## 2-1-1

### 2-1-1 ガス管メンテナンス箇所推定システムの効率と精度向上に関する検討

Study on improvement of efficiency and accuracy of estimation system for gas pipe maintenance location.

☆堀野由輔, 黒沢琢登, 及川靖広（早大理工）,
土井敏行, 小野寺文也, 中里直人(東京ガスネットワーク)

■従来のガス管メンテナンス箇所推定システム
- □3軸情報を用いて中心位置を推定
- □駆動音源はピンポン音を使用

■提案
- □3軸コイルアレイを2軸コイルアレイを3個へ変更
- □駆動音源の検討、比較

■実験結果
- □2軸情報を用いた推定では3軸情報を用いた推定に比べ精度が向上(Fig. 2 (a))
- □ピンポン音、ホワイトノイズ、ピンクノイズを駆動音源とした場合、コイルの周波数特性(Fig. 1)の影響からホワイトノイズで推定誤差は最小(Fig. 2 (b))

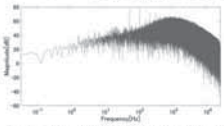


Fig. 1 Frequency response of receiving coil

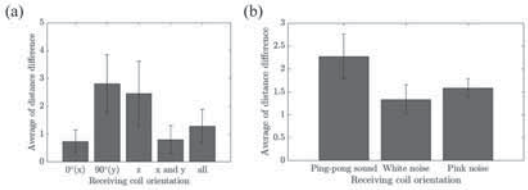



Fig. 2 Average of estimation error (a)axis (b)sound

## 2-1-2

### 2-1-2 Azure Kinect を用いて推定した発声者の指向特性に基づく音声補正の検討

A study of voice correction based on vocalist directional characteristics estimated using Azure Kinect.

☆倉本有紗, 及川靖広（早大理工）

◆これまで 2 台の Azure Kinect を用いて，音声収録時における発声者の頭部運動測定を行うシステムが構築された。

◆構築したシステムは発声者の音声と身体データを収録することができ，これを基に指向性の推定が行われた。

◆本研究では，推定した指向性を基に，身体を回転させた時の音声補正に取り組んだ。

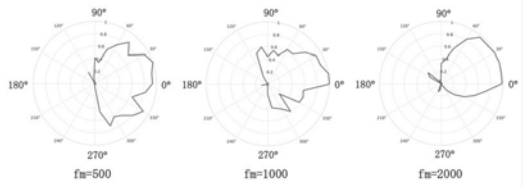



Fig.1 : The estimated directivity coefficient

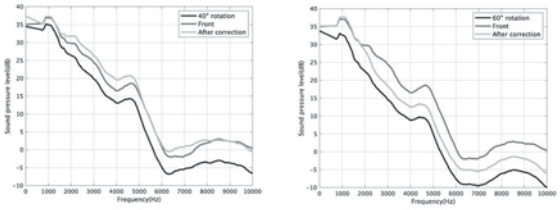


Fig.2 : Frequency response of corrected audio

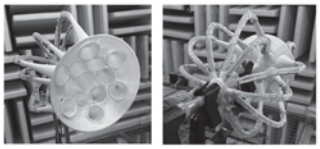
## 2-1-3

### 2-1-3 フレキシブルワイヤーによる焦点距離可変機構を備えた音響レンズ

Acoustic lens with control mechanism of focal length using flexible wires

☆村上龍斗（産技高専），高橋義典（工学院大），△高野邦彦（産技高専）

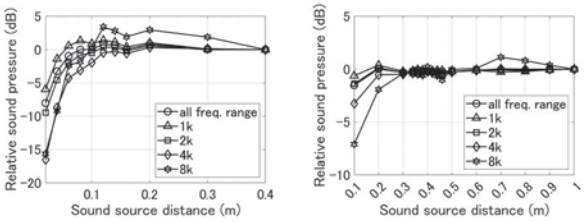
◆本研究では，フレキシブルワイヤーを用いて音響レンズを構成する複数の音響ホーンの向きを変える機構を提案することで，焦点距離が可変な音響レンズの検討をした(Fig.1).



(a) Front view　(b) Back view

Fig.1 Acoustic lens with control mechanism of focal length

◆音響ホーンの角度を変えることで，音響レンズの焦点距離を 0.1 m，0.2 m，0.3 m，0.4 m となるように調整し，距離選択性をそれぞれ測定した(Fig.2).




(a) Focal length of 0.1 m　(b) Focal length of 0.4 m

Fig.2 Observed sound pressure using acoustic lens normalized by the value of distance attenuation

## 2-1-4

### 2-1-4 環状障害物を用いた音響レンズによる平面波の屈折

Refraction of plane wave by acoustic lenses with annular obstacles

☆増川英寿，高橋義典（工学院大）

◆音響ホーン等から出力される平面波の拡散を目的として，2 種類の音響レンズの設計と FDTD シミュレーションを行った.

◆レンズに向かって平面波を発し，レンズの半径の円周上に 7 つの受音点を配置した(Fig. 1). レンズを配置した時と配置しなかった時の音圧と位相を比較した.

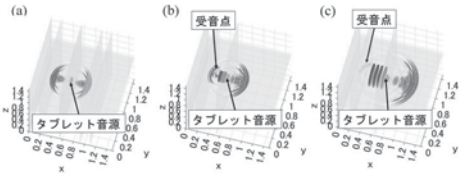



Fig. 1: Wavefront of plane wave (a), using acoustic lens of scattered sphere type (b), using acoustic lens of ring stack type (c).

◆パワ上昇では，ともに正面方向(0 度)より，周囲方向のパワが上昇していることが確認された(Fig. 2).

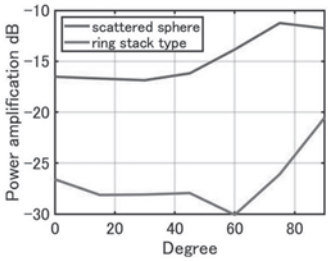



Fig. 2: Amplification of power with acoustic lens.

◆音の拡散の様子が確認できた.

## 2-1-5

### 2-1-5　音声認識システムを用いた日本語単語了解度の推定 －クラウド汎用 API と DNN-HMM による客観評価－

Estimation of Japanese word intelligibility using speech recognition systems -Objective Evaluation by Cloud GPAPI and DNN-HMM-

☆服部真稀，近藤和弘（山形大院・理工学研）

◆音声了解度の客観評価法として，音声認識を用いたノンリファレンス式の推定手法を検討

◆音声認識で主観評価手法である JDRT (Japanese Diagnostic Rhyme Test)を模擬することで了解度を推定

◆音声認識として Hybrid 型，汎用 API を利用

◆話者・雑音非依存モデルの比較では，SNR 0, 10，クリーン音声にて汎用 API による正答率は Hybrid 型を大きく上回る

◆SNR -15, -10 では未認識，誤答が増加

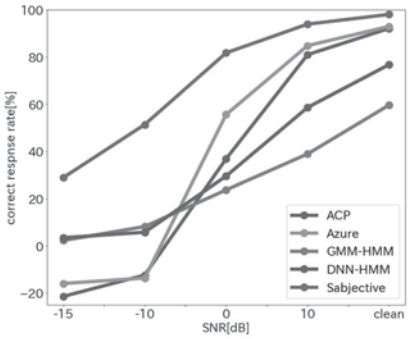



Fig.1：Correct response rate with objective and subjective vs. SNR

## 2-1-6

### 2-1-6　室内伝達特性に与えるそよ風の影響

Effect of gentle breeze on acoustic impulse response

☆森川輝久,山本斐世(阪産大),津田龍星,高橋徹(阪産大院),
江川琢真(立命大院),中山雅人(阪産大院)

**問題** 風の影響のみを調査するには無風状態との差に注目する必要あり
特定の室内配置に含む固有の伝達特性の影響を抑制する必要あり

**手段** 風環境下でインパルス応答測定(同期加算による風雑音の低減)
複数の室内配置を平均化し，固有の伝達特性の影響を抑制

**結論** 風の影響による直接波の到来時間の増減を確認

無風条件(No Wind)と到来時間を比較

- 追い風条件(Tail Wind)　→到来が 0.5 サンプル相当早い
- 向い風条件(Head Wind)　→到来が 0.5 サンプル相当遅い

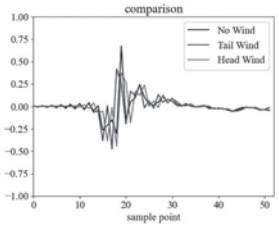



Fig.1：Direct sound arrival time difference

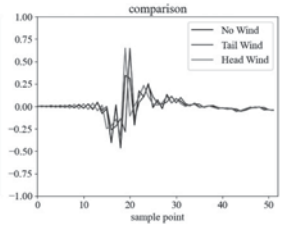



Fig.2：Direct sound arrival time difference with room environment switched

## 2-1-7

### 2-1-7　超音波空中ハプティクスの質感再現における摩擦音の特徴量に基づく振幅変調信号生成の検討

Study on Amplitude Modulation Signal Design Based on Features of Fricative Sounds in Texture Reproduction of Ultrasonic Mid-Air Haptics

☆上田明日斗，中山雅人，高橋徹（阪産大院）

◆ 従来の超音波空中ハプティクスでは，人に圧力を知覚させやすくするため，キャリア波を正弦波 200 Hz の包絡信号で振幅変調し，圧力を振動させていた．ここで，圧力を振動させると感覚が変化することに着目し，包絡信号を適切に設計することで質感を変化させることを検討する．我々は，振幅変調に用いる包絡信号に対して周波数変調を用いる手法で，瞬時周波数のスイープ周期や帯域幅を変えることで，質感を変化させる手法を提案している．本稿では，対象物の摩擦音を収録，分析し，その特徴量を振幅変調信号生成のパラメータに用いることでその対象物の質感を再現する手法を提案する．

◆ 提案手法の有効性を確認するために，質感に関する主観評価実験を行った．4 つの対象物の摩擦音からそれぞれ振幅変調信号を生成．被験者にその音源による超音波空中ハプティクスの質感と 4 つの対象物の質感を比べさせ，超音波空中ハプティクスの質感に近い対象物を選択させた．Fig. 1 に実験結果を示す．横軸は提案手法にて提示した質感、縦軸は被験者が選択した対象物、バブルの大きさは選択された数を示す。点線状のバブルが大きいほど正答率が高い．実験結果から，摩擦音から滑らかさの係数を取得し，それに合わせて粗さを変化させられることを確認した．

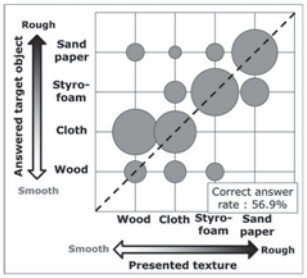



Fig. 1 :質感に関する主観評価実験結果

## 2-1-8

### 2-1-8　パラメトリックスピーカを用いた距離減衰と位相の補正に基づく音像ホログラム制御の検討

Acoustic hologram control based on distance attenuation and phase correction using parametric array loudspeakers

☆江川琢真(立命館大院/阪産大)，耿毓庭，西浦敬信(立命館大)，中山雅人(阪産大)

◆我々は PAL (Parametric-Array Loudspeaker)と EDL (Electro-Dynamic Loudspeaker) を用いた直間比制御に基づく錯覚型音像ホログラフィを提案している．本手法の方位制御では，左右のスピーカの出力制御により両耳間レベル差を制御し，距離制御では，PAL と EDL の出力の制御により直間比を制御していたが，両耳間時間差(ITD)と距離減衰を考慮することで更に性能を改善出来る可能性がある．そこで本稿では，PAL と EDL を用いた距離減衰と位相の補正に基づく音像ホログラム制御を検討する．提案手法では構築する音像の位置に応じた信号の振幅制御，距離補正では音像の距離に応じた信号の振幅補正，位相補正では音像の方位から左右の耳までの到達時間のモデルに基づく遅延を信号に付与し，PAL・EDL より出力することで音像の方位・距離を制御する．

◆評価実験を行い，音像の方位と距離の提示性能に関して評価した．Fig.1, 2 に ITD の絶対値誤差平均(EITD)とインパルス応答における最大振幅の絶対値誤差平均($\sigma_{ERROR}$)を示す．実験結果より，どちらも提案手法(PM)の値が最も低いことから提案手法の有効性が確認できた．

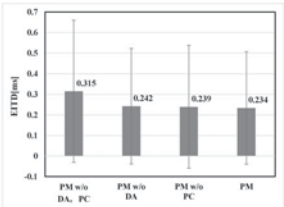



Fig. 1 Results of EITD for directions

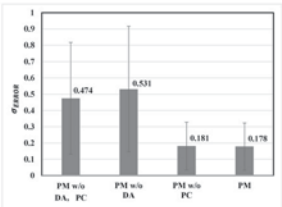



Fig. 2 Results of $\sigma_{ERROR}$ for distances

## 2-1-9

### 2-1-9　光音響計測における低周波数領域の計測雑音

Low-frequency measurement noise on optical sound measurement

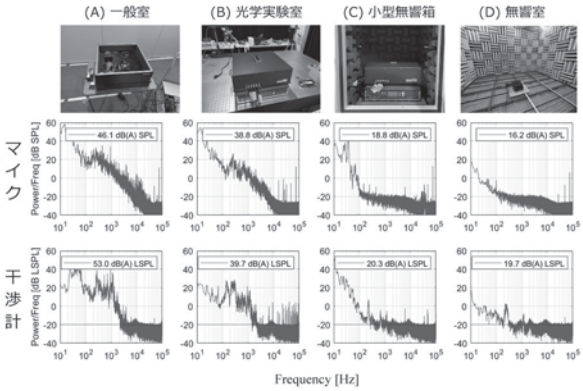
○石川憲治，△白木善史，守谷健弘(NTT)，石澤淳(日大)，日達研一，小栗克弥(NTT)

**Q.** 光音響計測は本質的に低周波数領域で雑音が大きいのか？

**A.** NO：雑音源は振動・暗騒音

振動対策を施した安定した光学系を用いて無響室で測定することで，100 Hz で -20 dB LSPL/Hz のショットノイズ限界を達成

測定環境・装置設計を工夫すれば，光音響計測は低周波数音の測定にも有効




## 2-3-1

### 2-3-1　Wave-U-Net Discriminator：敵対的生成ネットワークに基づく音声合成のための高速で軽量な識別器

Wave-U-Net Discriminator: Generative Adversarial Network-Based Speech Synthesis with Fast and Lightweight Discriminator

◎金子卓弘，亀岡弘和，田中宏，関翔悟(NTT)

◆近年，高品質な音声を高速に生成可能な**敵対的生成ネットワークベースのニューラルボコーダ**や**End-to-End 音声合成**が注目を集めている．

◆敵対的生成ネットワークで生成器と識別器の競合学習を成立させるためには，**識別器が入力音声を十分に表現**できることが必要である．

◆従来法では，**役割の異なる識別器を複数**用いることで，上記問題に対処していたが，**識別器の数の増加に伴い，識別器の総モデルサイズと処理時間が増大**することが課題であった．

◆提案する **Wave-U-Net Discriminator** では，**Wave-U-Net 構造を持つ識別器**を導入することで，**音声を波形の粒度で真偽判定**することを可能にし，**単一の識別器でも音声波形を十分に表現できる**ようにした．

◆実験では，提案法を **HiFi-GAN** と **VITS** に適用し，**音声品質を維持したまま識別器の高速化と軽量化**が可能であることを示す．

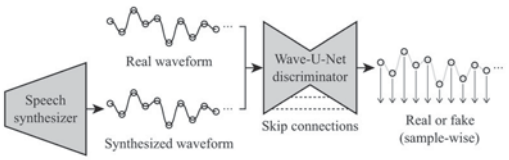



Fig.1：Overview of GAN training with a Wave-U-Net discriminator.

音声サンプル：
https://www.kecl.ntt.co.jp/people/kaneko.takuhiro/projects/waveunetd/

## 2-3-2

### 2-3-2　MISRNet：多入力単共有残差ブロックを用いた軽量なニューラルボコーダ

MISRNet: Multi-Input Single Shared Residual Blocks for Lightweight Neural Vocoder

◎金子卓弘，亀岡弘和，田中宏，関翔悟(NTT)

◆**軽量で高品質なニューラルボコーダ**として知られる HiFi-GAN は，複数の残差ブロックを並列に用いる**多受容野融合（MRF）**を用いることで，軽量なモデルで表現能力の向上を実現していたが，**残差ブロックの数に比例してモデルサイズが増える**ことが欠点であった．

◆提案する**多入力単共有残差ブロック（MISR）**では，残差ブロックの数ではなく，**カーネルサイズ 1 の軽量な畳み込み層**を用いて残差ブロックの**入力の種類を増やす**ことで，**モデルサイズの増加を抑制したまま表現能力の向上**を実現した．

◆評価実験では，軽量なニューラルボコーダである **HiFi-GAN** と **iSTFTNet** に提案法を導入し，**音声品質及び処理速度を損なうことなく，モデルの軽量化（約 30%削減）**できることを示す．

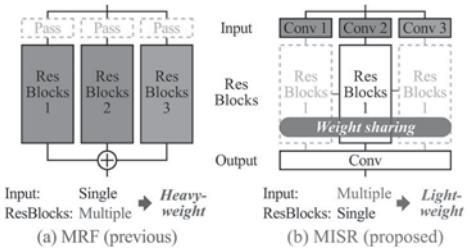



Fig.1：Comparison between MRF and MISR.

音声サンプル：
https://www.kecl.ntt.co.jp/people/kaneko.takuhiro/projects/misrnet/

## 2-3-3

### 2-3-3　複数サンプル・サブバンド信号の同時モデル化による高速なニューラルボコーダー

Joint modeling of multi-sample and subband signals for fast neural vocoding on CPU

◎金川裕紀，井島勇祐（NTT）

◆ニューラルボコーダーは、従来の信号処理ベースのボコーダーよりも高品質な音声を合成できるが計算量が大きく、高速な推論に GPU 等の並列プロセッサが必要であることが多い。

◆CPU での高速推論には順伝搬回数の削減が重要。
サブバンド WaveRNN を複数サンプル生成に拡張し高速化したが、4 サンプル以上で品質が大きく劣化。

◆同時生成サンプル数を増やしつつ品質を維持するため、隣接する複数のサンプル・サブバンド信号間の関係性を同時モデル化。

- ➢ 4 サンプル同時生成の場合、従来法の 2 サンプル同時生成から 1.47 倍の高速化を実現
- ➢ 上記の条件において、分析再合成・TTS の両タスクで自然性を損なわず動作

Table 1：Average real time factors (RTFs) @ Intel Core i7-8750H 2.20GHz. "Speed enhancement" denotes the improvement over CONVENTIONAL (M=2).

| Method | RTF | Speed enhancement |
|---|---|---|
| CONVENTIONAL (M=2) | 0.094 | - |
| PROPOSED (M=2) | 0.096 | 0.98x |
| CONVENTIONAL (M=4) | 0.055 | 1.71x |
| PROPOSED (M=4) | **0.064** | **1.47x** |
| PROPOSED (M=8) | 0.051 | 1.84x |

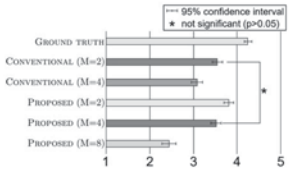



Fig. 1: Mean opinion scores of naturalness of synthetic speech. Acoustic features are predicted by FastSpeech2.

## 2-3-4

### 2-3-4　Harmonic-Net++: 基本周波数制御可能なメルスペクトログラム入力型高速ニューラルボコーダ

Harmonic-Net++: Fundamental frequency controllable fast neural vocoder with mel-spectrogram input

☆清水聡太 [1,2], 岡本拓磨 [2], 高島遼一 [1],
滝口哲也 [1], 戸田智基 [3,2], 河井恒 [2]
([1] 神戸大学, [2] 情報通信研究機構, [3] 名古屋大学)

◆従来法の Harmonic-Net+は合成速度が元の HiFi-GAN の 2 倍近く遅い

◆本研究では，合成品質を維持したまま Harmonic-Net+の合成速度を向上させるため，**メルスペクトログラム入力**の Harmonic-Net++を提案

◆Harmonic-Net+の合成品質と$f_0$の制御性能を維持したまま，**特徴量抽出を含めたリアルタイム合成**を達成

◆アップサンプリング層の最終層を Multi-stream 化し、合成速度をさらに向上

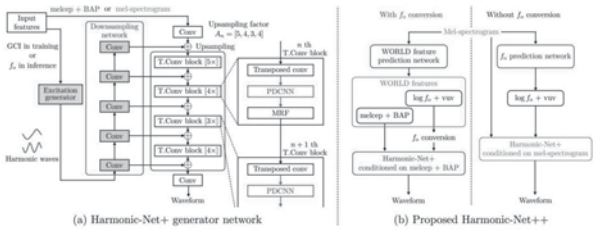



Fig.1 : Network architectures of (a) Harmonic-Net+ generator and (b) Harmonic-Net++.

## 2-3-5

### 2-3-5　SiFi-GAN：音源フィルタ構造に基づく HiFi-GAN

SiFi-GAN : Source-Filter-Architecture-Based HiFi-GAN

◎米山怜於（名大），呉宜樵（Meta Reality Labs Research），戸田智基（名大）

◆ニューラルボコーダにおける基本周波数（F0）制御性能はイントネーションやメロディパターンの柔軟な生成に不可欠

◆高速かつ高品質ニューラルボコーダ HiFi-GAN は F0 制御性能に乏しい（完全データ駆動学習ゆえの欠点）

◆HarmonicNet+ は HiFi-GAN に F0 駆動機構を導入し制御性能を改善

- ➢ 高 F0 生成の性能が不十分

◆提案手法 SiFi-GAN は HiFi-GAN に音源フィルタ構造を導入

- ➢ F0 駆動機構 ＋ 事前知識導入による帰納バイアス獲得
- ➢ 音声品質・F0 制御性能・推論学習速度を高い水準で両立

◆実験的評価から，提案手法は F0 駆動機構を導入した拡張 HiFi-GAN を大きく上回ることを実証

Table 1　Experimental results with 95% CI

| Method | V/UV ↓ | RMSE ↓ | MOS ↑ |
|---|---|---|---|
| $1.0 \times F_0$ | | | |
| Recording | – | – | $3.27 \pm 0.06$ |
| HiFi-GAN ex | 6 | **0.08** | $3.19 \pm 0.06$ |
| SiFi-GAN v1 | 6 | 0.09 | $3.28 \pm 0.06$ |
| SiFi-GAN v2 | 6 | 0.09 | $\mathbf{3.29} \pm 0.06$ |
| $0.5 \times F_0$ | | | |
| HiFi-GAN ex | 9 | 0.13 | $3.09 \pm 0.13$ |
| SiFi-GAN v1 | 9 | 0.13 | $\mathbf{3.16} \pm 0.13$ |
| SiFi-GAN v2 | 9 | **0.11** | $3.09 \pm 0.13$ |
| $2.0 \times F_0$ | | | |
| HiFi-GAN ex | 9 | **0.12** | $2.89 \pm 0.17$ |
| SiFi-GAN v1 | 7 | 0.15 | $3.09 \pm 0.16$ |
| SiFi-GAN v2 | **6** | 0.13 | $\mathbf{3.14} \pm 0.15$ |

## 2-3-6

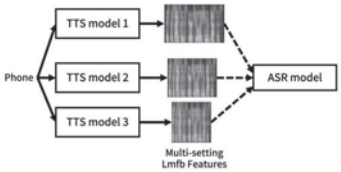
### 2-3-6　複数設定のスペクトログラムを用いた音声合成に基づく音声認識のデータ拡張

Data augmentation for speech recognition based on speech synthesis using multiple setting spectrogram

◎上乃聖，李晃伸（名工大）

◆ 音声合成を用いた音声認識のデータ拡張では，合成音声の自然性の向上以外にも，生成される音声の音響的な多様性が音声認識の性能改善に必要である．

◆ 本研究では，音声合成が生成するメルスペクトログラムの設定に着目し，複数の設定を持つスペクトログラムを生成し，音響的な多様性を持たせる手法を提案する．




◆ 複数個音声合成モデルを用いて，各音声合成モデルが生成するスペクトログラムは，nFFT，window size，hop size が異なるものを予測する．

◆ 実験結果により，一つだけの設定で生成した時と比較し，複数のメルスペクトログラムの設定を用いて音声認識のデータ拡張に用いた方が性能が良いことを示した．

Table 1：ASR performance (WER (%)) for LibriSpeech testset.
We used 10-hour paired data (speech and transcription) from Librispeech100.
We generated speech from 950-hour transcription.

| Model | test-clean | test-other |
|---|---|---|
| Real (960h) [oracle] | 2.90 | 7.14 |
| Augmented model (real: 10h, generated: 950h) | | |
| Setting 1 (used in ASR) | 5.89 | 23.36 |
| Setting 2 (used in TTS) | 5.72 | 22.82 |
| Setting 3 (used in TTS) | 6.12 | 23.02 |
| **Mixed setting [proposed]** | 5.30 | **21.25** |

## 2-3-7

### 2-3-7　メタ転移学習に基づく少数言語 End-to-End 音声認識における文字表現の効果分析

メタ転移学習、音声認識、ローリソース、アルファベット揃える

○周鋭，伊藤彰則，能勢隆（東北大）

◆ごく少量のデータを用いて音声認識を行うことを目標として，転移学習とメタ学習を組み合わせたメタ転移学習を提案した．その後，異なる言語のアルファベットを揃える方法を提案した．方法の有効性を証明するため，転移学習，メタ学習，メタ転移学習の三つの方法の実装と実験を行った

◆Tab.1 は異なる学習方法とアルファベットを揃える方法の結果です，テストは Bulgarian、Croatian、Swahili、Hausa 四つの言語を使って，毎種類言語のデータは 15 分間以下である。

◆結果により，三つの学習法いずれを使っても，約 15 分の微調整データにより，40%以下の CER を得ることができる。メタ転移学習により最も低い CER を得ることができた。最後に，アルファベットを揃えたら，対象言語がラテンアルファベット出ない場合に改善が見られた．

Tab.1 Character Error Rate of different learning methods

| | Bulgarian | Croatian | Swahili | Hausa |
|---|---|---|---|---|
| Transfer-Learning | 28.94 | 30.93 | 18.13 | 23.82 |
| Meta-Learning | 32.86 | 35.25 | 19.93 | 25.76 |
| Meta-Transfer-Learning | 28.26 | 30.95 | 17.93 | 22.03 |
| Alphabet-Unify | 27.15 | 30.64 | 18.17 | 22.12 |

## 2-3-8

### 2-3-8　End-to-End 音声要約のための テキスト要約コーパスを用いたデータ拡張

Data Augmentation Using External Text Summarization Corpora for End-to-End Speech Summarization

◎松浦孝平，芦原孝典，森谷崇史，田中智大，小川厚徳，
マーク・デルクロア，増村亮（NTT）

◆音声文書の要約を生成する音声要約技術が近年注目を集めつつある．

◆音声要約技術は従来 Cascade 型モデル（Fig. 1a）で実装されてきたが，音声認識誤りの伝播や非言語情報の欠落という問題がある．

◆音声から要約を直接生成する End-to-End 型モデル（Fig. 1b）は上記の問題の影響を受けにくいが，その学習に必要な音声と要約のペアデータの収集は一般的にコストが高い．

◆そこで，本研究では外部テキスト要約データセットを，合成音声や音素列を経由して End-to-End 型モデルへ与える手法を提案する．

◆How2 データセットにおける実験において，Table 1 のように提案手法（T-1, T-2）は Cascade 型（C-1）を上回る性能を達成した．

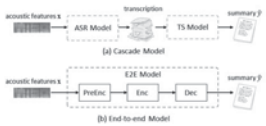



Fig. 1：Comparison between the (a) Cascade and (b) End-to-end (E2E) approaches for speech summarization. The cascade model consists of ASR and text summarization models while the E2E model unifies them.

Table 1：METEOR (MTR) and BERTScore (BS) scores of previous work (P-1) [Sharma+ 2022], our cascade and two end-to-end baselines (C-1, B-1, and B-2, respectively), and proposed models (T-1 and T-2).

| ID | # params | MTR (↑) | BS (↑) |
|---|---|---|---|
| C-1 | 201M+140M | 30.3 | 92.6 |
| P-1 | 104M | 28.8 | 91.0 |
| B-1 | 98M | 32.5 | 93.0 |
| B-2 | 203M | 32.8 | 93.2 |
| T-1 | 203M | **34.9** | **93.8** |
| T-2 | 203M(+7M) | 33.9 | 93.6 |

## 2-3-9

### 2-3-9　End-to-end 音声認識のための 後向き言語モデルの繰り返し shallow fusion

Iterative shallow fusion of backward language model for end-to-end speech recognition

○小川厚徳，森谷崇史，加茂直之，俵直弘，マーク・デルクロア（NTT）

◆End-to-end (E2E) 音声認識の精度向上には，大量のテキストデータを用いて学習された外部言語モデルの，メインの E2E 音声認識モデルへの shallow fusion (SF) が大きく寄与する．

◆SF では音声認識・外部言語の両モデルが出力するスコアを線形補間しながらビーム探索を行う．SF には前向き言語モデル (forward language model: FLM) が用いられる．

◆リスコアリングには FLM に加えて **後向き LM (backward LM: BLM)** もよく用いられ，互いの相補性により大きな効果を示す．

◆SF と同様に，E2E 音声認識におけるビーム探索に BLM を適用するための **繰り返し SF (iterative SF: ISF)** を提案する．更には ISF の計算コスト削減手法と，ISF に適した BLM の学習手法も提案する．

◆ESPnet の TED-LIUM2 レシピを用いた音声認識実験により，提案する BLM を用いた ISF の有効性を確認した．

Table 1. Comparison of the SF methods.

| Method | WER [%] |
|---|---|
| W/o SF | 11.2 |
| SF | 9.9 |
| ISF | 10.0 |
| SF + ISF | **9.1** |

## 2-3-10

### 2-3-10　Cyclic Partially-aligned Transformer for Visually Connected Speech-to-text Mapping

J. Effendi (NAIST), ○ S. Sakti (JAIST/NAIST), S. Nakamura (NAIST)

◆Proposes a **cyclic partially-aligned Transformer framework** to address the speech-to-text mapping problem given only a collection of **visually connected non-parallel speech-text data**.

- **Transformer-based VQ-VAE** is used to produce a discrete speech representation in a self-supervised manner.
- **Transformer-based sequence-to-sequence model** inside a **chain mechanism** is used to map from unknown untranscribed speech utterances

◆Our experiments showed improvement in comparison to the ASR baseline. Even with a limited dataset, our proposed method can still **effectively convey the semantics of a partially-aligned speech**.

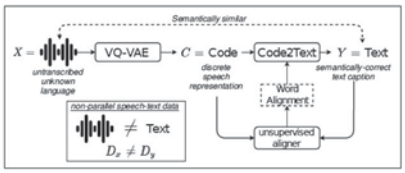



Fig.1：Overview of the Proposed System

Table 1：System Performance on Flickr8k

| Model | Sim% | BLEU | CIDEr |
|---|---|---|---|
| (Baseline) | | | |
| Random selection | 16.25 | 2.22 | 3.58 |
| ASR [8] | 16.94 | 4.69 | 7.08 |
| (Proposed) | | | |
| Code2Text | 35.79 | 15.04 | 31.66 |
| +Partial Code2Text | 40.94 | 16.80 | 36.86 |
| +Cycle Augmentation | **40.47** | **17.25** | **37.52** |

## 2-4-1

### 2-4-1　長時間ピッチ周波数変化量に基づく 日本語イントネーション平板化 評価尺度の提案

A study on measuring the degree of flattening of Japanese intonation based on pitch-frequency variation.

☆佐藤文哉，荒井秀一（都市大院）

日本語イントネーションの変化の 1 つに，発話の抑揚がなくなる平板化現象がある。日本語には同音異義語があるため，抑揚がなくなると発話内容を正しく伝えられない可能性が高くなる。そのため，円滑なコミュニケーションには平板化に注意する必要がある。しかし，イントネーションを平板化させるのは無意識のため，自分の平板化度合いを自覚するのは難しい。

また従来研究においても，話者毎の平板化度合いに焦点を当てた研究は存在せず，平板化の程度を測る尺度はない。現状の平板化を解析する方法は、発話内容を音素アライメントした結果とピッチ抽出結果からアクセントの型を推定する方法がある。しかし，発話ごとに，音声認識，音素アライメント，アクセント型解析を経た結果と，実際のピッチ周波数解析結果とを比較する必要があるため，この手法で安定した平板化判定を実現するのは困難であると考えた。そのため，発話内容に依存せずに平板化の度合いを測る方法を提案した。

## 2-4-2

### 2-4-2　宣伝音声と購買意欲の関係における感情と聴取者属性の影響：調整媒介分析を用いた検証

The influence of emotion and listener's attributes on the relationship between advertising speech and the willingness to buy: Verification using moderated mediation analysis

◎長野瑞生，井島勇祐，廣谷定男（NTT）

◆感情を媒介とする消費者行動モデル [R. Donovan+, 1982]

- ➢ 消費者の購買行動を外部刺激→感情→購買意欲の３ステップで表すモデル．音声刺激においてもこのモデルが適用可能であるが、感情の媒介効果は聴取者の属性によって異なることが示唆された [M. Nagano+, Front. Psychol., 2023]
- ➢ 前報[長野他, ASJ2022s]では聴取者の属性ごとに主観評価データを分けたうえで媒介分析による検証を行っており、分析に用いるサンプル数の減少や，属性間で比較するために媒介効果の割合の算出が必要等の問題があった

◆本稿では，上記問題を解決するために調整媒介分析を用いて検証

◆平均 F0, 話速, SDF0 を変更した宣伝音声を用いて，音声が購買意欲に影響を与えるときの感情の媒介効果に聴取者の属性が影響するかを大規模な主観評価実験によって調査

- ➢ 感情の媒介効果は聴取者の性別・年代によって異なる
- ✧ 男性よりも女性の方が感情の媒介効果が高い
- ✧ 年齢が高いほど快感情の媒介効果は高く、覚醒や支配の媒介効果は低くなる

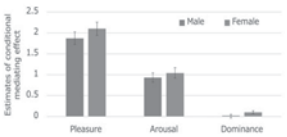



Fig.1: Gender-moderated mediating effect of emotional states.

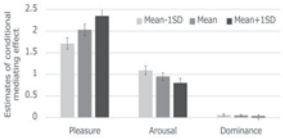



Fig.2: Age-moderated mediating effect of emotional states.

## 2-4-3

### 2-4-3　音声変換技術を用いた英語発話リズム訓練のためのｅラーニング教材の開発

Development of e-learning materials for English speech rhythm training using speech conversion technique

○廣谷定男，△亀井剛次，△服部正嗣（NTT CS 研）

◆多くの日本語母語話者は，発話した英語音声が英語母語話者に伝わりにくいという問題を抱えている．その理由の一つとして，英語は強勢拍リズムであるが，日本語はモーラ拍リズムであるため，日本語母語話者は英語をモーラ拍リズムで発話する傾向があることが考えられる．

◆我々はこれまでに，日本語母語話者が発話した英語音声の発話リズムのみを強勢拍リズムに変える発話リズム変換技術に基づく英語発話リズム訓練スマートフォンアプリを開発し，高校生を対象とした実証実験によりアプリの有効性の検証を行ってきた．

◆しかし，Android 端末のみで動作することや変換サーバとの Wi-Fi 接続が必要であったため，自宅等での長期訓練による効果の検証が困難であった．

◆本発表では，英語発話リズム訓練のためのｅラーニング教材を開発したので報告する．

Table.1: Web browser-based apps.

## 2-4-4

### 2-4-4　子音強調処理の改良による骨導提示音声の了解度改善

Improvement of word intelligibility of bone-conducted speech by modifying consonant emphasis method

☆王 思成（北陸先端大），上江洲 安史（北陸先端大），
鳥谷 輝樹（金沢大），鵜木 祐史（北陸先端大）

◆背景：骨導提示音声の了解度が低下するという問題に対し，先の提案法（CE）では，子音部を選択的に強調することで解決した．しかし，子音部の検出性能が低いため，了解度の改善効果が十分ではなかった．

◆目的：無声子音／有声子音の音響特徴に基づき，子音部の検出性能を大幅に改善することで，CE 法の問題を解決すること．

◆手法：無声子音／有声子音区間検出法を設計し，それぞれの結果の統合処理により子音区間を判定した後，検出した子音部の振幅を 12 dB 強調する処理を適用する（CE-IMP）．

◆評価：雑音環境下（55, 75 dB）で，４つの親密度ランクによる単語了解度試験を実施したところ，雑音レベル 75 dB において CE-IMP が他の強調法（CE と RT-FOE）よりも単語了解度を大幅に改善した．

◆結論：子音区間検出法を改良した子音強調処理を利用することで，先の問題を解決し，骨導提示音声の了解度を大幅に改善できた．

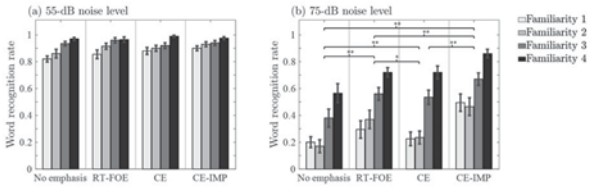



Fig.1: Mean word recognition rate derived in word intelligibility tests for each filter type and each word-familiarity rank under 55-dB and 75-dB noise level.

## 2-4-5

### 2-4-5　ささやき声の音調パターンの操作と疑問文の生成：雑音駆動音声のスペクトル変換によるピッチ制御

Pitch control by a spectral transformation of whispered voice.

○内田照久（大学入試センター 研究開発部）

**目的**　ささやき声のピッチ知覚を検討するため，スペクトルの変形によるピッチ制御の可能性を検討

**方法**　下降調の平叙文の音声から，上昇調の疑問文の音声を生成

(1) 通常音声で，時間長，音圧，$f_0$ を操作して上昇調に変換
(2) 通常音声から疑似ささやき声を合成，上昇調の生成も図る
(3) リアルなささやき声に，上記の変換法を適用して試作

**課題**　生成した音声を用いた聴覚実験による評価の実施
発声器官での生成面に留意したスペクトル変形法の改良

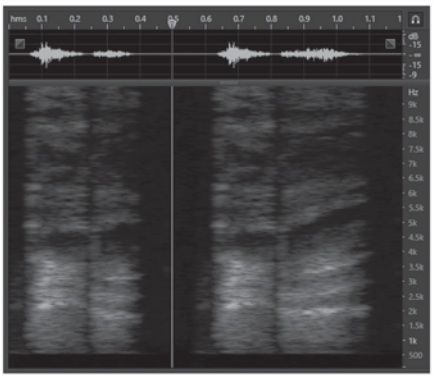


Fig. 1 Speech waveforms and spectrograms at the end of pseudo-whispered sentences.[M1: te-iru]
(The left figure shows the speech in descending pitch, and the right figure shows the speech converted to ascending pitch)

## 2-4-6

### 2-4-6　ポテトチップスの咀嚼音の音響解析と深層学習による食感推定

Acoustic Analysis of Chewing Sound of Potato Chips and Texture Estimation by Deep Learning

☆橋本卓己(神奈川工科大), 三浦貴大(産総研), 上田麻理(神奈川工科大)

◆前報では，音と食感の関係を明らかにするために 10 種類のスナック菓子を用いた咀嚼音の音響解析と深層学習による種類判別を行った。その結果，空気伝搬音の高周波成分が食感に違いを生み出している可能性があることが示唆されたが，スナック菓子の種類判別の精度が 30 %前後であり，データの見直しとモデルの改善が必要とする課題が挙げられた。

◆本稿では，その課題を踏まえてポテトチップス１噛み毎に１枚のメルスペクトログラム画像として3パターンのデータセットと4種類の学習手法を用いて，引き続き食感に基づいて種類を判別することが可能であるか試みたので報告する。

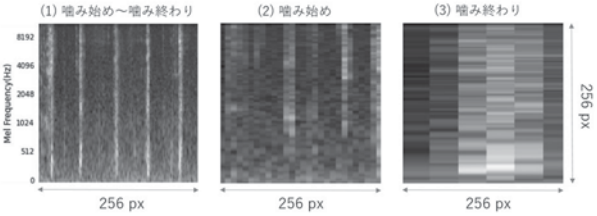



Fig.1 : Examples of data set (Mel Spectrogram)

## 2-4-7

### 2-4-7　謡曲の良さの評価手法についての検討

Study on evaluation method of the goodness of Noh singing

○木谷俊介(北陸先端大)

◆　謡曲の良さを評価する手法は確立されていない

◆　謡曲では幽玄であることが重要である（世阿弥“花鏡”, 1424）。

◆　謡曲の評価は、音響学的アプローチと美学的アプローチから研究されてきた。

◆　本研究では、

- 謡曲の良さを評価するための音響分析手法を提案（音響学的アプローチ）
- 謡曲の良さがこれまでにどのように評価されてきたのかを紹介（美学的アプローチ）
- 幽玄をどのように評価すべきかを考察(美学的アプローチ)する。

◆　今後は、音響学的アプローチと美学的アプローチを統合的に評価する手法の確立を目指す。

## 2-4-8

### 2-4-8　無音検出から無音弁別へ －実験と数理モデルの拡張－

From gap detection to gap discrimination: Experiments and extension of mathematical model

○森周司(九大)，△Cho Hyunsoo(CRI・ミドルウェア)，△Wong Willy(U. Toronto)

◆純音を用いた無音検出と弁別の実験を行い、無音検出の数理モデル（Mori et al. 2018）の弁別への拡張を試みた。

◆実験結果（Fig.1）は、周波数内（後続音 800 Hz）と周波数間（1600、2500 Hz）条件で、検出と弁別閾値が異なることを示した。

◆数理モデルに信号検出理論と神経発火の過渡的応答を導入することで、上記の結果を良く説明することが分かった。

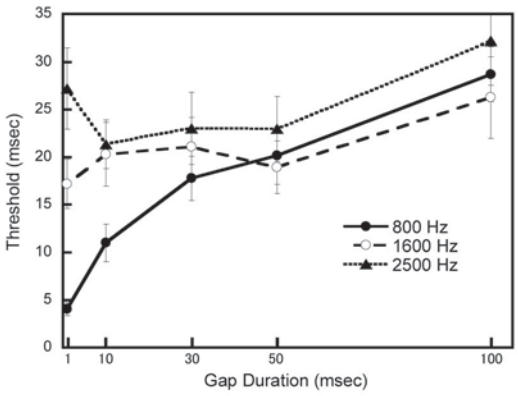



Fig.1 : Mean thresholds as a function of standard gap duration, separately for trailing marker frequency. Vertical bars represent standard errors.

## 2-4-9

### 2-4-9　選択的聴取能力と耳音響放射特性の関係性の検討

Study on the relationship between selective listening ability and otoacoustic emission

☆宮家一真，木谷俊介，鵜木祐史（北陸先端大）

◆背景　選択的聴取能力が聴覚末梢で生じるマスキングに関連するものと考えれば，耳音響放射（OAE）で選択的聴取能力を推定できる可能性がある．選択的聴取能力と OAE の関係を明らかにできれば，OAE を利用することで，選択的聴取能力の欠如といった聴覚情報処理の問題の早期発見に役立てることできる．

◆目的　選択的聴取能力と OAE の関係性を明らかにすること．

◆方法　選択的聴取能力（図と地課題）と TEOAE／DPOAE の測定実験を行い，それらの対応関係を調べる．

◆結果　Figure 1(a), (b) から右耳でTEOAE の測定値に応じて図と地課題の正答率が上昇することが分かった．Figure 1(c), (d) から，両耳で DPOAE の測定値と図と地課題の正答率が上昇することが分かった．

◆結論　OAE の測定値の高低から，選択的聴取能力の高低を推定できる．

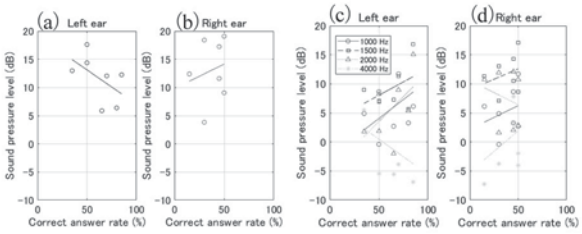



Fig.1 : Correct answer rate of auditory figure-ground task and OAE measured value (a) Left ear and (b) Right ear of TEOAE, (c) Left ear and (d) Right ear of DPOAE

## 2-4-10

### 2-4-10 Research on vibration characteristics of behind-the-ear cartilage conduction by comparing acceleration measurements between artificial and human pinnae

◎ Xiuyuan Qin, Sho Otsuka, and Seiji Nakagawa (Chiba Univ.)

- ◆ A novel type of headphone which transmits sound signal by cartilage conduction (CC) was proposed. The new position which is behind the ear (BTE).
- ◆ In this study, acceleration measurements were measured on artificial and human pinnae to investigate the relationship between hardness of the pinna and acceleration produced by BTE-CC.
- ◆ Although relationship of hardness and acceleration can be found in some specific conditions, other factors also influenced acceleration when BTE-CC was used.

Fig. 1 The self-made BTE CC headphone.

Fig. 2 The five points where an acceleration sensor was put for measurements

Fig. 3 The artificial pinnae used in the measurement. (from left to right : HATS-H, HATS-S, S1, S2)

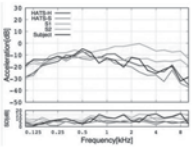



Fig.4.The acceleration results measured on four different artificial pinnae and subjects' pinnae at P1.

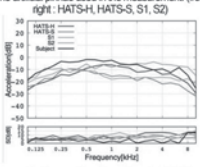



Fig.5 The acceleration results of four different artificial pinnae and subjects' pinnae at P2.

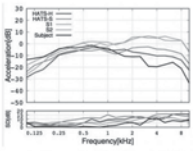



Fig.6. The acceleration results of four different artificial pinnae and subjects' pinnae at P3.

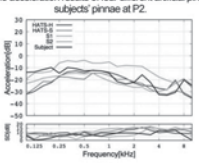



Fig.7. The acceleration results of four different artificial pinnae and subjects' pinnae at P4.

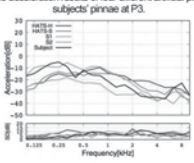



Fig.8. The acceleration results of four different artificial pinnae and subjects' pinnae at P5.

## 2-4-11

### 2-4-11 高齢者における音像定位に及ぼす音の立上り時間の影響

Effect of Rise Time on Sound Lateralization in the Elderly

○GUO YIJIE(中央大院), △森田和元, 戸井武司(中央大)

本研究では，高齢者の右左耳の違いを明らかにするために，異なる立上り時間(Rise time)と異なる両耳間時間差(ITD)を与えた音源を用いた音像定位実験を実施し，高齢者における音像定位に及ぼす音の立ち上り時間の影響を検討する。 被験者は，高齢者（男性 12 名，女性 12 名）と若年者（男性 12 名，女性 12 名）の計 48 名である。聴力検査と利き耳を調査し，Rise time と ITD を含めて実験結果に影響を与える要因を検討した。Rise time は音像定位の判断に影響を与えた。まず，図 1(a) ITD =±0.8 ms の場合，若年者の左右耳先行音に対する正解率の比較では，Rise time の増加とともに正解率が大きく減少し，右耳の正解率が左より高いことから，右耳優位性(REA)を示している。 次に，図 1 (b)に示す高齢者の左右耳先行音に対する正解率の比較では，REA は認められない。両図の比較より，高齢者の方が若年者よりも正解率が低下する傾向が見られる。

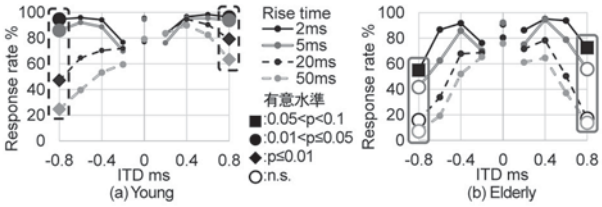



Fig.1 Differences of the percentage of correct responses regarding Rise time and ITD

## 2-4-12

### 2-4-12 骨伝導による音像定位を活用した方向認知精度の向上

Improvement of direction recognition accuracy using sound image localization by bone conduction

☆小澤風太（中央大院），△白石英孝，戸井武司（中央大）

歩行者用音響経路ナビゲーションとして，耳を塞ぐことのない骨伝導イヤホンを用いた音像定位技術に着目し，気導音に対して劣るとされる骨導音の音像定位精度の向上を目指した。

周波数帯域を制限した音源と，移動範囲と移動時間をそれぞれ変更した左右移動する音像を用いて音像定位実験を行い，方向認知精度がどのように変化するか調べた。

水平面方向の音像定位には(a)に示す 0-4 kHz 付近の周波数帯域が有効である可能性がわかった。左右移動する音像を用いた場合は，(b)に示す移動角度が 30°，(c)に示す往復時間が 1 s の場合が最も方向認知精度が高くなり，骨導音と気導音の定位精度が同等となった。このことから左右移動する音像を用いることで方向認知精度を向上させることが可能であると推察できる。

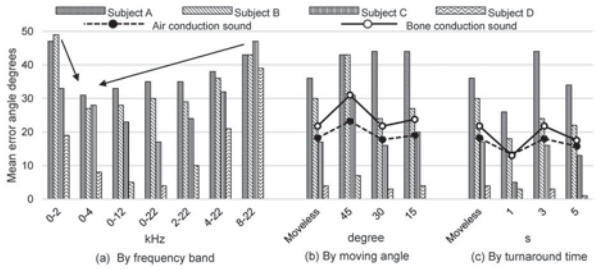



Fig. 1 Error of sound image localization direction from target direction when frequency band, moving angle and turnaround time is changed

## 2-4-13

### 2-4-13 背景に 3D スキャンした実験室を提示したときの腹話術効果

A study of ventriloquism effect
Presented 3D-scanned experimental room image as background image

◎河井薫，武藤憲司(芝浦工大)

- ◆ 本実験の目的は，視覚刺激の背景が腹話術効果に与える影響を明らかにすることである．我々は，腹話術効果の変化を明らかにするために，視聴覚刺激を水平にずらしたときにそれらのずれを知覚する角度を角度弁別閾として調査している．本研究では，3D スキャンした実験室を視覚刺激の背景として提示したとき(VR with a room image 実験，図 1 左)と，青空と地面という単純な背景を提示したとき(VR image 実験，図 2 右)の腹話術効果の差を調査した．
- ◆ 実験結果を以下図 1 右に示す．点線が VR image 実験，実線が VR with a room image 実験の実験結果である．全体的に VR image 実験と VR with a room image 実験の実験結果は同様の結果となった．

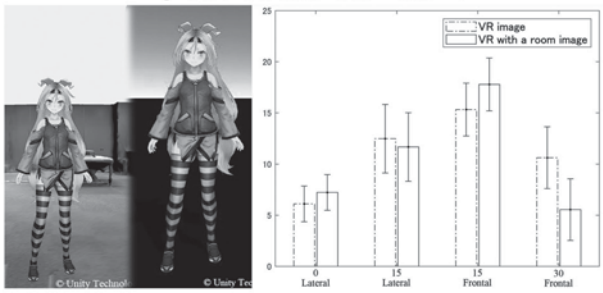



Fig.1: Visual stimulus in VR with a room image and in VR image, and results of VR and VR with a room image experiments.

## 2-4-14

### 2-4-14　騒音・評価音声の空間知覚特性の模擬による雑音環境下ハンズフリー拡声通話音の品質評価手法の検討

Study on robust speech quality evaluation method for hands-free communications with auditory spatial perception.

○栗原 祥子，鎌土 記良，千葉 大将，中山 彰
（日本電信電話株式会社 NTT コンピュータ＆データサイエンス研究所）

◆従来のハンズフリー拡声通話の音声通話品質評価法は，近端端末が雑音環境下にあることを考慮しておらず拡張が望まれている。

◆本稿では，雑音環境下における拡声通話音，周囲雑音の空間定位に着目した。耳掛型イヤホンから拡声通話音を，スピーカから周囲雑音を再生し，騒音の空間定位を模擬する新たな受聴試験を提案する。

◆提案法は，周囲雑音の S/N や発話条件に関わらず，静穏環境に近い評価傾向となることが明らかとなった。騒音の空間知覚特性は，通話音声の主観評価に少なからず影響を与えていると考えられる。

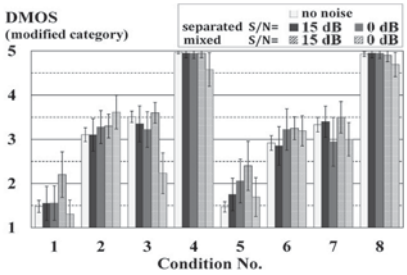



Fig.1： Result of listening test.

## 2-5-1

### 2-5-1　音響透過損失向上を目的としたヘルムホルツレゾネータによる音響メタマテリアル

Extraordinary sound transmission loss by acoustic metamaterial embedded with Helmholtz resonators

○山本 崇史（工学院大学）

ヘルムホルツレゾネーターを用いた音響メタマテリアルを用いて，一重壁構造のコインシデンスによる性能低下を抑制し遮音特性を向上させた事例を紹介する（図 1）．また，二重壁構造の共鳴透過による性能低下を抑制し遮音特性を向上させた事例についても紹介する（図 2）．

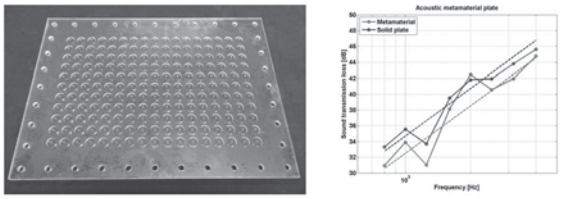


Fig.1： Acoustic metamaterial plate embedded with Helmholtz resonator.

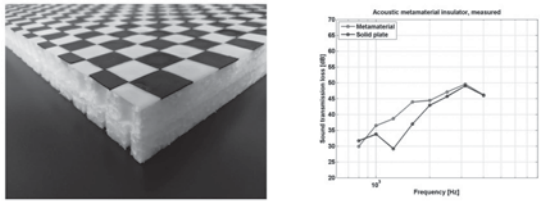


Fig.1：Acoustic metamaterial insulator embedded with Helmholtz resonator.

## 2-5-2

### 2-5-2　Locally resonant 型音響メタマテリアルシートの遮音特性

Sound insulation properties of locally resonant type acoustic metamaterials

○赤坂修一（東工大・物質理工），△中山真成（三菱ケミカル）

◆ばね部材と錘からなる構造体（スタブ）を音波の波長よりも短い周期で配列させた Locally resonant 型音響メタマテリアルシートを作製し、遮音特性とそのメカニズムについて、実験と数値計算により検討した。

◆スタブの固有振動数近傍（2 kHz 付近）で基板の振動が著しく低減し、質量則を超える高い遮音性能を観測した。また、3 kHz 付近では、スタブと基板の共振による透過損失の低下が見られた。

◆LR 型音響メタマテリアルの有効質量を算出し、上記の遮音特性との関係を明らかにした。

◆スタブの一体シート化により、生産性、施工性は大幅に改善された。一方、シート厚の増加は、スタブの効果を減少させ、遮音性能を低下させるため、これらを考慮した設計が重要であることが分かった。

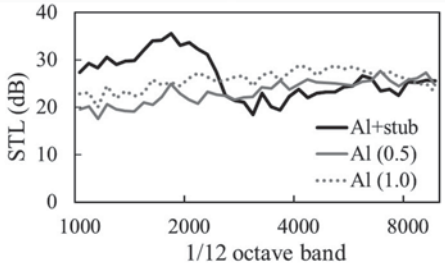



Fig.1：Sound transmission loss of acoustic metamaterial (stub: silicone rubber and Fe with 6 mm in diameter and 3mm in height, interval: 20mm, 10×14 pieces).

## 2-5-3

### 2-5-3　波動ブラックホールを含む構造の制振・防音特性について

Improvement of Vibration and Sound for Structures Including Acoustic Black Hole

○山口誉夫（群馬大・理工）

軽量で低振動低騒音を得るために Mironov らは波動ブラックホールを，提案している。波動ブラックホールは，板のエッジにおいて，パネルの板厚 $h(x)=\varepsilon x^m$ を急激に減少（伝播距離 $x$ の高次べき級数）させる構造をとる。板厚を端部で減少させていくと，曲げ波の波長が板厚の減少に伴って短くなる。短い距離で波動の波数が多くなり，波が長距離にわたり進行することと等価になる。その板厚の減少（$m$）が 2 次関数以上のべき級数であると，無反射とできることが理論的に示された。Krylov らは波動ブラックホールの表面に薄い粘弾性減衰材を積層することで，効果的に減衰が得られ，ブラックホールの板先端を短くできることを示した。減衰層の効果は（減衰層の板厚/鋼板の板厚）$^2$ に比例することが Oberst の理論から知られている。波動ブラックホールでは，先端の鋼板の板厚が極めて薄く，減衰層の効果が高い。これよりブラックホールの先端の長さが短くできる。本論では、制振構造物の減衰特性を効率的に解析できる解析手法である「モード歪みエネルギ法」を用いて、いくつかの波動ブラックホールについてその構造と制振・防音解析事例を説明する。

残渣板厚の効果やクレータ型波動ブラックホールの制振効果について示す。波動ブラックホール部の変形が他の部分の変形に比べて相対的に大きいモードほど高減衰が得られることなどを明らかにした。

## 2-5-4

### 2-5-4　完全バンドギャップを持つ音響メタマテリアル梁

Acoustic metamaterial rods with perfect band gaps

○松田 理，△藤田 健太郎，△小笠原 央，△友田 基信，△ライト オリバー
（北大院工）

物質中に共振構造を配置することで、物質を伝播する音響波の振舞いを人為的に制御できる。このような構造は音響メタマテリアルと呼ばれ、特定の振動数領域の音響波が伝播できない音響バンドギャップを設けることができる。本講演では、伝播するすべての音響モードをある特定の周波数範囲で完全に遮断する梁形状音響メタマテリアルの設計、およびこれに基づいて実際に作製した構造の音響波伝播実験結果について紹介する。

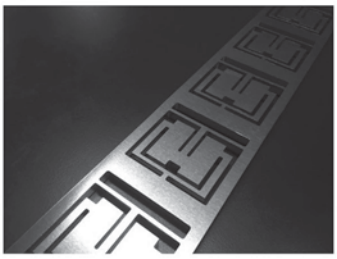


Fig. 1: A metamaterial rod made of aluminum with the size of 80 mm x 8 mm x 800 mm. It stops all acoustic modes propagating along the length direction (=800 mm) around 1 kHz.

## 2-5-5

### 2-5-5　音響問題におけるトポロジー最適化の展開の試みについて

Topology optimization in acoustic problems

○山田崇恭（東京大・工）

構造最適化は，設計者の勘や経験に頼らずに，力学的，数学的根拠に立脚して構造物の最適な形状を自動的に設計する方法である．構造最適化分野は，最適制御理論の研究分野として応用数学者を中心として，理論的研究が行われてきた．その後，計算機性能の発達とともに，有限要素法などの力学シミュレーションを援用した設計技術として，主に構造力学関連の工学分野の研究者も加わった．その後，その応用展開の研究を含めて世界中の多くの研究グループから成果発表がされるようになった．音響問題への最初の展開は，Wadbro らによる研究である．弾性体の固有振動問題，調和振動を仮定した弾性振動問題も含めれば，トポロジー最適化の研究が始まった頃から多数の研究成果報告がある．狭い意味での音響メタマテリアルの設計問題，すなわち，局所共振現象を利用した音響材料の設計問題への展開は，2010 年頃から報告されるようになった．メタマテリアルに関しては，負のポアソン比や負の熱膨張係数の設計に始まり，電磁波伝搬問題への展開など，デンマークの研究チームが先進的な研究成果を発表してきている．現在までに，多くの展開事例が報告されているが，本講演では，その中でも，音響問題，特に音響メタマテリアルへの展開について紹介する．最初に，トポロジー最適化の基本的な考え方を述べ，音響問題への展開方法について議論する．次に，研究グループにおいて取り組んだ事例について紹介する．

## 2-5-6

### 2-5-6　膜振動とヘルムホルツ共鳴を利用した音響メタマテリアルの有限要素解析

FE analysis of acoustic metamaterials using membrane vibration and Helmholtz resonance

○黒沢良夫（帝京大・理工）

◆本研究では、1 枚の PP（ポリプロピレン）のシートを折りたたみ（折り紙構造）、六角形の断面形状のハニカムの繰り返し構造を作り、そこに PP や PE（ポリエステル）からなる薄いフィルムを接着することにより、フィルムの面外振動により音響エネルギーを吸音する音響メタマテリアルを用いた。

◆フィルムに穴を開けることでフィルムの面外振動による吸音に加えてヘルムホルツ共鳴による吸音効果を付与した構造について、有限要素法を用いてテストピースのモデルを作成し(Fig.1)、実験結果との比較や(Fig.2)、数値計算により構造や材料データを変更した際の吸音特性の変化を解析した結果を報告する。

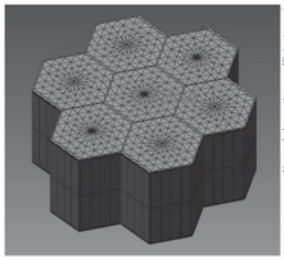


Fig.1 FE model for test piece

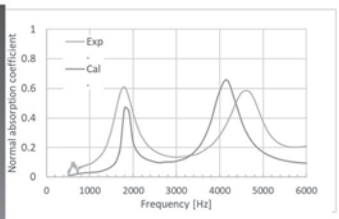



Fig.2 Comparison between experimental result and calculation result

## 2-5-7

### 2-5-7　複数回折り畳まれた管を応用した音響メタサーフェスの吸音特性

Sound absorption of acoustic metasurface applied multiple folding tubes

○中西伸介（広島工業大）

◆Fig.1 に示す，複数回折り畳まれた管を応用したユニットセル（12 mm ×12 mm×17.2 mm）を組み合わせた音響メタサーフェスを試作し，吸音特性を測定した．

◆Fig.2 は，従来の研究と同様に 200 Hz から 630 Hz で吸音特性の広帯域化が得られたことを示している．

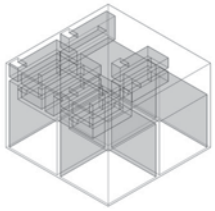


Fig.1 Super cell with multiple folding tubes for an acoustic metasurface.

◆さらに，Fig.2 は，空洞の深さ 43.2 mm・総厚 49.6 mm の音響メタサーフェスでは，100 Hz から 400 Hz で 0.1 以上の吸音特性を示し，低周波数化も得られたことを示している．

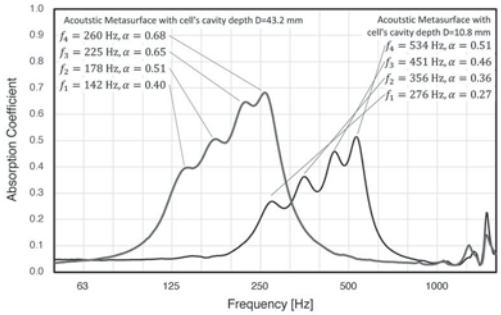



Fig.2 Absorption coefficient of acoustic meta-surface consisted of resonators.

## 2-5-8

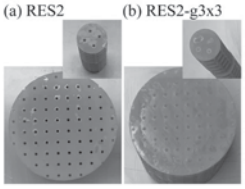
### 2-5-8　周期・非周期構造を利用した建築音響材料の開発

Development of architectural acoustic metamaterials using periodic and non-periodic microstructures

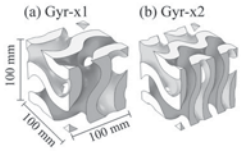
○菅原彬子（近畿大）

著者のグループがこれまで取り組んできた，音響メタマテリアル開発に向けた検討を紹介する．

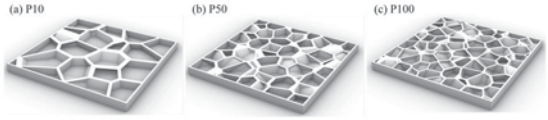
◆【吸音】共鳴器×多孔質材
➡低・中周波数帯域を広く吸音
➢ 共鳴器を単位格子とする微細周期構造




◆【遮音】フォノニックバンドギャップ
➡対象の周波数帯域を遮音
➢ 蝶の羽などにみられるジャイロイド構造




◆【拡散】非晶質構造
➡広帯域を多方向に拡散
➢ 3 次元ボロノイ分割を利用したランダム構造




## 2-5-9

### 2-5-9　音響メタマテリアル技術を用いた六角形型透明硬質吸音材の設計方針と特性評価

Design policy and evaluation of characteristics for a hard and transparent hexagonal acoustic panel using acoustic metamaterial technologies

○五味蔵酒，平良優大（ピクシーダストテクノロジーズ）

◆オフィス向けの吸音材として一般的には多孔質材が用いられるが，ガラスパーテーション等の透明部材や硬質な空間には適さないという問題がある。この問題の解決策として，我々は音響メタマテリアル技術を用いてガラスに貼り付け可能な六角形型透明硬質吸音材（iwasemi HX-$\alpha$, Pixie Dust Technologies 及び ITOKI）を開発した。

◆設計においては，会話周波数帯域のボリュームゾーンである 1/3 オクターブバンドの中心周波数 500 ~ 1000 Hz をターゲットとし，Fig1(a) の外形寸法の制約のもと複数の共鳴器の穴パラメータ（穴径，穴数，厚み）と背後空気層を最適化した。材質は透明性のためポリカーボネートを選択し，射出成形で作製した。

◆残響室法吸音率の測定では，設計帯域において平均吸音率 86.9%と意図したとおりの高い吸音率を実現できた（Fig.1(b)）。

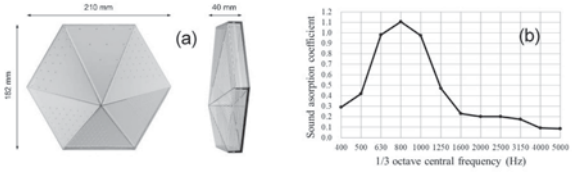



Fig.1 : (a) Design of iwasemi HX-$\alpha$ . (b) Measured reverberant sound absorption coefficient spectrum of the mass product iwasemi HX-$\alpha$.

## 2-5-10

### 2-5-10　透明な音響メタマテリアル吸音材を貼り付けたガラス張り会議室の音響特性と音声明瞭度の評価

Evaluation of acoustic properties and speech intelligibility in a glass-enclosed conference room with transparent acoustic metamaterial sound-absorber

○長谷芳樹，五味蔵酒，平良優大（ピクシーダストテクノロジーズ），
△林万由弓，△上野浩道（トヨタコネクティッド），△塩田久樹（ヒトバデザイン），
△畝忠孝，△辻未津高（ピクシーダストテクノロジーズ）

◆**目的:** 意匠性を重視した会議室などでは，オープンな雰囲気を演出可能なガラスで仕切ることなども多いが，それらの環境に充分な量の吸音材を配置すると意匠性が阻害されてしまう。そこで我々は，透明素材を用いた音響メタマテリアル吸音材をガラス張り会議室に貼り付けることで，その空間の意匠性を損なわずに会議室内外の音環境を向上する試みをおこなった。

◆**方法:** 透明なポリカーボネート製の音響メタマテリアル吸音材（iwasemi HX-$\alpha$, Pixie Dust Technologies 及び ITOKI）をガラス張り会議室の内側に計 128 個貼り付けた（Fig.1(a)）。（設計吸音帯域は約 440 Hz から 1120 Hz：吸音材の詳細は演題 2-5-9 参照）

◆**結果:** 透明吸音材を貼り付けた結果，室内の残響時間の大幅な低減が確認できた（Fig.1(b)）。これは，会議室内の意匠性（壁面アートなど）を損なわない状態で音環境を改善できたことを示している。
加えて，室外への漏れ出し音の 4 dB 程度の低減も確認できた。

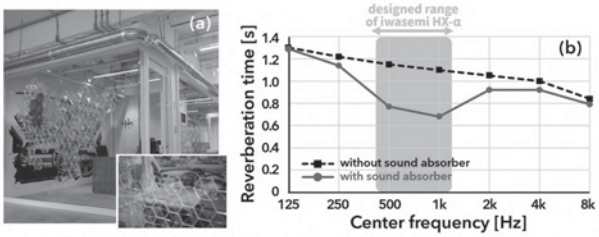



Fig.1: (a) The conference room equipped with the transparent sound absorbers *iwasemi* and (b) the reverberation time of the conference rooms.

## 2-5-11

### 2-5-11　パッチ型吸音構造の吸音特性（共鳴周波数について）

Resonant frequency of Patch-type sound absorbers

○眞田　明，△藤本望夢（岡山工技セ），△小野裕行（三乗工業）

◆多孔質吸音材料の表面に非通気のパッチを貼り付けることにより，薄い材料でも低周波数から比較的広い帯域で吸音効果が得られるパッチ型吸音構造を提案している。本研究では，吸音ピーク周波数の決定原理について，数値解析と理論的なモデルにより検討を行った。

◆パッチ型吸音構造の場合，ヘルムホルツ共鳴器と比較し，基材の厚さに対する共鳴周波数変化が小さいことが明らかになった。

◆パッチと反射壁面の間にできる矩形領域（Fig. 1 の斜線部）をスリットと見立て，スリットの開口端補正長の理論式を当てはめた結果，数値解析結果と基材厚さに対する共鳴周波数の変化の傾向が一致した。

◆パッチ型吸音構造の共鳴周波数が基材厚さにより変化する原因が開口端補正長にあることが明らかになった。

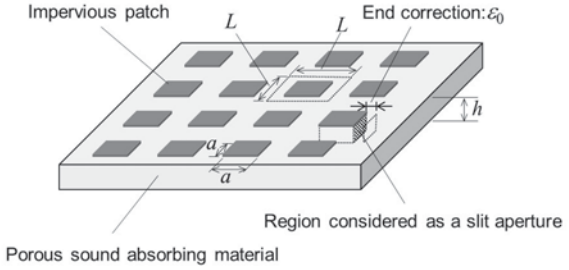



Fig.1 : Patch type sound absorber

## 2-5-12

### 2-5-12　矩形試験室における低周波域の音響透過損失測定に関する実験的検討

Experimental study on low-frequency measurement of sound transmission loss in rectangular test rooms

○西沢 啓子（都産技研），井上 尚久（前橋工大），佐久間 哲哉（東大・工）

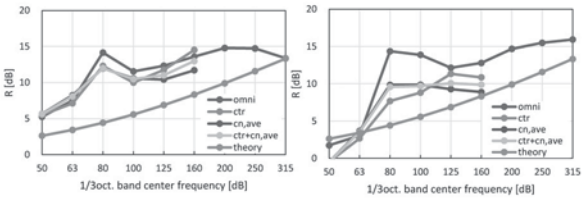
◆建築部材の遮音性能の実験室測定法は拡散音場を仮定した JIS A 1416 が規定されているが，低周波域では拡散音場の仮定が成立せず，現行規格では 100Hz 未満を適用対象外としている。

◆ISO 16283-1 では小規模室内かつ低周波領域の測定法（コーナー法）が示されているが，遮音実験室に適用した報告は見当たらない。

◆膜材料など遮音性能が低い試料では二室の独立を前提とした理論が成立せず，主に低周波域で音響透過損失の過大評価が生じている。

◆本報では，膜材料を対象に低周波域における従来法（JIS A 1416）とコーナー法（ISO 16283-1）による音響透過損失の実測値を比較した。

◆試料寸法の影響を検証するため，室間全面および調整壁設置状態の配置の異なる 3 カ所の小開口に試料を設置して同様の測定を行い，質量則との対応を検証した。




Sound reduction index estimated from the central and corner zones: (left) large sample, (right) small sample at bottom-right

## 2-5-13

### 2-5-13　乾式二重床構造における空気抜きの影響に関する理論的・数値解析的検討

Theoretical and Numerical study on effects of air gaps on dry-type double floor system

◎曹達（東大・工），會田祐（長谷工技研），
井上尚久（前橋工科大），佐久間哲哉（東大・工）

◆空気抜きの設置長さの影響に関する実験を行なった。空気抜きは，支持脚伝達力への影響は小さいが，空気層内音圧は大幅に低減される。

◆理論解析では，空気抜きをオリフィス絞りと想定し，Maxwell モデルによるモデル化を提案した。数値解析では，インピーダンス境界による空気抜きモデルの基礎的検討を行なった。

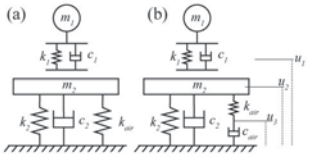



Fig. 1: Theoretical models: (a) w/o air gaps, (b) with air gaps

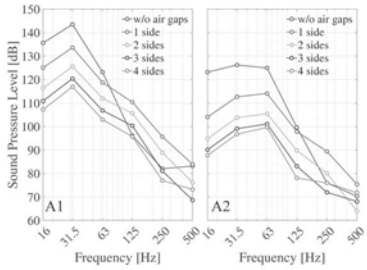



Fig. 2: Calculated sound pressure levels

## 2-5-14

### 2-5-14　壁式構造実験室における乾式二重床と畳の床衝撃音に関する検討 ―その 3 乾式二重床の違いによる床衝撃音レベル低減量―

Examination of floor impact sound of floating floor system and tatami in the reinforced concrete wall construction testing device – Part3 Reduction of transmitted impact sound due to differences in floating floor system –

○冨田隆太（日大・理工），△阿部今日子（日大・芸術）

◆筆者らは，子ども・高齢者にとって安全で快適な，かつ子どもの動作による床衝撃音の問題がクリアできる，「遮音性能と安全性に配慮した畳の開発及び設計指針の策定」に関する研究を行ってきた。

◆本報では，面材の異なる乾式二重床上に，畳を施工時の床衝撃音レベル低減量について述べる。

◆Fig.1 に畳の効果として，乾式二重床に対するレベル差を示す。図を見ると，面材の違いによらず，乾式二重床のみに比べて性能が向上していることがわかる。

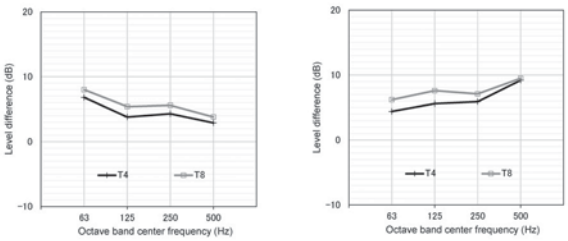



Fig. 1Level difference with respect to floating floor system (Heavy-weight impact sound) (Left: floating floor system 1, right: 2)

## 2-5-15

### 2-5-15　新型コロナウイルス感染症によるステイホーム中の音環境アンケート調査 ―2020 年と 2022 年の比較―

Sound environment questionnaire survey during stay home due to COVID-19 ―Comparison between 2020 and 2022―

○冨田隆太（日大・理工），△阿部今日子（日大・芸術）

◆筆者らは，はじめて緊急事態宣言が出され，解除された直後の 2020 年 6 月に共同住宅の音環境に関するアンケートを実施した結果を既報にて報告した。

◆本報では，2022 年 6 月に既報の回答者を対象とした追跡調査を行った結果を報告する。Fig. 1 に他の住戸や外からの音について聞こえる程度を質問した結果を示す。

◆2020 年に比べて，2022 年は「聞こえる～非常に聞こえる」などの指摘が減少している傾向が見られた。

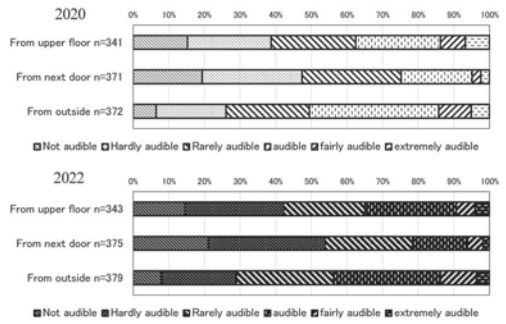



Fig. 1　Hearing level during stay home

## 2-6-1

### 2-6-1　ダミーヘッド(HATS)を用いた VHF 領域の耳介周りの音響計測(その 7)
### －計測環境改善と水平面・正中面の考察－

Acoustic measurement around the pinna in the VHF region using a dummy head –Part7, Improvement of measurement environment and discussions about HRTF on the horizontal and median planes.

☆春澤恒輝(神奈川工科大), 廣江正明(小林理研), 石井要次(積水化学工業), 神崎晶(東京医療センター), 上田麻理(神奈川工科大)

◆ 前報では, 音源に対する水平面内の頭の向きの影響を明らかにするために, TSP 信号を用いたインパルス計測を応用し, HATS の水平面内における VHF 領域の広帯域に亘る HRTF を求めた.

◆ しかし, 水平面全体でインパルス計測時に十分な SN 比を確保できていない. また, 正中面に対する HRTF も未だ把握できていない.

◆ そこで, 本研究では引き続き, VHF 領域の HRTF を求めるための計測環境の改善と, 水平面・正中面の HRTF の算出を行った.

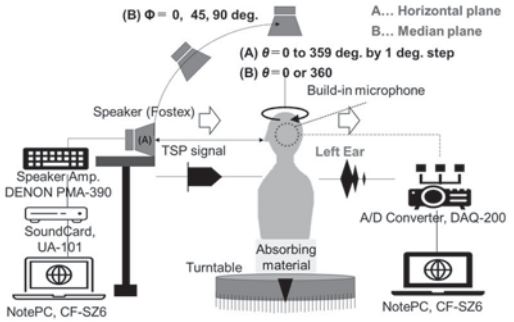



Fig.：Block diagram of measuring systems.

## 2-6-2

### 2-6-2　HATS を用いた聴覚保護具の遮音性能特性測定の試み

Attempt of sound attenuation measurement of protectors using HATS

○稲永 潔文(サザン音響)

◆作業環境における騒音暴露により聴力損失を防止するため、様々な遮音性能を持つ耳栓や聴覚保護具が用いられ、その性能特性の評価には Real Ear Attenuation at Threshold（以下 REAT)と呼ばれるヒトの聴覚閾値の差を測定する主観評価法が採用されている

◆一方で、近年ノイズキャンセル型ヘッドホンおよびイヤホン（以下NCHP）が普及し、これらの騒音低減量等の性能評価法が JEITA RC-8142A として規定されている。

◆従来、官能検査で行われてきた聴覚保護具の遮音性能特性測定を、HATS を用いた物理測定で行うことを試みた。

◆聴覚保護具には、ヘッドバンドタイプの 3M 社“3M™ H9A（H520A）”を、HATS にはサザン音響 “SAMAR Type 4500”を、また測定音場して、RC-8142A の規定を満たすサザン音響“小型疑似拡散 Box Type1020”を用いた。

Fig.1a Externals of HATS used for measurement without hearing protector & 1b with hearing protector

◆測定の結果、従来測定法と近似した結果が得られ、多くのメリットを有することが明らかになった。

◆一方、主観的な官能検査と物理測定の閾値の相違により生じていると思われる不整合な現象も明らかとなった。

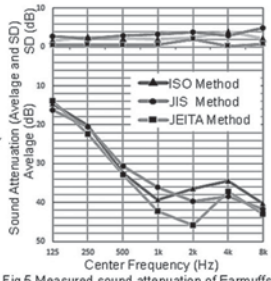



Fig.5 Measured sound attenuation of Earmuffs "3M™ PELTOR™ H9A(520A)" by JEITA RC-8142A and by current JIS and ISO standards.

## 2-6-3

### 2-6-3　オージオメータによる耳栓フィットテストの試み

Fit test for sound attenuation of earplugs applying audiometer

○横山 栄, 小林知尋(小林理研), 上松 葵, 中市健志(リオン)

現在、耳栓やイヤーマフなどの聴覚保護具の遮音性能測定法については JIS や ISO 等の規格で規定されているが、実際にユーザが聴覚保護具を着用した際の遮音性能は、着用状態や個人の外耳道の形状等が影響する。特に職場騒音による聴力損失防止の観点から、個人の遮音値を把握することが重要である。ANSI に準じた簡易測定器（3M™ E-A-Rfit™）等も利用可能であるが、決して安価な機器ではなく、騒音職場における健康診断に用いられるオージオメータ（RION, AA-K1A；耳覆い型イヤホン）を用いる手法を試みた。耳栓の遮音性能を測定した結果、本手法により、上述の ANSI に準じた機器と同等の性能が得られることがわかり、さらに騒音職場等への導入を視野に、測定時間の短縮、測定場所の暗騒音の影響等、引き続き、本手法の適用可能性を検討する予定である。

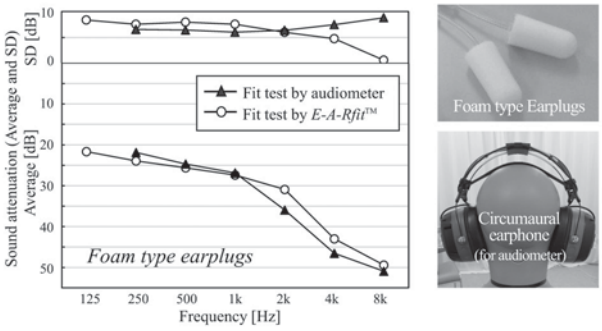



Fig.1：Sound attenuation of earplugs by fit test applying audiometer.

## 2-6-4

### 2-6-4　プレゼンテーションコンテスト「わたしのまちの音風景」開催の試みと展望

Trial and prospects for a presentation contest "Soundscape of my town"

○高野佐代子、土田義郎（金沢工大）

Sayoko TAKANO, and Yoshio TSUCHIDA (Kanazawa Institute of Technology).

◆ 中高生などにおける音響学への興味の展開を目指し、広報活動の一環としてプレゼンテーションコンテスト「わたしのまちの音風景」を企画し、実施したので報告する。

◆ 2022 年 11 月 14 日にオンライン形式でプレゼンテーションコンテストを実施した。審査員は対面会場に集合した。1 チーム 5 枚のスライド、写真以外にも動画、イラスト、加工を含む写真等とし、発表の制限時間は、4 分 30 秒以上 5 分以内とした。8 組の応募があり、金賞・銀賞・銅賞およびその他 2 件の個人賞が選ばれた。

**Figure 1. A part of the poster for presentation contest.**

## 2-6-5

### 厚木市における防災無線放送の長期観測 その6: ～風雨の音が及ぼす影響～

Long-term observation of disaster prevention radio broadcasting in Atsugi City Part 6: Effects of wind and rain sounds

☆仲村雄大，橋本卓己（神奈川工科大），
三浦貴大（産総研），
小川喜道，上田麻理（神奈川工科大）

前報では防災無線放送の季節の違いにおける聞き取り易さに及ぼす影響の解明を目的として，防災無線放送の音源を季節毎に分類し，1/3オクターブバンド分析による音響解析を行い，その結果，季節の違いによって周囲の音へ影響を及ぼし、防災無線放送の聞き取り易さにも影響している可能性を報告した.

本稿では,防災無線放送の聞き取り易さが風雨によってどのように変化するか，音響解析により調べることにした.

Figure: Image of listening to disaster prevention radio broadcasting

## 2-6-6

### 2-6-6 波動音響数値解析による屋外拡声システムの聴取性予測に向けた基礎的検討

Basic study on speech intelligibility prediction of outdoor public address system by wave-based computational acoustics.

○大嶋拓也（新潟大・工）

- 回折，地表面吸音，気象影響など屋外音響伝搬における諸現象を考慮した屋外拡声システムの聴取性予測には，発展性のある波動音響数値解析の適用が望まれる。
- 実在地域に基づく 380 m 四方の解析対象において，WE-FDTD 法により聴取性指標 $U_{50}$ を面的に算出し（Fig. 1），解析負荷に直結する時空間離散化幅の影響を検討した。
- 空間離散化幅0.05 m，0.1 m，0.2 mとしたCases 1–3のうち，1，2間は二乗平均平方根誤差1.30 dBで良好に対応した。1，3間は2.04 dBで，$y = x$からやや外れた（Fig. 2）。

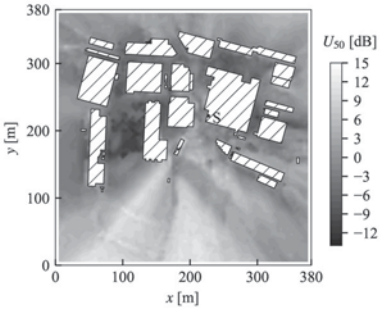



Fig. 1 Distribution of $U_{50}$ with source S for Case 1.

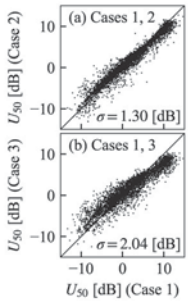



Fig. 2 Comparison of $U_{50}$.

## 2-6-7

### 2-6-7 日欧の騒音予測モデルの特性に関する検討 −2 種地表面上の半無限長厚み障壁問題における超過減衰量の比較−

Study on characteristics of Japanese and European road traffic noise models: Comparison of the excess attenuation by a semi-infinite length thick barrier on two types of ground surfaces

☆本間響（新潟大院・自然研），大嶋拓也（新潟大・工）

◆半無限長厚み障壁問題において，我が国の ASJ Model (ASJ), 欧州の Harmonoise (HN) 及び CNOSSOS-EU (CN) の等価騒音レベル$L_{\mathrm{Aeq}}$ [dB]を比較した (Fig. 1)。全モデルで$L_{\mathrm{Aeq}}$は，全地表面を柔らかい畑地としたケース (Case 1) より，音源側地表面を密粒舗装にした２種地表面のケース (Case 2) の方が大きくなった (梗概参照)。

◆Case 2 の障壁裏の予測点に対して，ASJ Model では，最も寄与の大きな音源から地表面超過減衰が生じ始める距離$r_c$ [m]が地表面上の伝搬距離$r$ [m]を上回るため (Fig. 2)，地表面超過減衰量が 0 dB となり (Fig. 3)，欧州 2 モデルと$L_{\mathrm{Aeq}}$の分布に差が生じた。

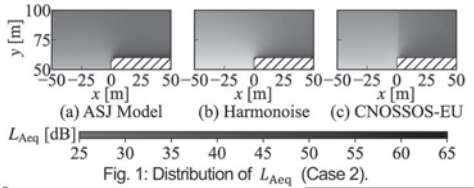



Fig. 1: Distribution of $L_{\mathrm{Aeq}}$ (Case 2).

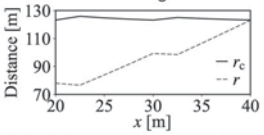



Fig. 2: Comparison of $r$ and $r_c$.

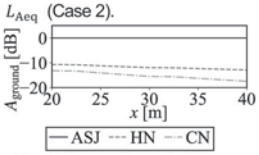



Fig. 3: Comparison of $A_{\mathrm{ground}}$.

## 2-6-8

### 2-6-8 保育器のためのリアクティブ型消音器の適用効果に関する実験的検討

−器内設備音に対するオリフィス形状の遮音効果 −

Experimental Investigation of Application Effects of Reactive-type Sound Absorber for Incubators: Insulation Effectiveness of Orifice Shape to Facilities Noise in Incubators.

☆奥 彩華，穴井 謙（福岡大）

◆保育器内の設備騒音の静音化のため，保育器の実機にオリフィス型消音器を設置し，適用効果を実験的に検討した。

◆器内設備音の周波数特性を測定し，315 Hz 帯と 1k Hz 帯が相対的に大きいことを把握した。遮音効果を見る際両者の周波数帯に着目した。

◆実験によって，オリフィス型消音器を設置することで遮音対象の周波数帯で遮音効果が得られることを確認した。315 Hz では 3.3 dB，1k Hz では 4.5 dB の効果が得られ，O.A.では 1.8 dB の低減量となった。(Fig.1)

◆消音器を取り付けるとノーマル状態と比較して僅かに温度上昇に時間を要するが，到達後は一定の温度を保ち，また，湿度も上限到達後一定の範囲を保った。空気の流れを大きく妨げずに音を低減した。

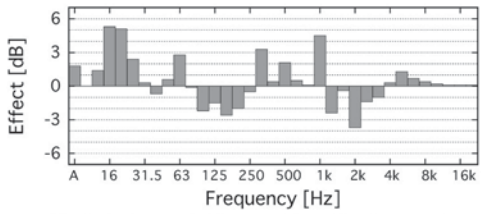



Fig. 1 : Insulation effectiveness of the orifice shape sound absorber.

## 2-6-9

### 2-6-9　オープンプラン型オフィスにおける音環境とワーカーの印象評価の関係 ーその１　調査結果の概要ー

Relationship between sound environments and worker's impression evaluation in open-plan offices -Part 1: Summary of investigation results-

☆荒木拓己，冨樫建五（近畿大院），原田和典（岡山県立大），
菅原彬子，△長澤康弘，平栗靖浩（近畿大），△岩切幸伸（コクヨ）

◆ 近年,ワーカー同士のコミュニケーションがとりやすいという点から,知的生産性向上に有効として,オープンプラン型オフィスが普及している.しかし既往研究で用いられる実験室の環境と実環境とでは乖離があるため,実環境での有効性がわかっていない.

◆ 本研究では実環境下での音環境とワーカーの印象評価の関係を明らかにすることを目的とする.

◆ 内容　筆者らがこれまでに開発した室内音環境計測システムを用いA社のオープンプラン型オフィスにおいてオフィス内の音環境とワーカーの印象評価を同時に調査する. 本報では,調査の概要および調査結果について報告する.

◆ 騒音レベルがある程度高い状態が生産性に有効であるという傾向がみられた.

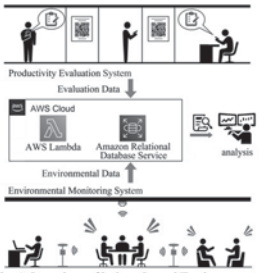



Fig. 1 Overview of indoor Sound Environment Measurement System

Fig. 2 Measuring scenery

## 2-6-10

### 2-6-10　オープンプラン型オフィスにおける音環境とワーカーの印象評価の関係 ―その２　１人作業と複数人作業による印象評価の違い―

Relationship between sound environments and worker's impression evaluation in open-plan offices　―Part 2　Differences in impression evaluation between one-person and multi-person work.

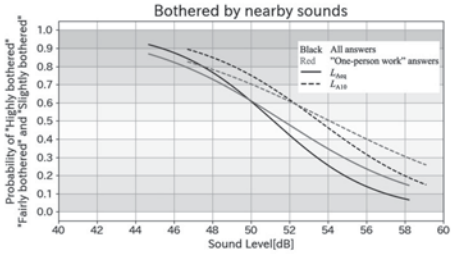
☆冨樫建五，荒木拓己（近畿大院），原田和典（岡山県立大），
菅原彬子，△長澤康弘，平栗靖浩（近畿大），△岩切幸伸（コクヨ）

◆背景　　前報では、オフィスの音環境とそこで働くワーカーの音環境に対する印象評価を同時計測するシステムの開発と、システムをＡ社オープンオフィスで稼働させ、実測調査し、その調査概要と単純集計結果について報告した。

◆目的　　本報では、実測調査のうち、作業種類（１人作業と複数人作業）による、オフィスの音環境とワーカーの印象評価の関係を明らかにする事を目的とする。

◆内容　　印象評価の回答を目的変数、音環境指標(A 特性等価騒音レベルなど)を説明変数として、ロジスティック回帰分析を行った。

◆結果　　作業種による影響は「周囲の音への気になりやすさ」で顕著に表れていることがわかった。




## 2-7-1

### 2-7-1　180 度指向性超音波スピーカの開発

Development of 180-degree directional ultrasonic speaker

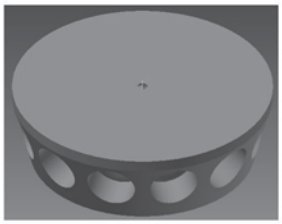
○今村俊樹，福士和義，池田賢二（セコム）

◆空中放射用途の超音波音源は，一般的に波長の短さを利用した指向性の狭いデバイスがほとんどで，広い空間をセンシングするには音源を複数個所に設置する必要があった。

◆今回，１台の超音波センサで８畳間程度の居室をセンシングできるよう，指向性の広い超音波スピーカを新たに開発した。

◆本スピーカは，入手しやすい汎用の超音波トランスデューサをベース音源とし，構造の工夫により，指向性の広さ（放射面側で均一）と，ベース音源と同等の正面音圧レベル[dB]を実現した。

◆FEM によるシミュレーションに基づき設計・開発したスピーカは，実測でもその効果が確認できた。



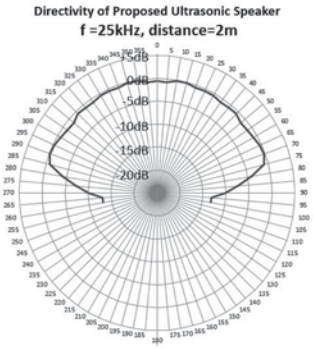



Fig.1 : appearance and directivity of proposed ultrasonic speaker

## 2-7-2

### 2-7-2　曲げ振動板間の乗り継ぎ特性

―近距離場音波浮揚による非接触ステッピング搬送(19)―

Transit characteristics between bending plates,
-Noncontact stepping transportation using near-field acoustic levitation(19)-

☆西雄太郎，青野浩平，孔徳卿，梶原秀一，青柳学（室蘭工大・院）

◆Fig.1 に振動板および振動モードを示す。振動板の構造は振動部と積層アクチュエータ(MPA)埋込部とそれぞれを繋ぐ結合部で構成されており，積層アクチュエータ(MAP)埋込部で曲げ応力を発生させ曲げ振動を振動部に励振する。

◆Fig. 2(a)，(b)は間隔１mm の解析結果と速度特性および搬送位置の実験結果である。Lx は浮揚物体が搬送先の振動板に重なる位置である。解析および実験から曲げモードで乗り継ぎ搬送が可能であることを示した。また，実験から浮揚物体の揺動が大きいことが確認されたため揺動を抑えるような構造を検討する必要がある。

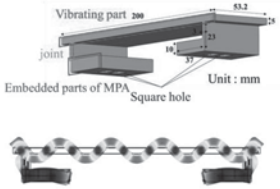



Fig.1 Vibration plate and bending vibration mode.

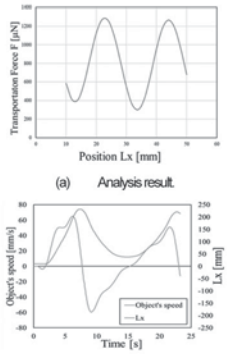



Fig.2 Transit transportation.

## 2-7-3

### 2-7-3 Evaluations of a lead-free BNBT15-BNM disc transducer for underwater acoustic propulsion system

☆Yimeng Wang[1], Deqing Kong[1], Yutaka Doshida[2], Minoru Kuribayashi Kurosawa[3],Manabu Aoyagi[1] ([1]Muroran Institute of Technology; [2]Ashikaga University; [3]Tokyo Institute of Technology)

◆The underwater acoustic propulsion system with a novel lead-free $(Bi_{0.5}Na_{0.5})_{0.85}Ba_{0.15}TiO_3$-$(Bi_{0.5}Na_{0.5}Mn)O_3$ (BNBT15-BNM) disc transducer at 4.45 MHz in thickness-vibration mode was designed and the underwater performances were investigated.

◆Zero-speed propulsion (ZSP) and no-load speed (NLS) were the two main metrics to evaluate the performance of the underwater propulsion system, and the results are shown in Fig. 1 and Fig. 2 respectively. When the driving voltage was 70 $V_{pp}$, ZSP and NLS could reach 3.2 mN and 194.98 mm/s, respectively.

◆The BNBT15-BNM transducer is expected to be applied in underwater propulsion applications like drug delivery systems in the near future.

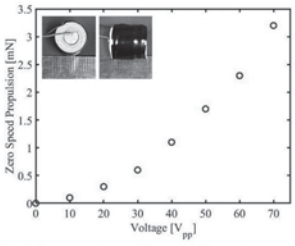



Fig.1 Zero-speed propulsion vs. driving voltage.

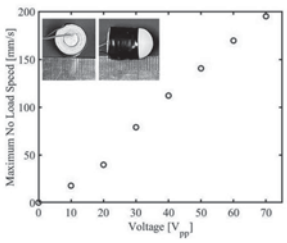



Fig. 2 Maximum no-load speed vs. driving voltage.

## 2-7-4

### 2-7-4 Design and experimental evaluation of a piezoelectric actuator with a wide driving frequency domain and high positioning resolution

☆Hao Yun, Deqing Kong, and Manabu Aoyagi (Muroran Institute of Technology)

◆In the advanced precise positioning field, the demand for piezoelectric actuators with the features of high accuracy, quick response, and large stroke is increasing.

◆We proposed a piezoelectric actuator that is capable of macro-micro-submicron motion with non-resonance and resonance drive, as shown in Fig. 1. A structure with multilayered piezoelectric elements and the compliant mechanism was designed and a prototype was fabricated.

◆As a result of experiments, the proposed actuator can realize the maximum velocity of 28.6 mm/s at a frequency of 23 kHz and a step displacement of 2.25 μm without backward motion. Moreover, the positioning resolution of 144 nm was obtained.

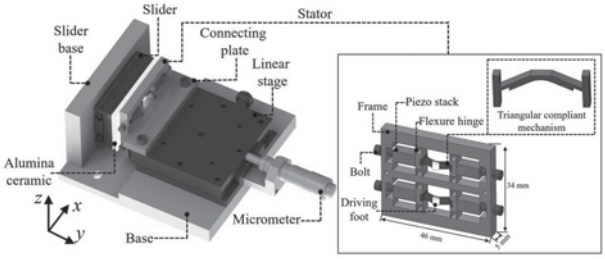



Fig. 1. Configuration of the proposed actuator.

## 2-7-5

### 2-7-5 超音波接合に用いる細長い接合チップを取り付けた複合振動源の特性

Characteristics of a complex vibration source attached with elongated welding tip used for ultrasonic welding

☆水野隼佑，淺見拓哉，三浦　光（日大・理工）

◆本研究では，超音波複合振動源の接合可能範囲を拡大するために，振動源に取り付ける細長い接合チップの寸法についてシミュレーション解析によって検討した。次に，製作した接合チップを超音波複合振動源に取り付けた場合の振動分布などについて測定した。

◆矩形断面の細長い接合チップの両端の変位の比が，ステップホーンの共振周波数を用いて片端を加振した場合の振動振幅の理論式から得た 1.414 倍となる寸法を，有限要素法解析にて検討した．その結果，長さ 100 mm の場合，断面は縦振動方向が 6.0 mm で，ねじり振動方向が 4.5 mm の時に上述の条件を満たすことがわかった．

◆両共振周波数における接合チップの振動変位分布を測定した．その結果が Fig.1 である。図より，ねじり振動方向のたわみ振動は 1 波長が約 51 mm で，接合チップ全体で約 2 波長の振動となっており，縦振動方向のたわみ振動は 1 波長が約 37 mm で，接合チップ全体で約 2.5 波長の振動となっていることが分かった．

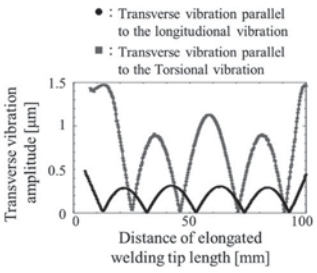



Fig. 1. Vibration distribution of elongated welding tip.

## 2-7-6

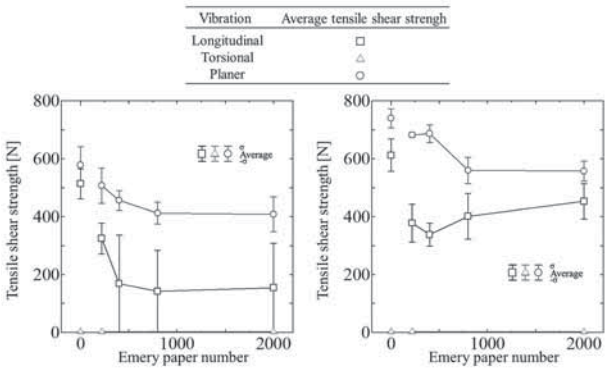
### 2-7-6 超音波複合振動源を用いた試料の表面粗さによる接合強度

Welding strength due to surface roughness of sample using ultrasonic complex vibration

☆山崎梨菜，淺見拓哉，三浦　光（日大・理工）

◆本検討では，試料の表面粗さが超音波接合に与える影響を検討するため，アルミニウム板及び銅板ともにエメリー紙で表面を粗くした場合について接合実験を行い，引張せん断試験の結果から接合強度の比較検討を行った。

◆図(a)は接合時間 0.2 s の場合の接合実験の結果，図(b)は接合時間 0.6 s の場合の接合実験の結果である。両図より，面状振動を用いた接合の引張せん断強度は，縦振動やねじり振動と比べて，表面粗さに関わらず高いことが分かった。




(a) at weld time of 0.2 s.　(b) at weld time of 0.6 s.

Tensile shear strength of aluminum and copper plates

## 2-7-7

### 2-7-7　直径 50 mm のボルト締めランジュバン型振動子を用いた超音波溶融による衝撃軽減特性

Impact-reduction characteristics by applying ultrasonic vibrations when a 50-mm-diameter bolt-clamped Langevin-type transducer was used

鈴木厚行，☆田中悠希，丸田 昂平，
藤川 大輝，木村 翼，則安 聡至（徳山高専）

超音波を金属や樹脂などに印加すると変形抵抗が低下する。本研究の目的は超音波の印加によって瞬間的に物体を軟らかくすることで，衝撃を軽減することである。直径 50 mm のボルト締めランジュバン型振動子 (BLT) 用いてホットメルト接着剤 (HMA) に超音波を印加し，衝撃力軽減特性などについて検討した結果，超音波の印加によって衝撃力が低減されることを確認した。測定した範囲では駆動電圧 80 $V_{rms}$，溶融試料の高さ 20 mm のときに低減率は最も高くなり，27% になった。

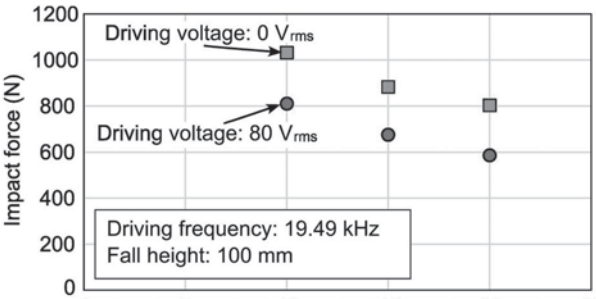



Fig.1：Relationships between sample height and impact force without and with the application of ultrasonic vibrations.

## 2-7-8

### 2-7-8　直径 50 mm のボルト締めランジュバン型振動子を用いた耐熱ナイロンの超音波溶着機の開発

Development of ultrasonic welding system for heat-resistant nylon using a 50-mm-diameter bolt-clamped Langevin-type transducer

☆則安聡至，鈴木厚行（徳山高専），橋本芳視（シーエス）

直径 50 mm のボルト締めランジュバン型振動子を用いた耐熱ナイロン用超音波溶着機を構成した。本研究では耐熱ナイロンの中でも溶着が難しい表面が毛焼きされたフェルトタイプの耐熱ナイロンの超音波溶着を図り、静荷重や溶着時間を変化させたときの溶着強度を測定した。Fig.1 に静荷重と溶着時間を変化させたときの引張強度試験の結果を示す。測定した範囲では最大で引張強度 1163 N が得られた。静荷重が 260 N を越えたあたりから溶着強度のばらつきが減ったようにみえる。溶着時間と溶着強度に大きな関係性は見いだせなかった。

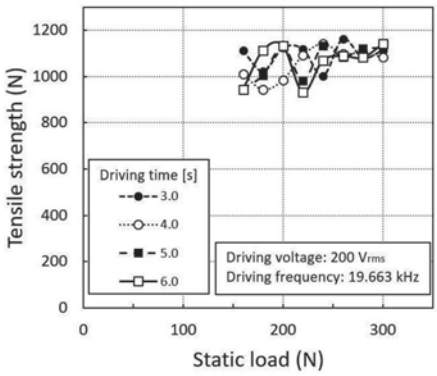



Fig.1：Relationships between static load and tensile strength.

## 2-7-9

### 2-7-9　面状振動する超音波振動源を駆動させるための発振器の開発 ー共振周波数追尾プログラムを設計するためのシミュレーションー

Development an oscillator for driving ultrasonic vibration source that vibrates in planar vibration locus - Simulation to design resonance frequency tracking program -

☆青木耀佑，淺見拓哉，三浦　光（日大・理工）

◆筆者らは，２つの周波数で同時入力することで面状振動を生成する超音波振動源によるワイヤボンディングの検討を行っている．本稿では，そのような振動源を駆動させるために必要な共振周波数追尾を行うための発振器に関して，シミュレーションを用いた検討を行った．

◆本検討では，超音波振動源は $F_H$ と $F_L$ の２つの共振周波数を持つ電気等価回路としてシミュレーションを行った．Fig. 1 は高い共振周波数 $F_H$ に対する入力電圧の周波数 $f_H$ の割合 $f_H/F_H$ を示した結果である．図より，$f_H/F_H$ は約 7 ms でほぼ 1.00 に一定となり，共振周波数追尾ができることが分かった．また，低い共振周波数 $F_L$ においても，同様に約 3 ms で共振周波数追尾ができることが分かった．

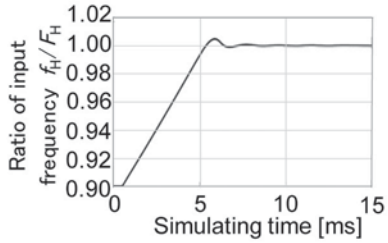



Fig.1. Ratio of input voltage frequency to resonance frequency $f_H/F_H$.

## 2-7-10

### 2-7-10　水中定在波音場における水出し珈琲の抽出促進 ―入力電力との関係―

Promotion of extraction of cold brew coffee in the underwater standing wave field -Relationship with input electric power-

☆渡辺啓介，淺見拓哉，三浦　光（日大・理工）

◆周波数 20 kHz の水中定在波音場を用いて，水出し珈琲抽出中に超音波を照射した場合について，入力電力を変化させた場合の成分抽出効果の違いについて検討した．

◆ 成分抽出量をペーパークロマトグラフィーの結果から画像処理を施し，簡易的に可視化した．

◆Fig. 1 に一例として，超音波なしの場合と，入力電力 2 W，照射時間 300 s の場合について，Green 要素と Blue 要素の間の散布図を示した．

◆RGB 要素の散布図だけでは傾向が分からないため，RGB の各要素の平均値を求め，超音波なしとの差を比較した．

◆その結果，抽出量は入力電力 2～3 W，照射時間 600 s の場合に多い傾向が見られた．

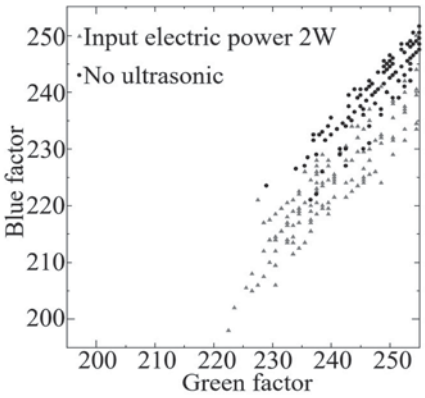



Fig.1. Scatter plot of Green-Blue factor

## 2-7-11

### 2-7-11　２台の空中強力超音波源を用いた濡れた布の乾燥 —布の大きさを変えた検討—

Drying of wet cloth using two powerful aerial ultrasonic sources
-Study of changing the cloth size-

☆大和田知花，伊藤美桜，淺見拓哉，三浦　光（日大・理工）

◆筆者らは，空中強力超音波源を用いた濡れた布の乾燥について検討を行っている。本稿では濡れた布の大きさをいくつか変えた場合の乾燥効果について検討を行った。

◆２台の空中超音波源の振動板面が平行かつ向かい合うように設置し，板間に強力な定在波音場を形成した。定在波音場内の音圧の節の位置に濡れた布を設置し，その布の大きさを変えて測定を行った。Fig. 1 は乾燥実験の結果である。図より，乾量基準含水率は時間の経過に伴って減少していることがわかる。また，Sample A, B, C の順に布の大きさを大きくすると乾燥時間が増加し，単位面積当たりの乾燥時間は短くなることがわかった。

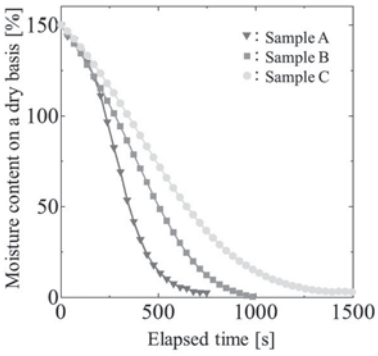



Fig. 1. Moisture content on a dry basis during ultrasound irradiation. Input power is 30 W.

## 2-7-12

### 2-7-12　たわみ振動板型超音波源の振動板に直接滴下した水の微粒化

Atomization of water dropped directly on the transverse vibrating plate type ultrasonic source

☆五十嵐隆一，淺見拓哉，三浦　光（日大・理工）

◆本研究では，たわみ振動板型超音波源を用いて振動板上に直接滴下した水を微粒化するための装置を作製した。そして超音波源の入力電力を変化させた時の微粒化量について検討した。

◆Fig.1 は微粒化実験の結果で，入力電力を 0 W から 25 W まで 5 W 刻みで変化させた場合の微粒化量との関係をまとめたものである。図より，微粒化量は超音波源への入力電力が大きくなるほど多くなることがわかる。また，超音波照射中に滴下した水が 150 ml であることから，25 W でほとんど微粒化できていることもわかる.

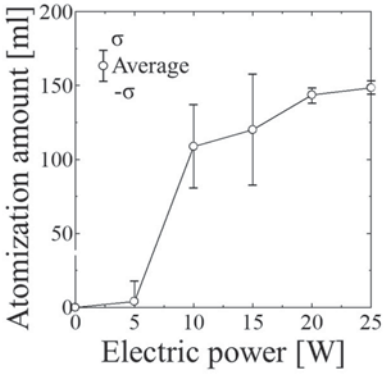



Fig.1. Relationship between electric power and atomization amount

## 2-7-13

### 2-7-13　非線形空中超音波励起による薄板からの漏洩波挙動

Behavior of leaky wave from thin plate excited by nonlinear airborne ultrasounds

☆濱田郁哉，清水鏡介（日大院・理工），大隅歩，伊藤洋一（日大・理工）

◆非線形空中超音波と受信器アレイを用いた金属薄板状構造物の高速非接触非破壊検査について研究を行っている.

◆本検討では，受信器アレイ面が欠陥イメージングに及ぼす影響について，実験的に検討を行った.

◆その結果，受信器アレイ面のイメージング結果への影響はほぼないことを確認した.

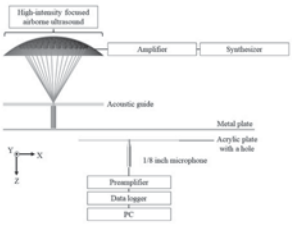



Fig. 1 Schematic of experimental equipment.

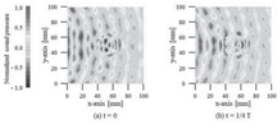


Fig. 2 Experiment results.

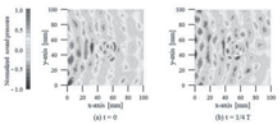


Fig. 3 Experiment results with acrylic plate.

## 2-7-14

### 2-7-14　振動面近傍に配置した小径穴付き円筒からの噴出流の流入出部分離による流量計測

Measurement of the flow rate of the jet generated from a small hole in a cylinder close to a vibrating surface by separating inflow and outflow part

○青野浩平，孔德卿，松本大樹，青柳学 (室蘭工大・院)

◆Fig.1 のように振動面近傍に小径穴付き円筒を配置することにより噴出流を発生させ，噴出流をアクリルパイプに流入させることでパイプ出口の流速を流速センサーにより測定した。

◆Fig.2 の流速分布から平均流速を算出し，Fig.3 のように中心流速に対する平均流速の近似式を求めた。流量は平均流速とパイプ断面積の積により求めた。

◆流量は振動振幅増加および振動面－円筒間ギャップが減少することにより増加し，最大で 28L/min に達した。(Fig.4)

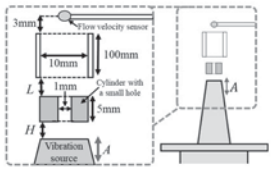



Fig.1 Measurement method of flow velocity.

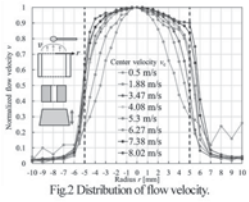


Fig.2 Distribution of flow velocity.

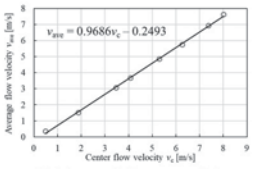



Fig.3 Average velocity vs. center velocity.

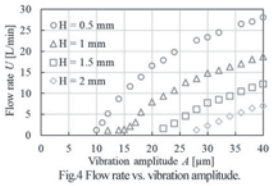



Fig.4 Flow rate vs. vibration amplitude.

## 2-7-15

### 2-7-15　近距離場音波浮揚する円柱に設けた小径穴への吸着および円柱の回転現象

Adsorption to a small-diameter hole in a near-field acoustic levitating cylinder and rotation phenomenon of the cylinder

☆唐木澤直樹，青野浩平，孔徳卿，青柳学（室蘭工大・院）

◆Fig. 1 のように振動子と土台の間に小径穴を有する円柱を配置し，音圧分布解析および噴出流の確認を行った。

◆Fig. 2 に音圧分布解析結果を示す。側面の穴の内部で強い音圧が確認できた。

◆Fig. 3 に小径穴への吸着現象を示す。噴出流は確認できず，浮揚する穴付き円柱側面の穴に平棒を近づけると穴への吸着が確認された。

◆Fig. 4 に円柱の回転現象を示す。側面の穴に近づけた発泡スチロール球は穴に吸着され，円柱に時計回りの回転が生じた。

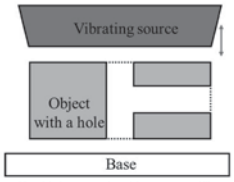



Fig. 1 Experiment and analysis model.

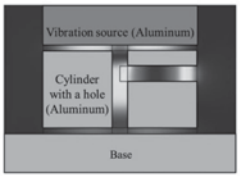



Fig. 2 Analysis result of sound pressure distribution.

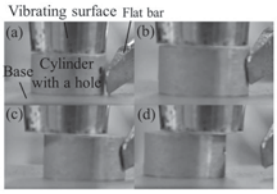



Fig. 3 Absorption phenomenon. (a)Contact. (b)Absorption. (c)Pulling. (d)Separation and swinging.

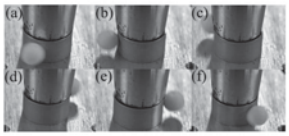



Fig. 4 Rotation phenomenon.

## 2-7-16

### 2-7-16　定在波音場中における反射を考慮した鉄球の荷重の変化

Changes in the load of an iron ball considering reflection in a standing wave sound field

☆桶田大晟（愛工大），畑中信一（電通大），安井久一（産総研），鎌倉友男（電通大），豊田昌弘（本多電子），小塚晃透（愛工大）

◆超音波振動子を複数取り付けた音源を 2 つ向かい合わせに配置し、定在波音場をつくる。その際に超音波の反射が起こらないよう、音源に様々な角度を付け、鉄球にどのような荷重の変化が起こるのか評価を行った。

◆Rayleigh の式を用いた音圧シミュレーションを行い、実験結果で得られたデータを基に考察を行う。

◆Fig.1 に各音軸の交差角度を変化させた場合の鉄球の荷重の実験結果を示す。音軸の交差角度が 170° から 155° まで 5° 間隔で垂直方向の力を計測したところ徐々に荷重が減少し、155° で 0 g となった。超音波の反射により荷重が変化すると考えたが、音源ユニットに角度をつけた分だけ斜め方向に力が働き、垂直方向の力が水平方向に分散された。これにより水平方向の計測が不可能である電子天秤では正しい値の計測はできない。

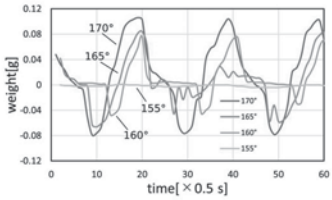



Fig.1 Time variation of the load on an iron ball for crossing sound axis angles of 170° ,165° ,160° and 155° at frequencies of 40 kHz and 39.999 kHz.

## 2-7-17

### 2-7-17　空中超音波によるチップ部品の浮揚に関する粘性音場解析

Viscoacoustic analysis on electronic chip parts levitation using airborne ultrasound

○和田有司，中村健太郎（東工大・未来研）

◆超音波浮揚による電子部品や精密部品、薬剤の非接触搬送は部品損傷や清潔性、歩留まりの回避の面で利点がある。

◆これまでチップ部品の浮揚について 20～40 kHz の振動系では小さなチップが捕捉できず、高周波の 70～100 kHz の振動系では保持可能であるという実験事実を得ている。

◆非粘性の線形音場解析で求めた音響放射力では上記を説明できない。

◆浮揚物の周囲に発生する Schlichting 音響流の影響を考慮するために熱粘性音響解析によりレイノルズ応力を算出した。

◆Fig. 1 に示す軸対称解析モデルに半波長モードの定在波を励振して、チップ表面に粘着境界条件を設定して、音響放射力を算出する。

◆Fig. 2 に計算されたチップサイズと音響放射力と重力の比の関係を示す。チップサイズ a が境界層圧 $\delta=(\nu/\omega)^{1/2}$ の 20 倍を下回ると、音波の振幅に関わらず浮揚困難となることが明らかになった。

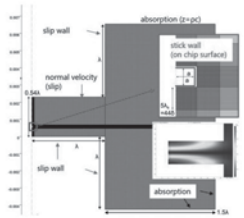



Fig.1 : Axisymmetrical model for ultrasonic levitation of a chip

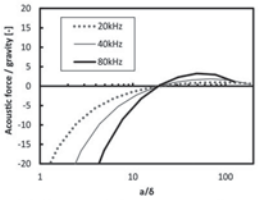



Fig.2 : Acoustic force / gravity dependence on chip size

## 2-8-1

### 2-8-1　日本語の語中破裂音発話における母語話者による有声・無声対立および中国語を母語とする日本語学習者との比較

Japanese voicing contrast in word-medial plosives produced by native speakers and comparison with Chinese-native Japanese learners

○李瑋昱（上智大），守本真帆（上智大・学振），溝口愛（前橋工科大），荒井隆行（上智大）

◆日本語母語話者と中国語を母語とする日本語学習者を対象に、語中の日本語破裂音の録音実験を行い、母語話者と学習者の発話の時間的特徴を調べた。

◆日本語母語話者２名と学習者２名が発話した/ata/と/ada/について報告する。

◆母語話者は有声と無声破裂音で、閉鎖区間の持続時間を優先して各部分が発話全体における割合と先行母音に続く声帯振動の停止するタイミングを調整していた。

◆学習者は母語の影響で破裂後の帯気の有無と声帯振動が開始するタイミングを優先して調整していた。

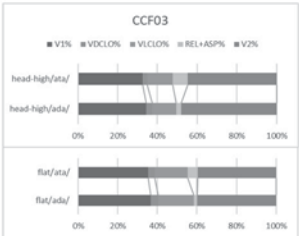



Fig.1 : Percentage of duration of utterances by native Japanese speakers

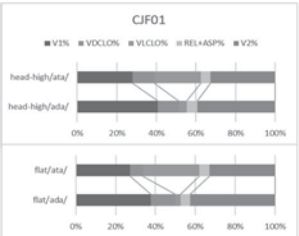



Fig.2 : Percentage of duration of utterances by Chinese-native Japanese learners

## 2-8-2

### 2-8-2　日本人学習者と母語話者による中国語第三声の生成

Comparative study of Mandarin Tone 3 production between native speakers and Japanese learners

○呉琪（神戸大）

中国語第三声の発音は学習者にとって困難であると多数指摘されているが，日本人学習者の第三声の生成に関する研究は，誤用の調査を行ったものが多く，音響的特徴と発声様式を考察したものは僅かである。そこで，本研究では，F0 に加え，声質の音響関連量である H1*-H2*にも焦点を当て，日本人学習者と母語話者による中国語第三声の相違点を比較した。その結果，F0 に関して，母語話者の F0 曲線は下降上昇調を示すのに対し，日本人学習者による第三声（誤用率が 100%のグループ）の F0 曲線は上昇調を示す。また，母語話者の場合，音節の後半から F0 が上昇する傾向が見られた。さらに，女性話者において，母語話者の F0 は日本人学習者より低く，且つ狭いピッチレンジで実現される。H1*-H2*に関して，女性話者の場合，母語話者の H1*-H2*は日本人学習者より低く，H1*-H2*の最小値は音節の後半に現れた。このことから，女性母語話者の第三声では，きしみ声の程度が学習者より強く，きしみ声が音節の後半で生じる可能性があると考えられる。

## 2-8-3

### 2-8-3　中国語を母語とする日本語学習者による促音の発話における時間的特徴 —アクセント型による違い—

Durational characteristics of Japanese geminates produced by speakers of Standard Chinese: The effect of pitch accent type

○守本真帆（上智大・学振），李瑋昱（上智大），溝口愛（前橋工科大），荒井隆行（上智大）

◆中国語を母語とする日本語学習者と日本語母語話者による発話実験を行い，異なるアクセント型における促音および非促音の発話における持続時間について調べた。

◆学習者と母語話者 5 名ずつによる HL 型・LH 型の/hata/および/hat:a/の発話について報告する。

◆学習者にとって，HL 型よりも LH 型の方が子音長の対立の習得が困難である可能性が示唆された。

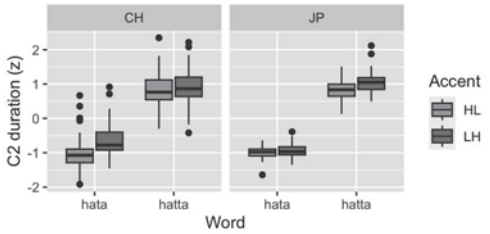



Fig. 1 Consonant duration (z-score) of singleton and geminate /t/ for each pitch-accent pattern, produced by native speakers of Standard Chinese and native speakers of Japanese.

## 2-8-4

### 2-8-4　中国語母音 a, o, e, i, u に見られる二重母音性の観察—日本語母音との比較

Are Chinese vowels a, o, e, i, u monophthong or diphthong? - A comparison to Japanese vowels

☆孫静（神戸大院），林良子（神戸大）

◆日本語（東京方言）には「ア，イ，ウ，エ，オ」の 5 つの単母音があるとされている。一方，中国語の単母音数については諸説がある。本稿では中国語母語話者を対象として，ピンイン一字で書かれる母音がどのように発音される母音なのかについて，日本語の単母音と比較しながら検討を行なう。

◆中国語の a, i, u, y は日本語の 5 母音のようにフォルマントが安定していて，音声的に単母音として扱ってよいと考えられる。o と e は母音発話中の舌の位置が動き続けており，二重母音性を持っている。

◆中国人学習者は日本語の/a, i/を中国語の a, i で代用して発音する傾向があり，/u/を中国語の[u]に近い音で発音する様子も観察された。ただし，中国人学習者が/a, i/を中国語の a, i で代用して発音してもコミュニケーション上の支障は特に生じない程度の母語干渉であると考えられる。

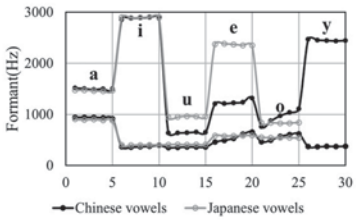



Fig. 1　Mean formant values of Japanese and Chinese vowels

◆以上をまとめると，音声的実現から見ると，中国語の単母音は a, i, u, y の 4 つで，o, e は二重母音と呼べる音であることが分かった。

## 2-8-5

### 2-8-5　シラブルを単位とした発音空間のエントロピーに着眼した母語・非母語発音分析

Entropy-based analysis of L1/L2 syllabic productions using an acoustic word embedding space

☆勝瑞雄介，峯松信明，齋藤大輔（東大・工），中西のりこ（神戸学院大）

◆本研究では日本語・米国人英語・日本人英語の音節単位での発音差異を，音響単語埋め込みの技術を用いて分析した.

◆英語の音素数 > 日本語の音素数　であり，発声の単位も，日本語のモーラに対し英語の音節の方が複雑な構造である.

◆そのため英語の方が発音空間での多様性が多いと考えられる.

◆本研究では教師なし音声分割で得られた音節音声に 対して，以下の２つの指標を用いて発音多様性の定量化を目指した.

- ➢ 発音エントロピー（提案手法）
  - ✧ 埋め込み空間上に GMM を作成し，各ガウス分布を事象と捉えエントロピーを算出する
- ➢ 散布疎度 [Kyo,1994]
  - ✧ 全データ対の埋め込み空間上での距離をデータ数で正規化

◆いずれの指標でも，英語 > 日本語，米国人英語 > 日本人英語 という傾向が見られた.

## 2-8-6

### 2-8-6　顔の動きをフィードバックする発話訓練支援システムの検討

Vocal exercise support system using feedback of facial movement.

☆設楽 郁巳，△関 和弘，北村 達也(甲南大学)

◆本研究では発話改善のため，発話中の顔ランドマークの動きの大きさをリアルタイムに表示するシステムを開発し，予備的な評価実験を行った.

◆ 画面中央にはWebカメラで撮影した利用者の顔が表示され画面右上部にランドマークの動きの大きさを表すメーターが表示される. 利用者はこのメーターが高い値を示し続けるよう意識しながら発話訓練を行う.

◆ 評価実験を行い，主観的な訓練効果について分析を行った. 実験参加者はシステムを使用した場合, 使用しない場合と比べ上手く音読ができたという自覚を持つ傾向が見られた.

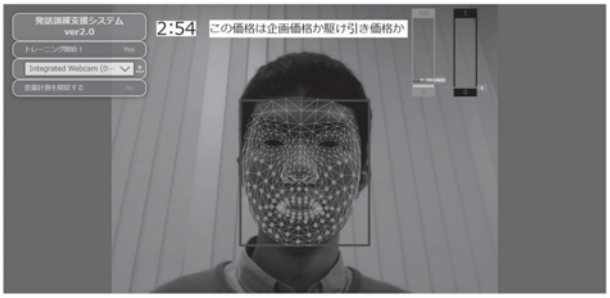



Fig.1：Screenshot of the proposed vocal exercise support system.

## 2-8-7

### 2-8-7　Using Curriculum Learning to Optimize the Response of Spoken Dialog System

☆ Xuecheng Niu, Akinori Ito, and Takashi Nose (Tohoku Univ. , Grad. Eng.)

◆Training a dialogue strategy by reinforcement learning (RL), which usually requires collecting experiences via direct interaction with real users, needs a costly data acquisition procedure. Fortunately, the user simulator has been applied to mimic real users, and thus the training becomes affordable. A user simulator chooses a user goal randomly and communicates with the dialogue agent to generate a simulated experience. Nonetheless, one criticism of this random sampling strategy is that it defies the laws of human learning. Therefore, this research sought to explore the effects of different schedules on agent's behavior and propose a rational one. The current experiments show that training dialogue agents using specific schedules affects the success rate. Moreover, a new metric, referred to as the matching rate, assesses the agent's behavior.

## 2-8-8

### 2-8-8　CALL システムによる異なる字幕呈示がシャドーイング練習に及ぼす影響に関する研究

A study on the effect of different caption presentations by a CALL system on shadowing practice.

☆方漱玉，伊藤彰則，能勢隆（東北大学）

◆「シャドーイング（Shadowing）」は言語学習時に，音声を聞きながらそれをまねして発音する訓練法であり，聞こえてくる音声のすぐ後ろを影（shadow）のように追いかけるのがポイントである。

◆先行研究で，リスニング練習に対して，部分字幕は字幕無しと完全字幕より有効であると示されている。

◆字幕のリアルタイムでサポートできる利点を維持しながら，字幕に集中しすぎという問題を避けるため，部分字幕を見ながらシャドーイングする方法を提案する。

◆提案した練習システムを構築して，部分字幕(P)，完全字幕(F)，字幕無し(N)三つ条件で実験を行った。結果によると，3 つのグループの被験者の能力は最初は同じであったが，30 日間の練習により，P グループが N グループと F グループより高い能力になったことが示された。

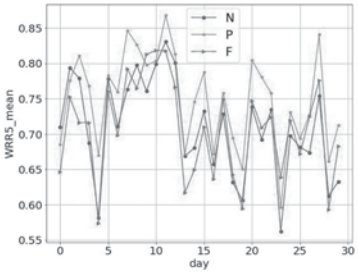



Fig.1：30-day change in shadowing ability.

## 2-8-9

### 2-8-9　real-time MRI 調音運動データからの end-to-end 音声合成

End-to-End Speech Synthesis from rtMRI Articulatory Movement Data

☆大谷祐人，澤田隼，大村英史，桂田浩一（東京理科大）

real-time MRI (rtMRI)で収録した調音運動データから音声を合成する end-to-end 型の深層学習モデルを提案する．rtMRI 動画像を入力とする音声合成の従来手法では声道形状パラメータのみを推定しており，基本周波数などの情報は別途与えている．そこで本研究では EfficientNetV2 を用いた CNN-BiLSTM モデルを用いて rtMRI からメルスペクトログラムを推定した後に，HiFi-GAN ボコーダを用いることで，end-to-end で音声合成を行う．提案モデルの概略を Fig. 1 に示す.

実験では単一男性話者による ATR 音素バランス 503 文の朗読を rtMRI で収録したデータセットを用いて，合成音声の音質と基本周波数の推定精度を評価した．音質の主観評価に近い評価指標である PESQ は 1.64 と十分に高いスコアを得られなかった．一方，基本周波数 $F_0$ の RMSE は 26.7 Hz と良好であり，rtMRI から基本周波数を含む音響パラメータ全体の推定が可能であることを確認した．今後は，rtMRI から $F_0$ を良好に推定できた理由を検証する予定である.

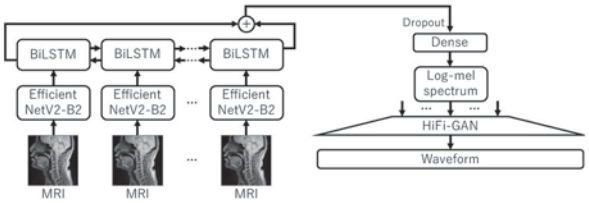



Fig. 1: Overview of the proposed model for speech synthesis from rtMRI

## 2-8-10

### 2-8-10　声帯物理モデルを用いた非定常流下の声門内圧測定実験と理論解析

Theoretical analysis of intra-glottal pressure measured from physical model of the vocal fold under unsteady flow

☆井上卓哉，中村和史，△中川栞里，徳田功（立命館大）

◆アクリル製 M5 型声帯物理モデル（Fig.1）を用いて、声門内圧計測実験および理論解析を行った。

◆非定常流を考慮したベルヌーイの定理により、声門内圧力分布を理論計算し、定常流の場合と比較した。

◆定常流下と非定常流下における声門内圧の差は、声門下圧の一次および二次の時間微分と、声門位置などに依存することが分かった。

◆声門下圧の一次の時間微分への依存性については、実験データでも同様の傾向が見られた。それ以外の依存性については、実験データと部分的に整合した。

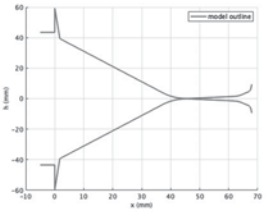



Fig. 1 Configuration of M5 model

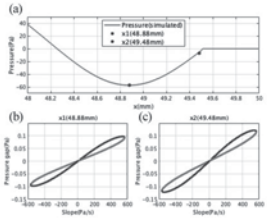



Fig. 2 (a) Simulated intra-glottal pressure distribution, (b, c) Dependence of pressure difference between steady and unsteady flows on slope of sub-glottal pressure at x1 and x2.

## 2-8-11

### 2-8-11　頭頸部の正中矢状面における rtMRI 動画への歯列補填

Tooth superimposition on rtMRI videos in the midsagittal plane of the head and neck region

☆常盤朔也，△一岡昴輝，△大谷幸聖，竹本浩典（千葉工大），前川喜久雄（国語研）

◆われわれは「リアルタイム MRI 調音運動データベース」を構築中で，これを用いて調音運動の定量分析を行っている。

◆しかし，MRI では歯列と口腔が同じ輝度値（黒）であるため歯茎摩擦音の分析などで問題となり，動画に歯列を補填することとなった。

◆本研究ではその手法を考案し，補填した歯列が周囲にはみ出した面積の割合（誤差面積率）を用いて精度を検討した。

◆1490 動画すべてに補填したところ，目視では良好な結果となった。また，誤差面積率を検討した結果，誤差は最大でも平行移動・回転移動の成分としてそれぞれ１ピクセル・１度程度であると推定された。

Fig.1：歯列補填した MRI 動画フレーム

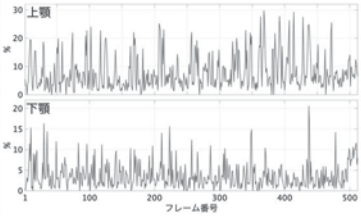



Fig.2：誤差面積率の変動

## 2-8-12

### 2-8-12　日本語標準語話者の正中面における子音/s/の声道形状の分析

Analysis of the vocal tract shape of consonant /s/ in the median plane for standard Japanese speakers

☆藤澤流以，天野沢海，竹本浩典（千葉工大），前川喜久雄（国語研）

◆われわれは「リアルタイム MRI 調音運動データベース」を構築中で，現在 22 名の話者データを試験公開中である。そしてデータベースの動画から調音器官の輪郭を抽出する研究を行ってきた。

◆本研究では，日本語標準語話者４名の声道形状を分析し，子音/s/の声道形状の変異を確認した。また，先行母音と後続母音のどちらがより強く子音/s/に影響するのかをそれぞれの子音/s/の調音器官の形状のばらつきを比較して検討した。

◆その結果，/sa/，/si/，/su/，/se/，/so/のうち/si/の/s/のみ形状が異なっていることが確認できた。また，先行母音より後続母音が同一の場合に/s/の調音器官の形状のばらつきが小さいことから，後続母音の方が/s/の形状に大きな影響を与えていることが明らかになった（Fig. 1, 2）。

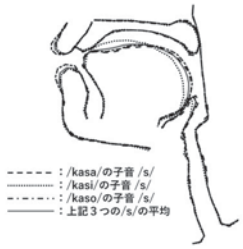



Fig.1：先行母音が/a/の場合の子音/s/の平均形状

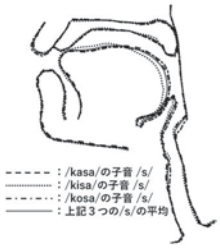



Fig.2：後続母音が/a/の場合の子音/s/の平均形状

## 2-8-13

### 2-8-13　日本語母語話者による英語母音知覚について：音響的類似性および語彙知識の影響

Perception of English vowels by Japanese listeners: Acoustic similarity and effects of lexical knowledge

◎渡丸嘉菜子（東京理科大）

◆一般的に、外国語音声は正確な判別や同定が困難である。とりわけ、日本語母語話者の英語音声知覚については、日本語母語話者は英語の/r/と/l/の判別及び同定が困難である、など広く知られた例も多い。

◆外国語音声をどのように知覚するか、については、幅広く議論されている。その中でも、Best & Tyler (2007)は、外国語音声と母語音声の調音的類似度を比較して知覚していると仮定し、モデル化している。

◆このモデルは、外国語音声知覚において、母語音声との調音的類似度を査定することを仮定している。本報告では、調音の類似以外にも知覚同化に影響を与える要因として、音響的類似と母語の語彙情報を挙げ、先行研究(Tomaru, 2009)を再検証しながら論じる。

◆先行研究(Tomaru, 2009)では、日本語母音話者による英語母音の知覚同化に、日本語の語彙知識の影響があることを論じている。しかし、音響的類似度を検証していないなど、不十分な点も多く、知覚同化における母語の語彙情報の影響の有無について不明な点も残る。

◆本報告では、英語母音と日本語母音を第１および第２フォルマントの二次元で比較し、音響的類似度を調査した。その上で、外国語音声と母語音声とが音響的に十分に類似していない場合、母語の語彙情報にもとづいて知覚同化が行われる可能性を示唆した。

◆本報告での分析結果をもとに、外国語音声知覚における母語の語彙情報の影響と、その度合いについて考察する。

## 2-8-14

### 2-8-14 残響環境下における合成音声/ba/-/wa/の識別ーフォルマント遷移に焦点を当ててー

Identification of /ba/-/wa/ in reverberant environments with synthesized continuum by changing formant transitions.

☆川井友子(上智大), 大澤恵里(昭和大),
溝口愛(前橋工科大), 荒井隆行(上智大)

◆残響環境下では，先行する母音や子音が後続する母音や子音をマスクするoverlap-maskingや，各音素内で時間波形が崩れエネルギー分布が不鮮明になるself-maskingなどにより，音声知覚が影響を受ける。

◆本研究では，フォルマント遷移の持続時間が連続的に変化する合成音声刺激/ba/-/wa/を用い，フォルマント遷移の変化に伴い残響環境下においてどのような影響が音声知覚に生じるか調査した。

◆フォルマント遷移に関しては，持続時間が30～70 msへと5 msずつ変化し，全部で9ステップ用意した。これに対して，残響環境下と非残響環境下の両条件で，聞こえた音声刺激が /ba/ あるいは /wa/のどちらであるか回答を求める識別課題を行った。

◆残響環境下では非残響環境下と比較して，より短い遷移時間の刺激ステップで/wa/の応答率が増加した。

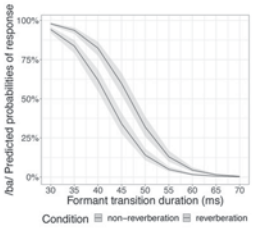



Fig. 1 Predicted probability of /ba/ (Left line : reverberation Right line : non-reverberation )

## 2-8-15

### 2-8-15　国立民族学博物館での特別展における声道模型展示とワークショップ

Physical models of the human vocal tract at a special exhibition of the National Museum of Ethnology, Japan and a related workshop activity

○荒井隆行(上智大・理工)

◆国立民族学博物館で開催された特別展「しゃべるヒト」において，声道模型を中心とする展示を担当させていただいた.

◆また，展示に合わせて開催されたワークショップにて，声道模型の工作を含むアクティビティーを行う機会を得た.

◆展示では，同様のふいご機構に異なる3パターンの声道模型セットを組み合わせた.

◆3つめの声道模型セットでは，Fig. 1の5母音の模型が使われた.

◆ワークショップでは，風船やペットボトル，ストロー，お菓子用の筒などを用い，3つのアクティビティーを実施．好評であった.

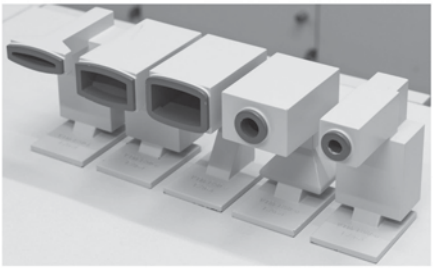


Fig. 1: VTM-B50 (with lips) of /i/, /e/, /a/, /o/, and /u/ (from left to right)..

## 2-8-16

### 2-8-16　発話訓練経験による文章発話時の顔ランドマーク変位の違い

Effects of voice training experiences on facial landmark movements during reading sentences.

☆安田奈央, 北村達也(甲南大)

◆発話時の調音運動と顔の動きの間には関連がある事が知られている。発話中の顔ランドマークの動きと舌運動の相互作用や、Clear Speechと通常の発話における顔の動きに差異があることも示されてきた。

◆本研究では、発話訓練経験の異なる話者が文章を読み上げた際の顔ランドマークの動きの差異を調査し、発話訓練への応用を検討する。顔ランドマーク座標の取得はMediaPipe Face Meshで行い、468点の顔ランドマークの座標から、各点1秒当たりの移動距離を求めた。

◆Fig.1に「北風と太陽」を読み上げた時の(a)プロのナレーター、(b)発話訓練経験のある大学生、(c)発話訓練経験のない大学生の各顔ランドマークのXY平面上の平均変位を示す。発話訓練の有無や発話スタイルの違いにより、発話時の口唇周辺、下顎、頬の領域に差異が表れ、(a)(b)は(c)と比較して口唇及び下顎の動きが大きいことが確認された。

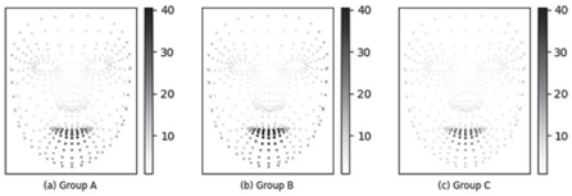



Fig.1 : Mean displacement of each facial landmarks on the XY-plane for Groups A, B, and C.

## 2-9-1

### 2-9-1　音楽家と音楽音響の接点

A relation between Musician and Musical Acoustic Research

○三浦 雅展（国立音楽大学）

◆音楽音響分野の研究が音楽家にとって有益な成果を上げることが期待されるものの，これまで積極的なコラボレーションは見られない。

◆音楽家に有用かもしれない音楽音響研究について紹介している。

◆ピアノ単音に含まれる倍音系列を１つ１つ分離し，再合成するデモ音が構築されている。音楽家にとって１つ１つの倍音を実際に聞くという体験は実際の音楽現場では容易ではない。

◆モーションキャプチャシステムを用いた演奏動作の可視化，演奏時の空気流出に関する様子についても説明している。オンライン会議システムを通した演奏音の送信における音質変化とその印象の関係についても説明している。

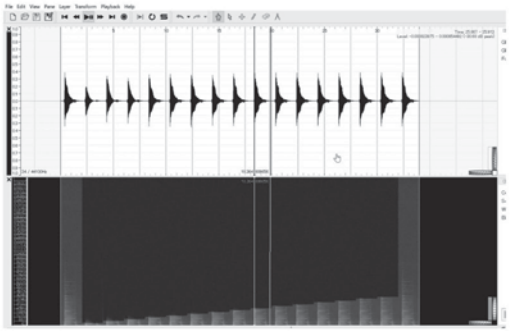


Fig.1 : Harmonic components' effect to the perception of piano sound

## 2-9-2

### 2-9-2　音楽大学学生による音響研究事例の紹介

Research Examples in Acoustics by Students in Faculty of Music

○丸井淳史（東京藝大）

◆音楽実技系あるいは実技を重視する学術系の専攻に所属する学生が，自身が専門とする楽器や興味のある分野の音楽を音響的に研究したいときに，どのように調査を行うのか，個人的な経験の事例を紹介する。

◆1 例目：3D プリンタを使用したレプリカ尺八の製作と評価
- ➢ 尺八専攻修士２年生
- ➢ 録音音源の一対比較と多次元尺度法による音色印象の定量化
- ➢ 半構造化インタビューによる音色印象・吹鳴感の質的調査

◆2 例目：クラリネットのリガチャの素材の違いが吹奏感に及ぼす影響
- ➢ 音楽学専攻学部４年生
- ➢ 自由演奏による語彙抽出　→評価尺度の作成
- ➢ 評価尺度を用いた評定　→因子分析

◆音楽大学には，学生自身が長く演奏を学び楽器と演奏を深く理解していること，同じ楽器を専門的に学んでいる者が近くにいて実験への参加依頼をしやすい（そして的確な回答を得やすい）環境にいること，などの特徴があり，それを活かした研究が行われている。

## 2-9-3

### 2-9-3　音響を学ぶ学生への音楽教育上の工夫と課題

How to teach music to students studying acoustics

○川上央，駒澤大介，三戸勇気（日大芸術）

◆ 楽器演奏においては、その響きが広がる空間は芸術性を高めるものとして重要であるが、マイクロホンで録音をした音楽の空間性を作り出すことが 20 世紀以降に新たに出現した。

◆ 音響を学習する学生は、楽譜が対象ではなく、音になった音楽を対象としている。そのためには演奏された音楽、つまりは音に対しての経験値を積むことが重要となる。

◆音楽学科では、200 人程度のホール(Fig. 1)と、小規模なスタジオ(Fig. 2)、無響室に近い施設(Fig. 3)がある。これらの 3 つの施設を利用して楽器の録音をし、残響がいかに楽器の音色の美しさに影響を及ぼすかという体験をさせている。

◆幼少期から音楽の専門的な学習を行っていない学生にとっては、大学では音楽理論や演奏を学ぶことに意識が向きがちであるが、音響を専門とする学生に一番必要なことは、音楽とは空間を伴った作品であるということである。楽器構造やホール形状など、楽譜からは読み取れない情報を体験学習することによって、その能力は向上する

Fig.1 Music Hall

Fig.2 Recording Studio

Fig.3 Anechoic Chamber

## 2-9-4

### 2-9-4　リアルタイム MRI を用いたソプラノ歌手の高音発声に関する研究

Research on singing at high pitches for a soprano using real-time MRI

☆戸田菜月，竹本浩典（千葉工大），高橋純（大阪芸大）

◆オペラ歌唱は高音でも厚みのある歌声が特徴である。これを実現するために，例えばソプラノ歌手は高音域で基本周波数<=声道の第 1 共鳴周波数を保つように声道形状の特殊な制御を行っていることが知られているが，その詳細は明らかではない。

◆本研究では，リアルタイム MRI で撮像した歌唱中のソプラノ歌手の声道形状を輪郭抽出し，主成分分析することで，高音で音高を変化させて発声する際，声道形状をどのように制御しているのか検討した。

◆E♭5 (622 Hz)，F5 (698 Hz)，G5 (783 Hz)，A♭5 (830 Hz)，B♭5 (932Hz) の順で上行した後，折り返して下行を行う課題を行った。すると，上行では，上を向いて口が開き，咽頭が広がり，頸椎が後弯し，喉頭が上昇した。一方，下行では口は閉じていくが，それ以外の部位はほとんど動かないことから，上行と下行では声道の制御が異なっていた。(Fig. 1・Fig. 2)。これは，音の上行では声道形状と声帯張力を，下行では声帯張力のみを制御していることを示唆すると思われる。

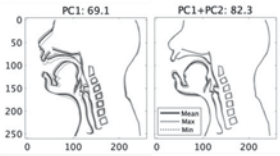



Fig.1：Results of principal component analysis of vocal tract shape

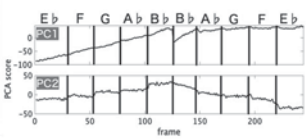



Fig.2：Principal component analysis score

## 2-9-5

### 2-9-5　音声分析と体内運動の観測に基づく歌唱指導の可能性

Possibility of singing instruction based on voice analysis and observation of internal body movements

◎高橋純（大阪芸大），戸田菜月，竹本浩典（千葉工大）

◆クラシック音楽において，優れた歌手の歌声とはどのようなものかという問いに対して明確な答えはない。

◆スポーツの分野では，卓越した身体技能のメカニズムがスポーツ科学によって探究され，怪我や故障のリスクを避けながら効率のよい訓練方法が提案されている。

◆近年，科学技術の発展により歌声の音響的特徴の解析や，歌唱時における体内運動の観測が可能となり，「歌唱行動」の実態を少しずつ明らかにすることが可能となってきている。

◆本招待講演では，今日の歌唱指導の現状を踏まえ，これまで筆者らが行ってきた研究について概説し，音声の分析とリアルタイム MRI 動画を用いた体内運動の観測による客観的な資料に基づく歌唱指導の可能性について検討し提案する。

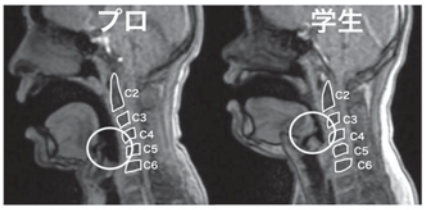



Fig.1：プロと学生の歌唱時の声道の比較

## 2-9-6

### 2-9-6　音響設計を学ぶ学生への音楽教育上の工夫と課題

Ingenuity and challenges in music education for students studying acoustic design

○西田紘子，城一裕，△ゼミソン ダリル，△宮崎由紀子，△管谷怜子，山内勝也，河原一彦（九州大・芸工）

◆九州大学芸術工学部音響設計コースの主に音文化学分野を対象に，音響設計家育成の場における音楽系教育の工夫と課題を紹介する。

◆これにより，音楽大学や芸術大学における音楽音響教育とは異なる立場から，音楽音響教育の実践例を提供する。

◆九州芸術工科大学設立時より音響設計コースで実施されている音楽の理論・演奏・創作に関する授業の方法，他分野との授業連携について概観し，音大等とは異なる教授方法を分析した。

◆理論では実践重視，ピアノ・レッスンでは選曲，作曲では文化相対主義，音響芸術では技術方法論の重視の点で工夫がみられた。

◆工学的・科学的思考や音響の学生特有の感性に沿った工夫がなされている反面，芸術的思考や制作実践の修得をめぐる課題が抽出された。

Fig.1：Screenshot of live streaming the recital in the subject "music theory and expression" at the recording studio of acoustic research center.

## 2-9-7

### 2-9-7　音楽家に音楽音響との接点は要らない —その理解なしに音楽家にはなれないから—

Do we really have to reconsider the point of contact between musicians and musical acoustics? Music is science, so why aren]t musicians scientist

○津崎　実

◆*the art or science of combining vocal or instrumental sounds (or both) to produce beauty of form, harmony, and expression of emotion*

◆上記は英英辞典における music の定義である。

◆日本では何故かこの science の側面が欠けて「オンガク」が一般的にも理解されている様子が窺われる。

◆著者は 20 年弱を芸術大学の音楽学部で指導してきた経験からも，この科学とは別世界の「オンガク」観が音楽の演奏家並びにそれを目指す学習者にあることを実感している。

◆音楽には娯楽としての側面があることは事実であり，科学的視点の欠如は演奏者やエンタテナーとしての評価を左右するべきではない一方で，高等教育として音楽を学ぶ者，または教える者には科学的視点の欠如は致命的であると敢えて主張したい。

◆この発表は日本における現状を嘆くことを目的とはしていない。我々はまだまだ music を学び取る発展途上段階にあるに過ぎない。

◆これから先，音楽を真に musical なものにしていくための方向性を示しつつ，各参加者の率先した関与を呼びかけたい。

## 2-10-1

### 2-10-1　音声変形による雑音残響環境での音声了解度向上

○赤木　正人（北陸先端科学技術大学院大学）

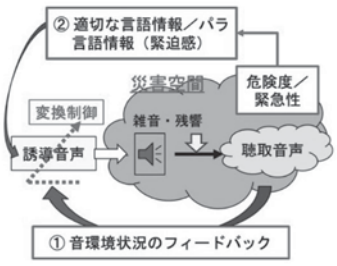
◆緊急時に安全な場所への適切な避難誘導を行うことは，被害の最小化を図る観点から重要である。この中でも，音声アナウンスによる避難誘導は，(1) 誘導サインを視覚で確認できなくとも有効，(2) 一度に多くの人に誘導指示を与えられる，ことから各所で使われている。

◆一方で，誘導音声がわかりづらい，聞き取りにくいとの声がある。

◆本研究の目的は，避難誘導時に，１）災害空間の音環境および避難者の状況がどうであろうと避難誘導音声がはっきり聞こえる、２）避難誘導音声から災害の危険性を察知し逃げる気になる，ような避難誘導音声を呈示するための音声システムを構築することである。

◆戦略：雑音残響を見越して呈示音声をあらかじめ変形することにより，より明瞭な声，通りやすい声として呈示する。雑音残響中でも明瞭な声，通りやすい声となるような音声を生成するために，雑音残響環境下の音声コミュニケーションにおいてヒトが意識的に／無意識のうちに行っている優れた能力（ロンバード音声，Clear Speech など）に着目し，これを模擬する。

◆講演では，雑音残響環境下で音声了解度を高めるための二種類の手法を紹介する。




## 2-10-2

### 2-10-2　効果的な避難呼びかけに関する考察

A study on effective evacuation instructions

○小笠原奈保美（群馬県立女子大・国際コミ）

◆水災害時に自治体が発令する避難指示文の効果的なあり方を追求するために，統語構造と呼びかけの口調に着目して研究を行った。多くの自治体の防災マニュアルで見られるような敬体と「避難してください」という依頼口調を用い，重文・複文・関係詞節等の複雑な統語構造を持つ「標準型」の避難指示文と，単純な統語構造と「避難せよ」という命令口調を用いた「単文・命令型」の避難指示文の２種類を作成した。オンライン調査により 20 代から 60 代までの男女 400 名にどちらか一方のアナウンスを聞いてもらい，内容の理解度の測定と「わかりやすさ」「危険が想像できたか」「危機感が高まったか」「避難の緊急性が感じられたか」の４項目の評価を行った。

◆内容の理解度については「標準型」と「単文・命令型」の間に有意差は見られなかったが，「避難の緊急性」についての評価では，「単文・命令型」の方が有意に高く，特に女性にその傾向が強いことがわかった。さらに，両方の避難指示文のスクリプトを読んでもらい，どちらがより「切迫感」があるか聞いたところ，男性・女性ともに「単文・命令型」の方を選んだ割合が高かった。ここでも男性より女性の方にその割合が高かった。

## 2-10-3

### 2-10-3　防災行政無線と災害情報伝達手段の多様化について

Disaster prevention administrative radio and diversification of disaster information transmission means

○天白成一（アルカディア）

◆近年、自然災害が各地で頻発しており災害情報伝達手段としての防災行政無線の役割が益々重要になってきている。

◆本稿では、防災行政無線の種別について解説するとともに、最近の傾向についても分類し説明する。

◆また、防災行政無線の抱えている問題点などについても議論するとともに今後の展望についても述べたい。

## 2-10-4

### （招待講演）避難を促すアナウンスの訓練システムに関する考え方と展望

Concept and prospects of an announcement training system to direct evacuation

○土田義郎，高野佐代子（金沢工大）

防災放送訓練システムの考え方

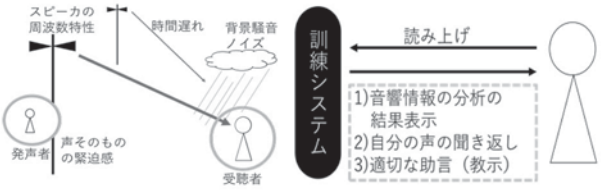
防災放送のアナウンス訓練システムとしては２つの段階がある。状況を理解するための段階と、実際に発声を適切に変える訓練のための段階である。

訓練システムの実装（＝TRANGIS＋防災ボイストレーナー）

放送の聞こえ方を模擬し、状況を理解させるシステムと、緊迫感をあたえる読み方となるような助言を与えて実際に訓練するシステムを融合したシステムを開発した。

今後の展望

助言をただ単に文字情報として与えるだけではなく、音声やVTuberのようなアバターがノンバーバル情報も含めて伝達するようなシステムとすることを予定している。




## 2-10-5

### 2-10-5　声の高さ・速さ・スペクトルを内挿した音声モーフィングによる避難放送の研究

A study of evacuation broadcasts by interpolated speech morphing; voice pitch, and speed

Sayoko TAKANO, and Yoshio TSUCHIDA (Kanazawa Institute of Technology).

○高野佐代子、土田義郎（金沢工大）

◆避難放送にける逃げる気になる音声「訴えるように」とそうでない音声「優しく」の２種類の音声を用いて、逃げる気になる防災放送の特徴を調べる。

◆音声分析の結果、基本周波数が「訴えるように」は「優しく」より高く、また第一フォルマントは「訴えるように」は舌が上方に位置し、また「優しく」では舌が下がるという傾向が見られた。

◆ ２ 種類の音声モーフィングの結果、声の高さ、速さ、スペクトルの順にすべてにおいて主効果において 1%で有意差が見られた。交互作用は見られなかった。

◆今後、逃げる気になる避難放送における発話運動の計測の必要がある.

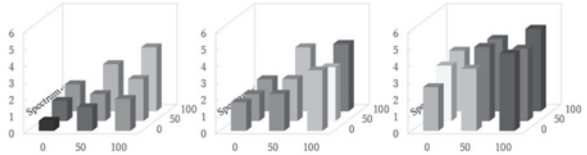


Fig. The visual feedback of the spoken sounds

## 2-10-6

### 2-10-6　地形及び気象が屋外拡声音の伝搬特性に及ぼす影響 ―無人観測ステーションを用いた長期観測によるケーススタディー

Effects of topography and weather on the propagation characteristics of outdoor acoustic mass-notification systems: A case study of long-term observation using unmanned observation stations

○佐藤逸人（神戸大院・工学研），栗栖清浩，松石遼太（TOA）

◆防災用屋外拡声システムによる拡声音の音圧レベルは，風に代表される外乱要因の影響でばらつきが大きくなることが考えられる。このばらつきについて調査するために，無人観測ステーションを構築して長期観測を行った。

◆谷間（Site 1）と海岸沿い（Site 2）の２つの地域で取得したデータを分析した結果，基準点と受音点の距離が長くなるほど，音の減衰量のばらつきは大きくなった（Fig. 1）。これは地形および気象の影響が積算されるためだと考えられる。

◆また，距離減衰と空気吸収による減衰以外の減衰量と風向・風速の関係についても分析し，ある程度の対応が見られることを確認した。

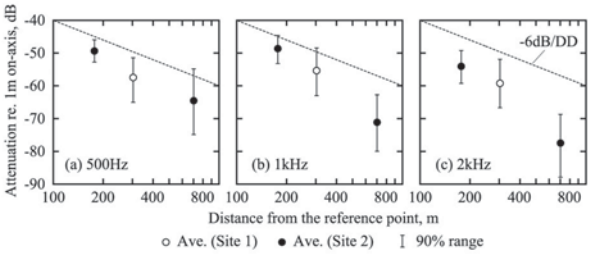



Fig.1:Sound attenuation relative to 1 m on-axis of the loudspeaker for each receiving point and each octave band.

## 2-11-1

### 2-11-1　自然音によるサウンドマスキングが記憶課題に与える影響

Effects of natural masking sounds on performance on a linguistic short-term memory task

○浅川香，細谷耕佑（三菱電機），亀川徹，丸井淳史（東京藝大），
陳薇雅，小川勇（三菱電機）

◆近年，働き方の多様化とともに，オフィス空間の環境の変動が大きくなり，音がどのような影響を及ぼすかは不明点が多い。

◆本研究では，他者の会話音声がある作業環境に適したマスキング音を明らかにするため，複数種類の自然音によるサウンドマスキングが単語記憶課題に与える影響を検討した。

◆実験では，会議音声とマスキング音（波の音，小川の流れ，森の中の鳥の鳴き声，カフェテリアノイズ，風の音）を聞きながら表示される４つの単語が，その後表示される単語と同じか否かを回答した。

◆実験の結果，小川の音（SM3）において正答率が高い傾向があった。個人差が大きかったためクラスタ分析したところ，音声の有無の影響（BN：暗騒音 vs SP：音声のみ）が大きい参加者（クラスタ 1）とそうでない参加者（クラスタ 2）の存在が示された（Fig1）。

◆サウンドマスキングにあたっては，これらの個人特性を考慮する必要がある。

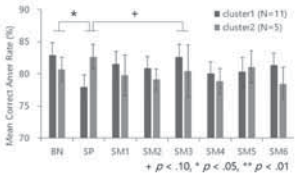



Fig1. Mean correct answer rate by clusters.

## 2-11-2

### 2-11-2　脳血流に基づく集中力向上のための音声情報マスキング効果の把握

Effect of masking speech information for improving concentration level based on cerebral blood flow

☆松本大祐（中央大院），山口雅夫（東芝テック），戸井武司（中央大）

◆ビジネスシーンや勉学など，知的生産性を高める環境創りが求められている。そこで本研究では，非侵襲的な機能的近赤外分光分析法(fNIRS)を用いた脳血流に基づく音環境の違いによる集中力評価を行い，どのような音環境で集中力が向上するか検討する。

◆唾液アミラーゼ及び脳血流を用いて，パープルノイズ，ホワイトノイズ，ブラウンノイズのストレス評価を行い，唾液アミラーゼの増加，Fig. 1 に示す右背外側前頭前野の領域での脳血流の減少から，パープルノイズにおいて最もストレスを感じていることがわかった。

◆唾液アミラーゼは各時点での離散的な評価となるが，脳血流では連続的な評価が可能であり，ストレスの時間的変化の評価に有用である可能性を示した。

◆Fig. 2 より音声情報をマスキングした際にウェルニッケ野の領域で脳血流反応の低下が起きる可能性を示した。マスキングにより音声情報が聞こえづらくなった際に集中力が向上することが考えられる。

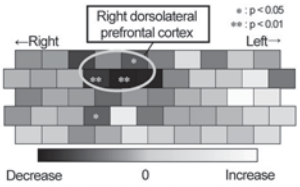



Fig. 1　Cerebral blood flow of purple noise relative to brown noise

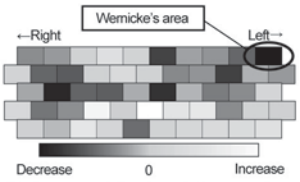



Fig. 2　Cerebral blood flow during presenting sound for masking of speech

## 2-11-3

### 聴覚刺激の時間変化に伴う知的集中の変化に関する調査

Investigation of time-series intellectual concentration associated with temporal changes in auditory stimuli.

○伊藤 弘章，野口 賢一，千葉 大将，加古 達也（NTT），
上田 樹美，下田 宏（京大）

◆聴覚刺激を伴う知的集中変化を定量かつ正確に測定

◆知的集中の低下を誘引した音の特徴を定量値として抽出可能か、について調査

◆非定常な音の発生直後に知的集中の急激な変化を確認
→聴覚刺激が要因となり、知的集中が低下する

◆ラウドネスを用いることで、知的集中の急激な変化を誘引する音の変化検出可能性を有することを確認

◆ラウドネスの突発的な変化がないことが、知的集中を低下させない条件であることも確認

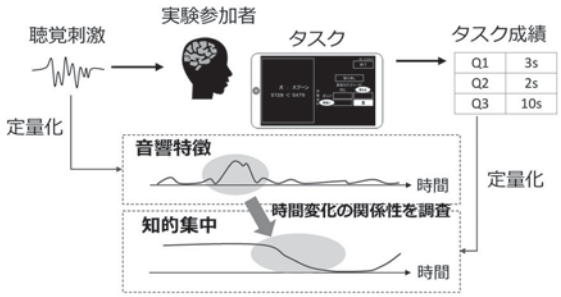



Fig.1：本調査の全体像

## 2-11-4

### 2-11-4　呼吸の同調効果を活用したエアコン動作音による知的生産性向上

Improvement of intellectual productivity by air conditioner operating sound utilizing respiratory synchronization effect

☆石川恭輔（中央大院），△LeeJangwoo，△KiSunghyun，
△LeeJuyoun（LG Electronics），戸井武司（中央大）

◆本研究では，振幅変動音の変調周波数や変調度を変更することで，呼吸統制と呼吸同調による知的生産性向上について検討する。

◆呼吸数の変化が知的生産性に影響を与えることを示すとともに，呼吸統制と呼吸同調の比較より，呼吸同調に有用性があることを示した。

◆振幅変動音の変調周波数を速くすると覚醒感は増加し，快適感は減少する傾向が見られた。一方，変調度を高くすると覚醒感は増加するが，快適感は一様に増加しない傾向を確認した。また，音源に関わらず変調周波数よりも変調度の方が覚醒感に与える影響が大きいことを見出した。

◆Fig. 1 に示す難易度に応じた各タスクの変調度の影響と，変調周波数の影響を把握した。各難易度における効果的な変調周波数や変調度が異なることを見出し，呼吸同調による振幅変動音を用いた知的生産性向上の有用性を示した。

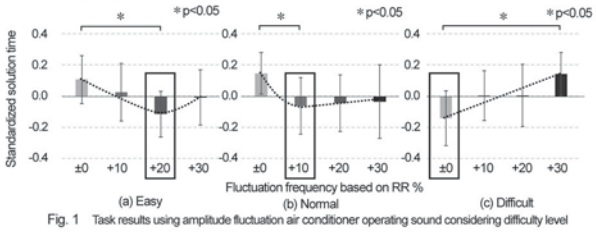



Fig. 1　Task results using amplitude fluctuation air conditioner operating sound considering difficulty level

## 2-11-5

### 2-11-5 保育施設周辺の音環境に対する施設運営者の意識

Questionnaire survey on attitude of administrators toward the sound environment around childcare facilities

○吉富佑樹，△川相裕吾（九州大・芸工），
片岡寛子(電通大)，高田正幸(九州大・芸工)

◆福岡市内の保育施設を対象とした活動状況に関するアンケート調査，保育施設園庭での音響測定を行い，保育施設で生じている音の現状や寄せられる苦情の傾向などについて分析を行なった。

◆苦情では「交通」「音」に関するものが全体の６割以上を占めた（Fig.1）。音に関する苦情では，子どもの声や行事の際に生じがちな音が元となるものが多く見られた（Fig.2）。

◆多くの保育施設は，屋外活動時や音響機器使用時の近隣への音の影響，地域住民とのコミュニケーションによる関係づくりを意識していた。

◆回帰分析の結果，保育施設に音に関する苦情が寄せられるかどうかと，保育施設の基本属性，日々の活動状況，音響機器の使用頻度や公開行事の実施状況には関係があるとは言えなかった。

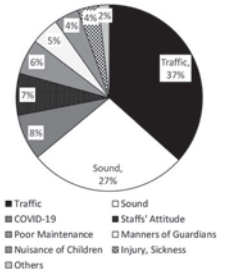



Fig.1: Response rates for various complaints received by childcare facilities

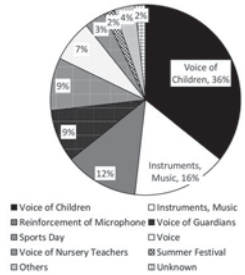



Fig.2: Response rates for complaints about various sounds from childcare facilities

## 2-11-6

### 2-11-6 音環境のアノニマス性その 5: 図書館の実環境における実験的検討

Anonymity of Sound Environment Part 5: An experiment in library environment

☆小橋侑生，安田駿（神奈川工科大），
三浦貴大（産総研），
田中哲雄，松本一教，上田麻理（神奈川工科大）

前報ではカフェにおけるアノニマス性の音響的構成要素の解明を目的として，カフェの実環境下でキーボード入力タスクを行った．その結果，他人の存在に伴う音環境のアノニマス性の一端を示した可能性が考えられた．本報では図書館におけるアノニマス性の音響的構成要素の解明を目的とし，アノニマス性の心理的／物理的要因を探るために図書館の実環境下でキーボード入力タスクを実施した．自己発生音への意識や打鍵速度の違いなどを実験的に調べることにした．

Figure: Example of application screen and application configuration

## 2-11-7

### 2-11-7 三次元音場再現システムを用いた音風景体験プラットフォーム

Platform for soundscape design using 3D sound reproduction system

☆大出隆晴，松尾綾子，上野佳奈子(明治大)，伊勢史郎(東京電機大)

◆本研究は 3D オーディオ制作システムを用いて音風景を制作・聴取する体験の価値を確認すること，および 3D オーディオ制作システムの活用の可能性を探ることを目的とした。そのために実験では被験者となる学生に，「卒業アルバム」というコンセプトで音風景を創作させた。

◆実験の手順としては，学生としての日常を特徴づけている音を収録させ，それらを三次元的に配置して，自身の学生生活の記録としての音風景を創作し，音響樽内で 3D 受聴させ，音風景の作り方，創作した音風景を聴取した印象について回答させた。

◆結果として日常生活で身近に聞いている音源を使用して三次元的な音風景として創作することによって新たな価値を提供でき，また音風景を創作・体験するプラットフォームとしての基礎的な知見が得られた。制作においては，音源の時間的・空間的な配置や動きを自由に操作できることが，音風景デザインの楽しさに繋がることが示唆された。

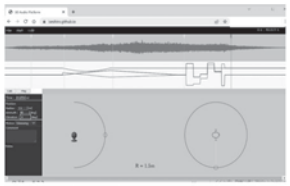


Fig.1:3D Audio Production System

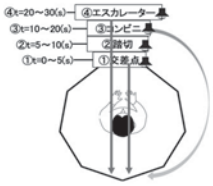



Fig.2: Image of soundscape design

## 2-11-8

### 2-11-8 深層学習を用いた自然環境音からの快適音創生手法における学習モデル比較

Comparison of learning models in the method for creation of comfortable sounds from various nature sounds by using deep learning

○星野博之（愛知工業大学）

◆目的　各種作業環境において人の状態を適切かつ快適に保つための音刺激を自然環境音から探索、創生することを目的としている。今回は、これまでに提案した、多数の自然環境音のスペクトログラムと快適さ主観値との関係の深層学習(DNN)モデルと非負値行列因子分解(NMF)を用いて自然環境音から快適音を生成する手法について、新たな快適音の創生および学習モデルの最適選択を目的とし、データを追加して快適さ推定深層学習モデルを新たに２つ作成して、モデルの違いによる生成音の特徴について検討した。

◆結果　Fig. 1 に 20 個の原音から２つのモデル M1,M2 により生成された快適音について、快適さ推定値と主観値との関係を示す。評価値 0.5 点以上の音が M1 で３音、M2 で４音生成されており、快適音の候補と考えることができ、高評価の２音は「虫 1(1.4)」「鳥 2_ウグイス(0.8)」からの生成音であった。一方でモデル比較では、２つのモデルでの生成音の主観値全体に大きな傾向の違いはなかったが、同じ原音からの２つの生成音の主観値に１点以上の差が３音であり、モデルの違いで生成音の主観値が大きく変わることがある。

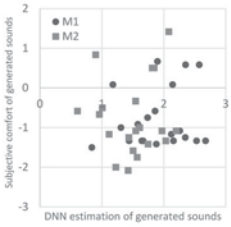



Fig.1 Relation between DNN estimation and subjective comfort rating of generated sounds by M1 and M2

## 2-11-9

### 2-11-9　骨導音及び気導音を考慮したシャンプー音の感性評価

Sensory evaluation of shampoo sound in consideration of bone conducted and air conducted sound

☆SONG WENZHANG（中央大院），△紺本 香織（株式会社マンダム），
△田辺 総一郎，戸井 武司（中央大）

近年，シャンプーの評価は主機能である頭髪や頭皮の汚れを洗浄するだけではなく，使用時の音についても自律感覚絶頂反応（ASMR：autonomous sensory meridian response）という観点から注目されている。しかしながら，実際に自分で洗髪する際に聞こえる体感音（以下，シャンプー音）は，第三者に聞こえる気導音だけではなく，骨導音も含んだ音として聞こえる。また，シャンプーの洗浄方法による擦る強さと泡立ちなどの様々な因子（以下，制御因子）があり，シャンプー音の音質に影響を与えることが経験的に分かっている。

そこで本研究では，骨導音と気導音の特性を把握することで，シャンプー音の再現方法を確立する。また，シャンプー音の音質と，複数の制御因子との関係性を明確化し，Fig. 1 に示す感性価値の異なるシャンプー音の感覚マップを構築する。

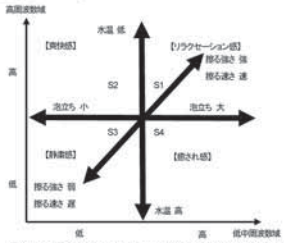



Fig. 1 Sensory map of shampoo sound

## 2-11-10

### 2-11-10　美味しそうなイメージを喚起させる咀嚼音の評価構造の分析

Evaluation Structure Analysis of the Imagery of Deliciousness Evoked by Biting Sounds

☆安藤慎一郎，高田正幸（九州大・芸工）

◆「美味しそう」というイメージに関連する評価の構造や音の特徴を明らかにするため，６種類の食品の咀嚼音を対象に，評価グリッド法を用いたインタビュー実験を行った。

◆評価グリッド法の手続きにより得られた「美味しそう」と判断した理由などの回答にテキストマイニングを適用し，評価構造の定量的な分析を行った。

◆回答に「刺激」や「ポジティブ／ネガティブ（な文脈での回答）」などのタグを付け（外部変数化して），対応分析を行ったところ，「ポジティブ」な文脈の回答に現れた語として，音の物理的特徴や音源（食品）の特徴を表す語，程度を表す語（副詞等）などが抽出された（Fig.1）。

◆共起ネットワーク分析を行い，刺激ごとに語の共起の傾向を把握した。

◆これらの結果から，「美味しそうな」というイメージと関連する語や，評価構造の一端を明らかにすることができた。

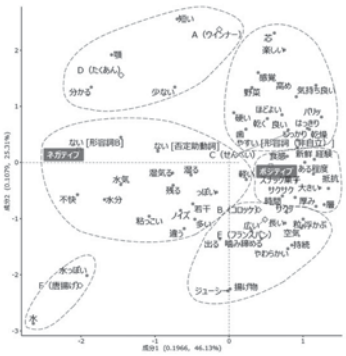



Fig. 1 Result of correspondence analysis for items including 60 major words, 6 stimuli, and positive/negative tags

## 2-11-11

### 2-11-11　和音の協和性と警告感の関係の検討

A study on the relationship between degree of consonance and impression of warning

☆阿部理花子，大塚翔，中川誠司（千葉大）

◆これまで，警告効果の高い警報音の設計を目指して，和音の協和感が警告感に及ぼす影響を検討してきた．しかしながら，すべての被験者に共通する傾向が見いだされたとは言い難い．刺激音によって各被験者に惹起される協和感の差異やバラツキの評価を目的に，三和音の構成周波数比が音の「濁り感」および「不快感」に及ぼす効果を調べた．

◆周波数の異なる３つの純音，および基本周波数の異なる３つの楽音で構成された三和音の構成音の周波数比を変化させて，それぞれ７つの刺激音を作成した（Table.1,2）．Scheffe の一対比較法を用いて，「濁った」と「不快な」の二つの形容詞尺度について各刺激音の聴感印象を調べた．

◆純音で構成された三和音については，全被験者において和音の構成周波数比の複雑化に伴う音の濁り感の増大が確認された．また，構成周波数比の複雑化に伴って音を「不快」と感じる傾向が過半数の被験者でみられたが，全体に共通する有意な傾向は認められなかった．

◆楽音で構成された三和音については，すべての被験者において，和音の周波数比の複雑化に従って音の「濁り感」，「不快感」が増大した．

◆純音，楽音に関わらず，三和音の構成周波数比が複雑になるほど協和感が大きくなることが改めて確認された．一方，「不快感」はより総合的な評価指標であり，ピッチ等の影響を受けた可能性が示唆された．

Table 1 Triads (pure tone)

| Sound | Frequency [Hz] | | | Ratio |
|---|---|---|---|---|
| A | 420 | 630 | 840 | 2 : 3 : 4 |
| B | 420 | 525 | 630 | 4 : 5 : 6 |
| C | 420 | 490 | 560 | 6 : 7 : 8 |
| D | 420 | 472.5 | 525 | 8 : 9 : 10 |
| E | 420 | 462 | 504 | 10 : 11 : 12 |
| F | 420 | 455 | 490 | 12 : 13 : 14 |
| G | 420 | 450 | 480 | 14 : 15 : 16 |

Table 2 Triads (piano sound)

| Sound | Fundamental frequency [Hz] | | | Ratio |
|---|---|---|---|---|
| A | 261.62 | 391.99 | 523.25 | 2 : 3 : 4 |
| B | 261.62 | 329.62 | 391.99 | 4 : 5 : 6 |
| C | 261.62 | 311.12 | 349.22 | 6 : 7 : 8 |
| D | 261.62 | 293.66 | 329.62 | 8 : 9 : 10 |
| E | 261.62 | 293.66 | 311.12 | 10 : 11 : 12 |
| F | 261.62 | 277.18 | 293.66 | 14 : 15 : 16 |

## 2-11-12

### 2-11-12　協和音付加による振幅変調に対する変調感の改善

Improvement of modulation feeling for amplitude modulation by adding consonance

☆広沢隼土（中央大院），△和久井康平，
松田道昭（荏原製作所），戸井武司（中央大）

◆機械動作音の純音成分が有する振幅変調は変調感を生み，不快感に影響する。本研究では，振幅変調する純音（以下，変調純音）への協和音付加により変調感を改善し，快適感の向上を目指す。

◆変調感および快適感に関し，協和音付加前後で７段階の相対評価を行った。協和音付加後の周波数特性を Fig. 1 に示す。協和音付加により，4 Hz の変動を感じる低速変調純音では変調度の小さい条件で，70 Hz の粗さを感じる高速変調純音では変調度の大きい条件でも変調感が改善された。

◆また，高速変調純音への協和音付加に対する快適感の評価値を Fig. 2 に示す。振幅変調しない定常純音へ付加した S と比べ，変調度 0.20(R1)，0.30(R2)，0.40(R3)の変調純音へ付加した場合，同等以上の快適感を得た。高速変調に起因した不快な変調感が改善されたことで，より快適感が向上したと考えられる。

◆変調純音に対する協和音付加は，特に高速変調純音に対して有効な手段となることを明らかにした。

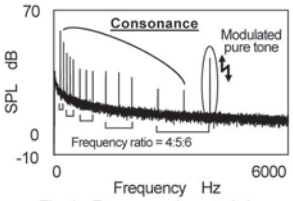



Fig. 1　Frequency characteristics

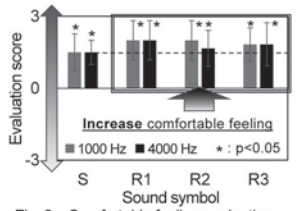



Fig. 2　Comfortable feeling evaluation results of high speed modulated pure tone

## 2-11-13

### 2-11-13　医療機器の呼吸変動音が覚醒に及ぼす影響把握

Influence of respiratory fluctuation sounds of medical devices on arousal

☆大林葵（中央大院），△島峰徹也（東京工科大），戸井武司（中央大）

持続陽圧呼吸療法(CPAP)の普及に伴い，機器本体の低騒音化が取り組まれ，マスクから発生する呼吸に伴い変動する気流音（以下，呼吸変動音）が相対的に目立っている。本研究では呼吸変動音による覚醒度の変化を，脳波分析を用いて把握する。

まず，呼吸変動音の周波数特性を Fig. 1 に示す。ピークを有する周波数をそれぞれ P1~P5 とすると，特に吸気の P3, P4, P5 の高周波が上昇していることがわかった。

次に，マスク着用時と同等の静圧負荷となる定常音と，呼吸変動音における覚醒度を示す AAC の変化について脳波分析，および主観評価を行う。覚醒度，および評価平均得点を Fig. 2 に示す。脳波分析より，呼吸変動音(Fluctuation)は定常音(Steady)と比較し覚醒度が上昇する傾向がみられた。また，主観評価より，呼吸変動音において快適感および眠りやすさの得点が低下した。したがって呼吸変動により覚醒度の変化，および主観的な印象の変化が生じ，睡眠を阻害することがわかった。また，呼吸音による変動感を有し，高周波ピークを低減することで，眠りやすく，快適感の向上が見込めることがわかった。

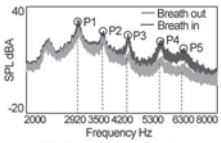


Fig. 1　Frequency Spectrum

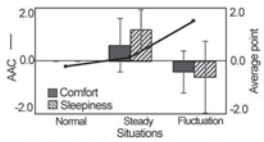


Fig. 2　Evaluation results of arousal level

## 2-12-1

### 2-12-1　ｅスポーツと教育

E-sports and education

○上田麻理，塩川茂樹(神奈川工科大)

本稿では，ｅスポーツ研究センターの活動の一環であるｅスポーツを介した地域活性のための行政や自治体との連携や大学での研究教育の一部を紹介する．

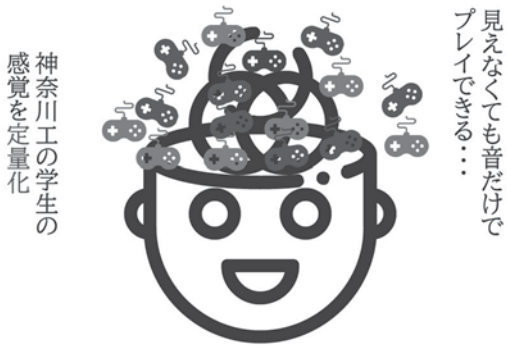


Fig. 神奈川工科大におけるｅスポーツの活用イメージ

## 2-12-2

### 2-12-2　プロスポーツクラブから見たｅスポーツの現在地

A view from professional sports clubs. Where e-sports is now and where it could go in the future

○武田裕迪(横浜マリノス株式会社ブランド戦略部)

常勝を使命とする横浜F･マリノスにおいて、ｅスポーツチームを保有する以上、競技力の向上は優先課題の一つである。速く正確な思考、判断、操作といったアウトプットの質を高めるために、視覚だけでなく聴覚、すなわち音の齎す情報における伸びしろ探求することは、ｅスポーツのジャンルによっては大いに有益であると考える。

Fig. 横浜F･マリノスｅスポーツのロゴ

## 2-12-3

### 2-12-3　静音領域定位の訓練手法

Training method of silent region localization

○岩谷幸雄（東北学院大学）

◆糖尿病などで視覚情報を失った場合に、聴覚を用いて３次元空間情報を得る必要がある。

◆音像定位能力は、晴眼者であっても日常生活で活用しているが、障害物や、開放空間の方向などは視覚情報に依るところが大きく、聴覚の能力として訓練する必要がある。

◆我々は、その能力を静音領域定位として考え、晴眼者に対する能力の気づきおよび訓練方法について検討している。

◆本発表では、実音源およびバーチャル音場を用いた静音領域定位能力の訓練結果について報告する。

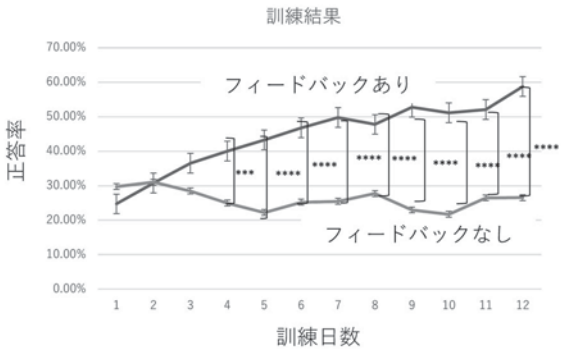


Fig.1 : Correct answer rate of training of silent region localization in virtual sound space.

## 2-12-4

### 2-12-4　FPSゲームにおける聴覚情報利用の重要性に関する検討その2

A basic study on the importance of using auditory information in FPS games part 2.

◎久我一輝，平塚大和，△鳴海颯一郎，三浦貴大（産総研），春日秀雄，塩川茂樹，上田麻理（神奈川工科大）

✧ FPS ゲーム（First Person Shooter）では，プレイヤーの足音や銃声などのゲーム内の音を聞くことでプレイヤーの位置を大まかに把握することができる。そのためゲームの音を聞いて情報を得ることが重要となってくる．本研究では，聴覚情報の利用がプレイヤーレベルの違いでどう変化するか，ゲームキャラクタごとの足音はどのように変化しているのかを調査，分析しどのようなトレーニング方法がより有効かを検討した．

No restricted visual experimentation

Fig1. Experimental Scene1

Restricted vision

Fig2.Experimental Scene2

## 2-12-5

### 2-12-5　対戦格闘ゲームにおけるコントローラーの操作音から得られる情報の検討その2 -音とプレイヤーコンディションの考察-

Examination of information obtained from controller sounds in fighting games part2 - Sound and player condition considerations

☆平塚大和，久我一輝，△鳴海颯一郎（神奈川工科大），三浦貴大（産総研），松本一教，田中博，田中哲雄，春日秀雄，塩川茂樹，上田麻理（神奈川工科大）

本研究は対戦格闘ゲームを対象とし，その中でプレイヤーのコントローラー操作音に着目した．そこでプレイヤーの最大音圧レベルの変化とコンディションを比較することによりコンディション客観的に計るための指標となりえるか明らかにしていく．実験参加者は 5 人で 1 日に 5 戦を計 5 回行ってもらい，ゲームプレイ時のコントローラー操作音を録音した．また、体調に関するアンケートを実験ごとに行った．録音した音データとアンケートやゲーム内結果を回帰分析で比較した結果，音圧が大きくなるほど体調とゲーム内結果は低くなる結果が得られた．

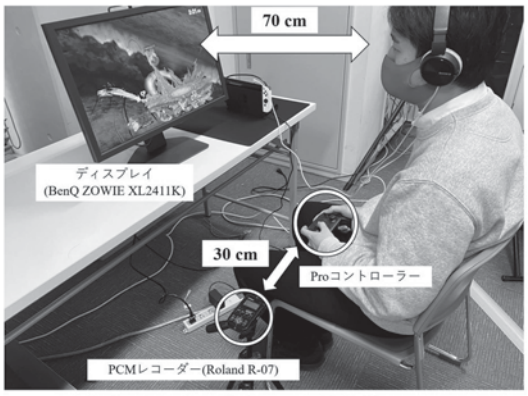



Fig.1: Experimental scene of fighting game

## 2-12-6

### 野球選手のモチベーション向上のための人工歓声付加システムの開発 その2

ーバッティング練習時の人工歓声付加の効果ー

Development of an artificial cheer addition system to improve the motivation of baseball players

○春日秀雄，二村亮平，△新田晃司，△一色正雄，上田麻理（神奈川工科大）

研究の目標

音を用いた新しいスポーツトレーニングの開発

◆目的

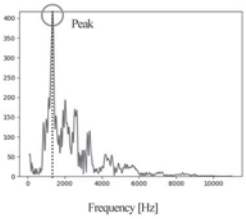
- 選手のモチベーション・やる気を向上させ、トレーニング効果を向上させるシステムの開発 ⇒ 声援を利用：声援には選手のモチベーションを高める効果がある

◆人工的な歓声を付加するシステムをバッティング練習時に使用

- ボールを打った際に選手を鼓舞するような声援を再生
- 特に良い当たりの際には大きな歓声を流す
  - ✧ 良い結果が出た際に『特別な報酬を得た感覚』を与えることができる
    ⇒ モチベーション向上につながる

◆様々なノイズのある環境で打球音を識別し、自動的に人工的な歓声を再生するシステムを構築

- 周波数スペクトルを分析し、ボールがバットに当たった際の音を判別
- 音圧レベルによって当たりの良さを判断し、状況に応じた音声を再生




バットでボールを打った際の音では 1200 ～ 1500 Hz の波形が強くあらわれる

## 2-12-7

### 2-12-7　スピーカの音質がダンスや体操の動作に与える影響

Effects of loudspeaker sound quality on dance and gymnastics movements

◎三春知史（神奈川工科大），三浦貴大（産総研），春日秀雄（神奈川工科大），△松石遼太（TOA），上田麻理（神奈川工科大）

◆スピーカの音質の良し悪しを判断する際，これまでは物理評価及び主観評価が主であった．音響分野の専門知識や音楽の経験等がない人は，スピーカの音質の良し悪しを耳で主観的に判断することは難しい場合もある．

◆しかしながら，園児が性能の異なる 2 種類のスピーカで同じダンスを踊った際に，安価なスピーカに比べ音質の良いスピーカは明らかに動きが活発になったという．

◆前報ではスピーカの音質の違いが運動動作に与える影響を調べた．その結果音質の差による影響は確認されなかった．その一因として，信号処理によって周波数特性を変えて音質の差を表現したが，実験参加者が音質の差を感じ取ることができなかった可能性がある．

◆そこで，今回は実験条件や対象者の改善を行いスピーカの異なる音質が主観的な感じ方に及ぼす影響と，それに伴う身体動作に与える影響を明らかにする．

Fig.1 Images of OpenPose bones.

## 2-12-8

### 2-12-8 視覚障害ゲーマーの e スポーツ参加を支えるアクセシビリティ技術の開発

Development of Accessibility Technology to Support Participation of e-Sports by Visually Impaired Gamers

○松尾政輝（筑波技大）

コンピュータゲームは今やｅスポーツという競技に昇華され，競技者の人口も増加している．同様に，視覚障害のｅスポーツプレイヤーも登場しているが，アクセシビリティ上の課題が存在しており満足にゲームをプレイできない場合がある．本稿では，そのような課題を解決するためのシステム開発の状況について紹介する．

◆***Audible Mapper*（オーディブルマッパー）**：全盲のゲーム開発者がゲーム内の地形を作成できるマッピングツールで，効果音情報を頼りにキーボードを用いて二次元図形を編集できる．

◆***ShadowRine***：視覚情報に加え，聴覚・触覚情報を用いてゲーム内の状況を伝達できるよう設計したインクルーシブ 2D アクション RPG．

◆***Planet Saga***：晴眼者と視覚障害者がともにプレイできる横スクロールアクションゲーム．正確かつ素早い操作が求められるリアルタイム性の高い設計．

インクルーシブアクションゲーム *Planet Saga*．左図：ビジュアルプレイ用の画面，右図：左図の画面を点図ディスプレイで表示．

## 2-12-9

### 2-12-9　ブラインドサッカーの蹴球音の音響特性と聴覚訓練システム

Acoustic characteristics of blind soccer kickball sound and auditory training system

☆梶竜也，二村亮平，春澤恒輝，（神奈川工科大），三浦貴大（産総研），
福永克己，坂尻正次（筑波技術大），上田麻理（神奈川工科大）

◆＜目的＞

HATS によりバイノーラル録音した蹴球音を用いて、ボールの距離・方向の判別を行うための訓練システムの開発・評価を行う。

◆＜評価実験＞

蹴る地点、蹴る方向、蹴り方の 3 条件を指定した 48 提示音条件を聴取し、仮想的なボールが自身の前後左右どの方向を通過するか、また、自身から近い距離、もしくは遠い距離を通過するか判断してもらった。その後、主観評価に回答してもらった。

◆＜結果＞

繰り返し回数が増加するごとに、正答率が増加することがわかった。また、前後の方向定位が困難であることがわかった。

Fig.1 : Recording of the presentation sound

## 2-12-10

### 2-12-10　聴覚障害者のためのモータースポーツレース現場の音を振動と光に変換して提供するシステムの導入事例

A case study of a system that converts the sound of a motorsport race into vibration and light for the hearing-impaired people

○長谷芳樹，△遠藤彰，△片山晴紀，△小崎貴輝，△高澤和希，
△三浦和希，△黒田藍子（ピクシーダストテクノロジーズ）

◆近年，“音”を気導音として以外の方法で聴取あるいは受容する方法の選択肢が拡がりつつある。この中で，我々が展開してきた SOUND HUG というシステムは，コンサートなどにおいて会場の音をマイクロホンで収音し，リアルタイムに振動と光で観客に届ける構成となっている。今回，このシステムを岡山国際サーキットにて開催された「第6戦 スーパー耐久レース in 岡山（通称 S 耐）」に導入し，聴覚障害者6名による観戦体験の評価をおこなったので報告する。

◆本実証の結果，例えば車が近付いてくる，あるいは遠ざかっている様子の伝達なども含め，特に振動情報の提供がレース観戦体験の向上に寄与できることが示された。また，全ての参加者が体験後にモータースポーツに対する興味が高まった，かつ，レース観戦自体は総合的に楽しめたと回答した。一方で，文字等によるレース状況などの情報提供も併せて必要であることもあらためて明確となった。

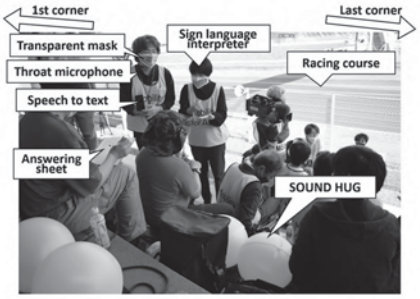



Fig.1: Experience of the motor race by hearing impaired people with holding SOUND HUG devices at Okayama International Circuit.

## 2-12-11

### 2-12-11　ヒアリングループおよび LED 照明補聴支援装置のスマートホンによる受信について

Receiving signals of hearing loop and LED lighting through smartphone

☆池田涼風，和田有司，中村健太郎（東工大）

- ヒアリングループや LED 照明補聴支援装置の信号をスマートホンで受信する方法を検討した。
- 前報で提案した LED 照明補聴支援装置とは、音声信号によるパルス幅変調によって LED 照明を点灯し、その照明下の限られた範囲に音を届ける補聴支援装置である（Fig. 1）。
- イヤフォンマイクのマイクロフォンを，ループ受信時にはピックアップコイルとし，LED 照明受信時にはフォトトランジスタとすることで信号を受信することが可能である（Fig. 2）。
- 外部マイクロフォンからの音声入力をイヤフォンに出力する既存アプリを使用することで、スマートホンで受信する。
- LED 照明補聴支援装置について、使用するフォトトランジスタの受光指向性を変えることで照明下での聴取可能範囲が変わることを確認した(Fig. 3)。

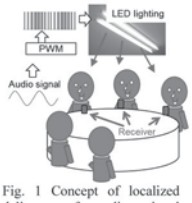



Fig. 1 Concept of localized delivery of audio signal through PWM-driven LED lighting.

Fig. 2 Connection of pickup coil (or photo transistor) and earphone to a smartphone.

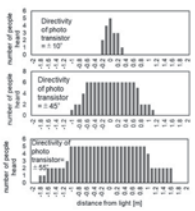



Fig. 3 Receiving area for different photo transistors.

## 2-12-12

### 2-12-12　一側性難聴者が伴奏下のメロディ聴取のために両耳聴より大きな音圧を必要とすることについて —残響の影響・方向性マスキング解除の検討—

Those with unilateral hearing loss require more sound pressure of melody than binaural listeners to hear out a melody from an accompaniment: Investigation of the effect of reverberation and spatial release from masking

☆辻慎也，荒井隆行（上智大・理工）

◆片側耳は正常だが，対耳側に難聴がある状態を一側性難聴という．

◆一側性難聴により，以下３つの場面で困難な状況が生じる．

１）難聴側聴取，２）騒音下聴取，３）音源定位

◆本稿では一側性難聴者・健聴者のモノラル受聴・両耳聴を対象に，残響下でのマスカ（伴奏・ノイズ）つきメロディ聴取能力を検討し，次の結果を得た（Fig. 1）．

１）残響下でのメロディ聴取のため，一側性難聴者は両耳聴と比べ，より大きな音圧を必要とした．

２）マスカが雑音のとき，一側性難聴者は健聴者のモノラル受聴と比べ，大きな方向性マスキング解除の効果を示した．

３）発症から長期経過した一側性難聴者のモノラル受聴におけるキューへの適応が示唆された．

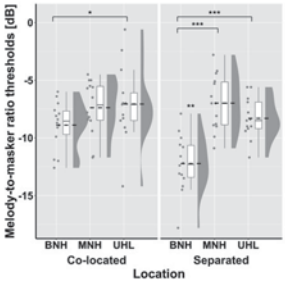



**Fig. 1** The melody-to-masker ratio (SMR) in noise for 3 hearing conditions; BNH (binaural normal haring, n=14), MNH (monaural normal hearing, n=14), and UHL (unilateral hearing loss, n=13). The solid and broken horizontal lines indicate the median and mean. SMR were measured in 2 conditions; target and masker were co-located (left) and separated (right). ***p < 0.001, **p < 0.01, *p < 0.05

## 2-12-13

### 2-12-13　疫学調査のための簡易聴こえ測定システム構築と評価

Construction and evaluation of simple hearing measurement system for epidemiological study

☆小田桐空大，△高瀬ゆり(神奈川工科大)，廣江正明(小林理研)，
△神崎晶(東京医療センター)，田中哲雄，上田麻理(神奈川工科大)

◆ ひさやま方式は，町と大学・医師らが共同で町民の健康管理を行いながら，生活習慣病予防健診のデータを活用した疫学研究をすることで，科学的根拠に基づいた町民の健康づくり活動を官学連携で展開するというものである．

◆ しかしながら，ひさやま方式だけでなく多くの場合，疫学調査において，聴力データは防音室などの特殊な部屋を有する．静けさの担保が必須であるため，聴力データは取得されていないのが現状である．さらに，自治体からは検査機器が高価であり購入が困難という意見が聞かれる．

◆ このような疫学調査時の聴力測定の困難さを解消するため，簡易聴こえ測定システムを構築したので報告する．

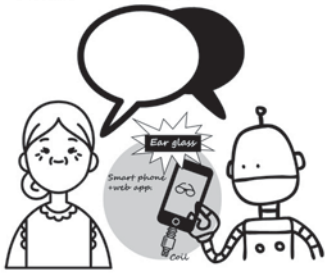



Fig.1：厚木市聴こえ特区プロジェクト（仮題）イメージ

## 2-12-14

### 2-12-14　スポーツ音響への応用に適したアンサンブル平均を用いた吸音特性の現場測定法　—測定モデルに関する再考察—

An in-situ measurement method using ensemble averaging technique for sound absorption of materials installed in athletic fields
- Reconsideration of the the EA method measurement model -

○大鶴　徹，富来礼次，岡本則子（大分大・理工）

◆スポーツ音響や建築音響をはじめ多くの分野で、芝や床・壁面などの吸音特性の現場測定が必要とされている

◆筆者らが提案してきたアンサンブル平均を用いる測定法（EA 法）は、現場での測定にも柔軟に適用可能である

◆理論上、材の吸音特性に入射角$\theta$に対する依存性が無ければ、$\theta$の生起確率分布に依存することなく、EA 法の適用で垂直入射吸音率の推定値が求められる

◆本稿では、測定モデルについて再考察を行った後、課題解決のための留意事項を試論的に整理した

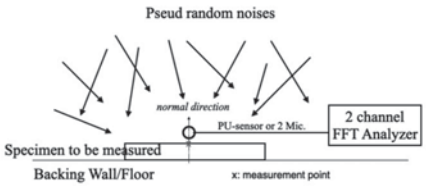



Fig.1: Measurement setup for EA method: $EA_{pp}$ method with two microphones and $EA_{pu}$ method with a PU sensor. In case a wall/floor is to be measured, no specimen is installed.

## 2-12-15

### 2-12-15　体操競技選手と一般大学生における身体活動時の内的発話に関する研究

A Study of Inner Speech during Physical Activity in Gymnasts and General University Students.

○久永将太（慶應大・総），浅野恵子（順天堂大・医）

◆「内的発話」とは自分自身への問いかけや指示出し、思考の整理や調整などに用い、音声を伴わない発話だといわれている。

◆これまでアスリートと一般大学生の身体活動時の心の状態と言語について比較調査した研究はされていないであろう。

◆体操競技選手(18~25 歳の男子選手 55 名)と一般大学生(18~24 歳の男女 80 名)の合計 135 名にアンケート調査を実施した。

◆質問項目：①日常生活時、②一般運動時の内的発話について、である。

◆「音声」、「映像」、「文字」による思考など、様々な思考のスタイルが存在していることが分かった。

◆一般大学生は「音声」、体操競技選手は「映像」を使用しており、後者は、自分自身を俯瞰する傾向にある。

◆一般大学生は「文字」での思考が体操競技選手と比べて少数であった。

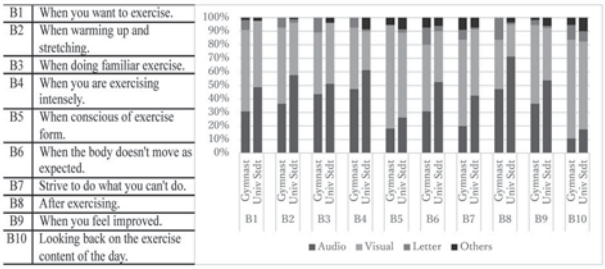



Fig.1：Questionnaire survey of inner speech during physical activity.

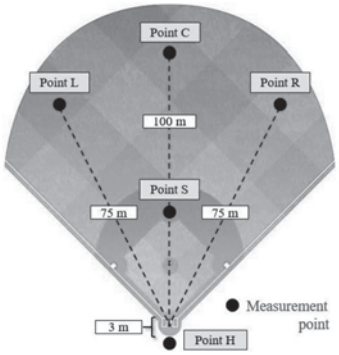
## 2-12-16

### 2-12-16 野球競技における打球音の音響学的特徴 その2 ー競技者による打球音判別と聴感評価ー

* Acoustical characteristics of hitting sounds in baseball games part 2 : Judge the sound of the ball hit by the athlete and evaluate his or her hearing

二村亮平，○春日秀雄，新田晃司(神奈川工科大)，
三浦貴大(産総研)，山口雄大(日本体育大)，上田麻理(神奈川工科大)

◆野球競技において競技者は，絶えず変化していく場面の中で適切なプレーをするため，常に周りの状況から情報を得ている．その中でも聴覚情報は，打球音や捕球音，人の声など様々な音を基に次の身体動作決定の他，ボールの挙動の予測に活用している．このことから，競技者は聴覚情報を活用することで競技力が向上すると考えられる．




◆本稿では，5 つの地点で打球音の音響計測を行い，打球種や打球方向，打球音の音響特徴の関係や競技力の差と熟練度の違いによる打球音の音響学的特徴の違いを明らかにするために音響解析を行った．

## 2-12-17

### 2-12-17 バレーボール競技の聴覚情報利用に関する基礎的検討その 3 ～プレーの良し悪しに関する音響的要因～

Examination of Auditory Utilization by Position in Volleyball Part III –Acoustic elements related to ” “good” or “bad” plays

☆髙橋隼人(神奈川工科大)，△山口雄大(日体大)，
三浦貴大(産総研)，上田麻理(神奈川工科大)

◆熟練した競技者が，普段言葉で感覚的に表現している競技状況を文脈ごとに明らかにすることを目的とし，本稿では競技者が感じるよい音，悪い音が何を示すのかを，アンケートにより調査した．

◆ Table. 1 の回答より，TF-IDF によって形態素解析したそれぞれの単語の重要度を分析したところ，トスにける「無音」の重要度が高いことが分かった．今後は様々な条件ごとにシチュエーションを想定した音の録音・分析を行う．

Table.1: Questions about how to feel sounds during the competition

| No. | 質問項目 | No. | 質問項目 | No. | 質問項目 |
|---|---|---|---|---|---|
| Q1 | 音全般に関する自由記述 | Q2 | 音の良し悪しに関する項目 | Q3 | 音を利用した判断基準に関する項目 |
| | | Q2-1 | 強打ジャンプサーブ | Q3-1 | (セッター用)<br>キャッチがＡパスかＣパスか |
| | | Q2-2 | 軟打ジャンプサーブ | Q3-2 | 相手のサーブボールに対してインかアウトか |
| | | Q2-3 | 強打フローターサーブ | Q3-3 | ブロックがワンタッチをしたかしていないか |
| | | Q2-4 | 軟打フローターサーブ | | |
| | | Q2-5 | トス | | |
| | | Q2-6 | 強打スパイク | | |
| | | Q2-7 | 軟打スパイク(フェイント) | | |
| | | Q2-8 | キャッチ | | |
| | | Q2-9 | ディグ | | |
| | | Q2-10 | ブロック | | |
| | | Q2-11 | スパイクを打つ際の助走 | | |

## 2-1P-1

### 2-1P-1 フレームごとの信号の大きさと窓の形状を考慮した反復型位相復元

Iterative phase reconstruction algorithm considering frame-by-frame signal magnitude and window shape.

☆小林朋記，田中僚郎（早大理工），矢田部浩平（農工大），及川靖広（早大理工）

◆ 窓の形状を考慮した位相復元アルゴリズム

同じ振幅スペクトログラムから位相により様々な波形が得られる

- 正しい位相により窓で変調された波形 Fig. 1(右上)
- 正しくない位相により乱れた波形 Fig. 1(右下)

乱れた波形を窓の形状で切り取ることで，窓の整合性の性質を活用

◆ 提案法

フレームごとに切り取る窓の高さを調整する

→ より多くのサンプルを変更し，小さな信号に対しても有効

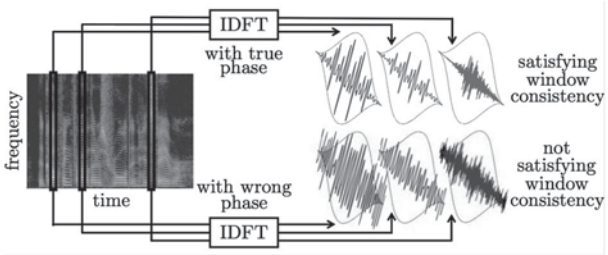



Fig.1: Illustration of window consistency.

## 2-1P-2

### 2-1P-2 ビジュアルマイクロホンにおける撮影距離拡張に関する検討

A study on extension of shooting distance in Visual Microphone

☆関根海渡，立蔵洋介（静岡大）

◆ **目的**：カメラ映像のみから音波を抽出するビジュアルマイクロホンにおいて，被写体とカメラ間の距離（撮影距離）に着目し，光学ズーム（焦点距離変更）による抽出可能な撮影距離の拡張を検討

◆**実験**：焦点距離２通り（焦点距離Ａ：短，焦点距離Ｂ：長）で撮影
撮影距離 10 cm から 100 cm まで 10 cm ずつ変更

◆**結果**：撮影距離が離れることで抽出精度の低下を確認
焦点距離を長くすることが抽出可能な撮影距離の拡張に有効
同じ撮影距離でも焦点距離長い方が提示音のパワーが増大

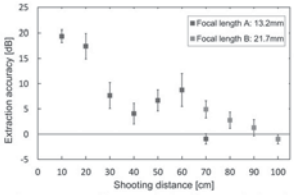



Fig. 1 Extraction accuracy of extracted signals at each shooting distance.

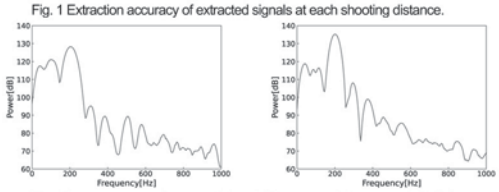



Fig. 2 Power spectra of extracted signal. (Target sound: Pure tone 200 Hz)
(a) Shooting distance: 70 cm, Focal length A　(b) Shooting distance: 70 cm, Focal length B

## 2-1P-3

### 2-1P-3　実際に使用する素材を用いた音響特性を多点同時計測する代替案の評価について

Comparative evaluation of alternatives for measuring acoustic environment using actual sound materials

○河原英紀（和歌山大），矢田部浩平（東京農工大），水町光徳（九工大）
榊原健一（北海道医療大），松崎博季（北海道科学大）

◆実際の音素材を用いて，講義や演奏などの最中に学生や聴衆の各々が経験している音環境を測定する方法について，実装する際の複数の代替案を提案するとともに，それぞれの代替案について暗騒音の影響を評価した。

◆公開されているインパルス応答のデータベースの情報を用いて数値実験を行ったところ，代替案は当初の狙いであったSNRによる劣化の軽減には応用できないことが明らかとなった。

◆効率と見通しの良い実装として，音響計測用のツールの改良と拡張に提案した方法の一つが応用できる可能性を示した。

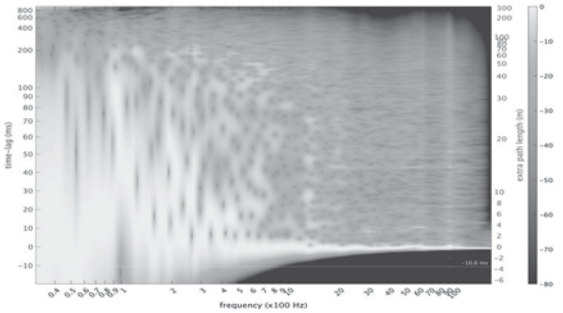



Fig.1：Visualization of an impulse response used in evaluation of alternative implementations. (Classroom impulse response of MIT IR database)

## 2-1P-4

### 2-1P-4　深層学習に基づく仮想マイク生成における学習基準の検討

Multi-task learning for neural network-based virtual microphone estimator

☆瀬川華子（筑波大），落合翼，マーク・デルクロア，中谷智広，池下林太郎，荒木章子（NTT），山田武志（筑波大），牧野昭二（早稲田大/筑波大）

◆研究背景と目的：

- 深層学習に基づく仮想マイク生成技術（NN-VME）
- 後段のアレイ信号処理が既知のとき、その出力を学習基準に利用
  → 信号処理性能をさらに改善できる仮想マイク信号を生成可能に
- 以下を組み合わせたマルチタスク損失関数 $L$ を提案
  $L = \alpha L_{\mathrm{VM}} + (1-\alpha) L_{\mathrm{BF}}$
  → $L_{\mathrm{VM}}$: 仮想マイク信号に対する損失関数
  → $L_{\mathrm{BF}}$: ビームフォーマ出力に対する損失関数

◆実験結果

- ビームフォーマ出力に対する損失関数のみを用いた場合 ($\alpha = 0.0$)
  → $\mathrm{SDR_{AP}}$（信号処理性能）が向上、一方でWERはやや低下
- マルチタスク損失関数を用いた場合 ($\alpha = 0.3$)
  → $\alpha = 0.0$ と同等の$\mathrm{SDR_{AP}}$，$\alpha = 1.0$と同等のWER を達成
  高い強調性能と認識性能を同時に実現

Table 1：$\mathrm{SDR_{VM}}$, $\mathrm{SDR_{AP}}$ [dB] (higher is better) and WER [%] (lower is better)

| RM | VM | $\alpha$ | $\mathrm{SDR_{VM}}$ | $\mathrm{SDR_{AP}}$ | WER | |
|---|---|---|---|---|---|---|
| 4,6 | 5(4,6) | 0.0 | 7.3 | 15.3 | 11.7 | $L_{\mathrm{BF}}$で学習 |
| 4,6 | 5(4,6) | 0.3 | 16.1 | 15.4 | 11.3 | 両方で学習 |
| 4,6 | 5(4,6) | 1.0 | 16.5 | 13.7 | 11.3 | |
| 4,6 | - | - | - | 10.9 | 11.7 | $L_{\mathrm{VM}}$で学習 |
| 4,5,6 | - | - | - | 14.4 | 9.5 | |

## 2-1P-5

### 2-1P-5　AVS 方式による距離のある音像定位現象に着目した音像定位距離制御の試み

A trial for distance control of sound image localization using AVS method

☆沖田和久，林友哉，福島学（NBU），△松本光雄（），柳川博文（arsl）

◆目的：音空間の創出

◆手段

距離のある音像定位現象を制御

AVS 方式

1/4 oct.バンドで ICTD と ICLD のみで側方定位する手法

◆結論：

両耳位置でのダイレクトパスとクロストークの合成 IR（Fig.2）と提示刺激（Fig.3）から，

反音像定位側のクロストークキャンセル量と時間範囲が重要.

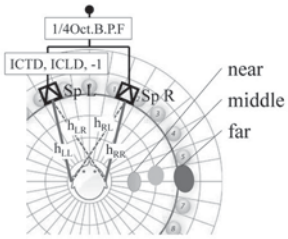



Fig.1 The AVS settings and The Sound Image Location

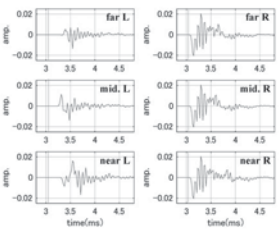



Fig.2 The difference of IRs based on the Sound Source Localization far / middle / near

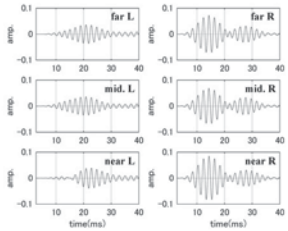



Fig.3 The Stimulus Sounds far / middle / near fc = 500 Hz, 1/4 Oct. band pulse

## 2-1P-6

### 2-1P-6　楽曲時間波形の包絡線に着目した振動刺激提示方法の一検討

Vibration stimulus display focusing on the envelope of time waveforms of musical signals

☆林友哉，沖田和久，△伊藤毅登，福島学（日本文理大学），風間道子（log a），柳川博文（arsl）

◆【目的】ライブ感ある楽曲提示のため，
低周波騒音とならないよう低音を再生したい.

◆【提案】低音を振動刺激として提示することを検討する.
振動刺激を楽曲の低周波成分ではなく，
楽曲時間波形の包絡線情報より得た振動タイミングとする.
Fig.1 に振動タイミングの刺激波形（2～5 段目）を示す.

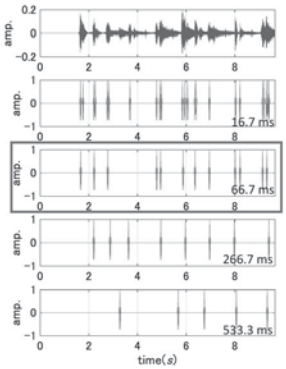



Fig.1 Stimulus waveform
(1st：Time waveform, 2nd：Designated section 16.7 ms, 3rd：Designated section 66.7 ms, 4th：Designated section 266.7 ms, 5th：Designated section 533.3 ms)

## 2-1P-7

### 2-1P-7 オープンイヤー型イヤホンにおける外音聴取最適化に向けたスピーカ配置に関する検討

Small Speaker Arrangement for Optimization of Outside Sound Listening in Open-Ear Type Earphones.

◎加古達也，千葉大将，野口賢一（NTT）

◆耳の負担を軽減し，周囲の環境音（外音）を取り込めるオープンイヤー型イヤホンの需要が増えているが，デバイスを耳近傍に置くことで外音の音質に影響を与えてしまう

◆市販デバイスの外音聴取影響の評価

- ➢ イントラコンカ型では3 kHz 以降で8 dB 音圧レベルが下がる
- ➢ 耳の上部（三角窩付近）に取り付ける場合9 kHz で8 dB 音圧レベルが下がり，骨伝導の様なこめかみ装着では4 dB 下がる

◆外音聴取性能と音質の両面を満たす適切な筐体配置の検討

- ➢ 耳の近傍の24 箇所にφ16 mm の筐体を配置し外部スピーカと筐体内蔵のスピーカの特性を測定 (Fig.1 )
- ➢ 2 変数を考慮した距離関数 $d$ から n=2,7,9 の配置が良い傾向

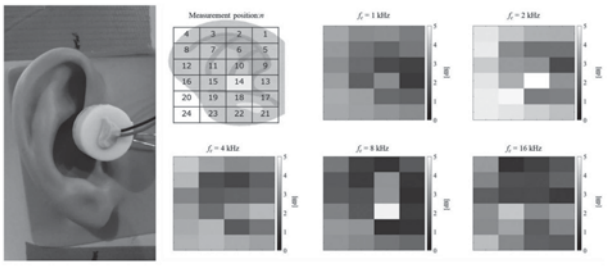



Fig.1 Results for Mahalanobis distance $d$ considering sound quality and outside sound listening performance at the enclosure locations around the ear.

## 2-1P-8

### 2-1P-8 超低周波数域におけるLS1 マイクロホン音圧感度のカプラ校正に関する実験的考察

Experimental study on coupler calibration of LS1 microphones in the infrasonic frequency

◎平野琴，高橋弘宜，山田桂輔（産総研）

◆産総研では20 Hz 以上の可聴域における音響標準として、カプラ校正法による標準マイクロホンの音圧感度校正を実施している。

◆災害対策において 20 Hz 以下の音の計測が活発化している状況を鑑み、カプラ校正法が適応可能な周波数の下限を検討するため、LS1 マイクロホンの音圧感度評価を0.2 Hz まで行った。

◆低周波数域では、Vent と Needle の隙間からの音圧漏れが問題となる。径の異なる Needle を用い音圧漏れが音圧感度に与える影響を調べた。

◆数 Hz 付近までは一致した音圧感度が得られたが、より低周波数域では Needle が細く隙間部分が大きい方が、振幅は小さく位相は大きい傾向が見られた。

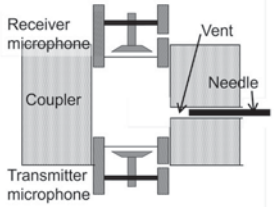



Fig.1. Setting of transmitter and receiver microphones in the coupler calibration method.

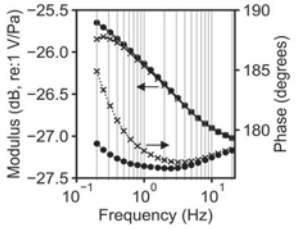



Fig.2. The pressure sensitivities of the LS1 microphone obtained using needles of different diameters (●: 0.745 mm, ×: 0.710 mm).

## 2-1P-9

### 2-1P-9 精密光干渉計によるパラメトリックスピーカ放射音の近距離場計測

Near-field measurement of sound radiation from parametric array loudspeaker

○石川憲治，△白木善史，守谷健弘（NTT）

パラメトリックスピーカ直近から 400 mm までの近距離場をレーザドプラ振動計を用いて 3 次元計測

搬送波・側帯波・復調波の 3 次元複素音圧分布の測定に成功

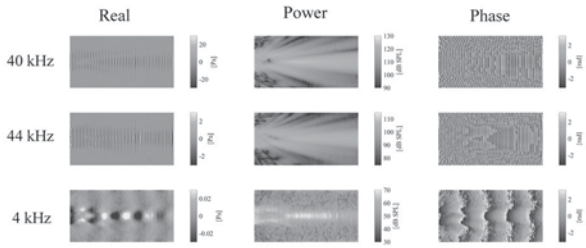



Fig.1 伝搬方向に沿った 2 次元断面（左端がスピーカ位置）

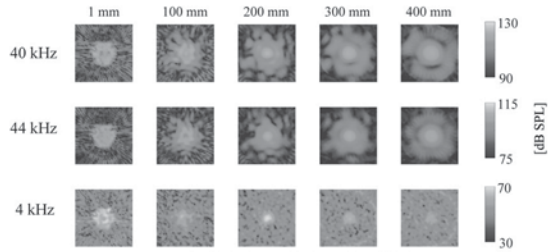



Fig.2 伝搬方向に垂直な 2 次元断面

## 2-1Q-1

### 2-1Q-1 音源・反射板間で生じる多重反射を用いた音源推定に関する検討 —距離スペクトルにおける対象距離外のピークについて—

Consideration on distance estimation from sound source to reflector using multiple reflection between sound source and reflector (peaks in range spectrum other than true peak)

☆大橋勇介，上保徹志（近畿大），中山雅人（阪産大），中迫昇（近畿大）

◆本稿では，音源と反射板の間に生じる多重反射に着目し，音響測距法によって算出された距離スペクトルにおいて生じた測定対象距離以外のピークについての検討を行う．

◆音源推定の実験において反射板と音源間sの距離が0.8mを超えると，距離スペクトル(Fig.2)には測定距離以外に大きなピークが現れ，観測信号(Fig.1)でも多重反射が起こる以前の時刻にピークが見られた．

◆測定対象以外のピークが床面からの反射によるものと仮定し，理論計算にて距離スペクトルのピーク位置を求めたところ，実験におけるピーク位置に近い値が算出された．

◆上記のことから，真の音源距離とは異なる距離が床からの反射によるものと考えることができる．逆に，このことを利用して音源の高さが推定可能であることが示唆された．

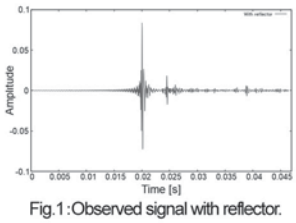



Fig.1:Observed signal with reflector.

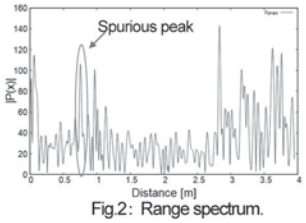



Fig.2: Range spectrum.

## 2-1Q-2

### 2-1Q-2　スピーカアレイを用いた回転する発話者の模擬性能

Evaluation of Spherical Loudspeaker Array Reproducing Speaker Rotation

☆横田康介，羽田陽一（電通大）

◆円滑な遠隔コミュニケーションを実現するためには，実際に人が話しているような臨場感を再現することが重要である。本報告では，12ch 球面スピーカアレイを用いて回転する発話者の再現を行い，回転方向の知覚と音色に関して評価を行った。

◆回転する音声を 1ch 直接再生(A), 4ch 直接再生(B), 12ch 直接再生(C), MVDR フィルタを用いた4ch 再生(D), MVDR フィルタを用いた12ch 再生(E)の５つの手法で再生し，受聴実験を行った。

◆回転方向の知覚については，指向性が確認できた MVDR フィルタによる再生手法の模擬性能が高いことが確認できた。

◆音色については，直接再生法の方が自然であったが，実際の人の自然な音声の再現は周波数特性の再現も含めて未だ不十分であり，音色の再現性に課題があることが分かった。

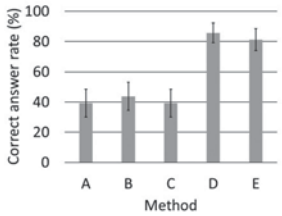



Fig.1：Correct answer rate of rotation direction

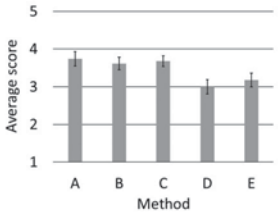



Fig.2：Results of timbre evaluation

## 2-1Q-3

### 2-1Q-3 仮想的焦点形成と瞬時スペクトル減算法による同一到来方向の複数雑音低減に関する考察

Consideration on multiple noise suppression in the same direction using virtual focus and instantaneous spectral subtraction

○小澤賢司（山梨大・工），渡邊貴治（秋田県立大・システム科技），坂本修一（東北大・通研）

◆直線状マイクロホンアレイ（マイクロホン数：4）で収録した音に位相回転を施すことで，仮想的に焦点を持つ円弧状アレイを考える。アレイの出力信号の２次元スペクトルを用いて，分析の時間フレームごと独立に減算すべき値を推定するため，瞬時スペクトル減算法と称している。

◆同一方向にあり距離が異なる複数音源の信号分離を，前報では雑音源の位置情報を既知という条件で行った。本稿では，その位置情報を未知のまま一括して雑音抑圧することを試みた。

◆仮想焦点を目的音（例えば $y$ 軸上１m）にあわせると，$y$ 軸上で３m 以上の音源に対して２次元スペクトルはほぼ同一となる。そこで，全ての雑音源が４ m 位置にあると仮定し，瞬時スペクトル減算を行うことで雑音抑圧を行った。Fig. 1 に示すとおり，３雑音源がある場合でも焦点から離れていれば 15 dB 程度の減衰が得られることを示した。

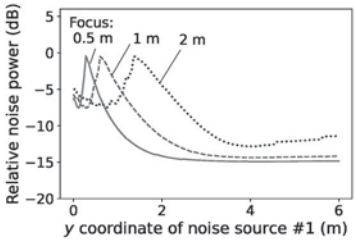



Fig.1 : Noise suppression of three noise sources when the target source position is set to the virtual focus. Noise sources #2 and #3 are fixed at 4.5 m and 5.5 m, respectively, and source #1 is moved along $y$ axis.

## 2-1Q-4

### 2-1Q-4　体導音を用いた手作業内容の認識に関する検討

A study on manual working recognition based on body-conducted sound

○佐宗晃，児島宏明，小木曽里樹，△長久保晶彦（産総研）

◆人口急減と超高齢化による労働力人口の加速度的な減少により、製造現場などでは協働ロボットの導入による省人化、生産の効率化、熟練技術の継承などが喫緊の課題となっている。このために作業行動を認識する技術の実現が不可欠である。

◆作業者が視覚的に遮蔽された状況下や周囲騒音が存在する環境下でも、手作業内容を頑健に認識可能とするために、手作業の振動が手の皮膚、組織、骨などを介して伝搬し、手首装着型のセンサーで計測される体導音に基づいた、手作業内容の認識手法を検討する。

◆ハイパーパラメータの調整が少なく高速に、また少量の収録データでも安定的にモデルの学習が可能な手法の構築を目指して、DRLDCC-ELM を提案している。

◆特定被験者１名が行った１３種類の手作業内容の認識に関する予備実験を行い、Transformer を用いたベースラインモデルと比較した結果、提案法の有効性を確認した。

## 2-1Q-5

### 2-1Q-5　複数点計測による腸蠕動音の空間特性の検討

A study of spatial characteristics of bowel sounds using multipoint measurements

◎早川朋花，大塚翔(千葉大)，
△堀創史，△本多通孝(福島県立医大)，中川誠司(千葉大)

◆ 腸蠕動音の聴診は腸機能の診断に用いられているが，蠕動音の発生部位や発生メカニズムの詳細には未解明な部分が多い.

◆ 腸蠕動音に係る先行研究の多くは１チャネルで計測を行ったもので，空間特性の検討はほとんどなされていない．本研究では腹部９カ所にセンサを配置して，空腹時および炭酸水摂取後の蠕動音の計測を行った.

◆ 炭酸水摂取後には RMS の大きな区間の割合が増加することがわかった(Fig. 1).

◆ 時間波形およびトポグラフィーによる解析の結果，ある箇所で蠕動音が発生した後，0.5 ~ 2 秒後に隣接した箇所で蠕動音が発生するパターンがしばしば観察された (Fig. 2).

◆ トポグラフィー解析から，腹部左上部と下部における蠕動音が発生がしばしば確認された (Fig. 3)．左上部は空腹による胃の収縮，下部は小腸と大腸のつなぎ目である回盲弁を通る際の腸内ガスの移動を反映していると考えることも出来る.

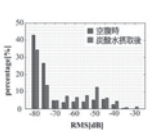



Fig. 1 各時間区間における RMS 値

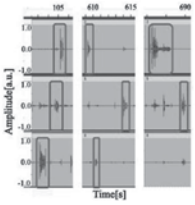



Fig. 2 腹部中央３点の時系列波形例

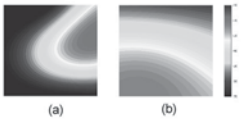



Fig. 3 蠕動音発生時のトポグラフィー例．左上部(a)，下部(b)に蠕動音発生源の存在が推定される

## 2-1Q-6

### 2-1Q-6 An Inexpensive 3D-Printed Bell Stethoscope Connected to a Smartphone for Low-Resource Settings

◎ IRWANSYAH, Sho OTSUKA, Seiji NAKAGAWA (Chiba University)

◆We have designed a stethoscope model that consists of a bell, cap, and acoustical resonator (see **Fig. 1**).

◆Our stethoscope can be acoustically optimized by using different resonators, each with a unique length of air column.

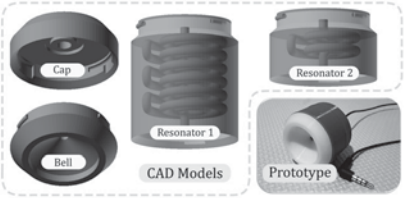



**Fig. 1** CAD models and the prototype of the proposed stethoscope

◆When we tested our stethoscope with Resonator 1, we found that it had a comparable frequency response to the selected acoustic stethoscope at low frequencies (see **Fig. 2**).

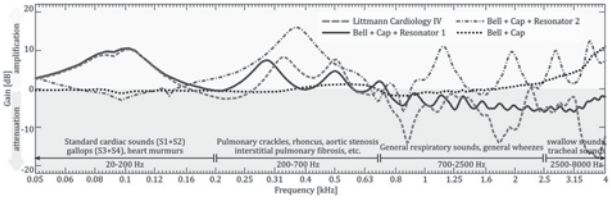



**Fig. 2** Stethoscope frequency responses

## 2-1Q-7

### 2-1Q-7 オープンイヤー型イヤホンにおける耳周辺のマイクを用いた音漏れの評価

Evaluation of sound leakage at the microphones around the ear in the open-ear earphone.

◎千葉 大将, 加古 達也, 野口 賢一（NTT）

◆オープンイヤー型イヤホンは音漏れが大きいため，耳周辺に周囲の音を収音するマイクを設置すると外音を利用する機能の性能や品質が劣化する恐れがある．そこで，オープンイヤー型イヤホンの耳掛け部にマイクを配置して周波数特性を測定し，耳周辺における音漏れの大きさを評価した．

◆2 kHz 以下では，逆相を放射して音漏れを打ち消す筐体は逆相を放射しない筐体（密閉型）よりも約 15dB 音漏れが小さいことを確認した．また，筐体形状や周波数帯域によって音漏れがより小さいマイク位置があることを確認した．

| 低周波 (20 – 2,000 Hz) | | | | | | | | | |
|---|---|---|---|---|---|---|---|---|---|
| マイク位置 [deg] | 0 | 45 | 90 | 135 | 180 | 225 | 270 | 315 | Mean |
| 密閉型 | -20.0 | -19.3 | -21.3 | -23.0 | -23.0 | -22.1 | -22.6 | -19.9 | -21.4 |
| 開放型 | -32.3 | -41.7 | -34.2 | -31.9 | -33.5 | -36.9 | -42.3 | -35.7 | -36.1 |
| 背面開口 d=4mm 相当 | -33.3 | -34.8 | -31.3 | -30.2 | -30.1 | -33.3 | -40.3 | -48.5 | -35.2 |
| 側面開口 d=2mm | -31.2 | -48.0 | -35.0 | -31.7 | -34.6 | -37.7 | -41.3 | -31.7 | -36.4 |
| 側面開口 d=3mm | -29.4 | -45.0 | -34.6 | -32.2 | -33.5 | -39.3 | -46.0 | -33.6 | -36.7 |
| 側面開口 d=4mm | -31.3 | -46.1 | -33.7 | -32.9 | -34.3 | -36.8 | -41.8 | -38.4 | -36.9 |
| 側面開口 d=5mm | -37.5 | -42.7 | -34.9 | -32.4 | -32.8 | -36.9 | -46.3 | -38.4 | -37.7 |
| 側面開口 d=6mm | -31.5 | -34.7 | -34.5 | -31.4 | -33.1 | -37.2 | -44.5 | -34.1 | -35.1 |

音漏れの大きさ [dB]　-50.0　-45.0　-40.0　-35.0　-30.0　-25.0　-20.0

| 高周波 (2,000 – 10,000 Hz) | | | | | | | | | |
|---|---|---|---|---|---|---|---|---|---|
| マイク位置 [deg] | 0 | 45 | 90 | 135 | 180 | 225 | 270 | 315 | Mean |
| 密閉型 | -22.4 | -17.5 | -20.3 | -21.7 | -27.7 | -30.7 | -27.1 | -22.7 | -23.8 |
| 開放型 | -16.8 | -17.4 | -19.4 | -22.5 | -28.3 | -26.8 | -23.7 | -18.8 | -21.7 |
| 背面開口 d=4mm 相当 | -19.4 | -19.9 | -24.0 | -25.4 | -31.6 | -32.8 | -26.2 | -20.9 | -25.0 |
| 側面開口 d=2mm | -21.7 | -22.4 | -24.5 | -25.5 | -33.3 | -29.2 | -25.4 | -20.2 | -25.3 |
| 側面開口 d=3mm | -19.9 | -20.6 | -25.0 | -25.3 | -32.9 | -29.3 | -26.6 | -20.5 | -25.0 |
| 側面開口 d=4mm | -19.3 | -19.3 | -23.7 | -26.3 | -29.5 | -27.7 | -27.4 | -20.3 | -24.2 |
| 側面開口 d=5mm | -17.3 | -17.6 | -22.6 | -25.4 | -29.4 | -27.3 | -25.5 | -19.6 | -23.1 |
| 側面開口 d=6mm | -16.3 | -16.5 | -21.6 | -23.8 | -26.3 | -25.0 | -22.0 | -17.2 | -21.1 |

音漏れの大きさ [dB]　-40.0　-35.0　-30.0　-25.0　-20.0　-15.0　-10.0

Fig.1:Mean sound leakage level at each enclosure and each microphone in the low frequency band (upper) and the high frequency band (lower).

## 2-1Q-8

### 2-1Q-8　光波マイクロホンによる音圧測定の SN 比の改善

The improvement of the signal-noise ratio in the measurement of the sound pressure using the optical-wave microphone

◎山田桂輔，平野琴，高橋弘宜（産総研）

◆光波マイクロホンは光を利用した音響測定法の一つであり、レンズのフーリエ変換効果を利用して音圧による屈折率の変化を検出する。

◆音圧による空気の屈折率変化は非常に微少なものであるため、光波マイクロホンによる音圧測定では SN 比の低さが問題になる。

◆本研究では光波マイクロホンシステムの感度理論式を基にシステムの感度向上方法を検討すると共に、その効果を実験的に検証した。

◆図 1 に光波マイクロホンシステムのレーザビーム径 $w_0$ を変化させた時の音圧感度を実験的に求めた結果を示す。3.15 kHz 以上では、ビーム径の拡大により感度は一定比率で増加し、SN 比の改善が期待できることが示された。

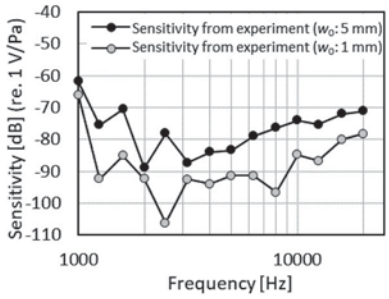



Fig.1:Experimental sensitivity of the optical-wave microphone. The sound pressure sensitivities (V/Pa) are represented by the decibel unit.

## 2-1Q-9

### 2-1Q-9　サーモホンと多入力型ΔΣ変調を用いた共通配線と素材のみの個別制御スピーカアレイ

Individually controlled speaker array constructed with only common wiring and materials using thermophones and multi-input ΔΣ modulator

☆田光史拓，武岡成人（静理工大）

◆3 次元音場の厳密な再生に向けて，多入力型ΔΣ変調(MIDS)を提案し，共通配線で多点を独立制御するスピーカアレイの研究を進めている。

◆MIDS 法を用いた個別制御スピーカアレイは出力部に乗算素子が必要となることが課題であった。

◆本報では，サーモホンの変換原理における乗算項に着目し，トランスデューサそのものに乗算出力機構を持たせることで，共通の電極と熱音響素材だけで構成できる個別制御スピーカアレイを提案する。

◆基礎実験として 16 チャンネルの個別制御スピーカアレイを試作し，各素子から個別の信号が出力されている様子が観察された。

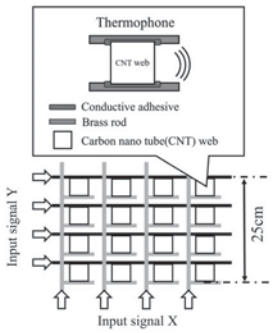



Fig. 1 Image of prototype 16ch speaker array

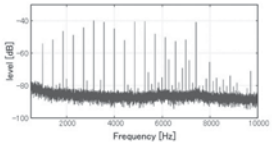



Fig. 2 Spectrum at center point

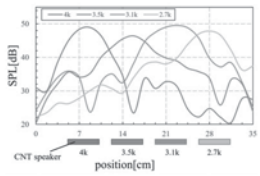



Fig. 3 Distribution of each frequency component

## 2-3P-1

### 2-3P-1　割り込み発話にも対応可能なリアルタイム話者交替システム

Real-time turn-taking system that can handle barge-in utterances

☆杉山雅和，西村良太(徳島大)，太田健吾(阿南高専)，北岡教英(豊橋技科大)

◆音声対話システムにおいて割り込み発話に対して話者交替判定を行うため，発話先頭 300ms のみを用いた話者交替システムを提案する。

◆提案システムでは入力されたユーザ発話が相槌，笑い声か否かを約 87%の精度で判定可能である。

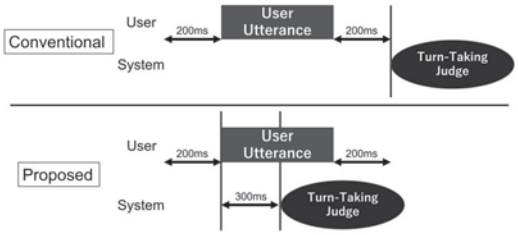



Fig.1：Timing of turn-taking judges in the proposed system.

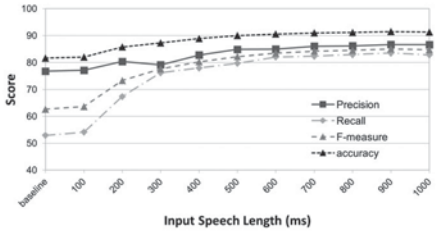



Fig.2：Relationship between input length and score.

## 2-3P-2

### 2-3P-2　Continuous Integrate-and-Fire を用いた音声区間検出とターン終了検知のマルチタスク学習

Multitask Learning with Voice Activity Detection and End-of-Turn Detection Using Continuous Integrate-and-Fire Model

○池口弘尚（名工大），東佑樹（AI Shift），上乃聖，李晃伸（名工大）

◆Continuous Integrate-and-Fire を用いて音声区間検出（VAD）とターン終了（EOT）検知のマルチタスク学習を行う手法を提案する.

◆判定タイミングを CIF で直接モデル化することで，EOT 検知の遅延低減を，VAD とのマルチタスク学習により精度の改善をねらう.

◆実験結果より VAD から EOT のカスケード型のモデルと比較し判定タイミングの遅延を低減することができた.

Table 1: EOT detection results (%) on CSJ lecture data. In the baseline model, we detect silence using a VAD model and then predict EOT based on the silence.

| | Precision | Recall | F1-score |
|---|---|---|---|
| ベースラインモデル | 81.49 | **70.26** | **75.46** |
| CIF(VAD の統合学習なし) | 79.72 | 65.15 | 70.42 |
| CIF(VAD の統合学習あり) | 80.09 | 49.17 | 60.13 |

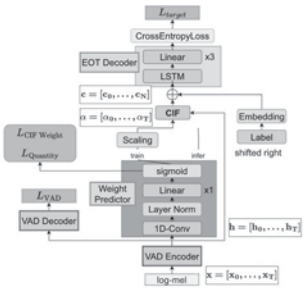



Fig. 1: Overview of the proposed architecture.

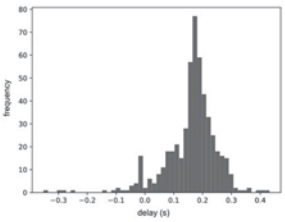



Fig. 2: Histogram of EOT detect timing (s). The average of the delay (0.16 s) is smaller than that of the baseline (0.2 s).

## 2-3P-3

### 2-3P-3　Transformer を用いた音声認識モデルにおける事前分布を用いた注意重みの平滑化の検討

Attention Weight Smoothing Using Prior Distributions for Transformer-Based End-to-End ASR

◎前角高史 (ヤフー)，藤田悠哉 (ヤフー)，
△Yifang Peng(カーネギーメロン大)，渡部晋治 (カーネギーメロン大)

◆注意機構は入力特徴系列間のグローバルな依存関係を考慮することで詳細な文脈情報を補足できるが、エンコーダの上位層の注意重みにおいて対角成分に大きな値が集中しやすい傾向がある。

◆注意重みの分布が急峻になると、特徴系列の局所的な文脈情報しか考慮されず、音声認識の性能を制限している可能性がある。

◆本稿では、注意重みに対してある事前分布を想定し、これを用いて注意重みを平滑化する手法を２つ提案する。

- 提案法 1. 学習可能な切断事前分布を用いた平滑化
- 提案法 2. 直前の層の注意重みを用いた平滑化

◆LibriSpeech と WSJ を用いた評価実験の結果、通常の注意機構を持つベースラインから最大で 7.9%の相対改善がみられた。

◆また、提案法によって上位層の注意重み分布の急峻なピークが緩和されることを確認した。

| | LibriSpeech | | WSJ |
|---|---|---|---|
| | test-clean | test-other | eval92 |
| Baseline | 8.3 | 20.5 | 11.7 |
| Proposed 1 | 8.3 | 20.1 | 11.4 |
| Proposed 2 | 8.2 | **19.9** | **10.9** |

Table.1: WER results

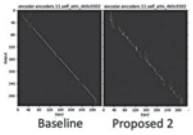



Fig.1: Heatmaps of Transformer model's self-attention weight

## 2-3P-4

### 2-3P-4　原言語テキストを補助入力とする Transformer 同時通訳音声認識

Transformer-based Automatic Speech Recognition with Auxiliary Input of Source Language Text for Transcribing Simultaneous Interpretation

☆谷口秀太，加藤恒夫，田村晃裕 (同志社大)，安田圭志 (マインドワード)

◆同時通訳訓練における音声認識は，音声に多くのフィラー，言い淀み，言い直しが含まれるため難しい反面，原言語情報を利用できる.

◆我々は Transformer に基づく音声認識に原言語テキストを補助入力とするエンコーダを追加したモデルを提案した.

◆ただし，大規模な原言語テキスト付き同時通訳音声コーパスは存在しないため，音声翻訳コーパス MuST-C を転用し，学習と評価を行った.

◆本研究では同時通訳音声を用いて提案モデルのファインチューニングと評価を行い，原言語テキスト補助入力の有効性を確認した.

◆また同時通訳音声においてフィラーをラベル付けし，認識精度を向上させる手法の有効性を確認した.

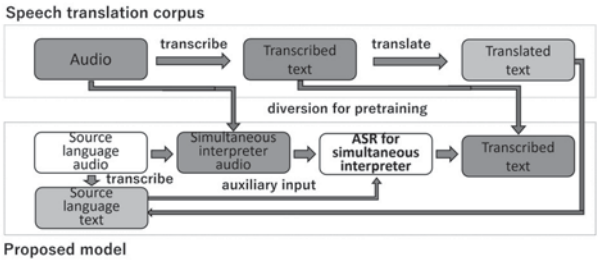



Fig.1：Proposed Model and Usage of Speech Translation Corpus for Pre-training Proposed Model

## 2-3P-5

### 2-3P-5　加齢により変化する音声の特徴に基づく話者年齢推定のためのデータ拡張の検討

Data augmentation for speaker age estimation base on acoustic features changed by aging

○北岸佑樹[†]，△小林奎太[‡]，△美馬亮太[‡]，俵直弘[†]，
小川厚徳[†]，浅見太一[†]，△森本健一[†]（NTT[†]，関西大[‡]）

◆DNN によって話者年齢推定の精度が高まりつつあるが，精度向上のためには幅広い年齢層で多数の話者の音声が必要

◆話者認識では speed perturbation や声道長正規化等による音声変換を用いた話者バリエーション増強法があり，話者年齢推定でも有効性が期待されるが，話速や声道長などは加齢に伴って変化すると報告される

◆このような手法を話者年齢推定で用いるには，変換後の音声の正解年齢を再定義することが望ましいが，厳密に定義することは困難

◆話者年齢と正の相関関係にある音声の対数パワーの標準偏差に着目

◆音声波形の対数パワーの標準偏差を増加させた音声の話者年齢を，元の音声と比べて老いて推定するモデル学習法を提案

◆実験の結果，年齢を変化させた音声を学習に加えて，かつ，元の音声と比べて推定年齢を高くするモデル学習によって，年齢推定精度改善

| モデル | 追加データ | 追加データの学習 | MAE |
|---|---|---|---|
| Baseline | なし | なし | 6.46 |
| Conventional | 音声波形の対数パワーの標準偏差を増加した音声 | 元の音声と同じように学習 | 7.18 |
| Proposed | | 元の音声より老いて推定 | 6.26 |
| Ablation | 元の音声のコピー | | 6.78 |

Table 1: Experimental results of speaker age estimation

## 2-3P-6

### 2-3P-6　朗読音声の発話スタイルを対象とした音響特徴量解析

Analysis of acoustic features for speaking style in recited speech

◎小口 純矢，森勢 将雅（明治大）

ROHAN コーパスの朗読音声データベースを用いて，発話スタイルによる音響的性質の違いを分析しました．

## 2-3P-7

### 2-3P-7　朗読における感情表現と発話様式プロファイル

Emotional Expression in Recitation and speech style profile

☆宮下隼（筑波大院），河原英紀（和歌山大），
榊原健一（北海道医療大学），寺澤洋子（筑波大）

◆本研究では朗読を行なった際に発話者が行う感情表現について、緊張ー弛緩の声質を示す声門開放時間率に着目して調査を行なった。

◆感情と緊張ー弛緩の関係性について調査することで、声色の変化が与える印象を明らかにし、感情音声に関する理解を深めることが本研究の目的である。

◆声門開放時間率を視覚的に比較できるようにするため、発話様式プロファイルである Valued Voice Range Profile（VVRP）を作成した。

◆結果として、音の高低や強弱といった韻律的特徴から類別が難しい感情同士を声門開放時間率の変化から類別できる可能性があることが示された。

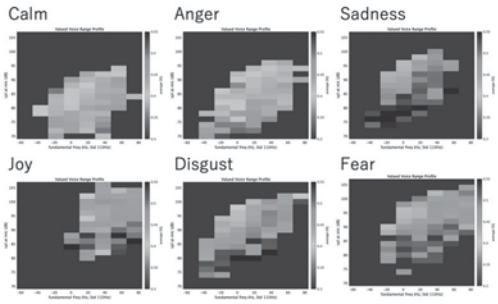



Fig.1: Valued Voice Range Profile for each of the six emotions

## 2-3P-8

### 2-3P-8　Transformer と GAN を用いた口唇動画音声合成

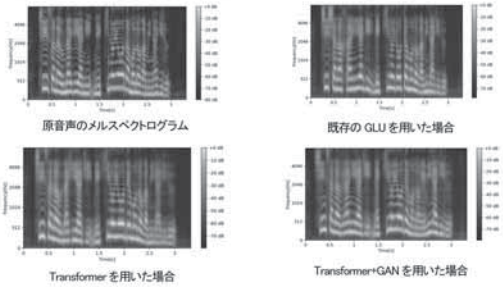
Lip-to-Speech Synthesis using Transformer and GAN Techniques

☆藤田直明，△南汰翼，鏑木時彦（九州大・芸工）

◆ 口唇動画から音響特徴量を予測する自己回帰モデルでは，了解可能な音声が生成可能であるが，イントネーション変化が乏しい，長時間の音声生成では誤差が累積するといった問題を抱えている。

◆ 既存手法の GLU を用いた自己回帰モデルでは，1フレーム前の予測音響特徴量の影響が大きいことから，生成した音響特徴量全体を考慮することができる Transformer を用いたモデルを構築した。この結果，音声評価指標において性能向上を達成した。

◆ また、scheduled sampling を用いて生成した音響特徴量と自己回帰的に推論した音響特徴量の差を低減するために，識別器を導入した学習を検討した。識別器は訓練時のアルゴリズムと推論時のアルゴリズムで生成した音響特徴量を識別する。GAN を用いることでさらにイントネーションの変化が大きい音声を生成可能となったが、音声評価指標は劣化した。条件づけや損失関数の見直しなどによる改善が求められる。



## 2-3P-9

### 2-3P-9 More differentiated pause insertion for phoneme-based multi-speaker TTS models

☆ † Dong Yang, ‡ Tomoki Koriyama, † Yuki Saito, † Takaaki Saeki, † Detai Xin, † Hiroshi Saruwatari
(† The University of Tokyo, ‡ CyberAgent, Inc.)

◆Pause insertion is an essential part of text-to-speech (TTS) systems because proper pauses significantly enhance the rhythm of synthetic speech. However, conventional pause prediction models ignore various speakers' different styles of inserting silent pauses, whose performance degrades when being trained with a multi-speaker speech corpus. Besides, most mainstream TTS models mainly use phonemes as input and treat all silent pauses as one phoneme, which leads to the duration of all silent pauses in synthetic speech not being sufficiently differentiated.

◆We propose a BERT-based pause insertion framework for multi-speaker phoneme-based TTS models, in which silent pauses are categorized by duration. Multi-task learning structure is used with injecting speaker embeddings to capture various speaker features.

◆The results of objective and subjective evaluations show that our pause insertion framework improves accuracy in predicting the position of respiratory pauses and makes the synthetic speech have better rhythm with more natural pauses.

## 2-3P-10

### ニューラルボコーダ合成音声の品質向上を狙った勾配法による後処理

Post-processing method for improving sound quality of speech signals generated by neural vocoder

☆高津航輝，矢田部浩平(農工大)，小泉悠馬(Google Research)

ニューラルボコーダの音声生成過程はモデルによってバラバラ

求 ニューラルボコーダの音声生成過程に依存しない合成音声の品質向上手法

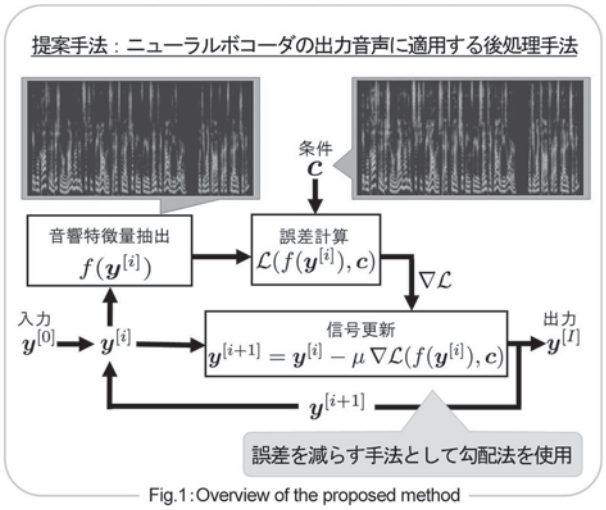



Fig.1：Overview of the proposed method

## 2-3P-11

### 2-3P-11 音声コーデックにおける前段までの量子化過程を考慮した残差ベクトル量子化器

Residual Vector Quantizer that takes into account the quantization process up to the previous stage in speech codecs

☆吉川晃登，立蔵洋介(静岡大)

目的：RVQ を用いた音声コーデックの性能向上

◆提案手法

RVQ の２つ目以降の量子化器の後段にポストネットを追加

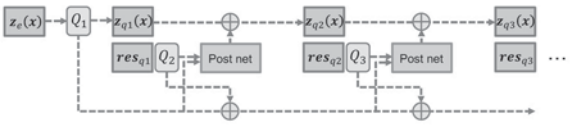



Fig.1：Overview of the proposed method.

◆MSE による評価

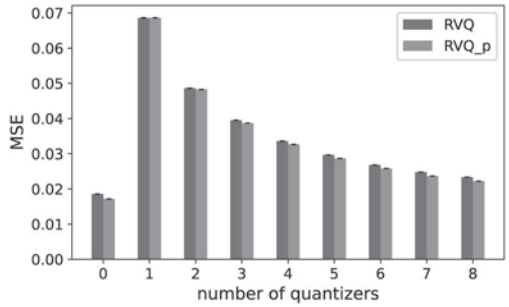



Fig.2：Comparative result of MSE.

従来手法と比較して量子化器を２つ以上用いた場合の性能が向上

## 2-3P-12

### 2-3P-12 音声合成における インターフェースを用いた 感情スタイル選択の試み

An Attempt to Emotional Style Selection Using an Interface in Speech Synthesis.

☆平川雄也，立蔵洋介(静岡大)

目的：感情音声合成の際の感情の制御性の向上

◆提案

感情スタイルマップを表示するインターフェースを追加した感情音声合成システム

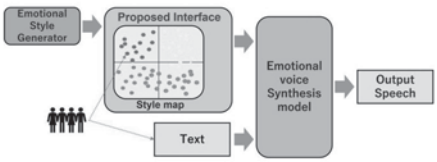



Fig.1 : Outline drawing of Proposed Interface.

◆主観評価実験：感情分類テスト

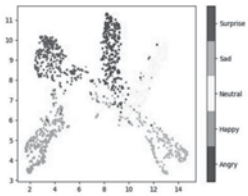



Table.1 : Classification Test Results

| Synthesis | Selected Emotion | | | | |
|---|---|---|---|---|---|
| Label | Neutral | Angry | Sad | Surprise | Happy |
| Neutral | 58.41 % | 10.23 % | 30.45 % | 0.68 % | 0.23 % |
| Angry | 59.00 % | 20.73 % | 16.91 % | 2.18 % | 1.18 % |
| Sad | 52.55 % | 7.73 % | 39.00 % | 0.45 % | 0.27 % |
| Surprise | 28.64 % | 25.45 % | 3.55 % | 27.09 % | 15.27 % |
| Happy | 19.36 % | 16.64 % | 10.45 % | 26.91 % | 26.64 % |

Fig.2 : Style Map on Proposed Interface

提案手法を用いた生成音声の分析を実施

個人間に生じる感情受け取りの差を確認，提案手法では対応が可能

## 2-3P-13

### 拡散確率モデルを用いた音声から顔画像への変換における改良・簡素化とその評価

Improvement and simplification of speech-to-face conversion using diffusion probabilistic models and their evaluation

◎加藤集平，橋本泰一（RevComm）

**提案手法**

- 音声からその話し手の顔を復元する試み（speech-to-face conversion）に**拡散確率モデル（diffusion probabilistic model; DPM）を導入**した。さらに，それを簡素化したシステムとあわせて性能を評価した。

**実験**

- 下記１・２の条件の組合せ（計４システム）を訓練・評価した。
  1. 入力音声の音響特徴量としてスペクトログラムを用いるか，Audio ALBERT を用いるか
  2. 音響特徴量から顔特徴量を予測する speech encoder を経由して顔特徴量をガイダンスとして DPM で顔画像を生成するか，音響特徴量を直接ガイダンスとして DPM で顔画像を生成するか

**結果**

- 生成画像と正解画像の顔の類似度・年齢および性別の再現度の観点から評価を行った結果，**Audio ALBERT & speech encoder ありが最も高い性能を示した。**

## 2-3Q-1

### 2-3Q-1　残差文埋め込みを用いた連続行動空間に基づく音声言語獲得エージェント

Continuous action space-based spoken language acquisition agent using residual sentence embedding

☆小松亮太，△木村友祐（東工大），岡本拓磨（NICT），篠崎隆宏（東工大）

◆本研究の目的は人間と同等の音声言語学習能力を備えたエージェントを計算機上で実現することである．

◆従来法では離散的な音声辞書から強化学習によって要素を選択して発話を行った．また，音声と画像を接地して音声辞書のうち視界に関連する要素に注意を向けた．しかし，発話方法は音声辞書を用いた録音再生であるため，単語を組み合わせて文生成できない制約がある．

◆そこで，本研究で提案するエージェントは複数単語を組み合わせるために発話生成に言語モデルを用いる．また，音声画像接地に基づく注意機構を言語モデルで実現するために残差文埋め込み法を提案する．

◆Fig.1 に示すように提案法は対話学習の効率が良いことが分かる．また学習時に現れない色タグと食べ物の組を提示された際にも平均約20%の割合で単語を組み合わせて発話できることを確認した．

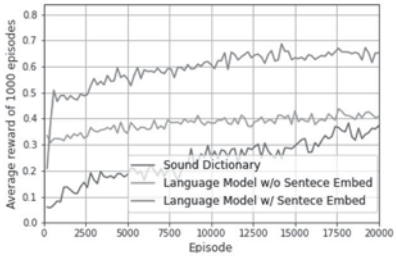



Fig.1 : Average reward in dialogue learning.

## 2-3Q-2

### 2-3Q-2　目的話者音声区間検出モデルの学習方法の検討

Investigation into the model architectures and training methods for personalized voice activity detection

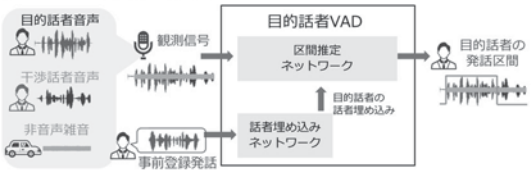
◎佐藤宏，品山健太郎，森谷崇史（NTT）

◆概要

・特定の話者の発話区間のみを推定する目的話者音声区間検出タスク

・モデル構造や学習方法の精度に対する影響を検証

・実データでの評価を実施




◆評価実験 (Average Precision [%])

・Conformer アーキテクチャの有効性を確認

・話者埋め込みネットワークの finetuning により性能が向上

| dataset | simulated | | real |
|---|---|---|---|
| | CSJ | Librispeech | Internal |
| LSTM | 93.2 | 90.3 | 74.9 |
| CNN+LSTM | 94.0 | 90.6 | 70.9 |
| Conformer | 94.4 | 91.2 | 82.6 |
| + fine tuning | 95.1 | 92.3 | 82.7 |

## 2-3Q-3

### 2-3Q-3　屋内音声通知システムを持続的に利用するための文章構造設計 −自治体からのメール内容の最適化の検討−

Sentence structure design for sustainable use of indoor voice notification system -Consideration of optimizing the content of e-mails from the municipality-

☆高橋稜，西山織絵（千葉工大院），苣木禎史（千葉工大）

◆本研究ではスマートフォンやスピーカー付きシーリングライトなどのスマート家電に付随するスピーカーおよびマイクロフォンを用いて屋内で音声として提示することを想定しており，国や自治体から配信されるメールをそのまま読み上げるだけでなく，人が確実に情報を理解できる文章構造および提示方法を研究している。

◆本稿では，メールの内容を要約して伝えるために文章設計について検討した。また，聞き取りやすさを考慮した文章の提示方法についても議論している。

◆人が聞き取れる，記憶に残りやすいといった観点から，音声で読み上げる文章の長さ，文章表現を制御する。

◆被験者を用いたリスニングテストの結果から一文に含まれる単語数は少ないほど聞き取り精度が高くなる傾向があると示された。

◆聞き取りやすさを考慮した文章の提示方法については，配信されるメールに含まれる日時や場所を絶対的な表現，相対的な表現の違いによる聞き取り精度，好みについてアンケート調査を実施した。

◆リスニングテスト，アンケート調査の結果，絶対的な表現か相対的な表現，もしくは両者のいずれかが好ましいのかは提示される情報によって異なるということが示唆された。

## 2-3Q-4

### 2-3Q-4 発話音声の感情を反映したテロップ画像の自動生成手法の検討

Automatic generation of subtitle text images reflecting emotions in speech utterances

☆中村史也（神戸大），相原龍（三菱電機），高島遼一，滝口哲也（神戸大），
△今井良枝（三菱電機）

◆ 人間は発話内容を理解する際，言語情報以外にも話者の表情や声の調子などから推定される感情情報を利用している。

◆ テレビや YouTube などの動画コンテンツでは，様々なフォントの字幕（テロップ）を用いて出演者の感情を表現することがしばしばある。

◆ 本研究では，発話に含まれる言語情報と感情情報を視覚的に表現するため，音声認識と音声感情認識による認識結果を GAN で学習したフォント画像生成モデルによりフォント文字画像として生成する手法（Fig.1）を検討する。

◆ 音声感情認識で得られる感情スコアにしたがってフォントの補間を行い，感情スコアが生成されるフォントに与える影響を調査した。

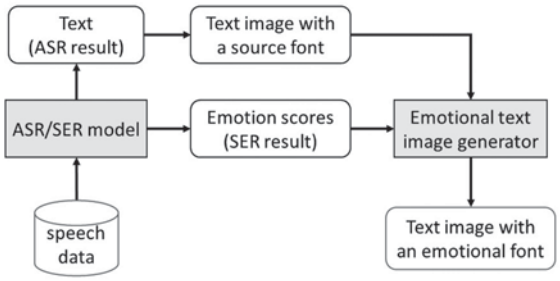



Fig.1 : Procedure for generating subtitle images from speech data.

## 2-3Q-5

### 2-3Q-5 感情音声認識における変調スペクトル特徴の雑音残響に対する頑健性の分析

Analysis of the robustness of modulation spectral features to noise reverberation in speech emotion recognition

☆郭 太陽，李 思俠，鵜木 祐史，岡田 将吾（JAIST）

◆背景　音声の振幅包絡線情報の変調成分（変調スペクトル）は人間の音声感情認識に重要な特徴である．著者らの検討から，この特徴が，雑音や残響に対して頑健であることがわかった（GUO et al., 2022）．

◆目的　**雑音残響環境下での機械による音声感情認識（SER）における変調スペクトル特徴（MSFs）の有効性と頑健性**を検討すること．

◆方法　雑音残響環境下で SVM を用いた SER 実験

- 刺激：富士通感情音声データ（5 感情）とその雑音駆動音声
- 雑音残響環境：SNR = −5 dB，残響時間 = 0.5，1.0，2.0 s
- 特徴量：単一の MSF，MSFs の組合せ，二つのベースラインモデルの特徴セット（IS09，eGeMAPS）

◆結果　**MSFs を利用した SER が雑音残響環境に影響を受け難いこと**
**MSFs の組合せを利用した SER が最も認識率が高かったこと**

◆結論　**MSFs が SER で重要な特徴であり，雑音残響環境で頑健な特徴である**ことを明らかにした．

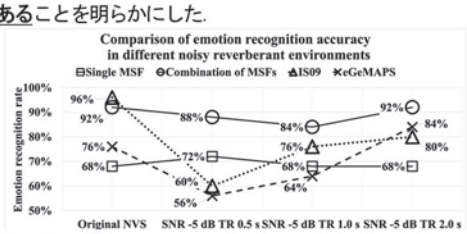



Fig.1 : Results of SER as reverberation time varied while SNR of −5 dB

## 2-3Q-6

### 2-3Q-6 ニューラルボコーダによる擬似ささやき声の生成に関する検討

A study on the generation of pseudo-whispering voice by neural vocoder

☆今井柊平，能勢隆，伊藤彰則（東北大）

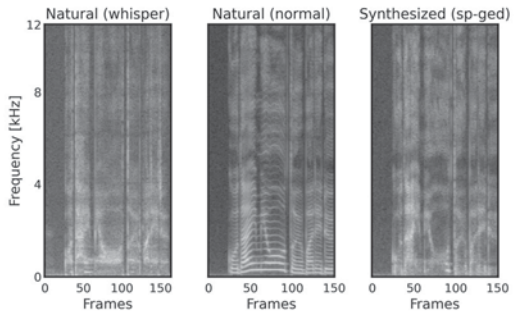
◆本研究では，通常音声で学習した TTS モデルにおいて推論時に擬似的にささやき声を生成することを目的とし，周期・非周期信号を用いたニューラルボコーダによる擬似ささやき声の生成を検討した．

◆TTS では通常，周期成分と非周期成分がどちらも含まれているメルスペクトログラムを音響特徴量として用いる．

◆擬似ささやき声のモデル化のためには，通常音声のメルスペクトログラムから周期・非周期成分を分離してモデル化することが必要となる．

◆本研究では，周期・非周期信号を用いるニューラルボコーダを GED 損失により学習することで，ガウスノイズから非周期成分をモデル化することを期待した．

◆実験により，自然性・ささやき声らしさともにリニアスケールのスペクトログラムに対し GED 損失を適用することの効果を確認した．




## 2-3Q-7

### 2-3Q-7 時変複素音声分析を用いた $F_0$ 推定における Hilbert 変換の検討*

A study of Hilbert Transform on $F_0$ estimation based on TV-CAR speech analysis

○舟木慶一（琉球大）

我々は、解析信号に対する $\ell_2$ ノルム正則化時変複素 AR(TV-CAR)音声分析を提案し、複素残差を用いた IRAPT 方式による $F_0$ 推定の性能向上のため、前処理として骨導フィルタ (Bone-Conduced filter: BCF)ならびに適応プリエンファシス(APE)の導入を行った。解析信号は実部に観測音声信号、虚部にその Hilbert 変換を有する複素信号であり、正の周波数領域でしかスペクトル成分を有しないため、歪みなく比 2 で間引きが行える。適応プリエンファシスは、池原らの方式で設計した(20,20)の ARMA フィルタを用いて Hilbert 変換を実現し、解析信号に対して行っている。しかし、適応プリエンファシスの場合、フレーム毎に 1 次の PARCOR 係数が変化するため、フレーム毎に Hilbert 変換を行い、解析信号を算出する方が理にかなっている。本稿ではフレーム毎に独立して解析信号を算出できる方式を用いて、適応プリエンファシスを行った音声信号に対して、フレームごとに解析信号を求め、TV-CAR 分析を行い、解析信号から複素残差を算出して、IRAPT 方式により $F_0$ 推定を実現し、従来手法との性能比較を行う。

## 2-3Q-8

### 2-3Q-8　最適輸送による GMM 補間を用いた中間属性の非実在話者生成

Generation of Mid-Attribute Non-existent Speakers by Gaussian Mixture Model Interpolation based on Optimal-Transport

☆渡邊亞椰，高道慎之介，齋藤佑樹，辛德泰，猿渡洋（東大院・情報理工）

◆話者生成は非実在話者の音声を生成するタスクである。話者生成を達成する既存手法として，多話者 TTS における話者埋め込み分布を性別等の属性に対応付けて GMM として学習し，学習した分布からのサンプリングをする TacoSpawn がある.

◆属性に対応した複数の学習済み話者埋め込み分布同士を最適輸送に基づいて補完した分布からは補間元となる属性の中間的な属性を持つ非実在話者の埋め込みが生成できることを確認した.

◆大規模な主観評価実験により，生成した中間的な属性の話者が一般的に中間的な属性を持つと知覚されること，ならびに補間前と比較して音質面で劣化が生じる可能性が低いことを確認した.

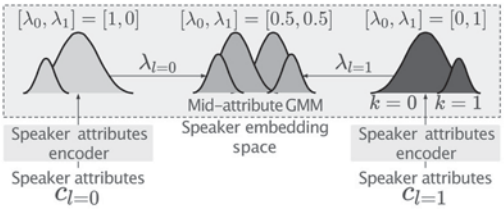



Fig.1 : Mid-attribute GMM, where L (number of speaker attributes) = 2, K (how many gaussian each GMM has) = 2. λ is interpolation rate.

## 2-3Q-9

### 2-3Q-9　小説オーディオブックの強調部分を学習に用いる抑揚制御可能な End-to-End 音声合成方式の検討

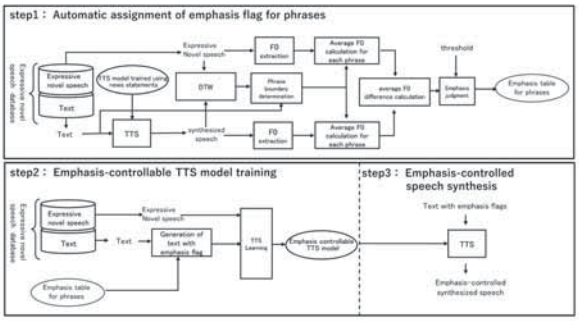
Emphasis Controllable End-to-End TTS Trained by Expressive Novel Speech Database

☆和田拓海，原直，阿部匡伸(岡山大院・HS 統合科学研)

◆本報告では，小説を朗読した音声とそのテキストを入力して合成した E2E-TTS 合成音声の比較を基に強調場所を決定し，この強調場所に強調フラグを付与して TTS を学習することで，音声合成時に強調したい場所を指定することができる TTS 方式を提案する.

◆客観評価実験では，強調フラグを付与したアクセント句の平均基本周波数は，付与していないアクセント句よりも高くなっており，強調されていることが明らかになった.

◆強調場所を特定する主観評価の結果，3 アクセント句，4 アクセント句の音声ともに最後のアクセント句を強調した場合の正答率が高いことが分かった.




## 2-3Q-10

### 2-3Q-10　インクリメンタル音声合成のための逐次読み・アクセント推定法の検討

Incremental prediction of pronunciation sequence and accent type for Japanese iTTS

○柳田智也（NAIST），中村哲（NAIST）

◆**インクリメンタル音声合成**

文より短い合成単位（単語やアクセント句）で音声を次々に合成
テキスト解析から音声合成までを逐次に実行

◆**問題点**

読み・アクセント型・形態素素性を逐次に予測する必要性

◆**逐次読み・アクセント推定法**

単一方向 LSTM により、表層単語系列から、
形態素情報とアクセント型を逐次推定

◆**実験条件および実験結果**

JSUT・JNAS・BTEC の翻訳データからテキストを取得
OpenJtalk より読みおよびアクセント型を取得

Table 1. Results of the prediction for morpheme

| Morpheme features | F score (macro) | Accuracy |
|---|---|---|
| Part-of-Speech tag | 0.914 | 0.977 |
| Reading style | 0.975 | 0.988 |
| Pronunciation | 0.882 | 0.967 |
| Accent type and the num. of mora | 0.417 | 0.895 |
| Concatenation rule of accent type | 0.984 | 0.965 |

Table 2. A result of the prediction for the accent phrase

| An accent phrase feature | F score (macro) | Accuracy |
|---|---|---|
| Accent type | 0.561 | 0.862 |

## 2-3Q-11

### 2-3Q-11　深層学習を用いた声質変換による自然対話中の笑い声合成

Laughter Synthesis in Human Dialogue by DNN Voice Conversion

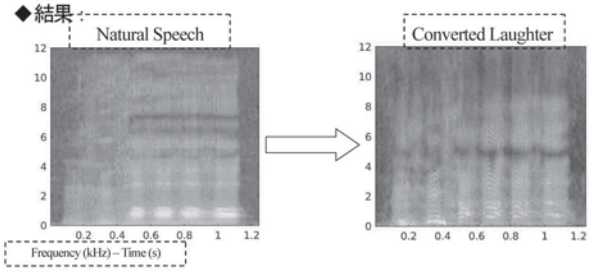
☆須藤智貴，近藤和弘（山形大・工）

◆背景：現在テキスト音声合成は広く利用されている。しかし合成音声を利用した対話を想定すると、非言語的な音声が多く含まれるためテキストに基づいた音声では不十分である。
特に笑い声は人間の感情表現や対話を円滑にするための行動として現れ、コミュニケーションを豊かにしている。

◆目的：声質変換によるアプローチで通常の発話音声から笑い声的な音声を合成することを試みる。
音声データは対話中に収録されたものを使用し、笑い声とそれ以外の音声の二つのドメイン間の変換を行う。

◆使用モデル：CycleGAN を使用した。笑い声を含む対話音声はパラレルデータを構築することが困難であり、ノンパラレルデータによるドメイン間のマッピングを学習するモデルが利用できると考えた。

◆結果：



2-3Q-12

# 2-3Q-12　サイクル学習を用いた注意機構付き VAE によるテキスト発話スタイル変換

Spoken text style transfer with conditional variational autoencoder and attention mechanism using cycle learning.

☆吉岡大貴, 安田裕介 (名大),
松永悟行, 大谷大和 (株式会社エーアイ), 戸田智基 (名大)

テキスト音声合成の研究において，感情やキャラクタ性などの「スタイル」を付与した音声を合成するために，音声だけでなく「発話するテキスト」のスタイルも制御可能な技術が有用であると考えられる．しかし，テキストのスタイル制御を行うテキストスタイル変換では，パラレルデータの作成コストが大きな課題となる．そこで我々は，ノンパラレルだがスタイル情報が付与されたデータを用いてテキストスタイル変換を行う CVAE + Attn + PE を提案した．このモデルでは Conditional VAE に「注意機構」と「位置埋め込み」を導入することにより，Bag-of-Words を用いる場合と比較して内容語保存性の改善に成功したが，一方で，スタイル変換性能に関しては未だ十分とは言えない．本稿では，人手によるパラレルデータを用意することなくスタイル変換性能を改善するため，モデルが生成した疑似パラレルデータを利用し，再構成と同時に循環再構成を行う「サイクル学習」を導入する．日本語のテキストスタイル変換タスクを対象とした実験的評価の結果から，本手法がコンテンツ保存性の劣化を最小限に抑えつつ，スタイル変換性能を改善できることを示した．提案モデルによる変換結果の例を Table. 1 に示す．

Table1. Disfluency style transfer example

| | |
|---|---|
| 変換前 | 文学研究としてはそれでも構わない訳ですが |
| ベースライン | 文学研究としてはもう構わない訳ですが |
| 提案手法 | 文学研究としてはえー構わない訳ですがまー |