

「IPtalk_s2t_Chrome」ウィンドの説明

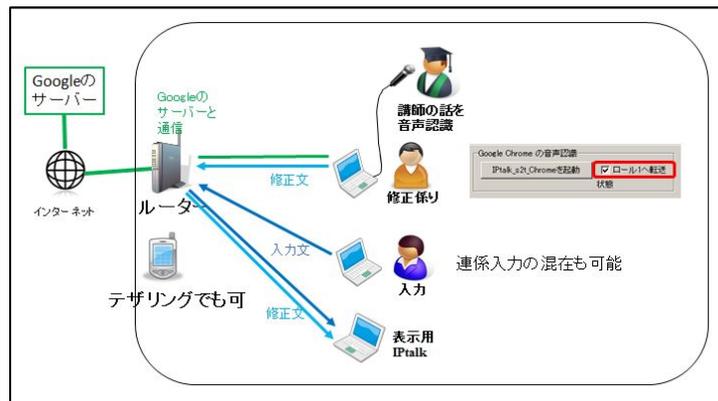
【概要】

- ・Chrome ブラウザーで音声認識した文を IPtalk に取り込むことができます。(Chrome ブラウザーが必要)
- ・Google の「Web Speech API」を使っていて、実用的な音声認識を無料で利用することができます。
- ・1 発話ごとにボタンを押す必要は無いのですが(連続認識が可能)、5 分～6 分で一度終わってしまうので、その時は、またボタンを押す必要があります。

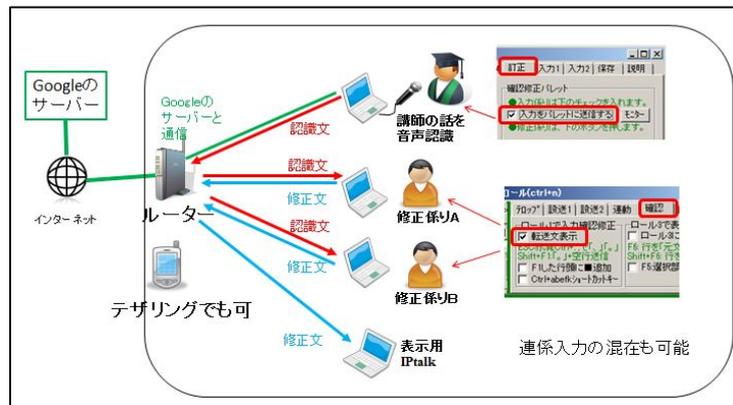
【使用イメージ】

インターネットに接続している必要があります。

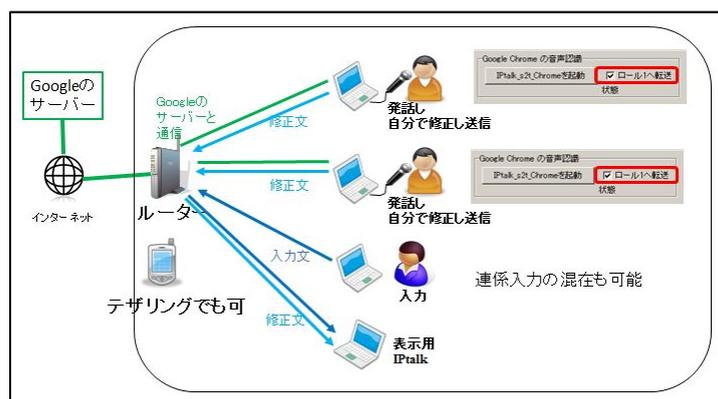
①講師の話をも音声認識して、修正してから表示に送信します。この時、通常の連携入力との混在も可能です。



②講師の話をも音声認識して、2 人の修正係りに送信します。前ロールに取り込めば、認識文の正しい部分だけを使い、他は連携入力も可能で、自由度の高い修正入力ができます。

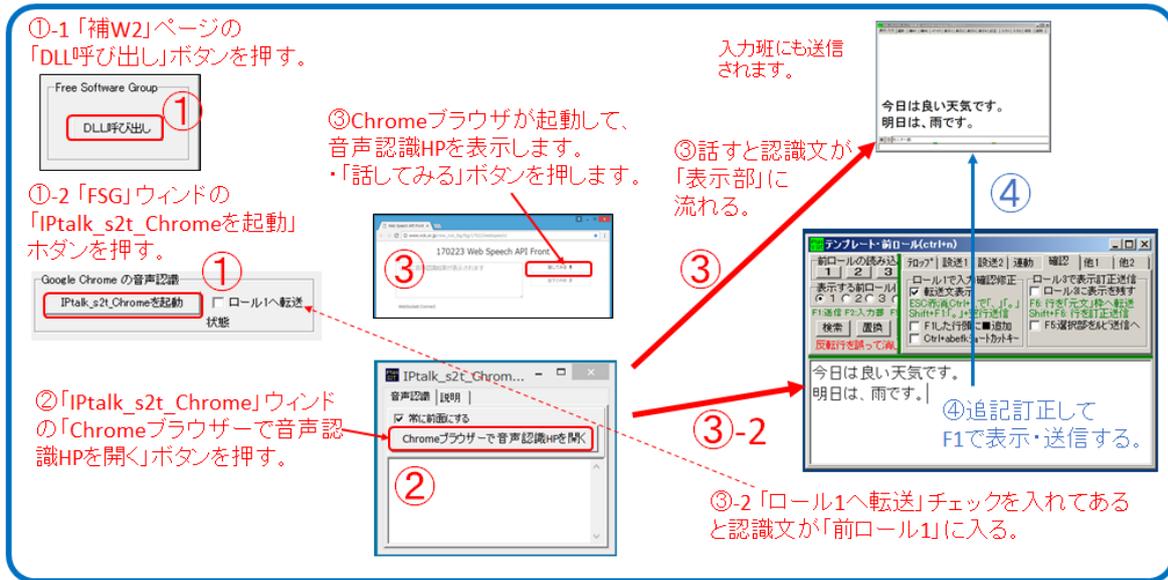


③会議で、自分の話を音声認識して、自分で修正して表示に送信します。各自がパソコンを持てば、他の発言の終わりを待たずに、音声認識で発言を準備することができます。



【操作概要】

Chromeブラウザの音声認識と修正係りのIPtalk

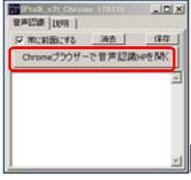
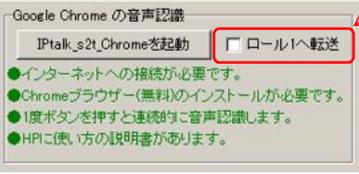


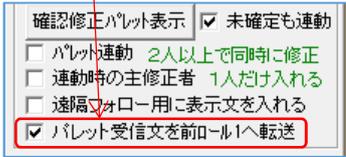
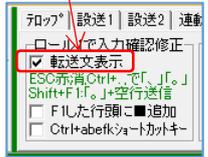
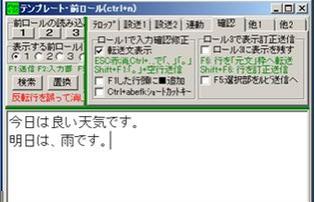
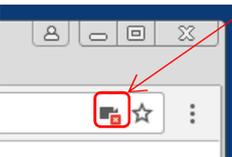
- ①「補 W2」ページの「DLL 呼び出し」ボタンを押し、「FSG」ウィンドの「DLL-2」ページの「IPtalk_s2t_Chrome」ボタンを押すと、「IPtalk_s2t_Chrome」ウィンドが開きます。
- ②「Chrome ブラウザーで音声認識 HP を開く」ボタンを押すと Chrome ブラウザーが起動して、NCK の音声認識ホームページが開きます。
- ③ここで、「話してみる」ボタンを押して、何か話すと、認識文が、表示部に流れます。話が終わったら「話すのを中断」ボタンを押します。
- ④「ロール 1 へ転送」チェックを入れれば、テンプレート前ロールのロール 1 に転送されます。

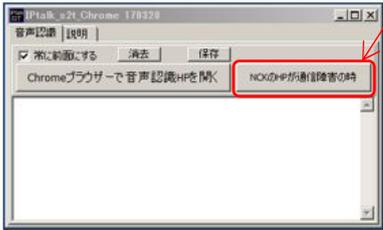
【簡単な試し方】

- ・パソコンはインターネットに接続している必要があります。
- ・Google Chrome ブラウザー(無料)がインストールされている必要があります。
<https://www.google.co.jp/chrome/browser/desktop/>
- ・マイクを接続しておいてください。

No	操作	反応
1	<p>①IPtalk9t66.exe を起動します。</p> <p>②「選択」ページの「上級者用(全機能)」のラジオボタンを入れます。</p>	<p>【注意】</p> <p>IPtalk のあるフォルダーに以下のファイルを置いてください。(同梱されているので解凍すれば、そのまま入ります。)</p> <p>・IPtalk_s2t_Chrome.exe</p>
2	<p>①「補 W2」ページの「DLL 呼び出し」ボタンを押し、「FSG」ウィンドの「DLL-2」ページの「IPtalk_s2t_Chrome」ボタンを押します。</p> 	<p>①「IPtalk_s2t_Chrome」ウィンドが開きます。</p> 

<p>3</p>	<p>①「Chrome ブラウザーで音声認識 HP を開く」ボタンを押します。</p>  <p>【注意】 「IPtalk 音声認識ページ」が開かない時の対策は後述します。</p>	<p>①Chrome ブラウザーが起動して、「IPtalk 音声認識ページ」が開きます</p>  <p>このページは暗号化して通信しています。</p>
<p>4</p>	<p>①「話してみる」ボタンを押します。</p>  <p>【注意】 「話すのを中断」ボタンが濃くならない時の対処は後述します。</p>	<p>①「話すのを中断」ボタンが濃く表示されます。</p>  <p>タブの赤い丸が点滅します。</p> 
<p>5</p>	<p>①何か話します。 右の3つのウィンドに同じ認識文が表示されます。</p> <p>【ヒント】 話が中断して一定時間経過すると「話してみる」ボタンに戻ってしまいます。 また、連続して話し続けた場合でも、最長で5分～6分程度で、認識を止めてしまいます。この場合、また、ボタンを押す必要があります。(回数の制限は無さそうです。)</p> <p>このため、音声認識が動作していることを監視できるように Chrome ブラウザーを配置することをお勧めします。</p> <p>【ヒント】 途切れなく話し続けると、Chrome ブラウザーの枠に長文を溜まってしまい、IPtalk に渡してくれません。 ある程度溜まると自動的に流れるのですが、即時性の点から問題があります。 強制的に表示に流したい時は、「話すのを中断」ボタンを押し、表示に流し、すぐに「話してみる」ボタンを押します。</p>	<p>①Google Chrome ブラウザーの「ここに音声認識結果が表示されます」枠に認識文が表示されます。</p>  <p>②「IPtalk_s2t_Chrome」ウィンドの下の枠に認識文が表示されます。(保存できます。)</p>  <p>③IPtalk の表示部に認識文が流れます。 入力班の全員に表示されます。</p> 
<p>6</p>	<p>認識文を「ロール 1」に転送します。</p> <p>①「FSG」ウィンドの「ロール 1 へ転送」チェックを入れます。</p> 	<p>①「テンプレート前ロール」ウィンドが表示されて「前ロール 1」を消去しますか？」と出るので「はい」を押します。</p> 

	<p>【ヒント】</p> <ul style="list-style-type: none"> ・認識文を前ロールに取り込んで、訂正・整文・要約を行うことをお勧めします。 <p>以下のマニュアルを参照してください。 「170422 ロール 1 修正機能と練習機能の簡単な説明.pdf」</p>	<p>「テンプレート前ロール」の「転送文表示」チェックと「訂正」ページの「パレット受信文を前ロール 1 へ転送」チェックが自動的に入ります。</p>  
7	<p>①何か話すと「前ロール 1」に認識文が入ります。</p>	<p>①認識文が「前ロール 1」に入るので、F1 で表示に流します。</p> 
8	<p>【トラブルの対処法】音声認識が始まらない時</p> <p>①Chrome ブラウザーの「リロード」ボタンを押してみる。</p> 	<p>①もう一度、「話してみる」ボタンを押します。</p> <p>・IEや FireFox など、Chrome 以外のブラウザが開いていないかも注意してください。</p>
9	<p>【トラブルの対処法】音声認識が始まらない時</p> <p>①Chrome ブラウザーの「マイクマーク」に赤い×がついている時は、「マイクマーク」をクリックします。</p>  <p>②上のラジオボタンを入れて完了ボタンを押す</p> 	<p>①「このページはマイクへのアクセスを禁止されています。」と出て来る。</p>  <p>②「マイクの許可を求めています」と出て来るので許可する。</p>  <p>もう一度、「話してみる」ボタンを押します。</p>
10	<p>【トラブルの対処法】</p> <p>「IPtalk 音声認識ページ」が開かない時</p> <p>Google Chrome ブラウザーがインストールされているかチェックしてください。</p>	<p>・Google Chrome ブラウザー(無料)がインストールされている必要があります。</p> <p>https://www.google.co.jp/chrome/browser/desktop/</p>

11	<p>【トラブルの対処法】</p> <p>「IPtalk 音声認識ページ」がタイムアウトする。 NCK のホームページの通信障害が考えられます。</p> <p>①ウィンドを横に広げると「NCK の HP が通信障害の時」ボタンが出てきます。</p> <p>②ボタンを押すと、IPtalk のホームページの「IPtalk 音声認識ページ」が開きます。 その後の操作は同じです。</p> <p>【注意】</p> <p>通常は、NCK の「IPtalk 音声認識ページ」を使ってください。</p> <p>・通信障害の頻度を知りたいのでメールで教えていただくと助かります。</p>	<p>①「NCK の HP が通信障害の時」ボタンが出て来る。</p>  <p>②IPtalk のホームページの「IPtalk 音声認識ページ」が開きます。</p>  <p>IPtalk のホームページの通信は暗号化していないので緊急用です。</p>
12	おわり	

【ヒント】

IPtalk9t66.exe があるフォルダーの中に、IPtalk_s2t_Chrome.exe を置く必要があります。

IPtalk の自己解凍ファイルに同梱されているので、そのまま解凍すれば自動的に入ります。

【ヒント】

「IPtalk_s2t_Chrome」ウィンドは、実は、IPtalk とは別ソフト (exe ファイル) になっています。

ただし単独では使えず、IPtalk から自動的に起動・終了するので、IPtalk のユーザーは、別ソフトであることを意識する必要はありません。

このようにした理由は、音声認識のプログラムは巨大になってしまうため、IPtalk にそのまま組み込んでしまうと、音声認識を使用しないユーザーには迷惑だろうと思ったためです。

そのため、配布用の自己解凍ファイルも音声認識機能用とパソコン要約筆記用の 2 種類を準備しました。

【ヒント】

IPtalk は、パソコンのレジストリやシステムファイルを全くいじりません。

自己解凍ファイルを実行してファイルを展開し、IPtalk を実行しても、パソコンの Windows は元の状態のままです。IPtalk を解凍したフォルダーにある exe ファイルと dll ファイルで動いています。

解凍したフォルダーのファイルを消してしまえば、パソコンは、完璧に元に戻ります。

今回の音声認識の機能を使っても同じです。

つまり、「ちょっと試してみたら、パソコンがおかしくなってしまうて困った」というような心配は全くありません。

(これは、IPtalk をずっと使っている方には常識だと思うのですが、この頃、質問メールの中に「IPtalk を再インストールしても・・・」というような表現を使う方が多いので、ちょっと気になっています。)

【IPtalk に音声認識を組み込んだ背景】

IPtalk のユーザーからパソコン要約筆記、PC テイクの代わりに音声認識ソフトが導入されるケースが増えていると聞いています。しかし、現状の認識率(安定性も含めて)では、訂正係りが必須だということは、IPtalk のユーザーには今さら言うまでも無いことだと思います。ところが、音声認識ソフトを導入する場合は、「ほとんど認識率 100%だ」「初心者も訂正係り 1 人で情報保障できる」と言われて半信半疑でいるということのようです。それは「デモを見た」という話だと思うのですが、「デモとは、そういうもの」で、すべての現場がそう行くほど甘くはありません。

問題と聞いているのは、話者や先生が認識文をチェックする訳ではなく、認識文に責任を持つ人がはっきりしないとか、利用者が不確かな音声認識の文を見て分かつたつもりになってしまうことが放置されがち、などです。もちろん、それは音声認識ソフトに責任があるわけではなく、使い方の問題です。

つまり、せっかくの最先端の技術・ソフトを導入しても、情報保障という観点での配慮、要約筆記の歴史の中でさんざん言われて来たいろいろな配慮が不足しているのだらうと思います。音声認識ソフトを導入する時に、長年、IPtalk で情報保障して来たみなさんのアドバイスが活かされれば、このような話にはならないだらうと思うわけです。

話を聞いてみると IPtalk のユーザーは、問題だと思う時もあるし、アドバイスしたい気持ちはあるが、自分が音声認識の経験が無いから強く言えないということが一番の理由のように感じました。

ちょっと試すには、音声認識ソフトは高過ぎると思います。それで無料の実用的な音声認識を組み込みました。

(他には、「ちんたら訂正するより、私が入力した方が速くて正確！」という意見も多いと思います、もちろん。)

みなさんに音声認識を使った入力を試してもらい、情報保障としての方法を考えていただき、上のようなケースに積極的に関与して、アドバイスできるようになっていただき、「音声認識ソフトを入れたから、これからはパソコン要約筆記や手話通訳は頼まないよ」などというような話を聞かないようになれば良いと思っています。

【お願い事項】

音声認識を試していただく時に、以下の点も検討していただけるとありがたいです。結果を教えていただければ、ホームページなどで共有させていただきたいと思っています。

①可読性(整文、要約)を重視した字幕のトライ

「ロール 1 確認修正機能」を使って、音声認識の利点である即時性は、リアルタイムで前ロールを作成する事に使い、入力者の負担が軽減した分を可読性や正確性に振り向けて、より質の高い文字通訳を実現できないか試していただけないでしょうか。

「170422 ロール 1 修正機能と練習機能の簡単な説明.pdf」を参照してください。

②少ない人数での入力の可能性のトライ

音声認識ソフトを導入する理由の一番が、入力者不足の解消、人件費の削減のようです。

「ロール 1 確認修正機能」を使って、音声認識ソフトを補助入力者と見て、2 人連係入力と同等の字幕を 1 人で可能か試していただけないでしょうか。

③音声認識を使った入力班の編成はどうあるべきか？

現場のマイクの状態や話者の話し方はさまざまですから、「その日」の「その現場」によって認識率は大きく変動します。

そのため、「入力した方が速い」から「そのまま表示できる」まで、いろいろなケースが考えられます。

最悪のケースを考えれば、通常連係入力の人数が必要になります。

最良のケースを考えれば 1 人で良いということになります。

会議などでは、A さんの発言は認識率が高く、B さんの発言は入力が必要というケースもあると思います。

入力班の編成は、連係入力と変わるのでしょうか？

④リスピークによる音声入力のトライ

スマホやパソコンの音声認識の普及で、この頃は、話者の音声の直接認識させる場合が多いようですが、音声認識といえば、一般的には、まだまだNHKのようなリスピーク方式です。

リスピークする人と訂正係りというように2人で行う方法が一般的と思いますが、例えば、リスピーク(音声入力)しつつ、自分で訂正して表示に流すという、1人で行う事は可能でしょうか？

また、1人をリスピーク、もう1人は通常のキーボード入力というような、異なる入力方式での連係入力は可能でしょうか？さらに、キーボード入力をしつつ、間に合わない部分を入力者がリスピークして音声認識させて前ロール1に入れておき、後で表示に流す、などというようなキーボード入力を主とし、補助入力的な利用は可能なのでしょうか？

⑤他人に頼らない字幕表示機器として使えるのか？

(これは、利用者が聴者の介在なしに、音声認識を利用する場合を想定しています。)

認識率100%が保証されるのなら、たぶん使えるだろうと思います。

しかし、現在の技術レベルは、認識率が安定せず、100%にはほど遠い状態です。

では、どの程度の認識率なら使えるようになると考えれば良いのでしょうか？

情報保障者の立場からは、「聴覚障害者は、誤変換があっても、それに気づくことができない。間違えた文を信じ込んでしまう。それが怖い」ということで、「100%でないダメ」という意見が一般的だと思います。

では、利用者が、「音声認識は誤変換がある。」と判った上で「自己責任で使いたい」と考えた場合はどうでしょう？

利用者に、「このような条件なら」とか「このような使い方はダメ」というアドバイスができればよいと思います。

聴者が、利用者の立場で試してみれば、音声と認識文を比較することができます。

あるいは、話者が音声認識に協力的だった場合はどうでしょう？

例えば、個人的な集まりで友人たちと雑談をする時、音声認識を使うことはできないのでしょうか？

まだまだ、難しいのでしょうか？

それとも使える場合もあるくらいに進歩しているのでしょうか？

⑥現場でのマイク設置などに関するノウハウの情報

認識率向上のポイントは、ノイズの無い音声をパソコンに入力できるかという点に尽きると思います。

いろいろな工夫があると思います。

情報をいただければ、HPで共有させていただきます。

⑦Googleの音声認識サーバーの混雑度の情報

入力現場で確実に音声認識が利用できることが最低限必要です。

有料の音声認識ソフトでも、認識サーバーが混雑していると問題が出るがあると聞きます。

私が使う限りは、そのような状況は経験していません。

みなさんには、実際に使ってみて、そのような現象がどの程度出るのかの情報を送っていただきたいと思います。

⑧クラウド型音声認識サービスを利用した情報保障の守秘義務の考え方

これは、Googleの音声認識だけの問題ではなく、他のソフト(クラウド型音声認識サービス)についても言えることだと思います。

・個人情報を含む音声と認識文をインターネットで通信します。これは、通信を暗号化すれば情報漏洩の心配はないと思います。(「IPTalk 音声認識ページ」はSSLの暗号化通信を使用しています。)

・Googleの音声認識サーバーは、サーバーログとして、それらを何らかの形で保存すると思われます。

この事については、入力者は、利用者・話者・主催者などに説明して、クラウド型音声認識サービスを利用する事の承認を得る必要があります。(ダメな時は、パソコンで動作する音声認識ソフトを使用するか、通常の連係入力になります。)

Google の情報の取り扱いの概要は以下のようです。

Google は、音声認識サービスなどを利用した場合の情報の取り扱いについて「Google 利用規約」に従うと言っています。「Google 利用規約」には

- ・Google のプライバシー ポリシーでは、本サービスの利用におけるユーザーの個人データの取り扱いとプライバシーの保護について記載しています。
- ・ユーザーは、本サービスを利用することにより、Google のプライバシー ポリシーに従って、Google がユーザーの個人データを利用できることに同意することになります。

と書かれています。

「Google のプライバシーポリシー」には、

- ・Google は、どの Google サービスから収集した情報も、そのサービスの提供、維持、保護および改善、新しいサービスの開発ならびに、Google とユーザーの保護のために利用します。

などと利用目的が書かれています。

「情報の保護」の項には

- ・Google は、Google が保持する情報への不正アクセスや、不正な改変、開示または破壊から、Google とそのユーザーを保護するよう努めています。

などと書かれています。

Google の音声認識サービスは、世界中から膨大な利用があるので、生の形や個人が特定されるような形で保存したり利用しているとはとても思えませんが、この文章をそのまま読むと、音声認識文や音声は Google 内部で利用される可能性があるということになります。また、情報漏洩に関しては、Google は「保護するように努める」と言っています。

Google の規定やポリシーの内容は、他のクラウド型音声認識サービスでもだいたい同じだと思います。

一方、今までの実績としては、Google の音声認識サービスから個人情報や機密情報が漏れてトラブルになったという話は聞きません。また、多くの人々が使用している Gmail も同じ規定・ポリシーですが、何の疑いもなく、個人情報を含むメールをしていると思います。

(ちなみに、Gmail でメールを解析して内容に関連するテキスト公告がページに表示されるということに関して、「プライバシー侵害ではないか?」という批判が上がったことに対して、Google は「このプロセスが自動で行われており、メールが不必要に開示されないよう完全に保護されている限り、電子メールの自動スキャンはプライバシーの侵害にはつながらない」としているとのことです。<https://ja.wikipedia.org/wiki/Gmail>)

これらのことを総合的に考えると、一般的な会議、講演、授業などであれば、利用可能と私は思うのですが、みなさんの意見はいかがでしょう?

関係者の承認を得ることが困難と思われるでしょうか?

「Google 利用規約」は以下です。

<https://www.google.com/intl/ja/policies/terms/>

「Google のプライバシーポリシーへようこそ」は以下です。

<https://www.google.com/intl/ja/policies/privacy/>

「Google Chrome のプライバシーに関するお知らせ」は以下です。

<https://www.google.com/intl/ja/chrome/browser/privacy/>

【今後の予定】

今回の IPtalk9t66 は、音声認識部分は Google 任せで、「音声認識を情報保障にうまく使う方法」の試行錯誤の環境を、全国の IPtalk のユーザーのみなさんに提供したいと考えて作りました。多くの人に使ってもらって、みなさんと考え、改良して行けたら良いと思っています。

また、もし、「Google の音声認識は情報保障に役立つ、Google の音声認識を本格的に使いたい」となった場合、現在の無料サービスの範囲で良ければ構いませんが、「5分に1度ボタンを押すのは面倒」という要望などが出てくると想像しています。

それらは、Google の無料サービスではなく、有料サービスを使えば、たぶん簡単に実現できると思います。

しかし、Google の有料音声認識サービスは、ベータ版で1時間使うと170円くらいの従量制の料金(注1)です。

全国で使った場合の金額はいくらになるか分かりませんが、私が個人的に負担するのはちょっと難しいと考えています。

IPtalk は無料にしたいので、その時は、方法を考えなくてはいけません。

とは言うものの、音声認識の料金については、私は楽観的です。

たぶん、5年もすれば、Windows や iOS に、無料の高性能な音声認識が標準的に組み込まれて(注2)、IPtalk などのソフトから利用できるようになるだろうと予想しています。

また、念のため言えば、遠い将来、たとえ無料で認識率100%になったとしても、パソコン要約筆記など、人の介在が無くなってしまふことは、福祉的な観点からも、無いと私は考えています。

その時のために、今から、いろいろと問題を整理したり、うまい使い方を考えておく必要があると思います。

また、他に無料で利用できる音声認識サービスがあるので、今後も、できる範囲で、IPtalk で使えるようにして行きたいとも思っています。

使った場合は感想などのフィードバックをいただくと助かります。

(注1)<https://cloud.google.com/speech/>

(注2)あまり知られていませんが、今でも、Windows と iOS(Mac)には、標準(無料)で音声認識が組み込まれています。

・Windows で音声認識を使用する方法

<https://support.microsoft.com/ja-jp/help/14213/windows-how-to-use-speech-recognition>

・Mac でテキストを音声入力する

<https://support.apple.com/ja-jp/HT202584>