

科学研究費助成事業 研究成果報告書

平成 27 年 6 月 26 日現在

機関番号：53601

研究種目：基盤研究(C) (一般)

研究期間：2012～2014

課題番号：24560727

研究課題名(和文) 話声の聴き取りやすさ評価・表示システムの開発

研究課題名(英文) Evaluation and display system for speech transmission performance.

研究代表者

西川 嘉雄(NISHIKAWA, YOSHIO)

長野工業高等専門学校・環境都市工学科・教授

研究者番号：20513807

交付決定額(研究期間全体)：(直接経費) 2,600,000円

研究成果の概要(和文)：音声の聴き取りやすさの指標として既存のSTI (Speech Transmission Index) に、実時間分析した話声の音圧レベルと暗騒音の音圧レベルを考慮した実時間STIの提案と実音場における主観評価実験方法の検討を行った。本研究により以下のことが確認できた。1)実時間STIはスピーチレベルと対応しており相関係数は0.984であった。表示システムのWebカメラの映像を見ることで実時間STIが低下する要因の確認ができた。2)実音場の主観評価実験方法の(聴き取りにくさ・聴き取り間違い)の検討により個人差を少なくできた。STIと実音場での主観評価の対応について示した。

研究成果の概要(英文)：Speech transmission performance is important acoustical functions of the classroom and lecture room. STI (Speech Transmission Index) is used performance evaluation of speech transmission. Real time STI (RT-STI) can be obtained from a real time analysis of speech and noise. Examine subjective evaluation (listening difficulty and intelligibility) in real sound field. The following presents the results
It became clear that there is a strong correlation between RT-STI and measurements speech level. By the proposal subjectivity evaluation experiment the difference among the individuals became smaller.

研究分野：工学 建築学 建築環境・設備

キーワード：音声伝送性能 STI 実時間STI スピーチレベル ノイズレベル 聴き取りやすさ 聴き取り間違い
主観評価

1. 研究開始当初の背景

学校教育では、高校大学ともに全入時代に入り、教育の質が重視されている。講義を受ける生徒や受講者などに講義者の話声が聞き取りやすい品質で伝わることは、講義室や教室が持つべき必要不可欠な性能である。話声が講義者から受講者に伝わる過程は、図-1のように音声伝達の系統としてまとめることが出来る。

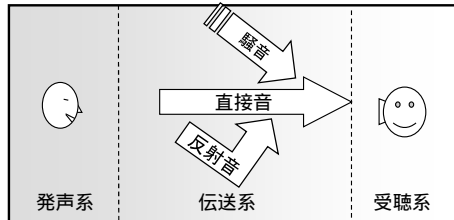


図-1 音声伝達の系統

教育機関は、話声の聞き取りやすさに対して、十分な性能を持った講義室・教室、拡声システム（伝送系）を整備し、聞き取りやすい話が出来る教員（発声系）を養成し、教育の質の向上に取り組むことが重要である。

講義や講演をする話者は、自分の話声が受講者に正確に伝わっているかの判断は、受講者を観察して行う。一方、受講者にとって話声が聞き取りにくい時に、話者に話し方の改善を要求するのは難しい。リアルタイムに話声の聞き取りやすさを話者に知らせることができれば、聞き取りにくくなった時に話者が自分自身で話し方などを改善することができる。また、受講者はつねに聞き取りやすい状況で話を聞くことができる。

日本建築学会から話声の聞き取りやすさの規準として「都市・建築空間における音声伝送性能評価規準・同解説」が2011年11月に出版された。この規準は、実音場のデータを多数収集したインパルス応答データベースのデータを用いて、無響室で行われた話声の聞き取りやすさに関する主観評価指標により決められた。実際の設計・測定では主観評価指標と対応の良い物理評価指標の STI (Speech Transmission Index) を用いて評価を行う。音声伝送性能のクラスは STI で 1st ~ 0.84, 2nd ~ 0.65, 3rd ~ 0.55, 4th ~ 0.45 が対応している。また、STI の聴覚的な閾値である JND (Just Noticeable Difference) は 0.03 である。

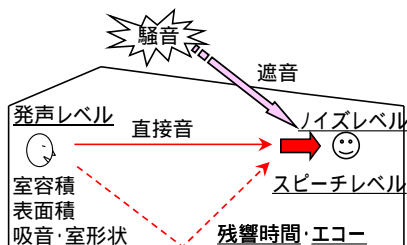


図-2 話声の聞き取りやすさの要因

STI は、放課後のような空席時にインパルス応答と音声の長時間スペクトル特性の疑似音声ノイズを測定したスピーチレベルと暗騒音により算出する。しかし、実際の講義では、図-2 に示す様な要因があり、話声をリアルタイムで評価するには、時々刻々と変化するスピーチレベルとノイズレベルを考慮する必要がある。

2. 研究の目的

リアルタイムで音声伝送性能を評価する実時間 STI を提案し、実音場での評価について検討を行う。実時間 STI の変動範囲やスピーチレベルとの対応などを明らかにする。また、スピーチとノイズの分離方法や表示システムによる音声伝送性能の低下要因についても確認する。

主観評価実験に用いる試験音について検討した後、実音場において精度よく主観評価指標を得る方法を検討する。さらに、実音場で得られた主観評価指標と STI の関係を確認する。

3. 研究の方法

本研究は(1)実時間 STI による評価と表示方法、(2)実音場における主観評価手法の検討、(3)実音場における主観評価指標と STI の比較の3項目について実施した。それぞれの項目の研究方法を以下に示す。

(1) 実時間 STI による評価と表示方法

実時間 STI の計測・表示システムの測定系統図を図-3 に示す。

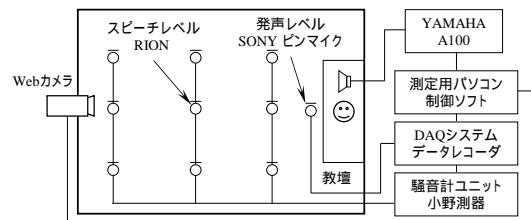


図-3 実時間 STI の測定系統図

実音場で計測した STI を使って、スピーチレベル及びノイズレベルにより STI がどの程度変化するかを確認する。また、実講義の話声を分析し、スピーチレベルとノイズレベルの状況について検討を行い、実時間 STI が提案可能か検討する。

講義室で実時間 STI を算出し、実時間 STI の時系列での変化を確認するとともに、スピーチレベルの関係を検討する。

話者のピンマイクの音圧レベルを使って、スピーチとノイズを分離する方法について検討を行う。

講義室で計測した実時間 STI の表示システムを作成し、実時間 STI の低下要因が確認できるか確認を行う。

(2) 実音場における主観評価手法の検討

講義室などの実音場において主観評価指標を精度よく得るための方法の検討を行う。

主観評価実験の測定系統図を図-4 に、測定状況を写真-1 に示す。試験音用単語の親密度と単語の音の大きさ（ラウドネス）と個人誤差を少なくするための実験方法の検討を行う。
 (3) 実音場における主観評価指標と STI 比較
 3 室において主観評価実験と STI を算出し音声伝送性能のクラスと比較検討を行う。

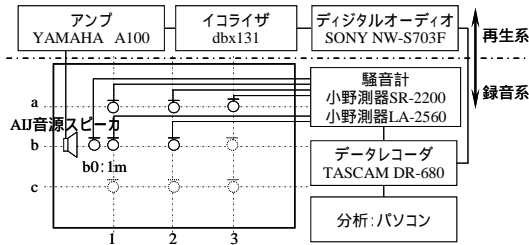


図-4 主観評価実験測定系統図



写真-1 主観評価実験風景

4. 研究成果

(1) 実時間 STI による評価と表示方法

STI に対するスピーチレベルとノイズレベルの影響

STI は、音声信号のパワー変動（包絡線情報）の保存性を表す MTF とスピーチレベルとノイズレベルにより算出される。スピーチレベルとノイズレベルを変化させることで STI の変化幅を確認し、実時間 STI の適用について検討を行う。

実測の STI のスピーチレベルとノイズレベルを個々に変化させた時の影響を確認する。スピーチレベルを -25 ~ 60dB でノイズレベルを -25 ~ 25dB の範囲で変化させたときの STI の変化を図-5 に示す。STI はノイズが減少するに従い少し大きくなり、ノイズが大きくなるにしたがって小さくなる。STI は、スピーチが小さくなるにしたがって小さくなり、スピーチが大きくなるにしたがって若干大きくなったのち次第に小さくなる。更に、ある程度以上大きくなると変化量は小さい。

SN 比で整理した結果を図-6 に示す。ノイズレベルを小さくし SN 比を大きくしても STI は変化しなくなる。スピーチレベルを大きくし SN 比を大きくすると、STI は算出過程の聴覚マスキングや閾値の影響で小さくなる。

実講義におけるスピーチレベルとノイズレベルの検討

一般的な講義室における 60 分間の講義を対象にスピーチとノイズに分離した。スピーチ

チレベルとノイズレベルの度数分布を図-7 に示す。ノイズは 40dB にスピーチは 63dB 付近にピークを持ちともにピークを持ち 35 ~ 80dB に分布している。ノイズとスピーチはそれぞれ異なった分布形状をしている。

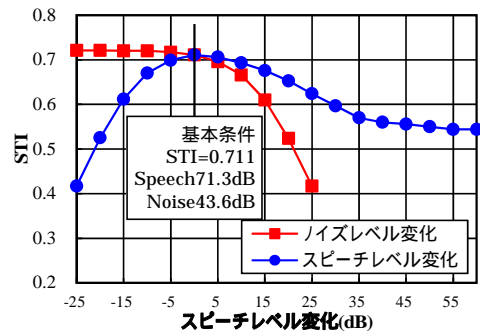


図-5 スピーチとノイズの個々に対する STI の変化

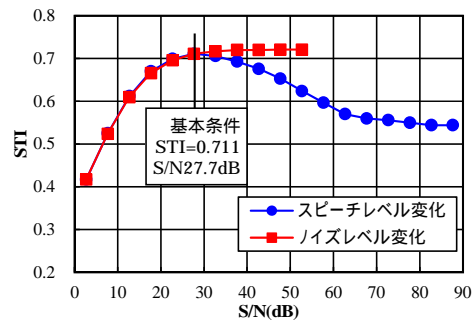


図-6 スピーチとノイズ変化における SN 比と STI の関係

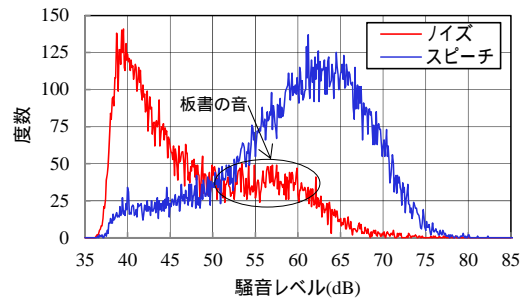


図-7 講義のスピーチレベルとノイズレベルの度数分布

スピーチレベルとノイズレベルの累積相対分布を図-8 に示す。スピーチが広く分布しているのに対して、ノイズは騒音レベルが低

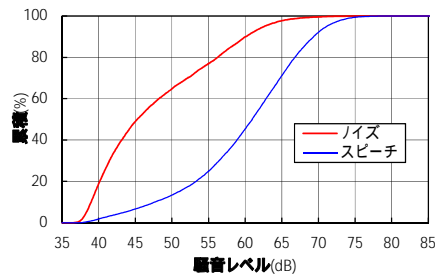


図-8 講義のスピーチレベルとノイズレベルの累積相対分布

い部分に多く分布していることが確認できる。SN は -30 ~ 30dB と推測される。講義室の講義は、SN 比が 0 ~ 30dB で変化する事から、STI は 0.3 程度小さくなると想定される。

講義室における実時間 STI

実際の講義室 (STI=0.788) の講義を対象に実時間 STI の 10 分間の分析を行った。図-9 に示す様に実時間 STI は 0.221 ~ 0.775 の間で変動している。想定より大きい幅で変化した。

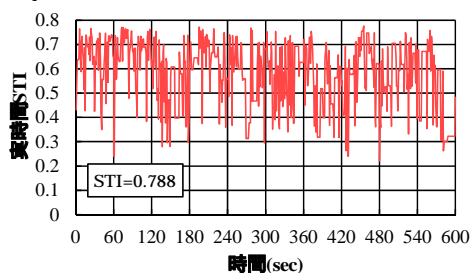


図-9 講義の実時間 STI

図-10 に、実時間 STI とスピーチレベルの関係を示す。実時間 STI とスピーチレベルの対応が良く、相関係数は 0.984 である。スピーチレベルが 70dB 以上になっても実時間 STI が変化しない。スピーチレベルが 60dB で実時間の STI は 0.48 ~ 0.75 に分布しており、ノイズ等の影響である。スピーチレベルとの対応が良く実音場の評価に対する有効性が確認できた。

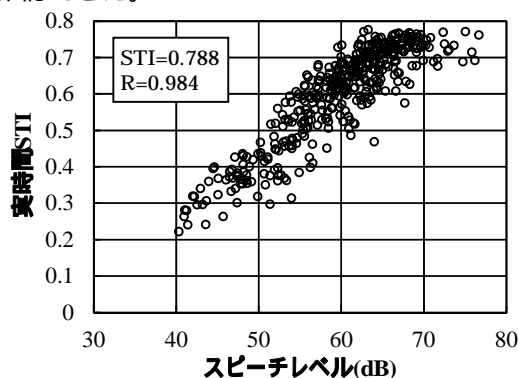


図-10 実時間 STI とスピーチレベルの関係

ノイズとスピーチの分離

受聴位置におけるスピーチレベルとノイズレベルが異なる分布をしていることから、更にスピーチが正確に確認できるピンマイクの音圧レベルを用いてスピーチとノイズを分離する。既存データの分析結果から L_s (69.5dB) 以上をスピーチとし 0.1 秒間の L_{eq} をスピーチレベルとする。ノイズは L_n (67.9dB) 以下とし実時間 STI を算出する直前までの音圧レベルの L_{50} をノイズレベルとした。スピーチと判定した範囲で 0.1 秒ごとに実時間 STI を算出した。結果の一部を図-11 に示す。この方法で算出した実時間 STI と聴感でスピーチとノイズを分離し算出した実時間 STI は $JND=0.03$ の範囲とすれば 72% 合致

していた。 L_s と L_n を適切に設定する事で分離の精度は向上する。

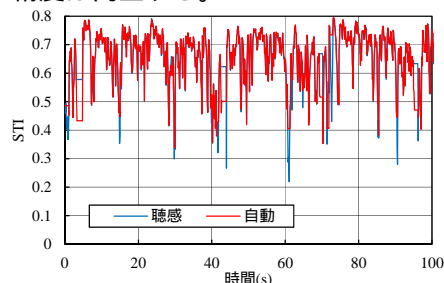


図-11 実時間 STI の算出方法の比較 (100 秒)

実時間 STI の表示

表示システムは、話者にピンマイクをつけ、講義室内は 9 点の無指向性マイクで実時間 STI を測定し、講義の風景を web カメラで録画する。図-12 に画面の表示例を示す。Web カメラは講義室全体と講義者を撮影する 2 種類とした。実時間 STI と建築学会のクラスに対応して画面の色を変え講義者に視覚的に知らせる。実時間 STI は時系列で変化を確認できる。

表示システムにより、実時間 STI の低下する要因を確認するため、実時間 STI が 0.65 を下回った部分に着目する。図-13 左は教科書やノートをめくるため音周辺騒音が大きくなり低下していた。図-13 右は講義者が、板書をしながら話し、発声レベルも下がったためにスピーチレベルが低下した例である。話者が下を向いたり、発生音が小さくったりすると実時間 STI は低下する。

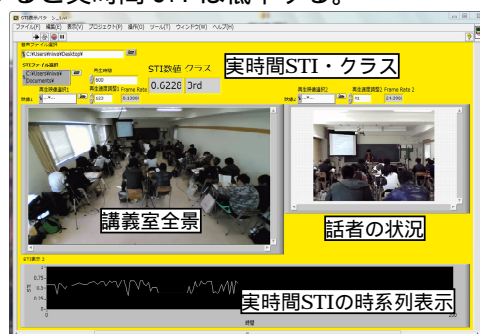


図-12 実時間 STI 表示システム



図-13 Web カメラの画像による実時間 STI 低下の要因検討

(2) 実音場における主観評価手法の検討

試験用単語の親密度

本科 1 年生 (16 歳) と 5 年生 (20 歳) の了解度試験の結果に差があった。聞き取れなかった単語を確認すると、5 年生が間違えていない単語を 1 年生が間違えていた (表-1)。低学年の学生には親密度の低い単語であった。年

年齢により親密度が異なるので注意が必要である。

表-1 年齢による了解度の誤答比較

	1年生(16歳)	5年生(20歳)
ヘダタリ	29%(11)	33%(2)
ハゲタカ	24%(9)	33%(2)
ハツマゴ	29%(11)	0%(0)
スミコミ	47%(18)	0%(0)

単語のラウドネス

被験者から聞き取りにくさの評価で単語の大きさ(ラウドネス)で判断していることが指摘された。主観評価実験単語の親密度とラウドネスの確認を行う実験をした。ヘッドホンで単語を再生し親密度とラウドネスを同時に評価した。単音節のラウドネス補正前後の単語を5段階で評価させた。図-14に結果を示す。3が標準であり、補正前後の評価値の分布から改善していることが確認できた。

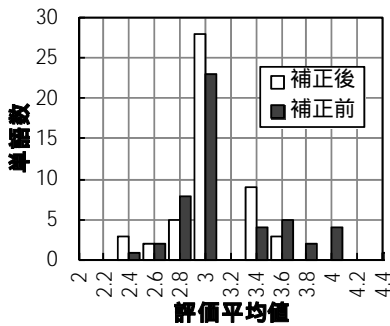


図-14 単語の音量評価 (3:標準)

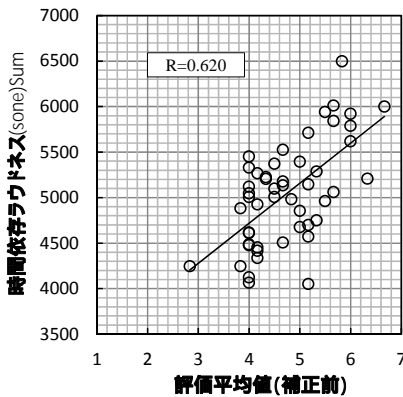


図-15 評価平均値と補正值の関係：時間依存ラウドネスによる補正

単音節のラウドネス補正では補正が不十分な単語があったため、時間依存ラウドネスによる補正をおこなった。単語の時間依存ラウドネスと評価の関係を図-15に示す。ラウドネス確認実験の聴感上のラウドネス(評価平均)と時間依存性ラウドネスに対応の関係があり、時間依存性ラウドネスが有効であることが確認できた。本研究の主観評価実験ではラウドネス補正した単語を使用する。

主観評価実験方法の検討

実験室で行う主観評価実験を実音場で、一般学生30名を対象に実施した。被験者からの意見とデータ分析の結果以下の問題点が指摘された。

- ・ 評価の判断基準がわからない。
- ・ カテゴリーの差がわからない。
- ・ 実験方法を十分理解していない。
- ・ 個人差が大きい。

改善前の評価実験の結果を図-16に示す。聞き取り間違いは、個々の回答が4%以下で平均は2%以下であった。聞き取りにくさは、測定位置の回答に差があった。これは、個人誤差の影響であった。

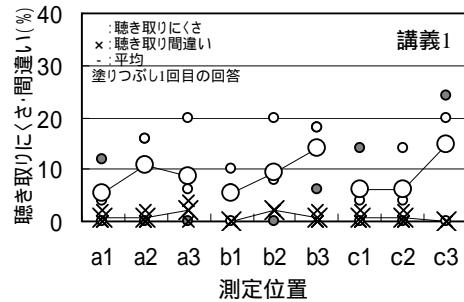


図-16 講義室における主観評価実験:改善前

上記の問題を解決するために以下の方法を取り入れた。

- ・ 実験方法と主旨を正確に理解させるため教示用紙を配布しゆっくり音読する。
- ・ 練習用音源に、響きを付加した非常に聞き取りにくい試験音を含める。
- ・ 実験は4回実施し被験者は前・中・後の席に移動する。
- ・ STI0.45以下の非常に聞き取りにくい試験音を含める。

図-17に改善後の実験結果を示す。b3で一度聞き取りにくさ20%を回答しているがそのほかは16%以下であった。また1-4回の平均値より2-4の平均値の方が聞き取りにくくない傾向があった。個人差を小さくする効果があった。

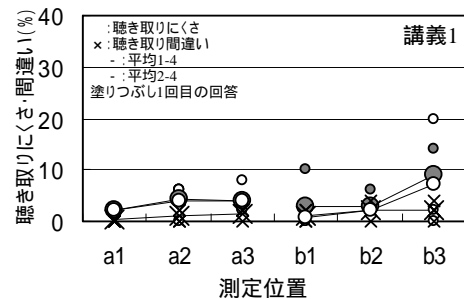


図-17 講義室における主観評価実験:改善後

(3) 実音場における主観評価指標とSTI比較
講義室など3室の主観評価性能とSTIと建築学会が提案するクラスと比較を図-18に示す。聞き取りにくさの音声伝送性能のクラス2ndは16%以下でSTIは0.60以上に分布して

いる。聞き取り間違い3rdの2%以下のSTIも0.60以上に分布している。実音場におけるSTIでは0.60以上を一つの目安とできる可能性がある。

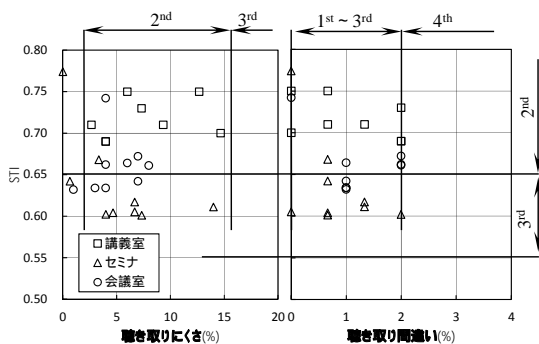


図-18 主観評価指標と STI と建築課学会の音声伝送のクラスの関係

図-19 に音声伝送性能のクラスと STI の関係に実測した主観評価指標と STI を測定点毎にプロットしたものを示す。聞き取り間違いは(破線)2%以下で対応が良い。聞き取りにくさは(実線)、講義室が概ね曲線の上にある。 세미나室と会議室は、STI では3rdの範囲にあるが、聞き取りにくさが16%を下回り良好な音声伝送性能である。

実時間STIで音声伝送性能の良否を判断するときは、建築学会のクラスの対応するSTIの数値より小さい値でも良い可能性がある。

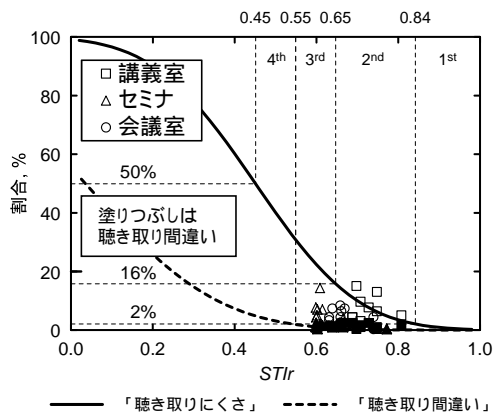


図-19 主観評価指標と STI と建築課学会の音声伝送のクラスの関係

(4) まとめ

本研究の結果以下が明らかとなった。

- 1) 講義室 (STI=0.788) における実時間 STI は 0.221 ~ 0.775 の範囲で変動した。
- 2) 実時間 STI はスピーチレベルと対応しており相関係数は 0.984 であった。
- 3) ピンマイクの音圧レベルでスピーチとノイズを分離すると聴感で分離した実時間の STI は JND0.03 の範囲で 72% 合致していた。
- 4) 表示システムは Web カメラの映像により実時間 STI が低下する要因について検討できる。

- 5) 年齢により単語の親密度が異なるので実験に使用する際は注意が必要である。
- 6) 聞き取りにくさの評価を単語のラウドネスで行う傾向が見られたため、実音場ではラウドネスをそろえた試験音を用いる。
- 7) 複数回の試験と、非常に聞き取りにくい試験音を含めることで、評価の個人差を少なくする効果があった。
- 8) 実音場の主観評価と STI の関係から実音場における聞き取りやすさ確保の目安として STI0.6 以上が必要となった。

5. 主な発表論文等

(研究代表者、研究分担者及び連携研究者には下線)

[学会発表](計6件)

西川嘉雄, 池部功一, 講義室における聞き取りにくさと聞き取り間違いの検討, 日本音響学会建築音響研究会, 2012年5月24日, キャンパスイノベーションセンター(東京)

西川嘉雄, 講義室における聞き取りにくさと了解度の検討, 日本建築学会北陸支部研究発表会, 2012年7月21日, 信州大学(長野市)

西川嘉雄, 丸山みず希, 講義室における音声伝送性能の主観評価実験方法の検討, 日本音響学会建築音響研究会, 2013年8月2日, 長野市生涯学習センター(長野市)

西川嘉雄, 実音場における音声伝送性能の主観評価実験方法の検討, 日本建築学会北陸支部研究発表会, 2014年7月13日, 富山大学(高岡市)

西川嘉雄, 教室における音声伝送性能評価, 日本建築学会第74回音シンポジウム, 2015年3月2日, 建築会館(東京都)

西川嘉雄, 実音場における音声伝送性能の検討 主観評価指標と STI の関係, 日本建築学会北陸支部発表会, 2015年7月12日, 長岡造形大学(長岡市)

6. 研究組織

(1) 研究代表者

西川 嘉雄 (NISHIKAWA YOSHIO)
長野工業高等専門学校・環境都市工学科・准教授
研究者番号: 201513807

(3) 連携研究者

橋本 典久 (HASHIMOTO NORIHISA)
八戸工業大学・工学系研究科・教授
研究者番号: 90265957