

ピッチの認識の相対的側面と絶対的側面

池 田 広 昭

Relative versus Absolute Pitch Recognition

Hiroaki IKEDA

Abstract

Absoluteness, as well as relativity, is sometimes observed to be involved in pitch perception. The absoluteness whereby high tones are distinguished from low tones may be ascribed to some sort of tension caused in the speech organs. Cantonese is one of the languages that utilize such absoluteness, while Japanese adopts highly relative pitch system.

日本語のアクセントも中国語やビルマ語、ベトナム語、タイ語などの声調もピッチが関与する現象であるから、その示差の特徴は相対的高低関係であると一般に考えられている。ピッチの相対性が共通しているということから日本語のアクセントと中国語などの声調を同列に扱うことも一部には行われている。ところが、声調言語を実際に調べたり、学んだりしてみると声調には必ずしも相対的でない側面もあることに気付く。また、日本語のアクセントと中国語などの声調は高低関係の相対性や他のいくつかの点に関して同一でない点が観察される。こういった点を考慮しつつ、主として広東語と日本語を例にとりながらピッチの関与する現象の認識にかかわる諸相に目を向けてみたい。

日本語のアクセントの定義と声調に関する代表的見方を引用し、その問隙点を考えるということから話を始めたい。日本語のアクセントは『音聲學大辞典』によれば次のように定義されている。

「アクセント節における拍相互間（方言によっては稀に拍内部）に認められる相対的高低関係のきまり」¹⁾

声調に関しては Kenneth L. Pike の *Tone Languages* に次のような箇所がある。

“Tone languages have a major characteristic in common: it is the relative height of their tonemes, not their actual pitch, which is pertinent to their linguistic analysis. It is immaterial

to know the number of vibrations per second of a certain syllable. The important feature is the relative height of a syllable in relation to preceding and following syllables. It is even immaterial, on this level of analysis (but not in the analysis of the linguistic expression of emotion), to know the height of a specific syllable in proportion to the general average pitch which the speaker uses. Rather, one must know the relationship of one specific syllable to the other syllables in the specific context in the particular utterance. A man and a woman may both use the same tonemes, even though they speak on different general levels of pitch. Either of them may retain the same tonemes while lowering or raising the voice in general, since it is the relative pitch of syllables within the immediate context that constitutes the essence of tonemic contrast.”²⁾

しばらく日本語のアクセントに関して言うべきことはないが、声調の説明についてはいくつか注意すべき点がある。まず第一に気の付く点は声調をになう単位を音節としていてその音節内部のピッチの動きには触れていないという点である。ここに引用した箇所には音節が声調をになう単位などということははっきり言っていないが文脈から明らかである。（*Tone Languages* の他のところでははっきり言っている。）次に声調の相対性をなんとなんの間にみとめているかということが注目される。

ひとつは個人対個人に関して相対性を認めている。すなわち、同じ「高」という声調でも男のほうが実際の声の高さは女より低いということである。これは男女間に限らずすべての個人対個人の関係について言えることである。さらにもうひとつ、同一個人内において、音節と音節の間に相対性を認めている。すなわち、ある音節が「高」か「低」かは基本周波数 (F_0 と以下表記する) の値によって決まるのではなく、直前や直後の音節との関係によって決まるということである。したがって、ある音節がその音節を発した人の一番普通の声の高さ (the general average pitch) に対して高いのか低いのかということも声調の認識には関係なく、あくまで前後関係が重要だということになる。第三の点としては、この引用箇所に限って言えばというただし書きが必要だが、声調の認識にかかわる要素としてピッチしか考えていないという点が注目される。また、ピッチにかかわるものとして F_0 以外のものを考えに入れていない点も着目すべきである。ここで誤解を避けるために言っておくと、ここに Pike を引用しているのは Pike の声調に関する説を紹介するためでもないし、批判するためでもない。あくまでこれから話をすすめる上での出発点とするために過ぎない。したがって、普通と違って、Pike が他で何を言っているのかは問題とせず、引用している箇所の記述だけを問題とする。それでも念のために付言すると、Pike は声調にはピッチの他に時間 (time)、強さ (stress)、声門閉鎖 (glottal stop) 等々が関与する場合があることを認めている。

以上 Pike の声調の説明に関して指摘した3つの点は声調の相対性に根本的にかかわる問題点であると思うので、これからはこの3点を中心に話をすすめて行くことになるが、その前に相対性の適用の範囲を明確にしておく必要がある。ものごとは焦点のしぼり方によって相対的と言えたり、絶対的と言えたりする。自然科学では相対性理論に行き着くし、ヒンズー教なら絶対的存在としてのブラフマンに行き着く。ヒンズー教でもこの世の現象は絶対より生じた相対の世界としている。ここまで極論しなくても、何に焦点をしばって相対的と言うのかということを明確にしておかないと水掛論になる恐れがある。

声調の相対性を論ずる場合、普通は Pike の引用にみられるように個人対個人、同一個人内における音節対音節に相対性の焦点をしばる。しかし、音節より下位の単位にまで焦点をしばることが可能である。上昇する声調を「低+高」というふうに分解したり、下降する声調を

「高+低」としたりして、音節全体に「上昇」とか「下降」という動きを認めるかわりに、音節より下位に「高」とか「中」とか「低」とかの段階のみを認めこの組合せによって声調を説明しようという方法である。確かに、動きの伴う声調 (以下 gliding tone という用語を用いる) は開始地点と終末部とに相対的高低変化があることは認めなければならない。こういった意味においてすべての gliding tone は音節内において相対的だと言える。gliding tone を音節よりさらに下の段階にまで分析することの是非は問わないにしても、音節内においては gliding tone に相対的な面があることは認めなくてはならない。

ところで、単独であらわれる場合の北京語の4つの声調は「 \square 」, 「 \vee 」, 「 \cup 」, 「 \searrow 」となっている。もちろん、文脈中ではこれ以外の形があるが、それはいま考慮外とする。今、このうちのどれかひとつを前後関係なしに発音したとする。この場合、聞き手は前後関係のない単独の発音でも4つのうちのどの声調であったかはつきり識別することができる。それは4つの声調の形態、つまり、動き方が全部違うからである。このように声調内部においては相対性が、上述のごとく認められるにしても、それより上の段階、すなわち、音節という段階になると相対的と言えない場合も出てくる。

架空の言語の声調体系として、「 \square 」, 「 \cup 」, 「 \searrow 」というすべて平らな声調よりなる体系を考える。こんどはすべて同じ形をしており、高さのみによって3つが対立している。このうちのどれかひとつを発音した場合、前後関係が与えられていないとすれば、聞き手はどの声調だったか識別できるであろうか。もし、できないとすれば、この言語においては、単独の音節の声調には絶対性が無い。この場合、音節と音節が接するときにはじめて声調の識別が可能になると言える。

分節音素 (segmental phoneme) も言語音として実現するときには話者の個人的特徴が付加される。聞き手は言語情報とこれらの情報を区別する能力がある。声調の場合の個人間の高低の違いなどはこの手のものである。ある人は 80 Hz ~ 250 Hz の話し声の声域をもち、ある人は 100 Hz ~ 450 Hz の話し声の声域をもつかも知れない。したがって、 F_0 の絶対値は声調の認識の決め手とはなり得ない。同じ 150 Hz でもそのもつ意味は人によって違う。そういう意味での声調の相対性は分節音素の実現の際の個人的特徴同様当然のことである。

以上のことをまとめると、声調を音節より下位の単位まで分解するとすべて相対的となり、個人間の相違まで

視野を広げるとふたたび、すべて相対的となる。一方、音節をひとつのまとまった単位として考えると、絶対的な場合も相対的な場合もある。すなわち、相対性を論ずる余地のあるのは、同一個人内における、音節という単位でみたときの声調である。音節内の相対性や個人対個人にみられる相対性は当然認められなければならないことであるから、これ以上触れないことにして、声調、そしてアクセントの相対性の焦点を同一個人内の音節(日本語のアクセントは拍あるいはモーラというのが正しいが、ここでは音節としておく)という音節対音節の関係にしばることにする。

前に単独の場合の北京語の声調にみたようにそれぞれの声調がすべて別の形態をしている場合、音節対音節の関係がなくてもどの声調が識別できる。このような声調は音節という単位において絶対的といえる。これからはこのような声調を絶対的声調と呼ぶ。実際の北京語は文脈があると少し事情が異なってくるが、単独の声調のみに話を限定すれば絶対的声調と言える。

また、架空のものとして考えた声調体系はその識別が不可能なというただし書き付きだが、音節という単位において相対的である。これを相対的声調と呼ぶことにする。

絶対的声調に関連して考えておくべきことがある。いままでは声調にかかわるものとしてピッチのみを考えていたが、以前少し触れたように、ピッチ以外のものも示差的特徴として声調の識別に関与している言語がある。たとえば、北部ベトナム語は6つの声調を区別するが、そのうちの3つは声門閉鎖、声門閉鎖+声門破裂、喉頭化(まとめて声門閉鎖と言ってよい)が示差的特徴となっている。こういった特徴が絶対的であることは論を待たない。ただし、北部ベトナム語の声調は絶対性と相対性のまざった体系である。

広東語に関して分析した結果次のような声調体系であることがわかった。すなわち高さを4段階区別する。形はのぼるものが2つ、常に平らであろうとするもの2つ、平らでも下ってもよいもの2つという区別がある。平らでも下ってもよいという声調が仮に平らとして実現すると平らな声調が4つあることになる。これは架空のものとして作った相対的声調の体系より複雑な体系である。こうなると単独であらわれたら4つのうちのどの声調が識別することはできないように思われるであろう。しかし、実際は識別しているのである。これは経験的事実である。はっきり意識してはいないにしても誰しも声調の高い低いを言うときにこういった経験が裏にあるの

表 1 Distinctive features of Cantonese tones

Tone	Height	Contour
T 1	high	-rising
T 2	high	+rising
T 3	middle	+level
T 4	extra low	-rising
T 5	low	+rising
T 6	low	+level

である。たとえば、広東語の[fen]という音連続には「分、粉、訓、焚、奮、份」などの形態素があてはまるが、これらは表1で示した第1声から第6声に相当する。順番に第1声、第2声、……となっている。このうちの第1声である「分」を言おうとして、普通に[fen]と言うと何度やっても「分」には聞いてもらえない。広東語話者からのコメントとして発音はそれほどまづくないがもう少し高く言わなければだめだというようなことを言われる。前後関係もなにもなく、単独で発音しても必ずこのようなことが起こる。声調は相対的なものだから前の音節や後ろの音節との関係においてのみどの声調が識別されると思いついていては納得が行かないことになる。

広東語に限らず、東南アジアの声調言語といわれるものにはみなこういった性格があるようである。理論的には最も相対性を主張し易い場合、すなわち、平らな声調がいくつも対立しているような場合でも、現実には絶対的認識が行なわれているのである。もし、相対的関係によってのみ声調が識別されたとすると広東語のように4段階も高さの区別のある声調体系においてはある音節の声調が決定できるまでに相当時間のかかる場合が出てきたりして不都合だからである。第1声が5音節も6音節も続いたりするとそれらが第1声とわかるには時間がかかることになり不便である。前後関係がなければ声調の識別ができないとすると以上のように不便なのである。東南アジアの言語の声調が一見相対的にみえるところでも現実には絶対的認識を行なっていることの背景には、これらの言語が基本的には単音節語であるということがありと思われる。1音節1音節が形態素あるいは語に相当しているわけで、辞書的意味に関係する声調がすぐに識別できないようでは伝達が進められに行かなくなるのであろう。多音節語なら必然的に前後関係が生じるわけだから必ずしも絶対的認識に頼る必要もないわけである。そのような言語における声調の認識は純粹に相対的に行なわれると推測することも無理とは言えない。冒頭

に引用した Pike の説明はそういう言語により適切であるとは思ふ。実際にそういった言語もあるようであるが筆者は直接分析する機会をもったことがないのでこれ以上のことは言えない。

前後関係なしに高いとか低いとかどのようにして判断しているのかというと、これはそれほど不思議な現象ではない。誰でも普通にやっていることである。これはちょうどある人が背が高いとか低いとかいうのに似ている。背丈が高いとか低いとかは本来相対的なことであるはずである。あの人よりこの人のほうが背が高いとか低いとか言うのは理屈に合っているが、いきなりあの人は背が高いと言うのは本当はおかしいと言える。何センチ以上は背が高く何センチ以下は低いという決まりもない。しかし、現実には 190 cm ぐらいの背丈の人に対し、他の人と比較せずに、あの人は背が高いといっても全然おかしくない。それは暗黙のうちに我々の頭の中に人間の平均的背丈はどのくらいかという基準があって、この基準に対して高いとか低いとか言っているからである。声調の場合もこれと少し似ている。我々はある人の発した音がその人にとって自然な高さか、無理をして高く言っているか低く言っているかすぐに判断することができる。日本語においてはこのことは語の辞書的意味の区別に役立たないが、声調言語においては重要な役割を果している。生味の体を持つ人間の声帯はある限られた範囲の音しか出すことができない。声帯の振動に関するこの範囲を声域という。声域に歌声の声域と話し声の声域があるがいまここで問題にしているのは話し声の声域である。声域の中間部の声は無理なく自然に聞こえるが、高い方の極限に近付けばある種の無理をしなければならなくなる。低い方の極限に近付く時も同じである。しかし、高い方の無理と低い方の無理はその仕方に相違がある。この無理の仕方の違いが高いとか低いとかの感覚を生む。この無理が具体的に発声にかかわる器官にどのようにあらわれるかいまのところわからないが、こういふことがある音を前後関係なしに聞いたときにでも低いとか中ぐらいとか高いとか判断することを可能にするのであろう。声域が 80 Hz~250 Hz の人の 150 Hz と声域が 130 Hz~400 Hz の人の 150 Hz とでは前者の 150 Hz は中ぐらいに、後者の 150 Hz は低く聞こえることになるであろう。声調言語を話す人はこういったことに敏感でそうでない人はそれ程敏感でないと言えると思う。

日本語のアクセントが広東語などの声調とちがう点は「高」と「低」の認識に上述のような絶対性がないということである。いま、「葉」と「齒」を単独で発音する

とする。もうすでに我々には音の高さを絶対的に認識する能力があることを指摘した後であるから、ただ単に高い、低いということばを用いていうと、「葉」という語は単独には高く言っても低く言っても意味に影響はない。「齒」も同様である。意味に影響がないというより、日本語には単独の 1 音節語をアクセントによって意味的に区別する習慣はない。そうではないと思う人は間違いであるからよく実験してみる必要がある。一方、両方とも「が」が後ろにつくと「葉」のほうは「が」より低く、「齒」のほうは「が」より高く言わなくてはならないというように高低関係が生じる。このように日本語は 2 モーラ以上に関して始めてアクセントという現象が起こる。この点、前後関係がなくとも必ずその声調固有の高さで言わなければならない広東語とは異なる。日本語のアクセントは相対的と言えるが、広東語の声調はみかけは相対的でも実は絶対的なのである。無論前にことわった通り音節という単位においての話であるが。

次のような漢字が並んでいたとする。

a) 医 胃 意

b) 氏 四 市 死 師 詩

これを読んでくれと言われたとしたら、日本人はどのように読むか。おそらく a) も b) もみな一様の高さで「イ、イ、イ」「シ、シ、シ、シ、シ、シ」のように単調に読んでいくことであろう。「が」がついたとき「胃」は「が」より低いから、また、「医」と「意」は「が」より高いからと言って「高、低、高」というアクセントで読むことはない。b) の場合も同様、「詩」だけを特別低く読むことはしない。これは単独という意識がはたらくからで、単独の場合には日本語にアクセントがないことをものがたっている。日本語のアクセントの相対性を裏付ける材料として扱うことができるであろう。

広東語で次のような漢字の並びを読むとする。

a) 威 委 畏 為 偉 謂 ([wɛi])

b) 衣 椅 意 而 以 易 ([ji])

日本語の場合とちがって平板に読むということはなく、必ず a) も b) も順に高く平らに高く始まり下降するかのどちらか、高いのぼり、中ぐらいで平ら、非常に低く平らに低い下りかのどちらか、低いのぼり、割合低く平らに読む。これは単独の場合でも声調が絶対性を保っていることを意味する。前から言っているように平らな声調も日本語のような高さの自由がない。

広東語は 4 段階もあるから仮に、第 1 声の中ぐらいから始めたとして第 6 声や第 4 声を言うことができないおそれがある。また、第 4 声の中ぐらいの高さから

始めたとなると第3声や第1声が言えなくて困ることになりかねない。したがって広東語の場合、同一個人においてはそれぞれの声調のとり Fo の値にかなりの制限があることになり、それぞれの声調はいつもだいたい同じ Fo をとろうとするということが考えられる。実際に Fo の値を測定してみるとそれぞれの声調のとりうる値の範囲はそれ程狭くないが、かといって自由ともいい難くその窮屈さがうかがえる。

日本語は2段階のアクセントしかなく、絶対性もないので「高」のとり値も「低」のとり値も本当に自由である。日本語は発話の全体に渡って、割合急に調子が高くなって次第に低くなるというパターンが観察されるが、この大きな流れの中に「高」「低」の関係があらわれる。一般に直前直後のモーラ同士の高低の差はかなり小さい。発話全体としての最高部の「低」と最低部の「高」とをくらべれば、最高部の「低」のほうが最低部の「高」よりよ程 Fo の値が高い。その最高部と最低部の差は広東語の第1声と第4声に相当する場合もある。しかし、日本語の場合この高低差は時間をかけて生じるのに対し、広東語の場合は一瞬にして生じる。このため、日本人には広東語の声調が非常に難しいし、広東語話者にとっては日本語のアクセントがなじめないということになる。

広東語の声調に関して他に気の付いた点を述べると、単独で声調を言う場合はその高低の差が普段より大きくなるという点がある。これは前後関係がなくても識別可能にするために絶対性を強めている現象と考えられる。また、文脈中においては高低の差は少くなる。しかし、前後の声調の高低の差は相対的にそれぞれの声調の組み合わせによってきちんときまっており、これがくず

れることはない。すなわち、文脈があるときは相対的低高が声調の識別にとって非常に重要になっている。その結果、ある環響の第1声とある環響の第3声の Fo の値が同じになるなどの現象がよくみられる³⁾。広東語は単独のときは絶対性が強く、文脈があるときは相対性が強いという2面をもちあわせていると言えると思う。

以上、高さの段階のみによって区別される声調にも絶対的側面があることを指摘してきたが、なぜそのようなになっているのかに思いをめぐらすと言語情報の冗長性 (redundancy) ということに思いあたる。言語は言語情報が雑音によって一部消される場合や聞きもらしなどによって伝達が不完全になるのを避けるため大変多くの重複を含んでいる。英語の3人称単数現在の s や不定冠詞などそういう典型的例である。声調そのものもおそらく分節音素を補う冗長部分なのであろう。その声調の高低の識別に発声器官の緊張の仕方を用いて絶対性を高めているなどというのも言語の一種の冗長性なのであろうと思う。

註

- 1) 『音聲學大辭典』日本音聲學會編、三修社。1976年6月。8ページ。
- 2) Kenneth L. Pike, *Tone Languages: A Technique for Determining the Number and Type of Pitch Contrasts in a Language, with Studies in Tonemic Substitution and Fusion*. Ann Arbor, Univ. of Michigan Press. 1964, p. 4.
- 3) 筆者の上智大学における修士論文 “Distinctive Features of Cantonese Tones: An Experimental Basis” 1979年を参照されたし。