

# デジタル 最前線

突撃レポート

漫画：なとみみわ  
編集：イシカワ



NEC  
<http://www.nec.co.jp/>

今月の  
テーマ

## 音声認識技術

日本電気株式会社 共通基盤ソフトウェア研究所 研究部長

佐藤研治さん

音声を言葉として機械に認識させる「音声認識」という技術がある。  
パソコンに文字入力ができたり、翻訳ができたりと、応用範囲の広い注目の技術だ。  
今回は、この技術を48年にわたって研究しているNECを取材した。





### 「音」の情報を手がかりに照合

話者適応	音響モデル
Aさん、Bさんなど、 人の声ごとの波形の特徴	「ア」「イ」「ウ」など、 言葉ごとの波形の特徴
Aさんの声  Bさんの声	「ア」の音  「イ」の音



デジタルの文字

「今日は快晴」

アナログの音声

「キョウハカイセイ」

### 「意味」の情報を手がかりに照合

言語モデル	単語辞書
単語のつながり方を 3単語前後の長さで記憶した統計情報	あらかじめ登録された 単語の読み方や表記
「今日」→「は」→「快晴」○ 「改正」×	「今日」 ○ 「快晴」 ○



#### 話者適応

話し手ごとの波形の特徴を事前に把握しておけば、発音がその人独特の発音だったとしても言葉をより正確に認識できます

これは Mさんの「ア」だ

音声認識

#### 音響モデル

例えば「ア」の音を持つ波形の特徴を事前に把握しておけば、同様の波形を持つ音が鳴った時にそれが「ア」だとわかります

「ア」だ あっ!!

どっかで聞いたような...

音声認識

うーん、何となくわかったんですが細かいところが今ひとつ...

では、一つずつ説明しましょう

まず前提として音をデジタル情報としてあつかう場合、上図のような波形のデータで表されます

音声認識

#### 言語モデル

前後の単語の並びからその言葉の意味を推定します。これにより、同音異義語でも正しい漢字で認識できます

「今日」→「は」→「快晴」○  
「改正」×

きっと「快晴」が正しいぞ

音声認識

#### 単語辞書

例えば「キョウ」という音を認識した時に「今日」や「京」などの単語として認識できます

「今日」  
「京」  
もしくは「キョウ」

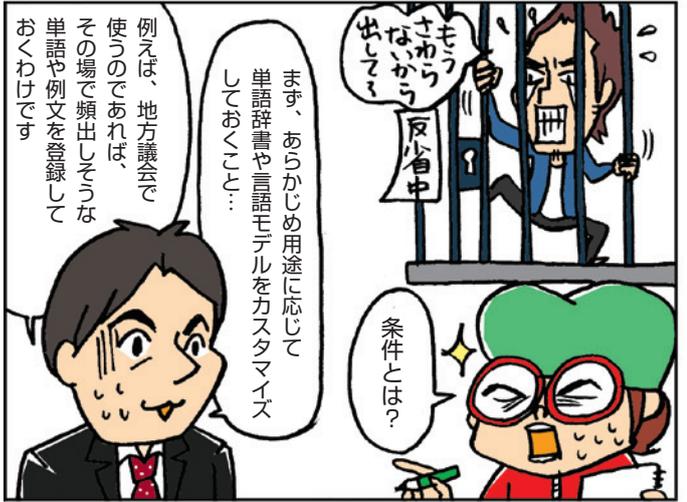
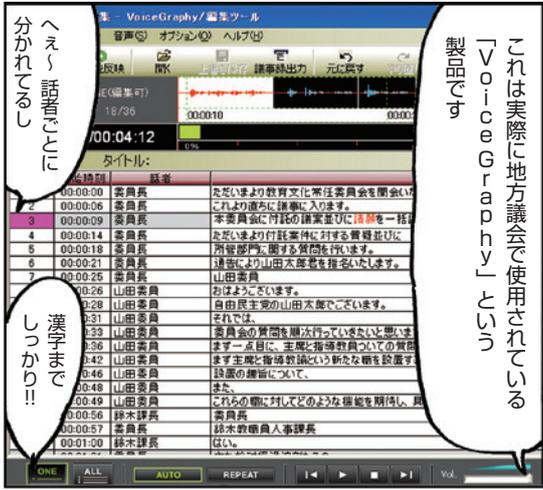
音声認識

では、「意味」の方は?

言葉の「音」だけを認識し、羅列してもそれだけでは意味が通じません

そこで、「単語辞書」と「言語モデル」も組み合わせると言語化します

音声認識





**日** 本における音声認識技術の研究は、1960年にNECと京都大学の共同研究として始まった。80年代以降の大きな技術革新を経て、現在は一部の地方議会やコールセンターなどで音声認識関連製品が活躍しているように、その認識率はすでに実用レベルにまで達している。次に期待したいのは、音声認識に「機械翻訳」と「音声合成」の技術を複合した「自動通訳」の実現だ。NECではこの技術の研究に1983年から取り組んでおり、2001年には世界初となるパソコン用通訳ソフトの製品化、2007年には携帯電話用通訳ソフトの試作版を広報している。携帯電話を翻訳機として使える日はそう遠くないはず!

取材後記  
勝手に未来予想!