

# 大脳トップダウン処理に注目した イントネーション研究 — 中・日対照 —\*

盧 嘉 (筑波大学大学院)

**キーワード:** 大脳トップダウン処理、文頭のイントネーション、実験音声学、音声情報レベルにおけるモダリティ、イントネーション類型

## 1 はじめに

実験音声学の研究方法は、近年急速に進展しており、とりわけパソコンの普及という点からは、音響音声学に見るべきものがある。しかし、それとは別個に脳神経科学の進展と、これを援用した聴覚情報処理系に注目した聴覚実験音声学の発展にも、これからの新領域としての新たな期待が寄せられている。

そのような情勢の中、本稿は城生佰太郎 (2001: 9-45) に着想を得て、大脳におけるトップダウン処理に着目することによって、川上夔 (1956) 以来の指摘である文頭位置におけるイントネーションを観察し、実験音声学の観点からそれに対する考察を加えるものとする。

## 2 問題の所在

### 2.1 大脳におけるトップダウン処理とは

---

\*本稿は、2005 年度に桜美林大学大学院に提出した修士論文(主査：城生佰太郎先生)をもとに、加筆・修正したものである。今回、あらたにデータ解釈をしておいた部分をはじめとして、内容面でも気づいた範囲で可能な限りの加筆・修正をおこなっている。

音声によるやりとりを、脳における情報処理という観点から捉えると、まず第1に気づく点は、反応時間である。つまり、相手の発話を最後まで聞いてから応答を考えていたのでは、例えば衛星回線による電話で応答するようなもので、タイミングずれを生じてしまい、スムーズな会話が成立しない。にもかかわらず、一般的に母語話者が互いになめらかなコミュニケーションを交わすことができるのは、音声情報を最後まで聞かないで、頭の部分の適当量を聞いただけで残りがどうなるのかを予め予測して聞いているからだ、と考えるのがもっとも自然である。

「大脳のトップダウン処理」とは、このような大脳における絶妙な働きを捉えた呼称だが、城生佰太郎 (2001: 31-32) に簡潔な説明があるので、以下に引用する。

私たちは、毎日数え切れないほどの言語音に囲まれて生活をしている。しかし、日常生活レベルでの音声は実験室で「純粋培養」されたようなものとは違い、高品位な例に出会うことはめったにない。街の雑踏や地下鉄など、うるさい騒音の中での会話だけでなく、時には調音明瞭度の著しく低い個人との対話を余儀なくされたり、母音や子音からアクセントやイントネーションの基本的タイプまでも異にする人たちともコミュニケーションを取らなければならない。従って、とうてい100%の聞き取りは望めないとしても、結構これで実生活の上では支障なくそれらをこなしているという事実がある。

結論的にいえば、これらの「芸当」が可能なのは、大脳における「トップダウン処理」のおかげであり、隣接領域の心理学などでは、こうした大脳の営みに注目してさまざまな実験が行なわれている。例えば、Yokosawa (1998) などは視覚情報処理過程に注目して、私たちが物を見るときに色・形態・奥行き・運動…など、いわゆる低次の属性分析を行なう「低次視覚」だけでも相当複雑な営みを行なっているのにもかかわらず、それら低次の各属性を統合してはじめて得られる「いつどこで何がどうした」という判断、すなわち「高次視覚」が一瞥しただけでも可能なのは、脳が経験則に基づいて実際には一部分しか見て

いない現象を推論して貼り合わせているからにほかならないとしている。

上にいう「トップダウン処理」とは、まさにこのような脳における推論機能をさすものであり、言語音の認知に関しても冒頭に述べた事実から考えて、こうした営みは十分にあてはまる。これを別の表現に改めれば、脳では耳から入った一音一音を逐一分析して捉えているのではなく、ある特徴的なキューに注目してこれを手がかりに言語音を推論して理解しているということにほかならない。

以上で明らかなように、私たちは日常における音声のやりとりの中で、少なからずこの「大脳のトップダウン処理」を用いて、相手の発話の結末をある程度推論しながら、お互いのコミュニケーションを円滑化していることが窺われる。つまり、卑近な言い方をすれば「探りを入れる」ということが、スムーズな会話のやり取りにおいては、不可欠だということが示唆される。

## 2.2 文頭のイントネーション

近年、イントネーションに関する実験音声学的研究はかなり盛んになってきているが、日本語を対象とした先行研究では主として文末の音調が扱われている。一方、日本語以外の言語を対象とした研究でも、例えば Greenberg (1963) は、言語の普遍性についての研究の中で、イントネーションを用いて疑問文をその対応する平叙文と区別する言語では、イントネーション上の差異が文頭においてではなく、文末において認められるという特性を普遍性の1つにあげており、やはり文末に注目している。

また、中国語と日本語を対照させた福岡昌子 (1998) の研究でも「学習早期の中国人母語話者は、文末のピッチの変動と勧誘の表現意図を結びつけることが難しく、勧誘を表す文末のピッチの上昇率が知覚上十分に習得されていないことがわかった」との指摘をしており、同様に文末のみ注目している。

たしかに、改めて指摘するまでもなく、文末の位置におけるイントネーションには注目すべき特徴が多く見られる。しかしながら、それ以外の要

素、すなわち

①位置としては、文末以外の位置

②要素としては、

- a. 強さ
- b. 長さ
- c. ピッチレンジ
- d. 韻質

などに対しても取り上げるべき点はあるであろうし、また、①②を総合的に捉えるという視点も必要であろう。

そのような中で、比較的早い時期に、文頭におけるイントネーションに着目した研究として、川上夔 (1956) が知られている。この研究では、単語の音調 (アクセント) の型の具体的な“姿”——音声学的レベルにおける呼び方——は本質的なものではなく、すべての現象は“弁別の特徴”——音韻論的レベルにおける呼び方——にこそその本質があるとするもので、後に展開される川上夔(1961) に主張される「句音調」という考え方の先駆けをなすものとして注目される。

この議論の最も重要な出発点は、アクセントの型のうち、下がり目を持たない「平板式・平板型」とイントネーションとの関係である。すなわち、「平板式・平板型」のアクセント型を持つ単語が、句や文などのようなさらに大きな単位に組み入れられると、もともとの単語レベルでは実現されていた“高さの姿”が変化してしまうという点を重視しているということにほかならない。なお、この現象を城生佰太郎 (1998: 157-8) が比較的ていねいに述べているので、以下に引用する。

さて、上に示した音調パターンは、いわゆる「辞書形 (dictionary form)」と呼ばれる形に相当するもので、ちょうど辞書を引くと数多くのさまざまな語形のうちから、見出し語としては代表形だけが載せられているというのと同じである。

従って、例えば表 4-1 にあてはめれば、3 拍語の平板式平板型に所属する「絵の具」という語の音調を、いろいろな脈絡に入れて詳しく調べてみると、

(1) エノグ (絵の具)

(2) アカイエノグ (赤い絵の具)

(3) アオイエノグ (青い絵の具)

(4) コノアカイエノグ (この赤い絵の具)

(5) コノアオイエノグ (この青い絵の具)

(6) トケタアカイエノグ (溶けた赤い絵の具)

(7) トケタアオイエノグ (溶けた青い絵の具)

(8) ワタシガモラッタコノアカイエノグ

(私がもらったこの赤い絵の具)

(9) ワタシガモラッタコノアオイエノグ

(私がもらったこの青い絵の具)

などとなり、ちょっと内省をただけでも(2)、(4)、(6)、(8)では「エノグ」のすべてが高く、逆に(3)、(5)、(7)、(9)ではすべてが低く調音されていることに気づく。さらに、例えば(8)などに関しては、途中でしっかりとした区切りをつければ、

(10) ワタシガモラッタ | コノアカイエノグ

(11) ワタシガ | モラッタ | コノ | アカイエノグ

(12) ワタシガ | モラッタ | コノ | アカイ | エノグ

ただし、| はしっかりとした区切りを示す。

と、音調の上がり目が何か所にもなってしまう。

川上 蓁 (1961) は、(8)のように途中に区切りをつけずに、ひとまとまりとして調音される音調単位を「句」と呼び、上野善道 (1989 など) は平板型における第2拍からのピッチ上昇は、単語に備わっている特徴が具現化したものとみている。

つまり、平板式の音調に共通する語レベルの特徴は、いかなる場合においてもピッチの下がり目を持たないということであり、先に「辞書形」として言及した単独で現れる「エノグ」が第2拍から上昇していたのは、句頭における句音調の影響によるものだというのである。

以上に述べたように、川上 蓁によれば、具体的にはさまざまな姿を取って具現化する平板型アクセントに代表されるようなプロソディー現象というものは、すべてイントネーションのレベルで扱うべきものであり、アクセントというものは、これよりも抽象化の度合いが高い音韻論的レベルで議論すべき性質のものである、という論調となる。

本稿でも、文頭位置に注目したイントネーション研究を行うが、着眼点および方法論は、上に述べた川上 蓁 (1956, 1961 など) とは異なる。すなわち、着眼点としては大脳における聴覚情報処理を、また方法論としては音響解析を中心とした実験音声学的方法を採択する、ということになる。具体的に述べれば、本稿では大脳における聴覚情報処理系の営みに着目して、特にその中でも「脳のトップダウン処理」に注意を払いつつ、文頭における1音節（結果的には、およそ100-200ms程度）の音声情報に焦点を絞り込んだ実験音声学的考察を加え、その短時間の音声刺激にもかかわらず、平叙文・疑問文・命令文の別をある程度は予測し得るのではないかということをも明らかにしたいと考える。

### 3 方法論

#### 3.1 被験者および分析資料

被験者には、日本語と中国語の男女1名ずつ、計4名のネイティブ・スピーカーを用いた。なお、それぞれの被験者に関するフェイス・シートは、次のとおりである。

日本語 1	J.H. 男性、1946年うまれ、東京
日本語 2	T.M. 女性、1970年うまれ、横浜
中国語 1	Y.T. 男性、1977年うまれ、長春
中国語 2	L.J. 女性、1979年うまれ、長春

次に、分析資料は表 3-1a, 3-1b に示すように、日本語と中国語における同一の文(一語文を含む)に対して、それぞれ①平叙文、②命令文、③疑問文の3種類を用意し、これを男女2名ずつ計4名の母語話者に調音してもらった。なお、日本語と中国語は同一の文意となるよう揃えた。

資料の選定に際しては、中国語では四声を考慮して第1声から第4声までのすべてのパターン(1~4は第1声、5~8は第2声、9~12は第3声、13~16は第4声で始まる)を、また日本語では、b) 起伏式頭高型(10~16)とa) それ以外(1~9)で始まる文を揃えた。

現在の中国語は、中華人民共和国の大陸地方と台湾半島、海南半島、およびシンガポールとマレーシアの一部で話されているが、当然のことながら地域による言語差が甚だしい。従って、本研究では対象とする中国語を、北京方言に限定する。北京方言は、中華人民共和国の首都北京市およびそれに属する華北地方、隣接の東北地方で一般的に話されている方言であり、中華人民共和国の標準語としての地位を保っている最も有力な方言である。

一方、これと対照させる日本語も、標準語としての地位を保っている東京方言と、これに準ずる横浜方言の母語話者を用いた。東京と横浜では、項目によっては多少の差異はあるが、本研究で目指している①平叙文、②命令文、③疑問文のイントネーションレベルにおける高さに関しては、特段の不都合はない。

表 3-1a 日本語文例<sup>1</sup>

a)	1.シメ'ル	b)	10.デ'ル
	2.サゲ'ル		11.マ'ツ
	3.ナオ'ス		12.タ'ッテ

<sup>1</sup>日本語資料の a) は低く始まるアクセント、b) は高く始まる頭高型アクセントのグループを示す。また、「'」は、いわゆるアクセント核を表わす。

- |           |              |
|-----------|--------------|
| 4. ハナ' シテ | 13. モ' ッテ    |
| 5. アソブ    | 14. ク' ルノ    |
| 6. ウメル    | 15. テ' オアラウ  |
| 7. アゲル    | 16. ア' トニツイテ |
| 8. アケル    |              |
| 9. ワスレテ   |              |

表 3-1b 中国語文例

- |                           |                           |
|---------------------------|---------------------------|
| 1. 关上 (guan shang) シメル    | 9. 等着 (deng zhe) マツ       |
| 2. 拎着 (lin zhe) サゲル       | 10. 举着 (ju zhe) アゲル       |
| 3. 跟上 (gen shang) アトニツイ   | 11. 洗手 (xi shou) テアアラウ    |
| 4. 出去 (chu qu) テル         | 12. 改过来 (gai guo lai) ナオス |
| 5. 拿着 (na zhe) モッテ        | 13. 让开 (rang kai) アケル     |
| 6. 别玩了 (bie wan le) アソバナイ | 14. 忘了 (jin qu) ワスレテ      |
| 7. 离开 (li kai) ハナシテ       | 15. 过来 (guo lai) クル       |
| 8. 埋起来 (mai qi lai) ウメル   | 16. 站着 (zhan zhe) タッテ     |

### 3.2 録音

録音は、パーソナル・コンピュータ (SONY, model PCG-971N) にマイクロフォン (TOSHIBA 指向性ダイミク・マイクロフォン, model ADM-1301US) を接続し、Wave 形式にてファイル化した。なお、サンプリング・レートは取り込み時点では 44,100Hz、量子化 16 ビットである。また、録音に際してはすべて指導教官である城生佰太郎教授にディレクターをお願いし、教授の監督のもとに適切な録音がとれるまで複数のテイクを行い、それらのうちから最終的にベストなテイクのみを解析に使用した。

### 3.3 解析方法および手順

解析は、パーソナル・コンピュータ (SONY, model PCG-971N) にインストールされた音声解析ソフト Multi-Speech 3700, ver. 2.4 (KAY 社製) によって行った。

解析手順は、以下のようになる。

- ①あらかじめデジタルファイル化されている表 3-1a, 3-1b の資料を波形表示で読み出し、
- ②これに声帯振動区間をマークする *voiced period marks* を振る。
- ③次に、これに対応するピッチを描記し、
- ④最後に、アルゴリズムによる基本統計量を求める。

なお、この結果は、ビットマップ形式にて画像ファイル化する。このようにして、まずは文全体のイントネーションに関する解析を行った後、第 2 段として、文頭の 1 音節のみをターゲットとする解析に移る。文頭の 1 音節に絞り込んだ解析は、上記①および②を遂行した後、

- ⑥*spectrogram* を描記させ、パターン図形との摺り合わせによってターゲットとする 1 音節を抽出する。
- ⑦その後、カーソルを立てて①の原波形とリンクさせ、当該区間のみに限定して、上記の②③④⑤を実行する。

という流れである。

なお、ピッチの計測に際しては、一貫して *voiced period marks* に依存した自動解析を採択し、手動による調整は行わなかった。

## 4 実験結果

Multi-Speech 上で実験資料を文全体のレベルで解析し、次に文頭第 1 音節に絞りこんだ解析結果の中央値 *Median* (Hz) を測定した。サンプルを図 4-1、4-2 に示す。

図 4-1 から明らかのように、平均値 (*Mean Frequency*) と中央値 (*Median*) とでは、若干値が異なる。一般的には、平均値を取ることが多いが、文頭の 1 音節のみをターゲットとする本研究では、そこに出現する音種によるマイクロ・プロソディー的影響を大きく受けるため、飛び抜けて高い値が出ることもしばしばある。このような場合に、平均値に頼ると、ほんのわずかな例外的なデータによって全体の数値が大きく影響されてしまうので不適当である。この点で、中央値はすべての数値の真中を取っているのも、

比較的例外的なデータによる影響を受けにくいという利点がある。従って、本稿ではこの中央値を用いた。

以下に、中国語1(全体、語頭)、中国語2(全体、語頭)、日本語1(全体、語頭)、日本語2(全体、語頭)、の順に命令文・疑問文・平叙文ごとの解析結果を基本周波数値によって示す。なお、単位はすべて Hz 表示である。

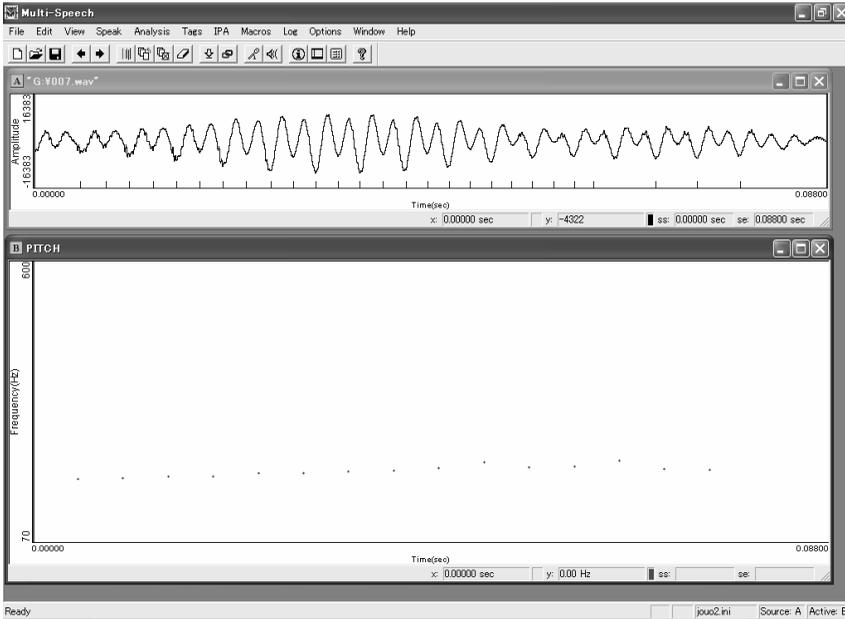


図 4-1 解析サンプル (1)

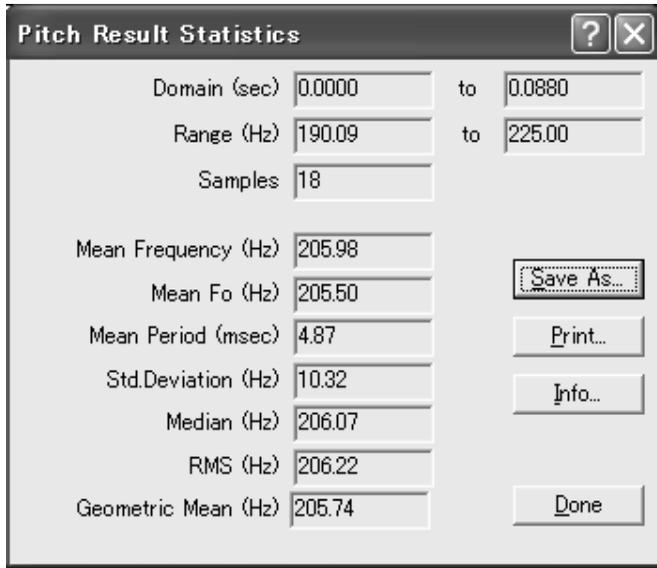


図 4-2 解析サンプル (2)

表 4-3 中国語 1 (全体)

文例	命令文	疑問文	平叙文
1	196	210	179
2	203	206	196
3	197	208	187
4	201	202	188
5	177	190	160
6	168	187	153
7	188	202	169
8	165	173	155
9	172	182	154
10	163	197	139
11	171	180	173
12	162	185	141
13	168	207	153
14	193	206	167
15	181	189	154
16	192	196	144

表 4-4 中国語 1 (語頭)

文例	命令文	疑問文	平叙文
1	230	215	200
2	218	211	199
3	247	210	187
4	253	228	217
5	168	164	148
6	201	181	152
7	205	195	175
8	199	192	172
9	247	196	158
10	162	153	111
11	254	206	159
12	184	142	136
13	153	149	137
14	238	230	204
15	245	225	163
16	254	220	138

表 4-5 中国語 2 (全体)

文例	命令文	疑問文	平叙文
1	279	294	250
2	264	294	249
3	272	304	253
4	233	288	205
5	225	227	210
6	206	221	196
7	226	260	215
8	206	222	196
9	199	223	190
10	200	223	190
11	190	217	170
12	201	232	188
13	229	275	222
14	254	264	238
15	237	254	233
16	238	249	232

表 4-6 中国語 2 (語頭)

文例	命令文	疑問文	平叙文
1	286	268	254
2	286	260	252
3	280	268	257
4	294	286	256
5	205	191	160
6	240	229	226
7	262	227	220
8	257	230	200
9	212	186	175
10	200	185	132
11	246	237	221
12	190	186	168
13	221	204	179
14	308	300	249
15	302	286	253
16	295	292	243

表 4-7 日本語 1 (全体)

文例	命令文	疑問文	平叙文
シメル	259	152	160
サゲル	252	147	220
ナオス	152	142	144
ハナシテ	210	160	196
アソブ	226	163	200
ウメル	304	196	259
アゲル	275	163	259
アケル	154	124	144
ワスレテ	154	131	140
デル	210	149	149
マツ	154	149	149
タッテ	187	142	163
モッテ	220	135	152
クルノ	166	131	152
テ'オアラウ	160	144	152

表 4-8 日本語 1 (語頭)

文例	命令文	疑問文	平叙文
シメル	140	126	131
サゲル	137	106	129
ナオス	114	101	110
ハナシテ	131	111	114
アソブ	122	102	113
ウメル	131	111	111
アゲル	114	99	114
アケル	113	95	113
ワスレテ	137	99	116
デル	176	140	160
マツ	163	147	149
タッテ	172	140	163
モッテ	187	142	157
クルノ	200	166	176
テ'オアラウ	187	160	176

表 4-9 日本語 2 (全体)

文例	命令文	疑問文	平叙文
シメル	245	205	238
サゲル	205	191	205
ナオス	215	191	205
ハナシテ	226	205	220
アソブ	259	200	238
ウメル	304	245	259
アゲル	339	226	238
アケル	294	200	245
ワスレテ	352	205	252
デル	245	205	220
マツ	267	220	267
タッテ	259	152	169
モッテ	245	191	210
クルノ	275	196	200
テオアラウ	284	215	215

表 4-10 日本語 2 (語頭)

文例	命令文	疑問文	平叙文
シメル	275	210	245
サゲル	252	187	215
ナオス	238	180	226
ハナシテ	259	205	232
アソブ	200	169	187
ウメル	252	205	232
アゲル	232	215	226
アケル	226	163	205
ワスレテ	238	172	215
デル	315	275	294
マツ	294	267	284
タッテ	339	238	315
モッテ	352	259	304
クルノ	383	245	339
テオアラウ	326	294	339

## 5 考察

### 5.1 中国語

まず、文全体の解析結果を見ると、

- ①中国語 1 でも、中国語 2 でも、ほぼ、疑問文>命令文>平叙文、という順にピッチが低くなっていることが確認された。
- ②最も高い基本周数の中央値は、第 1 声で始まる文であった。
- ③最も低い基本周波数の中央値は、第 3 声で始まる文であった。

となった。一方、文頭第 1 音節のみを計測した結果からは、

- ①文全体の解析結果とは異なり、中国語 1 でも、中国語 2 でも、命令文>疑問文>平叙文、という順にピッチが低くなっていることが確認された。
- ②最も高い基本周波数の中央値は、第 4 声で始まる文であった。
- ③最も低い基本周波数の中央値は、第 3 声で始まる文であった。

となった。

このうち、最も高い基本周波数の中央値を示したものは文全体を対象とした解析結果にせよ、文頭第1音節のみを計測した結果にせよ②であり、一方、最も低い基本周波数の中央値は同じく③であった。このことは、②では初頭音節に高平らのピッチである第1声があったためにその後も高さが持続し、全体として総じて高くなったものと解釈することができる。同様にして、③はこの逆に初頭音節に低平らのピッチである第3声があったために、その後も低さが持続し、全体として総じて低くなったものと解釈することができる。つまり、これらはプロソディーにおける順行同化現象である。

次に、一般的傾向として平叙文が文全体でも文頭1音節でも常に最も低いピッチを実現し、疑問文と命令文が文全体では疑問>命令、文頭の1音節では命令>疑問となっていたのは、情動的意味 (emotional meaning) との高い相関性を反映しているためであろう。すなわち、疑問文では総じて高いピッチが持続するのに対し、命令文では局所的な高まりはあるものの、全体を通じての持続性に欠けるため、文全体を通して見ると相対的に疑問文よりも起伏に乏しくなるということにほかならない。

一方、文頭1音節のみを捉えた場合、上の順番が逆転して命令>疑問となるのは、そもそも命令文というものが文末位置に向けて顕著な下降を遂げなければならないプロソディック・パタンを要求されているため、このピッチ・パタンを最も効率よく実現するためには、始発部分をできるだけ高く実現し、文末との落差を大きくすることではないかと考えられる。

## 5.2 日本語

中国語と同じく、まずは文全体の解析結果を見ると、日本語1、日本語2ともに、ほぼ、命令文>平叙文>疑問文という順にピッチが低くなることが確認される。さらに、この順番は文頭第1音節のみを計測した結果でも変わらなかった。ただし、当然のことながら、文頭位置ではアクセントの影響により、いわゆるアクセント核のない「低く始まるグループ」と、アクセント核のある「高く始まるグループ」とでは明瞭な差異が生じたが、相対的に見れば、上に指摘した傾向は変わらない。

ここで興味深い事実は、文末の位置だけを捉えれば、確かに従来指摘されてきたように疑問文>平叙文>命令文の順にピッチが低くなるのだが、文全体を見ても、文頭の1音節を見ても、従来の「文末」という特定の位置以外では、命令文>平叙文>疑問文という、正反対の結果が得られたことである。その理由として、例えば疑問文では、一般的に「か」「ね」など文末にモダリティーをあらわす助詞がくるため、そこだけを極端に高くする必要のあることが考えられる。

そこで筆者は、よりハッキリとその高さを伝えるためには、始発部分のもとより、全体的に総じてピッチを低く抑えておき、ここぞというモダリティーの部分で一気にこれを上昇させるという、発音上の効率を最大限に高める方策が取られたためではないか、という解釈を施した。すなわち、落差を大きくつけることによって、聴覚情報処理効率を高めるという認知レベルにおけるメカニズムが存在するのではないかという指摘である。

一方、命令文では、先に述べた中国語の場合と同様に文末に向けて顕著な下降を遂げなければならないので、始発部分のもとより、文全体をできるだけ高めておくのが効率の良い発声の方策であるということになる。また、平叙文では文末はほぼゆるやかに下降するので、始発部分のもとより、全体的に見ても疑問や命令の中間ぐらいに抑えられている、と解釈することが可能になる。

さらに、本研究では分析資料の選別に際して、モダリティーを必ずしも文末の助詞など、書き言葉における形式面のみに限定せず、音声情報レベルにおけるプロソディーにも同等の重み付けを付与している。つまり、たとえば上に指摘した疑問文に顕著な文末終助詞に見られるピッチ上昇は明らかな事実だが、問題はそれだけにとどまらず、そのような文末における急激なピッチ上昇が長年日本語母語話者の記憶痕跡に蓄積されることによって、文法的単位としてのモダリティーを形式面で持たないような疑問文においても、日本語では文末の位置で顕著なピッチ上昇が見られるのではないかと考えられる。すなわち、別の表現を用いるならば、これは、「音声情報レベルにおけるモダリティーに相当する現象」ということである。ちなみに、おそらくはこの現象こそが、城生佰太郎 (1985) などで主張されている「音法論」の研究対象となるのではないかと思われる。

以上をまとめれば、

日本語では、命令文>平叙文>疑問文という順にピッチが低くなることから、これら3種類の文の弁別が文頭のわずか100-200ms程度の短時間の情報処理によっても行われていることが示唆された。

ということになる。なお、脳神経科学を援用した脳波実験によってプロソディー現象と取り組んだ先行研究に城生佰太郎 (1997, 2005) などがあるが、城生佰太郎 (2005: 448) には、

筆者は N1 頂点潜時に注目した結果、いずれの場合にもピッチ下降を伴う有核アクセントタイプの方が、ピッチ下降を伴わない無核アクセントタイプよりも、相対的に潜時が早いことを示した。(以下略)

との指摘がある。すなわち、プロソディーに対する情報処理では、N1 頂点潜時(脳波の一種である「N1」と呼ばれる ERP 成分波の反応時間)に着目した場合、相対的なピッチの高低差が重要な聴覚的因子として情報処理されている可能性が示唆されているということである。

従って、本稿における上の指摘もそのような視点、すなわち「相対的なピッチの高低差が重要な聴覚的因子として情報処理されている可能性が示唆されている」という観点からは、妥当な結論であろうと考える。ただし、これを実証するためには将来における脳波実験が不可欠であることは論をまたない<sup>2</sup>。

### 5.3 中国語と日本語における問題点

以上の結果から、互いに文頭第1音節に注目すれば、中国語も日本語も命令文・平叙文・疑問文のそれぞれにおいて、明瞭にピッチ制御が異なっているということが確認された。ただし、細部を見ると、日本語ではピッ

<sup>2</sup>聴覚実験音声学の現状に照らして、ここに指摘したような将来への展望は今後とも旺盛に行われなければならないものと考えている。また、本稿では必ずしも脳波実験設備を備えていなくても、大脳における聴覚情報処理系の営みに関する課題に迫りうる方法があることを実践したという点にも、筆者の主張がある。

チの高さが命令文>平叙文>疑問文の順であったのに対し、中国語では命令文>疑問文>平叙文の順であった。

この理由を、筆者は次のように考える。

- (1) 命令文が、中・日両言語で最も高くなっているのは、ともに文末位置に向けて顕著な下降を遂げなければならないという、プロソディー面での制約を共有しているためである。
- (2) 疑問文が日本語で顕著に低くなっているのは、文末にモダリティーをになった助詞が来るためである。一方、中国語でピッチが日本語と比べて高めに制御されているのは、一部の文末疑問詞は実在するものの、総じて日本語ほどには文末における一極集中型のモダリティーがないためである。

しかしながら、中国語と日本語のイントネーションを、互いのピッチに見られる最高値と最低値との落差という点に注目して見直すと、両者の間にはさらに微妙な差異が見える。なお、このような問題はイントネーションの類型論で論じられることが一般的である。一例として、亀井孝・河野六郎・千野栄一 (1996: 83) によれば、イントネーションに関する類型論で知られる Cruttenden (1986) は、世界の言語を、(1)イントネーション言語、(2) 声調言語、(3) 高さアクセント言語の3種類に分類しており、英語、ドイツ語などヨーロッパの多くの言語は(1)、中国語、タイ語などアジアに分布する音節高さアクセント言語は(2)、日本語のような単語高さアクセント言語は(3)の、それぞれ典型的な例としている。

さらに同書では、(2)や(3)に属す言語は、イントネーションとは別に声調や高さアクセントなどが存在するために、イントネーションによる基本周波数の最高値と最低値の落差がある程度制限されるのに対し、(1)ではイントネーションにおける高さが、かなり大幅に変動することを許すとしている。つまり、(2)や(3)の声調言語や高さアクセント言語におけるイントネーションは、(1)のイントネーション言語に比べて、高低の落差が抑えられており、余り大きくは変動しないということである。

この指摘を、本研究における中国語と日本語で検証してみると、次に示す表 5-1 のようになる。なお、表中カッコ内の数値は周波数の高低落差を示す。

表 5-1 中国語と日本語におけるイントネーションの最高値と最低値  
(基本周波数の比較、単位は Hz)

中国語 1、全体	203-162 (41)	196-141 (55)	208-173 (35)
中国語 2、全体	279-190 (89)	253-170 (83)	304-217 (87)
中国語 1、語頭	254-153 (101)	217-111 (106)	230-142 (88)
中国語 2、語頭	308-190 (118)	257-132 (125)	300-185 (115)
日本語 1、全体	304-152 (152)	259-140 (119)	196-124 (72)
日本語 2、全体	352-205 (147)	267-169 (98)	245-152 (93)
日本語 1、語頭	200-113 (87)	176-110 (66)	166-95 (71)
日本語 2、語頭	383-200 (183)	339-187 (152)	294-163 (131)

上の表から明らかなように、中国語と日本語を対照させると、Cruttenden が「高低の落差が抑えられている」とした声調言語とアクセント言語との間にさえ、かなりの高低落差が認められる事例が実在することがわかる。すなわち、イントネーション類型に関しては今後とも音声学的レベルにおける更なる綿密な研究が不可欠であるということにほかならない。

## 6. 結論

以上に指摘したように、言語ごとに見られる個別のパラメータが存在するが、これを差し引くと、残る解は「平叙文・疑問文・命令文は、それぞれに出だしの部分のピッチが互いに異なる」ということである。この事実こそが、本稿で最も強調したかった部分である。

すなわち、われわれは会話において、音声情報を最後まで聞くことなしに、はじまりのほんの 100-200ms 程度の情報を手がかりとしてさえ、ある程度は結末を推論することが出来る。つまりは、発話の初頭部から、すで

に脳のトップダウン機能に依存したイントネーションの聞き分けを行なっている可能性が示唆された、ということである。

### 【参考文献】

- Cruttenden, A. (1986) *Intonation*. Cambridge: Cambridge University Press.
- 福岡昌子 (1998) 「イントネーションから表現意図を識別する能力の習得研究：中国 4 方言話者を対象に自然・合成音声を使って」『日本語教育』96: 37-48.
- Greenberg, J. H. (1963) *The languages of Africa*. Bloomington: Indiana University.
- 城生佰太郎 (1985) 『当節おもしろ言語学』講談社。
- (1997) 『実験音声学研究』勉誠社。
- (1998) 『日本語音声学』サン・エデュケーショナル。
- (2001) 「音声研究の方法」『コンピュータ音声学』9-45. おうふう。
- (2005) 『日本音声学研究：実験音声学方法論考』勉誠出版。
- 亀井孝・河野六郎・千野栄一 (1996) 『言語学大辞典第6巻 術語編』三省堂。
- 川上 稔 (1956) 「文頭のイントネーション」『国語学』25: 21-30.
- (1961) 「ことばの切れ目と音調」『国学院雑誌』62-5: 34-42.
- 上野善道 (1989) 「日本語のアクセント」『講座日本語と日本語教育 2 日本語の音声・音韻(上)』178-205. 明治書院。
- Yokosawa, K. (1998) ‘Attentional control setting for facilitation of pop out perception,’ *Selection and Integration of Visual Information* 40: 45-53.

# A Contrastive Study on the Chinese and Japanese Intonation with Special Attention to the Top-down Processing of the Cerebrum

LU Jia

I recorded the voice of four individuals, a Chinese male and female and a Japanese male and female, while they read aloud sentences in their own native tongue with three kinds of intonation patterns: assertive, interrogative and imperative. For each sentence, I measured the median pitch of the whole sentence as well as that of the first syllable using Multi-Speech 3700, ver. 2.4.

As a result, distinct pitch patterns for the three sentence types mentioned above were detected within the 100-200 ms from the beginning of each sentence. Both in Chinese and Japanese, the pitch of the first syllable of the sentence was the highest in the imperative sentences, and was the lowest in the interrogative ones. That of the assertive sentences falls somewhere in between.

This fact can be interpreted as follows:

- (1) The imperative sentence has a rapidly falling intonation toward the end, so its pitch needs to start out relatively high at the beginning.
- (2) The interrogative sentence has an abruptly rising intonation toward the end, so its pitch needs to start out relatively low at the beginning.
- (3) The assertive sentence has a gradually falling intonation toward the end, so its pitch tends to stay constantly somewhere in between the other two.

*Doctoral Program in Literature and Linguistics*  
*University of Tsukuba*  
*1-1-1 Tennodai, Tsukuba; Ibaraki 305-8571, Japan*  
*E-mail: [lujia1979@yahoo.co.jp](mailto:lujia1979@yahoo.co.jp)*