かったりして、不安定な音像定位を示す被験者もいた。このように被験者によって視覚刺激の影響に若干差異が見られたが、被験者全体の平均から見ると、腹話術効果と類似した傾向と捉えることができると考えられた。このようなことから、聴覚と視覚の感覚間の相互作用について見た場合、視覚は自己の意思で瞬時に目を閉じることができる。しかし、音を注意深く聴くときには目を閉じることがあり得るが、何かを注視するときには耳をふさぐことはまれである。このように視覚の影響は絶対ではなく、それが強く現れるためには、聴覚刺激と視覚刺激の時空間的な同一性が必要であると思われるが、本研究結果のように、刺激の物理的条件が全く同一であっても、被験者の注意度が影響することがあり、視覚刺激に注意度を向けた場合、腹話術効果が生じ易い。即ち、映像によって、音像が引き寄せられる現象で、音像の方向に注視点を動かすことによって、精度が高められるものと考えられ、感覚間相互作用に対し、心理学的な要因が潜められているものと考えられた。



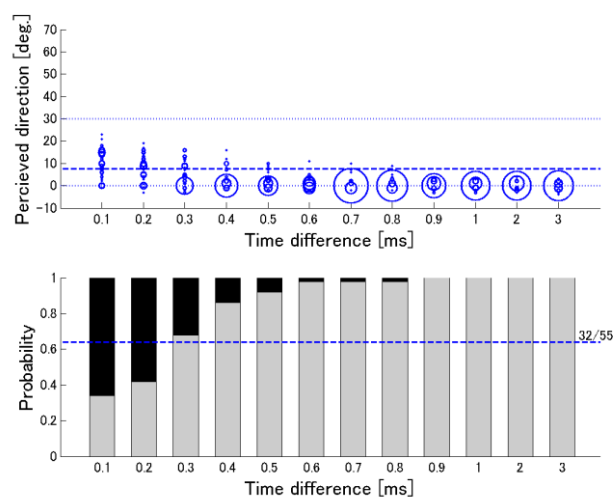
(A-1)スピーチ男声、映像なし



(A-2)スピーチ男声、映像あり

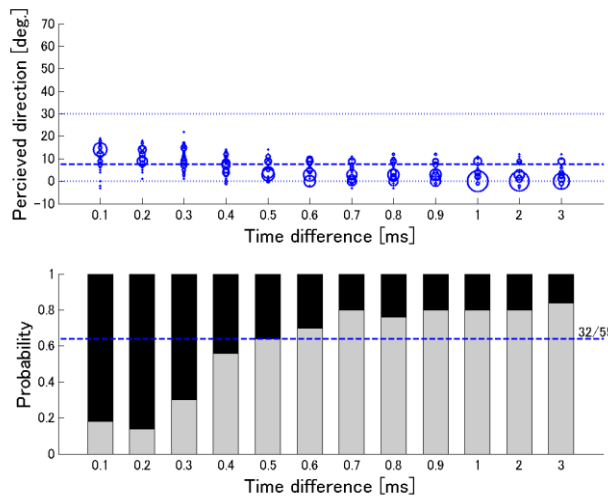


(B-1)スピーチ女声、映像なし

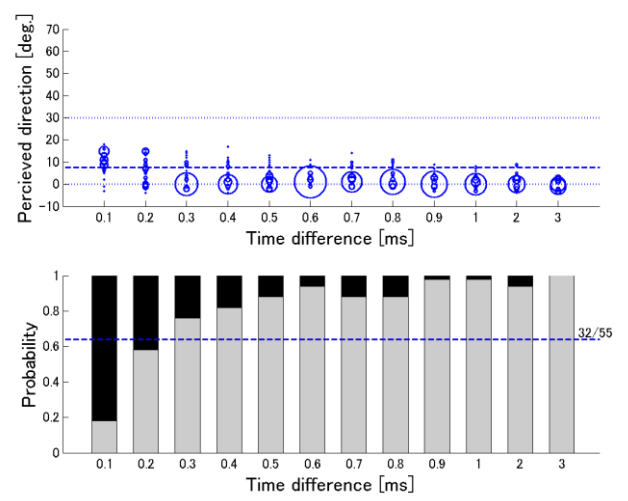


(B-2)スピーチ女声、映像あり

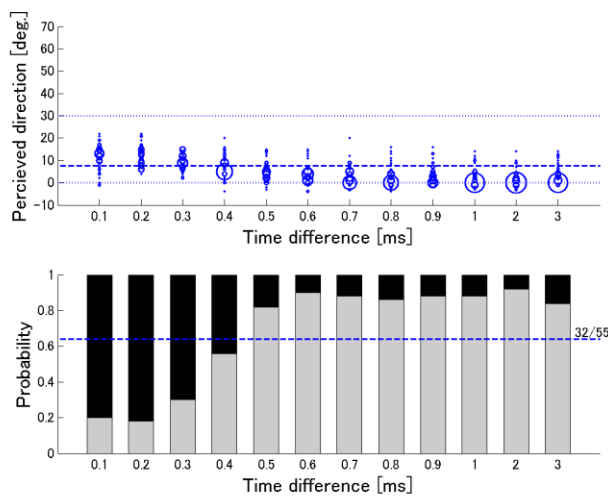
図 4.12 時間差による回答確率の変化, $\pm 30^\circ$, 映像の有/無, 5 名



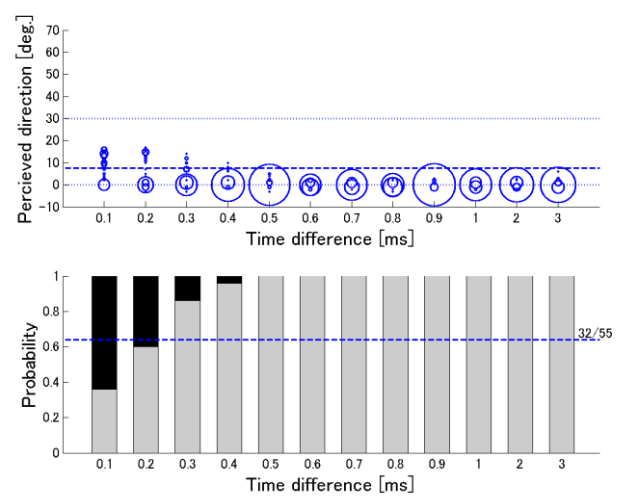
(C-1) ギター、映像なし



(C-2) ギター、映像あり

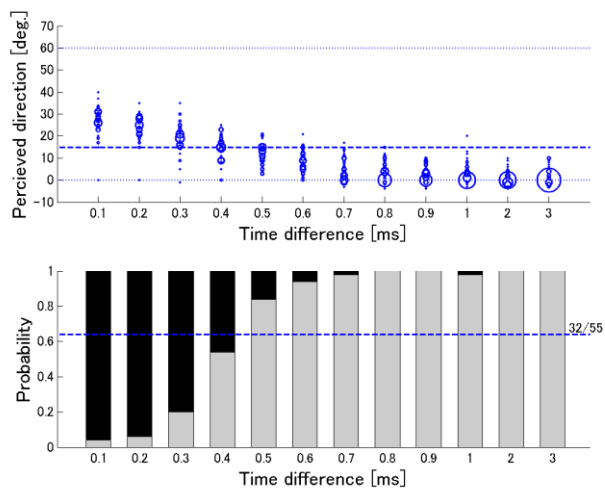


(D-1) 小太鼓、映像なし

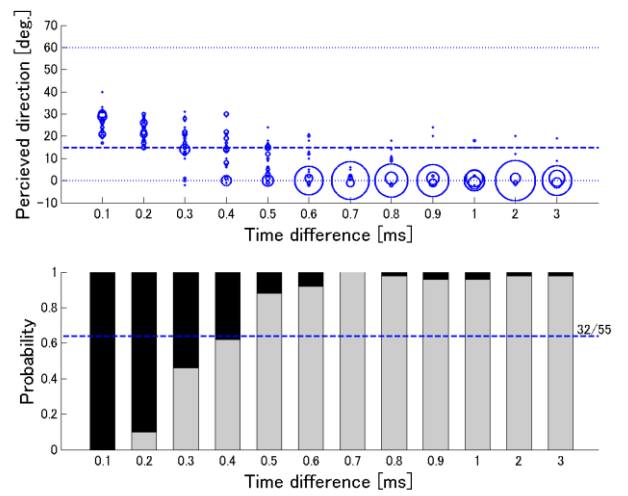


(D-2) 小太鼓、映像あり

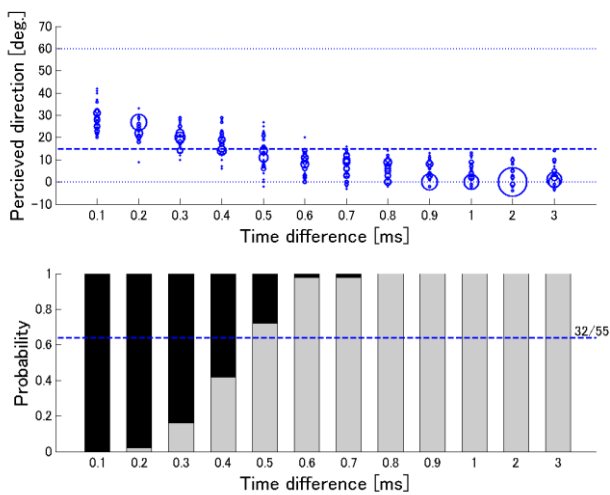
図 4.12 (続き) 時間差による回答確率の変化, $\pm 30^\circ$, 映像の有/無, 5 名



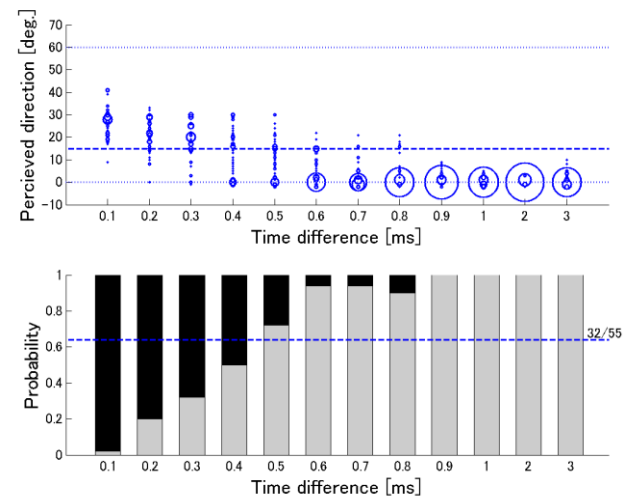
(A-1) スピーチ男声，映像なし



(A-2) スピーチ男声，映像あり

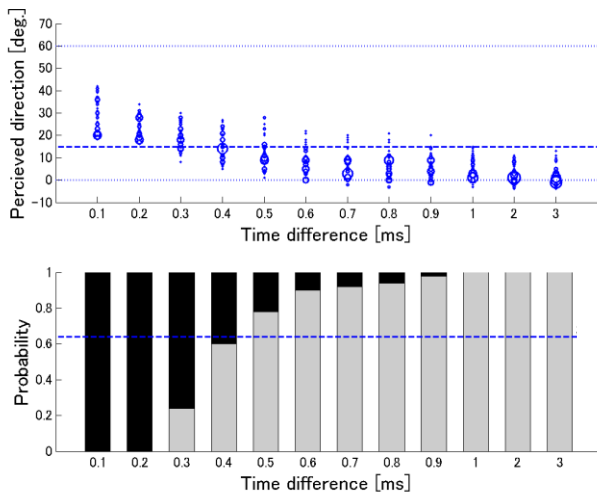


(B-1) スピーチ女声，映像なし

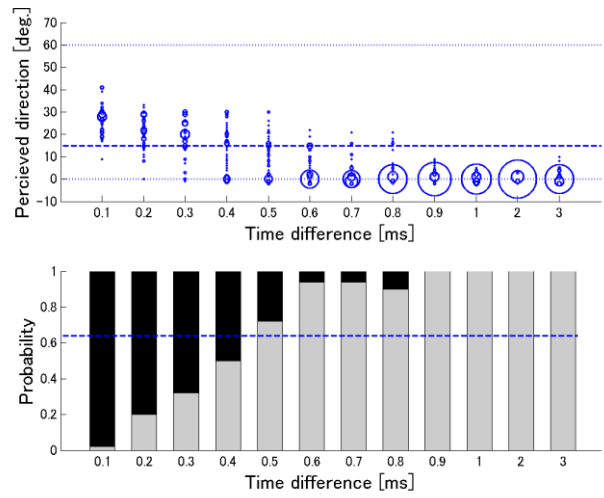


(B-2) スピーチ女声，映像あり

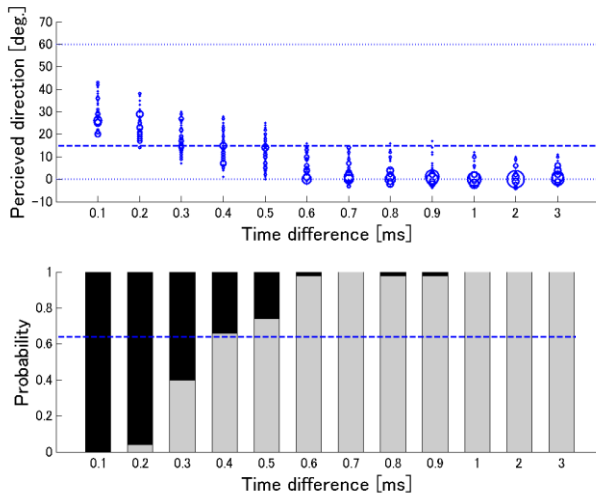
図 4.13 時間差による回答確率の変化， $\pm 60^\circ$ 映像の有/無，5 名



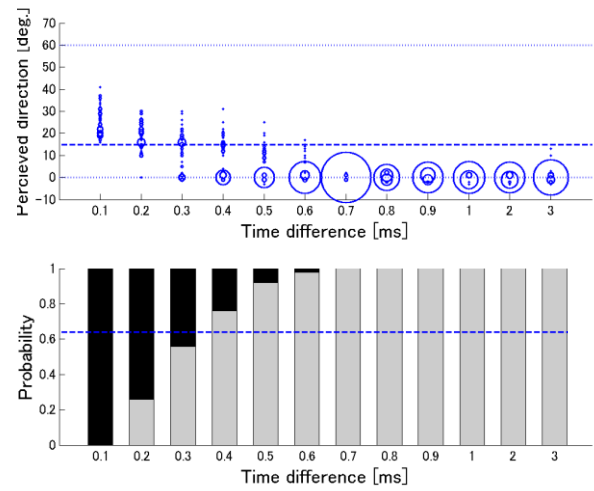
(C-1) 小太鼓, 映像なし



(C-2) 小太鼓, 映像あり

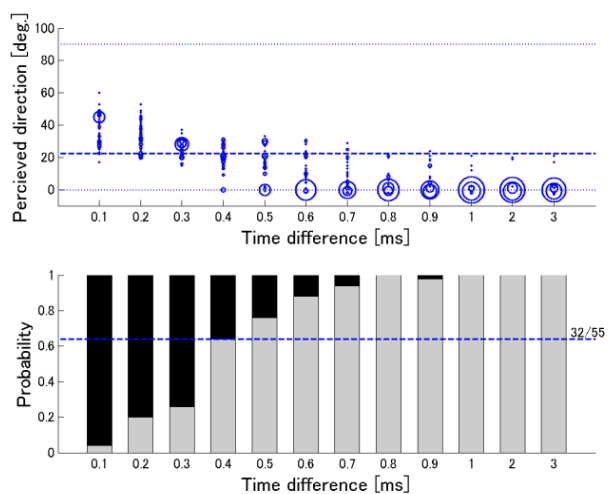


(D-1) ギター, 映像なし

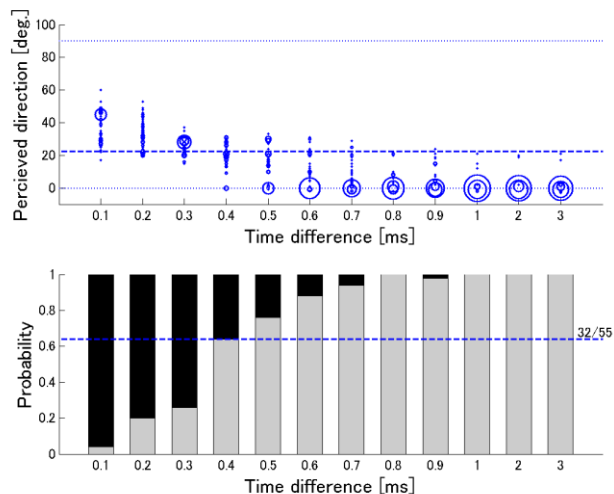


(D-2) ギター, 映像あり

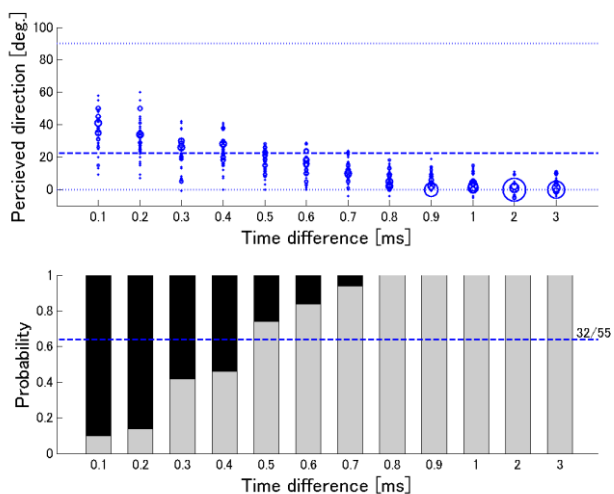
図 4.13 (続き) 時間差による回答確率の変化, $\pm 60^\circ$ 映像の有/無, 5名



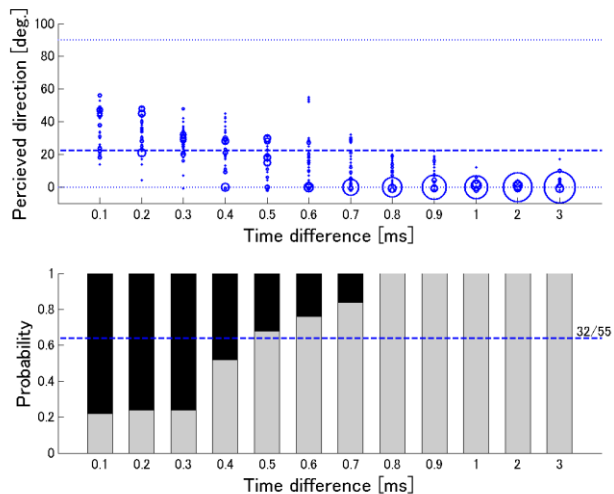
(A-1) スピーチ男声，映像なし



(A-2) スピーチ男声，映像あり

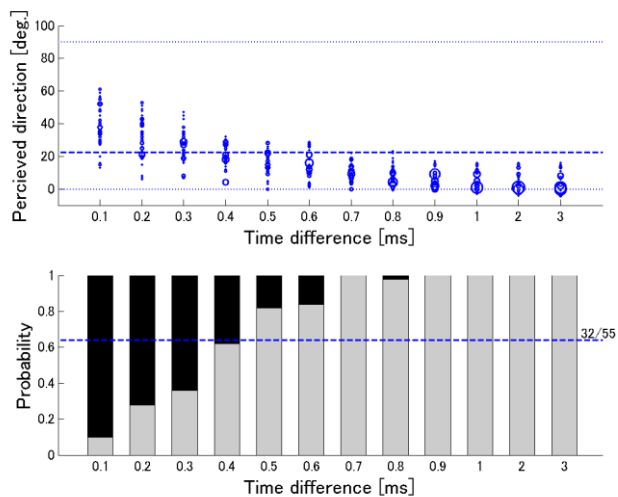


(B-1) スピーチ女声，映像なし

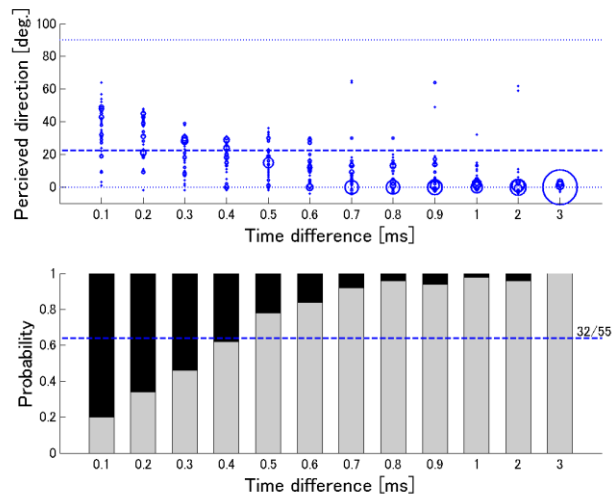


(B-2) スピーチ女声，映像あり

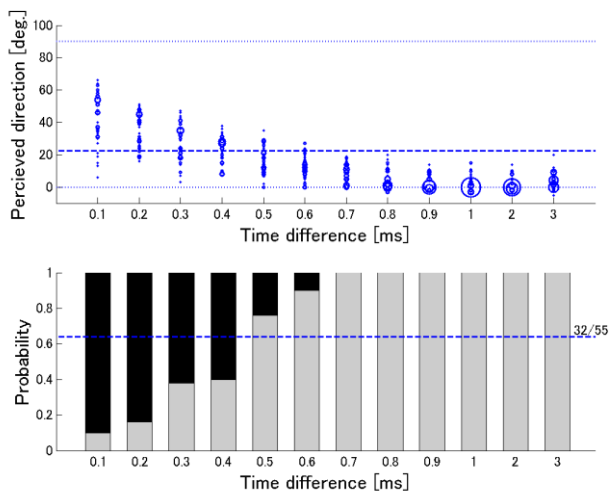
図 4.14 時間差による回答確率の変化， $\pm 90^\circ$ ，映像の有/無，5名



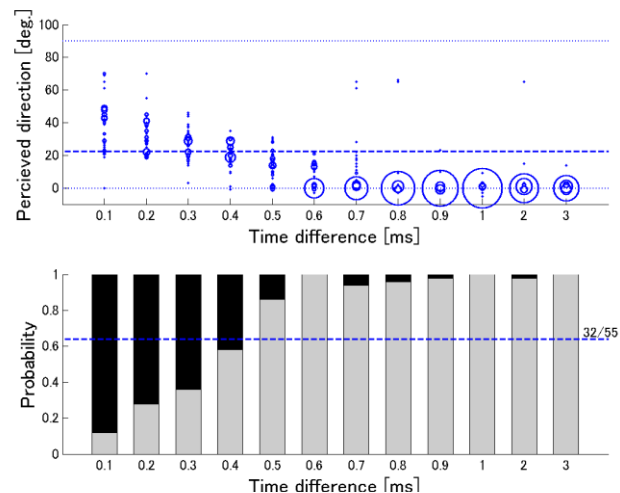
(C-1) ギター, 映像なし



(C-2) ギター, 映像あり



(D-1) 小太鼓, 映像なし



(D-2) 小太鼓, 映像あり

図 4.14 (続き) 時間差による回答確率の変化, $\pm 90^\circ$, 映像の有/無, 5 名

表4.2 刺激音の種類による映像 (ON)/(OFF) のSLT値

スピーカの配置	スピーカの開き角度 $\pm 30^\circ$			スピーカの開き角度 $\pm 60^\circ$			スピーカの開き角度 $\pm 90^\circ$		
映像の「OFF/ON」	OFF	ON	OFF-ON	OFF	ON	OFF-ON	OFF	ON	OFF-ON
スピーチ(女声)	0.5	0.3	0.2	0.5	0.4	0.1	0.4	0.4	0
スピーチ(男声)	0.4	0.3	0.1	0.5	0.5	0	0.5	0.5	0
小太鼓	0.5	0.3	0.2	0.5	0.5	0	0.5	0.5	0
ギター	0.5	0.3	0.2	0.4	0.4	0	0.5	0.4	0.1

被験者：SY, ST, S, O, W 単位：ms, 映像あり：ON, 映像なし：OFF. Main-speaker
とSub-speaker の音圧レベル：頭部中心位置で60dB, スピーカと被験者の距離：3m, 床
からスピーカまでの高さ(水平面上)：1.4m, (OFF－ON)＝視覚刺激の影響

表4.3 視覚刺激の影響 (被験者:5名)

Sub-speaker の方向	スピーチ(男声)		スピーチ(女声)		小太鼓		ギター	
	①	②	①	②	①	②	①	②
$\pm 30^\circ$	80	20	80	20	100	0	100	0
$\pm 60^\circ$	60	40	90	0	60	0	60	0
$\pm 90^\circ$	60	20	40	20	60	40	40	30

①：視覚刺激を加えた結果, SLT が短くなった被験者の割合 (%)

②：視覚刺激を加えた結果, SLT が長くなった被験者の割合 (%)

網掛けの部分は, 60%以上の被験者で SLT の変化が生じた視聴覚刺激の条件

表4.4 各被験者による視覚刺激の影響（被験者：5名 O, S, ST, SY, W）

音源方向	スピーカの開き角度±30° SLTと視覚刺激の影響					スピーカの開き角度±60° SLTと視覚刺激の影響				
被験者	O	S	ST	SY	W	O	S	ST	SY	W
ギター	off:0.4 on:0.2	off:0.5 on:0.4	off:0.7 on:0.4	off:0.5 on:0.2	off:0.5 on:0.3	off:0.4 on:0.4	off:0.6 on:0.6	off:0.9 on:0.6	off:0.5 on:0.4	off:0.5 on:0.4
off-on	+0.2	+0.1	+0.3	+0.3	+0.2	0	0	+0.3	+0.1	+0.1
スピーチ 男性	off:0.5 on:0.3	off:0.3 on:0.4	off:0.5 on:0.3	off:0.5 on:0.4	off:1.0 on:0.3	off:0.4 on:0.3	off:0.4 on:0.5	off:0.7 on:0.5	off:0.5 on:0.7	off:0.5 on:0.4
off-on	+0.2	-0.1	+0.2	+0.1	+0.7	+0.1	-0.1	+0.2	-0.2	+0.1
小太鼓	off:0.6 on:0.3	off:0.5 on:0.4	off:0.5 on:0.3	off:0.5 on:0.2	off:0.4 on:0.3	off:0.6 on:0.4	off:0.6 on:0.5	off:0.6 on:0.6	off:0.5 on:0.4	off:0.4 on:0.4
off-on	+0.3	+0.1	+0.2	+0.3	+0.1	+0.2	+0.1	0	+0.1	0

映像無し:off, 映像有り:on, off-on=視覚刺激の影響, 単位:ms

- (+) : 視覚刺激の付加により, S L Tが短くなった被験者,
 (−) : 視覚刺激の付加により, S L Tが長くなった被験者
 (0) : 視覚刺激の付加により, 影響はない被験者

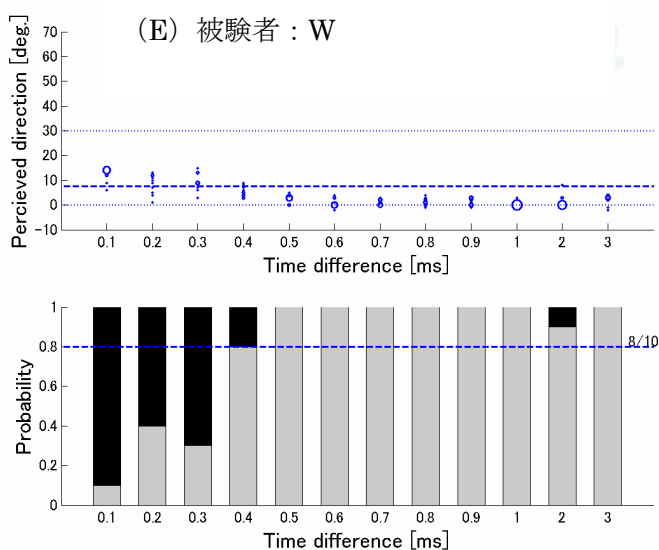
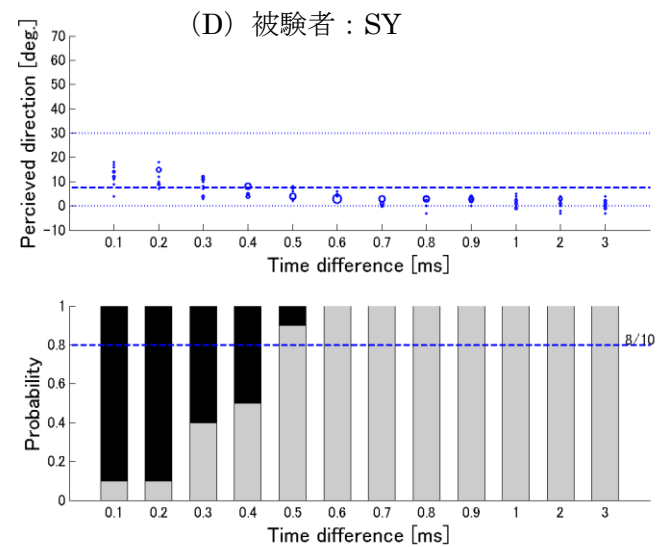
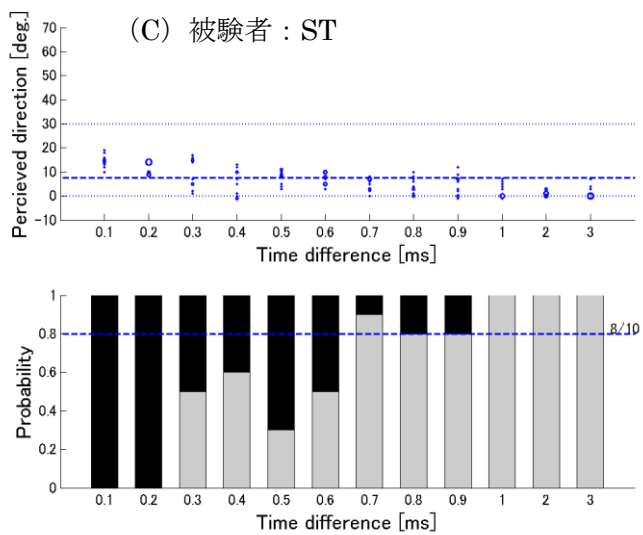
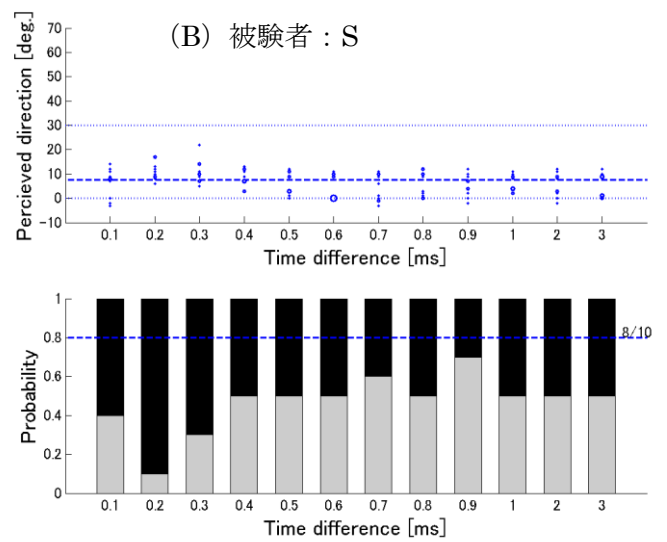
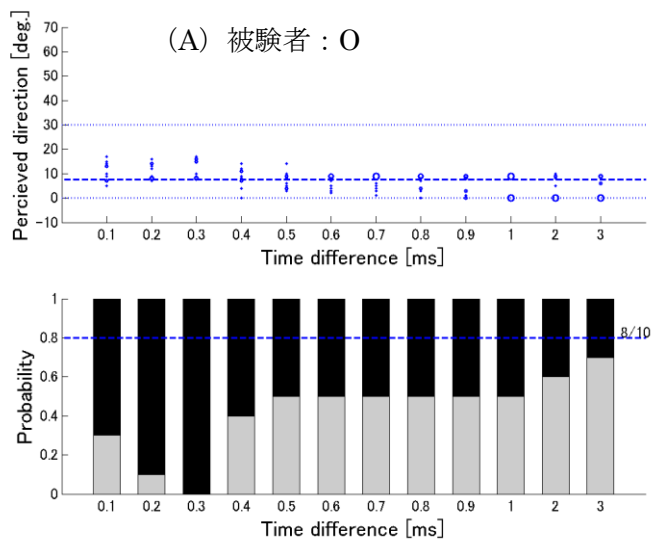


図 4.15 各被験者の先行音効果の発生した割合と SLT (ギター, 映像 OFF, $\pm 30^\circ$)

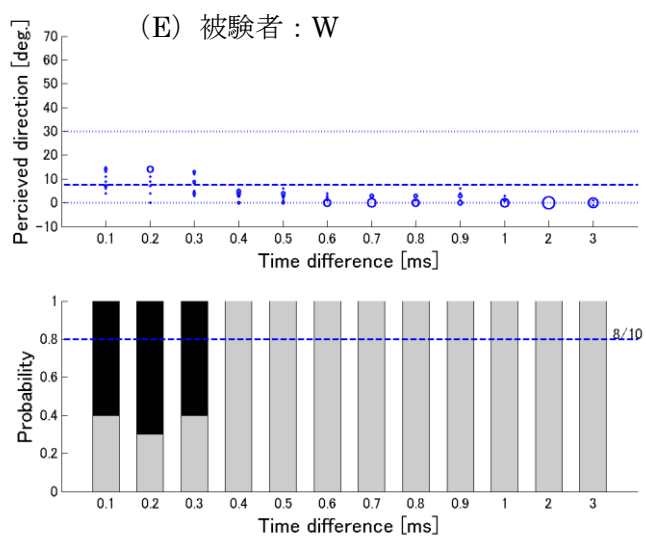
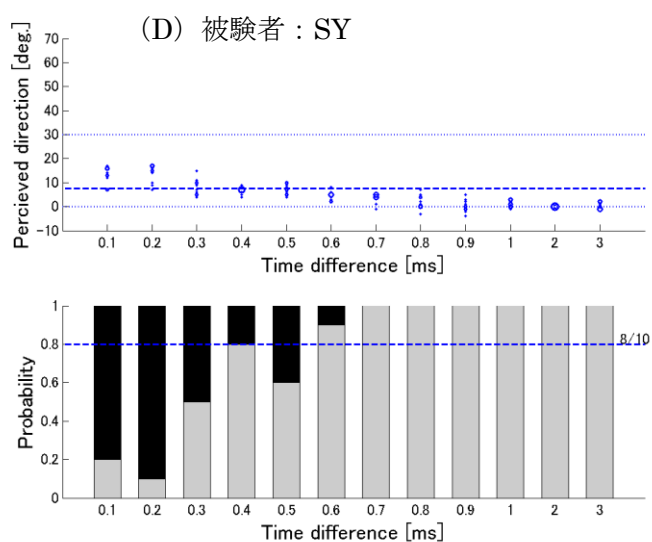
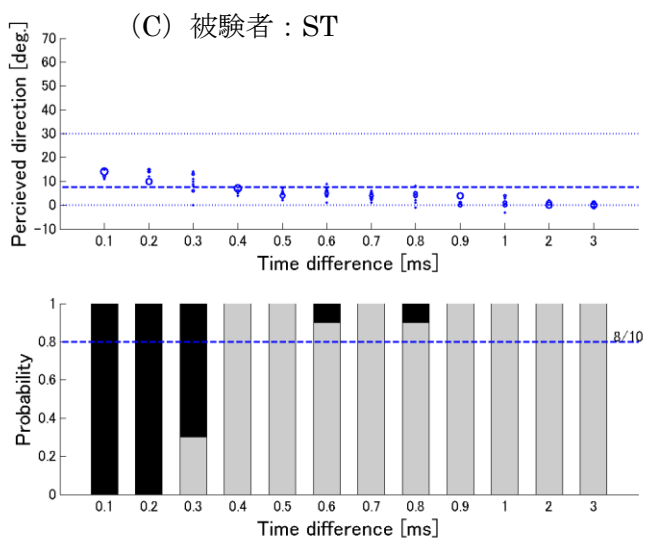
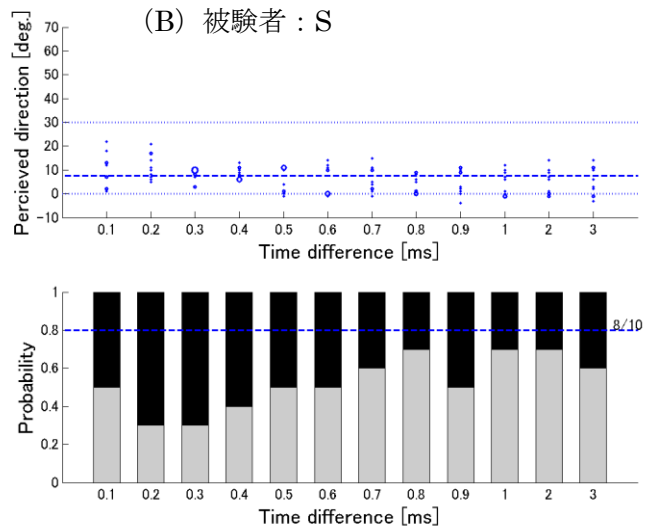
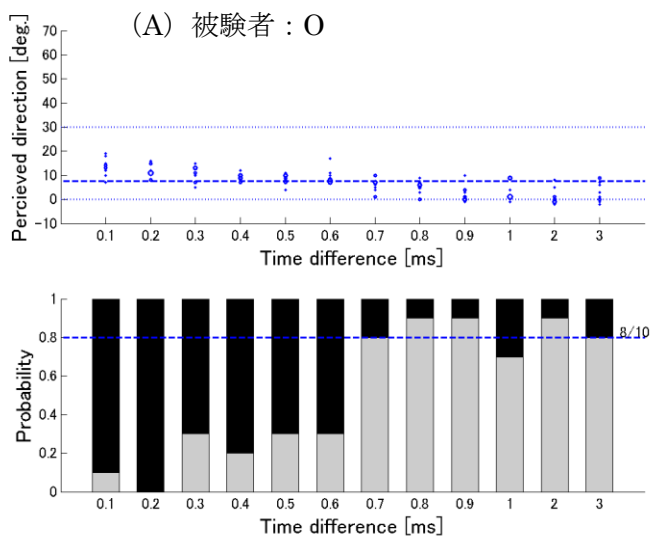


図 4.16 各被験者の先行音効

果の発生した割合と SLT (スピーチ女, 映像 OFF, $\pm 30^\circ$)

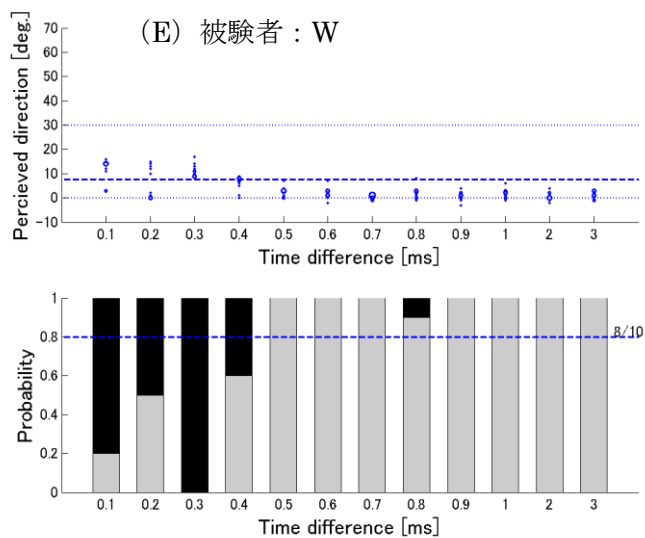
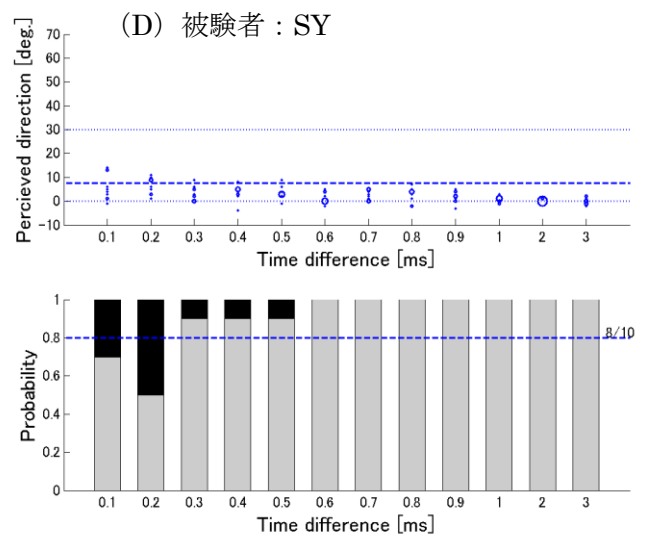
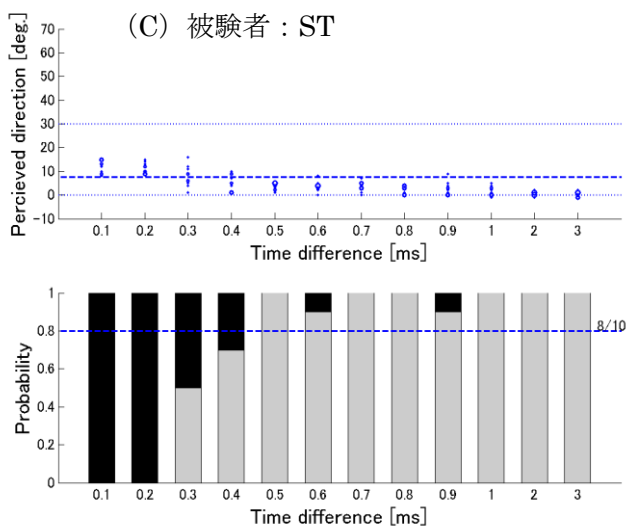
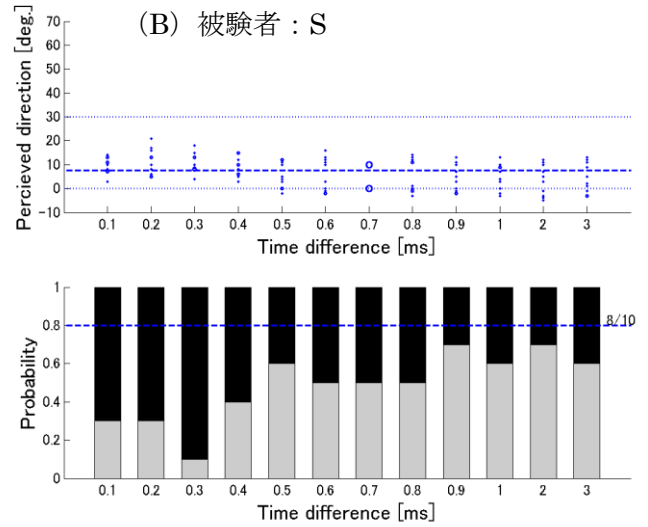
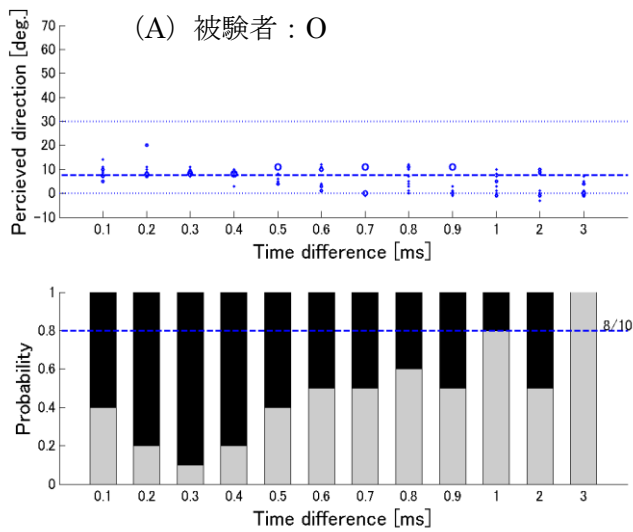


図 4.17 各被験者の先行音効果
の発生した割合と SLT（スピーチ
男，映像 OFF， $\pm 30^\circ$ ）

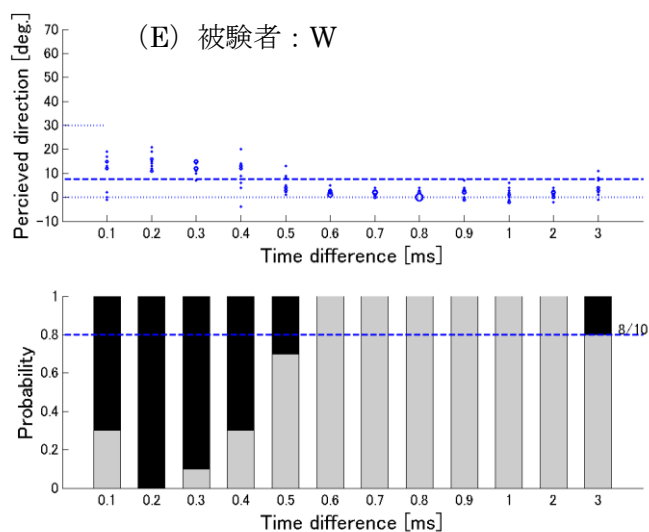
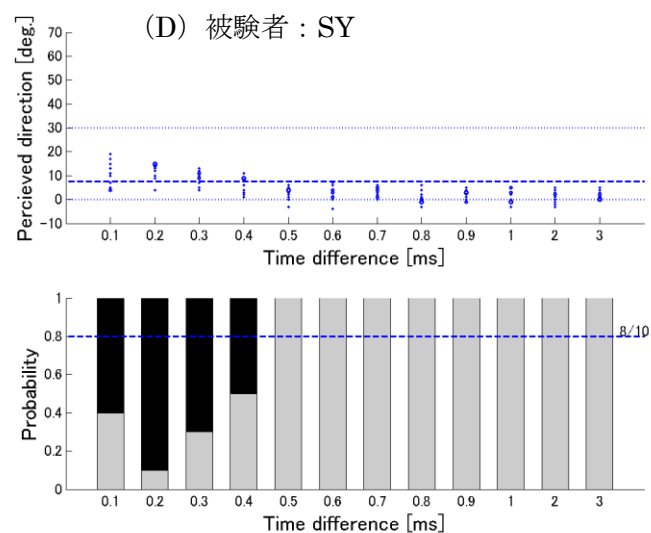
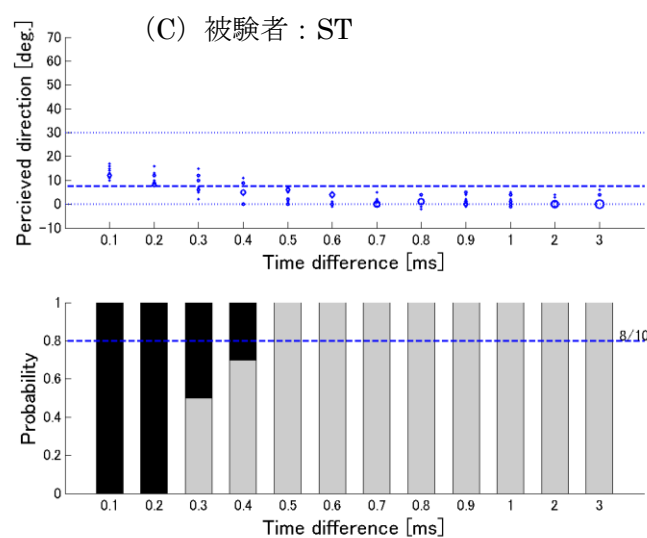
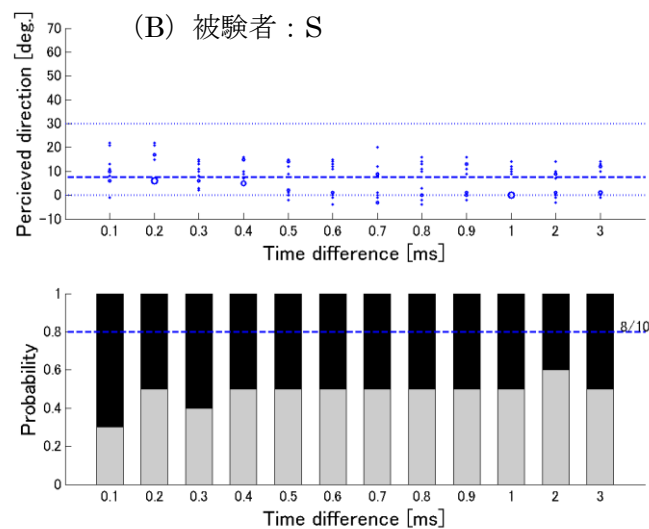
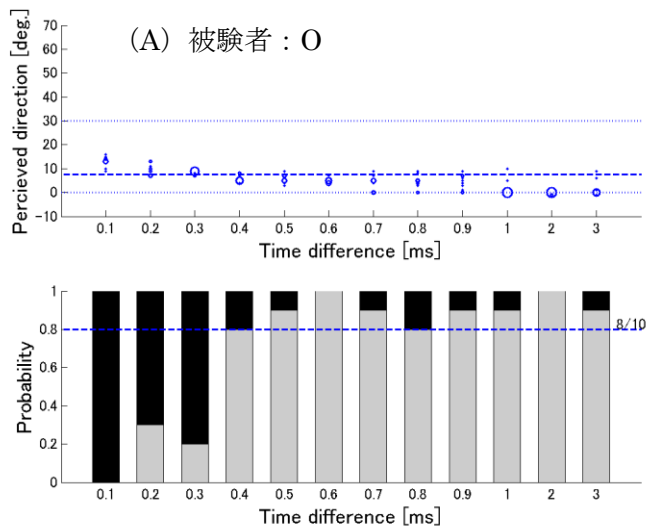


図 4.18 各被験者の先行音効果
の発生した割合と SLT（小太鼓，
映像 OFF， $\pm 30^\circ$ ）

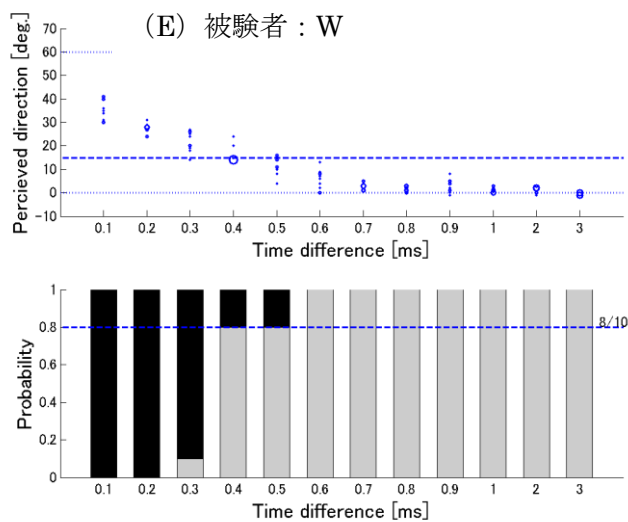
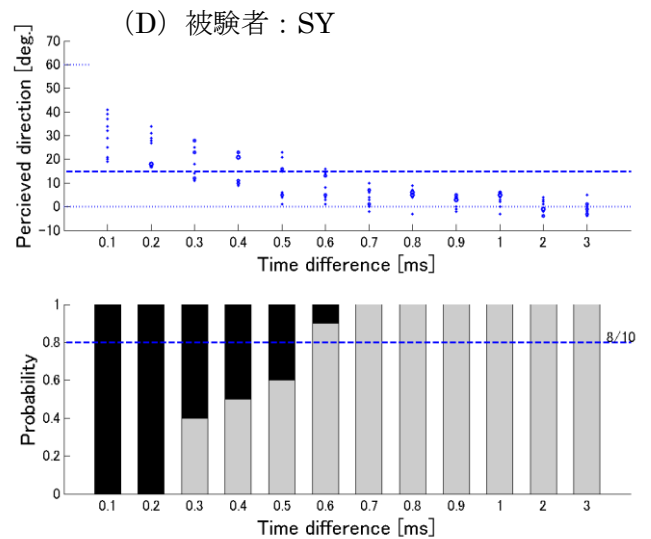
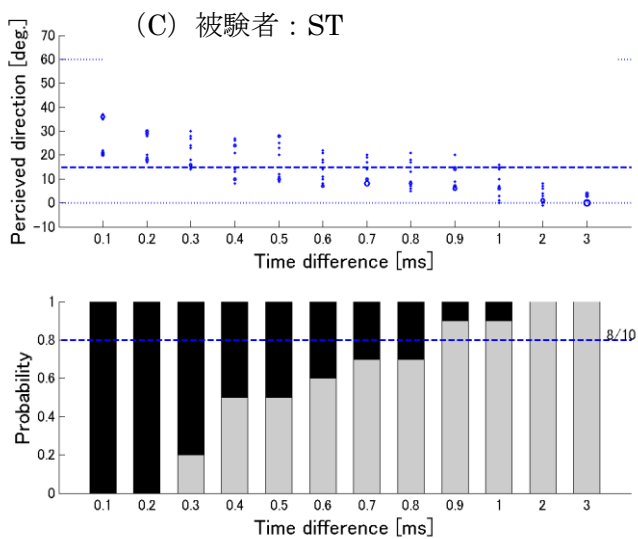
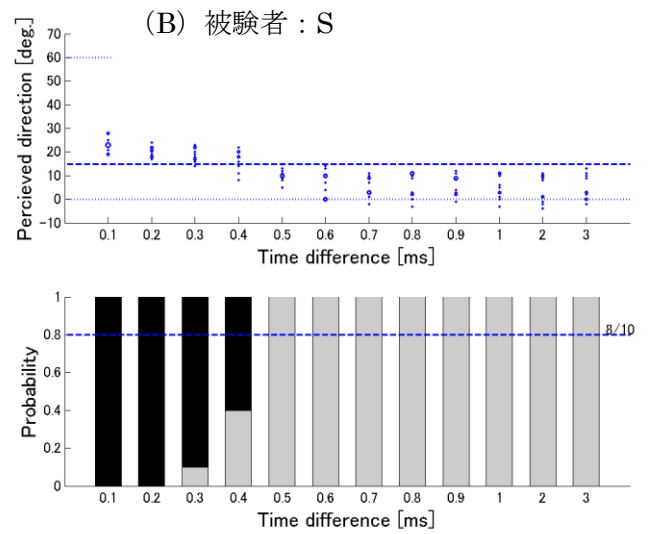
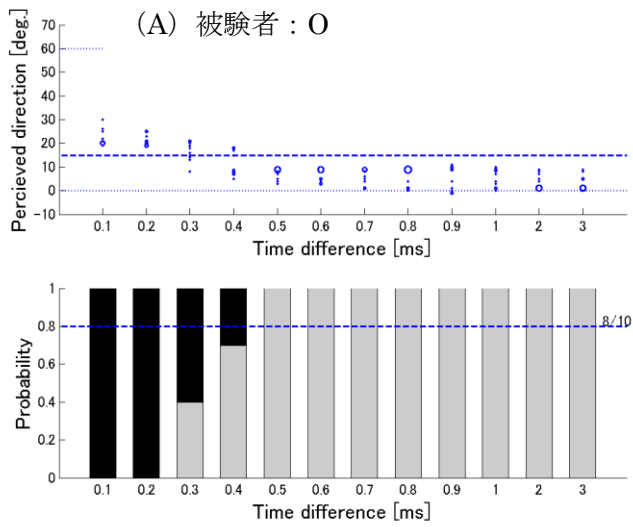


図 4.19 各被験者の先行音効果の発生した割合と SLT (ギター, 映像 OFF, $\pm 60^\circ$)

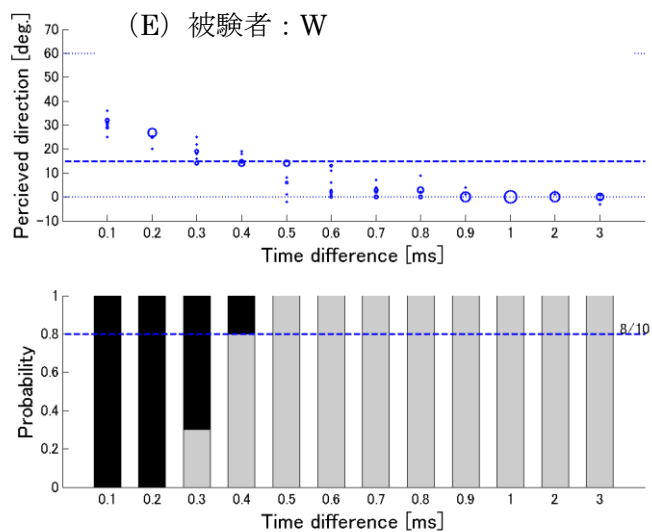
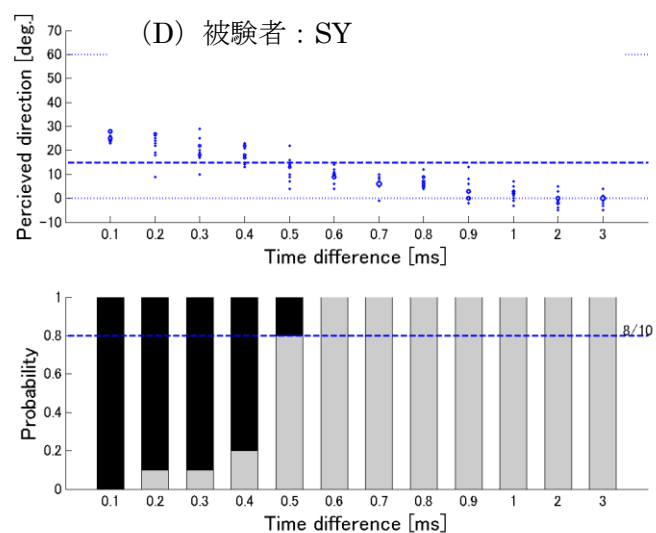
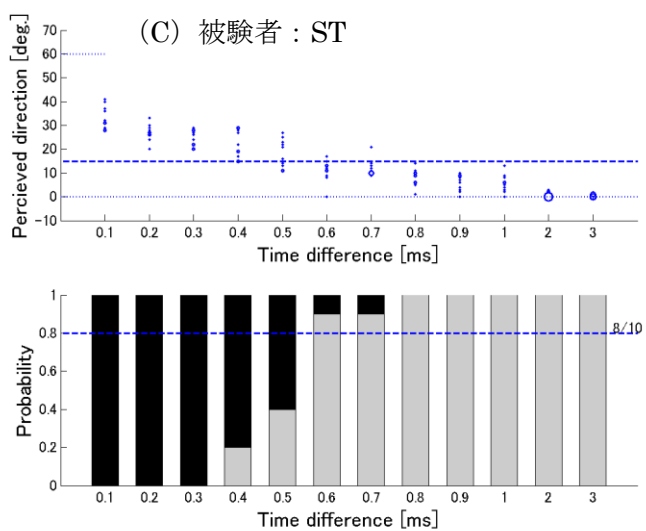
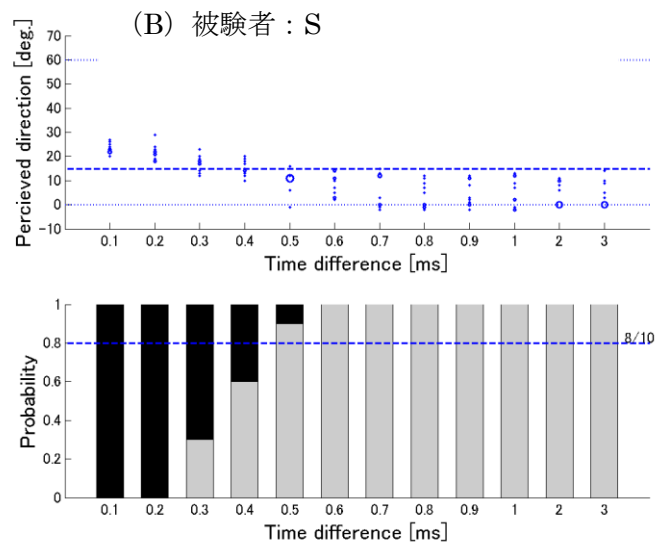
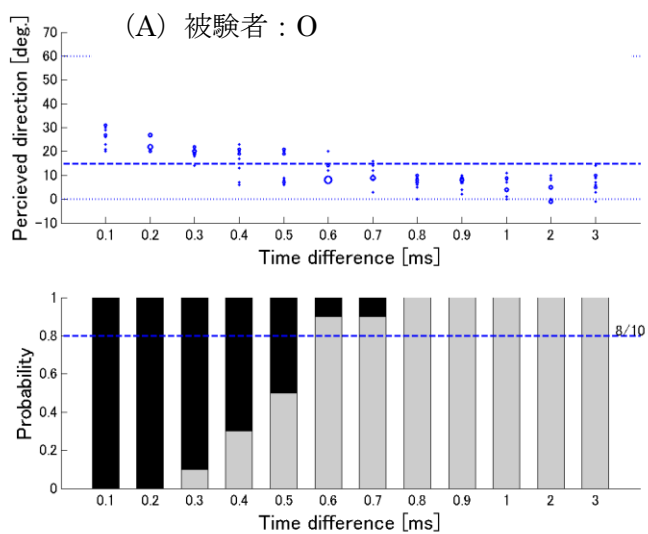


図 4.20 各被験者の先行音効果の発生した割合と SLT (スピーチ女, 映像 OFF, $\pm 60^\circ$)

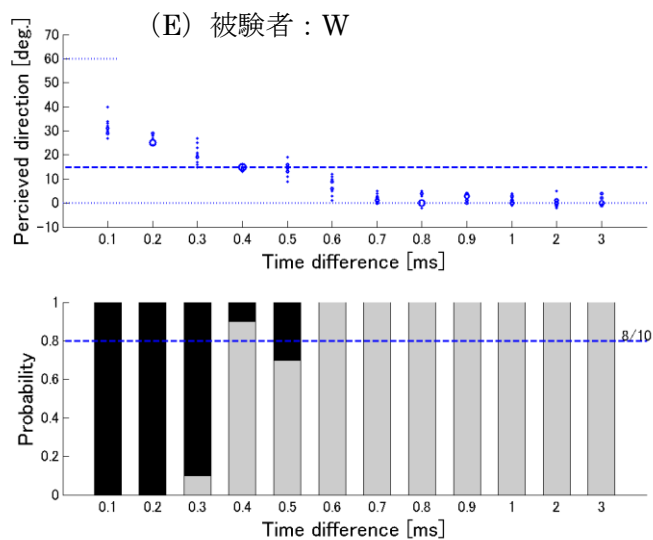
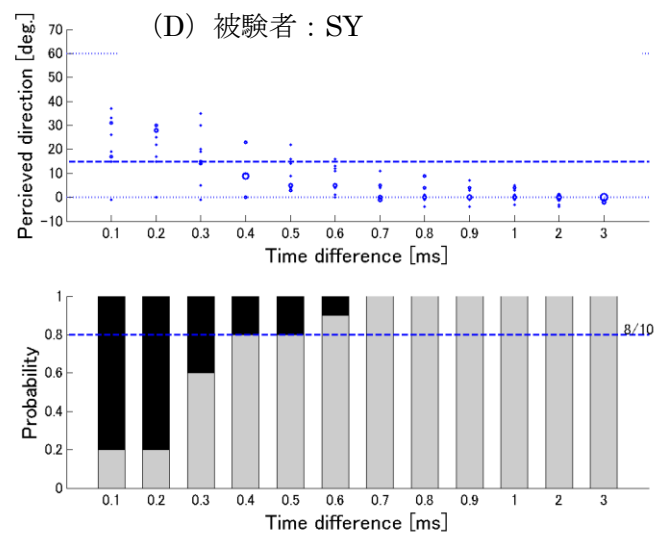
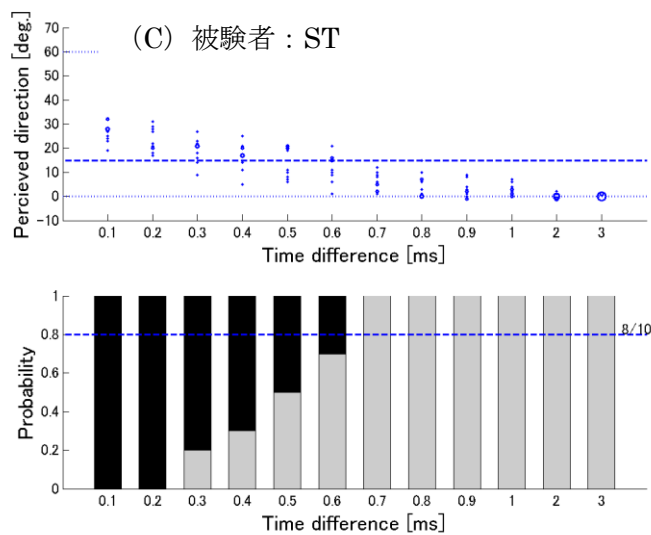
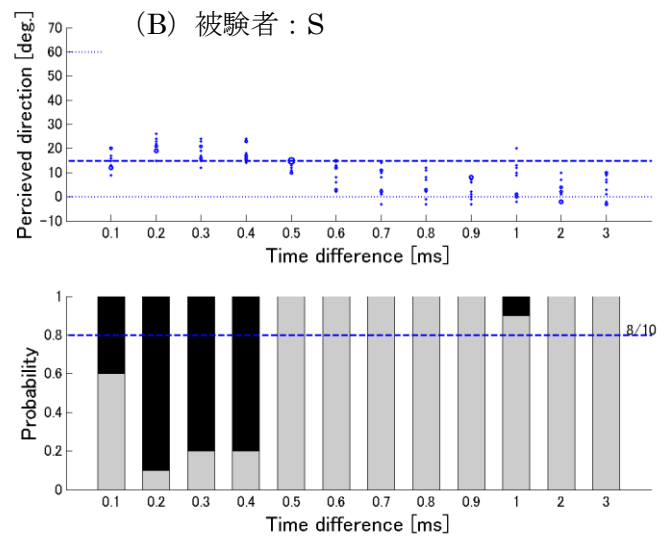
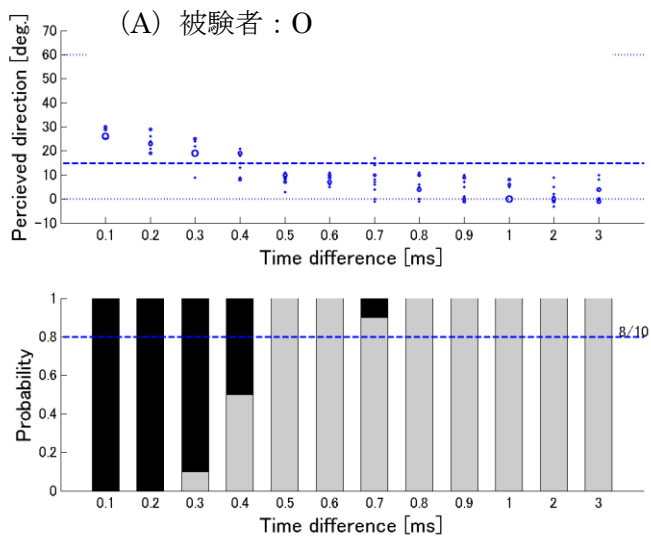


図4.21 各被験者の先行音効果の発生した割合と SLT (スピーチ男, 映像 OFF, $\pm 60^\circ$)

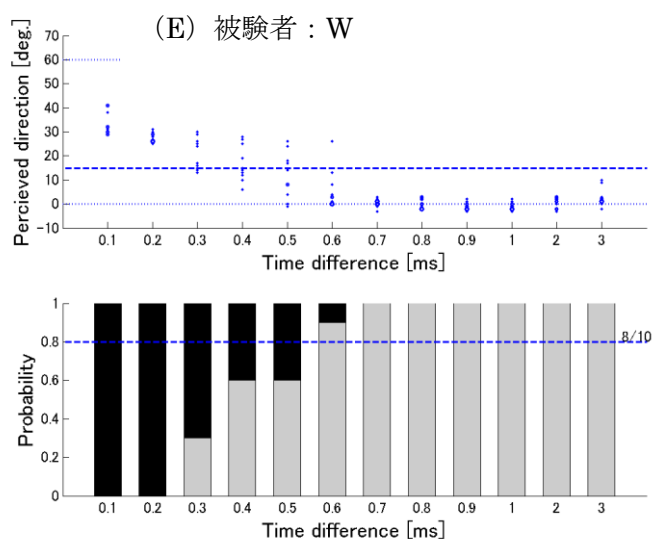
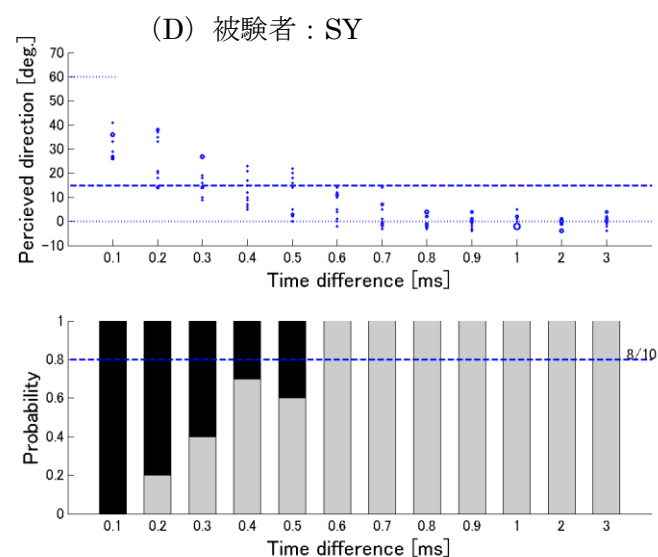
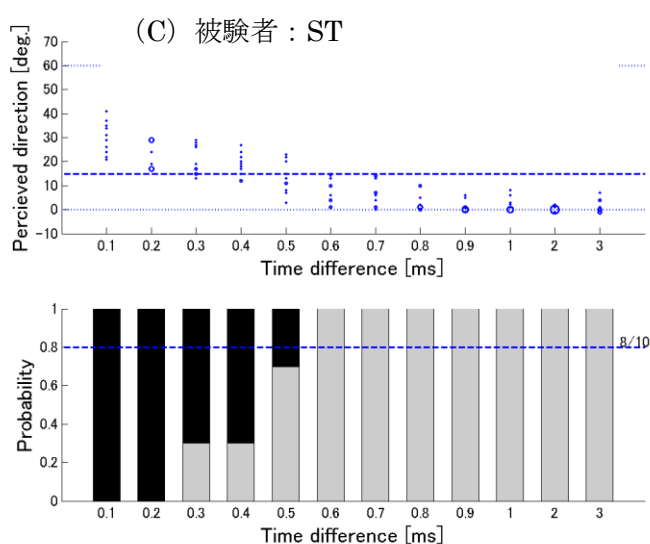
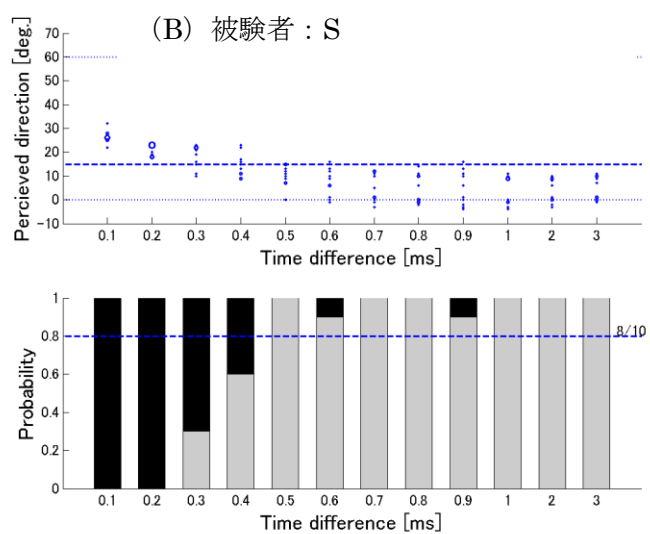
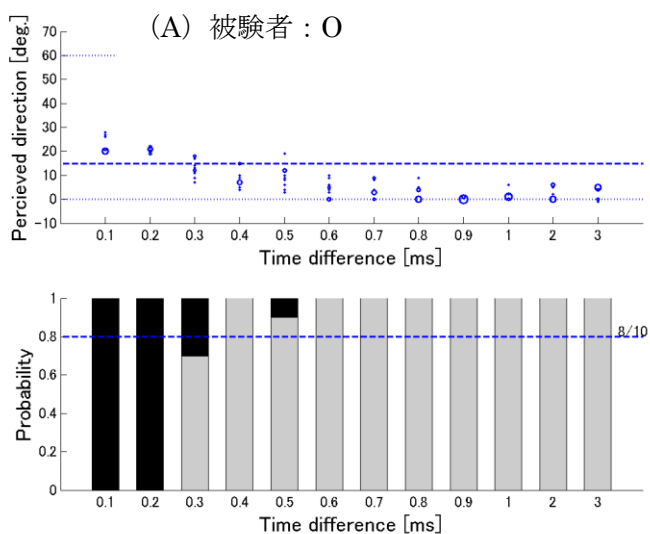


図 4.22 各被験者の先行音効果の発生した割合と SLT (小太鼓, 映像 OFF, $\pm 60^\circ$)

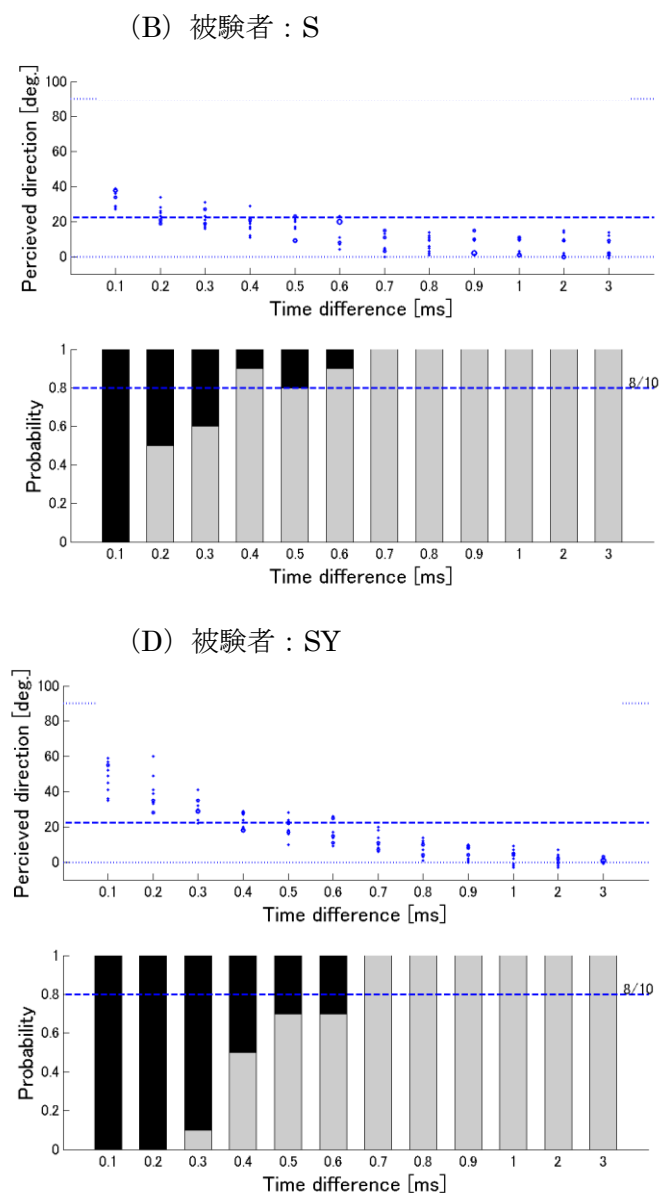
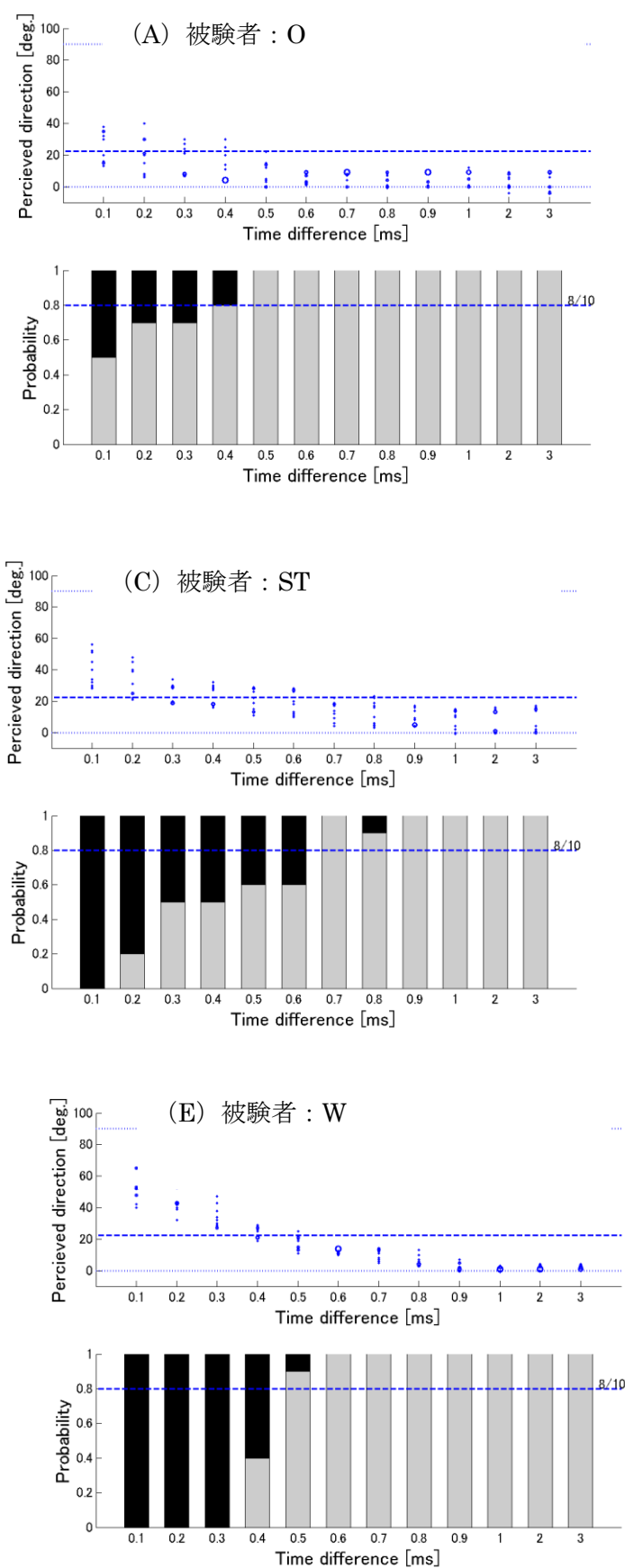


図 4.23 各被験者の先行音効果の発生した割合と SLT (ギター, 映像 OFF, $\pm 90^\circ$)

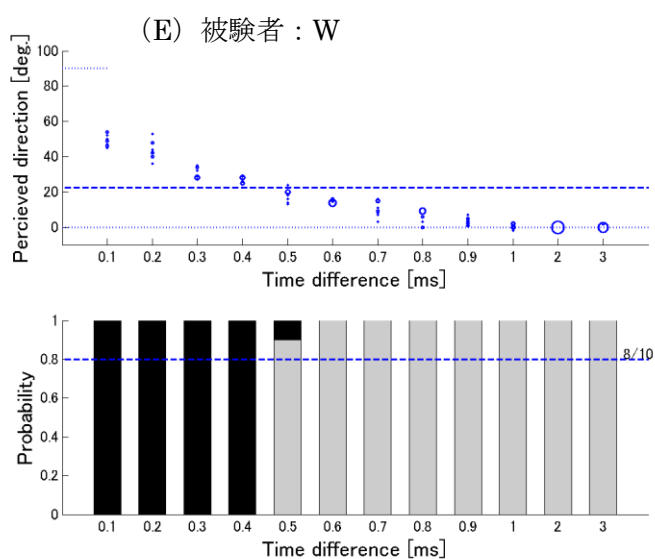
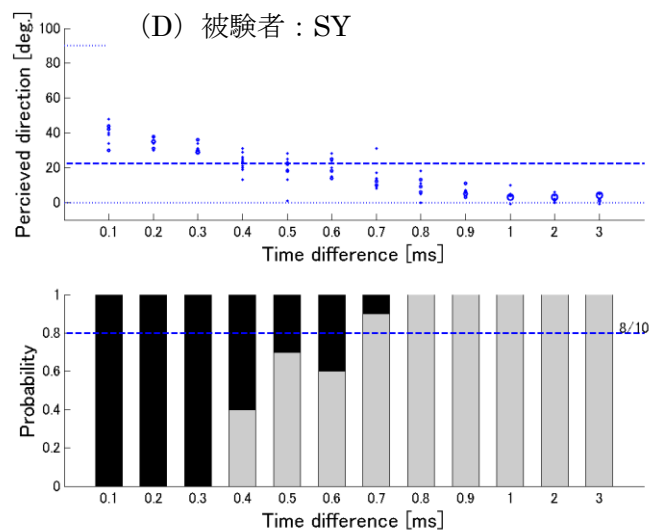
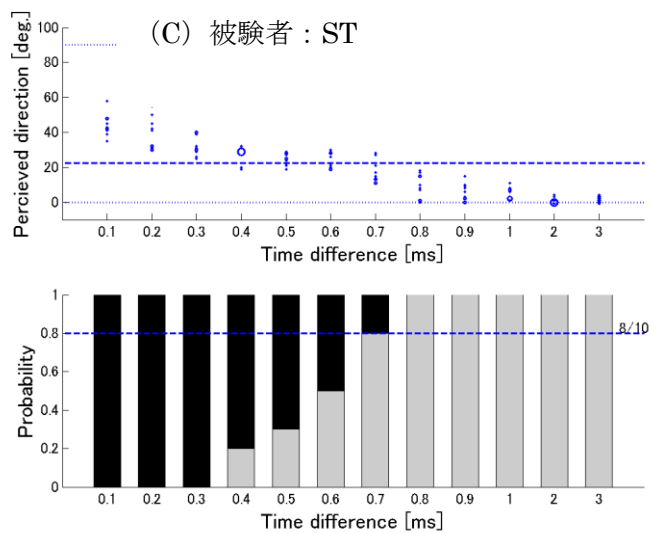
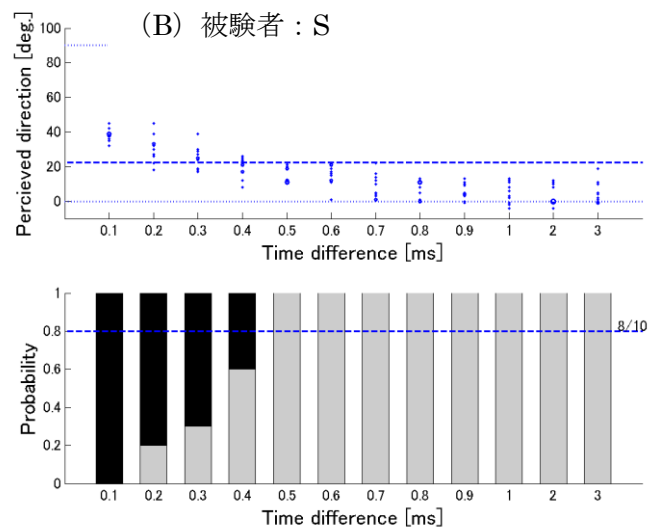
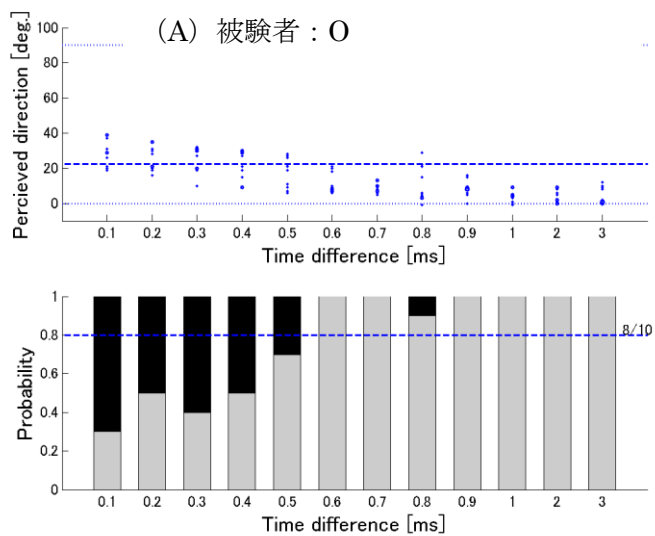


図 4.24 各被験者の先行音効果の発生した割合と SLT (スピーチ女, 映像 OFF, $\pm 90^\circ$)

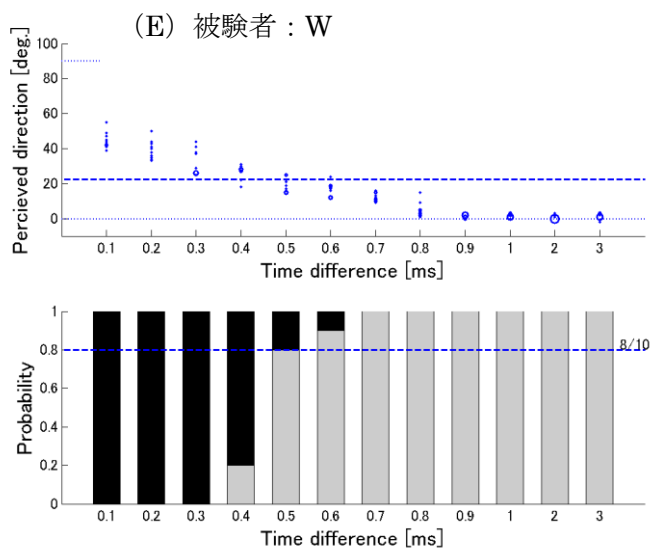
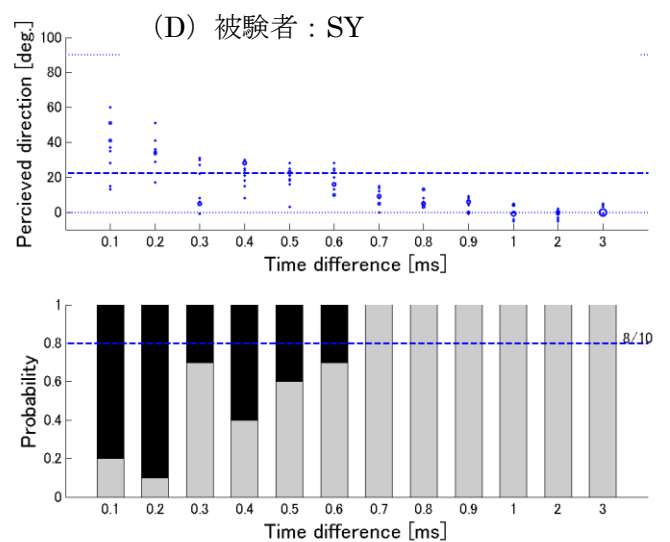
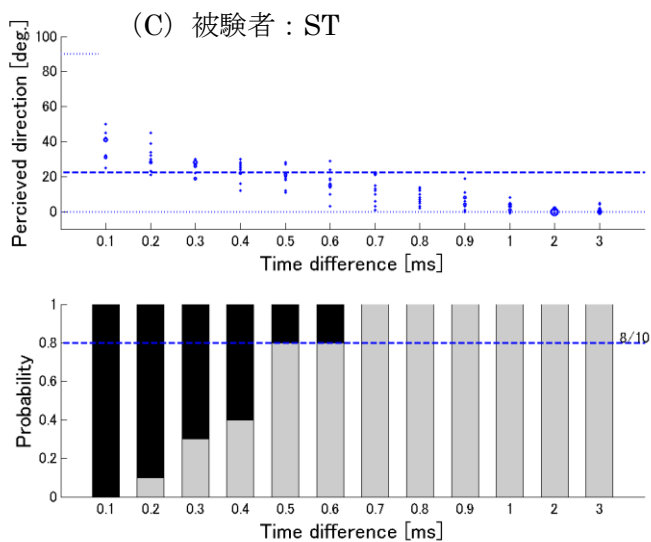
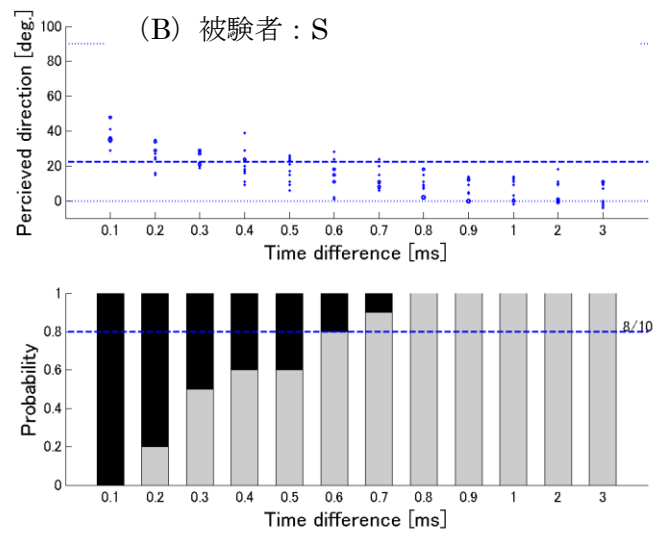
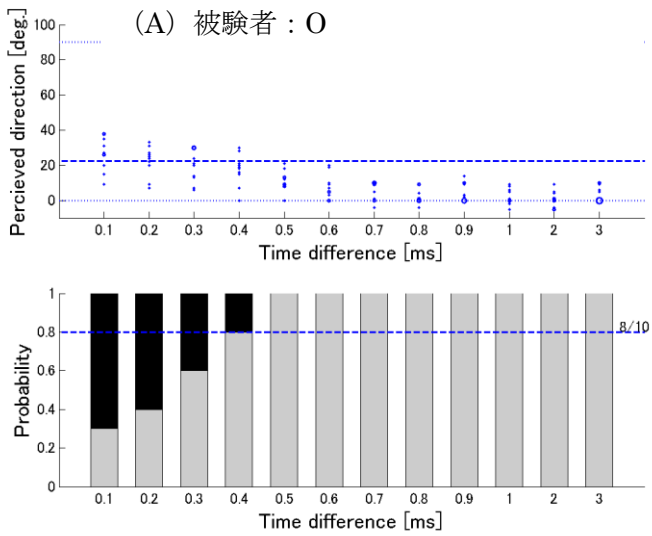


図 4.25 各被験者の先行音効果
の発生した割合と SLT（スピー
チ男，映像 OFF， $\pm 90^\circ$ ）

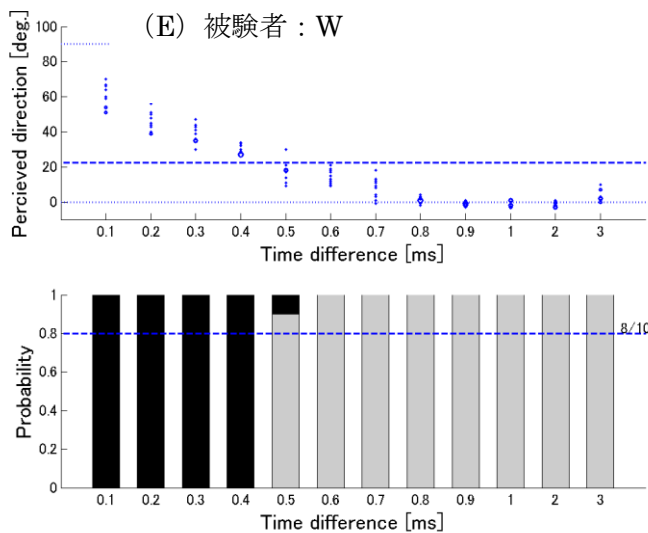
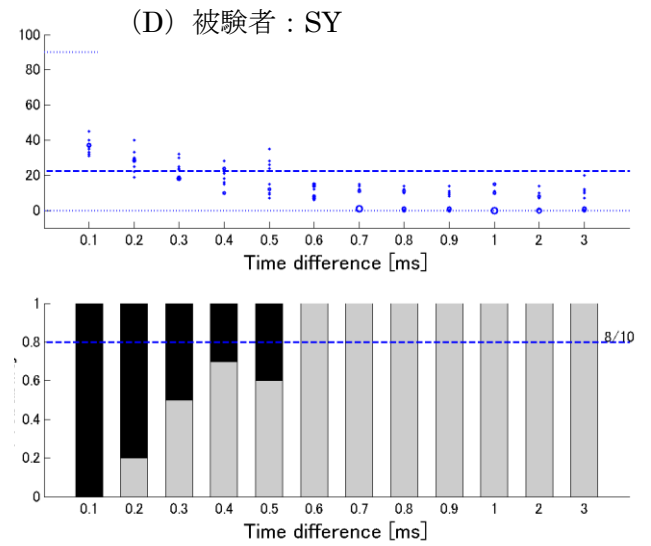
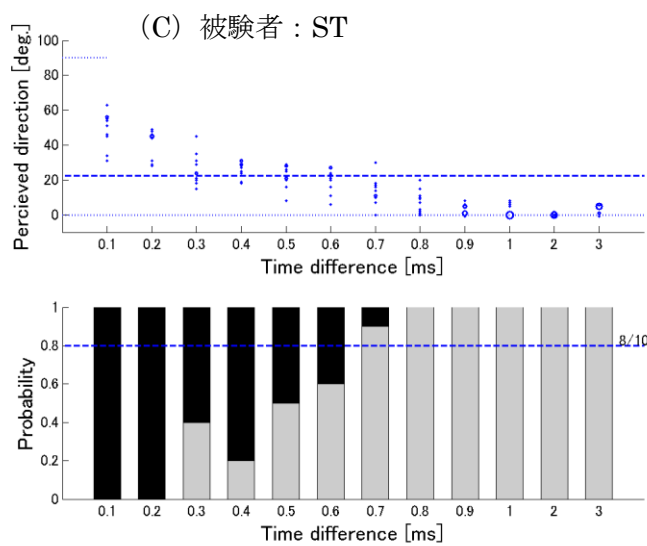
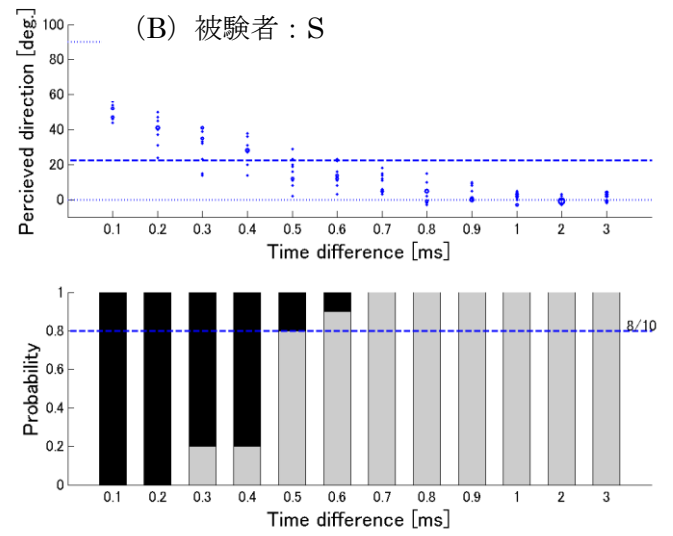
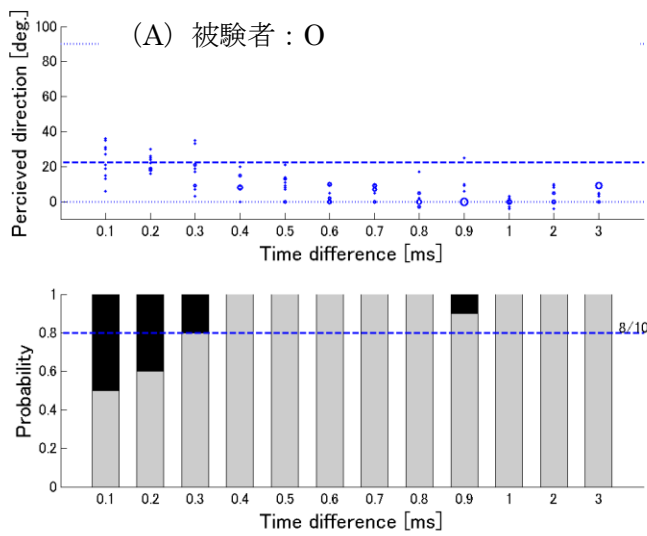


図 4.26 各被験者の先行音効果の発生した割合と SLT (小太鼓, 映像 OFF, $\pm 90^\circ$)

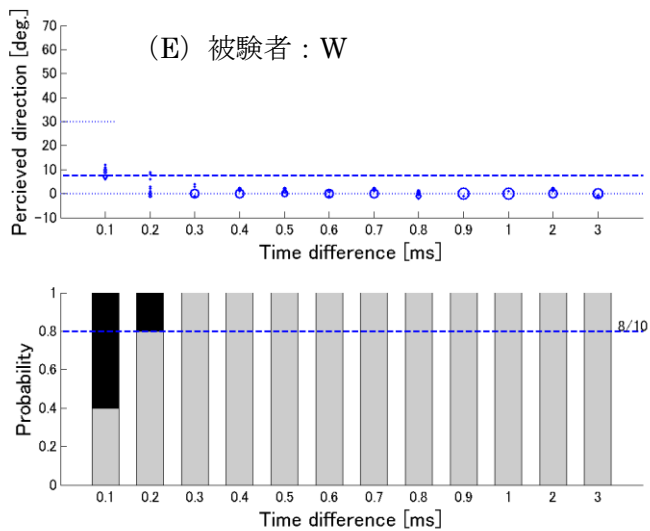
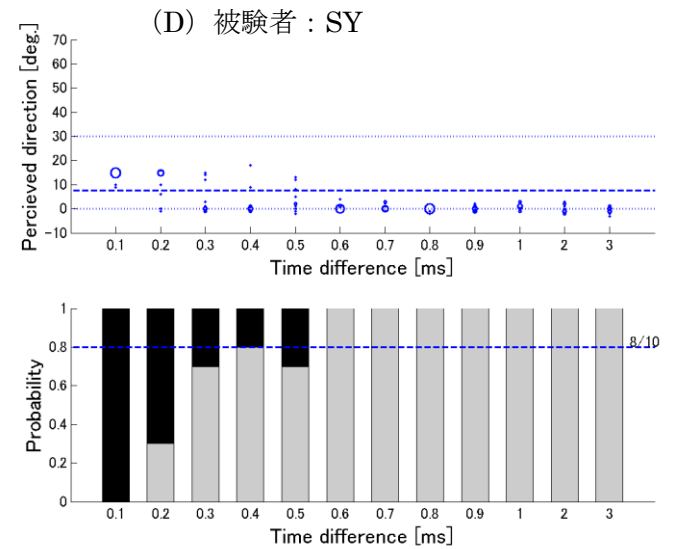
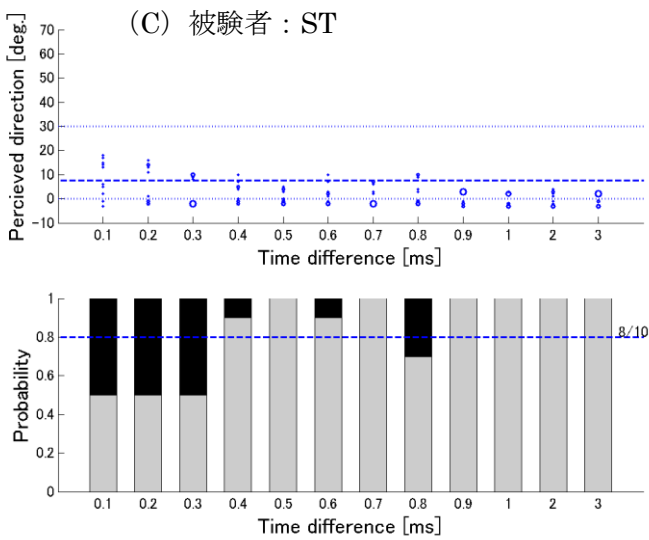
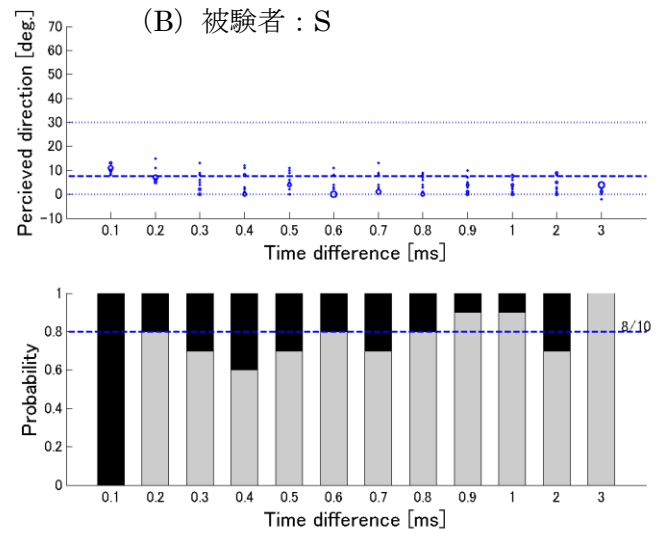
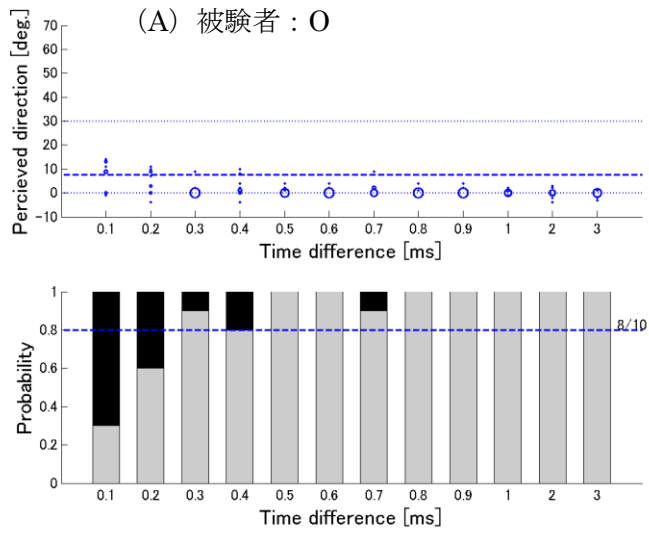


図 4.27 各被験者の先行音効

果の発生した割合と SLT (ギ

ター, 映像 ON, $\pm 30^\circ$)

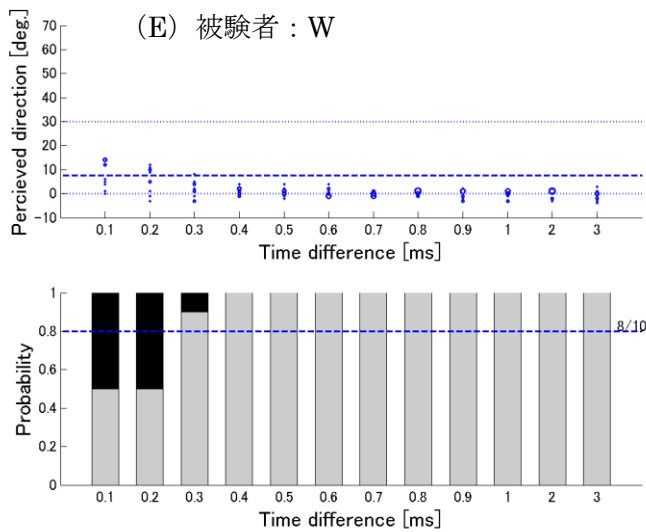
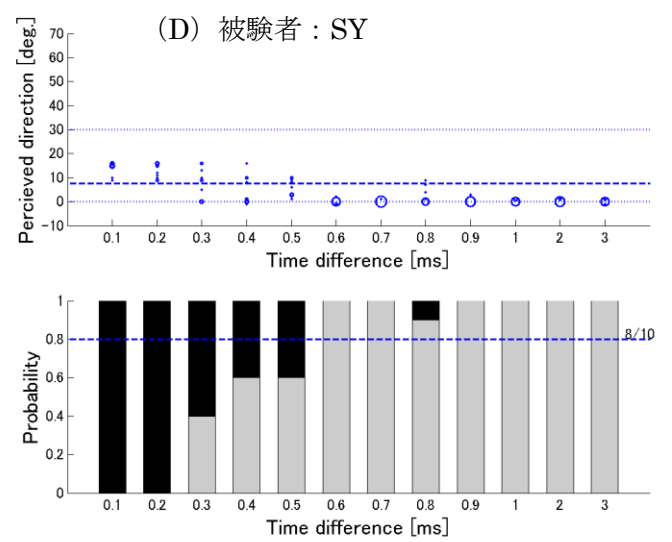
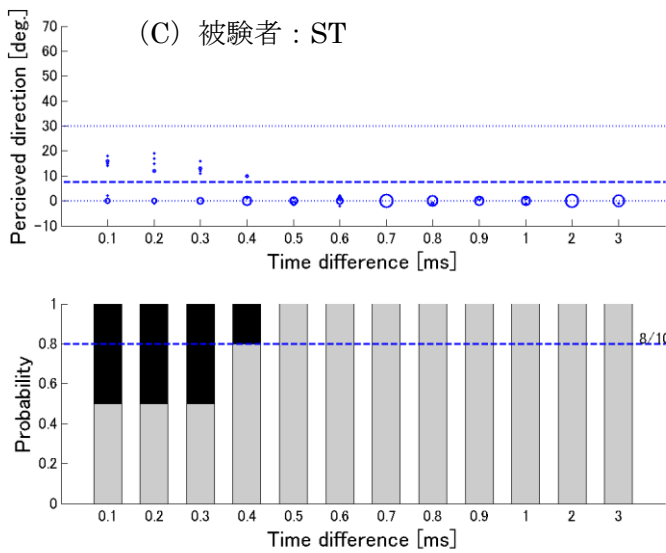
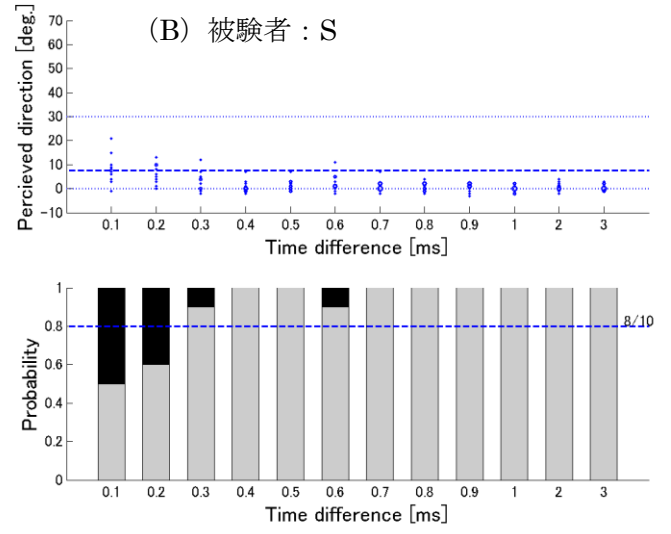
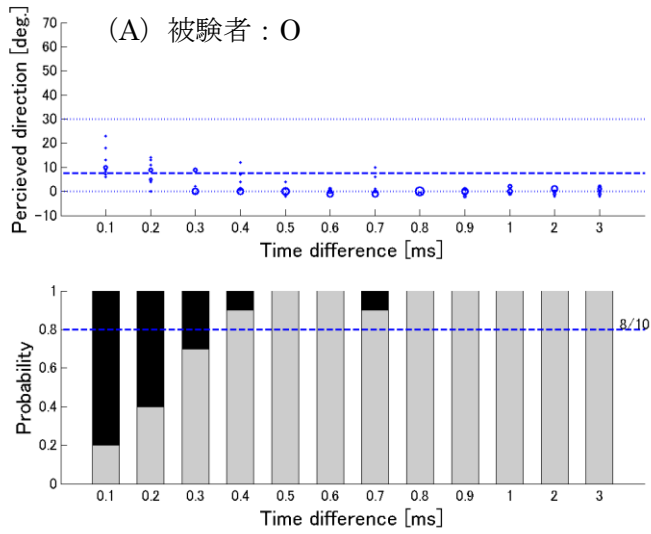


図 4.28 各被験者の先行音効果
の発生した割合と SLT (スピーチ
女, 映像 ON, $\pm 30^\circ$)

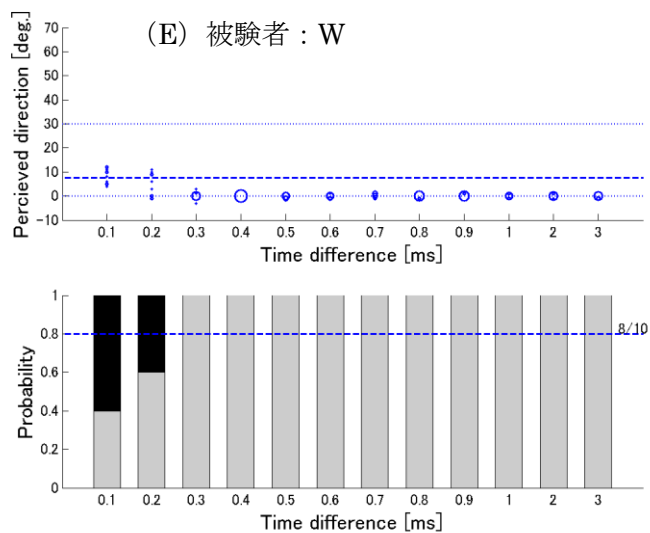
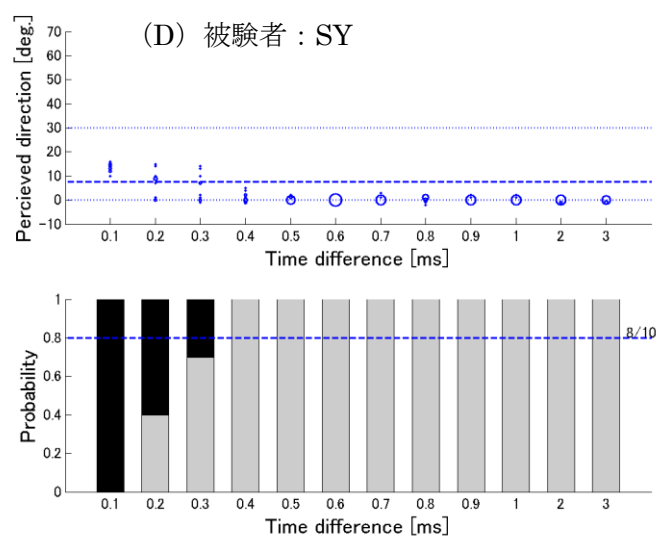
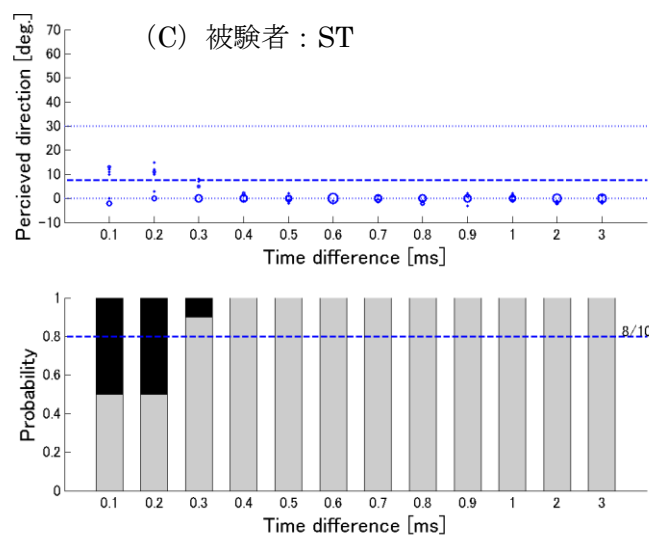
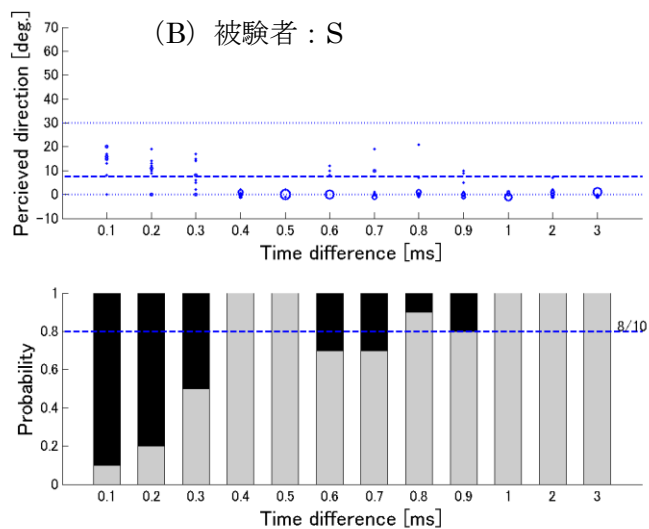
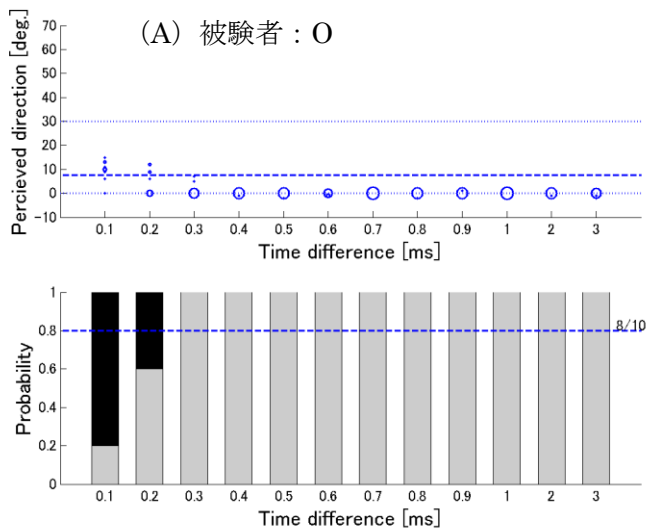


図 4.29 各被験者の先行音効果の発生した割合と SLT (スピーチ男, 映像 ON, $\pm 30^\circ$)

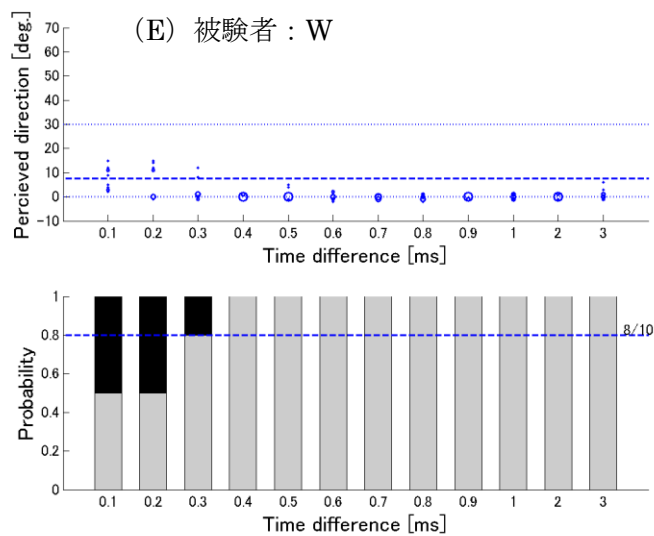
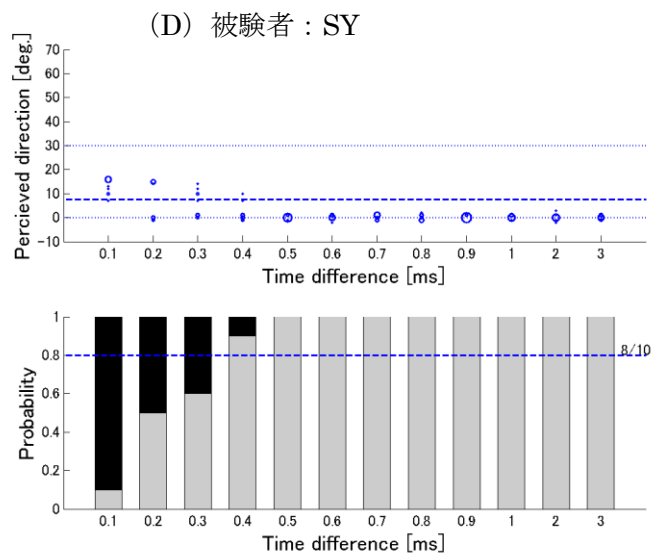
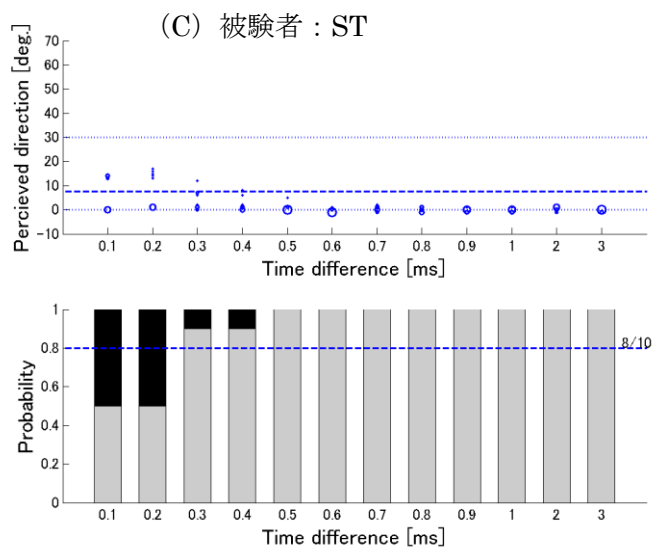
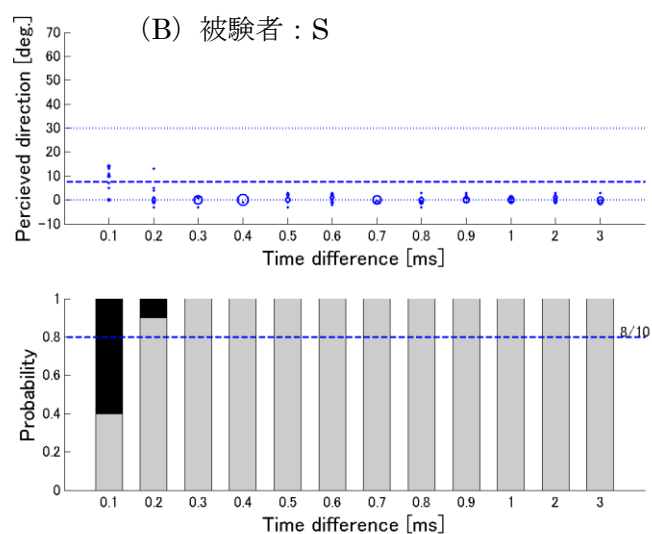
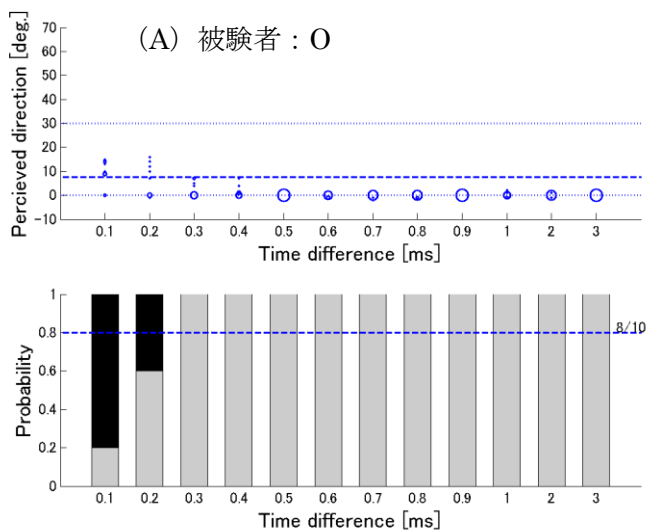


図 4.30 各被験者の先行音効果の発生した割合と SLT (小太鼓, 映像 ON, $\pm 30^\circ$)

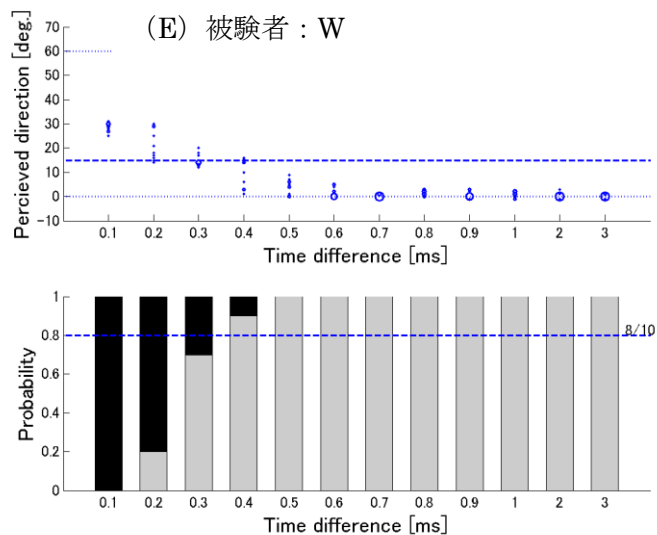
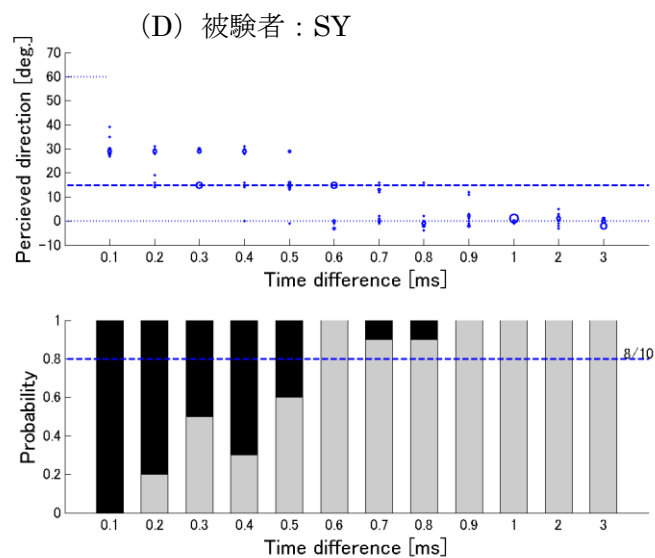
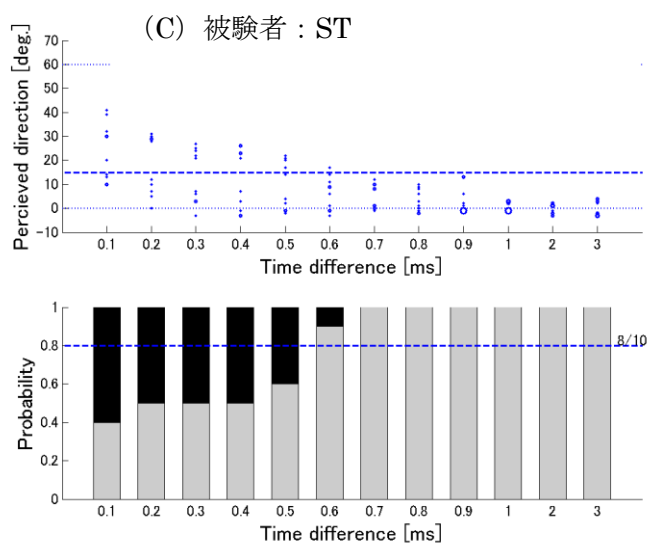
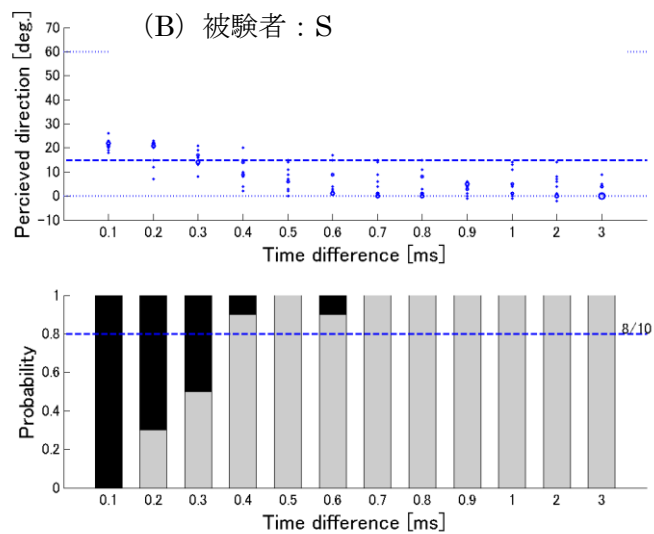
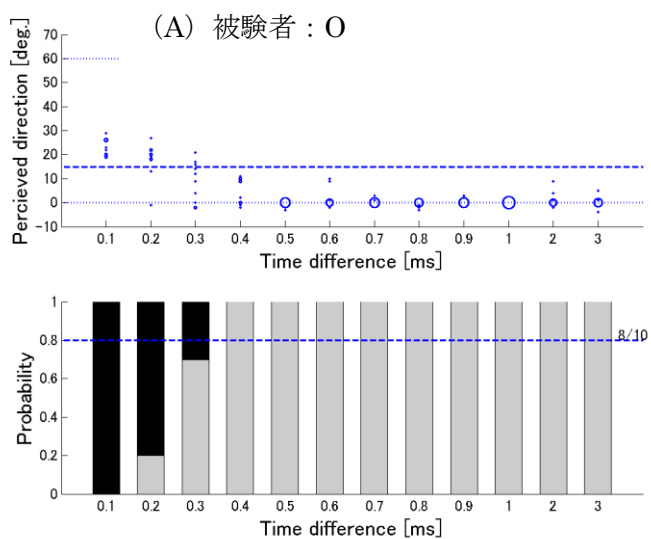


図 4.31 各被験者の先行音効果の発生した割合と SLT (ギター, 映像 ON, $\pm 60^\circ$)

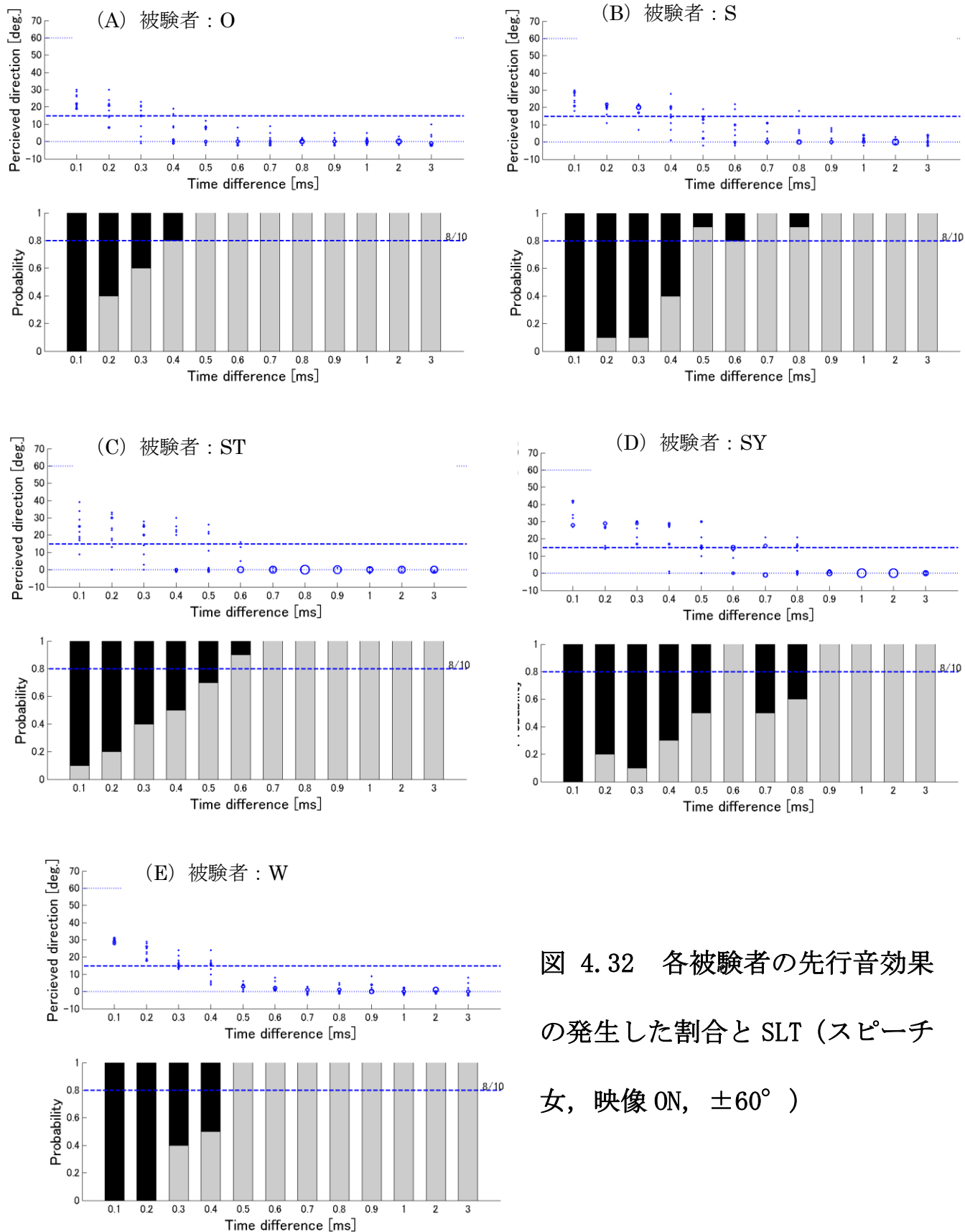


図 4.32 各被験者の先行音効果
の発生した割合と SLT (スピーチ
女, 映像 ON, $\pm 60^\circ$)

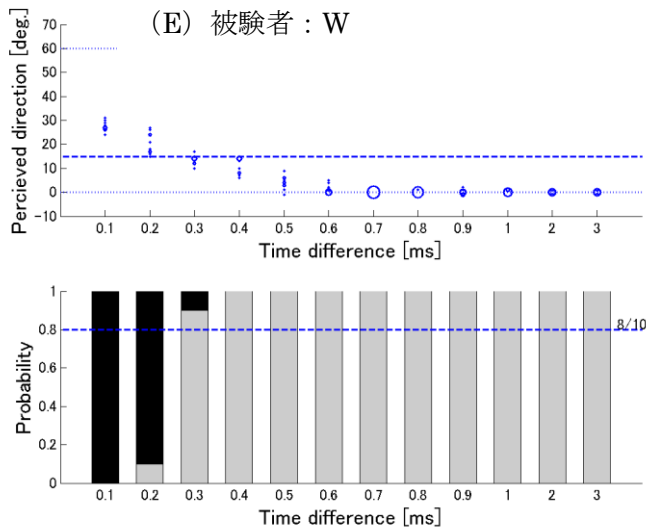
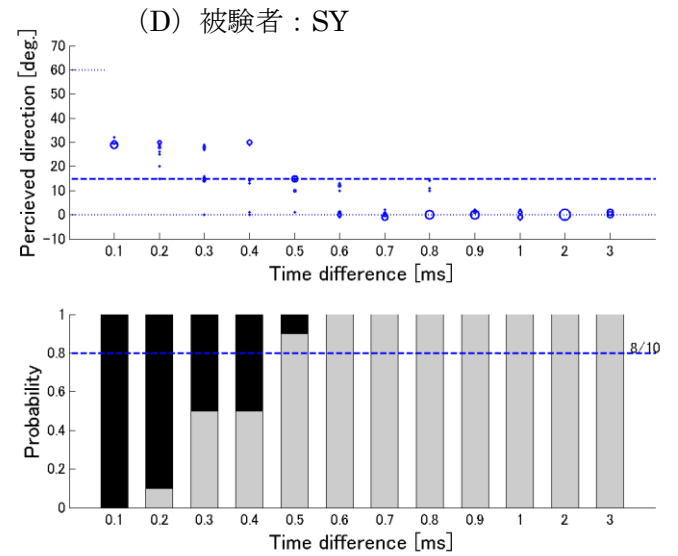
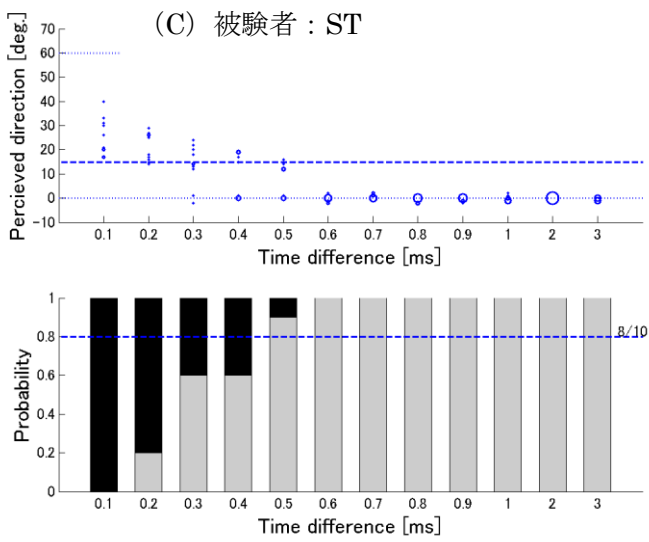
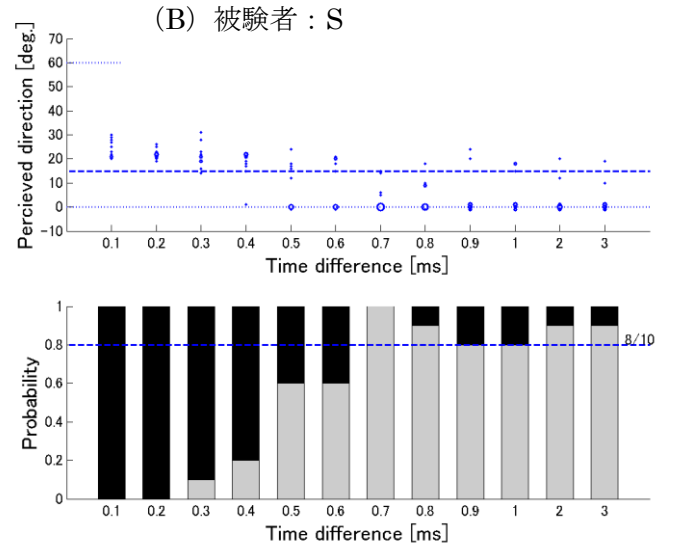
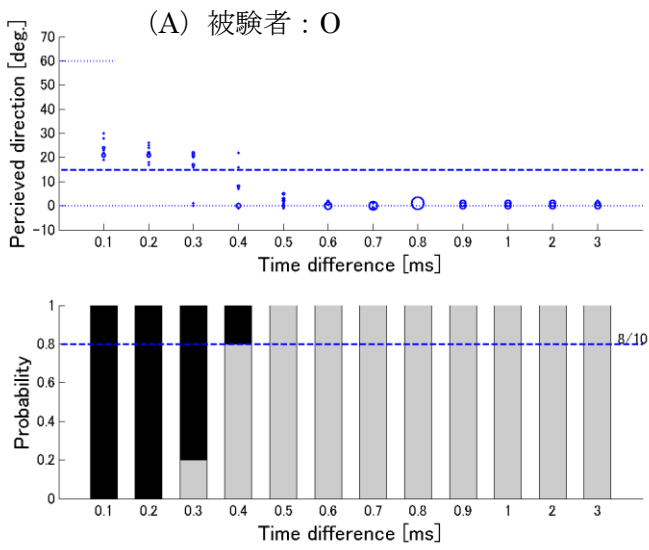


図 4.33 各被験者の先行音効果の発生した割合と SLT (スピーチ男, 映像 ON, $\pm 60^\circ$)

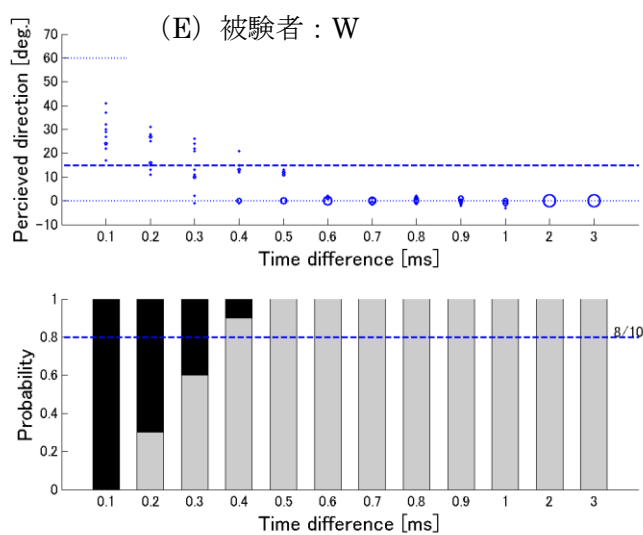
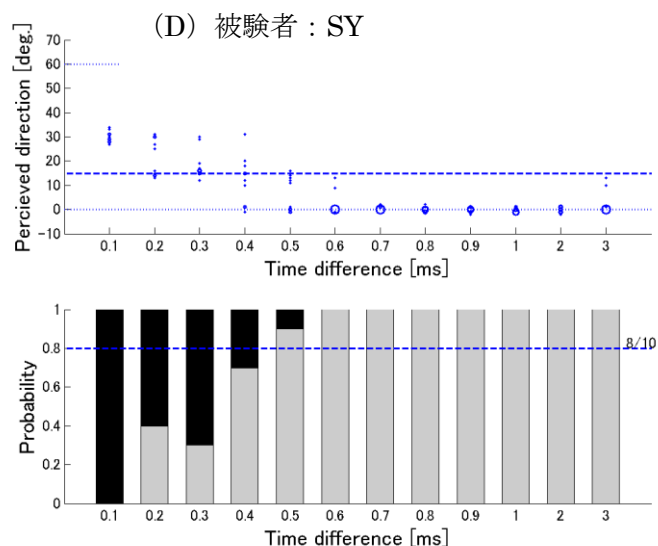
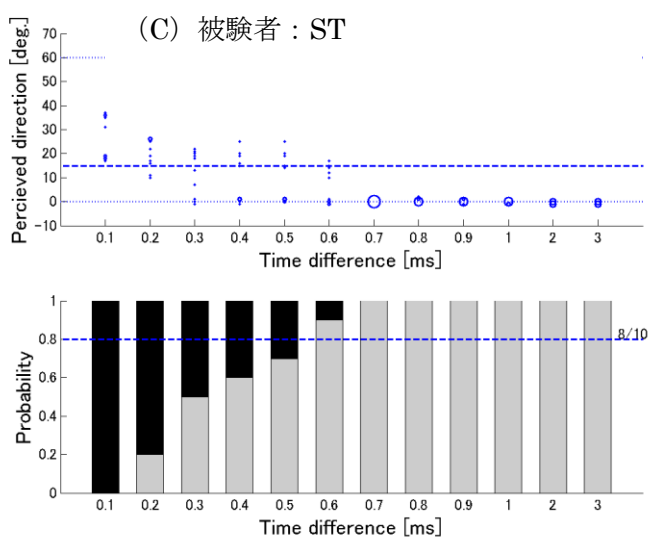
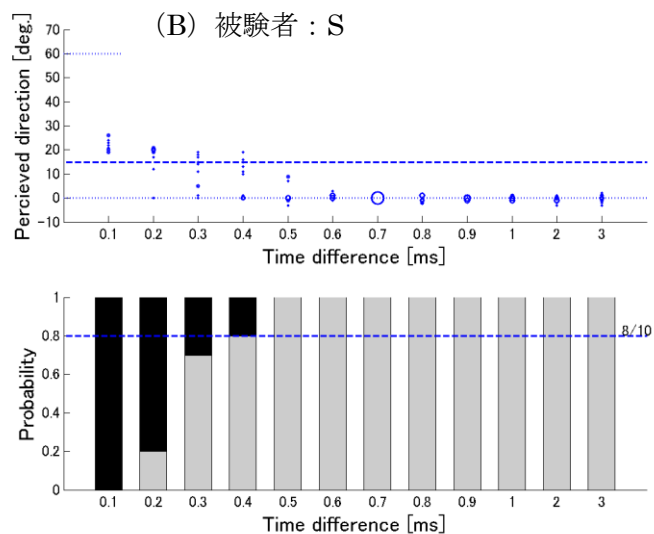
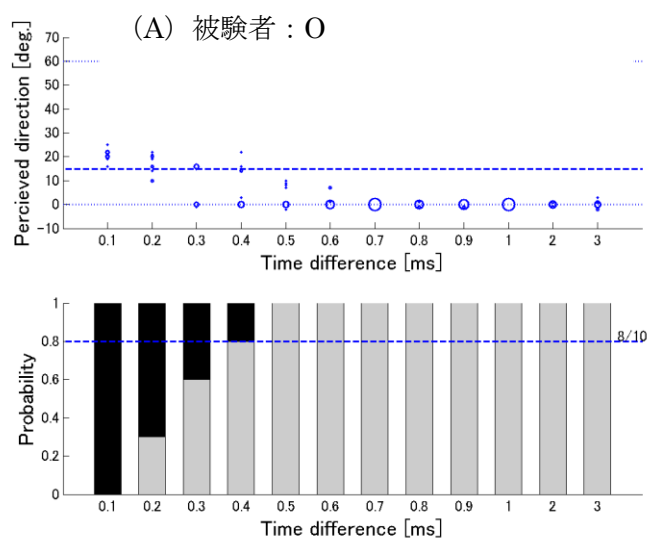


図 4.34 各被験者の先行音
効果の発生した割合と SLT
(小太鼓, 映像 ON, $\pm 60^\circ$)

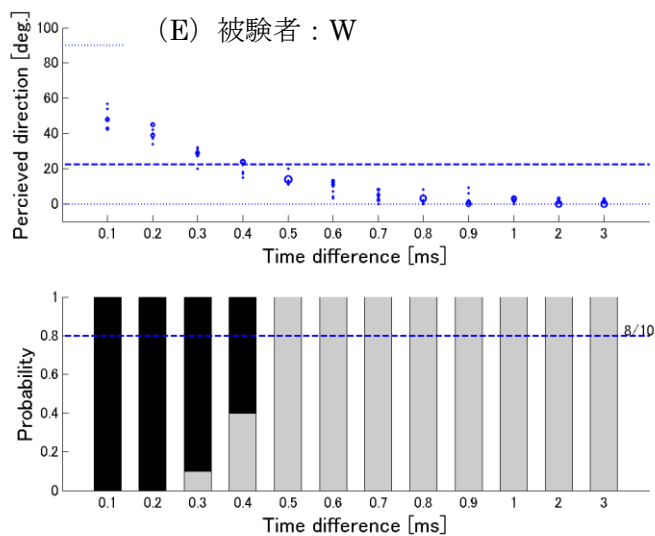
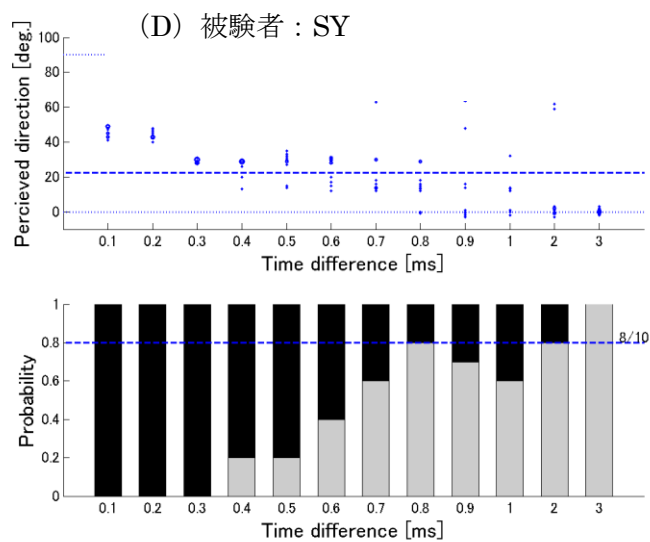
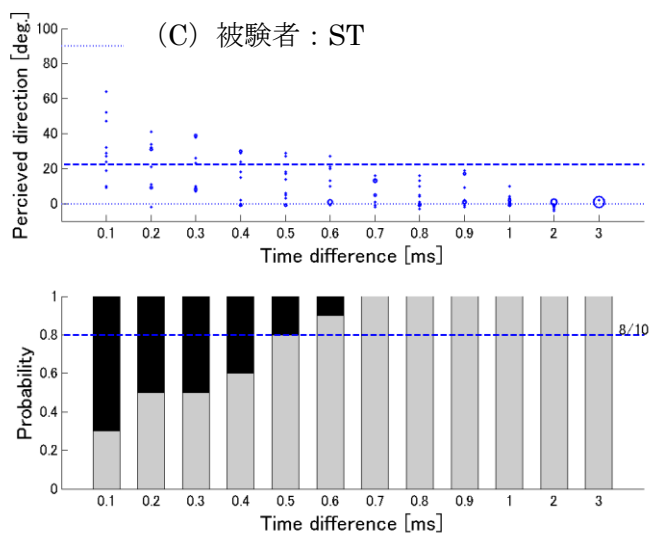
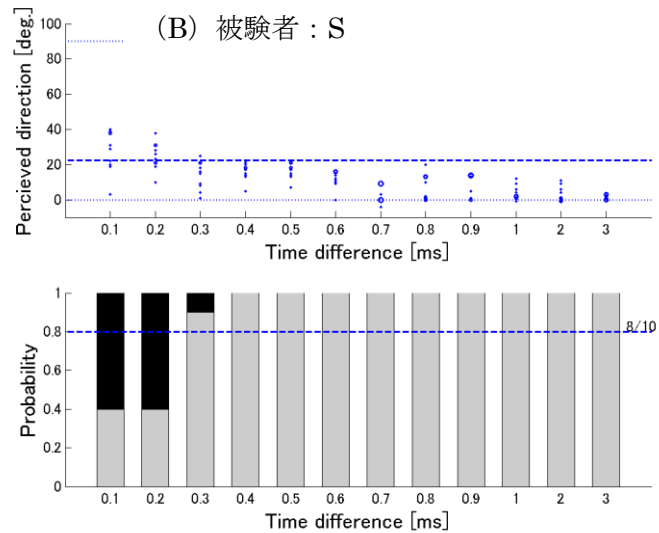
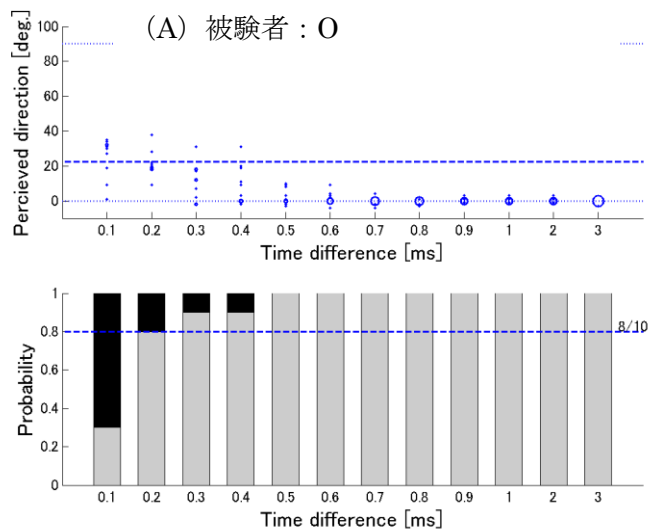


図 4.35 各被験者の先行音効果の発生した割合と SLT (ギタ一, 映像 ON, $\pm 90^\circ$)

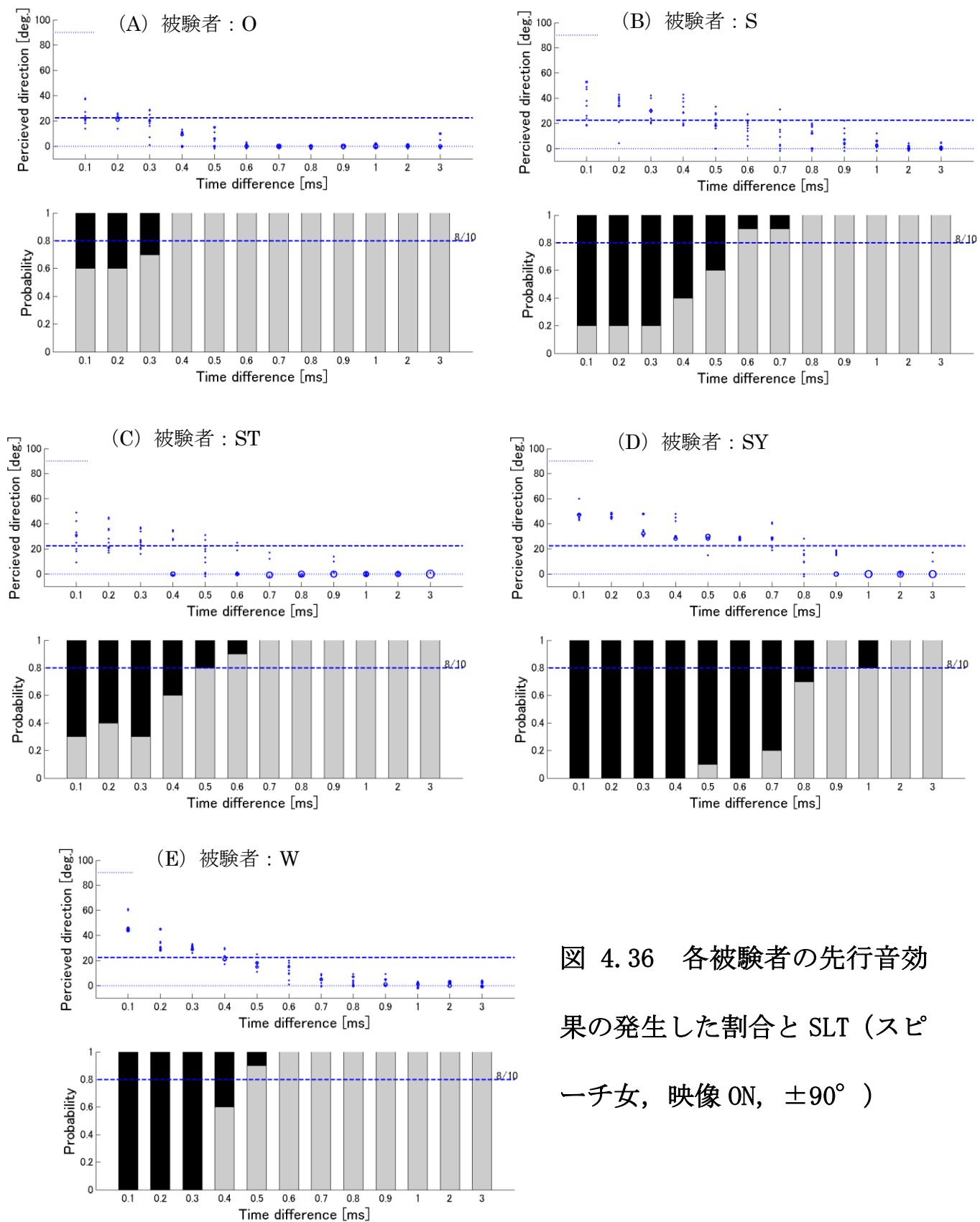


図 4.36 各被験者の先行音効
果の発生した割合と SLT (スピー
ーチ女, 映像 ON, $\pm 90^\circ$)

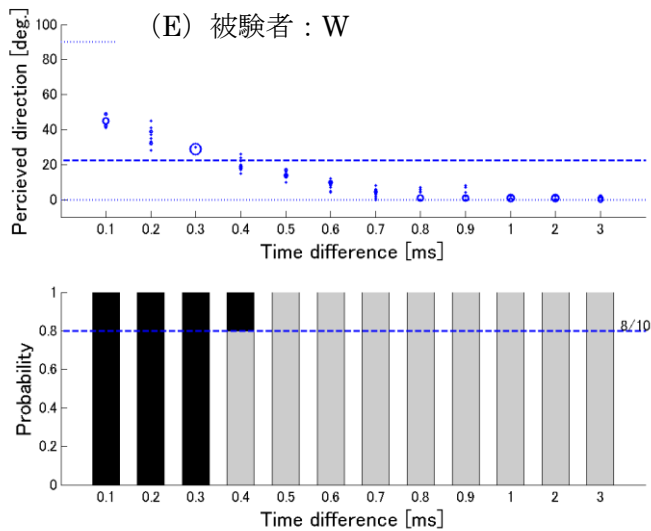
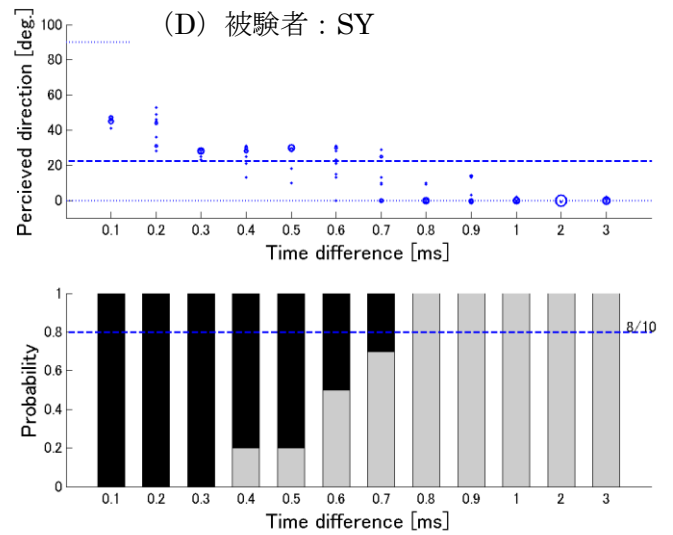
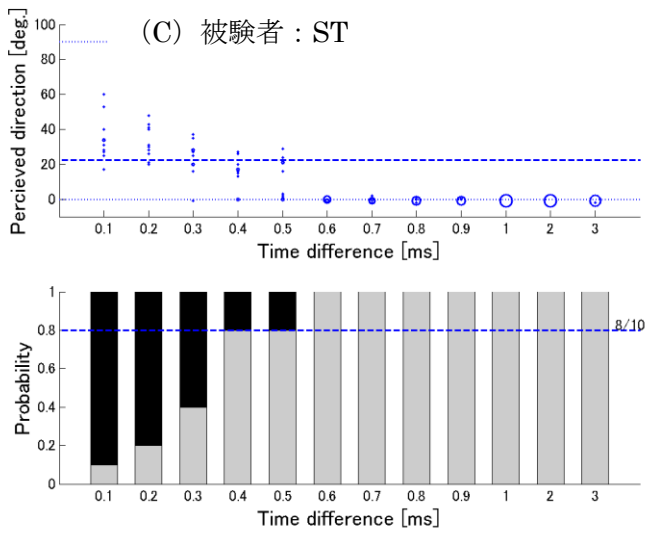
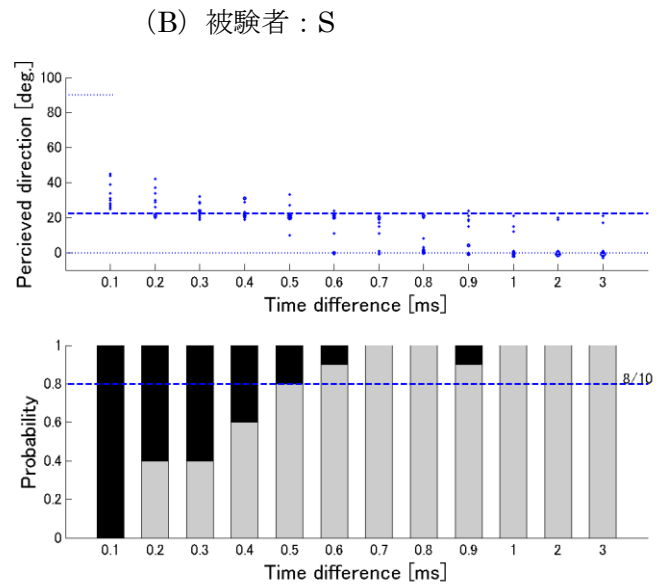
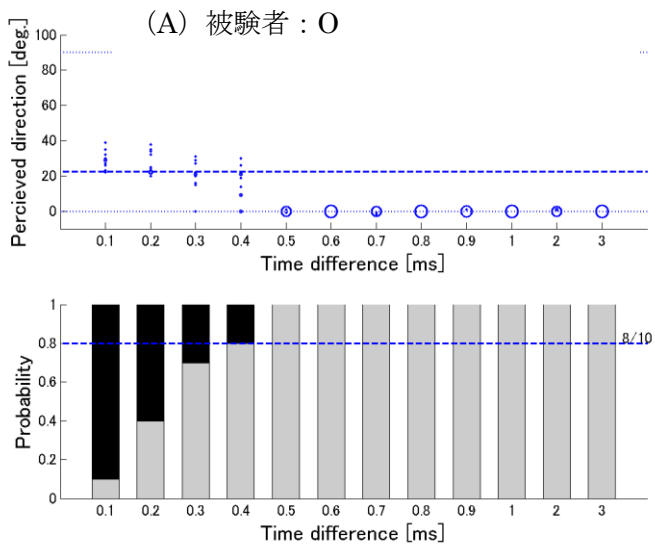


図 4.37 各被験者の先行音効果の発生した割合と SLT (スピーチ男, 映像 ON, $\pm 90^\circ$)

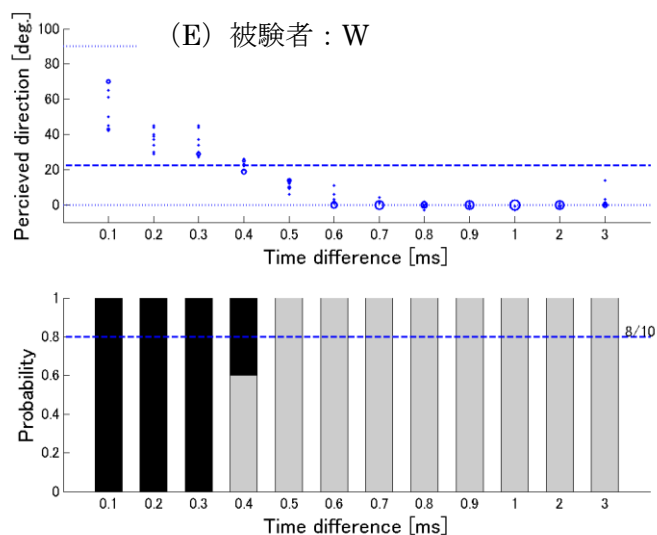
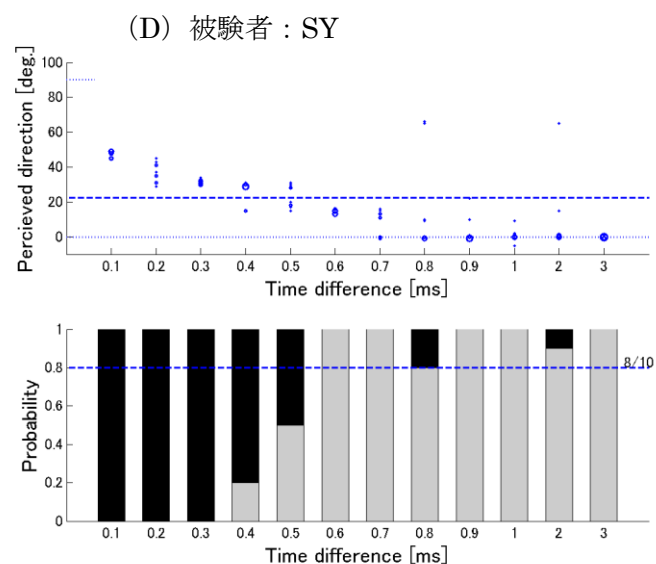
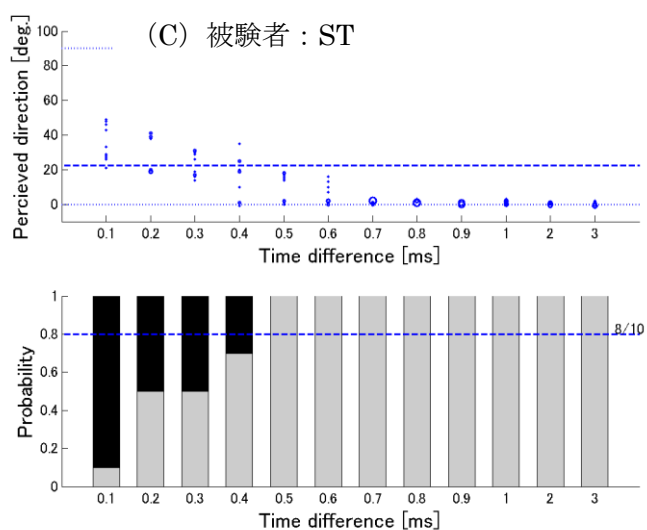
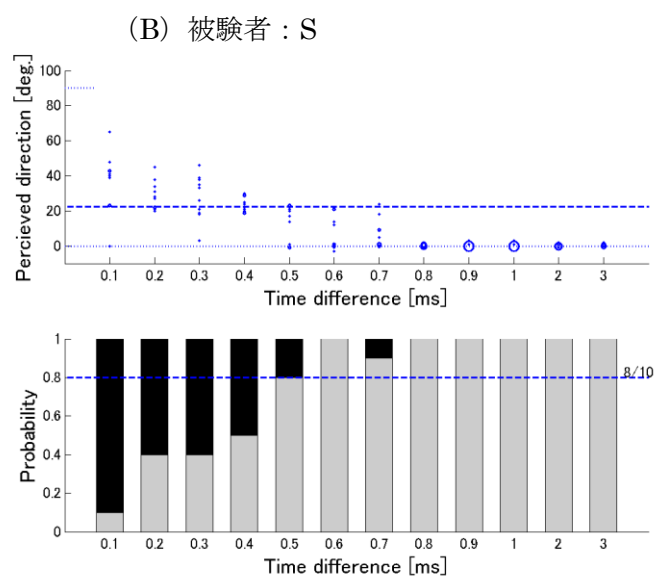
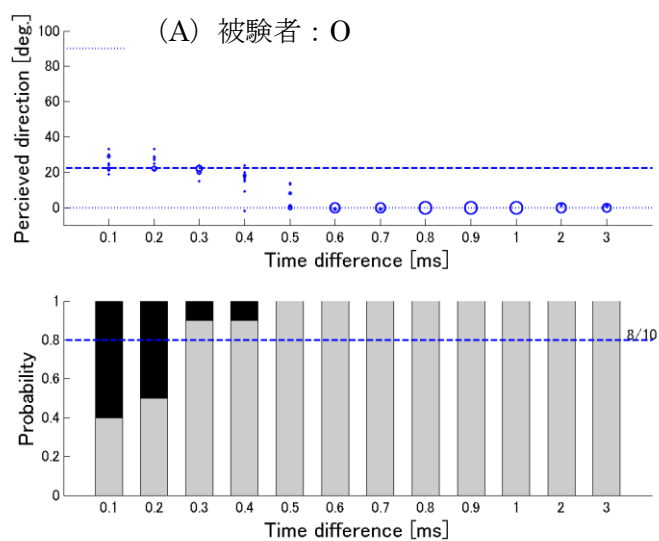


図 4.38 各被験者の先行音効果の発生した割合と SLT (小太鼓, 映像 ON, $\pm 90^\circ$)

4.5 ET実験

4.5.1 実験条件と回答方法

実験系や用いる視聴覚刺激は，SLT 実験と基本的には同じとなる．被験者も，SLT 実験に参加した 5 名である．ET 実験では，2 音源間に表 4.5 に示すような時間差を設けた．一番短い時間差である 1 ms は，明らかに音像が分離することなく，融合して知覚される時間差として決定した．一方，最も長い時間差である 30 ms は音像が分離して知覚されるであろう時間差として決定した．

ET 実験では，ランダム化最尤推定法を用いて ET を推定する．したがって，被験者には提示された音像が融合して一つの音像として知覚されるか，分離して二つの音像として知覚されるかを判定された．選択は，2 肢強制選択で，回答には SLT 実験にも使用したデータ収録装置を用いた．SLT 実験と同様に，データ収録装置の取り扱いと実験内容を口頭で説明し，練習を 10 回程度行った．回答は，音が 1 つに聞こえたら，収録装置の第一ボタンを押し，音が 2 つに聞こえたら，第二ボタンを押すよう指示した．各被験者の判断回数は，時間差を 30 条件で，その中に判断仮定が確定した刺激音（テスト音）を，10 条件を入れて，時間差の提示方向を 6 方向，刺激音 4 種類，映像の有無とした時の試行回数となる．被験者の回答回数が一人あたり， $30(\text{条件})+10(\text{テスト})\times 4(\text{刺激音の種類})\times 6(\text{スピーカの配置})\times 2(\text{映像「有／無」})=1920$ 回数となる．

表 4.5 ET の実験に用いた時間差

Delay times used in experiment : 30 values (ms)							
1.0	2.0	3.0	4.0	5.0	6.0	7.0	8.0
9.0	10.0	11.0	12.0	13.0	14.0	15.0	16.0
17.0	18.0	19.0	20.0	21.0	22.0	23.0	24.0
25.0	26.0	27.0	28.0	29.0	30.0		

被験者に提示した時間差：30 条件，単位[ms]，スタート基準は，1ms を先行音効果，30ms をエコーの領域と仮定して，その中間にある閾値を推定するランダム化最尤推定法を用いた．

4.5.2 ET実験の結果と考察

ET実験では，ETに対する視覚刺激の影響を検討した．被験者「O」による視覚刺激の「有／無」について，スピーカの開き角度左右 $\pm 60^\circ$ 方向の場合の実験結果の例を図4.39から図4.42に示す．横軸は，時間差[ms]を表し，縦軸は，時間差ごとに，音が1つか2つかに聞こえた割合を表す．図中の点線は音だけの場合で，実線が視覚刺激を付加した場合の結果を表している．具体的には，時間差ごとのすべての試行で音が1つに聞こえると回答した場合，縦軸の値は0と表示される．また，時間差ごとのすべての試行で音が2つに聞こえると回答した場合，縦軸の値は1と表示される．したがって，ここでは，縦軸の値が0.75（75%）になる遅延時間をETとした．図4.39 (a)からスピーチ(男声)に対し，視覚刺激無しのET_{off} は約4.2 ms，視覚刺激有りのET_{on} は約6.9 msであり，視覚刺激の存在によってETが約2.7 ms程度長くなっている．図4.40 (a)のスピーチ(女声)では，ET_{off} は約3.6 ms，ET_{on} は約17.1 msで，視覚刺激付加により，ET

は13.5 ms程度も長くなっている。図4.41 (a)の小太鼓では、 ET_{off} が約2.0 ms, ET_{on} が約4.3 msで、差は2.3 ms程度であった。また、図4.42 (a)のギターでは、 ET_{off} が約6.6 ms, ET_{on} は約6.3 msで、逆に短くなっているが、その差は0.3 msで非常に小さい。全体的に、被験者「O」では、視覚情報の存在によって、ETは長くなる傾向が見られる。被験者「ST」の場合は、スピーチ(男,女)と小太鼓について、被験者「O」とは逆に、 ET_{on} の結果が ET_{off} の結果より小さい遅延時間で起こり、視覚刺激の影響によって、ETが短くなる傾向を示したが、被験者「O, ST」ともに、左右の角度でほぼ対称的な映像の付加の影響が確認された。さらに、視覚刺激の付加による影響の差の有意性を、ETの平均値について t 検定で調べた。その一例として、被験者「O」を表4.6に、被験者「ST」を表4.7に、 t 検定を行った結果を示した。表中の有意性の記号は、視覚刺激を加えた結果、有意にETが「短い・長い」になった場合が○、視覚情報を加えても、有意の差が見られなかった場合を×で示した。ただし、危険率0.01として、例えば、表4.6は音源配置が 60° 方向で、スピーチ(男声)の場合は、 $t(49)_{0.01}=13.2$ である。自由度が49で t 値が13.2であることを示している。これは、標本の大きさを $n_1=n_2=50$ として、 t 分布表から自由度 $d.f.=49$ で危険率0.01の場合 2.678である。すなわち、 $2.678 < 13.2$ で境界線より外部にあり、映像ONと映像OFFとの差は、有意であることが示されて(○印)、有意に長くなっていることが分かる。また、○印で網掛けをしてある部分が、有意の差がある場合である。被験者「O」による結果では、 ET_{off} より ET_{on} に対応する遅延時間の方が大きく、視覚刺激の影響によりETが長くなる傾向が見て取れる。被験者「ST」では、スピーチ、小太鼓に対する ET_{on} は ET_{off} よりも短くなるが、ギターでは逆に長くなっている。このように、被験者によって、また、刺激の種類によって、視覚刺激付加の影響は若干異なる。図4.43～図4.58は、各被験者に対して、音源配置が $\pm 30^\circ$, $\pm 60^\circ$, $\pm 90^\circ$, 刺激音がスピーチ(男女)、小太鼓、ギターで、視覚刺激の「有／無」によってエコー検知限ET付近で音像が1つか2つかに聞こえたかを

判断した割合を図示したものである。横軸は2音源の時間差、縦軸に時間差ごとに、音が1つか2つかに聞こえた割合を示した。これらの結果から被験者5名について、表4.8に視覚刺激付加によるETの変化をグループ分けした結果を示した。表では、①映像付加でETが短くなる、②映像付加でETが長くなる、③映像付加の影響は見られないの三つの区分で分けた。この表をみると、スピーチ、小太鼓に対して、被験者「S, ST」では、視覚刺激を付加するとETが短くなり、被験者「O, SY」では長くなる。刺激音がギターの場合、被験者5名中4名において、視覚刺激付加によりETが長くなっている。一方、被験者「W」では、視覚刺激付加の影響は見られない。このように、視覚刺激の影響が有意と見られるとはいえ、被験者によってその効果が異なることは、SLT実験の考察でも述べたように、聴覚と視覚に向けられた能動的あるいは受動的に働く注意力が、被験者やその測定環境によって異なったのではないかと考えられ、聴覚刺激と視覚刺激のどちらに注意を集中させるかによって、エコーの知覚に違いが生じたものと推論される。スピーチや小太鼓では、音源の視覚的な動きだけでなく、音の動的节奏感に注意が引きつけられるということが起こり得るのではないかと考えられる。この場合は、音の立ち上がり部分を手掛かりとして判断することになり、それが視覚刺激とどうかかわりあうのかは、各被験者の注意力によって異なるということであろうか。森本ら¹⁴⁾は、刺激を音声(女声)、ソプラノリコーダ、アルトリコーダの3種類を用いて、音刺激と同時に視覚刺激を提示した場合の実験を行っている。視覚刺激の影響について、*t*検定を行った結果、両者の間に有意の差($p < 0.05$)はあるが、ETの値は、短くなる場合も長くなる場合もある不安定な状態で、明確な傾向は認められないとしている。この状況では、本実験と、ほぼ同様な傾向であるが、刺激音の違いによる影響については、詳細に述べられていない。本報告では、スピーチあるいは小太鼓が音源である場合、視覚刺激の付加でETは有意に短くなったり長くなったりした。この傾向は被験者によって異なるが、ギターの演奏音の場合には、ほとんどの被験者におい

て、視覚刺激の付加によりETが有意に長くなることが示された(6ms程度)。視覚刺激の有無によってETが大きくなるか、小さくなるかをSLT実験の際と同様に、各被験者ごとに集計した結果を表4.2.9に示す。表から分かるように、大半の条件において視覚刺激が有るとETが大きくなる傾向が示された。先行音効果に対する視覚刺激の影響は、SLTにおいては聴覚刺激と視覚刺激に近いほど先行音効果の生じる時間差は短く、速い時間差で先行音効果が起こる腹話術効果に類似している現象と考えられた。ETにおいては刺激音の種類に若干影響が有るものの、エコー検知限ETを長くする傾向が見られた。

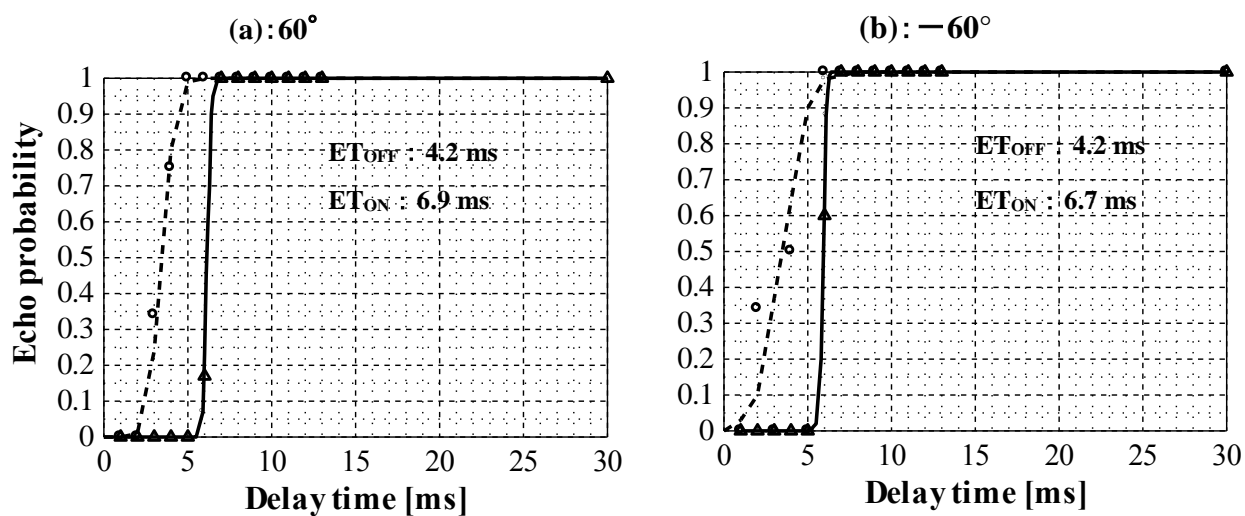


図 4.39 2 音源間の時間差とエコー知覚確率との関係

刺激音：スピーチ(男声)，被験者：O，映像の「有／無」による判断比較，
サブスピーカの開き角度：60°，-60°方向，点線：映像ない OFF，実線：映像ある ON.

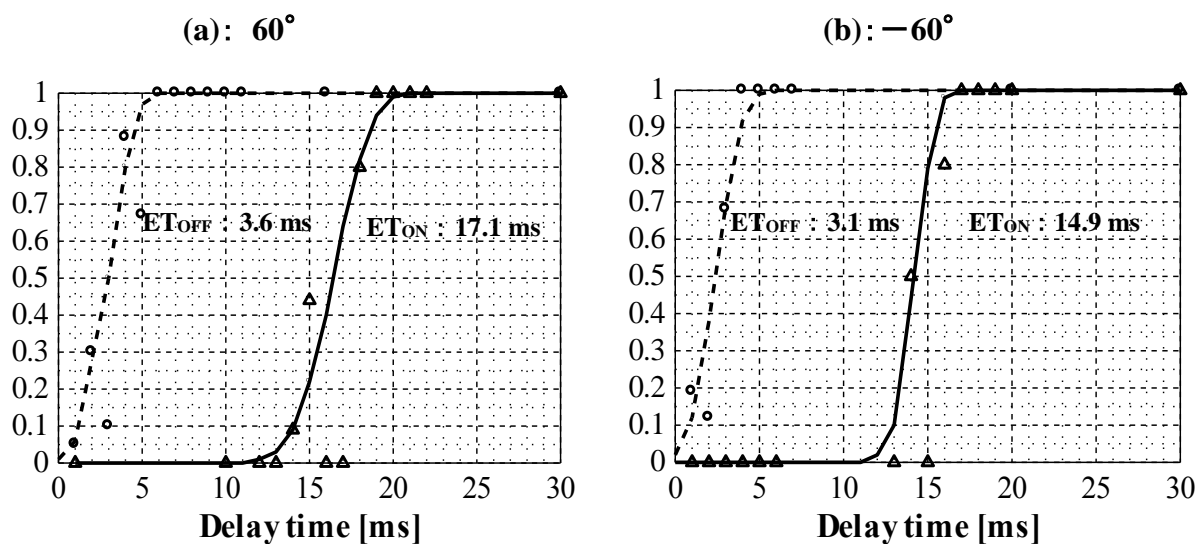


図 4.40 2 音源間の時間差とエコー知覚確率との関係

刺激音：スピーチ(女声)，被験者：O，映像の「有／無」による判断比較，
サブスピーカの開き角度：60°，-60°方向，点線：映像ない OFF，
実線：映像ある ON.

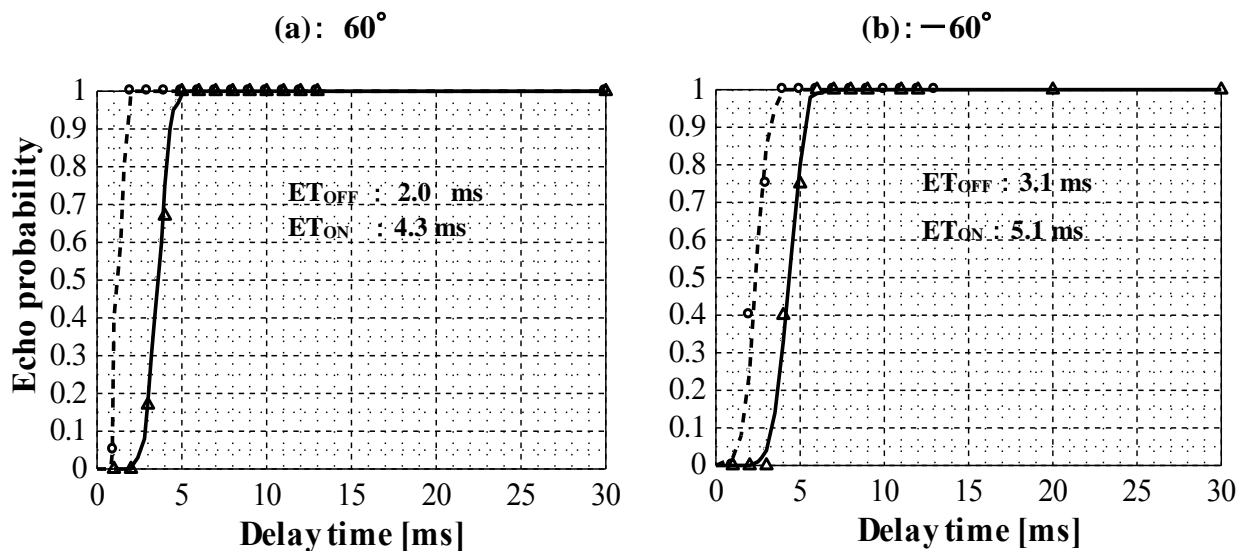


図 4.41 2 音源間の時間差とエコー知覚確率との関係

刺激音：小太鼓を叩く音，被験者：O，映像の「有／無」による判断比較，サブスピーカの開き角度：60°， -60°方向， 点線：映像ない OFF，実線：映像ある ON.

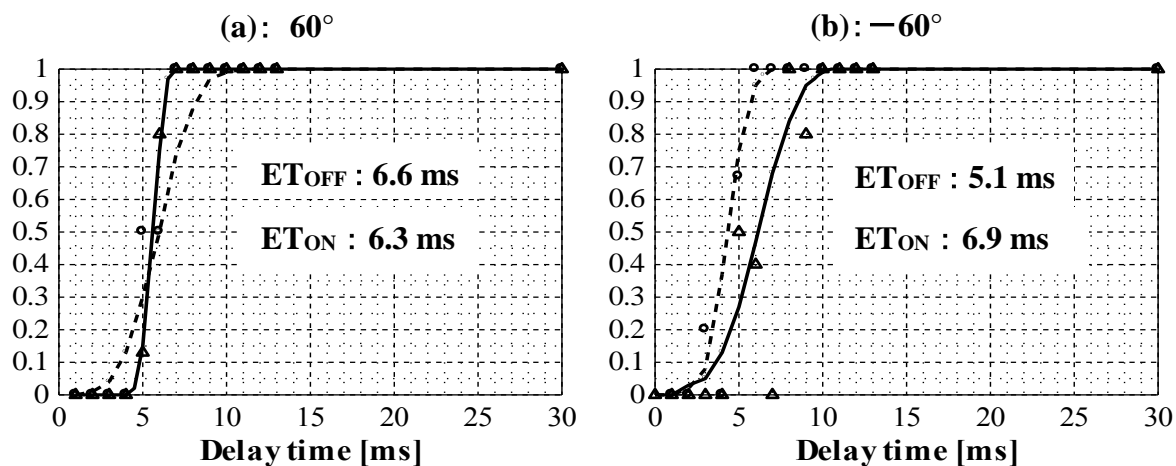


図 4.42 2 音源間の時間差とエコー知覚確率との関係

刺激音：ギターの演奏音，被験者：O，映像の「有／無」による判断比較，サブスピーカの開き角度：60°， -60°方向，点線：映像ない OFF，実線：映像ある ON.

表4.6 平均値差に対する有意差の t 検定 被験者:「O」

Direction of sub-speaker	Speech (male voice)			Speech (female voice)			Snare drum			Guitar		
	SD	OFF	ON	SD	OFF	ON	SD	OFF	ON	SD	OFF	ON
90°	○ t(49)= 7.5	5.3	7.8	○ t(49)= 31.4	4.8	11.2	○ t(49)= 4.4	2.6	3.5	× t(49)= 2.5	5.7	6.2
60°	○ t(49)= 13.2	4.2	6.9	○ t(49)= 66.1	3.6	17.1	○ t(49)= 11.3	2.0	4.3	× t(49)= 2.5	6.6	6.3
30°	○ t(49)= 7.8	5.1	6.7	○ t(49)= 58.3	5.6	17.5	× t(49)= 1.5	2.8	3.1	× t(49)= 1.5	4.1	4.8
-30°	○ t(49)= 12.3	4.7	7.2	○ t(49)= 65.2	4.9	18.2	○ t(49)= 14.7	3.1	6.1	○ t(49)= 8.3	4.7	6.8
-60°	○ t(49)= 12.2	4.2	6.7	○ t(49)= 57.8	3.1	14.9	○ t(49)= 11.8	3.1	5.1	○ t(49)= 8.8	5.1	6.9
-90°	○ t(49)= 5.4	5.2	6.3	○ t(49)= 42.1	2.8	11.4	○ t(49)= 4.9	4.0	3.0	× t(49)= 2.5	8.0	7.4

○：視覚刺激を加えた結果，有意にET付近が長くなった場合，×：視覚情報を加えた結果，有意の差が見られない場合

表4.7 平均値差に対する有意差の t 検定，被験者:「ST」

Direction of sub-speaker	Speech (male voice)			Speech (female voice)			Snare drum			Guitar		
	SD	OFF	ON	SD	OFF	ON	SD	OFF	ON	SD	OFF	ON
90°	○ t(49)= 11.3	7.6	5.3	○ t(49)= 21.1	10.1	5.8	× t(49)= 2.0	4.2	4.6	○ t(49)= 48.9	8.9	18.9
60°	○ t(49)= 9.8	7.4	5.4	○ t(49)= 12.7	10.1	7.5	○ t(49)= 9.8	7.1	5.1	○ t(49)= 40.2	6.4	14.6
30°	○ t(49)= 7.3	7.5	6.0	○ t(49)= 23.0	10.7	6.0	○ t(49)= 13.2	7.1	4.4	○ t(49)= 29.4	7.5	13.4
-30°	○ t(49)= 8.8	7.2	5.4	○ t(49)= 17.6	10.0	6.4	× t(49)= 1.5	4.4	4.7	○ t(49)= 7.4	7.9	9.4
-60°	× t(49)= 0.98	5.9	5.7	○ t(49)= 15.2	9.5	6.4	○ t(49)= 6.9	4.1	5.5	○ t(49)= 20.1	13.2	9.1
-90°	× t(49)= 0.0	4.8	4.8	○ t(49)= 25.0	10.1	5.0	○ t(49)= 5.9	4.9	3.7	○ t(49)= 40.2	19.7	11.5

○：視覚刺激を加えた結果，有意にET付近が短くなった場合，×：視覚情報を加えた結果，有意の差が見られない場合である。

表4.8 被験者ごとの視覚刺激のET付近に対する影響

刺激の種類	① 映像付加で短くなる	② 映像付加で長くなる	③ 映像付加で影響がない
スピーチ(男,女) 小太鼓を叩く音	S, ST	O, SY	W
ギターの演奏音	無し	O, ST, SY, W	S

①：映像を付加するとET付近が短くなる場合，②：映像を付加するとET付近が長くなる場合

③：映像を付加しても影響がない場合，被験者名：S, ST, O, SY, W の 5名

スピーチ(男性・女声)と小太鼓では，被験者「ST」は，被験者「O」とは異なり，視覚刺激によって，短い遅れ時間でET付近を知覚している．ギターの場合は，被験者「ST」と「O」が共に，視覚刺激によって，長い遅れ時間でET付近を知覚している．

表 4.9 被験者ごとの視覚刺激の影響

Sub-speaker の方向	スピーチ(男声)		スピーチ(女声)		小太鼓		ギター	
	①	②	①	②	①	②	①	②
±30°	20	70	20	80	40	60	20	80
±60°	30	60	30	60	30	60	40	60
±90°	40	60	50	40	20	80	10	80

視覚刺激の付加によるETの知覚, ただし①, ②でないものは除いた(単位%)

① 視覚刺激を加えた結果，ET が短くなった被験者の割合

②：視覚刺激を加えた結果，ET が長くなった被験者の割合

赤で囲んだ部分：ET の変化が生じた被験者が 60%以上

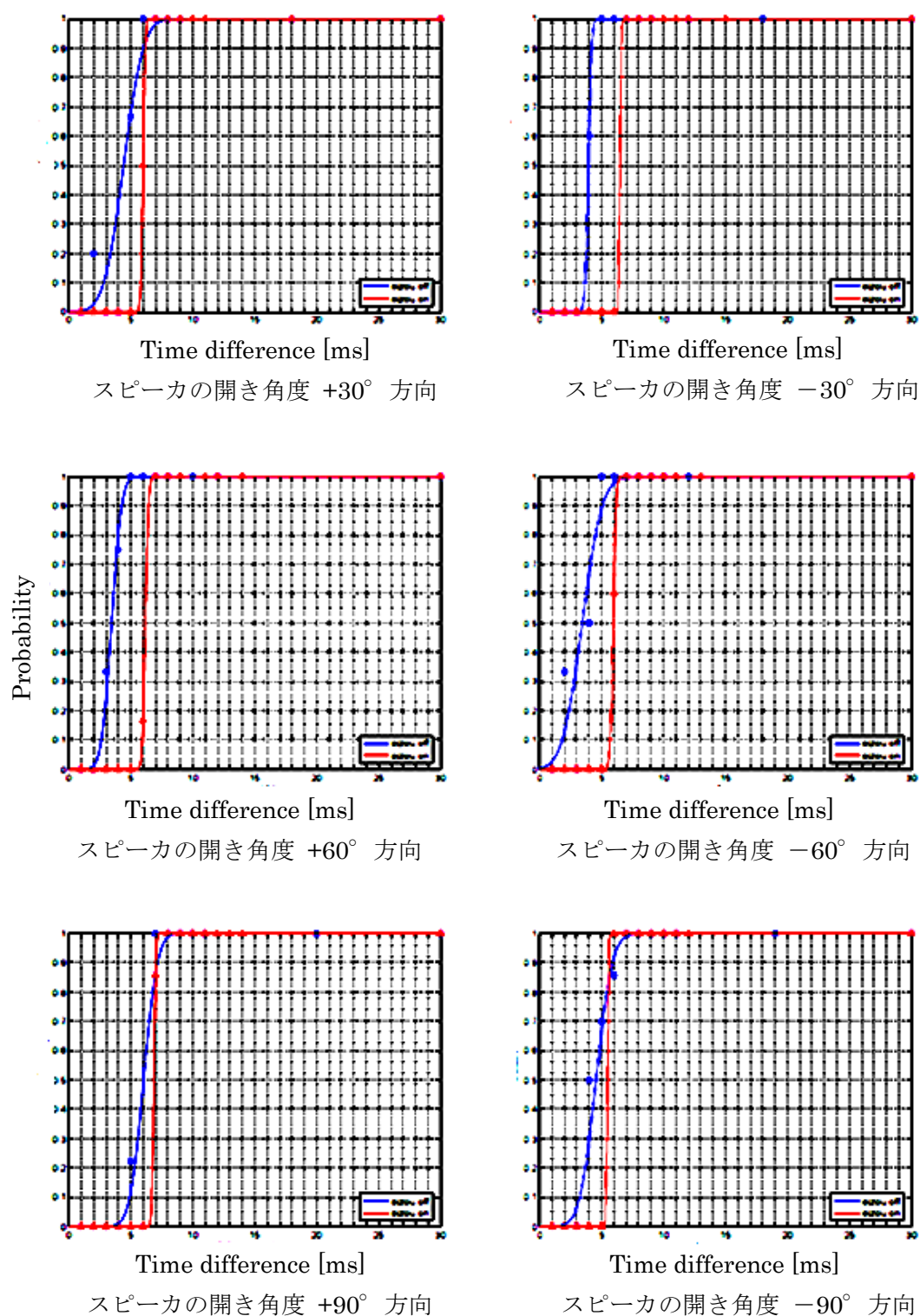


図 4.43 2 音源間の時間差とエコー知覚確率 (映像：有/無)

被験者：0, 青：映像無し OFF, 赤：映像有り ON, 刺激音：スピーチ(男声)

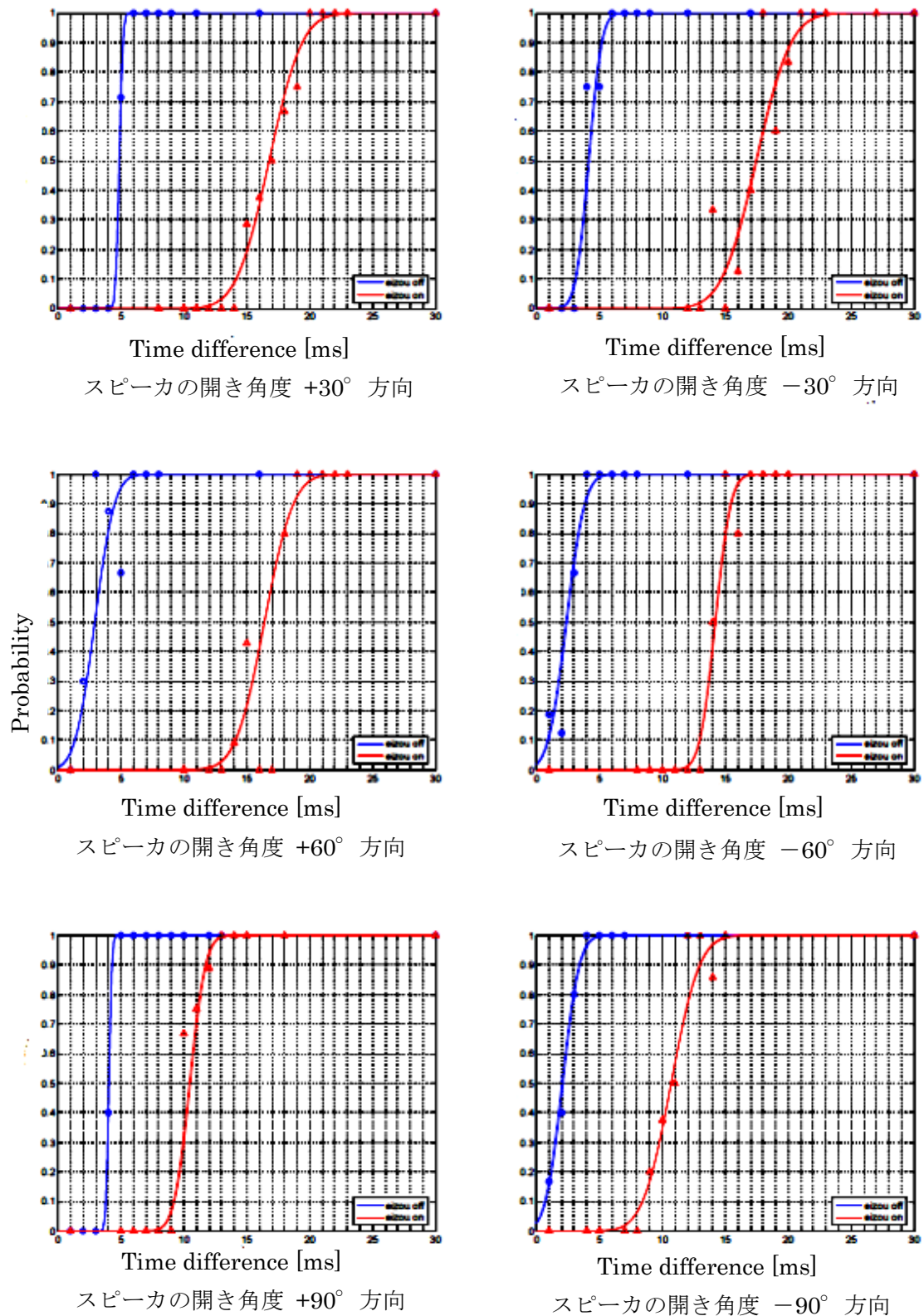


図 4.44 2 音源間の時間差とエコー知覚確率 (映像：有/無)

被験者：0, 青：映像無し OFF, 赤：映像有り ON, 刺激音：スピーチ(女声)

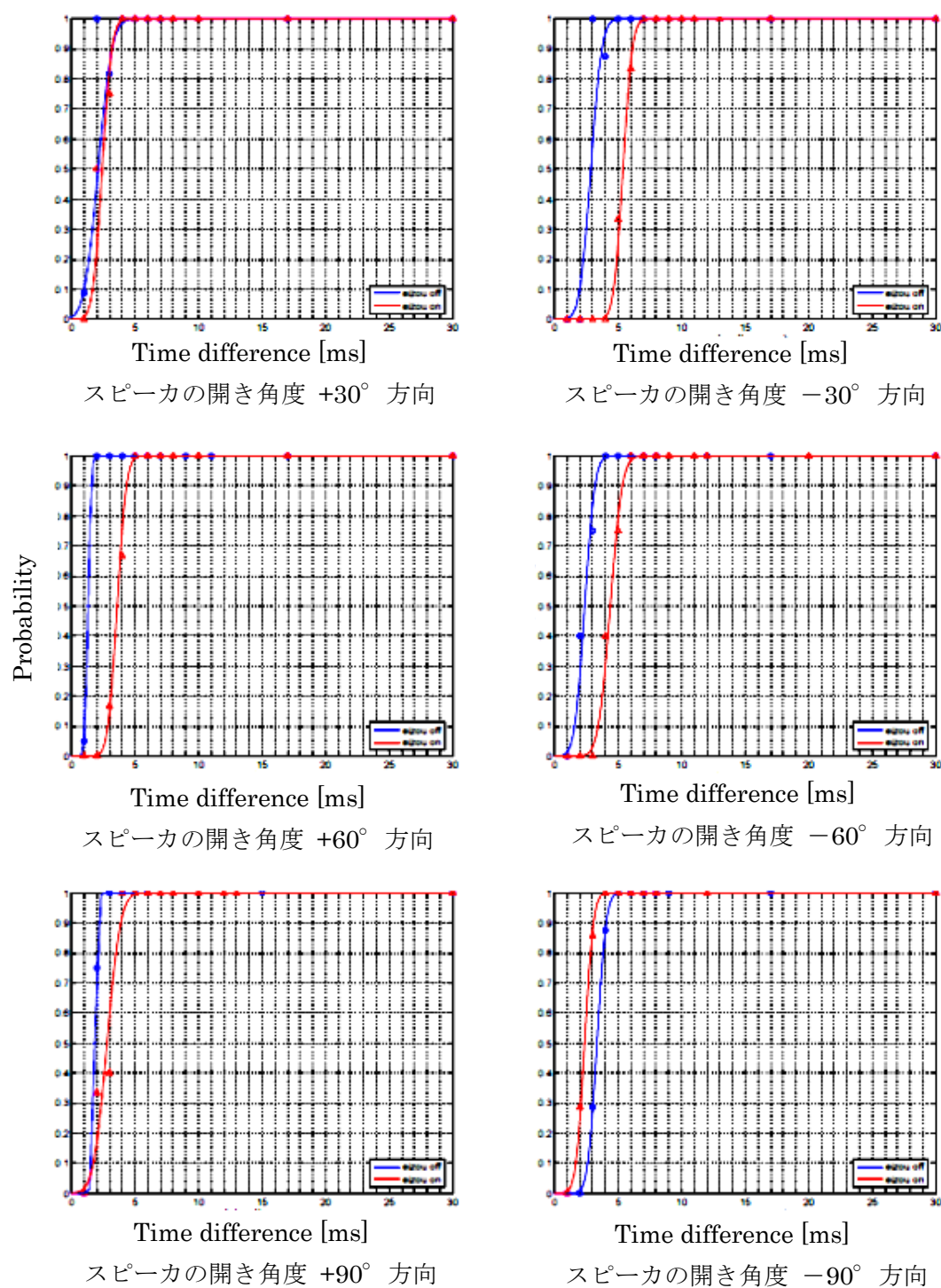


図 4.45 2 音源間の時間差とエコー知覚確率 (映像：有/無)

被験者：0，青：映像無し OFF，赤：映像有り ON)，刺激音：小太鼓

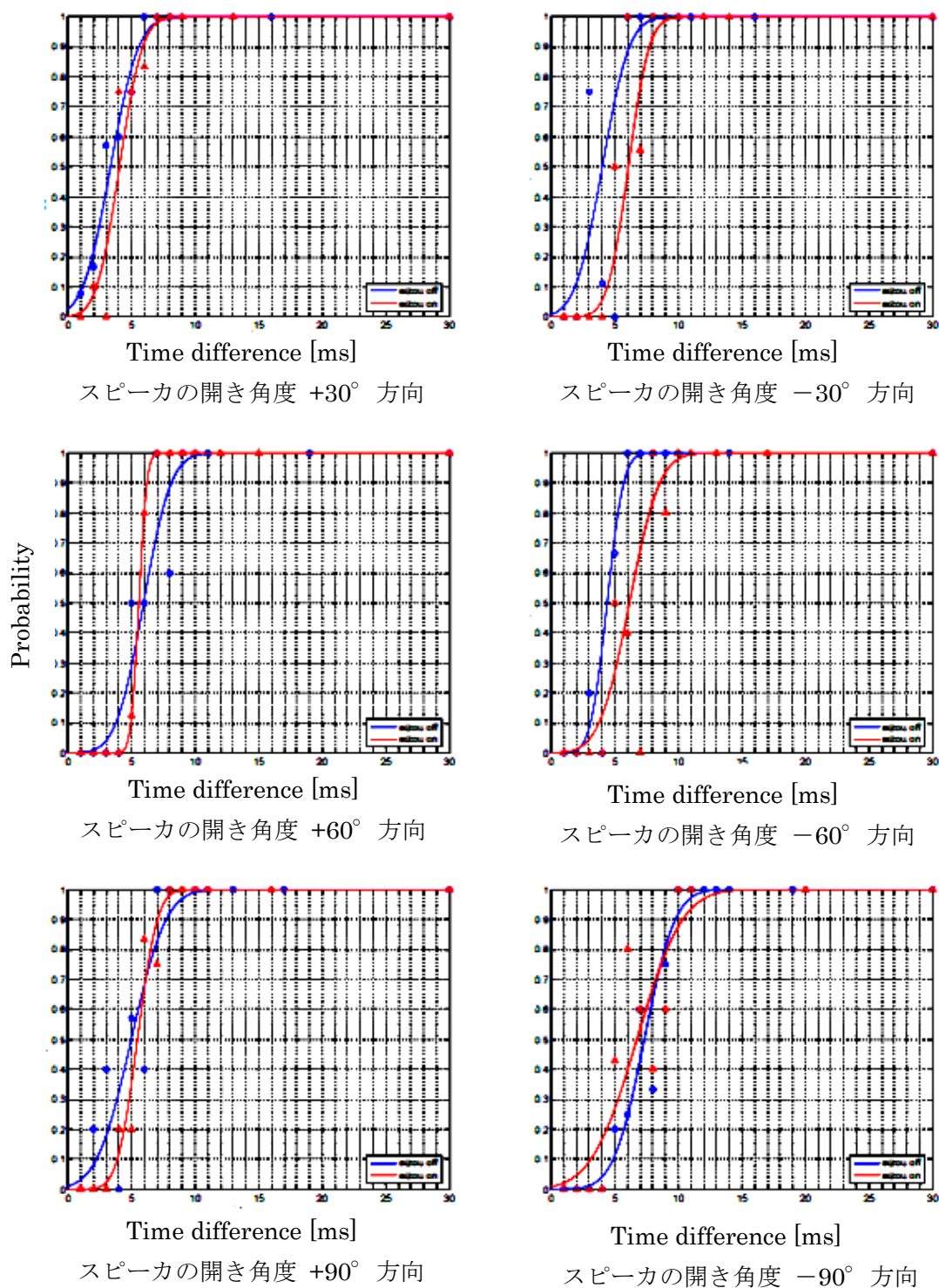


図 4.46 2 音源間の時間差とエコー知覚確率（映像：有/無）

被験者：0，青：映像無し OFF，赤：映像有り ON），刺激音：ギター

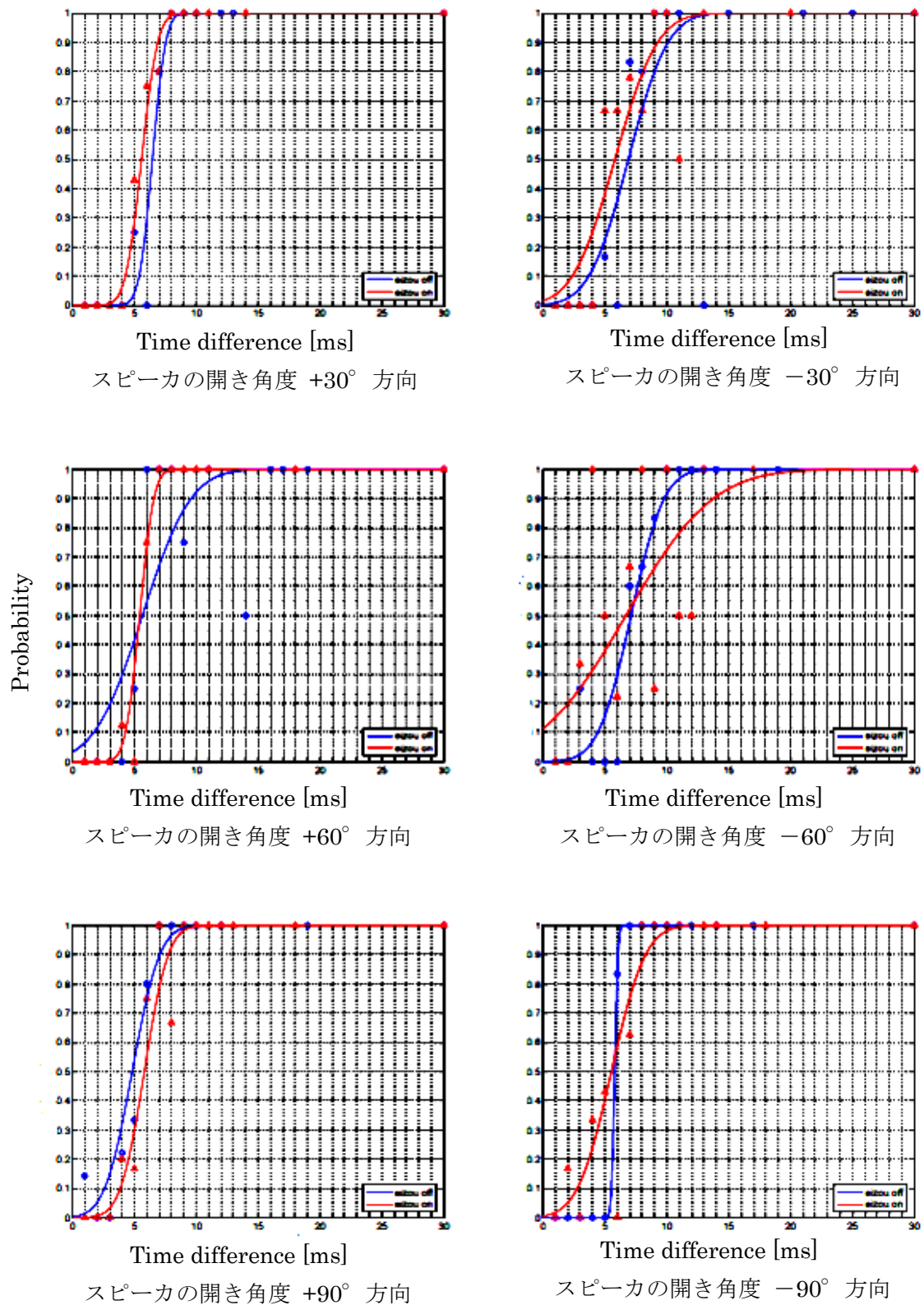


図 4.47 2 音源間の時間差とエコー知覚確率 (映像：有/無)

被験者：S. 青：映像無し OFF, 赤：映像有り ON, 刺激音：スピーチ(男声)

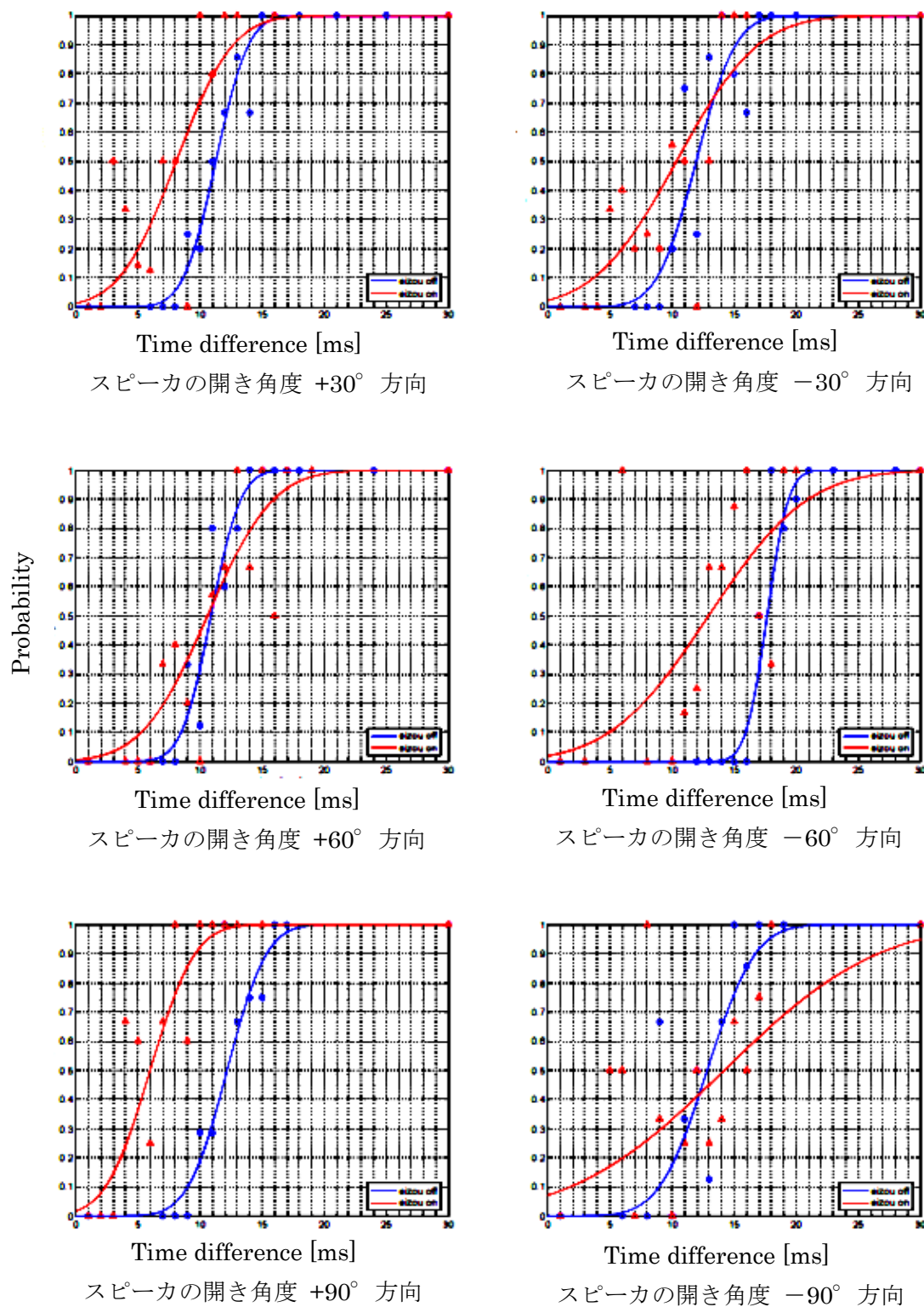


図 4.48 2 音源間の時間差とエコー知覚確率（映像：有/無）

被験者：S，青：映像無し OFF，赤：映像有り ON），刺激音：スピーチ（女声）

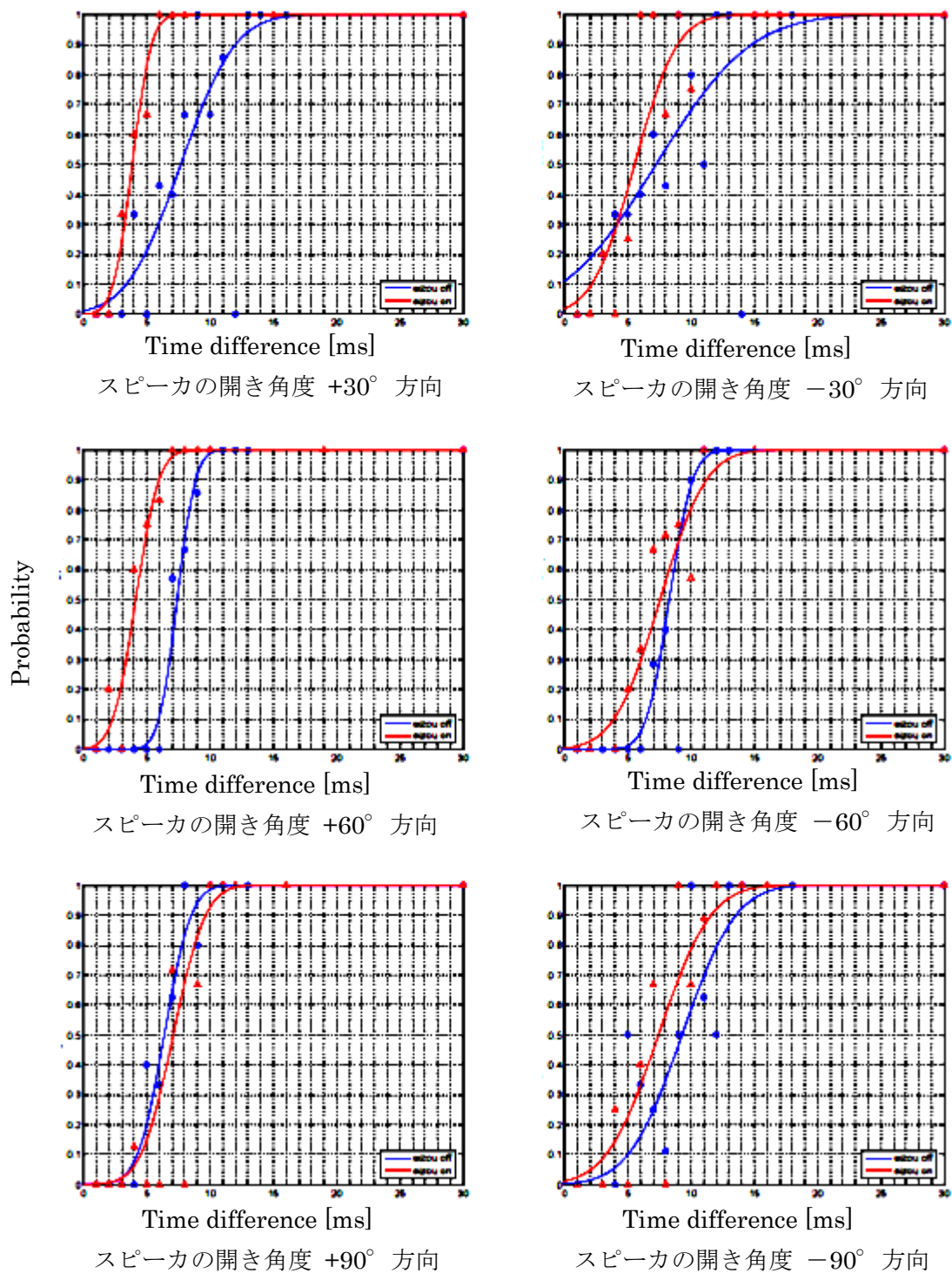


図 4.49 2 音源間の時間差とエコー知覚確率（映像：有/無）

被験者：S. 青：映像無し OFF. 赤：映像有り ON. 刺激音：小太鼓

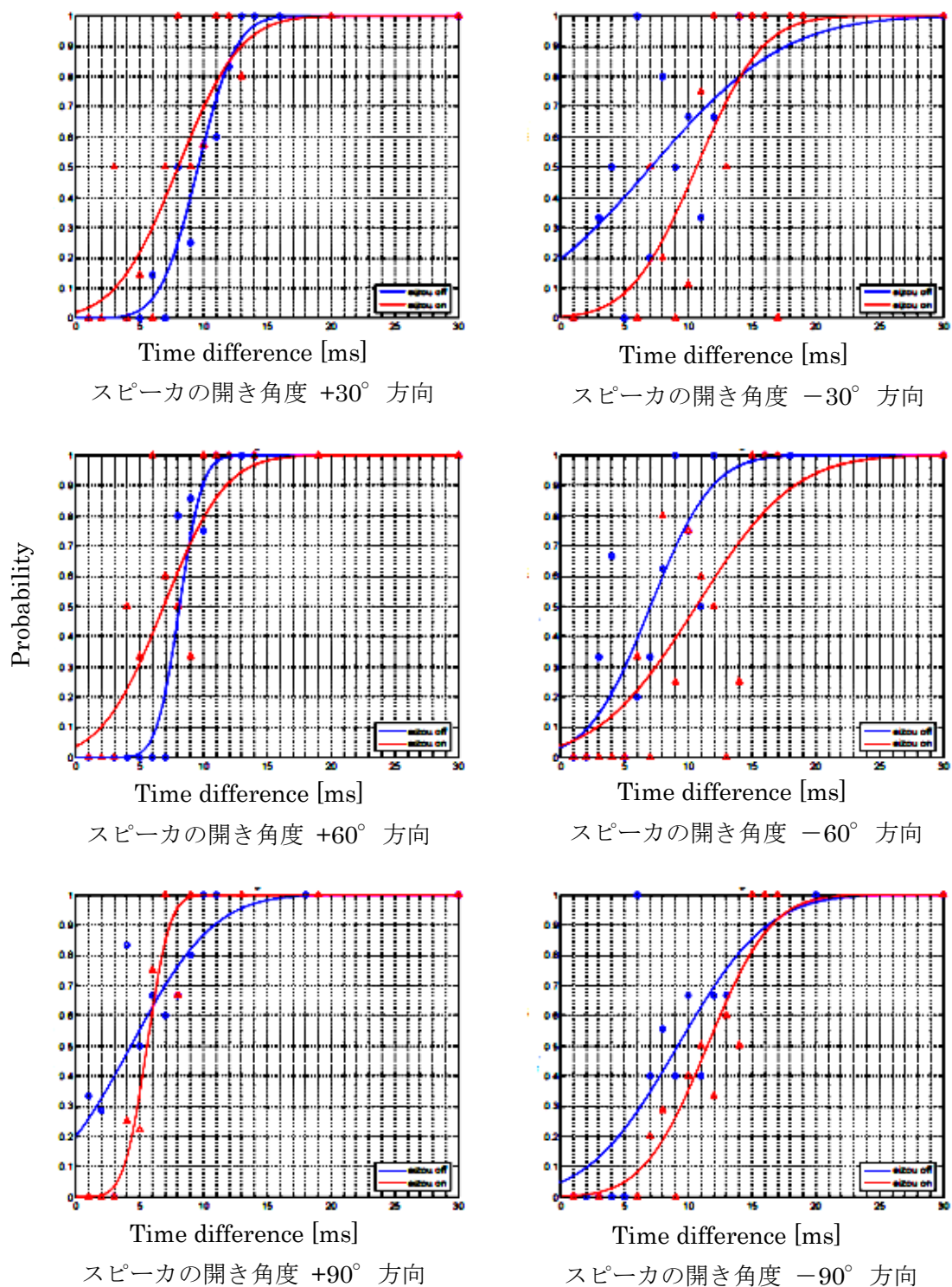


図 4.50 2 音源間の時間差とエコー知覚確率（映像：有/無）

被験者：S，青：映像無し OFF，赤：映像有り ON），刺激音：ギター

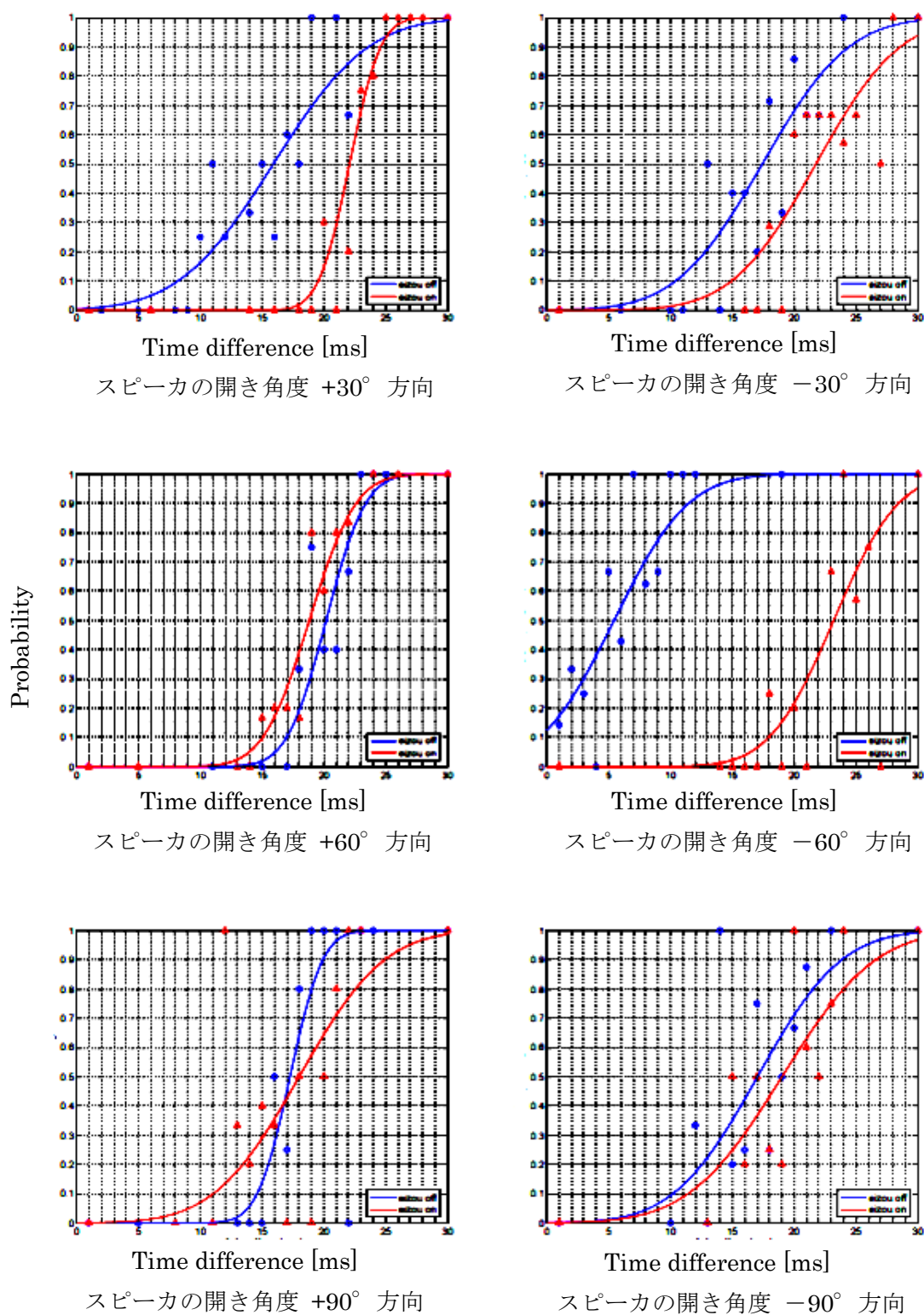


図 4.51 2 音源間の時間差とエコー知覚確率（映像：有/無）

被験者：SY，青：映像無し OFF，赤：映像有り ON），刺激音：スピーチ（男声）

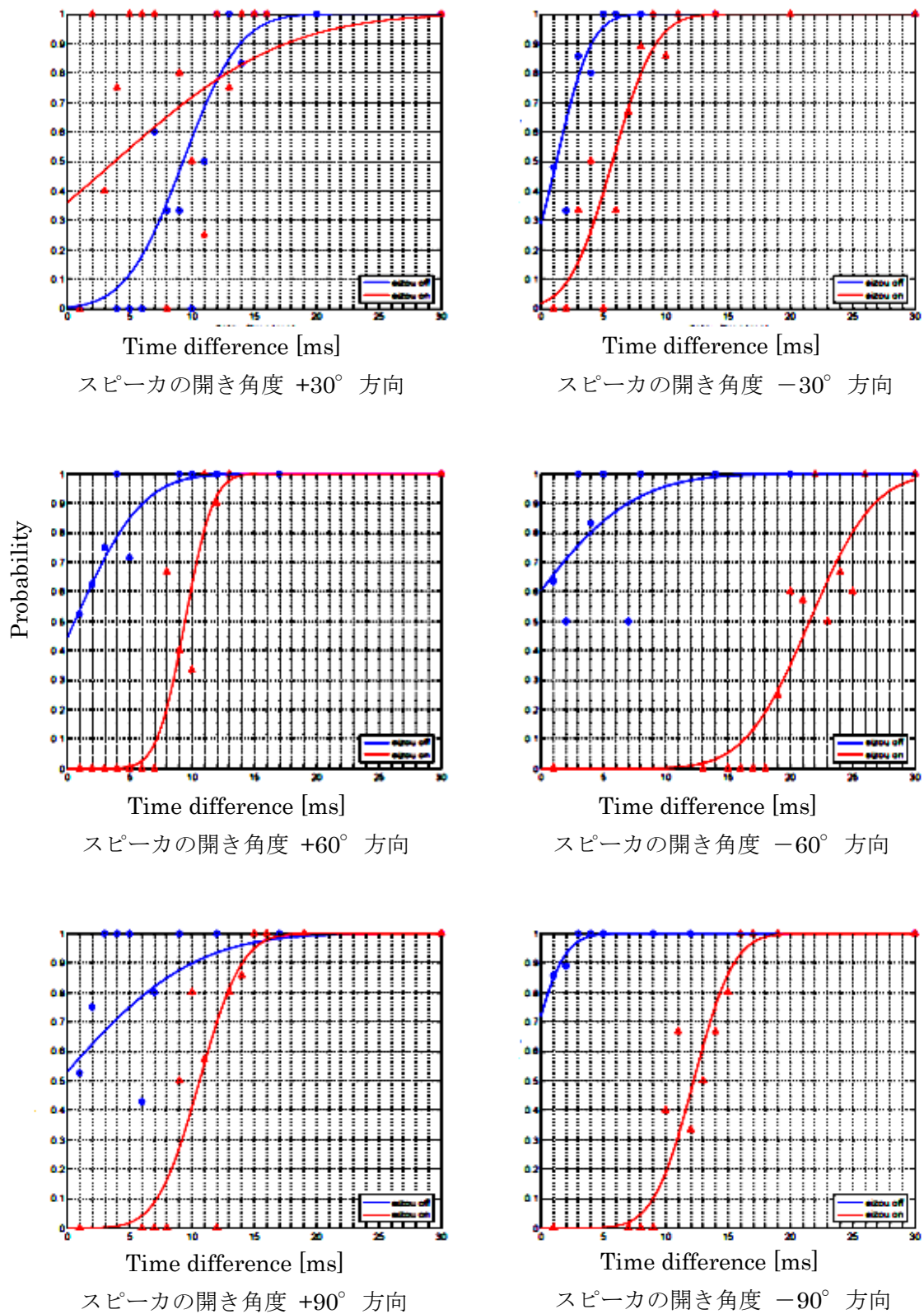


図 4.52 2 音源間の時間差とエコー知覚確率（映像：有/無）

被験者：SY，青：映像無し OFF，赤：映像有り ON），刺激音：スピーチ（女声）

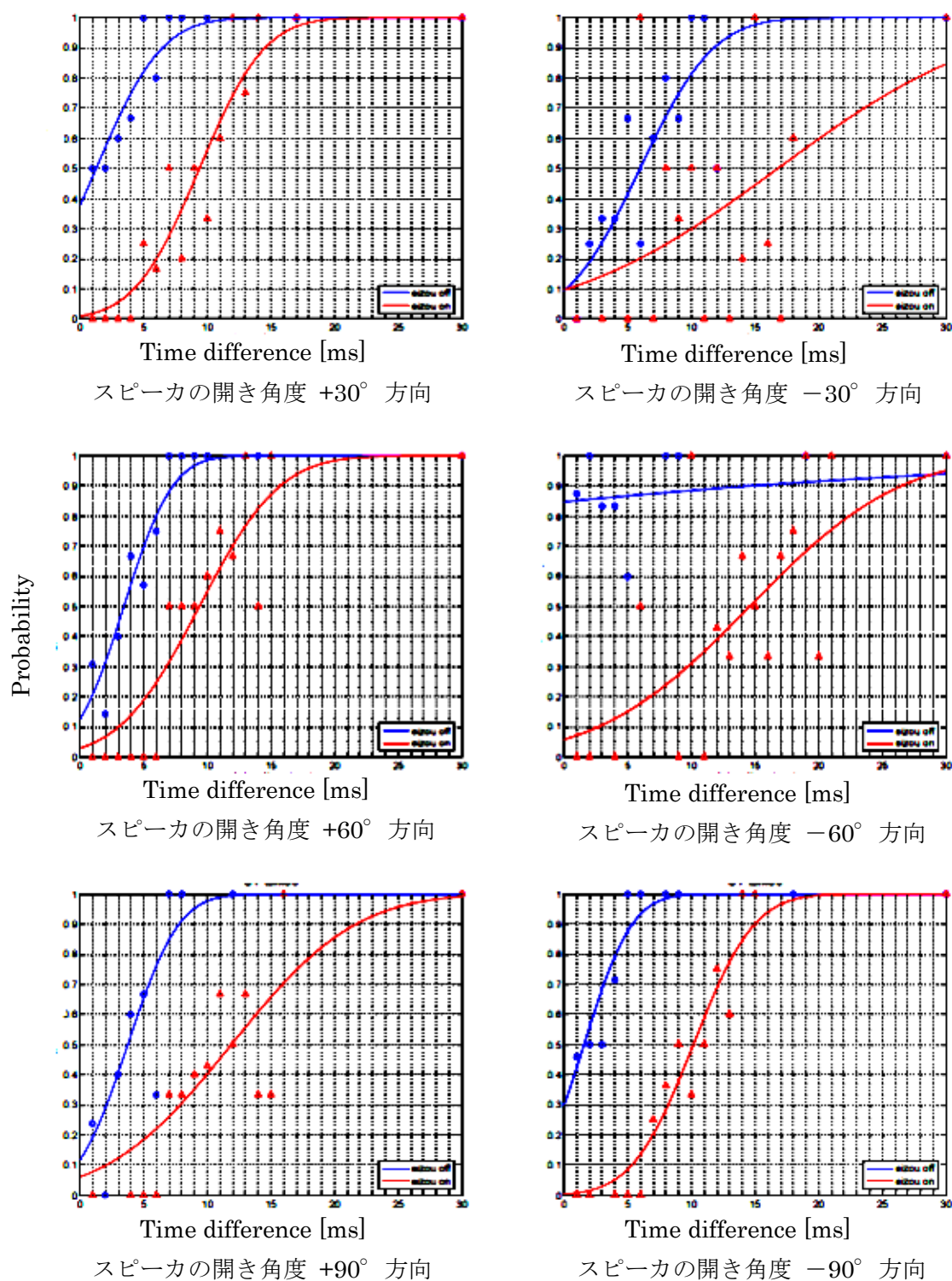


図 4.53 2 音源間の時間差とエコー知覚確率（映像：有/無）

被験者：SY. 青：映像無し OFF. 赤：映像有り ON). 刺激音：小太鼓

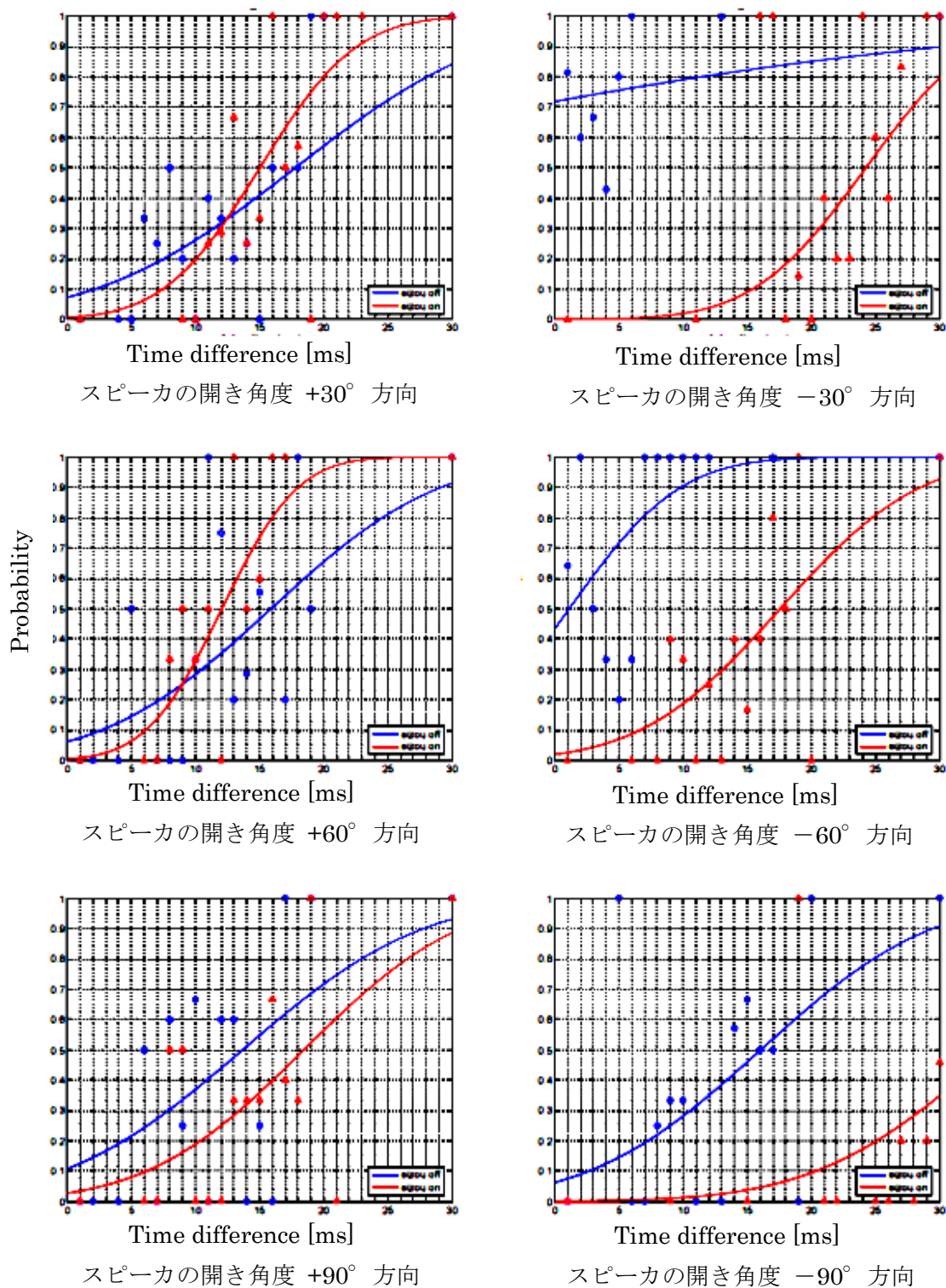


図 4.54 2 音源間の時間差とエコー知覚確率（映像：有/無）

被験者：SY，青：映像無し OFF，赤：映像有り ON），刺激音：ギター

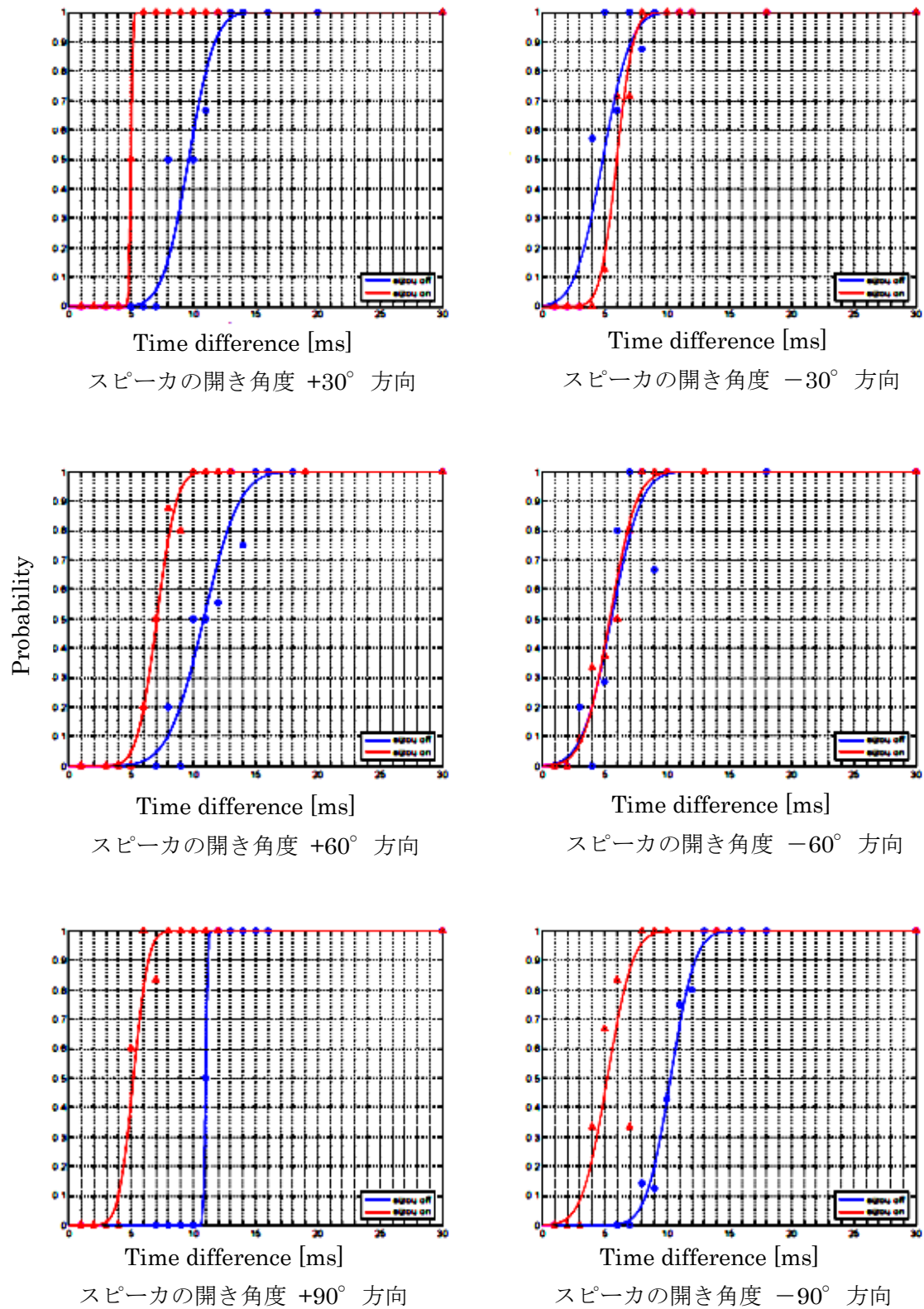


図 4.55 2 音源間の時間差とエコー知覚確率（映像：有/無）

被験者：W，青：映像無し OFF，赤：映像有り ON），刺激音：スピーチ（男声）

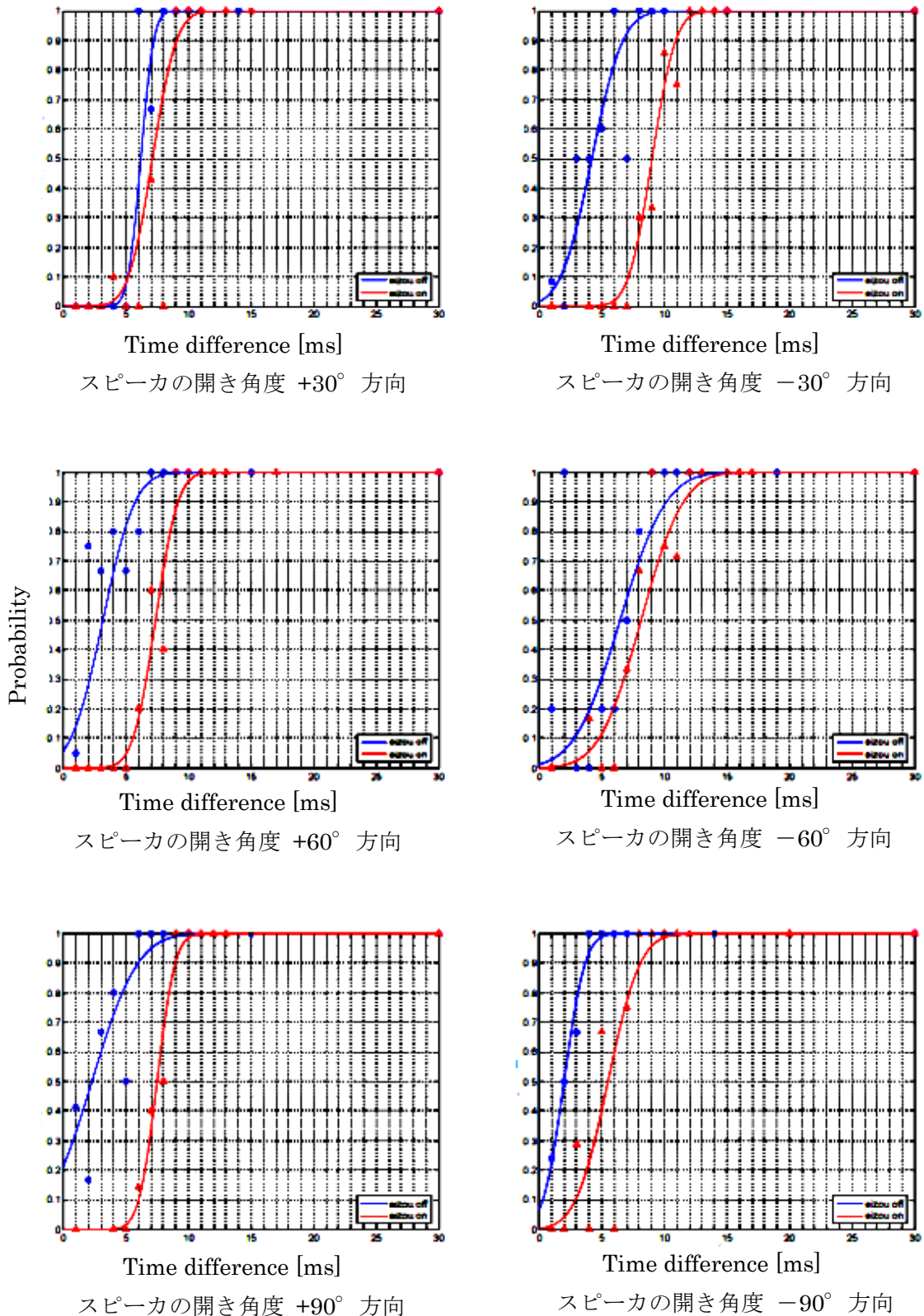


図 4.56 2 音源間の時間差とエコー知覚確率 (映像：有/無)

被験者：W，青：映像無し(OFF)，赤：映像有り(ON)，刺激音：スピーチ(女声)

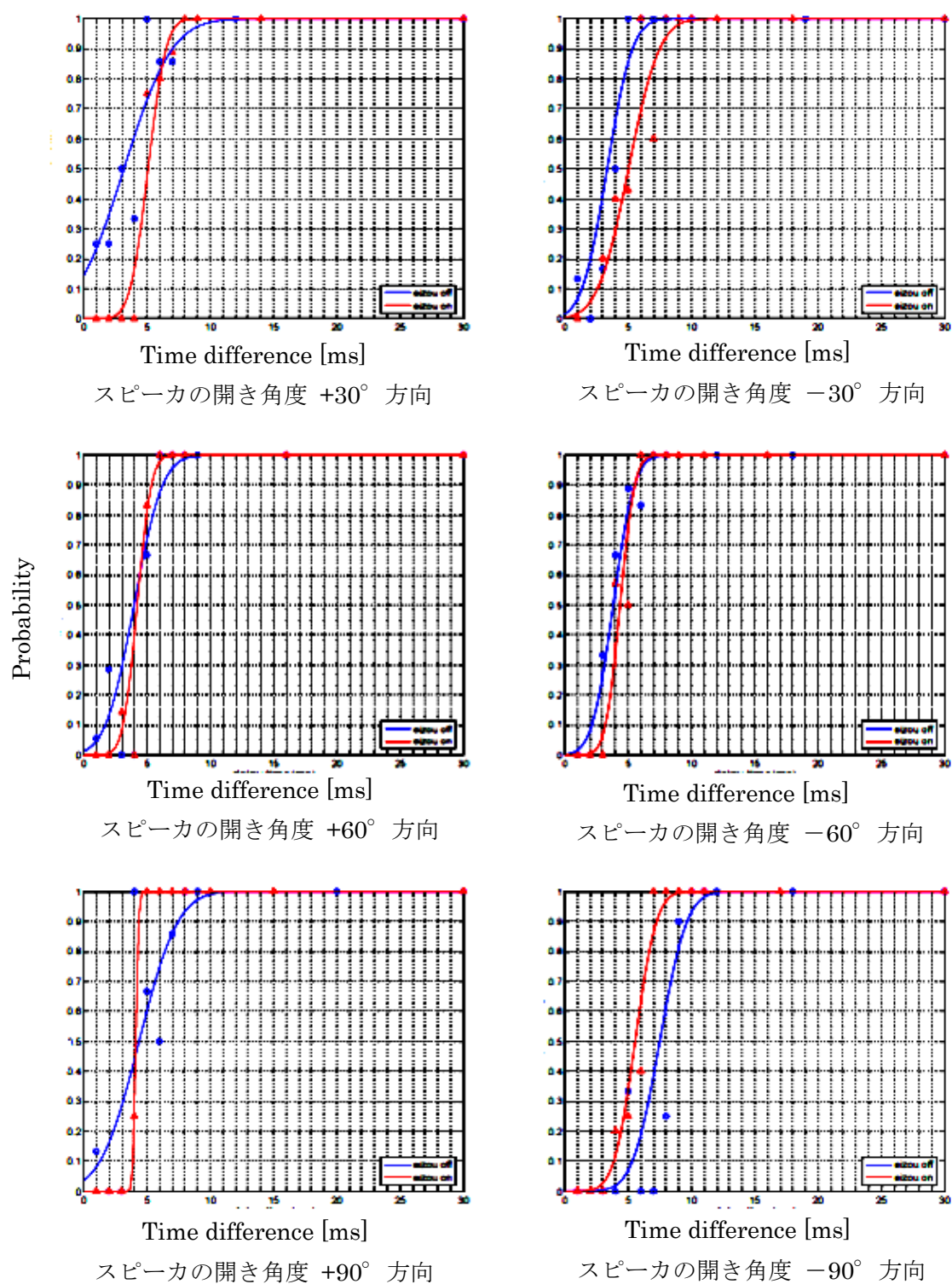


図 4.57 2 音源間の時間差とエコー知覚確率（映像：有/無）

被験者：W，青：映像無し(OFF)，赤：映像有り(ON)，刺激音：小太鼓

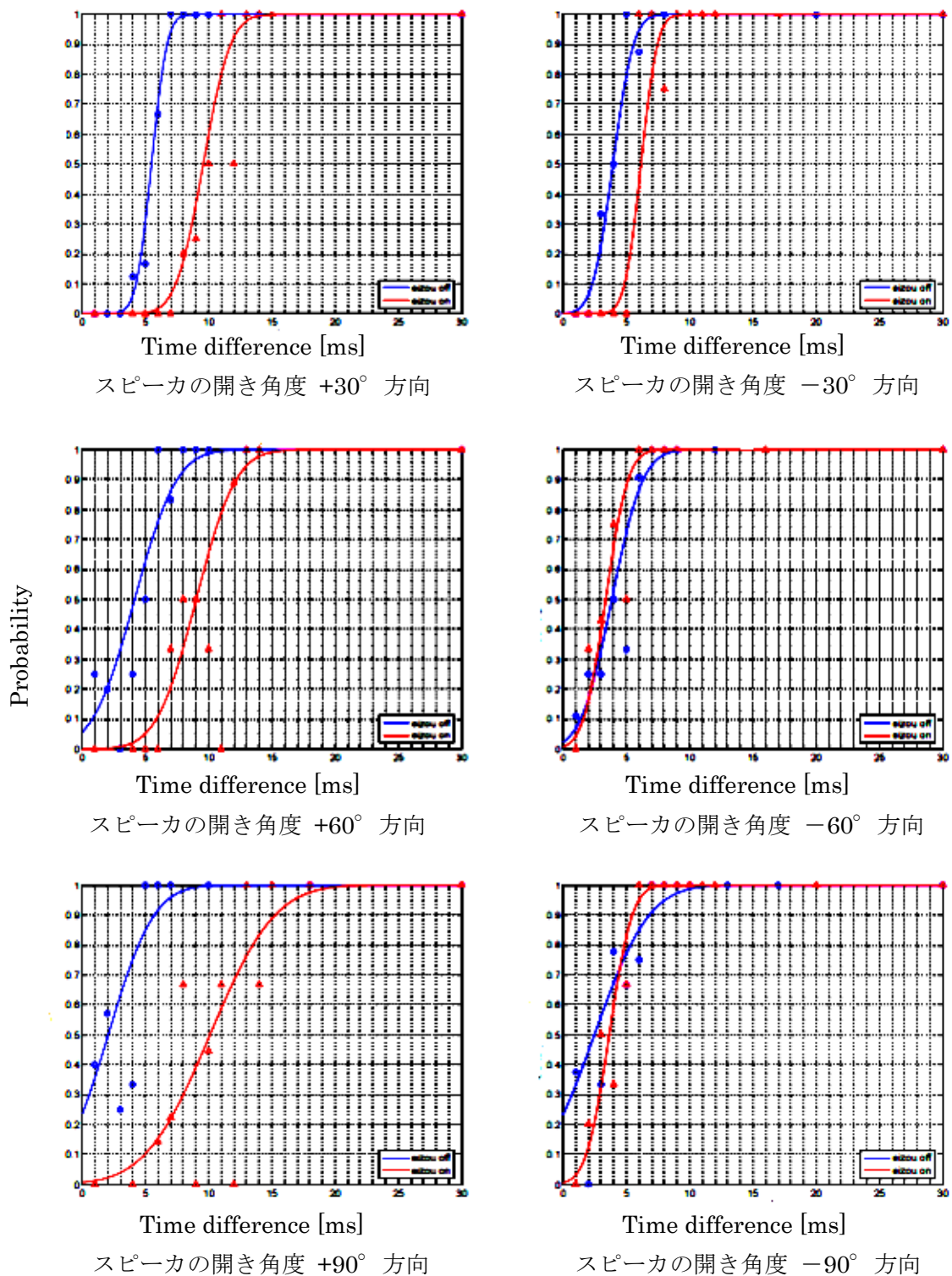


図 4.58 2 音源間の時間差とエコー知覚確率（映像：有/無）

被験者：W，青：映像無し(OFF)，赤：映像有り(ON)，刺激音：ギター

4.6 まとめ

音像定位における先行音効果に対する視覚刺激の影響について検討した。実験は、先行音に時間遅れを伴う後続音が加わった場合、先行音効果領域の下限 SLT と先行音効果領域上限の ET に対する視覚刺激の影響について実験を行った。その結果、次のことが分った。

- 1) SLTに対する視覚刺激の影響は、音源配置が $\pm 30^\circ$ で刺激音がスピーチ・小太鼓・ギターの場合に、視覚刺激の影響が現れ、SLTが短くなる傾向が見られた(0.2 ms 程度)。音源配置が $\pm 60^\circ$, $\pm 90^\circ$ では、視覚刺激の影響が若干見られるものの、その影響は小さい。
- 2) ETに対しては、ギター音の場合、視覚刺激の付加がET付近を有意に長くした(4 ms 程度)。スピーチ、小太鼓で有意な影響は見られたが、視覚刺激の付加がET付近を短くしたり長くしたり、被験者によって傾向が異なることが示された。
- 3) SLTは、視聴覚刺激の違いと動的な映像に強く引っ張られる傾向があり、有意に変化した。ETは、有意に変化するが、被験者ごとの構えによる、聴覚と視覚に対する注意の向け方などによって、一様ではないものと考えられた。

ここで明らかになったことは、聴覚刺激と視覚刺激が結びついた感覚間の空間的・時間的な一致と被験者の心理的な注意の要因(聴覚と視覚に対する注意など)が重要な役割を担っていると考えられることである。映像付加の効果は、刺激(音と映像の動的性質)の違いによって、先行音効果領域に概ね有意な差をもたらしたが、その程度や方向は、被験者によって異なり、一様ではない様相を示すなど、今後さらに検討が必要であると考えられる。しかし、ここで得られた聴覚刺激に対する視覚刺激の影響は、被験者ごとの傾向を見たものであり、特にSLTとETは定量的に求められた値で、室内空間における臨場感などを考える上で強固な基礎資料となり得るものと期待される。

第5章

結 論

本論文では、2 音源により生じる音像と先行音効果の関係を明確にするため、音像の方向と先行音効果の生起との関係、時間差・レベル差・音源の種類・音源の位置が先行音効果に及ぼす影響、視覚刺激が先行音効果の境界 SLT と ET にどのような影響を及ぼすかについて現実的側面を重視し、聴取実験を通して検討した。本論文の結果は、以下の通りである。

第 1 章では、研究の背景を明らかにするために、人が音の方向を知覚する要因を概観し、その上で、室内空間における先行音効果の聴覚事象について述べ、先行音効果の発生する聴覚の時間的な要因を考察した。更に、音の方向定位と先行音効果に関する従来の研究について言及しながら、先行音効果の基礎的な振る舞いを概観し、先行音効果が室内空間での音の聴取にどのように関わっているかを考察した。また、それらの知見を基に、先行研究者が行っていない事象について述べ、音像定位における先行音効果の関わりと問題点を整理し、本論文の目的を述べた。

第 2 章では、本論文の中心となる先行音効果の生起と知覚される音像の関係について聴取実験を通して検討した。それは、2 音源間によって生じる合成音像が、それらの音源の時間差やレベル差、配置によってどのように変化するのか、加法定位が生じると考えられる時間差から、先行して音が出力される音源側に移るまでの比較的短い時間差(定位方向が先行音側に移るときの境界を **SLT : Summing localization threshold** と名付けた)と時間差を更に増すと先行する音源側に片寄ってしまい、先行音効果の現象が生じ、更に時間差を増すとエコーになる。この音像の変化がどのよう

に起こるのかについて調べた。その実験方法として、2 音源から提示した時間差やレベル差によって音像の定位方向がどの方向にあるかを直接回答させる絶対方向判断の実験と 2 音源がごく短い時間差の場合、音像は 2 音源の中間に知覚される。これを音源とみなし、ダミースピーカを置き視認的に見えるように 3 つの音源として配置し、この 3 つの音源に知覚された音像がどれに最も近いかを強制的に選択させる 3 種強制選択実験で、共通な実験系と被験者で実施した。絶対方向判断の実験結果として、音源間のレベル差が 0, +3, +6 dB のいずれの場合においても、2 音源の時間差の変化にともなって、定位方向の平均値は 2 音源間を直線的に変化し、過去の研究の知見と対応が見られた。この傾向は、音源配置によらず共通であった。ただし、刺激音がクリックの場合は、この実験の中で絶対値の大きな時間差となる場合でも、音源方向には寄りきらず、その傾向は音源配置 60° の場合で顕著となった。また、定位方向が正面から離れるにしたがい、そのばらつきも大きくなっている傾向が見られた。3 肢強制選択の実験結果から、先行して音が出力された音源方向に定位していると見なせる 2 音源間の時間差の境界 SLT を、各音源配置について算出した。その時間差は、算出された条件で比べた限りでは、刺激音による違いは見出せなかった。また、その時間差の大きさ(絶対値)は、音源間のレベル差が 0 dB の場合で $-0.4 \sim +0.6$ ms となり、過去の研究結果と同等の値であり、レベル差が +3 dB では $-0.6 \sim +0.8$ ms で、+6 dB では、 $-0.8 \sim +1.5$ ms となり、レベル差の違いによる差異は若干見られたがほぼ近い対応が見られた。レベル差が大きくなると、その時間差の絶対値も大きくなった。さらに、音源配置が正面 60° のとき、正面方向の音源については、比較的小さな時間差があれば定位方向が移動するのに対し、側方 60° の音源については、大きな時間差が与えられないと先行音側に定位方向は移動しないことが分かった。また、刺激音がピンクノイズでレベル差が +6 dB の場合、その時間差の絶対値が算出できない条件があった。その原因は、その時間差が実験を行った時間差の範囲外に出てしまったと考え

られた。両実験の結果の比較から、3 肢強制選択の実験で算出された境界の時間差における音源の定位方向は、配置された音源方向から 10° 程度内側に寄ったところであった。このことは、先行して音が出力された音源に定位したとみなせる定位方向の範囲が、音源の大きさに比べて広いことを示している。それは、実際に設置した音源の距離 1.5 m では、 $1.5 \times 2 \times \tan(\pi/18) \approx 0.53$ [m]の幅に相当し、使用された音源に対して数倍程度の広さをもつとみなせる。この数値は、拡声においても、先行音効果を利用した定位方向の制御を考えたときに、有益な指標となり得る。ただし、本章の実験結果は、限られた音源配置における広帯域音による合成音像の知覚を調べたものであるため、実験結果のばらつきも大きいし、各個人のばらつきも大きくなってしまった。先行研究者の Litovsky *et al.* は、先行音効果を定量化する 3 つの指標、“Fusion,” “Localization dominance,” “Discrimination suppression”の傾向を共通の 6 名の被験者で比較する実験を行っており、大きな個人差のあることを報告している。ただし、視認的に 3 つの音源を配置した実験で算出された境界の時間差は、先行して音の出た音源の方向に、ほぼ確実に定位が移っているとみなせることが分かった。このことにより、2 音源による音像は音源間の時間差が増加するにしたがい、確実に 2 音源の間方向から先行音源方向へ移動し、先行音源方向への音像の収束は音源配置に若干影響を受けることが分かった。3 肢強制選択実験により SLT が求まり、絶対方向判断実験との比較で、先行音効果が生じ始める音像の範囲が求められた。それは 2 音源の方向から開き角度の 4 分の 1 内側が先行音効果と加法定位を分ける音像方向の境界付近であることが明らかとなった。これらの結果から被験者全体の傾向をまとめた本章の結果は十分な価値をもつものであると考えられる。

第 3 章では、第 2 章を踏まえて、刺激音の種類とレベル差が先行音効果に及ぼす影響について、吟味を加える意味から聴取実験を行なっている。前章でも述べたが、日常生活において、我々が聞く音のほとんどが、時間的に変動し、独自のパターンと音色を持

つ、これらは、音の強さ、周波数成分並びにその時間的变化などの、音を構成する上で重要な役割を果たし、音の認知と識別や方向定位においても、これらは重要な手掛かりとなる。このようなことから、室内で音が発せられた場合、直接音に続くいくつかの反射音を伴って聞こえる。そこでは、最初に到達した直接音の方向情報が優位性を持ち、その音の種類の違い(周期的な時間的配列の速さや音の出現速度)などと音のレベル差の違い(音の到達経路で生じるレベル差)などが方向定位にどのように関わっているのか明らかにする必要があった。さらに、室内において残響時間が極端に長い場合や直接音と初期反射音の時間間隔がかなり長いなどのような、拡声設備が適切に設定されていない場合では、音の方向が正しく認識できなくなることも考えられる。この場合、室の音響特性を考慮して、聴覚現象の先行音効果を有効に利用できるように音響装置を設計することができれば、音の方向は違和感なく把握できるはずである。このような設計を行うためには、どのような条件下で先行音効果が生じるのか、また、どのような範囲で先行音効果が生じているのか明らかにする必要があった。

本章では、刺激音の種類が先行音効果に与える影響についての聴取実験と音圧レベル差が先行音効果に与える影響について、聴取実験を行なった。刺激音の種類の実験では、刺激音の継続時間が、臨界継続時間を超える刺激音を用い、時間の長い音声(スピーチ)で、ある日の天気概況を朗読したものを用い、継続時間を約 15 秒間とした。また、楽器音として、「サキソフォンの演奏音」(継続時間約 12 秒間)と「ギター演奏音」(継続時間約 9 秒間)、さらに衝撃的な音として、「ピンクノイズのトーンバースト音」(on 200 ms-off 500 ms の 10 回 繰返し、継続時間約 6.5 秒間)の 4 種類の日常生活でよく聞く刺激音を用いている。その結果、同じ継続時間を持つ連続した刺激音(トーンバースト音)では、先行音効果が見られる遅延時間の範囲は狭く、変動の大きい音(例えば、スピーチや楽器音)に対しては、先行音効果の成立範囲は広くなる傾向が見られた。それは、スピーチや楽器音のような時間的系列のテンポ間隔が短く、周波数

帯域が狭い刺激音では、先行音効果の成立する範囲は広くなる傾向があり、ピンクノイズバースト音のような、テンポ間隔が長く、周波数帯域が広い場合には、先行音効果の生じる範囲は狭くなる。特に、時間的系列のテンポ間隔の違いが先行音効果の上限(エコー検知限 **ET**)に大きい影響を与えていることが分った。本章での結果を刺激音の種類の影響についてまとめると、刺激音の種類による SLT への影響は比較的小さい。刺激音の種類による ET への影響は比較的大きい。先行音効果の成立する条件には、刺激音の到来方向に対する依存性はあまり見られない。

先行音効果に及ぼす 2 音源の提示方向とレベル差の影響については、聴衆者の状況を考えて、直接音(第 1 音)より客席に近いスピーカからの第 2 音を優位に聴くことになり、第 2 音が第 1 音より高いレベルになる可能性もある。そこで、第 2 音のレベルが、第 1 音より 0, 3, 6, 9 dB と高くなった場合、先行音効果の現象にどのような影響が現れるのかを検討した。刺激音は、Pink noise burst (on 200 ms～off 500 ms)の一種類とし、2 音源の時間差を被験者の正面を基準として、正面以外の音の遅延時間差を 0～50 ms まで 40 段階と設定した。レベルは、基準レベルを 60 dB とし、その他のスピーカのレベルを 60, 63, 66, 69 dB のいずれかとし、先行音効果の生起にどのように影響するかを調べた。その結果、刺激音の種類と音源の配置には、関係なく、ほぼ同様な傾向が見られた。それは、レベル差が +6 dB 以内ならば、遅延時間の影響が優勢となり、音源配置 $0^{\circ} \sim \pm 60^{\circ}$ と $0^{\circ} \sim \pm 120^{\circ}$ 方向の平均値で、約 1.0 ms 付近に加法定位と先行音側に変移する境界 SLT があり、約 1.0～7.4 ms の時間差で先行音効果の生じる範囲があるように見える。約 7.4 ms 付近にエコー検知限 **ET** が存在するが分かった。また、レベル差による先行音効果の成立する割合は、レベル差が等しい場合、88%で、+3 dB では 75%となり、+6 dB 高ければ 57%で、+9 dB は生じない、遅延音のレベルが増すにつれて、先行音効果の生じる割合が徐々に減少することが明らかになった。つまり、レベル差は、刺激音の種類や音源配置にそれほど影響が

ないが、+6 dB 以上では、レベル差が優位になってしまうことが分かった。Pinheiro and Tobin らは、刺激音がホワイトノイズやローパスノイズの場合、音像の幅やレベル差定位の弁別限が 8~10dB のレベル差であると報告している。本研究で求められた先行音効果の生じるレベル差が+6dB 以内であれば先行音効果の現象が発生することが明確になった。本章のレベル差の影響についてまとめると、2 音源間のレベル差は先行音効果の生起に影響を与える。レベル差が増加するにつれ先行音効果の生じる範囲は徐々に減少するが、レベル差が約 6 dB 以内であれば先行音効果は発生する。このことは従来の研究結果より若干短い時間差となったが、これは、刺激音の種類の違いと被験者が知覚した音像方向の判断基準の差によるものと思われ、本研究の結果とほぼ同様な傾向であると考えられた。ここで求めた結果は、刺激音の種類とレベル差の違いを先行音効果の下限の時間差の境界 SLT と上限のエコー検知限 ET との範囲を定量的に示すことができた。また、刺激音の種類によって ET に影響をもたらし、レベル差は刺激音の種類にはあまり依存しないことが明らかになった。

第 4 章では、聴覚刺激に視覚刺激が付加されたとき、聴覚現象の先行音効果にどのような影響があるかについて調べた。まず、本章では、加法定位から先行音効果に音像が変化する時間差の境界 SLT と、先行音効果からエコーに音像が変化する時間差の境界 ET において、視覚刺激を音と同期した映像で、発声者のほぼ等身大の動画像を前方スクリーンに投影したとき、先行音効果の現象にどのような影響があるのかを検討した。まず、2 音源の配置を $\pm 30^\circ$ 方向、刺激音は、スピーチ(男声、女声)、小太鼓、及びギターを用い、時間差の条件を 0.1~3.0 ms に設定して行った。その結果、すべての種類の刺激音で視覚刺激の影響が存在し、時間差の境界 SLT が短く知覚され、先行音側に音像が引き寄せられる傾向が見られた。この傾向は、映像との同期による被験者の集中度合いによって生じたのではないかと考えられる(約 0.2 ms 程度)。特に、SLT は刺激音の違いによるものではなく音源の到来方向に依存していると考えられ、腹話術効果に類

以したと捉えられる。次に、先行音効果からエコーに変化する時間差の境界 ET では、時間差の条件を 1~30 ms と設定して実施した。その結果、スピーチと小太鼓では、視覚刺激を付加したことにより、被験者によってエコー検知限 ET が早くなったり、遅くなったりする傾向があった、ギターでは、エコー検知限 ET が遅くなる傾向があった(約 5 ms 程度)。このように、スピーチ及び小太鼓で、被験者によって結果が異なることは、刺激音の波形に連続的な部分と途切れる部分が存在するため、視覚刺激によって個人の持つ音を識別する感覚を惑わして起こる現象ではないかと考えられる。このように、視聴覚刺激の種類によって先行音効果の生じる範囲が境界値 ET で差異が見られたことは、音の発生メカニズムや、これまでの経験から、時空間上での刺激音の特徴に基づいて、音の位置や量感等を勘案して判断しているのではないかと思われた。また、聴覚と視覚の感覚統合によって音像定位の対応が困難な場合には、視聴覚刺激の間に矛盾の無い、時間分解能に頼って優位な感覚が、音像の方向を決定しているのではないかと考えられた。

以上、第 2 章から第 4 章まで得られた結果をまとめると、

本研究では、加法定位と先行音効果の生起を分ける時間的境界を明確とするために SLT を定義し、加法定位から先行音効果の変移を明確にした。音像は 2 音源間の時間差の増加に伴い、徐々に先行音側へ移動する。音源配置でスピーカの開き角度の 4 分の 1 の近辺に SLT がある。また、SLT に対する影響は、音源配置などの影響を受ける。刺激音の種類による影響は小さい。2 音源間のレベル差による影響は小さい。視覚刺激の影響は存在し、短くなる傾向がある。更に、ET に対しする影響は、刺激音の種類により変化する。2 音源間のレベル差の増加に伴い徐々に短くなる。視覚刺激の影響は存在し、長くなる傾向がある。

本研究で得られた知見から、どのような利用や応用が可能なのかを挙げると、先行音効果を利用した音の避難誘導システムで、例えば、避難方向に複数のスピーカを配置し、

その時間差，レベル差で音像を移動させ，音の方向指示で避難誘導システムの構築を考える上で基礎資料となり得ることや PA システムへの応用，例えば，広いホールや会議室では，不特定の話者が各位置から発言する，その発言者の方向から音が聞こえてくる方向に音像の制御が可能な PA システムの構築のための基礎資料となり得る．更に，音像の実時間制御などの工学的な応用などで，音像制御の分野や聴覚刺激と視覚刺激が同期して得られる感覚間についての相互作用に関する研究の手掛かりとなり得るものであると期待される．

しかし，まだ，残された課題もある．特に，第 4 章の議論で不足していることは，先行音効果に関わる諸現象（視聴覚刺激の種類・音源配置・場所の環境など）に対する心理的な側面での検討である．例えば，視覚と聴覚の機能の間でどちらが優位性を持つかということは，はっきりとは言えず，恐らく，その場の実空間の環境によっても異なるものと推察され，情緒的な側面に対する時変的な事象を含めた心理学的なアプローチも必要であると考えられる．このように，今後の課題も残された．

最後に，著者は長い間，聴覚の知覚現象である音像定位について関心を寄せてきた．特に，先行音効果の現象であるが，これは，人の感覚によるもので，物理量では測定できない，いわゆる音の錯覚現象でもある．このような知覚された音像を見確認のファントムソースと呼んだこともある．また，それは，人間の持つ聴覚特性の 1 つでもあり，聴取実験を通してしか明らかにできない，人間主体の特性であると理解していたところでもある．このことも著者には興味を寄せるところでもあった．

ようやく，本論文をまとめることができた．これも多くの方々のご支援とご協力のおかげである．あらためてここに深く感謝申し上げる．

謝 辞

東北大学名誉教授で秋田県立大学名誉教授の曾根敏夫先生には、秋田県立大学大学院システム科学技術研究科後期課程の社会人学生として本研究に従事する機会を賜り、終始並々ならぬ御指導を頂いた。また、日頃の研究のみならず、研究を離れた場においても、研究に対する姿勢、見聞を広めることの重要性、教養の糧となる知識を与えて頂いた。ここに心から感謝申し上げる。

秋田県立大学システム科学技術学部佐藤宗純教授には、院ゼミなどの場を通じて、本研究に対する有益な助言を頂いた。また、秋田県立大学システム科学技術学部高根昭一准教授には、先行音効果に関する知見を得る機会を与えて頂き、本研究を進める上で数多く参考にさせて頂いた。ここに心から感謝申し上げる。

東北大学電気通信研究所鈴木陽一教授には、音像定位と聴覚機構に関する知見を頂き、研究を進める上で多くのご意見を賜った。ここに心から感謝申し上げる。

秋田県立大学システム科学技術学部安倍幸治博士、長南安紀博士、渡邊貫治博士の方々には、院ゼミを通して本研究を行う上での実験系構築・設定やデータの分析などに関して多くの有益な助言を頂いた。感謝申し上げる。

山梨大学小澤賢司教授には、マルチモーダル・コミュニケーションに携わる糸口を与えていただき、関連資料を提供して頂いた。また、東北大学電気通信研究所齋藤文孝技官には、聴覚実験に関する測定方法など多くの有益なご意見を頂いた。更に、秋田県立大学システム科学技術学部陳国躍教授には、研究を進める上で有益な助言を頂き、感謝申し上げます。

東北大学名誉教授で東北工業大学名誉教授の齊藤伸自先生、元東北工業大学通信工学科松谷栄一教授、渡辺富也教授、元東北工業大学情報通信工学科高野剛浩教授には、直

接上司としてご指導を頂いた。また、元東北工業大学(故)橋本廣一 助教授、元東北工業大学情報通信工学科水野尚教授には、秋田県立大学大学院入学に際して、推薦の労をおとり頂いた。

秋田高専 大島静夫教授には、社会人博士の同期入学生として院ゼミと研究などで共に学び、討論し、励ましあい、共に国際学会で発表するなど、折に触れてさまざまな配慮をして頂いた。

元東北工業大学情報通信工学科山田 弘技師には、同僚として、研究を進める上で協力して頂いた。

秋田県立大学平成 16 年卒業生佐々木太氏には、本研究の被験者となり協力して頂いた。また、秋田県立大学大学院生砂田大輔氏、伊藤伸也氏、志賀健太氏には、実験装置の設計・製作など多大の協力を頂き、聴取実験では、被験者となり長時間実験に付き合ってくれた。また、本研究の被験者となって頂いた佐藤研究室の学生の皆様に協力して頂き 感謝する。

本研究は、このように多くの方々のご指導とご援助があつて、ようやくまとめることができた。ここに重ねて感謝申し上げる。

参考文献

- 1) 難波精一郎, 桑野園子, 江端正直, 泉清人, 宮坂栄一, 森本政之, 寺西立年, “聴覚ハンドブック”, ナカニシヤ出版 (1984. 11. 10).
- 2) Jens Blauert, 森本政之, 後藤敏幸, “Spatial Hearing 空間音響”, 鹿島出版会 (1986. 7. 10).
- 3) Warneke, H., “Die Grundlagen der raumbeztiglichen stereophonischen Übertragung im Tonfilm” [The fundamentals of room-related stereophonic reproduction in sound films], Akust. Z. **6.**, 174-188 (1941).
- 4) Wallach, H., Newman EB, Rosenzweig MR., “The precedence effect in sound localization”, Am J Psychol., 62(3), 315-336 (1949).
- 5) H. Haas, “On the influence of a single echo on the intelligibility of speech,” Acustica, **1**, 48-58(1951).
- 6) Cremer, L., “Die wissenschaftlichen Grundlagen der Raumakustik” [The scientific foundations of architectural acoustics], S. Hirzel Verlag, Stuttgart., Vol. 1, (1948)
- 7) 永井健三, 佐藤利三郎, 岩崎俊一 “立体録音に於ける方向感の実験的検討” 日本音響学会誌 **10**, 149-155, (1954)
- 8) De Boer, K., “A remarkable phenomenon with stereophonic sound reproduction” Philips Tech, Rev. **9**, 8-13, also Frequenz 3, 24-25 (1949).
- 9) Wendt, K., “Das Richtungshören bei der Überlagerung zweier Schallfelder bei Intensitäts- und Laufzeitstereophonie” [Directional hearing with two superimposed sound fields in intensity-and delay-difference stereophony], Dissertation, Technische Hochschule, Aachen. (1963).

- 10) D. Queen, “The effect of loudspeaker radiation patterns on stereo imaging and clarity,” J. Audio Eng. Soc., **27**(5), 368-379 (1979).
- 11) Leakey, D.M., B.M. Sayers and C. Cherry, “Binaural fusion of low and high-frequency sounds” , J. Acoust. Soc. Amer., **30**, 322 (1958).
- 12) J. Braasch, J. Blauert and T. Tjelani, “The precedence effect for noise bursts of different bandwidths. I. Psychoacoustical data” , Acoust. Sci. & Tech., **24**(5), 233-241 (2003).
- 13) MacGurk, H. and MacDonald, J., “Hearing lips and seeing voices, ” Nature, 264, 746-748 (1976).
- 14) 森本政之, 石井正光, 前田純一, “エコー検知限に及ぼす視覚刺激の影響” 日本音響学会誌 46 巻 3 号 (1990).
- 15) V.J Bolivar, A.J Cohen and J.C. Fentress, “Semantic and formal congruency in music and motion pictures : Effects on the interpretation of visual action,” Psychomusicology **13**, 28-59 (1994).
- 16) S.D. Lipscomb and R.A. Kendall, “Perceptual judgement of the relationship between musical and visual components in film” , Psychomusicology **13**, 60-98 (1994).
- 17) Grimes, T., “Audio-visual correspondence and its role in attention and memory” , Educational technology and development, **38**, 15-25 (1990)
- 18) チャールズ スペンス, 全訳 : 五十嵐由夏, 北村智利 : 解説, “視聴覚総合” , 日本音響学会誌, 63巻, 2号, 83-92 (2007).
- 19) 岩宮眞一郎, “オーディオ・ヴィジュアル・メディアを通しての情報伝達における視覚と聴覚の相互作用に関する音と映像の調和の影響” , 日本音響学会誌, **48**, 649-657 (1992).

- 20) 岩宮眞一郎, “オーディオ・ヴィジュアル・メディアによる音楽聴取行動における視覚と聴覚の相互作用”, 日本音響学会誌, **48**, 3, 146-153 (1992).
- 21) 岩宮眞一郎, “視覚と聴覚の相互作用に及ぼす音響再生系の音質の影響 —オーディオ信号に帯域制限を加えた場合—”, JAB. Journal, **33**, 29-35 (1993).
- 22) 大串健吾, 中山 剛, 福田忠彦, “画質と音質の評価技術”, テレビジョン学会(1991).
- 23) Cherry, E. C., W. K. Taylor, “Some further experiments upon recognition of speech with one and with two ears”, J. Acoust. Soc. Amer., **26**, 554-559 (1954).
- 24) G. Steinke, “Delta Stereophony - A Sound System with True Direction and Distance Perception for Multipurpose Halls”, J. Audio Eng. Soc., **31**(7), 500-511 (1983).
- 25) T. Nakajima, T. Chono, K. Suzuki, Y. Yamada, “Development of a new directional sound reinforcement system and its utilization,” AIJ J. Technol. Des., No. 3, 171-175 (1996).
- 26) D. M. Leakey, “Some measurements on the effects of interchannel intensity and time differences in two channel sound systems”, J. Acoust. Soc. Am., **31**(7), 977-986 (1959).
- 27) R. Y. Litovsky, S. Colburn, W. A. Yost and S. J. Guzman, “The precedence effect,” J. Acoust. Soc. Am., **106**(4), Pt. 1, 1633-1654 (1999).
- 28) R. Y. Litovsky, B. Rakerd, T. C. T. Yin and W. M. Hartmann, “Psychophysical and physiological evidence for a precedence effect in the median sagittal plane,” J. Neurophysiol, **77**, 2223-2226 (1997).
- 29) 井上公子, 岩宮眞一郎他 “音響設計学入門 —音・音楽・テクノロジー” 九州芸術工科大学音響設計学科 発行所 (財)九州大学出版会 (2002).
- 30) M. Morimoto, M. Ishii and Z. Maekawa “Influence of a vision stimulus on

- echo-threshold,” J. Acoust. Soc. Jpn. (J), **46**, 229-235 (1990).
- 31) H. Wallach, E. B. Newman and M. R. Rosenzweig “The precedence effect in sound localization” Am. J. Psychol. **62(3)**, 315-336 (1949).
 - 32) D. M. Leakey and E. C. Cherry, “Influence of noise upon the equivalence of intensity differences and small time delays in two-loudspeakers systems,” J. Acoust. Soc. Am., **29**, 284-286 (1957).
 - 33) K. Abe, D. Sunada, S. Takane and S. Sato “Relationship between the summing-localization behavior and perceived width of sound image” , Acoust. Sci. & Tech., **31(4)**, 260-266 (2010).
 - 34) D. Queen “The effect of loudspeaker radiation patterns on stereo imaging and clarity” J. Audio Eng. Soc., **27(5)**, 368-379 (1979).
 - 35) E. Zwicker, G. Flottorp and S. S. Stevens, “Critical bandwidth in loudness summation” , J. Acoust. Soc. Am., **29**, 548-557 (1957).
 - 36) E. Port “Über die Lautstärke einzelner kurzer Schallimpulse” , Acustica, 13 (Heft 1), 212-223 (1963).
 - 37) H. Takeshima, Y. Suzuki, S. Kono, and T. Sone, “Growth of the loudness of a tone burst with a duration up to 10 seconds” , J. Acoust. Soc. Jpn. (E), 9, 6 (1988).
 - 38) Schubert, E. D., and J. Wernick, “Envelope versus microstructure in the fusion of dichotic signals” , J. Acoust. Soc. Amer., 45, 1525-1531 (1969).
 - 39) Kietz, H., “Das räumliche Hören [Spatial hearing]” , Acustica, 3, 73-89, (1953).
 - 40) 田原靖彦, 佐藤洋, 西谷修一, “エコーの時間的分離知覚限界に基づく聴覚積分特性の一検討” , 日本音響学会誌 61巻 1号14-23, 43.55.Hy (2005).
 - 41) Damaske, P., “Head related two-channel stereophony with loudspeaker

- reproduction”, J. Acoust. Soc. Am., **50**, 1109-1115 (1971b).
- 42) 砂田大輔, 播摩敏雄, 安倍幸治, 高根昭一, 佐藤宗純, “先行音効果に及ぼす視覚刺激の影響”, 聴覚研究資料 日本音響学会, 35(3), 161-166 (2005).
- 43) 丸山拓哉, “視感覚と聴感覚に現れる異系感性相互作用”, 心理学研究, 35, 204-216 (1964).
- 44) 丸山拓哉, “感覚間相互作用 聴感 第8章”, 東京大学出版会 (1969).
- 45) 盛永四郎, 野口薫, 和田陽平, 大山正, 今井省吾偏, “感覚+知覚心理学ハンドブック”, 誠信書房, 感性間の相互関連 (1969).
- 46) 安倍幸治, 小澤賢司, 鈴木陽一, 曾根敏夫, “視覚情報が環境音知覚に及ぼす影響についての考察”, 日本機械学会 第9回環境工学総合シンポジウム 講演論文集, 26-29 (1999).
- 47) 石宮眞一郎, “音楽と映像のマルチモーダル・コミュニケーション”, (財)九州大学出版会, 8-48 (2000).
- 48) 中林克己, “ステレオ音像とテレビ映像の相乗効果”, テレビジョン学会誌, 37, 984-991 (1983).
- 49) 二階堂誠也, “AとVの心理学[指向性からみた目と耳の役割]”, JASジャーナル10, 11, 12号 (1984).
- 50) 藤崎和香, “視聴覚の情報統合と同時性知覚”, 日本音響学会聴覚研究会資料, **35**, 9, H-2005, 93 (2006).
- 51) O. Klemm, “Lokalisation von Sinneseindrucken bei disparate Nebenreizen”, Psychol. Stud., **5**, 73-162 (1909).
- 52) C. E. Jack and W. R. Thurlow, “Effects of degree of visual association and angle of displacement on the “ventriloquism” effect”, Percept. Mot. Skills **37**, 967-979 (1973).

- 53) 小宮山 撰, “視覚と聴覚による音像知覚”, 「一音と映像の相互作用一」 日本音響学会誌 52巻1号, 46-50 (1996).

著者発表論文

- 1) Toshio Harima, Kenta Shiga, Koji Abe, Shouichi Takane, Kanji Watanabe and Sojun Sato, “Relation between perceived direction of a sound image and the behavior of the precedence effect” *Applied Acoustics*, 74, 10, p.1122-1135, (2013).
- 2) Toshio Harima, Koji Abe, Shouichi Takane, Sojun Sato and Toshio Sone, “Influence of visual stimulus on the precedence effect in sound localization.” *Acoust. Sci. & Tech.* Vol.30, No.4 p.240-248 (2009)
- 3) Toshio Harima, Koji Abe, Shouichi Takane and Toshio Sone. “A study on the limitation of precedence effect and echo threshold in sound localization” Twelfth International Congress on Sound and Vibration, ICSV 12 , p.1-8 (2005)
- 4) Toshio Harima, Hisashi Mizuno, Koji Abe, Shouichi Takane, and Toshio Sone, “An experimental study on the precedence effect using subsequent sound sources surrounding subject” 東北工業大学新技術創造研究センター紀要 EOS Vol.19 No.1, p.21-29 (2006)
- 5) 播摩敏雄, 安倍幸治, 高根昭一, 曾根敏夫. “先行音効果に及ぼす視覚情報の影響” 日本音響学会 騒音・振動研究会 N-2007-47, p.1-10, (2007.11.30)
- 6) 播摩敏雄, 山田弘, 水野尚, 曾根敏夫, “先行音効果に及ぼす視覚情報の影響” 東北工業大学紀要, 1理工学編, 第27号, p.133~139 (2007.3)
- 7) 播摩敏雄, 砂田大輔, 安倍幸治, 高根昭一, 曾根敏夫, “音像定位における先行音効果に及ぼす視覚情報の影響” 日本音響学会 騒音・振動研究会資料資料番号, N-2005-39, p.1-7, (2005)
- 8) 播摩敏雄, 安倍幸治, 高根昭一, 曾根敏夫, “音像定位における先行音効果とエコー知覚の限界に関する考察” 電子情報通信学会信学技報, ヒューマン情報処理, HIP 2004-75, p.13-18, (2004.12)

- 9) 播摩敏雄, 曾根敏夫, “正中面上に妨害音源がある場合の方向定位” 東北工業大学情報処理技術研究所紀要, EOS, Vol. 17 No. 1, p. 27-32, (2004)
- 10) 播摩敏雄, 曾根敏夫, “先行音効果による音像定位とエコー検知限に関する検討” 電気関係学会東北支部連合大会講演番号 218, p. 320, (2004)
- 11) 播摩敏雄, 曾根敏夫, “音の定位における先行音効果に関する考察” 日本音響学会騒音・振動研究会資料, 資料番号 N-2004-42, p. 1-7, (2004)
- 12) 播摩敏雄, 曾根敏夫, “音像定位における先行音効果の限界に関する検討” 日本音響学会講演論文集, 3-5-9, p. 549-550, (2004. 9)
- 13) 播摩敏雄, 曾根敏夫, “バイク用ヘルメット装着時における実音源の方向定位” 日本騒音制御学会誌, 騒音制御: Vol. 27, No. 5, p. 365-372, (2003. 10)
- 14) Toshio Harima, Yōiti Susuki and Toshio Sone, “Influence of time difference and/or level difference for synthetic sound image” Memoirs of the Tohoku Institute of Technology Ser. I: Science and Engineering, No. 22, p. 135-144, (2002. 5).
- 15) 播摩敏雄, 鈴木陽一, 曾根敏夫, “近接2音源による合成音像定位” 平成13年度電気関係学会東北支部連合大会 2D21, p. 137, (2001)
- 16) 播摩敏雄, 鈴木陽一, 曾根敏夫, “音像定位におよぼす両耳間時間差と音圧レベル差の相互作用に関する一考察” 東北工業大学紀要 I: 理工学編, 第20号, p. 225-231, (2000. 3)
- 17) 播摩敏雄, 鈴木陽一, 曾根敏夫, “合成音像に対する時間差とレベル差の影響” 日本音響学会講演論文集, 1-2-4, p. 301-302, (2000)
- 18) 播摩敏雄, 鈴木陽一, 曾根敏夫, “2音源による時間差と音圧レベル差の交換作用が音像定位に及ぼす影響” 平成12年度電気関係学会東北支部連合大会, 2A19, p. 34, (2000)
- 19) 播摩敏雄, 鈴木陽一, 曾根敏夫, “音像定位に及ぼす両耳間時間差と音圧レベル差の

- 相互作用に関する一考察” 東北工業大学紀要 1: 理工学編第20号, p. 225-231, (2000)
- 20) Toshio Harima, Yôiti Susuki Nobuji Saito and Toshio Sone. “Critical rotational frequency of sound image for rotating sound source” 東北工業大学情報処理技術研究所紀要 EOS Vol.12, No.1, p. 63-66 (1999)
- 21) Toshio Harima, Yôiti Susuki and Toshio Sone, “Sound image localization of two coherent sources in median plane” 東北工業大学情報処理技術研究所紀要, EOS Vol. 11, No.1, p. 61-64, (1998)
- 22) 播摩敏雄, 高根昭一, 鈴木陽一, 曾根敏夫, “正中面における合成音像制御に関する一考察” 日本音響学会講演論文集, 2-3-10, p. 437-439, (1997. 9)
- 23) Toshio Harima, Shouichi Takane, Yôiti Susuki and Toshio Sone. “Localization of virtual sound image generated by two sources located on the median plane” Memoirs of the Tohoku Institute of Technology Ser. I : Science and Engineering. No.17, p. 175-179, (1997.3)
- 24) Toshio Harima, Shouichi Takane, Yôiti Susuki and Toshio Sone. “Localization of virtual sound image generated by two sources located on the median plane” The Sixth Western Pacific Regional Acoustics Conference 19-21 November 1997, Hong Kong No. G4, p. 568-573, (1997)
- 25) 播摩敏雄, 阿部一任, 鈴木陽一, 曾根敏夫. “前方正中面における音の方向定位に関する一考察” 日本音響学会講演論文集, 2-7-3, p. 321-322, (1997)
- 26) Toshio Harima, Shouichi Takane, Yôiti Susuki and Toshio Sone. “Localization of virtual sound image generated by two sources located on the median plane” *J. Acoust. Soc. Jpn. (E)*, Vol.18 No. 4, p. 205-208 (1997)
- 27) 播摩敏雄, 鈴木陽一, 曾根敏夫, “正中面に置かれた2音源による合成音像定位” 東北工業大学紀要 1: 理工学編, 第16号, p. 209-214, (1996. 3)

- 28) 播摩敏雄, 鈴木陽一, 曾根敏夫, “正中面に置かれた2音源による合成音像定位” 東北工業大学紀要 I : 理工学編, 第16号, p. 209~214 (1996. 3)
- 29) Toshio Harima, Yôiti Susuki and Toshio Sone. “Localization and critical rotational speed of sound image for simulated rotating sound source” *Memoirs of the Tohoku Institute of Technology Ser. I : Science and Engineering*, No.15, p.151-157, (1995. 3)
- 30) 播摩敏雄, 山田弘, 松谷栄一. “正中面内のコヒーレントな2音源の方向定位” 東北工業大学紀要 I : 理工学編, 第15号, p. 159-164, (1995. 3)
- 31) Toshio Harima, Yôiti Susuki and Toshio Sone. “Critical rotational speed for a rotating sound image” *J. Acoust. Soc. Jpn. (E)*, Vol.15, No. 3, p. 207-209, (1994)
- 32) 播摩敏雄, 鈴木陽一, 曾根敏夫. “正中面におかれた二音源による音像定位” 日本音響学会講演論文集, 3-7-7, p. 565-566, (1994)
- 33) 播摩敏雄, 阿部一任, 鈴木陽一, 曾根敏夫. “正中面における音の方向定位に関する一考察” 東北工業大学紀要 I : 理工学編, 第13号, p. 199-208 (1993. 3)
- 34) 播摩敏雄, 山田弘, 松谷栄一, “疑似移動音源による音像の知覚” 東北工業大学紀要 I : 理工学編, 第13号, p. 209-214, (1993. 3)
- 35) 播摩敏雄, 阿部一任, 鈴木陽一, 曾根敏夫, “正中面における音の方向定位に関する一考察” 電子情報通信学会, 信学技報, EA92-46, p. 47-54, (1992. 8)
- 36) 播摩敏雄, 阿部一任, 曾根敏夫, “正中面における音の方向定位に関する一考察” 日本音響学会聴覚研究委員会, 資料番号 H-92-46, p. 1-8, (1992. 8)
- 37) 播摩敏雄, 阿部一任, 曾根敏夫, “自動二輪車用ヘルメットの遮音特性の測定” 一時間引き伸ばしパルスによるデジタル信号処理手法—, 東北工業大学紀要 I : 理工学編, 第12号, p. 229-234, (1992. 3)

- 38) 播摩敏雄, 阿部一任, 鈴木陽一, 曾根敏夫, “回転音源による音像軌跡” 日本音響学会講演論文集, 1-8-7, p. 363-364, (1991. 3)
- 39) 播摩敏雄, 山田弘, 大坂和博, 古賀秀昭, 松谷栄一, “自動二輪車用ヘルメット着用時の方向定位に及ぼす音圧レベルの影響” 東北工業大学紀要 I : 理工学編, 第11号, p. 153-159, (1991. 3)
- 40) 播摩敏雄, 山田弘, 松谷栄一, “二輪車用ヘルメット着用時における実音源の方向定位” 日本騒音工学会, 技術発表会, 2-1-16, (1990. 10)
- 41) 播摩敏雄, 山田弘, 大坂和博, 古賀秀昭, 松谷栄一, “自動二輪車用ヘルメット着用時における帯域音の遮音特性” 東北工業大学紀要 I : 理工学編, 開学25周年記念号, 第10号 p. 361-373, (1990. 3)
- 42) 播摩敏雄, 山田弘, 大坂和博, 古賀秀昭, 松谷栄一, “二輪車用ヘルメット着用時における帯域音実音源の遮音特性” 東北工業大学紀要 1 : 理工学編, 第10号, p. 361-373, (1989. 3)
- 43) 播摩敏雄, 山田弘, 大坂和博, 古賀秀昭, 松谷栄一, “自動二輪車用ヘルメット着用時における実音源の方向判定” 東北工業大学紀要 1 : 理工学編, 第9号, p. 189-203, (1989. 3)
- 44) 播摩敏雄, 山田弘, 大坂和博, 古賀秀昭, 松谷栄一, “ヘルメット装着時における実音源の方向判定” 日本音響学会聴覚研究会資料, 資料番号 H-88-20, p. 1-9, (1988. 5)
- 45) 播摩敏雄, 山田弘, 大坂和博, 古賀秀昭, 松谷栄一, “白色雑音中における実音源の方向判定” 東北工業大学紀要 1 : 理工学編, 第8号, p. 199-213, (1988. 3)
- 46) 播摩敏雄, 山田弘, 古賀秀昭, 松谷栄一, “白色雑音中における実音源の方向判定” 日本音響学会聴覚研究会資料, 資料番号 H-87-27, p. 1-9, (1987. 5)
- 47) 播摩敏雄, 山田弘, 古賀秀昭, 松谷栄一, “雑音中における実音源の方向定位” 東北工業大学紀要 1 : 理工学編, 第7号, p. 197-206, (1987. 3)

- 48) 播摩敏雄, 山田弘, 古賀秀昭, 松谷栄一, “移動音源の音像軌跡と方向定位” 東北工業大学紀要 1: 理工学編, 第6号, p197-206, (1986. 3)
- 49) 播摩敏雄, 山田弘, 古賀秀昭, 松谷栄一, “回転音源による音像軌跡” 東北工業大学紀要 1: 理工学編, 第6号, p. 79-86, (1986. 3)
- 50) 播摩敏雄, 橋本廣一, “ステレオ音場における時間差と強度差による虚音像の方向定位について(1)” 東北工業大学紀要 G 電子・通信工学編, 第19号, p. 1-6, (1979. 3)