

研究概要報告書【サウンド技術振興部門】

(/)

研究テーマ	発話者の骨導音知覚特性を考慮した気導音－自己聴取音の声質変換モデル構築	報告書作成者	鳥谷 輝樹
研究従事者	鳥谷 輝樹		
研究目的	<p>発話中、ヒトは自身で発した音声(自己聴取音)を知覚することで円滑な発話運動制御を実現する。そうした音声生成・知覚の相互作用メカニズムを解明するには、発話時の音声伝達経路の様相について詳細な理解が必要である。自己聴取音の伝達には、気導・骨導という二種類の経路が存在する。これまで、音声生成・知覚メカニズムについては気導音の伝達系を中心に議論が進められてきた一方、骨導音の詳細な伝達・知覚特性は不明である。発話者－聴取者間のコミュニケーションにおいて、骨導音の伝達・知覚特性は発話者のみに作用する特性であるため、発話者－聴取者間の声質(音色)の知覚には相違が生じる。この相違は、自己主体感の形成や非言語情報伝達に大きな影響を与える可能性がある。たとえば、日常生活において録音した自身の音声を聞くと普段の自身の声とは声質が異なり、違和感を抱くことがある。こうした現象はライフサイエンスの分野において注目されており、その要因の究明が求められている。しかし、骨導音伝達・知覚特性が、発話者にとっての自己聴取音の声質にどのような影響を与えるのかについて、心理音響学的観点からは十分に議論されていない。</p> <p>骨導音知覚に関する研究は、長らく音声コミュニケーションの場面とは独立して行われてきた。骨導音の伝達経路として、発話器官によって頭蓋骨や軟組織の振動が励振され、聴覚末梢系の中耳や内耳が直接振動振動する伝達経路(以後、経路 B と呼称する)と、発話器官の振動励振が外耳道壁から音となって漏れ出すことにより、気導音として伝達される経路(以後、経路 C と呼称する)の存在が知られている(別紙図 1)。発話者の骨導音知覚特性が、これらの伝達経路によってどのように説明できるかについては依然、明らかでない。</p> <p>そこで本研究では、発話時の骨導音知覚特性を心理物理的手法によりトップダウン・ボトムアップの両面から解析することで、自己聴取音の声質への寄与を議論することを試みる(参考:別紙図 2)。本研究の目的は、発話者の骨導音知覚特性が自己聴取音の声質に与える影響を解明することである。当該年度では、<u>Analysis-by-Synthesis</u> の考え方にに基づき、骨導音の知覚特性を考慮した気導音－自己聴取音の声質変換モデルを構築することを目的とする。本研究で構築したモデルあるいは得られた知見は、動画配信メディアにおいて配信者が収録コンテンツの音声聴取時に生じる違和感を段階的に低減させるシステム等に応用できる。また、あらゆる個人に対する自己聴取音を合成できれば、自身の自己聴取音を含む聴覚情景を他者に遠隔提示する VR システムに応用できる。</p>		

研究内容

気導音－自己聴取音の声質変換規則を議論するために、下記の項目を実施した。

① 骨導音知覚特性の心理物理的推定:

(1-1) 骨導経路の伝達特性について、音響インパルス応答計測法を基にした物理的手法により、主要な2つの骨導経路(頭部振動:以後、経路Bと呼称、および外耳道内放射:以後、経路Cと呼称)における周波数－振幅特性を実測した(別紙図1)。大学生10名が実験に参加した。口唇正面でマイクロホンにより気導音を収録すると同時に、加速度センサで側頭部の振動を収録し、プローブマイクロホンで外耳道内観測音を収録した。参加者には30単語からなる英文を音読させ、上記3観測点から発話中の音声信号をサンプリングクロック同期収録した。音声の有声区間において、約10msの短時間窓ごとに周波数解析し、各観測点での信号の長時間平均スペクトルを算出することにより、気導音を基準とした経路Bの伝達関数および、気導音を基準とした経路Cの伝達関数を観測した。

(2-2) 気導刺激と喉頭付近での骨導刺激を同時提示し(別紙図2)、両者が知覚的に抑圧される(音の大きさが最小となる)際の刺激間の振幅・位相条件を周波数の関数として求める測定実験を実施した。実験には、大学生6名が参加した。周波数250, 500, 1000, 2000, 4000, 8000 Hzの純音について、気導刺激と骨導刺激のラウドネスが互いに等しくなる振幅条件、および両刺激の同時提示時のラウドネスが最小となるような位相条件を求め、周波数の関数として参加者ごとに解析した。

② 自己聴取音の声質における骨導音知覚の寄与の検討:

①で推定された骨導音知覚特性をフィルタとみなして「疑似骨導音」を合成し、気導音と混合させる。この混合信号の主観的な音色について、参加者が発声した音声を参加者自身が聴取した際に感じられる音色との間の相違が最小となるような気導音－疑似骨導音の混合比を決定することで、自己聴取音と骨導音の間の関連性を推察した。

実験には、大学生10名が参加した。事前に、参加者の気導音/arigatou/を収録した。参加者には、収録した気導音/arigatou/と、(1-1)で測定した伝達関数をもとにフィルタリングした2種類の骨導模擬音声の混合音を提示し、図3に示すGUIに基づき、提示音が自身の自己聴取音声の音色に最も類似するように「操作1」と「操作2」のスライダーを調整することを求めた。「操作1」では2種類の骨導模擬音声の間の混合比を調整し、「操作2」では模擬骨導音声全体と気導音との間の混合比を調整した。

研究概要報告書【サウンド技術振興部門】

(/)

<p>研究のポイント</p>	<p>骨導音知覚に関する研究は、長らく音声コミュニケーションの場面とは独立して行われてきた。近年、自己聴取音の合成については国内を中心に関心が高まっている。しかし、これらは詳細な骨導音知覚特性の解明を意図したのではなく、自己聴取音の精緻な模擬という点では課題が残る。申請者らは、観測可能な骨導伝達特性の物理測定を通して、発話時の骨導音知覚特性のボトムアップ的な解明から試みてきたが、得られた特性と自己聴取音の声質との関連性を深く議論するには、振幅・位相特性の両面の貢献について検討を深めることと、物理計測(ボトムアップ)と心理物理的手法(トップダウン)を組み合わせたアプローチが有効であると考えた。</p> <p>そこで本研究では、①物理計測によるボトムアップ的知見(1-1)と心理物理的計測によるトップダウン的知見(2-2)の両面から経路 B・C の骨導伝達特性の振幅・位相特性の実測・推定を試みた。また、②骨導伝達特性の参加者間共通性(概形)および参加者間個人性(局所的形状)の両方を考慮したフィルタ設計と心理物理的評価実験により、両経路の知覚的寄与を議論した。</p>
<p>研究結果</p>	<p>【検討(1-1)】 物理的に実測された振幅特性における参加者共通の特性として、経路 B では概ね 2000 Hz 未満の成分を通過させる特性を有しており、経路 C では 2~3 kHz 付近の成分を帯域通過させる特性を有していた(別紙図 3)。経路 C の振幅特性については、約 1~4 kHz の帯域における局所的形状に最大約 20 dB の個人差があったことから、自己聴取音の声質の個人性に寄与しているものと推察される。【検討(2-2)】 心理物理的に推定された位相特性における参加者共通の特性として、250 Hz 付近の低域では約 90° の位相遅れが、1000 Hz 付近では約 90° の位相進みがあることが分かったが、その他の帯域では参加者間の一貫した傾向は認められなかった。本実験の段階においては、自己聴取音の声質における位相特性の影響は限定的であると結論付けた。【検討②】 経路 B および C の伝達特性の影響を考慮した模擬骨導音声の混合比について主観評価を行なった結果、参加者により評価結果に若干の個人差があったものの、評価の分布傾向としては経路 C の寄与が経路 B に比べ約 4 dB 大きい可能性があることが分かった。</p>
<p>今後の課題</p>	<p>研究全体を通して、モデル化する骨導伝達特性を経路 B・経路 C という 2 つの伝達過程に限定した点は本研究の考慮すべき点である。これら 2 つの伝達過程によるモデル化の妥当性について、骨導音知覚における大局的な(合成伝達系としての)特性に関するさらなるトップダウン的アプローチにより追検討を実施予定である。</p> <p>個々の検討に関する今後の課題として、【検討(1-1)】 骨導伝達過程のモデル化の際、発話器官での入力に主に喉頭音源であることを仮定すれば、室内音響計測等で多用される掃引法などのインパルス応答計測法を応用すれば骨導伝達過程の振幅・位相特性を同時に推定できるかもしれない。今後は、経路 B・C の両者に関して、振幅・位相特性の同時推定を実施していく。【検討②】 先の検討①において、自己聴取音の声質に対する位相特性の影響は限定的であると仮定したが、模擬骨導音声の混合比の主観評価の際には位相に関しても振幅とは独立に操作した上で、自己聴取音の声質に対する位相特性の貢献を検討できるかもしれない。</p>

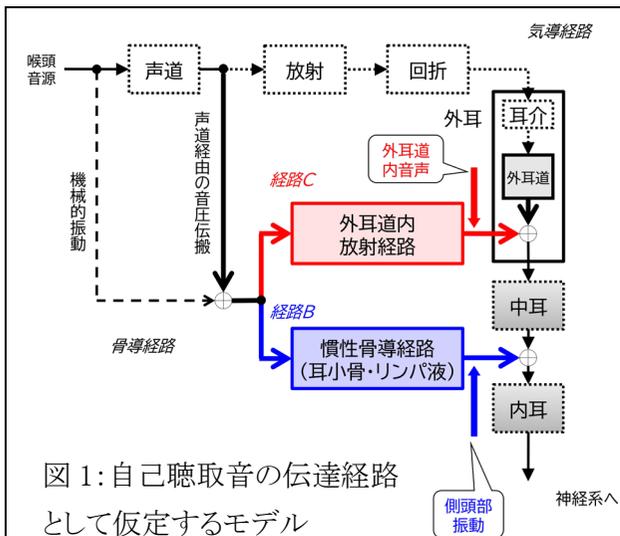


図 1: 自己聴取音の伝達経路として仮定するモデル

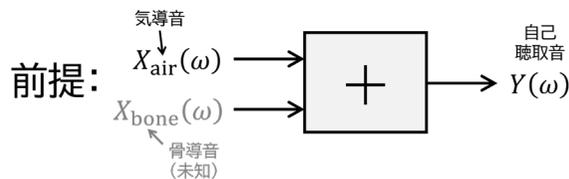


図 2: 本研究における検討項目①・②で明らかにする変量. 各変量はスペクトルであり, 角周波数 ω の関数.

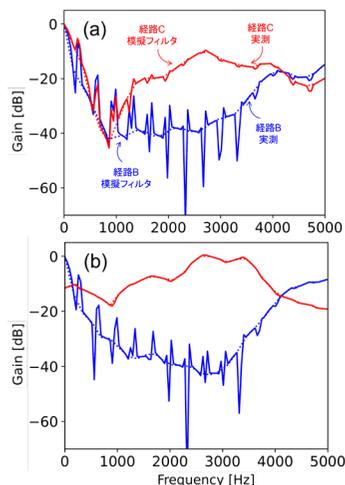
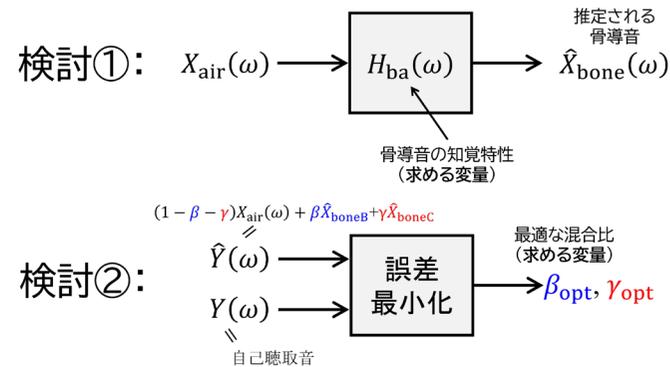


図 3: 実測された骨導伝達特性(経路 B・C)の振幅特性(異なる参加者(a)と(b))

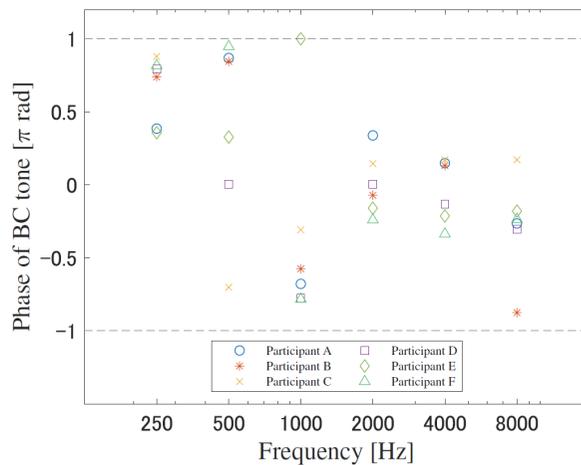


図 4: 心理物理的に計測された骨導伝達特性の位相特性

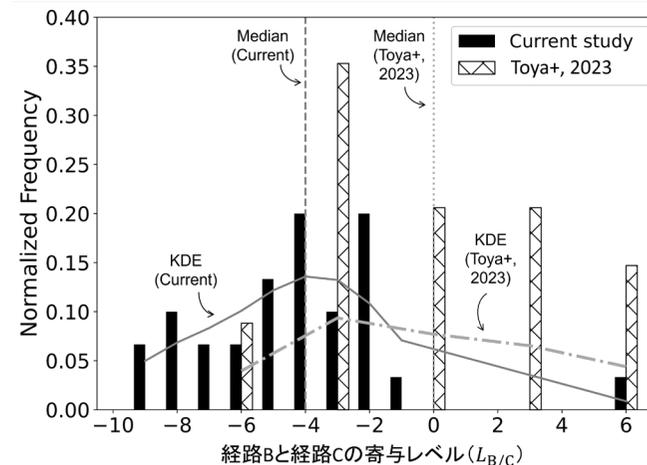


図 5: 主観評価実験で決定された経路 B と経路 C の振幅特性の寄与 (β と γ) のレベル ($L_{B/C}$) の度数分布