MatchEQ 機能を用いた骨導音の特徴の調査と自己聴取音の再現

著者名 宮坂優吾　武藤凛太

要旨

自分の声を自分で感じる自己聴取音と録音した音声を聞くのでは声が変わってしまうことに興味を持って、録音した声から自分の声を再現する方法を模索した。自分の声が録音した声と変わる原因は、自分の声が気導音と骨導音で構成されるのに対して、録音した声は気導音だけで構成されることだ。そのため、録音した声から自分の声を再現するために気導音と骨導音との関係の特徴を知る実験をおこなった。MatchEQ 機能を使って骨導音と気導音の関係を比べたところ、骨導音の特徴は個人差があり、発する言葉によってではなく誰が発声するかに元ずくことがわかった。そのため、一回骨導音から気導音への変換の仕方をこの MatchEQ 機能で学習させれば、その人の新しく録音した音声も骨導音へと変換することができ、これを用いて、自分が感じる自分の声を再現できることがわかった。

キーワード：骨導音

Ⅰ．はじめに

発声しているときに自分のなかで聞こえている声（以降「自分の声」）と、それを録音した声（以降「録音の声」）にギャップがあるように感じ、その原因を調べ、その違いについて研究することにした。また、それを元に録音した声から自分の声を簡単に疑似的に再現する方法を模索した。

自分の声が録音で変に聞こえるのは、普段は気導音（空気を伝わって届く音）と骨導音（骨から伝わる音）が混ざっているのに対し、録音は気導音のみだからということとを知った。

伏見ほか(2021)によると、マイクで録音した音声に音の高さや再生速度などの要素を編集して被験者に聞かせ、どれが自分の声らしく聞こえるか実験したところ音を低くしたものが自分の声らしいと判断する人が多かった。しかし、この研究では音の高さを低、中、高の 3 つでしか分けておらず、どの周波数をどのように強めれば録音の声を自分の声に再現できるかわかるためには不十分であった。

また落合ほか(2019)より骨導音と気導音のパワースペクトルという周波数ごとの音の強さは図 1 のようになっていた。

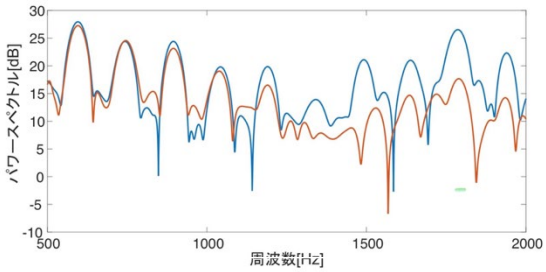


図 1　気導音と骨導音のパワースペクトル　青：気導音

赤：骨導音

図 1 からも、骨導音は低温が強く、骨導音と気導音からなる自分の声を再現するためには録音の声の低温を強めればいいとわかった。落合ほか（2019）はローパスフィルタという音の高音域をカットするフィルタに録音の音を通し、どの音域からフィルタをかけると自分の声に近づくかを調べる研究を行なった。その結果、800～1000HZ でローパスフィルタをかけることによって自分の声に近づくという結果が得られた。しかし、この研究では母音でしか実験が行なわれておらず、子音でも実験を行なう必要があると感じた。また、手軽に自分の声を再現するための手段として身近なイコライザーを用いたほうが良いと判断し、イコライザーを使って録音の声のどの周波数をどう

いじれば自分の声に近づくのか調べることにした。

しかしイコライザーで先行研究のように低音を強めたとしても自分の声に近づいているという感触を得られなかった。また、中間発表時の先生からの指摘により、そもそも骨導音の特徴は個人差があり、誰もがその通りにイコライザーをかけると自分の声を再現できるとは限らないとわかった。そのため、まずは骨導音の波形をとり、録音の声と比べることにより特徴をしらべようとした。

まず試したのは、録音した声を骨伝導イヤフォンで牛の骨に押し当てて再生し、跳ね返ってきた音を収録してみる手法だった。しかし、牛の骨を通して音声を収録することが難しく、断念した。

骨導音を収録できるマイクを探したところ、エレクトレットコンデンサーマイクというものが存在していることを知り、そのマイクを使用して自分の声と録音の声の違いを分析してみることにした。

II. 手法

自分の声を収録するための「収録環境」と、収録した音声を分析するための「ソフト環境」を用意した。それらの環境を使い、複数人でいくつかのパターンを発声し、収録し、比較を行なった。

収録環境

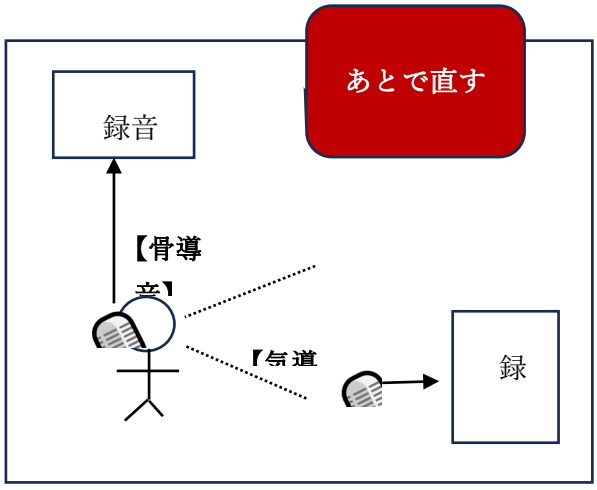


図２：収録環境（一回の発声を、２つの系統で録音できるようにする）

図２のような環境を用意し、一連の一回の発声に対して、２つの系統で録音する。

- 声を発する人：複数人を想定
- 骨伝導マイクで収録する系統（骨伝導マイク＋録音用 iPad）
  ※骨伝導マイク：SONY コンデンサーマイク ECM-TL3
- 普通のマイクで収録する系統（スマホのマイク）

収録パターン

- 声を発する人：A さん（男性）、B さん（女性）の２名
- 発声する内容：「あ、か、さ、た、な」、ヒューマンビートビートボックス（口を用いて破裂音を発しドラムセットの音などを模倣したもの。）

ソフト環境（WindowsPC 上）

収録された音声データを下記のソフト環境で分析する。

- 音楽制作ソフト：Cakewalk Sonar (Version 2015.12)
- MatchEQ 機能：iZotope Ozone11 Advanced
- 音量を整える機能：NoiseWorks DynAssist

実験手順

収録環境において、発話者が、各収録パターンを発話し、図１の２系統のマイクで、それぞれ気導音と骨導音を収録する。

ソフト環境において、収録された骨導音と気導音に対して以下の処理を行なう。

1. 骨導音と気導音の開始タイミングを合わせる【トリミング】
2. 骨導音と気導音の、各音の音量を一定の音量に合わせる【音量調整】
※後続の処理に悪影響を与えなくするため
※合わせる音量は「-12LUFS」(ラウドネスフルスケール)
3. MatchEQ 機能で、周波数カーブを計算する【周波数分析】
※Match EQ とは、主に音楽制作ソフト（DAW）のプラグイン機能で、参考音源（リファレンス）の周波数特性（音のバランス）を分析し、自分の音源に自動的に適用して似たトーンにする EQ（イコライザー）機能。気導音の「聞こえ」を骨導音の「聞こえ」に寄せるための周波数特性を計算し、別の音源に適用することもできる。
4. 気導音を、3.で計算した周波数カーブを適用して再生する

その音が、骨導音と似ているかどうかを、目（波形）と耳（聞こえ）で確認する。

III. 結果

- 発声者：男性
- 発話内容：「あ、か、さ、た、な」

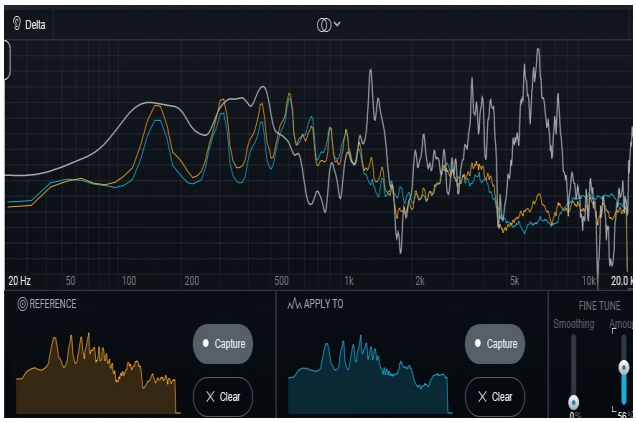


図３：男性による「あ、か、さ、た、な」と発した時の周波数特性カーブ

図３の周波数特性カーブの色分けの説明

- 白色：MatchEQ 機能で計算された、気導音を骨導音に寄せるための増減値
- オレンジ色：骨導音を MatchEQ 機能に読み込ませた分析結果
- 水色：気導音を MatchEQ 機能で読み込ませた分析結果

聞こえの違いを説明した動画：
https://youtu.be/P0rhNWbMpMs

結果２

- 発声者：男性
- 発話内容：ヒューマンビートボックス（人間が、パーカッションのような音を出した音）

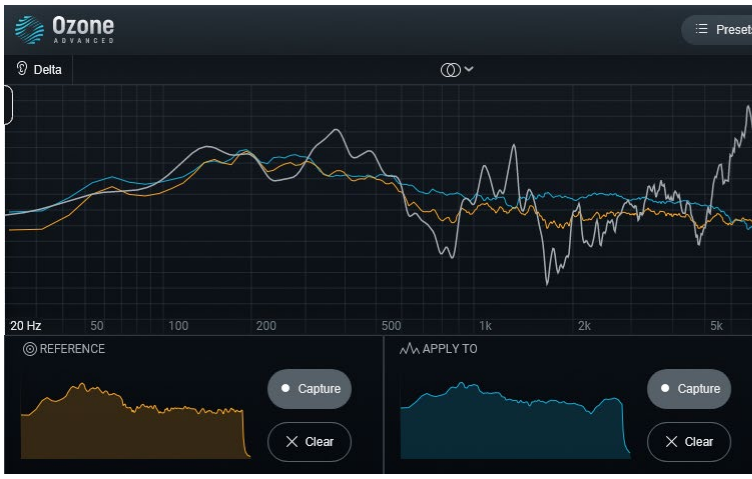


図４：男性による、ヒューマンビートボックスを発したときの周波数特性カーブ

図４の聞こえの違いを説明した動画：
https://youtu.be/ugFdJDnKD74

結果３

- 発声者：女性
- 発話内容：「あ、か、さ、た、な」



図５：女性による、「あ、か、さ、た、な」を発したときの周波数特性カーブ

図５の聞こえの違いを説明した動画：
https://youtu.be/qE-JQbBoJzg

MatchEQ 機能の計算のさせ方を含む説明動画の再生リスト：
https://www.youtube.com/playlist?list=PLuVM8CEl1hYqVDVsrkIPYwAhjQCIVPFCh

Ⅳ．考察

図３、図４、図５を見ると、骨導音は、気導音に比べ、200Hz 付近の低音がブーストされていることが分かった。中音域や高音域は、ブーストされる周波数帯もあれば、カットされる周波数帯もあることが分かる。

図３と図４の、MatchEQ 機能で分析された白い波形を見比べると、同じ人が「あ、か、さ、た、な」と話すときと、ヒューマンビートで（声ではないが音）を発した時の、骨導音と気導音の周波数の違いの曲線は、ある程度、似ている。図５は別の人の結果だが、曲線に有意な違いが見られた。

骨導音と気導音の違いは、話す言葉ではなく、話す人に依存して決まってくることが読み取れた。

話す人によって音が変わってくる要素として考えられるのは、その人の頭の中の骨格構造や、骨の組成成分。文献
https://www.noe.co.jp/technology/29/29news5.html
によると、音響シミュレーションという分野があり、音が伝搬する空間構造や囲っている材質によって、聞こえの伝搬が変わってくる世界だということが分かった。

そのため、骨導音と気導音の違いは、人の頭の中の空間構造や骨の材質によって変わってくるものかと予想するが、そこまで踏み込むには、時間とコストがかかるため、今後の課題となった。

貝沼（2022）によると、頭蓋には個人特有の音響伝搬特性があることを立証している。本研究で複数の話者の周波数カーブが、同一の人の場合は、似たカーブになっていて、別の話者のカーブは有意な違いがあることが分かったが、その結論と通じている。

補足：骨伝導マイクで収録した音声には、骨導音だけでなく、気導音も若干含まれている可能性があるが、試行錯誤した結果、録音した音声から気導音を「引く」ことが困難だった（位相を逆転して引き算しようにも、何割減して引き算すればいいかがわからなかった）ため、骨伝導マイクで収録した音声は、「ほぼ」骨伝導音だけ、だと割り切って進めた。

Ⅴ．結論

自分の声と録音の声が違って聞こえ、その原因を解明したく、調査を始めた。

調査の結果、気導音と骨導音という音の伝わり方があり、それぞれ収録するマイクが存在して

いることが分かった。

それらのマイクと分析環境を準備して実験をした結果、自分の声と、録音の声の違いを、MatchEQ 機能で、波形として目で見ることに成功した。

発話者と話す内容のバリエーションを振った結果、聞こえの違いは、話す内容ではなく、話す人によって変わることが読み取れた。そのため、一回個人ごとに MatchEQ を用いて気導音から骨導音の再現を学習させれば、気導音から骨導音を再現することができ、それの含有率をへんかさせることで自分の声を再現できる。その含有率を調べることは今後の課題である。

ただ、日をまたいで実験しておらず、日によっても骨導音と気導音の関係に変化が生じてしまうのかということまでは調べられていない。実験を行なった人数が二人と少なくもう少し人数を増やして結果を比較することも今後の課題である。また、音響シミュレーションの技術を学び、頭の骨格構造や音の吸収・反射する周波数の分析をしていくことで、聞こえの違いの根本原因を掘り下げていきたい。

謝辞

助言をくださった堀先生など先生方,ご協力ありがとうございました。

文　献

廣瀬通孝,伏見遼平,鳴海拓志,谷川智洋,吉田成朗(2016):自分の声を知り,コントロールするための「自分声フィルタ」の提案

松本直樹,村上隆啓,落合祐貴,丹治寛樹(2019):低域通過フィルタを用いた気導音からの自己聴取音の再現.

貝沼遼太郎:骨伝導型個人認証デバイスによる個人毎の差異に関する検討