

携帯電話ブース設計のための音響評価基準

関西大学 環境都市工学部 建築学科

建築環境工学第1研究室

建 12-0086 前田 健吾

指導教員 豊田 政弘 准教授

目次

第1章	はじめに	3
1.1	研究背景	3
1.2	既往研究	4
1.3	研究目的	6
第2章	問題の原因	7
2.1	携帯電話音声処理	7
2.2	狭小空間の音響特性	9
第3章	弁別閾について	11
第4章	実験方法	12
4.1	実験情報	12
4.2	実験用音声について	13
4.3	実験手順	14
第5章	実験結果	16
第6章	結果の分析と考察	20
第7章	おわりに	21
7.1	まとめ	21
7.2	今後の課題	22
	参考文献	23

第1章 はじめに

1.1 研究背景

携帯電話ブースというものがある。病院や空港、図書館や美術館などに広く設置され、プライバシーと静かな環境を確保することができ、その利用は現在広がり続けている。

しかしながら、通話のために作られた空間であるにも関わらず携帯電話ブースとして設計された吸音の少ない 5 m^3 程度の狭小空間で携帯電話を使用すると、通話相手の電話口から再生される音声の聞き取りにくさが著しく増大することが報告されている[1]。したがって、十分な会話了解性を確保するための携帯電話ブースの寸法や内装に関する設計指針が必要とされている。

1.2 既往研究

矢入ら[1]は、携帯電話ブースとして設けられた狭小空間で携帯電話を使用する際に、相手口から再生される音声の聞き取りにくさが著しく増大する問題について、原因の究明を行っている。

携帯電話ブースは一般に、床面積が1~2 m²、空間容積が2~5 m³の狭小空間で、扉のついた密閉型(写真 1)や半開放型(写真 2)などがある。



写真 1：密閉型携帯電話ブース



写真 2：半開放型携帯電話ブース

既報[1]によれば、狭小空間において携帯電話で通話する際の聞き取りにくさの増大は、残響成分による時間的なマスキングではなく、音声そのものの変質など直接的な要因があるとされている。言い換えれば反射音成分の中でも、残響感への寄与が比較的少ない部分であるインパルス応答の極めて初期の部分、すなわち、狭小空間に特有の「遅れ時間が短くパワの大きい反射音成分」が携帯電話の情報処理過程に何らかの影響をおよぼす可能性が示唆されている。

既報[1]では「STI（音声伝達指標：0～1の値をとりその数値が大きいほど聞き取り易いと判断する）」、「男性・女性」「大空間・狭小空間」「実音場・携帯電話伝送系経由」で場合分けされた音声から無作為に選んだ2つの刺激を並べ、十分に暗騒音の小さい条件の下、普段携帯電話で通話する際に使う側の耳からのみに音声を提示し、聴き取りにくさを評価させる聴感実験を行った。携帯電話伝送系を経由しない刺激を提示した場合は、残響感と聴き取りにくさに対応が見られるが、携帯電話を経由すると、残響感は大空間のほうが大きいにも関わらず聴き取りにくさは狭小空間の方が大きくなる傾向がみられた。さらに狭小空間において、携帯電話伝送系を経由した刺激を提示した場合、発話空間のSTIが高いにもかかわらず、聴き取りにくさが著しく増大することが分かっている。

このことを踏まえ、片桐ら[2]はどのような空間で発話を行った際にどの程度受話音声の聞き取りにくくなるかを明らかにするため、十分に暗騒音の小さい条件の下、様々な空間で発話を行った際の受話音声をシミュレートしたものと、発話空間の影響を取り除いた基準受話音声との比較から聴き取りにくさを評価させる聴感実験を行った。その結果を基に聴き取りにくさの予測を試みたが十分な精度を得ることはできなかった。その後、矢入ら[3]は、片桐の実験データに対し、発話空間のインパルス応答から減衰成分を除去した応答を用いてSTIを算出することで、携帯電話の聴き取りにくさを高い精度で予測できる可能性を示唆した。本研究においてもこの減衰除去応答から算出されたSTIを用いる。

また、減衰を除去したSTIを用いて鈴木ら[4]は携帯電話ブースの設計指標を構築するため「室容積」と「吸音率」で場合分けされた音声による聴き取りにくさと単語了解度を評価する聴感実験を行った。それぞれの空間条件に対する聴き取りにくさ、単語了解度を算出し、十分な精度で聴き取りにくさを予測することが可能になった。

1.3 研究目的

矢入ら[3]、及び、鈴木ら[4]の研究である程度設計指針は構築されたのだが、本研究ではさらに一步踏み込み、STI の値が高い場合 (0.60)、中くらいの場合 (0.42)、低い場合 (0.21) の3つの場合における弁別閾を求める。弁別閾を明らかにすることによって、携帯電話ブースの音響設計をする際や音響上問題のある携帯電話ブースを改修する際に有用な指針を構築することができるものと思われる。

第2章 問題の原因

本章では 1.2 節で述べた携帯電話ブースにおいて相手口から発せられる音声聴き取りにくくなる問題の原因とした、携帯電話の音声処理メカニズムと極小空間の音響特性について補足的に説明する。

2.1 携帯電話音声処理

携帯電話の情報処理は、音声生成機構の構造的なモデルの知識を利用する音声特有のものである。音源と調音の分離を前提とする方式、すなわち、生成源符号化方式であり、それをしない波形符号化方式と大別される。

一方、音声波形の生成機構は本来生理的な器官の運動であるが、物理的な音響現象としても理解可能である。その生成過程は音源の生成と調音に分けられる。有声音源は声帯振動に対応するパルス列に近似され、そのピッチ周期が声の高さに、周期の時間的な変化がアクセントやイントネーションに対応する。無声音源は空気の乱入に対応する白色雑音に近似され、その平均エネルギーが強さに対応する。音源で発生した音波は舌、唇、顎などの調音器官で形成される音響的な管（声道）の共鳴器官によってフォルマントが形成される。これが調音である。

携帯電話の音声情報処理の概略を **Fig. 1** に示す。送信側のマイクロフォンに入力された音声は、線形予測分析によって、その標本値間に内在する相関が除去される。ここで固定されるシステム関数は全極型であり、その極がフォルマントに対応する。したがってこれらが除去された後の残差には、有声音部では声帯パルスが無声音部ではランダムノイズが含まれており、それらを自己相関処理して得られるピーク周期から、有声か無声かの判断、声帯パルス音源のピッチ周期やパワーの推定が行われる。

線形予測分析の逆すなわち残差を入力として標本値間に位関関係を付与していく過程が音声合成であり、合成された信号が受信側の携帯電話から音声となって出力される。

予測と合成は完全な逆操作であるから、分解して得られた残差をそのまま入力して与えれば源波形を忠実に再現できる。しかしそれでは情報圧縮にならないため、様々な簡略化が行われるが、最も情報圧縮に貢献するのは残差の処理過程にある。声帯パルスが除去された残差には話し手の個性が含まれており、これを合成に使用しない限り本人の声には戻らない。現在の携帯電話通信では様々なパターンが事前に登録されたコードブックから、その話し手の残差に近いものを選択する手法がとられており、そのようにして合成された音声は、十分に本人の声として認識できるほど高品位である。音声符号化の原理自体は古くから確立されてきた理論ではあるが、近年の携帯電話通信の発展を支える音声情報処理の技術躍進には、このコードブックの発展も一役を担っている。

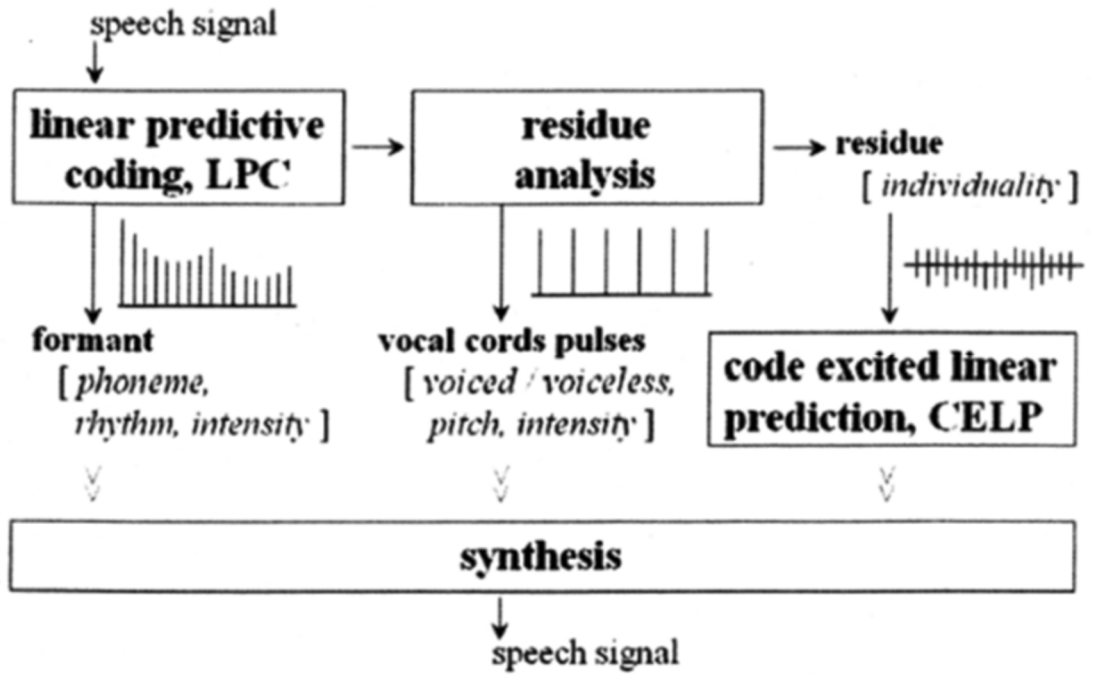


Fig. 1 携帯電話の音声処理の概略

2.2 狭小空間の音響特性

遅れ時間が短くパワの大きい反射音成分が影響をおよぼす可能性のある音声情報処理のステージは、**Fig. 1** に示した「線形予測分析」と「残差の処理過程」の 2 箇所であると考えられる。線形予測分析の時間窓は、一般的に 10~20ms であり、その中に入り込んだ反射音は、全極型システム関数の予測係数に影響を及ぼすため、フォルマントの推定に支障を与える可能性がある。これは、本来の有声音が別の有声音として誤認識される可能性があることを意味する。

また、この音声予測分析は、直接観測することのできない入力（音源）を一個のパルスまたは白色雑音と仮定してシステムを同定している。したがって、一個のパルス入力による出力の範囲では、残差は小さくランダムな予測誤差であるが、新しいパルスが加わった直後には、予測は大幅に狂い振れ幅の大きな残差を生じる。やがて系の出力がその応答と考えられる区間に入ると、残差は再び本来の予測誤差となり、振れ幅の小さなランダム波形となる。この変化を繰り返すことで、音声パルス列に同調したパルス列が抽出される。遅れ時間の短い反射音はこのピッチ周期にも影響を及ぼすと考えられる。推定した声帯パルス列のピッチ周期に誤りがあれば、声の高さや抑揚が変化してしまうだけでなく、本来無声音と判断すべきところに顕著なパルス列が観測されれば、無声音を有声音と誤認識する可能性も考えられる。

以上のように、実空間ではカラーレーションを生じるに過ぎない遅れ時間の短い反射音も、携帯電話を経由する音場では、著しい音声明瞭度低下に繋がる可能性がある。**Fig. 2** は、インパルス応答の初期部分を拡大して、狭小空間と大空間を比較したものである。それぞれ、STI が概ね同一の値を示すインパルス応答であるが、狭小空間のインパルス応答には、大空間のそれに比べ、直接音からの遅れ時間が非常に短い位置に顕著な反射音成分が観測される。

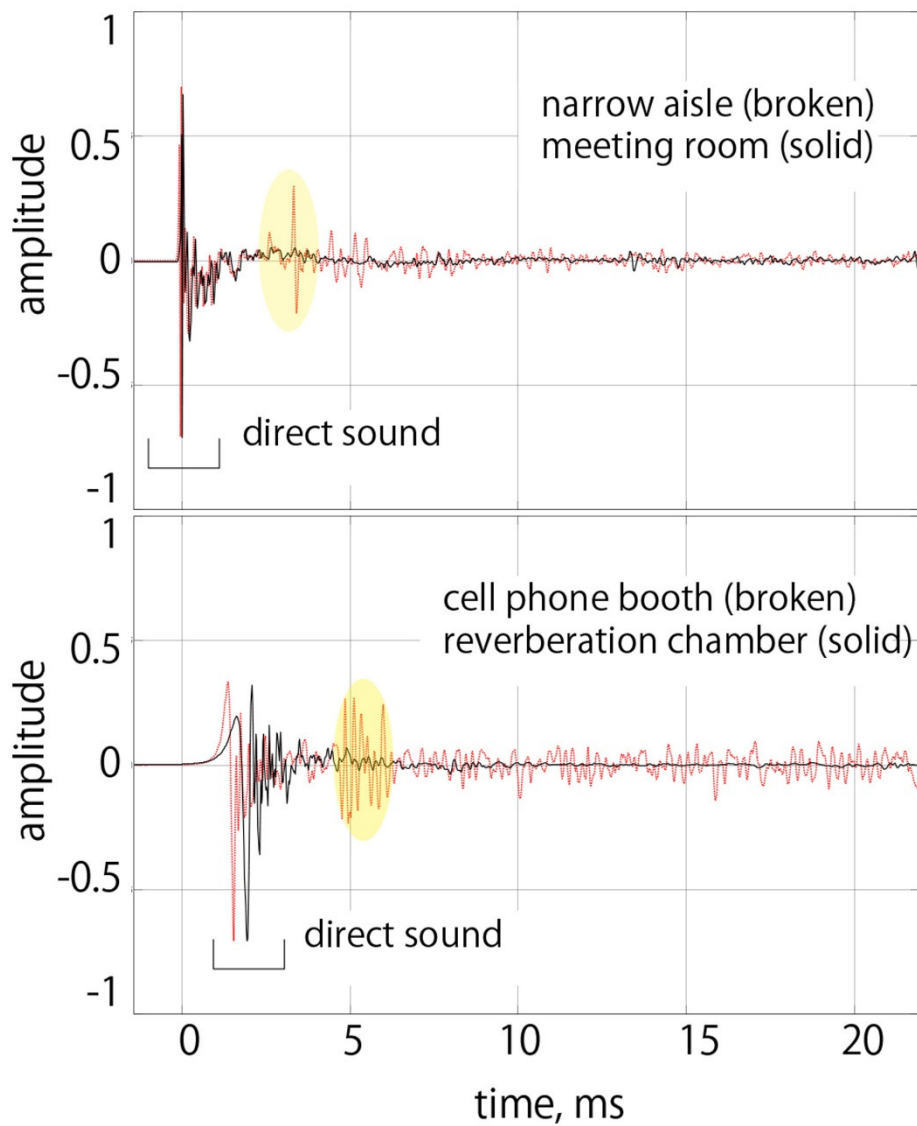


Fig.2 狭小空間と大空間におけるインパルス応答の比較図

第3章 弁別閾について

本研究で測定する弁別閾とは、基準となる刺激のある特定の要素を変えた時に変化したことがわかる最小の変化量のことである。人間の知覚と刺激の対応は必ずしも単純な関数関係にはならない。直前の、あるいは、これまでの体験や経験を基に刺激を感じ取っているからである。そのため実験の際は標準となる刺激と比較となる刺激を用意し、一対比較する。その際に求められた弁別閾は特定の標準刺激における弁別閾となる。

弁別閾のような精神物理を測定する代表的な方法として調整法、極限法、恒常法などがある。

調整法とは実験者または被験者自身が刺激を自由に变化させながら主観的等価（PSE : Point of Subjective Equality）などを主に測定する方法である。例えるなら、この調整法は日常生活においてラジオの音量を調節するようなやり方と酷似している。そのため実験に不慣れな被験者にも理解され易いメリットはあるが刺激変化の操作が被験者に一任されているので、操作の一義性、結果の再現性を欠いてしまう。また、被験者の意図的修正を受けてしまう可能性もある。

極限法とは実験者が刺激を一定の間隔で一方向に変化させ、各間隔で被験者の判断を求める方法である。適用範囲が広く実験の所要時間も短い。刺激の操作が実験者によって定められているので、一義性は保たれる。しかし刺激の変化方向が被験者に知られているので被験者の意図的修正を受けてしまう場合がある。

恒常法とは実験者が刺激の間隔を設定しランダムに提示する方法である。被験者の慣れや期待による誤差を取り除くことができ、一義性も保たれるが測定に時間を要する。現在最も正確で、かつ、最も選択範囲が広いとされている。本実験ではこの恒常法を用いる。

第4章 実験方法

携帯電話ブースの聴き取りにくさにおける STI の弁別閾を求めるため、恒常法を用いた一対比較によって刺激を評価させる聴感実験を行う。

4.1 実験情報

被験者	20 名
設問数	120 問
実験時間	54 分 (一問あたり 27 秒)
実験場所	無響室
音声信号	女声朗読 (夏目漱石「坊ちゃん」冒頭部)
ヘッドホン	SONY MDR-7506
出力	MacBook Air

4.2 実験用音声について

携帯電話ブースにおいて発話され、相手の電話口から発せられる音を、数値解析シミュレーションを用いて得たインパルス応答から、STI を算出するとともに、女声朗読（夏目漱石「坊ちゃん」の冒頭部分）のドライソースに畳み込み、実験音声を作成する。

STI の大きさによって弁別閾は異なる可能性があるため、STI が高い値を取る場合（0.60）、中くらいの値を取る場合（0.42）、低い値を取る場合（0.21）の3つを基準刺激とした。比較刺激はそれぞれ10個ずつとし、二者択一の対比較をした際の結果として0%から100%までを網羅できるように実験者の主観で選定した。また、実験で得られる評価の正確さを確かめるために、基準刺激を比較刺激に含めた。対比較であるため、実験が正確に行われれば、割合は50%に近い値をとるはずである。1つの問題につき、基準刺激と比較刺激の2回音声が表示される。音声は約9秒程度であり、問題番号の宣言と1回目の刺激との間、及び、1回目と2回目の刺激の間には0.5秒のインターバルを置き、問題と問題の間には5秒間の回答時間を設けた（Fig.3）。

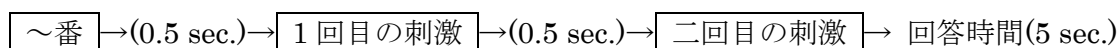


Fig.3 音声の流れ

比較刺激は10個ずつであり、基準刺激と比較刺激の順序を逆にしたものを用意し、それらを2回ずつ1回の実験で提示する。1つの基準刺激につき40個の評価を得ることになり被験者は、120問を回答する。被験者は20名であるから、1つの比較刺激につき80サンプルの回答を得ることとなる。

4.3 実験手順

被験者は正常な聴力を持っている男女 20 名とする。実験は前半と後半に分け、前半と後半の間には休憩時間を設けた。一人当たりの実験時間は全体で約一時間である。はじめに、教示文[Fig. 4]を読ませ、実験内容を理解させた後、ヘッドホン (SONY MDR-7506) を装着させた。十分に暗騒音の低い空間 (無響室) において、被験者が普段携帯電話で通話する耳の側からのみ刺激を提示し前後でどちらが聴き取りにくいかと問い、二者択一で回答させた。被験者には実際に評価してもらう聴感実験の前に、練習用として別途用意した一題を回答させた。再生音量は被験者ごとに調整させたが実験終了までは変更しないものとした。

音声の聞き取りにくさに関する評価実験

概要

この実験は携帯電話ブースを想定した評価実験です

設問数 120 問

実験時間 60 分 (25 分 休憩 10 分 25 分)

60 問終わったところで休憩を取ります

評価方法

音質を変えた夏目漱石の‘坊ちゃん’の冒頭部分が読まれます。

1 つの設問につき、音声は 2 回流され、前後で聞き取りにくかったほうの解答欄にチェックを入れてください。二回目の音声 flowed 後 5 秒間の解答時間があります。

なお回答が曖昧な場合があります。回答時間は五秒ですので直感で答えて頂いて構いません。

実験の流れ

～番 → 実験文 (前) → 実験文 (後) → 解答時間 (5 秒) → 次

あ [～番] → [実験文 (前)] → [実験文 (後)] → [解答時間 (5 秒)]

*注意事項

- ・ 実験は統計的に処理されるため個人の情報が表に出ることはありません
- ・ 途中で何らかの理由で実験を止めたいと思った場合には、中断したとしても何ら不利益を被ることはありませんので、遠慮なくお伝えください。
- ・ 実験中にお聞きいただく音に関しては一切質問にお答えできませんのでご了承ください。

Fig. 4 教示文

第5章 実験結果

Table 1~3 に基準刺激の STI の値が 0.21、0.42、0.60 の場合における結果を示す。ここで、割合とは、「基準刺激のほうが比較刺激より聴き取りにくいと評価した割合」を指すものとする。それらをもとに作成したグラフが **Fig. 5~7** である。本実験では弁別限の下限と上限を割合が 25% と 75% のときの近似直線の STI の値とする。したがって、グラフにおいて色で塗りつぶした範囲が弁別閾である。また R^2 は決定係数を表す。

基準 STI が 0.21 のとき

弁別限の下限の値は 0.178、上限の値は 0.4018 となり、弁別閾の値は 0.224 となった。

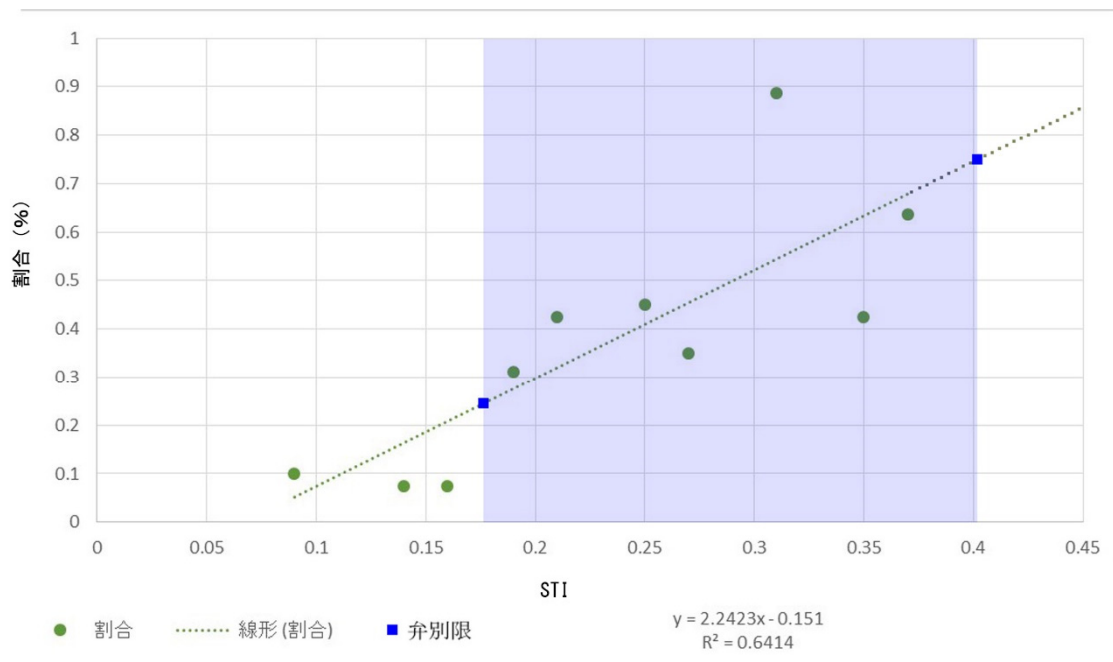


Fig. 5 STI 0.21 の場合の割合分布

比較音	割合
0.09	0.10
0.14	0.075
0.16	0.075
0.19	0.3125
0.21	0.425
0.25	0.45
0.27	0.35
0.31	0.8875
0.35	0.425
0.37	0.6375

Table 1 STI 0.21 の場合の割合算出結果

基準 STI が 0.42 のとき

弁別限の下限の値は 0.358、上限の値は 0.608 となり弁別閾の値は 0.250 となった。

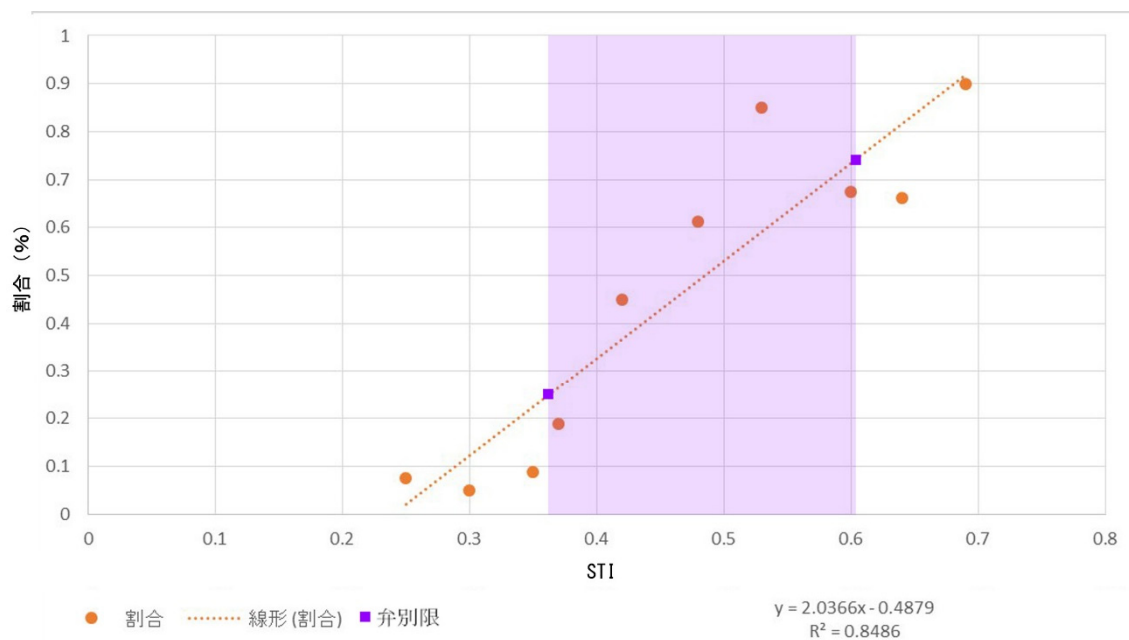


Fig. 6 STI0.42 の場合の割合分布

比較音	割合
0.25	0.075
0.30	0.05
0.35	0.0875
0.37	0.1875
0.42	0.45
0.48	0.6125
0.53	0.85
0.60	0.675
0.64	0.6625
0.69	0.90

Table 2 STI0.42 の場合の割合算出結果

基準 STI が 0.60 の時

弁別限の下限の値は 0.43、上限の値は 0.729 となり、弁別閾の値は 0.299 となった。

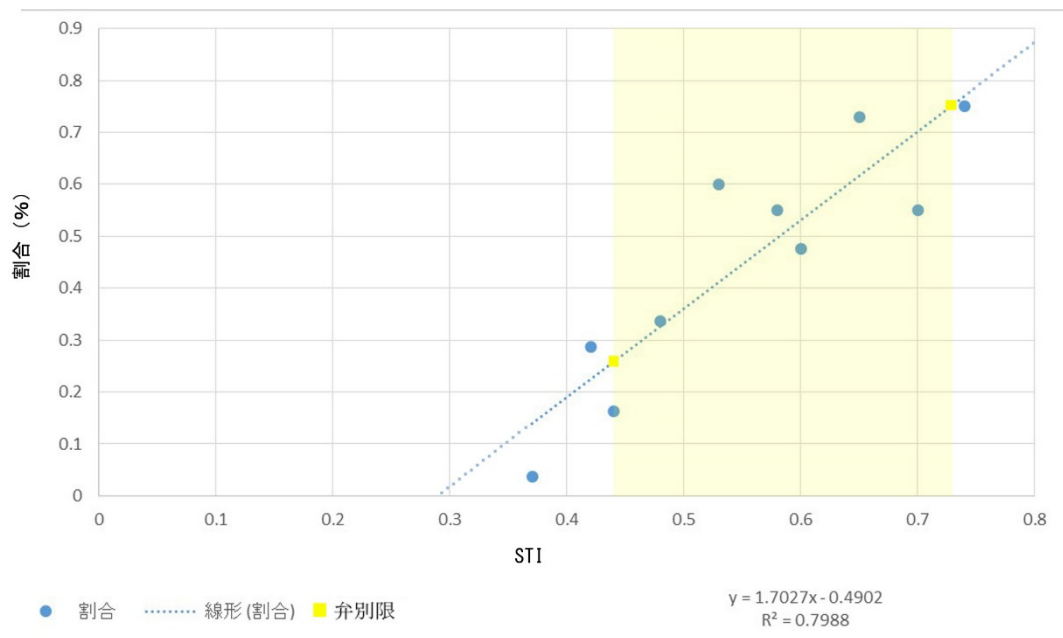


Fig. 7 STI0.60 の場合の割合分布

比較音	割合
0.37	0.0375
0.42	0.2875
0.44	0.1625
0.48	0.3375
0.53	0.60
0.58	0.55
0.60	0.475
0.65	0.7375
0.7	0.55
0.74	0.75

Table 3 STI 0.60 の場合の割合算出結果

第6章 結果の分析と考察

前章で示した結果より、基準とする STI によって弁別閾は変化しており、基準とする STI が高くなるにつれて弁別閾も $0.224 \rightarrow 0.250 \rightarrow 0.299$ と幅の広い値を取ることが分かる。人間の感覚として、大きな刺激を基準とした場合には細かな変化に気づきにくくなってしまふからだと推測できる。

決定係数が高い事、及び、同じ刺激での比較が 0.5 に近い事から、基準とする STI が 0.42 、 0.60 の場合については十分な精度で弁別閾を求めることができたと考えられる。しかし、基準とする STI が 0.21 の場合、割合の分布にばらつきが見られ、決定係数は十分高いとは言えず、同じ刺激での比較も 0.425 と 3 つの基準刺激の中では最低であるため、十分に信頼できる弁別閾を得ることはできなかったと判断する。

この原因のひとつは比較刺激の選定にあると考えられる。選定の際に想定した比較刺激の上限を低く見積もりすぎていた可能性がある。また、STI の低い刺激同士を比較する際は、判断が難しいものになるのではないかと推測される。

第7章 おわりに

7.1 まとめ

矢入ら[3]、及び、鈴木ら [4] の研究で十分な精度の設計チャートは完成している。しかし、問題とされる携帯電話ブースを改修する際や均一な音環境の携帯電話ブースを設計するためには、弁別閾が必須である。本研究はそういった携帯電話ブースを設計または改修する際の指針を得るために携帯電話ブースを想定した STI の弁別閾を測定した。本研究の結果、基準とする STI が 0.42、0.60 の場合については十分な精度で弁別閾を求めることができた。しかしながら、基準とする STI が 0.21 の場合については、十分な精度の弁別閾を得るには至らなかった。

7.2 今後の課題

基準とする STI が低い値を取る場合の再実験が必要だと考えられる。例えば、基準とする STI が 0.21 の場合であれば比較刺激の STI の上限を 0.37 から 0.45 程度に引き上げる必要があるものと思われる。

参考文献

- [1] 矢入幹記,「狭小空間の音響設計に向けて～携帯電話ブースにおける聞き取りにくさ～」, 日本音響学会建築音響研究委員会,建築音響研究委員会,建築音響研究会資料; AA2013-19 2013年5月30日
- [2] 片桐領太,“狭小空間での携帯電話による音声伝送性能の評価”,卒業論文(関西大学),2013
- [3] 矢入幹記,豊田政弘,「携帯電話音声の聞き取りにくさにおよぼす建築音響の影響 : 音声伝送性能評価指標による予測」,日本音響学会建築音響研究会資料; AA2014-36,東京,2014
- [4] 鈴木祐太,“狭小空間での携帯電話通話における聞き取りにくさと単語了解度の評価”卒業論文(関西大学),2014
- [5] 難波ほか,音の評価のための心理学的測定法,コロナ社,2001

謝辞

研究を進めていく中で、お忙しい中熱心にご指導いただきました豊田政弘先生には厚く御礼申し上げます。またご協力いただきました建築環境工学第一研究室の先輩や同期の皆様にも大変お世話になりました。心より御礼申し上げます。