

3-1-1

3-1-1 音声印象を操作可能な
Zero-shot 音声合成

Voice-Impression-Controllable zero-shot TTS

©藤田健一, 堀口翔太, 井島勇祐 (NTT)

- ◆ 音声印象を説明する形容詞対に関する低次元の音声印象ベクトルを導入し、音声印象を操作可能な zero-shot 音声合成手法を提案
 - 音声印象ベクトルは、各次元が音声印象に関する形容詞対の定量的なスコア (1 から 7 の連続値) に対応 (例えば 1:暗い-7:明るい)
- ◆ 音声印象の操作に関する主観評価実験
 - 互いの相関が低い「(B)男性的-女性的」「(E)迫力のある-弱々しい」「(F)若い-老けた」「(I)暗い-明るい」の4つの次元の値を操作前の値に対して-3 から +3 で操作
 - 主観評価では、印象制御前後の2音声聴取し、1つ目の音声に対する2つ目の音声を7段階で回答 (再生順はランダム)
 - ◇ スコアが4は制御前後で印象に差がなく、1, 7に近いほど操作後の音声に対応する印象に沿って変化していることを表す
 - 音声印象ベクトルを操作することで合成音声の印象を操作できることを確認

Table 1: 主観評価結果(1次元操作)

Dim	-3	-1.5	+1.5	+3
B	3.39±0.19	3.89±0.18	4.14±0.20	4.33±0.24
E	3.44±0.14	3.61±0.11	4.30±0.12	4.68±0.13
F	3.35±0.17	3.61±0.16	5.03±0.14	5.70±0.19
I	3.03±0.17	3.48±0.14	4.59±0.12	4.85±0.19

3-1-3

3-1-3 CycleSiFiNF-VC: ニューラルフォルマント合成を用いた音声制御可能なノンパラレル声質変換

CycleSiFiNF-VC: Voice-Controllable Non-Parallel Voice Conversion using Neural Formant Synthesis

©小林清流, 能勢隆, 伊藤彰則 (東北大), 小坂哲夫 (山形大)

- ◆ SSL モデルや LLM を用いた声質変換は、少量データで話者類似性の高い変換を実現する一方、声の明るさや音韻性など任意の音声特性を直観的に制御することは困難
- ◆ 本研究では、少量の音響パラメータを用いることで声質変換と音声変換を両立するノンパラレル声質変換手法 CycleSiFiNF-VC を提案
- ◆ Feature conversion network とサイクル学習により、声質変換しつつ、音響パラメータによる音声特性の制御を実現
- ◆ メルスペクトログラムを入力とすることで、事前に音響パラメータを抽出することなく、変換プロセス中の音声特性の制御を実現
- ◆ 結果として CycleGAN-VC2 と同等以上の変換品質と音響パラメータの独立した制御を達成

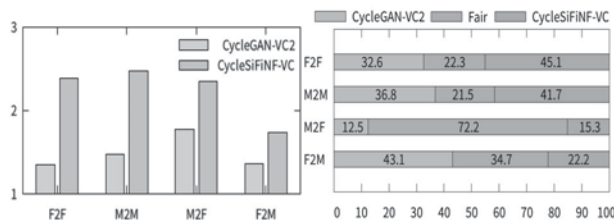


Fig.2: MOS for conversion quality

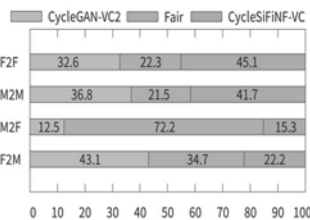


Fig.3: Results of XAB test for speaker similarity

3-1-2

相互情報量最小化による感情・音色の分離に基づく感情的音声合成

☆楊 家寧(東大院・情報理工), 李 勝, 篠崎 隆弘(東京科学大),
齋藤 佑樹, 猿渡 洋(東大院・情報理工)

従来、感情音声合成 (TTS) は参照エンコーダを用い、発話全体を代表するグローバルなスタイル・感情ベクトルを制御するアプローチが主流であった。しかしながら、この方式では参照音声に内在する微細な音響的特徴を捉えきれず、感情表現の精緻化に課題が残されていた。

本研究では、参照音声の固有属性を効果的に分離しつつ、音素レベルで詳細な感情埋め込みを予測可能とする新規な感情 TTS 手法を提案する。本手法では、スタイル分離フレームワークを導入し、音色抽出器と感情抽出器を相互情報量最小化の制約下で共同学習させることで、両者の特徴重複を排除し、異なるスタイル成分を明確に分離する。

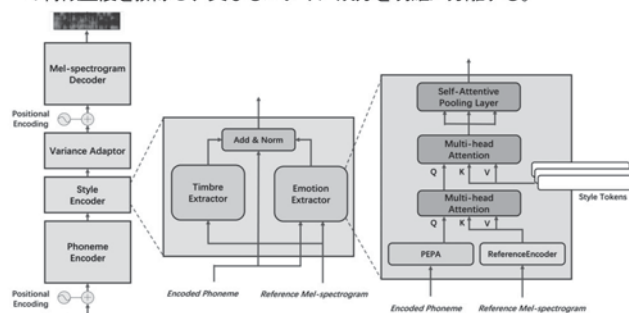


Fig. 1: Overview of the proposed TTS model

3-1-4

3-1-4 Any-to-any ストリーミング声質変換のための
離散音声ユニットを用いた自己教師あり表現学習
モデルからの知識蒸留

Knowledge distillation from self-supervised representation learning model with discrete speech units for any-to-any streaming voice conversion

©金川裕紀, 井島勇祐 (NTT)

- ◆ 自己教師あり表現学習(SSL)モデルに基づくノンパラレル声質変換

- テキストなしで学習でき、かつ頑健なコンテンツ抽出が可能 😊
- 非自己回帰モデル+高計算コストでストリーミング処理に不向き 😞

- ◆ 提案法

- SSL モデルの出力をベクトル量子化して得た離散音声ユニットを教師コンテンツとしてストリーミング生徒 VC モデルの学習に活用
- 下記の入出力を持つ3つのモジュールを end-to-end で学習
 1. スペクトル → コンテンツ (の事後確率)
 2. コンテンツ事後確率+話者ベクトル → 韻律 (F0、エネルギー、有声/無声フラグ)
 3. コンテンツ事後確率+韻律+話者ベクトル → 音声波形

- ◆ 実験結果の要約

ストリーミング処理動作しながら、教師である SSL モデルを用いたオフライン VC に迫る自然性や話者類似性を達成



Fig. 1: The training diagram for our streaming VC based on knowledge distillation from teacher content encoder using the SSL model

3-1-5

3-1-5 End-to-End 歌声合成における
楽譜情報を利用した韻律のモデル化の検討A Study on Modeling Prosody in End-to-End Singing Voice Synthesis
Using Musical Score Information

☆大野 友暉, 能勢 隆, 小林 清流, 伊藤 彰則 (東北大)

- ◆本研究では, End-to-End (E2E) 歌声合成の課題である韻律(ピッチ・継続長)推定精度向上を目指し, 代表的なE2E歌声合成モデルであるVISinger 2におけるピッチ及び継続長推定モデル改善を検討した.
- ◆具体的には, E2Eではない代表的な歌声合成モデルのSinsy, NNSVS等の知見を活用しVISinger 2に以下の改善を行った.
 - ピッチ推定精度向上のため, 楽譜音高からの差分F0予測, 学習データ拡張, 学習時のデータの確率的選択手法を導入した.
 - 継続長推定時の誤差累積問題の解決のため, 楽譜タイミングと実際の歌唱のずれをモデル化するタイムラグモデルを導入した.
- ◆RMSEを用いた評価により, ピッチ, 継続長共に推定精度改善を確認し, 提案手法の有効性が示された.

Table 1: Comparison of F0 RMSE [cent] for each method
across different pitch shifts [semitones]

	-12	-6	0	+6	+12
VISinger 2	841.1	212.4	84.4	541.6	1372.4
Differential F0 Prediction + Data Augmentation	177.4	125.4	88.5	132.9	215.3
+ Probabilistic Selection	105.8	88.1	88.8	87.0	122.8

3-1-7

3-1-7 Attentive Pooling を用いた話者照合における
話者埋め込みのアンサンブルの検討

Ensemble of Speaker Embeddings in Attentive Pooling-Based Speaker Verification

☆田中康平, 塩田さやか(東京都立大学)

- ◆【背景・目的】事前学習モデル(PTM)を利用した話者照合において, Attentive Poolingのみで構成される軽量のバックエンド(MHFA)が提案されている. 本研究は, MHFAを拡張しPTMの全ての層の情報を有効に活用することで従来法の性能向上を図る.
- ◆【提案法】複数のMHFAモジュールを並列構成し, 各モジュールがPTMの異なる層から話者情報を抽出してアンサンブルする. 層選択にはマスク付き重みと多様性ペナルティを用いる. 話者埋め込みを各モジュールで分割して抽出し, 従来法の軽量を維持する(Fig.1).
- ◆【実験・結果】VoxCeleb1データセットで評価を行った. 従来のMHFAと比較して約10%のパラメータ増加のみで約13%のEER改善を達成した. PTMの中間層(4-6層)からの情報が最も有効であることを確認した. PTMの各層からの情報がアンサンブルで有効活用されることを確認した(Table1).

Table 1 提案法のベースラインとの比較		
Backend model	Params	EER [%]
Baseline		
MHFA	4.34M	1.83 ± 0.08
MHFA[+hidden256]	8.63M	1.76 ± 0.08
MHFA[+attn128]	8.59M	1.93 ± 0.19
Proposed		
MHFA4[Fixed group]	4.78M	1.63 ± 0.12
MHFA4 ($\lambda = 0.0$)	4.78M	1.70 ± 0.10
MHFA4 ($\lambda = 10.9$)	4.78M	1.60 ± 0.07
MHFA2 ($\lambda = 13.0$)	4.49M	1.76 ± 0.17
MHFA8 ($\lambda = 7.8$)	5.37M	1.64 ± 0.10

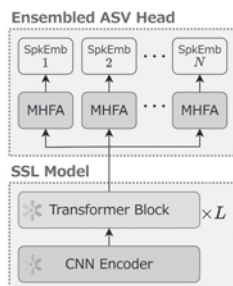


Fig. 1 提案法の概要

3-1-6

3-1-6 目的話者特徴量抽出における
非目的話者バイアスの抑制

Reducing bias from non-target speakers in guided speaker embedding

©堀口翔太, 芦原孝典, デルクロアマーク, 安藤厚志, 俵直弘(NTT)

- ◆目的話者特徴量抽出は, 発話の重複を含む複数話者音声から特定の話者の特徴量を精度よく抽出することを可能にした一方で, 重複の少ない場合に性能が若干低下してしまうという課題があった. 一般的な会話における発話の重複率は高々20%程度であり, この劣化による影響は無視できない.
- ◆本稿では, この劣化が非目的話者のみの発話区間が目的話者特徴量抽出に過剰な影響を及ぼしていることに由来することを確認した. これを解決するため, 目的話者特徴量抽出器のエンコーダにおいて, 系列全体から計算する統計量に基づいて計算を行う各モジュールにおいて, 目的話者の発話区間のみから計算した統計量を代わりに用いることで, 非目的話者の過剰な影響を防ぐ手法を提案した.
- ◆話者照合及び話者ダイアライゼーションにおける評価により, 提案手法の有効性を示した.

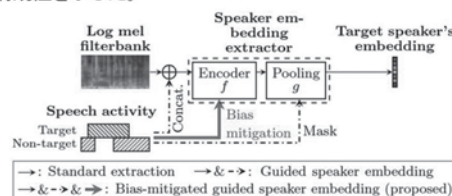


Fig. 1: Block diagram of standard single-speaker embedding extraction, guided speaker embedding extraction [Horiguchi+, ICASSP25], and the proposed bias-mitigated guided speaker embedding extraction.

3-1-8

3-1-8 PUNSER: 複数感情ラベル統合に基づく大規模
教師あり学習モデルによる音声感情認識PUNSER: Speech Emotion Recognition Using Large-Scale Supervised
Learning Models Based on Multi-Label Emotion Integration

©林崎由, 能勢隆, 小林清流(東北大), 深山寛(産総研), 伊藤彰則(東北大)

- ◆本稿では, ラベル形式が異なる10種類の感情音声コーパス(15万発話・226時間)を統合し, 大規模な教師あり学習を行うことにより, 事前学習(PT)済みモデルPUNSERを構築した.
- ◆事前学習にはカテゴリ感情ラベルと次元感情ラベルの両方を活用するマスク付きマルチタスク学習を採用し, ラベルの有無を表すマスクを導入して, ラベルが欠損している場合も学習を可能にした.
- ◆提案モデルは, 目標ドメインデータでファインチューニング(FT)することにより, カテゴリ感情認識と次元感情認識の両方で従来法を上回る認識率となり, SOTAの性能を達成した.
- ◆さらに, カテゴリ感情認識において提案モデルは, 目標ドメインデータを全く用いないドメイン外条件にも関わらず, 目標ドメインデータを用いた従来のSSLベース手法に迫る性能を示した.

Table 1 Comparison of emotion recognition accuracy on IEMOCAP corpus.

Method	Categorical				Dimensional		
	WA4	UA4	WA8	UA8	CCCv	CCCA	CCCD
emotion2vec base	0.587	0.521	0.409	0.253	0.561	0.426	0.453
emotion2vec+ large	0.452	0.377	0.362	0.209	0.413	0.128	0.289
WavLM large	0.667	0.592	0.558	0.362	0.682	0.527	0.527
DST	0.533	0.451	0.378	0.236	-	-	-
ShiftFormer	0.654	0.588	0.532	0.34	-	-	-
XEUS	0.672	0.604	0.555	0.359	0.682	0.54	0.531
PUNSER (PT)	0.579	0.555	0.417	0.328	0.167	0.027	0.231
PUNSER (FT)	0.672	0.618	0.584	0.401	0.685	0.581	0.542

3-1-9

3-1-9 音声コーパス構築のための
仮定を追加した発話区間検出法の改良An improved method for voice activity detection
with additional assumptions for speech corpus construction

☆俣野文義, 森勢将雅(明治大)

- 発話区間の最初と最後だけを推定する発話区間推定法を改良 (Fig. 1)
 - 時間に対して両方向にノイズ低減することによりクリーンに
 - 計算された二乗平均平方根に重みをつけることで収録開始直後終了直前に発生するノイズを更に低減
 - ゼロ交差数に上限を設けてクリップさせることで推定精度向上
- 実環境に近いピンクノイズを用いて精度を検証
 - 比較対象は信号処理ベース1種類, 機械学習ベース3種類
 - ノイズなしから SNR -5 dB までの全てで最も高精度
 - ただし音声がない場合も本来ない音声区間も判定してしまう

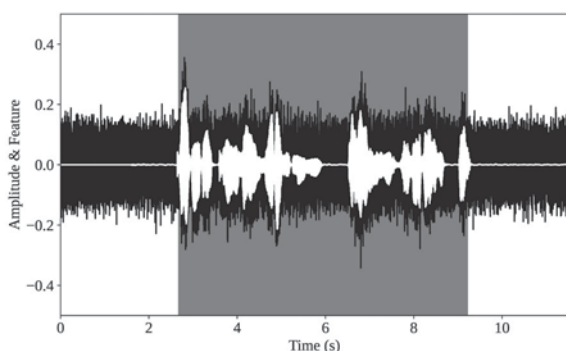


Fig.1: Clean signal (white line) and the signal after additive noise (black line), and speech-active intervals estimated from the features (gray shaded areas).

3-2-2

3-2-2 正中面の頭部伝達関数の空間的
卓越帯域が上昇角知覚に及ぼす影響

Effects of spatially dominant band in the head-related transfer function in the median plane on the perception of vertical angle of a sound image

©中村風香(千葉工大・院), 飯田一博(千葉工大・先進工)

HRTFの卓越帯域が上昇角知覚の手がかりとして寄与するか否かを音像定位実験により検証した。その結果を以下にまとめる。

- 正中面7方向のHRTFを分析した結果, 0°から上昇角が増加するに従って卓越周波数は4 kHz付近から9 kHz付近まで変化した。180°では1 kHz付近に出現した(Fig. 1)。
- 卓越帯域を除去したHRTFの音像定位精度は低下した(Fig. 2)。
- 平均上昇角誤差が本人HRTFよりも有意に大きくなった被験者の割合は, 60, 90, 120, 180°では80%以上であった。これは, 卓越帯域が上昇角知覚の手がかりとして寄与することを支持する結果であると言える。一方で, 30, 150°では有意な差が認められた被験者は33%に留まり, 検討の余地を残している(Table 1)。

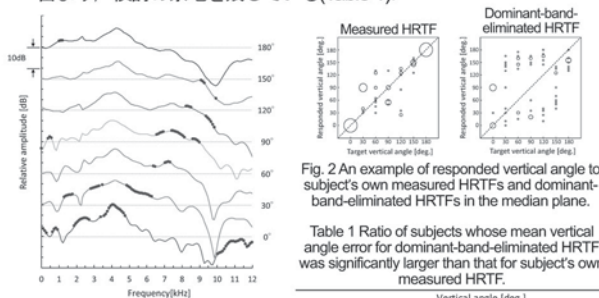


Fig. 1 Examples of HRTF and dominant frequency (closed blue circle) in the upper median plane for the left ear of a subject.

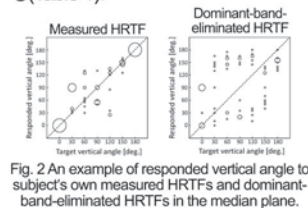


Fig. 2 An example of responded vertical angle to subject's own measured HRTFs and dominant-band-eliminated HRTFs in the median plane.

Table 1 Ratio of subjects whose mean vertical angle error for dominant-band-eliminated HRTF was significantly larger than that for subject's own measured HRTF.

Vertical angle [deg.]	0	30	60	90	120	150	180
significant difference	3/6	2/6	5/6	5/5	4/5	1/3	5/6

3-2-1

3-2-1 エネルギーの時間積分モデルと騒音下での
純音の反応時間

Time Integral Model of Sound Energy and Reaction Time of Pure Tones in Noisy Conditions

○下倉良太(大阪大), 添田喜治(産総研)

- ◆ Shimokura and Soeta (2022)は時間積分される音響エネルギーと反応時間との関係性を示した。
- ◆ そこで本研究は対象を純音とし, その包絡を3種類 (Fig. 1) に変化させることで時間積分の立ち上がりを調整し, この音響エネルギー-時間積分モデルが妥当であるのか検証した。またホワイトノイズとバンドパスノイズでこの純音をマスクし, 騒音環境下でもこのモデルが機能するのどうか検証した。
- ◆ 静寂条件では提示音圧レベルを10, 20, 30 dB SLと変化させ反応時間を計測し, ある音圧レベルに積分されるまでの時間 (t_{int}) と反応速度を比較したところ (Fig. 2), 精度良く相関した ($r=0.96$)。
- ◆ 騒音条件では純音へのマスキング量をもとに t_{int} を定義し, 反応速度と比較したところ, 静寂条件と同程度の相関を示した ($r=0.90$)。

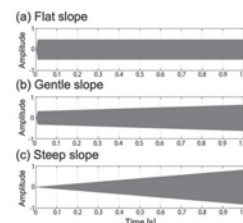


Fig. 1 Pure tones modulated in (a) Flat, (b) Gentle Slope and (c) Steep slope

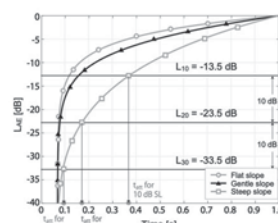


Fig. 2 Time-integrated sound energies of pure tones modulated in Flat, Gentle and Steep slopes

3-2-3

3-2-3 一側性難聴者の患側のメロディ聴取が
残響により妨げられることについて

Reverberation interferes with melody perception on the impaired ear side in individuals with unilateral hearing loss

©辻慎也, 荒井隆行(上智大・理工)

- ◆ 適度な残響により音楽の印象が向上することが知られているが, 高齢者や聴覚障害者では残響が音楽にもたらす効果は健聴者と違っていることが議論されている。
- ◆ 一側性難聴者 (UHL)・その発症直後の模擬 (MNH)・両耳聴 (BNH) を対象に, 無響/残響下でメロディの明瞭度 (MMR), その SRM を測定し, 次の結果を得た (Fig. 1).
- 1) BNH と比べ UHL は低いメロディの明瞭度を示した。一側性難聴の発症により音楽が「不明瞭」に感じられることに対するエビデンスが得られた。
- 2) BNH では残響の影響は有意でなかった一方で, MNH・UHL では残響が SRM 値を下げる有意な要因であり, その影響は患側で特に顕著であった。
- 3) MNH と比べ UHL では音の呈示位置による明瞭度への影響が小さく, 音楽の聞こえでは健側 / 患側で対称的な成績に近付くような適応が生じることが示唆された。

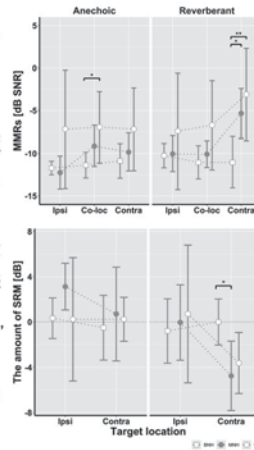


Fig. 1 Melody-to-masker ratio thresholds (MMRs; top row) and spatial release from masking (SRM; bottom row) for three hearing conditions: BNH (n=8), MNH (n=8), and UHL (n=7). Symbols and error bars indicate the mean and 95% confidence intervals, respectively, for each hearing condition, as described in the legend. **p<0.01, *p<0.05.

3-2-4

3-2-4 両耳提示された気導音の抑圧に向けた
骨導提示方式の検討Consideration on a setup/procedure of bone-conduction stimulation for
suppression of binaural air-conducted sound鳥谷輝樹(山梨大院・総研), 三木歩, 幸田怜奈(山梨大・工),
小澤賢司(山梨大院・総研)

- ◆背景: 耳を塞がない騒音抑圧の実現に向けた、骨導提示音による気導音抑圧の可能性を検討中 (Inoue et al., 2023).
→ 両耳純音刺激の抑圧量は約 5 dB に留まった (幸田ら, 2024).
- ◆目的: 気導提示純音抑圧に効果的な骨導提示方式の模索
- ◆方法: (1) 従来までの経皮骨導提示 (装置 A) に加え、軟骨導提示 (装置 B) による両耳提示純音の抑圧量を評価. (2) 左耳のみの純音提示に対しても抑圧量・外耳道内音圧 (ECSP) を評価し、両耳間の振幅・位相条件の相違の影響を推察.
- ◆結果: (1) 両耳提示において装置 A では平均 6.8 dB, 装置 B では平均約 4.4 dB で個人差も大 (Fig. 1). (2) 単耳提示では装置 A, B ともに抑圧量 8 dB 以上で、ECSP の減少も確認された (Table 1).
⇒ 両耳提示の際にも単耳での精緻な振幅・位相条件決定の必要性が示唆された. 軟骨導提示方式の有効性も示唆された.

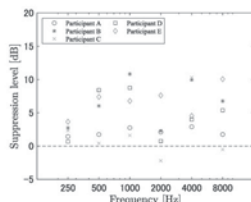


Figure 1: Mean suppression level for each participant using a transducer B.

Table 1: Suppression level (L_S) and decrement of ear canal sound pressure ($DECSP$) for 1000- and 2000-Hz pure tones.

	(1000 Hz)	L_S [dB]	$DECSP$ [dB]
Transducer A		8.2	6.5
Transducer B		9.9	8.6
	(2000 Hz)	L_S [dB]	$DECSP$ [dB]
Transducer A		15.0	14.1
Transducer B		15.0	11.0

3-3-2

3-3-2 物理制約ニューラルネットワークによる
波動方程式およびオイラーの運動方程式・
連続方程式を用いた音波伝播推定Physics-Informed Neural Networks Estimation of Acoustic Wave
Propagation Using the Wave Equation and the Euler Equations

©大久保翔太, △今野 智明, 堀内俊治 (KDDI 総合研)

- ◆音波伝播の時間領域解析において物理法則をニューラルネットワークに統合する Physics-Informed Neural Networks (PINNs) の適用が期待されている.
- ◆オイラーの運動方程式ならびに連続方程式を統合した PINNs および波動方程式を統合した PINNs を用いて二次元音場の音波伝播推定を行い FDTD 法との比較検証を実施した.
- ◆実験結果から、オイラーモデルが波動方程式モデルと比較して収束が速く、音波伝播推定の精度が高いことが確認された.

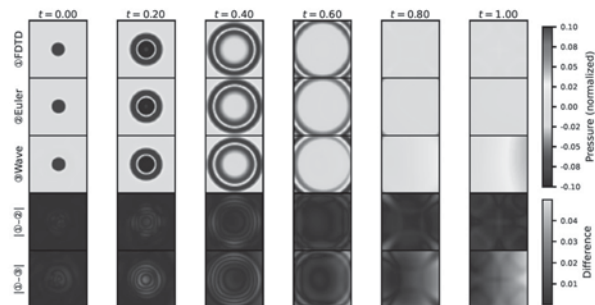


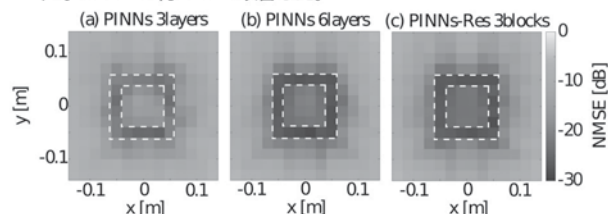
Fig.1: 固定端条件下のFDTDおよびPINNsによる音波伝播解析比較

3-3-1

3-3-1 残差接続と物理法則を用いた深層学習
による 3 次元室内インパルス応答推定Estimation of 3D room impulse responses
using physics-informed neural networks with residual connections

☆倉田駿, 佐藤元, 津國和泉, 池田雄介(東京電機大)

- ◆背景
・物理法則を用いた深層学習(PINNs)を用いて限られた測定信号から多数の信号を推定する手法が提案されてきた
・従来手法の PINNs の場合、モデルの表現力を上げるために隠れ層を深くした際、勾配消失が起き、推定精度が低下してしまう恐れがある
・本研究では、従来の PINNs に残差接続を導入したモデルを利用して 3 次元室内インパルス応答推定することで推定精度の改善を試みる
- ◆シミュレーション実験
従来手法(a), (b)と提案手法(c)を用いて 80 点の測定信号から $14 \times 14 \times 4$ (784) 点の信号を推定した結果、従来手法よりも提案手法は推定精度が改善した。特に Fig. 1(b), (c)を比較すると、内挿部分において平均 NMSE が約 1.1 dB 改善した.

Fig.1: NMSE map in the x-y plane ($z = -0.03$ m) (White lines indicate microphone area)

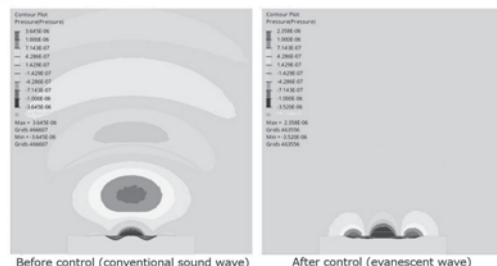
3-3-3

3-3-3 エバネッセント波を用いた
個別音響再生システムの開発Development of a Personalized Acoustic Reproduction System Using
Evanescent Wave

○佐野浩司, 志水達也, 宮本泰憲, 藤本圭祐(パナソニック)

近年、車室内やオフィス空間において、プライベートな音響環境へのニーズが急速に高まっている。本研究では、従来のエバネッセント波生成技術が抱えていた装置の大規模化や性能限界といった課題に対し、「流体力学的短絡現象」を活用した新しいアプローチにより、小型かつ高性能な個別音響再生システムを開発した。

本システムは、単一平板上に複数のアクチュエータを配置したシンプルな構造でありながら、高精度なエバネッセント波の生成を実現する。本技術は、既存の音響機器に容易に組み込める構成と制御手法を備えており、実用化に向けた準備が整っている。自動車・航空機・オフィス・家庭など、必要な人に必要な音だけを届ける多様な応用分野への展開が期待される。

Fig.1: Comparison of sound pressure distribution
(Participants □: Spring, ■: Fall, Papers ○: Spring, ●: Fall)

3-3-4

3-3-4 時間領域における深層学習を用いた音場推定と音場再現への応用

Sound field estimation using deep learning in time domain and its application to sound field reproduction

☆堀越光樹, 佐藤元, 津國和泉, 池田雄介(東京電機大)

◆研究背景

- Pressure Matching 法は高周波数の制御には多数の制御点が必要となり測定が困難である
- 単一の周波数領域において、少数の測定点から制御点における伝達関数を推定する深層学習手法が提案されてきた

◆提案手法

- 全層畳み込みネットワークを用いて、少数の測定点で取得した室内インパルス応答(RIR)から多点の RIR を推定する手法の提案

◆シミュレーション実験

- 9 (3×3)点の測定点から 169 (13×13)点のインパルス応答を推定した結果、提案手法の推定精度は平均 NMSE -43.7 dB 得られた
- 推定音場を用いて PM 法で音場を再現した結果、従来の PM 法よりも高精度に再現できており、スパース等価音源法(C - ESM)を用いた PM 法とほぼ同等の再現精度が得られた

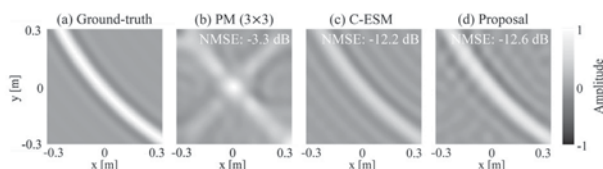


Fig.1: Desired and reproduced sound fields at 4.9 ms

3-3-5

3-3-5 スピーカアレイを用いた近接4点法のデータに基づく平面波合成

Plane wave synthesis using loudspeaker array based on the data from closely located four points microphone method

☆江里口裕月, 坂口智弘, 及川靖広(早大理工)

- ◆従来の音場再現技術は直接音と反射音の分離が困難であったり、反射音一つ一つを厳密に再現できていなかったりすることがある。
- ◆これまでに、近接4点法で取得した仮想音源分布から反射音の到来方向を求め、仮想音源が非常に広い領域に分布していることから、受聴領域には平面波が到来すると仮定し、それらを合成することで音場生成を行う手法を提案した。
- ◆今回はこの提案手法に基づき、スピーカアレイを用いた平面波合成を行った。MATLAB を用いて波面の時間変化の確認や、これまでの提案手法で生成した音場との比較を行っている。

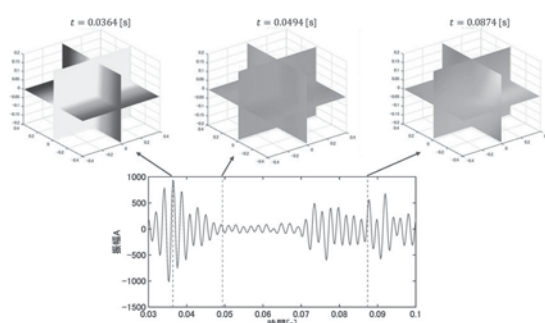


Fig.1 Time variation of wave front

3-3-6

3-3-6 CSP 重み付け全帯域使用 MUSIC 法に基づく音源数未知状況下の音源方向推定

Sound Source Direction Estimation under Unknown Number of Sources Based on CSP-weighted All-Bands MUSIC

☆篠山俊介, 志賀芳則, 岡山健仁(東京電機大院)

- ◆一般的な音源方向推定技術は音源数が既知であり、マイクロホン数未達の音源数に限られる。
- ◆本研究では、2マイクロホンで複数の音源方向を推定する。更に、音源数が未知である状況を想定する。
- ◆提案手法では、MUSIC 法と CSP 法を併用する。全帯域を使用した MUSIC スペクトルに対し、CSP 係数で作成した CSP 重みを付加することで、信頼性の高いスペクトルのみを使用して推定を行う。提案手法の妥当性を示すため実環境実験を実施した。Fig.1 に CSP 重み付け全帯域使用 MUSIC スペクトルの例、Fig.2 に全帯域使用 MUSIC と CSP 重み付け全帯域使用 MUSIC の音源数推定結果を示す。

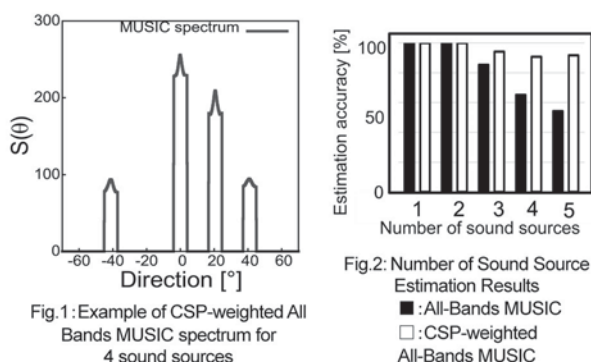


Fig.1: Example of CSP-weighted All Bands MUSIC spectrum for 4 sound sources

Fig.2: Number of Sound Source Estimation Results
■: All-Bands MUSIC
□: CSP-weighted All-Bands MUSIC

3-3-7

3-3-7 周波数帯域の取捨選択による音源定位性能改善

Improvement of Sound Source Direction Estimation Performance by Frequency Band Selection

☆橋本唯利, 志賀芳則, 岡山健仁(東京電機大院・工学研)

- ◆音源定位手法として GRODMIC が提案されている。この手法はマイクロホン間の伝搬遅延特性を近似し推定を行う。しかし、無音帯域で、近似対象が消失し、推定精度が劣化する問題がある。この問題に対し、有音帯域のみでの推定が必要であるとの主張に至った。本研究では、周波数帯域の取捨選択による有音帯域のみを使用した GRODMIC を提案する。
- ◆提案法ではフィルタバンクにより信号を分割し、分割した信号から有音信号のみを使用し推定を行う。
- ◆提案法の妥当性を示すため、実環境実験で従来の GRODMIC と比較を行った。Fig.1 に推定誤差を示す。
- ◆Fig.1 より提案法による大幅な性能改善が確認でき、主張の妥当性が示された。

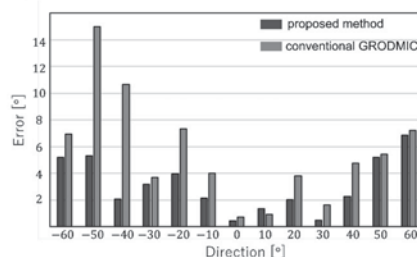


Fig.1: Errors of estimation using conventional GRODMIC and proposed method

3-3-8

3-3-8 少素子スピーカレイ群を用いる
音源位置・放射方向の提示

Presentation of sound source position and radiation direction
using small-number loudspeaker arrays

☆岸村 颯馬 (筑波大・院・シス情工),

若槻 尚斗, 善甫 啓一, △前田 祐佳 (筑波大・シス情系)

- ◆ビデオ通話において発声の向きを聴覚的に表現するにあたり、従来手法である波面合成よりも簡易で少素子な手法を目指す。
- ◆その一検討として、2 個のスピーカから成るスピーカレイが 2 対あるとき、その中間に位置する仮想音源の放射方向を 180° の範囲で提示することを試みた (Fig. 1)。
- ◆結果として、放射方向の提示の精度は少素子であることとトレードオフであると確認された。また、その妥協点を音圧分布および両耳間レベル差の数値シミュレーションから求められることが示された (Fig. 2)。

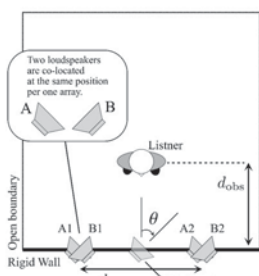


Fig. 1: Proposed loudspeaker array system and 2D sound field for simulation.

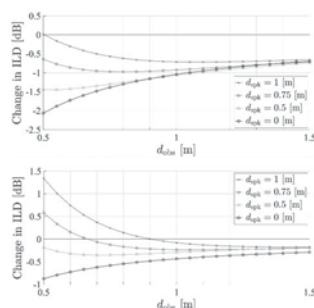


Fig. 2: Changes in ILD between directions to present.

3-4-1

3-4-1 主観評価実験による人声のステアリングの
汎用化に向けた特徴量の抽出

Extracting features for generalizing steering of the human voice based on
subjective evaluation experiments

○木下光太郎, 杉本岳大 (NHK)

- ◆AR / VR などの視聴時に音源の放射特性を再現することで、音源の向きに依存した音色の変化を表現するステアリングを検討している。
- ◆ステアリングでは事前に話者の放射特性を測定しておく必要があるが、それには特殊な測定環境が必要であり、コンテンツ制作の度に放射特性を別途測定することは汎用性に乏しい。
- ◆そこで、事前に放射特性を測定せずとも、これまで構築してきた他者の放射特性データベースを活用することを考えた。
- ◆本稿では、主観評価実験をとおして話者本人および他者の放射特性で生成したステアリング音声の比較聴取を行い、他者の放射特性を適用したときの品質を確認した。
- ◆また、主観評価実験の結果から放射特性の汎用化に資する特徴量を抽出し、ステアリングの汎用化手法を検討した。

Table 1 Correlation coefficient between spectral difference of filter
and average evaluation value

Direction condition	Correlation coefficient
($180^\circ, 0^\circ$)	-0.82
($90^\circ, 0^\circ$)	-0.85
($135^\circ, -30^\circ$)	-0.89
($45^\circ, 45^\circ$)	-0.81

3-3-9

3-3-9 特定の到来方向成分を強調する音場の
変換処理とその応用

Directional Filtering of Sound Fields for Emphasizing Specific Directions of
Arrival and Its Applications

☆村上涼, 植野夏樹 (熊本大学 大学院自然科学教育部)

- ◆空間的な音響情報を保持しつつ、特定の方向から到来する成分を強調する音場の変換処理手法を提案する。
- ◆提案手法は、到来方向に基づく重み付けを行う作用素として定式化される。
- ◆音場を球波関数の重ね合わせとして表現した場合において、von Mises-Fisher 型の重み関数のもとで、提案手法の閉形式表現を導出する。
- ◆提案手法の応用として、マイクロフォンアレイを用いて推定された音場に対して提案手法を適用する。
- ◆数値実験により、音源位置などの情報を含まない音場に対しても、特定の方向成分を強調できることを確認する。

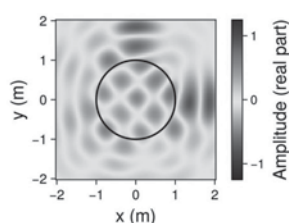


Fig. 1: The real part of the estimated sound field.

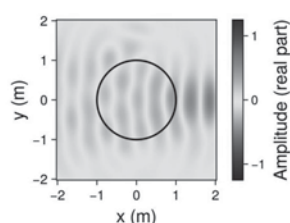


Fig. 2: The real part of the estimated sound field after applying the proposed method.

3-4-2

3-4-2 発話明瞭度の評価が異なる話者の
調音運動の差異

Differences in articulatory movements between speakers with
different speech intelligibility ratings

☆小松梨華, 竹本浩典 (千葉工大), 能田由紀子, 前川喜久雄 (国語研)

- ◆発話明瞭度の異なる話者 4 名が「北風と太陽」を朗読した際の舌と下顎の調音運動を Wave センサで計測し、その違いを検討した。
- ◆各センサの軌跡を正中矢状断面へ射影した結果、発話明瞭度の高い話者では下顎中切歯に装着した LI センサの動きが大きかったが、発話明瞭度の低い話者では小さかった (Fig. 1)。
- ◆下顎の動きをキャンセルした舌上のセンサ T1~T3 の移動ベクトルと LI の移動ベクトルのなす角 θ を $0-45^\circ$, $45-90^\circ$, $90-135^\circ$, $135-180^\circ$ の 4 区間に分類し、その分布の比率を算出した。
- ◆その結果、 $0-45^\circ$ の比率は発話明瞭度の高い話者では約 16%, 低い話者では約 27% となり、下顎の移動方向に対する舌の移動方向の独立性は発話明瞭度の高い話者で高かった。
- ◆これらの結果は、明瞭度の高い話者では下顎を大きく動かし、下顎と舌の動きが独立しているという先行研究と一致した。

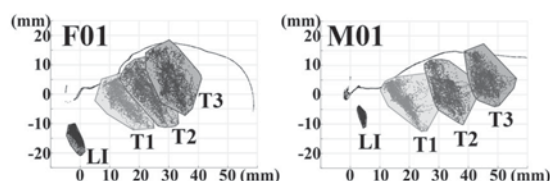


Fig. 1: 発話明瞭度の高い F01 と低い M01 のセンサ軌跡

3-4-3

3-4-3 wav2lip を用いた唇動画像が音声
理解度に与える影響の分析Analysis of the effect of lip-syncing videos generated
by wav2lip on speech intelligibility

○鈴木 基之, △佐藤 優多(大阪工大・情報)

研究の背景と目的

生成 AI を用いて音声データから生成した唇動画像を提示すること
で、音声を聞き難い状況において理解度がどの程度向上するか検証

音声理解度の計測実験

- 模擬難聴音声を作成し、聞き取り実験を実施
- 以下の3種類の試料を提示
 - － 音声のみ(“ao”)
 - － wav2lip で生成した唇動画像を同時に提示(“vv”)
 - － 話者の唇動画像を同時に提示(“rv”)

聞き取り実験の結果

- 提示方法によって、聞き取れない単語数に差
 - － “ao” は “rv” の2倍程度
 - － “vv” は “rv” ほどの性能はなく、効果に差
- 唇動画像の提示は母音の聞き取り(cv→c'v)に有効
 - － “vv” は “rv” と遜色ない効果

Method	Misheard words	Substituted morae		
		cv→cv'	cv→c'v	others
ao	14.7%	14.3%	30.1%	55.7%
vv	11.4%	8.5%	42.6%	48.9%
rv	7.5%	11.5%	38.5%	50.0%

3-4-5

3-4-5 避難する気になる
ボイストレーニングシステム構築に向けて
-Deep Learning による評価のフィードバック-Developing a Voice Training System to Promote Evacuation:
Feedback from Deep Learning-based Scoring○高野佐代子、長塚全、土田義郎
(金沢工大、Zen Voice Factory、金沢工大)

Sayoko TAKANO, Zen NAGATSUKA and Yoshio TSUCHIDA

- ◆ これまでに開発した VTuber (アバター) による教示のシステムに加えて、Deep Learning を用いて、危機感の有無を評価し、即時にフィードバックが得られるトレーニングシステムを開発する。
- ◆ トレーニング後はより危機感を感じる音声になり、話者6人中5人はフィードバックありの11回目の方がなしの11回目と比べて、有意に危機感が感じられる音声に改善した。

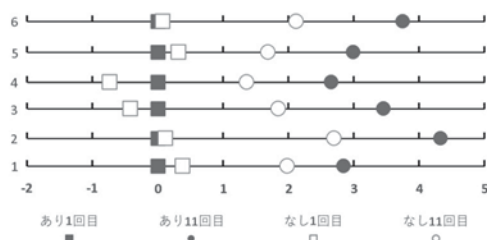


Fig. 1. The results of the speech training with/without Deep Learning based scoring.

3-4-4

3-4-4 ポップアウトボイス生成時における
声帯振動の検討Investigation of factors contributing to vocal fold vibration in pop-out voice
production

☆相馬巧海, 竹本浩典(千葉工大)

- ◆ 背景雑音などの妨害音がある環境下でも聞き取りやすい音声をポップアウトボイスと呼ぶ。本研究では、声帯振動がポップアウトボイスの生成に与える影響を検討するために、6名の男子大学生(M1~M6)を対象として、通る声(PV)と通らない声(NV)で単母音を発声した際に音声の録音とEGGを用いた声帯の開閉パターンの計測を実施し、音声波形とEGG波形それぞれから計算したfoとjitterを精度の検証をした上で比較し、またOQ, MA, HNRを比較して分析した。
- ◆ 音声波形とEGG波形から計算したfoとjitterでは、foには有意差が見られなかったが、jitterは音声波形から計算した方が有意に小さく(t検定, p<0.05)、音声波形、EGG波形それぞれでPVの方が有意に小さかった(t検定, p<0.05)(Table 1)。
- ◆ さらに、EGG波形の振幅やOQの分析から、PVの喉頭動態は大きな声を発声する際の特徴と一致していたことから、PVは小さな声でも大きな声と同様の喉頭動態を持つ声である可能性が示唆された。

Table 1 音声波形(左)とEGG波形(右)から抽出したjitterのPV-NVの値(%)。

灰色は負である値、黒色はEGG波形で音声波形と正負が逆転している値。

	/a/	/i/	/u/	/e/	/o/		/a/	/i/	/u/	/e/	/o/
M1	-0.35	0.14	-0.05	0.02	0.04	M1	-0.06	0.20	0.10	0.09	0.02
M2	-0.37	-0.24	-0.27	-0.61	-0.73	M2	-0.43	-0.27	-0.19	-0.62	-0.88
M3	-0.10	-0.14	-0.03	-0.09	-0.06	M3	-0.06	-0.16	0.02	-0.06	-0.04
M4	-0.19	-0.10	-0.17	0.01	0.04	M4	-0.08	0.29	0.32	0.15	0.10
M5	-0.04	-0.06	-0.22	-0.15	-0.28	M5	-0.10	-0.06	-0.24	-0.05	-0.29
M6	-0.12	-0.25	0.00	0.03	-0.23	M6	-0.15	-0.27	0.14	0.05	-0.25

3-4-6

3-4-6 発話開始時の構えが直音と拗音に
与える影響Effects of initial articulatory posture on palatalized and non-palatalized
consonants

☆朝倉麗仁, 竹本浩典(千葉工大), 前川喜久雄(国語研)

- ◆ 本研究では、拗音「キャ」の生成で硬口蓋のどの部分と舌が接触して破裂することが本質的に重要を検討するため、4名の話者が直音(/ka/), 拗音(/kja/), 「カ」の構えから「キャ」(/k>kja/), 「キャ」の構えから「カ」(/k>ka/)を発話する際の調音運動をfMRIで記録し分析した。
- ◆ /k>ka/の子音の構えでは舌は軟・硬口蓋に接触していたが、舌は前方より口蓋から離れ、軟口蓋のみと接触した状態で破裂した(Fig. 1)。
- ◆ /k>kja/の子音の構えでは舌は軟口蓋と接触していたが、舌は前方へ変位して硬口蓋後部と接触してから破裂した。
- ◆ 以上の結果は4名で共通していた。
- ◆ すなわち、拗音/kja/の生成では、硬口蓋全体に舌を接触させる必要はなく、硬口蓋後部に舌を接触させて、そこで破裂させることが本質的に重要であることが示唆された。

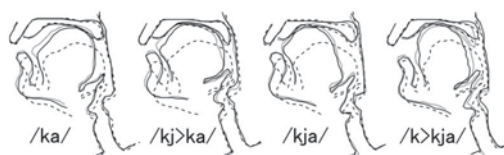


Fig. 1: /ka/, /kja/、/kja/、/k>kja/の調音運動
(点線: 構え, 実線: 破裂直前, 破線: 母音)

3-4-7

3-4-7 雑談における発話速度と発話量の関係
—50代女性は本当におしゃべりか—

The Relationship Between Speaking Rate and Speech Quantity in Chat

— Are Women Aged 50–59 Truly More Talkative? —

◎梁 辰, 倉片 憲治(早稲田大・人間科学)

- ◆年齢別・男女別の発話速度を測定した筆者ら以前の研究から、50代女性は雑談において特異的に速い発話速度を持つことが明らかになった。
- ◆本研究では引き続き『日本語日常会話コーパス』を用いて、50代女性の発話量における特徴について考察を行った。
- ◆その結果、50代女性は雑談において、発話単位数が多く(図1)、すなわち頻繁に口を挟むと同時に、会話に占める発話時間が長く(図2)、かつ、1回の発話時間が長い傾向が示された。

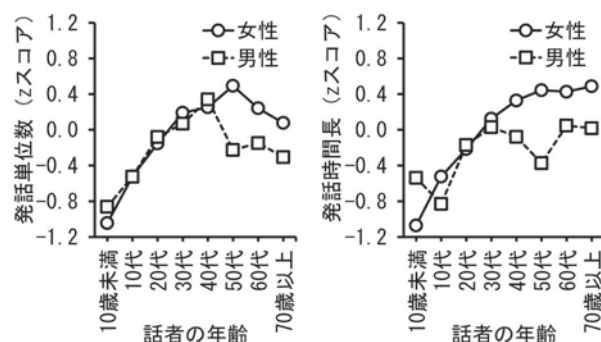


図1 男女別・年齢別の発話単位数

図2 男女別・年齢別の発話時間長

3-4-9

3-4-9 参画型コンピュータ支援言語学習システム
を用いた構文訓練における
修正フィードバックの効果検証Effect of Prompt Corrective Feedback in Learning Syntactic Form Using
Triologue-based Computer-Assisted Language Learning System

☆田中龍弥, 山本誠一, 加藤恒夫, 田村晃裕, 菅原真理子(同志社大学)

- ◆第二言語(L2)によるコミュニケーション能力の獲得には十分な訓練が必要であるが、教室形式では十分な時間を確保できない。
- ◆その不足を補う手段として対話型コンピュータ支援言語学習(Dialogue-Based Computer-Assisted Language Learning)システムが注目されている。
- ◆教室の訓練では、表現形式に焦点を当てた修正フィードバック(corrective feedback: CF)が重要な役割を果たす。
- ◆本研究では、二体のキャラクタと会話訓練を行う CALL システムに Whisper 音声認識と GPT-4o を用いて修正フィードバックを提示する仕組みを実装した。
- ◆日本人大学生 55 名を対象に修正フィードバックの有無による学習効果と定着度を比較評価したところ、訓練直後から約 1 か月間はフィードバック群の正答率がフィードバック無し群を有意に上回った。



Fig.1 : Image of CALL system

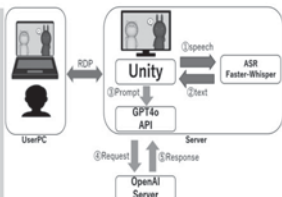


Fig.2 : System configuration

3-4-8

3-4-8 発話のしにくさと性格特性に関する日中比較調査

A Comparative Study of Speech Difficulty and Personality Traits between
Japanese and Chinese Speakers

◎孫静(神戸大/甲南大), 北村達也, 川上翔太(甲南大)

- ◆本研究では、日本語母語話者(JNS)大学生 423 名と中国語母語話者(CNS)大学生 453 名を対象に、Big Five 尺度短縮版を用い、発話のしにくさの自覚と性格特性の関連を調査した。
- ◆発話に関する質問では発話に関する調査では、JNS は CNS に比べて発話のしにくさを強く、自覚する傾向がみられた(問 3-1: JNS 52.96% vs CNS 18.94%, 問 3-2: JNS 34.98% vs CNS 23.79%)。一方、声の自己評価では CNS が圧倒的に肯定的で(問 3-3: CNS 73.34% vs JNS 17.73%)、両群の発話に対する意識に明確な差を示している。
- ◆発話のしにくさは情緒不安定性の高さや外向性の低さと強く関連し、開放性とは中程度に関連したが、誠実性や調和性との関連はほぼ見られなかった。性格特性は発話時の心理プロセスに異なる影響を与えることが示された。

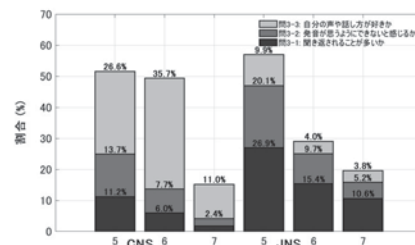


図1: 問3-1, 問3-2, 問3-3の結果

1: 全く違うと思う~7: 強くそう思う。

3-4-10

3-4-10 音響分析カードゲーム
「声のまようを読んでみよう入門編」の作成Creating an Acoustic Analysis Card Game:
Let's Read the Patterns of Voices Introductory Edition:○竹内京子(順天堂大), 青木直史(北大), 荒井隆行(上智大),
△大金さや香(国際医療福祉大), △鈴木恵子(おぎはら耳鼻咽喉科),
世木秀明(千葉工大), 秦若菜, △村上健(北里大), 安啓一(筑波技術大)

- ◆ことばのリハビリを行う職業である言語聴覚士養成校では、必修科目として音響学がある。しかしながら、音響学の授業は、学生の苦手な科目の筆頭であり、その改善が求められている。
- ◆音響学の授業では、臨床における音声の音響分析のため、サウンドスペクトログラムの読み方を学ぶ。今回は、サウンドスペクトログラムを学ぶ教材として、カードゲーム「声のまようを読んでみよう入門編」を作成した。作成経緯と、実際に作成したゲームの特徴を紹介する。
- ◆このゲームは、友達と競いながら基礎知識がなして遊ぶ、同時に音声のおもしろさを学び、さらには、言語聴覚士の仕事の紹介にもなるものを目指した。
- ◆本ゲームは、73 枚のカードを使い、モーラごとのサウンドスペクトログラムの区別ができるようになるものである。
- ◆作成上の工夫は、「あ」「い」「う」「え」「お」とそれ以外を区別する記号をつけ、F1, F2 の記号をつけたこと。バーストやボイスパーなどがサウンドスペクトログラム上で不鮮明な場合は、記号をつけて補ったことである。
- ◆これらのゲームは、全国の言語聴覚士養成校、都道府県言語聴覚士会、関連学会で配布中であり、ゲームのルールや遊び方も作成中である。

3-5-1

3-5-1 コンサートホールにおける
柱状拡散体の効果

The effect of columnar diffusers in a concert hall

○大山 宏, 山下 晃一, △高島 和博, △崎山 安洋(日本音響)
△斎藤 英夫(ヒビノ), △小宮山 雄太, 亀川 徹(東京藝大)

- ◆東京藝術大学音楽堂にて、コンサートホールにおける柱状拡散体(図中 CD と表記)の効果を演奏と録音、音場測定から検証した。
- ◆柱状拡散体のステージ上での設置場所を変更し実際に演奏する際、演奏者自身はその効果を顕著に感じていた。
- ◆録音されたデータを後日スタジオで演奏者自身が試聴したことにより、実際に演奏者が自身のパフォーマンスを客席にどのように届けることができるかを確認する機会ができ、演奏者へのインタビューは示唆に富む大変有意義なものであった。
- ◆柱状拡散体の設置による残響時間の変化はほとんど見られなかったが、筆者らが提案する残響減衰変動には変化が観測され、柱状拡散体を設置した時に一貫して値が小さくなっていることが分かった。

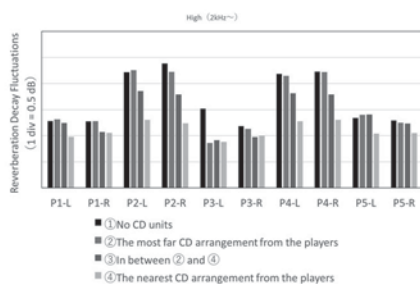


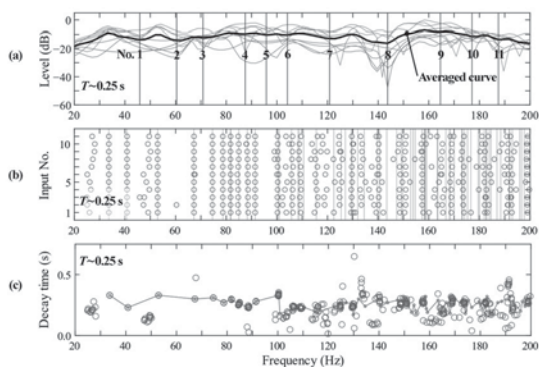
Fig.1:reverberation decay fluctuations (High).

3-5-3

3-5-3 音響モードの減衰パラメータ推定
に関する基本検討Fundamental study on estimating damping parameters of
acoustic modes

○山田祐生, 後藤耕輔(竹中技研)

- ◆小部屋での低音の残響時間の測定では、音響モードの影響により、減衰曲線に振動や屈曲が生じ、分析が困難となり得る。
- ◆このため、ISO 等に規定された標準的な帯域分析に依らず、個々のモードの減衰時間を把握するのが合理的とする考え方もある。
- ◆音響モードの減衰パラメータ推定に関する基本検討として、正弦波励振時の減衰信号より、固有周波数とモード減衰時間を同定するアプローチの適用性について調べた。

Fig.1 (Excerpts): Results of modal damping identification.
The blue and red lines represent the excitation frequencies and the specified parameter values, respectively.

3-5-2

残響時間予測における室形状補正係数の
一考察A consideration of room shape correction factors for reverberation time
prediction

○西原法子(竹中技研)

吸音材を部屋に設置したときの残響時間の予測精度は、(1)材料の吸音率と(2)室形状(拡散性)に依存する。(1)は残響室で求められる吸音率が不確かなこと、(2)は従来の残響式では室容積(V)と室表面積(S)のみしか反映されないことによる。

本報告は、小規模空間の平均自由行路に着目し、音線法による自由行路長(FPL)と $4V/S$ の比によって、Sabine 残響時間を補正する一考察を示す。

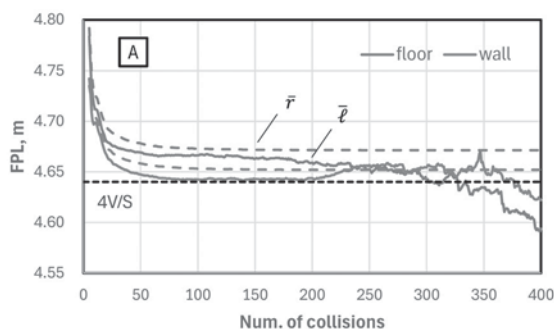


Fig.2より抜粋:Plot of free path length (FPL) vs. number of collisions.

3-5-4

3-5-4 室内音場におけるモードの縮退による
周波数変動係数の変化Changes of the frequency-domain coefficient of variation due to mode
degeneracy in a room sound field

○鈴木諒一, 星和磨, 羽入敏樹(日大・短大)

- ◆本研究の目的は、音楽スタジオのような小規模空間において問題となるモードの影響を、実音場で定量的に評価できる指標を確立することである。
- ◆筆者らはインパルス応答から算出できる指標指標である周波数変動係数に着目している。
- ◆本報では、モードの縮退の発生頻度の異なる直方体室のインパルス応答をモード合成法により求めた。求めたインパルス応答から周波数変動係数を算出し、周波数変動係数における音場のモード密度やモードの縮退の影響について考察した。
- ◆その結果、周波数変動係数が室内音場のモードの影響を定量的に評価する指標として有効である可能性が示唆された。

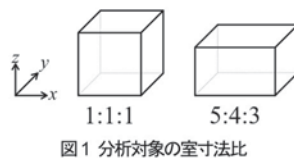
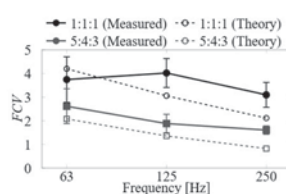


図1 分析対象の室寸法比

dimension ratio	L_x [m]	L_y [m]	L_z [m]
1:1:1	4.000	4.000	4.000
5:4:3	5.109	4.087	3.065

表1 直方体室の各辺の長さ

図2 各室寸法比の直方体室における
インパルス応答から算出した周波数変動係数(Measured)とその理論値
(Theory)

3-5-9

3-5-9 建築改修における工事騒音低減のための
工具工法の開発Development of Tools and Methods for Reducing Construction Noise
in Building Renovation Projects

○坂巻一弥, 高木一昌, 前田貴俊(RIA), 坂本慎一, 矢野博夫(東大生研)

- ◆一般社団法人リニューアブルイノベーション協会(以下 RIA)は、建築改修工事における騒音問題を社会課題として位置づけ、その解決を通じて持続可能な都市環境と建設産業の発展に寄与することを目的として設立された。
- ◆本論文では、RIAが認定する消音仮設壁(Fig. 1)および低振動型の機材を対象とし、実際のテナントビルの改修工事現場における音響特性の測定を通じて、各工法の騒音低減効果を定量的に評価した。
- ◆工事エリアの出入りに消音仮設壁を設置した場合、90 dB(A)強の工事音に対し、入居者エリアにて、暗騒音と同等まで音圧レベルがさがることを確認した。聴感評価では、作業音は聞こえるが、騒音には感じないことを確認した。しかし、打撃を伴う工具では固体伝搬音が発生し、消音仮設壁の効果は減少した。
- ◆これらの結果から、建築改修における騒音対策には、消音仮設壁の設置に加え、発生源となる工具の選定も重要であることが示唆された。消音標準化システムの社会実装に向け、今後も多様な現場条件での検証と、聴感評価を含む総合的な研究を継続する予定である。

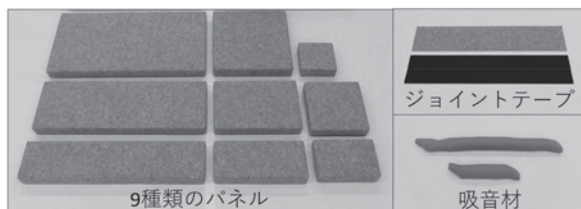


Fig.1 Soundproof temporary wall components

3-5-11

3-5-11 木造床に設置した乾式二重床における
標準衝撃源の床衝撃音レベル差
に関する実験的検討Experimental study on floor impact sound level difference of standard impact
sources in a dry-type floating floor on a wooden floor.

○富田隆太(日大・理工)

- ◆本稿では、標準衝撃源及び標準重量衝撃源の落下高さを変化させ、木造床及びその上に乾式二重床を施工した場合の床衝撃音レベル差について検討を行った。
- ◆その結果、タイヤ加振の場合には落下高さが高くなる、すなわち加振力が大きくなると床衝撃音レベル差が小さくなる傾向であった。一方、ゴムボール加振の場合には、落下高さによる床衝撃音レベル差の違いは見られなかった。
- ◆また、木造床における乾式二重床の床衝撃音レベル差の推定方法について検討を行い、今後の課題を明らかにした。

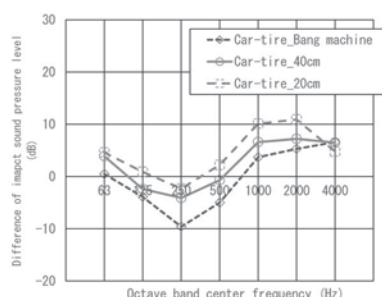


Fig.1: Difference of impact sound pressure by the car-tire

3-5-10

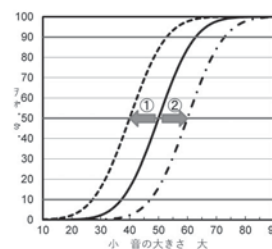
3-5-10 建築物の騒音問題は何故減らないのか？
一音の大きさと

音環境の不満足度に関する一考察—

Why do acoustic problems in buildings remain unresolved? — An analysis of
noise levels and dissatisfaction with the acoustic environment

○平光厚雄(建研), 佐藤洋(産総研), 平川侑(国総研)

- ◆建築物の音環境性能は向上してきていることは考えられるものの、建築物(特に、集合住宅)では、音環境性能に起因する問題、トラブルが発生しており、音環境に関する不満率は下がっていない。
- ◆住宅の音環境の満足度などに関するアンケート調査のレビューを行い、音環境の満足度(不満度)の年代による推移などを把握した。
- ◆音の大きさと不満率の関係は、Fig. 1ようになる。音の大きさが小さくなると、不満率が下がるが、アンケート調査結果では大きく下がっていない。これは、年々居住者の要求性能が上がっている、騒音に対する社会的受容度が低くなっていることなどが考えられる。
- ◆図中の実線から①ではなく、②のようにシフトすれば、不満率を下げるができる。
- ◆音環境の問題、トラブルが発生するメカニズムが把握できれば、よりよい住環境となるとともに、基準値の設定する際にも有用となるであろう。

Fig.1: Relationship between noise
levels and dissatisfaction with the
acoustic environment.

3-5-12

3-5-12 製造者が異なるゴムボール衝撃源の衝撃
力特性に関する実験的検討*An Experimental Study on the Impact Force Characteristics of Rubber Ball
Impact Sources from Different Manufacturers

○平川侑(国総研), 平光厚雄(建研)

- ◆昨年ごろから、国外の2製造者によりゴムボール衝撃源が販売されはじめた。本論ではこれらの製造者が異なるゴムボール衝撃源の衝撃力特性や重量床衝撃音遮断性能を比較した結果を報告する。
- ◆どのゴムボール衝撃源も JIS、ISO 規格に記載されている衝撃力特性には合致していた。しかし、同じ実験棟の同じ加振・受音位置でも測定結果が全く同じにならなかった。

		A社	B社	C社
CLT	床衝撃音レベル等級	60(60)	60(61)	60(60)
	最大A特性床衝撃音レベル	61	62	61
	逆A特性重みつき床衝撃音レベル	57	58	57
2×4	床衝撃音レベル等級	70(70)	70(71)	70(70)
	最大A特性床衝撃音レベル	67	68	67
	逆A特性重みつき床衝撃音レベル	65	66	65
RC	床衝撃音レベル等級	60(58)	60(58)	55(57)
	最大A特性床衝撃音レベル	57	56	57
	逆A特性重みつき床衝撃音レベル	53	53	53

Fig 1 Evaluation results of a single number rating for each impact
source

3-5-13

3-5-13 歩行等を対象とした加振力計測に基づく床衝撃音の可聴システム

An Audible System for Floor Impact Sound Based on Excitation Force Measurements from Walking and Related Activities

○富永大祐, 會田祐, △室 裕希(長谷工技研)

- ◆集合住宅において、音をめぐる不満やトラブルは依然として存在しており、住まい方における配慮意識に対する啓蒙が図られている。
- ◆生活音の伝わり方の体験的な理解が、配慮意識の向上に有効であると考え、歩行等を対象とした床衝撃音の可聴システムを開発した。
- ◆会議室等の一般的な室内空間に容易に設置可能であり、体験者が動作を行った際に、同一空間内において下階に生じる床衝撃音を聴取可能とするシステムとした。
- ◆加振力波形と建物の応答波形を畳み込み(畳み込み音)、スピーカーで再生することで床衝撃音を再現する。
- ◆実験建物での実測音と畳み込み音を比較し、システムの再現性を検証した。

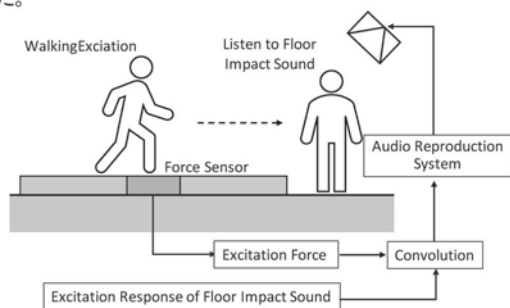


Fig. 1: Overview of the Auditory System

3-5-15

3-5-15 CLT 建築物におけるパネル接合部と標準衝撃源の位置関係が床衝撃音レベルに及ぼす影響 —その2 パネル接合部付近の衝撃位置に関する検討—

Effect of the relative position between panel joints and a standard impact source on floor impact sound pressure levels in CLT buildings
—Part 2: Investigation of impact positions near panel joints

◎兵藤伸也, 林敬祐, 小林真人(飛島建設),

小林彩(神奈川大・院), 劉金雨, 安田洋介(神奈川大・建築)

- ◆CLT の床スラブにおける接合部付近に測定点及び加振点を線状に設定し、床スラブのインピーダンスと下階での重量床衝撃音を測定した (Fig. 1)。測定結果よりパネル接合部付近におけるバングマシンとゴムボールの衝撃位置が床衝撃音へ及ぼす影響について検討した。
- ◆ゴムボール加振では、加振点が梁上であれば梁の直交方向での衝撃位置の違いは室内平均最大音圧レベル $L_{k,Fmax}$ にほとんど影響しないことが分かった (Fig. 2)。
- ◆バングマシン加振では、加振点の反対側に衝撃源を設置した場合はゴムボール加振と類似の傾向となるが、加振点側に配置した場合は衝撃位置によるばらつきが大きいことが確認された。

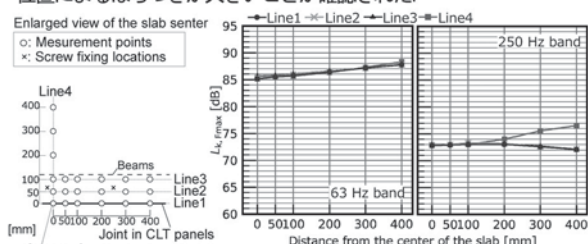


Fig. 1: Measurement point locations on floor slab.

Fig. 2: Relationship between impact position and average maximum sound pressure level $L_{k,Fmax}$ by rubber ball excitation.

3-5-14

3-5-14 CLT 建築物におけるパネル接合部と標準衝撃源の位置関係が床衝撃音レベルに及ぼす影響 —その1 衝撃源の設置方向に関する検討—

Effect of the relative position between panel joints and a standard impact source on floor impact sound pressure levels in CLT buildings
—Part 1: Investigation of the installation direction of the impact source

☆小林彩(神奈川大・院), 劉金雨, 安田洋介(神奈川大・建築),

兵藤伸也, 林敬祐, 小林真人(飛島建設)

- ◆CLT 建築物を対象に、そのパネル接合部付近を加振する際の衝撃源の位置が床衝撃音レベルに及ぼす影響について実大試験体を用いた実測を行った。その1として、本報では**衝撃源の設置方向の影響について検討する**。
- ◆パネル接合部かつ小梁上の加振点のみに着目した場合、バングマシンでは設置方向による影響がみられた (Fig. 1)。
- ◆上記加振点をバングマシンで加振した場合、CLT スラブではRC スラブに比べて 63 Hz 帯域での $L_{k,Fmax}$ の標準偏差が大きいことが示された。また、値のばらつきの要因はバングマシンの荷重以外にあると考えられる。

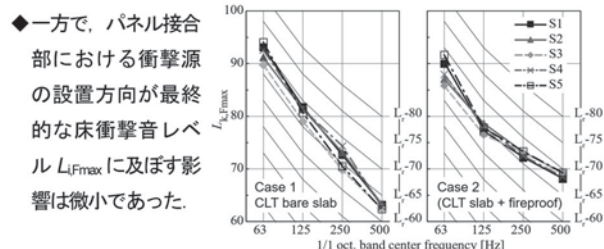


Fig. 1: $L_{k,Fmax}$ under bang machine excitation at the point P1-3 with different installation orientations.

3-6-1

3-6-1 Estimation of aircraft noise level at immission side using sound power level determined by field measurement

☆Xinyi ZHANG (Grad. Schl., Eng., The Univ. of Tokyo), Makoto MORINAGA (Daido Univ.), Shinichi SAKAMOTO (IIS, The Univ. of Tokyo)

- ◆This study estimated aircraft noise at the immission side using sound power levels of various aircraft types determined by previous on-site measurements.
- ◆In the estimation, the aircraft was modeled as an omnidirectional point source, and effects of geometrical spreading and atmospheric absorption were considered; flight paths were obtained from ADS-B data.
- ◆Based on these inputs, the A-weighted sound pressure levels at the observation site were calculated and compared with measurement.

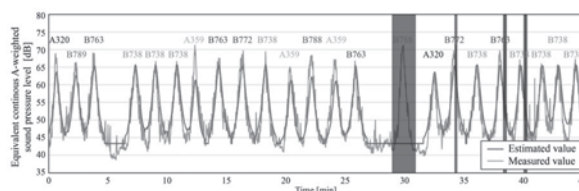


Fig. 1: Comparison between estimated and measured $L_{Aeq,1s}$ during the observation period.

3-6-2

3-6-2 九州新幹線全線における
新幹線鉄道騒音の面的推計Noise map estimation of Shinkansen Super Express Railway
along the Kyusyu Shinkansen whole line

☆安武朋希(熊本大院), 丸山直也(熊本大院),

原田和典(岡山県立大), 平栗靖浩(近畿大), 川井敬二(熊本大院)

- ◆欧州では、ノイズマップの作成によって新幹線騒音が面的に評価されている一方、日本では、代表地点における最大騒音レベルによる評価がされており、騒音の面的な広がり把握することができていない。
- ◆九州新幹線全線を対象として、GPS ログから車両速度を推計して音源データベースを構築し、それに基づき騒音レベルを算出することで、沿線地域における騒音曝露状況を面的に推計することを目的とする。また、推計結果を実測値と比較検討することにより、その精度を検証する。
- ◆推計の精度として、自治体の実測値と比較すると、113/138 点(82%)の予測点で ± 5 dB 以内の精度で予測ができており、二乗平均誤差(RMSE)は約4.3 dBであった。



Fig. 1: Example of Noise Map

3-6-4

3-6-4 日欧における鉄道車両の騒音基準について

On the noise regulation of railway vehicle in Japan and in Europe

○高野靖(京都大学)

- ◆交通騒音に関する基準は音源基準と環境基準に分けられる。自動車や航空機は、国際的に共通の音源基準により認証され、世界中で使用が可能である。一方、鉄道車両は国により異なる基準が設けられ、車両メーカーも個別に設計を行っている。
- ◆日本ではの鉄道騒音に関し音源基準は設けられず、軌道中心より 25 m ないし 12.5 m 離れた標準測定点での、市街地を対象とした環境基準を流用している。市街地でなくとも、環境保全のニーズは高いため、日本でも鉄道車両に対する音源基準を別途設けることが有効と考える。
- ◆欧州では、鉄道システムの共通仕様書(TSI)により音源基準を定め、その測定法を EN 規格にて定めている。EN 規格と同じ規格が ISO としても発行されている。ISO に準拠した試験条件での測定は、試験設備の整備などに多大なコストを伴うため、欧州以外での実施は困難な場合が多く、欧州への鉄道車両輸出の障壁にもなっている。
- ◆今後、日本でも実施可能な測定の結果から、ISO に対応する値を推計する手法を確立し、ISO の付属書等へ記載を目指すことが課題である。

3-6-3

3-6-3 実物大音響試験による上部に傾斜形状を有する鉄道向け防音壁の騒音低減効果の評価

Full-scale acoustic evaluation of inclined-top railway noise barriers

○佐藤大悟(鉄道総研), 阿部幸夫, 山本健次郎, 中村祐太(日鉄建材)

- ◆新幹線高架橋では、老朽化した既設防音壁の取替えが予定され、取替え前後での沿線騒音を同等以下とすることが求められている。防音壁構造の違いにより騒音低減効果は異なるため、それらの関係性を評価しておくことは重要である。
- ◆半無響室内にて実物大の防音壁模型を用いた音響試験を実施し、防音壁の断面形状等の違いが沿線騒音への寄与が大きい車両下部騒音の低減効果へ与える影響について評価を行った (Fig. 1)。試験対象の防音壁形状は、営業線における既設防音壁、および本研究で新たに実用化を目指す防音壁より選定した。
- ◆試験の結果、既設防音壁よりも高い騒音低減効果を有する防音壁形状の傾向が明らかとなった (Fig. 2)。今後、具体的な構造の開発に向けた取り組みを進める予定である。

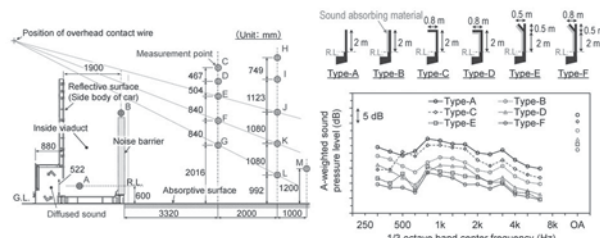


Fig. 1: Positional relationship on the central cross-section in the lateral direction Fig. 2: Measurement results at point M

3-6-5

3-6-5 Determining parameter values for prediction model of insertion loss of buildings for higher source positions

☆Qiyuan Wang (The Univ. of Tokyo), Ken Anai (Fukuoka Univ.),
Hiroo Yano, Shinichi Sakamoto (IIS, The Univ. of Tokyo)

- ◆The ASJ RTN-Model provides an effective prediction model for road traffic noise (RTN) propagation behind buildings, characterized by the insertion loss of buildings ($IL / \Delta L_B$). However, the current model is limited to a fixed height of the noise source at ground level ($h_s = 0$ m).
- ◆In this work, we report the attempt of adjusting prediction models to fit for higher source positions based on scale model experiments.
- ◆The modified model largely maintains the original structure while we adjust the determination method as well as values of the parameters within the prediction equation, including both variants and constants.
- ◆We examine the performances of the adjusted prediction model under various cases.

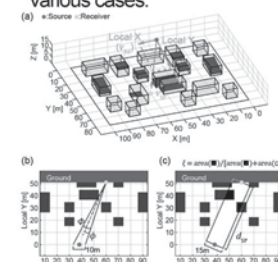
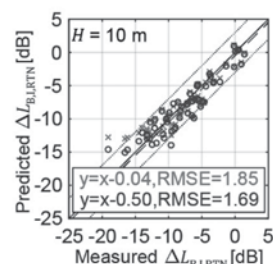


Fig. 1 Schematic illustrations of the adjusted prediction model.

Fig. 2 Comparison of measured IL and predicted IL at $h_s = 10$ m, $H = 10$ m (H is the building height).

3-6-6

3-6-6 遮熱性舗装上の自動車走行騒音の実測調査

Actual field measurement of road vehicle noise on heat-shielding pavements

☆齊藤稜河(東大院), 坂本慎一(東大生研)

- ◆ヒートアイランド現象の対策として導入が進む遮熱性舗装は、表面に遮熱材が塗布されており、従来舗装と比べて騒音特性が異なる可能性がある。
- ◆本研究では、東京都内の一般道路に敷設された遮熱性舗装8地点を対象に、定常走行中の単独車両に対する騒音実測調査を実施した。
- ◆音響パワーレベル L_{WA} と周波数特性をASJ RTN-Model 2023に基づき算出し、従来舗装と比較した。
- ◆その結果、遮熱性舗装上を走行する車両の L_{WA} が従来舗装より小さいこと、周波数特性が従来舗装と異なることが確認された。
- ◆遮熱材が車両通過や経年により剥離していくことで、騒音特性が変化する傾向が確認された。

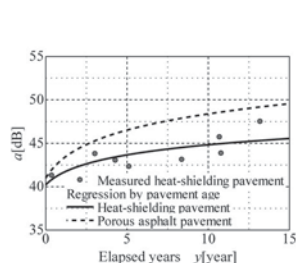
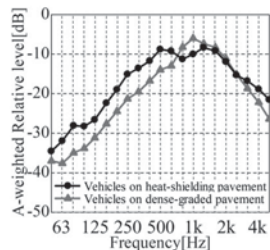
Fig. 1: Relationship between elapsed years and constant a 

Fig. 2: Frequency characteristics of heat-shielding and dense-graded pavements

3-6-8

3-6-8 ロードノイズ起因の窓ガラス振動寄与分析

Influence of Window Glass Vibration on Road Noise

○高宮壮輔, 吉田準史(阪工大・工), △宮石 晨

- ◆本研究では、ロードノイズを対象に窓ガラス振動の影響を実稼働TPA法を用いて把握した上で、ガラス部の振動特性変更によるロードノイズ低減効果について明らかにすることを目的とした。
- ◆実験ではロードノイズを想定した足回り加振による実稼働実験を行った。得られたデータに実稼働TPA法を適用したところ、244 Hzの車内音ピーク部において最大ではないものの、右ガラス中心部の寄与が他の周波数に比べて高いことが明らかになった。
- ◆そこで、振動特性の異なる窓ガラスを自動車模型に取り付けて同様の実験を行ない、着目周波数帯における右窓ガラス中心部の振動の低下が確認された。それに伴い、車内音の低減も確認された。

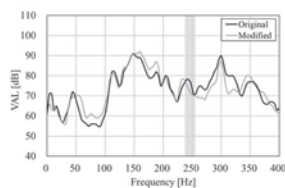


Fig. 1: Comparison of the VAL at the center of right-glass between original glass and damping glass.

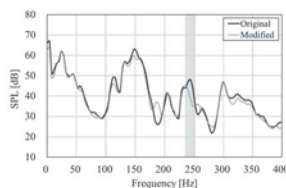


Fig. 2: Comparison of the interior noise between original glass and damping glass.

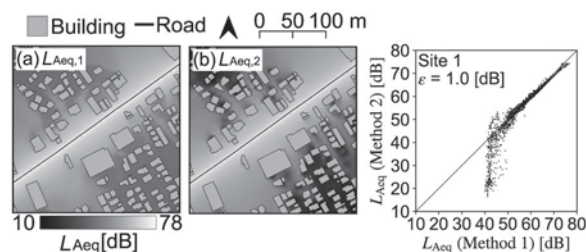
3-6-7

3-6-7 ASJ RTN-Modelに基づく騒音予測手法の特性に関する検討—実用計算法及び2世代の回折補正手法による市街地の騒音予測結果の比較—

Evaluation of noise prediction methods based on the ASJ RTN-Model: Comparison between the practical method and the diffraction correction methods of two generations in urban sites

☆浅野蘭丸(新潟大院・自然研), 大嶋拓也(新潟大・工)

- ◆道路交通騒音予測モデルASJ RTN-Modelに基づく3つの騒音予測手法から、新潟県内2箇所の道路交通騒音を予測し、等価騒音レベル L_{Aeq} [dB]を比較した。Method 1の L_{Aeq} の減衰は、Method 2に比べて小さい(Fig. 1, 2)。音源から離れた建物が入り組む部分では、Method 1の建物群による減衰に関する補正量 ΔL_{BB} が低下し、下限となることで、Method 1の L_{Aeq} が低下しなくなるためと考える。
- ◆一部の予測点で、Method 3の L_{Aeq} は、Method 2に比べて高い。Method 3の回折補正量 $\Delta L_{dif,tb}$ は障壁間隔 5m 以上で、Method 2の回折補正量 $\Delta L_{dif,dd}$ より5 dB 高くなるためと考える。

Fig. 1: Distribution of L_{Aeq} in Site 1.Fig. 2: Scatter plot of L_{Aeq} .

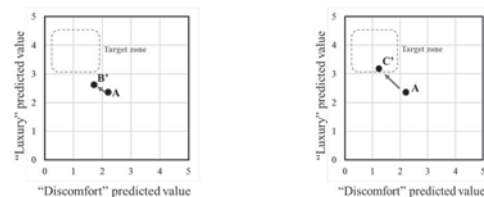
3-6-9

3-6-9 自動車用電動シート稼働音の音質改善手法について

Sound Quality Improvement Procedure for Power Seat Operating Noise.

○中村理人(阪工大), 吉田準史, △山内健斗

- ◆本研究では、自動車用電動シートの稼働音を対象に、先行研究で提案された音質評価手法を用いて音質向上指針の検討を行なうことを目的とした。
- ◆分析の結果、対象とした電動シートの音質改善には6 Bark (570 ~ 700 Hz) 帯の音圧レベル低減と4 Bark (400 Hz) 以下の音圧レベル増大が必要であり、音質改善にはフレーム下部の重点的な対策が必要であるとわかった。
- ◆そこで、シートフレームに質量付加による対策を施したところ、音質改善傾向を確認できたが、目標音質に設定した領域には達しなかったためモータ変更とフレーム対策を行なった際の音質改善結果を算出したところ、目標音質を達成できる見込みが得られた。



(a) Frame countermeasure.

(b) Frame and motor countermeasure.

Fig. 1: Sound quality improvement result.

(A : Original, B' : Frame countermeasure, C' : Frame and motor countermeasure)

3-6-10

3-6-10 野外コンサート会場における
音響伝達特性の分析

Investigation of Acoustic Transfer Characteristics at Open-Air Concert Area

○角井明日佳, 吉田準史(阪工大・工), △松井怜央

- ◆野外コンサートの騒音問題対策をするにあたり、アーティストや観客、近隣住民の双方が満足できる会場の上限音量を検討することや、屋外でありながらも会場内外の木々や丘陵部の調整による楽曲の伝達を抑えるための検討が必要となる。本研究では、その初期段階の検討、調査として会場内上限音量の検討に必要な会場内から居住区までの距離減衰を正確に、そして安定して把握することを目的とした。
- ◆環境騒音などの影響で厳密な音響伝達特性の把握が困難になる場合を考慮して、様々な箇所で開催会場からの伝達音を正確に計測・分析することを試みた。
- ◆計測地点のデータから距離減衰を求め、コンサート会場周辺の音響伝達特性を把握した。その結果、万博記念公園近隣の居住区では減衰量が概ね40 dB以上(例: FOHで95 dBの場合には居住区で55 dB以下となる)となっている可能性が高いことがわかった。

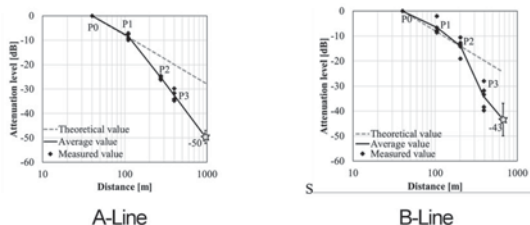


Fig.1 : Estimated distance attenuation.

3-6-12

3-6-12 出口方向を示す音響誘導設備が
避難行動に与える影響

「先導・迷走・追従」行動を考慮したマルチエージェントシミュレーション

Effect of acoustic guidance systems indicating exit directions on evacuation behavior: a multi-agent simulation considering leading, wandering, and following Actions.

☆平井健太郎(近畿大院), 菅原彬子, 平栗靖浩(近畿大)

- ◆災害時に出口を示す放送を行う音響誘導設備が、避難効率にどのような影響を与えるのか、マルチエージェントシミュレーション(MAS)を用いて検討を行った。
- ◆MASの妥当性向上を目的に、災害時における先導・迷走・追従行動をモデル化し、それらの避難行動を再現する設定を行った。
- ◆Fig.1に示す各ケースで150回シミュレーションを行った結果、音響誘導設備の有無で平均避難完了時間に差が生じた(Fig.2)。

また、設置数を増やすにつれて避難効率が向上した。

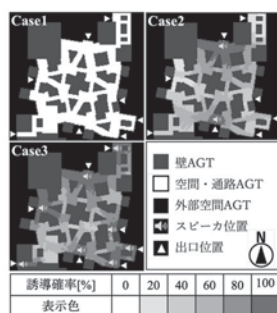


Fig.1 Simulation cases

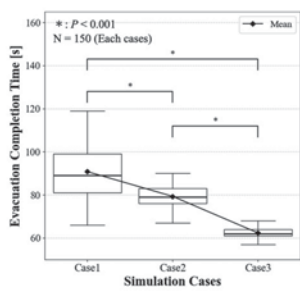


Fig.2 Variation in evacuation completion times

3-6-11

3-6-11 振幅スペクトル形状に着目した妨害音下における
音声了解度の非侵襲的推定法に関する検討

A study on non-intrusive estimation of speech intelligibility under disturbing sounds focusing on amplitude spectrum shape

○佐藤遼人, △西本拓矢(神戸大院・工学研)

- ◆時間包絡線に基づく非侵襲的音声了解度推定法の限界を補うため、声帯振動由来の振幅スペクトル特性に着目した指標(HS-SWIPE, HS-CEP)を提案した。
- ◆通常の音声と雑音駆動音声の品質の違いと、騒音と残響音の影響に着目して上記の指標の有効性を検証した。
- ◆HS-SWIPEとHS-CEPは雑音駆動音声のような周波数領域の劣化を評価可能であった。また、騒音の影響についても適切に評価できた。
- ◆一方、残響音の影響はある程度の評価が可能であったが、評価対象の音源に含まれる調波構造が直接音と残響音のどちらに由来するかを分離できないため、特に残響音が強い条件で音声了解度を過大評価する可能性があることを示した。

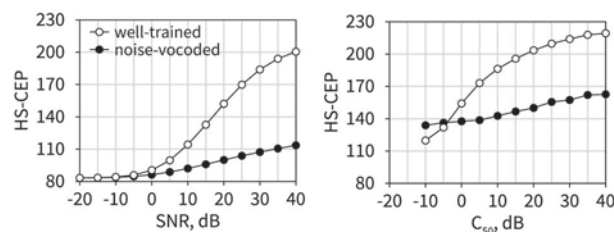


Fig.1: HS-CEP as a function of relative intensity of disturbing sound. Left and right panels show noise-added conditions and reverberation-added conditions, respectively.

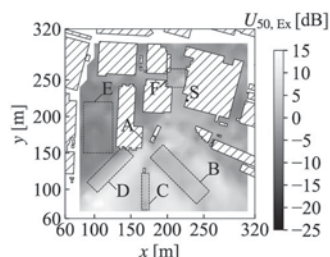
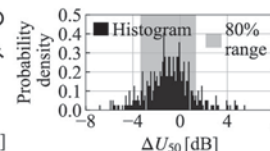
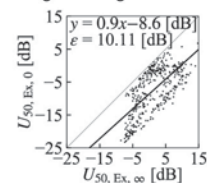
3-6-13

3-6-13 実地形及び建物群における屋外拡声システム
の聴取性の数値予測 一暗騒音の影響を考慮した音響縮尺模型実験との比較一

Numerical prediction of speech intelligibility of outdoor loudspeaker systems in real-life topography and buildings: comparison with an acoustic scale model experiment with background noise.

☆羽賀柊人(新潟大院・自然研), 大嶋拓也(新潟大・工), 石塚崇(清水建設技研)

- ◆実地形及び建物群における音響縮尺模型実験と波動音響数値解析による解析結果から算出した聴取性指標 $U_{50,i}$ [dB]を比較した (Fig. 1)。両結果の $U_{50,i}$ の分布の特徴は、良好な一致を示した。両者の差は、概ね ± 2 dBの範囲に80%の受音点が収まる (Fig. 2)。
- ◆暗騒音による $U_{50,i}$ の変化を検討した (Fig. 3)。暗騒音を導入すると、 $U_{50,i}$ が低下する受音点と低下しない受音点が存在し、前者は、暗騒音が大きいくほど低下幅が大きい。この差は、音源から到達する音の総エネルギーの差による。

Fig. 1 Distribution of $U_{50,Ex}$ Fig. 2 Histogram of ΔU_{50} Fig. 3 Comparison of U_{50} between Cases $N-\infty$, $N-0$

3-6-14

3-6-14 室内吸音設計規準に関する音声伝送の聴感実験
—模擬音場における聴き取りにくさ評価—

Auditory experiment on speech transmission for indoor sound absorption design standards - Evaluation of listening difficulty in simulated sound fields

☆陳科吉, 佐久間哲哉(東大・工), 佐藤逸人(神戸大・工)

- ◆ AIJES として「建築空間の吸音設計規準・指針」の室区分Ⅱの残響時間推奨値に関して、幾何音響解析により模擬音場を用いて聴き取りにくさの聴感実験を行い、推奨値の妥当性を検証した。
- ◆ 幾何音響解析により大空間の配置では床以外の初期反射音は到達しないため、明瞭度の低い条件になっていると推察された。
- ◆ 残響時間推奨値と「聴き取りにくくはない」以外の回答割合を表す LDR の対応について、中規模空間までは音声伝送クラスとの対応が確認されたが、大空間では1~2 ランク低下する傾向が見られた。
- ◆ 音声明瞭度 C_{50} と LDR の間には比較的確かな対応が確認され、室規模により理論値と解析値に系統的な差が生じることも確認された。

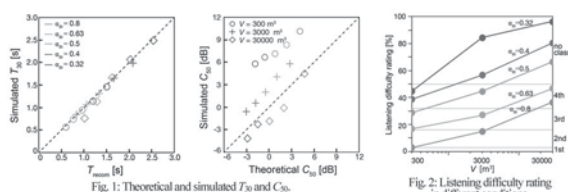


Fig. 1: Theoretical and simulated T_{30} and C_{50} .

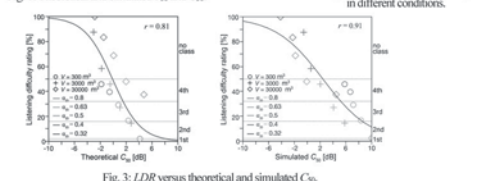


Fig. 3: LDR versus theoretical and simulated C_{50} .

3-8-2

3-8-2 クラリネット用リードの演奏適性と
演奏音における高調波成分の関係

The Relationship Between Playability of Clarinet Reeds and Harmonic Components in Performance Sound

☆鈴木真琴(国立音大・院), 三浦雅展(国立音大), 松谷晃宏(東京科学大)

- ◆ クラリネットの演奏には葦材のリードを用いる。リードは10枚一組で市販されているものの、演奏適性すなわち演奏のしやすさにばらつきがあり、その音響的な調査はこれまで見られない。関連研究として、卓越した演奏家が学生の楽器を演奏したのちに再度学生が演奏すると音質が向上し、高調波成分の3f0と5f0が増大する[1]。
- ◆ リードの演奏適性が演奏音に与える影響を明らかにするために、高調波成分を調査する。試奏の評価で演奏適性が高いリード(GR1, 2)と低いリード(BR1, 2)を用いて複数の音高で演奏音を録音し、奏者が演奏音からリードの演奏適性の聞き分け可能性をシェフエの一対比較法により検証した。その結果、クラリネット奏者は演奏適性が高いリードと低いリードの演奏音を概ね聞き分けすることができた。
- ◆ 演奏音の高調波成分におけるエネルギーバランスについて、シャリュモー音域ではf0と3f0が同程度のエネルギーであることが、奏者が良いと感じる要因の一つである可能性が示唆された。さらに、4f0のエネルギーが音の評価に影響している可能性が新たに示唆された。

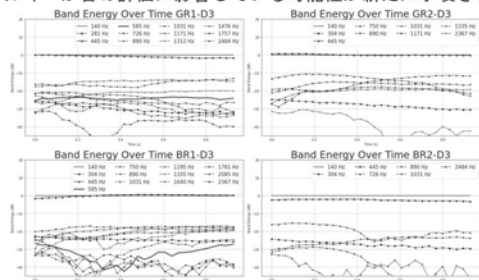


図1 D3における高調波成分の時間変化

[1]三浦雅展, "卓越した演奏家の吹鳴による吹奏楽器の状態変化", 日本音響学会音楽音響研究会資料 Vol.42, No.2, pp.17-22, (2023)

3-8-1

3-8-1 演奏支援に向けたバイオリン空間放射
特性の可視化法

Visualization of spatial radiation characteristics of violin for supporting musical performance

○牧 勝弘(愛知淑徳大), 饗庭 絵里子(電気通信大)

- ◆ バイオリンの空間放射特性には、奏者のポウイング技能が表れることが明らかにされている。
- ◆ バイオリン未熟練者のポウイング技能向上を支援するため、視認しやすい形で空間放射特性を可視化する方法について検討した。
- ◆ 聴覚と視覚の時間分解能を考慮してバイオリン音のデータ処理を行い、放射方向および放射指向性の強さを強調して音圧レベル分布を可視化した。
- ◆ その結果、熟練奏者に特有の空間放射特性を演奏動画として効果的に可視化することができた。
- ◆ 熟練奏者の空間放射特性を模範として、未熟練者が自身の演奏状態を視覚的に確認しながら自己評価することで、右手のポウイング技能の向上が期待できる。

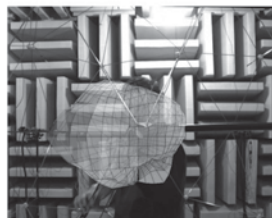


Fig.1 Visualization of the spatial radiation characteristics of a violin performance. A single frame from the performance video is shown. The analysis frequency is 902 Hz. At this frequency, strong radiation directivity is observed in the frontal direction.

3-8-3

3-8-3 シングル・リード楽器熟達奏者の
吹奏時の口唇圧に関する考察

A consideration of lip pressure in single-reed instrument playing by highly-skilled players

○柳下柚子(科学大・聖徳大)

- ◆ 本研究では、サクソフォンとクラリネット吹奏時の口唇圧をピエゾ抵抗圧力センサを用いて測定し、熟達奏者に共通する特徴を分析した。
- ◆ 両楽器の熟達奏者は、Fig. 1のようにロングトーンのコア部で一定の圧力を維持していた。下顎からの圧力が適切にコントロールされ、均一な力で演奏していることを示している。
- ◆ アタックでは3種類、リリースでは2種類の圧力の変化が見受けられた。
- ◆ 圧力の強さはサクソフォンに比べてクラリネットが40%程度と弱い傾向にあったが、両楽器共に個人差が大きかった。
- ◆ 音域や音量による圧力の変化は、サクソフォンでは低音域での圧力低下、クラリネットでは高音域での圧力上昇が見受けられた。これ以外に過半数に共通する特徴は見受けられず、音域や音量によって奏法がどのように変化するかは個人差が大きかった。

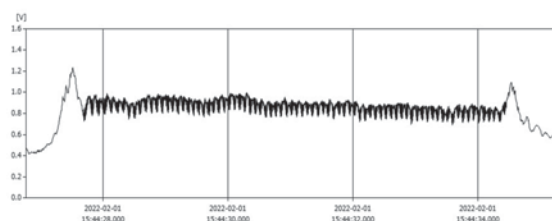


Fig.1: Lip pressure of a saxophonist

3-8-4

3-8-4 楽器演奏の計測とヤマハでの事例紹介

Measurement of Musical Instrument Performances and Case Studies at Yamaha

○小幡哲史(ヤマハ株式会社)

- ◆演奏動作や楽器の物理的反応を定量的に計測・分析することで、演奏メカニズムの理解を深めようとする試みが進められている。
- ◆一方で、演奏の計測結果を演奏者の技術向上や教育支援などに応用するためには、いくつかの課題が存在する。
- ◆演奏計測は大きく「演奏音」「楽器の挙動」「身体動作」の3つに分類されるが、計測結果の理解や解釈には専門的な知識や経験が求められる場合がある。
- ◆スポーツの分野では、個人が運動記録を振り返り、コーチがデータをもとに指導に活用する取り組みが広がりがつつある。
- ◆楽器演奏においても、奏者が自身の現状を把握し、指導者がより芸術的・表現的な指導に専念できるような計測の環境を整えることが望ましいと考えられる。

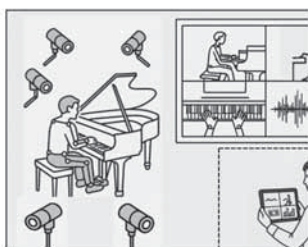


Fig.1: Measurement of Musical Performance and Examples of Data Application

3-8-6

3-8-6 モーションキャプチャを用いたピアノ演奏スキルと演奏動作について*

Motion Analysis of Grand Piano Performance Using Motion Capture

○三戸勇気(日大・芸術), 塚田湧太, 柴田直樹, 山口達也(日大・理工), 山口優芳, 駒澤大介, 川上央(日大・芸術), 篠田之孝(日大・理工)

筆者らは光学式モーションキャプチャシステムを用いて感情表現による鍵盤楽器の演奏動作の動作解析の研究を実施している。従来の研究では上半身のみで反射体であるマーカーを取り付けていたが、下半身の動きも計測するため、全身にマーカーを取り付けている。

本文は演奏スキルの影響を検討するため、音楽専攻の大学生、プロピアニストの各2名の演奏動作を計測し、身体各部の関節部位などの軌跡長を算出し、感情表現によるグランドピアノの演奏動作解析の検討を行った。

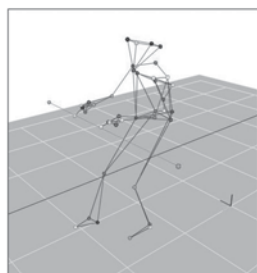


Fig.1 Marker adhesion position

3-8-5

3-8-5 グランドピアノ演奏音の受聴位置による聞こえ方の違い

Perceptual difference in grand piano sound from various listening positions

☆佐藤匡介(東京藝大院), 丸井淳史, 亀川徹(東京藝大)

- ◆グランドピアノの正面側と鍵盤側での演奏音の聞こえ方について、Fig.1に示す位置で録音した演奏音を用いて非類似度評価と評価語による評価をおこない、違いを調べた。
- ◆受聴位置の違いは主にピアノからの距離に基づいて判断されており、評価語評価ではピアノ近傍で明瞭さや迫力、遠方でやさしさや残響感といった属性が対応したが、ピアノ近傍の受聴位置で比較すると、正面側よりも鍵盤側の方が明瞭で近い音であると評価されていた。
- ◆主観評価の結果と音響特徴量の相関を調べたところ、打鍵ノイズを含む低周波数帯域のレベルが高い場合にピアノからの距離が近いと判断される傾向がみられた。

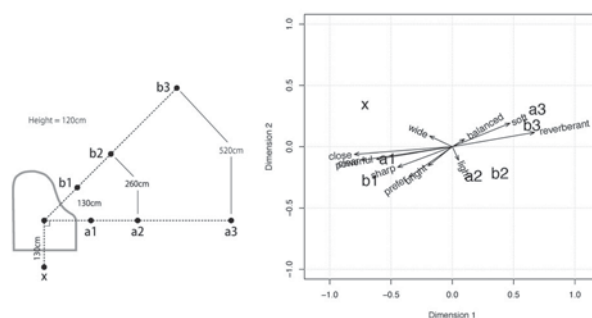


Fig.1: Microphone positions relative to the piano.

Fig.2: INDSCAL common space with projections of auditory attributes.

3-8-7

3-8-7 ピアノ演奏における音楽表現と身体動作に関する研究
—legatoの表現を対象とした事例分析と検証実験を通して—A Study on Musical Expression and Body Movement in Piano Performance: Through case analysis and verification experiment on *legato* expression

○黒宮可織(ソニー-CSL, NeuroPiano)

- ◆本発表は、ピアノ演奏における legato 表現と、それを生み出すタッチ、身体動作との関連を明らかにした研究の概要を紹介する。
- ◆まず、レッスン観察より、演奏指導実践において指導の頻度が高い表現の1つに「legato」の表現があり、それに対して①腕の使い方(A1)、②指先と手首の位相(A2)、③指の使い方(A3)の3つの身体動作による指導方法の候補を同定した。
- ◆次に、実験において、ピアニストの表現意図による最適な legato と、上述 A1~A3 の動作を導入した際のタッチや身体動作を検証した。
- ◆その結果、ピアニストが意図した最適な legato の表現、および指導方法の動作を導入すると、離鍵の速度が遅くなり、その際には肩関節の速度が大きくなることが明らかになった。

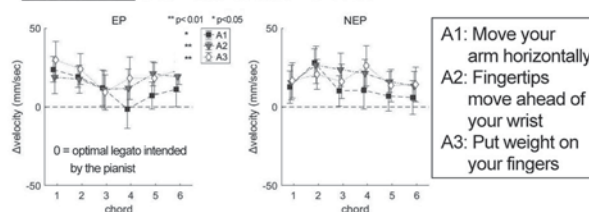


Fig.1: Speed of key release (mute_vel) when intervening with body movements for a performance intended to be optimal legato.

3-9-1

3-9-1 表情分析の活用による騒音の大きさに
対する顔表情と主観的印象の対応Correspondence between facial expression and subjective impression for
the loudness of noises using facial expression analysis

◎宮入徹(都産技研)

- ◆本研究では、表情分析ソフトウェアを用いて、顔表情による騒音の騒々しさ評価について検討した。
- ◆実験方法：3条件の音量 ($L_{Aeq} = 45, 60, 75$ dB) に調整したホワイトノイズ提示時の表情分析および主観評価を実施した (Fig. 1)。
- ◆75 dB 提示時において、「平常心」の表情確信度において有意な減少が確認され (Fig. 2)、主観評価との対応が見られた。

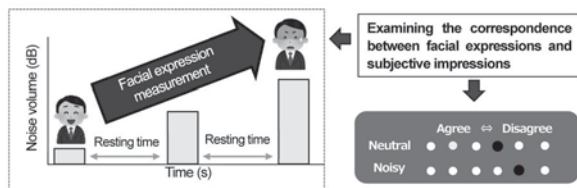


Fig. 1 Experimental outline

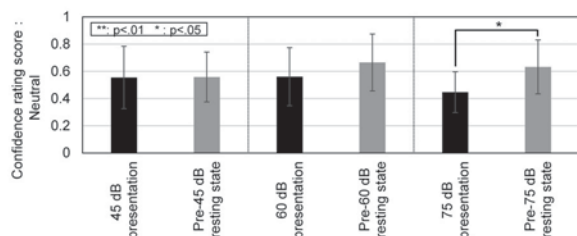


Fig. 2 Facial neutral scores

3-9-3

3-9-3 共用打ち合わせスペースにおける
ワーカーの評価構造と音環境の関係性
—半構造化面接手法による分析—Relationship between office worker's evaluation structure and sound
environment in a shared meeting space: Analysis using semi-structured
interview techniques☆遠藤環(茨城大院・理工), △横田健治, △佐土島哲次, △前田光,
△坂井剛(パナソニック ホールディングス), 辻村壮平(日大・工)

- ◆半構造化面接手法の一つである評価グリッド法によるインタビュー調査を行い、共用打ち合わせスペースにおけるワーカーの要求性能を整理した。
- ◆被験者毎の評価構造を統合して作成した全体評価構造図では、「アイデアが浮かぶ」、「効率的である」が最上位項目にみられた。
- ◆一般的な会議室と共用打ち合わせスペースの評価構造を比較した結果、相違点として共用打ち合わせスペースでは「おもてなし感」が重視されていることがわかった。
- ◆「自然音」がある環境や「静かすぎない」といった音環境が「集中できる」、「リラックスできる」などの心理的印象を介して知的生産性に寄与する構造が明らかとなった。
- ◆適度な環境音は「上質な雰囲気」や「おもてなし感」の形成にも関与し、空間ブランディングやホスピタリティ創出における音環境の重要性が示唆された。

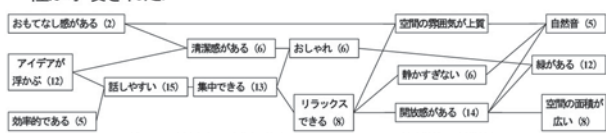


Fig. 1: Evaluation structure of participants: Excerpt

3-9-2

3-9-2 背景騒音および他者の存在を考慮した
オフィス製品の衝撃騒音に対する印象評価Subjective Evaluation of Impact Noise from Office Products
under the Influence of Background Noise and the Presence of Others

◎中村史香, 宮入徹, △服部遊(都産技研)

- ◆製品の付加価値の一つとして、静音性が様々な製品に取り入れられている。しかし、製品音の印象は使用環境により変化するため、静音性に関する明確な指標が存在しないという現状がある。
- ◆本稿では、使用環境として背景騒音と他者の存在・位置関係に着目し、これらが製品音に対する印象に及ぼす影響の把握を目的とした。
- ◆既報では定常音（ハンディファン）に対して検証を行ったが、本稿では衝撃性のある非定常音（ボールペンのノック音）を対象とした。
- ◆主観評価の印象と音源の評価量の関係を確認したところ、マスクトラウドネスにおいて高い相関が確認できた。
- ◆実験参加者と他者の位置関係を複数設定して音を提示したところ、位置関係が変わることで、印象が有意に変化することが確認できた。

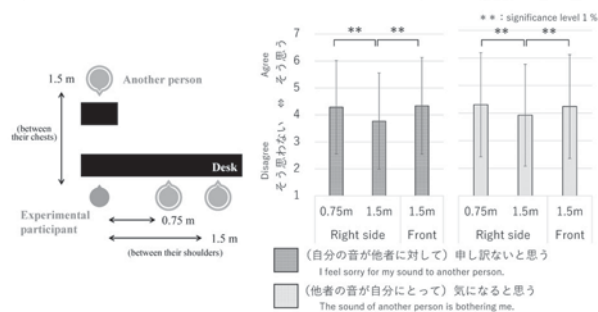


Fig. 1. Experimental conditions and averaged evaluation results

3-9-4

3-9-4 早稲田大学図書館における
ラーニングコモンズ設置の音環境への影響The effects of placing a learning common area on the soundscape of
Waseda University library

☆太田裕人, 及川靖広(早大理工)

- ◆早稲田大学中央図書館もラーニングコモンズ (LC) を導入し快適かつ目的別に利用できる空間を設計している。
- ◆騒音計で音環境を測定し、LC を設置していない理工学・戸山図書館の結果と比較することで LC の利用実態とその設置が音環境に与える影響を調査する。利用者数についても計測を行い、騒音レベルとの相関を調査した。
- ◆中央図書館は LC エリアを含め、他 2 館よりも騒音レベルは低く、LC は従来と変わらない利用のされ方であった。
- ◆中央図書館では利用者数と騒音レベルに相関があったが、他 2 館では有意な相関は得られなかった。

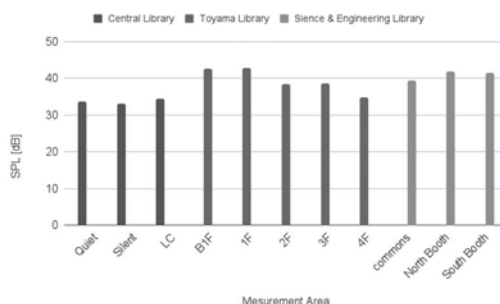


Fig. 1: SPL of each Measurement Area

3-9-5

3-9-5 ISO19029 から 10 年をみんなと振り返って

I looked back at the 10 years since ISO 19029 with everyone.

○川上央(日大・芸術)

- ◆鉄道駅: 駅員警笛 → 「ジリリリ」ベル → 発車メロディ → 廃止・統一化
- ◆家電品: 報知音 → 音声案内・ユーザインターフェース → 海外メーカの報知音・AI コンセルジュ
- ◆サイン音: 「ピンポン」 → 「ピヨピヨ」 → GPS・スマホ
- ◆ハートビル → バリアフリー → ユニバーサルデザイン → UI・UX → SDGs → ダイバーシティ

2016 年 ISO 19029 発行

2016 年 日本でSDGs実施指針

2017 年 日本で Alexa 搭載 Echo が発売

2020 年 コロナ・パンデミック

2021 年 東京オリンピック・パラリンピック

2023 年 Eye Navi リリース

2024 年 インバウンド

2025 年 大阪万博・生成AI

3-9-7

3-9-7 講無駄なダッシュをしたくない学生がデザインした鉄道駅改札階での到着サイン音には駆け込み乗車防止効果がありました

The Design of Sound Signature, Which Tells Passengers on the Ticket Gate Floor: The Behavior Observation Survey on the Effects of Preventing Passengers from Rushing.

○船場ひさお(駿河台大)

本研究は、相鉄いずみ野線・緑園都市駅に導入された到着サイン音「アンダンテ」の駆け込み乗車防止効果を検証したものである。「アンダンテ」は、電車の進行方向と種別を音の高さやテンポ、楽器の違いによって直感的に伝える音サインであり、フェリス学院大学と相鉄ホールディングスの共同研究により 2016 年に導入された。

本研究では、緑園都市駅と構造や乗降客数が類似する駅と比較し、行動観察によって駆け込み率および誤認駆け込み率を調査した。その結果、緑園都市駅は他駅と比較して両指標が有意に低く、サイン音が乗客の行動に与える影響が確認された。今後は、音の意味を理解しやすい利用者への配慮や視覚的案内との併用など、多様な利用者に対応する工夫が求められる。

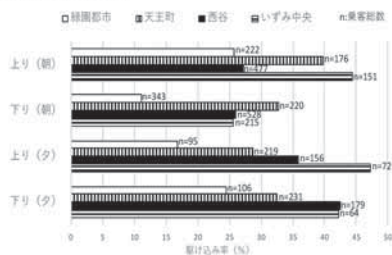


Fig.1: Percentage of people running on to the train

3-9-6

3-9-6 高齢者に優しい鉄道駅を目指して
最適な案内放送の提示レベルと発話速度を
提案しました

Proposal for the reproduction level and speech rate of public announcement at a railroad station with consideration for the elderly people

○辻村壮平(日大・工)

- ◆高齢者に優しい鉄道駅の音声案内放送を実現するために、これまでの著者らの鉄道駅の案内放送の制御手法に関する一連の研究結果を包括的に整理し、情報を聴き間違えない最適な案内放送の提示レベルと発話速度について検討した。
- ◆「聴き取りにくさを感じることはあるが聴き取り間違いは生じない性能」を満たす案内放送の最小 S/N 比と、その S/N 比による提示レベルにおいてうるささを感じる割合を算出した。さらに、聴き取りにくさも不自然さも感じない発話速度の範囲を明らかにした。
- ◆鉄道駅の案内放送の適切な制御には、駅空間の吸音力を高めることが非常に重要である。

Table: Comprehensive summary of the research results presented in our study.

Acoustic treatment	Absorptive	Reflective
Acceptable minimum S/N ratio	+2 dB or more	+6 dB or more
Percentage range of perceived noisiness by the announcement reproduced at the above S/N ratio	0 to 8 %	9 to 40 %
Optimum speech rate	6.5 - 7.5 mora/s	

3-9-8

3-9-8 大阪・関西万博開催の機運を醸成するために大阪メトロ中央線の列車接近・発車メロディをアップデートしました

New train arrival and departure melodies for the Osaka Metro Chuo Line have been updated to enhance the momentum of the Osaka-Kansai Expo

○岩宮 真一郎(日大ノ九大), 土田 義郎(金沢工大), 和氣 早苗(同女大), 藤本 由紀夫(京芸大), 西谷 喜久(soundplant), 小原 良夫(リブル)

- ◆大阪メトロ中央線のメロディ刷新へ! 「夢洲」への延伸に合わせて「中央線列車接近・発車メロディ制作委員会」が発足しました [1]
<https://www.youtube.com/watch?v=dnVfMncswLcs>
- ◆中央線「新メロディ」の試作3案を公開!
メロディの試聴とお客さまアンケートを実施します [2]
https://subway.osakametro.co.jp/news/news_release/20240412_chuo_line_melody.php
- ◆「新メロディ」プロジェクトにテレビが潜入!
どのメロディがいい? 深夜の試聴会と午前4時に及ぶ委員会 [3]
<https://www.youtube.com/watch?v=7-vqj9nkZUI&list=TLPQMjUwOTIwMjI5SGhH4HsBA3DA&index=4>
- ◆新メロディが完成! 事前に公開します: 分かりやすく親しみやすい大阪の歴史の上に花開く明るい未来が感じられる音 [4]
https://subway.osakametro.co.jp/news/news/subway/20241018_chuo_line_melody_kansei.php?_ga=2.187787502.714401820.1729228907-789350232.1725945993&_gl=1*gaevrc*_ga*Nzg5MzUwMjI5SGhH4HsBA3DA&index=4
- ◆ついに完成! 新メロディがデビューしました [5]
<https://www.youtube.com/watch?v=06nlwDgeso4&t=13s>



3-9-9

3-9-9 「声のようを読んでみよう入門編」
ルールと遊び方を考えるLet's Read the Patterns of Voices:
Developing Rules and Gameplay

○竹内京子(順天堂大), 青木直史(北大), 荒井隆行(上智大),

△大金さや香(国際医療福祉大), △鈴木恵子(おぎはら耳鼻咽喉科),
世木秀明(千葉工大), 秦若菜, △村上健(北里大), 安啓一(筑波技術大)

- ◆ことばのリハビリを行う言語聴覚士という職業がある。その養成校では、音響学が必修科目。しかしながら、学生が最も苦手で嫌いな科目であり、改善が求められている。
- ◆音声の音響分析を楽しく学ぶためのカードゲーム「声のようを読んでみよう入門編」を作成し、その使い方を考えた。その例を紹介する。
- 1) それゆけ! もんもんマスター: 3名から5名, 小学生以上向け, 全員でカードを囲み, 歌を歌いながら, 「あ」「い」「う」「え」「お」のサウンドスペクトログラムの違いを学ぶ。
- 2) もんもんかるた取り: 全カードを使用ではなく, 例えば, 力行とサ行だけなど, カードを制限し, 読まれた札を皆で取り合う。
- 3) もんもん単語ゲーム: 上級者向け。カードのサウンドスペクトログラムで単語を作成する。連続音ではないが, 分ち読み単語をどこまで当てられるかを競う。
- 4) もんもん50音並べ: カードを50音図の通りに並べる。「あいうえお」の区別を学び, 例えば, ウ段だけ並べ, 子音部が共通したその他のカードを, サウンドスペクトログラムだけを見て, 並べる。
- ◆ 今後は, 学生や使用者の意見も踏まえ, 内容を改善していきたい。さらに, その他の遊び方も広く募集していきたい。

3-9-11

3-9-11 Ambisonics を応用した
合奏練習コンテンツの生成

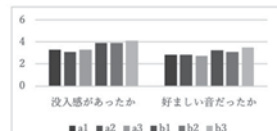
Mixing for ensemble exercise source applying Ambisonics

☆喜多叶奈子, 高橋義典(工学院大・情報)

- ◆合奏練習では他パートとのタイミングやバランスを把握することが重要。練習中の合奏の録音を聴いて個人練習する際、自分のパートの音を取り除くことで、本来の合奏に近い練習ができる。
- ◆本研究ではマンドリン五重奏を Ambisonics マイクロホンで収録後、練習者パートを弱めた練習用コンテンツを作成した。チャンネルミキシングによるビームフォーミングと頭部伝達関数によるバイノーラル合成によって、任意方向の楽器音(練習者パート)を弱めた「ターゲット除去コンテンツ」と、全方位の音を含む「全方位コンテンツ」を作成し、各コンテンツを聴きながら演奏する実験を実施した。
- ◆9名の奏者に演奏実験を実施し、2種類のコンテンツについて印象評価を行った結果、楽器ごとに処理の影響に差はみられたが、全体的には顕著な差はなかった。一方、ターゲット除去コンテンツでは、音質の不自然さや一体感の低下が課題として指摘され、減衰処理の改良が必要であることが示唆された。今後は音源分離精度の向上や音質の改善により、より自然な合奏体験の実現が期待される。



Fig.1: Recording condition

Fig.2: Experience results
blue: target removed, red: all direction

3-9-10

演奏環境の変化がドラム演奏における
テンポキープの精度に及ぼす影響An effect of varying room acoustics on
the accuracy of keeping tempo in drum performance

☆上田恭平, 辻慎也, 荒井隆行(上智大)

- ◆本研究では、非残響下と残響下でドラム演奏におけるテンポをキープする精度がどのように変化するのかを検討するため、電子ドラムを用いた演奏課題を行う実験を実施した。
- ◆演奏課題では、(1) 残響時間の長短、(2) 音色は似ているが音の持続時間が異なるドラム音源 A と音源 B、(3) Beat Per Minutes (BPM) の大小を条件とした。また、演奏データとして打叩時刻を収集した。
- ◆テンポキープの精度 (打叩時刻ずれ) を以下の式により定量化した。

$$\text{打叩時刻ずれ} = |t_i - T_i|$$

* t_i : i 回目の理想タイミング; T_i : 実際の打叩時刻

- ◆Aligned Rank Transform ANOVA の結果、RT・Sound、それらの交互作用が打叩時刻ずれに有意に寄与する要因であった。それらの条件について、平均値と標準偏差を Table 1 に示す。

- ① 残響時間が長い場合で打叩時刻ずれが有意に大きくなる。
- ② 音源 B を用いた場合に打叩時刻ずれが有意に大きくなる。
- ③ 音源により残響の影響度合いに有意差がみられた。音源 B で残響時間が長いとき、最も打叩時刻ずれが大きくなった。

Table 1 Mean and standard deviation of data

RT	Sound	mean	SD
2.1 s	A	0.375 s	0.374
2.1 s	B	0.377 s	0.339
0.1 s	A	0.247 s	0.281
0.1 s	B	0.342 s	0.400

3-9-12

3-9-12 親しみやすい防災放送訓練システムの
ための画面構成に関する検討A Study on Disaster Prevention Broadcast Training System
with User-Friendly Screen Structure

○土田 義郎, 高野 佐代子(金沢工大)

訓練は継続して行うことでその効果が定着する。そのためには無味乾燥な画面よりも、見やすく親しみやすい画面構成が求められる。今回、教示者のアバターなどを使用した 13 種類の画面構成を用意し、評価グリッド法によって相互を比較した。その結果、親しみやすさのためには下記の条件が必要であることが明らかとなった。

- ◇教示動画が大きく、アバターが見やすい
- ◇教示動画下部に字幕が入っている
- ◇構成要素がバランスよく枠内に収まっている

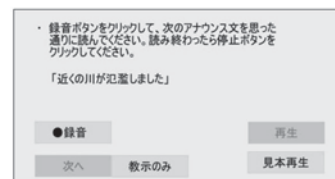


Fig.1 従来の訓練システム画面



Fig.2 好ましいとされた訓練システム画面の例

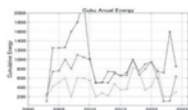
3-11-1

3-11-1 中部海域の深部震源域における発震の太陽引力との関係

Relationship between generation in deep hypocenter of the central ocean and solar gravitational forces

○菊池年晃(防衛大)

中部海域には地震を頻発する震源域が幾つか存在する。その震源震動を地震のエネルギーに基づいたパラメータ、累積エネルギー等を用いて定量化解析を行った。この定量化により異なる震源域や時期の震動構造を比較し得ることができる。この海域の震源域は二分化構造で、深部震源域はエネルギーの発震機能を有するが浅部震源域はエネルギーの滞留、蓄積の機能を有する。中部海域の震源、東海沖(青)、紀伊半島沖(赤)及び壱州灘(緑)の震源域の年間累積エネルギーの年次変化を示す。これらの3曲線は類似した年次変化をしており同一プレート上に位置していることが分かる。また2004年から2025年にわたる変化は周期性を包含している。その発端は、2004年9月に、紀伊半島南東沖震源域の浅部震源域でM7.4を始めとする大きな地震が頻発した。深部震源域はエネルギーの発震機能を有するが浅部震源域はエネルギーの滞留、蓄積の機能を有する。従って、2004年の大きな地震の頻発により、それ以前の蓄積エネルギーは消失したと考えられると、その後の、震源域で発生した累積エネルギーは重要な特性を持つ。その出発は2005年に発生したM5.4の単発パルスで、その後約1年間隔で4年継続した。これらのパルスの発震深度はプレートの最先端であり、プレートの新たな動きと考えられる。従って、図に示した震動変動は周期的な性質をもつので、今後の予測に役立つことができる。



3-11-2

3-11-2 複数 AUV 協調群制御実海域試験における音響測位結果に関する検討

Examination of acoustic positioning results in multi-AUV cooperative at-sea group control tests

○太田和彦, 松本さゆり(PARI), △佐藤匠, △金岡秀, △岡本章裕, △稲葉祥梧, △篠野雅彦, △藤原敏文(NMRI)

- ◆総合科学技術・イノベーション会議のSIP(戦略的イノベーション創造プログラム)「海洋安全保障プラットフォームの構築」(研究推進法人: JAMSTEC)に基づき、令和6年10~11月に駿河湾で2機のAUVを用いた相対位置制御に関する試験を実施した(Fig.1)。AUVは海底から一定高度で航行しつつ搭載ソナーの測位結果に基づく自動制御(Reference-Follower 群制御)により、両機間の距離は約150~250mでほぼ一定に保たれることが確認された。
- ◆本稿はこの測位に係る伝搬環境について、ガウシアンビームを用いた音場シミュレーションコードBel lhopによりAUV間の伝搬状況を解析すると共に音響的観点から制御可能域に関する検討を行った。

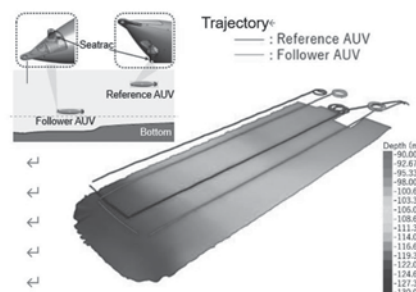


Fig.1: Experimental configuration.

3-11-3

3-11-3 南海トラフ海域ケーブル型海底地震計による鯨類鳴音の検出

Whale vocalizations detected with cabled ocean bottom seismometers in the Nankai Trough region

○岩瀬良一(JAMSTEC)

- ◆2024年に防災科学技術研究所が南海トラフ海域に設置した海底ケーブル型地震・津波観測網(N-net 沖合システム)を構成する地震計の波形データの解析により、鯨類鳴音と推定される信号を検出した。
- ◆東日本の太平洋沿岸では、各種の海底ケーブル型観測システムを構成する地震計によるナガスクジラ鳴音の検出が先行研究により報告されているが、東海地域より西側の海域については、海底地震計網による鯨類鳴音の検出報告は既往の文献に見当たらない。
- ◆解析対象期間は2025年3月8日から10日までの3日間であり、N-net速度型地震計(速度計)の波形データの鉛直成分を、併設の加速度計で計測される重力の方向を基に算出して使用した。
- ◆まず全観測点の速度計波形について1時間ごとにスペクトログラムを作成し、鯨類鳴音の特徴を有する信号の存否を目視により確認した。
- ◆その結果検出した鯨類鳴音は、過去の文献に示されているナガスクジラ鳴音のスペクトログラムと形状が似ていることから、ナガスクジラの鳴音である可能性が高い。
- ◆次に14~31 Hzのバンドパス・フィルタを適用した各観測点の波形を調査した。その結果、各観測点における信号の出現タイミングから、音源はほぼ単一個体が発出した鳴音と推定され、音源の個体は、本解析対象期間の3日間にN-net観測網を西から東、つまり九州東方沖から四国の室戸岬沖までを移動したものと推定される。

3-11-4

3-11-4 日本海北西部の季節変化する音速プロファイルが音波伝搬に与える影響

Influence of seasonally varying sound speed profiles on sound propagation in the northwestern Japan Sea

○鶴ヶ谷芳昭(三陽精工), 菊池年晃(防衛大)

- ◆日本海には日本固有水と呼ばれる冷たい水塊が存在しており、年間を通して水温は変化しない。一方、300mより浅い深度では、冬季の冷却と寒流、夏季は太陽輻射の増大による水温の上昇と暖流の勢力増大とダイナミックに変化する。その結果、夏季においては海面付近の垂直音速傾度が増大し、音波は海底方向への急速な屈折を生じるため、収束帯(CZ)伝搬となるが海底への透過が増大し、距離方向伝搬は小さくなる。一方、冬季、サウンドチャンネル軸深度が上昇し、浅い深度の音源においては、浅い深度に音波の集中を生じる。音源深度(SD)が深くなると音波伝搬に対する季節の影響はほとんどなくなる。

- ◆Fig.1にSDが10mにおける音波伝搬の比較を示す。周波数は1200

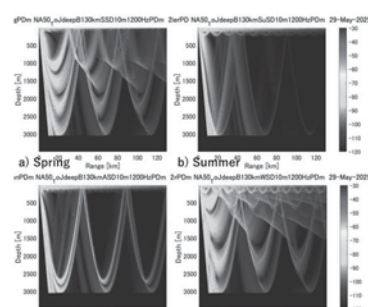


Fig.1 Seasonal sound propagation profiles

Hzである。夏季(右上)、CZ伝搬を生ずるが2ndCZは明確でなくなる。冬季(右下)、海面付近に上向き屈折による音波の収束が生じる。秋季、春季は季節の移行形態と考えることができる。

3-11-5

3-11-5 U-Net による血管壁セグメンテーションと組織内腔境界の検出

Vascular wall segmentation and detection of tissue lumen boundaries using U-Net

☆濱井悠里, 大村眞朗, 長岡亮, 長谷川英之(富山大・工)

- ◆シミュレーションデータに基づいた深層学習モデル(U-Net)を用いて頸動脈壁の自動セグメンテーションを行い、壁組織と血管内腔の境界を高精度に検出することを試みた。
- ◆組織と内腔境界の検出には、振幅包絡波形の立ち上がりに着目した。
- ◆Fig. 1は、in vivo データおよびU-Net の出力確率像 (Predict) に対し、同一健康被験者の頸動脈における B モード像上に組織内腔境界の位置をプロットしたものである。
- ◆in vivo データの検出と比較して、学習により推定された確率から境界を抽出することで、スムーズな境界が得られた。

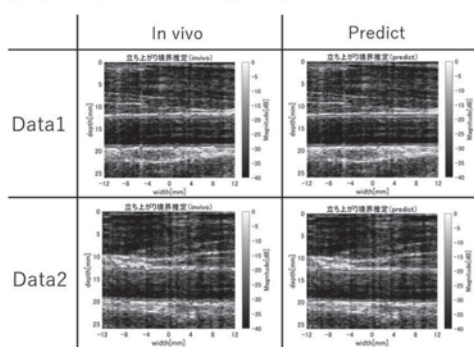


Fig.1: Detection of the boundary between the vessel wall and lumen in in vivo data and deep learning prediction results

3-11-7

3-11-7 3D U-Net による血流イメージングの基礎検討

Preliminary Investigation of Blood Flow Imaging Using 3D U-Net

☆王 洪鵬, 李 加一, 大村眞朗, 長岡 亮, 高 尚策, 長谷川英之(富山大・工)

- ◆空間および時間的情報を同時に学習可能な 3D U-Net を用いて、超音波血流画像のクラッタ抑制を検討した。
- ◆健康者男性の頸動脈から、7.5 MHz のリニアアレイプローブを用いて 2 回分の心拍の RF データを取得し、DAS ビームフォーミングにより処理してモデルの入力とした。入力データを SVD フィルタ処理したデータを教師データとして使用した。
- ◆Fig. 1 は、3 層のエンコーダとデコーダを含む 3D U-Net モデルである。
- ◆エンコーダにおいては 1 チャンネルの $192 \times 192 \times 8$ サイズの入力データは 32 チャンネルに拡張され、 $3 \times 3 \times 3$ の畳み込みを経て、最大プーリング($2 \times 2 \times 1$)で空間分解能を半減させ、最深層で 256 チャンネル、サイズ $24 \times 24 \times 8$ となる。その後、デコーダ処理を行い、最終出力は 1 チャンネルの元のサイズ $192 \times 192 \times 8$ に戻され、特徴を学習する。

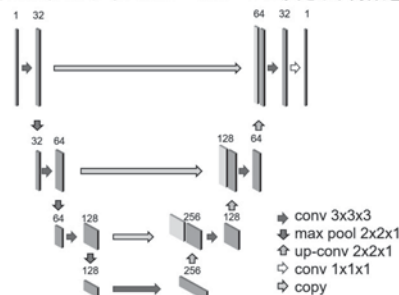


Fig. 1 Architecture of the 3D U-Net in this study.

3-11-6

3-11-6 15 MHz 超音波による血流速度推定の精度検証

Accuracy Verification of Blood Velocity Estimation by 15 MHz Ultrasound

☆南端 佑都, 大村 眞朗, 長岡 亮, 長谷川 英之(富山大・工)

- ◆高周波超音波の血流測定の利点として、空間分解能が向上し、血流信号の信号対雑音比も向上するため、浅部の血流を詳細に解析できる。
- ◆高周波リニアプローブのシミュレーション環境で得られたデータに対し、ブロックマッチング法における窓サイズの違い (Table 1) が層流分布の血流速度の推定精度に与える影響を比較した。
- ◆Fig. 1 はフレームごとに各空間位置の相対誤差を取り、フレーム間の平均および標準偏差を表示した図である。
- ◆窓サイズを小さくしていくと、相対誤差の平均は 10% 以上、エラーバーも平均 $\pm 10\%$ ほどで大きくなることが確認できる。
- ◆窓サイズを小さくすることで局所の速度分布を詳細に解析できる可能性が考えられるが、それに伴い空間のばらつきや相対誤差が Win1 と比べて平均最大 3% ほど上昇してしまうことが確認された。

Table 1 Condition of window size.

	lateral [mm]	depth [mm]
win1	3.0	1.50
win2	2.4	1.25
win3	2.0	1.00
win4	1.5	0.75
win5	1.0	0.50

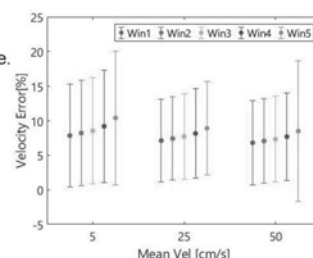


Fig. 1 Relative error in each velocity setting.

3-11-8

3-11-8 赤血球集合度評価のための血液の超音波伝播減衰計測の検討*

Investigation of ultrasound attenuation measurement of blood for evaluating red blood cell aggregation

☆新山晴斗, 政氏伸夫(東北大),

△八代 諭, △石垣 泰(岩手医大), 荒川元孝(東北大)

- ◆我々は、非侵襲かつ定量的な赤血球集合体サイズ推定法に対する研究を進めている。本手法では、計測スペクトルの補正のために、ヒト血液の減衰係数を計測する必要があるが、従来の計測法^[1]ではその精度に課題があった。この原因は、従来法が血流の影響を考慮していないためと考え、その影響を考慮した減衰係数計測法を検討した。
- ◆伝播経路の一部を水で置換し、血管内腔の浅部と深部に超音波ビームの焦点を合わせた計測を 2 回行い、それらのパワースペクトルの差分から血管内腔径方向の平均的な減衰係数を算出する置換法を適用した。
- ◆ヒト手背静脈に対する in vivo 実験の結果を Fig. 1 に示す。置換法で文献値^[2]に近い血液の減衰係数が得られた。この結果から、血流の影響を考慮することの妥当性が確認できた。

[1] M. Arakawa, et al., Front. Phys., 11, 1077696, 2023.

[2] B. E. Treeby, et al., Ultrasound Med. Biol., 37, 289-300, 2011.

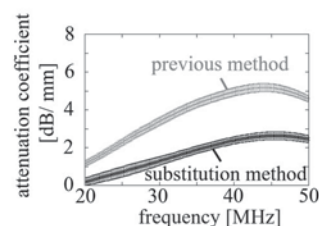


Fig. 1: Attenuation coefficients of blood calculated using the substitution and previous methods.

3-11-9

3-11-9 硬膜外麻酔補助のためのステアリング
超音波の送受信条件に関する検討

A study on transmission and reception conditions for steering-transmitted ultrasound for assistance of epidural anesthesia

☆副島湧介, △大西詠子, △山内正憲, 荒川元孝(東北大学)

【目的】我々は硬膜外麻酔の穿刺位置である胸椎間隙抽出の鮮鋭化を目指し、ステアリング超音波の送受信条件を検討してきた。本報では、骨の鏡面反射方向に存在する素子の振幅平均と受信全素子におけるそれとの比を係数 $A(x, z)$ として算出し、その2乗を出力画像に乗ずる手法の効果を評価した。

【実験】20代男性被験者1名に対して、麻酔針穿刺位置となる胸椎の傍正中縦断面を計測した。ステアリング角 ψ は -35° ～ 35° (5° 間隔)として超音波を送信し、全素子で受信した。

【結果】Fig. 1に、(a)一般的な手法($\psi = 0^{\circ}$)、(b)これまでの提案手法(波線を境目に左側 $\psi = -5^{\circ}$ 、右側 $\psi = 10^{\circ}$ とし、指向性補正とビームフォーミングに使用する素子を変更)、(c) $A(x, z)$ の2乗を(b)の画像に乗じる手法によるB-mode像を示す。(a)、(b)と比べ、(c)では骨が強調され、間隙の視認性が向上した。(c)における骨と間隙のCN比は左右の画像の平均で(a)に対して13.5 dB、(b)に対して8.1 dB上昇し、この手法によって描出能が向上する可能性を示した。

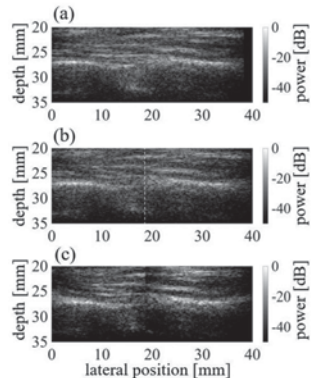


Fig.1: B-mode images obtained by (a) the conventional method ($\psi = 0^{\circ}$), (b) previously proposed method ($\psi = -5^{\circ}$ and 10°), and (c) multiplying the amplitude ratio factor $A(x, z)^2$ by (b).

3-P-2

3-P-2 全身振動と低周波音を同時に暴露可能な
実験装置の試作

Prototype experimental device that allows simultaneous exposure to whole-body vibration and low-frequency noise

◎笹森咲依, 岩永景一郎, 土肥哲也(小林理研)

◆全身振動(鉛直振動)と低周波音を同時に暴露可能な実験装置を試作(Fig. 1)

◆4本のエアシリンダで支持された座面に加振器で全身振動を暴露

◆低周波音は上部のスピーカから暴露可能

◆8 HzでVL=55~70 dB、高調波振動とのS/Nは20 dB以上(Fig. 2)

◆振動と低周波音の同時暴露の実験例(Table. 1)

8 Hz, VL=55 dB(固定)で40 Hz純音の音量を変化



Fig. 1: Prototype experimental device

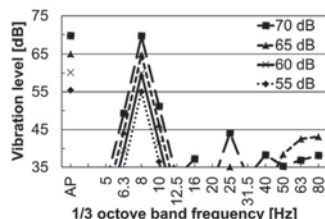


Fig. 2: Frequency response of vibration

Table. 1: Example of simultaneous exposure

Vibration		Sound	
Frequency	Vibration level	Frequency	Sound pressure level
8 Hz	55 dB	40 Hz	70 dB
			80 dB
			90 dB

3-P-1

3-P-1 背後空気層を持つ有孔板を用いた
家屋内低周波騒音の低減対策
-実家屋における「吸音ベッド」の実験結果-

Measures to reduce low-frequency noise in the actual house using sound-absorbing materials

○土肥哲也, 岩永景一郎(小林理研)

◆家屋内で簡便に実施可能な低周波騒音の低減対策方法を模索する目的で、家屋内に吸音体を後から設置する方法で家屋内の低周波騒音がどの程度低減するか実家屋を用いた実験的手法により検討した。

◆吸音体として背後空気層を持つ有孔板(高さ90 cm, 横幅90 cm, 奥行き約30 cm, 6面体の前面のみに直径6 mmの孔: 図1)を3個製作し、室内に「吸音ベッド」として配置した(図2)。就寝時における枕元付近の低周波騒音がどの程度低減するかを検証した。

◆建物外からスピーカで100 Hzの純音を建物室内に入射した。「吸音ベッド」を設置した場合の枕元のSPLは、無対策条件に比べて最大26 dB低減し、就寝時における低周波騒音が低減する可能性が示唆された。

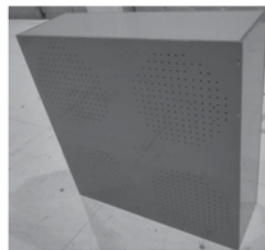


Fig.1: Sound absorber



Fig.2: Sound absorber bed

3-P-3

3-P-3 広範囲の螺旋角を有する垂直軸風車翼の
空力騒音調査

Aerodynamic noise investigation of vertical axis wind turbine blades with a wide range of helix angles

○真志取秀人(産技高専), 高橋義典(工学院大), △水牧真唯(北陸先端大・院)

◆垂直軸型風車の翼形状を螺旋変形させると空力騒音が低減することを、螺旋角150度まではこれまでの研究で確かめていた。

◆しかし、より大きな螺旋角条件を適用したときに風車から発生する空力騒音特性については、まだ計測されていなかった。また同じ螺旋角条件でも、風車翼枚数が変化するとその空力騒音特性に変化が生じることが予想された。

◆本研究では、螺旋角条件を360度まで拡張し、より広範囲に螺旋変形させた場合の垂直軸風車の空力騒音について、また同時に風車翼枚数を2枚から4枚へと変化した時にどのような特性を示すかについて、調査する実験を行った。

◆風車の空力騒音は、すべての螺旋角条件で低減された。さらに、螺旋翼を適用した場合の空力騒音低減効果は、風車翼枚数の増加に伴い顕著になった。この傾向は、垂直軸風車の螺旋角を360度まで変化させた場合でも変化しなかった。



(a) Straight blade



(b) Helical blade

Fig.: Shape of vertical axis wind turbine

3-P-4

3-P-4 無響室内での垂直軸風車翼の
空力騒音調査

Aerodynamic noise investigation of a vertical axis wind turbine blade in an anechoic chamber

○真志取秀人(産技高専), 高橋義典(工学院大), △水牧真唯(北陸先端大・院)

- ◆垂直軸型風車の翼形状を螺旋変形させると空力騒音が低減することを、これまで実験で検証していた。しかし、これらの結果は無響室内における無風状態での計測であった。
- ◆風車と風洞を組み合わせて風車翼の空力騒音を測定する場合、測定値に空力騒音以外の雑音が含まれることは否定できない。一方、無風環境では、通常運転とは異なり、風車をモーターで回転させる必要があるが、残響の影響を低減することができる。そしてこれらの結果についての比較検証はまだ行われていなかった。
- ◆本研究では、無響室と風洞における、螺旋変形翼を備えた垂直軸風車の空力騒音計測の違いを実験的に検証した。
- ◆その結果、無響室で実施した実験では、風車翼の螺旋変形による空力騒音の変化が、風洞実験時より明確に示された。
- ◆この結果は、風洞実験時には風車周囲の風向によって空力騒音の発生領域が実質限定されるためと推定しており、今後検証予定である。



Fig.1: Shape of vertical axis wind turbine

3-P-7

3-P-7 時間領域有限要素法による室内音場解析
の吸音材のモデル化に関する研究

—音圧および粒子速度の減衰の実測値と解析値の比較—

Modeling of absorption material on sound field analysis in rooms by time-domain finite element method

-Comparison of decay of sound pressure and particle velocity between analytical and measurement results-

○富来礼次, 岡本則子, 大鶴徹(大分大),

△長東直大, 菅野将郎, 川岸右京(大分大院)

- ◆RC 矩形室に吸音材を設置した音場を対象に、吸音材の比音響インピーダンス比として、EA 法から測定された値($z_{n,EA}$)と管内法で測定された値($z_{n,Tube}$)を使用した時間領域有限要素法 (TDFEM) による解析を実施し、得られた音圧および粒子速度の減衰を実測値と比較した。
- ◆他の音場で実施した既往の研究同様、 $z_{n,EA}$ を用いた TDFEM による解析で得られた音圧レベルの減衰が、 $z_{n,Tube}$ を利用した場合に比べ、実測値と近似した。
- ◆一方、粒子速度レベルの減衰では、Fig. 1 のように、吸音材と法線方向(x 方向)では、 $z_{n,Tube}$ を用いた解析値が実測値と近似し、その他の方向(y, z 方向)の吸音材近傍(Receiving point 4, 8)では $z_{n,EA}$ を用いた解析値が実測値と近似することを示した。

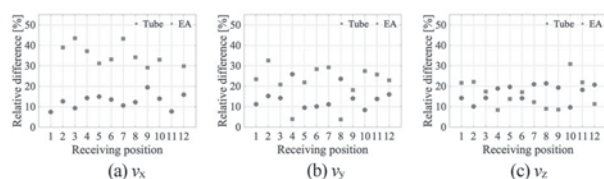


Fig.1: Difference of decay curves of particle velocity level between measurement and FE-analysis

3-P-5

3-P-5 運河沿いの室内における
船舶航行音の分析

Analysis of boat navigation noise in a room along a canal

☆山本晴貴, 武藤憲司(芝浦工大)

- ◆運河沿いの集合住宅室内で計測された小型船舶の航行音を分析した。
- ◆屋間のカーテンの開いた室内における船舶騒音分析を行った。
- ◆24 時間の騒音レベル推移を Fig.2 に示す。運河を航行した船は 73 艘で、タグボート、モーターボート、屋形船の 3 種が航行していた。この日、室内で計測された船舶騒音の最大値は、朝 5 時 53 分のタグボートで 59.1 dB であった。これは 17 時に記録された防災無線のチャイム 60.1 dB と同等の騒音レベルであった。



Fig.1 Indoor measurement location along the canal

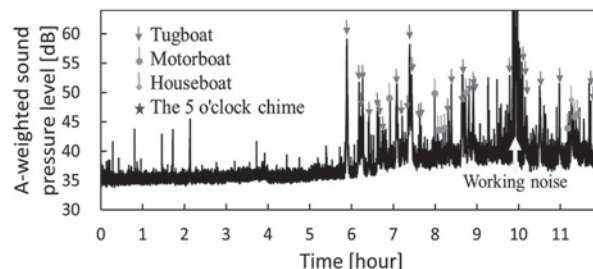


Fig.2 Transition of boat noise level indoor along the canal 24-hour

3-P-8

3-P-8 編物の吸音特性に関する基礎的研究
—統計的手法による予測理論の構築—

Sound absorption characteristics of knitted fabrics

- Modeling by a statistical method -

☆高橋響, 阪上公博(神戸大・工)

- ◆編物の吸音特性に関する研究は用途の限定されたものが多く、先行研究が乏しいため、建築内装材としての利用を目的とした実験的検討と、統計的手法を用いたモデル構築を行った。
- ◆多孔質型吸音の予測理論である Delany-Bazley モデルを参考に、多数の編物試料の特性インピーダンスと伝搬定数を音響管で測定し、各実部・虚部を冪乗則で表したときの定数を、最小二乗法を用いて求める。
- ◆分析の母集団とする編物試料の組み合わせについて、A~G の異なる 7 パターンで計算を行った。試料数が多く、異なる素材や編み方の組み合わせを持つパターン D,E,G で吸音率が互いに近くなっていることから、素材や編み方などによらず、共通の傾向を示す固有の吸音特性の特徴があるのではないかと考えられる。

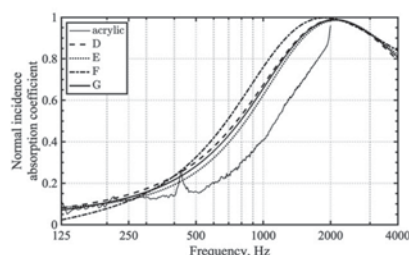


Fig.1: Prediction result(D,E,G) (acrylic (7), single crochet, back cavity 20mm)

3-P-9

3-P-9 エアダンピングを活用した 膜振動・ヘルムホルツ共鳴連成型広帯域 吸音音響メタマテリアルを用いた 空間共鳴低減評価

The resonance reduction evaluation of developed sound absorber using membrane vibration and Helmholtz resonator with air damping by mesh

◎ 後藤 達彦(株) 東芝

- ◆ 提案するエアダンピングを活用した膜振動・ヘルムホルツ共鳴連成型広帯域吸音音響メタマテリアルを複数個用いた音響評価を実施
- ◆ 小空間内に生じる共鳴モードが、メタマテリアル吸音帯域(2種合わせて450Hz~1100Hz)にて、10dB程度抑制されることを示した

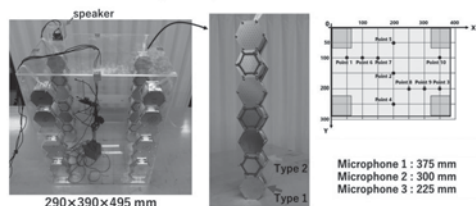


Fig.1 : Measurement conditions.

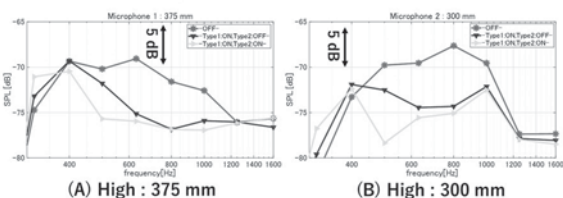


Fig.2 : Measured attenuation effect (point 3) .

3-P-11

3-P-11 深層学習を用いた矩形音場における 複数領域の吸音率推定の基礎的検討

Fundamental study on estimation of sound absorption coefficients in multiple boundary regions for a rectangular sound field using deep learning

☆大河佑里子, 佐藤元, 池田雄介(東京電機大)

◆ 背景

- 近年、測定と数値シミュレーションを併用した、深層学習ベースの吸音率推定手法が提案されてきた
- これまでの研究では、平均吸音率や各壁面を推定する手法が提案されている

◆ 提案手法

- 深層学習を用いた吸音率推定において、1つの壁面を複数の領域に分割し、それぞれの領域において吸音率を推定する。
- 1壁面を2分割にして8境界、4分割にして16境界の吸音率をそれぞれ推定

◆ シミュレーション実験

- 2分割で相対誤差4%を達成。境界が増えると相対誤差が悪化

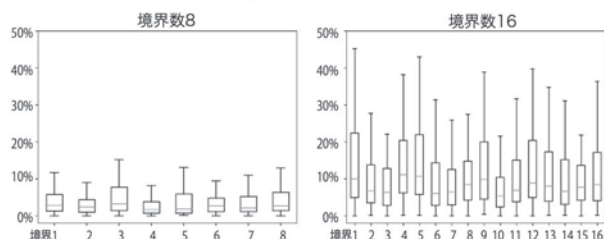


Fig.1: Relative error in sound absorption coefficient for each boundary

3-P-10

3-P-10 折紙テッセレーションを用いた形状可変な 吸音材料の試作

Prototyping deformable sound-absorbing structures using origami tessellation.

☆小林直生(慶大), △吉野弘一(TOPPAN デジタル),
矢田部浩平(農工大), △鳴海 結也(慶大)

本稿では形状可変な吸音材として、フェルト材を用いたミウラ折り構造を試作した。また、残響時間を調査し、吸音効果を確認した。

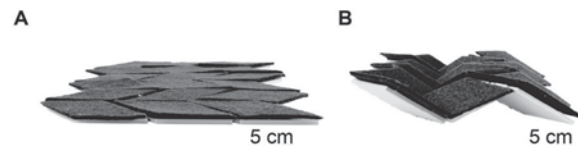


Fig.1: Prototyping deformable sound-absorbing structures.
A: developed state, B: folded state.

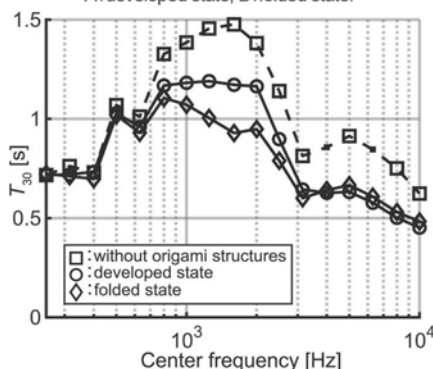


Fig.2: Average T_{30} in a simplified reverberation chamber at 1/3 octave-band center frequencies (250 - 10000 Hz).

3-P-12

3-P-12 プレナムドアの遮音性能に対する 開口位置の影響と空気層内障害物の効果

Effect of opening position and obstacles in the air layer on the sound insulation performance of plenum doors

☆兒玉和紗, 阪上公博(神戸大院・工)

- ◆ プレナムドアの遮音性能について、パネル上の開口位置を変化させた際の効果を比較した。開口面積は一定で、開口位置の違いについてのみ検討している。
- ◆ 開口位置による遮音特性の差は、低中域におけるピークやディップの周波数と音響透過損失の差として表れたが、最適な条件は一概に判断できず、用途や環境に応じた選択が必要といえる。
- ◆ プレナムドアの空気層内に円柱や角柱を設けた場合の遮音性能を検討した。また、円筒型 MPP や角柱型 MPP についてもその効果を検討した。
- ◆ 円柱、角柱の効果は高域を中心に確認できた。また、円筒型 MPP、角柱型 MPP の場合は目標とした周波数と、円柱や角柱を挿入した際に効果が見られた周波数で性能が向上しており、MPP の吸音効果と障害物としての反射による影響を併せ持った結果が得られた。

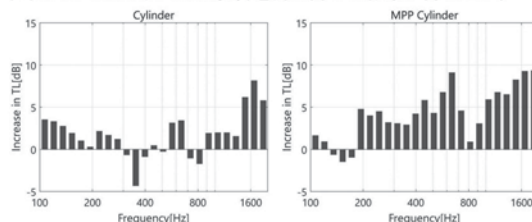


Fig.1: Increase in sound transmission loss due to insertion of cylinders and MPP cylinders in the plenum door along the z-axis direction.

3-P-13

3-P-13 バンド構造解析を用いた 階層型ソニック結晶のバンドギャップ評価 Evaluation of Band Gaps in Hierarchical Sonic Crystals Using Band Structure Analysis

☆和泉誠大, (近畿大院), 菅原彬子, △堂寺知成, 平栗靖浩(近畿大)

- ◆環境騒音は大きな課題をもたらし、効果的な遮音が求められている。ソニック結晶 (SCs) は遮音と換気の両立が可能な建築材料となる可能性を持つが、ソニックバンドギャップ (SBG) が狭く、異方性の問題がある。
- ◆本研究では、SCs に階層構造を導入した2次元三角格子型階層型ソニック結晶 (HSCs) の遮音特性を、有限要素法による数値解析とバンド構造解析による理論解析により検討した。それぞれの解析結果から、階層構造の採用により等方的な SBG が形成されることが分かった。
- ◆また HSCs の構造に変化を加えることで、周期構造としての対称性を崩すことにつながり、遮音性能が大幅に向上することが確認できた。本検討よりの結果から、HSCs は入射角依存性の緩和と SBG の低域化に寄与する可能性が示唆された。

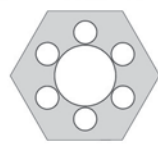


Fig.1 Study Model.

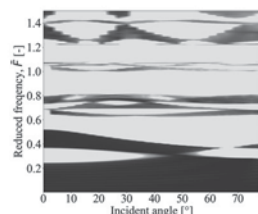


Fig.2 Variation in IL owing to incident angle of sound waves.

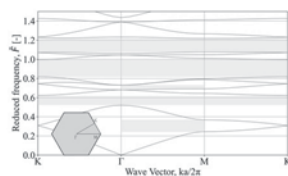


Fig.3 Band structure.

3-P-26

3-P-26 TODA BUILDING ホールの音響特性 —音響拡散を考慮した矩形のホール— Acoustic characteristic of the halls in TODA BUILDING —Rectangular halls in consideration of sound diffusion,

○土屋裕造, 佐脇真平(戸田建設)

- ◆2024年10月に竣工した TODA BUILDING の4階に、コンベンション、展示会など多用途に活用できる2つのホールが設置された。
- ◆壁面の音響拡散を考慮するため、非周期構造でありながら対称性のあるデザインが特徴的なペンローズタイル模様の音響拡散性能を有する音響パネルを採用した。実際に採用するに当たり、全体の響きと共に、ホールの設計要件に合わせた調整を行った。
- ◆その結果、これらの矩形のホールにおいて、音響拡散を考慮したこのパネルを採用することにより、良好な室内音響性能を実現することができた。



Fig.1 ホールA内観



Fig.2 音響パネル正面

3-P-14

3-P-14 木造床及び乾式二重床上に施工した 4.5 畳の防振量による床衝撃音対策 Floor impact sound reduction by tatami considering vibration isolation of four and a half mats installed on wooden floor structure and dry double floor.

○富田隆太(日大・理工), △奥山雄太(大林組/元日大院・理工),
△斗米真生(日大院・理工), △阿部今日子(日大・芸術)

- ◆本稿では既報に続く検討として、木造床及びその上に施工した乾式二重床に対する防振量の床衝撃音対策効果について実験的検討を行った。
- ◆その結果、バングマシンによる防振量の効果としては、重量床衝撃音で重要な周波数帯域である63 Hz 帯域では、木造床に比べて乾式二重床に防振量を施工することで、防振量の効果が大きくなることがわかった。

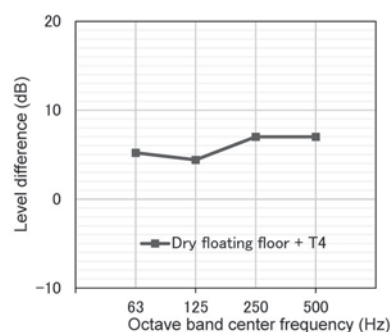


Fig.1: Floor impact sound level difference of vibration-proofed tatami mats with respect to dry floating floor using bang machine (car-tire)

3-P-27

3-P-27 全日本海員組合本部 「地下大会議室」の改修プロジェクト —モダニズム建築の保存と用途拡大を両立させる音響設計— Renovation Project of the Underground Conference Room at the All Japan Seamen's Union Office and Hall: Acoustic Design Balancing Modernist Architecture Preservation and Functional Expansion

○大木 大夢, 宮崎 秀生(ヤマハ)

- ◆昭和のモダニズム建築を今に伝える「全日本海員組合本部会館」の保存・復元と機能更新を両立する改修工事が実施された。
- ◆地下大会議室の音響設計では、建築的意匠を保存しながら、遮音性能や残響特性、電気音響設備の性能向上を図り、講演会から音楽利用まで柔軟に対応できる多目的ホールとして再構築した。
- ◆本発表では、保存・再生を前提とした音響設計の検討内容および改修後の音響測定結果を報告し、音響設計による空間性能の向上と価値再定義の可能性について考察する。

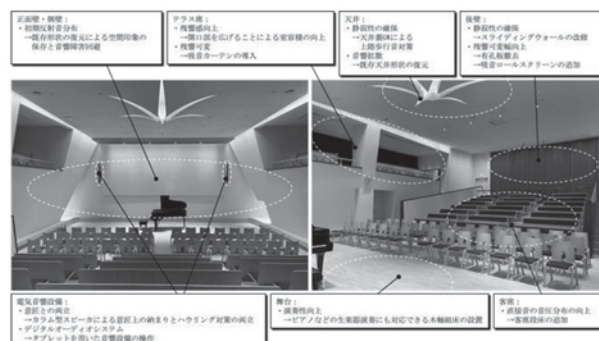


Fig.1: Overview of Acoustic Design

3-P-28

3-P-28 あなぶきアリーナ香川(香川県立アリーナ)の音響(その1)ー大規模アリーナの音響設計と音響性能ー

Acoustics of Anabuki Arena Kagawa (Kagawa Prefectural Arena)

(Part 1)-Acoustical design and acoustics of the large multi-purpose arena-

○清水寧(S/F Design Lab.)、日根野翔太、高橋顕吾(ヤマハ)、△山本力矢(SANAA)

- ◆あなぶきアリーナ香川(香川県立アリーナ)は高松市に県の施設として建設されたスポーツ、軽音楽イベント、展示会、などを行う多目的アリーナであり、メインアリーナ、サブアリーナ、武道施設から構成されている。騒音制御、遮音計画、室内音響計画を行った。
- ◆プロスポーツ、軽音楽イベントにも使用されるメインアリーナは大空間で生じやすい明瞭度の低下、反射音の障害、などの音響的な問題が生じないよう検討を行った。
- ◆競技エリアとその外側の外部に面した交流エリアの両空間は音響コンセプト(Fig.1)で示すように壁で区画されない一体の空間とし、適度な残響と反射音による音響障害の低減を意図した。
- ◆基本モデルによる幾何音響シミュレーション結果は競技エリアに直接戻る反射音の経路が少なくなり、時間遅れが大きい反射音も減少している。(Fig.2)
- ◆竣工後の測定結果から反射音による障害も低減し、目標性能を満たす結果が得られた。

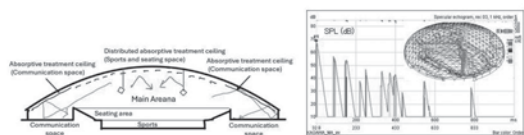


Fig.1: Schematic diagram of acoustical design on the main arena.

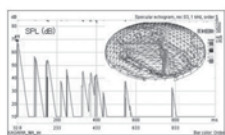


Fig.2: Modeling with geometrical acoustics method

3-P-29

3-P-29 あなぶきアリーナ香川(香川県立アリーナ)の音響(その2)ー明瞭性の評価と遮音性能の確保に向けた音響検討ー

Acoustics of Anabuki Arena Kagawa (part 2)

-Acoustic considerations to evaluate clarity and ensure sound insulation-

○日根野翔太、高橋顕吾(ヤマハ)、清水寧(S/F Design Lab.)、

越智栄三郎(大林組)、山本力矢(SANAA)

- ◆大規模なイベント空間で問題となりやすい明瞭性の評価と、遮音性能確保に向けた音響検討について報告
- ◆明瞭性は、響きの長さやエコーに着目し評価を行った。
- ◆パルス音源に対する残響感、T30が示す値よりも短く知覚されており、明瞭性を阻害するような響きではない可能性が示唆された(Fig.1)。
- ◆エコーによる妨害感が生じにくく、明瞭性は良好に確保されているものと推察される(Fig.2)。
- ◆遮音性は、外壁内側及び通路に遮音カーテンを設置し、遮音性向上を図った。遮音カーテンについては、モックアップによって性能確認を行い、採用するものを決定した。
- ◆所定の遮音性能が確保されており、遮音に関して問題は生じてない。

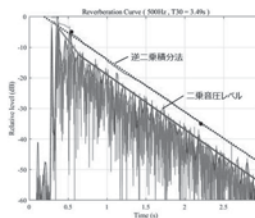


Fig.1: Energy decay characteristics

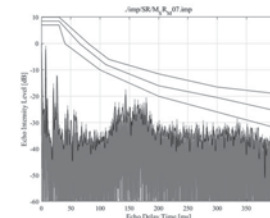


Fig.2: % Echo Disturbance Curve

3-P-30

3-P-30 福岡市民ホールの音響設計ー初期反射音制御を基軸とした馬蹄型大規模多目的ホールの設計手法ー

Acoustic design of Fukuoka Civic Hall - Design method for a horseshoe-shaped large-scale multipurpose hall based on early reflection control -

○日根野翔太、宮崎秀生(ヤマハ)、岸永伸二

- ◆本施設は、福岡市民会館大ホールの機能を継承しつつ、より大規模な興行にも対応可能な大ホール(固定席:1,994席)を中心に、中ホール(固定席:799席)、小ホール(可動席:150席)、リハーサル室、練習室から構成されている。
- ◆大ホールはクラシックやオペラからポップス、演劇まで対応する多目的ホールとして計画された。
- ◆生演奏においては、大空間特有の音響上の課題が生じたため、初期反射音制御を基軸とした音響検討を実施した(馬蹄型・反射面最適化)。
- ◆拡声用途においては、電気音響設備の最適化(スピーカー配置・最適調整)により、明瞭性を確保した。
- ◆これらにより、大規模空間でありながらも、建築的課題(一体感や視認性)が解消され、生演奏用途では優れた音響性能を、拡声用途では十分な明瞭性を備えた多目的ホールの実現を目指した。



Fig.1: Acoustical Concept

3-P-31

3-P-31 声楽家のホール音場評価における後期反射音の効果

ー演奏経験の違いに着目した実験的検討ー

The effect of late reflections on vocalists' evaluation for concert hall acoustic

-An experimental study focusing on differences of performance experience-

☆堀口万葉、松尾綾子、上野佳奈子(明治大)、井上瑞紀、青木亜美(日建設計)

- ◆多目的中規模ホールのステージ周囲の反射面による初期反射音及び客席後壁からの後期反射音が演奏者の音場評価に与える効果を検討するため、三次元音場シミュレータを用いた演奏実験を行った。
- ◆演奏しやすさの評価には個人差があったが、演奏経験に基づいて演奏者を群分けすると、結果に共通の傾向がみられた。
- ◆演奏経験の豊富なプロは後期反射音が聞こえにくい条件でもそれを知覚し、「客席に音が届く」感覚を得ていたが、経験の浅い声楽者は後期反射音がより明瞭に聞こえる条件で同様の感覚を得ていた。
- ◆後期反射音の検出感度が演奏経験により異なり、それが音場評価に影響を与えている可能性が示唆された。

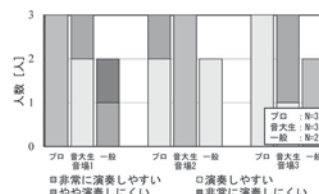


Fig.1: Evaluation of ease of performance

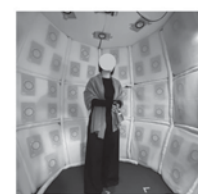


Fig.2: Performance experiment

3-P-32

2-P-32 パノラマ写真によりVR表示された建築空間の視覚情報に調和する残響時間の提案

Proposal of reverberation time to match the visual aspect of architectural space displayed in VR using panoramic photographs

☆千葉夢, 石川あゆみ(岐阜高専)

- ◆我々の研究室では、人が建築空間の視覚刺激(見た目)から予想した響きを“予想残響時間(以下 ERT)”と定義し、これを被験者実験で測定してきた。
- ◆先行研究では、パノラマ写真によりVR表示された建築空間の視覚刺激に調和する残響時間を求めるモデルとして、実験における対象空間の容積とそのERTの散布図が提案された。
- ◆本報では、このモデルの精度向上を目的に、先行研究のデータにオープンスペース(主の空間に他の1つ以上の空間が接続した空間)含む14空間のデータを加え、空間構造および空間の既知・未知が視覚印象およびERTに与える影響を明らかにする。
- ◆オープンスペースと閉空間の比較をした結果、閉空間よりもオープンスペースのERTの評価平均値の方が長いこと、「広さ」の評価平均値が閉空間よりオープンスペースの方が広く感じられやすいということが分かった。
- ◆空間の既知・未知の比較をした結果、未知の空間より既知の空間の方がERTの評価平均値が長いことが分かった。「広さ」「天井高さ」の評価平均値は概ね実際の容積・天井高に対応し、これらの視覚印象に空間の既知・未知が与える影響は少ないと考えられる。
- ◆対象空間を増やした結果、空間構造と空間の既知・未知が、ERTと視覚印象に影響を及ぼす可能性が示唆された。今後はこの影響を加味してモデルを提案する。

3-P-33

スパース性に基づく

特徴抽出を用いた聴感印象の識別

Auditory Impression Classification Based on Sparsity-driven Feature Extraction

☆村上寛名, 石川裕子, 石光俊介(広島市大), △江野峰弘, 藤本祐樹

- ◆本研究では環境の異なる4つの空間を対象とし、LASSO回帰によって3種類の次元数で評価の自動化と特徴量抽出を行った。「迫力感」の有無で「良い」「悪い」に分類した。
- ◆84次元の検証では、周波数特性の寄与が低くC80(明瞭度)とT60(残響時間)の寄与が高い傾向。
- ◆116次元の検証では、84次元同様に周波数特性の寄与が低く、時間的な変化に関連する特徴量が識別に有効であることが示唆。
- ◆30次元の検証では、84、116次元と比較すると、平均識別精度はやや低かったものの、116次元で抽出された特徴量は全て抽出され、さらに、時間変動に関する特徴量が1つ多く抽出された。

	Left	Right
RMS	●	●
Max Amplitude	●	●
Spectral Centroid		●
Spectral Slope	●	●
Spectral Temporal Variation	●	●
RMS 500 Hz	●	●
1000 Hz	●	
2000 Hz	●	●
4000 Hz	●	
8000 Hz	●	●

Table 1 Selected features from stimulus in the 116-dimensional set

3-P-44

3-P-44 音声避難誘導時における避難者の挙動解析システムの検証

Investigation of an Evacuee Behavior Analysis System for Voice Evacuation Guidance

☆一色暁光(日大院・理工), 伊藤洋一, 大隅歩(日大・理工)

- ◆音声を用いた避難誘導システムの研究を行っている。これまでは避難者の頭部運動が方向感特性に与える影響について十分に検討されていなかった。
- ◆本検討では、音声避難誘導システムでの避難者の頭部運動を取得するシステムをVIVE Focus 3, Unityを用いて構築し、仮想環境下で挙動を解析するシステムの有効性を検証した。
- ◆結果より、本システムが頭部運動を的確に捉え、音声避難誘導時の避難者挙動評価における有用性を確認した。

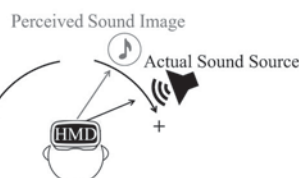
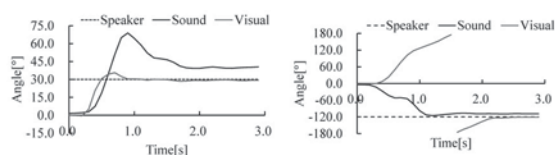


Fig.1: Evacuee Behavior Analysis System.



(a) Target Direction 30°. (b) Target Direction -120°. Fig.2: Head Movement under Auditory and Visual Guidance.

3-P-45

3-P-45 複合現実技術を用いた室形状情報に基づいたシミュレーション音場の可視化

Visualization of simulated sound fields based on room shape information using mixed reality technology

☆大沼翔矢, 内田彩芽, 池田雄介(東京電機大), 井上敦登(INSPIRED), 及川靖広(早稲田理工)

- ◆背景
 - 音場の可視化は音伝播の正確な理解と分析に役立つ。
 - これまでMRデバイスであるMicrosoft HoloLens 2やApple Vision Proを用いた音場の可視化システムを提案してきた。
- ◆提案手法
 - Vision ProのSLAM技術により取得した室形状情報に基づき、数値シミュレーション(BEM)を行い、その結果をVision Proを用いて実空間に重畳して可視化する。
- ◆結果
 - シミュレーション結果を実空間に表示させることができた。



Fig.1: Visualization results with Apple Vision Pro.(250Hz) (Left: direct sound, Right: reflected sounds)

3-P-46

3-P-46 廊下および階段室の音環境に関する有限要素解析

Finite element analysis of the sound environment in corridors and stairwells

☆渡部奈津実(大分大院), 富来礼次, 岡本則子, 大鶴徹(大分大)

- ◆廊下および階段室を対象に時間領域有限要素解析を行い、解析精度を検証するためインパルス応答を測定の実測値と比較するとともに、吸音材設置時の音環境改善効果を解析した。
- ◆仕器の影響により、解析と実測の減衰曲線に乖離が見られたが、仕器を撤去して再測定をした結果、Fig. 1に示すように4階および3階では非常に一致した。
- ◆廊下の天井面に吸音処理を施した場合の解析を行った結果、すべての階で1[s]程度残響時間が短くなり、音源付近の明瞭度が上昇した。

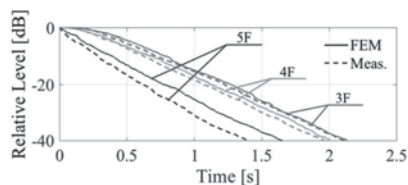
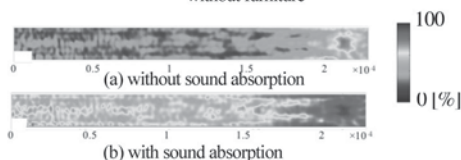


Fig. 1 Comparison of decay curves between TDFEM and measurements without furniture

Fig. 2 Distributions of D_{50} obtained by TDFEM

3-P-48

3-P-48 微細気泡 HIPE フォームの吸音解析モデルに関する基礎的検討

Fundamental Study on acoustic absorption analysis model of Microcellular HIPE Foams

☆竹村海渡(近畿大院), 菅原彬子, 平栗靖浩(近畿大),
△中村隼, 原口健二(JSP)

- ◆微細気泡 HIPE フォーム (HIPE-F) の吸音特性を実験により確認した。
- ◆実験結果から、気泡径の小さい HIPE-F のほうが低周波数帯域での吸音性能が優れていることが確認できた。
- ◆数値解析モデルに用いる単位格の大きさは、断面写真中の気泡径の平均値 (Mean) よりも近似式を用いて真の気泡径を推定した値を用いるほうが実験結果により近いことが明らかになった。



Fig. 1: Unit-cell model

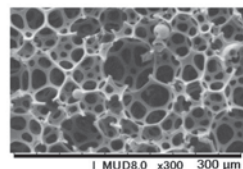


Fig. 2: SEM images

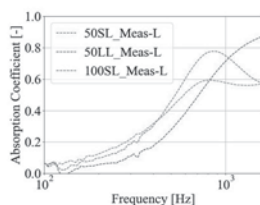


Fig. 3: Experimental results

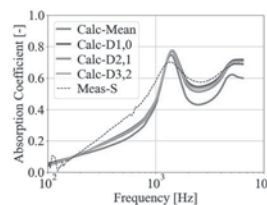


Fig. 4: Numerical analysis (50SL)

3-P-47

3-P-47 エLEMENTフリーガラキン法による室内音響数値解析

-その2 パラメータ設定の影響に関する数値実験-
Application of the Element-Free Galerkin Method to
Room Acoustic Simulation- Part 2: Numerical Experiments to Examine the Effects of Parameter
Settings -

☆佐々岡潤, 川井敬二(熊本大)

- ◆本研究は、「その1」で室内音響分野への適用可能性を確認したエレメントフリーガラキン法 (EFGM) について、パラメータ設定が計算精度と効率に与える影響明らかにすることを目的とする。
- ◆影響半径・バックグラウンドセルの分割数・修正積分則の適用の3点について数値実験を行い、解析解および有限要素法 (FEM) の結果と比較検討した。
- ◆影響半径: 影響半径の設定は近似関数の性質を決定するため、計算精度に大きく関与する。影響半径が過小または過大になると計算精度が低下するため、節点間隔を考慮した最適値を設定することが不可欠である。本実験では、節点間隔 0.010 m に対し、影響半径 0.015 m が最も良好な結果を示した。
- ◆バックグラウンドセル: バックグラウンドセルの分割数を減らすことで、計算精度を維持したまま全体行列の構築時間を大幅に短縮できることが確認された(各軸方向 100 分割で約 180 分に対し、50 分割では約 24 分)。
- ◆修正積分則: FEM では精度の向上が見られたが、EFGM では明確な精度の向上は確認できなかった。これは、形状関数の構築の方法や誤差特性が FEM と異なることに起因すると考えられる。

3-P-49

3-P-49 反射音を極限まで排除した無響室

Anechoic chamber with minimal reflection

○星野嗣人, 原健登, 矢入幹記(鹿島技研),
津金孝光(元 日本音響エンジニアリング)

△石垣充, 松尾浩義, △山田哲(日本音響エンジニアリング)

- ◆鹿島建設(株) 技術研究所にある無響室を改修した。
- ◆本改修では、反射音を極限まで排除することを目指した。
- ◆下記反射音を除去する工夫を施した。
 1. 測定対象物の固定・自重の保持は、ワイヤー等を用いて張力で保持する吊構造を採用。
 2. 反射物となる歩行面の床格子、天井照明はすべて撤去可能。
 3. 歩行面を支える床束は電動昇降式で格納できる構造。
- ◆測定結果から、反射物の除去によってインパルス応答後半のレベルが明確に小さくなることを確認した。

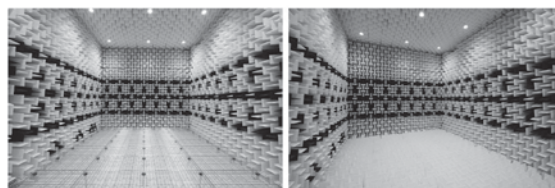


Fig. 1 An anechoic chamber in Kajima Technical Research Institute.

Left: With removable mesh floor panels.

Right: Without mesh floor panels.

3-P-50

3-P-50 残響成分を考慮した連続音場測定法による
明瞭性評価の試みAn attempt to Evaluate Intelligibility by Continuous Sound Field Measurement
Method Considering Reverberation Components○佐々木 渉吾, 立花 杜斗, 山名 一輝, 前田 和昭 (TOA株式会社),
△小林 由佳, 河原 一彦 (九州大・芸工)

- ◆網羅的な音響測定をするための M 系列信号を用いた連続音場測定法を提案しており、これまで精度検証と数値評価を行ってきた。
- ◆この測定法を用いた SII 評価では、明瞭性に悪影響を与えるはずの残響成分が有効な音声として扱われ、過大評価されてしまうことが課題として挙げられる。
- ◆本研究では、この課題を解決するために残響成分を考慮した数値評価をするため、アンビソニックマイクを用いて測定と数値評価を行った。
- ◆測定結果から全指向性成分と直接音による SII、MTF による補正をした SII で比較をした結果、残響成分を考慮した数値評価ができた一方で、直接音に基づく SII と MTF による補正の SII に差異が生じた。

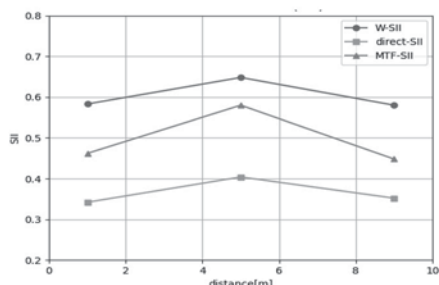


Fig. 1: SII and MTF at 3 points in the measurement path (2m)

3-Q-2

3-Q-2 Investigating Feature Characteristics of
Pseudo-Speaker Data
for Speaker Verification

☆Hengyi Zou, Sayaka Shiota (Tokyo Metropolitan University)

- ◆The performance of speaker verification systems is highly dependent on the diversity and volume of training data.
- ◆While pseudo-speaker augmentation via Vocal Tract Length Perturbation (VTLP) is widely adopted, its embedding-level characteristics remain insufficiently explored.
- ◆This study analyzes the distribution and separability of VTLP-generated pseudo-speaker embeddings to assess their impact on model robustness.
- ◆A selection and regeneration framework (as shown in figure 1) based on cosine similarity is introduced to ensure high-variability yet realistic augmentation.
- ◆Experimental results on the VoxCeleb1 dataset demonstrate that the proposed method enhances speaker verification performance and improves embedding space separability.

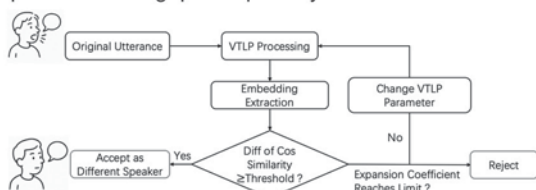


Figure 1: Proposed pseudo-speaker generation process

3-Q-1

Room Simulatorを用いたデータ拡張
によるNeural Speaker Diarization
モデルの実環境適応Adapting Neural Speaker Diarization Models to Real-world
Environments using Data Augmentation with a Room Simulator

○石塚賢吉, 大野正樹, 橋本泰一 (株式会社RevComm)

本研究では、Neural Speaker Diarization (SD) モデルを特定環境に適応させるための学習データのアノテーションコストの削減を目的とし、Room Simulatorを用いたデータ拡張手法を提案する。

事前学習済みのNeural SDモデルは、学習データに近い環境で録音された音声では高い精度を示すものの、音響特性や話者の特性が異なる音声の場合には精度が低下することがある。こうした状況においては、ターゲットとする実環境の音声データを用いてSDモデルをファインチューンすることで、精度向上が期待できる。Neural SD モデルをファインチューンするためには、「いつ、誰が話しているか」の情報がアノテーションされた学習データが大量にあることが望ましい。しかし、実環境で録音された対話音声に対して人手でこのアノテーションを行う作業には、多大な時間とコストがかかる。

そこで本研究では、Room Simulatorを用いて、ターゲット環境を想定した音響特性を持つ学習データを人工的かつ大規模に生成する。Room Simulationのソース音源としては、話者ごとに音声チャンネルが分離された、ビデオ会議などの実際の対話音声を使用する。

また本研究では、対面会議をスマートフォンで録音する状況をターゲット環境として定め、提案手法により生成したデータセットを用いて事前学習済みのSDモデルをファインチューンした。そして、人手でアノテーションした実環境録音データセットと、複数のスピーカー装置とスマートフォンで物理的に環境を再現したシミュレーションデータセットを用いて、得られたモデルの精度を評価し、提案手法が有効であることを確認した。

3-Q-3

3-Q-3 複素 CQT を用いたなりすまし音声の識別と
感情依存型話者照合の併用A Combined Approach of Emotion-Dependent Speaker Verification
and Spoofed Speech Detection using Complex CQT☆PHAM HOANG LONG, 西田昌史 (静岡大), 柘植寛 (大同大),
黒岩真吾 (千葉大)

話者照合は低コストかつ柔軟な生体認証技術として広く用いられているが、音声合成 (TTS) や声質変換 (VC) によるなりすまし音声はシステムの信頼性に脅威を与えている。本研究では、従来研究の感情依存型話者照合システムのなりすまし音声に対する頑健性を向上させるため、複素 CQT を用いたなりすまし音声検出モデルを統合した直列型および並列型システムを提案する。

本稿では、声質変換ツール seedVC を用いて、Japanese Twitter-based Emotional Speech corpus (JTES) よりなりすまし音声を生成し、評価した。なりすまし音声検出モデルの評価実験では、Table 1 より、正規化複素 CQT を特徴量として用いた手法の平均等誤り率 EER が 2.3% を示し、位相情報の活用による高い検出性能を確認できた。また、話者照合実験では Table 2 より並列型システムが EER 1.7% を達成し、従来の話者照合システム、感情依存型話者照合システム、および直列型システムの精度を上回った。

Table 1 EER of each spoofed speech detection model [%]

Model	EER
LFCC - CNN	18.1
MFCC - CNN	15.8
Complex CQT - Complex CNN	4.4
Normalized Complex CQT - Complex CNN	2.3

Table 2 EER of each speaker verification system [%]

	Anger	Joy	Sadness	Neutral	Average
SV System	27.9	28.5	19.1	25.1	25.1
Emotion-dependent SV System	18.7	19.9	14.2	17.9	17.7
Serial System	3.0	2.4	1.8	2.1	2.3
Parallel System	2.0	2.1	1.3	1.5	1.7

3-Q-4

3-Q-4 なりすまし音声検出に対する録音再生攻撃の収録条件に関する影響分析

Impact of recording conditions in replay attacks on spoofed speech detection

☆堀江涼花(都立大), 高道慎之介(慶応義塾大/東大), 塩田さやか(都立大)

- ◆近年、情報セキュリティへの関心が高まる中、生体認証技術が注目されている。声を使った生体認証技術である話者照合における課題としてなりすまし音声攻撃がある。本研究では、なりすまし音声攻撃の対策を目的として、検出が困難ななりすまし音声の作成を目指した。
- ◆図1に録音再生攻撃を行う収録状況を示す。再生機器としてBose、iPad、Mac、Sonyの4種類を使用し、攻撃収録環境として距離(遠距離または近距離)と音量(音量大または音量小)の組み合わせによる4条件で収録を行った。
- ◆作成したなりすまし音声になりすまし音声検出及び話者照合の性能に与える影響について分析した結果、図2のように攻撃収録環境や再生機器、使用したモデルが結果に影響を与えることが示された。

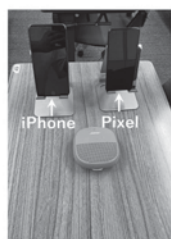


図1:収録の様子

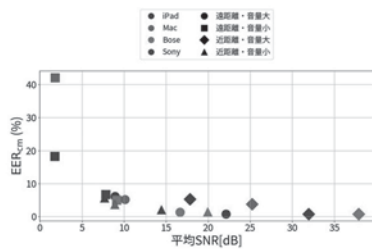


図2:攻撃収録環境ごとの平均SNRとEER_{GM}

3-Q-5

3-Q-5 録音再生攻撃環境を考慮したマルチタスク学習におけるなりすまし音声検出の検証

Investigation of Multi-Task Learning for Environment-Aware Replay Spoofing Detection

☆菅野滉大, 塩田さやか(東京都立大学)

- ◆他人の声を録音・再生する「なりすまし音声攻撃」が、音声認証の大きな脅威となっている
- ◆このような攻撃を検出する技術がなりすまし音声検出
- ◆録音再生の過程で含まれる微細なノイズのみが検出の手がかりとなる難しいタスクであり、検出精度や頑健性が課題となっている
- ◆そこで本研究では、マルチタスク学習の手法に着目し、各タスクの重み調整が、検出精度や頑健性に与える影響を調査
- ◆「なりすまし検出」に加え、「再生機器」と「再生距離」の分類をサブタスクとして同時に学習 (Fig.1)
- ◆マルチタスク学習を適用したモデルの埋め込みベクトルを可視化すると、追加したサブタスクのラベルに応じてクラスターを形成 (Fig.2)

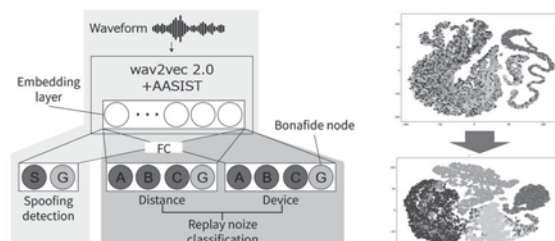


Fig.1:Model Architecture for Multi-task

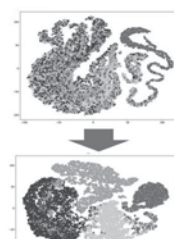


Fig.2:t-SNE

3-Q-6

3-Q-6 Detecting Multilingual Audio Deepfakes: A Robust and Generalizable Hybrid Model Approach

Candy Olivia Mawalim, Wang Yutong, ☆ Aulia Adila, Shogo Okada, Masashi Unoki (JAIST)

- ◆**Aim:** To investigate robust multilingual audio deepfake detection for cross-lingual generalization.
- ◆**Problem:** Highly realistic machine-generated voice necessitates robust deepfake detection systems capable of generalizing across diverse linguistic environments, and synthesis technique.
- ◆**Solution:** Developing hybrid detection model with end-to-end models (i.e., RawNet, AASIST) and the language-agnostic representations learned from multilingual self-supervised learning utilizing a multilingual and multisource corpus (JMAD dataset).
- ◆**Evaluation:** Based on the SAFE Challenge (completely unknown evaluation data), using balanced accuracy as the primary metric.
- ◆**Summary:** Our model showed high generalization (approx. 73% of accuracy) for unseen multilingual and synthesis algorithm.



Fig. 1: Our proposed hybrid model, involving two training stages.

3-Q-7

3-Q-7 単一英語話者による発話内の同一単語に着目したDeepFake 音声検出に向けた一検討

A Study on DeepFake Voice Detection Focusing on Identical Words in Speeches by a Single English Speaker

☆根本結菜(日本工業大院), 大田健紘(日本工業大・基幹工)

- ◆近年、音声合成や音声変換技術の発展より、DeepFake 音声現実的な脅威となっている。本研究では、ASVspoof 2021 のDeepFake タスクに含まれる音声データを用い、単一話者による同一単語に着目し、本物音声と偽物音声の識別に有効な音響特徴量を検討した。
- ◆音響特徴量として第1～第4 フォルマント周波数(F1～F4)および基本周波数(f0)を抽出し、単語単位で平均化した5次元の音響特徴ベクトルを求めた。これをもとにランダムフォレストによる分類を行い、特徴量の重要度を算出し、識別に有効と考えられる特徴量を調査した。
- ◆各音響特徴量の重要度を Table 1 に示す。F1 および F2 の重要度が高かった。これにより、母音の弁別に深く関与するフォルマント周波数がDeepFake 音声の識別において有効であることが示唆された。

rank	Features[Hz]	Importance [%]
1	F2	26.6685
2	F1	24.2456
3	f0	23.1383
4	F3	15.4763
5	F4	10.4713

使用した音声データのリンクを以下に示す。

ASVspoof2021 Challenge Speech Deepfake Database (URL: <https://www.asvspoof2021.com/>)

3-Q-8

3-Q-8 なりすまし音声検出に対する
話者適応を用いた音声合成攻撃Attacks Based on Speaker Adaptation-Based Speech Synthesis
Against Spoofing Speech Detection Systems☆古林嵯羽仁(東京都立大), 高道慎之介(慶応義塾大/東京大),
塩田さやか(東京都立大)

- ◆ 深層学習を用いたディープフェイク音声が悪用される事例が課題となっている
- ◆ 対策手法であるなりすまし音声検出や話者照合に関する研究では、実環境での攻撃者の立場を考慮した合成音声作成手法が少ない
- ◆ 本研究では、不正録音音声から目標話者を高精度に模倣する音声を作成することを目的として、日常的な環境での不正録音音声を収録した音声データセットである J-SpAW を用いて話者適応を含めたテキスト音声合成を行いなりすまし音声の作成を行った (Fig. 1)
- ◆ 一部条件下の合成音声ではなりすまし音声攻撃の成功可能性が先行研究と比べて上昇した (Table. 1)
- ◆ 不正録音音声においても、追加学習を行うことで目標話者へ話者性を近づけることが可能であることを報告する

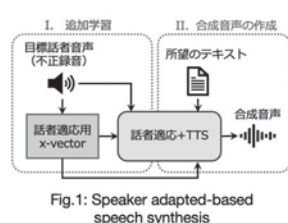


Table. 1: EERs of spoofing speech detection and UTMOS scores

環境	EER _{cm} (%) ↓		UTMOS ↑	
	なし	あり	なし	あり
All	2.50	2.13	2.52	1.49
E1+E4	2.56	2.13	2.69	1.73
E1	3.00	6.56	2.52	1.63
E4	2.13	2.50	2.47	1.58

講演取消

3-Q-10

3-Q-10 音声伝える感情を色彩情報で表現した
ラベルの収集とその分析

Collection and analysis of speech emotion labels represented by color

◎永瀬亮太郎, 高島遼一, 山下洋一(立命館大)

- ◆ 音声伝える感情を扱う研究
 - 代表的な感情表現であるカテゴリ感情と次元感情
 - ✓ 感情の強弱や繊細な違いなどを人間が直感的に理解可能でかつ定量的に示すことは困難
- ◆ 感情を色相と彩度、明度で表現したラベルを収集
 - 曖昧な感情を人間が視覚的に解釈可能になることを期待
 - クラウドソーシングを用いて色彩情報を既存データセットに付与
 - ✓ 色相と彩度・明度を分けて数段階で評価
- ◆ カテゴリ感情と色彩情報との関連性を分析
 - 色相は感情の種類、彩度は感情の覚醒度合い、明度は感情のポジティブ度合いと関連がある可能性有
- ◆ 音声から色彩情報を推定する回帰モデルを学習・評価
 - 各色彩情報を用いて Support vector regression モデルを学習
 - (明度について) 正解に対し中程度の相関がある予測が可能

Table 1: Regression results for each color attributes

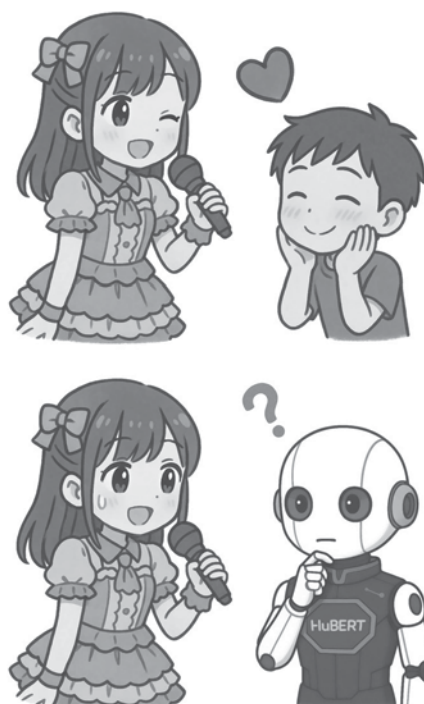
Fold	Hue	Saturation			Value		
	AE ↓	LCC ↑	SRCC ↑	CCC ↑	LCC ↑	SRCC ↑	CCC ↑
1	93.5	0.195	0.166	0.125	0.393	0.380	0.345
2	58.3	0.226	0.042	0.125	0.535	0.572	0.434
3	73.6	0.104	0.165	0.062	0.724	0.703	0.674
4	61.9	0.459	0.474	0.392	0.630	0.630	0.469
Ave.	71.8	0.246	0.212	0.176	0.571	0.571	0.481

3-Q-11

3-Q-11 大規模音声自己教師あり学習モデルに
もとづく音声好感度予測Voice likability prediction using
large-scale self-supervised speech models

◎須田仁志(産総研)、高道慎之介(慶應大/産総研)、深山覚(産総研)

SSLモデルで予測できるか？



人が感じる声の好感度は

3-Q-12

3-Q-12 音響ベクトル時系列の自己注意重み特徴量を用いた吃音検出モデルのゼロショット・クロスリンガル評価

Evaluating Zero-Shot Cross-Lingual Stuttering Detection Based on Self-Attention Weights of Temporal Acoustic Vector Sequence

☆宮原絃造, 加藤恒夫, 田村晃裕 (同志社大院)

- ◆ 吃音の症状は、音素や単語の反復、音の延長、発話途中の長いポーズ、フィラーの挿入として現れる。特定の音素に現れるなどの音響的特徴よりも、時間的な構造に特徴がある。
- ◆ 従来の吃音検出モデルは数十時間の吃音音声データセットを学習に用いていたが、それらの手法は日本語などの訓練用の吃音音声データセットが公開されていない言語には適用できない。
- ◆ 著者らが提案した自己注意重み特徴量は、音響的特徴よりも時間的特徴を効率的に捉え、音響的特徴の異なる言語間でモデルを転用できる。
- ◆ 英語と中国語で学習し、ドイツ語で評価するゼロショット・クロスリンガル評価で、全ての吃音クラスでF1スコアが40%以上となった。

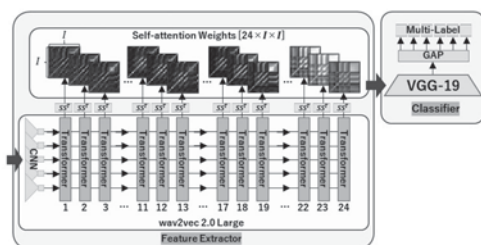


Fig.1: Structure of stuttering detection model using self-attention weight features of wav2vec2.0

3-Q-14

3-Q-14 Enhancing Speech Synchronization in Automatic Dubbing through Diverse Translation Phrasing

○ Jan Meyer Saragih, Faisal Mehmood, Sakriani Sakti

- ◆ We have explored a speech synchronization method for the improvement of dubbing tasks. Through machine translation model, multiple translation candidates were generated to maximize the synchronization between source and target speech.

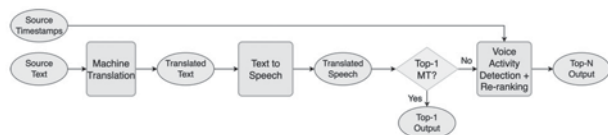


Figure 1: Proposed Framework

- ◆ The results showed that the proposed approach managed to improve the dubbing task through speech length compliance, while showing minimal drop in translation quality.
- ◆ The drop in translation quality was observed more on lexical-based metric i.e., BLEU compared to semantic-based metric i.e., XCOMET.

Table 1: Comparison of Traditional and Proposed Dubbing Approaches with Multiple Translations

System	Dataset	BLEU	XCOMET	SLC
Baseline	FLEURS	14.16	0.531	0.750
	IWSLT S.1	28.85	0.830	0.841
	IWSLT S.2	19.20	0.605	0.847
Top-1 MT	FLEURS	42.31	0.965	0.778
	IWSLT S.1	50.92	0.985	0.678
	IWSLT S.2	50.78	0.973	0.706
Top-5 MT	FLEURS	37.32	0.963	0.805
	IWSLT S.1	39.56	0.982	0.736
	IWSLT S.2	47.48	0.969	0.749

3-Q-13

3-Q-13 日本語音声認識における補助情報の認識方式の比較

Comparison of Recognition Formats for Japanese Speech Recognition with Auxiliary Information

小堀聡太, 藤江真也(千葉工大)

- ◆ 補助情報を検出する際の最適な認識方式について比較

➢ 補助情報にはフィラー: F, 言い直し: D, 感動詞: I を選出

- ◆ 発音形複合トークン方式 (提案手法)

➢ 補助情報ではない同音語に対して別のトークンとして認識することが可能

◇ 例: ウ+| ン+| エ+F ソ+D ソ ウ ダネ

- ◆ マルチタスク方式 (Fig. 1)

➢ 「トークン」と「各トークンの補助情報ラベル」(流暢: 0, フィラー: 1, 言い直し: 2, 感動詞: 3) を同時予測することが可能

◇ 例: [出カトークン]: ウン エー ソ ソ ウ ダネ

[出カラベル]: 3 3 1 1 2 0 0 0 0

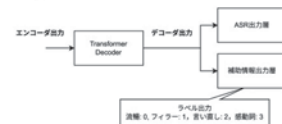


Fig. 1 Multi Task Model overview

- ◆ 提案手法がすべての補助情報に対して高いF1を達成 (Table.1)

Table.1 Experimental Results

補助情報	方式	音字形			発音形		
		適合率	再現率	F1 値	適合率	再現率	F1 値
フィラー	マルチタスク方式	50.73	33.52	40.37	54.44	43.14	48.13
	複合トークン方式	81.08	77.65	79.33	83.19	80.92	82.04
感動詞	マルチタスク方式	86.26	76.91	81.31	87.29	75.15	80.76
	複合トークン方式	86.85	85.92	86.39	87.67	86.66	87.16
言い直し	マルチタスク方式	20.80	8.80	12.37	23.76	7.76	11.70
	複合トークン方式	40.49	36.17	38.20	42.72	36.97	39.64

3-Q-15

3-Q-15 音声翻訳と機械翻訳の統合デコーディングに基づく半教師あり学習を用いた end-to-end 音声翻訳の検討

Semi-Supervised End-to-End Speech-to-Text Translation with Joint Machine Translation and Speech Translation Decoding

○田中 智大, 増村 亮, 庵 愛, 牧島 直輝, 河田 尚孝, 折橋 翔太, 鈴木 聡志, 山根 大河(NTT)

- ◆ 音声から翻訳後のテキストを直接生成する end-to-end 音声翻訳が注目を集めているが、音声認識や機械翻訳と比べデータの収集が困難
- ◆ データ不足の問題を解決するために、end-to-end 音声翻訳モデルを用いてラベルなしデータに擬似ラベリング付与する手法が存在
- ◆ しかし、少量データで学習された end-to-end 音声翻訳は、過学習により出現頻度の小さいトークンに対する精度が低下する課題がある
- ◆ 本稿では音声認識と機械翻訳のカスケードシステムと、end-to-end 音声翻訳の両者を用いて擬似ラベリングを行う手法を提案
- ◆ En→Ja, Ja→En のデータセットを用いてそれぞれ実験を行い、音声認識またはカスケードシステムを単体で用いて擬似ラベリングを行う場合と比べて高精度な翻訳が可能であることを確認

Method	Joint decoding in inference	En→Ja		Ja→En	
		dev	test	dev	test
ASR-T2TT	-	9.20	11.73	18.54	18.96
Initial S2TT	✓	7.99	9.32	18.38	19.32
		10.02	12.00	22.39	22.36
S2TT-PL	✓	8.11	10.40	21.54	22.15
		9.46	11.86	22.62	23.26
ASR-T2TT-PL	✓	8.79	11.27	21.98	22.31
		9.03	11.96	23.08	23.43
Proposed PL	✓	9.47	11.51	22.42	22.76
		10.13	12.20	23.33	23.71
Fully supervised	✓	11.16	14.00	27.67	28.28
		11.51	14.43	30.38	30.60

Table1. BLEU scores in different methods

3-Q-16

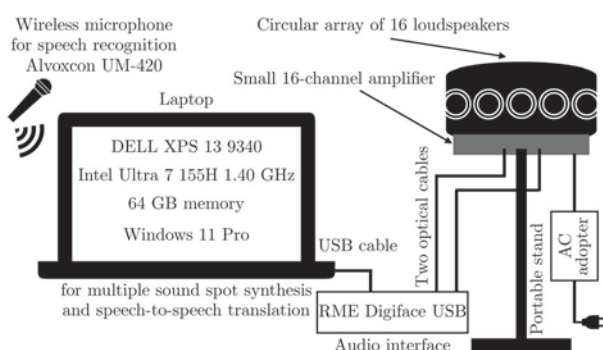
リュックサックで持ち運び可能な 3-Q-16 多言語同時通訳+音声マルチスポット再 生コンパクトデモシステム

Compact demo system integrating multilingual simultaneous speech translation and multiple sound spot synthesis carried out with a backpack

○岡本拓磨, 河野みちよ
(情報通信研究機構)

・ポイント

- ・デモシステム一式をリュックサック1つで持ち運び可能
- ・日→英中韓の多言語同時通訳をノートPC1つで実装



ポスター会場にて実機を用いた 同時通訳実演デモ(日→英中韓)

3-Q-18

3-Q-18 音声のジェンダー印象の差異が声質変換 後の音声特徴量に与える影響の分析

Analysis of the Effect of Gender Impression Differences on Speech Features after Voice Conversion

☆両角祐二, 坂野秀樹, 旭健作(名城大院・理工学研)

- ◆日本語母語聴取者 10 名を対象に、日本語・英語・タイ語の音声に対するジェンダー印象の主観評価実験を実施(男性は「男性性の強さ」、女性は「女性性の強さ」を 5 段階で評価)
- ◆全話者の平均基本周波数に揃うように基本周波数系列に倍率を乗じて変換し、同様の評価尺度で再度主観評価実験を実施
- ◆変換前：基本周波数とジェンダー評価に相関がある傾向が見られた
- ◆変換後：ジェンダー評価のばらつきが小さくなる傾向が見られた
- ◆基本周波数はジェンダー印象に大きく影響するが、聴取者の母語や文化的背景、韻律なども影響する可能性

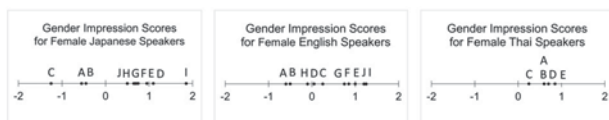


Fig. 1 Gender impression results for female speakers across languages

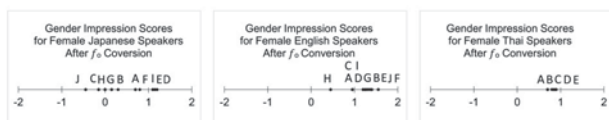


Fig. 2 Gender impression results for female speakers across languages after f_0 conversion

3-Q-17

3-Q-17 目標話者の印象と 音声変換の話者類似度の関係の分析

Analysis of relationship between impressions of target speakers in voice conversion and similarity between converted and target speakers

◎近藤祐斗, 亀岡弘和, 田中宏, 金子卓弘 (NTT 株式会社)

- ◆注目した音声変換の課題：目的話者間の変換性能の不均衡
- ◆本分析：音声変換における目標話者音声の印象と変換音声・目標話者音声間の話者類似度との間にみられる傾向について調査
- ◆本研究の意義の例：話者類似度が低く出る傾向にある印象に関して評価スコアを予測する主観定量化モデルを構築し、学習データ選択に活用して目的話者間の変換性能の不均衡を低減
- ◆結果：様々な表現語における評価と話者類似度との間に中程度の相関

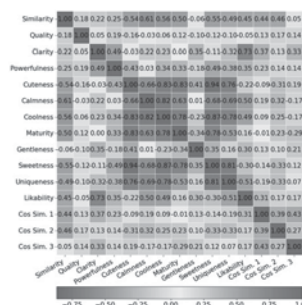


Figure 1 各質問項目の評価スコア、話者類似度の相関係数 Cos Sim. 1, Cos Sim. 2, Cos Sim. 3 はそれぞれ日本語話者 ECAPA-TDNN, VoxCeleb で学習された ECAPA-TDNN, VoiceGrad で用いた話者エンコーダで得た話者ベクトルから計算したコサイン類似度を表す。

3-Q-19

3-Q-19 男女間 DNN 声質変換における変換誤差の 静的・動的特性の分析

An analysis of static and dynamic characteristics of conversion errors in DNN voice conversion between male and female

☆松久直生, 坂野秀樹, 旭健作(名城大院・理工学研)

- ◆本研究では、DNN 声質変換モデルを用いて声質変換をした際の変換誤差の動的特性を分析する。
- ◆変換した音声と目標の音声でどの程度誤差があり、原音声の基本周波数差が変換にどの程度影響を与えるのかを動的特徴量誤差、基本周波数差、メルケプストラム距離の時系列から分析した。
- ◆その結果、変換方向によって品質が低下する要因がわずかに異なる可能性があるものの、明確な傾向は観察できなかった。

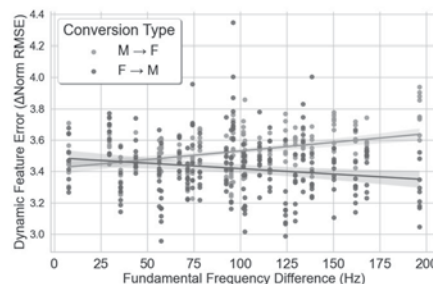


Fig.1: Fundamental frequency difference vs. Dynamic feature error in conversion types of $F \rightarrow M$ and $M \rightarrow F$.

3-Q-20

3-Q-20 自己聴取音声の生成を目的とした
声質変換処理の検討

Voice conversion processing for generating self-hearing voice

☆野々山大樹, 鈴木千文, 西野隆典(名城大)

- ◆ 自己の録音音声聴取すると違和感を覚える。これは、普段の発話時に聴取している音声(自己聴取音声)と録音音声の聴取では、音の伝達経路に違いがあるためである。
- ◆ 録音音声をもとに、声質変換処理にて、自己聴取音声の生成を目指す。本研究では、音声分析合成システム WORLD を用いて3つの処理を行い、主観評価実験による評価を行った。
 1. 基本周波数のシフト
被検者によって録音音声が高く感じられることや、高く感じられる場合があることへの対応。
 2. スペクトル補正
低周波数帯の増幅と高周波数帯の減衰といった、スペクトルの特徴を模倣。
 3. 非周期性指標の補正
自己聴取音声聴取時は、録音時に比べて摩擦音・息音などの非周期性成分が、抑制されて聞こえるといった印象の再現。
- ◆ 生成した音声と録音音声で、発話者毎に主観評価実験(5名)を実施した結果、全被験者に対して、録音音声よりも自己聴取音声に似ている音声を生産することができた。

3-Q-21

3-Q-21 列変換型声質変換モデルの
モバイル端末実装

Implementing Sequence-to-Sequence Voice Conversion Models on Mobile Devices

©山下陽生(1,2), 岡本拓磨(2), 高島遼一(3,2), 大谷大和(2), 滝口哲也(1), 戸田智基(4,2), 河井恒(2)

(1 神戸大学, 2 情報通信研究機構, 3 立命館大学, 4 名古屋大学)

- ◆ 声質変換モデルのモバイル実装には、軽量化と開発効率性が求められる。PyTorch では TorchScript 形式に変換することで、モバイル端末上での推論が可能となる。
- ◆ 本稿では、実装の簡易性を重視し、torch.jit.trace によって系列変換型 VC モデル WSAS-VC をモバイル端末へ導入した。
- ◆ Torch.jit.trace は分岐に非対応だが、修正が少なく既存モデルに適用しやすい。
- ◆ torch.jit.script との比較を表に示し、端末上での実行時間は平均 3 秒であった。

Aspect	torch.jit.script	torch.jit.trace
Control flow support	Supports conditional branches and loops. For example, if and for statements.	Records only the operations used during tracing. Control flow is ignored.
Type annotations	Required. Functions and variables must have explicit type definitions.	Not required.
Support for Python libraries	Not compatible with many standard libraries such as random, os, and numpy.	Such libraries are ignored. Only PyTorch operations are recorded.
Implementation complexity	High. Requires strict syntax, type definitions, and careful debugging.	Low. Works well with existing models with minimal modification.
Flexibility of input shapes	Can handle dynamic input shapes.	Input shape is fixed based on the example used for tracing.

Table1. Comparison between torch.jit.script and torch.jit.trace.

3-Q-22

3-Q-22 Speech-laugh の笑い出しを誘発する
音素コンテキストの傾向

Phone context to induce laughter onset in speech-laugh

☆瀬戸口 遼, 有本 泰子(千葉工大)

- ◆ 背景: Speech-laugh は口腔前方を狭めて調音する音素で発生しやすいが、speech-laugh 発生前の音素文脈が不明であり、どんな音素文脈が笑い出しを誘発するかはわからない
- ◆ 目的: Speech-laugh 発生前における音素の文脈情報を発話と比較し、speech-laugh を誘発する音素文脈を明らかにする
- ◆ 分析: Speech-laugh 開始時や開始 1 つ前、2 つ前の音素とその組み合わせを音素 unigram や音素 bigram で表し、尤度比検定で音素文脈が発話と同じ分布から生成されるかを検証
- ◆ 結果: 音素 unigram では speech-laugh 開始 1 つ前の音素文脈の分布が最も変化しており、発話開始直後や撥音、促音の発声後に speech-laugh は発生しやすいことがわかった。Speech-laugh が始まる前に促音が多いことから、speech-laugh の始まりやすさは speech-laugh が開始する直前の促音によるものである可能性がある

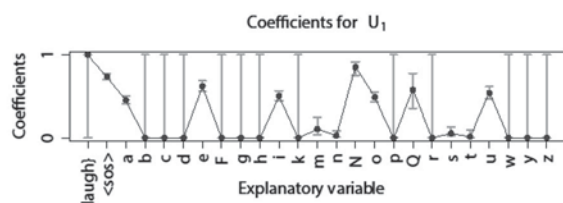


Fig.1: Estimated coefficients for the model that models the phone context before the beginning of speech-laugh

3-Q-23

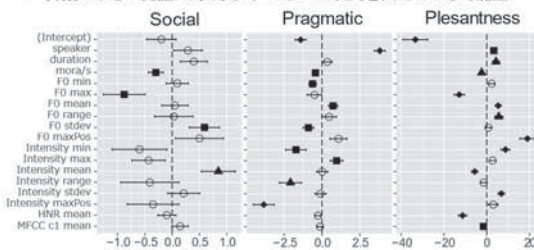
3-Q-23 笑い声の機能に対する先行発話の
音響的特徴の影響

Performance of acoustic features of preceding utterance in classifying subsequent laughter functions

☆黛大誠, 有本泰子(千葉工大)

- ◆ 背景: 対話の文脈に応じた表現を持つ笑い声の合成に向け、笑い声の機能を先行発話から予測するために、笑い声の機能ごとに先行発話の音響的特徴が異なるのか明らかにする必要があります。
- ◆ 目的: 先行発話の音響的特徴によって笑い声の機能が異なるかを検証する
- ◆ 分析: 目的変数を笑い声機能 (Pleasant, Social, Pragmatic, Pleasantness), 説明変数を先行発話の音響的特徴量とした多項ロジスティック回帰分析を実施
- ◆ 結果と考察

1. Social(気まずさの緩和)は Pleasant に比べ発話の強度平均が高い
➢ 大きな声で発生した気まずさを笑い声で緩和している可能性
2. Pragmatic(発話意味の変化)は Pleasant に比べ発話の強度変動が小さい
➢ 抑揚の小さい発話に対し笑い声によって皮肉を伝えている可能性

Fig.1: Partial regression coefficient predicted by multinomial logistic regression (■: $p < 0.05$, ▲: $p < 0.01$, ◆: $p < 0.001$)

3-Q-24

3-Q-24 顔と声の主観的類似度に基づいた話者埋め込み生成手法の検討

An Investigation of Speaker Embedding Generation Based on Subjective Face-Voice Similarity

☆宮本蓮, 齋藤大輔, 峯松信明(東大)

顔と声の「しっくりくる」をモデルに学習させる！

研究の目的

顔と声から受ける印象は人によって様々である。しかし、Face-to-Speech の従来手法はこの人間の知覚を反映できていない。

提案手法

1. 顔-顔, 声-声, 顔-声の3つのドメインで主観的類似度を測定 (Fig. 1)

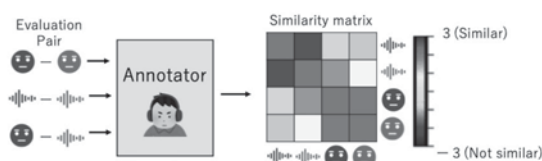


Fig.1: Overview of evaluation experiments measuring subjective similarity

2. 類似度行列をもとにモデルを学習

→これにより、人間の「似ている」という主観をモデルに学習させる！

結果

モデルの出力が人間の主観と高い相関。話者識別性能の向上も見られた。

3-Q-26

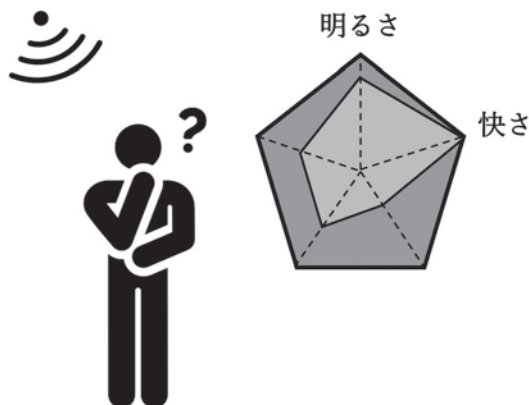
3-Q-26 環境音印象の多軸定量化タスクにむけたデータセット構築と印象予測モデルの初期検討

An initial study on developing dataset and prediction models toward quantifying multidimensional impressions of environmental sounds.

☆西島大史, 齋藤大輔, 峯松信明(東大)

従来の参照ベース評価から音そのものの絶対的な印象を評価する新しい枠組みを提案

- ◆環境音の印象を5次元(快さ・明瞭さ・明るさ・穏やかさ・臨場感)で評価するデータセットを構築
- ◆音響信号から印象を自動予測するモデルを開発し、7段階評価で誤差1段階以内の精度を達成



3-Q-25

3-Q-25 Audio Captioning モデルの発達のカリキュラム学習

Developmental Curriculum Learning for Audio Captioning Models

○稲垣 賢斗, 高道 慎之介(慶大)

オーディオキャプションングは、入力音に対してその内容を説明するキャプションを生成する、クロスモーダルな機械学習タスクである。音などの外部の情報と言葉を結び付ける作業は人間の言語発達の過程でも見られる。本研究では人間の言語発達に倣って段階的にデータセットを変化させる、発達のカリキュラム学習を提案する。日本語を母語とする幼児は、一語発話、二語発話といったように、段階的に言語表出の文長・文法複雑性が増していく。発達のカリキュラム学習では、そうした言語表出を模倣して段階ごとにデータセットを切り替えて学習する。本手法は、通常の学習と比べて異なる学習過程を経ることで、最終的なモデルの性能向上に寄与した。そして、機械学習モデルにおいて人間の学習戦略の応用が有効である可能性を示した。

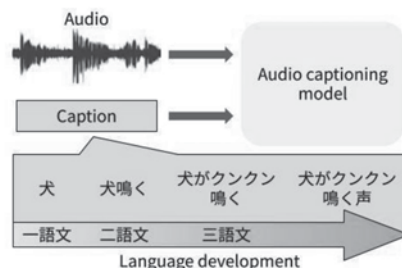


Fig.1: Overview of developmental curriculum learning in audio captioning

3-Q-27

3-Q-27 自己教師あり学習による音声表現を用いた聞き手頭部動作生成

Listener Head Motion Generation Using Speech Representations from Self-Supervised Learning

◎三河多聞(豊橋技科大), △藤井康寿(Google DeepMind), 若林佑幸(豊橋技科大), 太田健吾(阿南高専), 西村良太, 北岡教英(豊橋技科大)

- ◆本稿では、自己教師あり学習 (Self-Supervised Learning, SSL) の音声表現を用いた聞き手頭部動作生成モデルを提案し、log-Mel Filter Bank (FBank) を使用した場合などと比較評価を行った。
- ◆非言語的コミュニケーションにおいて、聞き手の頭部動作は対話の自然さ向上に不可欠であり、音声特徴と動作の関連性が注目されている。
- ◆従来はFBank 等の特徴量が主に用いられてきたが、近年ではSSLモデルから得られる音声潜在表現が持つ音韻情報の有効性が報告されている。
- ◆実験では、日本語マルチモーダル対話データセットを用い、自己回帰型のLSTM-Transformerモデルを採用し、動作発生タイミングの精度や人間の応答性パターンとの相関性等、複数観点からの分析を行った。
- ◆SSL特徴量を用いて生成された頭部動作はFBankを用いた場合より、実際の人間の動作と同期しており、また話し手の頭部動作に対する応答性の面でも改善がみられた。

Table 1. Evaluation Results. Boldface denotes the best value for each metric, and gray shading highlights our principal proposed methods.

Method	L2	FID	Var.	SI	P-FID	PCC	F1	Pre.	Rec.	Acc.	Act.
GT	-	-	15.87	1.97	-	-	-	-	-	-	0.45
Random	35.15	11.38	18.77	1.94	11.44	-2.21	0.59	0.44	0.88	0.57	0.64
Mirror	39.33	1.89	22.27	1.95	19.36	0.97	0.54	0.56	0.51	0.67	0.41
Delay Mirror	39.65	2.11	22.57	1.89	6.52	0.49	0.52	0.55	0.49	0.67	0.35
FBank	24.36	0.52	11.91	2.04	0.57	5.69	0.48	0.63	0.39	0.68	0.19
Ours(HB)	21.20	1.00	9.66	2.08	1.05	9.44	0.45	0.75	0.32	0.70	0.23
Ours(W2V2)	21.87	0.89	10.00	2.06	0.97	9.56	0.44	0.68	0.32	0.69	0.13

3-Q-28

3-Q-28 音声対話システムにおける フィラーによる心情表現のための システム発話の心情推定

Emotion estimation for sentences generated by dialogue GPT in order to emotion expression with fillers.

☆宮内雅望(大工大院・情報研), 鈴木基之(大阪工大・情報)

◆目的

- ・より人間に近い振る舞いをする音声対話システムの構築
→フィラーを用いた表現により、システムの心情を表現
- ・今回はシステムの心情を推定するモデルの精度向上が目標

◆手法

- ・LSTM モデルの学習を用いた心情の推定
- ・LSTM の入力の発話履歴の系列パターンや前発話数を変えて、心情の推定を行いより識別率の高いモデルを探す。
- ・入力の系列パターンは固定長・可変長・話者統合・前発話なし
- ・前発話数は0・3・5・7で実験

◆結果

- ・最も高い識別率は前発話なしの、LSTM 内の層1層で 32.1%
- ・入力の系列パターンの違いが識別性能に強く影響している

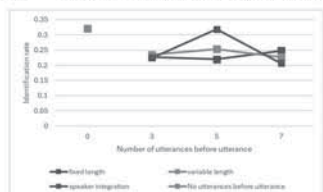


Fig.1: Recognition rate of emotion labels by number of previous utterances for each pattern

3-Q-30

3-Q-30 TV-CAR 分析に基づく ARX-LF 分析の検討

A study of ARX-LF analysis based on TV-CAR analysis

○舟木慶一(琉球大学)

音声信号の生成モデルを声帯音源波モデルで生成される声帯音源波と声道フィルタのソースフィルタモデルと仮定し、音声から LF モデル[1]と AR 係数を同時推定する ARX-LF 分析 [2]が提案されている。この方式は SEDREAMS 法[3]で GCI(声門閉鎖点)を推定し、IAIF 法[4]で推定された声帯音源から DyProg 法[5]により推定された LF モデルパラメータを予備推定値とし、その近傍に対する ARX モデル同定により LF モデルパラメータを本推定する方式である。一方、我々は解析信号に対する TV-CAR 分析[6]の応用として、TV-CAR 分析に基づく SEDREAMS 法[7]、IAIF 法[8]、DyProg 法[9]の提案を行なった。今回、それらに基づく ARX-LF 分析を構築したので報告する。

3-Q-29

3-Q-29 音声質問応答のための 日本語音声理解モデルの検討

Study on Japanese Speech Understanding Model for Spoken Question Answering

◎高島悠樹, 安藤厚志, 浅見太一 (NTT)

- ◆本研究では、大規模言語モデルに基づく日本語音声理解モデルのマルチモーダルな指示理解能力を評価するため、音声に基づく質問応答タスクに着目し、その性能を検証する。
- ◆音声質問応答・音声クエリ質問応答タスクでの評価を通じて、音声理解モデルが文脈の音声情報を適切に参照しつつ、指示文に基づいて正しい回答を生成できるかを明らかにする。
- ◆実験の結果、音声理解モデルは音声の発話内容を正しく認識できるが、指示文と統合的に処理する能力が不十分であることを示す。
- ◆また、音声質問応答データを単に用いたモデル学習においては、Table 2 に示すように、音声を参照せずに指示文のみを情報源として回答を生成する傾向があることを示す。

Table 1: Accuracy (%) on question answering tasks.

	JCommonsenseQA			JSQuAD		JaQuAD	
	TextQA	SQA	SQQA	Text	SQA	Text	SQA
1st step (single)	93.3	62.6	40.6	78.6	43.6	88.3	46.3

Table 2: Examples of incorrectly predicted answers for the SQA tasks on JCommonsenseQA.

Instruction (text)	Options (speech)	Ground-truth	Prediction
電子機器で使われる最も主要な電子回路基板の事をなんと言おう？	指示板, パソコン, マザーボード ハードディスク, まな板	マザーボード	プリント基板
音楽にあわせて踊る場所は？	ディスコ, 市街, イチゴ ショッピングセンター, マンゴー	ディスコ	ダンスホール

3-Q-31

3-Q-31 Voting に基づく基本周波数推定法の 誤差評価

Evaluation of voting-based fundamental frequency estimation methods

◎小口 純矢, 郡山 知樹 (サイバーエージェント)

Voting 法とは

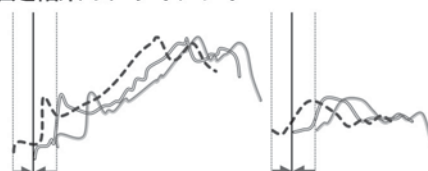
複数推定器の中央値 (V/UV は最頻値) で推定

本研究

①理論的考察: 誤差分散の低減, 真値への漸近性

②改良:

●推定結果のアライメント



●手法の選択



U/UV エラーの低減, 耐雑音性向上を達成

3-Q-32

3-Q-32 基本周波数に含まれる大局的変動の
大きさが自然性に与える影響の調査

A study on evaluation of naturalness by magnitude of global fluctuation of fundamental frequency in singing voice.

◎鈴木千文, 坂野秀樹, 旭 健作(名城大)

- ◆今回は基本周波数の大局的変動の大きさが聴感上与える影響を調査するために合成音を作成した
 - 基本周波数から回帰直線の傾きである微分信号を算出
 - 算出した微分信号の直流成分が一定ではないため、微分信号に対して回帰直線(Fig.1)を求め、それを引くことで直流成分を一定にする
 - 直流成分を一定にした微分信号からローパスフィルタで抽出した大局的変動の大きさを変える
 - ✦ Fig.2 の基本周波数の比較より回帰直線の成分を取り除くことでわずかに音高が変化した
- ◆基本周波数の大局的変動の大きさを変化させていない場合と0.5倍にした場合では、0.5倍にした方が自然性が低下した
 - 大局的変動の大きさを小さくすると自然性が低下する可能性あり

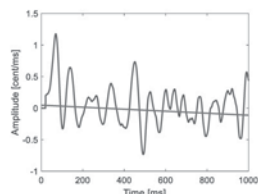


Fig.1. The derivative with regression line (Left).

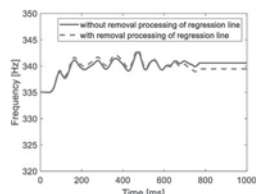


Fig.2 Comparison of fundamental frequency (Right).

3-Q-34

3-Q-34 声質変換を通じた実時間対話音声の
音響特徴分析と対人評価への影響

The acoustic feature analysis of real-time conversational speech through voice conversion and the impact on interpersonal evaluation

☆菅野聖真(豊橋技科大院), 松井淑恵(豊橋技科大)

- ◆声質変換を用いた実時間対話実験で録音された音声の音響特徴を分析し、対人評価に影響を与える音響特徴量の有無を調査した。
- ◆VC 条件と肉声条件で録音された協力者音声の発話区間を推定し、区間ごとに基本周波数、話速、HNR、jitter、shimmer を算出した。それぞれの分布がVCの有無で変化するかを観察した。
- ◆HNR と jitter がVCの利用によって変化したことが一部の話者で確認できたものの、これらの変化が対人評価の変化を直接的に引き起こした可能性は低いことがわかった。

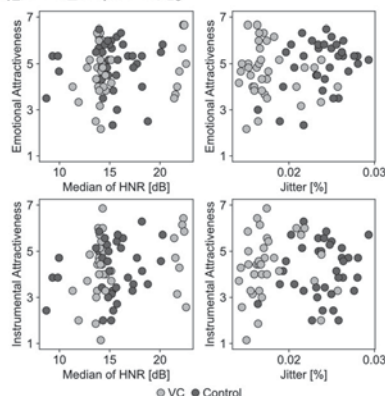


Fig. 1: Scatter plot of subjective evaluation values (top: emotional attractiveness, bottom: instrumental attractiveness) for HNR (left) and jitter (right).

3-Q-33

3-Q-33

音響特徴量による日本語ラップの分析
An Analysis of Japanese Rap Using Acoustic Features

☆鳥居広, 及川靖広(早大理工)

- ◆背景
ラップの歌唱法や響きを定量的に評価する研究は少ない。
筆者らは前回の研究においてフォルマントを比較する手法を提案した。
- ◆概要
2つの音声におけるフォルマントおよび RMS 振幅の差分を計算し、それぞれの類似度スコアを統合する手法を提案した。
日本語ラップ 30 曲分の楽曲中に含まれる韻を提案手法により評価し、検証を行なった。
- ◆結果
フォルマントだけでなく RMS 振幅も評価に加えることで、AUC が 0.776 から 0.894、AP が 0.182 から 0.423 と数値が向上した。

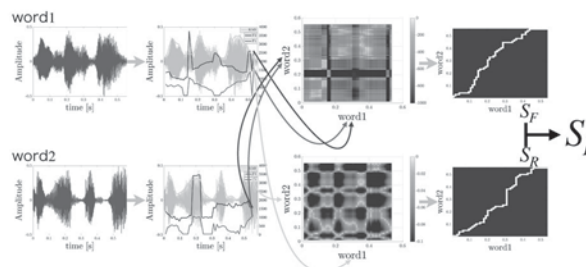


Fig.1: Example of Evaluation Method

3-Q-35

3-Q-35 日本語文章の疑問形が遅延聴覚
フィードバック条件下での発話に与える影響
Effect of Japanese Interrogative Sentences on Utterance
under Delayed Auditory Feedback

☆福留和杜, 水町光徳(九工大)

- ◆話し手の発話を話し手自身にわずかに遅らせて聞かせる遅延聴覚フィードバック (Delayed Auditory Feedback : DAF) は、個人差はあるものの発話に影響を与える。本研究では、日本語の肯定文と疑問文の両条件下で音読時間を比較し、文形式の違いが DAF の効果に与える影響を調べた。
- ◆音読文章として、肯定文を2文と、その疑問文(末尾に?を付加)を2文の計4文を準備し、それぞれの文章に対して100、200、300、400 ms の4種類の遅延時間条件下で発話時間長を計測した。
- ◆Fig.1 に示す通り、疑問文で発話時間が長くなった割合は平均で58.6% (SD : 20.5) であり、DAF 条件下において疑問文の方が肯定文より、発話時間が長くなる傾向が確認できた。

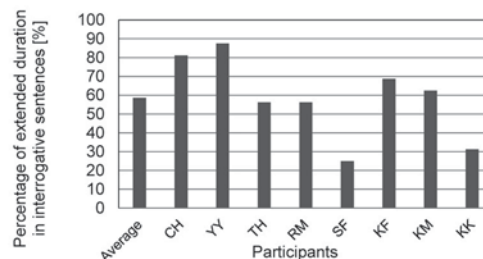


Fig. 1 Percentage of extended speech duration in interrogative sentences compared with affirmative sentences.

3-Q-36

プレゼンテーションにおける文頭プロミネンス指導による聞き手関心評価向上

The Effect of Initial Prominence Instruction on Listener Engagement Ratings in Presentations

○山下典子(早大人科・東京プロ塾), 菊池浩史・菊池英明(早大人科)

- ◆ボイストレーニングの現場では、量的なエビデンスに基づいた指導方法の確立が望まれている。そこで本研究では、プレゼンテーション指導において、文頭の F0 高低差に着目し、文頭にプロミネンスを実現させる話し方をすると、聞き手の関心が高められるかを検証した。
- ◆実験は、文頭の F0 高低差を加工したプロミネンスあり音声となし音声を刺激として用い、聞き手関心評価実験を実施した。
- ◆分析は、対応のある t 検定を行った。帰無仮説は「プロミネンスあり音声の聞き手関心評価は、なし音声と同等か低い」対立仮説は「プロミネンスあり音声の聞き手関心評価は、なし音声より高い」とした。
- ◆結果は、 p 値 0.0174 で有意水準 0.05 を下回ったため、帰無仮説を棄却し、対立仮説を採択した。つまり、文頭のプロミネンスが聞き手の関心評価を高めることが示唆された。

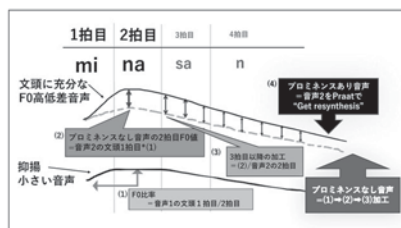


Fig. 1 文頭の F0 高低差の加工例

3-Q-38

3-Q-38 空間の印象と残響が
発話に与える影響の評価

Evaluation of the Impact on Speech Production under Different Spatial Impressions and Reverberation

◎加藤日花里, 大久保翔太, △今野智明, 堀内俊治(KDDI 総合研)

- ◆空間の印象と残響が発話に与える影響を明らかにすることを目的とし、特性の異なる 2 つの空間(大・小)において、自身の声に両耳室内インパルス応答(BRIR)や単位インパルスを実時間畳み込みした際の発話の印象に関する主観評価を行った。
- ◆各空間において、BRIR の条件に応じて「空間の広さ」、「響きの長さ」や「響きの豊かさ」の評価が変化した。
- ◆空間大・小ともに、いずれかの BRIR を用いた条件 I), II) では、III) 単位インパルスを畳み込んだ条件よりも「自然さ」、「空間と音の調和」が高く評価された。また、空間大では「話しやすさ」についても条件 I), II) で高く評価された。
- ◆以上のことから、BRIR の畳み込みによって空間の広さや響きの印象が変化し、空間と音の印象のギャップや話しやすさが改善することが示唆された。

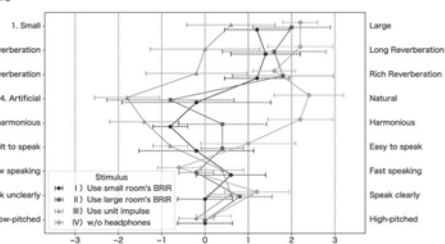


Fig.1: The evaluation scores in large room (mean ± standard deviation)

3-Q-37

3-Q-37 発話者の向きの知覚
—残響による影響の検討—

Perception of speaker's facing direction: An investigation of the effects of reverberation

☆松本汐莉, △佐藤美涼, 辻慎也, 荒井隆行(上智大),

杉本岳大, 木下光太郎, 中山靖茂(NHK)

- ◆非残響下での発話者の向きの知覚では、少なくともラウドネスがキューになることが先行研究[1]において示されている。本研究では、残響下で発話された場合の向きの知覚について調査した。
- ◆有響室でバイノーラル録音した音声と、無響室でモノ録音した音声に対して BRIR の畳み込みで残響を再現した音声を用いて、無響室での音声との聴取実験の回答結果 (Fig.1) を比較した。
- ◆バイノーラル録音では、後ろ向きの角度において無響室の音声のほうが正確な向き知覚が可能だった。有響室で録音した場合、放射方向の違いによるレベル変化がない間接音が直接音によるキューの聞き取りを妨げたと考えられる。
- ◆モノ録音では、残響の有無は話者の向き知覚に有意に影響しなかった。畳み込みによる間接音は直接音のレベル変化に比例し放射特性に依存するため、直接音によるキューを妨害しなかったと考えられる。

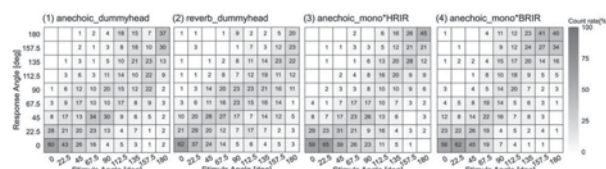


Fig. 1 Experimental response results

[1] Tsuji et al., Acoust. Sci. Tech., 46 (3), 237–241, 2025.

3-Q-39

3-Q-39 朗読音声の自分らしさが
聞き手の没入感に与える影響の検討

The Influence of Perceived Similarity to One's Own Voice on Immersion in Audiobook Listening

◎柳田耀, 長野瑞生, 藤田健一, 井島勇祐 (NTT)

- ◆先行研究では、アバターの音声や話者のアクセントにおける“自分らしさ”(聞き手と話者の類似性)が、聞き手の没入感に影響を与えることが示唆されている [Kao+, 2021; Dragojevic+, 2024].
- ◆“自分らしさ”は、聞き手の“没入感”(Transportation: 物語世界に入り込んだように感じる)に直接影響を及ぼすだけでなく、“同一化”(Identification: 登場人物の視点に立ち、思考や感情を自分のもののように感じる)を介して間接的にも影響するとされている [Kao+, 2021]. しかし、オーディオブックにおいてこのような影響が生じるかどうかは十分に検討されていない。
- ◆本研究では、自分および他人の合成音声を用いた大規模な主観評価実験を実施し、音声の“自分らしさ”が“同一化”や“没入感”に与える影響を分析した。
- ◆分析の結果、自分の音声は他人の音声よりも“自分らしい”と評価され、その評価が“同一化”を介して“没入感”に間接的に影響することが確認された。

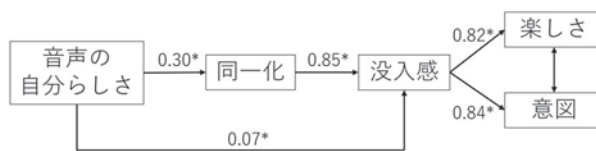


Fig.1: Path analysis of the hypothesized model

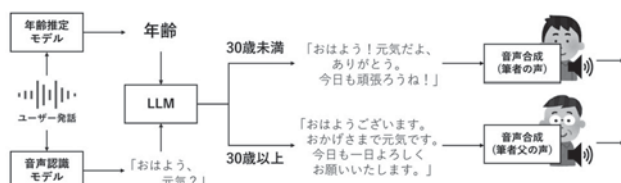
3-Q-40

3-Q-40 発話者の年齢を考慮する
音声対話システムの検討

A study on spoken dialogue systems that consider the speaker's age

八木 理一[†], 周 王子茜[†], 平野 雄太[†], 胡 宏偉[†], サクティ サクリアニ[†][†]東大寺学園高等学校, [†]奈良先端科学技術大学院大学

ユーザーごとに最適な言葉遣いや振る舞いは異なるため、音声対話システムにおいても話し手の属性の考慮が求められる。本研究では、話者年齢推定を音声対話システムに組み込むことで、ユーザーの年齢に応じた応答の実現を目指す。



- ・話者年齢推定には自己教師あり学習 (SSL) モデルを用い、推定精度の向上を図る。
- ・ASR のテキスト出力と推定された年齢をもとに、大規模言語モデル (LLM) が自然で年齢に応じた応答を生成します。
- ・生成されたテキストを YourTTS に入力し、30 歳未満なら筆者の声に、30 歳以上なら筆者の父に変換して出力する。
- ・以上により、ユーザーの年齢特性を踏まえた柔軟で適応的な音声対話システムの構築を目指す。

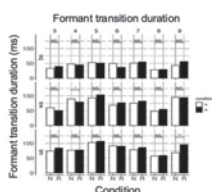
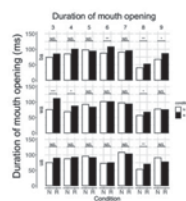
3-Q-42

3-Q-42 残響下での日本語における
破裂音/b/と接近音/w/の発話
—単音節における子音から母音への
フォルマント遷移に焦点を当て—

Production of plosive /b/ and approximant /w/ in Japanese under reverberant environments: A focus on consonant-to-vowel formant transitions in monosyllabic words

☆川井友子(上智大), 溝口愛(前橋工科大), 荒井隆行(上智大)

- ◆本研究は日本語の単音節/ba/と/wa/を発話する際、静音下と比較して残響下では発話変化が生じるか、フォルマント遷移の持続時間に着目して調査した。
- ◆残響下にて生じる発話変化には、生理学的反射である“Lombard Reflex”と“Clear Speech”による変化がどちらも含まれる。そこで本研究では、参加者に静音下と残響下条件にて“Clear Speech”による発話をしてもらった上で、フォルマント遷移の持続時間に変化が生じるのかを調査した。
- ◆フォルマント遷移の持続時間に変化が生じるか調査した結果、主に“Lombard Reflex”による発話変化では両条件間において有意な変化は生じなかった。

Fig. 1 フォルマント遷移の持続時間 (NS = 有意差なし, * $p < 0.05$, ** $p < 0.01$, *** $p < 0.001$)Fig. 2 ターゲット音節の発話に伴う開口持続時間 (NS = 有意差なし, * $p < 0.05$, ** $p < 0.01$, *** $p < 0.001$)

3-Q-41

3-Q-41 発音学習を目的としたモデル音声の
自己聴取音化に関する実験的検討

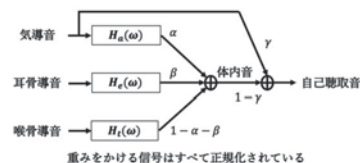
An Experimental Study on Self-Perceived Voice Conversion of Model Speech for Pronunciation Training

☆山中涼雅, 齋藤大輔, 峯松信明(東大院・工学系)

録音音声とは異なる「自己聴取音」を推定する試みがなされている

<従来の自己聴取音推定>

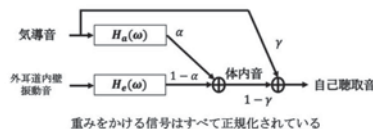
- ・自己聴取音 = 気導音 + 体内音 (体内を通して伝搬する振動)
- ・3種類の振動音をもとに体内音・自己聴取音を推定する



重みをかける信号はすべて正規化されている

<提案手法>

- ・体内の振動が外耳道の内壁を通じて生み出す空気振動に着目する



重みをかける信号はすべて正規化されている

従来の骨導センサを不要としつつ、精度と簡便性を両立した新たな自己聴取音推定手法の構築が可能かを検証しました！

3-Q-43

3-Q-43 音声スペクトル包絡からの基本周波数の
予測精度に関する検討

On the Accuracy of Fundamental Frequency Predicted from Spectral Envelopes of Speech

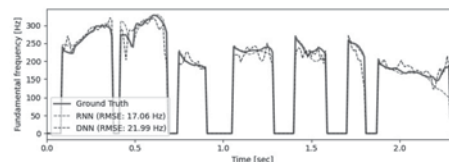
○今野 英明(北海道教育大)

◆音声のスペクトル包絡からの基本周波数予測

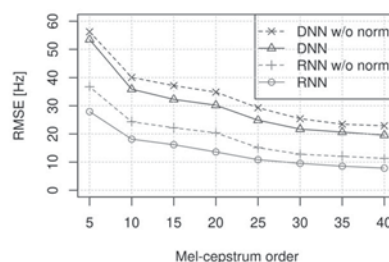
- LSTM を用いたリカレントニューラルネットワーク (RNN)
- 順伝播型の深層ニューラルネットワーク (DNN)
- 日本語 MULTEXT 韻律コーパス (男女各3名) の F0

◆次数の異なるメルケプストラムによるスペクトル包絡で精度検討

◆予測結果例：25 次メルケプストラムによるスペクトル包絡から



◆メルケプストラムの予測次数と予測誤差 (RMSE)



3-Q-44

3-Q-44 ヒトの仮声帯振動における
仮声帯間距離の影響Effect of ventricular fold distance on vocal-ventricular fold oscillations
in human voice☆水本直宏(立命館大), △谷良信(立命館大),
徳田功(立命館大)

- ◆ヒトの声帯と仮声帯の一体型物理モデルをシリコンで作成し、仮声帯間距離が仮声帯振動に及ぼす影響を調査した。
- ◆仮声帯間距離が異なる物理モデルを作成して吹鳴実験を行い、振動特性の変化を調べた。
- ◆仮声帯の内転が強いほど仮声帯振動の周波数が上昇し、音声信号は複雑化した。
- ◆特定の仮声帯間距離において、声帯と仮声帯の振動が2:1の整数比になった。

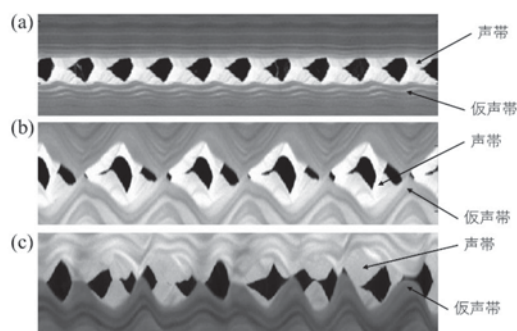


Fig.1: Kymagrams of (a) normal, (b) 1mm, and (c) 2mm models.

3-Q-45

3-Q-45 物理モデルを用いた声帯溝症が発声に
与える影響の調査

Physical modeling of sulcus vocalis and its effect on phonation

☆竹内理人(立命館大), △葉山奈南(立命館大), 徳田功(立命館大)

- ◆人間の声帯をMRIで撮影して得られたデータをベースに、声帯溝症を模擬した物理モデルを作製した。
- ◆溝のない声帯の物理モデルと、深さが異なる4種類の溝がある声帯の物理モデルの計5種類を用意した。
- ◆吹鳴実験を行い、各モデルについて声門下圧・音声・高速度映像を計測し、溝のある物理モデルと溝のない物理モデルを比較した。
- ◆溝がある物理モデルは溝のない物理モデルに比べて、基本周波数、Jitter、オンセット圧、オフセット圧のすべてが上昇する傾向を示した。

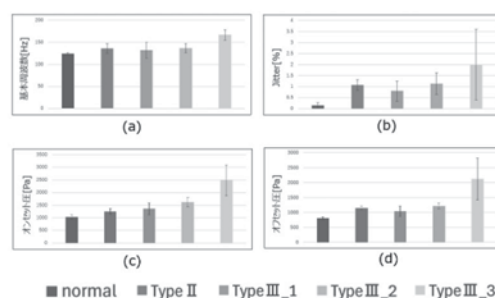


Fig.1: Experimental results of (a) fundamental frequency, (b) Jitter, (c) onset pressure, and (d) offset pressure.

3-Q-46

3-Q-46 一様型声帯物理モデルを用いた
声門内圧測定と数値解析評価Measurement of intraglottal pressure and numerical analysis
using a uniform-type vocal fold physical model

☆久米田員男, △上田凌生, △塩谷悠人, 徳田功(立命館大)

- ◆アクリル製 M5 声帯モデルを作成し、最小声門幅が 1.0 mm と 2.0 mm の一様型声門形状(Uniform 型)に固定した。
- ◆一様型声帯モデルの圧力分布を、実験・理論・数値解析の三手法で比較し、先行研究との形状比較から声帯振動のメカニズムを考察した。
- ◆理論モデルは上流側での傾向を良好に再現した一方で、声門狭窄部では誤差が生じた。CFD 解析は実験値と整合的な結果を示した。
- ◆声門形状の Convergent 型 → Uniform 型 → Divergent 型への遷移が効率的な声帯駆動に寄与していることが示された。

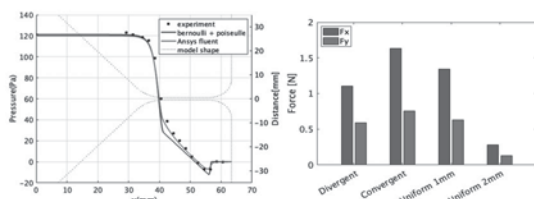
Fig.1: Intra-glottal pressure
($h=1.0$ mm)

Fig.2: Vocal cord force (experiment)

3-Q-47

3-Q-47 声帯膜の位置を変更した物理モデル
における声道の影響についてInfluence of vocal tract on physical models of vocal membranes with
variable positions

☆廣野凌, 松原大輔, 荒木駿佑(立命館大),

△西村剛(大阪大), 徳田功(立命館大)

- ◆声帯膜が声帯の外側から声門中心まで段階的に配置されているモデルに、声道を取り付け、吹鳴実験を通じて、声門下圧および音声信号を計測した。
- ◆声道を考慮した場合でも、声帯膜の位置は、発声閾値圧に同様の影響を及ぼすかを検証した。
- ◆声道の有無にかかわらず声帯膜が声帯に対して外側に位置するモデル(Lv. 1, Lv. 2)でオンセット圧が上昇し、声帯膜が声門中心に近接するモデル(Lv. 3, Lv. 4, Lv. 5)ではオンセット圧が低下した。
- ◆オフセット圧についても単体吹鳴時と声道取り付け時において、おおむね同様の傾向が示された。

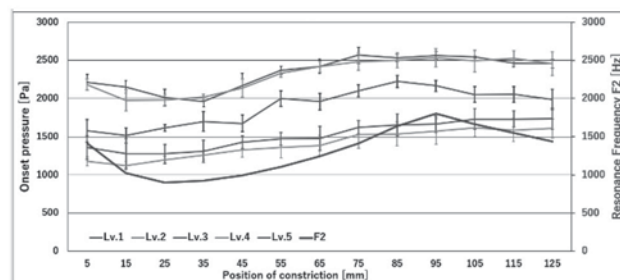


Fig. 1. Onset pressures of VM models with vocal tract.

3-Q-48

3-Q-48 サル類における声帯および
仮声帯組織のヤング率測定Measurement of young's moduli of vocal folds, vocal membranes, and
ventricular folds in rhesus macaques

☆杉江昂輝, 山本康旭, △山本憲隆(立命館大)

△西村剛(大阪大), 徳田功(立命館大)

- ◆アカゲザル喉頭から摘出された仮声帯、声帯、声帯膜に対して一軸引張試験を実施した。
- ◆応力-ひずみ線図から線形ひずみ限界 ϵ_1 とヤング率 E を算出した。
- ◆すべての個体で、仮声帯、声帯、声帯膜の順に、線形増大から指数関数的増大の領域に移行した。
- ◆他の組織に比べて仮声帯は、線形ひずみ限界 ϵ_1 が低く、ヤング率 E が高い傾向がみられた。
- ◆声帯膜は広い線形領域を持ち、ヤング率 E が低い値を示したことから、柔軟性が高いことが示唆された。

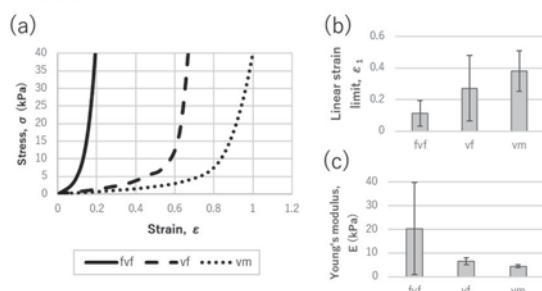


Fig. 1: Stress-strain curves from a rhesus macaque (a), and comparisons of linear strain limit (ϵ_1) (b) and young's modulus (c) between vocal folds (vf), vocal membranes (vm), and ventricular folds (fvf).

3-Q-50

3-Q-50 気嚢が声帯膜物理モデルの発声に与える影響について

The effect of air sacs on vocalization in a physical model of vocal membranes

☆松原大輔, 廣野凌, 徳田功(立命館大), 西村剛(大阪大)

- ◆チンパンジーの摘出喉頭のマイクロCT画像から作製した物理モデルと、気嚢付き声道モデルを使用し、気嚢が、声帯膜を持つ動物の発声に及ぼす影響について調べた。
- ◆気嚢がある場合とない場合で、物理モデルに与える流量を0~152 L/min (気嚢なしの場合では0~140 L/min) まで20秒間かけて増加させ、状態の変化を調べた。
- ◆オンセット流量および、声帯の振動モードが変化する流量に、気嚢の存在が影響することが確認できた。
- ◆声帯膜を持つ動物の発声において、気嚢が、発声パターンを制御する重要な役割を果たしている可能性が示唆された。

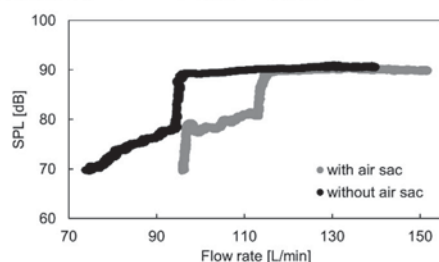


Fig. 1: Dependence of SPL on flow rate in chimpanzee model with and without air sacs.

3-Q-49

3-Q-49 仮声帯付きアカゲザルの声帯物理モデル
における整数比振動について

Vocal-ventricular fold vibrations in a physical model of macaque larynx

○山本康旭, 藤江真人(立命館大) △西村剛(大阪大), 徳田功(立命館大)

- ◆アカゲザル喉頭のマイクロCT撮像データに基づいて、仮声帯付き声帯物理モデルを構築した。
- ◆声帯の角度を 0° と 15° に設定した条件で吹鳴実験を行い、音声信号および高速度動画を計測し、比較を行った。
- ◆声帯が角度を持つ場合においては、仮声帯と声帯が1:2の周波数比で同期振動した。
- ◆サルの喉頭において1:2の同期振動が生じるのは、声帯膜の高周波数振動が、仮声帯の低周波数振動に引き込まれることが要因であることが示唆された。

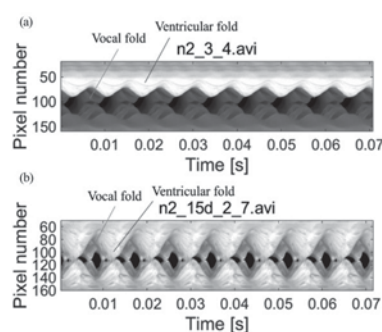


Fig.1 Kymograms of physical model (-2mm) with vocal fold angle of (a): 0° and (b): 15° .