

金沢大学サテライト・プラザミニ講演

日 時:平成20年8月30日(土) 午後2時~3時30分

会 場:金沢大学サテライト・プラザ 講義室

演題:「コトバの音声の諸相」

講師:柘植 洋一 (金沢大学 人文学類 教授)

はじめに

私は言語学を専門にしています。言語学というのは、コトバの仕組みを明らかにするという学問ですが、特に私はエチオピアの言語が専門です。エチオピアはアフリカの北東部にあり、アフリカの中では珍しい古くからの独立国です。この間のオリンピックで、残念ながらハイレ・ゲブレセラシエという選手はマラソンに出なかったのですが、アベベ・ビキラなど有名なマラソン選手を輩出した国です。エチオピアでは80ぐらいの言語が話されているのですが、その中の主要な言語であるアムハラ語と、エチオピアの中でもずっと南西部で話されている一群の言語を調査しています。今日はそのエチオピア言語の話ではなくて、コトバを調べる上で一番基礎になる音声の話をしていただきます。

I. コトバの最小の単位は「音」

1. コトバ=意味内容を伝達するための仕組み

世界では、6000~7000という非常にたくさんの数のコトバが話されています。日本では日本語やアイヌ語、それから周りにいきますと朝鮮語、中国語、タイ語、ベトナム語からはじまり、中学校で勉強する英語、それからドイツ語、フランス語などいろいろな言語があります。皆さんに知っているコトバを挙げてくださいますと、最初はいろいろ出てくるのですが、大体10か20ぐらいで終わりになってしまいます。しかし、実際には非常に多くの数の言語が話されています。それらの言語はそれぞれの仕組みは違っていますが、コトバという点ではすべて同じものです。

このコトバというのはいろいろな役割を果たしているのですが、その中でも最も大事なものはやはりコミュニケーションを行う手段だということです。コミュニケーションというのは何かをお互いに伝え合うことですが、具体的には、自分が思っていることを相手に伝えるために作り出されたものです。人間がいつごろからコトバを持つようになったかとい

うのは証拠がないのでよくわかりません。ネアンデルタール人になるとどうだったかわかりませんが、少なくとも現生人類、私たちの近い祖先が持っていたことは確かです。

私たちは、コトバでもって伝えたいことを相手に伝える、また相手からその相手が伝えようとしている内容をコトバでもって受け取るということをしています。それがどういう仕組みになっているかという、基本的には、意味と音を結びつけて利用しているということです。つまり、意味というのは伝えたい内容です。それを伝えるために、音（おん）を利用しているのです。例えば「私はおなかが空いた」ということを相手の人に伝えるときに、頭の中に浮かんだ「おなかが空いた」ということをそのまま相手にぼんと伝えられればいいのですが、それはできません。ですから、何らかの形を利用してその内容を伝えます。人間は、その際に音を選択しました。音を使って、その意味内容を伝えるという形でもって、内容の伝達を行っていきます。私たちは自分の伝えたい内容が相手に伝わらず、非常にもどかしい思いをすることも多いのですが、多くの場合は何らかの形でコミュニケーションが成立しています。ということは、自分の伝えたいことを、基本的にはコトバでもって相手に理解してもらっているわけです。

この場合に、コトバの仕組みをはどうなっているかという、まず、普通私たちは文の形で相手に伝えたい内容を伝えています。伝えたい内容には無限に近く、それを表すためにほぼ無限に近い文を作り出しています。その文というのは、分解していきますと、まず単語に分かれています。つまり、単語を組み立てて文を作り出しているのです。単語の数はどのぐらいかという、よく「この辞書は20万語収録しています」とか「8万語収録しています」と書いてあるように、誰もそれをきっちりと調べ上げた人はいないのですが、ほぼどの言語でも恐らく十万以上という数の単語を持っています。

その大量な数の単語をさらに見ていきますと、それらはすべて実際には数十の音に還元されるのです。ということは、音と意味が結び付いて、一つの単語を作り上げています。例えば犬なら日本語の「犬」という単語がありますが、その場合に音の面では「i-n-u」という音です。それから意味の面では、〈犬〉という意味が（〈犬〉とどういう意味かというややこしい議論は抜きにします）あり、この両者が結合して、全体として日本語の「犬」という単語が出来上がっています。

この音を見ていきますと、最終的には日本語を作り上げている音の種類は数十ですし、ほぼどの言語でも100以下しかありません。この音にある意味を結び付けます。それを幾つかつなげて今度は別の単語を作っていくということをやっていくわけです。そうすると、

一つの音に一つの意味を結び付ける場合、例えば 50 の音があるとすると 50 の単語を作ることができます。今度は音を二つ並べると 50×50 で 2500 の単語を作ることができます。このようにして、非常に少ない 100 以下の数の音を、一つあるいは二つ、あるいは三つとつなげ、そのそれぞれに意味を結び付けることによって、数十万という単語を作っていきます。その何十万という単語を、単語の並べ方の規則、つまり文法に従って並べることによって無限の文を作り出していき、無限の状況に対応しているのです。人間の言語は、このようにたった数十の音から数万、数十万の単語を作り、そして更に無限の文を作っているわけで、非常に経済的な仕組みから成っているのです。

さて、私たちはコトバを考えるときに、すぐに単語や文というものを思い浮かべ、文法あるいは語の作り方、語の意味を考えるのですが、今日のテーマは、音を見直してみようということです。

2. 「音」と「文字」

しばしば私たちは文字と音を混同することがありますが、文字というのは、あくまでも音を視覚に訴える形で固定したものに過ぎません。ですから、コトバの一番小さい単位として文字を考えるのは適当ではないのです。人間の歴史を見ても、文字が生まれたのはせいぜい 5000 年ぐらい前です。それに対して、人間がコトバを持ったのはそれよりもずっと前ですし、世界の言語を見ても、文字で書かれない言語はたくさんあります。

日本語は仮名を使っています。漢字も使いますが、幼稚園に入る前から平仮名、片仮名を覚えます。覚えると親に喜んでもらえ、自分もうれしいということで、一生懸命文字を勉強していきます。ですから、コトバと文字が一体化しているような感じを持つのですが、基本になるのはあくまでも音声、音なのです。それを具体的に見ていただきたいと思いません。

逆さコトバあるいは回文といって、上から読んでも下から読んでも同じコトバがあります。トマト、新聞紙 (シンブンシ)、竹やぶ焼けた (タケヤブヤケタ)、ダンスが済んだ (ダンスガスンダ) など、幾らでも作れます。「ダンスが済んだ」から「イカのダンスは済んだのかい (イカノダンスハスンダノカイ)」なども作れますし、英語でも同じような回文があります。ただ、これはあくまでも文字を見ているから、上からあるいは左から読んでも右から読んでも「トマト」は「トマト」、「シンブンシ」は「シンブンシ」というだけです。英語の例もそうです。「Madam, I' m Adam.」という回文がありますが、これも左から見て

右から見ても全くつづりからして同じになります。

では、実際の発音はどうかということを知りたいと思います。「トマト」という音声を実際に発音して、逆再生してみるとどのように聞こえるでしょうか。「トマト」は「トマト」だろうと思われる方もいらっしゃるかもしれませんが、そうではないという方もいらっしゃると思います。では、まず「トマト」を聞いてみます。私が先ほど録音したものです。

『トマト』

これが私の声です。これを逆再生してみます。そうするとどうなるでしょうか。

『オタモット』

もしトマトという文字が音の一番小さい単位であったならば、逆に流してみても同じになるのですが、実際はそうではありませんでした。これはどういうふうになっているかというと、ローマ字で書いてみるとお分かりになると思います。「トマト」というのは [t-o-m-a-t-o] という六つの音がこのような順番に並んで発音されたものです。日本語の仮名というのは、「トマト」のように一つの子音＋一つの母音を一つの文字で表すという音節文字というシステムです。従って、この文字はさらに小さい中身を持っているのです。ですから、「トマト」という内容はさらに [t-o-m-a-t-o] という中身に還元され、逆転再生するという事は右側から進むということです。そうすると、[o-t-a-m-o-t] となります。もう一度聞いてみましょう。

『オタモット』

ちょっと変に聞こえるのは、一つは音の高低があるからです。それから、実際の音声というのは [t] [o] [m] [a] [t] [o] というものをばらばらに発音しているのではなくて、連続して発音しているのです。そうすると、ある音からある音に移り変わる部分があり、そういうところが逆転再生すると耳慣れなく聞こえるというわけです。

もう一つ、聞いてみましょう。

『イシンーニシ』

これは何かというと、もうお分かりになった方もおられるかと思いますが。これを逆転再生してみますと「シンブンシ」と言っています。これはローマ字で書いても、ちょっと分かりにくいのです。[shinbunshi] を逆にすると、[ishnubnish] という発音なのです。

(注：sh と 2 つのローマ字で書いてあるのは実際は 1 つの音です。従ってここで、逆の場合も hs ではなく、sh と書いています。)

今度はこれを聞いていただきましょう。

『イカヤキ』

これはイカを焼いた「イカヤキ」と発音しました。これを逆に再生してみます。

『イカヤキ』

ちょっとごちないですが、ほぼ同じように聞こえます。「イカヤキ」というのは、トマトや新聞紙と違って、文字を逆に読むと「キヤカイ」になってしまうのですが、しかし、録音して逆転再生してみますと、お聞きいただいたようにほぼ同じようになります。これもローマ字で書いてみればお分かりになると思います。[ikayaki] は [ikayaki] となります。

もう一つ同じような例を聞いていただきます。

『アカサカ』

これも逆に読んだら「カサカア」になるのですが、逆転再生しますと、

『アカサカ』

これもイントネーションがちょっと変ですが、基本的には同じになります。このように録音して逆転再生してみると、コトバの一番小さい単位は音であるということをご理解いただけたかと思います。

Ⅱ. 「音声」とは？

1. 音声（言語音）

次に、音声とは何かということです。人間はいろいろな音を出します。例えばあくびをしたり、いびきをかいたり、くしゃみをしたり、おならをしたり、いろいろなところからいろいろな音を出します。その中でコトバに用いられる音を音声といいます。この音声は具体的にどのように出されるかという、皆さんが毎日コトバを使っているということは発音をしているということですが、どのように発音しているかということはあまり反省されたことはないのではないかと思います。

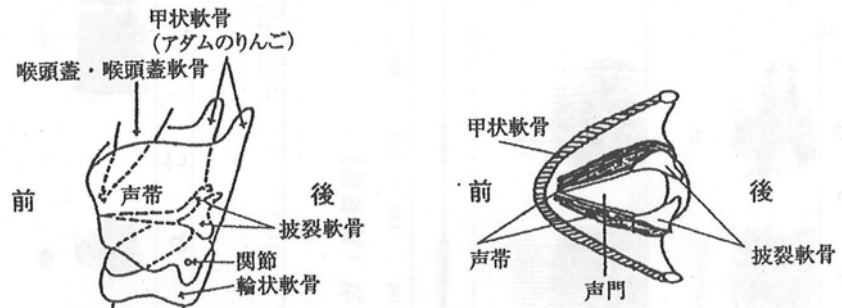
外国語を勉強するときには、例えば「この言語にはこのような音声があって、日本語にはないから注意なさい」と言われて注意を払うことはあると思いますが、自分が使っている日本語の発音、自分の発音をじっくり観察する、ましてや一つ一つの発音がどのような仕組みで作られているかということに思いをめぐらせた方はあまりいらっしゃらないかと思います。今日はそこをご理解いただければと思っています。

2. 音声はどのように作り出されるのか？

(1) 音声器官

人間は自分の体を使って音声を作り出しているのですが、音声を作り出すための人間の身体の部分を音声器官といいます。まず音声の元になるのは、人間が吐く息です。これが音声の元になります。ですから、一番元になる部分は肺から外に出される空気の流れです。人間は毎日呼吸をしています。しなければ死んでしまいます。吐く息を呼気、吸う息を吸気といいます。音声を作り出す際には呼気を利用しています。例えば息を吸いながら発音してみましょう。「カナザワ」を息を吸いながら発音してみましょう。これは、普通の音声の出し方ではなく、普通は吐く息を利用しています。

③喉頭と声帯

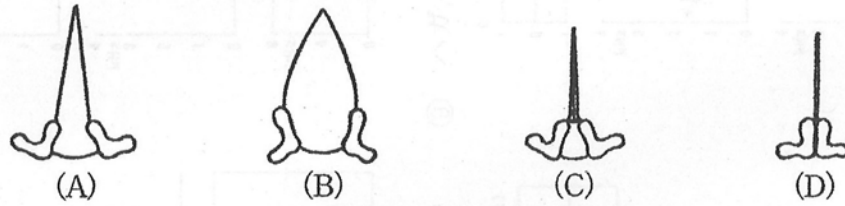


今石元久（編）『音声研究入門』（岩波書店、1999）

音声器官は、大きく三つの部分に分かれます。まず肺から外に空気を出す、吸って吐きます。その肺から出た空気は気管支から気管を通ります。気管から上がっていった空気は、今度は男の人ですと表から見えますが、のど仏に達します。ここが2番目に大事な働きをするところです。ここには声帯があります。まず、肺で空気の流れを作り、今度は、上がってきた空気は声帯を通ります。のど仏といわれるのは喉頭（こうとう）という器官で、甲状（こうじょう）軟骨と披裂（ひれつ）軟骨と輪状（りんじょう）軟骨という軟骨に囲まれたボックスになっています。このボックスの中に声帯があるのです（③喉頭と声帯・左図）。

右側に声帯の絵が描いてありますが、これは上から見た図です（③喉頭と声帯・右図）。そうすると、粘膜の二つ、一对のひだが前から後ろに向かって張られています。これが声帯で、ぴたっとくっついていてるのは、完全に声帯が閉じた状態で、いわゆる息を詰めた状態です。ちょっと皆さん、息を詰めていただけますか。「アー」と言って最後に息を詰めてみましょう。これが閉じた状態です。

④声門の状態



服部四郎『音声学』（岩波書店，1951）

さらに、声帯と声帯の間隙間を声門といいます。別に門があるわけではなくて空洞の部分ですが、これが空いている状態が（A）の状態です。（B）は特に息を大きく吸ったときで、非常に広く広がります。（D）は完全に声帯を閉じた状態です。（C）はほとんど声帯がくっついた状態で、ここを無理やり下から息が上がっていきこうとすると、声帯が震えます。このようにして出されるものを「声」（こえ）といいます。「アー」と発音して、のどのところに軽く手を当てていただきますと、震えているのがお分かりになると思います。あるいは、耳をふさいで「アー」と言うと頭の中でわーっと反響しているのがお聞きになれると思います。

それに対して、例えば歯と歯を閉じて息を吐いてみてください。そうすると、そのときにはのどのところは震えませんし、振動していません。これに対して先ほどのように、声帯が完全に閉じているのではないけれども、非常に密着に近い状態になって、そのときに下から空気が上がってくると声帯がぶるぶると震えるわけです。これが声であり、「アー」「イー」「ウー」「エー」「オー」という音の元になるものです。多くの音声は、ここで声が出ます。

今度は、声帯よりも上の部分です。声帯よりも上の部分というと、空洞の部分が三つあります。一つが咽頭（いんとう）です。それから鼻腔（びくう/びこう）です。もう一つは口腔（こうくう/こうこう）です。これらを含めて、声門よりも上にあるこのような空洞の部分をもとめて声門上腔（せいもんじょうくう）といいます。ここで今度はいろいろな種類の音を作り出します。いろいろな種類の音を作り出すことを「調音」（ちょうおん）といいます。医学関係では構音（こうおん）といいます。言語学あるいは音声学では調音といいます。調音の「調」というのは、調べるというよりもむしろピアノの調律などの整えるという意味で、音を整える、いろいろな種類の音を出すということです。音声器官は、このような働きから三つの部分に分かれるわけです。

例えば、声帯で声が出ると言いましたが、日本語には「ア」「イ」「ウ」「エ」「オ」とい

う五つの母音があります。これは、出る声は同じですが、口の中の形を変えたりすることによって、いろいろな違った音声を出しています。

模擬的に声の代わりをする機械があります。(デモンストレーション) この機械を口に当てて口の形を変えるだけで、違った種類の音が出るというものです。要するに、元になるもの(音)は同じで口の形を変える、特に口の中でも大事な働きをするのが舌なのですが、その動きを変えることによって、違った種類の音を作り出しているというわけです。

(2) 音の分類

①重要な区別

では、そのようにして出される音はどのような種類があるのでしょうか。まず大事な区別として、母音(ぼいん)と子音(しいん)という区別があります。それはどういう区別かというと、声道(声門上腔)から上のところで空気の流れが基本的に邪魔されないで出される音を母音といいます。どこかで邪魔されて出る音の子音といいます。ですから、「アー」「イー」「ウー」「エー」という音は、声門上腔から上の部分でどこにも邪魔されていません。それに対して[s][f][sh]というのはみんな子音です。どこかで邪魔されています。あるいは[r][h]という音もみんな子音です。これが一つ大事な区別です。

もう一つの大事な区別は、有声音と無声音です。声帯が震えるか震えないかということです。つまり、今言った[s][f][sh]というのは全然震えていません。このようなものを無声音といいます。日本語でいう濁音と清音の違いが、ほぼこれに当たると考えていただいて結構です。有声音が濁音、無声音が清音ということです。

ただ、気を付けていただきたいのは、日本語では例えば「カ」が清音で「ガ」が濁音とありますが、「カ(ka)」と言った場合には母音を含んでしまっているので、無声音や有声音といった場合には、この場合に無声なのは[k a]の最初の子音[k]、有聲なのは[g a]の最初の子音[g]についてなのだということです。何度も言いますが、「カ」というのは[k]という音と[a]という音から成り立っていて、これを続けて発音すると「カ」になります。「ガ」という音は[g]という音と[a]という音で、これを一緒に続けて発音しますと「ガ」という音になります。この違いをもたらしているのは、[k]と[g]の違いです。[a]というのは共通で同じです。

②子音の分類

今度は、母音と子音について簡単にどのように区別されるかということを見ておきます。いろいろな子音があるのですが、子音は三つを押さえれば分類することができます。ポイントは先ほど言いましたように、子音というのはのどから上の部分でどこかで空気の流れが邪魔されて出る音だということですが、どこで邪魔されるか、そしてどのように邪魔されるかという二つの違いが大事なのです。そしてもう一つ、無声か有声かということです。

まず、どこで邪魔されるかということですが、例えば「パ」と発音するときには、両方の唇がくっつきます。唇を閉じている間は何も音が出ておらず、唇を閉じて開いたときに音が出ます。

次に「タ」と発音してみてください。そうすると、今度は唇は関係ないですね。どこで空気の流れを邪魔していますか。そうです、舌の先の方が、上の歯の裏やあるいは歯茎辺りにくっついて、これをはがすときに「タ」という発音になるわけです。

では、「カ」はどうでしょう。「カ」と発音するときには鏡で見えますと、舌の後ろの方がわっと盛り上がるのがお分かりになると思います。その部分が口の中の上面にくっついて「カ」となります。

このように、「パ」「カ」「タ」の違いは、邪魔される場所が①上下の唇で空気の流れを閉じて、それを開くときに音を出すか、②舌の先と歯あるいは歯茎、あるいは③舌の後ろの方と軟口蓋という口の上の部分で閉鎖を作って開放するかという場所の違いです。

もう一つは、どのような邪魔の仕方かということです。もう一度「パ」と発音してみましょう。今度は「ファ」と発音してみてください。そうすると、「パ」「ファ」は両方とも唇が関係しているのですが、「ファ」の場合には完全にくっついて開くのではなくて、唇と唇が狭いすき間を作って「ファ」という音を作り出しています。このように空気の邪魔の仕方も、完全に空気の流れを止めて開くか、あるいは止めないで狭いすき間を通して空気を出させるか、このようなものを摩擦音といいます。例えば [f] [v] [s] [z] というのは摩擦音です。あるいは、舌や唇を震わせて発音する言語もあります。

もう一つは、有声か無声かということです。有声か無声かの違いは先ほどもお話ししました。「パ」と「バ」、「カ」と「ガ」、「ファ」と「ヴァ」、「ス」と「ズ」、あるいは「シ」に声帯の振動を加えると「ジ」になります。これも細かくご説明していると、きりが無いのですが、子音の違いは三つのポイントを押さえれば決まってくるということです。

③母音の分類

では母音はどうかというと、母音の違いには舌の動きが重要です。まず「イ」「エ」「ア」と発音してみてください。そうすると、だんだん下あごが下がってきます。あるいは、口に鉛筆を挟んであごを固定したままでも「イ」「エ」「ア」と発音できます。そのときに鏡でご覧になると、「イ」の場合には舌の前の方がずっと高くなります。「エ」は少し下がって、「ア」ですと舌がほぼ平らになります。「イ」「エ」「ア」の順で、舌の位置が高いところから低いところへ下がります。もう一つ、「ウ」「オ」という音の場合は、舌の後ろが高まります。つまり、舌の前の方が高くなる「イ」「エ」「ア」という母音と、舌の後ろの方が高くなる「ウ」「オ」という母音があるのです。舌のどの部分が高まるかによって、さらにその高まりの程度によって違ってくるといわけです。

もう一つは、唇の丸めがあるかないかということです。「イ」という発音のときには唇が平らです。それに対して「オ」ですと、唇がやや丸まった感じになります。この唇が丸まっているか、丸まっていないかという点も、母音を区別する非常に大事な違いになります。例えば「イー」と発音していて、これに唇に丸みをつけると「ユー」というような音になります。例えばドイツ語やフランス語を勉強すると、このような母音が出てきます。

子音については約 600 種類、母音については約 200 種類の音が使われているだろうという報告がありますが、基本的にそれぞれの母音・子音は三つずつの特徴によって分類がなされています。それにもう少し細かい特徴が加わると、さらにいろいろな種類の母音が作り出されるというわけです。

Ⅲ. 音声の記号化

1. 音声記号は怖くない

皆さんは中学校に入ったときに、英語の発音記号を習ったと思います。試験にも、「この単語の発音を発音記号で書け」といった問題があつて、嫌な思いをした方もいらっしゃるかもしれません。あるいは得意だった方もいらっしゃるかもしれません。しかし、発音記号というのは怖くも何ともありません。つまり、発音記号というのはあくまでも記号にし過ぎないということなのです。

いろいろな言語がいろいろな音を持っているのですが、これがどういう音であるかということ、それぞれの個別の言語に関係なく共通に表すシステムがあると便利だということとで考え出されたのが国際音声記号です。これは国際音声学会（国際音声学協会）という

組織で決めたもの、中学校の発音記号の親玉のようなものだと思います。資料は「1993年改訂、1996年修正版」とありますが、2005年になって新しい記号が追加されましたので、それも書いておきました。

IPA (国際音声記号) 1993年改訂・1996年修正版

子音 (肺臓気流)

	両唇音	唇歯音	歯音	歯茎音	後部歯茎音	そり舌音	硬口蓋音	軟口蓋音	口蓋垂音	咽頭音	声門音
破裂音	p b			t d		ʈ ɖ	c ɟ	k ɡ	q ɢ		ʔ
鼻音	m	ɱ		n		ɳ	ɲ	ŋ	ɴ		
ふるえ音	ʙ			ɾ					ʀ		
はじき音				ɽ		ɽ					
摩擦音	ɸ β	f v	θ ð	s z	ʃ ʒ	ʂ ʐ	ç ʝ	x ɣ	χ ʁ	ħ ʕ	h ɦ
側面摩擦音				ɬ ɮ							
接近音		ʋ		ɹ		ɻ	ɹ	ɻ			
側面接近音				ɻ		ɻ	ɻ	ɻ			

記号が2つ並んでいるものは、右が有声音、左が無声音 網かけは調音が不可能と考えられる部分

2005追加

子音 (肺臓気流以外)

吸着音	有聲入破音	放出音
⊙ 両唇	ɸ 両唇	ʔ 例:
歯 (後部) 歯茎	ɖ 歯 (茎)	p' 両唇
! 硬口蓋歯茎	ɟ 硬口蓋	t' 歯 (茎)
≡ 歯茎側面	ɟ 軟口蓋	k' 軟口蓋
	ɢ 口蓋垂	s' 歯茎摩擦

母音

記号が2つ並んでいるものは、右が円唇、左が非円唇

まずは、子音の表を見てみましょう。子音を決めるのは、一つは空気の流れがどこで妨害されるかです(横軸)。この表では、前の方(唇)からのどの方まで順番に並んでいます。もう一つは妨害の仕方です(縦軸)。一遍空気の流れを断ち切ってしまうのか、それとも空気の通り道を狭くするのか、あるいは震わせるのかといったことです。それと、例えば、枠の中に二つ記号が入っている場合がありますが、その場合には左側が無声音、右側が有声音という仕組みになっています。つまり子音が三つの基準で分類され、それを一つの記号で表したものが国際音声記号という発音記号だということです。

[p]という記号が何を示しているかという、妨害の仕方からすると破裂音です。破裂音というのは空気の流れを一遍止めて、それを開いたときに出る、[p][b][t][d]という音です。妨害される場所でいうと、両唇音です。両方の唇で空気の流れを一遍遮断して、開くときに破裂的に出る音ということです。これには[p]と[b]の2種類あって、そのうちの[p]は左側にありますから無声音です。このような中身を一つの記号で表しているのです。

あるいは[m]という記号は、妨害の仕方という鼻音になります。つまり鼻に抜ける音です。妨害される場所でいうと、[p]と同じ両唇音です。それから[m]というときに

は声帯が震えていますから、有声音だということです。

もう一つ、[p]の右側に[b]と書いてありますが、これは破裂音で両唇音で有声音です。このように、子音についていえば、音を作る場所、音の作り方（気流の妨害の仕方）、声帯振動の有無の三つの特徴を一つの文字で表したものが音声記号です。

さらに、右下に母音の表もあります。母音の場合は横軸が舌の前後を示し、縦軸は舌の高まりの具合を示しています。「i」という記号で書いてある音は、舌の前の方が一番高くなって出される音です。それから、[a]の音は舌が一番低く平らな「アー」という音です。このようにして見ればいいということです。細かい点はいろいろありますが、発音記号というのは意味のない記号ではなくて、このような中身を理解すれば発音記号は怖いものではないということです。それをお分かりいただければと思います。

2. 音の連続

今までは一つ一つの音のお話をしてきましたが、実際の発音の場合には、音というのは一つ一つ節目のようにばらばらに発音されるのではなくて、音節を作っています。音節というのは基本的に母音を中心にした音の集まりです。この音節が幾つか集まって、それが意味と結び付いて一つの単語を作っています。

ですから、例えば「あめ」という単語がありますが、これは[a][m][e]という三つの音から成り立っていますが、[a][me]が一つ一つの音節で、[ame]が音節のまとまりです。これだけではなくて、実際の日本語の発音ではアクセントがかかっています。例えば、東京方言ですと、空から降ってくる「雨」の場合は[a]が高くなっていて[me]が低いです。それに対して食べる「飴」の場合には[me]の方が高くなっています。これがアクセントです。

ですから、実際の発音においては、まず一つ一つの素材になる音がくっついて音節を作り、さらに音節にアクセントが加わって、そしてまた文になるとときにはイントネーションが加わります。そのような形で連続して発音されて、実際の発音になって出てきます。今日は音声の連続についてのお話はいたしません、一つ一つの音の上にかぶさる特徴があります。こういった特徴があるために、単純に逆転再生しても元のようにきっちりとした形では聞こえないわけです。

IV. 目で見る音声

1. 音声言語は音波である

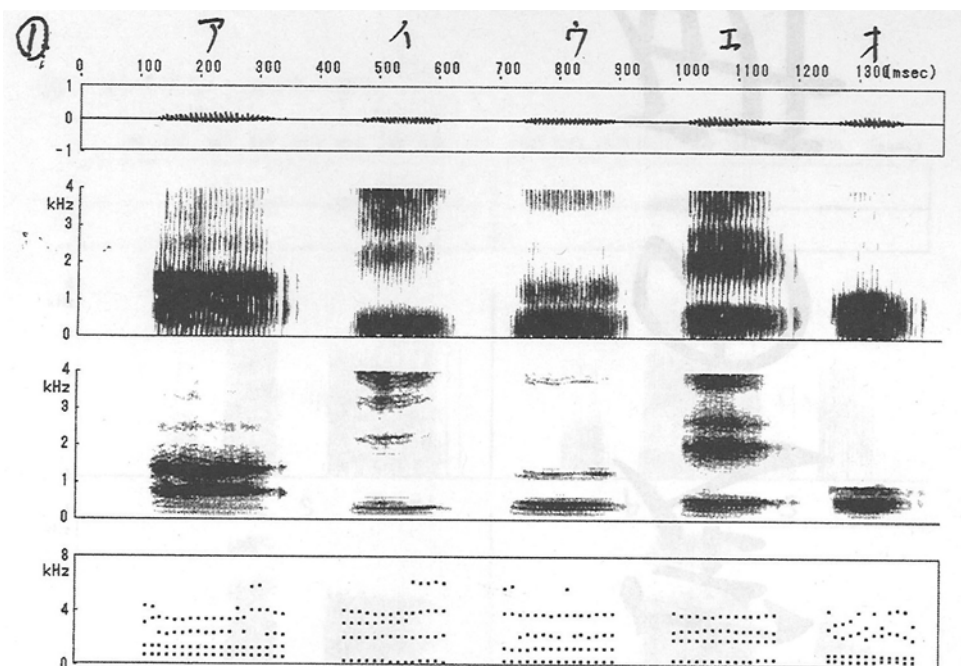
先ほど音声を実際にどのように発音して、それがどのように分類されるか簡単にお話ししましたが、今度は目で見てみましょう。具体的には、客観的に機械で音声を分析・観察してみるわけです。音声というのは、肺から上がってきた空気が声帯を震わす、あるいは狭いところを通ってくるによって、音波の形で出てきます。言語音声は音波であるということです。その声帯の振動でいえば、声帯が振動するたびに空気の流れが途切れます。そして、それが波の形で伝わっていきます。音声というのは、いろいろな成分から成り立っているのです、それを成分分析して見てみようということです。

2. 音響分析

まず、人間が聞き取れる音がどのぐらいの範囲かという、20Hz から 20kHz です。ですから、1秒間に20～2万回の振動まで聞き取ることができるわけです。声帯の振動数（これを基本振動数といいます）は、大体100～300Hzです。男が一番低くて、女の人が高くなり、子どもになるとさらに高くなります。これは声帯の長さの違いなどによります。

音声分析をするには、かつては非常に大がかりな機械が必要でした。例えば私たちの研究室でも、音響分析の機械を十数年前に入れたのですが、そのときにはまだ数百万円して、十数年予算申請をしてやっと通って非常にうれしかったのですが、そういう高額な機械でなければこのようなことができませんでした。しかし、今はパソコンで簡単に分析ができるようになりました。市販のソフトもありますし、フリーソフトもありますので、どなたでも音響分析を試してみることができるようになりました。今日は幾つかのサンプルを見ながら、実際にどのような音響的な特徴が見られるかということをお話したいと思いません。

①ア, イ, ウ, エ, オ

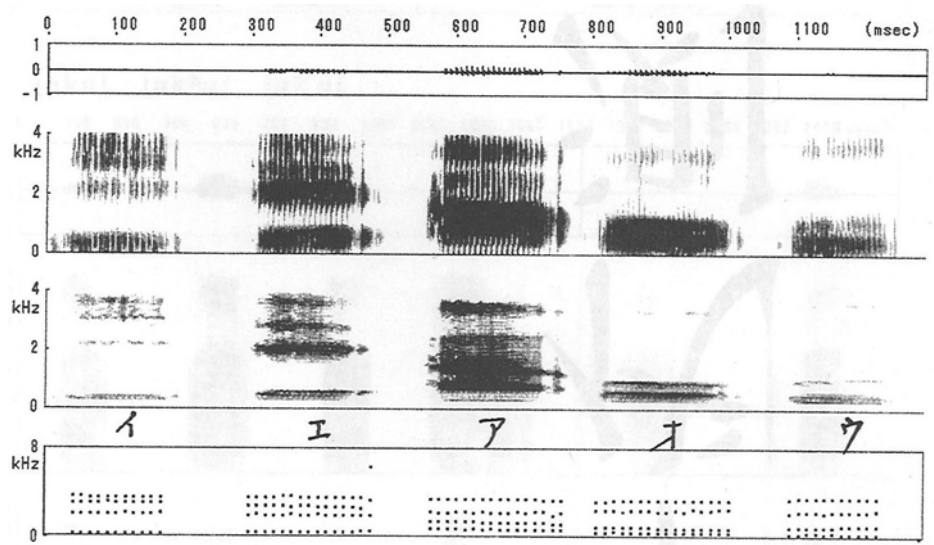


最初に母音から見ていきたいと思います。まず、はっきりと区切って「ア」「イ」「ウ」「エ」「オ」と発音した場合が①です。1段目にあるのが音声波形で、2段目と3段目がスペクトログラムというものです。母音というのは基本的な周波数の音に、その整数倍の周波数を持った多くの倍音が組み合わさってできている音ですが、このようにある音に含まれる周波数成分を分析して示したものがスペクトログラムです。

①の2段目が広帯域スペクトログラムというもので、横が時間軸で、縦じまが1回1回の振動です。中で特に色が濃いところと薄いところがありますが、色が黒いところはその周波成分が強いということを示しています。この母音を構成する周波数成分の中で強いものをフォルマントといいます。下から数えて一番目が第1フォルマント、第2フォルマントという呼び方をします。例えば「イ」の場合ですと、周波数成分の濃い部分が第1フォルマントですが、第2フォルマントは非常に高いところにあります。それから、「ア」や「イ」と比べて「ウ」も第1フォルマントと第2フォルマントの在り方が違っていることがお分かりになるとと思います。

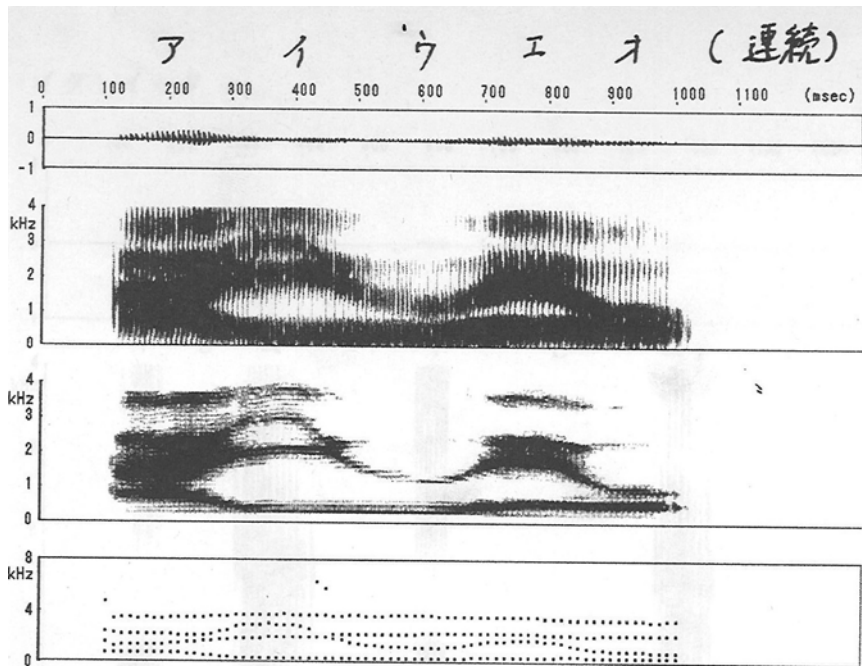
このように母音の違いというのは、音響的にはどの周波数成分が強いかという違いによって区別されるということがお分かりになるとと思います。

②イ, エ, ア, オ, ウ



今度はこれを「イ」「エ」「ア」「オ」「ウ」という順番に並び替えてみます。これを見ると、第1フォルマントは山のような形になっていますが、これは「イ」「エ」と前の方で発音される母音から、「オ」「ウ」のように後ろの方で発音される母音の違いに対応しています。第2フォルマントはだんだん右肩下がりになっています。これは舌の高さの違いに対応しているということが分かります。

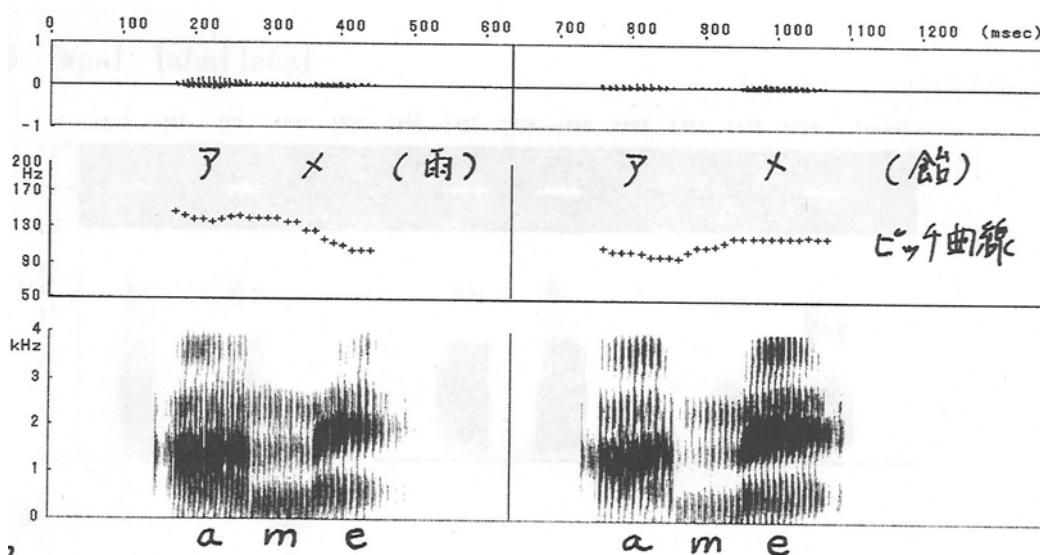
④アイウエオ (連続)



今までは独立して発音していましたが、続けて発音したものが④です。そうすると、全体としてもわっとした感じになっています。最初の部分を聞いてみますと「ア」と聞こえます。もう少し右は「イ」です。その中間は、「ア」でもないし「イ」でもないという音です。ちょうど真ん中はほぼ「ウ」に聞こえます。次が「エ」、その次が「オ」です。

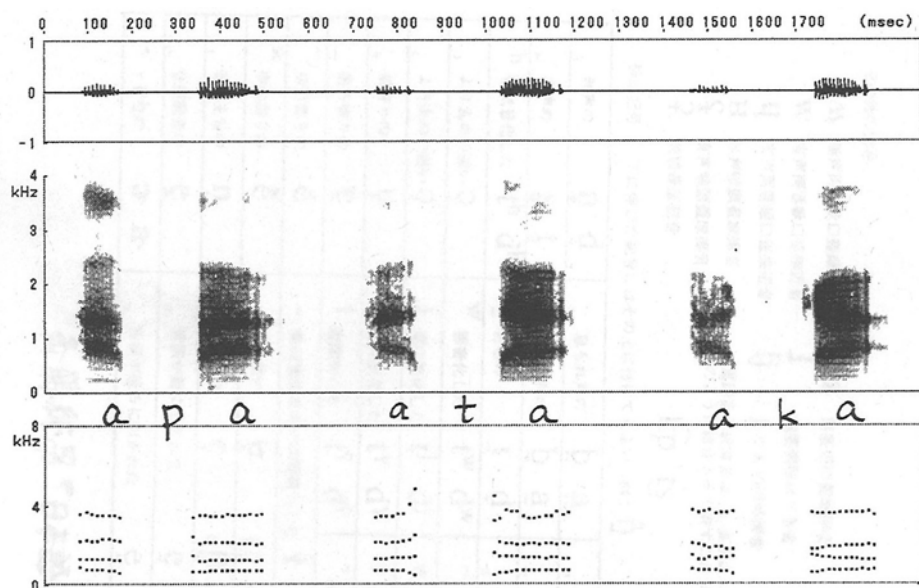
このように、コアになる部分は聞いていただいたとおりなのですが、先ほどから申しましたように、実際の音声というのは連続していて、ある音からある音に移り変わっていく部分があるわけです。ですから「ウ」と「エ」の間の部分を聞いてみると、「ウエ」というような感じになっています。ですから、実際にははっきりとしたフォルマントが恒常的に出てくるのではなく、移り変わりの部分がどうしても出てきます。

⑤雨と飴



次に、先ほどの「雨」と「飴」をスペクトログラムで見えます。波形は見てあまり違いが分かりませんが、2段目の図でピッチ（高さ）の変移を見てみますと、左側が「雨」で、右側が「飴」です。飴の方は、相対的に「メ」が「ア」に比べて高くなっていることがお分かりになると思います。

⑥アパ, アタ, アカ

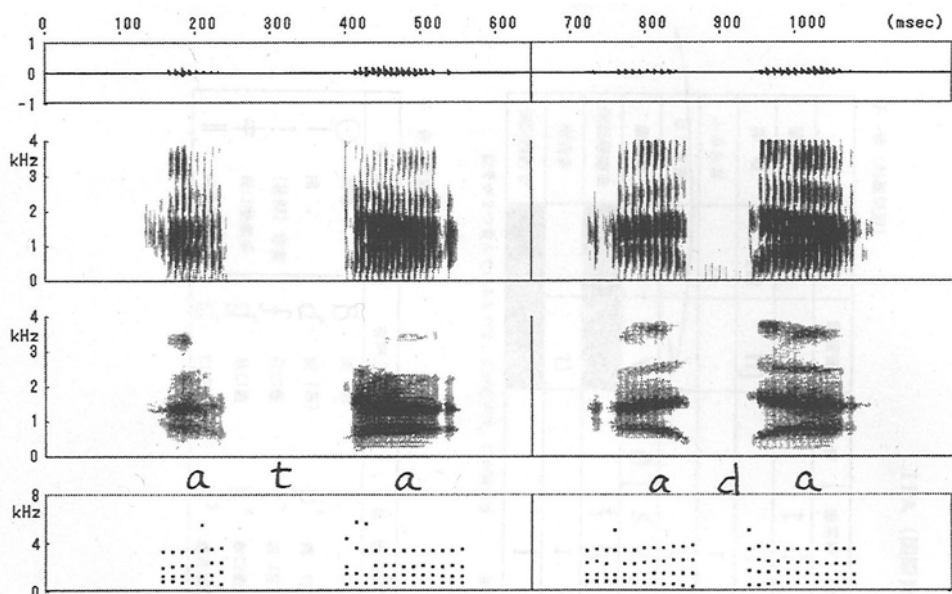


次は、「アパ」「アタ」「アカ」を発音しています。黒く出ているところは [a] で、真ん中の部分は何も出ていません。その部分を聞いてみても何も言っていません。つまり、これで見ると、[p] も [t] も [k] も、実際に音は出ていないのです。従って私たちは、音が出ていない部分から [a] に移るところで [p] という音を聞いているのです。

では、どうやって [p] か [t] か [k] かを聞き分けているのでしょうか。先ほどのフォルマント曲線を見てみます。一番下がフォルマントの移り変わりを示したもので、実際には音は出ていないのですが、子音から母音に移っていくときの移り方が違うのです。例えば「アタ」というときには、第1フォルマントは最初は少し低くなって、そして普通の高さになっています。それに対して第2フォルマントは少し高いところから入って安定しています。このような違いが、私たちの実際の聞き分けにかかわっているのではないかと考えられています。

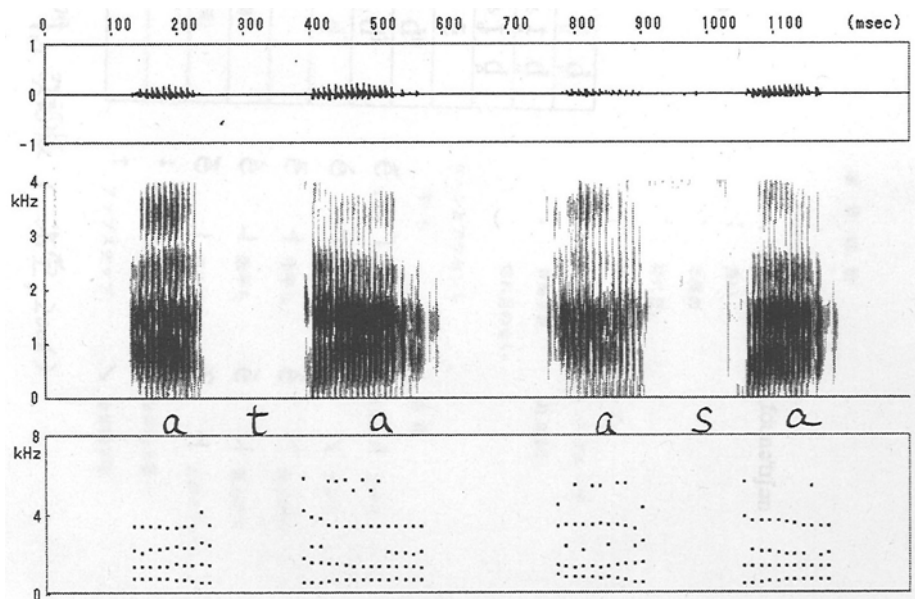
ですから、[p] の中核部分は何にも音がありません。実際に聞いているのは、あくまでも [p] の音から [a] の音、あるいは [t] の音から [a] の音、[k] の音から [a] の音に移る部分で初めて音を聞き取っているということが分かるわけです。

⑦アタ, アダ



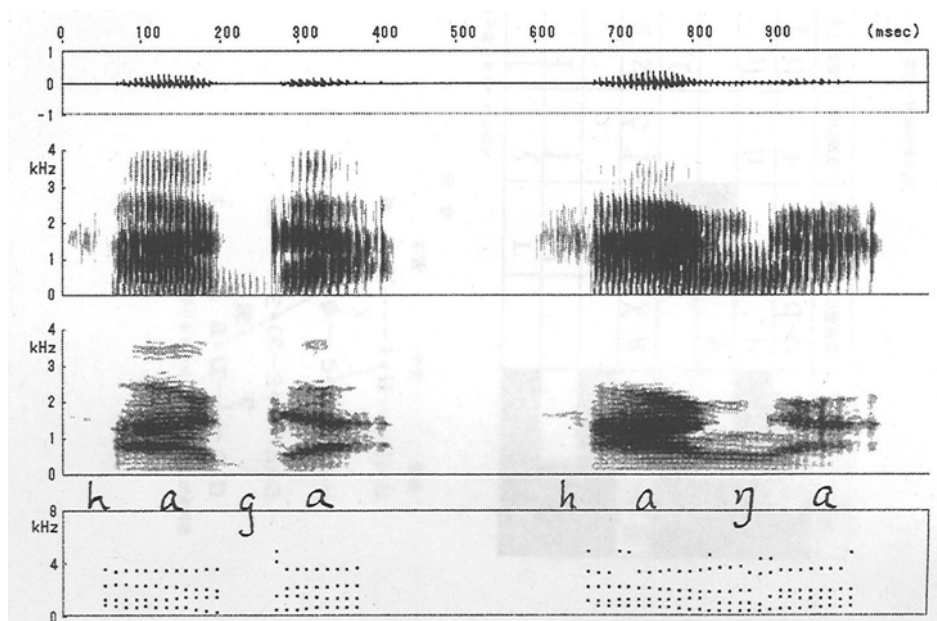
今度は「アタ」と「アダ」を見ていただきます。左側が「アタ」で、右側が「アダ」です。どこが違うかという点、若干長さも違うのですが、それより大事な点は、空白の部分が「ダ」の方は非常に低い部分に少し縞があります。つまり、この部分があるということは、声帯が振動していることを示しています。「タ」の場合は [t] が声帯の振動を伴わないので、何も音が出ていないこととなります。[t] と [d] というのは同じところで同じように出されるのですが、[t] は声帯が震えない、[d] は声帯が震えるという点が、音響的にも反映されているということです。

⑧アタ, アサ



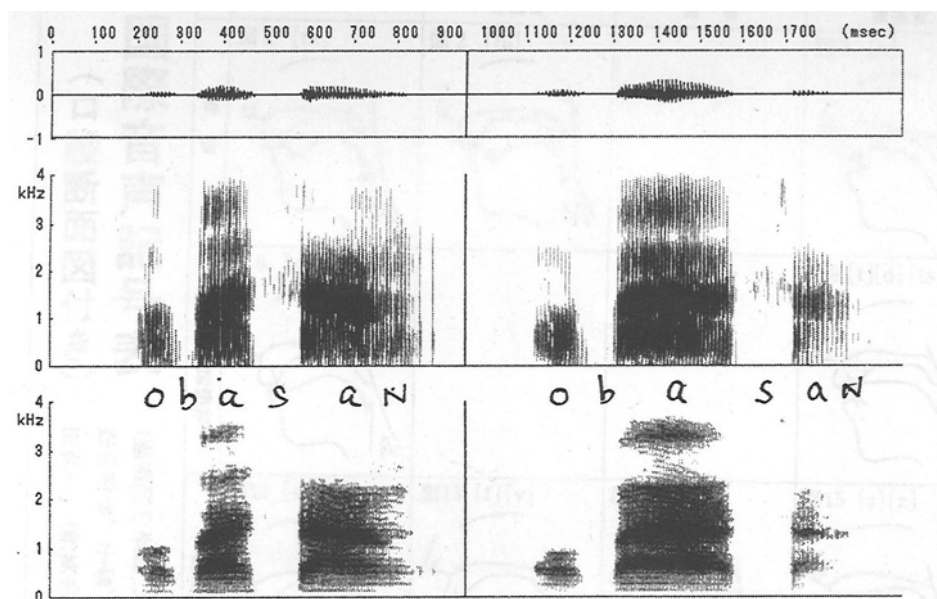
今度は「アタ」と「アサ」です。[s]という音は[t]同様声帯が震えません。この[s]という音は、舌の前の方と口の先の間にすき間ができて、そこを空気が非常に強いスピードで通り抜けるときに出る音です。このように音響的に分析しますと、比較的高いところに周波数成分が見られます。これが摩擦音といわれる[s][z][f][v]という音の特徴です。[t]の音は何も聞こえませんが、[s]の部分は空気の音が聞こえます。

⑨ハガ, ハガ (鼻濁音)



次に、いわゆる「が行」の鼻濁音があります。語中の「ガ」「ギ」「グ」「ゲ」「ゴ」という音です。金沢方言も、もともと鼻濁音のある方言ですが、若い世代ではだんだん失われつつある音です。これを音響的にスペクトログラムで見ると、最初の部分は [h] の音で、次が [a]，次が [g]，次が [a] です。一方で「ハガ（鼻濁音）」の方は、[g] が [ŋ] という音になっています。これは、要するに鼻に抜ける音なのですが、鼻に抜ける音は前後の母音の特徴が反映されたような形で見えてきます。

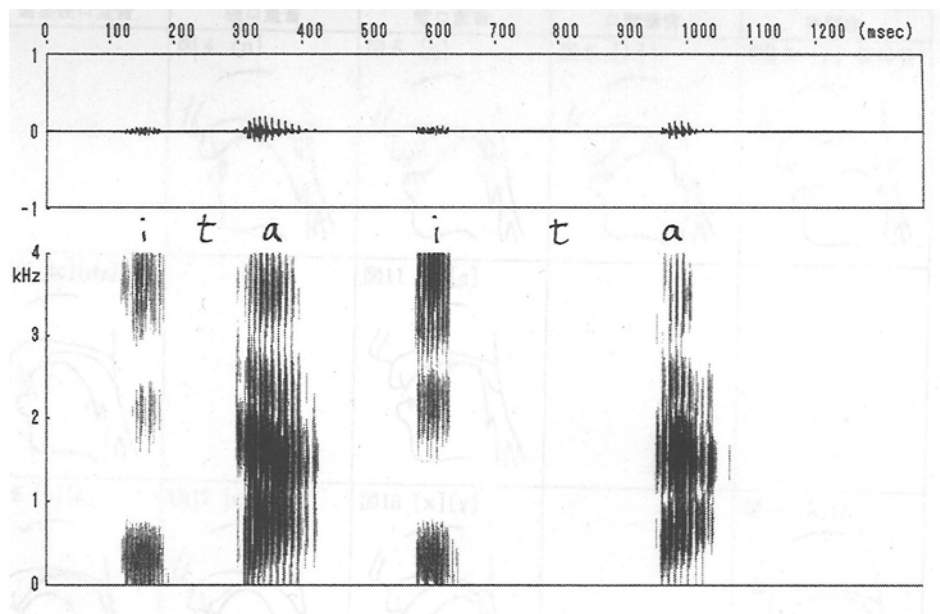
⑩オバサン， オバーサン



日本語は、母音の長短がコトバの意味を区別する非常に重要な特徴としてあります。例えば「オバサン」と「オバーサン」で、実際にどのような具合か見ていただきます。最初が「オバサン」の [o]，次の空白の部分が [b]，次が「バ」の [a]，次が [s]，そして [a] [n] になります。[a] と [n] の間にははっきりしません。右側が「オバーサン」の [o] [b] [a] です。

ここで [b] と [s] の間を測ってみますと、「オバサン」の [a] の場合には非常にラフに測って約 138m秒，それに対して「オバーサン」の場合は 288m秒で，約 2 倍です。日本語で書くと「オバサン」「オバーサン」と「バー」と二つ数えますが，ほぼそれに対応しています。

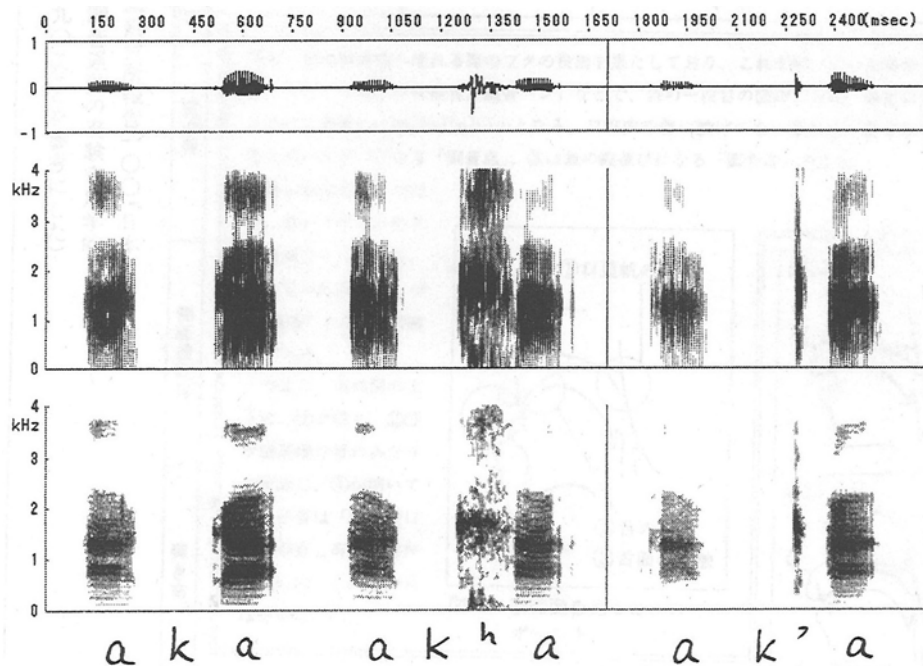
⑪ イタ, イッタ



先ほどの長い・短いの違いは子音にもあります。例えば「イタ」と「イッタ」で長さを比べてみますと、「イッタ」の方がかなり長くなっています。先ほどの2倍どころではなくて、「イタ」の方は129m秒ですが、「イッタ」の場合ですと311m秒で2.5倍以上の長さを取っています。私たちは俳句を作ったりするとき、「イタ」は2音、「イッタ」は3音と数えますが、実際に測ってみると2倍どころではなくてかなり長い秒数がかかっています。

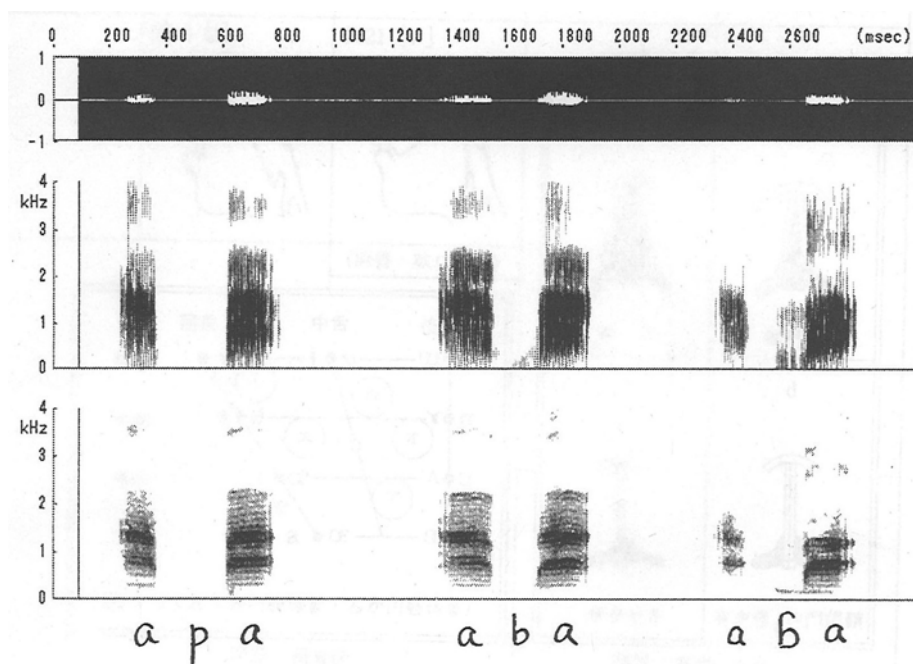
これを実験的に、2倍だったらどうなのか、1.8倍だったらどうなのか、機械で自由に空白部分を操作できますので、それを聞いてもらって、大体どの辺までの長さだったならば「イタ」と聞いて、どこまでだったら「イッタ」と聞けるのかという実験も簡単にすることができます。

⑫ [aka] [ak^ha] [ak'a]



残りの耳慣れない発音を聞いていただこうと思います。[aka] は日本語の「赤い」などの普通の「アカ」です。その次は [ak^ha] です。息が出る発音、帯気音といますが、中国語や韓国語などでは息が出る音か出ない音かというのは非常に重要な違いとしてあります。そういった場合には、[k] があって次に息の部分である [h] が続いて [a] という感じになります。最後の発音はあまり耳慣れないと思うのですが、私がやっている言語などに見られる [ak'a] というものです。これは [k] で一遍破裂するのですが、無音の部分があって [a] が出てくるという発音です。

⑬ [apa] [aba] [aba]



最後に有声音の同じような発音を聞いてみましょう。[apa] [aba] [aba] です。[b] というのは入破音（にゅうはおん）という、飲み込むような音です。スペクトログラムに特徴的な模様があるので、普通の [b] とは違うと区別されます。

以上、実際に音声を特にスペクトル分析、音声を周波数成分分析してみると、ということがそこから言えるかということをお話ししました。また、そこから今申しましたように、いろいろな機械的な操作ができます。例えばアクセントにしてもピッチが出てきましたが、それを機械的に変えることが今は簡単にできますので、それをいろいろ変えていくことによって大体どこからどこまでの高低の違いだったら「雨」と聞いて、どのぐらいまで上がれば「飴」と聞かれるかということも簡単に実験することもできるようになっています。市販のソフト以外にも、簡単にただで手に入るソフトもありますので、皆さんも一度マイクとコンピューターさえあればできるのでやってみて、いろいろな音声を目で見て楽しんでいただければと思います。

今日はこのお話の機会を与えていただきまして、日ごろあまり振り返ることのない音声について、実際に人間が自分の体の部分をどのように使って出しているのか、また音響的に見るとどうかということをお話しさせていただきました。

質疑応答

(質問者1) 初めてこのようなお話を聞いて、びっくりしました。音声研究が人間社会やロボットなどと、具体的にどのようにかかわっているのか教えてください。

(柘植) 例えば、私たちはATMで操作するときなど、いろいろなところで人工的な音声などを聞いています。そのときにも人間の出す音声を音響的に分析して、その中でどういう特徴が大事かということを取り出して、それを新しく合成して作り出し、人間がしゃべらなくても済むようにしています。

(質問者1) では、お医者さんが人の声を治すこととも関係があるのですか。

(柘植) もちろん、お医者さんの世界でもこのような知識が基礎にあって、その上でいろいろな症状に対応して診断されたりしていると思います。非常にいろいろな広がりがあります。例えば手術で喉頭を摘出された方がそれに替わる音声をどう出すのかという点などにもかかわってくるものです。

今日のお話は先端の研究でも何でもなくて、非常に基礎的な研究で、こういった研究を基にして、いろいろな新しい研究がどんどん進んでいる状態です。私が知っているのは言語学という狭い分野ですが、かなり研究が進んでいる分野だと言えます。

(質問者2) 音が出ないのは [p] [b] [t] [d] だけですか。

(柘植) いえ、もっとあります。音声記号の表で見てくださいと、あまりご覧になつたことのない記号があると思いますが、[p] [b] [t] [d] のほかにも、一番上の段の音声がそういう仕組みで出される音声です。

(質問者2) 要するに破裂音ですか。

(柘植) そうです。その構えをしている間は全然何も音が出ていないので、出る瞬間に次の音に移っていくときにそれを私たちは聞いているということです。