

錯視・錯聴コンテスト 2021

## 両耳でしか聞こえない声

新美亮輔（新潟大学人文学部） niimi at) human.niigata-u.ac.jp

作品の紹介と一般向けの説明の動画：

<https://youtu.be/UybdhYlrG40>

### 1. どんな錯聴か

- ・ステレオヘッドホン（イヤホンも可）で聴いてください。スピーカーではうまくいきません。
- ・ノイズの中に人の声が聞こえます。しかし、ヘッドホンの片耳を外して、もう一方の耳だけで聞くと、声は消えてしまいます。右耳・左耳どちらでも同じです。
- ・すなわち、左の音も右の音も、それぞれ単独では声を含みません。ただのノイズです。しかし両耳聴取すると、存在しないはずの声が知覚されます。

### 2. 説明

両耳性ピッチ（binaural pitch, dichotic pitch）を声に拡張したものです。両耳性ピッチは、右耳と左耳のノイズの間の関係を操作することで、存在しないピッチが両耳聴取される現象です。これまでの手法で作られる両耳性ピッチは純音様の単音だけでしたが、任意の音声を両耳性ピッチにできないかと考えて作ったのが今回の作品です。

左右両耳に同一の広帯域ノイズを同時提示するとき、一部の狭い周波数帯域だけ左右で位相差をつけると、その帯域に該当するピッチが両耳性ピッチとして知覚されます（Huggins Pitch）。[1,2] 両耳間位相差をつけることは、両耳間相関（interaural correlation）を下げることであります。そこで、やはり左右両耳に同一のノイズを提示するとき（つまり、両耳間相関が1）、たとえば右耳の一部の狭い帯域だけ別のノイズの同じ帯域の成分に置き換えると、この帯域だけ両耳間相関0となり、やはりこの帯域に該当するピッチが両耳性ピッチとして聞こえます。[3,4] 両耳間で位相を $\pi$ ずらす（位相反転）場合には、両耳間相関を-1に下げていることとなります。いずれの手法でも、片耳の音単独では単なるノイズであり、ピッチに関する情報が全く含まれていないことが特徴です。両耳マスキングレベル差とはこの点が異なります。

今回の作品は、この両耳性ピッチを聞こえやすくする試みから生まれました。まず従来手法で両耳性ピッチを作りました。白色ノイズAを用意し、中心周波数440 Hzの狭帯域（たとえば420–460 Hz）を、別の白色ノイズBの同じ帯域の成分に置き換えます。これをAbとします。Abを左耳に、元のノイズAを右耳に同時提示すると、440 Hz相当のピッチ（「ラ」）が両耳性ピッチとして聞こえます。[5] ただ、ノイズに比べてあまり強く聞こえません。

ここで、置き換える帯域の幅を広げれば両耳性ピッチ知覚の強度は増しますが、広げすぎるとピッチが損なわれます。そこで、帯域を広げるかわりに、倍音成分を加えました。つまり、中心周波数880 Hz、1320 Hzの帯域もノイズBに置き換え、相関を0にしました（動画中で紹介しているメロディの例 [6]）。従来手法よりも、かなり聞こえやすくなりました。

このように倍音構造が両耳性ピッチ知覚にも反映されるのなら、置き換える帯域（両耳間相関を下げる帯域）を自由に操作できれば、任意の音声を両耳性ピッチで再現できそうです。それを実現したのが、今回の作品です。

作成方法の概要は Figure 1 のとおり。まず声（モノラル）を用意します。今回はアナウンス

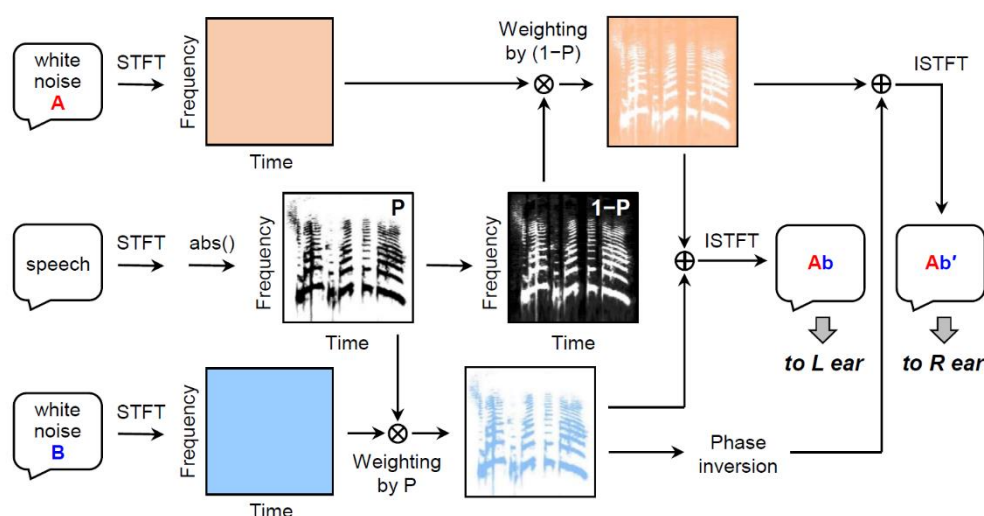


Figure 1. Schematic illustration of how to make the binaural speech stimulus.

音声素材 [7] を使いました。この声を STFT（短時間フーリエ変換）して絶対値をとり、時間つきパワースペクトル  $P$  を得ます（要はスペクトログラム）。 $P$  は最大値が 1 となるように正規化します。次に、声と同じ長さの白色ノイズ A と B を用意します。B を STFT し、これに  $P$  をかけて重みづけをし、白色ノイズ B のうち声と同じ帯域・時間だけの成分  $b$  を取り出します。 $b$  に  $-1$  をかけて位相反転した成分  $b'$  も作ります。つまり、 $b$  と  $b'$  は音波にすると相関  $-1$  です。

白色ノイズ A も STFT します。そして  $(1-P)$  をかけ、声と同じ帯域・時間の成分を取り除きます。これに  $b$  を加えて逆 STFT し、音波  $Ab$  を作ります。 $b$  ではなく  $b'$  を加えてから逆 STFT した音波  $Ab'$  も作ります。 $Ab$  を左耳、 $Ab'$  を右耳に同時提示します。元の声の帯域・時間に該当する成分は両耳間相関が  $-1$  ( $b$  と  $b'$ ) ですが、他の成分は両耳間相関が  $+1$  (左右とも A) です。そのため、両耳性ピッチによって元の声に似た音が知覚されます。しかし、 $Ab$ 、 $Ab'$  とともに、それ単独ではただのノイズにすぎません。なお、元の声の帯域・時間に該当する成分の両耳間相関を 0 にしても同様のデモを作成できますが、効果は若干弱いようです。

実際には、より効果が大きくなるよう、種々の調整をしています ( $P$  の形状など)。また、聴き取りやすくするため、声はゆっくりのスピード (75%) に変換してから用いています。

### 3. どのような意義があるのか

両耳性ピッチ自体は既知の現象ですが、任意の音声にまで拡張できたことは新しい点です。人の声のような複雑な音も両耳性ピッチによってある程度知覚できることは、両耳間統合についての新しい知見かもしれません。一方で、両耳性ピッチによる声は、やや聞き取りにくいようです。両耳性ピッチ知覚の時間解像度の低さが原因ではないかと推測しています。

また、倍音成分を加えることで、より聞き取りやすいデモを作れることもわかりました。両耳性ピッチは、人間の音の知覚が耳だけによっているのではなく脳の中で起こっていることを端的に体験できる現象です。体験しやすく効果量の大きいデモは、より多くの人に聴覚心理学への興味を持ってもらうことに役立つかもしれません。

[1] Cramer, E. M., & Huggins, W. H. (1958). Creation of pitch through binaural interaction. *Journal of the Acoustical Society of America*, 30, 413–417. <https://doi.org/10.1121/1.1909628>

- [2] 他にも、両耳間位相差を操作することにより両耳性ピッチを作る手法がいくつか知られています (Klein & Hartmann, 1981, *Journal of the Acoustical Society of America*, 70, 51–61; Hartmann & McMillon, 2001, *Journal of the Acoustical Society of America*, 109, 294–305)。
- [3] Akeroyd, M. A., & Summerfield, A. Q. (1999). A binaural analog of gap detection. *Journal of the Acoustical Society of America*, 105, 2807–2820. <https://doi.org/10.1121/1.426897>
- [4] Fan, L., Kong, L., Li, L., & Qu, T. (2021). Sensitivity to a break in interaural correlation in frequency-gliding noises. *Frontiers in Psychology*, 12, 692785. <https://doi.org/10.3389/fpsyg.2021.692785>
- [5] 動画中のデモでは、A も Ab も、最後にフィルタにより低域 (200 Hz 以下) と高域 (2000 Hz 以上) をカットしています。単に、ノイズのうるさを抑えて聞きやすくするためです。
- [6] 実際には、帯域置き換え加工による意図せぬノイズ等が片耳だけに出てしまう可能性を除くため、まず A, B, C の 3 つの異なる白色ノイズを用意し、A の一部帯域を B で置き換えたもの Ab と、同様に C で置き換えたもの Ac を作り、Ab と Ac をそれぞれ左耳と右耳に流しています。
- [7] こえやさん <https://www.koeyasan.com/> より 2021 年 8 月 31 日ダウンロード (CV: 日下真尋)。  
なお、スピードとピッチを変化させて使用しています。