

なぜ「射[↑]た？」は「居[↓]た？」に化けないか

馬場良二

About the combination of pitch accent and intonation on the end of sentence in Japanese

Baba Ryoji

0. はじめに

ことばを発するということは、音の高さや強さリズムといった調子を発するということである。そして、日本語ではどんな語であっても、それを声に出して言おうとすると、必ず高い低い調子を伴って発音される。例えば、東京方言の場合「射[↑]た」だったら「射」は高く、「た」は低く、そして、「居[↓]た」だったら「居」は低く、「た」は高く発音され、その高低を記号で示すなら、「射[↑]た」「居[↓]た」のようになる。

ひとつかあるいはいくつかの語が集まってできている文を言おうとすると、イントネーションと呼ばれる声の調子が必要となる。問いかけだったら、文末で上昇するイントネーション、それに対する断定の答えだったら、下降のイントネーションが代表的である。ここでは、語の音調のアクセントも文の音調のイントネーションも言語音の高低、つまり、ピッチととらえている。そして、実際の発話でアクセントとイントネーションが重なると、「射[↑]た？[↑]」「射[↑]た。[↓]」、「居[↓]た？[↓]」「居[↓]た。[↓]」と発音される。が、高い低いアクセントに低い高い文末イントネーションがかぶさっても「射[↑]た？[↑]」は「居[↓]た？[↓]」にも「居[↓]た。[↓]」にもならない。それはどうしてか、という疑問がこの研究の発端である。ここでは、同じピッチであるアクセントとイントネーション、それぞれの本質を実験音声学的に解明することを目的としている。

1. 分析対象

さて、私は外国人に日本語を教えることを職業としている。その立場からすると考察する対象というのは自然な言語でなければならない。問いかけとそれに対する返事という形式をとったのは、たとえば、天気を気にしながら

「あゝめ?」「あゝめ。」と言ったり、話題になっている新刊書がどれか知りたくて「あゝれ?」と問いかけ、「あゝれ。」と返事をするといったぐあいに、この形式が日常生活の会話で実際によく使われる自然な形式だからである。また、「射ゝた」や「居ゝた」といった2拍語を取り上げたのも、1語でできた文を取り上げたのも事象をできるだけ単純化してとらえたかったからである。2拍語のうち「射ゝた」と「居ゝた」を取り上げたのは、今後、「煎ゝった、行ゝった」のような促音、「イゝゝだ、胃ゝだ」のような有声音、「イゝー(E)だ、いゝいだ(飯田)」のような長い母音、これらと対比させてみたいからである。

「射た」「居た」を International Phonetic Alphabet (IPA) で表記すると [ita] となり母音ふたつの間に無声の閉鎖子音 [t] が入ることになる。無声子音には声はないわけで、必然的にピッチも存在しない。このような音声はピッチの分析がしにくいので、まず、有聲の子音しか含まない語で、その音韻的構造の単純な「もゝも(腿)」と「もゝも(桃)」を取り上げた。

2. 実験方法

話し手は東京語話者で、「もゝも(腿)?ゝ」「もゝも(腿)。ゝ」、「もゝも(桃)?ゝ」「もゝも(桃)。ゝ」という問いかけ、返事のやりとりをひとりでした。それをアメリカ KAY 社のソナグラフ DSP5500 で分析し、そのデータをパーソナルコンピュータに送り、同じく KAY 社製の音声分析ソフト LPC Parameter Manipulation & Synthesis Program でピッチ抽出を行った。その数値をある関数にあてはめて、その関数でのパラメータを割り出した。

3. ソナグラフ DSP5500による分析

図1はソナグラフ DSP5500による「もゝも(腿)?ゝ」「もゝも(腿)。ゝ」の分析結果、図2は同様に「もゝも(桃)?ゝ」「もゝも(桃)。ゝ」の結果である。

両図の上の段はピッチ(pitch)、波形(wave form)、強さ(amplitude)、ゼロクロッシング(zero crossing)のよっつの分析結果が表されている。ピッチは、声帯の振動数で「声の高さ」を示す。グラフでは階段状の短い横棒で現われている。波形はひょうたん形の図形で、横軸に時間を、縦軸に振幅を示す。横軸ひと目盛りが100ms、100/1000秒である。強さは盛り上がった曲線で、その音声の物理的な強さである。ゼロクロッシングは基線に

ソナグラフ DSP5500による分析結果 1

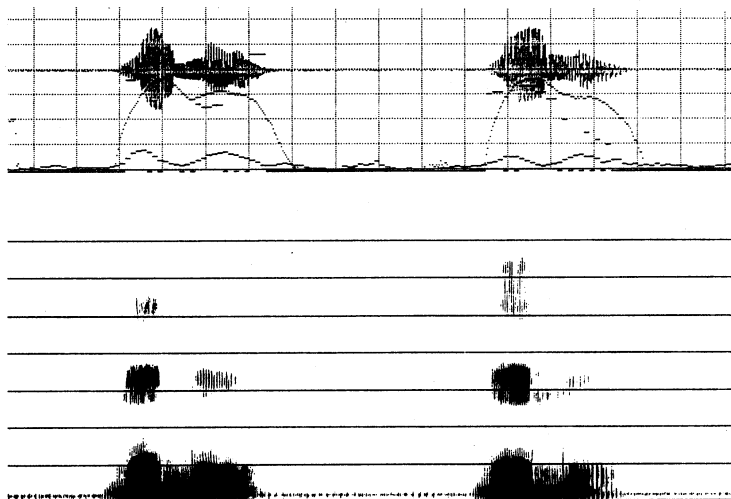


図1 問いかけ「もつも(腿)？」 返事「もつも(腿)。」

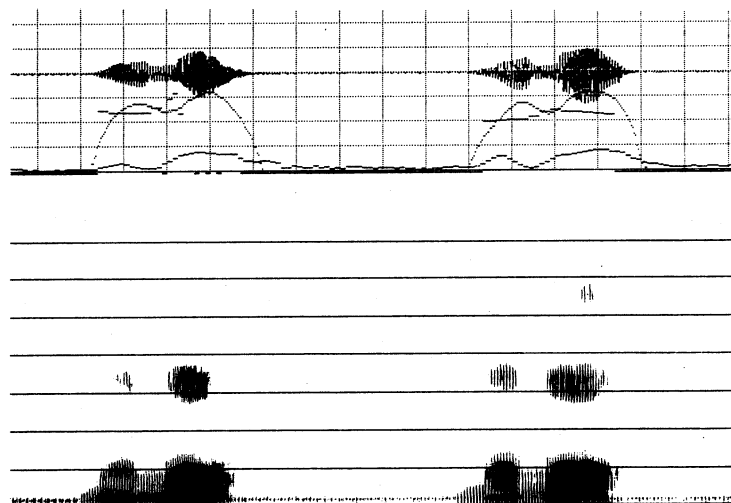


図2 問いかけ「もつも(桃)？」 返事「もつも(桃)。」

近い曲線で、単位時間に何度波形が基線を通じたかを示しており、[s] [ʃ] といった高い周波数の音を見るのに便利なものである。下の段はソナグラフである。横軸が時間、縦軸が周波数で、強さを濃淡で示している。濃いほどその時点のその周波数成分が強いことを示す。横軸の時間は上段のグラフと同じで、縦軸はひと目盛りが 1000Hz となっている。

「ソナグラフ DSP5500による分析結果1」の図1、2の上段に見える波形で、ひょうたんがよつつあることになるが、すべて語音は [momo] で同じであり、同じであるからよつつともへちまとかトマトとかではなくひょうたんとなっている。

「ソナグラフ DSP5500による分析結果2」の図3、4の上段に見える図形はタワシと三角形となっている。「居「た。ゝ」だけは三角形でなくつばつきの壺のような形となっている。いずれにしる「ソナグラフ DSP5500による分析結果1」に見られるよつつの図形とくらべれば、これらはすべて似ている。こちらの語音は [ita] である。すでに述べたように [t] は無声子音でその閉鎖部では声帯が振動しない、つまり、声がない。そこで、波形のふれもなくなり、[ita] 全体が [i] の部分と [a] の部分とのふたつにわかれるのである。一方、[momo] の [m] は [t] と同様子音ではあっても、有声の鼻音であり、よって、波形も現れる。だから、「ソナグラフ DSP5500による分析結果1」のよつつの図形はふたつにわかれることなく、つながってひょうたんとなるのである。

無声部分はすべて [t] なのかということ、そうとも言えない。すべての無声部分が [t] であるなら、同じ無声閉鎖子音を含む [ika] [ipa] がすべて識別できないことになる。[ita] [ika] [ipa] の聞き分けは、子音の閉鎖部へわたる部分の [i]、また、子音の閉鎖部からわたる部分の [a] それぞれでなされる。何の音も出ていないところでその音を識別するなどはできるはずもなく、音の出ているところで識別するのである。図3、4の下段のソナグラフを見ると [i] [a] の母音部に横縞ができていのがわかる。特定の母音はある特定の周波数の音が強く、その組み合わせによってそれぞれの母音の音質が形成されるからである。この横縞のことをフォルマントと言い、下から第1、第2、第3フォルマントと数えていく。母音の音質の特定のためには第1と第2フォルマントを見れば充分であると言われている。[i] の第1フォルマントは500Hzより下、第2フォルマントは2000Hzより上に位置している。一方、[a] の第1フォルマントは1000Hz近くで、第2フォルマントは2000Hzより下のようである。その [a] のフォルマントを見ると、

ソナグラフ DSP5500による分析結果 2

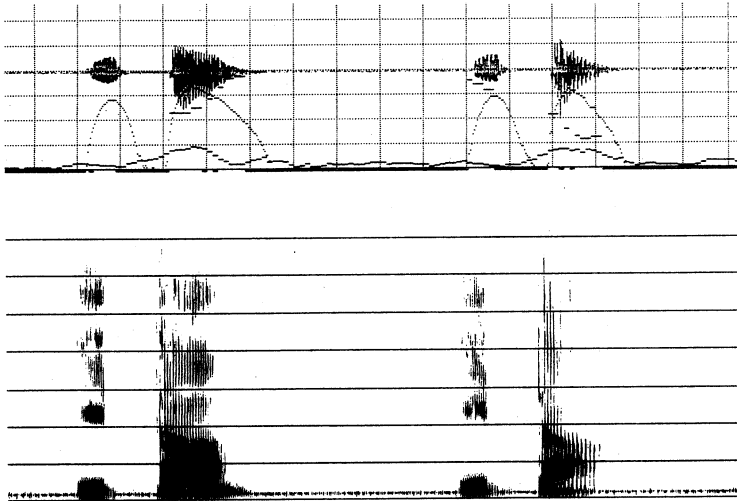


図3 問いかけ「射つた？」 返事「居つた。」

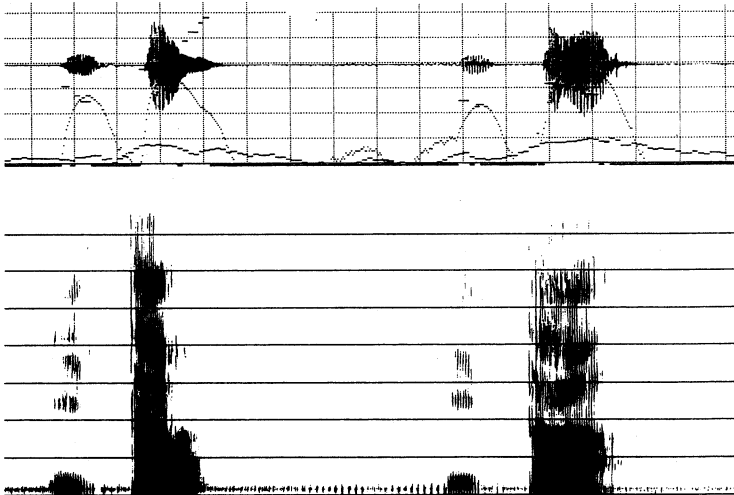


図4 問いかけ「居つた？」 返事「居つた。」

子音の閉鎖部からわたりの部分では第1、2フォルマントの横縞が開いており、数10msで閉じるかっこうになっている。この、開いている形から閉じる形へ移るところの [a] のフォルマントの部分が [t] を知覚させていると考えられる。均質な [a] であれば各フォルマントは一定の周波数のまま変化しないのであり、同じ [a] でありながらもこの部分は [t] を知覚させる [a] なのである。もしくは、[a] を知覚させる [t] なのである、と言ってもいいかもしれない。横軸ひと目盛りが100msだから、この [t] の無音部も大体100ms、1/10秒ということになる。

図1のひょうたんは頭でっかちだが、図2のひょうたんはどちらかというところと下ぶくれである。これは、高い拍では強さも増す傾向があるため、アクセントが高低の「腿」ではひとつ目の「も」が強く発せられ、アクセントが低高の「桃」ではふたつ目の「も」の方が強く発せられるからである。

図2の左、下から3マス目の中ほどから横に伸び、隣りのマスの端のあたりからふたマスぶち抜きでぐんぐん上へ伸びている階段状の短線がピッチである。右のピッチは下からふた目盛り目から始まってたらたらと横に伸びている。ピッチの始まりに高さの違いがあるのは、左の問いかけの「桃?」の方が発声の出だしが少しばかり高く、一方の返事の「桃。」は返事だけあっていくぶん断定的におさえた高さで始まっているからである。「桃」のアクセントは低高で「も」であり、そこに問いかけの上昇イントネーションが続くとピッチは一方向的に上がっていく。図2左の問いかけの「も」ではそのアクセントとピッチがダブった形で上昇が急である。返事の「も」の方には上昇イントネーションはないものの、低高のアクセントはある。だから、ピッチにも低から高への変化が現れていいはずであるが、ほぼ平坦になっている。音韻論的にははっきり認められる第1拍低から第2拍高へのアクセントの変化も、実際の発音ではあまりはっきりなされないことが多く、また、返事、断定の口調ではアクセントとイントネーションとのふたつの音調が相殺しあった形で、よけいにそのアクセント、低から高へのピッチの上昇はおさえられるのであろう。

図1ではピッチの始まりでの高さの違いが大きい。そして、問いかけの方の「もも(腿)?」では、ピッチが一度さがり、それから文末にむかって上昇していることがわかる。一度下がるのは「もも」の高低アクセントの音声的な実現で、文末にむかって上昇しているのは問いかけを表すイントネーションの実現である。一方、返事の方の「もも(腿)。」では、立ちあがったあとピッチは下降するだけである。

図3の「射つた？」には、図1ではっきり見られたピッチの下降はなく、[t]の無音部のあと[a]の始まりから見られる上昇だけである。ピッチが下降していくところは音響学的には存在しないわけだが、[i]を聞き取ったあと無音部をはさんでガクンと下がった点からのピッチを聞けば、ピッチが下降したんだということはわかり、我々は[i]は高く、[a]は低いと知覚する。「射つた。」では、アクセントの高から低への変化、そして、文末の下降イントネーションが続いていることがよくわかる。図4でも、「居つた？」における上昇、「居つた。」におけるピッチの平らな状態は「ソナグラフDSP5500による分析結果1」で見られるものと同じである。ひとつよく分からないのが[i]におけるピッチの下降である。図3「射つた」の[i]におけるピッチの下降は、第1拍が高であり、第2拍が低であることを無音部の前の[i]の音のうちから示しておくという働きがあるのかもしれないが、「居つた」の[i]では一度ピッチを下げる必然性がない。[t]の調音の際に伴う生理的な機構なのであろうか。

4. ピッチ曲線の関数化

「関数のグラフ1」「関数のグラフ2」にある合計やっつのグラフの○はソナグラフから数値化したピッチで、「ソナグラフDSP5500による分析結果1」「ソナグラフDSP5500による分析結果2」にあるピッチの値をそのままうつつしてある。点線は第2節で「ピッチ抽出を行なった。その数値をある関数にあてはめて、その関数でのパラメータを割り出した。」と述べた「ある関数」による曲線である。○は、[momo]を語音とする「関数のグラフ1」ではひとつながりで長く、0.2秒以上の長さを持ち、[ita]を語音とする「関数のグラフ2」ではふたつに途切れている。

点線の示す「ある関数」というのは以下のとおりである。

$$\text{式(1)} \quad \ln F_0(t) = \ln F_{\min} + A \phi_1(t - T_0) + A a \{ h_2(t - T_1) - h_2(t - T_2) \} \\ + A l h_2(t - T_3)$$

これは、藤崎博也氏による式(2)、(3)をもとに設定した(藤崎博也「日本語の音調の分析とモデル化」『講座日本語と日本語教育第2巻』1989、明治書院、参照)。

関数のグラフ 1

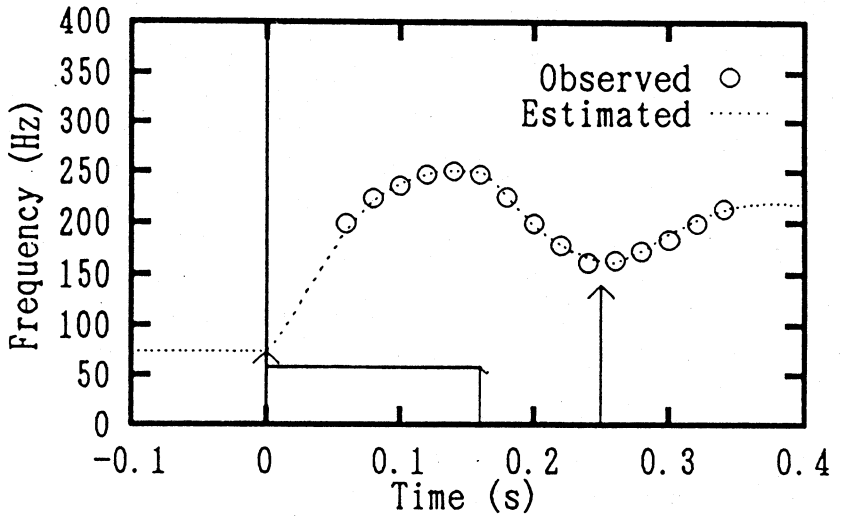


図5 問いかけ「もつも (腿) ?フ」

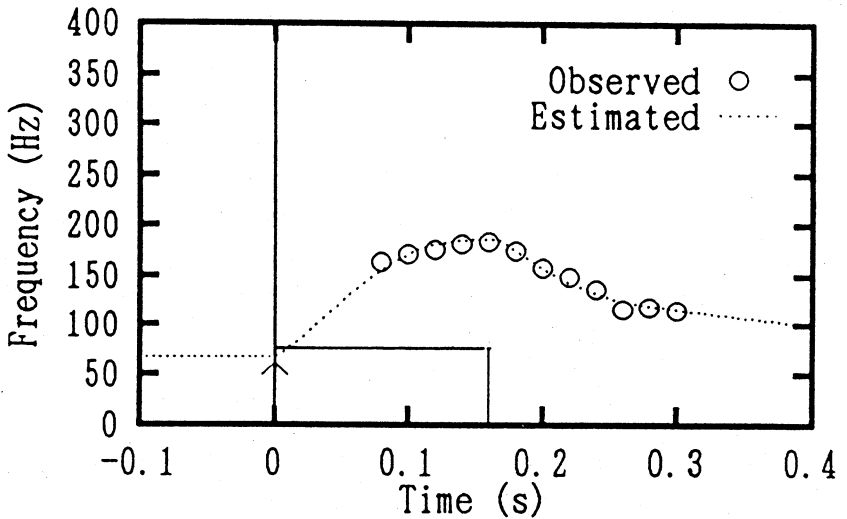


図6 返事「もつも (腿)。〜」

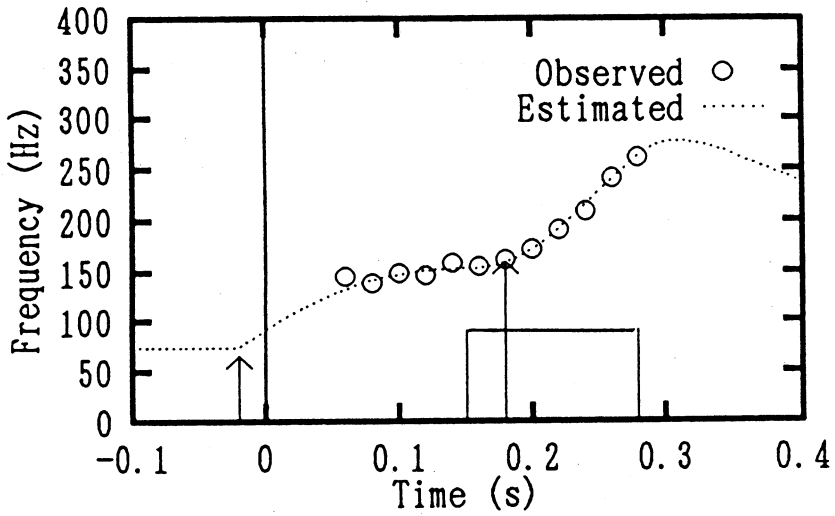


図7 問いかげ「もも(桃)？」

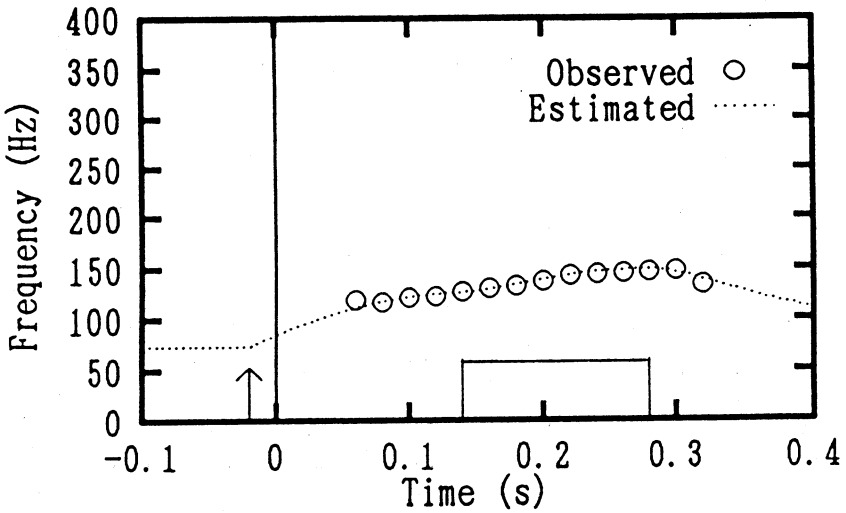


図8 返事「もも(桃)。」

関数のグラフ 2

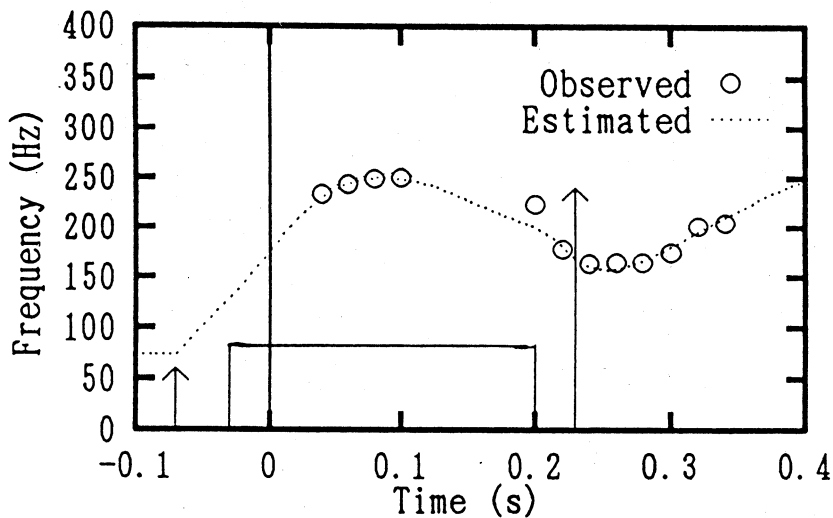


図9 問いかけ「射った？」

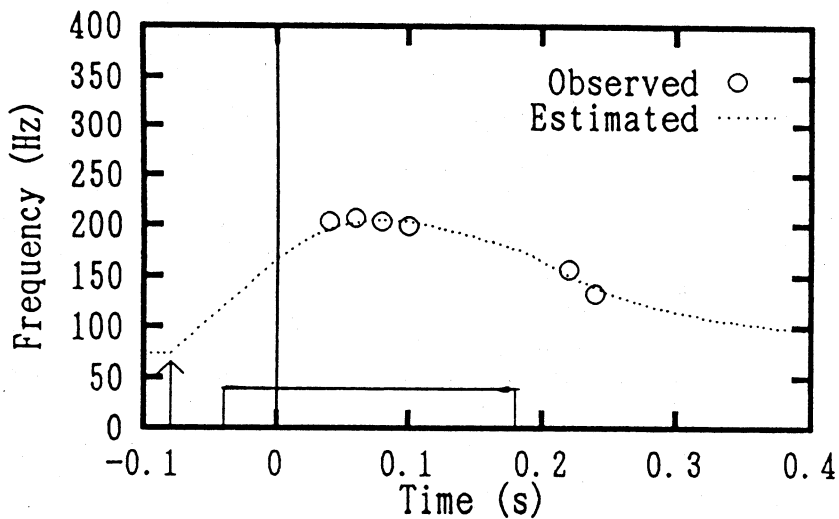


図10 返事「射った。」

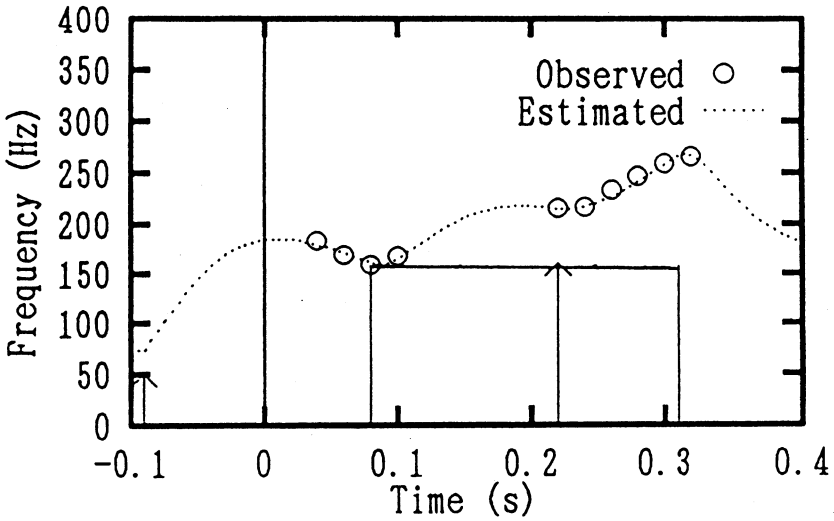


図11 問いかけ「居た？」

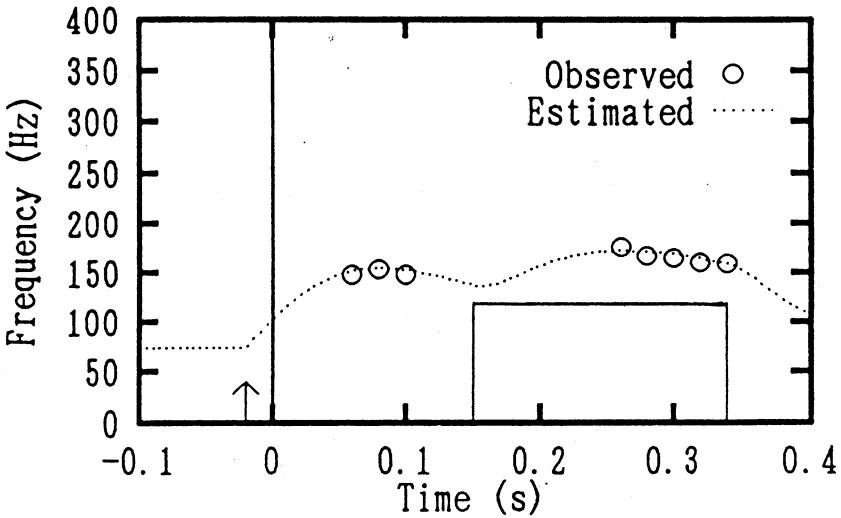


図12 返事「居た。」

$$\begin{aligned} & \alpha^2 t e^{-\alpha t} & t \geq 0 \\ \text{式(2)} \quad h_1(t) = & \begin{cases} 0 & t < 0 \\ \min \{1 - (1 + \beta t)e^{-\beta t}, \theta\} & t \geq 0 \end{cases} \\ \text{式(3)} \quad h_2(t) = & \begin{cases} 0 & t < 0 \end{cases} \end{aligned}$$

式(2)はフレーズ成分に相当する系のインパルス応答、式(3)はアクセント成分に相当する系のステップ応答を示している。インパルス応答というのは、瞬間的に立ち上り、その後減衰していく外力に対する応答であり、ステップ応答というのは、時刻0までは振幅0で、それ以降は振幅1をもつ外力に対する応答である。フレーズというのは文の中でひとつまとまりの語句といった単位で、1語文では文全体と同じことになる。式(2)、(3)におけるパラメータ α 、 β を割り出し、○の列に式(1)の関数が重なるようにしたのがグラフの中の点線である。式(1)の A_p はフレーズ・イントネーションを、 A_a 、 A_I はそれぞれアクセント成分、文末イントネーションを示している。 $\ln F_{min}$ は固定値で、ここでは54Hzくらいに設定している。これは、声帯がまったく緊張していなくてもその筋肉が固有のものとして持っているだろうと仮定した数値である。 $A_p h_1(t - T_0)$ の部分が式(2)に相当し、1語文の場合は、この部分で返事の方の文の断定の文末イントネーションを含んだ音調を示している。そして、 $A_a \{h_2(t - T_1) - h_2(t - T_2)\}$ が式(3)に相当し、語のアクセント、 $A_I h_2(t - T_3)$ が問いかけの文の上昇イントネーションを示している。藤崎氏の式(2)はひとつまとまりの句全体にかかわるピッチの変化を示し、式(3)は語アクセントによるピッチ変化を示す。ここでは、文末の上昇イントネーションに式(3)をあてはめて使っている。

各グラフのTime 0はソナグラフDSP5500による分析結果から、音声が始まったと考えられる時点である。ピッチの始まりが重ならないのは、音声が始まってもピッチが抽出可能なまでに音量が上がっていないためである。ピッチ抽出のために閾値を下げることもできるが、そうすると分析の精度が落ちる。

問いかけの文では↑が2本、返事の文では1本見られる。この↑は式(2)で表わされる瞬間的なインパルス応答が加えられる時点を示す。式(1)の T_0 がこの時点である。式(2)の緊張は $\ln F_{min}$ で表わされる最

も基本的な緊張の上に乗る形になる。さらにその上にかぶさるのが式(3)で表わされるステップ応答である。これはグラフの中では長方形の箱で示されている。長方形のたち上がり時点が式(1)の T_1 で、反対の端が T_2 である。どのグラフでも0.1秒から0.2秒の間持続していることが分かる。返事の文では、以上のふたつの応答だけでピッチの上昇および減衰が示されるが、問いかけの文ではさらに文末の上昇のための緊張を示す関数が必要となる。それが、2本目の↑で、このインパルス応答が加えられる時点は式(1)の T_3 である。

図9の問いかけ「射つた？」のグラフを見ると、音声の始まりの0.07秒ほど前から声帯の筋肉の緊張が始まり、ついで、語アクセントを形成する筋肉の緊張が追加され、その直後音声が始まり、0.2秒後に語アクセントを形成する筋肉の緊張は解かれ、ピッチが下がったあと文末イントネーションを上昇させる緊張が加わっていることが分かる。

低高のアクセントを持つ語の問いかけの文では最初の緊張、音声の開始、語アクセントのための緊張、その緊張の加えられている時間内での文末イントネーション上昇のための緊張の追加、こういった過程が成されている。このことが、図7、図11のふたつのグラフを見くらべることによって分かる。

5. わかったこと

以上の実験と分析により、次のことがわかった。

1. 「もつも(腿)？」「射つた？」、あるいは、「あつめ(雨)？」といった2拍で高低のアクセントを持つ語の1語でできている文は問いかけの上昇イントネーションを伴って発音されても、「もつも(桃)？」、「居た？」、「あつめ(飴)？」にはならない。それは、語の音調の下降が認められてから文末イントネーションの上昇が始まるからである。
2. 藤崎氏の「アクセント成分に相当する系のステップ応答」式(3)は2拍語疑問文の文末イントネーションにあてはめることができる。

6. 日本語教育への応用

外国人に対する日本語教育では、書く時も話す時も、男でも女でも、いつでもどこでもつかえるようにと「ですます体」を教える。この文体を日本語教育ではよくフォーマルという。また、友達同士の会話などで使われる文体をインフォーマルといい、インフォーマルからフォーマルへ行くのはやさし

くないが、フォーマルからインフォーマルへは難しくない、ともいわれる。これが事実か、あるいは、最良の方法かはともかく、日本にいる留学生は日本語の教育を受けているにもかかわらず、または、受けているからこそ、インフォーマルな言い方は弱く、「も₁も₂?」とか「見₁る₂?」のような、アクセントで一度下がって、そのあと、文末の上昇イントネーションでまた上がるという音調ができない。日本で日本語を学習し、その後、数年日本にいたような学生なら友人もいようし、アルバイトもしているだろうからインフォーマルな日本語であってもかなり流暢なのであるが、文末のごく短い間に下がって上がってというのは容易ではないようだ。

が、手本を示してやらせてみると、意外なほどによくできる。「居₁る₂?」「射₁る₂?」、「買₁らう₂?」「飼₁らう₂?」などとアクセントだけが違うミニマル・ペアを並べてもおもしろいし、「た₁らべ₂る₃?」「お₁き₂る₃?」といった3拍の動詞を提出しても、「あ₁め(雨)?」「き₁の₂う₃?」といった名詞で練習してもいい。しかも、学生たち自身、日頃話しているインフォーマルな日本語が日本人と自分とでどこか違うことは分かっている、その違っているところのひとつが実感されるらしく、嬉々として練習に取り組む。教師が指示しなくても学生同士で言い合ったりする。

「も₁も(腿)?」、「射₁た₂?」が「も₁も(桃)?」、「居₁た₂?」に化けないのは前者のふたつでは一度下がってから文末イントネーションの上昇があるからだ、ということは内省だけでも分かる。が、それを実験音声学によって客観的に目に見える形にして確認することには大きな意義があるし、また、内省できない外国語話者への日本語教育のためにはなくてはならない作業である。

今後も、日本語話者にはあまりにあたりまえで改めて考察することもない言語音声の諸事実を実験音声学の手法を使い、研究していきたい。

参考文献

1. 藤崎博也「日本語の音調の分析とモデル化」『講座日本語と日本語教育 第2巻』1989、明治書院
2. Osamu Mizutani, Nobuko Mizutani AURAL COMPREHENSION PRACTICE IN JAPANESE 1979, The Japan Times, Ltd.