

2017年度 博士論文

音楽の明瞭さの評価要因と音場の物理量の対応

The relationship between evaluation terms of clarity of
music and acoustical parameters

2015年入学

学籍番号 2315917

今村秀隆

Hidetaka Imamura

2017年10月31日

東京藝術大学大学院音楽研究科音楽文化学専攻
音楽音響創造研究分野

Research Field of Creativity of Music and Sound
Tokyo University of the Arts

要旨

音楽聴取において、人が明瞭さを感じる要因となる音場の物理量を3つの心理実験を実施することで調査した。

先行研究より、コンサートホールを想定した響きの多い音場における演奏音聴取の条件では、初期反射音の音圧レベルが高いほど明瞭だと評価されることが明らかになっている。本研究では、響きの少ない小空間での再生音の聴取における初期反射音の影響を調査することで、スタジオ施設のコントロールルームやリスニングルームにおいて明瞭さと対応する音場の物理量の解明を目指した。更に、明瞭さという言葉が指す感覚が抽象的であることに着目し、明瞭さの具体的な評価要因の解明にも取り組んだ。心理実験のために、実験室において22台のスピーカーを用いて反射音を模擬し、時間的・空間的特徴の異なる複数の音場における音楽聴取を再現した。

実験1では、音場の室容積が $60\text{ m}^3\sim 400\text{ m}^3$ 、かつ後期残響音が存在しない条件下で、明瞭さに最適とされる $30\text{ ms}\sim 80\text{ ms}$ の初期反射音の音圧レベルの調整を被験者に求めた。聴取音源として、スピーカー聴取を想定したステレオ市販音源、および演奏音を想定したモノラル無響音源の2種類を用いた。被験者に着目を求める感覚を「音色の明瞭さ」と「音像の明瞭さ」の2つとした。実験の結果、明瞭さに最適な初期反射音の音圧レベルは、聴取する音源、そして着目する明瞭さの評価要因によって異なることが明らかになった。無響音源の聴取においては、音色の明瞭さは直接音に対して約 -10 dB 、音像の明瞭さは約 -20 dB の初期反射音の音圧レベルが最適とされた。一方で、市販音源の聴取においては、音色および音像の明瞭さのどちらも直接音に対して約 -23 dB の初期反射音の音圧レベルが最適とされた。最適な初期反射音の音圧レベルには個人差があり、約 $10\text{ dB}\sim 20\text{ dB}$ の差が見られた。更に、高い音圧レベルを明瞭だと評価する被験者は、初期反射音の到来方向の偏りにも着目して判断を行うことが示唆された。

実験2では、人が具体的にどのような要因に着目して音楽の明瞭さを評価しているかについて、レポートリーグリッド法に基づくインタビュー調査を行った。調査の結果、音場の室容積が 125 m^3 および 512 m^3 の条件下では、無響音源の明瞭さの主な評価要因は「こもり具合」、「距離感」、「響きの量の適切さ」、「聴き取りやすさ」といった音色の明瞭さに関するものであることが分かった。一方、市販音源の聴取においては、音色の明瞭さに加え、「輪郭」や「定位の分かりやすさ」、「音像の広さ」等も評価要因とされた。更に、ある被験者のグループは「音色の明瞭さ」、「音像の広さ」、「響きの量の適切さ」、もう一方のグループは「輪郭」や「定位の分かりやすさ」に特に着目し、それぞ

れ別の観点で音場の明瞭さを評価した。小空間での市販音源の聴取においては、明瞭さは複数の要因によって評価されることが明らかになった。それぞれの評価要因と対応する音場の物理量を明らかにするため、実験3を実施した。

実験3では、実験2から得られた明瞭さの評価要因について、音場の物理量との対応を調査した。調査の結果、音場の室容積が $60\text{ m}^3\sim 224\text{ m}^3$ という条件下で市販音源を聴取した場合、音楽の明瞭さは「音色の明瞭さ」、「音像の明瞭さ」、「音像の広さ」の3つに大別されることが明らかになった。全ての評価要因において、室容積が大きいくほど、そして初期反射音の到来方向の偏りが小さいほど明瞭と評価された。「音像の明瞭さ」はこの傾向に加え、初期反射音が直接音と同じ方向から到来する場合により明瞭と評価された。更に、音像の広さは1-IACC（両耳間相関度）が高いほど広いと評価された。

以上の結果より、生楽器の聴取を目的とした演奏空間においては、その室容積に関わらず、直接音に対して約 -10 dB の高い初期反射音の音圧レベルが明瞭とされることが示唆された。演奏空間の音楽の明瞭さの評価要因は、直接音の音質や距離感、聴き取りやすさといった音色の明瞭さに関するものである。一方、スタジオ施設におけるコントロールルーム等の小空間においては、音色の明瞭さに加え、音像の明瞭さや音像の広さにも着目して明瞭さの評価がされることが示唆された。小空間における明瞭さおよび音像の明瞭さは、直接音に対して約 -23 dB の低い初期反射音の音圧レベルが適切とされ、到来方向の偏りが小さいほど明瞭だとされることが明らかになった。

目次

第1章	序論	6
1.1	室内音響設計とは	7
1.1.1	響きの知覚	7
1.1.2	要素感覚に基づく室内音響設計	8
1.2	先行研究と課題	9
1.2.1	コンサートホールにおける音楽の明瞭さの研究	9
1.2.2	小空間における音楽の明瞭さの研究	12
1.2.3	課題	15
1.3	本研究の目的	17
1.4	論文の構成	17
第2章	音楽の明瞭さの多次元性	19
2.1	コンサートホールにおける音楽の明瞭さの多次元性	20
2.2	音楽再生システムにおける音楽の明瞭さの多次元性	21
2.2.1	ITU-R 勧告 BS.1116 における明瞭さの記述	21
2.2.2	IEC 60268 における明瞭さの記述	23
2.2.3	その他の文献における明瞭さの記述	24
2.3	筆者による先行研究	25
2.3.1	一次反射音が明瞭さに与える影響	25
2.3.2	小空間における音楽の明瞭さの多次元性	28
2.4	結論	29
第3章	音場の可聴化	30
3.1	仮想音源分布を用いた音場の可聴化	31
3.2	本研究で用いたシミュレーション手法	35
3.2.1	初期反射音の到来方向の設定	35
3.2.2	生成したインパルス応答の畳み込み	36

第 4 章	実験 1：明瞭さに最適な初期反射音の音圧レベルの調査	37
4.1	実験方法	38
4.1.1	調整法	38
4.1.2	実験刺激の作成	38
4.1.3	初期反射音レベルの調整方法	44
4.2	分析	46
4.2.1	前処理	46
4.2.2	4 要因の分散分析	46
4.2.3	クラスター分析によるグループ分け	56
4.2.4	グループ別の音源の主効果	59
4.2.5	グループ別の室容積の主効果	61
4.2.6	グループ別の評価語 × 到来方向の交互作用	63
4.3	考察	65
4.3.1	着目する明瞭さ、および聴取音源の違いが最適な初期反射音レベルに与える影響	65
4.3.2	室容積の違いが最適な初期反射音レベルに与える影響	68
4.3.3	初期反射音の到来方向の違いが最適な初期反射音レベルに与える影響	68
4.4	結論	69
第 5 章	実験 2：音楽聴取における明瞭さの評価要因	70
5.1	実験方法	71
5.1.1	レパートリーグリッド法	71
5.1.2	実験刺激の作成	71
5.1.3	インタビュー調査	76
5.2	分析	76
5.2.1	音場の明瞭さの順位	76
5.2.2	自由記述の分析	77
5.2.3	聴取音源に基づく評価要因の分類	79
5.2.4	室容積に基づく評価要因の分類	82
5.2.5	初期反射音の到来方向に基づく評価要因の分類	83
5.2.6	再生音圧の補正有無に基づく評価要因の分類	85
5.2.7	被験者グループによる評価要因の分類	85
5.3	考察	89
5.3.1	音楽の明瞭さの評価要因	89
5.3.2	聴取音源の違いが評価要因に与える影響	90

5.3.3	初期反射音の到来方向の違いが評価要因に与える影響	92
5.3.4	評価要因の個人差	93
5.3.5	再生音圧の補正が評価要因に与える影響	93
5.4	結論	93
第 6 章	実験 3：小空間における明瞭さに対応する音場の物理量	95
6.1	実験方法	96
6.1.1	シェッフェの対比較法	96
6.1.2	実験刺激の作成	96
6.1.3	比較評価の方法	101
6.2	分析	102
6.2.1	7つの評価要因の評価値の分析	102
6.2.2	個人差を考慮した INDSICAL	105
6.3	考察	108
6.3.1	音楽の明瞭さの評価要因	108
6.3.2	「音色の明瞭さ」の評価要因と物理量との対応	108
6.3.3	「音像の明瞭さ」と「音像の広さ」の物理量との対応	112
6.4	結論	115
第 7 章	総括	116
7.1	音楽の明瞭さに最適な初期反射音の音圧レベル	116
7.2	音楽の明瞭さの評価要因	117
7.3	小空間における明瞭さに対応する音場の物理量	118
7.4	室内音響設計への応用と今後の研究の課題	119
謝辞		121
発表論文等一覧		122
参考文献		123
付録 A	インパルス応答の生成に使用したソースコード	129

第1章

序論

音場の響きが音楽の聴取に影響することは広く知られている。我々が感じる音楽のダイナミクス、音の持続感、距離感等の様々な感覚が響きによって変化し、音楽の総合的な印象を形成する。

これらの感覚の中でも、「明瞭さ」は特に重要な要素である。響きが楽音の明瞭な聴取を阻害してしまうと、音楽の印象は大きく損なわれる。これまで、明瞭さに焦点を当てた音場の設計に関する研究は数多く行われてきた。しかし、その多くはコンサートホール等の大規模な空間における演奏音の聴取に限定されており、現実の環境を十分に反映していない。現在、市場に流通する音楽の多くは、スタジオ施設のコントロールルームやリスニングルームといった小規模な空間でスピーカーを用いて制作、聴取されている。

本研究は、人が音楽聴取において明瞭さを感じる音場の物理量の解明を目指す。特に、小空間に特徴的な初期反射音の影響に焦点を当て、スタジオ施設のコントロールルームやリスニングルームの音響設計に貢献することを目指す。

本章では、研究の導入として以下の項目を解説する。

1. 室内音響設計とは
2. 先行研究と課題
3. 本研究の目的
4. 論文の構成

1.1 室内音響設計とは

1.1.1 響きの知覚

室内で音が生じると、音源から聴取者に直接届く直接音に加えて、壁や天井で反射した複数の反射音が遅れて到達する。我々は音楽を聴く際、楽器や声、スピーカー等の音源から発せられる楽音だけでなく、これら反射音の集まりで形成される響きを共に聴いている。響きの特徴は反射音の空間的および時間的特徴によって定義される。

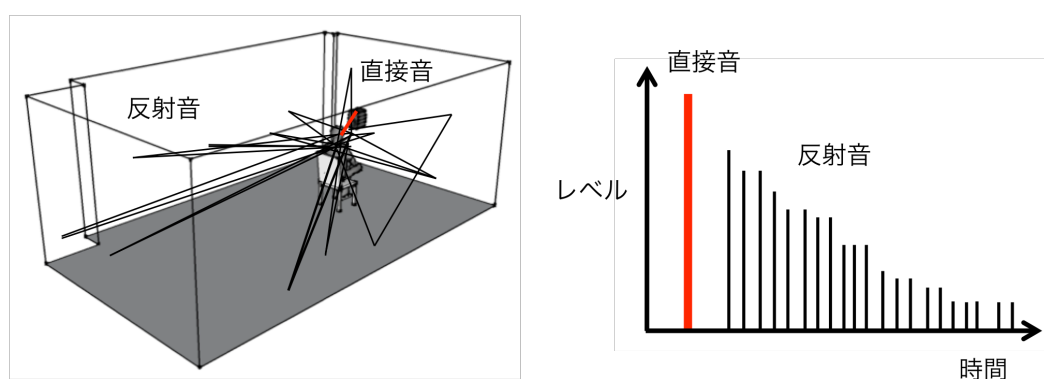


図 1.1.1 室内空間における直接音と反射音の空間的（左）・時間的（右）特徴の概念図

図 1.1.2 は、人の音楽聴取のプロセスを示す (森本, 1990)。音源から発せられた音は壁や天井で反射し、空間伝達関数で表される伝送路を経て、聴取者の位置に到達する。その後、頭部伝達関数の影響を受け聴取者の外耳道入り口に達し、心理空間の入力である音刺激となる。心理空間において、聴取者は音刺激をまず心理的事象である音像として知覚する。この音像に対する感覚は要素感覚と呼ばれ、個人の好みや欲求等によって評価に大きな差が生じない、基本的な感覚であるとされている。現在明らかになっている要素感覚の例としては、残響感や音の持続感 (Jordan, 1981)、音像の広さや音による包まれ感 (Ando, 1985)、距離感や奥行き感 (亀川・丸井, 2016)、ラウドネスや明瞭性等が挙げられる (ISO3382-1, 2009)。最終的に、聴取者は音像の様々な要素感覚を自身の好みや欲求に照らし合わせ、主観的な評価を行う。「好ましさ」や「心地よさ」といった抽象的な印象は、聴取の目的や用途、個人によって異なる総合主観印象となり、要素感覚の組み合わせによって構成されると言われている。

Olive and Martens (2007) による研究では、音楽を聴取する空間が異なると、スピーカーに対する嗜好度の評価も変化することが示され、音の好みまでもが響きに影響されることを明らかにした。また、Pätynen and Lokki (2016) はコンサートホールにおける側方反射音が音楽のダイナミクスを深め、演奏から引き起こされる情動を強くすることを明らかにした。このように、響きが音楽聴取に与える影響は無視できないほど大きい。

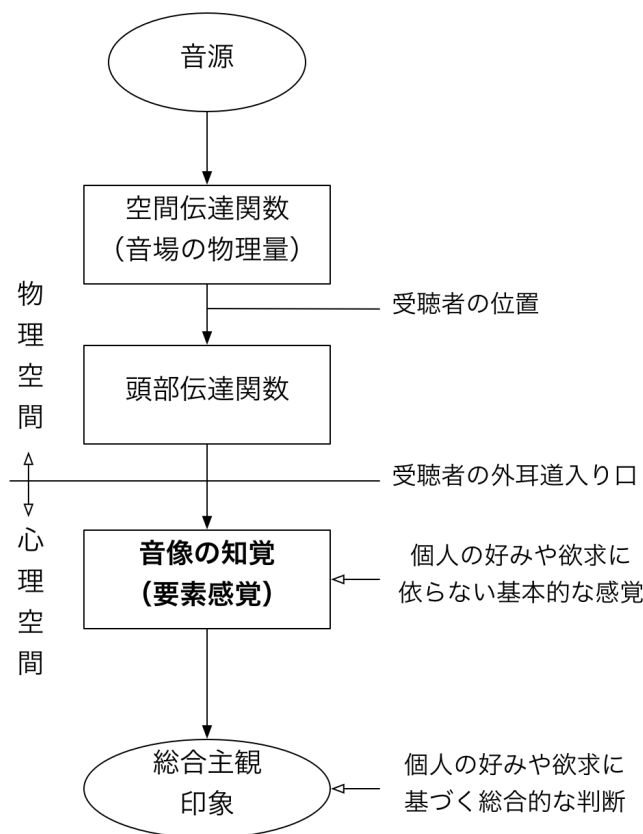


図 1.1.2 人の音楽聴取のプロセス。森本 (1990) の図に一部加筆。

1.1.2 要素感覚に基づく室内音響設計

音楽聴取のための音場には唯一無二の最適な響きが存在するわけではなく、人間が音楽を聴く際の目的によって最適だと考えられる特徴は異なる。例えば、コンサートホール等の演奏空間においては楽音が豊かに響くことや、音に包まれるような感覚が求められる場合が多い。一方、音楽制作が目的であるスタジオ施設のコントロールルームやリスニングルームでは、音の周波数特性が平坦であることや、詳細な音の変化が聴き取れることが求められる。異なる目的には異なった検討、設計がされることが望ましい。

過去には、残響時間等の物理量と、好ましき等の総合主観印象との対応を調べる室内音響研究が行われていた。しかし、(Barron, 1988) によって音場の物理量と総合主観印象の間には明確な対応関係が存在せず、総合主観印象は個人によって異なる評価がされるという研究結果が報告されて以来、物理量と要素感覚との対応関係を調査する研究が主流となっている。要素感覚を制御できれば、任意の音量感や残響感を持つ響きの設計が可能となり、抽象的な設計目標ではなく「どのような特徴を持った音を目指すか」という要素感覚に基づく設計が可能となる。

室内音響設計は、用途や目的、対象とする聴取者に合わせて達成すべき要素感覚を決定し、その

要素感覚を制御するために物理的な工夫を凝らす行為であると言える。このような室内音響設計のためには、要素感覚と対応する物理的要因を明らかにしてゆくことが必要である。

1.2 先行研究と課題

本研究では、様々な要素感覚の中でも特に明瞭さに着目する。音楽聴取において高い明瞭さを実現することは非常に重要である。スタジオ施設のコントロールルームやリスニングルームでは、エンジニアや音楽家が細かな音色変化や音の空間配置の違いを明確に聴き取れることが要求される。また、コンサートホール等の演奏空間においては、演奏内容が十分に聴き取れるための明瞭さが確保される必要がある。更に、スピーカー、ヘッドホン等の再生機器に関しても、オーディオ雑誌や製品レビューでは明瞭さ、クリアさ、透明感、解像度といった言葉が頻繁に使われる。音楽家やエンジニアだけでなく、一般聴取者にも音楽聴取における明瞭さの重要性が認識されていることが明らかである。

本節では、代表的な音楽聴取環境としてコンサートホールおよびスタジオ施設の 2 つに焦点を当て、音楽の明瞭さの向上を目的とした室内音響研究について概観する。更に、本研究で取り扱う課題についても述べる。

1.2.1 コンサートホールにおける音楽の明瞭さの研究

コンサートホールは一般的に響きが長く、楽音がその響きによって阻害されてしまう場合がある。このため、音楽の明瞭さに関する研究も盛んに行われている。

音楽聴取のための響きの科学的研究は、1900 年の Sabine による残響式 (Sabine, 1900) によって、響きの長さを定量的に説明することから始まったとされる。その後、響きの研究は主にコンサートホールや劇場を対象に発展してきた。18 世紀以降、特にヨーロッパを中心に建設されたコンサートホールは、大きな室容積が特徴であり、約 $15000 \text{ m}^3 \sim 25000 \text{ m}^3$ の容積を持つ (Beranek, 2004)。図 1.2.1 には、本学のコンサートホールである奏楽堂の写真を、図 1.2.2 には本学音楽学部 4 号館第 6 ホールの写真を示す。

残響時間は響きの長さを示す指標であり、室内において音を止めてから残響音が 60 dB 減衰するのに要する時間として定義されている。Sabine の残響式 (式 1.1) によれば、残響時間 RT (Reverberation Time) は室容積 V 、室総表面積 S と室の平均吸音率 $\bar{\alpha}$ によって求められる。多くのコンサートホールが約 1.5~3.0 秒の響きを持つことが知られている (Beranek, 2004)。Sabine による残響式の提案以降、残響時間は長らく音場の響きを代表する物理量として扱われてきた。国際規格である ISO3382-1 (2009) には、コンサートホールにおける残響時間の測定方法や音響指標が記載されている。

$$RT = 0.161 \frac{V}{S\bar{\alpha}} \quad (1.1)$$

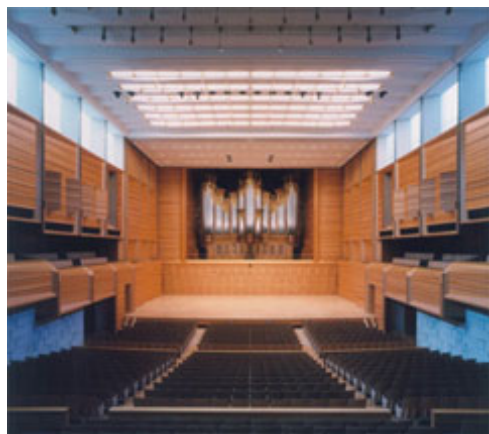


図 1.2.1 東京藝術大学 奏楽堂



図 1.2.2 東京藝術大学 音楽学部4号館第6ホール

一方、Beranek (1962) による 58 のホールを対象とした測定および指揮者、演奏家、音楽評論家へのアンケート調査の結果、残響時間がほぼ同じであってもコンサートホールの評価には大きなばらつきが生じることが明らかになった。これは、残響時間以外の指標がコンサートホール設計においても重要であることを示唆している。Beranek は、評価に最も関係する物理量はホールの横幅であり、横幅の狭いホールが高い評価を受けると主張した。横幅が狭いホールでは、側壁間で生じる反射音の伝搬距離が短く、直接音に対して遅延時間が短い反射音が多く客席に届く。この遅延時間が短い反射音を初期反射音と言い、Beranek は初期反射音がコンサートホールの評価に重要な貢献をしていると主張した。この結果を受け、コンサートホールにおける初期反射音に着目した研究が数多く行われるようになった。

初期反射音の影響を調査した基礎的な研究の一例は、Haas (1951) による音声を用いた研究である。Haas は 2 つのスピーカーを用いて、一方からの音声に対してもう一方の音声を様々に減衰または増幅させ、エコー障害が知覚される割合を調べた。図 1.2.3 に示される実験結果には、直接音に対する遅延音の遅延時間と音圧レベル、そして被験者がエコー障害を知覚した割合の関係が示されている。Haas の実験結果によると、遅延時間が 80 ms 未満で音圧レベルが直接音に比べて約 -3 dB 低い場合、ほとんどの被験者はエコー障害を知覚せず、遅延時間が 30 ms 以内であれば、音圧レベルが直接音に対して 10 dB 大きくてもエコー障害は感じられない。Haas の結果は Muncy et al. (1953) により追加検証され、音声ではなく音楽を聴取する場合は、音声のケースと比べてエコー障害が知覚される遅延音の遅延時間は長くなり、音圧レベルも高くなることが明らかにされた。

Haas の研究結果を受け、Thiele (1953) は直接音からの遅延時間が 50 ms 以内である初期反射音が音声の明瞭性を向上させる効果があることを明らかにし、音声の明瞭さを定量化する室内音響指標 D_{50} (Deutlichkeit: ドイトリヒカイト) を提案した。この指標は、エコー障害を引き起こさない 50 ms 以内の初期反射音が音声の音量感を補強するという考えに基づいている。Thiele は心理実験を通じて、 D_{50} 値が大きいほど音声の明瞭度が高いことを検証した。また、Beranek and Schultz (1965) は、 D_{50} が音声だけでなく音楽に対しても有効であると主張した。

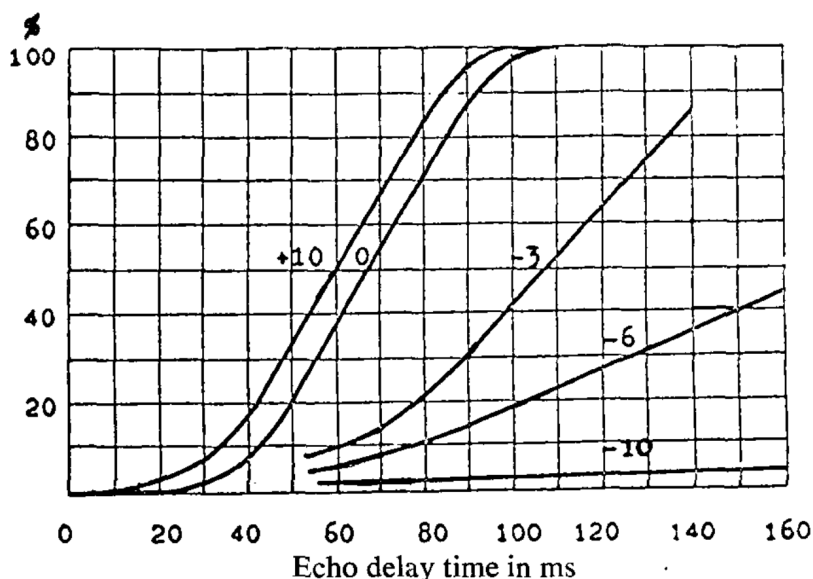


図 1.2.3 Haas (1951) による実験結果を示す。縦軸は遅延された音声によって遅延されていない音声に妨害されたと感じられた割合、横軸は遅延されていない音声と遅延された音声との遅延時間を示す。曲線の数字は、遅延されていない音声に対する遅延された音声の音圧レベルを示す。

Reichardt et al. (1975) は、コンサートホールにおける音楽の明瞭さに関する室内音響指標として、式 1.2 で表される C_{80} を提案した。これは、直接音の到来後 80 ms までに到来する初期反射音はエコー障害とは知覚されず、音楽聴取における楽音を補強し、その後到来する残響音は直接音の聴取を阻害するという考えに基づいている。 C_{80} は D_{50} を基にしており、値が大きいほど音楽の明瞭さが高いとされる。Reichardt は実験室でスピーカーを用いた心理実験を行い、異なる C_{80} 値を持つ音場を模擬して、知覚される音楽の明瞭さと C_{80} 値との高い相関を明らかにした。また、 C_{80} 値を用いてコンサートホールにおける音楽の明瞭さを 3 段階に分類し、 C_{80} 値が -1.6 dB 以下を低明瞭度 (bad clarity)、 -1.6 dB から 1.6 dB を境界領域 (border category)、 1.6 dB 以上を高明瞭度 (good clarity) と定義した。

$$C_{80} = 10 \log_{10} \frac{\int_0^{80} p^2(t) dt}{\int_{80}^{\infty} p^2(t) dt} [\text{dB}] \quad (1.2)$$

Höhne and Schroth (1995) によれば、コンサートホールでの音楽聴取における C_{80} の丁度可知差異 (JND: Just Noticeable Difference) は 2 dB から 3 dB である。現在、 C_{80} は ISO3382-1 (2009) によってコンサートホールで知覚される音楽の明瞭さを表す響きの物理量として標準化されている。

Beranek (2004) は、コンサートホールにおける音楽の明瞭さは、更に水平方向の明瞭さ (Horizontal definition) と垂直方向の明瞭さ (Vertical definition) の 2 つの感覚に区別されると主張している。水平方向の明瞭さは連続して演奏される音が区別できるかどうか、垂直方向の明瞭さは同時に演奏される音が区別できるかどうかに関するものである。しかし、Kahle and Bruneau (1995)

は、これら 2 つの明瞭さはコンサートホールにおける響きの変化においては評価の相関が高いため、区別することにはあまり意味がないと主張している。現在、 C_{80} はこれら 2 つの明瞭さの両方に対応する値として扱われている。

Reichardt による C_{80} の提案以降、コンサートホールのような大空間でのオーケストラ聴取において、 C_{80} 値が聴感上の明瞭さと良好に対応することが多くの研究で確認されている (Barron, 1993) (Miśkiewicz et al., 2012)。一方で、Marshall (1994) は明瞭と判断されるために必要な C_{80} の値や、直接音を補強する初期反射音の遅延時間が音楽のジャンルや楽器によって異なる可能性を指摘している。また、Marshall は C_{80} がコンサートホールのような大空間での演奏音の聴取にのみ有効であり、異なる空間では別の検討が必要であるとも述べている。しかし、聴取する音源や空間の容積による影響についてはほとんど検討されていない。これは、コンサートホールにおいては一般的にオーケストラや楽器等の生楽器が聴取され、室容積も前述のように約 15000 m^3 以上であることが多いためだと言える。

1.2.2 小空間における音楽の明瞭さの研究

スタジオ施設のコントロールルームやリスニングルーム等、コンサートホールと比較して室容積が小さく、かつ生演奏の聴取を目的としない空間は、小空間 (Small Rooms) や再生環境 (Reproduction Environment) として、前述した演奏空間と区別して取り扱われている (Toole, 2009) (ISO3382-2, 2008)。これらの小空間は、室容積が約 100 m^3 前後とコンサートホール等の大空間と比較して極端に小さく、残響時間も約 $0.2 \sim 0.6$ 秒と短いことが特徴であるが、音響的に小空間と大空間を区別する有効な定義は存在していない。小空間として小ホールのような演奏空間を扱っている研究も存在しているため、現状ではコンサートホールや劇場以外の空間が小空間と呼ばれている。本研究では混同を避けるため、小空間の中でも生演奏の聴取を目的としない空間、すなわちスタジオ施設のコントロールルームやリスニングルームといったスピーカーを用いて音楽聴取を行う再生環境を、生演奏の空間を大空間と呼ぶのに対して小空間と定義する。

小空間の響きの特徴は大空間とは大きく異なり、聴取する音源も生楽器ではなくスピーカーを用いた市販音源である。現在、このような小空間において知覚される音楽の明瞭さと対応する響きの物理量はほとんど明らかにされていない。一方、スピーカーから発せられる直接音の聴取がなるべく自然であるようにするための多くの実践的な検討や研究が行われてきた。これらの先行研究の中には、音楽の明瞭さと響きの物理量との関係が見出だせるものが多い。

小空間の室内音響研究の代表例としては、スタジオ施設のコントロールルームに関する研究が挙げられる。コントロールルームの研究は、1947 年に業界初の独立した録音スタジオが設立されてから、盛んに行われるようになった (Newell, 2003)。コントロールルームは、音楽制作過程における録音のモニタリングやミキシングを行う場所で、一般に左右対称の構造を持ち、前方には複数のスピーカーが設置されている。精密なモニタリングを実現するために、エンジニアは前方のスピー

カー二台を結ぶ線を底辺とする正三角形の頂点に相当する位置に座り、スピーカーからの音を聴きつつ、音素材に対して様々な音響編集を行う。近年では、コンピュータ技術の進展や DAW (Digital Audio Workstation) の普及により、セミプロやアマチュアのエンジニア、音楽家も一般的な住居内等の日常的な環境においてこのような音楽聴取を行っている。コントロールルームの例として、図 1.2.4 に本学千住キャンパス内のコントロールルームを、図 1.2.5 に本学の音響制作スタジオを示す。

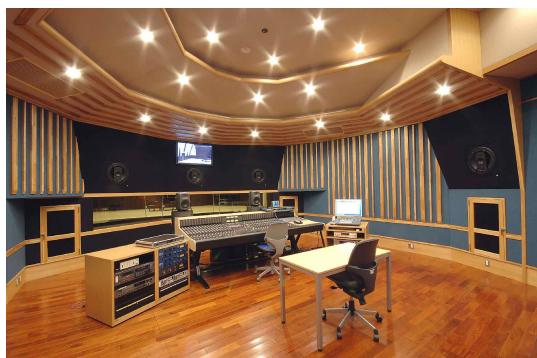


図 1.2.4 東京藝術大学千住キャンパス コントロールルーム

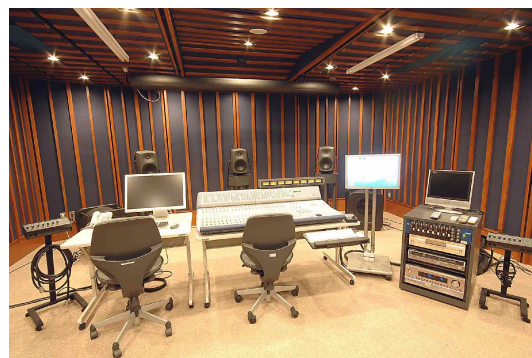


図 1.2.5 東京藝術大学千住キャンパス 音響制作スタジオ

録音やミキシングは原盤制作の過程の中でも重要な作業であるため、スピーカーから再生される音の振幅周波数特性を聴取位置においてなるべく平坦にすることが課題とされてきた。平坦な周波数特性とは、横軸を周波数、縦軸を音圧とした場合に、得られた音圧が周波数によってばらつかず、一定の幅を保っている状態である。これは、多くの場合スピーカーの音色に対する室内の反射音による影響を最小限に抑えることを意味する。

聴取位置での周波数特性のばらつきは、特に低域における定在波の影響が大きい。本研究では定在波による影響を詳細には扱わないが、定在波により音が増幅される部分や打ち消し合って減衰する部分で音を聴取した場合、人が知覚可能なピークやディップが生じることが知られている (Toole, 2006)。定在波への対策としては、部屋の固有モード解析等の手法を用いた設計や、吸音材を用いた処置が行われることが多い (中原他, 2009)。

1970 年以降は、コントロールルームにおける初期反射音が中高域の周波数特性に与える影響の調査が行われてきた。この研究分野では、小空間に特有の初期反射音の特性に注目されている。小空間では初期反射音の伝搬距離が短く、その結果、コンサートホールと比べて距離減衰によるエネルギー減衰が小さく遅延時間も短くなり、聴取位置では直接音と高く相関することとなる。こうして、小空間における初期反射音はコムフィルタ現象を引き起こし、周波数特性に影響を及ぼす。この現象への対策として、Hidley はスタジオ設計におけるノンエンバイロメントルーム (Non-environment Room) の概念を提案した (Newell, 2003)。このアプローチでは、広帯域の吸音処理を施し、コントロールルーム内の反射音を削減することが目指される。室内における反射音の挙動を示す概念図として、図 1.2.6 に一般的な小空間、図 1.2.7 にノンエンバイロメントルームにおける直接音、初期反射音、後期残響音の関係を示す。

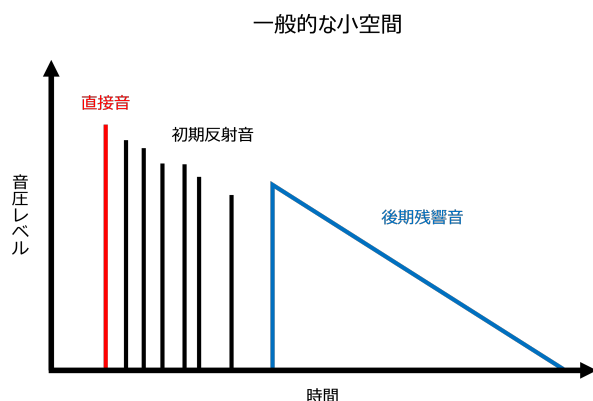


図 1.2.6 一般的な小空間における直接音、初期反射音、後期残響音の関係を表す概念図。初期反射音および後期残響音のどちらも高い音圧レベルを持つ。

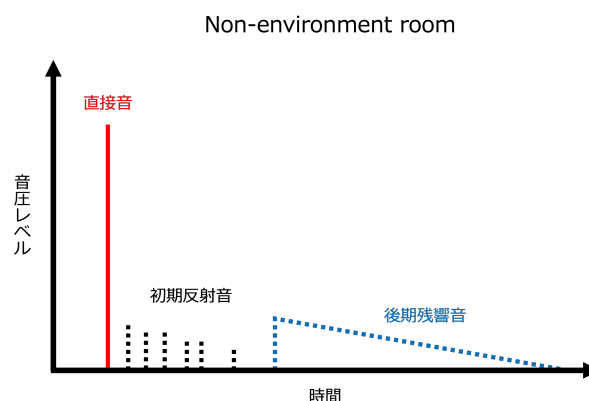


図 1.2.7 コントロールルームに広帯域の吸音処置を施したノンエンバイロメントルームの直接音、初期反射音、後期残響音の関係を表す概念図。室内を吸音することで、初期反射音および後期残響音のどちらの音圧レベルも低くなっている。

エネルギーの大きい初期反射音は直接音の音色だけでなく、距離感 (Griesinger, 2009) や奥行き感 (Theile, 2001)、空間の広がり感 (Blauert, 1996) 等の空間印象に影響することも明らかにされている。一方、室内が無響室のように過剰に吸音されると、周波数特性のばらつきは抑制されるが、エンジニアたちはこのような環境での音楽聴取を「好ましくない」や「不自然」と評価した報告がある (Rumsey, 2008) (尾本他, 2001)。これは、エンジニアはある程度の響きが必要であると感じており、無響室はこうした聴取に最適な環境ではないことを示している。この背景に基づき、周波数特性のばらつきを抑えつつ音楽聴取に適した響きを提供するための、コントロールルームにおける初期反射音の扱いが検討されてきた。その中で、1980 年代に初期反射音を吸音し後期残響音を響きとして残すリフレクションフリーゾーン (RFZ) が提案された (D'Antonio and Konnert, 1984)。図 1.2.8 に、RFZ における直接音、初期反射音、後期残響音の関係を示す。RFZ では、初期反射音の音圧レベルを低く保ちつつ、拡散された後期残響音を響きとして利用する。

更に、2006 年には Massenburg らによって直接音の到来直後に拡散された響きのみを提供する Ambient (響き) + Anechoic (無響) の造語であるアンベコイック (Ambechoic) が提案された。このアプローチに基づいて Blackbird Studio の Studio C が設計されたことがよく知られている (中原他, 2009)。図 1.2.9 に Ambechoic における直接音、初期反射音、後期残響音の関係を示す。直接音が到来した直後に拡散および吸音された残響音が加わる構成が示されている。

このように初期反射音を狙って吸音するアプローチが存在する一方で、前述した Haas の実験結果に基づき、コントロールルームにおける初期反射音は直接音のエネルギーを補強する点で有効であるという主張も存在する (Olive and Toole, 1989)。これは、初期反射音の取り扱いに関してさらなる検討の余地があることを示唆している。小空間における初期反射音が聴感印象に影響することは明らかなものの、設計における初期反射音の扱いには様々なアプローチが存在している。

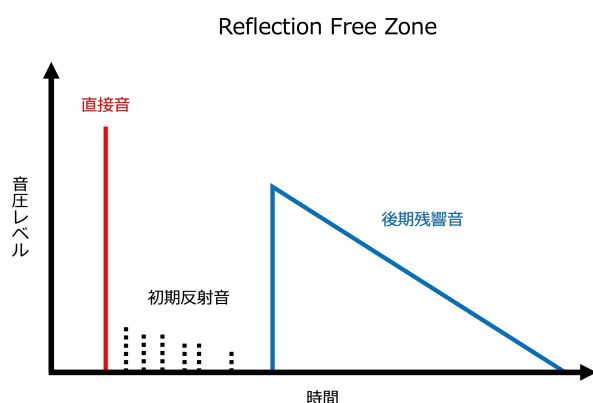


図 1.2.8 リフレクションフリーゾーン (RFZ: Reflection Free Zone) における直接音、初期反射音、後期残響音の関係を表す概念図。初期反射音が吸音され音圧レベルが低く、拡散された後期残響音が付加されている。

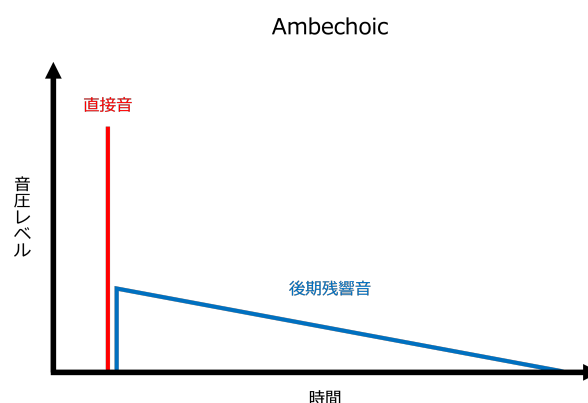


図 1.2.9 アンベコイック (Ambechoic) における直接音、初期反射音、後期残響音の関係を表す概念図。直接音の到来後、すぐに拡散および吸音された残響音が付加されている。

Tervo et al. (2014) は、ミキシングエンジニアによるコントロールルームの好ましさの評価に明瞭さの印象が寄与することを報告している。小空間において明瞭さの感覚が重視されていることは第2章で詳細に述べるが、Newell による著書 (Newell, 2003) や、オーディオシステムの音質評価のための室内に関する ITU-R 勧告、BS.1116-3 (ITU-R, 2015)、IEC 60268 (1997)、Nakayama et al. (1971) による文献からも確認できる。これらの文献から、小空間における音楽の明瞭さを向上させることの重要性が強調される。

1.2.3 課題

コンサートホール等の大空間では、 C_{80} 値が聴感上の音楽の明瞭さに対応することがよく知られている。一方、スタジオ施設のコントロールルームやリスニングルーム等の小空間において明瞭さと対応する物理量は明らかになっていない。本節では、本研究で取り扱う以下の課題について述べる。

1. 小空間における明瞭さに最適な初期反射音の音圧レベルの解明
2. 音楽の明瞭さの評価要因の解明
3. 初期反射音の時間的・空間的特徴による影響の解明

小空間における明瞭さに最適な初期反射音の音圧レベルの解明

Reichardt et al. (1975) や Beranek (2004) による C_{80} に関する研究は、主にコンサートホールでのオーケストラ演奏の聴取に焦点を当てていた。これらの実験では大空間が前提とされ、オーケストラ演奏や楽音等、コンサートホールでよく聴かれる音源が実験刺激として使用されていた。大

空間では直接音に対してエコー障害となる後期残響音の音圧レベルが高い。更に、客席から演奏者までの距離が遠く、直接音の音量感が不足しがちである。このことから、初期反射音によって直接音を補強することが、音量感の向上や響きのかぶりの軽減といった観点で明瞭さの向上に繋がっていると考えられる。

一方、小空間は容積が小さく、コンサートホールとは響きの特徴が大きく異なる。前述したように、小空間における初期反射音はむしろ直接音の音色や空間印象に影響する。小空間では、明瞭さは音量感や響きのかぶりという観点ではなく、音色や空間印象の劣化に着目される可能性が高い。この場合、初期反射音の音圧レベルが高ければ高いほど明瞭であるとは考えにくく、どの程度の初期反射音の音圧レベルが音楽の明瞭さに最適とされるのかを明らかにする必要がある。

更に、Marshall (1994) は、明瞭さの向上に貢献する初期反射音の区間や音圧レベルは、音楽のジャンルや楽器によっても異なる可能性を指摘している。小空間で主に聴取されるのは生の演奏音ではなく録音編集される市販音源であり、生演奏と録音された音源では、初期反射音の音圧レベルの聴感上の最適値は異なると予想される。音源による影響を調査することも必要である。

音楽の明瞭さの評価要因の解明

「音楽の明瞭さ」という評価語は言葉の抽象度が高く、具体的にどのような感覚を指しているのかが曖昧である。音楽の明瞭さが複数の感覚で構成されていることは第2章で詳しく述べるが、Beranek (2004) や BS.1116-3 (ITU-R, 2015) 等の文献からも支持される見解である。音楽の明瞭さが複数の異なる感覚を指す多次元的な概念である場合、明瞭さは要素感覚ではなく総合主観印象と近い要素となり、物理量との対応関係の解明は困難になる。

コンサートホール等の演奏空間においては、恐らく音量感の観点で明瞭さが評価されていることが予想されるが、小空間での聴取においても同様に音量感に着目されるのか、あるいは音質や音像等に基づいて評価されるのかを明らかにする必要がある。

筆者は東京藝術大学の学士課程にてレコーディング、ミキシングを専攻し、様々な制作現場を経験した。制作において、音楽家やエンジニアが音楽の明瞭さについて言及する際に、「こもった感じ」等の音色に関係するものと、「音像のぼやけ具合」等の音像に関係するものとを異なる感覚として区別している場面に多く遭遇してきた。筆者自身、それらの感覚を異なるものとして捉えて制作していた。小空間においては音楽の明瞭さは複数の異なる感覚に分けられるのではないかという問いは、筆者の経験からも生じた研究の動機となっている。

初期反射音の時間的・空間的特徴による影響の解明

大空間における初期反射音は伝播経路が長いため、時間的な密度が小さくエネルギー減衰は穏やかである。一方で小空間では伝播経路が短く、時間的な密度が大きくエネルギー減衰が急である。初期反射音の音圧レベルだけでなく、この室容積に起因する初期反射音の密度とエネルギー減衰の差が音楽の明瞭さにどのように影響するかの解明が必要である。

加えて、コンサートホール等の大空間では拡散音場が仮定されている。拡散音場は、室内のどの点でも反響のエネルギーが均等であり、方向性がない音場を意味する。Gover et al. (2004) による研究では、残響時間が 0.36~0.4 秒の典型的なリスニングルームにおいて球面マイクロフォンアレーを用いて測定を行い、小空間においては初期反射音が強い方向性を持つことを明らかにした。これは、小空間における初期反射音の空間的方向性が音楽の明瞭さに影響を及ぼす可能性を示唆している。初期反射音の到来方向が明瞭さの評価にどのように影響するか調査も必要である。

筆者による先行研究によって、直接音の到来後 30 ms までに到来する単発の一次反射音の到来方向の違いが既に音楽の明瞭さに影響することが明らかになっている。本研究では、直接音の到来後 30 ms~80 ms の間に到来する反射音を初期反射音として扱い、空間的特徴の違いによる影響を解明する。

1.3 本研究の目的

本研究は、人が音楽聴取において明瞭さを感じる音場の物理量の解明を目指す。特に、小空間に特徴的な初期反射音の影響に焦点を当て、スタジオ施設のコントロールルームやリスニングルームの音響設計に貢献することを目指す。このために、小空間で特徴的である強いエネルギーを持つ初期反射音がどのように明瞭さに影響するかを心理実験によって明らかにする。

スタジオ施設のコントロールルームやリスニングルーム等の小空間では、エンジニアや音楽家が細部にわたる音の違いを明瞭に聴き分けることが必須である。これらの環境では、細かい音色の変化や音の空間配置の違いが十分に聴き取れることが求められる。小空間での音楽の明瞭さに関連する物理量を特定することで、より精密な音楽制作環境の設計が可能となり、スタジオ施設や個人の制作環境の改善に繋がる。また、リスニングルームのような一般的な音楽聴取環境においても、聴取者が個人の好みや目的に合わせて明瞭さを調整できるようになるだろう。明瞭さは聴取の満足度と高い相関を示す可能性があるため、本研究は音楽の明瞭さを維持しつつ、響きの豊かさや包み込まれるような感覚を高めるような音響設計手法の開発にも繋がることを期待される。更に、音楽制作で用いるリバーブエフェクターのような音響効果の開発にも役立つことが期待される。

本研究によって特定された物理量を実空間でどのように達成するかについては多くの課題がある。コンピュータや信号処理技術の発展により、将来反射音の高度なアクティブコントロールが実現可能となることで、あらゆる音場で任意の響きの設計が可能になることを期待したい。

1.4 論文の構成

本章では、人の聴感印象と物理量の関係の解明が、音楽聴取の質を高めるために重要であることを述べた。更に、コンサートホールでは直接音の到来後 80 ms までの初期反射音が明瞭さに寄与し、それ以降の後期残響音が明瞭さを損ねるとされる C_{80} 値が、音楽の明瞭さの評価指標として用いら

れてきたことを述べた。一方、小空間での音楽聴取における明瞭さの指標に関する研究は限られており、初期反射音による影響の詳細な検討が必要であることを述べた。

第 2 章では、音楽の明瞭さという言葉の抽象度が高く、具体的にどのような感覚を指しているかが曖昧である点に着目し、音楽の明瞭さが多次元的な感覚であることを示唆する先行研究について述べる。また、音楽の明瞭さが少なくとも「音色の明瞭さ」と「音像の明瞭さ」の二つに区別されることを筆者の先行研究を通じて明らかにする。

第 3 章では、本研究の心理実験を支える室内音響シミュレーションによる音場可聴化の技術について述べる。

第 4 章では、明瞭さに最適な初期反射音の音圧レベルを調査する実験 1 によって、聴取する音源、そして着目する明瞭さが異なる場合、初期反射音の音圧レベルの最適値が変わることを示す。

第 5 章では、人が音楽を明瞭と感じる要因を調査した実験 2 について述べ、音楽の明瞭さが「音色の明瞭さ」、「こもり具合」、「輪郭」、「定位の分かりやすさ」、「音像の広さ」、「響きの量の適切さ」、「聴き取りやすさ」といった複数の要因によって評価されることを明らかにする。また、市販音源と無響音源の聴取では、着目される明瞭さの評価要因が異なることも示す。

第 6 章では、市販音源の聴取時の明瞭さの各評価要因と、音場の物理量との関係を調査する実験 3 の結果を述べ、音楽の明瞭さが「音色の明瞭さ」、「音像の明瞭さ」、そして「音像の広さ」という三つの要素に分類されること、及びそれぞれの要素に対応する物理量を明らかにする。

第 7 章では、本研究の総括を述べる。

図 1.4.1 に本論文の研究フローを示す。

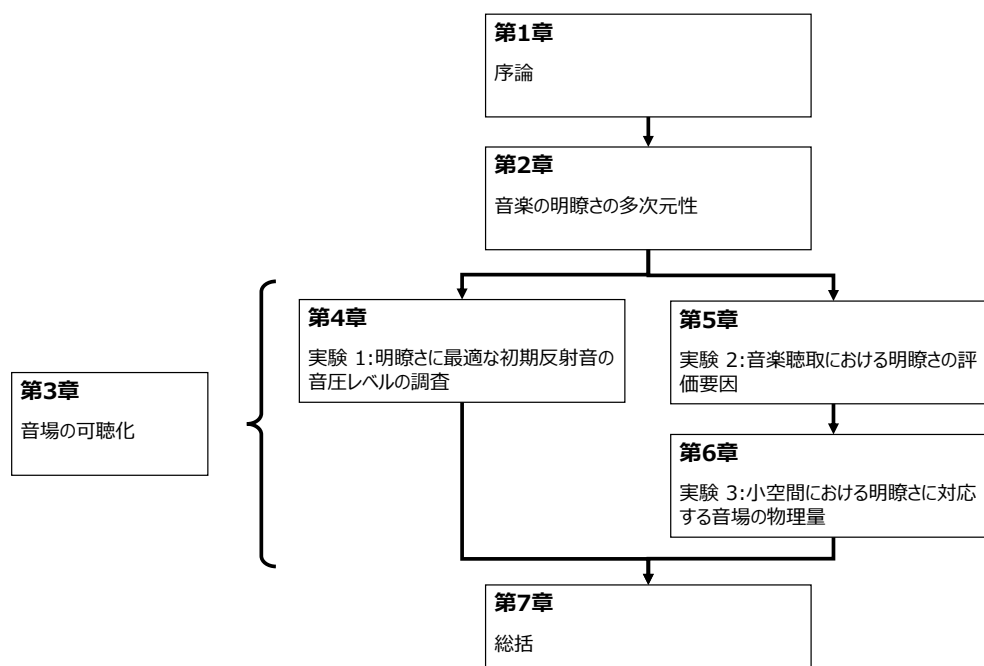


図 1.4.1 本論文の研究フロー

第2章

音楽の明瞭さの多次元性

音楽の明瞭さという言葉は抽象度が高く、具体的にどのような感覚を指しているのかが曖昧である。音声の場合は、発話内容が聴き取れるかどうかといった明確な基準に基づいて評価することが可能であり、実際に音声了解度のような指標が存在する(戸井田, 1995)。音楽においても、例えばボーカルの歌詞がどれだけ聴き取れるかによって明瞭さを判断することができるかもしれないが、ボーカルの無い楽曲の場合、何を基準に評価すべきか判断が難しい。

本章では、音楽の明瞭さが多次元的な感覚であることを示唆する先行研究について述べる。人が明瞭であると感じる要因が、仮に音量感、音質、音像の輪郭や定位感等複数ある場合に、これらを一括りにしてしまうと、個人によって指し示す感覚が異なってしまう。この場合、物理量との対応関係の解明は困難になる。明瞭さの評価要因を分解し、それぞれについて対応を見ていくことが必要である。

2.1 コンサートホールにおける音楽の明瞭さの多次元性

Kahle and Bruneau (1995) はコンサートホールの響きをスピーカーを用いて模擬し、ホールの印象評価実験を実施した。この研究により、コンサートホールの印象は響き (Reverberance) と空間 (Spaciousness) の二つの印象に大別されることが示された。また、音楽の明瞭さについても、これらの印象に応じて二つの異なる側面があるとした。一つは音色の明瞭さ、もう一つは同時に演奏される複数の音の区別がつく明瞭さであり、前者を水平的な明瞭さ、後者を垂直的な明瞭さと呼んで区別している。

これを受け、Beranek (2004) は、音楽の明瞭さを以下のように定義している。

The terms “definition” and “clarity” are synonyms for the same musical quality. They name the degree to which a listener can distinguish sounds in a musical performance. Definition is discernable in two forms: horizontal, related to tones played in succession, and vertical, related to tones played simultaneously.

“Definition” と “Clarity” という言葉は、共に音楽的な質を表す同義語である。これらは、聴取者が演奏における音を知覚、識別できるかどうかの尺度である。Definition は 2 つの形態に区別できる。1 つは水平的な明瞭さであり、これは連続して演奏される音と関係している。もう 1 つは垂直的な明瞭さであり、これはある瞬間に同時に演奏される音に関係するものである。(和訳筆者)

Beranek によれば、明瞭さは「聴取者が演奏における音を知覚、識別できるかどうかの尺度」である。明瞭さは二つに分けられ、「水平方向の明瞭さ (Horizontal definition)」は連続して演奏される音が区別できるか、もう一つの「垂直方向の明瞭さ (Vertical definition)」は同時に演奏される音が区別できるか、と定義されている。これらの水平方向、垂直方向は、譜面 (五線譜) を想起すると理解しやすい。譜面上の水平方向は時間軸を表し、垂直方向はある瞬間に同時に鳴る音を示している。

Beranek が指摘する水平方向の明瞭さは、ある瞬間に鳴る音が前の音の残響により影響され、濁ったり聞き取りにくくなるか否か、すなわち音色や響きの明瞭さを意味する。図 2.1.1 に、Beranek (2004) による、ある瞬間に鳴る音が前の音の残響によってマスクされる現象の概念図を示す。図の横軸は時間を、縦軸は音の強さを表し、斜線は響きの減衰を示している。上部の (A) では、残響が長い音場において、最初に発せられた音の響きが二つ目の音をマスクしている様子が示されている。一方、(B) は残響が短い音場において、最初に発せられた音の響きが二つ目の音よりも小さく、マスクされていない状況を示している。

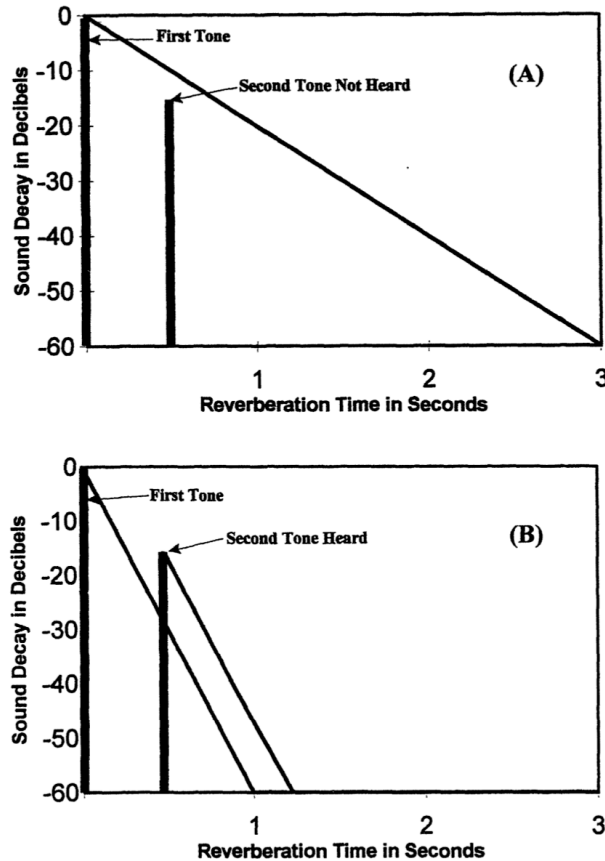


図 2.1.1 前の音の残響によって後続の音がマスクされる様子を示す概念図 (Beranek, 2004)。横軸は時間、縦軸は音の強さを、斜線は響きの減衰を示す。(A) は残響が長い音場において、最初に発せられた音の響きが二つ目の音をマスクする状況を示し、(B) は残響が短い音場において、最初の音の響きが二つ目の音よりも小さく、マスクされない状況を示している。

垂直方向の明瞭さは、ある瞬間に同時に演奏されている複数の楽器を区別できるかどうか、すなわち空間的な明瞭さや音像の明瞭さと解釈できる。

一方、Kahle はコンサートホールにおけるこれら二つの明瞭さの評価が高い相関を持ち、区別することに意味がないとも指摘している。これは、残響時間が長く、初期反射音の音圧レベルが低い音場における比較では、二つの明瞭さの評価が似たり寄ったりになる結果に基づく。しかし、初期反射音の影響が顕著な小空間や、異なる音源を聴取する場合に、これらの感覚が同様に評価されるかどうかは検証が必要である。

2.2 音楽再生システムにおける音楽の明瞭さの多次元性

2.2.1 ITU-R 勧告 BS.1116 における明瞭さの記述

ITU-R 勧告の BS.1116-3 (ITU-R, 2015) には、「劣化が少ない高品質のオーディオシステムの主観音質評価」のための聴取環境について、推奨される部屋の大きさや形状等が記載されている。ま

た、モノラル再生システム、ステレオ再生システム、マルチチャンネル再生システム、そして高さ方向のスピーカーを含むより高度な再生システムの評価項目として、明瞭さに関するいくつかの異なる感覚について記載されている。

前方の音像の質 (Front image quality)

マルチチャンネル再生システムの評価項目として、前方の音像の質 (Front image quality) が、以下のように定義されている。

This attribute is related to the localization of the frontal sound sources. It includes stereophonic image quality and losses of definition.

前方の音像の質の属性は、正面の音源の定位に関連している。これは、ステレオ音像の質と、明瞭さの劣化を含む。(和訳筆者)

ここでは、“losses of definition” という、知覚される音像の劣化に関する言及が見られる。

音色の質 (Timbral quality)

より高度な再生システムの評価項目として、音色の質 (Timbral quality) が以下のように定義されている。

The attribute of timbral quality may be described by two sets of properties:

- The first set of timbral properties is related to the sound colour, e.g. brightness, tone colour, coloration, clarity, hardness, equalization, or richness.
- The second set of timbral properties is related to the sound homogeneity, e.g. stability, sharpness, realism, fidelity and dynamics. These properties may be descriptive of the timbre of the sound, but may also be descriptive of other characteristics of the sound.

音色の質に関する属性は、以下の 2 つの特性の組によって記述される：

- 1 つ目は、音色に関連するものである。明るさ、音色、カラーレーション、明瞭さ、音の硬さ、音の均質さ、または豊かさを含む。(和訳筆者)
- 2 つ目は、音の均質性に関連する。音の安定性、鋭さ、現実感、忠実度とダイナミクスを含む。これらの特性は、音の音色を表すものであってもよいが、音の他の特性を説明するものであってもよい。(和訳筆者)

ここでは、音色に関する感覚としての “clarity” への言及が見られる。

定位の質 (Localization quality)

高さ方向のスピーカーを含む、より高度な再生システムの評価項目として、定位の質 (Localization quality) が以下のように定義されている。

This attribute is related to the localization of all directional sound sources. It includes stereophonic image quality and losses of definition. This attribute can be separated into horizontal localization quality, vertical localization quality and distant localization quality. In case of the test with accompanying picture, these attributes can be also separated into localization quality on the display and localization quality around the listener.

定位の質の属性は、全ての指向性を持つ音源の定位に関連するものである。これは、ステレオ音像の質と、明瞭さの劣化を含む。この属性は、水平的な定位の質と、垂直的な定位の質、そして距離の定位の質に分けることができる。映像を伴う評価の場合、これらの属性はまたディスプレイ上の定位の質と、聴取者周辺の定位の質とに分離することもできる。(和訳筆者)

ここでは、明瞭さの劣化、そして水平的な定位の質と、垂直的な定位の質、そして距離の定位の質に関する言及が見られる。

2.2.2 IEC 60268 における明瞭さの記述

IEC 60268 (1997) では、スピーカーの評価項目の一つである音像の定位感 (Image localization) の中に、以下のような明瞭さに関する記述が見られる。

- Image localization: perceived spatial location of a reproduced sound source. The image may be well defined or blurred.
 - Image stability: perceived location of the reproduced sound source, may change with pitch, loudness, or timbre. It may also change as a function of listener position, head rotation, or other normal movements. If these effects are small, the image will be stable.
- 音像の定位：再生される音源の知覚される空間的な定位。音像が明確に定義されているか、またはぼやけているかである。(和訳筆者)
 - 音像の安定性：再生される音源の知覚される定位は、ピッチ、ラウドネス、または音色によって変化する可能性がある。聴取者の位置、頭の回転、または他の通常の動きの関数として変化することもある。これらの効果が小さい場合、音像は安定する。(和訳筆者)

2.2.3 その他の文献における明瞭さの記述

Nakayama (1967); Nakayama et al. (1971) は、無響室におけるマルチチャンネル再生環境の違いが聴感印象に与える影響を一連の研究で調査した。1 から 8 チャンネルまでのスピーカー数、およびスピーカー配置の違いを含めた 13 の再生環境を用い、被験者に対して類似度と好ましさの評価を求めた。多次元尺度法による類似度データの分析の結果、マルチチャンネル再生環境の評価は、Fullness、Clearness、Depth の 3 つの属性によることを明らかにした。

Letowski (1989) は、音楽聴取において知覚される要素感覚を分類し、MURAL (MULTilevel auditoRy Assessment Language) を提案した。図 2.2.1 は、Letowski による MURAL の図を示す。知覚される音像 (AUDITORY IMAGE) は、まず音色 (TIMBRE) と空間印象 (SPACIOUSNESS) に分けられ、区別できるかどうか (DISTINCTIVENESS) は、音色と空間印象にまたがる感覚であると表されている。音色に関する区別ができるかどうかについてを、音質の明瞭さ (CLARITY OF SOUND TEXTURE) および密度 (BLEND) の 2 つに分類し、空間印象に関する区別できるかどうかについてを、方向感 (DIRECTIONAL SENSITIVITY) としている。

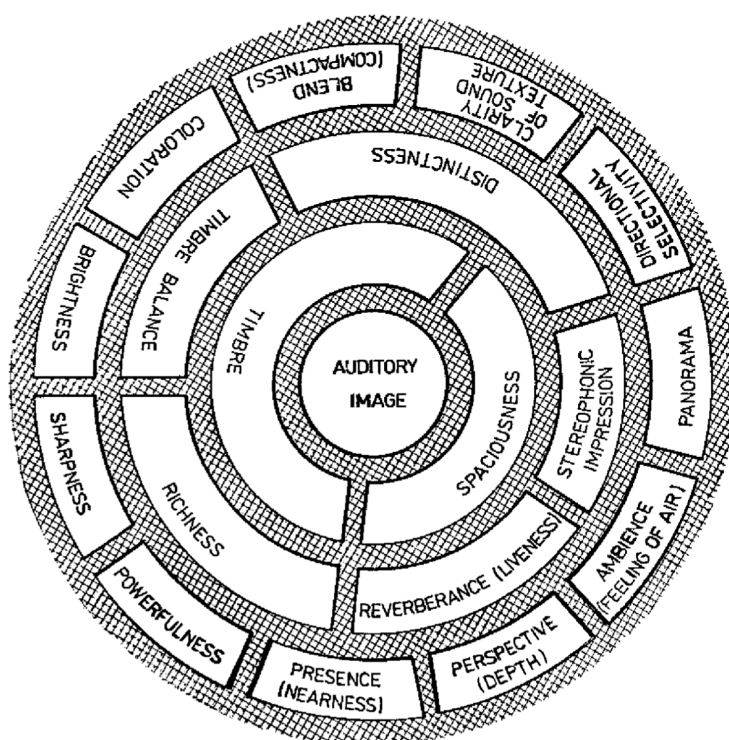


図 2.2.1 Letowski (1989) による MURAL の図

Guastavino and Katz (2004) は、様々な種類の音源の聴取において、1 次元、2 次元、3 次元のスピーカー配置を含む 3 種類の音楽再生システムの評価を被験者に求めた。実験では、各再生システムの音響的な質について口頭で自由に表現を求めた。得られた言葉の分析の結果、再生シス

テムの音響的な質は主に 6 つの感覚 (Presence、Coloration、Readability、Timbre、Localization、Stability of the image) で構成されていることを明らかにした。更に、6 つの感覚について被験者に再生システムの詳細な評価を求めた。ここで、Guastavino は Coloration の判断を「こもった - 明瞭な (Muffled - Clear)」という尺度で求めている。また、聴き取りやすさである Readability の判断を「ぼやけた - はっきりした (Poorly defined - Well defined)」という尺度で求めている。更に、この 2 つの感覚は、再生システムや聴取音源によって区別されることも明らかにした。この研究結果は、音楽再生システムの評価において、音色の明瞭さと空間的なぼや具合とが明確に区別されることを示している。

2.3 筆者による先行研究

2.3.1 一次反射音が明瞭さに与える影響

筆者による先行研究 (Imamura et al., 2014) では、直接音の到来後 30 ms 以内に到来する一次反射音の到来方向と到来時間が、知覚される「音色の明瞭さ」と「音の輪郭」に与える影響を調査した。実験では簡易無響室を用い、モノラルの直接音を再生するための 1 台のスピーカーと、一次反射音を再生する 4 台のスピーカーを配置し、直接音に続く一次反射音のみが存在する音場の聴取環境を構築した。実験の刺激として、女性ボーカルを含むポップス楽曲の市販音源を使用した。図 2.3.1 は、簡易無響室におけるスピーカー配置と刺激の作成方法を示したダイアグラムを示す。

実験の説明変数として、一次反射音の到来方向 (前方、側方、後方)、および到来時間 (直接音のみ、遅延時間 10 ms、遅延時間 30 ms) の 2 つに着目した。刺激は被験者に 2 つずつ対として提示し、シェッフェの一対比較法に基づいて、各刺激ペアを比較し判断するよう求めた。実験の指示は被験者に対して以下の通り行われた。

- 音色の明瞭さ (Clarity) : どちらの刺激の音色が明瞭であるか。ある瞬間に再生されている音が前の音の残響等に影響され、濁っていたり聴こえづらくなっていないか。
- 音の輪郭 (Spatial Definition) : どちらの音の輪郭がはっきりしているか。ある瞬間に再生されている音が、同時に再生されている他の音と空間的に区別できるかどうか。

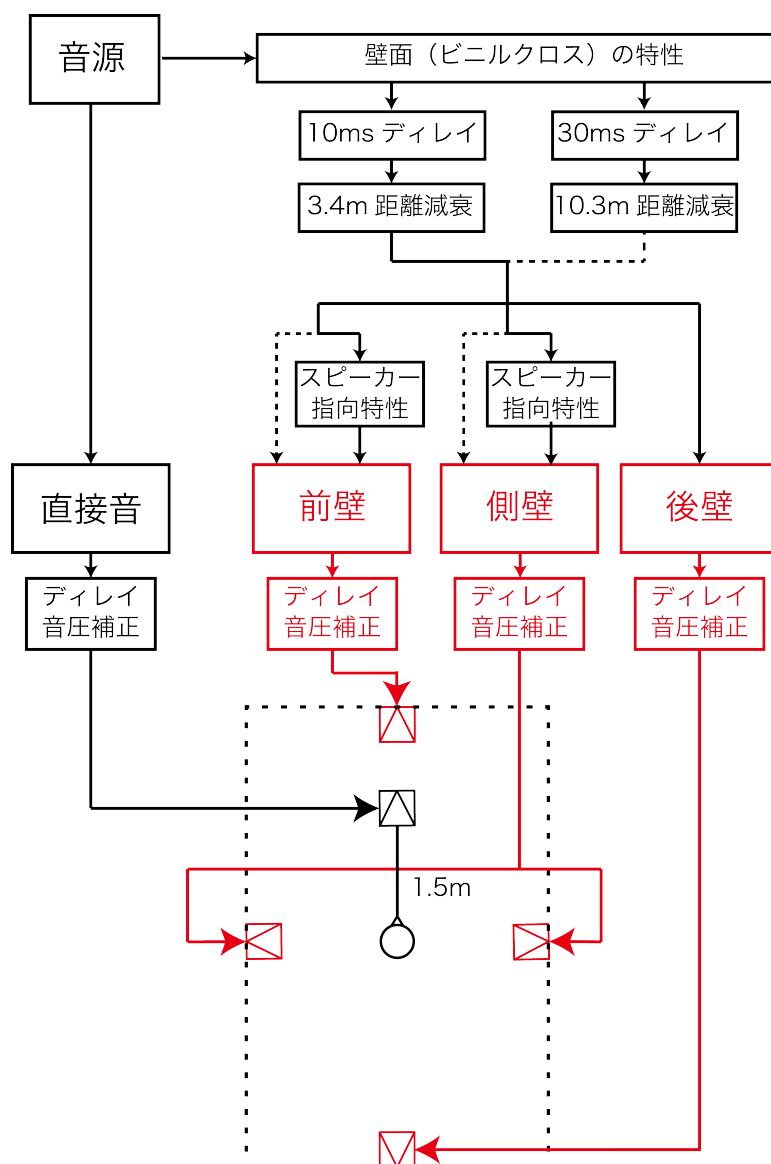


図 2.3.1 筆者による先行研究 (Imamura et al., 2014) で実施した簡易無響室における実験のスピーカー配置および刺激作成のダイアグラム

聴取実験から得られた結果を図 2.3.2 および図 2.3.3 に示す。音色の明瞭さおよび音の輪郭の両方において、一次反射音が明瞭さを有意に低下させることが分かる。一方、30 ms の遅延で一次反射音が前方から到来する場合にのみ、影響は小さい。Begault and Anderson (2001) は音源と同じ方向から到来する単発の反射音は、他の方向から到来する反射音よりも知覚されにくいと述べており、このためだと考えられる。更に、音の輪郭については到来方向による影響が顕著に表れた。一次反射音が 10 ms で到来し、聴取者の左右から来た場合に特に低く評価された。側方からの強い一次反射音が音の輪郭の低下を引き起こすことが明らかとなった。

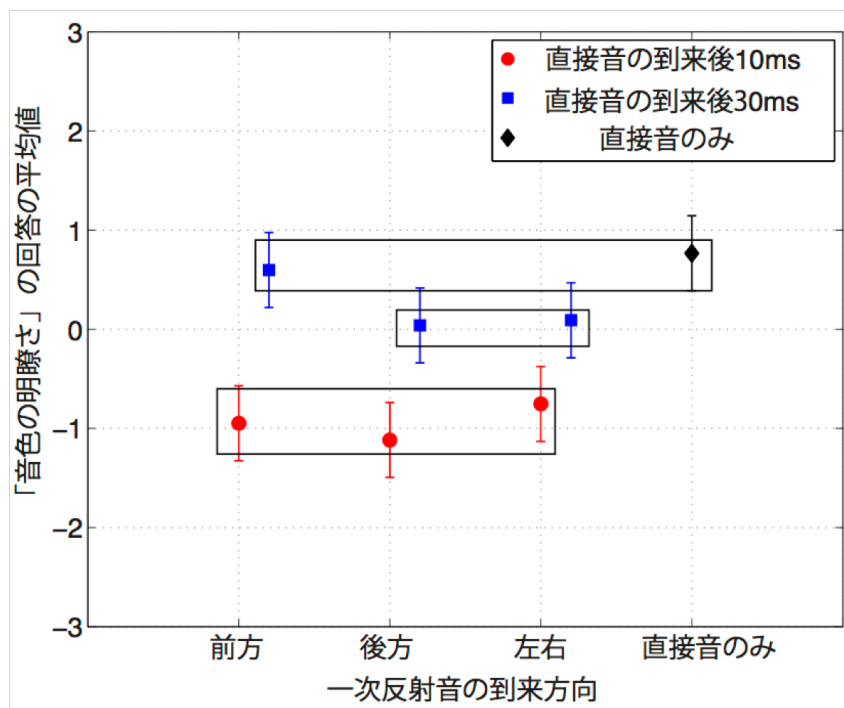


図 2.3.2 音色の明瞭さの評価の平均値と 95 %信頼区間。有意差が見られなかった刺激同士が黒の枠で囲まれている。

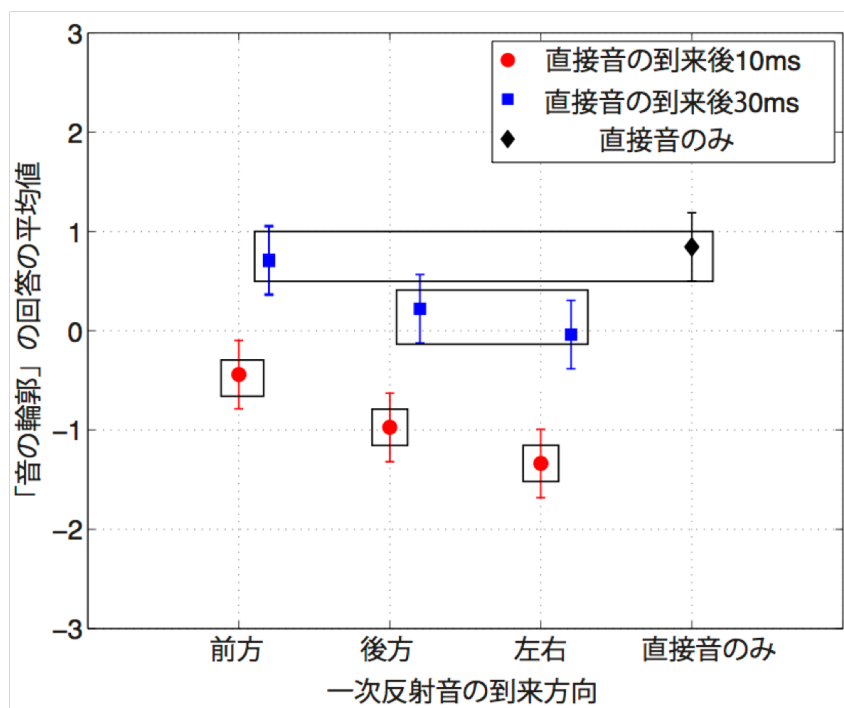


図 2.3.3 音の輪郭の評価の平均値と 95 %信頼区間。有意差が見られなかった刺激同士が黒の枠で囲まれている。

以上の結果から、明瞭さについて評価を求めても、音色の明瞭さに焦点を当てるか、空間的な音の輪郭に焦点を当てるかによって、評価の傾向が異なることが示された。ここで教示を行った音の輪郭は、Beranek (2004) の述べた垂直方向の明瞭さに相当する。一次反射音の到来方向が異なる小空間の場合、これらの感覚は有意に区別されることがこの研究で明らかにされた。

2.3.2 小空間における音楽の明瞭さの多次元性

筆者による先行研究 (今村他, 2016) では、小空間における吸音処置の違いが明瞭さに与える影響を調査した。幾何音響シミュレーションソフトウェア CATT Acoustics V8 を使用し、直方体の空間 (幅 5 m × 奥行き 6 m × 高さ 5 m) におけるステレオでの音楽聴取を、聴取位置でのバイノーラル信号を生成することによって可聴化した。吸音面を「床および天井」、「側壁」の 2 種類、それぞれの壁面の吸音率を 0.01、0.25、0.50、0.75、0.90 の 5 段階に変化させ、合計 10 種類の刺激を作成した。表 2.3.1 は実験で使用した 10 種類の刺激を、図 2.3.4 は可聴化を行った部屋のイメージ図を示す。これらの刺激を、ヘッドホンを用いて 10 名の被験者に一対比較で提示した。被験者は提示されたペアの刺激の明瞭さがどの程度似ているかについて 7 段階で評価した。実験には女性ボーカルのポップスの市販音源を用いた。

表 2.3.1 実験で使用した 10 種類の刺激のそれぞれの壁面に設定した吸音率。最上段は刺激名を示す。

	floor1	floor25	floor50	floor75	floor90	side1	side25	side50	side75	side90
床・天井	0.01	0.25	0.50	0.75	0.90	0.90	0.90	0.90	0.90	0.90
側壁	0.90	0.90	0.90	0.90	0.90	0.01	0.25	0.50	0.75	0.90
前後壁	0.90	0.90	0.90	0.90	0.90	0.90	0.90	0.90	0.90	0.90

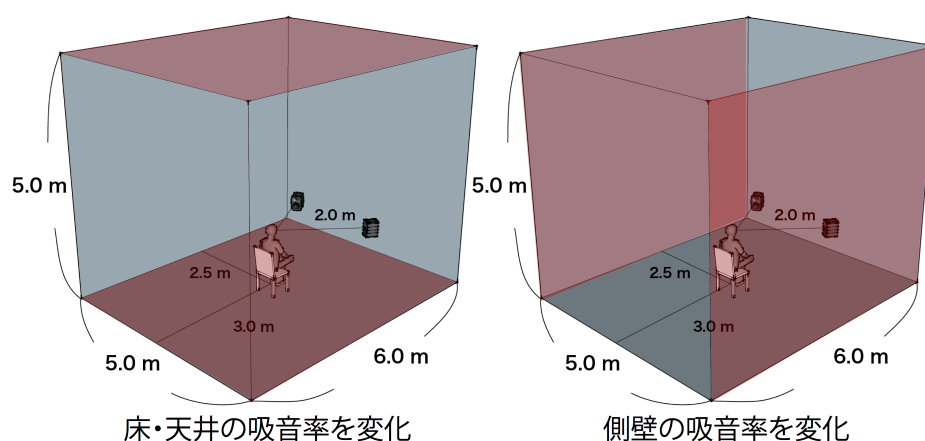


図 2.3.4 CATT Acoustics よりシミュレーションした部屋のイメージ図

聴取実験から得られたデータに対して、個々の被験者間での違いを考慮する多次元尺度法である INDSCAL (個人差尺度法) に基づく分析を行った。図 2.3.5 に示されるシンボル間の距離は、明瞭

さに関する印象の非類似度を表す。同一吸音率の刺激群は第 1 次元において類似していると評価されたが、第 2 次元では吸音面の位置による違いが顕著に反映された。8 名の被験者は両方の次元をほぼ等しい重み付けで判断していた。これにより、壁面の吸音率及び吸音が施される位置が異なる小空間を比較すると、音楽の明瞭さは少なくとも 2 つの異なる側面で区別されることが示された。

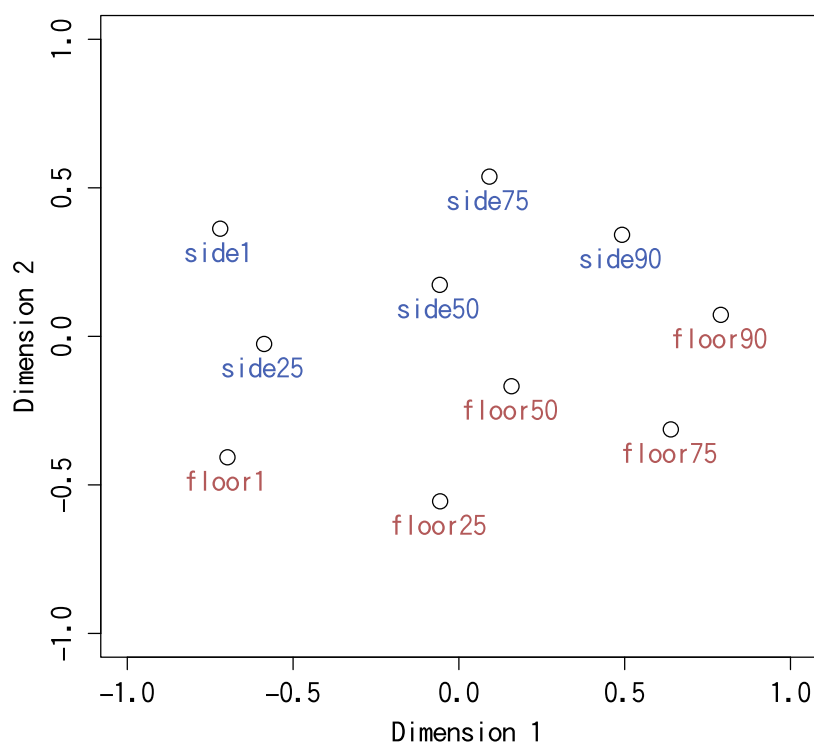


図 2.3.5 筆者による先行研究 (今村他, 2016) の INDSCAL を用いた分析結果。シンボル間の距離は、明瞭さに関する印象の非類似度を表す。

2.4 結論

先行研究では、コンサートホールにおける音楽の明瞭さは少なくとも音色の明瞭さ（水平的な明瞭さ）および音像の明瞭さ（垂直的な明瞭さ）に大別されることが主張されている。更に、音楽再生システムの評価に際しては、音楽の明瞭さに関連する多様な評価要因が多数の研究で言及されている。音色の質やカラーレーション、音像や定位感、方向感に関する明瞭さがその例である。

筆者の先行研究によって、直接音の到来後 30 ms 以内に到来する一次反射音の音圧レベルが高いほど、市販音源の聴取では音色の明瞭さと音の輪郭が低下することが明らかになった。更に、一次反射音の到来方向が評価に影響し、音色の明瞭さと音の輪郭は有意に区別されることが明らかにされた。

第3章

音場の可聴化

本研究では、人が感じる明瞭さの定量化のために、様々な音場における音楽聴取を模擬した音刺激を作成する必要がある。しかし、反射音の時間的構造や空間的分布の影響を詳細に調査するには、それらが複雑に影響しあって変化してしまう実空間の収録データを用いた可聴化は望ましくない。

本章では、反射音の特性を任意に制御するために採用した、羽入らによる仮想音源分布を用いた音場の可聴化技術 (羽入, 2002) とシミュレーションの手法について述べる。

3.1 仮想音源分布を用いた音場の可聴化

任意の音場での音楽聴取を模擬するには、いくつかの手段が存在する。実際の空間で演奏される音楽を直接収録する方法や、マイクロフォンやダミーヘッドで収録したインパルス応答に音楽信号を畳み込んで音刺激を作成し、その刺激をスピーカーやヘッドホンで提示する方法等が挙げられる。しかし、このように実際の音場で収録する手法では、室の形状が変わると反射音の時間的構造や空間的分布、音圧レベルが複雑に影響しあって変化する。これらの要因の影響を個別に調査するには、反射音の特性を任意に制御できる手法が求められる。

そこで、本研究では羽入らによる仮想音源分布を用いた音場の可聴化の原理 (羽入, 2002) (岸他, 2012) に基づく音場の可聴化技術を採用した。羽入らによる仮想音源分布を用いた音場の可聴化では、室内の全ての反射音を仮想空間内の 3 次元座標および音圧レベルの情報を持つ仮想音源として捉える。各仮想音源は再生環境のスピーカー配置に応じて分割され、時系列応答に合成されることで、受音点におけるスピーカーごとのインパルス応答が生成される。

図 3.1.1 は、仮想音源分布の概念図を示す。受音点に到来する反射音は、受音点から見て壁面の向こう側にある音源から来た音と考えることができる。この仮想音源は一つの反射音に対して一つずつ対応し、反射音の到来方向と受音点との距離を表す。この距離が遠いほど反射音の到来が遅れることを意味する。仮想音源分布は、特定の空間で発生する反射音群の空間的・時間的特徴を表す。

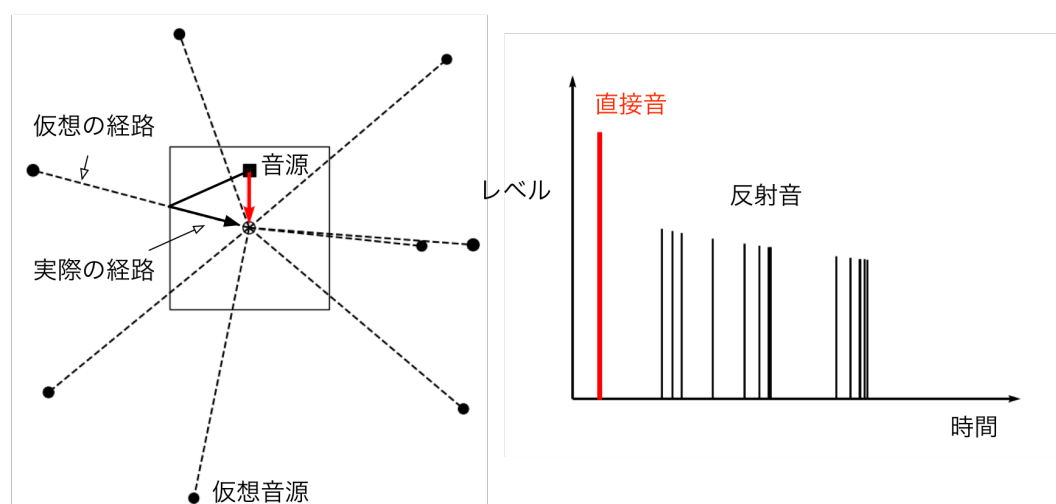


図 3.1.1 室内の反射音群の空間的・時間的特徴を表す仮想音源分布の概念図

次に、仮想音源分布に基づくインパルス応答の生成について述べる。音速を c 、室の残響時間を T としたとき、半径 cT の仮想的な球内に N_{total} 個の仮想音源を一様に配置する。室内での音源発生から t 秒後までに生じる仮想音源数 $N_r(t)$ は、式 (3.1) から求められる。ここで、 h は羽入らにより定式化された室形状の複雑さ係数 (羽入・星, 2008)、 V は室容積である。

$$N_r(t) = \frac{h(ct)^3}{V} \quad (3.1)$$

複雑さ係数 h は、室形状の複雑さと反射音の密度や音圧レベルとの関係を定式化したものであり、式 3.2 から求めることができる。 V に対して S が大きくなり、室形状が複雑になるほど大きな値となる。室形状が相似であれば V が異なっても同じ値を示す。

$$h = \frac{S^3}{64V^2} \quad (3.2)$$

残響時間 T を持つ室内における仮想音源数を求めたい場合は、式 (3.1) の t に残響時間 T を代入すれば良い。残響時間 T は Sabine の残響式 (1.1) から求める。残響式より、残響時間 T は室容積 V 、室総表面積 S と室の平均吸音率 $\bar{\alpha}$ によって求められる。

仮想音源の生成には、乱数を用いて n 個の仮想音源全てに 3 次元座標 x_i 、 y_i 、 z_i を設定する。仮想の球内の原点、つまり受音点から各仮想音源までの距離 r_i は $\sqrt{x_i^2 + y_i^2 + z_i^2}$ から求められ、各仮想音源に対応した反射音の到来時刻 t_i は r_i/c となる。次に、個々の仮想音源に式 (3.3) で表す音響出力 W_i を与える。ここで、 W_0 は実音源の音響出力である。

$$W_i = \frac{4\pi W_0}{3h} \quad (3.3)$$

各仮想音源の音響出力 W_i および到来時刻 t_i を元に、全ての仮想音源について距離減衰 $4\pi(ct_i)^2$ 、および室の平均吸音率 $\bar{\alpha}$ を考慮して反射音のエネルギー I_i を計算する。 I_i は式 (3.4) から求める。ここで、平均吸音率 $\bar{\alpha}$ は実空間では周波数によって異なるため、 I_i も周波数依存する。周波数帯域ごとに $\bar{\alpha}$ を設定すれば周波数別のインパルス応答が生成でき、壁面の素材などにより異なる響きを生成できる。しかし、本研究では平均吸音率 $\bar{\alpha}$ の周波数依存は考慮せず、全周波数帯域で同じエネルギーを持つものとして扱った。また、空気減衰の影響は検討から除外した。

$$I_i = \frac{W_i}{4\pi(ct_i)^2} \exp\left(\ln(1 - \bar{\alpha}) \frac{cS}{4V} t_i\right) \quad (3.4)$$

以上をまとめると、羽入らによる仮想音源分布を用いた音場の可聴化の原理では、音速 c 、室容積 V 、室総表面積 S 、室の平均吸音率 $\bar{\alpha}$ をパラメータとして室の仮想音源数 N_{total} を算出し、乱数を基に仮想音源数 N_{total} 分の仮想空間内の座標を算出していることが分かる。式 3.1 より、室容積 V が小さいほど反射音数が多くなる。また、式 3.4 より、平均吸音率 $\bar{\alpha}$ が等しい場合は、室容積 V が小さいほど反射音レベルの減衰が早くなる。また、室の複雑さが増すと反射音数が増す一方で、個々の反射音のレベルが減少することが分かる。室容積 V と室の複雑さが等しい場合には、平均吸音率 $\bar{\alpha}$ が大きいほど反射音レベルの減衰が早い。実空間における室容積、室形状、平均吸音率の違いによる反射音の時間的特徴の変化が式によって再現されていることが分かる。

最後に、全ての仮想音源のエネルギー I_i および到来時刻 t_i を考慮して時系列応答に合成すること

によって、受音点におけるインパルス応答を求める。この際、再生を行う環境のスピーカー配置に合わせて全ての仮想音源を分割する。図 3.1.2 は、5 チャンネルの再生環境に合わせて仮想音源を分割した様子を示す (岸他, 2012)。上部の図の円は各仮想音源を示す。分割した各領域に存在する仮想音源を時系列に合成して、方向別の応答を作成する。各仮想音源のスピーカーへの割り当ては、それぞれの仮想音源の 3 次元座標から最も近いスピーカーを選択すれば良い。

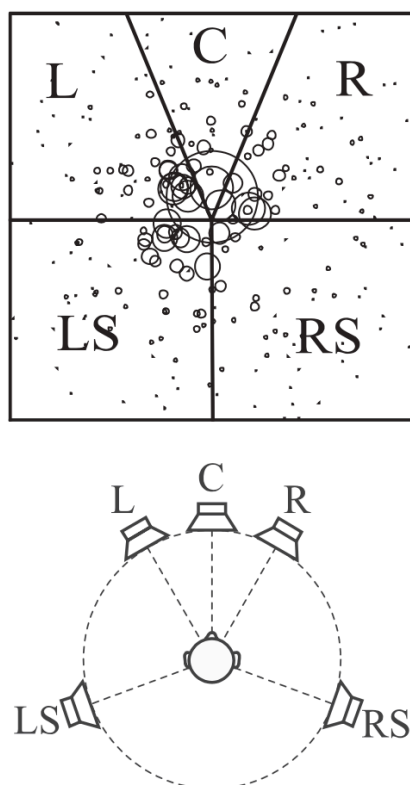


図 3.1.2 5 チャンネルの再生環境に合わせて仮想音源を分割した様子を示す (岸他, 2012)。上部の図の円は各仮想音源を示す。

本研究では、再生システムとして東京藝術大学千住キャンパススタジオ B における 22.2 マルチチャンネル音響システムを使用した。本システムは、ITU-R BS.2051-1(ITU-R, 2017)「番組制作における高度音響システム」において、チャンネル配置の 1 つとして標準化されている Sound System H のスピーカー配置に準拠している。図 3.1.3 は本学スタジオ B の写真を、図 3.1.4 はスタジオ B におけるスピーカー配置を示す。なお、本研究では 2 台のサブウーファーは使用していないため、実験に用いたスピーカーは 22 台であった。スタジオ B は、幅 6.8 m、奥行き 6.8 m、高さ 4.5 m であり、残響時間は約 0.27 秒である。聴取位置からスピーカーまでの距離は約 2.5 m であった。使用したスピーカーは、22 台全て KS Digital 社アクティブ同軸モニタースピーカーの C5-Coax である。 $-19 \text{ dBFS}_{\text{rms}}$ のピンクノイズを規定状態の実験系に接続し、聴取位置に設置した騒音計でチャンネル毎の A 特性音圧レベルが 79 dB となるように再生音圧を調整した。



図 3.1.3 本学スタジオ B

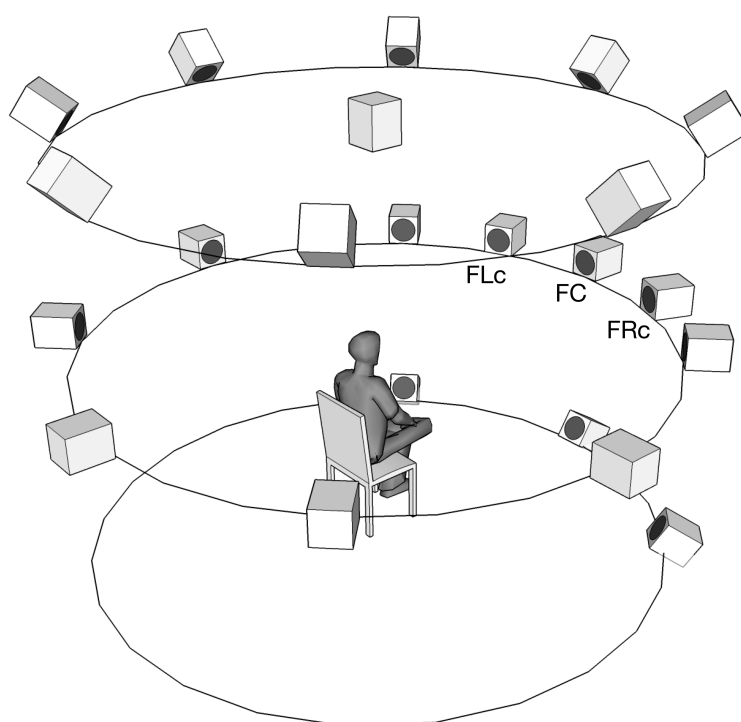


図 3.1.4 本学スタジオ B におけるスピーカー配置。ITU-R BS.2051-1(ITU-R, 2017)「番組制作における高度音響システム」における Sound System H のスピーカー配置に準拠している。本研究において直接音の再生に使用した 3 台のスピーカーについては、ID を記載している。

3.2 本研究で用いたシミュレーション手法

本研究では、前述した羽入らによる仮想音源の可聴化の手法に基づき、実験ごとに任意の音場の仮想音源分布を作成した。仮想音源分布は前述した本学スタジオ B のスピーカー配置に基づき、22 の方向別に分割され、方向別の評価音場のインパルス応答が生成された。インパルス応答を聴取音源に畳み込むことで実験刺激を作成した。インパルス応答の作成および畳み込みにはプログラミング言語 Python3.5 を使用した。インパルス応答の生成は前節で述べた手法に則しているが、本研究においては到来時刻が 80 ms までの初期反射音を表す仮想音源に関しては、3 次元座標を乱数ではなく、恣意的に偏らせることで方向性を持たせた。評価音場のインパルス応答の生成に使用した Python3.5 のソースコードを付録 A に記載する。

本節では以下の項目について解説を行う。

- 初期反射音の到来方向の設定
- 聴取音源とインパルス応答の畳み込み

3.2.1 初期反射音の到来方向の設定

前節で解説した手法では、乱数を用いて全ての仮想音源の 3 次元座標 x_i 、 y_i 、 z_i を設定した。しかし、本研究においては到来時刻 t_i が 80 ms よりも短い仮想音源については、到来方向の変化が聴感印象に与える影響の調査を行う必要がある。そこで、仮想音源を分布させる仮想的な球の水平面を放射状に 16 等分し、乱数による 3 次元座標の生成範囲を 16 のブロックから選択可能にした。分割した 16 のブロックから任意のブロックを選択することで、選択されたブロックの範囲にのみ仮想音源が生成されることとなる。本研究においては、垂直面の違いは検討せず、仮想音源の到来方向の違いは水平方向のみを考慮した。図 3.2.1 の左図は、仮想的な球の水平面を放射状に 16 等分し、球を上部から俯瞰した様子を示す。番号はブロックの ID を示す。例として、4 番、5 番、12 番、13 番を選択し、側方に仮想音源が分布するよう設定した様子を選択範囲を黒く塗ることで表す。図 3.2.1 の右図は、生成範囲を設定された t_i が 80 ms よりも短い仮想音源が 3 次元の仮想空間内で分布している様子を示す。

なお、羽入らによる手法では、受音点から各仮想音源までの距離 r_i は、生成された 3 次元座標 x_i 、 y_i 、 z_i を基に算出する。しかし、本研究では先に N_{total} 個の距離 r_i を乱数を用いて生成した。そのうち、 t_i 、つまり r_i/c が 80 ms よりも短い仮想音源の 3 次元座標のみ、前述した手法で方向性を持たせ、80 ms よりも長い仮想音源については方向性を持たせず生成した。座標の設定には極座標系を用い、乱数を用いて 2 個の偏角 θ 、 ϕ を生成した。以上より、仮想音源の生成範囲として選択するブロックの数に関わらず、生成される仮想音源数および受音点から各仮想音源までの距離は等し

くなる。そのため、初期反射音の到来方向が異なっても聴取する響きの音圧レベルは一定となる。

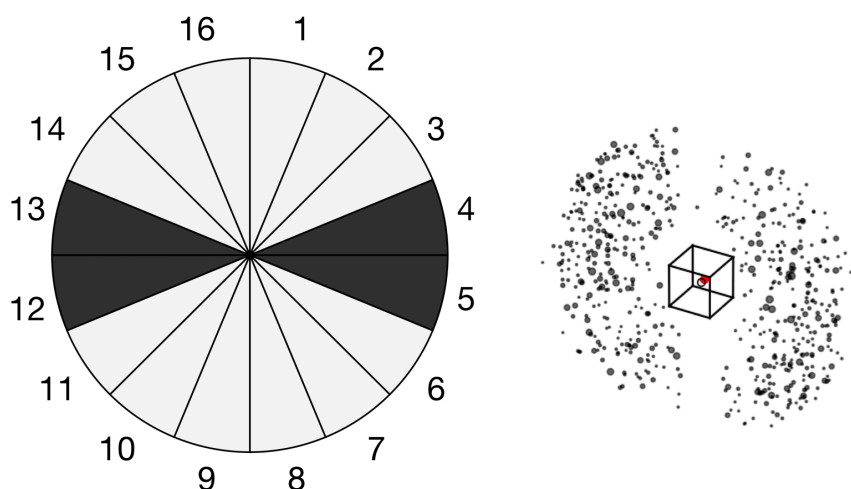


図 3.2.1 左図は、仮想的な球の水平面を放射状に 16 等分し、球を上部から俯瞰した様子を示す。例として、4 番、5 番、12 番、13 番を選択し、側方に仮想音源が分布するように設定した場合の 3 次元の仮想音源分布を、右図に示す。

3.2.2 生成したインパルス応答の畳み込み

羽入らによる原理では全ての反射音を仮想音源として捉えたが、直接音に関しても同様に仮想音源として考え、インパルス応答を生成すれば良い。受音点から直接音までの距離を r_{direct} とした時、実音源の音響出力 W_0 に対して距離減衰 $\sqrt{W_0/4\pi r_{\text{direct}}^2}$ を考慮し、直接音に音圧レベルを与える。本研究では、全ての実験で音響出力 W_0 を 1 とした。次に、直接音の到来時刻 t_{direct} を r_{direct}/c から算出すれば良いが、本研究では直接音の到来時刻に対する反射音の到来時刻を制御する必要がある。各仮想音源の到来時刻は受音点に対して求められているため、全ての実験において直接音の到来時刻は 0 秒としてインパルス応答に合成した。

次に、直接音を再生するスピーカーを選択する。本研究では、直接音はモノラル音源およびステレオ音源の 2 種類を使用した。モノラル音源は、本学スタジオ B におけるスピーカー配置を示す図 3.1.4 の FC スピーカーから、ステレオ音源は FLc および FRc スピーカーから再生した。

第4章

実験1：明瞭さに最適な初期反射音の音圧レベルの調査

小空間は室容積が小さく、初期反射音が聴感印象に与える影響は大きい。この場合、ホールのように初期反射音の音圧レベルが高ければ高いほど明瞭だとは知覚されない可能性がある。

本章では、小空間での音楽聴取において、明瞭さに最適な初期反射音の音圧レベルを調査した実験1について述べる。実験に際して次の3つの仮説を立て、検証した。

1. 音楽の明瞭さに最適な初期反射音の音圧レベルは、初期反射音の時間的・空間的特徴によって異なる
2. 着目する明瞭さの違いによって、音楽の明瞭さに最適な初期反射音の音圧レベルは異なる
3. 聴取を行う音源の違いによって、音楽の明瞭さに最適な初期反射音の音圧レベルは異なる

実験では、被験者に対し初期反射音の時間的・空間的特徴の異なる複数の音場での音楽聴取を求め、最も明瞭だと感じる値まで初期反射音の音圧レベルの調整を求めた。聴取音源としてステレオの市販音源、およびモノラルの無響音源の2種類の音楽を使用し、評価の基準としては音色の明瞭さ、および音像の明瞭さの2つに着目を求めた。それぞれの評価において、どの程度の初期反射音の音圧レベルが明瞭とされたかを分析した。

なお、論文内では「初期反射音の音圧レベル」を簡潔に表現するため、「初期反射音レベル」と記述する。また、本実験で言及する初期反射音とは、直接音の到来後 30 ms～80 ms の間に到来する反射音を指す。

4.1 実験方法

4.1.1 調整法

調整法は精神物理学的測定法の一つで、被験者または実験者が刺激を自由に变化させながら、刺激閾や弁別閾、主観的等価点などを測定する方法である(難波・桑野, 1998)。本実験では、調整法に基づき、被験者が最も明瞭だと感じる値まで自ら初期反射音レベルを調整することを求めた。調整法は被験者にとって分かりやすく自然な方法であるため、特に主観的等価点の測定法として優れていると言われており、これを採用した。主観的等価点とは、任意の基準の刺激量に対して心理的に等しいと感じる刺激量のことである。

本来、調整法で主観的等価点を測定する際、被験者には基準となる刺激を提示する。本実験であれば、被験者に初期反射音レベルの調整を求める際、基準として「最も明瞭である音場の初期反射音レベル」を提示し、基準と同等になるよう調整を求めることが望まれる。しかし「最も明瞭である音場の初期反射音レベル」は未知のため、「被験者本人の中にある明瞭さ」を意識することを求め、基準を提示せずに初期反射音レベルの調整を求めた。調整の手順については後述する。

4.1.2 実験刺激の作成

羽入らによる仮想音源分布を用いた音場の可聴化の手法に基づき(第3章を参照)、室容積 V が 60 m^3 、および 400 m^3 の2種類、初期反射音の到来方向が7種類の合計14種類の音場を再現するインパルス応答を生成した。室容積は1辺が4mほどのリスニングルーム、および7mほどのリスニングルームを想定した。どちらの室も複雑さ係数 h は等しいため、相似形である。室の平均吸音率 $\bar{\alpha}$ は0.4で統一した。

初期反射音の到来方向は、それぞれ全方向から到来するもの(ALL)、前方広範囲から到来するもの(W_F: Wide_Front)、後方広範囲から到来するもの(W_B: Wide_Back)、側方広範囲から到来するもの(W_SI: Wide_SIde)、前方から到来するもの(N_F: Narrow_Front)、後方から到来するもの(N_B: Narrow_Back)、側方から到来するもの(N_SI: Narrow_SIde)の7種類とし、初期反射音に該当する仮想音源の生成範囲を限定した。ALLの音場は、初期反射音の生成範囲を分割した16ブロック全て(図3.2.1を参照)、W_Fは1~4番および13~16番、W_Bは5~12番、W_SIは3~6番および11~14番、N_Fは2~3番および14~15番、N_Bは6~7番および10~11番、N_SIは4~5番および12~13番のブロックに限定した。直接音の距離減衰は考慮せず、直接音のエネルギーは音響出力 W_0 と等しい1としてインパルス応答に合成した。

なお、30ms以内の反射音および80ms以降の後期残響音はインパルス応答から除外し、初期反射音のみが聴取可能な状態にした。単発の一次反射音が音楽の明瞭さに与える影響については、筆者による先行研究(Imamura et al., 2014)によって明らかになっている。さらに、音場の反射音は

組み合わせによって複雑に影響し合うため、本実験では初期反射音のみの影響に着目することとし、これらを除外した。

表 4.1.1 に、生成した 14 種類の音場の名称および各生成パラメータを示す。音場の名称は「室容積 到来方向」となっており、室容積 V が 60 m^3 のものを S (Small)、室容積 V が 400 m^3 のものを L (Large) で示す。図 4.1.1～図 4.1.42 に、それぞれの音場の 3 次元仮想音源分布、2 次元仮想音源分布（仮想音場を上方から俯瞰した図）、そして反射音の時系列応答を示す。仮想音源分布の図における点は各仮想音源を示し、中央の立方体および四角は仮想の聴取空間を示す。見やすさのため、3 次元および 2 次元の仮想音源分布の図には、仮想音源の総数 N_{total} 個から半数をランダムに選択し、表示した。

表 4.1.1 本実験で生成した 14 種類の評価音場の名称および生成パラメータ

音場の名称	室容積 [m^3]	室総表面積 [m^2]	複雑さ係数 h	平均吸音率 $\bar{\alpha}$	初期反射音の到来方向
S ALL	60	94	3.6	0.4	全方向
S W_F	60	94	3.6	0.4	前方 (広)
S W_B	60	94	3.6	0.4	後方 (広)
S W_SI	60	94	3.6	0.4	側方 (広)
S N_F	60	94	3.6	0.4	前方 (狭)
S N_B	60	94	3.6	0.4	後方 (狭)
S N_SI	60	94	3.6	0.4	側方 (狭)
L ALL	400	340	3.6	0.4	全方向
L W_F	400	340	3.6	0.4	前方 (広)
L W_B	400	340	3.6	0.4	後方 (広)
L W_SI	400	340	3.6	0.4	側方 (広)
L N_F	400	340	3.6	0.4	前方 (狭)
L N_B	400	340	3.6	0.4	後方 (狭)
L N_SI	400	340	3.6	0.4	側方 (狭)

作成したインパルス応答を以下の 2 つの音源に畳み込み、実験刺激とした。

1. 市販音源 (ステレオ) : Norah Jones, “Come Away With Me” CD (2002) Track1 : Don't Know Why
2. 無響音源 (モノラル) : Archimedes Project : シロフォン (Hansen and Munch, 1991)

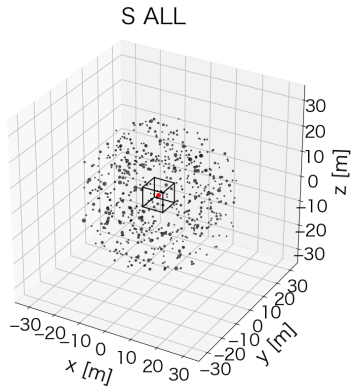


図 4.1.1 S ALL の音場の 3次元仮想音源分布

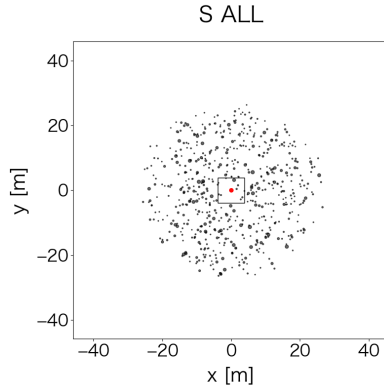


図 4.1.2 S ALL の音場の 2次元仮想音源分布

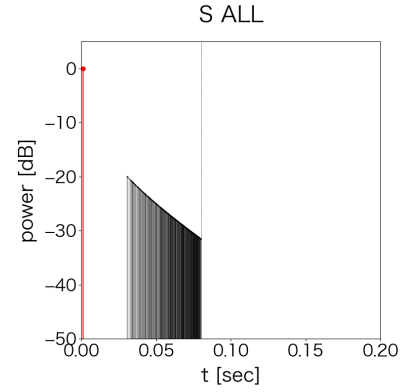


図 4.1.3 S ALL の音場の反射音の時系列応答

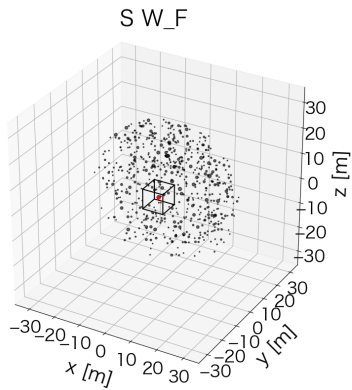


図 4.1.4 S W_F の音場の 3次元仮想音源分布

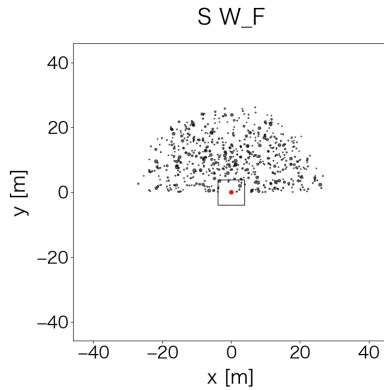


図 4.1.5 S W_F の音場の 2次元仮想音源分布

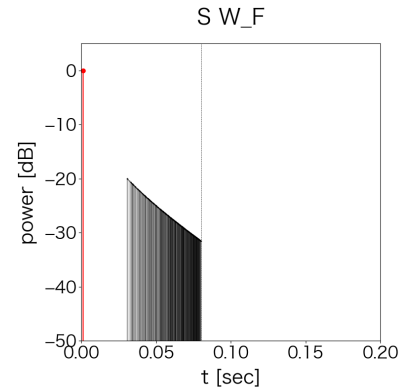


図 4.1.6 S W_F の音場の反射音の時系列応答

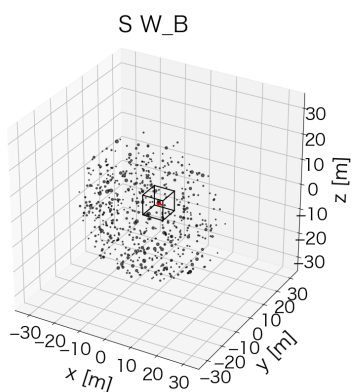


図 4.1.7 S W_B の音場の 3次元仮想音源分布

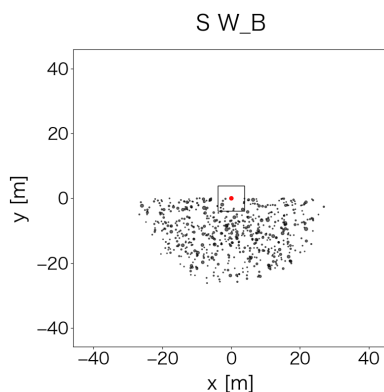


図 4.1.8 S W_B の音場の 2次元仮想音源分布

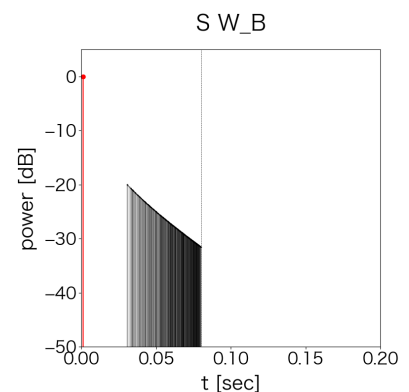


図 4.1.9 S W_B の音場の反射音の時系列応答

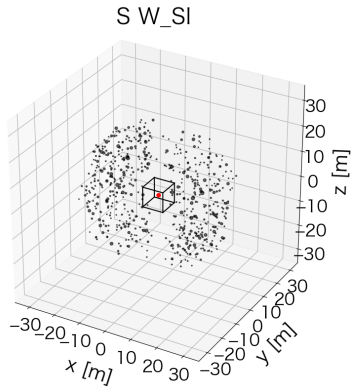


図 4.1.10 S W_SI の音場の 3次元仮想音源分布

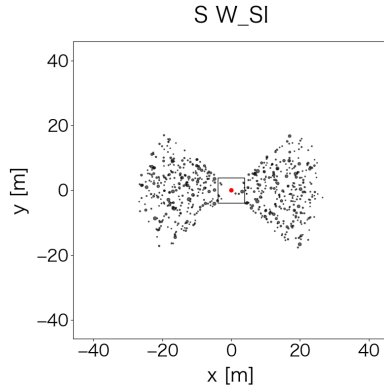


図 4.1.11 S W_SI の音場の 2次元仮想音源分布

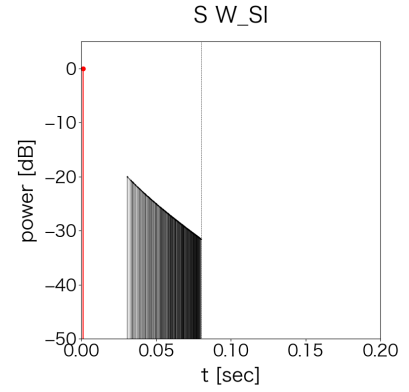


図 4.1.12 S W_SI の音場の反射音の時系列応答

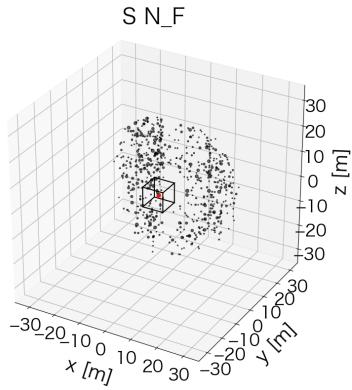


図 4.1.13 S N_F の音場の 3次元仮想音源分布

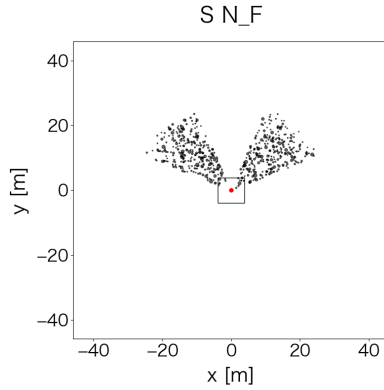


図 4.1.14 S N_F の音場の 2次元仮想音源分布

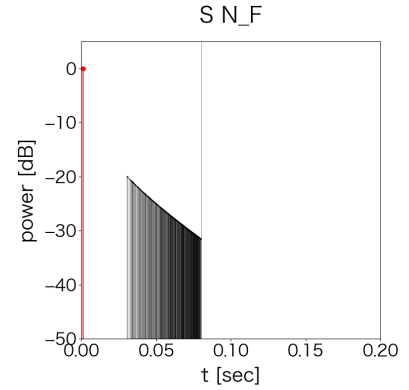


図 4.1.15 S N_F の音場の反射音の時系列応答

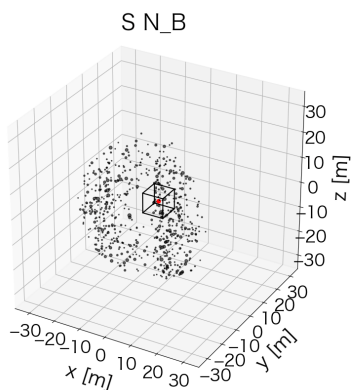


図 4.1.16 S N_B の音場の 3次元仮想音源分布

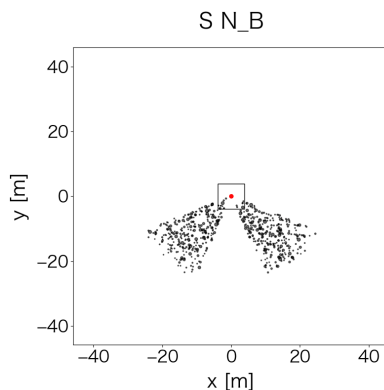


図 4.1.17 S N_B の音場の 2次元仮想音源分布

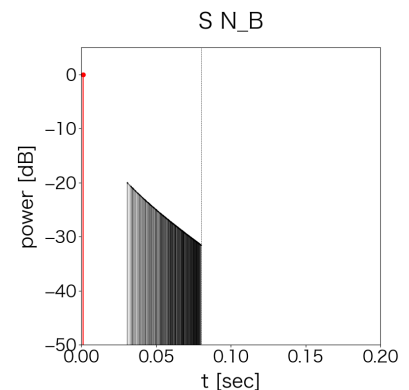


図 4.1.18 S N_B の音場の反射音の時系列応答

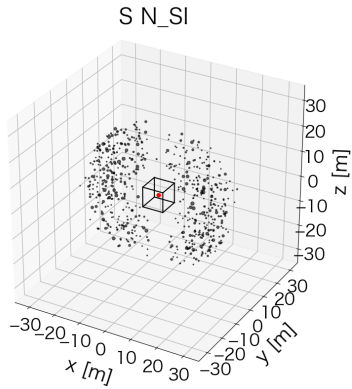


図 4.1.19 SN_SI の音場の 3次元仮想音源分布

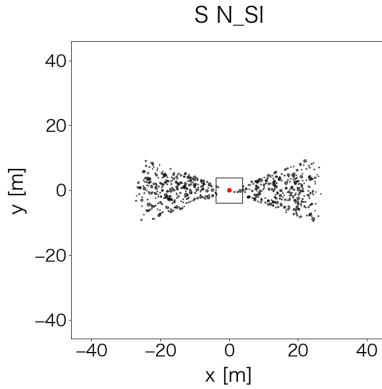


図 4.1.20 SN_SI の音場の 2次元仮想音源分布

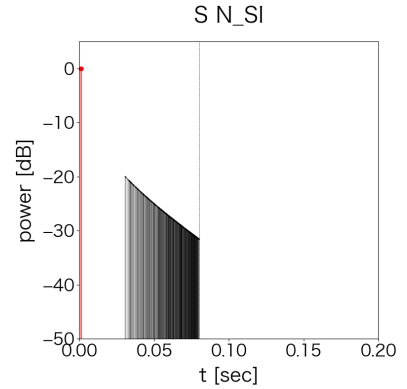


図 4.1.21 SN_SI の音場の反射音の時系列応答

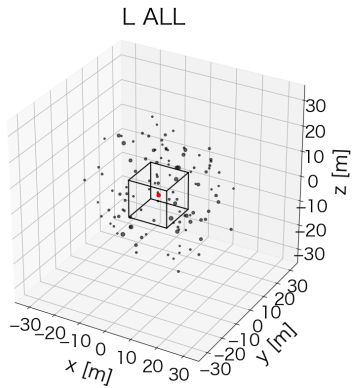


図 4.1.22 L ALL の音場の 3次元仮想音源分布

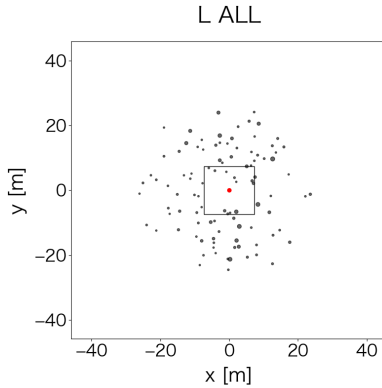


図 4.1.23 L ALL の音場の 2次元仮想音源分布

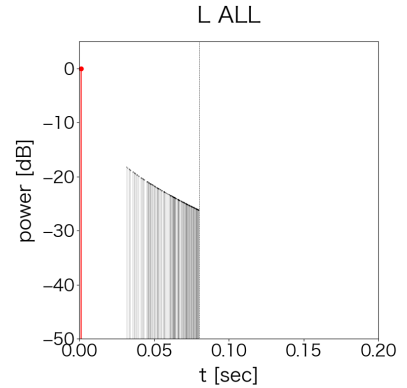


図 4.1.24 L ALL の音場の反射音の時系列応答

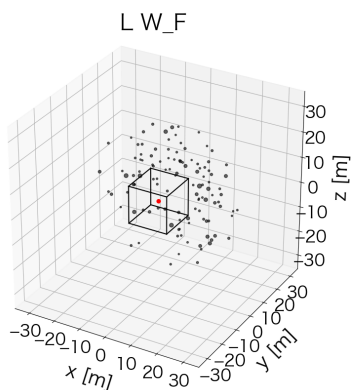


図 4.1.25 L W_F の音場の 3次元仮想音源分布

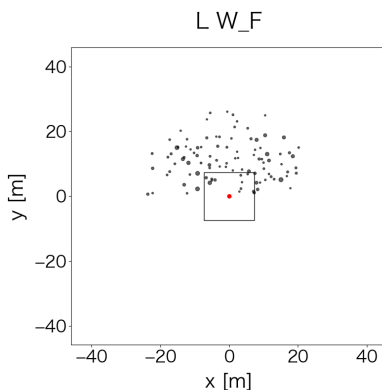


図 4.1.26 L W_F の音場の 2次元仮想音源分布

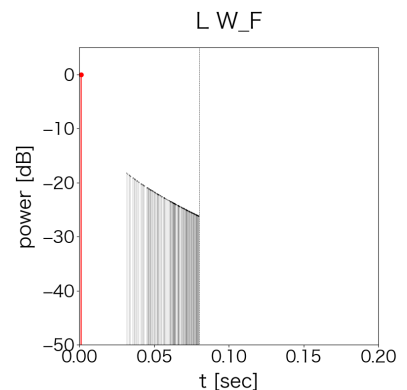


図 4.1.27 L W_F の音場の反射音の時系列応答

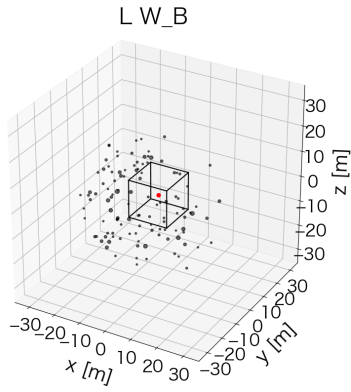


図 4.1.28 LW_B の音場の 3次元仮想音源分布

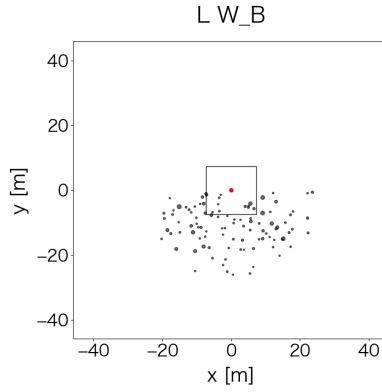


図 4.1.29 LW_B の音場の 2次元仮想音源分布

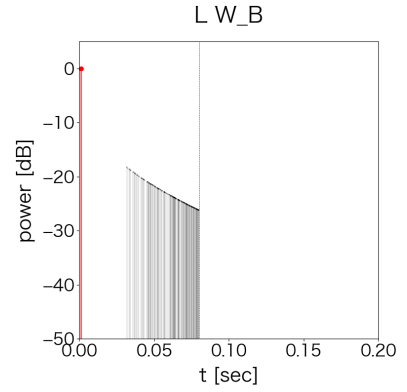


図 4.1.30 LW_B の音場の反射音の時系列応答

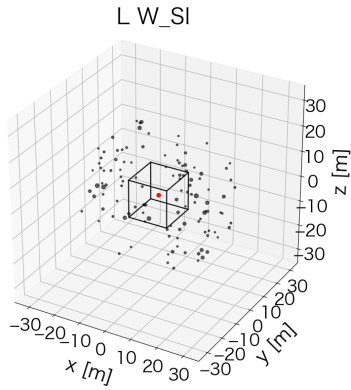


図 4.1.31 LW_SI の音場の 3次元仮想音源分布

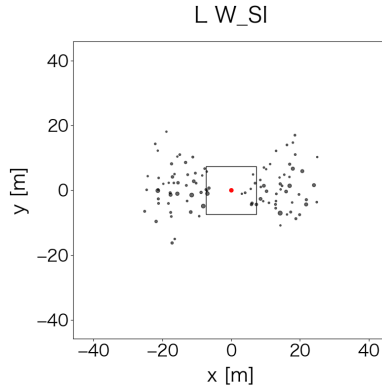


図 4.1.32 LW_SI の音場の 2次元仮想音源分布

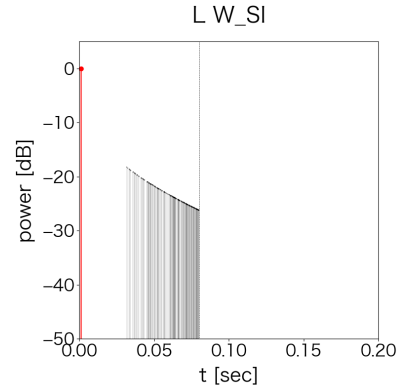


図 4.1.33 LW_SI の音場の反射音の時系列応答

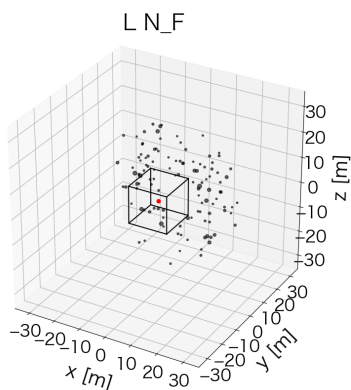


図 4.1.34 LN_F の音場の 3次元仮想音源分布

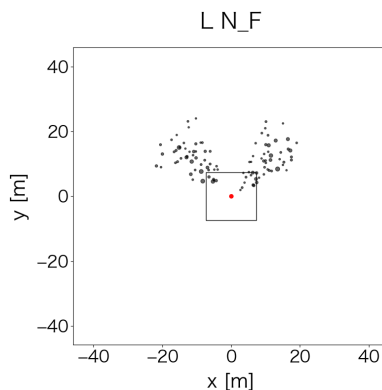


図 4.1.35 LN_F の音場の 2次元仮想音源分布

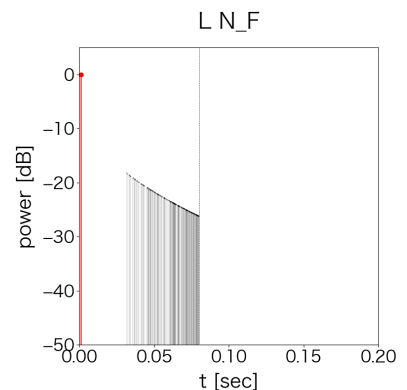


図 4.1.36 LN_F の音場の反射音の時系列応答

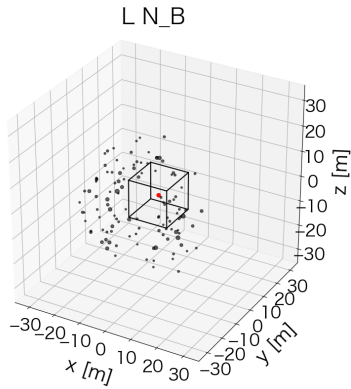


図 4.1.37 L N_B の音場の 3次元仮想音源分布

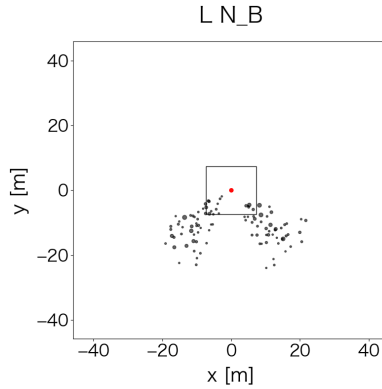


図 4.1.38 L N_B の音場の 2次元仮想音源分布

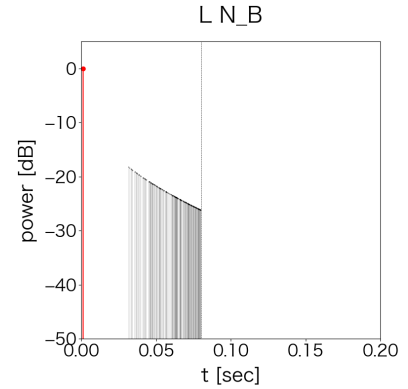


図 4.1.39 L N_B の音場の反射音の時系列応答

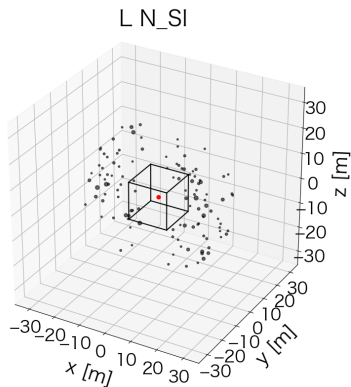


図 4.1.40 L N_SI の音場の 3次元仮想音源分布

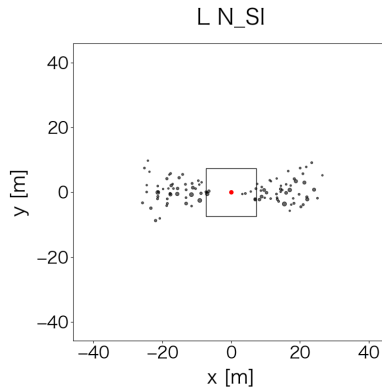


図 4.1.41 L N_SI の音場の 2次元仮想音源分布

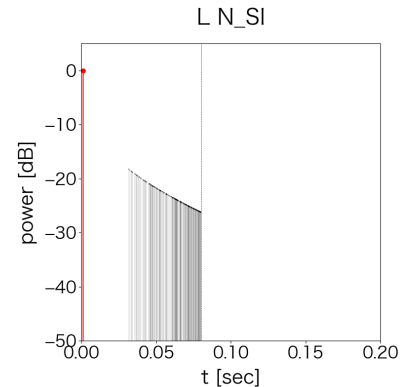


図 4.1.42 L N_F の音場の反射音の時系列応答

4.1.3 初期反射音レベルの調整方法

被験者に対し、音色の明瞭さ、および音像の明瞭さの 2 種類の明瞭さへの着目を求め、それぞれが最も明瞭だと感じられる値まで初期反射音レベルの調整を求めた。Beranek (2004) による水平的な明瞭さ、および垂直的な明瞭さの定義を参考に、それぞれの明瞭さについて以下のように教示した。

- 音色の明瞭さ：ある瞬間に再生されている音が前の音の残響等に影響され、濁っていたり聴こえづらくなっていないかどうか、音色がこもっていないか、響きのかぶりがないか。Beranek (2004) による水平的な明瞭さを参照した。
- 音像の明瞭さ：ある瞬間に再生されている音が、同時に再生されている他の音と空間的に区別できるかどうか。輪郭がはっきりしているかどうか。Beranek (2004) による垂直的な明瞭さを参照した。

被験者は手元の Griffin Technology 社の USB 接続ノブコントローラー (NA16029 PowerMate)

を使用し、ノブを左右に回すことで初期反射音レベルを調整した。初期反射音レベルの調整範囲は、直接音に対して ± 0 dB から -60 dB の間とした。ノブコントローラーには目盛りが存在せず、ノブの回転に上限および下限も無い。初期反射音レベルが ± 0 dB に達してからもノブは右に回し続けることができ、その場合は初期反射音レベルは変化しなかった。調整の上限や下限が存在することにより中央の値を選んでしまう中心化バイアスや、目盛り上の値を無意識に選んでしまうといったバイアスを避けるために上記のように設定した。

Cycling'74 社のビジュアルプログラミング言語 Max7 を使用し、実験刺激の初期反射音の再生音圧レベルをノブの値に基づき制御するアプリケーションを作成した。なお、インパルス応答は直接音部と初期反射音部を別のファイルとして生成した（付録 A を参照）。本実験における初期反射音レベルは、生成された初期反射音部のインパルス応答の音響エネルギー値に基づき、直接音部の音圧レベルを ± 0 dB とした場合のデシベル値として記録した。

実験に参加した被験者は 14 名で、それぞれに 14 刺激 \times 2 音源 \times 2 評価項目の合計 56 試行の調整を求めた。被験者には「最も明瞭さが高いと感じた音圧レベル」に加え、これより高い値に調整した際に明瞭さが低下したと感じた音圧レベル、さらに、これより低い音圧レベルに調整した際に明瞭さが低下したと感じた音圧レベルの調整も同時に求めた。それぞれ、本実験では「最適レベル」、「上限レベル」、「下限レベル」と呼ぶ。

被験者には以下のように教示を行った。教示は、回答の進行状況と共に被験者前方のディスプレイに常に表示した。

1. 聴取している音楽が最も明瞭に聴こえる位置にノブを合わせ、ボタンを押して回答して下さい。（最適レベルの調査）
2. 1. の設定から右にノブを回し、明瞭さが低下したと感じた位置にノブを合わせ、ボタンを押して回答して下さい。（上限レベルの調査）
3. 1. の設定から左にノブを回し、明瞭さが低下したと感じた位置にノブを合わせ、ボタンを押して回答して下さい。（下限レベルの調査）
4. 上記の 2. および 3. において、ノブを回しても明瞭さが低下しないと感じた場合は、申告ボタンを押して次の問題に進んで下さい。

4.2 分析

4.2.1 前処理

上限レベルあるいは下限レベルの回答の際、ノブコントローラーを回しても最適レベルの値から明瞭さが低下しなかったと申告された試行があった。本実験の全試行数は14名×56で784試行だが、上限レベルの調整では100試行、下限レベルでの調整では296試行の申告があった。これらの試行については、上限レベルは最も初期反射音レベルの高い±0 dB、下限レベルは最も初期反射音レベルの低い−60 dBと回答されたものとして値を修正し、分析に用いた。なお、上限レベルの調整で申告があった試行の最適レベルの平均値は直接音に対して−7.4 dB、下限レベルの調整で申告があった場合の最適レベルの平均値は−32.1 dBであった。

4.2.2 4要因の分散分析

図4.2.1は、14名全員の結果から得られた初期反射音の最適レベル、上限レベル、下限レベルの平均値および95%信頼区間を示す。

回答された明瞭さに最適な初期反射音レベルが、聴取する音源、着目する明瞭さ、音場の室容積および初期反射音の到来方向によって異なるかどうかを分析するため、最適レベル、上限レベル、下限レベルそれぞれについて4要因の分散分析を行った。分散分析とは、3群以上のデータ間の平均値を比較する統計手法の一つで、データの値の変化に影響すると考えられる要因による効果の有無の検定ができる。本実験においては、扱う要因は、「評価語」、「音源」、「室容積」、「初期反射音の到来方向」の4つである。これら4つの要因からなるデータの分散分析を行うことで、各要因の主効果、また要因と要因の組み合わせである交互作用によって、調整された初期反射音レベルが有意に異なったかどうかを検定する。検定における有意水準は $\alpha = 0.05$ を採用した。

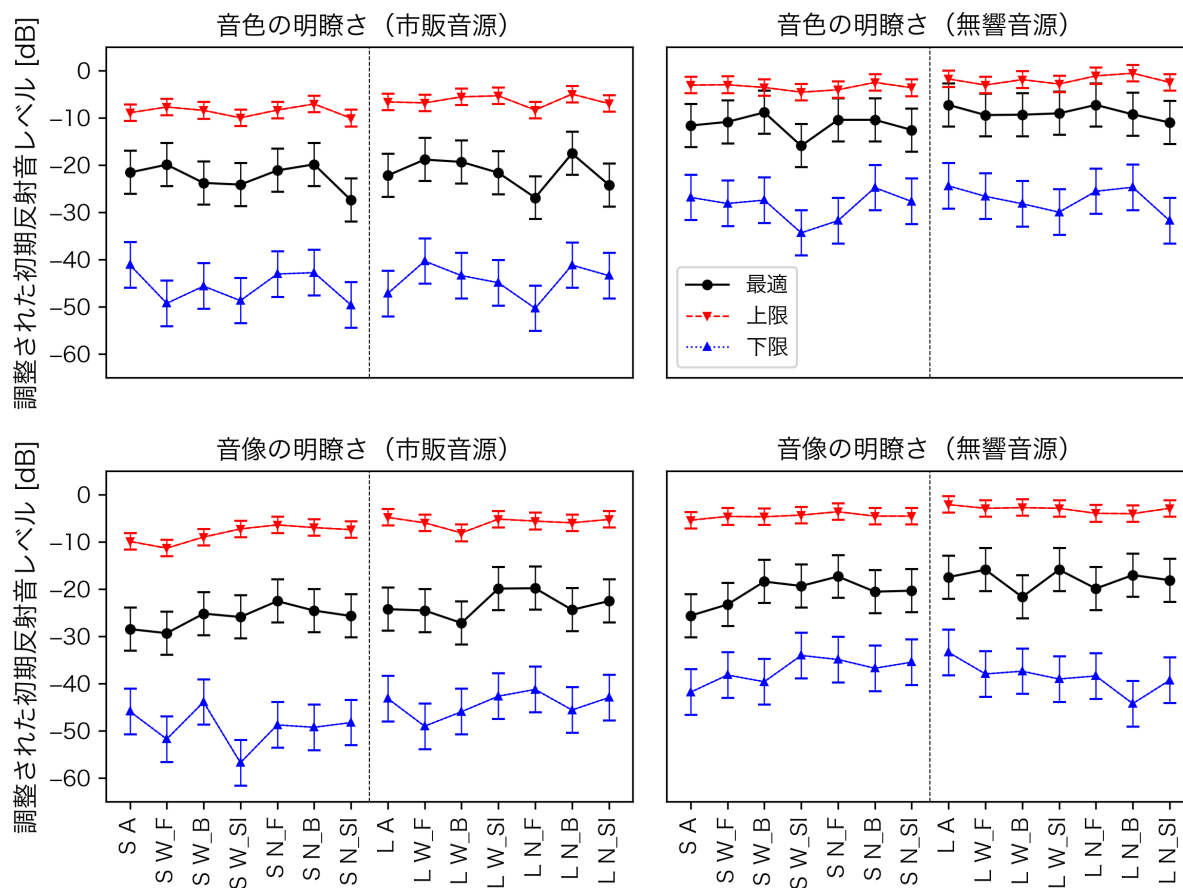


図 4.2.1 調整された初期反射音の最適レベル、上限レベル、下限レベルの平均値および 95% 信頼区間

表 4.2.1 は、調整された最適レベルの 4 要因の分散分析の結果を示す。表に示す p -value が 0.05 を下回った場合には、それぞれの要因によって調整された最適レベルに有意な差があったと言える。音源の主効果 ($F(1, 13) = 7.26, p = 0.02$)、室容積の主効果 ($F(1, 13) = 26.27, p < 0.001$)、評価語 × 到来方向の交互作用 ($F(6, 78) = 2.57, p = 0.03$) で有意差が見られる。また、評価語の主効果 ($F(1, 13) = 3.81, p = 0.07$)、評価語 × 音源の交互作用 ($F(1, 13) = 3.27, p = 0.09$) で有意傾向が見られる。

表 4.2.1 最適レベルの 4 要因の分散分析結果。

Source	SS	df	MS	F-ratio	p-value	
s	72338.50	13	5564.50			
評価語	6712.29	1	6712.29	3.81	0.07	+
s × 評価語	22880.71	13	1760.05			
音源	14246.13	1	14246.13	7.26	0.02	*
s × 音源	25519.44	13	1963.03			
室容積	1000.37	1	1000.37	26.27	0.00	***
s × 室容積	494.96	13	38.07			
到来方向	446.84	6	74.47	1.09	0.38	ns
s × 到来方向	5340.91	78	68.47			
評価語 × 音源	2102.88	1	2102.88	3.27	0.09	+
s × 評価語 × 音源	8356.91	13	642.84			
評価語 × 室容積	40.41	1	40.41	0.74	0.41	ns
s × 評価語 × 室容積	709.52	13	54.58			
評価語 × 到来方向	1239.37	6	206.56	2.57	0.03	*
s × 評価語 × 到来方向	6274.75	78	80.45			
音源 × 室容積	28.70	1	28.70	0.55	0.47	ns
s × 音源 × 室容積	674.86	13	51.91			
音源 × 到来方向	104.18	6	17.36	0.20	0.98	ns
s × 音源 × 到来方向	6882.50	78	88.24			
室容積 × 到来方向	738.76	6	123.13	1.63	0.15	ns
s × 室容積 × 到来方向	5900.20	78	75.64			
評価語 × 音源 × 室容積	31.04	1	31.04	0.58	0.46	ns
s × 評価語 × 音源 × 室容積	700.24	13	53.86			
評価語 × 音源 × 到来方向	558.03	6	93.01	1.16	0.34	ns
s × 評価語 × 音源 × 到来方向	6279.93	78	80.51			
評価語 × 室容積 × 到来方向	415.85	6	69.31	1.22	0.30	ns
s × 評価語 × 室容積 × 到来方向	4413.97	78	56.59			
音源 × 室容積 × 到来方向	242.35	6	40.39	0.58	0.75	ns
s × 音源 × 室容積 × 到来方向	5454.97	78	69.94			
評価語 × 音源 × 室容積 × 到来方向	476.69	6	79.45	0.76	0.60	ns
s × 評価語 × 音源 × 室容積 × 到来方向	8147.77	78	104.46			
Total	208754.04	783	266.61			

+ $p < 0.10$, * $p < 0.05$, ** $p < 0.01$, *** $p < .001$

表 4.2.2 は、上限レベルの4要因の分散分析の結果を示す。音源の主効果 ($F(1, 13) = 11.85$, $p < 0.01$)、室容積の主効果 ($F(1, 13) = 42.34$, $p < 0.001$) で有意差が見られる。評価語 × 到来方向の交互作用 ($F(6, 78) = 2.11$, $p = 0.06$) で有意傾向が見られる。

表 4.2.2 上限レベルの4要因の分散分析結果。

Source	SS	df	MS	F-ratio	p-value	
s	7093.09	13	545.62			
評価語	23.25	1	23.25	0.19	0.67	ns
s × 評価語	1612.84	13	124.06			
音源	3188.22	1	3188.22	11.85	0.00	**
s × 音源	3498.95	13	269.15			
室容積	734.02	1	734.02	42.34	0.00	***
s × 室容積	225.36	13	17.34			
到来方向	81.24	6	13.54	0.90	0.50	ns
s × 到来方向	1168.72	78	14.98			
評価語 × 音源	112.51	1	112.51	1.01	0.33	ns
s × 評価語 × 音源	1447.22	13	111.32			
評価語 × 室容積	0.29	1	0.29	0.02	0.90	ns
s × 評価語 × 室容積	207.80	13	15.98			
評価語 × 到来方向	175.71	6	29.28	2.11	0.06	+
s × 評価語 × 到来方向	1083.14	78	13.89			
音源 × 室容積	38.17	1	38.17	2.76	0.12	ns
s × 音源 × 室容積	180.10	13	13.85			
音源 × 到来方向	50.03	6	8.34	0.56	0.76	ns
s × 音源 × 到来方向	1157.44	78	14.84			
室容積 × 到来方向	81.41	6	13.57	0.95	0.47	ns
s × 室容積 × 到来方向	1115.77	78	14.30			
評価語 × 音源 × 室容積	0.93	1	0.93	0.03	0.86	ns
s × 評価語 × 音源 × 室容積	393.09	13	30.24			
評価語 × 音源 × 到来方向	109.37	6	18.23	1.20	0.32	ns
s × 評価語 × 音源 × 到来方向	1184.52	78	15.19			
評価語 × 室容積 × 到来方向	150.56	6	25.09	1.82	0.11	ns
s × 評価語 × 室容積 × 到来方向	1073.47	78	13.76			
音源 × 室容積 × 到来方向	50.96	6	8.49	0.83	0.55	ns
s × 音源 × 室容積 × 到来方向	793.58	78	10.17			
評価語 × 音源 × 室容積 × 到来方向	64.13	6	10.69	1.12	0.36	ns
s × 評価語 × 音源 × 室容積 × 到来方向	741.97	78	9.51			
Total	27837.87	783	35.55			

+ $p < 0.10$, * $p < 0.05$, ** $p < 0.01$, *** $p < .001$

表 4.2.3 は、下限レベルの 4 要因の分散分析の結果を示す。評価語の主効果 ($F(1, 13) = 7.91$, $p = 0.01$)、音源の主効果 ($F(1, 13) = 25.9$, $p < 0.001$) で有意差が見られる。また、音源 × 室容積の交互作用 ($F(1, 13) = 4.82$, $p = 0.05$)、音源 × 室容積 × 到来方向の交互作用 ($F(6, 78) = 2.48$, $p = 0.03$)、評価語 × 音源 × 室容積 × 到来方向の交互作用 ($F(6, 78) = 2.46$, $p = 0.03$) で有意差が見られる。評価語 × 音源の交互作用 ($F(1, 13) = 3.84$, $p = 0.07$)、評価語 × 音源 × 室容積の交互作用 ($F(1, 13) = 3.66$, $p = 0.08$) では有意傾向が見られる。

表 4.2.3 下限レベルの 4 要因の分散分析結果。

Source	<i>SS</i>	<i>df</i>	<i>MS</i>	<i>F</i> -ratio	<i>p</i> -value	
s	121210.41	13.00	9323.88			
評価語	6648.07	1	6648.07	7.91	0.01	*
s × 評価語	10927.37	13.00	840.57			
音源	33033.06	1	33033.06	25.89	0.00	***
s × 音源	16584.67	13.00	1275.74			
室容積	491.26	1	491.26	1.51	0.24	ns
s × 室容積	4235.73	13.00	325.83			
到来方向	801.05	6	133.51	1.07	0.39	ns
s × 到来方向	9700.88	780	124.37			
評価語 × 音源	3253.07	1	3253.07	3.84	0.07	+
s × 評価語 × 音源	11019.80	13.00	847.68			
評価語 × 室容積	8.37	1	8.37	0.07	0.80	ns
s × 評価語 × 室容積	1563.29	13.00	120.25			
評価語 × 到来方向	1299.08	6	216.51	1.71	0.13	ns
s × 評価語 × 到来方向	9885.08	780	126.73			
音源 × 室容積	451.56	1	451.56	4.82	0.05	*
s × 音源 × 室容積	1217.09	13.00	93.62			
音源 × 到来方向	212.05	6	35.34	0.33	0.92	ns
s × 音源 × 到来方向	8322.59	780	106.70			
室容積 × 到来方向	497.04	6	82.84	0.60	0.73	ns
s × 室容積 × 到来方向	10698.89	780	137.17			
評価語 × 音源 × 室容積	466.87	1	466.87	3.66	0.08	+
s × 評価語 × 音源 × 室容積	1658.22	13.00	127.56			
評価語 × 音源 × 到来方向	543.90	6	90.65	0.71	0.64	ns
s × 評価語 × 音源 × 到来方向	9911.10	780	127.07			
評価語 × 室容積 × 到来方向	575.89	6	95.98	0.89	0.51	ns
s × 評価語 × 室容積 × 到来方向	8424.82	780	108.01			
音源 × 室容積 × 到来方向	1627.59	6	271.26	2.48	0.03	*
s × 音源 × 室容積 × 到来方向	8525.91	780	109.31			
評価語 × 音源 × 室容積 × 到来方向	1572.78	6	262.13	2.46	0.03	*
s × 評価語 × 音源 × 室容積 × 到来方向	8320.50	780	106.67			
Total	293688.01	783	375.08			

+ $p < 0.10$, * $p < 0.05$, ** $p < 0.01$, *** $p < .001$

室容積の主効果

分散分析より、最適レベルおよび上限レベルで室容積の主効果に有意差が見られた。図 4.2.2 は、室容積 S (60 m^3) および室容積 L (400 m^3) の調整された初期反射音レベルの平均値および 95% 信頼区間を示す。

■最適レベル 室容積 L の最適レベルの平均値は直接音に対して -15.3 dB 、室容積 S の最適レベルの平均値は -17.9 dB であり、室容積 L が 2.6 dB 高い最適レベルに調整されている。

■上限レベル 室容積 L の上限レベルの平均値は -3.5 dB 、室容積 S の上限レベルの平均値は -5.5 dB であり、室容積 L が 2.0 dB 高い上限レベルに調整されている。

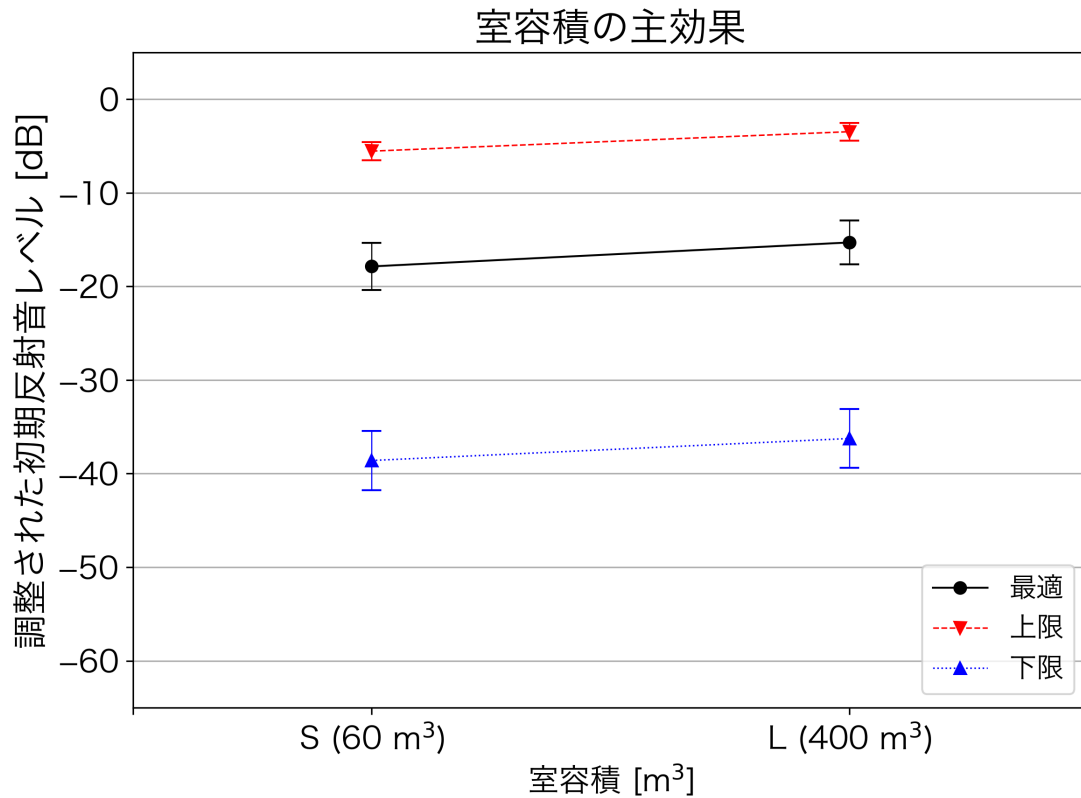


図 4.2.2 室容積 S (60 m^3) および室容積 L (400 m^3) の調整された初期反射音レベルの平均値および 95% 信頼区間

評価語 × 音源の交互作用

分散分析より、最適レベル、および下限レベルで評価語 × 音源の交互作用に有意傾向が見られた。

分散分析で有意差が見られた場合、多重比較と呼ばれる検定法を用いて、どの群間に有意差があるかを調べることができる。しかし、検定を繰り返すことで第一種過誤率が大きくなってしまうため、比較する仮説の数に応じて p 値を補正する Holm の方法による多重比較を用い、後述する検定を実施した。

図 4.2.3 は、音色の明瞭さ（市販音源）、音色の明瞭さ（無響音源）、音像の明瞭さ（市販音源）、音像の明瞭さ（無響音源）それぞれの初期反射音レベルの平均値および 95% 信頼区間を示す。多重比較より、最適レベルでは音色の明瞭さの評価における聴取音源の違いに有意差が ($F(1, 13) = 15.92, p < 0.01$)、無響音源の聴取において、評価語の違いに有意傾向 ($F(1, 13) = 4.66, p = 0.05$) が見られた。同様に下限レベルについても、音色の明瞭さの評価における聴取音源の違いに有意差が ($F(1, 13) = 26.32, p < 0.001$)、無響音源の聴取における評価語の違いに有意傾向が ($F(1, 13) = 7.60, p = 0.02$) 見られた。無響音源の聴取における音色の明瞭さで、有意に初期反射音レベルが高い。

■最適レベル 音色の明瞭さの評価において、市販音源の最適レベルの平均値は直接音に対して -22.0 dB、無響音源の最適レベルの平均値は -10.2 dB であり、無響音源が 11.8 dB 高い最適レベルに調整されている。音像の明瞭さの評価において、市販音源の最適レベルの平均値は直接音に対して -24.6 dB、無響音源の最適レベルの平均値は -19.3 dB であり、違いに有意差は見られなかった。

■下限レベル 音色の明瞭さの評価において、市販音源の下限レベルの平均値は直接音に対して -45.0 dB、無響音源の下限レベルの平均値は -28.0 dB であり、無響音源が 17.0 dB 高い下限レベルに調整されている。音像の明瞭さの評価において、市販音源の下限レベルの平均値は直接音に対して -45.0 dB、無響音源の下限レベルの平均値は -37.9 dB であり、違いに有意差は見られなかった。

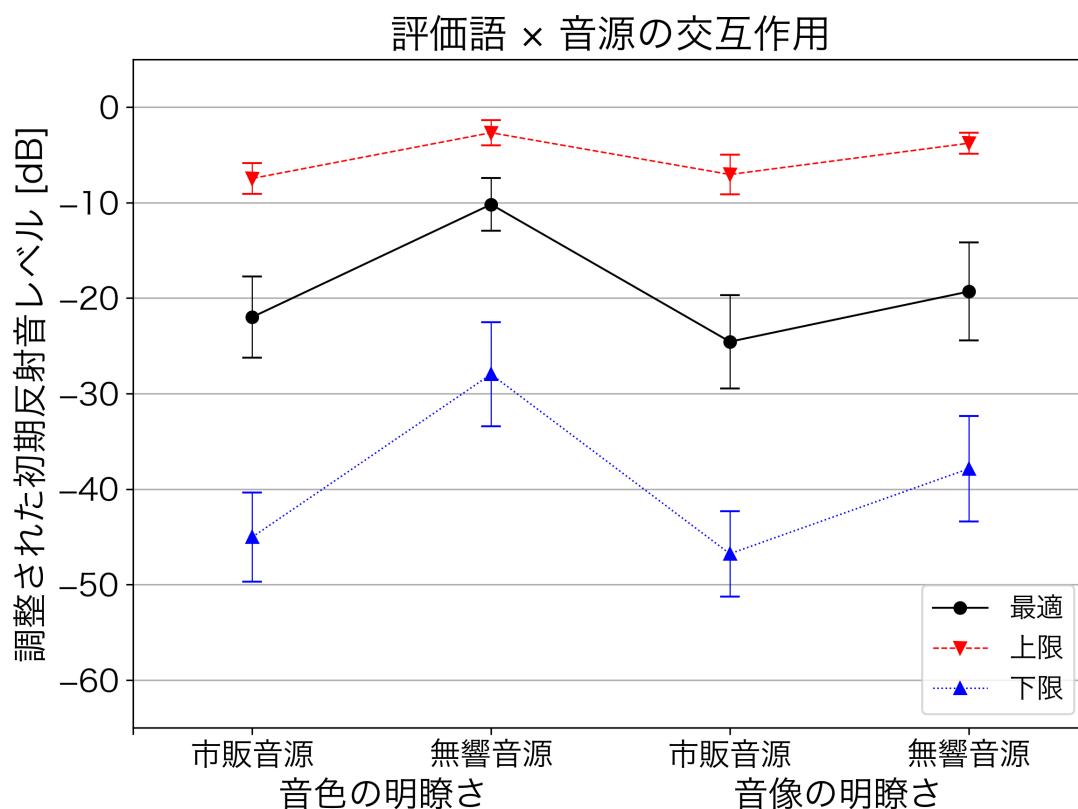


図 4.2.3 音色の明瞭さ（市販音源）、音色の明瞭さ（無響音源）、音像の明瞭さ（市販音源）、音像の明瞭さ（無響音源）それぞれの初期反射音レベルの平均値および 95% 信頼区間

評価語 × 到来方向の交互作用

分散分析より、最適レベルおよび上限レベルで評価語 × 到来方向の交互作用に有意傾向が見られた。図 4.2.4 は、音色の明瞭さおよび音像の明瞭さの到来方向別の初期反射音レベルの平均値および 95% 信頼区間を示す。多重比較より、最適レベルでは到来方向が ALL、W_F、W_B、W_SI の音場の聴取において、評価語の違いに有意差が見られた。上限レベルでは、音色の明瞭さにおける到来方向の違いに有意傾向が見られたが ($F(6, 78) = 1.85, p = 0.10$)、各対同士の刺激の間に有意差は見られなかった。

■最適レベル ALL、W_F、W_B、W_SI の 4 つの音場の最適レベルの音色の明瞭さの平均値は直接音に対して -15.8 dB、音像の明瞭さの平均値は -22.9 dB であり、音色の明瞭さが 7.1 dB 高い最適レベルに調整されている。

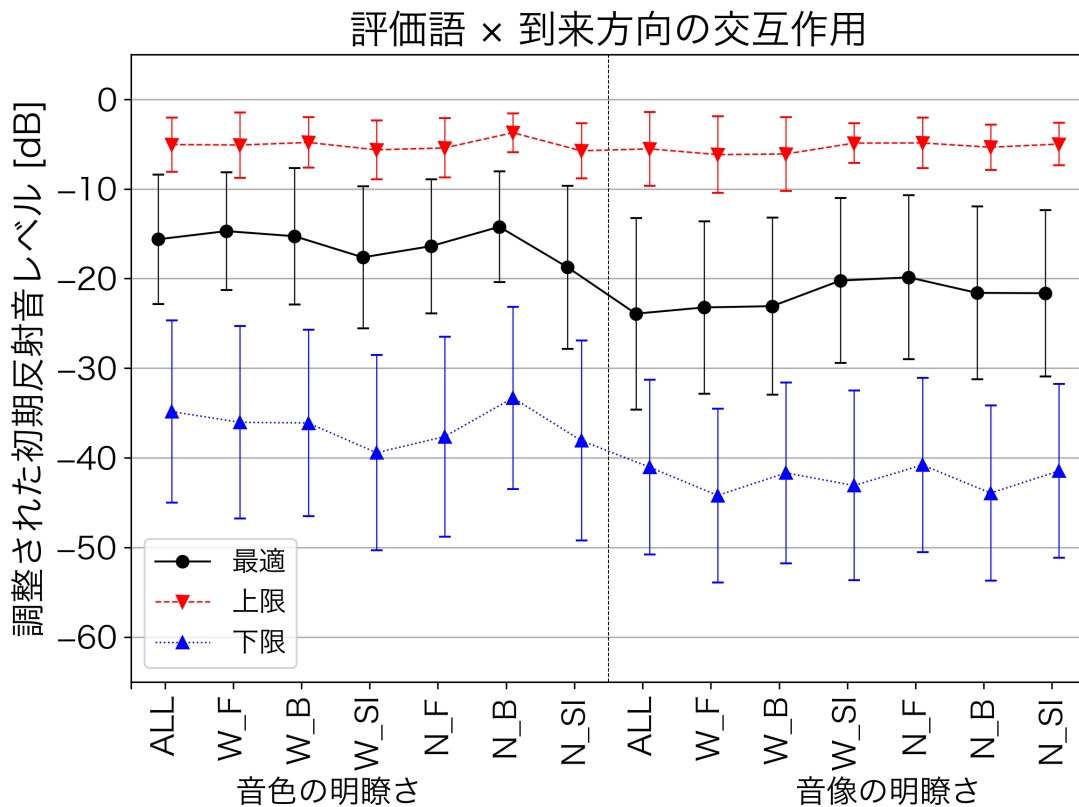


図 4.2.4 音色の明瞭さおよび音像の明瞭さの到来方向別の最適レベルの平均値および 95% 信頼区間

音源 × 室容積の交互作用

分散分析より、下限レベルで音源 × 室容積の交互作用に有意傾向が見られた。図 4.2.5 は、市販音源 × S、市販音源 × L、無響音源 × S、無響音源 × L の初期反射音レベルの平均値および 95% 信頼区間を示す。多重比較より、下限レベルでは室容積 S および室容積 L のどちらの室容積においても音源の違いに有意差が見られた。また、市販音源の聴取における室容積の違いに有意差が見られた ($F(1,13) = 5.27, p = 0.04$)。

■下限レベル 室容積 S において、市販音源の下限レベルの平均値は直接音に対して -45.9 dB、無響音源の下限レベルの平均値は -31.3 dB であり、無響音源の方が 14.6 dB 高い下限レベルに調整されている。室容積 L において、市販音源の下限レベルの平均値は直接音に対して -42.0 dB、無響音源の下限レベルの平均値は -30.5 dB であり、無響音源の方が 11.5 dB 高い下限レベルに調整されている。

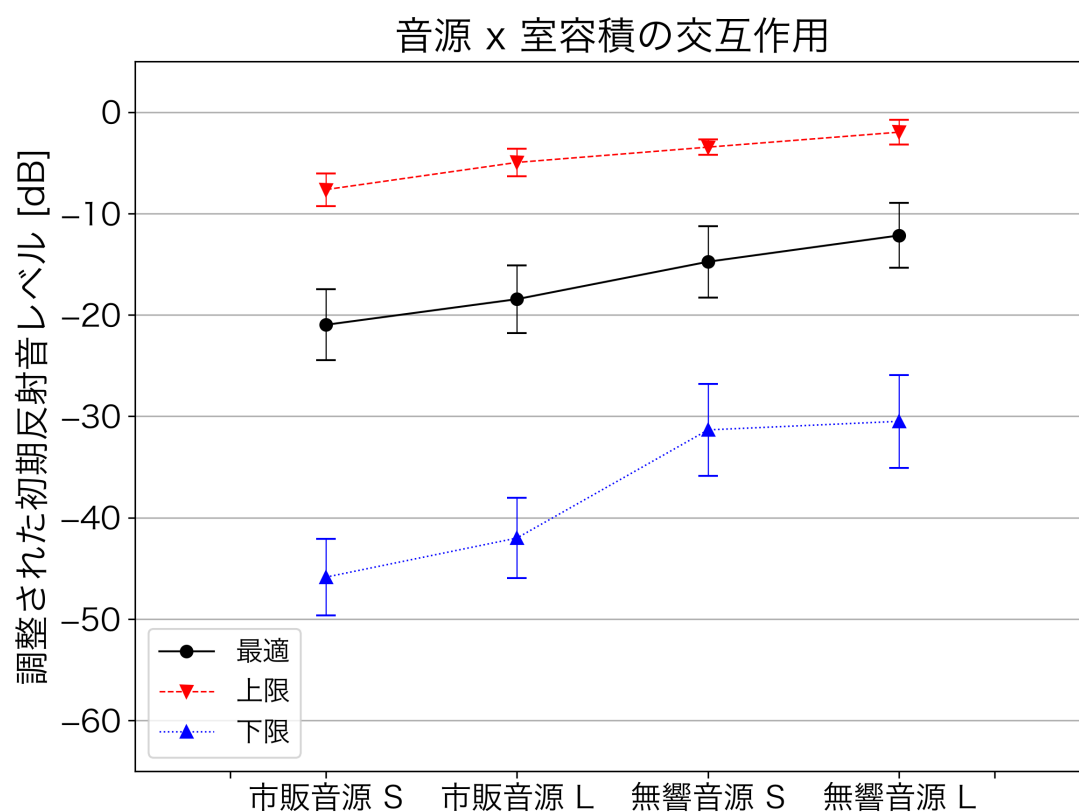


図 4.2.5 市販音源 (S)、市販音源 (L)、無響音源 (S)、無響音源 (L) の平均値および 95% 信頼区間

4.2.3 クラスタ分析によるグループ分け

個人によって回答傾向に違いが見られるかどうかを検証するため、ウォード法に基づく階層型クラスタ分析を実施した。クラスタとは群れ、集団を指し、クラスタ分析とは、分析の対象となる個体を互いの類似度に従っていくつかのグループに分割する手法の総称である。クラスタ分析には、クラスタ間の距離をどのように定義するかに応じて多様なアルゴリズムが存在する。本実験で採用した階層型クラスタ分析のウォード法は、クラスタ内での分散が最も小さくなるようにクラスタ間の結合を順次決定する。計算量が多いが分類感度が良いと言われている。

クラスタ分析には、被験者ごとの全ての試行における最適レベル、上限レベル、下限レベルの回答レベルを分析対象に用いた。図 4.2.6 は分析結果の樹形図を示す。アルファベットは被験者の ID を示す。A、B、C、F、J、L の 6 名（以下、グループ 1）と、D、E、G、H、I、K、M、N（以下、グループ 2）の 8 名に大別できることが分かる。

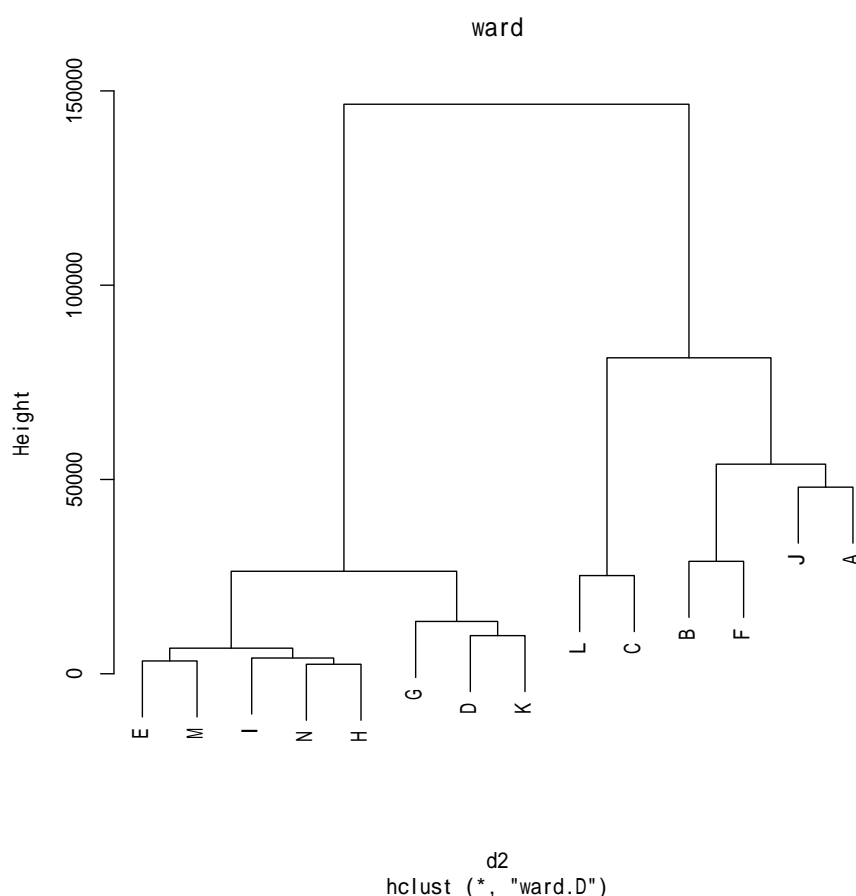


図 4.2.6 ウォード法に基づく 14 名の被験者のクラスタ分析の樹形図

図 4.2.7 は、グループ 1 の 6 名の被験者から得られた初期反射音レベルの平均値および 95% 信頼区間を示す。分散分析の結果、最適レベルでは室容積の主効果に有意差が見られた ($F(1, 5) = 6.61, p = 0.05$)。上限レベルでは室容積の主効果に有意差が見られた ($F(1, 5) = 7.51, p = 0.04$)、音源の主効果に有意傾向が見られた ($F(1, 5) = 6.07, p = 0.06$)。下限レベルでは音源の主効果に有意差が見られた ($F(1, 5) = 7.99, p = 0.04$)。

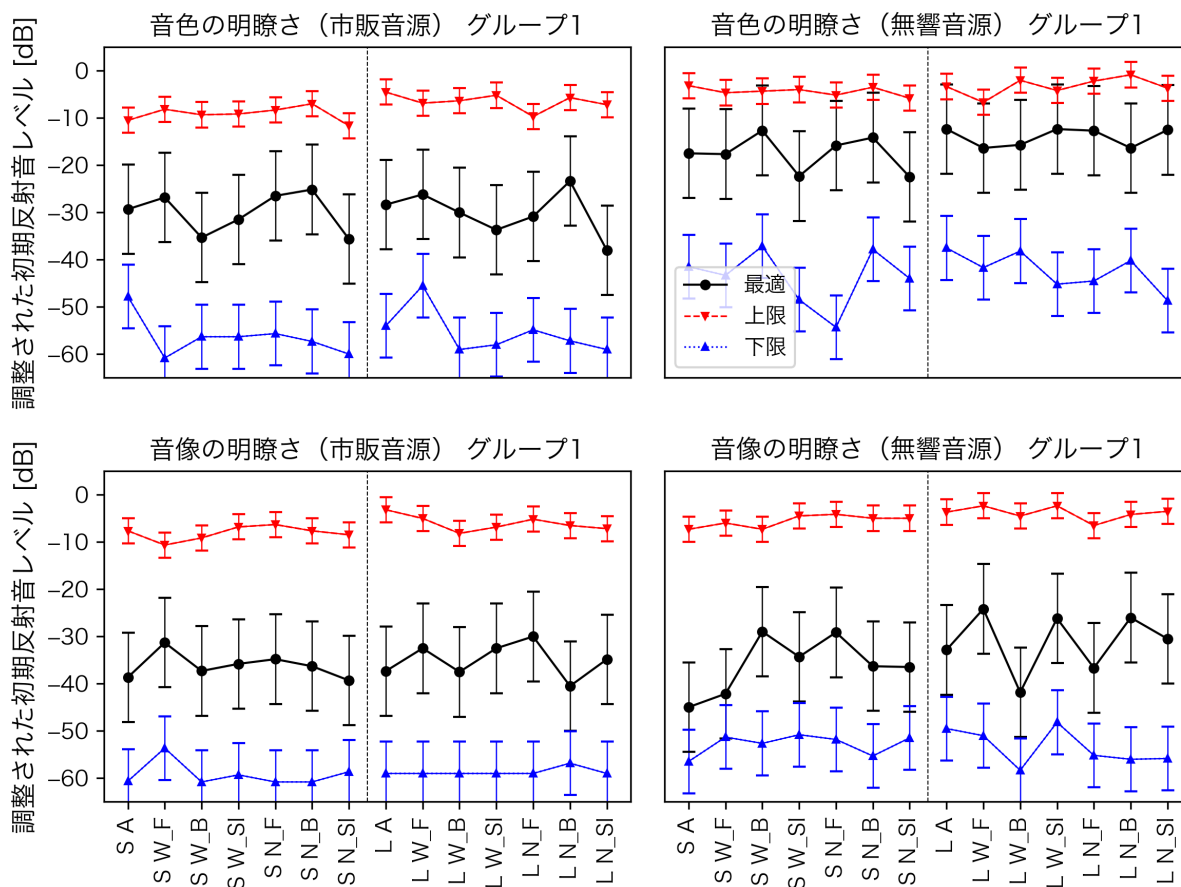


図 4.2.7 グループ 1 の 6 名の被験者から得られた初期反射音レベルの平均値および 95% 信頼区間

図 4.2.8 は、グループ 2 の 8 名の被験者から得られた初期反射音レベルの平均値および 95% 信頼区間を示す。分散分析の結果、最適レベルでは音源の主効果 ($F(1, 7) = 8.34, p = 0.02$)、室容積の主効果 ($F(1, 7) = 27.4, p < 0.01$)、「評価語」×「到来方向」($F(1, 7) = 6.41, p < 0.001$)の交互作用に有意差が見られた。上限レベルでは音源の主効果 ($F(1, 7) = 6.39, p = 0.04$)、室容積の主効果 ($F(1, 7) = 119, p < 0.001$)に有意差が、「評価語」×「到来方向」($F(1, 7) = 2.20, p = 0.06$)の交互作用に有意傾向が見られた。下限レベルでは 4 つ全ての要因の交互作用で有意差が見られた ($F(6, 42) = 3.02, p = 0.02$)。

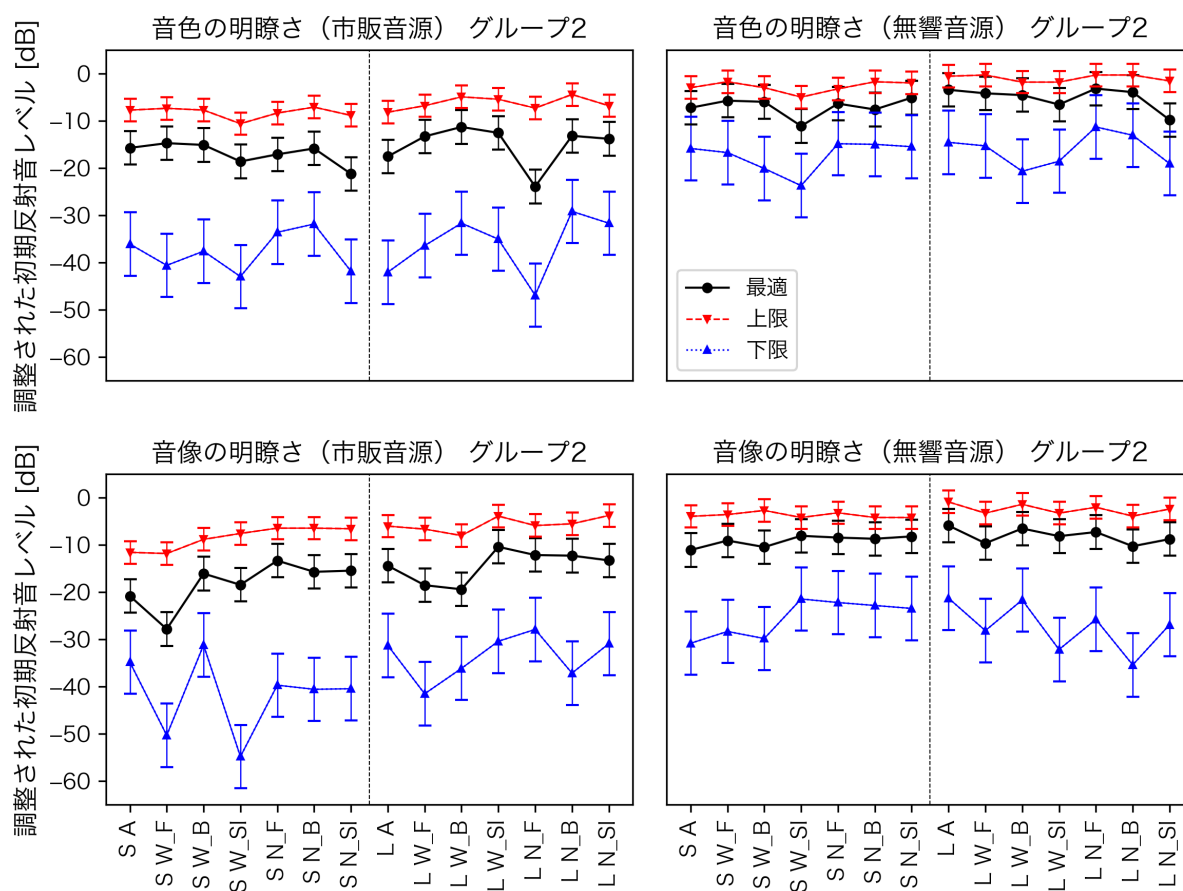


図 4.2.8 グループ 2 の 8 名の被験者から得られた初期反射音レベルの平均値および 95% 信頼区間

4.2.4 グループ別の音源の主効果

図 4.2.9 は、被験者グループ 1 における市販音源および無響音源の初期反射音レベルの平均値および 95% 信頼区間を示す。分散分析より、上限レベルで音源の主効果に有意傾向が、下限レベルで音源の主効果に有意差が見られた。

■上限レベル 市販音源の上限レベルの平均値は直接音に対して -6.4 dB、無響音源の上限レベルの平均値は -3.7 dB で、無響音源の方が 2.7 dB 高い上限レベルに調整されている。

■下限レベル 市販音源の下限レベルの平均値は直接音に対して -55.3 dB、無響音源の下限レベルの平均値は -43.5 dB で、無響音源の方が 11.8 dB 高い下限レベルに調整されている。

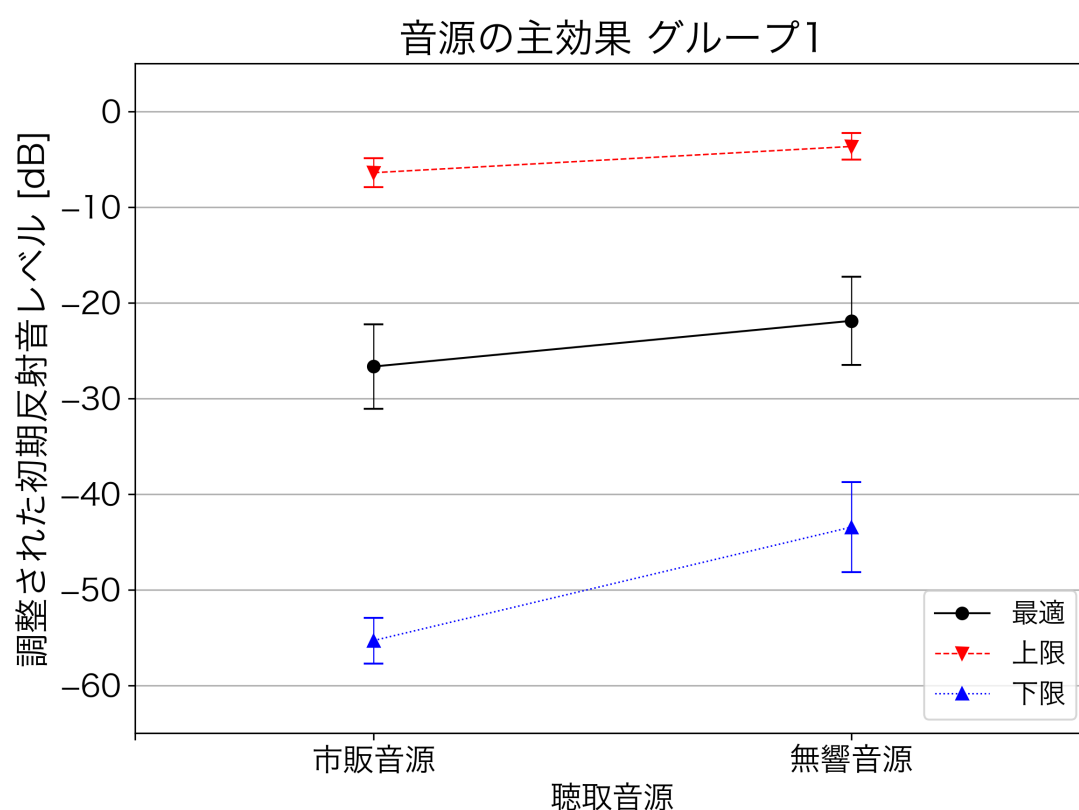


図 4.2.9 被験者グループ 1 における市販音源および無響音源の初期反射音レベルの平均値および 95% 信頼区間

図 4.2.10 は、被験者グループ 2 における市販音源および無響音源の初期反射音レベルの平均値および 95% 信頼区間を示す。分散分析より、最適レベル、上限レベル、および下限レベルの全てで音源の主効果に有意差が見られた。

■最適レベル 市販音源の最適レベルの平均値は直接音に対して -14.5 dB、無響音源では -7.1 dB で、無響音源の方が 7.4 dB 高い最適レベルに調整されている。

■上限レベル 市販音源の上限レベルの平均値は直接音に対して -6.2 dB、無響音源の上限レベルの平均値は -1.2 dB で、無響音源の方が 5.0 dB 高い上限レベルに調整されている。

■下限レベル 市販音源の下限レベルの平均値は直接音に対して -35.4 dB、無響音源の下限レベルの平均値は -21.5 dB で、無響音源の方が 13.9 dB 高い最適レベルに調整されている。

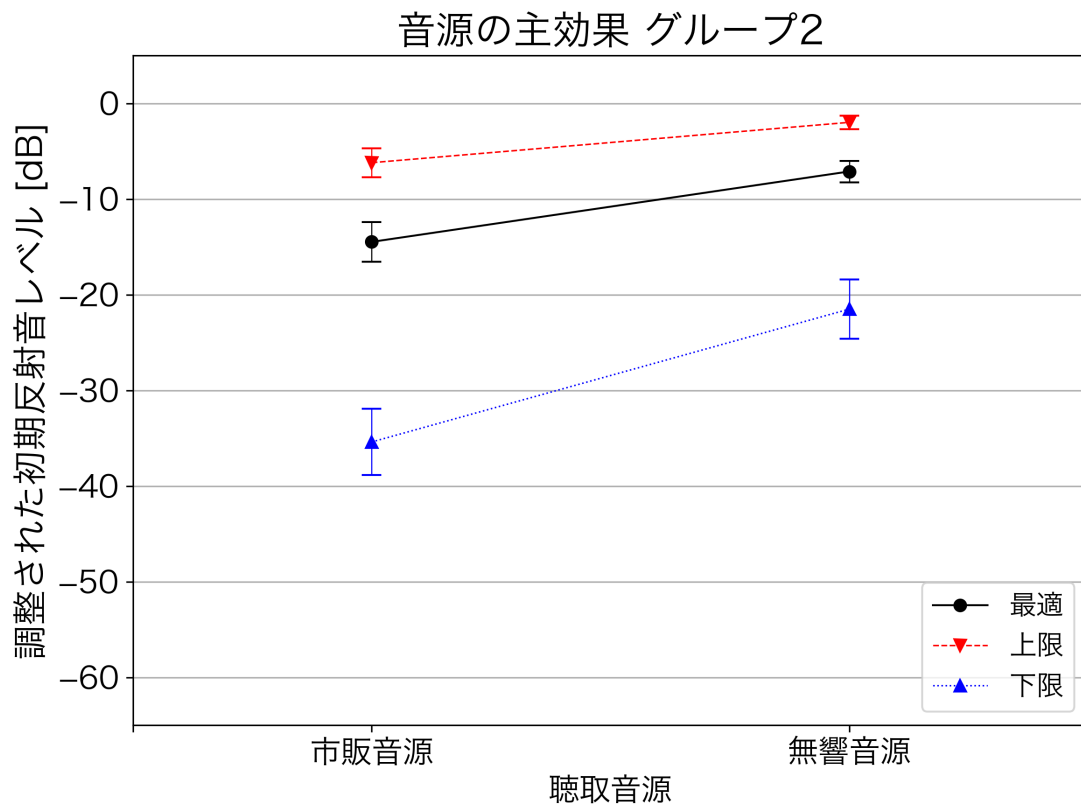


図 4.2.10 被験者グループ 2 における市販音源および無響音源の初期反射音レベルの平均値および 95% 信頼区間

4.2.5 グループ別の室容積の主効果

図 4.2.11 は、被験者グループ 1 における室容積 S (60 m^3) および室容積 L (400 m^3) の初期反射音レベルの平均値および 95% 信頼区間を示す。分散分析より、最適レベル、および上限レベルで室容積の主効果に有意差が見られた。

■最適レベル 室容積 S の最適レベルの平均値は直接音に対して -25.8 dB 、室容積 L の最適レベルの平均値は -22.8 dB であり、室容積 L の方が 3.0 dB 高い最適レベルに調整されている。

■上限レベル 室容積 S の上限レベルの平均値は直接音に対して -6.1 dB 、室容積 L の上限レベルの平均値は -4.0 dB であり、約 2.1 dB 高い上限レベルに調整されている。

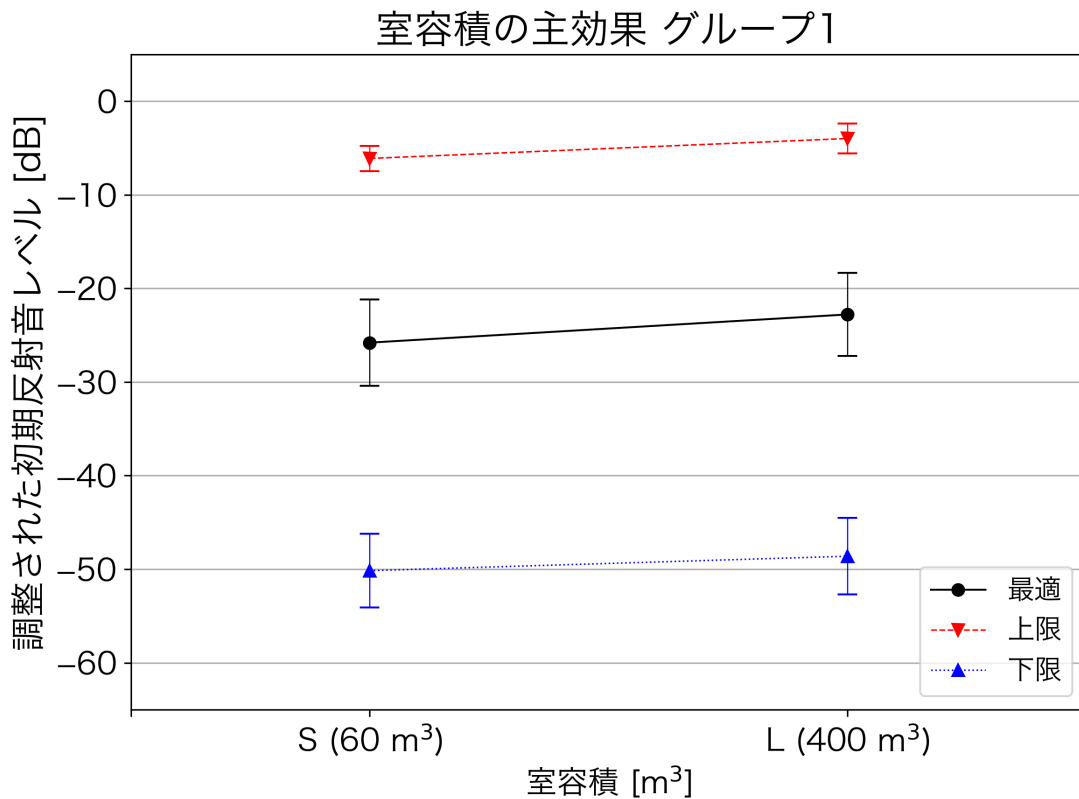


図 4.2.11 被験者グループ 1 における室容積 S (60 m^3) および室容積 L (400 m^3) の初期反射音レベルの平均値および 95% 信頼区間

図 4.2.10 は、被験者グループ 2 における室容積 S (60 m^3) および室容積 L (400 m^3) の初期反射音レベルの平均値および 95% 信頼区間を示す。分散分析より、最適レベル、および上限レベルで室容積の主効果に有意差が見られた。

■最適レベル 室容積 S の最適レベルの平均値は直接音に対して -11.9 dB 、室容積 L の最適レベルの平均値は -9.7 dB であり、室容積 L の方が 2.2 dB 高い最適レベルに調整されている。

■上限レベル 室容積 S の上限レベルの平均値は直接音に対して -5.1 dB 、室容積 L の上限レベルの平均値は -3.0 dB であり、室容積 L の方が 2.1 dB 高い上限レベルに調整されている。

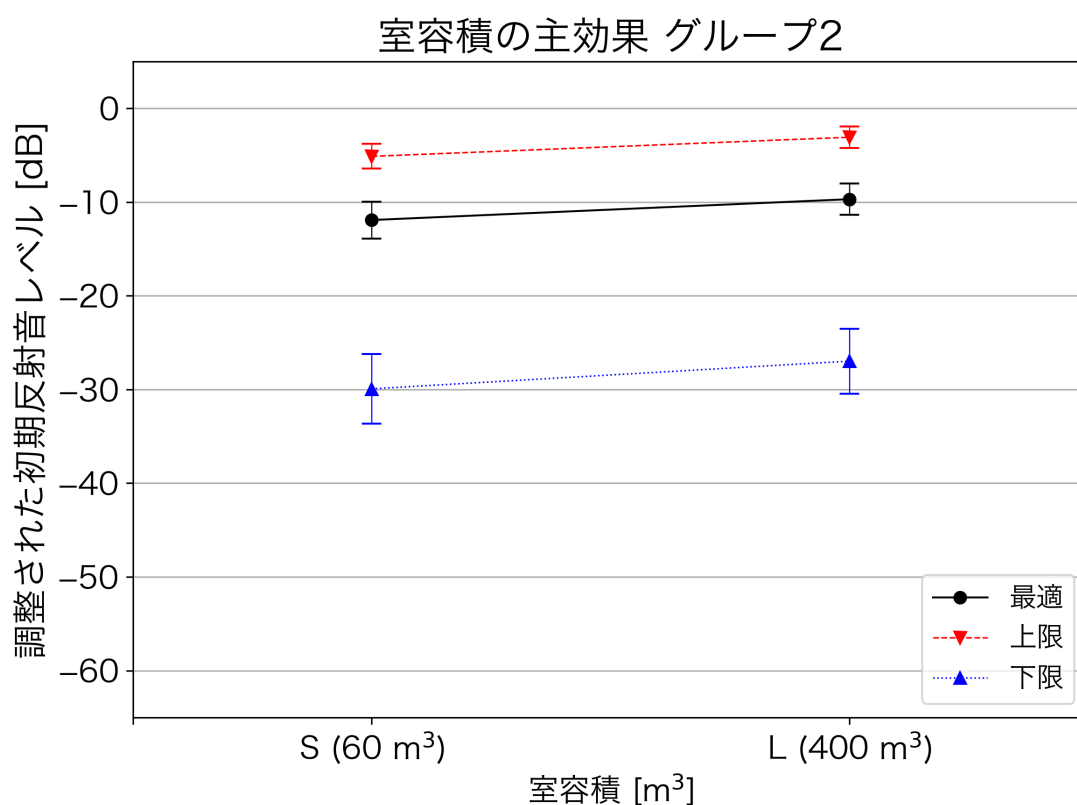


図 4.2.12 被験者グループ 2 における室容積 S (60 m^3) および室容積 L (400 m^3) の初期反射音レベルの平均値および 95% 信頼区間

4.2.6 グループ別の評価語 × 到来方向の交互作用

図 4.2.13 は、被験者グループ 1 における音色の明瞭さおよび音像の明瞭さそれぞれについて、7 つの到来方向の初期反射音レベルの平均値および 95% 信頼区間を示す。グループ 1 では、これらの評価語 × 到来方向の交互作用に有意差は見られなかった。

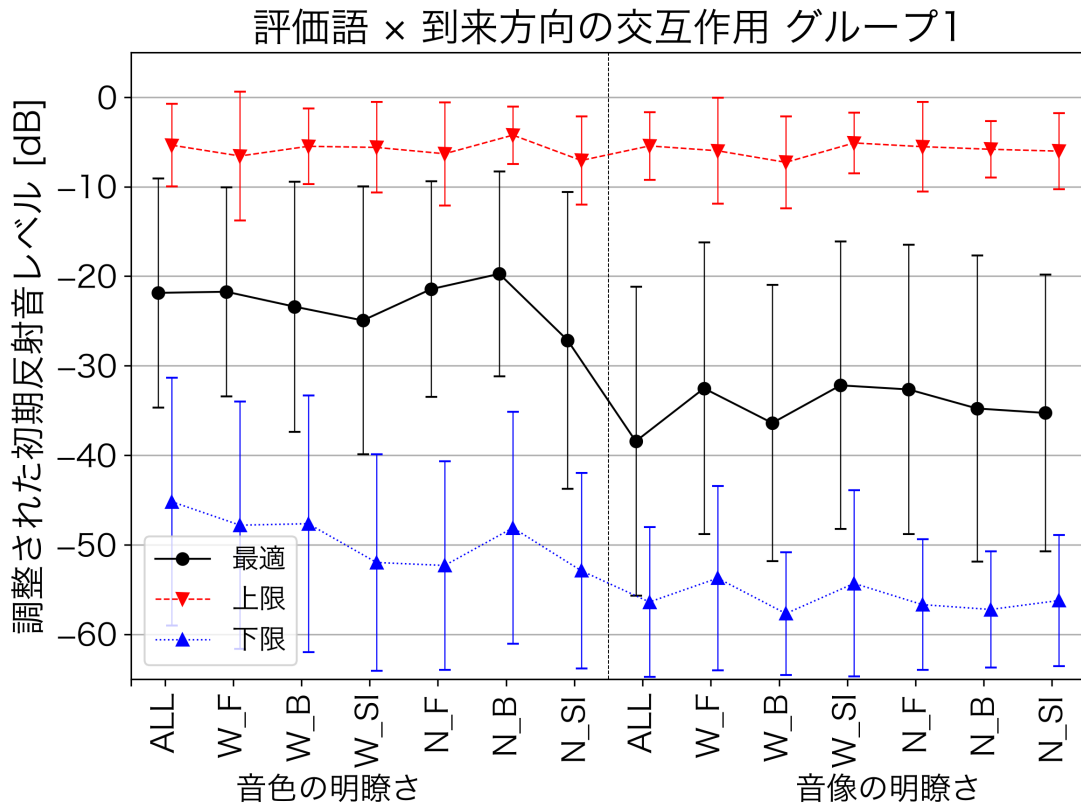


図 4.2.13 被験者グループ 1 における音色の明瞭さおよび音像の明瞭さそれぞれの 7 つの到来方向の初期反射音レベルの平均値および 95% 信頼区間

図 4.2.14 は、被験者グループ 2 における音色の明瞭さおよび音像の明瞭さそれぞれについて、7 つの到来方向の初期反射音レベルの平均値および 95% 信頼区間を示す。グループ 2 では、最適レベルで交互作用に有意差が、上限レベルで交互作用に有意傾向が見られた。

■最適レベル 多重比較より、最適レベルでは W_F で評価語の違いに有意差が、W_B で評価語の違いに有意傾向が見られた。更に、どちらの明瞭さにおいても到来方向の違いに有意差が見られたが、続く多重比較では各対の間に有意差は見られなかった。

■上限レベル 多重比較より、上限レベルでは音色の明瞭さにおいて到来方向の違いに有意差が見られた。しかし、最適レベル同様に、続く多重比較では各対の間に有意差は見られなかった。

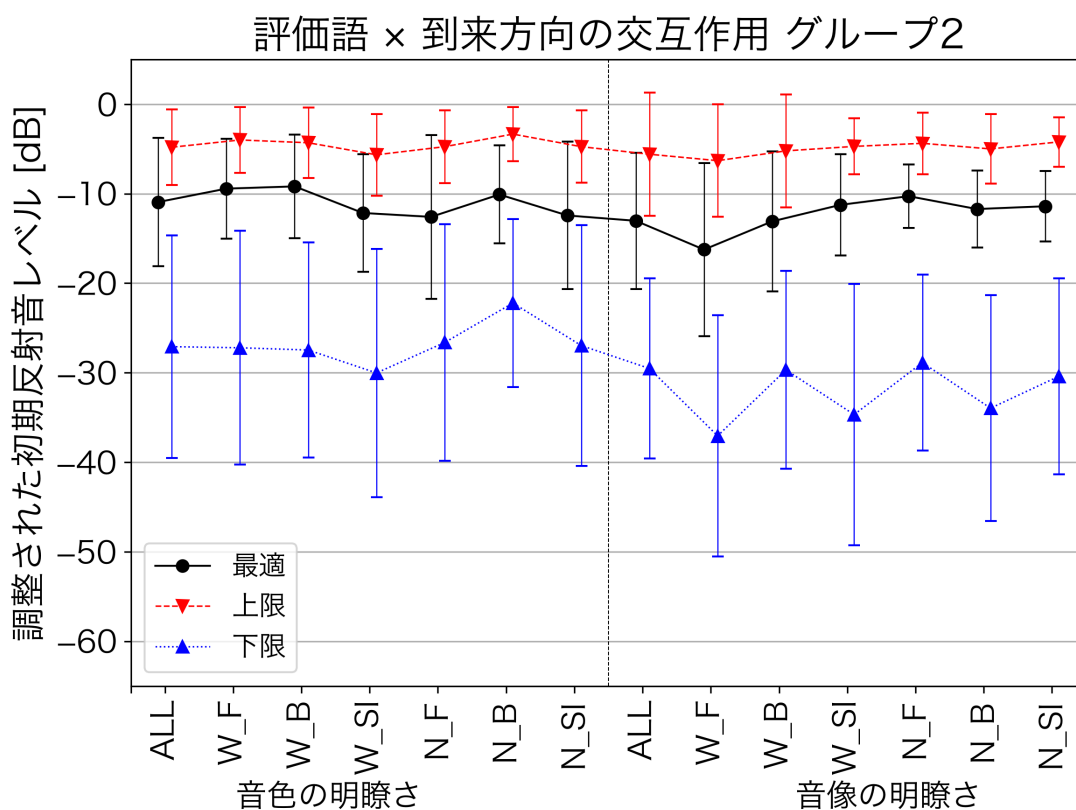


図 4.2.14 被験者グループ 2 における音色の明瞭さおよび音像の明瞭さそれぞれにの 7 つの到来方向の初期反射音レベルの平均値および 95% 信頼区間

4.3 考察

4.3.1 着目する明瞭さ、および聴取音源の違いが最適な初期反射音レベルに与える影響

被験者14名から得られた初期反射音レベルの分析の結果、音源 × 評価語の交互作用に有意差が見られた。表4.3.1は、聴取音源ごと、評価語ごとの初期反射音レベルを示す。表内の数値は最適レベルの平均値を、括弧内は上限レベルの平均値、および下限レベルの平均値を示す。表より、音色の明瞭さに着目して無響音源を聴取した際の初期反射音レベルがその他の条件よりも高くなっていることが分かる。音色の明瞭さに着目した際の市販音源の最適レベルは -22.0 dBである一方、無響音源では -10.2 dBに調整されており、約 12 dBの差が見られる。また、上限レベルにおいては約 5 dB、下限レベルにおいては約 17 dBの差が見られる。聴取する音源、および着目する明瞭さの種類によって、最適な初期反射音レベルは有意に異なることが示唆された。同時に、初期反射音レベルは高ければ高いほど明瞭という訳ではなく、最適値が存在する可能性も示唆された。

表 4.3.1 被験者14名全員の結果に基づく、評価語ごと、聴取音源ごとの初期反射音の最適レベルの平均値。括弧内は下限レベルの平均値～上限レベルの平均値を示す。図4.2.3と対応している。

	音色の明瞭さ	音像の明瞭さ
市販音源	-22.0 dB (-45.0 dB ~ -7.5 dB)	-24.6 dB (-45.0 dB ~ -7.5 dB)
無響音源	-10.2 dB (-28.0 dB ~ -2.7 dB)	-19.3 dB (-37.9 dB ~ -3.8 dB)

音源 × 室容積の交互作用の分析の結果、室容積が S (60 m³)、 L (400 m³) のどちらの場合でも、無響音源が市販音源よりも約 10 dB高い初期反射音レベルで明瞭だと知覚されることが明らかになった。聴取音源によって最適な初期反射音レベルが異なった理由の1つとして、自然な聴取のために音源そのものが響きを必要とするかどうかに関係している可能性がある。本実験で用いた無響音源は、それ自体に響きを含まない単体の楽器音である。コンサートホールはこのような楽器音に響きを付加することを目的に設計されており、我々はそうした音に慣れていると言える。無響空間で楽器を聴取すると、往々にして不自然で音量感が不足しがちと感じるのはこのためだと言える。一方で、市販音源は制作の過程で響きが付加されていることが多く、音源自体に響きを含む。つまり、無響音源と比較して自然な聴取に必要とされる響きは少ないと考えられる。

あるいは、音源のエンベロープ変化が理由である可能性も考えられる。入交(2013)は、音量変化の激しい楽曲ほど残響を感じやすい傾向があることを明らかにした。楽曲のアタックが強かったり減衰が激しいなど、楽曲信号のエンベロープ変化が大きい場合は、一定の残響が付加された場合でも残響音が楽音によってマスクされないため、残響音の知覚上のラウドネスが大きくなるとしている。楽曲によって残響音が異なるラウドネスとして知覚されることは、Griesinger(1995a,b)も

同様に指摘している。本実験で使用した無響音源は、無響室で録音されたドライソースである。図4.3.1に、無響音源の信号から0～3秒の範囲を抜き取り、二乗平均波形をデシベルで表記したグラフを示す。アタックが強く音の減衰が急峻で、楽曲信号のエンベロープ変化が大きいことが分かる。ある瞬間に鳴った楽音は、前の楽音に付帯する残響音の影響を受けづらいつけることができる。これによって、無響音源においては初期反射音レベルが高い場合でも、Beranek (2004) の水平的な明瞭さを損ねること無く直接音が補強されたと考えることができる。一方、音像の明瞭さは同時に鳴っている音の区別に関する感覚である。このため、楽曲信号のエンベロープに関わらず一定の低い初期反射音レベルが最適とされたと考えられる。しかし、本実験で使用した音源は2種類のみであり、信号のエンベロープ変化の違いも極端である。更に、直接音の想定がモノラルとステレオという違いも存在するため、結果の一般化のためには更なる検討が必要である。

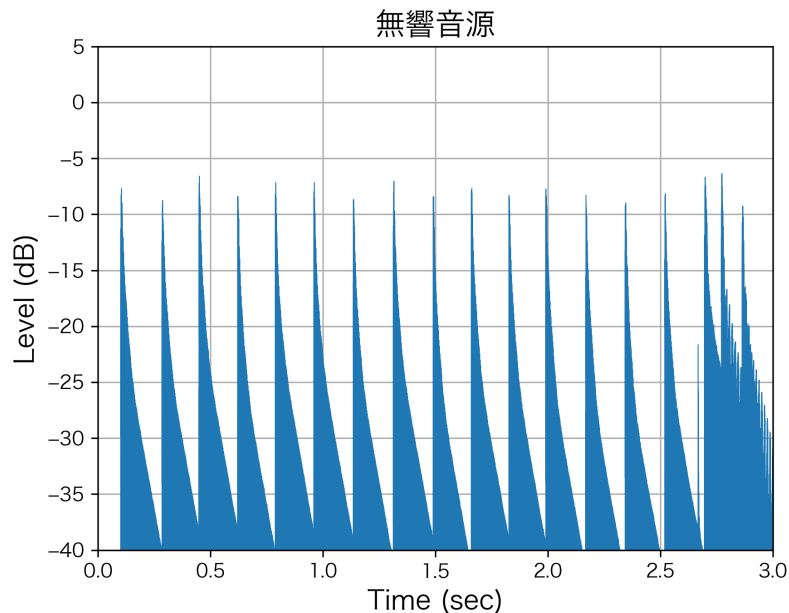


図 4.3.1 本実験で使用した無響音源の信号から0～3秒の範囲を抜き取り、二乗平均波形をデシベル表記で表したグラフ。

図 4.3.2 に、本実験で使用した市販音源の信号から0～3秒の範囲を抜き取り、二乗平均波形をデシベルで表記したグラフを示す。音の減衰は穏やかであり、楽曲信号のエンベロープ変化は小さい。楽音と楽音の間の時間差は小さく、ある瞬間に鳴っている楽音は、前の楽音に付帯する残響音によって影響されやすいと言える。これにより、市販音源は無響音源と比べて、初期反射音レベルが高くなると水平的な明瞭さが損なわれてしまったと考えることができる。

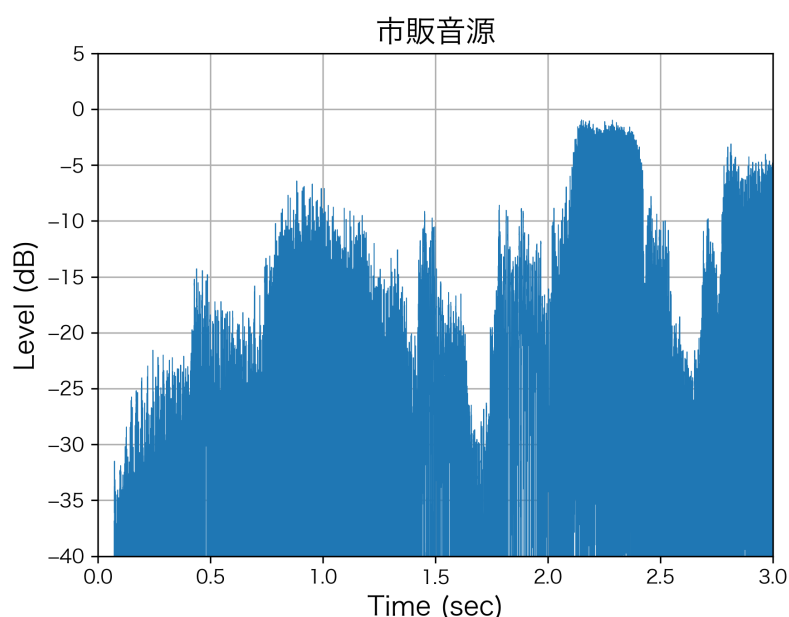


図 4.3.2 本実験で使用した市販音源の信号から 0～3 秒の範囲を抜き取り、二乗平均波形をデシベル表記で表したグラフ。

表 4.3.2 および表 4.3.3 は、被験者グループ 1 およびグループ 2 の結果に基づき、評価語ごと、聴取音源ごとの調整された初期反射音レベルを示す。どちらのグループも前述した全体結果とほぼ同じ傾向を示しており、無響音源の聴取で音色の明瞭さに着目した場合、他の条件よりも高い初期反射音レベルが最適とされている。一方、グループ 1 は総じて最適レベルが低く、グループ 2 はグループ 1 よりも約 10 dB～20 dB、最適レベルが高い。これは、明瞭さのために最適だと感じられる音圧レベルには個人差が存在する、あるいは個人によって着目している感覚が異なる可能性を示唆している。例えば、同様に音色の明瞭さに着目を求めても、グループ 1 は音色の周波数特性に着目し、グループ 2 は音量感や距離感に着目して判断した可能性もある。この場合、明瞭さの指標の提案のためには、周波数特性に関する明瞭さと、音量感や距離感に関する明瞭さを分ける必要が出てくる。

表 4.3.2 グループ 1 の 6 名の結果に基づく評価語ごと、聴取音源ごとの初期反射音の最適レベルの平均値。括弧内は下限レベルの平均値～上限レベルの平均値を示す。

	音色の明瞭さ	音像の明瞭さ
市販音源	-30.0 dB (-55.9 dB ～ -7.8 dB)	-35.6 dB (-55.8 dB ～ -7.8 dB)
無響音源	-15.8 dB (-43.0 dB ～ -3.8 dB)	-33.6 dB (-53.1 dB ～ -4.7 dB)

表 4.3.3 グループ2の8名の結果に基づく評価語ごと、聴取音源ごとの初期反射音の最適レベルの平均値。括弧内は下限レベルの平均値～上限レベルの平均値を示す。

	音色の明瞭さ	音像の明瞭さ
市販音源	-15.9 dB (-36.9 dB ~ -7.2 dB)	-16.3 dB (-36.9 dB ~ -7.2 dB)
無響音源	-6.0 dB (-16.7 dB ~ -1.8 dB)	-8.6 dB (-26.4 dB ~ -3.1 dB)

4.3.2 室容積の違いが最適な初期反射音レベルに与える影響

最適レベルおよび上限レベルで、室容積 L が室容積 S よりも約 3~5 dB 高い初期反射音レベルが明瞭とされた。本実験で使用した刺激には 80 ms 以降の後期残響音が含まれていないため、室容積の違いは初期反射音の密度および減衰の違いに関係していると考えられる。

この結果は、室容積が小さいほど、つまり初期反射音の密度が大きく減衰が急であるほど、直接音に対する初期反射音の影響が強いことを示唆している。室容積が小さい場合は 80 ms の中でも遅延時間の短い反射音のエネルギーが大きくなるため、室容積が大きい場合と比較して聴取に対する影響が高いものと思われる。一方で、室容積 S (60 m³) および室容積 L (400 m³) の違いに対して、最適とされる初期反射音レベルは約 3~5 dB ほどの違いであったため、その影響は聴取する音源や着目する明瞭さによる違いほど大きくはない。

4.3.3 初期反射音の到来方向の違いが最適な初期反射音レベルに与える影響

到来方向が ALL、W_F、W_B、W_SI の音場の聴取において、最適レベルに対する評価語の違いに有意差が見られた。ALL、W_F、W_B、W_SI の 4 つの音場の最適レベルの音色の明瞭さの平均値は -15.8 dB、音像の明瞭さの最適レベルの平均値は -22.9 dB であり、着目する評価語による違いは約 7.1 dB であった。ALL、W_F、W_B、W_SI の 4 つの音場は、その他の音場と比べて初期反射音の到来方向の偏りが小さく、広範囲から初期反射音が到来する音場である。初期反射音の到来方向の偏りが小さい場合に、音色の明瞭さと音像の明瞭さの区別がされやすい可能性が示唆された。

被験者グループごとに評価語 × 到来方向の交互作用を分析した結果、グループ 2 は W_F で評価語の違いに有意差が、W_B で評価語の違いに有意傾向が見られた。これは、グループ 1 が初期反射音の到来方向に着目して評価をしないグループであり、グループ 2 は到来方向にも着目して評価するグループであることを示唆している。または、前述したようにグループ 2 は高い初期反射音レベルを明瞭だと判断しており、聴取する初期反射音レベルが比較的高かったために到来方向の違いを区別することができた、とも考えられる。以上より、個人によって影響の度合いは異なるが、初期反射音の到来方向の偏りや空間分布によっても明瞭さは影響を受けることが明らかになった。

4.4 結論

本実験では、被験者に対し初期反射音の時間的・空間的特徴の異なる複数の音場における音楽聴取を求め、最も明瞭だと感じる値まで初期反射音レベルの調整を求めた。聴取音源としてステレオの市販音源、およびモノラルの無響音源の 2 種類を使用し、評価の基準として音色の明瞭さ、および音像の明瞭さの 2 つに着目を求めた。明瞭さのためにどの程度の初期反射音レベルが最適とされたかを分析した結果、以下の結論が得られた。

1. 聴取する音源によって、最適な初期反射音レベルは異なる。
2. 着目する明瞭さの種類によって、最適な初期反射音レベルは異なる。
3. 市販音源の聴取においては、無響音源と比べて約 10 dB 低い初期反射音レベルで明瞭だとされる。
4. 無響音源の聴取においては、音色の明瞭さは音像の明瞭さよりも約 10 dB 高い初期反射音レベルで明瞭だとされる。
5. 最適な初期反射音レベルは室容積が小さいほど低くなる。
6. 最適な初期反射音レベルには約 10～20 dB 程の個人差が存在する。
7. 高い初期反射音レベルを明瞭だと評価する被験者は、初期反射音の空間的な分布や到来方向の偏りにも着目して明瞭さの評価を行う可能性がある。

第5章

実験2：音楽聴取における明瞭さの評価要因

音楽の明瞭さという言葉は抽象度が高く、具体的にどのような感覚を指しているのか曖昧である。音楽の明瞭さが複数の異なる感覚を指す言葉である場合、明瞭さは要素感覚として扱えず、物理量との対応関係の解明は困難になる。

本実験では、音楽聴取において人が明瞭だと感じる要因を明らかにすることを目的とし、複数の被験者に対しレパートリーグリッド法を参考にしたインタビュー調査を実施した。被験者に複数の音場での音楽聴取を模擬した刺激の比較を求め、明瞭に感じた刺激の選択、またその理由についての自由記述を求めた。得られた音楽の明瞭さに関する自由記述から、形容詞句の評価語として音楽の明瞭さの評価要因を抽出した。

5.1 実験方法

5.1.1 レポートリーグリッド法

レポートリーグリッド法は、Kelly (1992) によって提唱された人間の認知構造を把握するための手法である。レポートリーグリッド法では被験者に刺激の比較評価を求め、その判断理由を尋ねることで刺激の評価項目を被験者自身の言葉によって抽出する。音響の分野での適応事例も多く、代表的なものとしては、Berg and Rumsey (1999, 2006) によるオーディオ聴取における空間品質評価の要因を調査した一連の研究が挙げられる。Berg らは複数の異なる録音・再生手法による音楽刺激を 3 つの組で被験者に提示し、似ている刺激の組とその類似点について、またそれらと異なる刺激とその相違点について聞くことで、音楽聴取において録音・再生手法が異なる場合に起こりうる印象の変化要因を明らかにした。彼らの実験の結果、naturalness、source localization、envelopment、depth 等の音楽聴取における空間印象に関する評価項目が抽出され、その後これらの項目は多くの研究者によって用いられている。

5.1.2 実験刺激の作成

羽入らによる仮想音源分布を用いた音場の可聴化の原理に基づき、室容積が 125 m^3 (総表面積 150 m^2) および 512 m^3 (総表面積 384 m^2) の 2 種類、 $30 \text{ ms} \sim 80 \text{ ms}$ までの初期反射音の到来方向が前方、後方、側方の 3 種類、室の平均吸音率 $\bar{\alpha}$ が 0.4 の合計 6 種類の評価音場を再現するインパルス応答を生成した。

室容積は、1 辺が 5 m ほどのリスニングルーム、および 8 m ほどのリスニングルームを想定した。これらの室容積は、リスニングルームに推奨される部屋の大きさや形状に関する記載のある ITU-R BS1116-3 (ITU-R, 2015) において、 $20 \text{ m}^2 \sim 70 \text{ m}^2$ の床面積が推奨されていることから決定した。どちらの室も複雑さ係数 h は等しいため、相似形である。初期反射音の到来方向が前方の音場は、初期反射音の生成範囲を 2～3 番および 14～15 番のブロック (第 3.2.1 項を参照)、後方の音場は 6～7 番および 10～11 番、側方の音場は 4～5 番、12～13 番のブロックに限定した。直接音の距離減衰は考慮せず、直接音のエネルギーは音響出力 W_0 と等しい 1 としてインパルス応答に合成した。

実験 1 では 80 ms 以降の後期残響音を除外したが、本実験では一般的な小空間の音場の変化によって引き起こされる、なるべく多様で現実に即した音楽の明瞭さの印象の違いに着目を求めたかったため、 80 ms 以降の後期残響音もインパルス応答に含めた。これにより、室容積が異なる音場の刺激は、反射音密度の違いにより算出された反射音の音圧レベルが異なった。音量感以外の明瞭さの評価要因にも着目してもらうため、 125 m^3 の音場のインパルス応答の実効値を 512 m^3 のものと等しくなるまで 2.0 dB 下げた音場を加え、評価音場は 9 種類とした。

表 5.1.1 に、本実験で生成した 9 種類の音場の名称および生成パラメータを示す。音場の名称は

「室容積_到来方向_再生音圧の補正の有無」となっており、室容積 V が 125 m^3 を S (Small)、室容積 V が 512 m^3 を L (Large) で示す。再生音圧を 2.0 dB 下げた音場は、マイナス記号 $-$ を添えている。表 5.1.2 に、生成された評価音場の物理量を示す。物理量は実測値ではなく、生成したインパルス応答から直接算出したものである。図 5.1.1～図 5.1.27 に、それぞれの音場の 3 次元仮想音源分布、2 次元仮想音源分布（仮想音場の上方から俯瞰した図）、音場の反射音の時系列応答を示す。仮想音源分布の図における点は各仮想音源を示し、中央の立方体および四角は仮想の聴取空間を示す。見やすさのため、3 次元および 2 次元の仮想音源分布の図には、仮想音源の総数 N_{total} 個から半数をランダムに選択し、表示した。

表 5.1.1 生成した 9 種類の評価音場の名称および生成パラメータ

刺激名	室容積 [m^3]	室総表面積 [m^2]	複雑さ係数 h	平均吸音率 $\bar{\alpha}$	初期反射音の到来方向	再生補正音圧 [dB]
S_Front	125	150	3.38	0.4	前方	0
S_Back	125	150	3.38	0.4	後方	0
S_Side	125	150	3.38	0.4	側方	0
S_Front-	125	150	3.38	0.4	前方	-2.0
S_Back-	125	150	3.38	0.4	後方	-2.0
S_Side-	125	150	3.38	0.4	側方	-2.0
L_Front	512	384	3.38	0.4	前方	0
L_Back	512	384	3.38	0.4	後方	0
L_Side	512	384	3.38	0.4	側方	0

表 5.1.2 本実験で生成した 9 種類の評価音場の名称および物理量。物理量は生成したインパルス応答に対して算出した。

刺激名	残響時間 [秒]	直接音に対する初期反射音レベル [dB]	直接音に対する後期残響音レベル [dB]	C_{80} [dB]
S_Front	0.26	-19.7	-31.4	31.5
S_Back	0.26	-19.7	-31.4	31.5
S_Side	0.26	-19.7	-31.4	31.5
S_Front-	0.26	-19.7	-31.4	31.5
S_Back-	0.26	-19.7	-31.4	31.5
S_Side-	0.26	-19.7	-31.4	31.5
L_Front	0.42	-21.7	-31.4	28.2
L_Back	0.42	-21.7	-31.4	28.2
L_Side	0.42	-21.7	-31.4	28.2

作成したインパルス応答を以下の 2 つの音源に畳み込み、実験刺激とした。

1. 市販音源（ステレオ）：Norah Jones, “Come Away With Me” CD (2002) Track1 : Don't Know Why

2. 無響音源（モノラル）：Archimedes Project：シロフォン (Hansen and Munch, 1991)

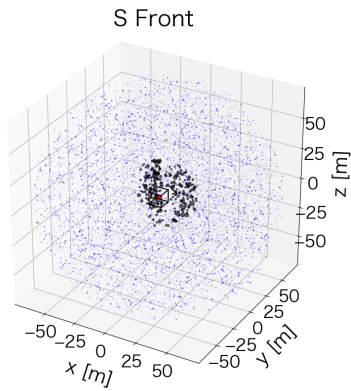


図 5.1.1 S_Front の音場の 3次元仮想音源分布

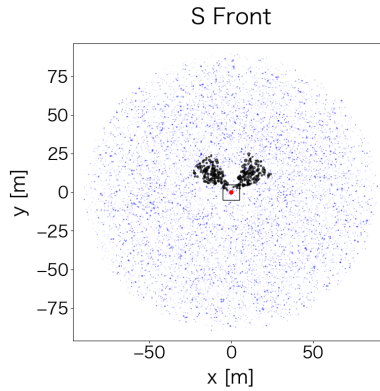


図 5.1.2 S_Front の音場の 2次元仮想音源分布

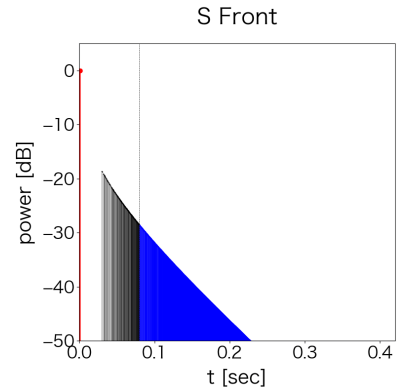


図 5.1.3 S_Front の音場の反射音の時系列応答

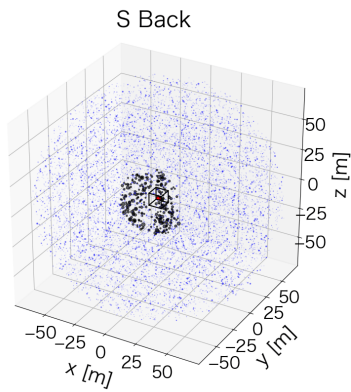


図 5.1.4 S_Back の音場の 3次元仮想音源分布

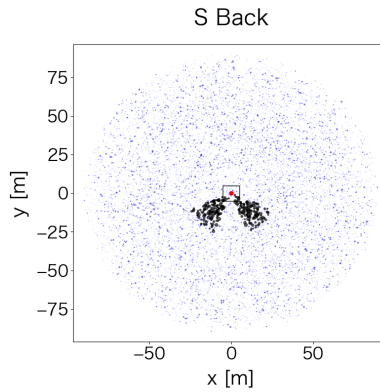


図 5.1.5 S_Back の音場の 2次元仮想音源分布

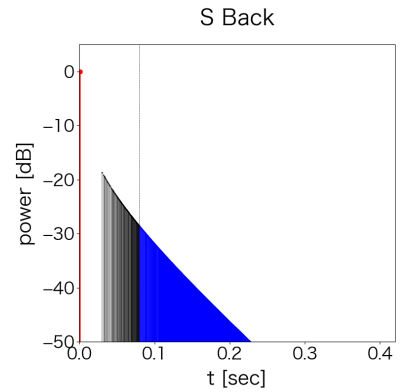


図 5.1.6 S_Back の音場の反射音の時系列応答

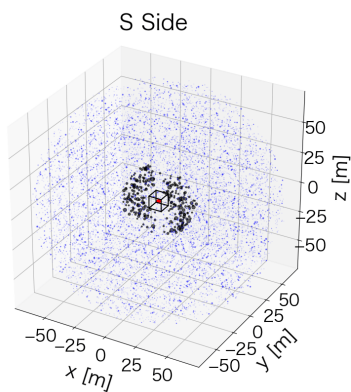


図 5.1.7 S_Side の音場の 3次元仮想音源分布

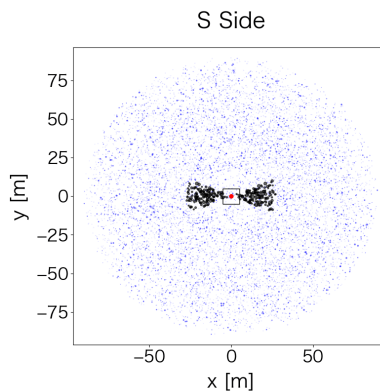


図 5.1.8 S_Side の音場の 2次元仮想音源分布

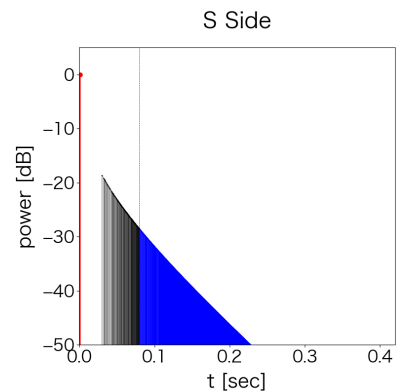


図 5.1.9 S_Side の音場の反射音の時系列応答

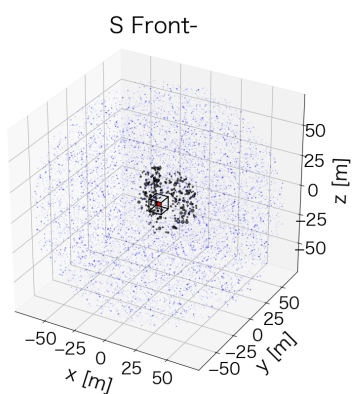


図 5.1.10 S_Front- の音場の 3次元仮想音源分布

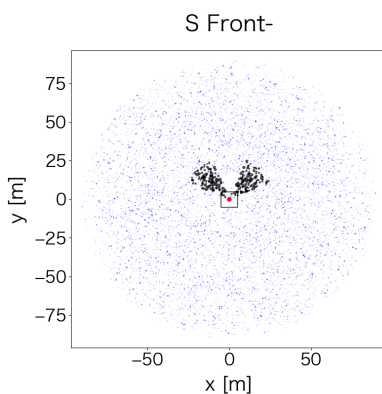


図 5.1.11 S_Front- の音場の 2次元仮想音源分布

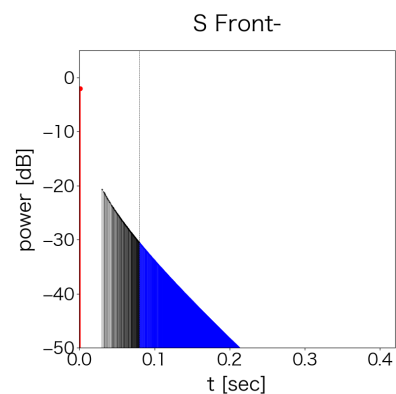


図 5.1.12 S_Front- の音場の 反射音の時系列応答

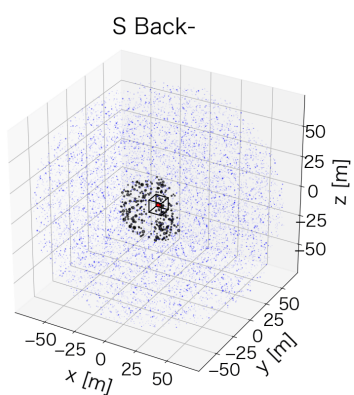


図 5.1.13 S_Back- の音場の 3次元仮想音源分布

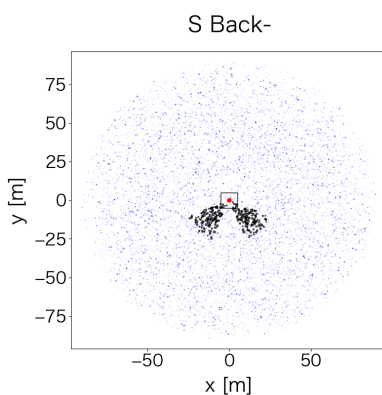


図 5.1.14 S_Back- の音場の 2次元仮想音源分布

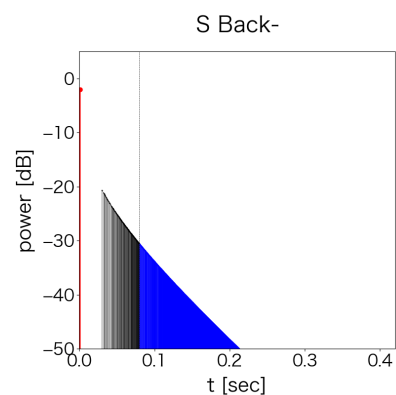


図 5.1.15 S_Back- の音場の 反射音の時系列応答

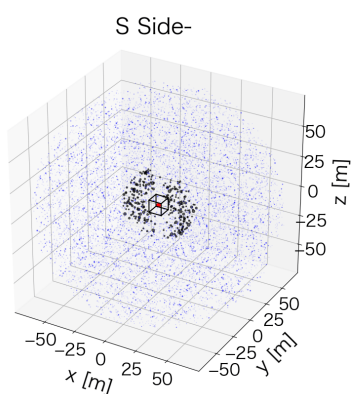


図 5.1.16 S_Side- の音場の 3次元仮想音源分布

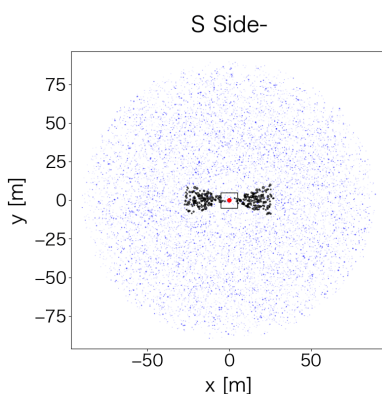


図 5.1.17 S_Side- の音場の 2次元仮想音源分布

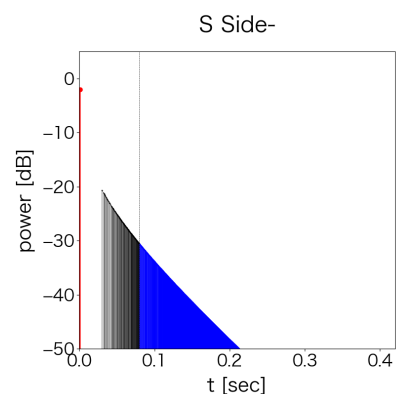


図 5.1.18 S_Side- の音場の 反射音の時系列応答

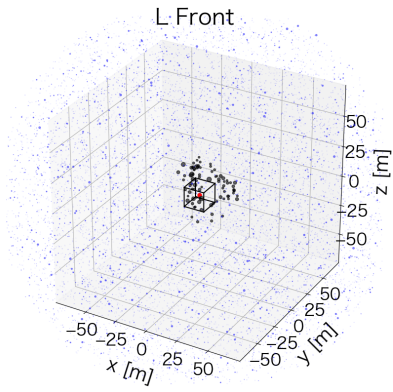


図 5.1.19 L.Front の音場の 3次元仮想音源分布

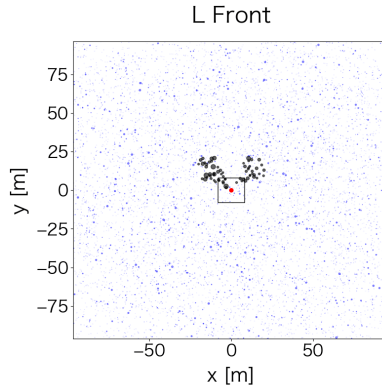


図 5.1.20 L.Front の音場の 2次元仮想音源分布

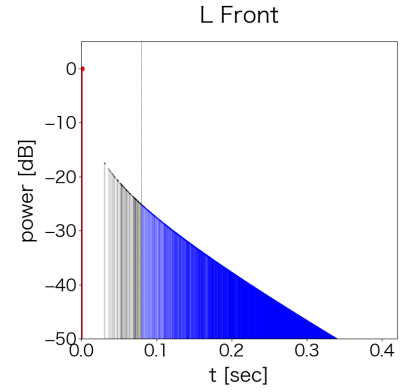


図 5.1.21 L.Front の音場の反射音の時系列応答

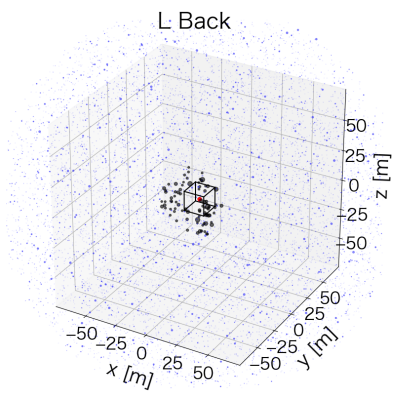


図 5.1.22 L.Back の音場の 3次元仮想音源分布

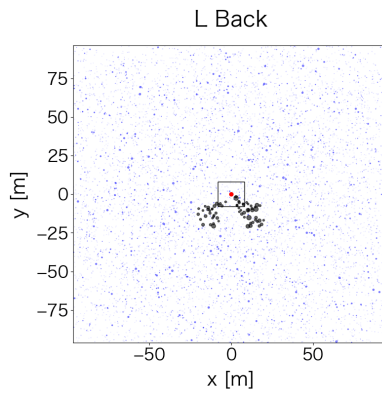


図 5.1.23 L.Back の音場の 2次元仮想音源分布

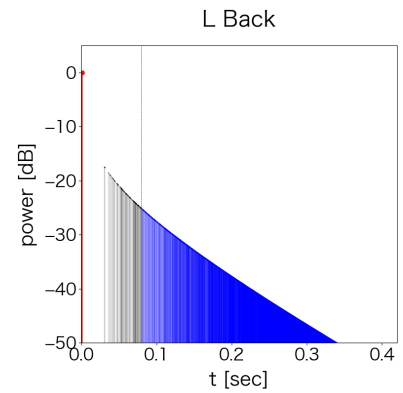


図 5.1.24 L.Back の音場の反射音の時系列応答

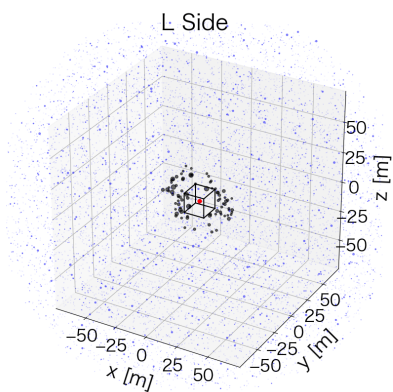


図 5.1.25 L.Side の音場の 3次元仮想音源分布

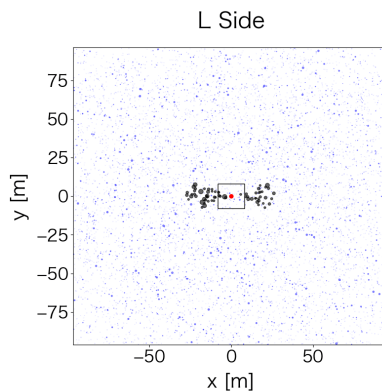


図 5.1.26 L.Side の音場の 2次元仮想音源分布

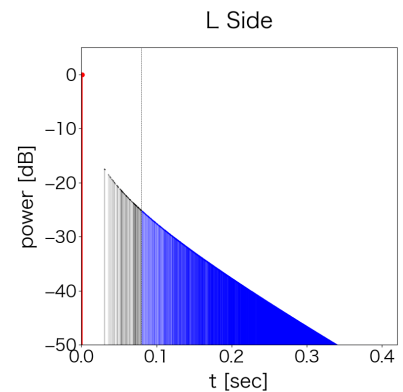


図 5.1.27 L.Side の音場の反射音の時系列応答

5.1.3 インタビュー調査

生成した 9 種類の評価音場の中から 2 つの音場の組み合わせをランダムに選択し、被験者に「刺激 A、刺激 B、刺激 A あるいは B」の 3 つの刺激の比較聴取を求めた。

本来、レパトリートグリッド法では、刺激空間を網羅する評価項目を収集するため、被験者に刺激同士の類似点と相違点の両方の記述を求める。一方、本実験は「なぜ明瞭だと判断したか」の調査が目的であるため、以下の手順に従い、レパトリートグリッド法よりも簡略化したインタビュー調査を実施した。

1. 3 つの刺激の中から、他の 2 つと異なって聴こえる刺激の回答を求めた。
2. 異なって聴こえた刺激と残りの 2 つの刺激を比較し、より明瞭に感じられた刺激の回答を求めた。
3. 一方を明瞭だと判断した理由について、自由記述を求めた。

被験者に提示した 3 つの刺激は、Cycling'74 社 Max7 で作成されたアプリケーションを操作することで、任意のタイミングで再生、停止および刺激の切り替えを行うことができた。刺激はループ再生された。8 名の被験者が市販音源を使用した実験に参加し、そのうち 7 名が無響音源を使用した実験に参加した。

5.2 分析

5.2.1 音場の明瞭さの順位

提示された刺激の組のどちらがより明瞭に感じたかについて回答を求めた結果から、サーストンの一対比較法に基づき 9 種類の音場の明瞭さの順位づけを行った。サーストンの一対比較法とは、被験者に対して刺激を 2 つずつ提示し、特定の判断基準に基づいてどちらかの刺激を選択させる方法である。選択された度数を観測度数として、サーストンの比較判断の法則から、刺激間においての特定の判断基準に対する間隔尺度を見出すことができる。

サーストンの一対比較法により、市販音源では 36 試行×8 名で 288 試行、無響音源では 36 試行×7 名で 252 試行の結果が得られた。市販音源では全試行中 7 試行、無響音源では 8 試行で、同じ刺激を対で提示したにも関わらず異なる刺激だと判断されたため、これらの試行結果については提示した A、B 両方の刺激の選択度を 0 とすることで対処し、分析した。

表 5.2.1 は、市販音源および無響音源それぞれの明瞭さに関する順位づけの結果を示す。市販音源では室容積の大きい音場の明瞭さの順位が高い一方、無響音源では室容積の小さい音場の順位が高い傾向が見られる。さらに、室容積に関わらず初期反射音の到来方向が前方、側方、後方の順で明瞭さの順位が低下する傾向が見られる。再生音圧による差はほとんど見られない一方で、市販音

源の聴取では S_Side- のみ、S_Side と比べて順位が高い。

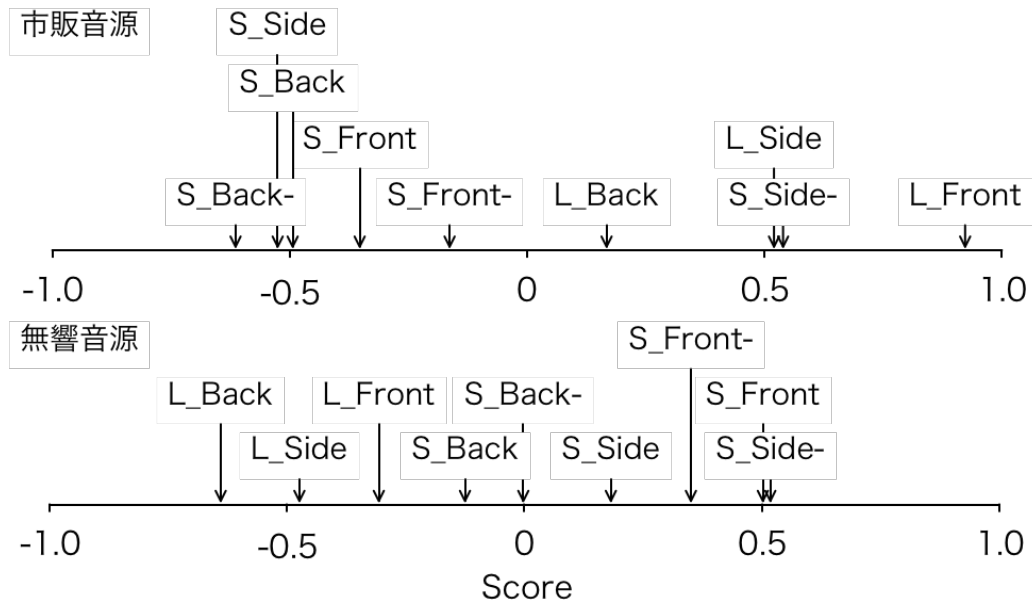


図 5.2.1 サーストンの一対比較法による 9 種類の音場の明瞭さの順位づけの結果。ラベルは表 5.1.1、表 5.1.2 の各音場の名称に対応している。Score の値は明瞭さの順位と対応している。

5.2.2 自由記述の分析

明瞭さを判断した理由について得られた自由記述から、明瞭さの評価要因を分析した。全ての回答が形容詞句の評価語に変換され、筆者によりほぼ同じ意味を指すと考えられたものは一つにまとめられた。例えば、「広がりがあり、楽器が分離して聴こえる」といった回答は、「音像が広い」、「分離している」と変換し、「空間的に広がっており、各楽器のかぶりが少ない。各周波数がまんべんなくきれいに聴こえる。」といった回答からは、「音像が広い」、「かぶりが少ない」、「バランスが良い」と変換した。また、「聴き取りやすい」、「聴きやすい」は同じ語句に、「こもった」、「くもった」等も同じ語句にまとめた。これらのまとめと変換は、筆者の主観や実験後に実施した被験者へのヒアリングを元に行った。一方、筆者が元の意味から大きく外れる可能性が少しでもあると判断したものは、異なる語句として残した。また、同じ刺激を異なると判断された比較から得られた回答については、分析から除いた。音場が明瞭であった理由については 42 個の評価語（回答数 495）が、不明瞭であった理由については 42 個の評価語（回答数 390）が得られた。

表 5.2.1 は、市販音源（8 名）および無響音源（7 名）の聴取から得られた、明瞭さの評価要因を示す。括弧内は回答数を示す。表の上から、筆者が「音色に関する評価語」と判断したもの、「音像に関する評価語」と判断したもの、「響きに関する評価語」と判断したもの、「これらどれにも分類されない評価語」と判断したものである。1 評価語あたりの回答数の平均値は 10.4 回であったため、10 回以上言及された言葉については見やすさのため太字にしている。

表 5.2.1 市販音源および無響音源の聴取から得られた自由記述の分析から得られた、明瞭さの評価要因となる評価語。括弧内は回答数を示す。上から音色に関するもの、音像に関するもの、響きに関するもの、その他の評価語である。10 回以上言及された言葉を太字で示す。

明瞭さの要因	不明瞭さの要因
<p>音色が明瞭 (35)、高域がしっかりした (22)、やわらかい (17)、アタックが強い (17)、音色が明るい (15)、バランスが良い (13)</p> <p>音色が自然 (6)、芯がある (5)、タッチがはっきりしている、倍音が綺麗、倍音が多い</p>	<p>こもっている (48)、高域がきつい (14)、つまった (14)、低音が出過ぎている (10)</p> <p>アタックが強すぎる (9)、高域が少ない (8)、音がかたい (8)、音色が不明瞭 (6)、コムフィルタがかかったような (5)、中音域が足りない (3)、やわらかい (3)、低域が弱い (2)、バランスが悪い (2)、コンプがかかったような (2)、芯がない (2)、劣化したような (2)、アタックが小さい、音色が不自然、倍音が少ない</p>
<p>輪郭がはっきりしている (62)、音像が広い (42)、定位が分かりやすい (21)、音が近い (15)、中心に音がまとまっている (14)、分離している (10)</p> <p>音像が狭い (9)、奥行きがある (6)、音像が前方に定位した (5)、包み込まれている (2)、分離している、横からの音がよく聴こえる</p>	<p>音が遠い (29)、中央に音がまとまりすぎている (26)、輪郭がぼやけた (15)</p> <p>音が近い (9)、音像が狭い (7)、音像が広い (6)</p>
<p>響きが少ない (31)、響きが自然 (12)</p> <p>かぶりが少ない (8)、低域の響きが豊か (5)、響きが多い (3)、響きが程よい (2)</p>	<p>響きが多い (53)</p> <p>響きが混ざっている (9)、響きが不自然 (9)、響きが少ない (8)、響きが分離している (7)、かぶっている (5)、部屋なりがある (3)、響きが色々な方向から聴こえる (3)</p>
<p>聴き取りやすい (38)、はっきり聴こえる (31)、すっきりしている (15)</p> <p>音量が大きい (6)、音が生っぽい (5)、伸びがある (4)、音が鋭い (4)、くっきりしている (3)、リアリティがある (2)、まろやかな (2)、まっすぐな、音に丸みがある、歯切れの良い</p>	<p>音がきつい (17)、聴き取りづらい (14)、うるさい (10)</p> <p>平べったい (7)、濁った (6)、痩せている (2)、つぶれた (2)、もやもやした (2)</p>

表 5.2.1 の最上段より、音色に関する評価要因として代表的なものには、「音色が明瞭 (35)、高域がしっかりした (22)、やわらかい (17)、アタックが強い (17)、音色が明るい (15)、バランスが良い (13)、こもっている (48)、高域がきつい (14)、つまった (14)、低音が出過ぎている (10)」等が見られた。これらをまとめると、音色の印象、アタック感、音色の明るさ、周波数的なバランスに関する要因だと言える。

表 5.2.1 の 2 段目より、音像に関する評価要因として代表的なものには、「輪郭がはっきりしている (62)、音像が広い (42)、定位が分かりやすい (21)、音が近い (15)、中心に音がまとまってい

る(14)、分離している(10)、音が遠い(29)、中央に音がまとまりすぎている(26)、輪郭がぼやけた(15)」等が見られた。輪郭がはっきりしているかどうか、定位が分かりやすいかどうかといった音像の明瞭さに関する要因と、音が近いか遠いかといった距離感、そして音像が広いかどうか、まとまり具合といったような、音像の広さに関する要因だと言える。

表5.2.1の3段目より、響きに関する評価要因として代表的なものは、ほとんどが「響きの量」に関するものであり、響きが少ないことが明瞭さの要因となり、多いことは不明瞭さの要因となっている。他にも、響きが自然かどうかといった響きの質に関する要因も見られる。

表5.2.1の最下段より、その他の要因として代表的なものには、「聴き取りやすい(38)、はっきり聴こえる(31)、すっきりしている(15)、音がきつい(17)、聴き取りづらい(14)、うるさい(10)」等が見られた。聴き取りやすさや音の大きさといった要因と、比較的抽象的な「はっきり」、「すっきり」といった要因が見られた。

5.2.3 聴取音源に基づく評価要因の分類

ソーストンの一対比較法による音場の順位づけの結果より、聴取する楽曲によって着目される要因が異なる可能性が得られたため、聴取音源ごとの評価要因の分析を行った。表5.2.2は市販音源の聴取から得られた明瞭さの評価要因を、表5.2.3は、無響音源の聴取から得られた明瞭さの評価要因を示す。

音色に関する評価語を比較すると、市販音源の聴取では「音色が明瞭、高域がしっかりした、バランスが良い、こもっている」等の、音色や周波数特性のバランスに関する要因が主な評価要因となっていることが分かる。一方、無響音源の聴取では、同様な傾向に加え、「アタックが強い、音色が明るい、やわらかい、つまった」といった、音色の質に関する要因が多く見られる。音像に関する評価語を比較すると、市販音源の聴取では「音像が広い、輪郭がはっきりしている、定位が分かりやすい、中心に音がまとまっている、中央に音がまとまりすぎている」等、音像の幅や輪郭に関わる要因が多く見られる。一方で、無響音源の聴取では「音が近い、音が遠い」といった距離感が主な要因となっていることが分かる。また、「中央に音がまとまっている」という要因は市販音源の聴取のみに表れている。響きに関する評価語では、市販音源、無響音源のどちらにおいても響きの量に言及されているが、無響音源の聴取の場合だけ、響きが少ないことが明瞭さの要因になっている。その他の評価語では、市販音源では「聴き取りやすさ」に関する要因が最も多く、一方で無響音源では「はっきり聴こえる」という回答が多い。

表 5.2.2 市販音源の聴取から得られた、明瞭さの評価要因となる評価語。括弧内は回答数を示す。上から音色に関するもの、音像に関するもの、響きに関するもの、その他の評価語である。10 回以上言及された言葉を太字で示す。

市販音源の明瞭さの要因	市販音源の不明瞭さの要因
<p>音色が明瞭 (23)、高域がしっかりした (12)、バランスが良い (11) やわらかい (7)、音色が自然 (4)、音色が明るい、アタックが強い、芯がある</p>	<p>こもっている (40)、低音が出過ぎている (10) 音がかたい (6)、音色が不明瞭 (5)、コムフィルタがかかったような (4)、低域が弱い (2)、バランスが悪い (2)、高域がきつい (2)、コンプがかかったような (2)、芯がない、つまった、劣化したような、アタックが強すぎる</p>
<p>音像が広い (38)、輪郭がはっきりしている (34)、定位が分かりやすい (14)、中心に音がまとまっている (10)、分離している (10) 音像が狭い (6)、奥行きがある (5)、音像が前方に定位した (5)、音が近い (2)、包み込まれている (2)、横からの音がよく聴こえる</p>	<p>中央に音がまとまりすぎている (26) 輪郭がぼやけた (8)、音が近い (8)、音が遠い (7)、音像が広い (5)、音像が狭い (5)</p>
<p>響きが少ない (7)、かぶりが少ない (7)、響きが自然 (6)、低域の響きが豊か (3)</p>	<p>響きが多い (27) 響きが混ざっている (7)、響きが分離している (7)、響きが不自然 (5)、部屋なりがある (3)、かぶっている (2)、響きが少ない</p>
<p>聴き取りやすい (30)、すっきりしている (12)、はっきり聴こえる (11) 音が生っぽい (3)、リアリティがある (2)、音量が大きい (2)、伸びがある (2)</p>	<p>聴き取りづらい (12)、音がきつい (14)、うるさい (10) 平べったい (2)、濁った、痩せている</p>

表 5.2.3 無響音源の聴取から得られた、明瞭さの評価要因となる評価語。括弧内は回答数を示す。上から音色に関するもの、音像に関するもの、響きに関するもの、その他の評価語である。10 回以上言及された言葉を太字で示す。

無響音源の明瞭さの要因	無響音源の不明瞭さの要因
アタックが強い (16)、音色が明るい (14)、音色が明瞭 (12)、高域がしっかりした (10)、やわらかい (10) 芯がある (4)、音色が自然 (2)、バランスが良い (2)、タッチがはっきりしている、倍音が綺麗、倍音が多い	つまった (13)、高域がきつい (12) こもっている (8)、アタックが強すぎる (8)、高域が少ない (8)、中音域が足りない (3)、やわらかい (3)、音がかたい (2)、アタックが小さい、芯のない、音色が不明瞭、コムフィルタがかかったような、劣化したような、音色が不自然、倍音が少ない
輪郭がはっきりしている (28)、音が近い (13) 定位が分かりやすい (7)、音像が広い (4)、中心に音がまとまっている (4)、音像が狭い (3)、奥行きがある、分離している	音が遠い (22) 輪郭がぼやけている (7)、音像が狭い (2)、音像が広い、音が近い
響きが少ない (24) 響きが自然 (6)、響きが多い (3)、響きが程よい (2)、低域の響きが豊か (2)、かぶりが少ない	響きが多い (26) 響きが少ない (7)、響きが不自然 (4)、響きが色々な方向から聴こえる (3)、かぶりがある (3)、響きが混ざった (2)
はっきり聴こえる (20) 聴き取りやすい (8)、音量が大きい (4)、音が鋭い (4)、すっきりしている (3)、くっきりしている (3)、音が生っぽい (2)、伸びがある (2)、まるやかな (2)、まっすぐな、音に丸みがある、歯切れの良い	濁った (5)、平べったい (5)、音がきつい (3)、つぶれた (2)、もやもやした (2)、聴き取りづらい (2)、痩せている

5.2.4 室容積に基づく評価要因の分類

サー斯顿の一対比較法による順位づけの結果、市販音源では室容積の大きい音場の明瞭さの順位が高い一方、無響音源では室容積の小さい音場の順位が高い傾向が見られた。この判断理由について検討を行うため、各音源ごとに、室容積のみが異なる刺激の比較から得られた評価語とその出現頻度を分析した。表 5.2.4 および表 5.2.5 は、市販音源および無響音源の聴取それぞれについて、S_Front- および L_Front、S_Back- および L_Back、S_Side- および L_Side の比較から得られた評価語を示す。表の太字は、2 回以上言及された評価語を示す。市販音源、無響音源のどちらでも「輪郭がはっきりしている」が最も多く言及されている。一方、市販音源では「すっきりしている」や「音像が広い」、そして「こもっていない」が、無響音源では「音が近い」や「響きが少ない」が明瞭さの理由とされている。また、どちらの音源においても「響きが少ない」が明瞭さの要因として、「響きが多い」が不明瞭さの要因として表れている。一方、市販音源で明瞭とされたのは室容積が大きい音場であり、無響音源で明瞭とされたのは室容積が小さい音場である。

表 5.2.4 室容積が異なる刺激の比較において得られた評価語とその出現頻度。市販音源の結果。2 回以上言及された言葉を太字で示す。

室容積が大きい刺激の明瞭さの要因（市販音源）	室容積が小さい刺激の不明瞭さの要因（市販音源）
輪郭が はっきりしている (6)、 すっきりしている (4)、音像が 広い (3)、聴き取りやすい (2)、かぶりが少ない (2)、音色が 明瞭 (2)、音像が 狭い (2)、響きが 自然 (2)、 バランスが良い (2) 高域が はっきりした 、中心に音が まとまっている	こも っている (6)、響きが 多い (5)、聴き取り づらい (4)、音が きつい (3)、うる さい (2)、中心に音が まとまりすぎている (2)、低音が 出すぎ (2) 音色が 不明瞭 、 コムフィルタがかかった ような、 にごった 、響きが 混ざっている 、響きが 不自然 、 面が感じられる

表 5.2.5 室容積が異なる刺激の比較において得られた評価語とその出現頻度。無響音源の結果。2 回以上言及された言葉を太字で示す。

室容積が小さい刺激が明瞭な要因（無響音源）	室容積が大きい刺激が不明瞭な理由（無響音源）
輪郭が はっきりしている (8)、 近い (5)、響きが 少ない (4)、 アタックが強い (2) 聴き取り やすい 、音像が 広い 、響きが 自然 、 定位が分かりやすい 、音色が 自然 、音色が 明るい 、 生っぽい 、音色が 明瞭	遠い (7)、響きが 多い (4) やわらかい 、 ぼやけた 、高域が 少ない 、 つまった 、音像が 狭い 、 にごった 、響きが 不自然

5.2.5 初期反射音の到来方向に基づく評価要因の分類

サーストンの一対比較法による順位づけの結果、室容積の違いに関わらず初期反射音の到来方向が前方、側方、後方の順で明瞭さの順位が低下する傾向が見られた。この判断理由についての検討を行うため、各音源ごとに、初期反射音の到来方向のみが異なる刺激の比較から得られた評価語とその出現頻度を分析した。表 5.2.6～表 5.2.11 は、市販音源および無響音源の聴取それぞれについて、初期反射音の到来方向のみが異なる刺激の比較から得られた評価語を示す。表の太字は、2 回以上言及された評価語を示す。初期反射音が前方から到来する音場の明瞭さの要因としては、「輪郭がはっきりしている」、「やわらかい」、「聴き取りやすい」、「音が中央に集まっている」、「音像が広い」、「響きが少ない」、「音色が自然」等が見られる。初期反射音が側方から到来する音場の明瞭さの要因としては、「広がりがある」、「音像が広い」、「奥行きがある」、「定位が分かりやすい」、「輪郭がはっきりした」が見られる。これらは全て音像に関する評価語であると言える。到来方向が側方の不明瞭さの要因としては、「こもっている」が見られる。初期反射音が後方から到来する音場の不明瞭さの要因としては、「こもっている」、「中央にまとまりすぎ」、「音がぶつかっている」、「濁った」、「聴き取りづらい」等が見られる。中央に音がまとまっていること、そして聴き取りづらいことが不明瞭さの要因となっている。

表 5.2.6 初期反射音の到来方向が前方、および側方の刺激の比較において得られた評価語とその出現頻度。市販音源の結果。2 回以上言及された言葉を太字で示す。

前方の刺激が明瞭な要因（市販音源）	側方の刺激が不明瞭な理由（市販音源）
輪郭が はっきりしている (3) 、 聴き取りやすい (2) まとまりがある、高域がとおった、響きの量が適切、 やわらかい、定位が分かりやすい	こもった、バランスが悪い、高域がきつい、ぼやけた、音が分離している、響きが不自然

表 5.2.7 初期反射音の到来方向が前方、および側方の刺激の比較において得られた評価語とその出現頻度。無響音源の結果。2 回以上言及された言葉を太字で示す。

前方の刺激が明瞭な要因（無響音源）	側方の刺激が不明瞭な理由（無響音源）
やわらかい (3) 、 音が中央に集まっている (3) 芯がある、鋭い、輪郭が はっきりした 、 定位が分かりやすい 、音が近い、音像が狭い、響きが自然、 聴き取りやすい	こもっている (2) かたい、輪郭がぼやけている、音が遠い、響きが不自然

表 5.2.8 初期反射音の到来方向が側方、および後方の刺激の比較において得られた評価語とその出現頻度。市販音源の結果。2 回以上言及された言葉を太字で示す。

側方の刺激が明瞭な要因（市販音源）	後方の刺激が不明瞭な理由（市販音源）
広がりがある (3)、奥行きがある (2)、定位が分かりやすい (2) 分離している、リアリティがある、高域が通った、すっきりした	こもっている (3)、中央にまとまりすぎ (2)、音がぶつかっている (2) 音像が狭い、低音が大きい、音がかたい、ひっこんだ、音色が不明瞭

表 5.2.9 初期反射音の到来方向が側方、および後方の刺激の比較において得られた評価語とその出現頻度。無響音源の結果。2 回以上言及された言葉を太字で示す。

側方の刺激が明瞭な要因（無響音源）	後方の刺激が不明瞭な理由（無響音源）
音像が広い (2)、輪郭がはっきりした (2) 音が近い	濁った (2) つまった、ハイカットされた、凝縮された、高域が減衰した、響きが多い

表 5.2.10 初期反射音の到来方向が前方、および後方の刺激の比較において得られた評価語とその出現頻度。市販音源の結果。2 回以上言及された言葉を太字で示す。

前方の刺激が明瞭な要因（市販音源）	後方の刺激が不明瞭な理由（市販音源）
音像が広い (3) まとまっている、すっきりしてる、かぶりが少ない、輪郭がはっきりしている、分離している、響きが少ない、音色が明瞭	聴き取りづらい (3)、中央にかたまっている (2) 音像が狭い、音色が不明瞭

表 5.2.11 初期反射音の到来方向が前方、および後方の刺激の比較において得られた評価語とその出現頻度。無響音源の結果。2 回以上言及された言葉を太字で示す。

前方の刺激が明瞭な要因（無響音源）	後方の刺激が不明瞭な理由（無響音源）
響きが少ない (3)、音色が自然 (2)、輪郭がはっきりしている (2) 高域が通った、明るい	つまった、響きが長い、こもった

5.2.6 再生音圧の補正有無に基づく評価要因の分類

サーストンの一対比較法による順位づけの結果、再生音圧による順位の差はほとんど見られなかった。一方で、市販音源の聴取でのみ、S_Side- の音場の順位が、再生音圧のみが異なる S_Side と比べて順位が高い。表 5.2.12 は、市販音源の聴取において、S_Side-、および S_Side の刺激の比較において得られた評価語とその出現頻度を示す。不明瞭さの要因として、「声がうるさい」の他に、「響きが多い」等が見られる。

表 5.2.12 S_Side-、および S_Side の刺激の比較において得られた評価語とその出現頻度。市販音源の結果。2 回以上言及された言葉を太字で示す。

S_Side- が明瞭な要因	S_Side が不明瞭な理由
映像が狭い、すっきりしている、響きが少ない、輪郭がはっきりしている	響きが多い (2) 、 声がうるさい (2) 響きが不自然、こもっている、聴き取りにくい

5.2.7 被験者グループによる評価要因の分類

実験 1 の結果、個人によって音色の明瞭さおよび映像の明瞭さの中でも異なる感覚に着目して判断が行われる可能性が示唆されたため、被験者ごとにサーストンの一対比較法に基づく分析を行い、評価傾向の違いを調査した。表 5.2.13 および表 5.2.14 は、市販音源および無響音源それぞれについて、被験者ごとにサーストンの一対比較法に基づく分析を行い、評価音場を上から順位順に並べたものである。アルファベットは被験者の ID を示す。無響音源の聴取には、市販音源で評価を行った被験者のうち H 以外の 7 名が参加した。表において、横線で区切られた音場は、順位のスコアが異なる音場である。同じ枠内の音場は、順位が同じ音場である。被験者 ID の横には、その被験者がどのような音場を上位に判断したかについて記載した。例えば、容積が小さい音場の順位が高い場合は「容積小」、初期反射音が側方の音場の順位が高い場合は「側方」と記載した。

市販音源の聴取では、A、C、E、F（以下グループ 1）の被験者が室容積の違いによって順位づけをしているのに対し、B、D、G、H（以下グループ 2）の被験者は到来方向によって順位付けをしていることが分かる。これは、明瞭さの判断基準として着目する点が個人によって異なることを示唆している。

更に、市販音源の聴取において、A、C、E、F の 4 名から得られた明瞭さの評価要因と、B、D、G、H の 4 名から得られた明瞭さの評価要因の分析を行った。表 5.2.15 にグループ 1 から得られた明瞭さの評価要因を、表 5.2.16 にグループ 2 から得られた明瞭さの評価要因を示す。表の上から、音色に関する評価語、映像に関する評価語、響きに関する評価語、その他の評価語を示す。太字は 10 回以上言及された評価語を示す。

音色に関する評価要因では、グループ 1 およびグループ 2 の両方で、こもっているかどうかに着

表 5.2.13 市販音源の聴取における、被験者ごとの評価音場の順位付けの結果。同じ枠内の音場は、順位のスコアが等しい音場である。被験者 ID の横には、その被験者がどのような音場を上位に判断したかについて記載した。

A：容積大	B：側方	C：容積大	D：前側	E：容積大	F：容積大	G：前後	H：前側
L_Front	S_Side-	L_Front	L_Front	L_Side	L_Side	L_Front	L_Front
S_Side-	L_Side	L_Side	S_Side-	L_Front	L_Back	L_Back	L_Side
L_Side	S_Side	S_Side-	L_Side	L_Back	L_Front	S_Back-	S_Side-
L_Back	L_Front	L_Back	S_Front	S_Side-	S_Front-	S_Front-	S_Front
S_Side	S_Front-	S_Side	L_Back	S_Front	S_Side-	S_Front	S_Side
S_Front-	L_Back	S_Back	S_Front-	S_Front-	S_Back-	L_Side	S_Back
S_Front	S_Back	S_Front-	S_Back	S_Side	S_Side	S_Back	S_Front-
S_Back	S_Back-	S_Back-	S_Back-	S_Side-	S_Front	S_Side-	L_Back
S_Back-	S_Front	S_Front	S_Side	S_BacL_	S_Back	S_Side	S_Back-

表 5.2.14 無響音源の聴取における、被験者ごとの評価音場の順位付けの結果。同じ枠内の音場は、順位のスコアが等しい音場である。被験者 ID の横には、その被験者がどのような音場を上位に判断したかについて記載した。

A：容積小	B：容積小	C：容積大	D：容積小	E：容積小	F：容積小	G：容積小
S_Front	S_Side-	L_Back	S_Side	S_Front	S_Front-	S_Side-
S_Back	S_Front	L_Front	S_Side-	S_Back-	S_Side-	S_Side
S_Side-	S_Front-	L_Side	S_Front	S_Front-	S_Back-	S_Back
S_Front-	S_Side	S_Front-	S_Front-	S_Side	S_Side	S_Back-
L_Front	S_Back	S_Back	S_Back	S_Side-	S_Front	S_Front
L_Side	S_Back-	S_Side-	L_Front	S_Back	L_Front	S_Front-
S_Back-	L_Back	S_Back-	S_Back-	L_Side	L_Side	L_Side
S_Side	L_Side	S_Front	L_Front	L_Back	S_Back	L_Front
L_BacL_	L_Front	S_Side	L_Back	L_Front	L_Back	L_Back

目されている。一方、グループ 1 では「音色が明瞭」という回答が多いが、グループ 2 では「高域がしっかりした」、「こもっている」の回答数が多い。グループ 2 は特に音色の周波数特性に着目して評価していると言える。音像に関する評価要因では、グループ 1 は「音像が広い」や「中心に音がまとまっている」等、主に音像の広さに着目している。一方、グループ 2 は「輪郭がはっきりしている」や「定位が分かりやすい」等、輪郭および定位についての言及も多い。グループ 2 は輪郭や定位にも着目するグループだと言える。響きに関する評価要因では、グループ 1 は「響きが多い」等の響きの量に関する回答が多い。一方でグループ 2 は、そもそも響きに関する言及が少ない。グループ 1 は、響きの量に着目するグループだと言える。その他の評価要因については、どちらのグループも聴き取りやすさに言及している。一方で、グループ 2 には「はっきり聴こえる」や「音がきつい」といった回答も見られる。グループ 2 は音像の輪郭や定位に着目していたので、この「はっきり聴こえる」が、音像の輪郭や定位を意味している可能性も考えられる。

表 5.2.15 市販音源の聴取において、被験者 A、C、E、F（グループ 1）から得られた明瞭さの評価要因となる評価語。括弧内は回答数を示す。上から音色に関するもの、音像に関するもの、響きに関するもの、その他の評価語である。

明瞭さの要因（グループ 1）	不明瞭さの要因（グループ 1）
音色が明瞭 (22) やわらかい (7)、バランスが良い (4)、音色が自然 (3)、音色が明るい	こもっている (11) 低音が出過ぎている (9)、音がかたい (6)、音色が不明瞭 (5)、コムフィルタがかかったような (4)、低域が弱い、バランスが悪い、アタックが強すぎる
音像が広い (11)、中心に音がまとまっている (10) 音像が前方に定位した (5)、奥行きがある (4)、分離している (3)、音像が狭い (3)、定位が分かりやすい (2)、音が近い (2)、包み込まれている (2)、横からの音がよく聴こえる、輪郭がはっきりしている	音が遠い (6)、中央に音がまとまりすぎている (4)、音像が広い (3)、音が近い (2)、音像が狭い
響きが自然 (5)、響きが少ない (3)、低域の響きが豊か (2)	響きが多い (25) 響きが分離している (7)、響きが混ざっている (6)、部屋なりがある (3)、響きが不自然 (2)、響きが少ない
聴き取りやすい (13) すっきりしている (4)、伸びがある (2)	うるさい (9)、聴き取りづらい (6)、音量が大きい (2)、平べったい (2)、つまった、濁った

表 5.2.16 市販音源の聴取において、被験者 B、D、G、H（グループ 2）から得られた明瞭さの評価要因となる評価語。括弧内は回答数を示す。上から音色に関するもの、音像に関するもの、響きに関するもの、その他の評価語である。

明瞭さの要因（グループ 2）	不明瞭さの要因（グループ 2）
高域がしっかりした (12) バランスが良い (7)、アタックが強い、芯がある、音色が自然、音色が明瞭	こもっている (29) 高域がきつい (2)、コンプがかかったような (2)、芯がない、劣化したような、バランスが悪い、低域が出過ぎている、低域が弱い
輪郭がはっきりしている (33)、音像が広い (27)、定位が分かりやすい (12) 分離している (7)、音像が狭い (3)、奥行きがある	中央に音がまとまりすぎている (22) 輪郭がぼやけた (8)、音が近い (6)、音像が狭い (4)、音像が広い、音が遠い
かぶりが少ない (7)、響きが少ない (4)、響きが自然、低域の響きが豊か	響きが不自然 (3)、響が多い (2)、かぶっている (2)、響きが混ざっている
聴き取りやすい (17)、はっきり聴こえる (11) すっきりしている (8)、音が生っぽい (3)、リアリティがある (2)、音量が大きい (2)	音がきつい (12) 聴き取りづらい (6)、痩せている、うるさい

5.3 考察

5.3.1 音楽の明瞭さの評価要因

音色の明瞭さと音像の明瞭さ

インタビュー調査から得られた評価語とその出現頻度の分析より、明瞭さの評価要因として、「音色が明瞭」、「やわらかい」、「音色が明るい」、「つまった」等の音質に関するものと、「こもった」、「高域がしっかりした」、「バランスが良い」等の周波数特性に関するもの、「輪郭がはっきりしている」、「定位が分かりやすい」等の音像の空間的な明瞭さに関する回答が得られた。この結果は、音色に関する明瞭さを含む“Timbral quality”と音像に関する明瞭さを含む“Localization quality”を区別している ITU-R BS.1116-3 (ITU-R, 2015) や、“Coloration”の明瞭さと空間的な“Readability”を区別している Guastavino and Katz (2004) の研究と対応する。これらの研究は音楽再生システムを対象としており、明瞭さの感覚に着目して実施されたものではないが、小空間においても、音楽の明瞭さの評価要因には少なくとも音質に関する明瞭さ、周波数特性に関する明瞭さ、そして音像の空間的な明瞭さが存在することが本実験によって明らかになったと言える。

音の距離感

「音が近い」や「音が遠い」等、音の距離感に関する回答も多く見られた。ほとんどの場合、「音が近い」ことが明瞭さの要因となっていた。なお、音の距離感に関する回答は無響音源の聴取のみで見られた。市販音源の聴取では音の距離感については 1 回も言及されておらず、距離感が明瞭さの要因となるのは無響音源の聴取時のみであると言える。

無響音源の聴取においては、ほとんどの場合に室容積が小さい音場、つまり初期反射音レベルの大きい音場が明瞭だとされていた。Griesinger (2000) および Pellegrini (2002) は、直接音の到来語 20 ms～50 ms で到来する初期反射音は知覚される距離感に影響すると述べている。無響音源の聴取では、初期反射音レベルが大きいことで距離感が近くなり、明瞭だと判断されたと考えることができる。

音像の広さ

「音像が広い」という回答が明瞭さの要因として表れていた。聴取音源ごとに明瞭さの評価要因を分析した結果より、音像の広さに関する要因は特に市販音源の聴取で多く見られた。表 5.2.2 より、市販音源の聴取では合計 90 回言及されている。一方、無響音源の聴取では 14 回しか言及されていない。音像の広さが音楽の明瞭さに影響することを述べた先行研究は筆者の知る限り存在しないが、Kim (2009) は音像の広さと音楽聴取における総合的な嗜好とに相関があることを報告している。明瞭さにおいても、音像が広いことで明瞭だと判断される可能性が得られた。

一方、不明瞭さの要因として「中心に音がまとまりすぎている」という回答もされており、純粹

な音像の広さを評価しているのではなく、「狭いことで楽器が中央で混ざって聴こえる」、「広いことで音像が分離して聴こえる」といった判断である可能性も考えられる。

響きの量

「響きが多い」、「響きが少ない」といった、響きの量の適切さに関する回答が明瞭さの要因として表れていた。コンサートホールにおける C_{80} は、Beranek (2004) の水平方向の明瞭さに対応すると考えられている。水平方向の明瞭さとは、ある瞬間に鳴っている音が前の音の残響等に影響され、濁ったり聴こえづらくなるかどうかである。表 5.2.4 および表 5.2.5 より、市販音源で明瞭とされたのは室容積が大きい音場であり、無響音源で明瞭とされたのは室容積が小さい音場であることが分かる。表 5.1.2 より、室容積 S の残響時間は 0.26 秒、室容積 L の残響時間は 0.42 秒であるため、市販音源では残響時間の長い音場が「響きが少ない」と判断されている。一方で、無響音源では、残響時間の短い音場がそのまま「響きが少ない」と判断されている。室容積 L は室容積 S と比べて初期反射音レベルが低く、減衰が遅い。市販音源では初期反射音の音圧レベルに着目して響きの量を判断している一方で、無響音源では初期反射音の減衰、あるいは後期残響音によって響きの量を判断している可能性が示唆された。

聴き取りやすさ

「聴き取りやすい」、「はっきり聴こえる」等の、聴き取りやすさに関する回答が明瞭さの要因として表れていた。表 5.2.2 および表 5.2.3 より、市販音源の聴取では「聴き取りやすい」および「聴き取りづらい」が合計 42 回言及されている。一方、無響音源では 10 回である。聴き取りやすさや聴き取りにくさといった感覚は、音声の明瞭さの評価項目とされてきた (Morimoto et al., 2004) ため、市販音源の聴取では女性ボーカルの歌詞に着目して、聴き取りやすさの判断がされたものと考えられる。

結論

上記の評価要因同士の関係や構造については、更なる主観評価実験による検証が必要であるが、得られた評価要因をまとめると表 5.3.1 のようになる。

5.3.2 聴取音源の違いが評価要因に与える影響

市販音源の聴取では室容積の大きい音場が、無響音源の聴取では室容積の小さい音場が明瞭だと回答される傾向が見られた。表 5.1.2 より、本実験における室容積 S の初期反射音レベルは直接音に対して -19.7 dB、室容積 L は -21.7 dB であった。

実験 1 では、無響音源の明瞭さの最適初期反射音レベルの平均値は直接音に対して -10.2 dB (音色の明瞭さ)、および -19.3 dB (音像の明瞭さ) であった。本実験では、無響音源は室容積の小さい音場が「音が近い」、「響きが少ない」という要因で明瞭だと判断されており、主に音色の明瞭さ

表 5.3.1 小空間での音楽聴取における音楽の明瞭さの評価要因

評価要因	明瞭さの要因	-	不明瞭さの要因
音色の明瞭さ	音色が明瞭	-	音色が不明瞭
こもり具合	高域がしっかりした、音色が明るい	-	こもっている
音の距離感	音が近い	-	音が遠い
輪郭	輪郭がはっきりしている	-	輪郭がぼやけている
定位の分かりやすさ	定位が分かりやすい	-	定位が分かりにくい
音像の広さ	音像が広い	-	音像が狭い
響きの量の適切さ	響きが少ない	-	響きが多い
聴き取りやすさ	聴き取りやすい、はっきり聴こえる	-	聴き取りづらい

に着目して判断したものと考えられる。音色の明瞭さの最適レベルである -10.2 dB に近い初期反射音レベルを持つ室容積 S の音場を明瞭だと評価したと考えられる。一方、市販音源の聴取における明瞭さの最適初期反射音レベルの平均値は -22.0 dB（音色の明瞭さ）および -24.6 dB（音像の明瞭さ）であった。本実験では、この最適レベルに近い室容積 L を明瞭だと判断したと考えられる。市販音源では室容積の大きい音場が「すっきりしている」、「音像が広い」、「こもっていない」と評価された。

聴取音源ごとに被験者から得られた評価要因を分析した結果を以下にまとめる。

市販音源の聴取における音楽の明瞭さの評価要因。音色、音像、音像の広さ等の様々な感覚に着目されている。

- こもり具合
- 音像の広さ、音のまとまり具合
- 定位の分かりやすさ
- 輪郭
- 響きの量の適切さ
- 聴き取りやすさ

無響音源の聴取における音楽の明瞭さの評価要因。主に音質および距離感に着目されている。

- こもり具合
- 音質の明瞭さ（アタックの鋭さ、明るさ、やわらかさ）
- 音の距離感
- 響きの量の適切さ
- 聴き取りやすさ

実験 1 で述べたように、本実験で用いた市販音源は楽曲信号のエンベロープ変化が穏やかであり、残響成分によって楽音が阻害されやすい。初期反射音レベルが低い場合には、楽音が阻害されなく

なるために音色、音像等の多くの評価要因が向上すると考えられる。一方で、初期反射音レベルが高くなると直接音が補強される効果よりも楽音を阻害する影響が大きくなり、音色の明瞭さが不明瞭だと評価されたと考えることができる。無響音源は楽曲信号のエンベロープ変化が急であり、楽音は残響成分によって比較的阻害されにくい。初期反射音レベルが高い場合には、楽音が阻害される影響よりも直接音が補強される効果が大きく、音質が向上したように感じられ、さらに距離感が近く聴き取りやすくなったと考えられる。

5.3.3 初期反射音の到来方向の違いが評価要因に与える影響

初期反射音が前方から到来する音場における明瞭さ

初期反射音が前方から到来する音場が他の方向から到来する音場と比べて明瞭だと評価された要因として、「輪郭がはっきりしている」、「やわらかい」、「聴き取りやすい」、「音が中央に集まっている」、「音像が広い」、「響きが少ない」、「音色が自然」等、多くの評価語が得られた。Begault and Anderson (2001) は、音源と同じ方向から到来する単発の反射音が知覚されにくいことを明らかにした。前方は直接音の到来方向であったため、前方から到来する初期反射音による影響は少なく知覚されたものだと考えられる。

初期反射音が側方から到来する音場における明瞭さ

初期反射音が側方から到来する音場が他の方向から到来する音場と比べて明瞭だと評価された要因として、「広がりがある」、「音像が広い」、「奥行きがある」、「定位が分かりやすい」、「輪郭がはっきりした」等が見られた。これらは全て音像に関する要因である。また、前述のように音像が広いことも明瞭さの要因の 1 つとなっており、初期反射音が側方から到来する音場は音像の広さに寄与したものと思われる。一方で、前方から到来する音場と比較して不明瞭だとされた要因としては、「こもっている」が見られた。これは前述したように、直接音から異なる方向から初期反射音が到来したことで、直接音と同じ方向から到来する場合よりも直接音が阻害されたように感じられたからだと考えられる。

初期反射音が後方から到来する音場における明瞭さ

初期反射音が後方から到来する音場が他の方向から到来する音場と比べて不明瞭だと評価された要因としては、「こもっている」、「中央にまとまりすぎ」、「音がぶつかっている」、「濁った」、「聴き取りづらい」等が見られた。音像が狭く、中央に音がまとまることで音がぶつかり、濁ったり聴き取りづらく判断されたと言える。これは、初期反射音が直接音とは異なる方向から到来したことによって、マスキングや音色の変化が大きくなり、側方から到来する音場と比較して更に音像も狭くなったためだと考えられる。

5.3.4 評価要因の個人差

被験者ごとの音場の順位付けの結果より、市販音源の聴取では、グループ 1 の 4 名は室容積の違いに着目して順位づけをしているのに対し、グループ 2 の 4 名は初期反射音の到来方向に着目して順位付けを行っていることが分かった。これは、明瞭さの評価の際に着目する点が個人によって異なることを示唆している。被験者のグループごとに市販音源の聴取から得られた明瞭さの評価要因を分析した結果、次のような特徴が見られた。グループ 1 は音色の明瞭さや音像の広さ、そして響きの量に着目しているが、グループ 2 は輪郭や定位の分かりやすさに着目しているのが特徴的である。

市販音源の聴取において、グループ 1 が着目した明瞭さの要因

- こもり具合
- 音色の明瞭さ
- 音像の広さ
- 響きの量の適切さ

市販音源の聴取において、グループ 2 が着目した明瞭さの要因

- こもり具合
- 輪郭
- 定位の分かりやすさ

5.3.5 再生音圧の補正が評価要因に与える影響

再生音圧による影響はほとんど見られなかったが、市販音源の聴取でのみ、S_Side- の音場の順位が S_Side と比べて高かった。S_Side- および S_Side の刺激の比較において得られた評価語の分析より、不明瞭さの要因として「声がうるさい」の他に、「響きが多い」等が見られた。再生音圧の違いを、響きの量の違いとしても判断されていたと言える。一方で、何故初期反射音が側方から到来する音場の比較においてのみこのような違いが表れたかについては、更なる調査が必要である。

5.4 結論

本実験では、小空間における音楽聴取において音楽が明瞭だと評価される要因を明らかにするため、レポートリーグリッド法を参考にしたインタビュー調査を行った。調査から得られた明瞭さに関する自由記述を分析した結果、以下の結論が得られた。

1. 音楽の明瞭さは、少なくとも「音色の明瞭さ」、「こもり具合」、「音の距離感」、「輪郭」、「定位の分かりやすさ」、「音像の広さ」、「響きの量の適切さ」、「聴き取りやすさ」等の複数の要因によって評価される。
2. 聴取する音源によって明瞭さの評価要因は異なる。市販音源の聴取に特徴的なものは、「音像の広さ」、「輪郭」、「定位の分かりやすさ」であり、無響音源の聴取に特徴的なものは「こもり具合」、「音の距離感」が挙げられる。市販音源の聴取においては室容積が大きい音場が明瞭だとされ、無響音源の聴取においては室容積が小さい音場が明瞭だとされた。
3. 初期反射音が側方から到来する場合は「音像の広さ」が主な明瞭さの要因となる。後方から到来する場合は「音像の狭さ」および「聴き取りづらさ」が不明瞭さの要因となる。初期反射音の到来方向が前方、側方、後方の順で明瞭さの順位が低下する。
4. 市販音源の聴取においては、個人によって着目する音楽の明瞭さの要因が異なる。グループ 1 は「音色の明瞭さ」や「音像の広さ」、「響きの量」に着目した一方、グループ 2 は「輪郭」や「定位の分かりやすさ」に着目した。

第6章

実験3：小空間における明瞭さに対応する音場の物理量

実験1および実験2より、無響音源の聴取では「音色の明瞭さ」に主に着目され、先行研究の結果と同様、比較的高い初期反射音レベルが求められることが明らかになった。一方、市販音源の聴取では、「輪郭」や「定位の分かりやすさ」、「音像の広さ」といった音像に関する要因によっても明瞭さが評価されることが明らかになった。更に、個人によって明瞭さの評価基準が異なり、ある被験者グループは「音色の明瞭さ」や「音像の広さ」、そして「響きの量」に着目し明瞭さを評価した一方で、他のグループは「輪郭」や「定位の分かりやすさ」に着目し、初期反射音の到来方向の違いに基づき明瞭さを評価した。市販音源の聴取では、明瞭さの指標としてこれらの感覚を別のものとして区別する必要があると言える。

小空間での市販音源の聴取において、上記それぞれの評価要因と対応する音場の物理量を明らかにするため、実験3を実施した。本実験では、実験2から得られた市販音源の聴取における明瞭さの評価要因について、シェッフエの一対比較法を用いて音場の比較評価を求めた。得られた結果から、市販音源の聴取における明瞭さの評価要因と音場の物理量との対応関係について調査した。

6.1 実験方法

6.1.1 シェッフエの対比較法

シェッフエの対比較法は、実験 2 で用いたサーストンの対比較法にカテゴリー判断を取り入れたものである。複数の刺激を 2 つずつ対にして被験者に提示し、その刺激を比べてどちらがどれだけ好きか等の判断を求める。2 つの刺激のみを比較するため、判断がやさしいという利点があり、結果の信頼性が高い (難波・桑野, 1998)。また、どれだけ好きかについて 7 段階評価等のカテゴリー判断を求めるため、より細かな被験者の判断の傾向を調査することができる。

シェッフエの対比較法では、刺激の提示法による順序効果の有無、そして各被験者が全ての比較評価を行うか否かによって、4 つの変法が存在している。順序効果とは、刺激の提示順序が評価に影響してしまうことを言う。本実験では、刺激の比較中に被験者が刺激を自由に切り替えることが可能であったため、順序効果は発生しない。更に、各被験者が全ての刺激について評価したため、シェッフエの対比較の中でも順序効果を考慮せず、個人差を考慮する中屋の変法を用いた。

6.1.2 実験刺激の作成

羽入らによる仮想音源分布を用いた音場の可聴化の原理に基づき、室容積が 60 m^3 (総表面積 94 m^2)、 126 m^3 (総表面積 162 m^2) および 224 m^3 (総表面積 232 m^2) の 3 種類、 $30 \text{ ms} \sim 80 \text{ ms}$ までの初期反射音の到来方向が全方向 (ALL)、前後方向 (FB: FrontBack)、左右方向 (SI: Side) の 3 種類、室の平均吸音率 α が 0.4 の合計 9 種類の評価音場を再現するインパルス応答を生成した。音楽聴取のための小空間として現実的な音場を想定するため、ITU-R BS.1116-3 (ITU-R, 2015) において推奨される室形状を元に、室容積および総表面積を設定した。推奨される室形状は、 l を長さ、 w を幅、 h を高さとした時、式 $1.1w/h \leq l/h \leq 4.5w/h - 4$ 、かつ床面積が $20 \sim 60 \text{ m}^2$ を満たすものである。それぞれ、室容積が 60 m^3 ($l:5 \text{ m}$ 、 $w:4 \text{ m}$ 、 $h:3 \text{ m}$)、 126 m^3 ($l:7 \text{ m}$ 、 $w:6 \text{ m}$ 、 $h:3 \text{ m}$)、 224 m^3 ($l:8 \text{ m}$ 、 $w:7 \text{ m}$ 、 $h:4 \text{ m}$) として総表面積を算出した。これより、室の複雑さ係数 h は室によって異なる。

初期反射音の到来方向が全方向 (ALL) の音場は初期反射音の生成範囲を 16 ブロック全て (第 3.2.1 項を参照)、前後方向 (FB) の音場は 1 番および 16 番、8~9 番、左右方向 (SI) の音場は 4~5 番、12~13 番のブロックに限定した。本実験では、実験 2 と同様、直接音の到来後 30 ms 以内の反射音は除外し、 80 ms 以降の後期残響音をインパルス応答に含めた。更に、より実際のスタジオ施設の聴取環境に近づけるため、音源は受音点から 3 m 離れた位置にあると仮定し、直接音に対して距離減衰を反映させた。距離減衰を考慮しない場合と比べて、直接音の再生音圧レベルは -10.2 dB となった。

表 6.1.1 に、本実験で生成した 9 種類の音場の名称および生成パラメータを示す。音場の名称は

「室容積_到来方向」となっており、室容積 V が 60 m^3 のものは S (Small)、室容積 V が 126 m^3 のものは M (Middle)、室容積 V が 224 m^3 のものは L (Large) で示す。表 6.1.2 に、生成された評価音場の物理量を示す。物理量は生成したインパルス応答から直接算出した。図 6.1.1～図 6.1.27 に、それぞれの音場の 3 次元仮想音源分布、2 次元仮想音源分布（仮想音場の上方から俯瞰した図）、音場の反射音の時系列応答を示す。仮想音源分布の図における点は各仮想音源を示し、中央の立方体および四角は仮想の聴取空間を示す。見やすさのため、3 次元および 2 次元の仮想音源分布の図には、仮想音源の総数 N_{total} 個から半数をランダムに選択し、表示した。

表 6.1.1 本実験で生成した 9 種類の評価音場の名称および生成パラメータ

刺激名	室容積 [m^3]	室総表面積 [m^2]	複雑さ係数 h	平均吸音率 $\bar{\alpha}$	初期反射音の到来方向
S_ALL	60	94	3.60	0.4	全方向
S_FB	60	94	3.60	0.4	前後方向
S_SI	60	94	3.60	0.4	左右方向
M_ALL	126	162	4.18	0.4	全方向
M_FB	126	162	4.18	0.4	前後方向
M_SI	126	162	4.18	0.4	左右方向
L_ALL	224	232	3.89	0.4	全方向
L_FB	224	232	3.89	0.4	前後方向
L_SI	224	232	3.89	0.4	左右方向

表 6.1.2 本実験で生成した 9 種類の評価音場の名称および物理量。物理量はシミュレーションによって生成したインパルス応答から算出して求めた。

刺激名	残響時間 [秒]	直接音に対する初期反射音レベル [dB]	直接音に対する後期残響音レベル [dB]	C_{80} [dB]
S_ALL	0.20	-9.8	-24.8	35.1
S_FB	0.20	-9.8	-24.8	35.1
S_SI	0.20	-9.8	-24.8	35.1
M_ALL	0.24	-10.0	-22.4	32.6
M_FB	0.24	-10.0	-22.4	32.6
M_SI	0.24	-10.0	-22.4	32.6
L_ALL	0.30	-10.5	-20.1	30.3
L_FB	0.30	-10.5	-20.1	30.3
L_SI	0.30	-10.5	-20.1	30.3

作成したインパルス応答を以下の音源に畳み込み、実験刺激とした。

- 市販音源（ステレオ）：Norah Jones, “Come Away With Me” CD (2002) Track1 : Don't Know Why

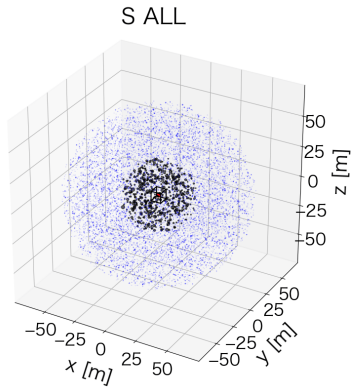


図 6.1.1 S_ALL の音場の 3次元仮想音源分布

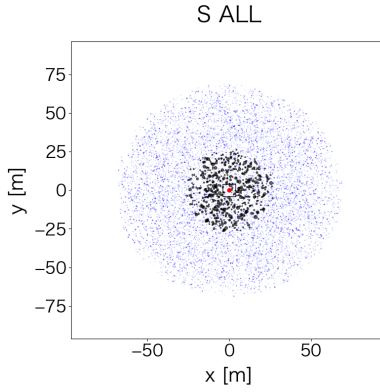


図 6.1.2 S_ALL の音場の 2次元仮想音源分布

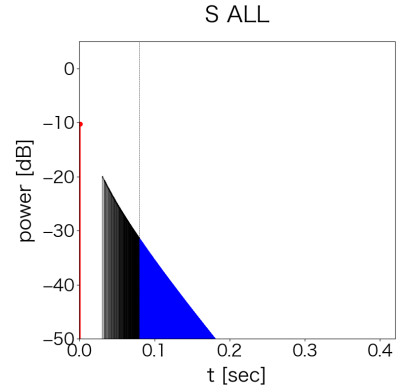


図 6.1.3 S_ALL の音場の反射音の時系列応答

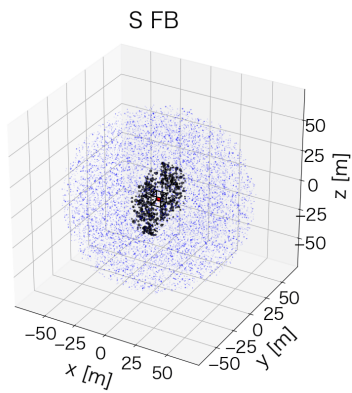


図 6.1.4 S_FB の音場の 3次元仮想音源分布

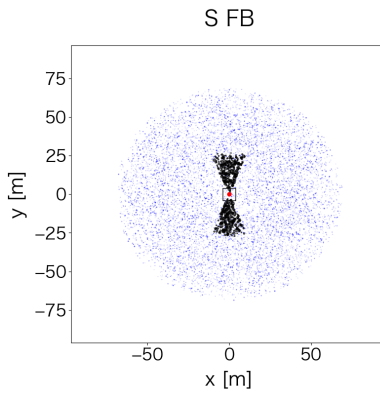


図 6.1.5 S_FB の音場の 2次元仮想音源分布

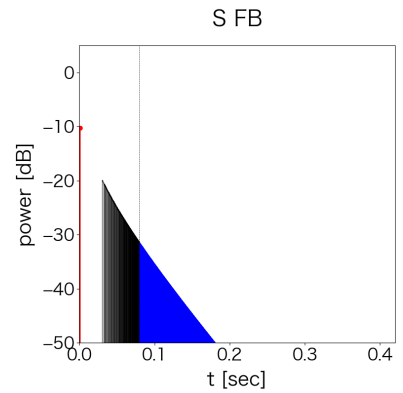


図 6.1.6 S_FB の音場の反射音の時系列応答

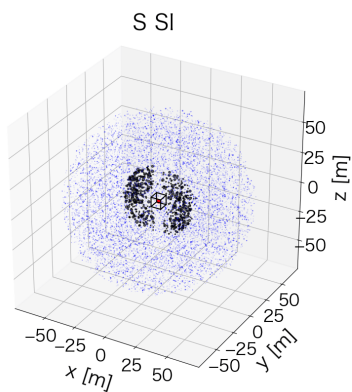


図 6.1.7 S_SI の音場の 3次元仮想音源分布

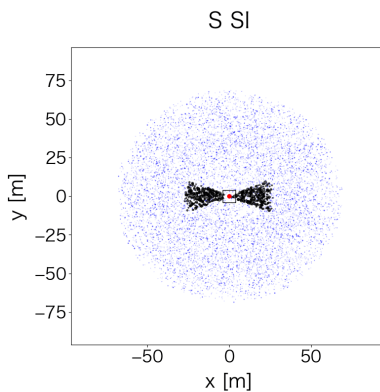


図 6.1.8 S_SI の音場の 2次元仮想音源分布

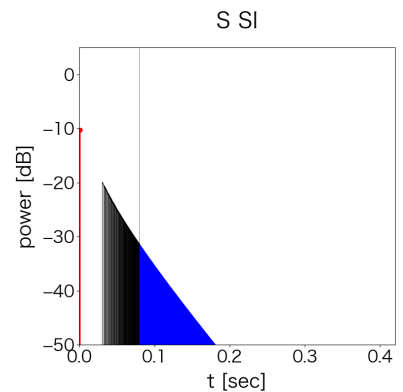


図 6.1.9 S_SI の音場の反射音の時系列応答

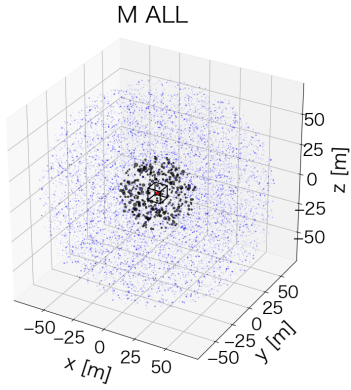


図 6.1.10 M_ALL の音場の 3次元仮想音源分布

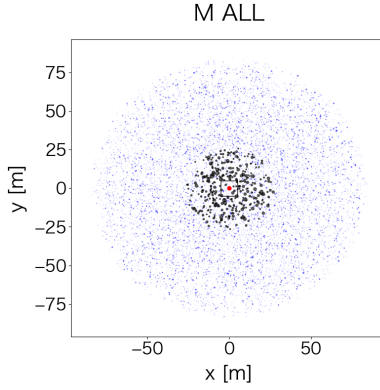


図 6.1.11 M_ALL の音場の 2次元仮想音源分布

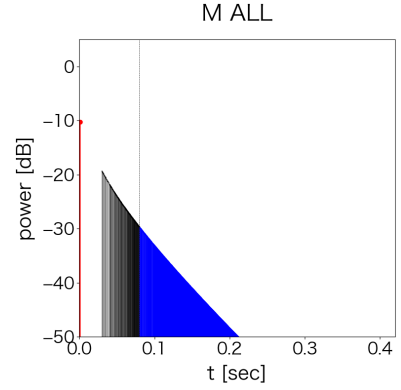


図 6.1.12 M_ALL の音場の反射音の時系列応答

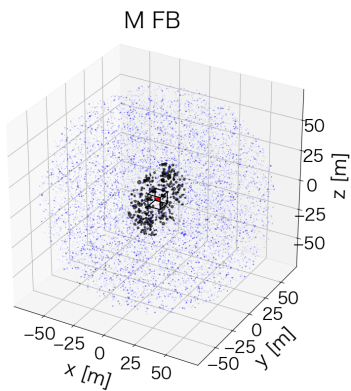


図 6.1.13 M_FB の音場の 3次元仮想音源分布

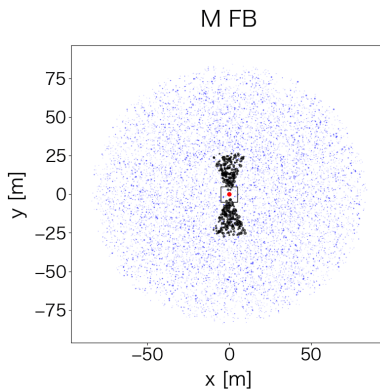


図 6.1.14 M_FB の音場の 2次元仮想音源分布

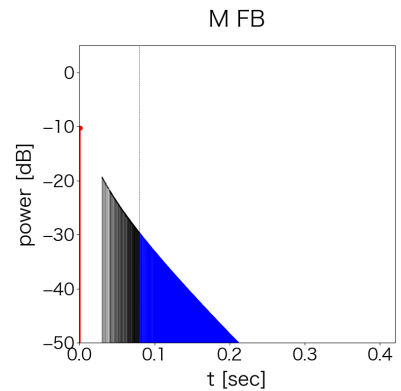


図 6.1.15 M_FB の音場の反射音の時系列応答

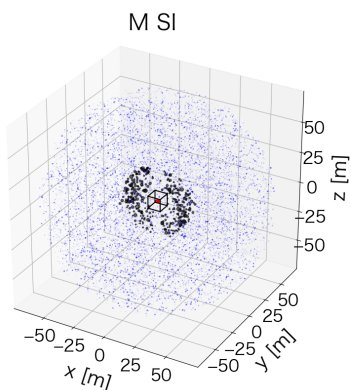


図 6.1.16 M_SI の音場の 3次元仮想音源分布

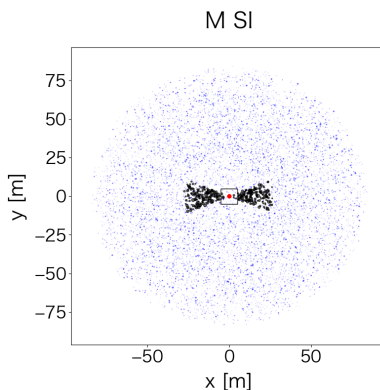


図 6.1.17 M_SI の音場の 2次元仮想音源分布

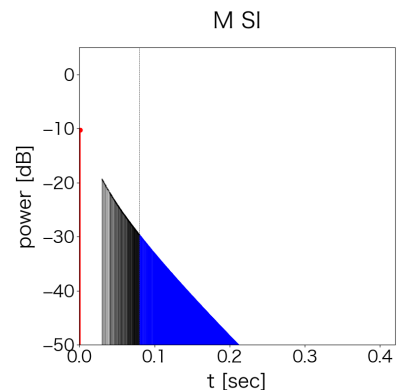


図 6.1.18 M_SI の音場の反射音の時系列応答

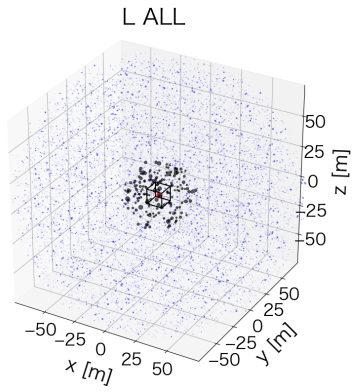


図 6.1.19 L_ALL の音場の 3次元仮想音源分布

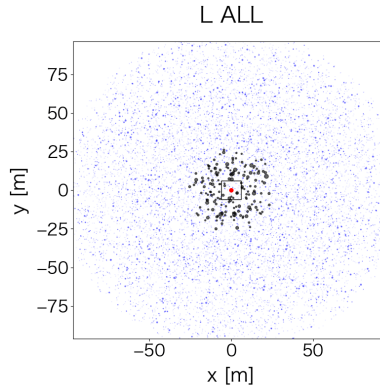


図 6.1.20 L_ALL の音場の 2次元仮想音源分布

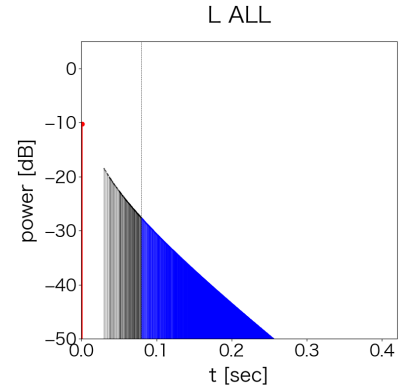


図 6.1.21 L_ALL の音場の反射音の時系列応答

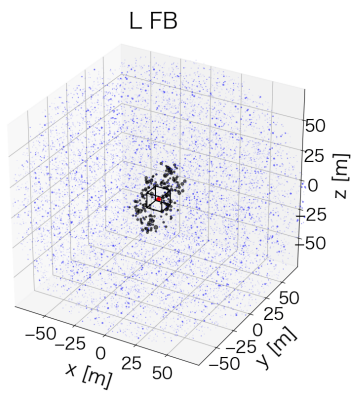


図 6.1.22 L_FB の音場の 3次元仮想音源分布

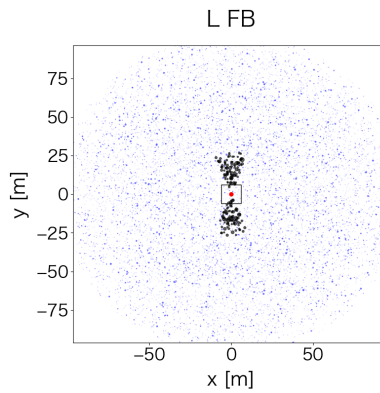


図 6.1.23 L_FB の音場の 2次元仮想音源分布

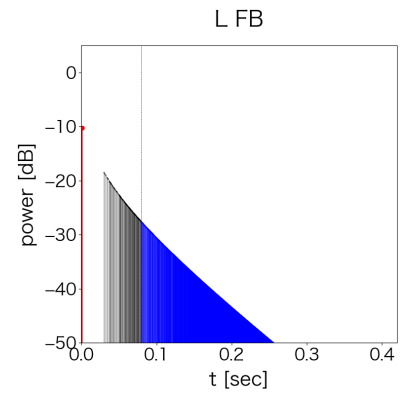


図 6.1.24 L_FB の音場の反射音の時系列応答

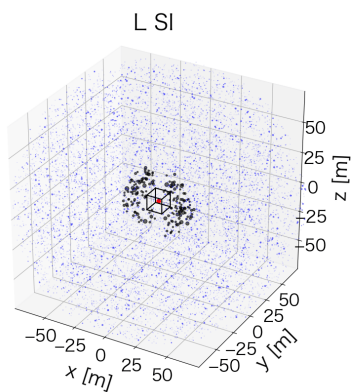


図 6.1.25 L_SI の音場の 3次元仮想音源分布

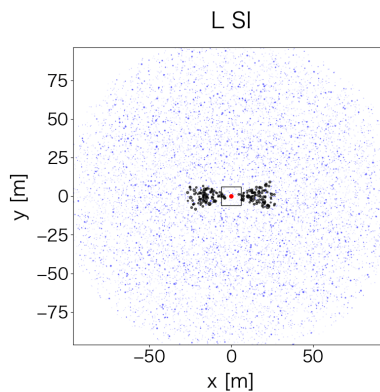


図 6.1.26 L_SI の音場の 2次元仮想音源分布

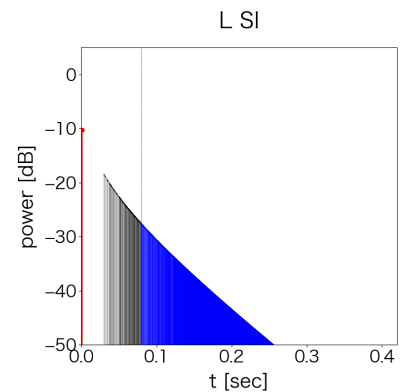


図 6.1.27 L_SI の音場の反射音の時系列応答

6.1.3 比較評価の方法

9 種類の評価音場の中から 2 つの音場の組み合わせをランダムに選択し、10 名の被験者に対しシェッフェの一対比較法に基づく心理実験を実施した。被験者には、実験 2 から得られた明瞭さの評価要因だと考えられる以下の 7 つの評価語それぞれについて、刺激 A と刺激 B のどちらがどの程度明瞭に感じたかについて、「どちらでもない」を含む 7 段階評価を求めた。

- 音色の明瞭さ (A の音色が明瞭 - B の音色が明瞭)
- こもり具合 (A がこもっていない - B がこもっていない)
- 輪郭 (A の輪郭がはっきりしている - B の輪郭がはっきりしている)
- 定位の分かりやすさ (A の定位が分かりやすい - B の定位が分かりやすい)
- 音像の広さ (A の音像が広い - B の音像が広い)
- 響きの量の適切さ (A の響きの量が適切 - B の響きの量が適切)
- 聴き取りやすさ (A が聴き取りやすい - B が聴き取りやすい)

提示された 2 つの刺激は、Cycling'74 社 Max7 で作成されたアプリケーションを操作することで任意のタイミングで再生、停止および刺激の切り替えを行うことができた。回答もアプリケーション上で行われた。また、刺激はループ再生された。図 6.1.28 は本実験で使用されたアプリケーションの GUI を示す。

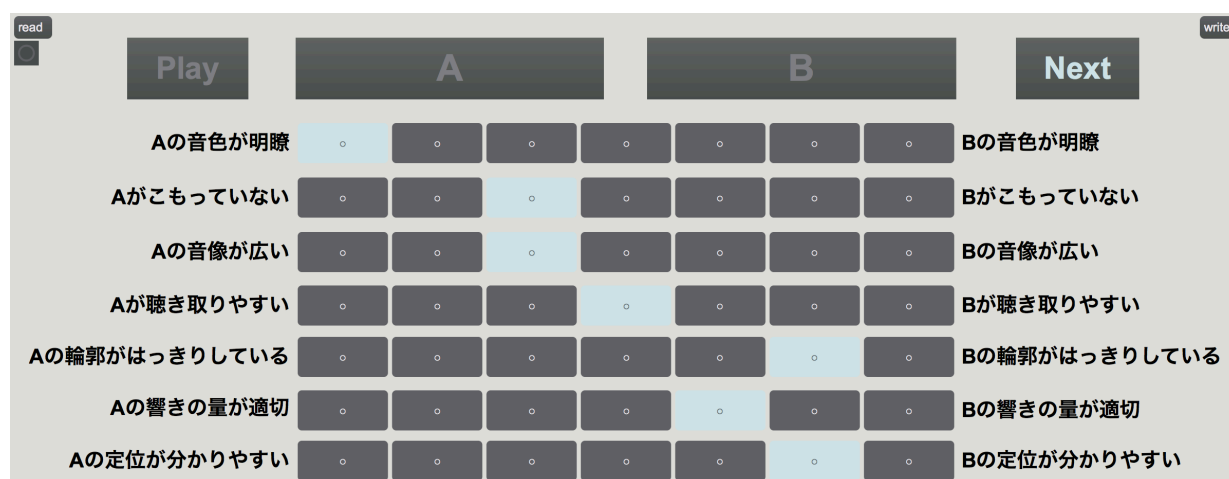


図 6.1.28 本実験で使用されたアプリケーションの GUI

6.2 分析

6.2.1 7つの評価要因の評価値の分析

10名の被験者から得られた「音色の明瞭さ」、「こもり具合」、「輪郭」、「定位の分かりやすさ」、「音像の広さ」、「響きの量の適切さ」、「聴き取りやすさ」のそれぞれの明瞭さの要因に基づく音場の評価値を分析した。図 6.2.1～図 6.2.7 は、7つの評価語についての音場の評価の平均値および 95% 信頼区間を示す。

7つ全ての評価語において、室容積 L の音場の評価が最も高く、次に室容積 M、最も低いものは室容積 S となっている。更に、初期反射音が全方向から到来する音場が、左右および前後から到来する音場と比べて高く評価されていることが分かる。一方、評価語によって、初期反射音が左右から到来する音場と前後から到来する音場の評価が異なる傾向が見て取れる。

図 6.2.1、図 6.2.2、図 6.2.3、図 6.2.4 は、それぞれ「音色の明瞭さ」、「こもり具合」、「響きの量の適切さ」および「聴き取りやすさ」の評価の平均値および 95% 信頼区間を示す。表 6.2.1 には、これらの評価語について、各室容積で、初期反射音が前後から到来する音場と左右から到来する音場の 2 群間に対応ありの t 検定を実施した結果を示す。全ての評価語、全ての室容積で有意な差は見られなかったが、「響きの量の適切さ」の室容積 M でのみ有意傾向が見られた。

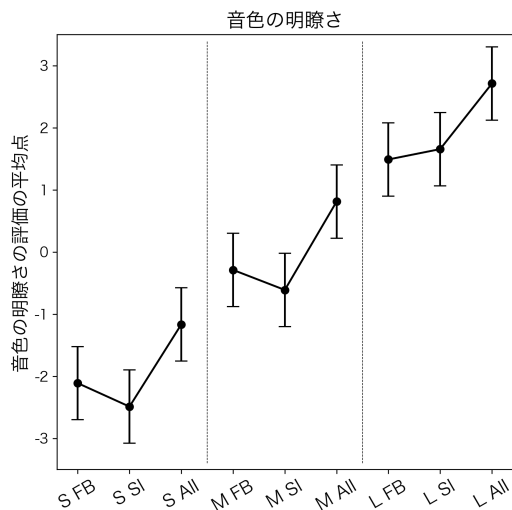


図 6.2.1 音色の明瞭さの評価の平均値および 95% 信頼区間

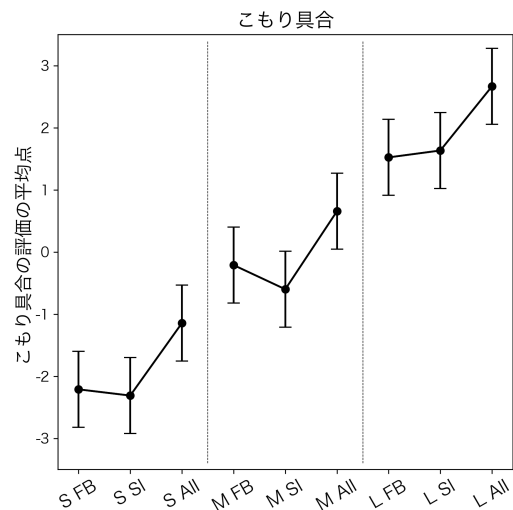


図 6.2.2 こもり具合の評価の平均値および 95% 信頼区間

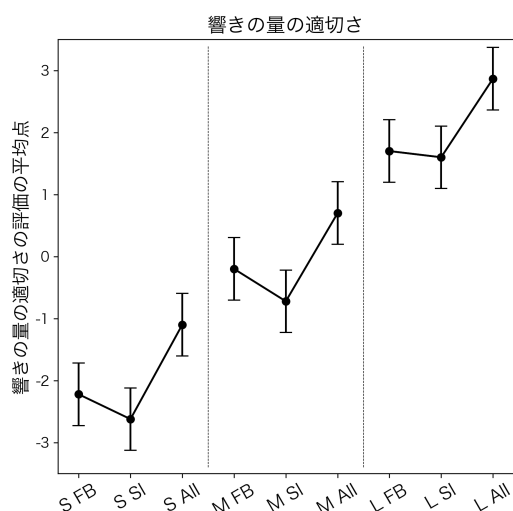


図 6.2.3 響きの量の適切さの評価の平均値および 95% 信頼区間

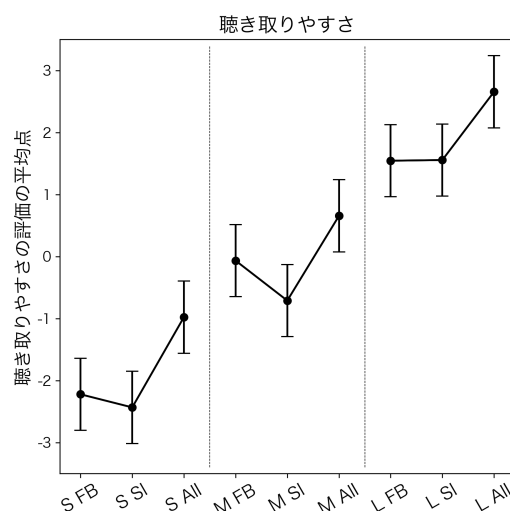


図 6.2.4 聴き取りやすさの評価の平均値および 95% 信頼区間

表 6.2.1 各評価要因、各室容積ごとの、初期反射音が前後から到来する音場と左右から到来する音場の評価値の対応ありの t 検定の結果。

評価要因	室容積	t-value	p-value
音色の明瞭さ	S	-0.89	0.40
音色の明瞭さ	M	-0.88	0.41
音色の明瞭さ	L	0.38	0.72
こもり具合	S	-0.32	0.76
こもり具合	M	-1.21	0.26
こもり具合	L	0.07	0.95
響きの量の適切さ	S	-1.55	0.16
響きの量の適切さ	M	-1.92	0.09 +
響きの量の適切さ	L	-0.42	0.68
聴き取りやすさ	S	-0.78	0.46
聴き取りやすさ	M	-1.84	0.10
聴き取りやすさ	L	-0.36	0.73

+ p < 0.10, * p < 0.05, ** p < 0.01, *** p < .001

図 6.2.5、図 6.2.6 は「輪郭」および「定位の分かりやすさ」の評価の平均値および 95% 信頼区間を示す。表 6.2.2 には、これらの評価語について、各室容積で、初期反射音が前後から到来する音場と左右から到来する音場の 2 群間で対応ありの t 検定を実施した結果を示す。全ての評価語、全ての室容積で有意な差は見られなかったが、「輪郭」の室容積 M、「定位の分かりやすさ」の室容積 S および室容積 M で有意傾向が見られた。前述した「音色の明瞭さ」、「こもり具合」、「響きの量の適切さ」および「聴き取りやすさ」と比べ、初期反射音が前後から到来した場合に左右から到来する場合よりも明瞭に評価される傾向が僅かに強い。

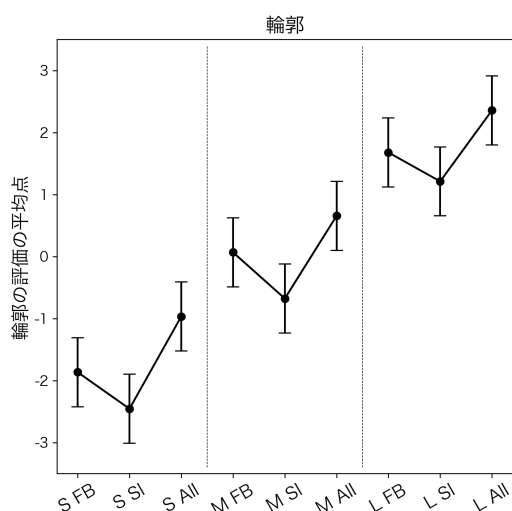


図 6.2.5 輪郭の評価の平均値および 95% 信頼区間

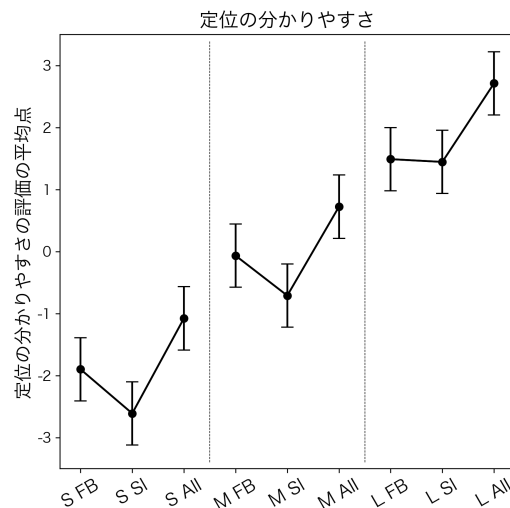


図 6.2.6 定位の分かりやすさの評価の平均値および 95% 信頼区間

表 6.2.2 各評価要因、各室容積ごとの、初期反射音が前後から到来する音場と左右から到来する音場の評価値の対応ありの t 検定の結果。

評価要因	室容積	t-value	p-value
輪郭	S	-1.32	0.22
輪郭	M	-1.93	0.09 +
輪郭	L	-1.51	0.17
定位の分かりやすさ	S	-2.07	0.07 +
定位の分かりやすさ	M	-2.01	0.08 +
定位の分かりやすさ	L	-0.31	0.76

+ p < 0.10, * p < 0.05, ** p < 0.01, *** p < .001

図 6.2.7 は「音像の広さ」の評価の平均値および 95% 信頼区間を示す。表 6.2.3 には、各室容積で、初期反射音が前後から到来する音場と左右から到来する音場の 2 群間で対応ありの t 検定を実施した結果を示す。全ての室容積で有意な差は見られなかったが、その他の評価語と比べると、初期反射音が左右から到来する場合に、前後から到来する場合よりも広く評価される傾向が僅かに強い。

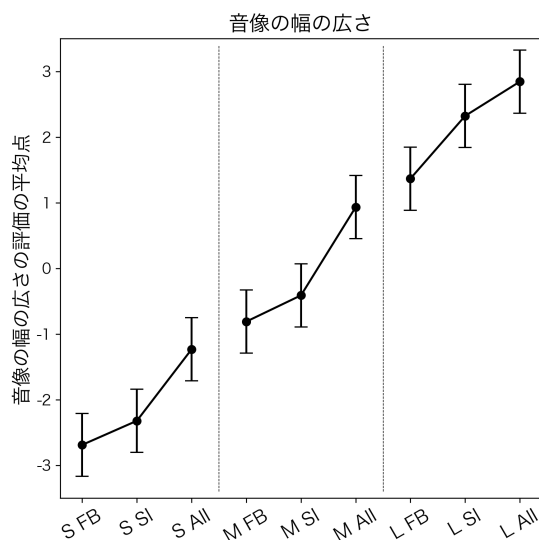


図 6.2.7 音像の広さの評価の平均値および 95% 信頼区間

表 6.2.3 音像の広さの、各室容積ごとの、初期反射音が前後から到来する音場と左右から到来する音場の評価値の対応ありの t 検定の結果。

評価要因	室容積	t-value	p-value
音像の広さ	S	0.47	0.65
音像の広さ	M	0.41	0.69
音像の広さ	L	1.63	0.14

+ p < 0.10, * p < 0.05, ** p < 0.01, *** p < .001

6.2.2 個人差を考慮した INDSCAL

被験者によって重視した評価要因に偏りが見られるかどうかを分析するため、個人間の違いを考慮した多次元尺度法のひとつである INDSCAL (INDividual Differences SCALing) による音場間の類似度の分析を行った。多次元尺度法とは、非類似性データの分析手法であり、対象物の関係を低次元空間における点の布置で表現する手法である。INDSCAL では、全被験者に共通する単一の低次元布置を考え、布置の各次元に与える重みに個人差があると考え (渡辺, 2002)。

9 つの音場に対する 7 つの評価項目それぞれの 10 名分の評価値を基に、被験者ごとに音場間のユークリッド距離を求めた。算出した音場間のユークリッド距離を基に 3 次元布置を作成する INDSCAL 分析を実施した。なお、3 次元布置でのストレス値は 0.097 であった。

図 6.2.8 のシンボル間の距離は音場の非類似度を示している。見やすさのため、音場の室容積ごとにシンボルの濃さを変え、更に次元 2 の値を反転してプロットしている。シンボル間の距離が近い音場は、似た明瞭さを持つ音場として判断されたと言える。

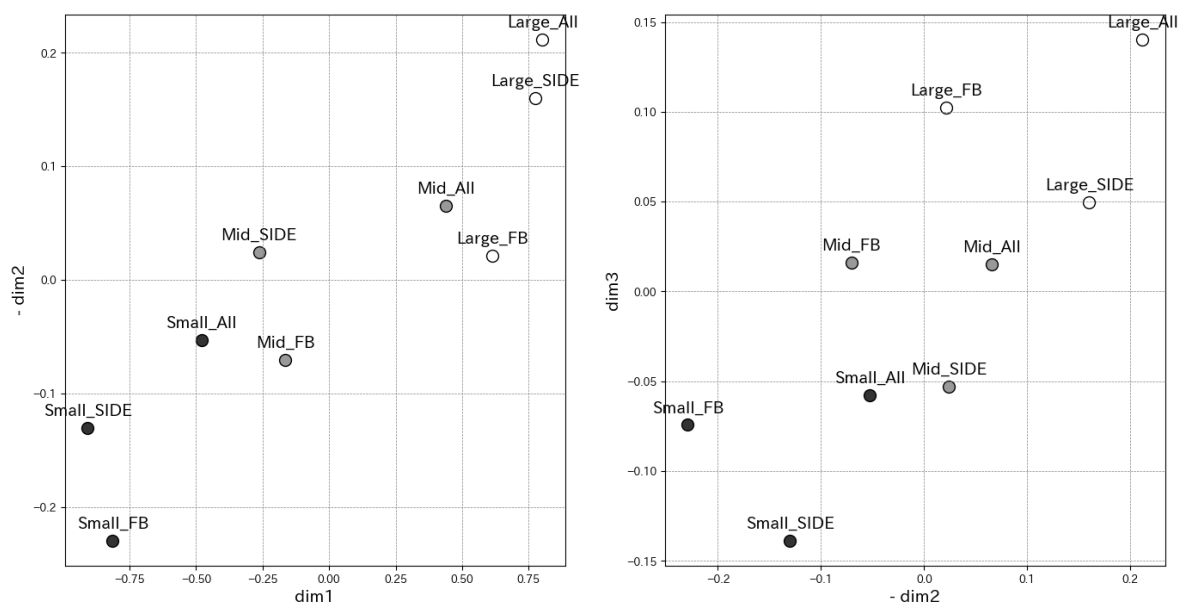


図 6.2.8 INDSCAL による音場の 3 次元布置。シンボル間の距離が音場間の非類似度を示している。見やすさのため、音場の室容積ごとにシンボルの濃さを変え、更に次元 2 の値を反転してプロットしている。

図 6.2.8 より、次元 1 は室容積が大きいほど値が高くなり、室容積が等しい場合は初期反射音の到来方向が全方向で高い傾向が見られる。これは 7 つ全ての明瞭さに共通する傾向である。次元 2 は室容積が大きいほど値が高くなり、室容積が等しい場合は初期反射音の到来方向が全方向、左右、前後の順で値が高い傾向が見られる。次元 3 は室容積が大きいほど値が高くなり、室容積が等しい場合は初期反射音の到来方向が全方向、前後、左右の順で値が高い傾向が見られる。次元 1 は約-0.8 0.8 の範囲を取り、次元 2 は約-0.2 0.2、次元 3 は約-0.15 0.15 の範囲を取るため、次元 1 の影響が最も支配的だと言える。

図 6.2.9 は、各次元に対する被験者の重み付けを表す。アルファベットは被験者の ID を示し、各被験者が 3 つの次元それぞれにどの程度の重みを持って評価したかについて布置で表されている。例えば、被験者 A は図 6.2.8 の次元 1 を約 0.8 倍、次元 2 を約 2 倍、次元 3 を約 4 倍に引き伸ばし、あるいは縮めて評価したことを示す。図 6.2.8 および図 6.2.9 より、被験者 C、J の 2 名は次元 2 を次元 1 よりも重視し、被験者 D、F、G、I の 4 名は次元 3 を次元 1 よりも重視し、被験者 A、B、D、E、F、G、I の 7 名は次元 3 を次元 2 よりも重視したことが分かる。

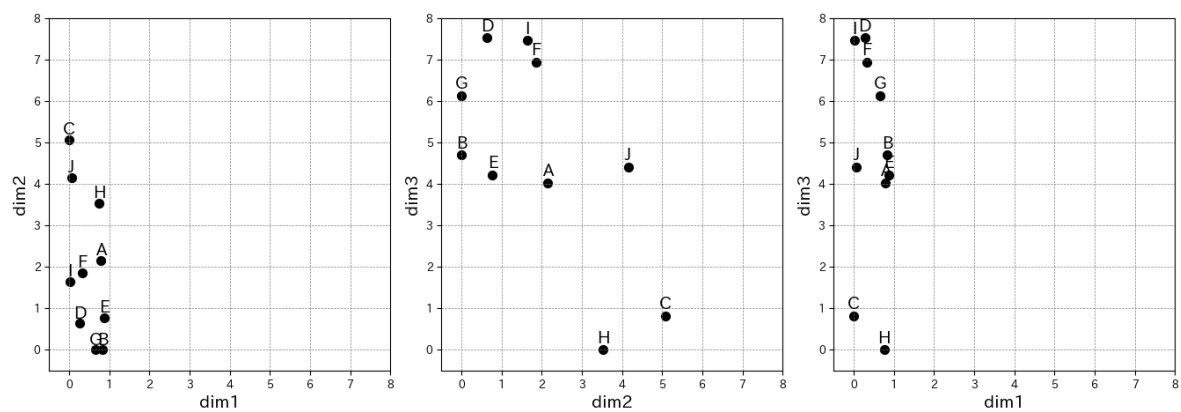


図 6.2.9 INDSCAL による被験者の各次元に対する重み付けの分析結果。アルファベットは被験者の ID を示す。

6.3 考察

6.3.1 音楽の明瞭さの評価要因

音楽の明瞭さの 7 つの評価要因の INSCAL 分析より、音楽の明瞭さの評価要因は、評価傾向の異なる 3 種類に分けることができると考えられる。

1 つ目は「音色の明瞭さ」、「こもり具合」、「響きの量の適切さ」および「聴き取りやすさ」を含む、音色の明瞭さである。これらの評価要因は Beranek (2004) によって定義されている「水平的な明瞭さ」に内容が類似している。水平的な明瞭さは、連続で演奏される音が区別できるかどうか、ある瞬間に鳴っている音が前の音の残響等に影響され、濁ったり聴こえづらくなるかどうかといった感覚である。これら 4 つの評価語は、室容積が大きく、初期反射音の到来方向の偏りが低い場合に明瞭と評価された。この傾向は、INSCAL 分析の結果の次元 1 の布置と対応している。次元 1 は 3 つの次元の中でも最も重視された感覚である。

2 つ目は「輪郭」および「定位の分かりやすさ」を含む、音像の明瞭さである。これらの評価要因は Beranek によって定義されている「垂直的な明瞭さ」に内容が類似している。垂直的な明瞭さは、同時に鳴っている音が区別できるかどうかといった感覚である。これら 2 つの評価語は、室容積が大きく、初期反射音の到来方向の偏りが低い場合に明瞭と評価され、更に初期反射音が前後から到来する場合に左右から到来する場合よりも明瞭だと評価される傾向があった。この傾向は、INSCAL 分析の結果の次元 3 の布置と対応している。10 名の被験者のうち 4 名は次元 3 を次元 1 よりも重視したため、人によっては非常に重要な要因だと言える。

3 つ目は「音像の広さ」である。上記の音色の明瞭さおよび音像の明瞭さと同様に、室容積が大きく、初期反射音の到来方向の偏りが低い場合に高く（広く）評価された。更に、初期反射音が左右から到来する場合に、前後から到来する場合よりも高く評価される傾向があった。これは、音像の明瞭さと逆の傾向にある。この傾向は、INSCAL 分析の結果の次元 2 の負の値と対応している。

6.3.2 「音色の明瞭さ」の評価要因と物理量との対応

「音色の明瞭さ」、「こもり具合」、「響きの量の適切さ」および「聴き取りやすさ」の評価要因を含む音色の明瞭さは、室容積が大きく、初期反射音の到来方向の偏りが小さい場合に明瞭と評価された。

まず、室容積の変化に対応する物理量を検討する。表 6.1.2 より、本実験で生成したインパルス応答の初期反射音レベルは室容積が小さい順から直接音に対して -9.8 dB、 -10.0 dB、 -10.5 dB である。これは実験 1 から得られた結果と対応しており、初期反射音の最適レベルである -22 dB に近いほど評価が高い。小空間において初期反射音は明瞭さに寄与しないことが改めて確認された。一方、音場間の音圧レベル差は 1 dB 以下と小さく初期反射音レベルにのみ着目して評価されたとは考え

にくい。ここで、第 1 章で言及した、初期反射音を吸音し後期残響音を響きとして残すコントロールルームの設計思想であるリフレクションフリーゾーン (RFZ) に基づき、音場の「初期反射音レベルに対する後期残響音レベル」を算出する。一定の後期残響音が「響きの量の適切さ」や「聴き取りやすさ」に寄与すると考える。後期残響音レベルは室容積が小さい順から -24.8 dB、 -22.4 dB、 -20.1 dB であり、音場間の音圧レベル差は約 2 dB である。算出した「初期反射音レベルに対する後期残響音レベル」の値は、室容積が小さい順から -15.0 dB、 -12.4 dB、 -9.6 dB となった。

次に、初期反射音の到来方向の偏りに対応する物理量を検討する。全方向の音場が最も明瞭とされた理由は不明だが、(Hanyu and Kimura, 1996) は残響音を左右 2 方向だけから提示した条件より、8 方向からバランスよく残響音を提示する条件の LEV が大きくなると報告している。メカニズムの解明には追加の調査が必要であるが、人の音の空間の分解能等が関係しているかもしれない。ここでは、全方向に均等にエネルギーが分散していれば空間バランスは良く、ある方向に偏っている場合は空間バランスは悪くなると考えることにする。初期反射音の到来方向が全方向の音場は、初期反射音の生成範囲を仮想球内の全範囲に設定したものである。前後あるいは側方から到来する音場は、初期反射音の生成範囲を仮想球内の 4/16 に限定したものである。ここで、分割された各範囲の初期反射音のエネルギーを算出し、平均エネルギーからの偏差を空間バランスの指標として算出する。まず、式 6.1 によって、各範囲のエネルギーを全体の総エネルギーで基準化する。ここで、 e_i は第 i エリアの元の音響エネルギー、 N はエリアの総数（この場合は 16）を示す。

$$e'_i = \frac{e_i}{\sum_{j=1}^N e_j} \quad (6.1)$$

次に、式 6.2 によって、基準化された音響エネルギー値に基づき初期反射音の到来方向の平均絶対偏差を算出することで、到来方向の空間的偏りを表す指標を計算し、これを仮に SV (Spatial Variability) と呼ぶ。 μ' は基準化された値の平均を示す。

$$SV = \frac{1}{N} \sum_{i=1}^N |e'_i - \mu'| \quad (6.2)$$

本実験で生成したインパルス応答に対して SV を算出したところ、初期反射音の到来方向が全方向の音場で 0、前後あるいは側方の音場で 0.094 となった。この値が大きいほど音色の明瞭さに影響すると考え、「初期反射音レベルに対する後期残響音レベル」における初期反射音レベルに対する重みとして導入し、式 6.3 に示すように指標化する。これを仮に、 C_{SR} (SmallRoom) と呼ぶ。

$$C_{SR} = 10 \log_{10} \left(\frac{\int_{80}^{\infty} p^2(t) dt}{\int_5^{80} p^2(t) dt (1 + SV)} \right) [\text{dB}] \quad (6.3)$$

図 6.3.1 は、本実験で生成したインパルス応答から直接算出した C_{SR} の値と、IDSCAL 分析から得られた次元 1 の値を示す。ピアソンの相関係数は 0.959 ($p < 0.000$) であり、聴感とよく対応し

ていることが分かる。

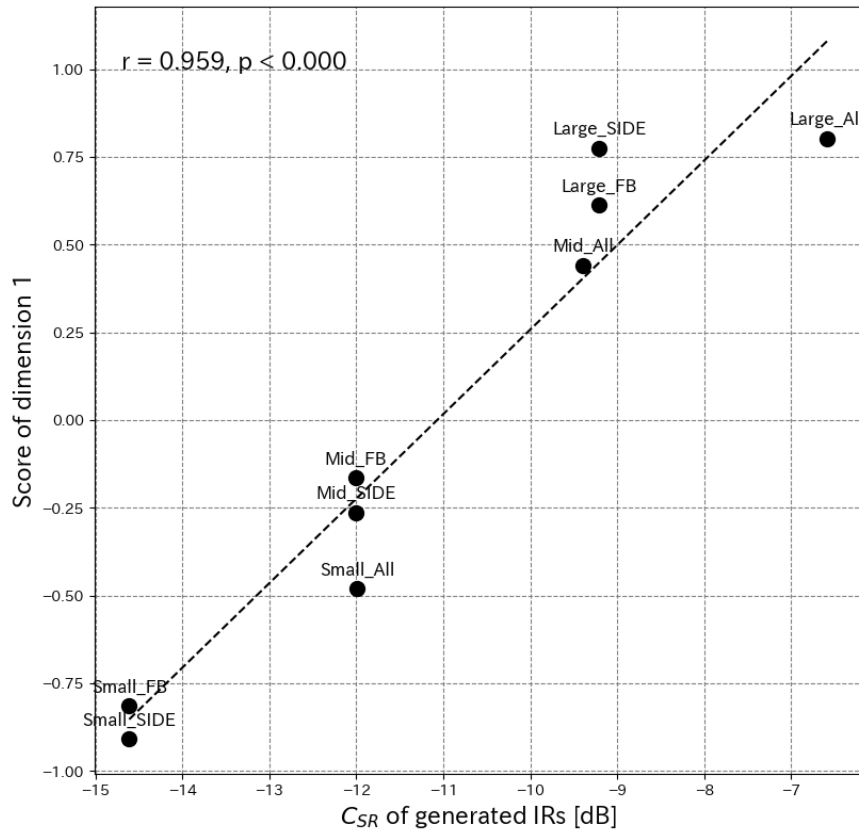


図 6.3.1 本実験で生成したインパルス応答から直接算出した C_{SR} の値と、IDSCAL 分析から得られた次元 1 の値を示す。

一方、本実験は無響室ではなく、響きの存在するスタジオにおいて実施したため、被験者は前述したインパルス応答のみを直接聴いて評価したとは言えない。 C_{SR} の聴感との対応を更に調べるため、実験環境の聴取位置にて、9 種類の評価音場のインパルス応答を測定した。測定は Log-Swept-Sine 信号 (Farina, 2007) を各評価音場のインパルス応答に畳み込み、第 3 章で述べた再生システムで再生することで行った。測定には SCHOEPS 社の CMC6 プリアンプと全指向性の MK2 カプセルとの組み合わせを用いた。到来方向の空間的偏りを表す SV (Spatial Variability) の算出のために、CMC6 プリアンプと双指向性の MK8 マイクとの組み合わせも用いた。双指向性マイクを用いた測定では、マイクを正面、側方それぞれに向けた 2 本を用いた。

図 6.3.2 には、9 種類の評価音場の 1/3 オクターブバンドの全指向性マイクの周波数特性を示す。図より、評価音場の違いは周波数特性にはほとんど影響していないことが分かる。一方で、4000 Hz 以上からは恐らくマイクの指向性による影響、そして 250 Hz 以下からは測定値にばらつきが見られるため、物理量の算出には、 C_{80} などの室内音響指標でも採用されている 500 Hz および 1000 Hz の帯域の算術平均値を用いた。

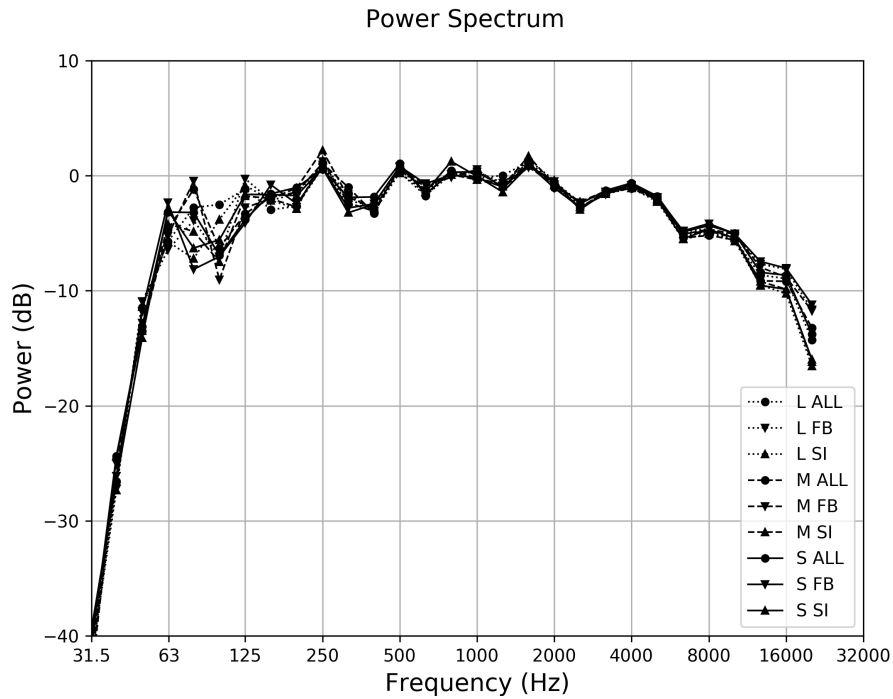


図 6.3.2 9 種類の評価音場のインパルス応答の 1/3 オクターブバンドの周波数応答。

図 6.3.3 は、実験環境にて実測したインパルス応答から算出した C_{SR} の値と、IDSCAL 分析から得られた次元 1 の値を示す。なお、双指向性マイクを 2 本のみ用いているため、空間の分割数は 2 で算出した。ピアソンの相関係数は 0.704 ($p = 0.034$) であり、実測値についても聴感と十分に対応があると言える。一方で、インパルス応答から直接算出した結果と比較すると相関は低下している。 C_{SR} の実測値の範囲は -9.5 -6.5 の狭い範囲にあり、生成したインパルス応答だけでなく実験環境そのものの響きも加わったことで、刺激の室容積の違いが十分に値に表れなかったからだと考えられる。

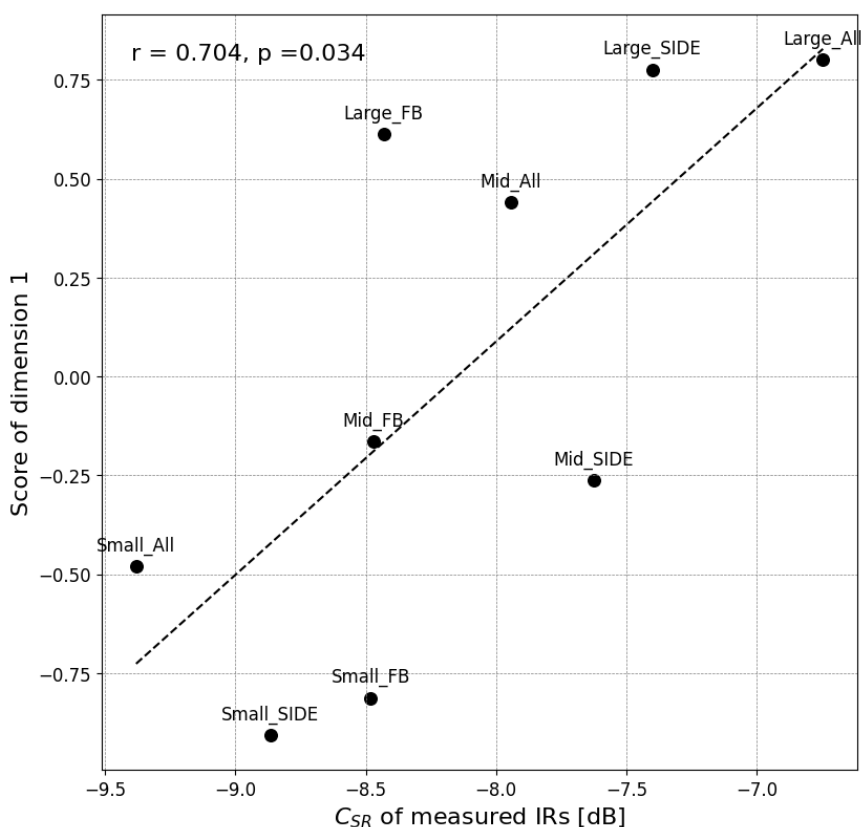


図 6.3.3 実験環境にて実測したインパルス応答から直接算出した C_{SR} の値と、IDSCAL 分析から得られた次元 1 の値を示す。

C_{SR} が実空間において感じる明瞭さをどの程度説明できるかについては、追加の検証が必要である。単純なエネルギー比に基づく C_{SR} は、大きすぎる後期残響音や小さすぎる初期反射音が明瞭さにマイナスに働くケースを表現できない。更に、実験 1 からは、同じ初期反射音レベルであっても室容積が小さいほど直接音に与える影響が大きい結果が得られている。初期反射音レベルだけでなく、密度や減衰などの要素も取り入れた指標が必要である可能性がある。

更に、本研究では C_{80} の指標に習い、直接音の到来後 30~80 ms 以内に到来する反射音を初期反射音として扱ったが、初期反射音として定義する区間についても検討が必要である。また、どのような音源であれば既存の明瞭度指標ではなく本研究で提案した指標を用いるべきかについても、明確ではない。これらを検証するには、様々な音源を用いた実空間での検証が必要である。

6.3.3 「音像の明瞭さ」と「音像の広さ」の物理量との対応

「輪郭」および「定位の分かりやすさ」の評価要因を含む音像の明瞭さは、音色の明瞭さと同様に、室容積が大きく、初期反射音の到来方向の偏りが小さい場合に明瞭と評価された。更に、初期反射音が前後から到来する場合に、左右から到来する場合と比べて明瞭だと評価される傾向が見られた。一方、「音像の広さ」の評価では、左右から到来する場合に前後から到来する場合と比べて明瞭だと

評価される傾向が見られた。

Begault and Anderson (2001) や Griesinger (2000) によって、直接音と同じ方向から到来する初期反射音は、その他の方向から到来する場合と比べて知覚されにくいことが明らかにされている。音像の明瞭さ、そして音色の明瞭さの一部の評価要因で前後が高い評価な傾向にあるのは、このことが理由だと考えられる。一方、音像の明瞭さで特にこの傾向が強いのは、左右から到来する反射音によって「音像が広がる」際に、単に音像が広がるのではなく、ぼけたように広がるからだと考えることができる。「音像の明瞭さ」と「音像の広さ」は、到来方向の空間的偏りの観点ではトレードオフの関係にある。

音像の広さは IACC (両耳間相関度) と対応することが明らかにされている (Bömer et al., 2011)。IACC とは Inter Aural Correlation Coefficient の略で、右耳と左耳に入力される音響信号の相関を示し、音の広がり感を表す物理指標の 1 つとして規格化されている (ISO3382-1, 2009)。IACC は 0 から 1 の間の数値で表される。9 種類の評価音場について、実験室の聴取位置における IACC を、Bruel Kjaer 社 Head and Torso Simulator Type 4128-C のダミーヘッドを使用し測定した。測定は前述した評価音場のインパルス応答の周波数応答の測定と同様に行った。IACC の値は、インパルス応答を 1/3 オクターブバンド (125 Hz~12000 Hz) に分割し、全帯域の平均を用いた。

図 6.3.4 は、9 種類の評価音場の INDSCAL 分析から得られた布置の次元 2 の値と、1-IACC (1/3 オクターブバンドの平均値) の平均値を示す。評価音場の次元 2 と、1-IACC の平均値のピアソンの相関係数は 0.468 ($p < 0.020$) であり、中程度の正の相関関係があると言える。しかし、IACC は室容積の違いを表現できていない。まずは C_{SR} の指標に基づき設計を行い、その後で IACC を基準に音像の明瞭さと音像の広さのバランスを考慮することが望ましいと言える。

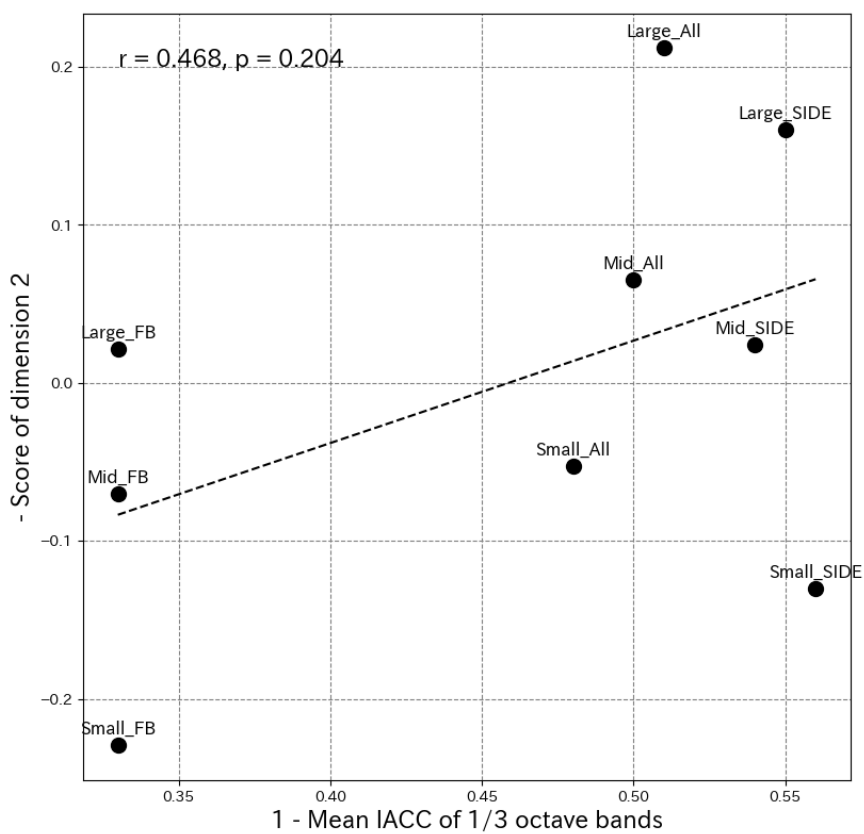


図 6.3.4 9 種類の評価音場の INDSCAL 分析から得られた布置の次元 2 の値と、1-IACC (1/3 オクターブバンドの平均値) を示す。

6.4 結論

本実験では、小空間での市販音源の聴取における明瞭さの評価要因だと考えられる 7 つの評価語について、音場の物理量との関係を調査した。被験者には室容積が 3 種類、30～80 ms までの初期反射音の到来方向が 3 種類の合計 9 種類の音場の評価を求めた。調査の結果、以下の結論が得られた。

- 小空間における市販音源の聴取では、音楽の明瞭さの評価要因は評価傾向の異なる 3 種類に分けることができる。それぞれ、「音色の明瞭さ」、「音像の明瞭さ」、「音像の広さ」である。
- 3 つ全ての明瞭さは、室容積が大きく、初期反射音の到来方向の偏りが小さい場合に明瞭と評価される。
- 小空間で適用できる音楽の明瞭さの指標として、「初期反射音レベルに対する後期残響音レベル」を算出したところ、聴感と良い対応が得られた。指標には、初期反射音の到来方向の空間的偏りに関する重みも導入した。
- 音像の明瞭さは、初期反射音が前後から到来する場合に左右から到来する場合よりも明瞭とされ、音像の広さとトレードオフの関係にある。
- 音像の広さは、IACC と中程度の正の相関関係が見られた。

第7章

総括

7.1 音楽の明瞭さに最適な初期反射音の音圧レベル

コンサートホール等の大空間においては、初期反射音の音圧レベルが高いほど音楽が明瞭だと評価されることが先行研究によって明らかになっている。しかし、筆者による研究 (Imamura et al., 2014) では、スタジオ施設等の小空間における市販音源の聴取においては、初期反射音の音圧レベルが高いほど明瞭さが低く評価される可能性が示唆された。

この結果を受け、本研究の実験1では、音場の室容積が $60\text{ m}^3\sim 400\text{ m}^3$ 、かつ後期残響音が存在しない小空間を模擬した統制条件で、明瞭さに最適とされる初期反射音の音圧レベルを調査した。聴取音源として、スピーカー聴取を想定したステレオ市販音源、および演奏音を想定したモノラル無響音源の2種類を用いた。被験者に着目を求める感覚を「音色の明瞭さ」と「音像の明瞭さ」の2つとした。実験の結果、明瞭さに最適な初期反射音の音圧レベルは、聴取する音源、そして着目する明瞭さによって有意に異なることが明らかになった。

更に、無響音源の聴取においては、音色の明瞭さは直接音に対して約 -10 dB 、音像の明瞭さは約 -20 dB の初期反射音の音圧レベルが最適とされた。一方、市販音源の聴取においては、音色および音像の明瞭さのどちらも、直接音に対して約 -23 dB の低い初期反射音の音圧レベルが最適とされた。聴取音源によって明瞭さに最適な初期反射音の音圧レベルが異なった理由には、その音源を聴取する文脈が関係している可能性がある。無響音源はそれ自体に響きを含まない単体の楽器音であり、自然な聴取のためには響きによる直接音の補強を必要とする。一方で、市販音源は制作の過程で響きが付加されていることが多く、聴取に必要とされる追加の響きは少ないと言える。あるいは、音源のエンベロープ変化が理由である可能性も考えられる。無響音源は音源のエンベロープ変化が激しく、ある瞬間に鳴った楽音は、前の楽音に付帯する残響音の影響を受けづらい。これによって、初期反射音の音圧レベルが高い場合でも、音色の明瞭さを損ねること無く直接音が補強されたと考えることができる。理由の解明には更なる検討が必要だが、演奏音の聴取と市販音源の聴取、つまりコンサートホールでの聴取の文脈と、コントロールルームやリスニングルームでの市販音源の聴取の文脈では、異なる初期反射音の音圧レベルが最適とされることが明らかとなった。

一方で、最適な初期反射音の音圧レベルには個人差があり、約 10 dB～20 dB の差が見られた。更に、高い音圧レベルを明瞭だと判断する被験者は、初期反射音の到来方向の偏りにも着目して判断を行うことが示唆された。

7.2 音楽の明瞭さの評価要因

個人によって着目する明瞭さの評価要因が異なる可能性が示唆されたため、実験 2 では、レパトリリーグリッド法に基づくインタビュー調査によって、人がどのような要因に着目して音楽の明瞭さを評価しているかを調べた。聴取音源は実験 1 と同じ市販音源および無響音源の 2 種類、評価音場の容積は 125 m³ および 512 m³ の 2 つとした。調査の結果、小空間での音楽聴取における音楽の明瞭さは、表 7.2.1 に示す複数の要因によって評価されることが明らかになった。

表 7.2.1 小空間での音楽聴取における音楽の明瞭さの評価要因

評価要因	明瞭さの要因	-	不明瞭さの要因
音色の明瞭さ	音色が明瞭	-	音色が不明瞭
こもり具合	高域がしっかりした、音色が明るい	-	こもっている
音の距離感	音が近い	-	音が遠い
輪郭	輪郭がはっきりしている	-	輪郭がぼやけている
定位の分かりやすさ	定位が分かりやすい	-	定位が分かりにくい
音像の広さ	音像が広い	-	音像が狭い
響きの量の適切さ	響きが少ない	-	響きが多い
聴き取りやすさ	聴き取りやすい、はっきり聴こえる	-	聴き取りづらい

聴取する音源によって、上記の中でも着目される要因が異なった。2 つの音源に共通して着目されていた要因は、「こもり具合」、「響きの量の適切さ」、および「聴き取りやすさ」であった。無響音源の聴取では更に「音色の明瞭さ」、「音の距離感」のに着目され、市販音源の聴取では「輪郭」、「定位の分かりやすさ」、「音像の広さ」に着目された。

無響音源の聴取においては、評価要因のほとんどが Beranek (2004) の音色の明瞭さ（水平的な明瞭さ）に関係していた。また、室容積が小さいほど、初期反射音の音圧レベルが大きいほど明瞭だと評価された。これは実験 1 の結果と対応する。無響音源の聴取の文脈で音色の明瞭さを評価する場合は、小空間においても C_{80} の考え方が有効であることが明らかになった。

市販音源の聴取においては、音色の明瞭さに加え、「輪郭」や「定位の分かりやすさ」、「音像の広さ」といった音像に関する要因によっても明瞭さが評価された。更に、個人によって評価要因が異なった。ある被験者グループは音色の明瞭さや音像の広さ、そして響きの量に着目し、室容積の違いに基づき明瞭さを評価した。もう 1 つのグループは輪郭や定位の分かりやすさに着目し、初期反射音の到来方向の違いに基づき明瞭さを評価した。このように、小空間での市販音源の聴取においては、明瞭さは複数の異なる要因によって評価されることが明らかになった。

7.3 小空間における明瞭さに対応する音場の物理量

実験3では、市販音源の聴取における明瞭さの評価要因である「音色の明瞭さ」、「こもり具合」、「輪郭」、「定位の分かりやすさ」、「音像の広さ」、「響きの量の適切さ」、「聴き取りやすさ」と、音場の物理量との対応を調査した。

調査の結果、これらの評価要因は「音色の明瞭さ」、「こもり具合」、「響きの量の適切さ」、「聴き取りやすさ」を含む音色の明瞭さ（水平的な明瞭さ）、および「輪郭」、「定位の分かりやすさ」を含む音像の明瞭さ（垂直的な明瞭さ）、そして「音像の広さ」の3つに区別されることが明らかになった。

- 音色の明瞭さ（水平的な明瞭さ）：音色の明瞭さ、こもり具合、響きの量の適切さ、聴き取りやすさ
- 音像の明瞭さ（垂直的な明瞭さ）：音像の輪郭、定位の分かりやすさ
- 音像の広さ：音像の広さ

3つ全ての音楽の明瞭さに対応する音場の物理的な要因は、「室容積」および「初期反射音の方向的偏り具合」であった。室容積は大きいほど、そして初期反射音の到来方向の偏りが小さい場合に明瞭だと評価された。

室容積の違いは、音場の様々な物理量に影響する。実験3で生成した9つのインパルス応答の分析より、室容積が大きくなるほど初期反射音レベルが小さく、かつ後期残響音レベルが大きくなることが分かった。そこで、初期反射音を吸音し後期残響音を響きとして残すコントロールルームの設計思想であるリフレクションフリーゾーン（RFZ）に基づき、音場の「初期反射音レベルに対する後期残響音レベル」の値を算出した。更に、初期反射音の到来方向の偏りに対応する物理量として、空間を複数の範囲に分割し、それぞれの範囲の反射音のエネルギーを求め、到来方向のエネルギーの平均絶対偏差をSV（Spatial Variability）として導出し、初期反射音のエネルギーに対する重みとして反映させた。ここでは、仮に指標を C_{SR} （SmallRoom）と呼び、心理実験から得られた心理量との対応を見たところ、生成したインパルス応答から算出した C_{SR} と心理量との相関は 0.959 ($p < 0.000$)、実験環境で実測したインパルス応答から算出した C_{SR} と心理量との相関は 0.704 ($p = 0.034$) と、良い対応が得られた。

一方で、 C_{SR} は単純なエネルギー比に基づくため、大きすぎる後期残響音や小さすぎる初期反射音が明瞭さにマイナスに働くケースを表現できない。更に、同じ初期反射音レベルであっても室容積違いが評価に影響した結果も得られているため、密度や減衰などの要素も取り入れた指標が必要である可能性がある。また、どのような音源であれば既存の明瞭度指標ではなく本研究で提案した指標を用いるべきかについても、明確ではない。様々な音源を用いた実空間での検証が必要である。

音像の明瞭さは、音色の明瞭さと同様に、室容積が大きく初期反射音の到来方向の偏りが小さい場合に明瞭と評価された。更に、初期反射音が前後から到来する場合に、左右から到来する場合と

比べて明瞭だと評価される傾向が見られた。一方、「音像の広さ」は、左右から到来する場合に前後から到来する場合と比べて明瞭だと評価される傾向が見られた。

音像の広さに関する明瞭さは、1-IACC（両耳間相関度）と中程度の相関が見られた。しかし、IACCは一般的に音場の側方反射音が多いほど低くなる傾向があるため、音像の明瞭さと音像の広さはトレードオフの関係となっている。まずは前述した指標に基づき全体の明瞭さの設計を行い、その後で空間の目的や個人の好みに照らし合わせ、音像の明瞭さと音像の広さのバランスを取ることが有効だと考えられる。

7.4 室内音響設計への応用と今後の研究の課題

本研究によって、スタジオ施設のコントロールルームやリスニングルームの音響設計に貢献する、いくつかの新たな知見が得られた。

室容積に関わらず、それが生楽器の聴取を目的とした演奏空間であれば、直接音に対して -10 dBほどの高いレベルの初期反射音を供給することで明瞭さが向上する。このような音楽聴取においては、初期反射音が付加されることによって直接音の音質や距離感、聴き取りやすさが向上し、音色の明瞭さが高くなる。

一方で、市販音源の聴取を目的としたコントロールルームやリスニングルームでは、直接音に対する初期反射音レベルを約 -23 dBと低くすることで明瞭さが向上する。こうした小空間においては、明瞭さの観点からは初期反射音を積極的に吸音することが必要である。また、一定の後期残響音が付加されることも明瞭さに寄与する。更に、一定の方向から到来する強い初期反射音を拡散する等、初期反射音の空間的な偏りを軽減することで、明瞭さが向上する可能性がある。

人によっては、音色だけでなく、音像の明瞭さや、音像の広さに着目した明瞭さを判断することがある。輪郭や定位の分かりやすさといった音像の明瞭さを重視したい場合は、直接音と異なる方向からの反射音をなるべく除去することが効果的である。音像の広さに着目する場合は、側方からの初期反射音を供給することで、明瞭さを向上させられる可能性がある。音像の明瞭さと音像の幅の広さはトレードオフの関係にあるため、まずは初期反射音レベルを低減し、反射音の偏りを避け、その上でIACCや側方エネルギーに着目した検討を行うと良いだろう。

音楽聴取における音場の明瞭さの向上に向けては、多くの課題がある。まず、聴取する音源による影響の調査が必要である。今回得られた結果は既存のコンサートホールでの結果と異なる傾向にあるが、ではどのような音源であれば生楽器の聴取、あるいは市販音源の聴取になるのか、明確に説明できていない。音源の特徴によって最適な初期反射音レベルが異なるとしたら、聴取する音源を基に最適な明瞭さを実現できる可能性がある。更に、明瞭さに対応する指標として、単純なエネルギー比ではなく、反射音の減衰や密度を考慮した上で適切な積分区間を検討するなど、より聴感と良く対応する指標を検討する必要がある。

音楽の聴取において、人が響きを知覚するプロセスは非常に複雑である。様々な時間的・空間的構造を持つ反射音と、複雑な知覚・心理構造を持つ人の主観的判断との関係性を定量的に扱うことは、挑戦的な課題であった。しかし、僅かながら本研究は実際の音楽聴取環境の向上に有効であり得る知見を得ることができたと思う。

音楽と音響は切り離せない関係にある。人が音楽を聴く時、そこには必ず響きが存在する。音響が、音楽家のイメージする音楽を見事に体現した時、我々は価値観さえ変わってしまうような、強烈な音楽体験をする。

本研究が、少しでも多くの人に、そのような素敵な体験をもたらす礎となることを願う。

謝辞

本研究を遂行、まとめるにあたり、多くの方々にご指導およびご協力を頂きました。

東京藝術大学音楽環境創造科、丸井淳史准教授には、終始親身にご指導、ご鞭撻を頂きました。丸井先生からは、実験計画や分析に関する知識だけでなく、真理を追求する研究精神、そして学問の素晴らしさを教えて頂きました。先生との議論から得られた学びは、今後の人生の羅針盤になると考えております。心より感謝申し上げます。

東京藝術大学音楽環境創造科、亀川徹教授には、常に深いご指導と激励を頂きました。制作を行う立場からの視点で、室内音響や音楽の主観印象に関する鋭いご意見を頂きました。心より感謝申し上げます。

株式会社ソナ、東京藝術大学非常勤講師の中原雅考先生には、関連研究を含め、多岐にわたる多大なるご助言を頂きました。筆者が東京藝術大学の学士課程で行った研究では、実験に関するご意見やご指導だけでなく、制作した吸音パネルの素材のご支援も頂きました。また、何よりも、魅力溢れる授業で、筆者が室内音響の研究に進んだきっかけ与えて下さったことに対し、深く感謝申し上げます。

ロチェスター工科大学の Sungyoung Kim 先生には、本研究内容に関する多くのご助言を頂きました。深く感謝申し上げます。

東京藝術大学音楽環境創造科、西岡龍彦教授には、本論文をまとめるにあたり、多くのご意見を頂きました。深く感謝申し上げます。

本研究を遂行した博士後期課程の3年間に渡り、公益財団法人尚志社から奨学金を受けました。また、尚志社社友の皆様からは常に激励を頂きました。ご支援に対し感謝申し上げます。

東京藝術大学音楽環境創造科、大学院音楽音響創造の学生の皆様には、心理実験の被験者としてご協力頂きました。ここに感謝申し上げます。

最後に、温かい励ましを常に送り続けてくれた家族に深く感謝します。

以上のすべての方々に、重ねて感謝いたします。

発表論文等一覧

A. 原著論文

題目	掲載誌	著者
小空間における音楽の明瞭さに関する 評価要因の調査	日本音響学会誌 (2018 年発行予定)	今村秀隆, 丸井 淳史, 亀川徹
An assessment of a spatial ear training program for perceived auditory source width	The Journal of the Acoustical So- ciety of America, 142(2), EL201- EL204 (2017)	S.Kim, H.Imamura

B. 国際会議プロシーディングス

題目	掲載誌	著者
Efficacy of a new spatial ear training program for “Ensemble width” and “Individual source width”	The Journal of the Acoustical So- ciety of America, 140(4), 2987- 2987 (2016)	H.Imamura, S.Kim
Height Loudspeaker Position and Its Influence on Listeners’ Hedonic Re- sponses	Audio Engineering Society In- ternational Conference on Sound Field Control (2016)	S.Kim, M.Indelicato, H.Imamura, H.Miyazaki
Influence of Directional Differences of First Reflections in Small Spaces on Perceived Clarity	Audio Engineering Society Con- vention 136, eBrief (2014)	H.Imamura, A.Marui, T.Kamekawa, M.Nakahara
Influence of First Reflections in Lis- tening Room on Subjective Listener Impression of Reproduced Sound	Audio Engineering Society Con- vention 134, eBrief (2013)	H.Imamura, A.Marui, T.Kamekawa, M.Nakahara

C. 学術講演

題目	掲載誌	著者
スピーカーを使用した音楽聴取における明瞭さの評価要因	音楽音響研究会資料, 35(4) (2016)	今村秀隆, 丸井淳史, 亀川徹
スピーカーを使用した小空間での音楽聴取における明瞭さの室内音響指標	日本音響学会講演論文集 (2015)	今村秀隆, 丸井淳史, 亀川徹
音像幅の知覚に関する聴能形成プログラムの提案	4th Audio Engineering Society Japan Conference (2015)	今村秀隆, キムソンヨン
小空間における一次反射音の到来方向の違いが音楽聴取時の明瞭さに与える影響	音楽音響研究会資料, 33(1) (2014)	今村秀隆, 丸井淳史, 亀川徹
Discovering Perceptual Characteristics of Multichannel Music with Height Ambiences for Japanese Listeners	13th International Conference on Music Perception and Cognition (2014)	H.Imamura, S.Kim, A.Marui, T.Kamekawa, R.King, W.Woszczyk
リスニングルームにおける一次反射音が音楽聴取の主観印象に与える影響	音楽音響研究会資料, 32(5) (2013)	今村秀隆, 丸井淳史, 亀川徹

参考文献

- Y. Ando (1985) *Concert Hall Acoustics*: Springer.
- M. Barron (1993) *Auditorium Acoustics and Architectural Design*: Spon Press.
- M. Barron (1988) “Subjective study of British symphony concert halls,” *Acta Acustica united with Acustica*, Vol. 66, No. 1, pp. 1–14.
- D. R. Begault and M. R. Anderson (2001) “Early Reflection Thresholds for Virtual Sound Sources,” in *International Workshop on Spatial Media, Aizu-Wakamatsu, Japan*, No. September.
- L. L. Beranek (1692) *Music, Acoustics, and Architecture*: John Wiley & Sons Inc.
- L. L. Beranek (2004) *Concert Halls and Opera Houses: Music, Acoustics, and Architecture*: Springer Science & Business Media.
- L. L. Beranek and T. J. Schultz (1965) “Some recent experiences in the design and testing of concert halls with suspended panel arrays,” *Acta Acustica united with Acustica*, Vol. 15, No. 1, pp. 307–316.
- J. Berg and F. Rumsey (1999) “Identification of perceived spatial attributes of recordings by repertory grid technique and other methods,” in *Audio Engineering Society Convention 106*.
- J. Berg and F. Rumsey (2006) “Identification of Quality Attributes of Spatial Audio by Repertory Grid Technique,” *Journal of the Audio Engineering Society*, Vol. 54, No. 5, pp. 365–379.
- J. Blauert (1996) *Spatial Hearing: The Psychophysics of Human Sound Localization*: The MIT Press.
- J. I. Bömer, S. Oode, and A. Ando (2011) “Effect of frequency bandwidth on interaural cross-correlation in relation to sound image width of reproduced sounds of a violin,” *Applied Acoustics*, Vol. 72, No. 9, pp. 623–631.
- P. D’Antonio and J. H. Konnert (1984) “The RFZ/RPG Approach to control room monitoring,” in *Audio Engineering Society Convention 76*.
- A. Farina (2007) “Advancements in impulse response measurements by sine sweeps,” *Audio Engineering Society Convention 122*.
- B. N. Gover, J. G. Ryan, and M. R. Stinson (2004) “Measurements of directional properties of

- reverberant sound fields in rooms using a spherical microphone array,” *The Journal of the Acoustical Society of America*, Vol. 116, No. 4, pp. 2138–2148.
- D. Griesinger (1995a) “Further investigation into the loudness of running reverberation,” in *AES 17th Proceedings of the Institute of Acoustics Conference*.
- D. Griesinger (1995b) “How loud is my reverberation,” in *Audio Engineering Society Convention 98*.
- D. Griesinger (2000) “The Theory and Practice of Perceptual Modeling - How to use Electronic Reverberation to Add Depth and Envelopment Without Reducing Clarity,” in *Preprint from the Nov. 2000 Tonmeister Conference*.
- D. Griesinger (2009) “The importance of the direct to reverberant ratio in the perception of distance, localization, clarity, and envelopment,” in *Audio Engineering Society Convention 126*.
- C. Guastavino and B. F. G. Katz (2004) “Perceptual evaluation of multi-dimensional spatial audio reproduction.,” *The Journal of the Acoustical Society of America*, Vol. 116, No. 2, pp. 1105–1115.
- H. Haas (1951) “Über den Einfluss eines Einfachechos auf die Hörsamkeit von Sprache,” *Acta Acustica united with Acustica*, Vol. 1, pp. 49–58.
- V. Hansen and G. Munch (1991) “Making Recordings for Simulation Tests in the Archimedes Project,” *Journal of the Audio Engineering Society*, Vol. 39, No. 10, pp. 768–774.
- T. Hanyu and S. Kimura (1996) “Influence of acoustic envelopment on spatial impression,” *The Journal of the Acoustical Society of America*, Vol. 100, No. 4Supplement, pp. 2802-2802.
- R. Höhne and G. Schroth (1995) “Zur Wahrnehmbarkeit von Deutlichkeits- und Durchsichtigkeitsunterschieden in Zuhörersälen,” *Acta Acustica united with Acustica*, Vol. 81, No. 4, pp. 309–319.
- IEC 60268 (1997) “Sound System Equipment – Part 13: Listening Tests on Loudspeakers,” International Electrotechnical Commission.
- H. Imamura, A. Marui, T. Kamekawa, and M. Nakahara (2014) “Influence of Directional Differences of First Reflections in Small Spaces on Perceived Clarity,” in *Audio Engineering Society Convention 136*.
- ISO3382-1 (2009) “Acoustics - Measurement of room acoustic parameters - Part1 : Performance spaces,” International Organization for Standardization.
- ISO3382-2 (2008) “Acoustics - Measurement of room acoustic parameters - Part 2: Reverberation time in ordinary rooms,” International Organization for Standardization.
- ITU-R (2015) “BS.1116-3: Methods for the Subjective Assessment of Small Impairments in Audio Systems Including Multichannel Sound Systems.”

- ITU-R (2017) “BS.2051-1 Advanced sound system for programme production.”
- V. L. Jordan (1981) “A group of objective acoustical criteria for concert halls,” *Applied Acoustics*, Vol. 14, No. 4, pp. 253–266.
- E. Kahle and M. Bruneau (1995) “Validation of an Objective Model of the Perception of Room Acoustical Quality in an Ensemble of Concert Halls and Operas,” Ph.D. dissertation, IRCAM.
- G. A. Kelly (1992) *The psychology of personal constructs. Volume 1: A theory and personality.*
- S. Kim (2009) “Recording piano in surround : discovering preferences , investigating auditory imagery , and establishing physical predictors .,” Ph.D. dissertation.
- T. Letowski (1989) “Sound Quality Assessment: Concepts and Criteria,” in *Audio Engineering Society Convention 87*.
- L. G. Marshall (1994) “An acoustics measurement program for evaluating auditoriums based on the early/late sound energy ratio,” *The Journal of the Acoustical Society of America*, Vol. 96, No. 4.
- A. Miśkiewicz, T. Rogala, T. Rościszewska, T. Rudzki, and T. Fidecki (2012) “Concert Hall Sound Clarity: A Comparison of Auditory Judgments and Objective Measures,” *Archives of Acoustics*, Vol. 37, No. 1, pp. 41–46.
- M. Morimoto, H. Sato, and M. Kobayashi (2004) “Listening difficulty as a subjective measure for evaluation of speech transmission performance in public spaces,” *The Journal of the Acoustical Society of America*, Vol. 116, No. 1607.
- R. W. Muncey, A. F. B. Nickson, and P. Dubout (1953) “The acceptability of speech and music with a single artificial echo,” *Acta Acustica united with Acustica*, Vol. 3, No. 3, pp. 168–173.
- T. Nakayama (1967) “Rotation and Interpretation of the Reference Axes of the Multidimensional Scales for Sound Quality Evaluation,” in *Acoustical Society of Japan*.
- T. Nakayama, T. Miura, O. Kosaka, M. Okamoto, and T. Shiga (1971) “Subjective Assessment of Multichannel Reproduction,” *Journal of the Audio Engineering Society*, Vol. 19, No. 9, pp. 744–751.
- P. Newell (2003) *Recording Studio Design*: Focal Press.
- S. E. Olive and W. L. Martens (2007) “Interaction between Loudspeakers and Room Acoustics Influences Loudspeaker Preferences in Multichannel Audio Reproduction,” in *Audio Engineering Society Convention 123*.
- S. E. Olive and F. E. Toole (1989) “The Detection of Reflections in Typical Rooms,” *Journal of the Audio Engineering Society*, Vol. 37, No. 7/8, pp. 539–553.
- J. Pätynen and T. Lokki (2016) “Concert halls with strong and lateral sound increase the emotional impact of orchestra music,” *The Journal of the Acoustical Society of America*,

- Vol. 139, No. 3, pp. 1214–1224.
- R. Pellegrini (2002) “A Virtual Reference Listening Room as an Application of Auditory Virtual Environments,” Ph.D. dissertation, Ruhr University of Bochum.
- W. Reichardt, O. A. Alim, and W. Schmidt (1975) “Definition und Messgrundlage eines objektiven Masses zur Ermittlung der Grenze zwischen brauchbarer und unbrauchbarer Durchsichtigkeit bei Musikdarbietung,” *Acta Acustica united with Acustica*, Vol. 32, pp. 126–137.
- F. Rumsey (2008) “Loudspeakers, Reflections, and Rooms,” *Journal of the Audio Engineering Society*, Vol. 56, No. 5, pp. 394–400.
- W. C. Sabine (1900) *Reverberation: The American Architect*.
- S. Tervo, P. Laukkanen, J. Pätynen, and T. Lokki (2014) “Preferences of critical listening environments among sound engineers,” *Journal of the Audio Engineering Society*, Vol. 62, No. 5, pp. 300–314.
- G. Theile (2001) “Multichannel Natural Music Recording Based on Psychoacoustic Principles / 1,” in *AES 19th International Conference*.
- R. Thiele (1953) “Richtungsverteilung und Zeitfolge der Schallrückwürfe in Räumen,” *Acta Acustica united with Acustica*, Vol. 3, No. 12, pp. 291–302.
- F. E. Toole (2006) “Loudspeakers and rooms for sound reproduction - A scientific review,” *Journal of the Audio Engineering Society*, Vol. 54, No. 6, pp. 451–476.
- F. E. Toole (2009) *Sound Reproduction*: Focal Press.
- 今村秀隆, 丸井淳史, 亀川徹 (2016) 「スピーカーを使用した音楽聴取における明瞭さの評価要因」, 音楽音響研究会資料, 第 35 巻, 第 4 号.
- 入交秀雄 (2013) 「無響室録音音源と電子残響を用いた、音源信号の特徴量と、残響の最適ミキシングレベルの関係の考察」, 博士論文.
- 尾本章, 中原雅考, 小野雅子, 藤原恭司 (2001) 「レコーディングスタジオにおける音響物理指標に関する基礎研究 -レコーディングエンジニアに対するアンケート調査-」, 日本音響学会研究発表会講演論文集, 767–768 頁.
- 亀川徹, 丸井淳史 (2016) 「空間音響再生方式と音楽的文脈が距離感・奥ゆき感に与える影響について」, 日本音響学会誌, 第 72 巻, 第 11 号, 684–695 頁.
- 岸智也, 小島健二, 中原雅考, 羽入敏樹, 星和磨 (2012) 「テレビゲームにおけるサウンドエフェクト: インタラクティブリバーブの開発 (<小特集>音楽制作を彩る音づくりの技術”エフェクタ”）」, 日本音響学会誌, 第 68 巻, 第 7 号, 362–368 頁.
- 戸井田義徳 (1995) 「空間内における音声情報伝達 (<小特集>音声の明瞭度と認識率)」, 日本音響学会誌, 第 51 巻, 第 4 号, 312–316 頁.
- 中原雅考, 佐竹康, George Massenburg, Peter D’Antonio (2009) 「スタジオの音響設計の現状 -サラウンドスタジオの設計例と解説-」, 日本音響学会誌, 第 65 巻, 第 2 号, 93–100 頁.

難波精一郎, 桑野園子 (1998) 『音の評価のための心理学的測定法』, コロナ社.

羽入敏樹 (2002) 「仮想音源分布を用いた音場の可視化・可聴化」, 音響技術, 第 31 巻, 第 1 号, 27-32 頁.

羽入敏樹, 星和磨 (2008) 「室形状の複雑さと平均自由行程および反射音密度の関係」, 日本音響学会研究発表会講演論文集, 1199-1200 頁.

森本政之 (1990) 「コンサートホールの音場評価について」, 音響学会建築音響研資.

渡辺洋 (2002) 『心理統計の技法』, 福村出版.

付録 A

インパルス応答の生成に使用したソースコード

```
1 import numpy as np
2 import soundfile as sf
3
4 # 室内音響パラメータの指定
5 c = 344 # 音速 [m/s]
6 V = 60 # 室容積 [m3]
7 S = 94 # 室総表面積 [m2]
8 A = 0.4 # 平均吸音率
9 r_D = 3.0 # 受音点から音源までの距離 [m]
10 W0 = 1 # 音源の出力
11
12 # パラメータの算出
13 RT = (0.161 * V) / (S * (-np.log(1 - A))) # 残響時間
14 h = (S ** 3) / (64 * (V ** 2)) # 室の複雑さ係数
15 N = int((h * (c ** 3) * (RT ** 3)) / V) # 総仮想音源数
16 W = (4 * np.pi * W0) / (3 * h) # 仮想音源の出力
17
18 # 仮想音源の受音点からの距離ベクトルを乱数で生成
19 vec_r = np.random.rand(N) ** (1 / 3) * (c * RT)
20
21 # 初期反射音部 (30ms~80ms) の距離ベクトル、偏角ベクトルの生成 (左右から
    到来する音場の例)
22 vec_r_ER = vec_r[(vec_r / c > 0.03) & (vec_r / c < 0.08)] # 距離ベクトル
23 N_ER = len(vec_r_ER) # 仮想音源数
24 vec_theta_ER = np.arcsin(np.random.rand(len(vec_r_ER)) * 2 - 1) # 偏角
    θ -1~1
```

```
25 vec_phi_ER = np.r_[
26     (7 / 8) * np.pi + 0.25 * np.pi * np.random.rand(N_ER)[:N_ER // 2],
27     # 左側偏角  $\phi$  [rad]
28     (-1 / 8) * np.pi + 0.25 * np.pi * np.random.rand(N_ER)[N_ER // 2:]
29     # 右側偏角  $\phi$  [rad]
30 ]
31 # 後期残響音部 (80ms $\sim$ ) の距離ベクトル、偏角ベクトルの生成
32 vec_r_LR = vec_r[vec_r / c > 0.08] # 距離ベクトル
33 N_LR = len(vec_r_LR) # 後期残響音の仮想音源数
34 vec_theta_LR = np.arcsin(np.random.rand(len(vec_r_LR)) * 2 - 1) # 偏角
35      $\theta$ 
36 vec_phi_LR = np.random.rand(N_LR) * 2 * np.pi # 偏角  $\phi$ 
37 # 仮想球内における22チャンネルのスピーカークのxyz座標の指定
38 xyz_speakers = [
39     [-0.707, 0.707, 0.0], # FL
40     [-0.383, 0.924, 0.0], # FLc
41     [0.0, 1.0, 0.0], # FC
42     [0.383, 0.924, 0.0], # FRc
43     [0.707, 0.707, 0.0], # FR
44     [-1.0, 0.0, 0.0], # SiL
45     [1.0, 0.0, 0.0], # SiR
46     [-0.707, -0.707, 0.0], # BL
47     [0.0, -1.0, 0.0], # BC
48     [0.707, -0.707, 0.0], # BR
49     [-0.612, 0.612, 0.5], # TpFL
50     [0.0, 0.866, 0.5], # TpFC
51     [0.612, 0.612, 0.5], # TpFR
52     [-0.866, 0.0, 0.5], # TpSiL
53     [0.0, 0.0, 1.0], # TpC
54     [0.866, 0.0, 0.5], # TpSiR
55     [-0.612, -0.612, 0.5], # TpBL
56     [-0.0, -0.866, 0.5], # TpBC
57     [0.612, -0.612, 0.5], # TpBR
58     [-0.612, 0.612, -0.5], # BpFL
59     [0.0, 0.866, -0.5], # BpFC
60     [0.612, 0.612, -0.5] # BpFR
61 ]
62 # 初期反射音部のxyz座標、遅延時間、音圧レベル、再生スピーカークの算出
63 ERs = []
64 for i in range(N_ER):
```

```

65     pos = [
66         vec_r_ER[i] * np.cos(vec_theta_ER[i]) * np.cos(vec_phi_ER[i]),
67             # x
68         vec_r_ER[i] * np.cos(vec_theta_ER[i]) * np.sin(vec_phi_ER[i]),
69             # y
70         vec_r_ER[i] * np.sin(vec_theta_ER[i]) # z
71     ]
72     t = vec_r_ER[i] / c # 遅延時間
73     power = np.sqrt(W / (4.0 * np.pi * ((c * t) ** 2)) * np.exp(t * np.
74         log(1.0 - A) * c * S / (4.0 * V))) # 音圧レベル
75     dist_LS = [np.sqrt((i[0] - pos[0]) ** 2 + (i[1] - pos[1]) ** 2 + (i
76         [2] - pos[2]) ** 2) for i in xyz_speakers] # 再生スピーカーまで
77         の距離
78     index_closest_LS = dist_LS.index(min(dist_LS))
79     ERs.append([pos[0], pos[1], pos[2], t, power, index_closest_LS])
80 ERs = np.array(ERs)
81
82 # 後期残響音部のxyz座標、遅延時間、音圧レベル、再生スピーカーの算出
83 LRs = []
84 for i in range(N_LR):
85     pos = [
86         vec_r_LR[i] * np.cos(vec_theta_LR[i]) * np.cos(vec_phi_LR[i]),
87             # x
88         vec_r_LR[i] * np.cos(vec_theta_LR[i]) * np.sin(vec_phi_LR[i]),
89             # y
90         vec_r_LR[i] * np.sin(vec_theta_LR[i]) # z
91     ]
92     t = vec_r_LR[i] / c # 遅延時間
93     power = np.sqrt(W / (4.0 * np.pi * ((c * t) ** 2)) * np.exp(t * np.
94         log(1.0 - A) * c * S / (4.0 * V))) # 音圧レベル
95     dist_LS = [np.sqrt((i[0] - pos[0]) ** 2 + (i[1] - pos[1]) ** 2 + (i
96         [2] - pos[2]) ** 2) for i in xyz_speakers] # 再生スピーカーまで
97         の距離
98     closestLS = dist_LS.index(min(dist_LS))
99     LRs.append([pos[0], pos[1], pos[2], t, power, closestLS])
100 LRs = np.array(LRs)
101
102 # 音源の読み込み
103 fs = 48000
104 lenIR = int(RT * fs) # 生成するインパルス応答の長さ
105
106 # 直接音のwavを生成

```

```
98 level_D = np.sqrt(W0 / (4.0 * np.pi * ((r_D) ** 2))) # 距離減衰を考慮
99 ir_D = np.append(level_D, np.zeros(lenIR - 1))
100 sf.write("ir_D.wav", np.transpose(ir_D), samplerate=fs)
101
102 # 初期反射音部のインパルス応答をスピーカーごとに生成
103 IRs_ER = []
104 for ind_speaker in range(len(xyz_speakers)):
105     t_ERs = ERs[np.where(ERs[:,5] == ind_speaker)]
106     out = np.zeros(lenIR)
107     if len(t_ERs) != 0:
108         for ER in t_ERs:
109             out[int(ER[3] * fs) - 1] = ER[4]
110     IRs_ER.append(out)
111 sf.write("ir_ER.wav", np.transpose(IRs_ER), samplerate=fs)
112
113 # 後期残響音部のインパルス応答をスピーカーごとに生成
114 IRs_LR = []
115 for ind_speaker in range(len(xyz_speakers)):
116     t_LRs = LRs[np.where(LRs[:,5] == ind_speaker)]
117     out = np.zeros(lenIR)
118     if len(t_LRs) != 0:
119         for LR in t_LRs:
120             out[int(LR[3] * fs) - 1] = LR[4]
121     IRs_LR.append(out)
122 sf.write("ir_LR.wav", np.transpose(IRs_LR), samplerate=fs)
```