

3-1-1

3-1-1 Multiple Vertical Panning を用いた
立体音響システムにおける
スピーカ配置の臨場感への影響Effect of the loudspeaker arrangement on sense of presence
in 3D audio system using multiple vertical panning

○木村敏幸(東北学院大)

- ◆著者はこれまでに大画面ディスプレイに適合する立体音響システムとして、今までとは異なる観点に基づいた新たな立体音響システム(Multiple Vertical Panning, 以降「MVP方式」と呼ぶ)を提案し、MVP方式が十分な性能を有することを実証してきた。さらに、スピーカをディスプレイ上部及び下部に5個ずつ配置しても視聴者は臨場感の違いを識別することができないということも示してきた。
- ◆しかしながら、先行研究におけるスピーカ配置は、ステレオホニックのような従来のスピーカ配置との互換を取るために必ずディスプレイの4隅にスピーカを配置していた。一方で、MVP方式を用いた遠隔通信会議システムを構築する場合、従来のスピーカ配置と互換をとる必要性がないため、ディスプレイの4隅にスピーカを配置する必要がなくなる。
- ◆そこで、本発表では、ディスプレイの4隅にスピーカを配置しないスピーカ配置が臨場感に及ぼす影響を視聴覚実験によって検討した。その結果、ディスプレイの4隅にスピーカを配置しなければ、必要なスピーカの数を6個(スクリーン上部及び下部にそれぞれ3個)にまで減らすことができるということが分かった。

3-1-3

3-1-3 IIR フィルタを使った頭部伝達関数の
近似による音空間レンダリングSound field rendering by IIR filter approximation
of head-related transfer function

☆片岡優太, 矢田部浩平, 及川靖広(早大理工),

Sadjad Siddiq (スクウェア・エニックス)

- ◆ビデオゲームなどのコンテンツで頭部伝達関数(HRTF)を使用する場合、計算資源が限られているため負荷の少ない処理が必要である。
➢ HRTFをIIRフィルタとしてモデル化することが有効である。
- ◆動的な音像を定位するために任意の方向のHRTFを補間する際、IIRフィルタを使った従来手法は計算コストが高くなる。
- ◆本研究では、貪欲法を使ってHRTFを双2次フィルタの直列接続として近似し、計算コストが低い線形補間に適したモデル化を実現した。
- ◆数値実験を通して、MIT Media Lab.のHRTFとモデル化したHRTFを比較し(Fig.1左)、MIT Media Lab.のHRTFと補間によって推定されたHRTFを比較した(Fig.1右)。左右の図から音像定位に影響がある15 kHz以下では誤差が小さくなることを確認した。

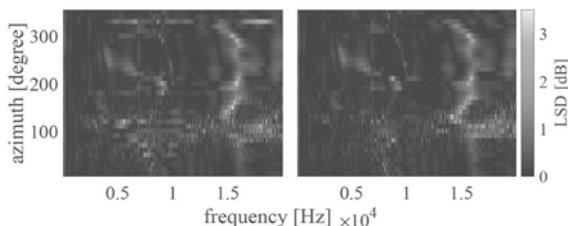


Fig. 1: The left figure shows the log spectral distortion (LSD) of HRTF modeled at 10 degree intervals and the right figure shows the LSD of the interpolated HRTF (elevation angle: 0 degree).

3-1-2

3-1-2 音響インパルス応答の測定における
マイクロホンホルダーの影響

Effect of microphone holder on acoustic impulse response measurements

☆寺島眞生, 森川大輔, モクタリ パーハム, 平原達也(富山県立大)

多点の頭部インパルス応答(hrir)を相反法によって同時に測定するためには、マイクロホンアレイが使用される。そのマイクロホンアレイに用いる小型エレクトレット・コンデンサー・マイクロホン(ECM)をアルミ角棒上に正確に配置するためのマイクロホン・ホルダーを3Dプリンタによって作成し(Fig.1)、それらの音響特性を測った。

裸の状態の小型ECMで計測したインパルス応答のスペクトルと各マイクロホン・ホルダーを装着した小型ECMで計測したインパルス応答のスペクトルの差 $\Delta|SPC_i|$ ($i:A, B, C$)をFig.2に示す。高域における $\Delta|SPC_i|$ の増加量が最小であり、材料の使用量が最も少ないのはマイクロホン・ホルダーCであった。また、アルミ角棒からの高さ h を30 mmおよび60 mmとして計測したインパルス応答波形の違いはほとんど無かった。

3つのマイクロホン・ホルダー用いて音響インパルス応答を計測し、それらの影響を確認した結果、マイクロホン・ホルダーCがマイクロホンアレイに最適であることがわかった。

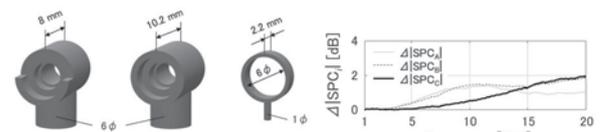


Fig.1 マイクロホン・ホルダー (A, B, C)

Fig.2 $\Delta|SPC_i|$ ($i:A, B, C$)

3-1-4

3-1-4 多角スピーカアレイを用いた
局所音場合成における音像定位評価Sound localization in local sound field synthesis
using polygonal loudspeaker array☆黒川翔瑠, 津國和泉, 池田雄介, 小坂直敏(東京電機大),
及川靖広(早大理工)

- ◆背景
再現領域を聴取者頭部周辺に限定する局所音場再現手法において、再現領域の広さや位置の違いが、音像定位に与える影響は不明
- ◆目的
両耳間レベル差(ILD)の測定によって、局所再現領域の広さと位置の違いが音像定位に与える影響を調査
- ◆結果
直径40cmの局所領域を中心から離れた位置で再現する場合、聴取者から近い位置のスピーカ方向から伝搬する平面波で、ILDが歪む

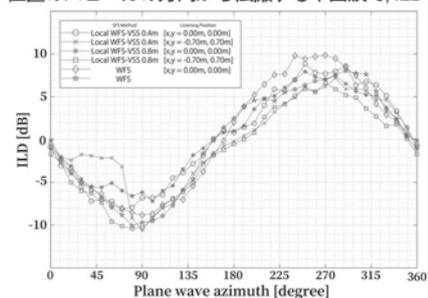


Fig. 1: ILDs at different listener's positions with WFS-VSS and conventional WFS. x and y indicate the coordinates of the listening position, respectively. The plane wave propagation azimuth is 0 degree in the front direction of listener, and the azimuth rotates counterclockwise.

3-1-5

講演取消

3-1-6

3-1-6 放射エネルギー最適化に基づくフォルマント強調を用いた音声強調型パラメトリックスピーカ

Speech-enhanced Parametric Array Loudspeaker Using Formant Enhancement Based on Emitted Energy Optimization

◎ 歌毓庭(立命館大院), 中山雅人(阪産大), 西浦敬信(立命館大)

パラメトリックスピーカ(PAL)は再生の原理上、低域の再生が困難である。我々はこれまでミッシングファンダメンタルに基づき、PALの楽器音と音楽再生音強調を提案してきた。そこで本稿では、音声の強調に注目し、今まで提案してきた放射エネルギー最適化やPALの逆特性で音質補償の処理に、フォルマント強調を加えて、音声強調型PALを提案する。そして、音声の出力音圧と了解度に関する評価実験で提案手法の有効性を確認する。

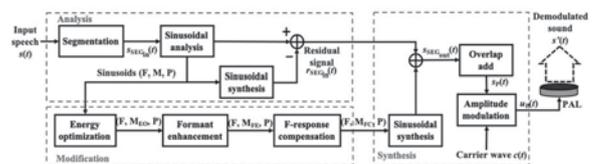


Fig. 1 Block diagram of proposed method

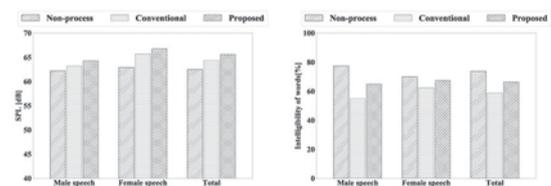


Fig. 2 Experimental results on sound pressure and intelligibility

3-1-7

3-1-7 パラメトリック/ダイナミックスピーカにおける回帰分析に基づく空間インパルス応答の直間比補正

DRR Correction of Spatial Impulse Response Based on Regression Analysis in Parametric/dynamic Loudspeaker

☆ 江川琢真, 中山雅人, 高橋徹 (阪産大院)

- ◆我々はPAL (Parametric-Array Loudspeaker)とEDL (Electro-Dynamic Loudspeaker)を用いたVSS (Virtual Sound Source)の距離制御に基づく新しい立体音響システムを提案している。本稿では、受聴者近傍において劣化する傾向があったDRR (Direct-to-Reverberant Ratio)の誤差を低減するため、実音源のDRRを計測し、その回帰分析に基づきVSSのDRRを補正する手法を提案する。
- ◆実環境においてスピーカ、パラメトリックスピーカ、無指向性マイクで室内インパルス応答を計測し、従来手法(補正前)と提案手法(補正後)の性能をDRRで評価した。Fig.1に従来手法のDRRと実音源のDRRを示す。Fig.1より近距離音場(0.5m以下)におけるDRRの誤差が大きいことがわかる。Fig.2に提案手法の補正関数を示す。Fig.3に近距離音場の条件における従来手法と提案手法の平均DRR誤差の評価結果を示す。実験結果より、従来手法に比べて提案手法の平均DRRの誤差が小さかったことから、提案手法の有効性が確認できた。

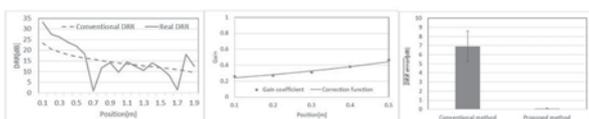


Fig.1: Conventional DRR and real DRR Fig.2: Gain coefficients and correction function Fig.3: Experimental result for average of DRR error

3-1-8

3-1-8 Acoustic Contrast を目的関数とした トランスオーラルシステム

Transaural system using acoustic contrast as its objective function

○ 鮫島俊哉(九大・芸工), △ 小林佳蓮(フォスター電機株式会社), 関口周(九大・芸工)

- ◆音場再生手法の一つであるトランスオーラルシステムは、バイノーラル録音された音源をスピーカを用いて聴取者の耳元に再生する手法であり、スピーカと耳元との音響伝達関数(再生系伝達関数)の逆フィルタを組み込むことによりそれを実現するものである。
- ◆その逆フィルタは、再生系伝達関数の等化とクロストークキャンセルの2つの目標を同時に満足するように設計されるが、それは両立することが物理的に困難なものである。
- ◆そこで、Acoustic Contrastを最大化するという制御器設計手法を応用することで、クロストークキャンセルのみを強調したトランスオーラルシステムを提案する。
- ◆提案システムの利点は、(1)小規模なシステムであっても所望の音像定位効果が得られること、(2)既往のトランスオーラルシステムにおける課題であった、ダイナミックレンジとS/Nの低下の問題が改善されること、である。

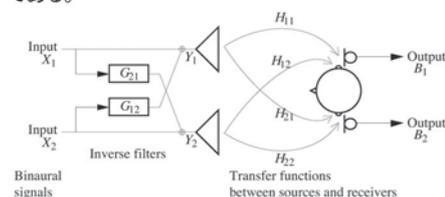


Fig.1: Transaural system using acoustic contrast.

3-1-9

3-1-9 ミッドフリンジロック干渉計による
シンプルかつ低雑音な光音響計測法
Simple and low-noise optical measurement of sound
using mid-fringe locked interferometer

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

課題：光では小さい音が計測できない

- ・ 空気中だと音による光位相変化が微小
- ・ 現状 70 dB が検出可能音圧レベルの目安
- ・ 光を使いたくても測定できない場合が多くある

↓

「音のための精密光計測」の研究に着手

- ・ NTT 物性研量子光デバイス研究グループと共同研究
- ・ 光音響計測の飛躍的な進歩を目指す

【本発表】

- ・ シンプルな光干渉計+フィードバック制御
- ・ 光学系を安定化し低周波雑音を相殺
- ・ 等価雑音量換算で 50 dB/Hz の性能を実現

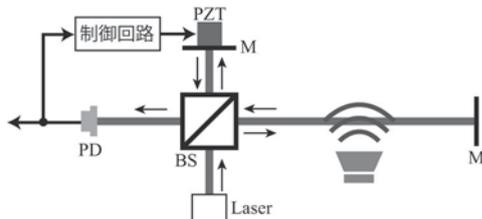


Fig.1 Schematic of the proposed measurement system.

講演取消

3-1-11

3-1-11 少数マイクロホンアレイによる
三次元初期インパルス応答の外挿

Extrapolation of three-dimensional early Impulse responses

☆津國和泉, 松橋遼, 黒川翔瑠, 池田雄介, 小坂直敏(東京電機大)

- ◆問題点
効率的な多点インパルス応答の測定は困難
- ◆目的
初期反射音を含むインパルス応答の効率的な測定
- ◆提案手法
等価音源法および虚像法を利用し、少数マイクロホンの測定データから、マイクロホン周囲のインパルス応答を推定
- ◆実験
- キューブ型、十字型、球型の3アレイを用いてマイクロホン形状による推定精度の比較
- 上記の3アレイを用いて提案手法のノイズ耐性を比較

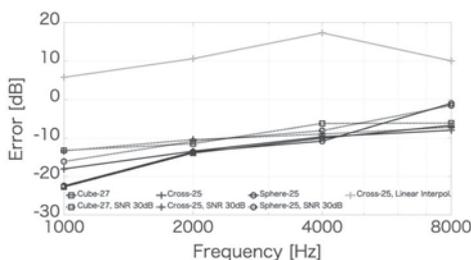


Fig.1: Estimated Error Mean. (Cube: □, Cross: +, Sphere: ○, Solid lines show no noise, dot lines show 30dB SNR, gray line shows Linear Interpolation.)

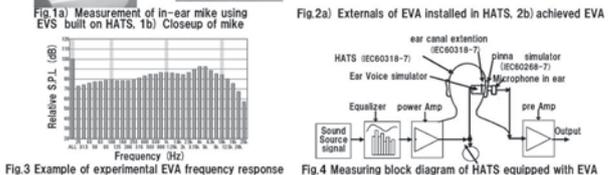
3-1-10

3-1-12 EVS (イヤーパーボイスシミュレータ)を
内蔵した HATS に関する検討と考察

Examination and consideration concerning HATS
with built-in EVS (Ear Voice Simulator)

○稲永潔文(サザン音響)

- ◆近年の音響機器、通信機器の小型化やウェアラブル化に伴い、耳殻近傍でピックアップ可能な種々の情報を、インターフェース用信号として活用することが試みられている。
- ◆特に耳内マイクでの收音は、口部放射音マイクの設置が不要で咀嚼音等も收音可能であり、またマイク筐体が遮音されていることから外部騒音の影響を受け難く、外部とのコミュニケーションや各種機器の制御などにも適している。
- ◆しかし、実際の耳内マイクを搭載したウェアラブル機器の検討や性能チェックを行うためには多人数の被験者が必要で、個人差の影響もあり一般的ではなかった。
- ◆そこで平均的な人間の耳内伝搬音を生成可能な、イヤーパーボイスシミュレータ(EVS)を搭載したHATSの開発を行った。
- ◆その結果、耳内收音用マイクを搭載したウェアラブル機器の性能チェックや検査を容易に行うことが可能な、イヤーパーボイスシミュレータ(EVS)を搭載したHATS測定系が実現可能となった。



3-1-13

3-1-13 位相共役光を用いた低周波振動に頑健な干渉計の構築と音計測への適用

Construction of interferometer that is robust for low-frequency vibration by using phase conjugate wave

☆田部井駿祐, 矢田部浩平, 及川靖広 (早大理工), 石川憲治 (NTT)

- ◇光学的音響計測では非接触に音場の計測が可能
- ◇光学系が外部の低周波振動の影響を受けてしまう
- ◇位相共役光を用いて低周波振動をキャンセルする位相共役干渉計を構築

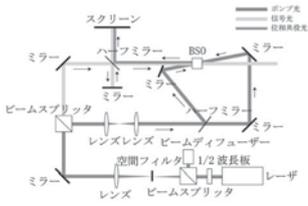


Fig1. 位相共役干渉計の概略

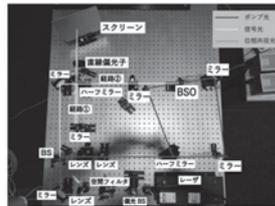


Fig2. 実際の位相共役干渉計

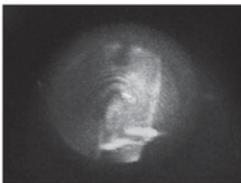


Fig3. スクリーン上に観察された干渉縞

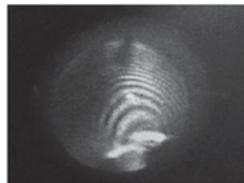


Fig.2の経路1または経路2に低周波振動の例として炎を置く、経路1の参照光に置く位相変化がそのままスクリーンに映る(左図)。対して経路2では位相共役光を返して位相変化をキャンセルするので干渉縞の乱れがない(右図)。

3-1-14

3-1-14 感染症対策が音声の伝達特性に与える影響の測定

Measurement of the effect of prevention of infectious diseases on transfer characteristics of speech

○高橋慧, 中家諒, 山名一輝(TOA)

- ◆感染症対策として普及しつつあるマスク、フェイスシールド、パーティションの影響を測定した。
- ◆マスクの影響は種類によって違いがあり、伝達特性にあまり影響を与えない場合や、2 kHz以上のエネルギーが減衰する場合等が見られた。
- ◆フェイスシールドや透明マスクを装着すると、エネルギーが減衰するだけでなく増幅される周波数帯域が存在することが明らかになった。
- ◆パーティションは厚さによって影響が異なることを確認した。厚さの薄いパーティションを使用することで、伝達特性に与える影響を抑えられることが示唆された。
- ◆マスクの影響はスピーカから再生する音源を使用した測定だけでなく、被験者に発話させての測定も行なった。結果、異なる傾向が確認された。マスクによって口の動きが制限されるため、実際の使用状況での影響を明らかにする上では被験者の発話を測定することは重要であると考えられる。

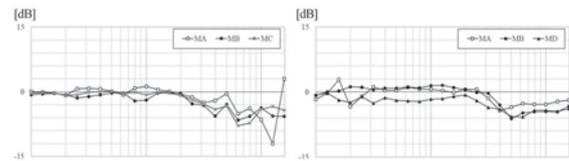


Fig.1: Difference of frequency response from "No Mask." (Left : Measured by log-swept sine signal, Right : Measured by speech)

3-2-1

3-2-1 聴取者ごとの感情知覚モデルに基づく音声感情認識

Speech Emotion Recognition based on Listener-wise Perception Model.

○安藤厚志, 森岳至, 小橋川哲(NTT), 戸田智基(名大)

- ◆従来の音声感情認識の多くは、聴取者ごとの知覚感情のうち多数派であった感情(多数派感情)を直接推定する手法を用いる。しかし、感情の知覚基準は聴取者ごとに異なると考えられるため、従来手法は様々な感情知覚基準を単一のモデルで学習していることに等しく、このようなモデルを学習することは困難である。
- ◆本稿では、聴取者ごとに異なる感情知覚基準の獲得を目指し、聴取者ごとの感情知覚モデル(特定聴取者モデル)を構築し、複数の特定聴取者モデルを用いて多数派感情を推定する手法を提案する。
- ◆提案手法は Fine-tuning に基づく手法と、聴取者コードに基づく手法の二つのアプローチにより特定聴取者モデルを構築する。いずれも、音声認識や音声合成における話者・ドメイン適応を参考としている。
- ◆実験の結果、多数派感情および聴取者ごとの知覚感情の推定において提案手法は従来手法に比べて高い推定精度を示す傾向がみられた。

Table 1: Estimation results of the major / listener-wise perceived emotions.

	Major	Listener 1		Listener 2		Listener 3				
		WA	UA	WA	UA	WA	UA			
Baseline	65.1	63.8	65.1	59.5	66.0	63.7	61.5	64.7		
Proposed	Listener indep.	64.9	66.7	63.5	59.7	66.5	67.3	58.5	65.5	
	Fine-tuning	(Lis.1)	64.0	61.7	62.3	66.8	63.0	65.8		
		(Lis.2)	64.3	59.7	67.1	65.4	61.1	64.7		
		(Lis.3)	62.8	57.7	67.0	63.4	63.7	65.5		
	Avg.	65.1	65.2	-	-	-	-	-	-	
Listener code	68.4	67.2	61.3	60.5	67.9	67.3	73.3	61.7		

3-2-2

3-2-2 ニュートラル音声の特性を考慮したソフトラベルによる感情分類モデル学習

Speech Emotion Classification using Soft Labels based on the Characteristics of Neutral

☆大澤 まゆ子(立命館大), 井本 桂右(同志社大), 山西 良典(関西大), 福森 隆寛, 山下 洋一(立命館大)

- ◆目的: 感情音声分類の精度向上を目指す。
 - ◆問題: ニュートラル感情がラベルとして付与された音声には、他の感情の特徴が含まれることがある。
 - ◆提案手法: ニュートラル音声に他の感情の特徴を含むことを考慮した2種類のソフトラベルを用いてニュートラル音声を表現する手法を提案する。提案手法のソフトラベルには以下の式を応用している。
- $$q'(k, e) = (1 - \alpha)\delta_{k,e} + \alpha u(k, e)$$
- ここで、発話の正解感情 e を、感情クラス k とすると、 $q'(k, y)$ は学習時に正解とする分布を表す。 α は平滑化係数で、 $\delta_{k,e}$ は $u(k, e)$ と独立した分布である。
- ◆分類精度: Fig.1 に平滑化係数 α に対する感情音声の分類性能を示す。

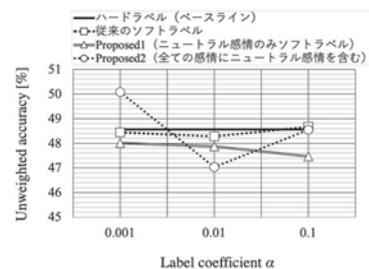


Fig.1: Accuracy comparison between the baseline and the proposed method

3-2-3

3-2-3 音声の時系列特徴量と統計量の融合による感情認識

Emotion recognition by fusion of time series features and statistics of speech

☆羽田優花, △櫻井美咲, 加藤正治, 小坂哲夫 (山形大)

- ◆従来の日本語音声感情認識の研究では、フレームごとに計算した時系列の特徴量から音全体の特徴を捉えるために統計量の計算をしたものを特徴ベクトルとして認識モデルの学習、認識を行うことが多い。
- ◆本稿ではDNN 音声感情認識についてフレームごとの特徴量から統計量を計算したものを特徴ベクトルにした場合と、フレームごとの時系列特徴をそのまま用いた場合の比較と、これらを後期融合させた際の認識率について報告する。
- ◆以下の式で後期融合を行う。対象感情emoは怒り、喜び、悲しみ、平静。融合後の尤度をL、統計量の尤度をS、時系列特徴量の尤度をDとする。αは重み調節のパラメータである。

$$L_{emo} = \alpha S_{emo} + (1 - \alpha) D_{emo}$$

- ◆融合の結果を Figure 1 に示す。融合により認識率が向上した箇所があり、α = 0.3, 0.4 の際に最高である74.5%を得た。単体での認識率の高かった時系列特徴量への重みをやや強くした場合に良い結果が得られたと言える。

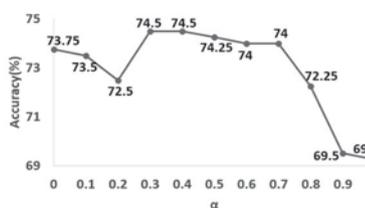


Figure 1 Result of late fusion

3-2-5

3-2-5 感情音声合成のための Dirichlet VAE

Dirichlet VAE for emotional speech synthesis

☆藤本崇人, 高木信二, 橋本佳, 大浦圭一郎, 南角吉彦, 徳田恵一(名工大)

本稿では、感情音声合成のための Dirichlet variational auto-encoder に基づく音響モデルを提案する。提案モデルでは Dirichlet 分布に従う潜在変数を用いることで、音声の細かい感情の強度を表現する。音声毎により適切な感情強度を推定するため、音響モデルをより精度良くモデル化することが可能となる。また、提案モデルでは特定の感情やその強度も制御可能である。実験結果からモデル化性能の向上とより強く感情を表現可能であることが示され、提案モデルの有効性を示した。

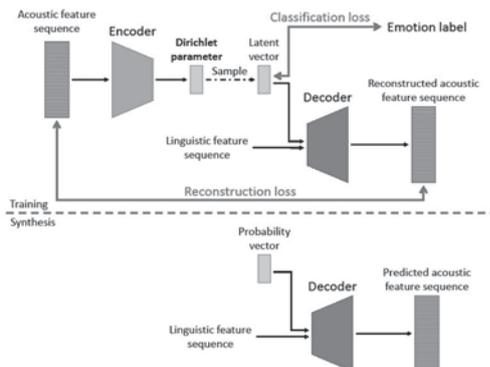


Fig. 1: 提案法の概要

3-2-4

3-2-4 笑い声の感情次元知覚に対する呼気音・吸気音の影響

Influence of exhalation and inhalation on emotional dimension perception of laughter

☆今西利於, △小山俊樹, 有本泰子(千葉工業大学)

- ◆音声合成技術を用いて会話をするロボットが広く親しまれている
- ◆音声合成器から出力される発話は発話内容が正しく聞き取れる言語音として成立しているが、笑い声を合成音声で生成すると人間が発する笑い声とは異なる不自然なものになってしまう
- ◆上記の背景から本研究では笑い声合成への応用を念頭に、笑い声の構成要素が笑い声の感情知覚に与える影響を分析した
- ◆主観評価実験によって、笑い声に対して「快-不快」、「覚醒-睡眠」、「支配-服従」の3つの感情次元に関する評価値を分析した結果、全ての感情次元で有声吸気が他の構成要素よりも有意に高い値を示すことがわかった

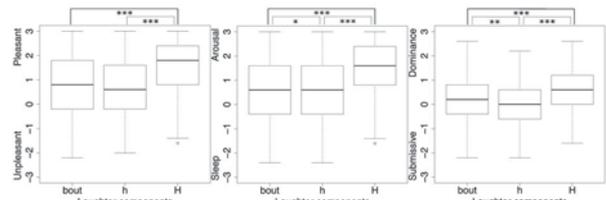


Fig. 1: boxplot of emotional for each laughter component (*p < 0.05, **p < 0.01, ***p < 0.001)

3-2-6

3-2-6 階層化多重粒度生成モデルを用いた表現豊かな音声合成

Expressive speech synthesis using hierarchical multi-grained generative model

◎法野行哉^{1,2}, 坪井一菜², 沢田慶², 橋本佳¹, 大浦圭一郎¹, 南角吉彦¹, 徳田恵一¹ (1名工大, 2Microsoft)

本稿では、表現豊かな音声合成のための階層化多重粒度生成モデルを提案する。提案モデルは、発話単位、フレーズ単位、単語単位の3つの異なる時間解像度の潜在変数をモデル化する多重粒度VAE、条件付き事前分布、多層コンバータの3つの要素で構成される。与えられたテキストを考慮しながら、潜在変数を持つ言語的階層構造や時間的依存関係をモデル化することで、単語単位の潜在変数の適切なサンプリングが可能となり、表現豊かな音声の合成を実現した。また、提案モデルでは、合成時の発話スタイルを容易に制御することも可能である。

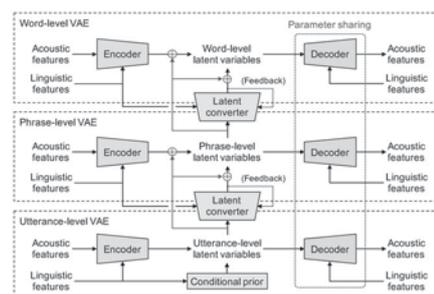


Fig. 1: Overview of the proposed system

3-2-7

3-2-7	WOZ との比較による自律型アンドロイド ERICA の傾聴対話システムの評価
Comparison of WOZ and autonomous system for attentive listening with android ERICA	
井上昂治, Divesh Lala, 山本賢太, 中村静, 高梨克也, 河原達也 (京大・情報学)	

■ 自律型アンドロイド ERICA の傾聴対話システム

- 傾聴対話：相手の話に耳を傾けて聞く
 - 高齢者の話し相手 (傾聴ボランティア)
- 提案システム
 - 話を理解・共感している感覚を醸したす多様な聞き手応答
 - 相槌, 繰り返し, 掘り下げ質問, 評価応答, 語彙的応答

■ WOZ との比較によるシステムの評価

- 人間による傾聴 (WOZ: Wizard-of-Oz) との比較
 - 提案システムと同じ種類の聞き手応答を人間が生成・発話
 - 20名の被験者 (70~80歳代)

【結果1】提案システムの各発話の評価

- ほとんどが破綻せずに, 被験者の肯定的な反応を引き出した
- ただし, 「聞き手応答として適切」と判定されたのは約6割

【結果2】提案システムと WOZ の比較 (主観評価)

- 基本的な傾聴スキルに関する評価は人間と遜色なし
 - 真面目に / 集中して / 積極的に話を聞く
- 深い理解を要する高度なスキルにはギャップがある
 - 話を理解する, 興味・共感を示す

3-2-9

3-2-9 グループワーク音声認識のためのマルチチャンネルVADの反復処理による精度改善

Iterative multi-channel VAD for transcription of group discussion

☆中野魁人(滋賀大), △中山隆弘(東京大・高大セ), △白水 始(国教研), 市川治(滋賀大)

- ◆新しい教育手法としてグループワークなどアクティブラーニングが注目されている。グループワークの学びの過程を可視化するために音声認識を行いテキスト化を行っている。しかし、雑音に強いとされる接話マイクロフォンを用いても、出力の音声データに隣接する生徒の発音が混入することが頻発してしまう。
- ◆我々はこの問題に対処するために、グループ内の複数の接話マイクロフォンの出力を総合して、対象話者の発声区間の判定を行うマルチチャンネルVADを提唱した。
- ◆グループ内の話者から、対象話者でない人を1人ずつ選びペアとする。横軸に対象話者、縦軸に隣接話者のフレームごとの音声パワーのプロットをとる。音声に対象話者からの発声であれば、下側にプロットされるため、決定境界を引くことができる。全ペアについて下側であれば、そのフレームの音声は対象話者からの発声と判断できる。
- ◆マイクロフォンゲインのバラつきを補正するために、各話者共通と考えられる背景雑音のレベルで音声パワーを基準化することを行った。
- ◆前回の報告では、反復を行わない学習により、決定境界を決定した。本報告では、反復を行うことにより、決定境界の学習を改善することを目指す。
- ◆実際の中学生のグループワーク授業のデータを用いて、VADにより得られた発話区間情報と手作業により得られた正解の発話区間情報を比較する評価実験を行った。反復ありの学習方式の性能が一番よく、次に反復なしの学習方式、そして学習無し(対角線方式)が一番悪かった。

3-2-8

3-2-8 SPOKEN LANGUAGE ACQUISITION BASED ON REINFORCEMENT LEARNING AND WORD UNIT SEGMENTATION

☆Shengzhuo Gao, Wenxin Hou, Tomohiro Tanaka, Takahiro Shinozaki (東工大)

- ◆The process of spoken-language acquisition has been one of the topics of greatest interest to linguists for decades
- ◆We propose a new framework for simulating spoken-language acquisition by combining reinforcement learning and unsupervised learning methods
- ◆We work in the direction of Skinner's idea, which states that children acquire language based on behaviorist reinforcement principle, and provide a proof of concept by running a computer-simulated experiment
- ◆We show that a spoken language can be acquired considerably faster by identifying potential word segments from collected ambient sounds in an unsupervised manner

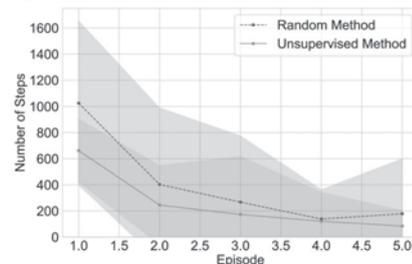


Fig.1: Comparison between Random baseline and Unsupervised-accelerated proposed method

3-2-10

3-2-10 構音障害者音声認識における発話辞書適応の検討

An investigation of adaptation of a pronunciation dictionary for dysarthric speech recognition

☆澤田祐哉, 高島遼一, 滝口哲也, 有木康雄 (神戸大)

- ◆ 構音障害者は筋肉の不随意運動により、健常者とは異なるスタイルの発話をする。
- ◆ 音声認識で一般に用いられている発話辞書は、構音障害者の発音を反映していないため、構音障害者音声認識に一般辞書を用いることは必ずしも適切ではない可能性がある。
- ◆ 本研究では、構音障害者の音素認識結果から音素の誤認識パターンを分析し、発話辞書を修正するルールを決定する。
- ◆ 決定したルールに基づいて発話辞書に発音を追加することで、構音障害者に発話辞書を適応させる。
- ◆ 適応辞書を用いた単語認識実験から、適応辞書を用いた場合には単語の誤認識率が低減することが分かった。

Table 1 The extracted substitution rules with their occurrence rates of each dysarthric speaker.

Speaker	1	2	3	4	5
SPK1	z→d	f→s	ts→sh	ts→ch	s→y
	0.190	0.143	0.143	0.107	0.098
SPK2	f→s	p→t	z→d	ch→k	e→i
	0.294	0.250	0.185	0.179	0.169

3-2-11

3-2-11 構音障害者音声認識における認識モデルの比較評価

A comparative evaluation of acoustic models on the task of dysarthric speech recognition

©高島遼一, 有木康雄, 滝口哲也(神戸大)

- ◆従来の構音障害者音声認識の研究では、それぞれ異なる音声認識モデルをベースラインとして、各自の提案手法を評価している。
- ◆本稿では、構音障害者音声認識に適したベースラインモデルを明らかにすることを目的として、複数の音声認識モデルの比較評価を行う。
- ◆実験では、DNN-HMM ベースのモデルとしてクロスエントロピー(CE) 基準, sMBR 基準, lattice-free MMI (LFMMI)で学習したモデルを, end-to-end モデルとして CTC, attention-s2s, hybrid CTC/attention モデルを比較評価した。
- ◆実験の結果, end-to-end モデルは多少学習データが不足していたとしても, DNN-HMM より優位であり, attention-s2s は CTC とのマルチタスク学習により attention matrix の推定を安定化することが重要であることが明らかとなった (Table 1)。

Model	PER [%]	α	w/o RNN-LM	w/ RNN-LM
CE-DNN-HMM	48.4	1.0 (CTC)	26.8	26.1
sMBR-DNN-HMM	44.8	0.8	27.4	27.3
LFMMI-TDNN	35.6	0.6	26.3	26.1
CTC	26.8	0.5	25.7	25.7
Attention-s2s	86.5	0.4	28.2	28.2
		0.2	30.5	30.6
		0.05	55.4	57.6
		0.0 (Attention-s2s)	86.5	86.4

Table 1: Phone error rates (PERs) of compared models (left) and PERs of hybrid CTC/attention models (right)

3-2-13

3-2-13 単語の表記と素性を同時出力する end-to-end 音声認識

End-to-end ASR for simultaneously estimating grapheme and text features

©大町基, 藤田悠哉 (ヤフー), 渡部晋治, Xuankai Chang (JHU)

課題

- ◆End-to-end 音声認識 (E2E ASR)は、発話内容の表記列を出力するものが多い。しかし、実用面を考えると発話内容に含まれる単語の読みや品詞などの素性情報もまた重要な情報となる。

提案法と特長

- ◆発話内容に含まれる個々の単語の表記と素性を並べた単一の系列を出力する E2E ASR を提案
 - > 条件付き独立の仮定を置くことなく、表記と素性の同時確率が最大となるようにモデルを学習するため、各系列の推定性能が改善すると期待できる
 - > 個々の系列の対応関係を保持したまま出力するため、単語分割やアライメントなどの処理も同時に実現できる

結果

- ◆CSJ を用いた音声認識実験：各系列の推定精度が改善 (Table 1)
- ◆単語分割、読み・品詞推定：言語情報のみを用いた手法 (KyTea) の性能を上回る (Table2, 3)。提案法は、言語情報のみを使った手法では解くことが難しい、同型異音語や言い淀みを含む文に対しても頑健に推定できることを確認

System	eval1		eval2		eval3	
	G	Y	G	Y	G	Y
Conv-G	7.4	—	5.3	—	7.4	—
Conv-Y	—	4.5	—	2.9	—	3.2
Prop-GY	6.9	4.4	4.7	2.6	6.1	3.7
Prop-GYP	6.4	4.1	4.7	2.7	5.2	3.0

Table 1: CER of CSJ. G and Y denote Grapheme and Yomi, respectively. Prop-GY and Prop-GYP are proposed method.

System	Precision	Recall	F-value
KyTea	99.67	99.62	99.64
Prop-GYP	99.80	99.80	99.80

Table 2: Evaluation of word segmentation. Kytea is one of the NLP-based method.

System	Y	P	Y+P
KyTea	99.50	99.01	98.68
Prop-GYP	99.77	99.37	99.21

Table 3: Evaluation of feature estimation. P denotes part-of-speech (POS).

3-2-12

3-2-12 音楽重畳音声認識のための時間領域音源分離と sequence-to-sequence 音声認識の同時学習

Joint training of time-domain source separation and sequence-to-sequence ASR with background music interference

☆Woo Jeongwoo, 三村正人, 吉井和佳, 河原達也 (京大)

- 映画やドラマなどの音声は、演出のために背景音楽と混合されるが、音声認識の性能の低下の原因となる。
- 右図のような時間領域音声音楽分離モデル(SEP)と音声認識モデル(ASR)の同時学習を提案する。分離モデルには Conv-TasNet を使用する。
- 日本語話し言葉コーパス(CSJ)とアニメーションの背景音楽の混合音源を用いた音源分離・音声認識実験において提案手法の有効性を確認した。

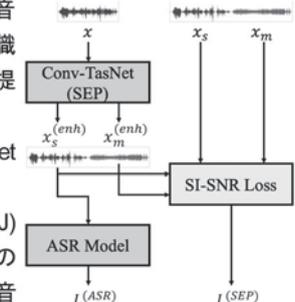


Fig.1: An overview of the proposed method

Table.1: The word error rates (%) in different SNRs

Model	Clean	5dB	0dB	-5dB	avg
Clean	11.25	46.62	78.96	93.33	72.97
Mixture	12.29	15.26	19.63	31.57	22.15
Time Domain SEP + Clean ASR (cascade)	11.40	14.35	18.23	28.59	20.39
Time Domain SEP + Clean ASR (joint)	12.52	13.63	15.41	19.90	16.31

3-2-14

Sequence-to-sequence モデルによる日本語臨床会話における

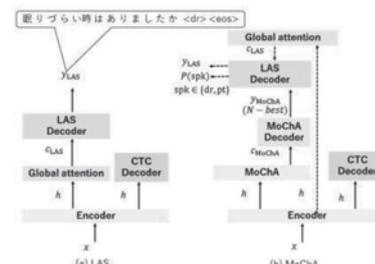
日本語臨床会話における

音声認識・話者ダイアライゼーションの同時実現
Simultaneous realization automatic speech recognition and speaker diarization using sequence-to-sequence on Japanese medical conversation

○小島淳嗣, 小黒玲(アドバンスト・メディア)

- 医者と患者による日本語臨床会話から発話内容と話者を推定
- LAS: 文字列と発話が医者が患者が示す話者ラベル (i.e., dr, pt) を出力
- MoChA: MoChA で N-best を予測し、LAS デコーダでリスコア + 話者ラベルを予測

モデル	文字誤り率	話者分類精度
LAS	28.22	96.0
MoChA	54.28	-
+ LAS リスコア	50.62	52.0



3-2-15

3-2-15 方言ラベルを補助特徴量とした End-to-End 日本語方言音声認識 Using Dialect Label as Auxiliary Feature for End-to-End Japanese Dialect Speech Recognition

☆今泉遼(都立大), 増村亮(NTT), 塩田さやか, 貴家仁志(都立大)

- ◆標準語で作られた音声認識モデルに地域特有の方言音声を入力すると標準語に比べて認識性能が低下することが知られている。
- ◆解決法の1つに方言と標準語のデータベースを合わせて学習することがあげられる。ただし各方言データより標準語のデータが圧倒的に多いためドメイン依存性が薄まり、認識性能が下がる可能性がある。
- ◆本研究では、方言ラベルを補助特徴量として用いて入力音声のドメイン依存性を高めることで上記問題の解決を提案する。
- ◆Fig.1に実験で用いたTransformerに基づくencoder-decodeモデルを表す。デコーダ部に方言ラベルをトークンとして入力することで提案法を適応する。
- ◆提案法によって方言認識性能(CER)が上がったことをTab.1に示す。

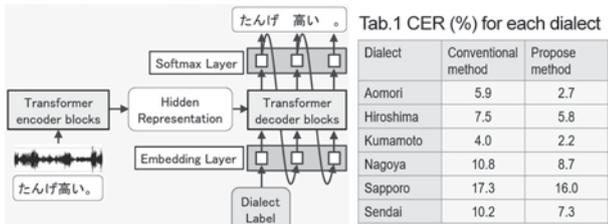


Fig.1 Transformer encoder-decoder model using dialect label as auxiliary feature
Input : Aomori Dialect

3-2-17

3-2-17 Transformer を用いた日本語テキストからの読み仮名・韻律記号列推定 Phonetic and Prosodic Information Estimation From Japanese Text Using Transformer

☆懸川直人, 原直, 阿部匡伸(岡山大・HS 統合科学研), 井島勇祐(NTT)

- ◆日本語 End-to-End テキスト音声合成には読み仮名, 韻律情報の推定が必要となる。本稿では, Transformer を用いた日本語テキストからの読み仮名と韻律記号列の推定を提案する。
- ◆従来システムを用いて約500万文のニュース記事から読み仮名, 韻律情報を推定し, これを正解として学習と評価をおこなった。
- ◆評価尺度として, 読み仮名がすべて正しく推定できた文の割合(S-P-accuracy), 読み仮名・アクセント型がすべて正しく推定できた文の割合(S-PB-accuracy), そして読み仮名・韻律情報がすべて正しく推定できた文の割合(S-PBN-accuracy)を算出した。
- ◆Fig. 1より, S-PBN-accuracyに関してBiLSTM(Bidirectional Long Short-Term Memory)を用いたseq2seqのモデルより提案方式の方が14.9ポイント良い性能が得られることが明らかとなった。

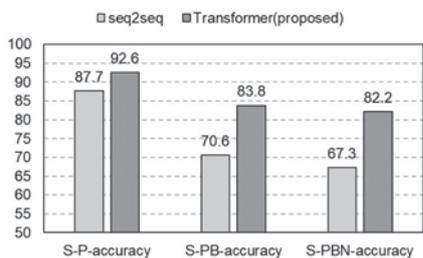


Fig.1: Correct estimation ratio in sentence based evaluation.

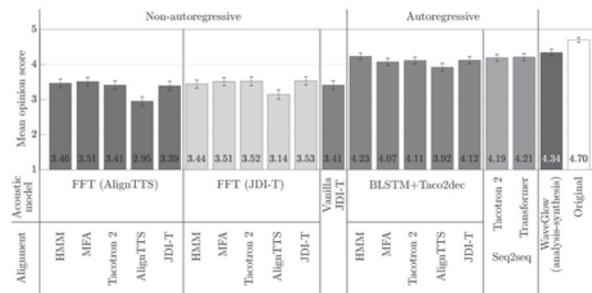
3-2-16

3-2-16 フルコンテキストラベル入力を用いたFastSpeech型ニューラルTTSモデルの比較

Comparison of FastSpeech-based neural TTS models with full-context label input

○岡本拓磨¹, 戸田智基^{2,1}, 志賀芳則¹, 河井恒¹
(¹情報通信研究機構, ²名古屋大学)

- ・FastSpeech : 音素入力(English)
 - ・非自己回帰モデルでありながら高音質, 教師モデルが必要
 - ・教師モデル不要な高音質FastSpeech型ニューラルTTSモデル
- ・AlignTTS : 文字列入力(English)
- ・JID-T : 音素列入力(Korean)
- ・FastSpeech型TTSモデルにおけるアライメント方式は様々
 - ↓
 - ・日本語TTSのためのフルコンテキストラベル入力での比較
 - ・Feed-forward Transformerモデル : AlignTTS, JDI-T
 - ・5種類の音素アライメント方式を比較



・結果 : 高速生成可能であるが非自己回帰モデルには及ばない

3-4-1

3-4-1 耳介各部の微小変形が伝達関数に与える影響の検討

Effects of small deformation of the pinna on its transfer function (16pt)

☆川村海斗, 竹本浩典(千葉工大)

- ◆微小変形による耳介伝達関数(PRTF)の変動量が大きな耳介部位を特定するため, 12人の左耳周辺の形状から輪郭線を等間隔・等数の点群としてトレースして平均耳介を作成した(Fig. 1a)。
- ◆平均耳介から各耳介への変形を薄板スプラインで記述し, 22のトポロジカルに同じ点(変形点)を決定した(Fig. 1abの●)。
- ◆元の耳介形状と, 各変形点から半径3mmにある耳介表面を1ボクセルだけ膨張・収縮させた耳介形状から計算により矢状面上の仰角36方向(10度間隔)のPRTFを求めた。
- ◆変形前後のPRTFで周波数ビンごとに差の絶対値をとって平均してスペクトル距離(Spectral Distance: SD)を計算した。
- ◆各変形点で全耳介・全仰角のSDの平均を求め(平均SD), 平均耳介にマップした(Fig. 2)。
- ◆その結果, 耳介切痕(★)と耳甲介舟(▼)の微小変形がPRTFに最も大きな変動を与えることが明らかになった。

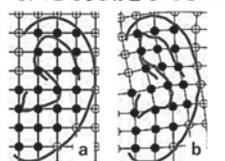


Fig.1 (a) 平均耳介と直交格子 (b) ある耳介に薄板スプラインを用いて写像した格子点

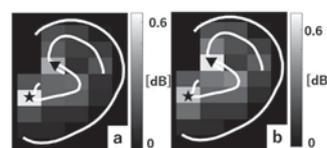


Fig. 2 平均SDの平均耳介へのマップ, (a)膨張時, (b)収縮時

3-4-2

3-4-2 耳介の側面観で陰になる部分が伝達関数に及ぼす影響

Effects of shadow parts of the pinna from lateral view on transfer functions
 ☆金子祐太郎, 竹本浩典(千葉工大)

- ◆ 正中面における音像定位の手掛かりとして、頭部伝達関数の第1,2ピーク(P1, P2)と第1, 2ノッチ(N1, N2)は重要である。
- ◆ 側方から耳介をカメラ等で撮影して計測した形状パラメータから重回帰分析により頭部伝達関数を生成する研究が進められているが、奥行や耳介の陰となる部分は考慮されていない。
- ◆ 本研究では、側方より耳介を撮影した際に陰となる部分が伝達関数に及ぼす影響を検討した。
- ◆ 被験者8名の3次元耳介形状モデル(ORG)を用いて、側面観で耳介の陰となる腔の部分埋めたモデル(SHD)を作成した。
- ◆ ORGとSHDの伝達関数をシミュレーションし、仰角7方向(仰角0~180°, 30°間隔)でP1P2N1N2を抽出した。
- ◆ 耳介の陰となる部分を埋めると、仰角7方向でP1P2N1N2の周波数は上昇し、特にP2周波数は顕著であった。なお、正面方向のN1N2周波数の変動量は半数の被験者で弁別閾を超えていた(Table.1)。
- ◆ また、P1レベルは減少し、N1N2レベルは増加してノッチが浅くなる傾向があった(Table.2)。

Table. 1 Changes in P1P2N1N2 frequency [oct] at vertical angle of 0° for 8 subjects. bold: N1 or N2 >= 0.2 oct.

Subject	F1	F2	F3	F4	M1	M2	M3	M4
P1	0.070	0.130	0.060	0.070	0.220	0.080	0.070	0.07
P2	0.320	0.260	0.190	0.180	0.510	0.190	0.600	0.15
N1	0.270	0.230	0.080	0.080	0.440	0.440	0.040	0.08
N2	0.150	0.210	0.130	0.130	0.340	0.080	0.200	0.10

Table. 2 Changes of P1P2N1N2 level [dB] at vertical angle of 0° for 8 subjects.

Subject	F1	F2	F3	F4	M1	M2	M3	M4
P1	-1.5	-2.0	-1.2	-1.2	-1.0	-2.5	-1.1	-1.7
P2	6.6	-1.1	1.6	6.0	-2.8	-4.5	-2.8	1.6
N1	7.9	-1.3	5.3	-1.6	1.2	0.0	-2.6	-11.5
N2	7.0	6.0	1.0	8.6	12.4	4.2	9.0	5.5

3-4-4

3-4-4 耳介形状による個人の頭部伝達関数のノッチ周波数のカテゴリ推定

Estimation of the category of notch frequency of the individual head-related transfer functions using anthropometry of the pinna

☆西山織絵(千葉工大・院) 飯田一博(千葉工大・先進工)

個人化頭部伝達関数を生成するために、N1, N2が含まれる周波数binのカテゴリを耳介形状から推定する方法を提案し、その推定精度を検証した。以下のことが明らかになった。

- 1) 76 耳介のうち 71 耳介の N1 は中心周波数 6.3 kHz から 8.4 kHz の 6 つの周波数 bin に分布した。同様に 72 耳介の N2 は中心周波数 9.0 kHz から 11.9 kHz の 6 つの周波数 bin に分布した。
- 2) 上昇角知覚に関する N1, N2 周波数の弁別閾によると、6 つの N1, N2 周波数 bin はそれぞれ 2 つのカテゴリに分類できた。
- 3) ナイブな耳介の N1, N2 周波数 bin カテゴリ的的中率はそれぞれ 78.9%, 77.8% であった。

Table 1 Percentages of correct discrimination (PCD) of two categories of N1 and N2 frequency bin using three anthropometric parameters.

	X ₁ , X ₂ , X ₆	Estimated		PCD
		Category 1	Category 2	
N1	Observed	Category 1	20	76.9%
		Category 2	9	80.0%
	Sum			78.9%
	X ₁ , X ₄ , X ₆	Estimated		PCD
		Category 1	Category 2	
N2	Observed	Category 1	21	65.6%
		Category 2	5	87.5%
	Sum			77.8%

3-4-3

3-4-3 両耳間時間差・音圧差による音像の分離知覚の検知限

Detection limit of sound images split by interaural time and level difference
 ☆酒井翼, 森川大輔, 平原達也(富山県立大)

- ◆ 両耳間時間差(ITD)または両耳間音圧差(ILD)によって生じる音像の分離知覚の検知限と検知限に刺激音の周波数帯域が与える影響を明らかにするために、帯域雑音を用いて音像の分離知覚実験を行った。
- ◆ ITD を変化させる条件の音像の分離知覚の検知限は 0.25 ~ 1 kHz の刺激音(ITD-Low)と 0.25 ~ 8 kHz の刺激音(ITD-Wide)に差がなく、1 kHz 以上の周波数帯域の影響を受けない。
- ◆ ILD を変化させる条件の音像の分離知覚の検知限は 2 ~ 8 kHz の刺激音(ILD-High)と 0.25 ~ 8 kHz の刺激音(ILD-Wide)に差があり、2 kHz 以下の周波数帯域の影響で低くなる。

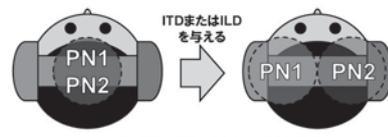


Fig. 1 被験者が知覚した分離音像

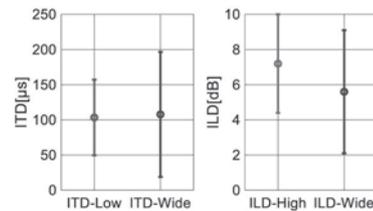


Fig. 2 刺激音毎の検知限

3-4-4

3-4-4 耳介形状による個人の頭部伝達関数のノッチ周波数のカテゴリ推定

Estimation of the category of notch frequency of the individual head-related transfer functions using anthropometry of the pinna

☆西山織絵(千葉工大・院) 飯田一博(千葉工大・先進工)

個人化頭部伝達関数を生成するために、N1, N2が含まれる周波数binのカテゴリを耳介形状から推定する方法を提案し、その推定精度を検証した。以下のことが明らかになった。

- 1) 76 耳介のうち 71 耳介の N1 は中心周波数 6.3 kHz から 8.4 kHz の 6 つの周波数 bin に分布した。同様に 72 耳介の N2 は中心周波数 9.0 kHz から 11.9 kHz の 6 つの周波数 bin に分布した。
- 2) 上昇角知覚に関する N1, N2 周波数の弁別閾によると、6 つの N1, N2 周波数 bin はそれぞれ 2 つのカテゴリに分類できた。
- 3) ナイブな耳介の N1, N2 周波数 bin カテゴリ的的中率はそれぞれ 78.9%, 77.8% であった。

Table 1 Percentages of correct discrimination (PCD) of two categories of N1 and N2 frequency bin using three anthropometric parameters.

	X ₁ , X ₂ , X ₆	Estimated		PCD
		Category 1	Category 2	
N1	Observed	Category 1	20	76.9%
		Category 2	9	80.0%
	Sum			78.9%
	X ₁ , X ₄ , X ₆	Estimated		PCD
		Category 1	Category 2	
N2	Observed	Category 1	21	65.6%
		Category 2	5	87.5%
	Sum			77.8%

3-4-5

3-4-5 頭部伝達関数を用いて提示される骨導音の方向知覚の計測

Measurement of perceived direction of bone-conducted sound using head-related transfer function

○大野悟志, 岩城護(新潟大・自然)

- ◆ 本研究は、気導音と骨導音の方向知覚の違いを明確にすることを目的とした。
- ◆ 音源方向が後方の場合を加え、気導ヘッドホンと骨導ヘッドホンによる全方位での方向知覚特性の測定を試みた。
- ◆ 正面や真後ではどちらも音源方向の違いを正確に判断することができた。しかし、側方での方向知覚特性は気導音も骨導音あいまいになることがわかった。
- ◆ 気導ヘッドホンと骨導ヘッドホンで知覚される方向には違いがあることがわかった。
- ◆ しかし、方向知覚の違いに明確な規則性は見いだせなかった。

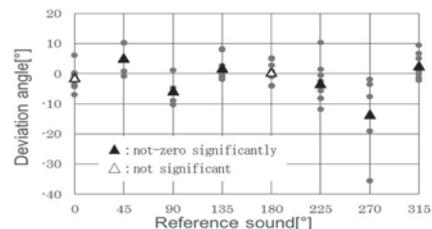


Fig.1 Average deviation of perceived direction when listening to the reference and test sounds (Reference sound was presented by air conduction headphones, while test sound by bone conduction headphones)

3-5-1

3-5-1 S造建物の重量床衝撃音レベルに関する検討 -駆動点インピーダンスに関する実験的検討-

A Study on heavy weight floor impact sound level of steel frame building. -Experimental study on driving point impedance-

○黒木拓, 大脇雅直(熊谷組)

近年、S造のホテルにおける音環境に関する要求性能の向上やS造の共同住宅の増加に伴い、重量床衝撃音遮断性能の予測計算が設計時に求められる場合が増えてきている。S造建物の重量床衝撃音遮断性能を把握するため、実現場において重量床衝撃音レベルおよびインピーダンスレベルの測定を行い、S造建物の重量床衝撃音レベルおよび梁による拘束の影響について検討した。

- ◆大梁のスラブ端部のインピーダンスレベル上昇量は2~8dB程度、小梁は1~4dB程度であった。梁の大きさは同程度であったが、インピーダンスレベル上昇量は異なっていた。鉄骨梁とデッキの納まりの違いによる影響が考えられる。
- ◆全時間応答インピーダンスレベルには共振による大きな低下は見られなかった。梁による拘束の影響が小さかったことが要因の一つと考えられる。
- ◆今回測定を行った建物のスラブは山型デッキプレートスラブ180mmであり、一般的なRC造共同住宅に比べ薄かったが、タイヤの重量床衝撃音遮断性能は $L_{i(Fmax)(H)}$ -45~55であった。スラブの共振によるインピーダンスレベルの低下が小さかったことが要因の一つと考えられる。

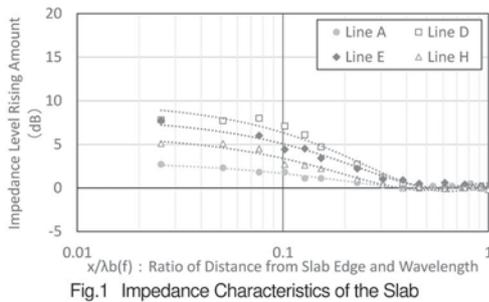


Fig.1 Impedance Characteristics of the Slab

3-5-3

3-5-3 遮音性能を有する長屋又は共同住宅の界壁及び天井の構造方法に関する告示改正に向けた実験的検討

Experimental study for the amendment of the Notification on the structure method of separation wall and ceiling between each unit of row houses or apartment houses

○平光厚雄(国総研)

- ◆長屋又は共同住宅の界壁の遮音性能を規定している建築基準法、建築基準法施行令および同告示が改正され、長屋又は共同住宅の界壁については、小屋裏又は天井裏に達しなくてもよい要件が追加された。
- ◆防火上の規制緩和の観点から、様々な法令の改正が行われており、界壁については、小屋裏又は天井裏に達しなくてもよい「強化天井」に関する規定が追加された。

- ◆吸音材が不要の強化天井でも、遮音の規定により100mm厚以上の吸音材が必須であったため、吸音材を全面に使用しない強化天井の遮音性能に関する実験的検討を行なった。
- ◆開口の数が増える程、性能が低くなる傾向がみられ、天井開口部に設置した吸音材の効果が確認された。そのため、天井裏全面に吸音材を敷設しない強化天井の場合は、天井部分の開口率制限が必要であることがわかった。(Fig.1)

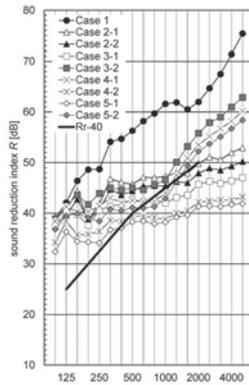


Fig.1: Measurements of sound transmission loss

3-5-2

3-5-2 壁式構造実験室におけるBOX床構造の重量床衝撃音レベル低減量の検討

Examination of reduction of transmitted heavy-weight impact sound of box floor structure in the reinforced concrete wall construction testing device

○富田隆太(日大・理工), 岡庭拓也(戸田建設/日大・理工), △阿部今日子(フリーランス)

- ◆筆者らは、これまで小規模保育所を対象として、床衝撃音の実用的な対策方法について検討を行ってきた。RC造の実建物に、畳付収納家具を連結して、四畳半程度の面積を施工した。その結果、タイヤ加振時では、 L_r -60から L_r -50へと性能が向上した。
- ◆本研究では、これまでの知見を応用し、重量床衝撃音対策効果の高い床仕上げ構造の提案を目的に実験的検討を行っている。
- ◆壁式構造実験室におけるBOX床構造の重量床衝撃音レベル低減量について基礎的検討を行った。

- ◆その結果、右図に示すように、質量付加の程度によるが、密閉時でも全周波数帯域で性能を向上させ、63Hz帯域で7dB向上させることができた。

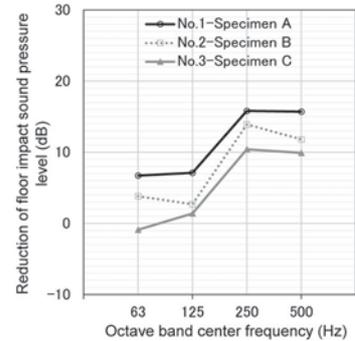


Fig. 1 Reduction of transmitted heavy-weight impact sound due to difference in mass addition

3-5-4

3-5-4 パルプモールドの吸音特性に関する研究 - 背後空気層の影響 -

A Study on sound absorption characteristics of molded pulp : The effect of the back cavity

○小泉穂高, 栗木茂, 山本拓弥(戸田建設)

- ◆大きな背後空気層を有するパルプモールドについて残響室法吸音率を測定し以下の知見を得た。
- (1) 背後空気層が大きくなることで、吸音率のピークは低周波数側にシフトする。
- (2) パルプモールドに中空層を設けた場合、背後空気層の影響により卓越する吸音周波数とは別の周波数帯において、吸音率のピークを発生させることができる。

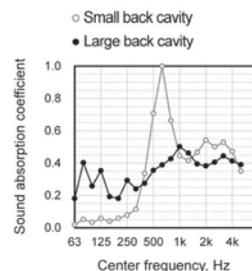


Figure 1: Sound absorption coefficient of molded pulp

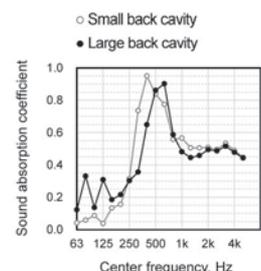


Figure 2: Sound absorption coefficient of double leaf molded pulp

3-5-5

3-5-5 重力方向に弾性率勾配を有する 粒子状材料の吸音特性 —吸音面が鉛直または水平な場合—

Sound absorbing characteristics of granular material with elastic gradient along gravitational direction
- for vertical and horizontal sound absorbing surface -
◎霧羽琢元(竹中技研), 大谷真, 高野靖(京大院・工学研)

- ◆ 粒子状材料の固体部分の弾性率は、重力を受けて鉛直方向に増加する。そのような粒子状材料を吸音材として壁や天井のような鉛直または水平な面に設置することを想定し、面の向きが吸音率に与える影響を調査した。基礎的な条件として、鉛直および水平な吸音面に対して垂直に音波が入射する場合について検討した。
- ◆ 実験結果から、垂直入射吸音率は吸音面の向きによって変化することが示された。加えて、重力方向に弾性率の勾配を有する粒子状材料の垂直入射吸音率を計算するモデルを示した。

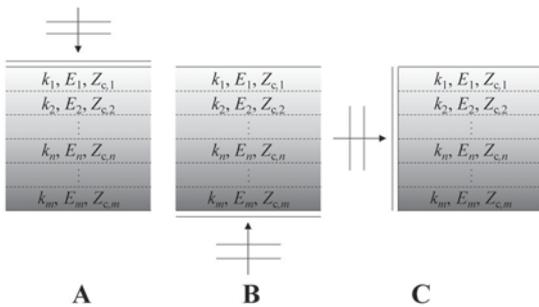


Fig.1 Multi-layered model of granular material with elastic gradient along gravitational direction. (sound incidence [A] from above to horizontal surface, [B] from below to horizontal surface, [C] from side to vertical surface)

3-5-7

3-5-7 減衰密度を用いた残響室法吸音率測定 —低周波域における分布形状の分析—

Measurement of sound absorption coefficient in a reverberation chamber using damping density —analysis of the shape of distribution in low frequency—
◎後藤耕輔, 山田祐生(竹中技研)

- ◆ 筆者らは残響室法吸音率測定の精度改善のために、減衰密度より算出した残響時間を利用する手法を検討している。本報では減衰密度の形状を混合正規分布モデル(GMM)で近似する手法を提案する。
- ◆ 本報では下記の2項目により、提案手法の検討を行った。
 - 室内インパルス応答モデルを対象とする基礎的な検討
 - 実際の残響室法吸音率の測定による検討
- ◆ 両検討において正規分布での近似(従来手法)よりもGMMでの近似(提案手法)は減衰密度の形状と良い当てはまりを見せた。(Fig. 1)
- ◆ 残響室法吸音率の測定では空室時と比較して試料設置時における減衰密度の多峰性が明確となった。
- ◆ 提案手法と従来手法Bの結果はほぼ同値であり、従来手法Aよりも125~250 Hzにおいて理論値に近づいている。(Fig. 2)

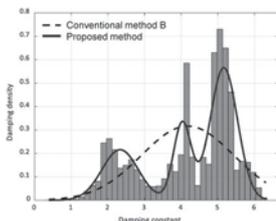


Fig.1 Result of approximating the damping density.

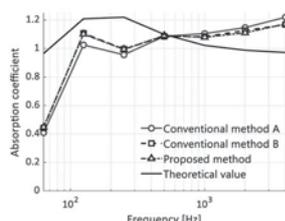


Fig.2 Measured result of sound absorption coefficient.

3-5-6

3-5-6 3軸音圧—粒子速度センサーを用いた音響測定に関する基礎的検討

Basic investigation on acoustic measurement using three-dimensional pressure-velocity sensor
◎富来礼次, 大鶴徹, 岡本則子(大分大・理工),
△藤田和宏(大分大院・工学研)

- ◆ 音圧と3軸方向の粒子速度を測定可能な音圧—粒子速度センサ(以下、3軸センサ)を用いた音響測定のための基礎的検討として、センサの校正に関する実験を行った。
- ◆ 既報で使用してきた音圧と1軸方向の粒子速度を測定可能な音圧—粒子速度センサ(以下、1軸センサ)の校正で用いてきた音響管を流用し、3軸センサの校正を行った。
- ◆ 得られた補正值を用いて、音響管内の音圧および粒子速度を測定し、校正が効果的に行われていることを確認した。また、アンサンブル平均を利用した材料のin-situ測定法にも適用し、3軸センサで得られた測定値相互および1軸センサで得られた測定値と比較した。

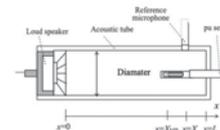


Fig.1: Acoustic tube for calibration and insertion length X_{tip} of pu sensor

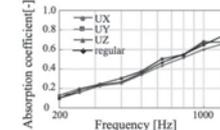


Fig.2: Absorption coefficients of glass-wool measured by PU regular and USP regular

3-5-8

3-5-8 残響室内の温度勾配と 残響時間の関係について

Relation between temperature gradients and reverberation time in reverberation room
◎豊田恵美, 横山栄, 杉江聡(小林理研)

- ◆ 残響室内に生じる高さ方向の温度勾配の変化が残響時間に与える影響について検討するため、空室状態の残響室において、1年を通して温湿度を計測するとともに、30分毎に残響時間の定期測定をおこなった。
- ◆ 低周波数帯域においては、温度勾配の変化にともない残響時間も変化する。温度勾配が大きくなると残響時間は短くなり、温度勾配が小さくなると残響時間は長くなる傾向がある (Fig.1)。
- ◆ 温度勾配の変化による残響時間の変化が、残響室法吸音率の測定結果に与える影響についても検討した。

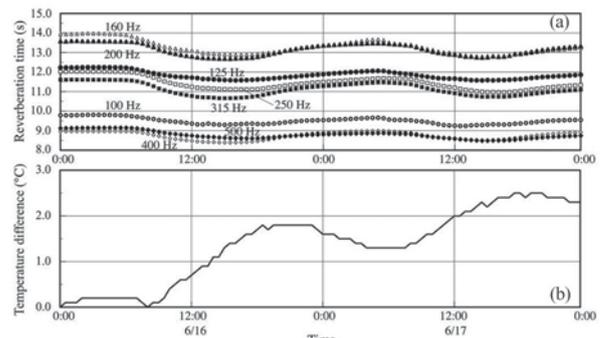


Fig.1: Relation between temperature gradients and reverberation times. (a) Measurement results of reverberation times (b) Calculation result of vertical temperature difference

3-6-1

3-6-1 非常用屋外拡声システム調査研究委員会の設立主旨と活動方針

Launching and action plan of the research committee on outdoor acoustic mass-notification systems for emergency use

○佐藤洋(産総研), 栗栖浩浩(TOA), 佐藤逸人(神戸大)

- ◆非常用屋外拡声システム調査研究委員会(以下当委員会)は、災害等非常時屋外拡声システムのあり方に関する技術調査研究委員会(以下前委員会)の後継組織として、委員の公募を経て2019年11月に日本音響学会に設置された。
- ◆当委員会の活動の目的は、屋外拡声システム全般に対する調査・研究を進め、その成果を用いて前委員会が作成した日本音響学会が公開した「災害等非常時屋外拡声システムの性能確保のための規準(第1版)」(以下ASJ屋外拡声規準)を拡充するとともに、この規準を社会に普及させる活動を通して特に運用者の視点を取り入れることである。
- ◆具体的には以下の活動を行っている。
 - (1) 音源に関連する研究の推進: 言語学からみた避難伝達文の分析、単語親密度やポーズ長の影響、緊迫感のような非言語情報の影響、妨害音の存在する音場に適応した発話方法といった様々な観点からの研究を推進し、ASJ屋外拡声規準を拡充する。
 - (2) 運用者の視点を取り入れるための活動: 屋外拡声システムの発注者および運用者となる自治体関係者などにもASJ屋外拡声規準を普及させる活動を順次進める。
 - (3) 測定データの収集: 特に多種多様な要因が存在するスピーカから聴取者までの経路について、実測データを蓄積することにより、それらの要因を屋外拡声システムの設計で考慮する方法を検討する。

3-6-3

3-6-3 (招待講演)災害時のためのアナウンス訓練システムに関する研究の現状

Current status of research on announcement training systems for disasters

○土田義郎, 高野佐代子(金沢工大)

kw:防災無線、アナウンス、訓練、避難、正常性バイアス

防災行政無線による屋外拡声子局からの音声放送は地域住民への避難の呼びかけが可能な重要な手段である。アナウンスは、きちんと聞こえるだけでなく、適切な緊迫感を与えて、避難行動を実際に誘起するアナウンスであることもまた求められている。ここでは災害時のアナウンスのトレーニングシステムとして、2016年に発表した「AEGIS」と、開発中の「防災ボイストレーナー」について概観する。

- * AEGIS: MAX7 ベースの屋外伝搬音のシミュレーションを行うソフト。自分の発した音声を受聴者にどのように聞こえているかを可聴化する。ロングパスエコーやノイズによって聞きづらいということに気づき、より適切な発話を心掛けるように促すことを意図している。
- * 防災ボイストレーナー: MATLAB ベースの音声分析を行うソフト。自分の発した音声の「高さ」「大きさ」「速さ」を確認できる。アドバイスに基づいて、発話と分析を繰り返すことができる。



Fig.1 AEGIS



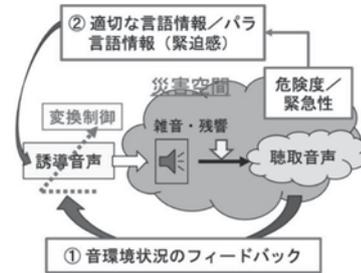
Fig.2 Voice Trainer for Emergency

3-6-2

3-6-2 確実に情報を伝える音声避難誘導システム ~避難誘導にふさわしい音声~

○赤木 正人 (JAIST)

- 災害回避に向けた避難誘導時に、音声により避難誘導情報を誰にでも確実に伝える、すなわち
 - ① 災害空間の音環境および避難者の状況がどうであろうと避難誘導音声がよく聞き聞こえる。
 - ② 避難誘導音声から災害の危険性を察知し逃げる気になる。
 ような避難誘導音声を提示するための仕掛けが必要である。
- 筆者らのグループでは、この問題の解決に向けて次のシステムを提案している。
 - (a) ヒトが残響雑音環境で話すのと同様に、音声を常にモニタリングし雑音残響に負けないように音声の重要な特徴を強調し提示する
 - (b) 聴取者に対してより注意喚起が行えるように、状況に合わせて音声のバラ言語情報を適応的に付加する。



講演では、システムに関する現在までの到達点および応用可能性を紹介する。

3-6-4

3-6-4 積雪寒冷地の漁村集落における屋外拡声器を用いた防災行政無線の長期測定

Long-term Field Survey of in the Disaster Management Radio System Installed in a Fishing Village in a Snowy Cold Region

◎飯泉元気, 廣田誠一(道経研北総研)

- ◆本研究は積雪寒冷地において防災行政無線を有効に活用するための基礎資料の整備を目的としている。
- ◆本報では津波被害が予想されている北海道の漁村集落をフィールドに、防災行政無線の音の伝わり方を長期測定した結果を報告する。
- ◆積雪期の測定として、自治体が日常的に防災行政無線を用いて放送している音楽放送(「エーデルワイス 音声なし(12時)」及び「夕焼け小焼け 音声なし(17時)」を対象とした調査の結果、ベクトル風速に正の相関が見られ、10~15 dBの減衰が見られた(Fig.1)。
- ◆Swept-Sine 信号を毎15時に放送する試験を、積雪期と非積雪期に実施した結果、単位時間当たりのエネルギーとベクトル風速の関係に、音楽放送と同様に正の相関が見られた。(Fig.2)
- ◆気象条件によっては集落内にシャドーゾーン(死角)を生じる危険性があることが明らかになった。

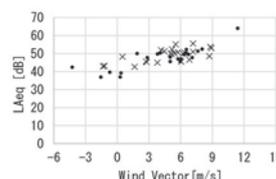


Fig. 1 Relationship between L_{Aeq} and wind vector

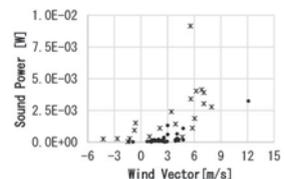


Fig. 2 Relation between vector wind speed and energy of impulse response

3-6-5

3-6-5 バイノーラル録音による 屋外拡声スピーカ音の主観評価

Subjective evaluation of outdoor PA system using binaural recording

◎小林洋介(室蘭工大), 野口啓太(室工大/TOA),
△鳴野寛(室蘭工大), 栗栖清浩(TOA)

- ◆屋外拡声システムの放送品質は日常生活にとどまらず、避難判断などにも利用される重要な情報源であり高品質で高明瞭度であることが望ましい。しかし、拡声品質の主観評価のために実際の拡声フィールドに多数の被験者を集めて評価を行うことは現実的ではない。
- ◆この問題を解決するために実際の拡声音声をダミーヘッドを用いたバイノーラル録音による音源を用いて主音声の聴き取りにくさ尺度の主観評価を行った。
- ◆本研究では北海道東部の牧草地において5条件の拡声パターンを設定し、3種のスピーカ(従来型のレフレックスホーンスピーカ単体、ホーンアレイスピーカ、スリムタイプアレイスピーカ)の合計15条件を比較した。
- ◆録音した音源を防音室内でヘッドホンを用い、同時に計測した騒音計で記録した音圧レベルになるように提示した被験者24名分の評価結果を分析した。
- ◆評価結果を分散分析したところ、バイノーラル録音した音源で比較したい録音地点とスピーカ種について、特に従来型のレフレックスホーン型スピーカと新しい2種のアレイスピーカの違いには有意差があり、拡声条件の品質序列について評価可能であることが示された。

3-6-6

3-6-6 都市街区における屋外拡声システムの 最適配置に関する基礎的研究 -音線法による検討-

A basic study on optimal arrangement of outdoor public address system in urban area using sound ray-tracing method.

○佐藤逸人(神戸大院・工学研)

- ◆都市のような高層建築物が密集している地域において屋外拡声システムを用いた場合、建築物間で音の反射が繰り返されることにより残響音が生じ、さらに音声が聞き取りにくくなると予想される。
- ◆本研究ではモデル化した都市街区を対象とした幾何音響シミュレーションにより、子局スピーカの設置高さ、指向性、複数の子局スピーカを同時に用いる場合の遅れ時間が拡声音の明瞭性に及ぼす影響を定性的に検討した。
- ◆複数の全指向性スピーカから同時あるいは遅れ時間をつけて音を再生すると、受音点によってはエコー障害が発生してしまう。一方、道路方向に軸を向けた狭指向性スピーカから適切な遅れ時間で再生をすれば、広い範囲で拡声音の明瞭性を確保できる(Fig. 1)。

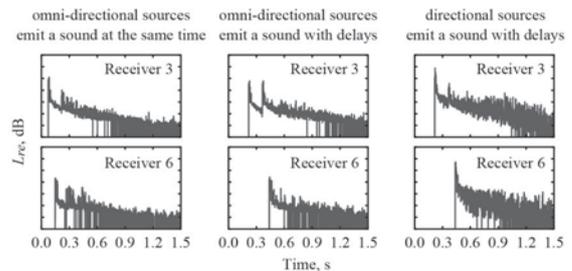


Fig. 1: Examples of echo time pattern obtained in the present study.

3-7-1

3-7-1 等音速層で覆われた浅海域における水中 音速構造の変化による音場への影響

Influence on sound field by the change of underwater sound speed structure in shallow-water covered with iso-velocity layer

○鶴ヶ谷芳昭(山陽精工), 菊池年晃(防衛大), 水谷孝一(筑波大)

- ◆基盤となる海底層の上に等音速の堆積が生じる場合があり、波長に比較して比較的薄い層が形成される。等音速層がある場合、当然、堆積物上面からの音波の反射・透過を生じる。その結果、等音速層(ML)の有無、等音速層の厚さによって水中の音場は影響を受ける。
- ◆Fig. 1に周波数600 Hz、音源深度100 m、海底上のMLの厚さが3 mの場合の水中の音速構造の違いによる音場を示す。左図は水中も等音速の場合の音場、中央の図は海面から海底まで音速が10 m/s増大する正傾度(positive gradient)の場合、右図は10 m/s減少する負傾度(negative gradient)の場合である。等音速の場合、音波は直線で伝搬するが、正傾度の場合、音波は上向きの屈折、負傾度の場合、下向きの屈折で伝搬する。その結果、同じ角度で音源から放射された音波が海底に到達したときの角度が異なる。したがって、水中の音速構造の違いによるML上面での音波の反射・透過も異なる。ML上面での反射が卓越するが、MLの厚さが減少する(1m以下)と海底層の影響が増大する。

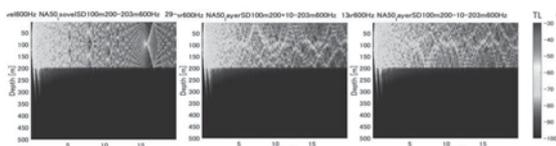


Fig.1: Comparison of sound field by the difference of seawater sound speed structure. IVL: 3 m, freq.: 600Hz. left: iso-velocity middle: positive gradient (+10 m/s) right: negative gradient (-10 m/s)

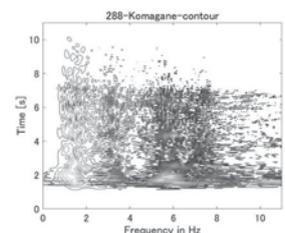
3-7-2

3-7-2 活断層振動の時間周波数解析と その方位性

Time-frequency analysis of active fault vibration and its orientation

○菊池年晃(防大), 水谷孝一(筑波大)

- 我々は、地震の活断層内の振動構造を音響学的知見から調べている。活断層からの振動放射は方位的に双極性である。その中心軸上には活断層内の振動が反映される点が存在する。今回は、2012年1月28日に富士山付近で発生したM5.4の地震に対して検討した。まず、観測信号を特異値分解法で処理し、更にタイムリバーサル法を適用して伝搬の地域性を低減した。その結果、分布特性を二分する明確な双極性が得られた。その軸に位置する駒ヶ根で受信した信号に時間周波数解析を適用し、その結果を下図に示す。横軸は周波数で、縦軸は時間である。主な周波数成分は1.5, 3.5, 5.9, 7.8及び10.5 Hzである。そして、それら振動成分はそれらの周波数を変えずに時間的に継続している。各成分とも他成分との変調はなく一定の周波数を維持しており、いわゆる音源の本来の振動を表す縦型パターンである。駒ヶ根を中心とした方位角210°から10°の範囲では周波数は異なるがほぼ縦型パターンである。一方、方位角が10°から210°では散乱パターンになる。これは活断層内の進行波と反射波が混合した状態を表している。即ち、双極性特性の緩やかな分布が形成される主因である。



3-7-3

3-7-3 JAMSTEC 海底ケーブル型 観測システム・レガシーデータの救出

Legacy data rescue of cabled seafloor observatories in JAMSTEC

○岩瀬良一(JAMSTEC)

- ◆ 深海底における変動現象の多面的な長期連続観測を目的として、1993年に設置した「深海底総合観測ステーション」(初島沖ステーション)以来、海洋研究開発機構(JAMSTEC)では複数の海底ケーブル型観測システムを設置・運用してきた。このうち、アナログテープや2.6 GB DVD-RAMカートリッジに収録されたハイドロフォンデータを初め、古いデータについては、再生装置の入手が困難となり、未解析のデータも多い。これらの中には深海環境や鯨類鳴音等の変動の解明にも資するデータが含まれている可能性もある。筆者は長期観測という当初の目的に沿うべく、これらのレガシーデータの救出を試みている。
- ◆ 初島沖ステーションのハイドロフォン信号は、1999年7月までTEAC製カセットデータレコーダXR-5000WBを用い、記録媒体のVHS160分テープの走行速度を0.6 cm/sとし、テープ1本に1週間の連続記録を行っていた。約300本のテープに収録されたが、再生手段が当該レコーダに限られることに加え、関係者の異動等により収録形態の情報も失う状況となり、ほとんど再生の機会がなかった。近地地震波形等の良好な再生結果が今回得られたものの、再生作業開始直後に当該レコーダの故障に直面した。使用可能な同レコーダに関する情報があれば筆者までお寄せ頂けると幸いである。
- ◆ 1999年に設置された北海道釧路十勝沖観測システムのハイドロフォン信号については、2.6GB DVD-RAMカートリッジに収録されていたレガシーデータのほぼ全てを今回救出した。

3-7-5

3-7-5 超音波後方散乱特性解析による 赤血球集合体サイズ推定に用いる 周波数帯域の影響の検討

Influence on frequency range in estimation of red blood cell aggregated size by analyzing ultrasonic backscattering characteristics

☆東山京平, 深瀬晶予, 森 翔平, 荒川元孝(東北大),

八代 諭, 石垣 泰(岩手医大), 金井 浩(東北大)

- ◆ 非侵襲的かつ定量的な血液性状評価のため、超音波計測により求めた赤血球の散乱特性を単一球散乱体の理論特性と整合することで赤血球集合体サイズを推定する。その際に用いる周波数帯域の変化によりサイズ推定結果に生じる影響を検討した。
- ◆ 中心周波数40 MHz(周波数帯域27-45 MHz)に加えて、中心周波数30 MHz(周波数帯域19-40 MHz)の超音波プローブを用いた計測により、赤血球集合体を模擬した直径20 μmのマイクロ粒子のサイズを推定した。
- ◆ 30 MHzのプローブを用いて計測された粒子の散乱特性は、より高精度に理論特性と整合され(Fig. 1)、より真値に近い値に推定された。

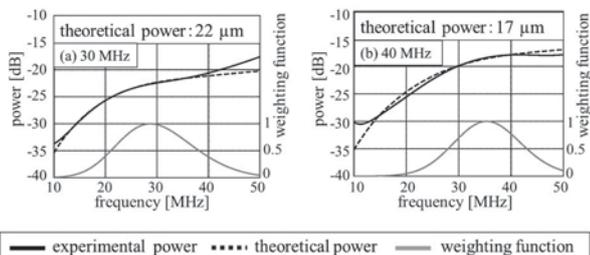


Fig.1: Fitting results of measured and theoretical power

3-7-4

3-7-4 後方散乱特性解析による リンパ浮腫の重症度分類の評価

Evaluation of classification performance of lymphedema using ultrasonic backscatter analysis

○大村眞朗(富山大・工, 千葉大・CFME), 西東若菜(千葉大・院融合),

秋田新介(千葉大・院医), 吉田憲司, 山口匡(千葉大・CFME)

- ◆ ヒト両下肢を対象として、一般的な臨床検査によりリンパ浮腫の重症度(ISLステージ)を分類し、後方散乱特性解析による定量評価指標を用いた分類法と比較した。
- ◆ リファレンスファントム法により減衰および後方散乱係数を、振幅包絡統計解析法によりhomodyned-K(HK)分布の統計パラメータを算出した。後方散乱係数に対してガウシアン散乱モデルを仮定した数式による近似を行い、散乱体サイズ・音響濃度も推定した。
- ◆ 評価指標は、真皮厚、皮下組織の結合織割合、真皮および皮下組織の減衰係数、散乱体サイズ、音響濃度、HK分布の統計パラメータとし、ロジスティック回帰による分類(ISLステージ0対1以降)を行った。
- ◆ 音響濃度を用いた際の分類結果例が図1、評価指標を個別および複合的に用いた際の分類器の性能が表1であり、評価指標を複合的に用いることで、感度71.8%、特異度100%を示した。

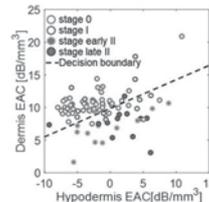


Fig. 1: Scatter plots and decision boundary of effective acoustic concentration (EAC).

Table 1: Results of classification performance

	AUC	Sensitivity	Specificity
Dermal thickness	0.893	66.7	100
Fiber ratio	0.737	64.1	79.2
Attenuation coefficient	0.600	23.1	93.8
Scatterer size	0.752	66.7	81.3
Acoustic concentration	0.782	56.4	95.8
HK-μ	0.774	74.4	79.2
HK-k	0.711	87.2	56.3
All parameters	0.933	71.8	100

3-7-6

3-7-6 双方向超音波加振による 音場・変位分布計測と 異方性対象物の粘弾性特性推定

Measurements of acoustic field and displacement distribution to estimate anisotropy in viscoelastic properties for an object using dual ultrasound excitation

☆福澤徳美, 森 翔平, 荒川元孝, 金井 浩(東北大)

- ◆ 我々の研究グループでは、Fig. 1で示すような、双方向超音波加振を用いた粘弾性推定手法を研究開発している。
- ◆ 双方向から超音波を照射することで生じる音響放射力の分布ならびに等方性・異方性ファントム表面上の変位分布を計測した。音響放射力の分布では、指向性と局所性を持つことを確認した。
- ◆ ウレタンゴム束をウレタン樹脂で固めた異方性ファントムに関して、変位 $D(\omega)$ と音響放射力 $F(\omega)$ の周波数特性を計測し求めた伝達関数 $H(\omega) (= D(\omega)/F(\omega))$ を、Fig. 2に示す。計測方向の違いによって剪断弾性率の推定値が異なり、本手法によって粘弾性とその異方性を推定できたと考えられる。

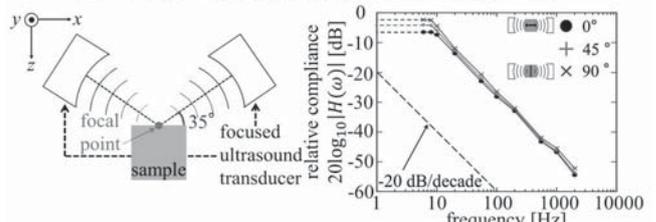


Fig.1: Dual ultrasound excitation. Fig.2: Estimation of viscoelastic properties of an anisotropic object.

3-7-7

3-7-7 単一プローブによる橈骨動脈粘弾性特性変化の計測における Pulse Transit Time (PTT) 法の適用に関する基礎検討

A basic study on application of pulse transit time method to measurement of viscoelastic properties of radial artery with a single ultrasound probe

☆庄司悠人, 森翔平, 荒川元孝,

△大庭茂男(東北大), 小林和人(本多電子), 金井浩(東北大)

- ◆我々は動脈硬化症の早期診断を目指し、血流依存性血管拡張反応(FMD)中の血管の粘弾性計測を行ってきたが、従来の手法では血圧の絶対値が計測条件の変化に影響されやすいという課題があった。その解決のため、血圧計測への脈波伝播時間(PTT)法の応用を検討した。
- ◆PTT 法による血圧計測では取得される電圧の絶対値にかかわらず収縮期血圧 p_{sys} 、拡張期血圧 p_{dias} を推定できる。本報告では、PTT 法で推定した p_{sys} と p_{dias} により、計測回ごとに血圧の校正を行った。
- ◆Fig. 1 に FMD 計測中の血圧変化を提案法と従来法それぞれで校正した結果を示す。提案法による校正によって、駆血解放後の血圧が、直後低い値をとった後、次第に安静時の値に近づいていく様子が見られ、推定精度が向上したと言える。

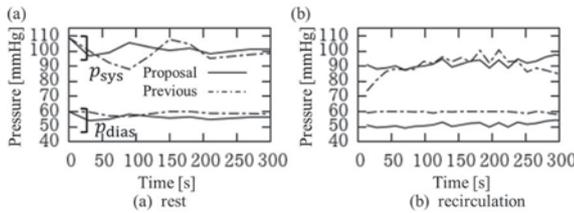


Fig. 1 Changes of systolic and diastolic blood pressures during flow mediated dilation

3-7-9

3-7-9 筋線維の高分解能観察を目指した Wiener filter 設計過程における伝達関数の推定法に関する検討

Examination on transfer function in Wiener filter for higher resolution observation of muscle fiber

☆川又健太, 森 翔平, 荒川元孝, 金井 浩(東北大)

- ◆筋線維の高分解能な観察を目指し、受信信号を広帯域化するための Wiener filter の設計・適用を行っている。本報告では、その設計過程において推定する超音波送受信系の伝達関数 $H(\omega)$ の位相について検討を行う。計測対象は腓腹筋とした。
- ◆ハイドロフォンで計測した伝達関数の位相 $\angle H_h(\omega)$ により設計した Wiener filter と、伝達関数の位相 $\angle H_r(\omega)$ に一様乱数を仮定して設計した Wiener filter をそれぞれ適用した結果を比較する。また、位相に注目するため、フィルタ適用後の振幅特性を全周波数帯域で 1 とした。
- ◆それぞれのフィルタを適用した後の中心ビームの振幅特性と位相特性、および B モード画像を Fig. 1, 2 に示す。これらより、振幅特性が計測対象の信号情報を失っても、位相特性が計測対象の信号情報を持つことで、腓腹筋の画像情報を取得できることが確認でき、伝達関数の位相の推定の重要性が再認識された。

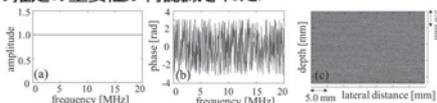


Fig. 1: Results of applying Wiener filter using $\angle H_h(\omega)$ as phase of transfer function.

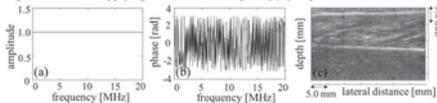


Fig. 2: Results of applying Wiener filter using $\angle H_r(\omega)$ as phase of transfer function. (a) amplitude, (b) phase, (c) B-mode image)

3-7-8

3-7-8 多周波位相差トラッキング法における検出最大速度に関する検討

Measurable maximal velocity in multiple-frequency phased-tracking for measurement in the heart wall

☆小原優, 森翔平, 荒川元孝, 金井浩(東北大)

- ◆【目的】心臓壁の局所微小速度推定を目指し、多周波位相差トラッキング法(提案法)における、検出最大速度と窓幅に関する検討を行う。
- ◆【原理】単一周波数 f_c の位相差を用いた速度推定法(従来法)では、フレーム間の変位が周波数 f_c 成分の半波長を超えると、エイリアシングが生じる。提案法では、フレーム間の変位が窓幅の半分以下であれば、エイリアシングが生じずに速度を推定できる。
- ◆【結果・考察】従来法と提案法における検出最大速度とファントム実験(50 Hz 正弦波)によって得られた速度推定の誤差を Fig. 1 に示す。提案法では、クロススペクトルの帯域に対して十分な周波数分解能を持つ窓幅であれば、従来法でエイリアシングが生じる速度でも計測できた。

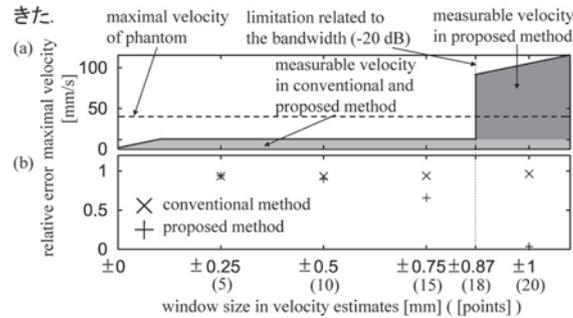


Fig. 1: (a) Measurable maximal velocity in multiple-frequency phased-tracking, (b) relative error in velocity estimates.

3-7-10

3-7-10 超音波を用いた生体内の音速推定を目指した対象散乱体の大きさの影響に関する検討

Study on effect of target scatterer sizes in ultrasonic sound velocity estimation

☆中山碧惟, 森翔平, 荒川元孝, 金井浩(東北大)

- ◆我々は、プローブ各素子における点散乱体からの散乱波の受信時刻差を用いた音速推定手法を開発している。これまでに、提案する音速推定法では、散乱体径が大きくなるにつれて音速が速く推定されることが明らかになっている。そこで、本報告では、この原因について検討するため、模擬実験を行い、水槽実験と比較した。
- ◆散乱体が大きさを持つ場合は、超音波が点ではなく、面で散乱され、面上の複数点からの散乱波が重なる。このとき、点散乱体の場合よりも早く受信される散乱波が存在する。この散乱波の影響により、中心素子以外では、点散乱体の場合よりも受信時刻が早くなり、音速が速く推定されたと考えられる。そこで、散乱体の表面上の複数の点からの散乱波の重ね合わせを模擬した模擬実験を行った。
- ◆水槽実験と同様の計測条件で模擬実験を行い、音速を推定した。水槽実験と比較した結果、水槽実験と同様に散乱体径が大きくなるにつれて、音速が速く推定された。
- ◆今回の実験では、ほとんどの径において、水槽実験での音速推定結果が模擬実験での音速推定結果よりも遅く推定された。今後は、模擬実験での音速推定値に比べて、水槽実験での音速推定値が遅く推定された原因について詳細に検討する。

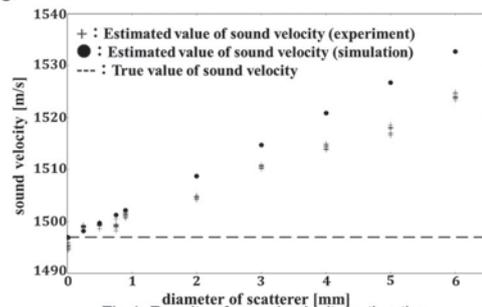


Fig. 1: Results of sound velocity estimation.

3-7-11

3-7-11 硬膜外麻酔補助のための骨表面の描出を目指した超音波反射特性に関する理論的・実験的検討

Theoretical and experimental investigations on ultrasound reflective properties for depicting bone surface to assist epidural anesthesia

☆橋本拓実, 森 翔平, 荒川元孝, △大西詠子, △山内正憲, 金井 浩 (東北大学)

- ◆硬膜外麻酔には、「穿刺位置の正確な特定が難しい」という課題がある。麻酔の補助として用いられる超音波診断装置によっても、B-mode 断層像ではその位置を容易には特定できない。この問題を解決するため、超音波プローブの各素子で得られる素子データを用いた胸椎描出法を目指している。
- ◆胸椎は背中表面に対して傾いていることから、傾いた面反射体に対する反射特性を検討する必要がある。超音波診断装置を用いて、骨のような面反射体(アクリル板)からの受信信号を計測し、プローブ表面に対する反射体の傾き ϕ の変化による反射特性の変化を検討した。基礎実験と同様の条件の模擬実験も行った。B-mode 像のある1点において、点散乱体を仮定して、各素子の包絡線振幅から反射の角度特性 $R(\theta)$ を求めた。
- ◆反射特性の計測値を Fig. 1(a)、模擬実験値を Fig. 1(b)に示す。計測値と模擬実験値ともに、 ϕ が変化すると反射特性が 2ϕ 左にシフトした。今後は、取得した反射特性から対象物の角度を推定する方法を検討していく。

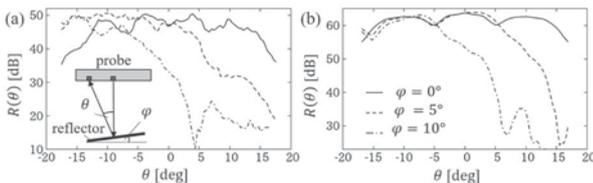


Fig. 1: Results of reflection property. (a) measurement, (b) simulation

3-7-13

3-7-13 複合現実技術と移動マイクロホン計測による音波面の可視化

Visualization of sound wavefront using mixed reality and moving microphone

☆渡辺靖明, 池田雄介(東京電機大), 片岡優太, 及川靖広(早大理工), 小坂直敏(東京電機大)

- ◆これまで、我々は、複合現実技術と手持ちマイクロホンを用いた音場の可視化システムを提案してきた。
- ◆特に、近年、時間変化を可視化する目的で空間インパルス応答の計測/可視化システムを提案しており、本発表では、提案システムとその実測例を紹介する。
- ◆実験では、マイクロホンを移動しながら計測した信号から、グリッド上のインパルス応答を復元し、Microsoft HoloLens を用いて実空間上にCGアニメーションとして可視化した (Fig.1)

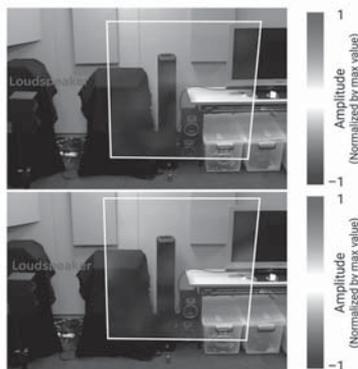


Fig. 1 Examples of MR visualization of impulse responses (Upper part: 1.67ms, Lower part: 2ms)

3-7-12

3-7-12 偏光高速度干渉計を用いた水中の音波の可視化

Visualization of underwater sound wave using high-speed polarization interferometer

☆吉田颯汰, 及川靖広(早大理工), 谷川理佐子(パナソニック)

- ◆水中の音波の測定は一般的にハイドロホンを用いて行われるが、扱いに注意が必要であることや音場が乱れる恐れがある。
- ◆偏光高速度干渉計を用いた音場の計測は、音場に光を照射し音波の情報を得るため、計測音場に非接触に計測が可能である。
- ◆非常音場の計測も可能であるため、水中の音波の計測にも有効であると考えられる。
- ◆提案手法を用いた実験によって、水中の音波の可視化が可能であることが確認された (Fig.1)。

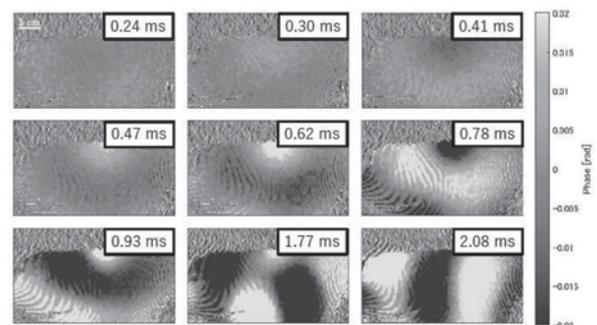


Fig. 1: Underwater sound wave visualization image

3-7-14

3-7-14 空中超音波フェーズアレイと複数受信器を利用した金属薄板中の欠陥イメージング

Defect Imaging in Metallic Thin Plate Using Aerial Ultrasonic Phased Array and Multiple Receivers

☆清水鏡介, 大隅歩, 伊藤洋一(日大理工)

- ◆空中超音波フェーズアレイ(Airborne Ultrasound Phased Array : AUPA)と弾性波源走査法を組み合わせた非破壊検査方法について研究を進めている。
- ◆本報告では、AUPA と複数受信器を用いた金属薄板内の欠陥のイメージングを試みている。
- ◆結果より、センサ数が増加するにつれて、全体的にアーチファクトが低減され、減肉部がより精細に可視化されていることが確認できる。

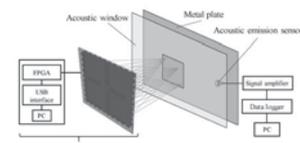


Fig.1 Experimental device

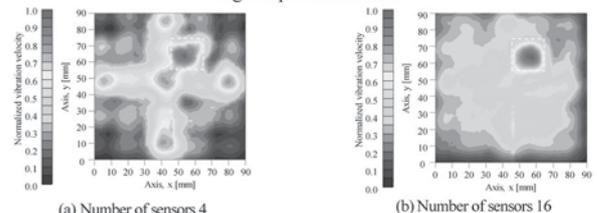


Fig.2 Averaging result of peak distribution

3-7-15

3-7-15 音波照射加振による 非接触浅層地中探査法の基礎研究

-珪砂土壌を用いた場合-

Basic study of noncontact shallow underground exploration method by acoustic irradiation induced vibration, -When using silica sand soil-

○杉本恒美, 杉本和子, 中川 裕, 上地 樹(桐蔭横浜大・院),
△歌川紀之, 黒田千歳(佐藤工業)

- ◆比較的浅い地中を探査する手法としてはレーダ探査, 磁気探査, 電気探査および弾性波探査といった方法が主に用いられている。
- ◆しかしながら, どの手法も計測用機材を地表面に接触もしくは近づける必要があり, 遠隔から非接触で探査を行うことは困難であった。
- ◆一方で, 音波照射加振を用いた非接触音響探査法では, 測定面の振動速度分布を計測できれば良いので, 10 m 以上の離隔があったとしても, 非接触で浅層地中探査ができる可能性がある。
- ◆今回は実際に音波照射加振により非接触浅層地中探査が可能かどうかについて, 音源を固定台に載せた状態での基礎的な技術検討を行い, 粒径の揃った珪砂土壌では検出可能であることを確認した。

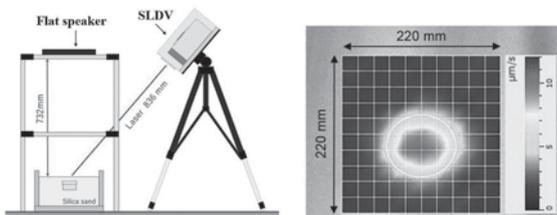


Fig.1: Experimental Setup. Fig.2: Vibration velocity distribution (149 Hz). White Circle indicates the position and size. (Hollow container, dia.100mm, depth 50mm).

3-7-17

3-7-17 TSP 信号とマルチマイクロフォンによる打音 検査診断の精度向上

Study of hammering test improved by time-stretched pulse and multi-microphone

☆花房万由子(滋賀大),市川治(滋賀大)

- ◆社会インフラの点検のために, 打音検査が広く行われている。
- ◆打音検査を, 熟練技術者の聴覚に頼るのではなく, 接触型のスピーカとマイクロフォンを用いて, コンピュータによって異常を診断する方法が開発されている。
- ◆本報告では, コンピュータによる診断の精度を向上するため, 以下の2つの手法を試す。
 - 再生音として打撃を模した音の代わりに Time Stretch Pulse (TSP) 信号を用いる
 - マイクロフォンを複数 (今回は2個) に増やす
- ◆ただし, ひび割れなどの異常のある金属やコンクリートを入手することができなかったため, 裏に間柱がある場合と無い場合の内装壁を対象とした。
- ◆機械学習のモデルとしては, 時間解像度を高めた音声認識の音響モデルのフレームワークを適用した。

3-7-16

3-7-16 非接触音響探査法のための 空間スペクトルエントロピーによる 共振周波数識別を使用した欠陥検出

Defect detection based on resonance frequency identification by spatial spectral entropy for non-contact acoustic inspection

○杉本和子, 杉本恒美(桐蔭横浜大), △歌川紀之, △黒田千歳(佐藤工業)

- ◆非接触音響探査法では, 遠隔から非接触で複合材料の内部欠陥を, 音響加振とレーザドップラ振動計を用いて, 検出・映像化する。
- ◆空間スペクトルエントロピー (SSE: Spatial Spectral Entropy) により, コンクリートの内部欠陥の共振周波数帯を検出し, 映像化周波数帯を絞ることによって, より明瞭な音響欠陥画像を得られた。
- ◆コンクリート内部の円形剥離欠陥に対して, 実験結果と円板の固有振動数モデル (周辺固定) の理論値との対比より, 円形剥離欠陥の直径・剥離欠陥の埋設深さと, 共振周波数の関係を検討した。

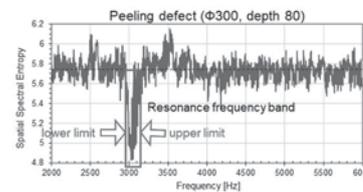


Fig.1: Resonance frequency detection of a circular peeling defect using SSE analysis.

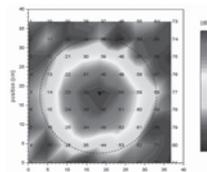


Fig.2: Acoustic visualization of a circular peel defect (Φ300, depth 80 mm).

3-8-1

3-8-1 言語聴覚士養成校の音響学教育を考える

Consideration from the aspect of acoustics education at Speech-Language-Hearing Therapist training schools.

○竹内京子(順天堂大), 青木直史(北大), 荒井隆行(上智大),
△鈴木恵子(北里大), 世木秀明(千葉工大), 安啓一(筑波技術大)

- ◆ことばに関するリハビリの専門職である言語聴覚士の養成校の必修科目「音響学」は, 臨床上の多くの専門科目の基礎であり, 言語聴覚療法には, 欠かせない科目にもかかわらず, 養成校の学生にとって苦手な科目の筆頭にあげられる。また, 「音響学」を担当する教員は, 非常勤講師がほとんどである。さらに, 他校同科目の教師との連携や, 学生が音響学を学ぶための教材も少ない。
- ◆今回, 我々は, 以下のようなプロジェクトを立ち上げ, 言語聴覚士の音響学教育について考えることにした。この発表は, プロジェクトの現在の経過報告である。
 - 1) 言語聴覚士養成校の「音響学」の現状 (教師, 養成校の言語聴覚士教員, 学生による) のアンケート調査を実施する。
 - 2) 各校の「音響学」(聴覚心理学も含む) 教師のネットワークを作り, 教材, 授業内容について検討する。
 - 3) 音響学教師と言語聴覚士教員との連携も視野に入れた, 授業ガイドラインならびに, 学生向け教材を実際に作成し, 全国の教師間, 養成校学生間で共有できるコンテンツを構築する。
- ◆現在は, 養成校の向けアンケート調査項目の検討, 音響学担当教師のメーリングリスト作成, 研究成果のサイト作成を行っている。
 - ◆本発表は, 科研費番号 20K03074・18K02988 の成果である。

3-8-2

3-8-2 緊急事態宣言下の 言語聴覚士養成校の音響学の授業

How acoustics classes should be at Speech-Language-Hearing Therapist training schools under the declaration of a state of emergency

○竹内京子(順天堂大), 青木直史(北大), 荒井隆行(上智大),
△鈴木恵子(北里大), 世木秀明(千葉工大), 安啓一(筑波技術大)

- ◆ことばのリハビリを行う言語聴覚士の養成校では、「音響学」が必修科目である。しかしながら、2020年の新型コロナウイルスの影響で、他校同様、遠隔授業や分散登校の対面授業などを余儀なくされている。本発表では、実際の養成校の音響学の授業の実践報告と、それを踏まえ、今後の対策としてすべきことについて議論したい。
- ◆多くの養成校は、大学などの大規模校を除き、普段から学習管理システム(LMS)を採用していない場合が多い。授業の配布資料も電子ファイルではなく、紙媒体の配布がほとんどである、コンピュータ環境も十分ではない。加えて、実習を含む科目が多く、対面の授業を大切にしていることも特徴である。
- ◆講義形式の場合は、オンデマンド録画映像による授業が、実習形式の場合は、Zoomなどのオンライン授業、または、Zoom 配信併用の分散登校または、一斉登校による対面授業が行われており、それぞれの形式にあった工夫を行っている。
- ◆今後の対策としてすべきことは、1) オンライン授業で使える教材の作成、2) 音響学の教員ネットワークを作成し、教材や情報を共有すること、3) コミュニケーション空間の変化に対応した授業内容にすることの3点が考えられる。

本発表は、科研費番号 20K03074・18K02988 の成果である。

3-8-4

3-8-4 オンライン授業に対応した 聴能形成のシステム構築と運用

Development and management of training system of Online Technical Listening Training

○河原 一彦, 高田 正幸(九州大・芸工),
△森尾 謙一, △藤原 賢(日本音響)

- ◆2020年のいわゆる新型コロナウイルス感染拡大の社会的背景により、教育機関はオンライン授業の対応に迫られた。九州大学においても、2020年度春学期は、オンラインの授業のみが行われることとなった。そのような社会背景において、九州大学芸術工学部音響設計学科では、音の感性教育である、聴能形成IIを、オンライン会議サービス等を用いて行った。本報告では、授業実現のためのシステムの構築と運用について述べる。
- ◆真耳 Online と zoom および Listento を利用することで、ある程度の音の配信品質を保って、授業をおこなうことができるようになった。

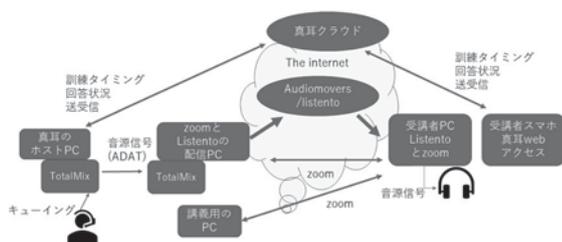


Fig.1: System Diagram of Online Technical Listening Training

3-8-3

3-8-3 小学校家庭科教科書における 「音と生活との関わり」の内容の分析

The analysis description of "connection between sound and life" in elementary school home economics textbooks

○豊増美喜(大分大), △鈴木佐代(福岡教育大)

- ◆小学校では、令和2年度(2020年度)から、新しい学習指導要領が全面实施となり、中学校家庭科で扱ってきた「音と生活との関わり」が小学校家庭科の内容となった。本報では、今回の改定で新しくなった小学校家庭科教科書の「音と生活との関わり」の記載の内容について、これまでの中学校教科書の音の記述を参考に、内容を分析、考察する。
- ◆小学校教科書の音の内容は、「涼しく快適に過ごす住まい方」や「夏を涼しくさわやかに」の章と、「共に生きる地域での生活」や「あなたは家庭や地域の宝物」に関連する章で取り扱われており、音の章や単元として独立していない。
- ◆中学校の教科書の小見出しや図表等で示されている、防音の工夫(音は振動として伝わり、厚手のカーペットや防振マットを敷くことで防音を工夫する内容等)はない。
- ◆今後も、コロナ禍による在宅時間の増加が考えられることもふまえ、児童が生活の中で音と上手につきあっていくために、小学校第3学年理科で学習する「音の性質」との関連を図り、音の性質を利用した防音の工夫の教材の提案が望まれる。

3-8-5

3-8-5 この音何デシベル? -大学生の音の大きさに対する理解度を まずは測る試み、オンライン版の開発-

What is this sound in dB?

-We tried to determine students' comprehension of sound levels, developed an online ver.-

○上田麻理, 田中哲雄(神奈川工科大)

筆者らはこれまで、数式を使わずに音の大きさの理解度を測る試みとして、音の大きさチャート・キットを作成し、音響学等の授業で実践してきた。昨今の新型コロナウイルスの影響を受けて、大学の授業がオンラインになったため、同教材のオンラインバージョンを開発した。

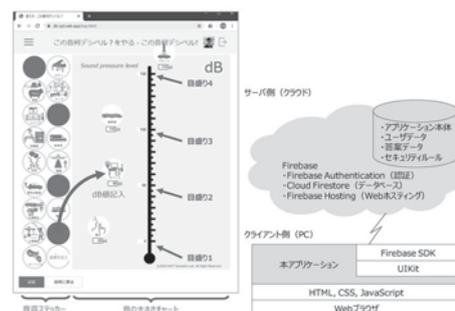
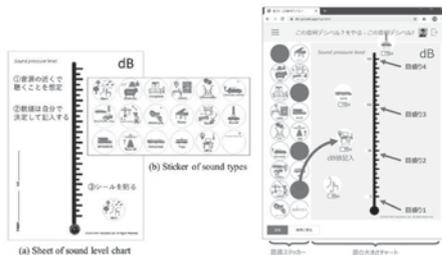


Figure: Online version of the sound level chart kit (left) and application configurations(right).

3-8-6

3-8-6 この音何デシベル？
 -オンライン版・オフライン版の教材比較と評価-
 What is this sound in dB?
 - Comparison and evaluation of teaching materials between online ver. and offline ver.-
 ○上田麻理,田中哲雄(神奈川工科大), 高橋義典(工学院大)

昨今の社会情勢により、大学においても授業のオンライン化が波及している。オフライン・オンライン授業のメリット・デメリットを定量的に把握することは重要な課題の一つと言える。そこで、教材のメリット・デメリットを把握するために、音の大きさチャート・キットのオフライン版及び、オンライン版を用いて教材の評価を行ったので報告する。



Which would you prefer?

Figure: Offline version of the sound level chart kit. (left) and Online version (right).

3-8-8

講演取消

3-8-7

3-8-7 家庭用品で作る簡易声道教材のオンライン授業での活用報告
 Report of utilization of simple vocal tract teaching materials made with household items in online classes
 ○児玉明日夏, 浜田史楓, 笠原美左和, 真志取秀人(都立産技高専), 高橋義典(工学院大学)

身近にある家庭用品を用いた工作によって、オンライン授業でも利用可能な簡易声道教材を提案する。簡易声道教材の構造は柔らかいプラスチックボトルで作られた人工声道の部分と、ストローに切り取ったクリアファイルをリードとして貼り付けた人工声帯の部分で構成されている。Fig.1に使用した材料を、Fig.2に完成した教材を示す。

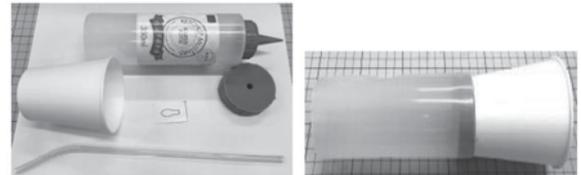


Fig.1 household items Fig.2 Simple vocal tract teaching aid

大学生9人を対象にしたオンライン授業にて提案教材を使用した。その結果、全員が20分程度で組み立てられ、母音に似た音を出すことができた。子供向けのオンライン・ワークショップにおいても、大人のサポートを推奨することで、未完成となる受講者を防ぐことができると考えられる。さらに、講師のサポートがなくても、事前に用意した説明書に従って組み立てられることから、コロナ休校中の親子で取り組む、オンデマンドな「おうち教材」としてもおおいに期待できる。

3-8-9

3-8-9 ストロー発声時の口唇周辺の皮膚振動とストロー壁面振動との関連性

Correlation between skin vibration around mouth and straw wall vibration during straw phonation
 ○川村直子(姫獨大), 北村達也(甲南大)

◆音声障害者が行うストロー発声訓練では、訓練時のフィードバックとして発声時にストローの先端、あるいは口唇周辺(皮膚)の振動を患者が自覚できることが重要であるとされている。しかしながら、口唇周辺の皮膚振動とストロー壁面振動との関連性は検証されていない。



Fig.1: Straw phonation

◆そこで本研究では、ストロー発声時の口唇皮膚の振動とストロー壁面の振動、および音声を同時計測し、これらの相関について検証した。

◆その結果、ストロー発声時の口唇皮膚の振動とストロー壁面の振動には強い正の相関がみられた。さらに、ストロー先端側よりもストロー口唇側の壁面振動の方が、ストロー発声時の口唇皮膚振動および音声の振幅と対応が良いことが示された。

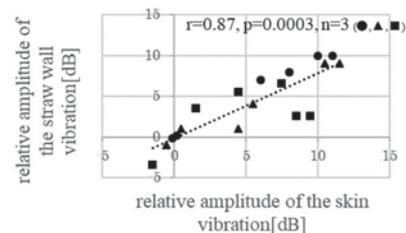


Fig.2: Correlation between skin vibration of upper lip and straw wall vibration for condition A. Marker symbol corresponds to each participant.

3-8-10

3-8-10 児童の日本語母音発声時の口唇運動に関する動画記録による検討

Examination of lip movements during production of vowels in Japanese children by video recording

○白勢 彩子(東京学芸大), 北村 達也(甲南大)

目的・概要 日本語の長・短母音のモーラリズムの生成機構の成立過程の解明を目的に、共通話話者の児童9名と成人2名を対象に長・短母音対の発声生成実験を行ない、口唇運動を動画記録して検討した。

手続 カメラが常に話者の正面に位置するよう、前面にアームでWebカメラを固定したスポーツ用ヘルメットを用いて収録した。動画中の一部フレームから正中面上の赤唇縁と皮膚部の境界(上下), 左右の口角の4点を手動で計測して学習データを作成し、アンサンブル回帰木により推定して口唇の開きの高さや幅を求め、対象母音区間の最大値を得た。

結果 母音イの結果をFig.1に示す。計測値は、孤立発声「イ」の口唇幅により正規化した。横軸は、話者の年齢を示し、例えば「6;8」は6歳8ヶ月(実験時)を意味する。「adl」は成人話者である。特定の話者や年齢層に固定した結果がなく、一貫した特徴が見受けにくいものの、児童では短・長母音間の差が少ない傾向が強い。長母音における口唇運動の拡大は、長さを音韻的に識別するための必須の動作ではないことや、発達に伴い、音韻に遅れて口唇運動の相違が成立することが考えられる。

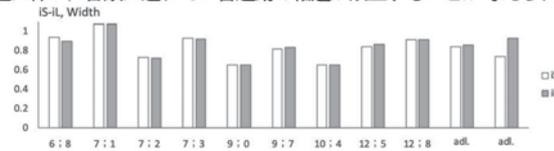


Fig.1: Width of mouse openings in short and long vowels production. iS: short vowel (/i/), iL: long vowel (/i:/).

3-8-12

3-8-12 日本人英語学習者の様々な文末調イントネーションパターンの知覚認識

The Perception of English Intonation Patterns in Various Question Styles by Japanese L2 learners of English

○浅野恵子(順天堂医)

- ◆日本の英語教育では「コミュニケーション重視」の傾向にあり、イントネーション関連の英語特有リズム習得を優先的に指導推奨。
- ◆疑問文と付加疑問文の文末調のイントネーションの文科省指針はあるが、実際の教示は現場に依存であることが現状。
- ◆英語教職課程を履修している学生がどのように質問形式文末調イントネーションを認識しているかを調査し、英語音声学の授業にどのように反映させればよいかを検討。
- ◆Q1,2のアンケート調査の結果、おおよそ、1つのパターンで疑問文文末調を認識している。Q3は回答に明確な差無。
- ◆他の文末調パターンがあると回答した場合の理由を自由記述形式で回答した結果を、情報、感情、態度、状況、質問形式、言語別、その他の項目に分類 (Table)。
- ◆「質問形式」の項目で、特徴が見られ、付加疑問文は文字通り、疑問文であるとの認識。

Table Categorized items of free description in each question

	Q1 Yes/N0	Q2 Wh /How	Q3 Tag: Rising	Q3Tag: Falling
情報:Information	0	4	10	17
感情:Emotion	2	4	5	2
態度:Attitude	1	0	0	0
状況:Differences	1	0	1	2
質問形式: Question Style	4	1	7	4
言語:Language	0	1	0	0
その他:Others	0	0	1	0

3-8-11

3-8-11 オペラ歌手が非声楽的な発声を模擬したときの横隔膜などの制御

Control of diaphragm and other speech organs when opera singers imitate non-operatic singing style

☆戸田菜月, △工藤理佐子, 竹本浩典(千葉工大), 高橋純(大阪芸大)

- ◆オペラ歌手には熟達した歌唱法(声楽的な発声)が求められるが、歌唱中に横隔膜や声道をどのように制御しているか明らかではない。
- ◆そこで、本研究では、プロのオペラ歌手と声楽を学ぶ学生が声楽的な発声をしたとき、非声楽的な発声をしたときの横隔膜と声道をrMRIで撮像した。
- ◆横隔膜の変位をカイモグラフで分析した結果、声楽的な発声では、プロは音高に応じて横隔膜の変位を細かく制御していたが、学生ではそのような制御は見られなかった。非声楽的な発声では、プロも学生もそのような制御を行っていなかった。
- ◆また、音高の上昇にともなう声道形状の変化をトレースして分析した結果、声楽的な発声では、プロも学生も喉頭を下降させた。しかし、プロは口の開きと咽頭後壁の後方への変位を増大させたが、学生は舌の扁平度合を増加させた (Fig.1)。

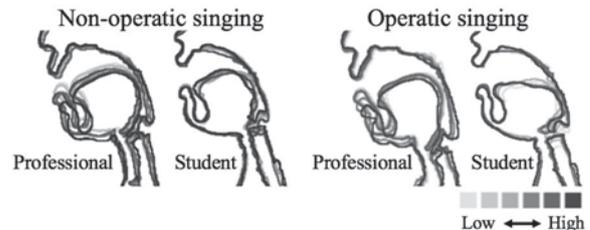


Fig.1: Changes in vocal tract shape at high pitches

3-8-13

3-8-13 英語コミュニケーション力向上のための自己学習教材の開発

—感情表現を中心として—

Development of an effective self-studying method to acquire expressiveness in speaking Environment

石井朱美(芝浦工業大学)

- ◆母語のみならず外国語においても感情を表現できるようになることは重要である。
- ◆本研究では、まず工学系大学生166人に英語に対する意識調査を行った。調査の結果、学生は、「英語は好きだが苦手意識がある。しかし、話せるようになりたいと思っている。」ことがわかった。
- ◆英語母国語話者が感情を込めて発声した音声や映像を用いた自己学習教材と方法について検討し、見本音声無し、音声見本あり、音声付き映像のどの方法が良いか検証した。
- ◆学習実験の結果、映像を見本にした方法が一番良いことがわかった。学習者の音声練習後の録音音声は感情ごとに異なっていた。

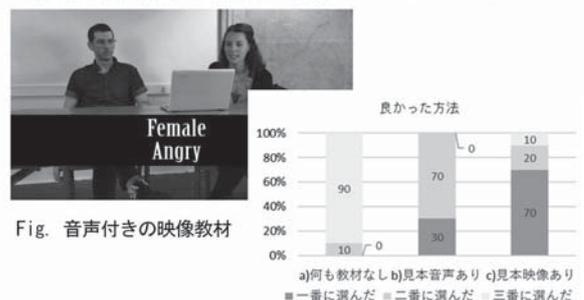


Fig. 音声付きの映像教材

Fig. 学習者のアンケート

3-8-14

3-8-14 Shadowability-based Analysis of Behaviors of Listeners with Diverse Language Backgrounds when Listening to Japanese English Utterances

☆Chuanbo ZHU, Ryo TAKASHIMA, Shintaro ANDO, Zhenchao LIN,
Daisuke SAITO, Nobuaki MINEMATSU(The University of Tokyo),
Noriko NAKANISHI(Kobe Gakuin University)

- ◆ This study conducted experimental analysis to examine how one's language background influence his/her listening comprehension while listening to L2 English.
- ◆ Speakers were Japanese university students, while listeners were native and non-native but proficient speakers of English. In each group of the listeners, some were learners of Japanese and the others were not.
- ◆ Their listening comprehension was observed online with reverse shadowing and script-shadowing method. Shadowing and script-shadowing utterances were compared with the Dynamic Time Warping (DTW) method, which can quantify brokenness of articulation and additional pausing in shadowings.
- ◆ We made a pedagogical discussion on how learners should talk in English and how they should improve their pronunciation.

3-9-2

3-9-2 楽器音の指向性を反映した演奏練習のためのホール残響模擬システム

Hall reverberation simulation system reflecting the sound directivity of musical instrument for practicing musical performance
○長谷芳樹(ピクシーダストテクノロジーズ), 饗庭絵里子(電通大)

- ◆ 楽器の持つ指向性やホールの残響音の到来方向を意識して演奏練習をおこなえるシステムの構築を試みた。
- ◆ ステージの縁の部分に複数のマイクロホンを設置し、ステージ中央に設置したスピーカーから広帯域信号を送波して、ホールの各方向からの残響音(インバルス応答)を求めた。直接音は除去した。
- ◆ グランドピアノを対象とした。ソフトウェア音源を用い図1に示す位置に仮想マイクロホンを設置し演奏音を収録した。この演奏音に、対応する方向の頭部伝達関数を畳み込んで呈示音とした。聴取者はコンサートホールでのグランドピアノ演奏経験のある演奏者4名とした。
- ◆ 演奏者へのヒアリングの結果、右方向に広がっている空間を感じた、等のコメントが得られた。一方で、広めの空間は感じるものの、通常のグランドピアノ演奏者の位置ではなくステージ上で聞いている音に近い、といったコメントもあった。今後、詳細な検討が必要である。

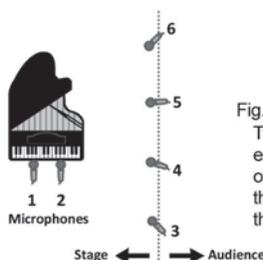


Fig.1: Setup of the virtual microphones. Two microphones are placed at the ear positions of the player, and the other four microphones are placed at the frontal edge of the stage facing the instrument.

3-9-1

3-9-1 グランドピアノ演奏における感情表現と身体部位の軌跡長の関連について*

* Relationship between emotional expression and trajectory length of body parts in grand piano performance
○三戸勇気(日大・芸術), 幸田優志(日大・理工), 川上央, 駒澤大介(日大・芸術), 三浦雅展(国立音楽大学), 山口達也, 篠田之孝(日大・理工)

我々は、現在まで鍵盤楽器の演奏動作に関する研究を行ってきた。本研究では、複数のピアニストによるグランドピアノの演奏動作の特徴について抽出することを目的としている。今回はプロのピアニスト3名の方にご協力いただき、感情表現に対する演奏動作の違いについて検討した。各ピアニストの演奏動作をモーションキャプチャシステムにより計測し、演奏者による演奏動作の違いを見出すための検討を行った。今回は、感情表現による身体各部の重心位置の高さ方向の軌跡長に着目して分析を行った。

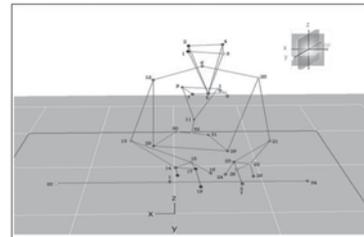


Fig.1 Marker adhesion position

3-9-3

3-9-3 7.1ch 立体音響再生システムにおける天井スピーカ配置の違いが聴こえに及ぼす影響

○竹上力, 脇田由実(大工大), 入交英雄(株式会社 WOWOW)

- ITU-Rは、家庭内使用を視野に入れた7.1サラウンドシステムを提案しているが最適スピーカ配置はまだ決めていない。今回は天井におけるスピーカ配置場所の違いと臨場感の関係を報告する。
 - 天井スピーカ配置の違いにより、臨場感がどう変化するのか3つの配置を一対比較法にて評価実験を行った。実験には3曲と4つの評価項目を使用した。
 - 包み込まれ感のシェッフェの一対比較法の結果をFig.1とする。包み込まれ感ではBizet 以外では後方スピーカの上下の位置の違いは聞こえに影響が少なかった。理由として楽曲特徴の差が関係している可能性があり、今後、楽曲特徴の違いとスピーカ配置の関係も明らかにしたい。
- また、包み感以外の評価項目は耳の高さのスピーカ数に依存しており、一定数以上を配置すると一定レベルの広がり感を得ることができることを示唆していると考えられる。

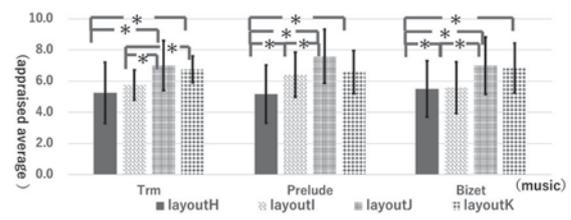


Fig.1 viewing experiment result (Envelopment)

3-9-4

3-9-4 メロディー輪郭の変化に対する ミスマッチ・フィールドの脳左右半球差 Interhemispheric differences in melody contour perception assessed by mismatch field

○☆内藤柚菜, ヤップ・ゲック・シエン(千葉大), 柴玲子(慶應大),
大塚翔(千葉大), 湯本真人(東京大), 中川誠司(千葉大)

- ◆音楽経験者と非音楽経験者を対象に、メロディーの輪郭(音程の上下の変化パターン)および音高の変化に対するミスマッチ・フィールド(MMF)を計測した。また、MMFの脳左右半球差を検討した。
- ◆連続する5音で構成される標準刺激の第3音の音高を変化させ、5種類の逸脱音列を作成した。全頭型脳磁界計測装置を用いて、逸脱音列に対して誘発されるMMFを観察した。
- ◆音楽経験の有無にかかわらず、右半球に比べて左半球のMMF潜時の短縮が認められた。
- ◆非音楽経験者においては第3音の音高変化が大きいほどMMFが増大したが、音楽経験者では輪郭変化(山形→谷形)に対してより大きなMMFが出現した。
- ◆音楽経験者では輪郭変化に対するMMFは左半球優位であった。また、非音楽経験者に比べてMMF潜時の短縮も認められた。
- ◆音楽経験により、音列を群化して捉える能力が高まり、メロディー輪郭の弁別能力が向上したこと、およびメロディー処理に係る速度と左半球優位性が亢進したことを示している。

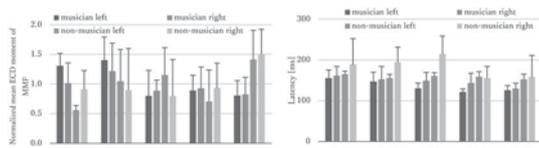


Fig.1: ECD moment of MMF

Fig.2: Peak latency of MMF

3-9-6

3-9-6 モーラ相対位置を用いたEncoder-Decoder モデルに基づく逐次複数歌詞候補生成の検討 Sequential Multiple Lyrics Generation Based on Encoder-Decoder Model Using Mora Relative Position

☆高橋力哉, 能勢隆, 伊藤彰則(東北大)

- ◆本研究では、従来のN-gram言語モデルに基づく作詞補助システムにおける長期的な歌詞の文脈を考慮できないという問題点を解決するため、Encoder-Decoderモデルを用いた歌詞生成手法の検討を行った。
- ◆複数フレーズ生成とフレーズスコアを用いた並べ替えにより、自然なフレーズが生成できることを確認した。
- ◆モーラ相対位置により、フレーズの自然性を保ちながら生成フレーズのモーラ数の制御ができることを確認した。
- ◆入力3単語において、主観評価で生成フレーズ自体、入力フレーズとの接続ともに入力2単語よりも高い自然性を確認した。
- ◆今後は、さらに長い文脈を考慮した歌詞候補生成、歌詞のトピックを考慮した歌詞候補生成を課題として取り組む予定である。

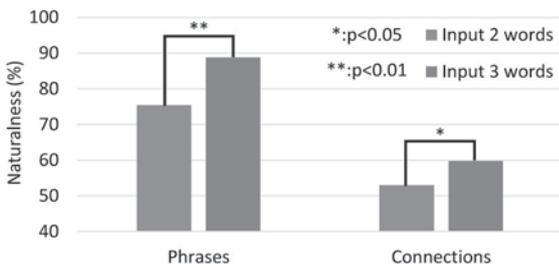


Fig.1: Naturalness of phrases and connections

3-9-5

3-9-5 一側性難聴における音楽感受、 および音楽活動への影響・順応の検討 ——質問紙調査を中心として——

Music appreciation and adaptation for those with unilateral hearing loss:
a questionnaire survey

☆辻慎也, 荒井隆行(上智大・理工)

- ◆片側耳は正常であるが、対耳側に難聴がある状態を一側性難聴という。
- ◆著者らは過去に後天性の一側性難聴者について：
 - ・難聴の発症直後、音楽が「不自然」、「不快」、「不明瞭」に感じられ、音楽を聴く時間、音楽の楽しさ、演奏頻度…などが低下し、小編成の音楽を好むようになったこと
 - ・発症後、音楽を聴く時間、音楽の楽しさ、演奏頻度…などが増加 → 順応の過程がみられたこと

- ◆本稿では先天性の一側性難聴者について：
 - 1) 幼少期から青年期にかけて、後天性難聴と共通する順応の過程があること
 - 2) 後天性難聴と比べ、難聴発症直後にあたる幼少期における困難感が少ない傾向にあること
 - 3) 一方、先天性難聴であっても高い困難感を示すケースもあり、個別検討が必要であること

3-S1-1

3-S1-1 軟骨伝導における呈示部位と聴覚閾における 外耳道内音圧の関係

Relationships between transducer placements and ear-canal-sound
pressure at the threshold of the cartilage conduction

☆菅原優聖, 大塚 翔, 中川誠司(千葉大)

- ◆“軟骨伝導”は、装用に係る骨伝導の問題点を解決する方法として提案され、外耳道閉鎖症用の補聴器やスマートホンに活用されている。
- ◆耳介の一部や耳珠のみに振動子を呈示した先行研究では、気導成分と外耳道内放射成分が支配的であることが報告されている。しかしながら、実際の軟骨伝導デバイスでは、軟骨の周囲の頭骨部にも振動面が接触する機会が多い。振動子の接触部位の変化は、伝搬成分および知覚特性の変化を生じるが、その多くは未解明のままである。
- ◆振動子の呈示部位が軟骨伝導の伝搬成分や知覚に与える影響を定量的に検討するために、軟骨伝導の聴覚閾およびその際の外耳道内音圧を呈示部位ごとに推定した。
- ◆耳介および耳前部に呈示した場合は、乳様突起と比較して2000 Hz以下における聴覚閾が10~20 dB程度下回った(Fig. 1)。低域が増強される外耳道内放射成分の寄与によるものと考えられる。
- ◆聴覚閾における外耳道内音圧は、1000 Hz以下では乳様突起よりも耳介は5~10 dB程度小さいものの、逆に4000 Hz以上では耳介の方が5 dB程度大きくなった(Fig. 2)。乳様突起呈示時には、高周波域において、本来の聴覚閾の差を補う以上の慣性骨導成分、もしくは圧縮骨導成分が生じること、それらの作用のために聴覚閾が5 dB程度低下することを示している。

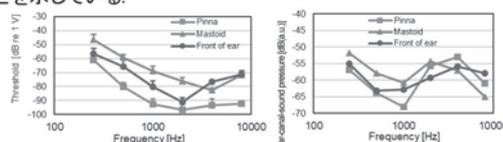


Fig. 1 Threshold for each placement.

Fig. 2 Ear-canal-sound pressure at the threshold for each placement.

3-S1-2

3-S1-2 遠位呈示 AM 骨導超音波の単音節明瞭度・単語理解度と調音素性情報伝達特性の分析

Estimations of word intelligibility, monosyllable articulation, and transmission characteristics of phonetic features of distantly-presented bone-conducted ultrasound

☆土井 公一朗, 大塚 翔, 中川 誠司(千葉大)

- ◆骨導超音波(骨導呈示された 20 kHz 以上の高周波音)は明瞭な音として知覚可能であり、振幅変調することで音声伝達も可能である。
- ◆骨導超音波は首や体幹、上肢等の頭部から離れた部位(遠位)に呈示した場合でも知覚される。
- ◆遠位呈示骨導超音波の音声伝達性能を明らかにするため、単語理解度と単音節明瞭度を測定した。また、SINFA (Sequential information analysis) によって調音素性情報の伝達特性を分析した。
- ◆頭部や鎖骨に呈示した場合も、気導条件や従来の骨導呈示部位である乳様突起に呈示した場合と遜色ない正答率が得られた (Fig. 1)。
- ◆母音部の伝達は子音部よりも良好であった。また、拗音や鼻音などでは、呈示部位に依存した伝達率の変化が確認された。伝搬経路の長さや周波数特性が、音響特徴の伝達に影響を与えることが示唆される。

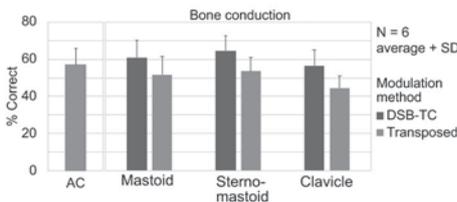


Fig. 1 Scores of the monosyllable-articulation tests for each stimulus type and stimulus placement.

3-S1-4

3-S1-4 分離音効果の周波数特徴に関する研究

A Study on the Frequency Characteristics of the Split-Off Effect

○山崎 祐 (九大芸工) 中島 祥好 (サウンド株式会社) 上田 和夫 Gerard B. Remijn (九大芸工)

- ◆対数周波数軸上で直線変化する二つのグライド音を時間的な重なりを持たせて呈示すると、長い連続音とその時間的中央付近に位置する分離音が知覚されることがある。この錯聴を「分離音効果」と呼ぶ。
- ◆第一グライド音の終わりりと第二グライド音の始まりりが、時間的、周波数的に近接していることで知覚的統合が生じて、本来は存在しない分離音が知覚される。
- ◆分離音周辺のトーンクロマ情報が上手く知覚できない際でも、知覚的統合が起こるのか調べることを目的とし、参加者(n=9)には刺激の印象を4段階のスケールで評価させた。
- ◆分離音周辺のトーンクロマ情報が少ない時の方が、分離音が知覚されやすくなる結果が得られた(Fig.1)。トーンクロマが明確でなくなったことで二つの音が区別しづらくなり、このような結果が得られたと考えられる。

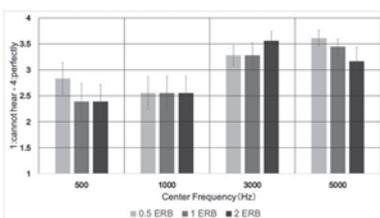


Fig. 1: Average rating values (for n=9) per stimulus for the question "Did you hear the "Middle tone" in the split-off stimulus?" Error bars show the standard errors.

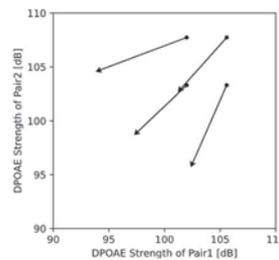
3-S1-3

3-S1-3 DPOAE を用いた蝸牛での抑圧現象の方向性の検証

Testing of the direction of tone suppression in the cochlea using DPOAE

☆藤匠海, 村上泰樹(大島商船高専)

- ◆蝸牛では複合音が呈示されると、純音が呈示された時と比べ音響振動が減少する抑圧現象が知られている。
- ◆抑圧現象の方向性として、高強度の音と低強度の音を同時に呈示した時、低強度の音のみ抑圧される一方向的な抑圧と高強度の音と低強度の音の両方が抑圧される双方向的な抑圧が予測されている。
- ◆DPOAE を生じる 2 音ペアを 2 つ用い、単独ペアでのみ観測される DPOAE 音圧に対し他方のペアが存在するときの DPOAE 音圧から抑圧量を測定した。測定結果はベクトル表示により解析した(図1、表1)。
- ◆結果として 2 つのペアの音圧が等しい時は双方向的な特性を抑圧現象は示した。
- ◆2 つのペアの音圧レベルに差があるときは一方向的な特性を抑圧現象は示した。



Level of Pair 1 [dB]	Level of Pair 2 [dB]	Position	Angle [deg]
70	70	Upper right	49.12
70	65	Lower right	66.97
65	70	Upper left	21.21
65	65	Lower left	44.39

Fig. 1: Vector expression of suppression in DPOAE

3-S2-5

3-S2-5 同一話者による平静音声と怒り音声のモーフィングに対する感情知覚

Emotion Perception in Morphing of Neutral and Angry Voices by the Same Speaker

☆坂本拓也, 坂下尚史(豊橋技科大)

河原英紀(和歌山大), 松井淑恵(豊橋技科大)

- ◆平静と怒りの感情音声をモーフィングしたときに人がどの程度怒りの感情を知覚するのか検討した。
- ◆「平静」と「怒り」の感情で発話された男声と女声それぞれの「[ま]」を原音声として、STRAIGHT モーフィングを利用し中間的な音声を作成した。怒りの割合を 0, 20, 40, 50, 60, 80, 100[%]に設定。
- ◆2 つのモーフィング音声を聞き、どちらがより怒りの感情に聞こえるか GUI 上で実験参加者に判定させた。
- ◆怒りの割合の応じて怒りが知覚される割合は単調増加したが、その増加具合には個人差が存在することが示唆された。

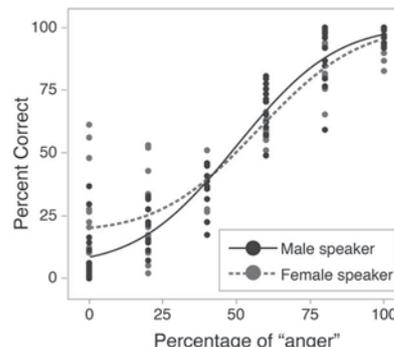


Fig. 1 Result of morphing speech listening experiment. The horizontal axis is the morphing ratio of the comparative stimulus (100 is angry speech only, 0 is neutral speech only). The vertical axis is the ratio of feeling angry compared to 50% angry speech. The small circles are the results for each participant.

3-S2-6

3-S2-6 両耳性の遅延追加呈示パルス列刺激の物理的特徴とオクターブ・シフトとの関係

The relation between physical characteristics of binaural addition of two pulse trains with delay and octave shift.

☆渡邊 菜, 津崎 実 (京都市芸術大学)

- ◆ 同一の周期を持つふたつのパルス列が周期の1/2に相当する時間的なずれを持って加算呈示されると原周期よりも1オクターブ高く知覚される(オクターブ・シフト)。
- ◆ このようなふたつのパルス列に対し、それぞれが異なる位置に定位されるような両耳間レベル差 (interaural level difference: ILD) あるいは両耳間時間差 (interaural time difference) を与えて加算呈示すると、原周期のピッチよりも1オクターブ高くなったと判断される頻度が低下し、オクターブ・シフトが起きにくくなる。
- ◆ この刺激音を単耳で聴取した場合と両耳で聴取した場合を比較すると、オクターブ・シフト推定値に差が現れなかったことから、上記のオクターブ・シフト低下の原因は、両耳間差を与えたことによって単耳に呈示される音響信号が厳密な周期構造から逸脱したためである可能性が示唆された。
- ◆ 上記を受けて、本研究では単耳のみで聴取される刺激の周波数分析と、聴覚イメージモデルを用いた刺激の周期性の分析の点から、物理的特徴とオクターブ・シフトの生じ方についての関連性について検討した。
- ◆ その結果、刺激に含まれる周波数成分や聴覚イメージモデルによって検出された周期性の面でも、オクターブ・シフトの低下が説明可能であることが明らかとなった。

3-S2-8

3-S2-8 耳栓装着時の骨伝導音知覚に及ぼす刺激呈示の効果: 単音節明瞭度の比較

Effect of the vibrator placement on perception of bone-conducted sound during earplugging: Comparison of mono-syllable articulation

☆信夫 大志, 大塚 翔, 中川 誠司 (千葉大)

- ◆ 機械工場等の強大騒音下では作業者は耳栓を装着することが多いが、円滑なコミュニケーションが損なわれてしまう。
- ◆ 一方、骨伝導音は耳栓装着下でも明瞭に聴取される。労働者が装着するヘルメット等に骨伝導スピーカを内蔵させることで、強大騒音下でも利用可能な音声コミュニケーション・デバイスが実現できる。
- ◆ 強大騒音下、耳栓装着時において、頭皮を含む頭部の様々な部位から呈示された骨伝導音の単音節明瞭度試験を実施した。
- ◆ 明瞭度は顎状突起(Con; 下顎骨と頬骨の接続部)で最も低い結果となったが、呈示部位間での有意差は観察されなかった。
- ◆ 各呈示部位で生じる異聴傾向には耳介付近の部位と頭皮上部位では異なる子音カテゴリの異聴が確認された。また、特に蝸牛からの距離が大きい前頭極正中部と後頭正中部にて特徴的な異聴が観察されていることから、頭部内組織の伝搬過程において比較的大きな音声特性の変化が生じていると思われる。

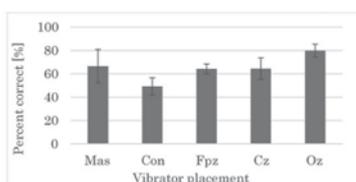


Fig. 1 Percent correct of the mono-syllable articulation test.

3-S2-7

3-S2-7 耳介への押付呈示圧が軟骨伝導知覚に与える影響

Effects of pressing pressure of a vibrator the cartilage conduction

☆実川 徹, 大塚 翔, 中川 誠司 (千葉大)

- ◆ 骨伝導における振動子の装用性の改善を目的として“軟骨伝導”が提案され、外耳道閉鎖症用の補聴器やスマートホンに活用されている。
- ◆ 軟骨伝導では、振動子の接触面積や接触圧などに応じて聞こえが変化するが、呈示条件による聞こえの差異が詳細に検討された例はない。
- ◆ 振動子の押付呈示圧が軟骨伝導の聞こえや末梢伝搬過程に及ぼす影響の解明を目的として、耳介への押付呈示圧を変化させた際の外耳道内音圧を計測した。また、顎状突起に呈示した場合の計測も行った。
- ◆ 耳栓装着下において耳介へ刺激呈示した場合は、呈示圧の増大に伴う外耳道内音圧の有意な上昇が認められた (Fig. 1)。
- ◆ 対して、顎状突起への呈示においては、外耳道内音圧の有意な変化が認められなかった (Fig. 1)。耳介軟骨と比較して押付呈示圧の影響が小さいと考えられるが、耳介軟骨の変形も考慮した詳細な検討が必要である。

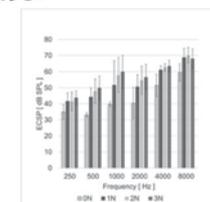


Fig. 1 Ear canal sound pressure for the pinna-presentation (earplugged)

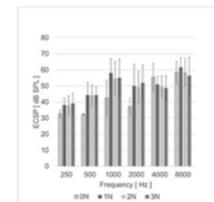


Fig. 2 Ear canal sound pressure for the condyle-presentation (earplugged)

3-S3-9

音列の規則性がオリーブ蝸牛束と大脳皮質の活動に与える影響

Effect of rhythm-based prediction on activities of medial olivocochlear fibers and cerebral cortex

○石坂 勇毅, 大塚 翔, 中川 誠司 (千葉大)

- ◆ 脳幹から外有毛細胞へは、オリーブ蝸牛束と呼ばれる遠心性の神経投射が存在し、音響刺激に反応して、外有毛細胞の増幅を抑制し、雑音下での聞き取りの改善する役割や内耳を強大音から保護する機能を果たしていると考えられている。この反応はオリーブ蝸牛測反射 (MOCR) と呼ばれている。
- ◆ 著者らはこれまでに先行する音列の規則性からターゲットとなる音を予測することで MOCR が変化することを報告した。
- ◆ 本研究では、その神経的メカニズムを解明するための予備的な検討として、予測的処理に関与していると考えられている δ 波帯域の脳活動の律動と MOCR を同時に計測した。
- ◆ 音列の規則性を調節するために、先行する音列の音同士の間隔にジッターを加えた。そのジッターの量に応じた MOCR、および、 δ 波の変動を比較した。
- ◆ ジッターが大きいくまど、MOCR 強度、及び、 δ 波の位相同期度が減少する傾向が見られた (Fig. 1)。この結果は、MOCR と脳活動の律動がともに、規則性に基づいた予測的処理に関与している、あるいは、その影響を受けることを示唆している。

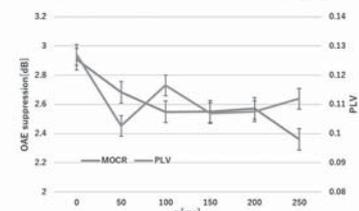


Fig. 1 Mean MOCR strength and phase locking value (PLV) of delta wave as a function of jitter (alpha).

3-S3-10

3-S3-10 単純計算作業による精神疲労がオリーブ蝸牛束反射に与える影響

Effect of Mental Fatigue on Medial Olivocochlear Reflex

☆鈴木裕登, 大塚翔, 中川誠司(千葉大)

- ◆ 脳幹から外有毛細胞へは、オリーブ蝸牛束と呼ばれる遠心性神経の投射がある。オリーブ蝸牛束は音に反応し、外有毛細胞の増幅を抑制することで内耳を保護している。そのため、この反応(オリーブ蝸牛束反射: MOCR)の強度を測定することで、騒音性難聴のリスクを事前に評価できる可能性がある。
- ◆ しかし、MOCR強度は測定の変動するため、騒音性難聴の正確なリスク評価は困難である。これまでに、注意などの認知機能や覚醒度の変動によって、MOCRが変化することが報告されている。
- ◆ 本研究では、認知機能の低下をもたらす要因の一つである精神疲労に着目し、単純な計算作業の繰り返しに伴う精神疲労の増加と、MOCR強度の測定ごとの変動の関係を調べた。
- ◆ 精神疲労は、フリッカー検査(CFF)によって客観的に評価し、MOCRは耳音響放射(外有毛細胞の増幅を反映)を測定することで評価した。
- ◆ 計算課題の繰り返しに伴い、臨界フリッカー周波数(CFF)は減少する一方で、MOCR強度は増加した(Fig. 1, 2)。CFFの減少は、精神疲労の増加を反映していることから、この結果は、精神疲労がMOCRによる内耳保護機能に影響を与え、騒音性難聴のリスクに影響を与えることを示唆している。

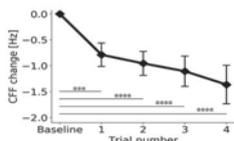


Fig.1 CFF changes relative to its baseline. ***p<.005, ****p<.001

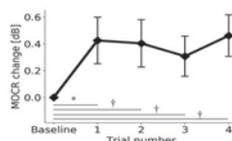


Fig.2 MOCR changes relative to its baseline. † p<.01, *p<.005

3-S3-12

3-S3-12 管楽器の音色の評価に視覚情報が与える影響

The Effect of Visual Information on the Evaluation of Wind Instrument Timbre

☆荒井優社(九大芸工), 中島祥好(サウンド(株)),

上田和夫, Remijn Gerard B.(九大芸工)

- ◆ 本研究では視覚情報が聴覚情報から得られる楽器音の音色の印象に影響を与えるのかを調べることを目的とする。
- ◆ トランペットの音色の印象を12個の評価語について1~7の7段階で評価するという実験を行った。
- ◆ 実験は2つの条件に分かれており、聴覚刺激のみを呈示して評価を行うグループA(Audio, n=13)と、聴覚刺激をトランペットの写真を用いた視覚刺激と一緒に呈示して評価を行うグループAV(Audio and Visual, n=13)の2条件であった。
- ◆ 実験の結果、12個の評価語のうち8個の評価語でグループ間の評価値に有意な差が見られ、視覚情報が付加されることで特に金属性因子に関する評価語に関して音色に対し、より強い印象を抱くことが明らかになった。



Fig.1 Stimulus conditions used in the experiment: Audio only (Group A); audio and visual information (Group AV).

3-S3-11

3-S3-11 可聴周波数帯における遠位呈示骨伝導の基礎知覚特性の検討

A study on basic perceptual characteristic of distantly-presented bone-conducted sounds in the audible frequency band.

☆石川大夢, 大塚翔, 中川誠司(千葉大)

- ◆ 従来の骨伝導デバイスの装着性の問題を解決するために、骨伝導音を頭部から離れた位置(遠位)に呈示する方法が提案されている。しかしながら、可聴周波数帯の骨伝導音を遠位呈示した場合は骨伝導成分のみの評価が難しく、知覚特性を検討した例は極めて限定的である。
- ◆ 本報では、可聴周波数帯における遠位呈示骨伝導音の知覚特性の解明を目標とし、気導成分を十分に遮断した状態で検出閾および周波数弁別閾(DLF)の測定を行った。
- ◆ 刺激呈示部位が頭部から離れるほど検出閾が上昇し、4000 Hzまでは周波数の増大に伴い、距離減衰が増加することが確認された(Fig. 1)。
- ◆ 乳様突起及び胸鎖乳突筋の全ての周波数、さらに最も頭部から離れた鎖骨における500-2000 Hzでは、気導音と同等の周波数分解能力が確認された(Fig. 2)。
- ◆ 周波数弁別閾では部位間の差異が確認されなかったことから、体内伝搬に伴う周波数情報の劣化は大きくないことが示唆された。

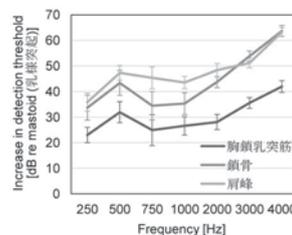


Fig.1 Relative threshold of the distantly-presented bone-conducted sounds. The threshold for the mastoid process in each participant served as the reference (0 dB).

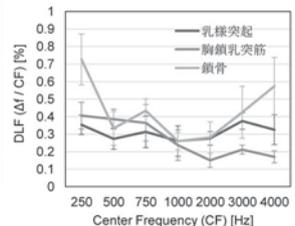


Fig.2 Frequency difference limens (DLF) for each stimulus placement

3-S3-13

3-S3-13 移動音の方向を知覚するために必要な移動角度の推定

Estimation of moving angle needed to identify a direction of moving sound

○工藤彰洋, 牧下涼亮(苫小牧高専), 武居周(宮大工)

- ◆ 移動音の方向を知覚するために必要な最小の移動角度の推定を目的として、種々の移動速度(2 - 32°/s)と移動角度(1 - 64°)について、水平面上において被験者と等距離で円運動する音像の移動方向を回答させる音響心理実験を行なった。
- ◆ 実験は被験者ごとに測定したHRTFを用いたヘッドホン再生のバーチャルサウンドにより実施した。
- ◆ 実験結果から、移動角度 $\Delta\theta$ に対する正答率の特性を被験者ごとに算出した。標準累積分布関数の逆関数で正答率にZ値変換を施すことによって、非線形の特性を直線へと変換し、最小二乗法によって正答率が75%になる時の移動角度 $\Delta\theta_{min}$ を導出した。
- ◆ 被験者ごとの $\Delta\theta_{min}$ の算出結果をTable 1に示す。表中でn/aと記載している箇所は、正答率が低かったために $\Delta\theta_{min}$ を導出できなかったことを表している。導出結果から、 $\Delta\theta_{min}$ はMAMAと同様の傾向を示すことや、その値はMAMAの1.5倍程度であることが示唆された。

Table 1: Calculated minimum moving angle needed to perceive a direction of moving sound.

ϕ [°]	S1	S2	S3	S4	S5	Ave. ± Std.Dev.
0	3	2	6	5	n/a	4 ± 1.5
45	15	9	31	6	n/a	15 ± 9.6
90	16	22	103	20	34	39 ± 32

3-T1-1

3-T1-1 SMOC コーパス: 対話システムのための大規模マルチモーダル対面雑談コーパス

SMOC corpus: A large-scale multimodal one-on-one chat-talk corpus for dialog systems.

☆山崎善啓(東北大), 千葉祐弥(NTT CS 研), 能勢隆, 伊藤彰則(東北大)

- ◆音声・表情・ジェスチャーなどを考慮したマルチモーダル対話システムはこれまでに提案されているが、深層学習の適用はユーザー状態の推定などに限定される。
- ◆原因の一つとして、大規模かつ高品質なマルチモーダル対話コーパスの集積が十分でないことが挙げられる。
- ◆本研究では、対話システムの構築に汎用的に利用可能なデータとして、大規模なマルチモーダル対面雑談コーパス (Spontaneous Multimodal One-on-one Chat-talk コーパス, SMOC コーパス) の構築を行う。
- ◆これまでに全510対話の収録を実施し、発話時間長で計56時間分の対話データが得られた (Fig. 1)。



Fig. 1: Example of collected dialog data.

3-T1-3

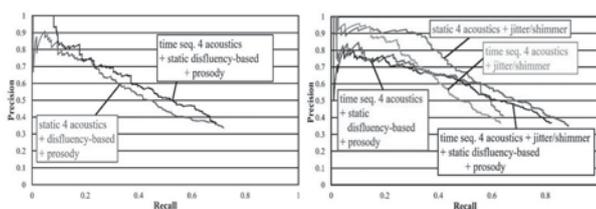
3-T1-3 Comparison of With/Without Time Sequential Features in Automatic Fluency Detection of Spontaneous Speech

話し言葉の流暢性・非流暢性の識別における時系列特徴量の有無の比較

☆Huaajin Deng, Takehito Utsuro (Univ. Tsukuba),

Akio Kobayashi (Tsukuba Univ. Tech.), Hiromitsu Nishizaki (Univ. Yamanashi),

- ◆Fluent speech detection:
 - Time sequential four acoustic features with static disfluency-based and prosody features achieve the best performance.
 - Jitter and shimmer features have almost no contribution.
- ◆Disfluent speech detection:
 - Static four acoustic and jitter/shimmer features show the best performance.
 - Either static or time-sequential jitter/shimmer features improve the precision a lot.



(a) Fluent speech detection (b) Disfluent speech detection

Fig.1 Comparison of static and time-sequential features

3-T1-2

3-T1-2 Cycle-Consistency を利用したマルチモーダル音声強調システムの各種ノイズに対する効果の検証

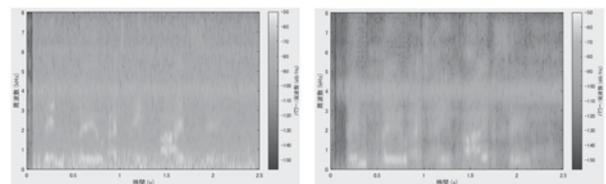
Effect of Multimodal Speech Enhancement Using Cycle-Consistency on Various Types of Noises

☆池上凌(東京理大院・理工), △大村英史, 桂田浩一(桂田研)

音声強調とは、雑音混じりの音声から目的の音声を強調して聞き取りやすくする技術を指す。この音声強調の品質を高めるために動画を音声に加えて利用したものがマルチモーダル音声強調である。筆者らはニューラルネットワークで構成された従来の音声強調システムの学習に Cycle-Consistency を導入することによって未知の種類ノイズに対して性能を改善することに成功した。本稿では入力音声に重畳する雑音音声と出力される強調音声のそれぞれの周波数成分を解析し比較を行い、提案手法の各種ノイズに対する効果を検証する。実験の結果、Figure1, Figure2 に示すように、提案したモデルは0~1kHz 帯のノイズをよく削減する一方で、3~5kHz 帯のノイズ低減は上手くできていないことが分かった。また先行研究と比較して、雑音成分の除去と話者成分の再現がより優れていることが分かった。

Figure 1 Noise Added

Figure 2 Enhanced Audio



3-T1-4

3-T1-4 音のビット列の特徴分析

Analysis based on Bit Sequence Representation of Audio Waveform

☆大川正暉, 王禱康, 西崎博光(山梨大・工),

- ◆音波形のビット列表現を用いた音分類の分析
 - 音波形のビット列表現を入力とした End-to-End 手法
 - ◇ ビット列の表現方法と振幅の変化の分類精度を比較
- ◆実験: 各種入力の分類精度の評価結果
 - 使用データ
 - ◇ GTZAN データセット (D1)
 - ◇ 日本で放送されたラジオデータ (D2)
 - ◇ 発話音声の Google Speech Commands データセット (D3)
 - 音波形のビット列表現を変化させ分類を評価
 - ◇ 全てのビット列表現方法は MFCC と同等の性能
 - ◇ 音楽・音声分類、発話単語分類に適用可能
 - 音波形の振幅最大値を揃えたビット列の分類を評価
 - ◇ 振幅最大値を揃えると分類性能向上

Table 1: Classification Accuracy for Input Format [%]

入力形式	Dataset		
	D1	D2	D3
Int16	96.5	89.6	85.0
Float32	94.7	88.7	84.3
Float16	96.5	87.8	84.9
Raw wave	85.1	88.7	22.7
MFCC	93.9	90.4	91.1

3-T1-5

3-T1-5 雑談コーパスを用いた双方向 LSTM に基づくフィラー予測の検討

Filler prediction based on a bidirectional LSTM using a chat-talk corpus
 ☆山崎善啓(東北大), 千葉祐弥(NTT CS 研), 能勢隆, 伊藤彰則(東北大)

- ◆ 音声対話システムにはフィラーのような発話の非流暢性も重要な要素であるとされている。
- ◆ フィラーを表出する音声対話システムを実現するために、文に対してフィラーを挿入するフィラー予測モデルに着目する。
- ◆ 我々は双方向 LSTM に基づいて文全体を考慮し、フィラーの挿入位置と種類を同時に最適化するモデルを提案した (Fig. 1)。
- ◆ ベースラインモデルとの性能の比較の結果、提案モデルはより高精度なフィラー予測が可能であることを示した (Table 1)。

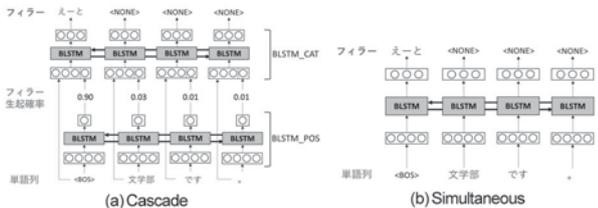


Fig. 1: Overview of the proposed filler prediction model.

Table 1: Experimental results.

Condition	Precision	Recall	F-measure
Baseline (CRF and N-gram, random, W = 15,484)	0.075	0.063	0.067
Cascade (re-trained all layers)	0.345	0.197	0.249
Simultaneous	0.471	0.112	0.141

3-T2-7

3-T2-7 音声感情認識の認識精度に対する発話末尾の音響特徴量の影響

Improvement of accuracy of speech emotion recognition by using acoustic features at the end of speech

☆高山和明(滋賀大院・DS 研), 市川治(滋賀大・DS)

- ◆ 研究概要
音声感情認識において、発話全体の音響特徴量に加えて発話末尾の音響特徴量を用いることにより認識精度が向上したことを報告する。
- ◆ 提案手法
発話全体の LLD (Low Level Descriptor) から計算した統計量 (INTER-SPEECH 2010 の 1,582 次元の特徴量; IS10 特徴量) に加えて発話末尾 0.5 秒分の統計量 (IS10 特徴量) を入力として DNN および LightGBM の分類器 (喜び, 怒り, 悲しみ, 平静) の学習を行う。
- ◆ 音声コーパス
JTES (Japanese Twitter-based Emotional Speech)
- ◆ 結果
DNN モデル, LightGBM モデルともに提案手法で認識精度 (Accuracy) が向上した。LightGBM モデルでは発話末尾の特徴量を 1,582 次元から 76 次元へと大幅に減らした場合でも認識精度が向上した。

Table 1: Accuracy of classification

Model	Method	Without Tail*	With Tail*	Num of features (tail)
A	DNN	62.8 %	65.1 %	1,582
B	LightGBM	63.8 %	67.4 %	1,582
C	LightGBM	63.8 %	66.1 %	76

*Tail: Features from the end (tail) of speech.

3-T1-6

3-T1-6 マルチモーダルデータを用いた話者年齢推定法の検討

Study of using multi-modal data for speaker age estimation

○俵直弘, 小川厚徳, 北岸佑樹, 神山歩相名 (NTT)

- ◆ 話者認識コーパス VoxCeleb に年齢ラベルを付与した新たなマルチモーダル年齢コーパス AgeVoxCeleb を構築
- ◆ 特徴: 年齢推定で従来用いられてきた音声コーパス NIST-SRE と比べて、より幅広い年齢層を含み (Fig.1), 音声と顔動画の両方が利用可能 (Fig.2)
- ◆ AgeVoxCeleb を用いた音声・顔画像年齢推定
 - 現在の SOTA の顔年齢推定モデルと音声年齢推定モデルの性能を本コーパスで比較することで、顔情報を利用したほうがドメインの違いに対して頑健に年齢推定できることを明らかにした (Tbl.1)
 - 顔年齢推定で得られた年齢ラベルで音声年齢推定モデルを学習する cross-modal 学習が可能であることを示した (Tbl.2)

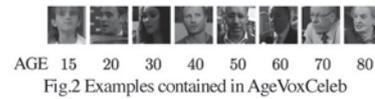


Fig.2 Examples contained in AgeVoxCeleb

Tbl.1 Mean absolute error (MAE) of face / speech age estimation

Modality	Test condition	
	In-domain	Out-domain
Face	7.10	7.59
Speech	5.80	11.05

Tbl.2 MAEs of cross-modal experiment

	Grand truth	Cross modal
	8.19	8.82

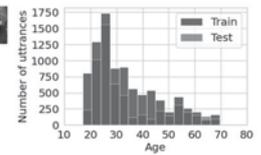


Fig.1 Age histogram of (upper) conventional SRE and (lower) proposed AgeVoxCeleb

3-T2-8

3-T2-8 話者性別推定とのマルチタスク学習による話者年齢推定

Speaker age estimation by multi-task learning of gender estimation

◎北岸佑樹, 神山歩相名, 安藤厚志, 俵直弘, 森岳至, 小橋川哲 (NTT)

- ◆ 性別を考慮した話者年齢推定のため、**話者性別と年齢のマルチタスク学習と推定誤差を許容する範囲を性別依存とする手法**を提案
- ◆ 従来の課題: 多くの話者年齢を直接推定する手法では、性別による音響特徴量の違いが考慮されていない
- ◆ 提案手法
 - マルチタスク学習による話者の性別を考慮した年齢推定
 - 推定誤差を許容する範囲 (σ) を性別に応じて可変として最適化
- ◆ 実験: σ を固定として性別依存 (GI), 性別非依存 (GD), マルチタスク学習 (MTL) の比較と, σ を学習可能として GI, MTL の比較
- ◆ 実験結果
 - MAE (mean absolute error) と相関係数 ρ を Tab.1 に記載
 - 提案手法によって MAE が改善
 - σ を学習可能とした時, GI と比べて MTL の方が改善幅大

Method	σ	Male		Female	
		MAE	ρ	MAE	ρ
GI	固定	4.94	.78	4.27	.87
GD	固定	4.85	.78	4.32	.87
MTL	固定	5.02	.77	4.10	.88
GI	学習	4.91	.79	4.18	.88
MTL	学習	4.83	.78	4.01	.88

Tab.1: Results of Age Estimation Experiments

3-T2-9

3-T2-9 ビブラートの自然さを計測する知覚モデルの誤差評価

Error evaluation of perceptual model for measuring vibrato naturalness.

☆宮崎嵩大(山梨大), 森勢将雅(明治大)

- ◆ビブラートの構成パラメータには、ビブラートの速さ・深さが存在する。本研究はこれらのパラメータとビブラートの自然さの関係について、行われた主観評価実験を用いて知覚モデルを構築する。
- ◆主観評価実験のデータ群は歌手2人で構成されており、2人分とそれぞれの歌手の計3種類のデータでモデル構築した。
- ◆知覚モデルは複数のガウス分布による非線形フィッティングに基づいて構築した。
- ◆歌手ごとのデータによる知覚モデルの評価誤差の分布がそれぞれ異なることを確認し、歌手が自然さの要因となる可能性を示唆した。
- ◆歌手ごとのデータによる知覚モデルの性能が全歌手のものと比較して低く、誤差関数の算出法を工夫する必要性を示唆した。

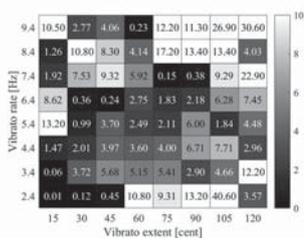


Fig. 1 : Heat map of singer 1's error.

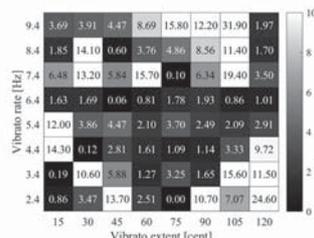


Fig. 2 : Heat map of singer 2's error.

3-T2-11

3-T2-11 Human-in-the-loop 型歌唱デザインのためのDNNを活用した歌声パラメータ生成法の提案

Study on singing parameter generation using DNN for human-in-the-loop singing design

☆小川樹(山梨大), 森勢将雅(明治大)

- ◆本研究では、人間がデザインした結果に基づき、DNN が歌声パラメータを生成する Human-in-the-loop 型歌唱デザインを提案する。
- ◆Fig. 1 のような流れで、歌声合成システムの中に人間が介在し、人間がデザインしたパラメータを基に、歌声を合成する。
- ◆本稿では、デザインする対象を F0 に限定し、加工した F0 を基に音響特徴量を予測する F0 条件付け音響モデルを構築した。
- ◆推定した F0 は Fig. 2 のように、入力とほとんど差はなかったが、この僅かな差が音質に影響するか、今後主観評価実験で明らかにしていきたい。

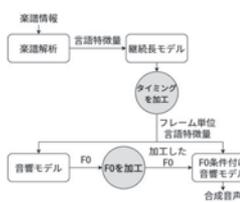


Fig. 1 提案手法の流れ

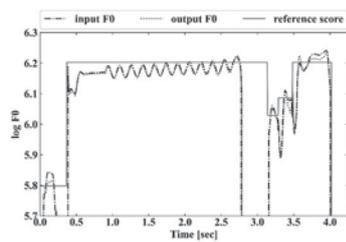


Fig. 2 推定した F0 の結果の比較

3-T2-10

オンライン・オフライン編集機能を統合したビブラートデザインインタフェースの評価

Evaluation of a vibrato design interface combining online and offline editors.

☆小野雄大(山梨大), 森勢将雅(明治大)

- ◆前報では、オンライン・オフライン編集機能を統合する手法を提案し、実時間歌唱デザインツール - Parrot V3 を開発した。本研究では、Parrot V3, オフライン編集, オンライン編集におけるビブラートデザインのユーザビリティを評価する比較実験を行った。
- ◆Table 1 より、各インタフェース間でデザインした音声と手本音声とのビブラート特徴の誤差は、同程度だった。Figs.1, 2, 3 より、Parrot V3 の作業時間と作業回数, SUS のスコア値は、オフライン編集とオンライン編集それぞれの良い結果と同程度だった。
- ◆Parrot V3 は、オフライン編集とオンライン編集のメリットを享受でき、ビブラートデザインの更なる省力化に寄与する可能性を示した。

Table 1: RMS errors of vibrato features between the target and designed voices

パラメータ	Parrot V3	オフラインエディタ	オンラインエディタ
振幅 [cent]	16.668	19.042	21.028
周波数 [Hz]	0.783	0.583	0.724
長さ [%]	12.192	9.577	11.815
周波数傾斜 [Hz/s]	0.017	0.012	0.012

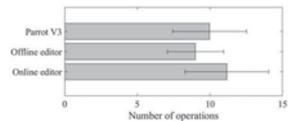


Fig.2 Number of operations in each interface

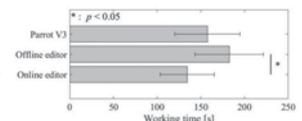


Fig.1: Working times in each interface

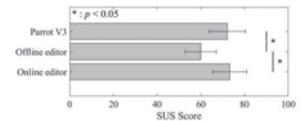


Fig.4: SUS scores in each interface

3-T2-12

3-T2-12 歌声と背景楽曲のタイミングとノート持続時間との関係性の解析

Analysis of the effect of duration on perception of shift length between singing and background music

☆小林礼奈(山梨大), 森勢将雅(明治大)

- ◆本研究では、歌声と背景楽曲のタイミングを知覚できる時間について、楽曲による差があるのかを調査するための音声評価実験を行った。
- ◆評価には、テンポが異なる10曲を用いた。各曲サビ1フレーズを切り取り、背景楽曲から時間シフトさせた音声を作成した。
- ◆被験者は、歌声と背景音楽のタイミングがシフトされていることを認識した上で、聴取した楽曲の違和感について評価した。
- ◆結果として、歌声と背景音楽のシフト時間の許容範囲(図中の縦軸が概ね対応する)と楽曲のノートの個数、ノート単位の平均持続時間との相関が示唆された。
- ◆音符の長さが短い楽曲ほど、許容範囲が低くなることが考えられる。

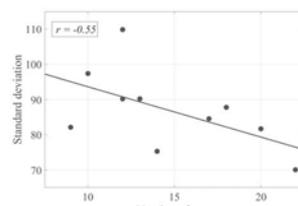


Fig.1: Relationship between standard deviation and number of notes

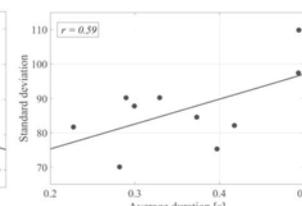


Fig.2: Relationship between standard deviation and average duration

3-T3-13

3-T3-13 発声の量と発音の質に関する特徴を用いた日本人英語の主観流暢さ予測の改善

Improved Prediction of Perceived Fluency of Japanese English using
Quantity of Phonation and Quality of Pronunciation

☆瀧陽, 安カ川彩乃, 齋藤大輔, 峯松信明(東大), 齊藤一弥(UCL)

Recent and rapid globalization requires teachers and engineers to provide language learners with online tools for assessing their oral proficiency. Out of various assessment metrics, we focus on fluency. In our previous study, a picture description task was imposed on 90 Japanese students and 10 native speakers, and their utterances were manually rated in terms of fluency by 10 native raters. Speech features were manually extracted and their correlations to the averaged fluency over the raters were examined. This paper realizes automatic prediction of perceived fluency using the same speech corpus and automatically extracted speech features, which are related to quantity of phonation and quality of pronunciation. As all the utterances in the corpus are very noisy, we build DNN-based acoustic models through multi-condition training. With the models, the utterances are converted to their posteriorgrams with multiple granularities. This hierarchical speech representation is shown to be very effective, and a high correlation of 0.925 is obtained between human scores and machine scores, which is higher than the average of inter-rater correlations, which is 0.873.

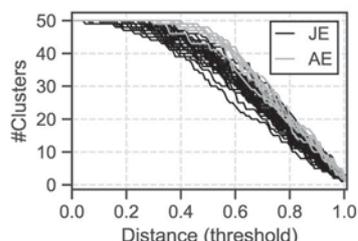
3-T3-15

3-T3-15 Acoustic Word Embeddingを用いたL2単語音声の語彙密度分析

Analysis of lexical density of L2 word utterances based on acoustic word embeddings

◎安藤慎太郎, 齋藤大輔, 峯松信明(東大・工)

- ◆ Acoustic Word Embedding (AWE)
 - … 可変長単語発話→固定長ベクトル (Embedding)
 - ◇ 単語の識別を目的として
検索語抽出などの分野で検討が進められている
- ◆ 近年はAWEと人間の認知との対比も検討されている
 - ◇ AWEの単語識別傾向 ↔ 人間の認知傾向
 - ◇ ゼロリソースAWEの学習過程 ↔ 幼児の言語獲得
- ◆ 本研究では対象言語を英語とし、
AWEモデル→成人母語話者の単語識別・認知モデル
と解釈してL2単語音声の語彙密度を分析する
- ◆ AWEによるL2発音分析が可能な一例を示した



Hierarchical clustering of AWEs for each speaker.
JE=Learner (Japanese), AE=Native (American)

3-T3-14

3-T3-14 Frame-based shadowability annotation using shadowing and script-shadowing L2 utterances

○ Zhenchao Lin, Ryo Takashima, Daisuke Saito, Nobuaki Minematsu
(The University of Tokyo), Nakanishi Noriko (Kobe Gakuin University)

- ◆ Language teachers often claim that the goal of speech training should be intelligible enough pronunciations, not native-sounding ones, because some types of accented pronunciations are intelligible or comprehensible enough.
- ◆ However, if one aims to provide a technical framework of automatic assessment based on intelligibility or comprehensibility, s/he has to be faced with a big technical challenge. That is collection of L2 utterances with annotations based on these metrics.
- ◆ Further, learners always want to know which parts (words, morphemes, or syllables) in their speech should be corrected. This means that data collection needs a valid method of intelligibility annotation with fine granularity.
- ◆ In our previous studies, a new metric of shadowability was introduced, and it was shown experimentally to be highly correlated to perceived intelligibility or comprehensibility as well as it was explained theoretically to be potential to give annotations with fine granularity.
- ◆ In this paper, shadowability annotation with fine granularity is examined experimentally, and a new and more valid method of collecting shadowing utterances is introduced. Finally, we tentatively derive frame-based shadowability annotation for L2 utterances.

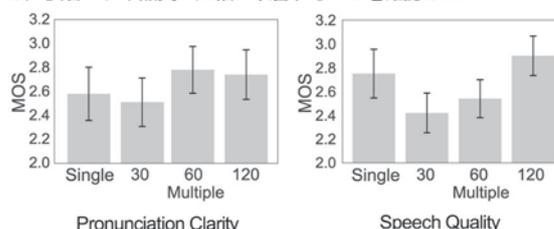
3-T3-16

3-T3-16 多数話者健全音声を用いた CycleGAN に基づく構音障害音声の明瞭性改善

Improving Pronunciation clarity of Dysarthric Speech
Using CycleGAN with Multiple-Speaker Healthy Speech

☆今井柊平, 能勢隆, 金垣葵(東北大), 渡辺聡(HTS), 伊藤彰則(東北大)

- ◆ 本研究では、構音障害者の音声の明瞭性を改善するために CycleGAN に基づくノンパラレル声質変換を応用して構音障害者の音声を健全者の音声へと変換する手法を検討した。
- ◆ 構音障害者にとって発話は負担が大きい行為であるため、大規模データデータセットを作成することは難しい。
- ◆ 本手法ではノンパラレル声質変換を用いることで少量の構音障害音声と多量の健全音声で学習を行うことを可能とした。
- ◆ 健全音声として単一話者の長時間音声を用いた場合に比べて、多数話者の音声を用いてデータ量を増やすことで効果的に明瞭性・品質を改善できることを示した。
- ◆ 声質変換手法としてボコーダパラメータを使用した CycleGAN-VC2 とスペクトログラムとニューラルボコーダに基づく Scyclone を比較し、後者により品質が大幅に改善することを確認した。



3-T3-17

3-T3-17 文章の複雑さを考慮した 雑談対話音声からの認知症傾向検出

Dementia Detection from Chat Dialogue considering complexity of sentences
☆石原颯人, 入部百合絵(愛知県), 北岡教英(豊技大)

- ◆徳島市, 木更津市, 鈴鹿・四日市市に住む高齢者 44 名の雑談対話音声を集集, 音声からの言語情報を用い, 認知症傾向を検出した. 本研究では特に認知症傾向を捉えるため平均最大係り受け距離を用いた.
- ◆被験者の認知症傾向の有無は, 音声収録時に実施した. 改定長谷川式簡易知能スケールに基づいて判断しており, 認知症傾向ありが 25 名, 認知症傾向なしが 19 名であった.
- ◆構文解析器 MeCab を利用して, 品詞割合や Type Token Ratio(TTR) などの語彙に関する特徴量に加え, 係り受け解析器 CaboCha を利用して, 平均最大係り受け距離を抽出した.
- ◆複数の識別器を用いた予備実験の結果, 精度の最も高かった SVM の RBF カーネルを利用して識別実験を行った. 評価方法は 10 分割交差検証である.
- ◆語彙に関する特徴量と平均最大係り受け距離を各々単独で用いるよりも組み合わせた識別結果が最も高く, Precision が 82%, 認知症傾向ありの Recall が 92%の結果を得た.

Table 1. Classification results

	Precision	Recall	F-Measure
Dementia	0.79	0.92	0.85
w/o Dementia	0.87	0.68	0.76
Weighted Average	0.82	0.82	0.81

3-U1-1

3-U1-1 音響シーンを用いて検出誤りの深刻さを 考慮したイベント検出の評価指標

Evaluation metric of sound event detection considering severe misdetections by scenes

◎砺波紀之(立命館大・情報理工), 井本桂右(同志社大学),
福森隆寛, 山下洋一(立命館大・情報理工)

- ◎音響イベント検出に向けた新たな評価指標の提案
→ イベントとシーンの関係に基づいて深刻な誤りの考慮
- ◎誤検出を深刻さの異なる二種類に分割
→ 軽度の誤り (FP₁): ターゲットのシーンで発生するイベントの誤検出
e.g., 「street」というシーンで誤検出された「car」
深刻な誤り (FP₂): ターゲットのシーンで発生しないイベントの誤検出
e.g., 「office」というシーンで誤検出された「car」

従来: イベントとシーンの
関係を考慮しない指標

$$Fscore = \frac{2TP}{2TP + FP + FN}$$

$$IoU = \frac{2TP}{2TP + 2FP + 2FN}$$

提案: イベントとシーンの
関係を考慮した指標

$$HM = \frac{2TP}{2TP + FP_1 + 2FP_2 + 2FN}$$

2種類の誤検出に異なる重みを付与
→ 深刻な誤りに大きな罰則

Table 1: sound event
detection results for "car"

Metric	Model 1		Model 2	
	office	street	office	street
# TPs	0	200	0	200
# FP ₁ s	200	0	0	200
# FN ₁ s	0	50	0	50
# TNs	250	250	250	250
F-score	61.54%	61.54%		
IoU	44.44%	44.44%		
HM	47.06%	61.54%		

TP: true positive
FP: false positive
FN: false negative
TN: true negative

3-T3-18

3-T3-18 発声支援のための口形素列によるフレーズ認識方式の検討

Speech Enhancement System Using Lip-reading

☆中原智典, 福山晃平, 松井謙二(阪工大 R&D 工学部),
中藤良久(九工大), 加藤弓子(聖マリアンナ医科大学)

- ◆本研究では, PC, スマホによる低コスト, 周囲の視線が気にならない外観, カメラ, ディスプレイ連携による使いやすいUIなどを特徴とする読唇方式で発声支援実現のための基礎検討を行った. 今までの報告では, 発声するフレーズや単語を学習する場合に少量のデータで使用者に最適化する手法の提案を行った. しかし, 語彙数, 認識精度に課題があったので今回は語彙の拡大と学習用データ作成手法の改良を行った.
- ◆モデルの学習には, 各母音(A, I, U, E, O)と閉唇状態(X)の6種類の口形と口形を2音節ずつ組み合わせた計36種類の動画を使用した. 認識対象語彙の特徴量系列データに対してノイズ付加, 時間軸適応, 移動平均処理等によって学習用データ生成の工夫を行った. Table 1に示すように20単語認識で良好な認識精度が得られた.

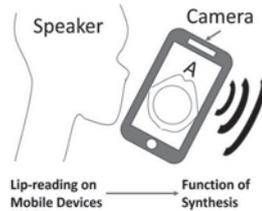


Fig. 1 Image of the speech enhancement system

Table 1 Word recognition results

True	Result	
	First	Second
1 A-RI-DA-TO	A-RI-DA-TO	O-RI-TO
2 I-I-E	I-I-E	KI-I-RO
3 O-RI-TO	O-RI-TO	A-I-O
4 O-RI-TO	A-RI-TO	SI-ME-TO-TO
5 O-RI-TO	O-RI-TO	O-RI-TO
6 O-RI-TO	O-RI-TO	I-I-E
7 O-RI-TO	O-RI-TO	O-RI-TO
8 O-RI-TO	O-RI-TO	O-RI-TO
9 O-RI-TO	O-RI-TO	O-RI-TO
10 O-RI-TO	O-RI-TO	O-RI-TO
11 A-I-O	O-RI-TO	A-I-O
12 O-RI-TO	O-RI-TO	O-RI-TO
13 O-RI-TO	O-RI-TO	O-RI-TO
14 O-RI-TO	O-RI-TO	O-RI-TO
15 O-RI-TO	O-RI-TO	O-RI-TO
16 I-I-E	I-I-E	I-I-E
17 O-RI-TO	O-RI-TO	O-RI-TO
18 O-RI-TO	O-RI-TO	O-RI-TO
19 O-RI-TO	O-RI-TO	O-RI-TO
20 O-RI-TO	O-RI-TO	O-RI-TO

3-U1-2

3-U1-2 雑音・残響下での異常音検知を目的とした 機器動作音識別の検討

Study on identification of operating sounds
for detecting abnormal sounds under noise and reverberation

☆田中一希, 宮崎亮一(徳山高専)

- ◆近年, DNN を用いた異常音検知の研究が進む中, 雑音や残響が存在したりするように, 実プラント内での運用には多くの問題が発生する.
- ◆本研究では実プラント内を模擬した環境において DNN を用いて機器の動作音から正常音や異常音を識別することを扱う.
- ◆識別対象機器以外の動作音や雑音を含み, 識別対象の機器動作音をビームフォーミングで強調されたデータを学習データとする.
- ◆実プラント内を模擬したシミュレーション実験により, 残響下においての構築されたモデルの識別精度の検証を行う.
- ◆残響毎の識別予測確率 (Fig.1) より本研究で示した学習方法は残響下においても頑健な異常音検知を行えると考えられる.

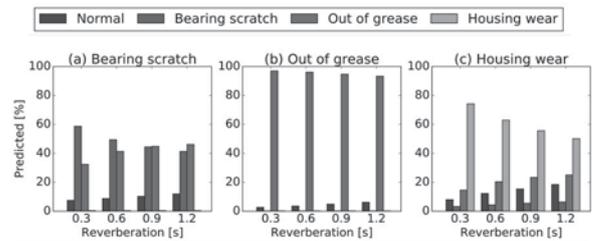


Fig.1: Identification of Motor sound with varying reverberation time.
(a) Bearing scratch, (b) Out of grease, and (c) Housing wear.

3-U1-3

3-U1-3

調波量み込みの高速化

Harmonic Lowering for Accelerating Harmonic Convolution for Audio Signals

○竹内博俊(東大院・情報理工), 大石康智(NTT),
柏野邦夫(NTT), 猿渡洋(東大院・情報理工)

調波量み込み: CNNで調波構造を捉えることのできる量み込み手法

◆調波量み込みの問題点: 計算速度が遅かった

- 近似なし高速化手法を提案。
- 速度・メモリ使用量ともに従来法より効率的!
- 通常の量み込み(下限)と比べても遜色なく速い!

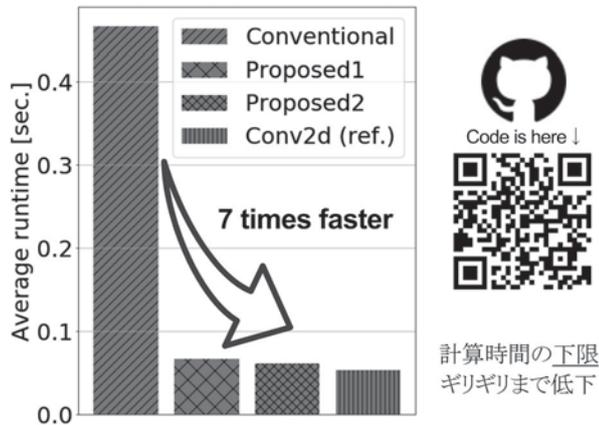


Fig.1: Performance Comparison by Runtime

3-U1-5

3-U1-5 音源分離のための周波数間相関を考慮した多変量複素 Gauss 分布に基づく深層学習による分散共分散行列推定の検討

Estimation of frequency-covariance matrix based on multivariate complex Gaussian distribution by deep neural networks for audio source separation

☆成澤 直輝, 高宗 典玄(東大), 北村 大地(香川高専),
中村 友彦, 猿渡 洋(東大)

- ◆教師あり音源分離手法である独立深層学習行列分析(IDLMA)[Makishima+, 2019]では、音源の複素スペクトログラムの周波数間相関を陽にはモデル化していない。一方、教師なし音源分離手法である独立半正定値テンソル分析[keshita, 2018]の研究では、周波数間相関を陽にモデル化することによる性能の向上が示されている。
- ◆本研究では、IDLMAを拡張させて周波数間相関を陽にモデル化する教師あり音源分離手法の構築のための前段階として、音源の従う周波数方向の分散共分散行列を推論するDNNについて検討する。特にDNN出力の正定値性を保証するため、分散共分散行列 R_{jn} のCholesky分解行列 L_{jn} を推論するDNN及びその誤差関数を設計した(Fig.1参照)。
- ◆DNN自体の推論精度を評価するため、周波数方向の多チャネルWienerフィルタによる音源分離実験を行った。

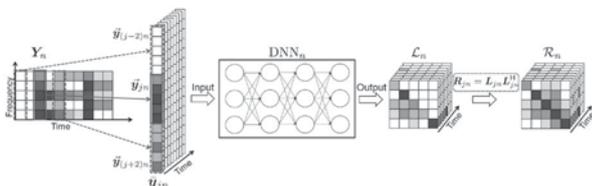


Fig.1: Proposed DNN for frequency-covariance matrix estimation.

3-U1-4

3-U1-4 残響劣化音に対する推定原音声を用いたノンレファレンス音声了解度推定

On Non-Reference Speech Intelligibility Estimation For Reverberant Speech Using Estimated Clean Speech

☆中澤和司, 近藤和弘(山形大)

- ◆残響の影響を考慮するために「ここにはOOと書いてある」といった文章に単語を埋め込んだテスト用音声を録音。
- ◆残響劣化させたテスト用音声に対するDRT試験を7名に実施し了解度を測定。
- ◆Bi-LSTMを用いて残響劣化音から原音声を推定。
- ◆原音推定によってSTOIの値がすべての残響時間で改善、最大で約0.35改善。
- ◆推定した原音声と残響劣化音を用いてフルレファレンスを模擬した了解度推定のためのDNNを学習。
- ◆残響劣化音に対する了解度をRMSE、相関係数それぞれ0.00353、0.99945の精度で推定可能。

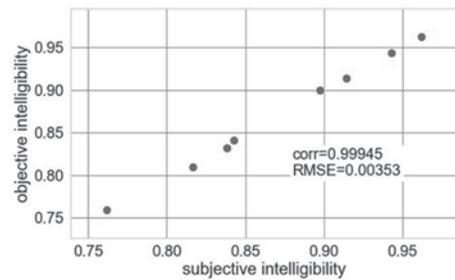


Fig.1 Distribution of subjective vs. estimated speech intelligibility

3-U1-6

3-U1-6 スマートフォン収録された自動車走行音からのDeep Neural Networkによる自動車距離推定

Estimation of distance from car running sound recorded by a smart phone based on Deep Neural Network

△中島 悠貴, ○實廣 貴敏(愛工科大)

- ◆目的: 自動車の接近を歩行者のスマートフォンで検知できるよう、自動車走行音から自動車の距離推定を行う。
- ◆収録: Apple社製iPhone 8、Apple Watch Series 3を用い、トヨタ社製プリウスPHV(2012年製)、ハイエース(1996年製)の走行音を収録、iPhoneなどを設置した前後20mの距離で測定。
- ◆手法: 距離ラベルをつけた学習データを用い、Deep Neural Networkを学習。(1)音声分析。32kHzサンプリング周波数とし、1フレームあたり80次元メル周波数スペクトル、(2)中心フレームの前後5フレームずつをまとめ、880次元メル周波数スペクトルを1ベクトルとする。(3)DNNによりフレームごとに距離推定を行う。
- ◆実験: 2車種の走行音に対して9種類の距離ラベル(-20~+20mを5m間隔に分割、それ以外)を与え、学習、評価を行なった。DNNは3階層で隠れ層は100ユニットとした。Fig.1に正解率(Accuracy)を示す。どの場合も高い精度が得られた。

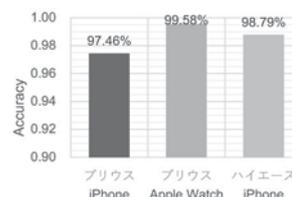


Fig.1 Accuracy of estimated distance by DNN for running sound of Prius PHV (iPhone), Prius PHV (Apple Watch), and Hiace (iPhone).

3-U2-7

3-U2-7 深層自己符号化器に基づく音響特徴量の離散符号化

Discrete encoding of acoustic features based on deep autoencoder

◎木下裕磨, 小野順貴(都立大)

- ◆本研究では、我々が研究開発を進めている音光変換デバイス「プリンキー」によりセンシングされた音信号から時間周波数特徴を抽出し、この特徴量をデジタル通信する仕組みを検討する (Fig. 1 参照).
- ◆プリンキーには、音響情報を光により伝送する過程で信号が劣化してしまう課題がある。デジタル通信により劣化に頑健となることを期待できるが、プリンキーによるデジタル通信では、単位時間あたりに送信可能なデータ量が極めて小さい。
- ◆そこで、本研究では、音響特徴量を深層自己符号化器に基づき小さいデータ量の離散符号として表現し、この離散符号により音の時間周波数特徴をどの程度伝送できるか調査する。
- ◆自己符号化器の学習時には、階段関数を近似した微分可能な関数を用いることで、量子化の影響を含めた学習を実現する。
- ◆実験により、提案する離散符号化法は、量子化なしの通常の自己符号化器と同等の性能で音響特徴量を再構成できることが確かめられた。

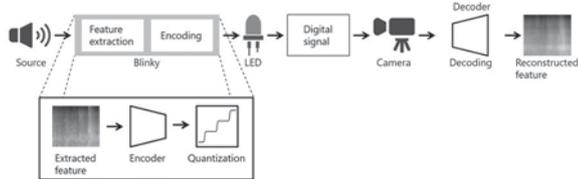


Fig.1: Pipeline for proposed discrete encoding of acoustic feature by using deep autoencoder

3-U2-9

3-U2-9 ベイズ最適化を用いたシャント音識別アルゴリズムのハイパーパラメータ探索

Hyperparameter search for shunt murmur discrimination algorithm using Bayesian Optimization

☆野田史也, 西島恵介, 古家賢一(大分大)

- ◆腎臓病の患者は日常的にシャントの機能を管理する必要があるが、判定には正しい知識と経験が必要のため、シャント機能を自動的に検査するシステムの実現が望まれている。
- ◆本報告では、従来手法で用いられたグリッドサーチではなく、ベイズ最適化を用いてシャント音識別アルゴリズムのハイパーパラメータ探索を行う。シャント音識別アルゴリズムとしてランダムフォレストを用いた検討を行う。
- ◆実験結果より、提案手法を用いることで精度の向上が確認された。また、ハイパーパラメータ探索時間の短縮も確認された。

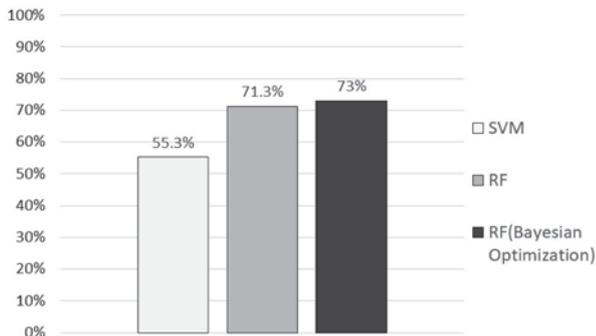


Fig.1: Results of the accuracy rate.

3-U2-8

3-U2-8 非定常音を含む単変量時系列に対する動的モード分解を用いた特徴量抽出及び可視化の検討

Feature extraction and visualization using dynamic mode decomposition for univariate time series with non-stationary sounds

◎伊藤弘章, △村田伸, 大石康智, 原田登 (NTT)

- ◆非定常音を含む単変量時系列に対する動的モード分解 (Dynamic Mode Decomposition : DMD) により、観測信号を指数減衰及び増加成分と周期振動成分をもつ基底の和として分解することで、振幅の時間変化定数と周波数を物理単位にもつ特徴量として抽出し、可視化する手法を提案する。

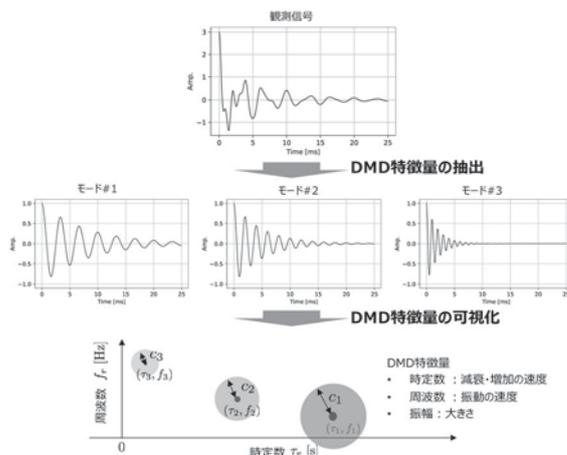


Fig.1: An illustration example of feature extraction of dynamic mode decomposition and its visualization

3-U2-10

3-U2-10 開腹手術後の時間経過を考慮した蠕動音特徴量の探索

Bowel-sound-feature selection considering postoperative elapsed time

☆齊藤慎之介, 大塚翔, △前佛聡樹(千葉大),

△堀創史, △本多通孝(福島県立医大), 中川誠司(千葉大)

- ◆腸蠕動音は腸機能の評価に用いられているが、診断に有用な蠕動音特徴量を客観的に比較・検討した例は極めて少ない。開腹手術後の時間経過に伴って蠕動音発生が活発化することを利用して、腸機能評価に有用な蠕動音特徴量の探索を行った。
- ◆術後1および2日 (n=27) と術後4日以上 (n=71) の腹部聴取音を取得した
- ◆腹部聴取音から16個の特徴量を抽出し、類似度の低い9個の特徴量を選択した (Fig. 1).
- ◆術後日数を目的変数、9個の特徴量を説明変数とし、ランダムフォレストを適用、accuracyが72%、precisionが76%の分類器が得られた
- ◆分類器から特徴量重要度を算出した (Fig. 2)。現在の蠕動音の診断基準にも関連する Silence Duration, Zero Crossing Rate が最も重要と判定された。また、現在の診断基準との関連の低い、いくつかの特徴量が有用である可能性が示唆された。

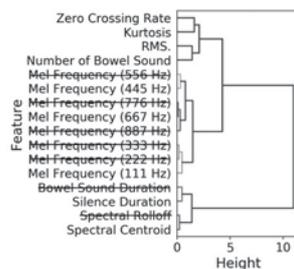


Fig. 1. A dendrogram of features

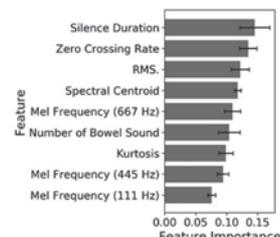


Fig. 2. Importances of each feature

3-U2-11

3-U2-11 データ拡張を用いた複数位置のシャント音特徴量による狭窄の識別

Classification of stenosis using multi positions shunt sound features using data augmentation

○西島恵介, 古家賢一(大分大)

- ◆血液透析患者はバスキュラーアクセス(患者と透析装置との間で血液循環するための仕組み)を管理し, 狭窄の早期発見が必要である。
- ◆狭窄する位置は様々なため狭窄位置に影響されない識別手法が必要である。
- ◆複数位置の聴診から得られたシャント音を用いて, 時間周波数領域において, ある位置のパワースペクトルから残りの位置のパワースペクトル平均を減算する特徴量による狭窄の識別を試みた。
- ◆データ数の少なさに起因する過学習への対応や精度向上を図るために, 時間軸方向の移動・伸縮によるデータ拡張を提案した。
- ◆時間軸方向の前後0.01秒の移動, 0.01秒の伸縮によるデータ拡張と元データにより学習モデルを作成し識別(提案法)とデータ拡張なしの識別(従来法)を比較した。
- ◆識別にはサポートベクターマシンを用い, 最大6箇所まで計測された20症例(正常7, 狭窄13)のデータを用いた。
- ◆従来法の精度60%に対し, 提案法は精度65%と向上した。

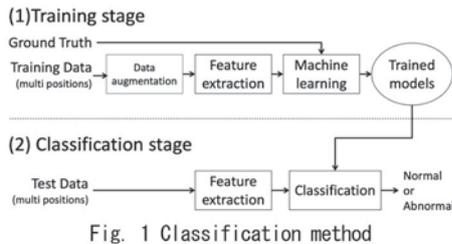


Fig. 1 Classification method

3-U2-13

3-U2-13 光レーザマイクロホンのための深層学習による多段階音声強調の検討

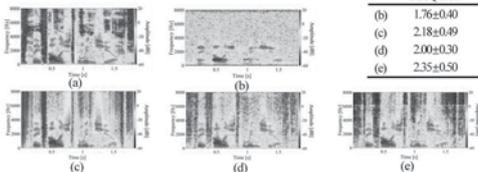
Speech enhancement based on multi-stage process used deep neural network for optical laser microphone

☆CAI Chengkai(立命館大院), 岩居健太, 西浦敬信, 山下洋一(立命館大)

光レーザマイクロホンはレーザ光を利用し, 物体の振動を計測することにより音声を受音するシステムであり, マイクロホン周囲にある雑音の影響を受けずに遠方からの音を受音することができる。しかし, 光レーザマイクロホンによる観測音声は, 被照射物体の形状及び振動特性により, 雑音の混入と高域成分の欠落が発生し, 音質が劣化する。従来, 深層学習による音声強調手法が提案されているが, これらの手法は多種類の歪みを含む光レーザマイクロホンの観測音声に対しては十分な性能を示すことができない。本稿では, 時間波形に基づいた深層学習による光レーザマイクロホンのための多段階音声強調手法を提案する。まず, 畳み込みニューラルネットワーク(CNN)により雑音除去を行う。次に, 雑音除去した音声を用いて回帰型ニューラルネットワーク(RNN)により高域情報を復元する。提案法の有効性を確認するために, 強調後の音声を PESQ, LSD, STOI の3つの指標による評価実験を実施した。実験結果(Fig. 1, Table 1)より提案法の有効性を確認した。

Table 1 Evaluation results of each method.

	PESQ score	LSD [dB]	STOI score
(b)	1.76±0.40	1.62±0.10	0.85±0.04
(c)	2.18±0.49	1.09±0.09	0.90±0.02
(d)	2.00±0.30	1.21±0.09	0.93±0.03
(e)	2.35±0.50	1.11±0.08	0.94±0.03



(a),(b),(c),(d),(e) are the clean speech, observed speech, and the results of conventional spectrogram-based DNN, conventional waveform-based DNN, proposed waveform-based DNN

Fig. 1 Enhanced speech with each method.

3-U2-12

3-U2- Deep Image Prior に基づく事前学習不要な12 ミュージカルノイズ除去の性能評価

Performance evaluation for removal of musical noise without pre-training based on Deep Image Prior

☆藤村拓弥, 宮崎亮一(徳山高専)

- ◆非線形雑音抑圧手法は高いノイズ除去性能を誇るが, ミュージカルノイズが発生する問題がある。
- ◆我々はミュージカルノイズを除去する手法として Deep Image Prior に基づく手法を提案したが, 十分な性能評価は行われていない。
- ◆本稿では, 従来手法であるメディアンフィルタに基づく手法との比較や主観評価実験を通して DIP に基づく手法の性能評価実験を実施する。
- ◆性能評価実験では, 損失関数としてL1 ノルムを用いることで
(i) 音質を維持したままミュージカルノイズを除去でき (Table 1, 2),
(ii) 主観的な品質も改善できることが明らかになった (Fig. 1).

Table 1: Result of musical noise reduction

α	10^{-3}		
	10 dB	20 dB	30 dB
SNR	0.405 (0.025)	0.310 (0.019)	0.269 (0.013)
$L_{1.5}$	0.305 (0.005)	0.251 (0.002)	0.237 (0.001)
L_2	0.258 (0.002)	0.240 (0.001)	0.242 (0.001)
median	0.431 (0.008)	0.422 (0.007)	0.426 (0.007)

Table 2: Result of PESQ

α	10^{-3}		
	10 dB	20 dB	30 dB
original	3.663 (0.017)	3.904 (0.012)	4.084 (0.008)
L_1	3.701 (0.025)	3.754 (0.023)	3.778 (0.027)
$L_{1.5}$	3.279 (0.023)	3.303 (0.024)	3.310 (0.025)
L_2	3.145 (0.020)	3.163 (0.025)	3.175 (0.030)
median	3.535 (0.029)	3.552 (0.030)	3.561 (0.030)

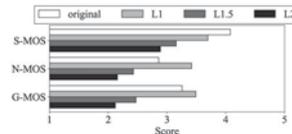


Fig. 1: Result of subjective evaluation.

3-U3-14

3-U3-14 音場シミュレーションを用いたゲームにおける空間表現方法の検討

A study of spatial representation method for games with sound field simulation

○櫻井赴, △松下翔, 小島健二(カブコン)

- ◆ゲーム内の仮想空間上で, 音の波動性を考慮した, 現実と同様な聞こえを再現する方法を検討した。
- ◆ゲーム内地形上に設置した音源による音波の伝搬を, FDTD 法を用いた数値解析により事前計算した。
- ◆数値解析によって求まる瞬時音響インテンシティベクトルから, 解析領域内の参照点における音の到来方向を, 球面調和関数で表現した。
- ◆聴取点と, 聴取点を囲う4つの参照点が結ぶ四面体の体積比を利用して, 解析領域内の任意の聴取位置における球面調和関数を補間した。
- ◆補間した球面調和関数を用いて, 任意の聴取位置に到来する音波の聞こえをアンビソニックス再生によって再現した。

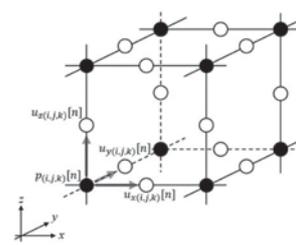


Fig. 1: Sound pressure and velocities at reference point (i, j, k)

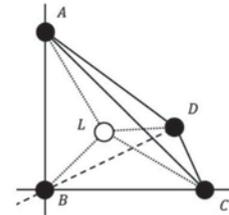


Fig. 2: Listener position L and reference points A, B, C, D surrounding the player

3-U3-15

3-U3-15 パラメトリックスピーカとラウドスピーカを 組み合わせたサラウンドシステムの 配置に関する検討

A study on the arrangement of a surround system
combining parametric speaker and loudspeaker.
☆加藤弘士朗, 上原正志, 大川茂樹(千葉工大)

- ◆パラメトリックスピーカ(Parametric Array Loudspeaker: PAL)による指向性の鋭い放射音を受聴者の後方で反射させ、サラウンドシステムにおける後方スピーカとして利用。
- ◆居室においてPALとラウドスピーカのインパルス応答測定から配置に関する検討。
- ◆反射点によって差はあるものの、PALの反射音が最大振幅となった。
- ◆最大振幅の前に直接音成分も含まれるが微小なものに留まった。

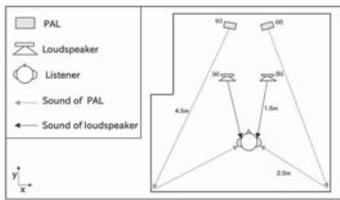


Fig. 1: Overview of the proposed system

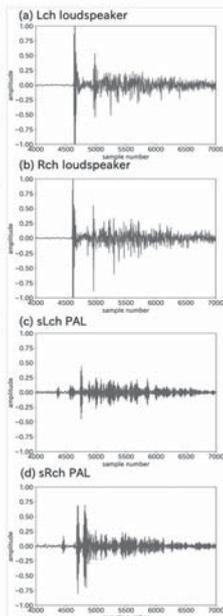


Fig. 2: Impulse response of PAL and loudspeaker

3-U3-17

3-U3-17 パラメトリックスピーカの ビーム制御の検討

Study on Beam Control of Parametric Array Loudspeaker
☆岡野愛, 梶川嘉延(関西大)

- ◆指向性の高い音を放射するパラメトリックスピーカ (PAL) のビーム制御を行うにあたりサンプリング周波数により放射角度の制限があった。そこで、本研究ではサイドバンドと搬送波を分離してビーム制御を行う分離型フェイズドアレイ技術について検討を行う。
- ◆解析では放射角度5度に設定して従来のフェイズドアレイ技術、従来の分離型フェイズドアレイ技術、新たに提案した分離型フェイズドアレイ技術の3種類のビーム制御について放射方向の違いを比較検討を行った。
- ◆解析の結果、従来法と比べて提案法を用いた場合の方が適切な方向により高い音圧かつより鋭い音響ビームの形成できることがわかった。これにより今回提案した分離型フェイズドアレイの有効性を示すことができた。

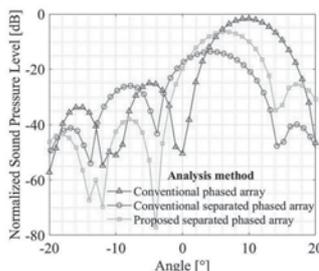


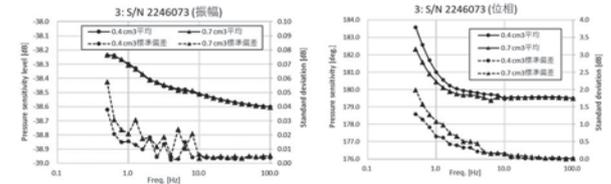
Fig. 1: Comparison of directivity among several phased array techniques.

3-U3-16

3-U3-16 カプラ校正法による 標準マイクロホン音圧感度校正の 下限周波数に関する実験的考察

Experimental investigation on lower limit frequency of the coupler calibration method in determining LS microphones' pressure sensitivity
○高橋弘宜, 山田桂輔, 堀内竜三(産総研)

- ◆超低周波音は防災・減災の観点で有用視され、その研究開発の進展に伴い、超低周波音を計測する装置の計測信頼性が重視されている。
- ◆可聴域を含む低周波数側の音響標準として、標準マイクロホンの音圧感度の振幅を1 Hzから20 kHzの周波数域で供給している。防災・減災目的の超低周波音の信号解析は時間領域で行われることもあり、下限周波数として0.1 Hz、音圧感度の振幅と位相の両方が必要である。
- ◆カプラ校正法でLS2マイクロホンの音圧感度の振幅と位相を校正したときの低周波数側の下限を、容積の異なる2つの音響カプラの整合性の観点から実験的に検討した。
- ◆0.5 Hzから100 Hzにおいて、音圧感度の振幅は異なる容積の音響カプラを使っても0.01 dB以内で一致、一方位相は周波数低下に伴ってずれが大きくなり、0.5 Hzでは1.3°に達することを確認した。



Pressure sensitivities determined by the coupler calibration method with 0.4 cm³ and 0.7 cm³ couplers.

3-U3-18

3-U3-18 エリア收音による遠方の目的音源強調のための 低周波数帯域 PSD 推定値の補正方法の検討

Compensation of power spectral density in low frequencies to improve the performance of enhancement of sound sources located within distant two-dimensional area.
◎千葉大将, 野口賢一, 中川朗(NTT)

- ◆干渉雑音源に囲まれた目的音源のみを收音するためのリアルタイム音源強調技術として、互いに離れた位置に配置した2つの小型マイクロホンアレイを用いて、ビームフォーマと局所パワースペクトル密度(PSD)推定に基づくポストフィルタ処理を組み合わせた手法がある。
- ◆この手法では、目的音源がアレイの遠方に位置するとビームフォーマの指向性が鈍い低周波において音源強調性能が劣化する。
- ◆ビームフォーマの指向性が鋭い周波数帯域のPSD推定値を用いて低周波のPSD推定値を補正し音源強調性能を改善する手法を提案する。
- ◆目的音源と2つのアレイ中央の距離 d が0.5~3.5 mの条件で、提案法による音源強調性能の改善を確認した。

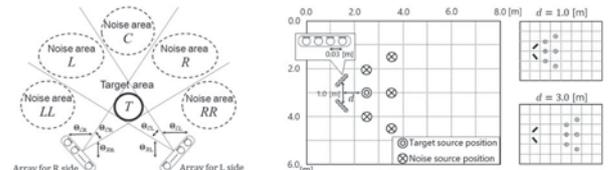


Fig. 1: Definition of Areas.

Fig. 2: Positions of two microphone arrays and sources.

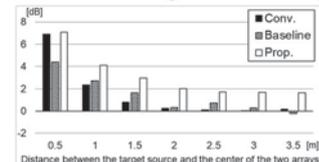


Fig. 3: Source enhancement performance (SDR).

3-U3-19

3-U3-19 近接 2ch 観測間のクロススペクトル法を用いた音響測距における環境雑音の影響の基礎的検討

Fundamental study on the effect of environmental noise on Acoustic Distance Measurement between adjacent 2ch observations using cross-spectral method

☆山門 樹, 篠原 寿広, 上保 徹志, 中迫 昇(近畿大)

- ◆様々な工学分野において対象物までの距離情報は重要であり、送信波と反射波の位相干渉に基づく音響測距法が提案されている。
- ◆実環境において、雑音の混入は不可避である。雑音を低減する手法として、加算平均が挙げられるが、複数回の測定が必要となる。
- ◆測定系の影響による雑音を除去するための手法として、近接2ch観測間のクロススペクトル法を用いた音響測距法を用いる。
- ◆本稿では、雑音が距離推定に及ぼす影響を調べるために、様々なS/Nの白色ガウス雑音を近接2chの観測信号に加えて、音響測距法を適用するシミュレーションを行った。
- ◆Table 1 にマイクロホン 1, 2 にそれぞれガウス雑音を加えて S/N を -10~10dB の範囲で変化させたときの測距精度を示す。理論上、Mic1 の雑音を低減させることができれば、高雑音下において音響測距法による距離推定が可能となることを確認した。

Table 1 Estimation accuracy v.s. S/N.

Mic.1 [dB]	Mic.2 [dB]	Accuracy [%]
10	10	100
10	-10	88.91
-10	10	61.11
-10	-10	3.63

3-U4-21

3-U4-21 瞬時スペクトル推定法による同一到来方向の複数音源の分離に関する考察

Consideration on separation of multiple sound sources in the same direction by instantaneous spectral estimation

○小澤賢司(山梨大・工), 渡邊貴治(秋田県立大・システム科技), 坂本修一(東北大・通研)

- ◆マイクロホンアレイからの出力信号を 2 次元画像と見なし、2 次元 FFT により得たスペクトル上で、複数音源についてのスペクトルを推定する。この推定を、分析の時間フレームごとに独立に行うため、瞬時スペクトル推定法と称している。
- ◆4 マイクロホンによるアレイの出力について、連立方程式を解いて、3 音源のスペクトルを推定した。前報では振幅特性のみを推定することで雑音抑圧を行ったが、十分な性能が得られなかった。ここでは、位相特性も推定することで雑音抑圧に加えて、音源分離まで行った。
- ◆目的音源 (Source #1) を 2 m, 雑音源 (Source #2) を 3 m に固定し、雑音源 (Source #3) を移動させて雑音抑圧量を調べた計算機実験結果を Fig. 1 に示す。前報に比べて良好な雑音抑圧を達成した。

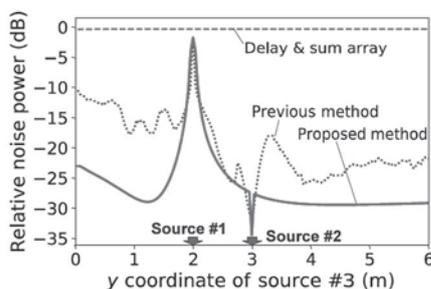


Fig.1: Noise suppression as a function of distance of one noise source.

3-U3-20

3-U3-20 複合現実技術と FDTD シミュレーションを用いた音響インテンシティの可視化

Visualization of Sound Intensities with Mixed Reality and FDTD Simulation.

☆大川祐貴子, 渡辺靖明, 池田雄介(東京電機大),

片岡優太, 及川靖広(早大理工), 小坂直敏(東京電機大)

- ◆目的: 複合現実技術を用いて音場の数値シミュレーション結果を可視化し、実測結果と比較
- ◆提案システム: SLAM による室形状情報に基づいた 2D-FDTD シミュレーションを行い、時間信号から音響インテンシティを算出し、複合現実技術を用いて実測結果とともに可視化
- ◆実験:
 - ・3.1m×7.7m×2.8mの小会議室で可視化
 - ・壁面の垂直入射吸音率を0.1, 0.3, 0.5で音響インテンシティを比較



Fig. 1: Visualization of sound intensities with the proposed system in the small conference room.

3-U4-22

3-U4-22 瞬時スペクトル減算法による到来方向が異なる 2 雑音の抑圧に関する考察

Study on suppression of multiple noises with different directions of arrival using instantaneous spectral subtraction

☆竹中幸輝, 小澤賢司(山梨大・院・医工農学総教)

- ◆マイクロホンアレイの観測信号を 2D FFT することによりスペクトルを得る。その直流成分には、(1) アレイに直交する方向 (0°) から到来した目的音の全スペクトルと、(2) その他の方向から到来した雑音の遅延和アレイ出力に相当する成分が混在している。ニューラルネットワークにより、(2) の成分を推定し減算することで雑音抑圧を行う。
- ◆単一雑音を想定した前報のシステムでは、複数の雑音源に対応することは不可能であった。その原因が、空間スペクトルの形や変化の系統性の差異によるものであると考察し、2 音源の信号による 2 次元スペクトルを学習データとすることで、複数雑音源への対応を図った。
- ◆目的音 (0°) と 2 つの雑音抑圧 (-90° ~ 90°) の計 3 種類の白色雑音を入力した際の雑音抑圧量を Fig. 1 に示す。同側と対側それぞれにおいて、ある程度までの抑圧が可能になったといえる。

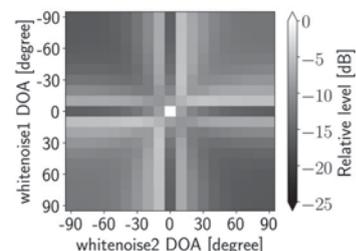


Fig.1: Noise suppression for two white noise inputs.

3-U4-23

3-U4-23 複数の搬送波周波数を利用する パラメトリックスピーカの音波伝搬解析

Sound Wave Propagation of Parametric Array Loudspeaker Using Multiple Carrier Frequency.

☆今井健太, 梶川嘉延(関西大)

- ◆複数の搬送波周波数を利用するパラメトリックスピーカ (PAL) を用いることで極小領域オーディオスポットを構築する方法がこれまで検討されてきている。
- ◆本研究では数値解析を通じて、複数の搬送波を利用する PAL における最適なエミッタ配置について検討を行う。
- ◆解析には MATLAB のツールボックスである k-wave を利用する。音軸方向の音圧分布を比較することで Length-Limited の効果を検討する。
- ◆解析結果を Fig. 1 に示す。ここで case2 は case1 に比べ内側と外側のエミッタの比率が近くなるように想定している。解析の結果、内側と外側のエミッタの比率を近づける程 Length-Limited の効果が顕著になることがわかる。

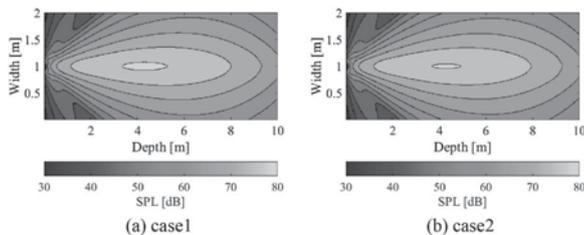


Fig. 1: Comparison of sound pressure distribution of fundamental frequency 2 kHz.

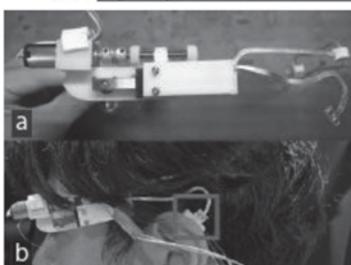
3-U4-25

3-U4-25 耳甲介舟裏への軟骨伝導提示における音響特性へのトランスデューサ押し付け力の影響の評価

Evaluation of the effect of applied force of a transducer behind the back of cymba conchae on cartilage conduction hearing

☆野崎悦, △高木健, △川原圭博(東京大)

- ◆軟骨伝導は耳の軟骨に骨伝導トランスデューサを当てることで音の聴取を可能にする手法である。
- ◆耳介の窪みである耳甲介舟の裏側への音提示は、一般的な骨伝導に比べトランスデューサの押し付け力が小さく、外観上も目立ちにくい。ため、常時装用に適した音声インターフェースに適している。
- ◆軟骨伝導による音の伝搬特性はトランスデューサを当てる位置や押し付け力によって大きく変化することが知られているが、これらを安定させた状態で音の伝搬特性を測定することは難しかった。
- ◆そこで本稿ではトランスデューサを当てる位置と押し付け力を眼鏡型測定治具を用いて正確に調節することを可能にした。
- ◆また、耳甲介舟裏における軟骨伝導の 1 kHz での最小可聴値を測定した結果、直接気導経路以外の経路の存在と 0.4 N の押し付け力における最小可聴値の飽和が示唆された。



る最小可聴値の飽和が示唆された。

(a) The proposed eye-glasses jig for measurement. (b) A person wearing the jig.

3-U4-24

3-U4-24 縮小ダミーヘッドを用いた 頭部伝達関数の計測

Head-related transfer function measurement with a downscaled dummy head

○森川大輔, 堀田元輝, 寺島眞生, 平原達也(富山県立大)

- ◆縮小ダミーヘッドを用いて頭部伝達関数を計測した。
- ◆実寸で計測した頭部伝達関数と遜色のない頭部伝達関数を計測できることがわかった。
- ◆縮小によって距離が近くなるために、実寸で計測を行う場合よりも低域の SN が高くなることがわかった。

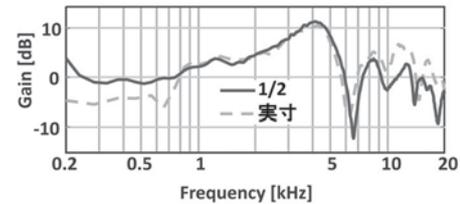


Fig. 1: Head-related transfer function in the forward (0°) direction.

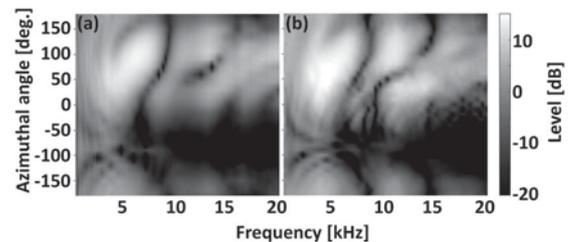


Fig. 2: Head-related transfer function in the horizontal plane. (a) a down scaled dummy head and (b) real head.

3-U4-26

3-U4-26 等価回路と有限要素法を併用した マイクロスピーカーの 周波数応答解析の一検討

A Study on Frequency Response Analysis of Micro-speaker Using Equivalent Circuit and Finite Element Method

☆南恵介, 梶川嘉延(関西大)

- ◆スマートフォンなどの小型音響機器は小型化並びに複雑化が進んでおり、従来の音響等価回路解析法による設計では実測特性が正確に再現できないという問題点がある。そこで本稿では、電気-機械等価回路と有限要素法(FEM)を組み合わせた周波数応答解析を検討する。
- ◆マイクロスピーカーに電圧を入力したときの振動現象を電気-機械等価回路で、気室や音孔といった音響構造部分を有限要素法による数値解析で表す。機械系に流れる電流はスピーカユニットの振動板における振動速度に対応し、有限要素法の周波数応答解析に代入する。
- ◆周波数特性の比較を Fig. 1 に示す。Fig. 1 より、提案法では第一ならびに第二共振周波数が実測と近く、他の手法と比べて正しく再現できることがわかった。今後は振動板の減衰を考慮し、構造をより精密にすることで実測特性の再現を目指す。

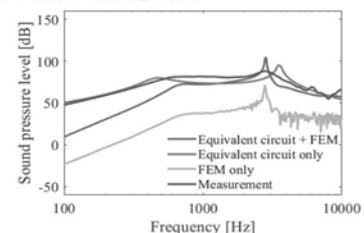


Fig. 1: Comparison of frequency characteristics

3-U4-27

3-U4-27 フレキシブル超指向性スピーカーを用いた可変オーディオスポットの検討

Towards a Variable Audio Spot Using Flexible Parametric Array Loudspeakers

☆宮川剛志, 近藤和弘(山形大院・理工学研)

- ◆フレキシブル超指向性スピーカーを用いることでオーディオスポットの範囲を対象に合わせて変化させる。
- ◆角度変化によるフレキシブル超指向性スピーカーの指向性の変化を確認。
- ◆純音 21 個を 3 分割し、分割した音波をそれぞれ別の超指向性スピーカーから出力する。
- ◆音圧測定によりオーディオスポット内の音圧が高いことを確認。
- ◆フレキシブル超指向性スピーカーの形状に応じオーディオスポットの範囲の変化を確認。
- ◆オーディオスポット外の音圧が約 47dB であり、オーディオスポット内と比べ 23dB 差と十分小さい。

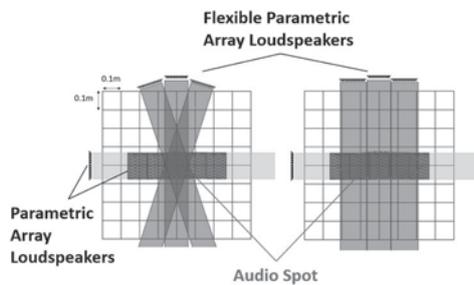


Fig.1: Variable Audio Spot

3-V1-2

3-V1-2 熊本城ホールの音響計画

Acoustic design of Kumamoto-Jo Hall

○青木亜美(日建設計)

- ◆熊本城ホールは 4 つの異なる用途の空間が積層した構成となっており、高度な防振遮音性能と各室の室内音響計画が要された。
- ◆大音量拡声音楽イベントが想定された 2F シビックホールは防振遮音構造とし、①防振遮音層を 70 kg/m²、②固定—防振の離隔を 600 mm 以上、③浮床ゴムピッチを粗にするなど、低音域遮音性能向上の工夫を行った。その他の室も各所で固定遮音層の強化や防振遮音層の設置を行った上、マスキングノイズとしての空調騒音の下限管理も行った。
- ◆4F メインホールでは拡声音楽イベント時に一体感と臨場感が得られるよう、中高音域の散乱や後壁の吸音+散乱を行った。さらに、バルコニー下の席でも同様の一体感が得られるよう、バルコニー席床を壁から離し、音が籠らないよう試みた。

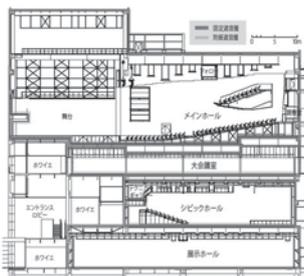


Fig.1: Section of Kumamoto-Jo Hall and specification of sound insulation

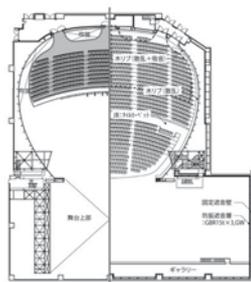


Fig.2: Plan of Main Hall

3-V1-1

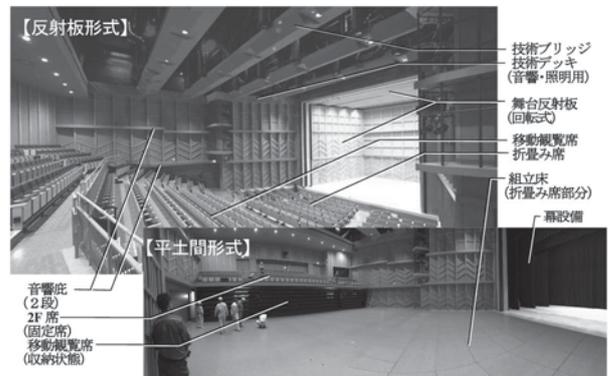
3-V1-1 三条市体育文化会館の音響設計

—音楽・講演会～運動利用まで「重ね使い」可能なマルチホール—

Acoustical design of the Sanjo City Sports and culture hall - The multi-hall that can be multiused from music/lectures to exercise -

○高橋顕吾, 日根野翔太(ヤマハ)

- ◆本施設はアリーナ棟とホール棟からなる複合施設であり、マルチホール（最大 500 席）は音楽・講演会から運動利用までスポーツと文化の枠を超えた様々な市民利用（＝重ね使い）に対応するために回転式反射板、幕設備、移動可動席・折り畳み席、可動間仕切壁、組立床、等の多数な可変機構を導入している。
- ◆特に反射板形式では音楽リサイタルが可能な専用ホール並みの音響性能が求められ、上記の機構や建築・設備条件により生じる舞台開口部や客席中通路前後での段差により音場が不連続になることが音響上の課題となった。
- ◆設計段階での音響解析による内装形状の検討とともに、施工最終段階での測定や試奏による音場調整（回転式反射板の角度調整）を行い、所期の音響性能（反射板形式 RT=1.5 秒、C80=0dB、LE=20%前後）が実現されている。



3-V1-3

3-V1-3 戸田建設新音響実験棟の音響性能について —音の体験ができる実験施設—

Acoustic characteristics of new acoustic laboratory of Toda Corporation
-Experimental facility where visitors can experience various acoustic phenomena-

○佐藤真平, 土屋裕造, 山内崇, 小泉穂高(戸田建設)

- ◆戸田建設筑波技術研究所に新たに作られた音響実験施設について、施設の概要と音響特性の一部を報告する。
- ◆新施設は実験施設であると同時に建物で発生する様々な音や響きを体験できる見学施設としての機能を併せ持つことを特徴とする。
- ◆シミュレーター室は 36.6 ch サラウンドシステムによる立体音場再現施設であるが、床板を撤去することで無響室としても利用することができる。シミュレーター室の音響特性を測定したところ、逆二乗特性をはじめとする種々の音響特性で良好な結果を得た。

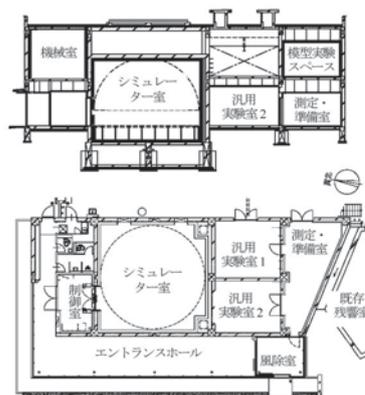


図1 断面図・1階平面図

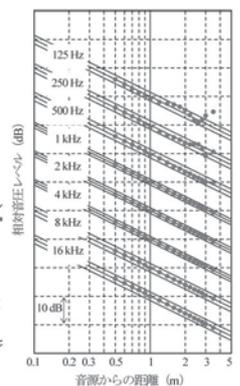


図2 逆二乗則特性

3-V1-4

3-V1-4 神楽殿の舞台音響に関する研究

Research of stage acoustics on Kagura theatre

○下倉良太(大阪大)

- ◆神楽殿の舞台音響の特徴を把握するため、松江市内にある美保神社・熊野大社・佐太神社の神楽殿を対象に音響調査を行った。
- ◆調査方法はインパルス応答測定法で、神楽殿の舞台上に12面体スピーカー、受音点にバイノーラルマイクを装着した実験者と3Dマイクを配置した。また受音点は、舞台上、舞台周辺、本殿周辺に分類した。
- ◆美保神社は、毎朝夕と音楽に合わせて巫女が鈴を鳴らしてお供えをする朝御饗祭・夕御饗祭が行われるため、神楽殿(舞殿)の舞台上、その周辺で他の神楽殿よりも残響時間が長かった。
- ◆熊野大社は、壁4面全てが開放するため、舞台周辺でも音が届いて残響時間が長く、周囲で観覧する観客に配慮した建築になっていることがうかがえた。
- ◆一方、佐太神社の神楽殿は舞台が本殿のほうを向いており、本殿の基礎となる石垣と、舞台の高さが近いため、本殿付近に音が届きやすく、残響時間も長かった。



Fig. 1 Kagura theatres in (a) Miho shrine, (b) Kumano shrine and (c) Sata shrine

3-V2-6

3-V2-6 ペンローズタイル型拡散体を設置した 小会議室の室内音響特性

Acoustics characteristic of the small meeting room to set up Penrose-tiling-type diffusers

○土屋裕造(戸田建設), 佐久間哲哉(東京大学)

- ◆非周期構造で高い拡散性と反射方向の均一性を有するペンローズタイル型拡散体を開発し、実用性を考慮した形状で有孔板の吸音機構を付加した拡散体を実際の小会議室に設置したので、その室内音響特性について調査した。
- ◆残響時間の低減および残響時間が短くなりすぎないこと、またフラッターエコーの抑制、といった効果が得られたことを確認した。

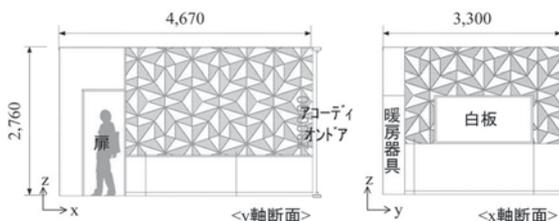


図1 小会議室 ペンローズタイル拡散体を設置した状態

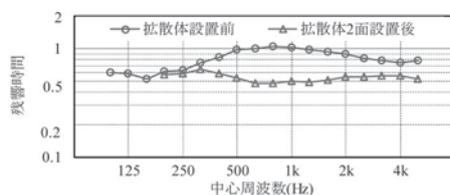


図2 小会議室の残響時間 拡散体設置前後

3-V1-5

3-V1-5 音場支援システムの初期反射音制御手法 -高吸音性空間における空間印象の制御 に着目した比較検討-

Method to control early reflections for acoustic enhancement system: Comparative investigation focused on controlling spatial impression in highly absorptive space

◎橋本梯, 渡辺隆行, 宮崎秀生(ヤマハ)

- ◆ヤマハの音場支援システム Active Filed Control (AFC)は室内音響特性を柔軟に制御するため、初期反射音の制御に In-line 方式, 残響音の制御に Regenerative 方式を採用している。
- ◆制御中に実空間での音響測定データに基に生成した FIR フィルタの畳み込み処理を行っており、初期反射音の FIR フィルタは空間印象を決定づける重要な役割を果たす。
- ◆空間印象の制御に着目し、効果的な初期反射音の制御手法を検討するため、3つの初期反射音の制御手法を同一の高吸音性空間に構築した。
- ◆500席(幅15.5m, 奥行き26m, 高さ5.5m)の講堂で音響測定を実施し、その測定データを用いて実験室(幅7.2m, 奥行き11.8m, 高さ8m)に音場支援システムを構築した。
- ◆室内音響測定および主観評価実験により各制御手法の効果と特徴をまとめた。



Fig.1: An auditorium (left) and a laboratory (right) used for the experiment.

3-V2-7

3-V2-7 音源方向判定システムの 開発とその適用例

Development and Application of Sound Source Direction Determination System.

○池上雅之(大林組技研)

- ◆利用者のいる建物では、きしみなどの異音・響きすぎなどの音響障害・音漏れなど様々な音の問題がしばしば発生する。騒音計では到来方向が分からないため音源を特定しにくい。またまれに生じる異音のような現象に人を貼り付けておくことも困難などの課題があり、工事現場では対応に苦慮していた。
- ◆今般、これらの問題の調査ツールとして音源方向判定システムを開発した。本システムは、到来音の方向判別を可能とするC-C法と呼ばれる分析手法を利用して、音の到来方向を判定してマーカーで表現し、映像と重ね合わせて表示するシステムである。音の問題の調査のため、工事現場などに適用し始めたので、その概要を紹介する。

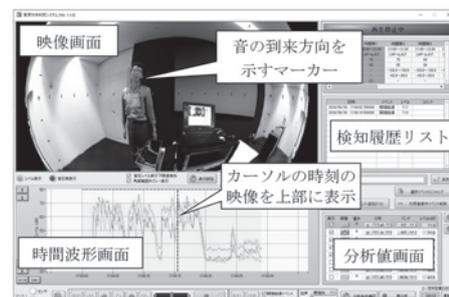


Fig.1: Display of This System (Reacting to the Speaking Mouth).

3-V2-8

3-V2-8 室内音響測定のための 自律移動ロボットの検討

Study of an Autonomous Mobile Robot for Room Acoustic Measurement

☆石川直生, 上原正志, 大川茂樹(千葉工大), 福島学(日本文理大)

- ◆室内音響測定を自律移動ロボットが行うためのアルゴリズムの設計, 実装したシステムで測定実験を行った
- ◆Q学習による測定点および経路の獲得
- ◆測定実験

研究室内で20点の測定 所要時間約28分 音場マップの生成

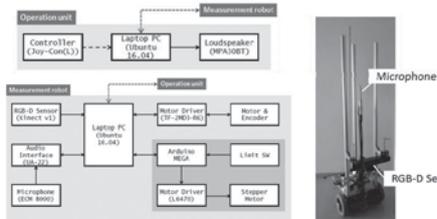


Fig.1 Measurement diagram using the mobile robot

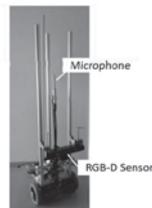


Fig.2 Appearance of the measurement robot

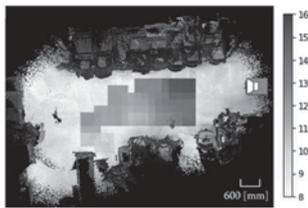


Fig.3 Visualization result of C50

3-V2-9

3-V2-9 移動ロボットを用いた 3次元室内音響測定

3-D acoustic measurement by using mobile robot

☆上原正志, 石川直生(千葉工大), 福島学(NBU), 大川茂樹(千葉工大)

- ◆移動ロボットを用いた室内音響測定の効率化
- ◆マイクロホン鉛直軸上に駆動
 - 車輪による平面移動+ボールねじ機構による高さ方向の移動
- ◆SLAM(環境地図作成, 自己位置推定)の利用
 - 室内音響指標の空間分布を環境地図上に可視化

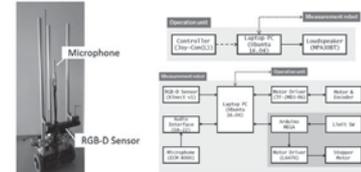


Fig.1 mobile robot with microphone for acoustic measurement



Fig.2 Visualization result of C50[dB] distribution

3-W1-1

3-W1-1 重子音に関わる先行母音末の フォルマント遷移について —日本語とイタリア語の比較—

Formant transitions at the end of vowels preceding geminate consonants in Japanese and Italian

☆角田正路, 岩上恵梨, 荒井隆行(上智大), 守本真帆(国立国語研究所)

- ◆日本語とイタリア語に共通する特徴に、単語中に重子音(日本語:促音, イタリア語:geminazione)が現れることがある。重子音判定の第一の手がかりとされる子音の長さをはじめ、先行母音長や後続母音長、調音器官の動きは両言語とも重子音と単子音の発話における区別や知覚における判別に関与している。
- ◆日本語の重子音においては、発話で先行母音末尾にフォルマント遷移がみられ、知覚で先行母音末尾のフォルマント遷移の有無が影響することが明らかとなった(柳澤・荒井 2015)が、イタリア語の重子音においては、フォルマント遷移についての報告がない。
- ◆イタリア語学習経験のある日本語母語話者と日本語学習経験のあるイタリア語母語話者に日本語とイタリア語の両方で重子音を発話してもらい分析した結果、いずれの話者・言語においても、先行母音末尾にフォルマント遷移がみられた。L1としてのイタリア語の重子音の発話でもフォルマント遷移がみられたことで、知覚にフォルマント遷移が影響を与えている可能性が考えられる。
- ◆フォルマント遷移の傾きの値を定量化した結果、その程度には母語による違いが示唆されるが、L1, L2による差はあまりない。このことから、L2のフォルマント遷移についてはL1のフォルマント遷移を転移していることが考えられる。

3-W1-2

3-W1-2 英語発音評価における文の 発音難易度推定のための特徴量の検討

Analysis of sentence-based features to estimate difficulty of sentences on English pronunciation evaluation

☆内城 沙月, 伊藤 彰則, 能勢 隆(東北大)

- ◆本研究では、話者本来の発音能力を測るような自動英語発音評価を目指すために、従来の発話単位の発音評価における文の発音難易度の影響や発音難易度を表す特徴量の検討を行った。
- ◆話者と文を因子とした二元配置分散分析では、発音評価に対し、話者の発音能力と文の発音難易度はどちらも影響することが示された。
- ◆分散分析の結果を踏まえて、発音評価スコアから文の発音難易度に起因するスコア(文スコア)の抽出を行い、文スコアを表す特徴量の検討を行った。
- ◆特徴量としては、音素の種類数・調音特徴・単語親密度・Surprisalが文スコアと相関があり、特徴量として有用であることがわかった。
- ◆特徴量から重回帰分析による発音難易度推定を行った結果、文スコアとの相関係数は0.48となった。

Table 1: Correlation coefficients of sentence-based features

特徴量	相関係数	偏相関係数
文の長さ	-0.24	0.02
音素の種類数	-0.37	-0.30
調音特徴	0.33	—
単語親密度	-0.32	—
Surprisal	-0.29	—

3-W1-3

3-W1-3 3Dプリンタによる声道モデルの音響特性の計測実験に対する3次元FDTD法を用いた検証

Verification of acoustic characteristic measurement of 3D printed physical models of the vocal tract by using the 3D FDTD method
☆後藤康泰, 坂野秀樹, 旭健作 (名城大院)

- ◆3Dプリンタを用いて声道モデルを作成し、その音響特性を計測する
 - 荒井が復元した千葉、梶山による日本語5母音の声道モデル
- ◆2種類のシミュレーション結果との比較を行う
 - 1次元FDTD法によるシミュレーション
 - ◇計測結果より高い周波数になる傾向がみられた
 - ◇3kHz以上(第4フォルマント以降)で顕著になった
 - 3次元FDTD法によるシミュレーション
 - ◇計測結果と概形がよく一致した
- ◆以前に報告した測定結果とシミュレーション結果の差異は1次元FDTD法の計算時に生じる誤差に起因すると考えられる

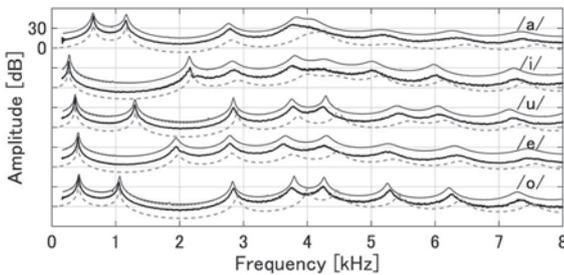


Fig.1 Acoustic characteristic of physical models (bold lines) and transfer functions computed by 1D FDTD method (chain lines) and 3D FDTD method (thin lines)

3-X1-1

3-X1-1 F0パターンの操作に基づく自身の声らしさの表現と個人差に対する検討

The expression of one's own voice based on the F0 manipulation and individual differences

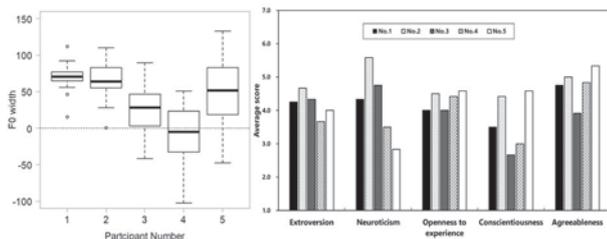
☆守井成輝(香川大・工), 佐藤敬子(香川大・創造工)

- ◆目的

自分の声らしさの表現における個人差について検討するため、性格特性と自身の声の知覚との関係を調査した。
- ◆実験結果

自身の声に最も近いと評価されたF0の調整幅は、参加者ごとに異なることが示された。性格特性の観点からは、録音音声よりF0を低く調整した参加者は、外向性と勤勉性の得点が低い結果となった。
- ◆今後の課題

参加者を増やして傾向を見るときも、女性を参加者に加え、性差も含めたF0の調整幅の違いの調査を行う。



Results of F0 adjustment (Left side) and personality traits (Right side)

3-W1-4

3-W1-4 スマートフォンのピッチ角による感情レベル指定型音声コミュニケーション支援アプリの提案

Voice communication assist application with the control of the degree of emotion by smartphone pitch angle

○河瀬卓也, 岩城護(新潟大・自然研)

- ◆構音障害者のコミュニケーションにおいて、感情表現の一助となることを目的としたスマートフォンアプリの開発を行った。
- ◆従来のTTS(Text to Speech)機能に、スマートフォン本体のピッチ角の角度によって音声合成時の韻律パラメータを変化させる機能を加えることによって、感情レベル毎に応じたテキストの読み上げを可能とした。
- ◆アプリケーションの機能は、自由入力型TTSとフレーズ選択型TTSの2種類を搭載することで利便性の向上を図った。
- ◆各感情レベルに合わせて音声を合成するために、感情評定値付きオンラインゲーム音声チャットコーパス(OGVC)の音響的特徴量を解析し、TTSの編集パラメータを決定した。

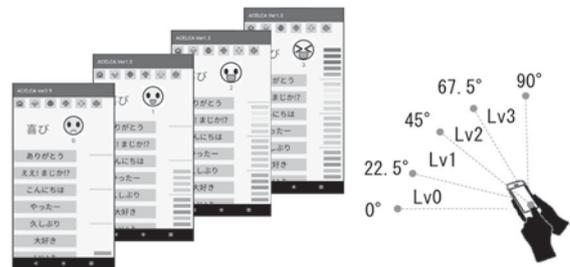


Fig.1: Phrase & emotion level selection (e.g. Emotion: JOY, Level: 0~3)

3-X1-2

3-X1-2 振動覚刺激による有声および無声両唇閉鎖音の知覚への影響

Effect of vibro-tactile stimuli on the perception of voiced and voiceless bilabial stop consonants

◎小野明日香, △中谷桃子, △中根愛(NTT サービスエボリューション研究所), △渡邊淳司, 廣谷定男(NTT コミュニケーション科学基礎研究所)

- ◆近年、音声知覚における聴覚と体性感覚の統合についての報告がなされている。
- ◆本研究では、これまでほとんど検討されてこなかった有声および無声両唇閉鎖音の知覚における振動覚刺激の影響を調べた。
- ◆実験の結果、有声子音を含む1モーラ音声の声帯振動と合致するような時間長を持つ振動覚刺激を手を与えることで、両唇閉鎖音が有声に知覚されやすくなった。一方で、声帯振動よりも十分に短い振動覚刺激は効果が認められなかった。
- ◆本結果は、振動覚刺激が音声知覚に寄与する可能性を示唆している。

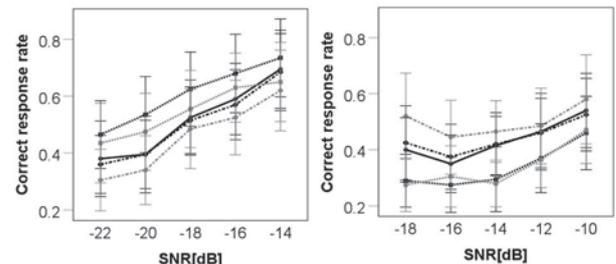


Fig.1: Effect of vibration presentation and SNR in dB on correct response rate for speech stimulus /ba/ (left) and /pa/ (right). Error bars indicate standard errors. (Vibration presentation Solid line: Speech-only, Black dotted: With long vibration, Gray dotted: With short vibration, Black dotdash: With long vibration "sound", Gray dotdash: With short vibration "sound")

3-X1-3

3-X1-3 反復単語変形効果における実験試行回数
の影響の検討

The examination of the influence of the experiment trial number in the verbal transformation effect

○船津誠也(県立広島大), 藤本雅子(早稲田大)

- ◆反復単語変形効果の生起条件が、初めてこの種の実験に参加する被験者(ナイーブな被験者)と以前に実験に参加した被験者との間で異なるかどうかについて検討を行なった。
- ◆被験者はナイーブな被験者7名と2度目の参加となる被験者5名である。ISIを変化させ、知覚遷移が生じるまでに知覚した刺激音声の数を計測した。刺激音声は前回までと同様男性話者1名が発話した/banana/である。
- ◆ANOVAによる検討を行なったところISIによる主効果($F(4, 48)=0.744$ ($p=0.5671$)), 実験参加回数による主効果($F(1, 48)=0.719$ ($p=0.4007$))に有意差は生じなかった。交互作用も有意ではなかった($F(4, 48)=0.779$ ($p=0.5441$))。

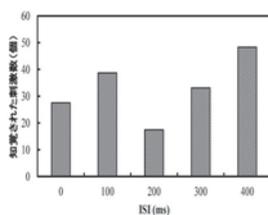


図1 ナイーブな被験者の結果

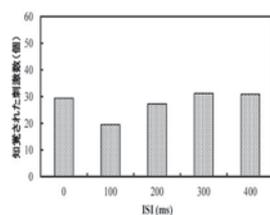


図2 2回目参加者の結果

3-X1-4

3-X1-4 耳内嚥下音の評価方法について

Evaluation of intra-aural swallowing sounds

○羽石英里(昭和音大), 河原英紀(和歌山大), △城本修(県立広島大), 萩原かおり(昭和音大), 三浦優佳(昭和音大)

- ◆嚥下のタイミングをとらえる指標として推奨される耳管の開大音を、耳管にマイクロホンを挿入して録音した。
- ◆これを、先行研究にあるPraatを用いた分析方法に加え、新たに本発表のために開発した嚥下音抽出ツール swallowingChecker を用いて分析した。
- ◆swallowingChecker では、耳管の開閉に伴う1.5 kHzから10 kHzに至るエネルギーをとらえ、上位5%のエネルギーを有するピークを選択してイベント候補を抽出し可視化する。利用者は、GUIを用いて対話的に候補を確認し、判定結果を記録する。
- ◆Praatによる原波形・スペクトグラム表示と聴取による方法では嚥下音の同定が困難だったケースも、swallowingChecker ではイベント近傍の波形を拡大した表示など、複数の情報提示と聴きやすく処理されたイベント音を聴取することにより、判断可能なケースがあった。

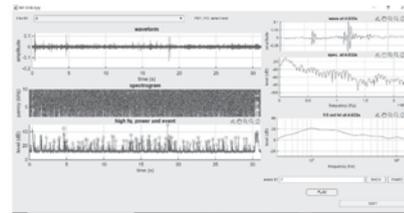


Fig.1: swallowingChecker