



静かな環境でPC出力音声を文字起こしする方法（イヤホン使用時）

課題の概要

静かな場所ではPCのスピーカーを使えないためイヤホンで音声を聞きたい一方で、PCが再生する音声をアプリで取得してエスペラントに文字起こししたいという課題があります。通常、PCの再生音声はステレオミックス（Stereo Mix）という仮想的な録音デバイスを使うことで取得できます。しかし一部のPCでは、イヤホン出力に切り替えるとステレオミックスが音声を拾えなくなる問題があります^①。つまり、「イヤホンに音が出ている状態でPCの再生音声をどうキャプチャするか」が大きなポイントです。また可能であれば自分のマイク音声も同時に取り込んで、相手の音声と合わせて文字起こしできると理想的ですが、これも簡単ではありません。以下、WindowsおよびUbuntu(Linux)それぞれでの解決策を、パソコンに詳しくない方にも分かるように説明します。

WindowsでPCの出力音声を取得する方法

1. **ステレオミックスを有効にする:** まずお使いのWindows PCにステレオミックス機能がある場合は、それを有効にして入力ソースに指定する方法があります^②。ステレオミックスとは、PC上で再生される全ての音を仮想的なマイク入力として扱う機能です。具体的には、「サウンド設定」→「録音」タブで空白部分を右クリックし「無効なデバイスを表示」を選択すると、Stereo Mix（ステレオ ミックス）というデバイスが現れることがあります。これを右クリックして「有効化」し、必要に応じて「既定の録音デバイス」に設定します^③。アプリ側でこのステレオミックスを入力ソースとして選択できれば、イヤホンから音を聞いていてもその音声をPC内部でキャプチャできます。

しかし、ご指摘通り一部のPCではイヤホンを挿すとステレオミックスで音を拾えない場合があります^①。これはPCのオーディオドライバの仕様によるもので、イヤホン接続時にスピーカー出力が物理的にオフになるとステレオミックスも無音になることがあるためです。その場合は次の方法を検討します。

2. **仮想オーディオデバイスを利用する:** ステレオミックスが使えない、またはうまく動作しない場合、VB-Audio Virtual CableやVoiceMeeterといった仮想デバイスを利用する方法があります^④。これはソフトウェア上で仮想的な音の出入口（ケーブル）を作り、PCの再生音声を一度その仮想ケーブルに出力してから入力として取り込む仕組みです。具体的な手順の一例を示します：

- **VB-Cableのインストール:** VB-Audio社が提供する無料のVirtual Cableをダウンロード・インストールします。インストール後、Windowsの音声デバイスに「CABLE Input」「CABLE Output」という仮想デバイスが追加されます^④。
- **出力デバイスの切り替え:** Windowsのサウンド設定で、普段使っているスピーカー/イヤホンの代わりに既定の出力を「CABLE Input」に変更します。これでPCの音声は一旦仮想ケーブルに流れます。
- **仮想ケーブルからイヤホンへ音をモニタ出力:** 音が聞こえなくなるので、録音デバイス一覧から「CABLE Output」を探し、右クリックして「このデバイスを聴く」を有効にし、再生先にお使いのイヤホン（スピーカー）を指定します。こうすることで、仮想ケーブルに流れた音声がリアルタイムでイヤホンにも出力され、引き続き音が聞こえる状態を保てます。
- **アプリで仮想ケーブルの音声を取得:** アプリ側の録音ソースに「CABLE Output」（仮想ケーブルの出力側）を指定します。これにより、PC上の再生音がそのままアプリに入力されます^③。先ほど設定した

モニタリングによって自分にはイヤホンで音が聞こえつつ、同じ音声データをアプリが取得できるわけです。

仮想オーディオデバイスを使えば、イヤホン使用中でも確実にPCの再生音を捕捉できます。特にVB-Cableはシンプルで遅延も極めて小さいため、文字起こし用途に適しています。VoiceMeeterというソフトもありますが、こちらは高度なミキサー機能を備えています。その分設定が複雑になるため、まずはステレオミックスかVB-Cableでの解決を試すことをお勧めします。

Ubuntu/LinuxでPCの出力音声を取得する方法

Linux (Ubuntu) には、Windowsのステレオミックスに相当する機能が標準で組み込まれています。多くのLinuxディストリビューションで使われているPulseAudioやPipeWireでは、各出力デバイスに対して「モニター」(monitor)ソースが自動生成されます^⑤。モニターソースとは、その出力デバイス経由で現在再生されている音声を拾う仮想的な入力デバイスのことです^⑥。つまり追加のソフトを入れなくても、PCの再生音声を録音するための「見えないマイク」が用意されているイメージです。

UbuntuでPC音声を文字起こしアプリに渡すには、録音ソースをモニターに切り替える必要があります。手軽な方法として、PulseAudio Volume Control (pavucontrol) というツールを使った設定が便利です：

- UbuntuソフトウェアセンターやAPTで pavucontrol をインストールし、起動します。
- アプリを文字起こしモードで起動した状態で、pavucontrolの「録音」タブを開きます。すると現在音声を入力しようとしているアプリ（プロセス）が一覧に表示されるはずです。
- そのアプリの入力ソース項目をクリックし、プルダウンメニューから「<出力デバイス名>のモニター」と書かれた項目を選択します。例えば内蔵オーディオを使っているなら「Built-in Audio Analog Stereoのモニター」といった名前です。
- これで、そのアプリはマイクの代わりにPC出力のモニターから音声を取得するようになります。あとは普段どおりイヤホンで音を聞いていても、再生中の音声がアプリに取り込まれエスペラント文字起こしが行える状態になります。

PulseAudioを利用している限り、このモニター機能で遅延なく出力音声を録音できます。またPipeWire環境でも概念は同じで、互換のため pavucontrol で同様の操作が可能です。Linuxの場合追加の仮想ケーブルを入れる必要は基本的になく、標準機能だけで実現可能のがメリットです。

自分のマイク音声も同時に取り込むには（オプション）

PCの出力音だけでなく自分の声（マイク入力）も合わせて文字起こししたい場合、音声ミキシングの工夫が必要です。通常、ビデオ会議アプリ等では自分の声はスピーカー/イヤホンから出ず直接相手に送られるため、前述の方法では自分の声が含まれません。これを解決するには以下のアプローチがあります。

1. Windowsの場合 – VoiceMeeterでミックス: VoiceMeeterは仮想オーディオミキサーで、複数の音源（デバイスやアプリの音声）をまとめて別の出力に混合することができます^⑦。具体的には、VoiceMeeter上でハードウェアインプットにマイクを、VAIO(仮想)入力にシステム音声を割り当て、両者をミックスした上で仮想の出力(B1など)として提供します。そしてアプリ側ではその仮想出力を録音デバイスとして指定すれば、自分の声+PC音の両方が含まれた音声ストリームを取得できます。加えてVoiceMeeterではミックスした音声を自分のイヤホン（物理出力A1等）にもモニタリングできるため、相手の声も引き続き聞こえます。設定はやや複雑ですが、公式ガイドやチュートリアル動画も多く公開されています。

2. Windowsの簡易策 – マイクの再生を有効化: 簡単な方法として、Windowsサウンド設定でマイクをスピーカーに聞かせる（マイクモニター）設定を使う手もあります。録音デバイスのマイクプロパティで「このデバイスを聴く」をオンにすると、自分の声がリアルタイムにイヤホンから聞こえるようになります。その状態で

前述のステレオミックスやVB-Cableを使えば、**自分の声もPC出力に混ざった状態でキャプチャできる可能性**があります。ただしマイク音をスピーカー出力に乗せるとハウリング（音のループによるノイズ）のリスクがある点や、自分の声が遅れて聞こえてくる違和感がある点には注意が必要です。

3. Ubuntu/Linuxの場合 - モジュールの活用: PulseAudioでは `module-loopback` という仕組みを使い、マイク入力を出力デバイス（スピーカー/イヤホン）へリアルタイムにループバックできます。例えばターミナルで `pactl load-module module-loopback source=<マイク> sink=<イヤホン出力>` とロードすれば、以降マイクの音がスピーカー（イヤホン）にそのまま流れるようになります。これで先ほど設定したモニターソースには**相手の音声+自分の声**の両方が含まれることになり、アプリでまとめて文字起こしが可能になります。PipeWireの場合も類似の方法が可能ですが、コマンドや設定は若干異なります。

以上のように、自分の声も含めた同時録音は技術的に可能ですが、**音質や遅延への影響**、設定の難易度などの問題があります。会議の記録を手元で残す程度であれば、まずは相手の声（PC出力音声）のみを確実に文字起こしできるようにして、必要に応じて自分の発言は別途メモするなどの運用も検討してください。

まとめ

静かな環境でイヤホンを使いながらでも、工夫すればPCが再生する音声をアプリに取り込んでエスペラント文字起こしすることは**十分可能**です。Windowsではステレオミックスの有効化がシンプルですが、機種によって制限があるため、VB-Audioケーブルなどの仮想デバイスを用いる方法が確実です^{3 4}。UbuntuなどLinuxでは標準のモニター機能を利用するだけで実現できます⁶。自分の声も含めたい場合は追加のミキシング設定が必要となり難易度は上がりますが、VoiceMeeterのようなツールを使えば不可能ではありません⁷。まずは**PC出力音声のキャプチャ**を確立し、その上で余裕があればマイク音声の混合にも挑戦してみてください。必要な設定を一つ一つ試していくけば、ほんのりんごの環境でも目的の文字起こしシステムを構築できるはずです。頑張ってください。

¹ Stereo Mix: Works with Speakers, Doesn't work with Headphones - Microsoft Q&A

<https://learn.microsoft.com/en-us/answers/questions/4259935/stereo-mix-works-with-speakers-doesnt-work-with-he>

² setup_audio_loopback_windows.ps1

https://github.com/Takatakatake/jako-esperantlingva-paroltransskribo-kaj-interpretado-japana-korea-windows/blob/80b8e219707825eb408733b2d420284f27bffeee/scripts/setup_audio_loopback_windows.ps1

³ ⁴ README_en.md

https://github.com/Takatakatake/jako-esperantlingva-paroltransskribo-kaj-interpretado-japana-korea-windows/blob/80b8e219707825eb408733b2d420284f27bffeee/README_en.md

⁵ ⁶ Tutorial - Recording Computer Playback on Linux - Audacity Manual

https://manual.audacityteam.org/man/tutorial_recording_computer_playback_on_linux.html

⁷ VB-Audio VoiceMeeter

<https://vb-audio.com/Voicemeeter/>