

熊本方言イントネーションの知覚

国立国語研究所 前川 喜久雄 熊本短期大学 吉岡 泰夫

PERCEPTION OF KUMAMOTO INTONATION

MAEKAWA Kikuo(NLRI) and YOSHIOKA Yasuo(Kumamoto Junior College)

Abstract: Perception experiment was conducted in order to test the validity of an intonation model of Kumamoto accentless Japanese, dialect which has no lexical accent. Naturalness of 20 synthesized F0 contours, including both grammatical and agrammatical ones, were judged by 29 native speakers of the dialect. Contours conforming to the model were judged to be much more natural than those violating the model. The result revealed the overall validity of the model.

1 熊本イントネーションのモデル

無アクセント方言である熊本方言には、語レベルでのピッチ指定は存在しないが、文ないしは談話レベルでのピッチ制御は存在する。疑問詞/単純疑問文の差異、名詞句内部ならびに補語・述語間の修飾関係の差異等は熊本方言においても東京方言などの有アクセント方言と同様にイントネーション(ピッチ形状)の差異に組織的に反映される[1-3, 7, 8]。また談話上の各種要因がイントネーションにあたる影響も報告されている[5]。

Fig. 1 にしめすのは熊本方言の疑問詞疑問文「ナンノミユットネ」(=何がみえるの?)と単純疑問文「ナンカミユットネ」(=何かみえるの?)のイントネーションである。こうしたイントネーションは、Fig. 2 にしめすような抽象的な韻律構造から一群の音声実現規則によって生成されるものとかがえられる[6]。

この表示は2段階の階層構造をなしている。

あらゆる発話はただひとつの Ut (utterance) によって支配される。Ut を構成するのは複数個の Ip (intonational phrase) である。(東京などの有アクセント方言では Ip のしたに語彙アクセントを指定するための階層が必要になるが無アクセント方言ではそうならない。) Ut と Ip は音韻論的トーンを導入する。トーンは H か L かのいずれかの値をとる。Ut の左には L% が、右には L% か H% かのいずれかが導入され、Ip の左右には H- L- がそれぞれ導入される。(ここで % はトーンの所属をしめす補助記号である。) その結果、Fig. 2 A では L% H- L% というトーンの連鎖が生じている。Ip 所属のトーンは談話上の要請によって削除されることがあり、Fig. 2 A では L- が削除されている。H- も削除されることがあるが、それは二番目以降の Ip においてであり、発話始端の Ip の H- はかならず実現されるようである[4]。

Fig.1 F0 contours of Kumamoto WH(panel A) and YES-NO(panel B) questions. Frequency scale is in logarithmic.

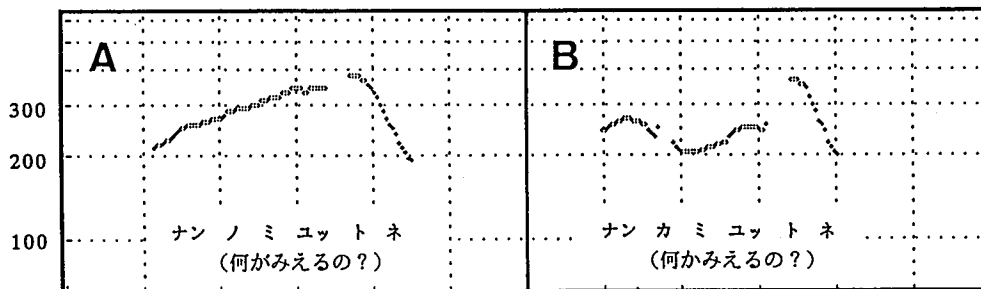
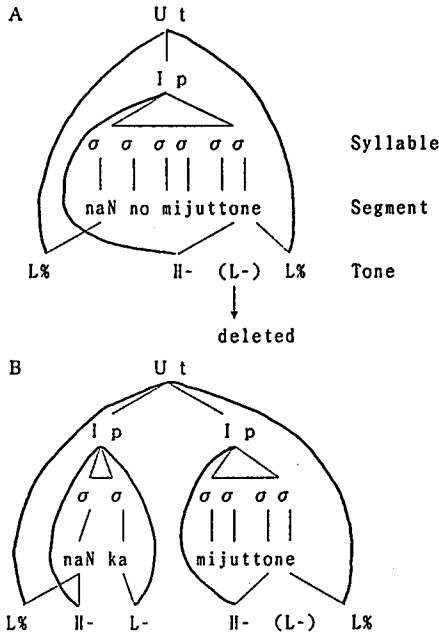


Fig. 2 Prosodic Structures of Kumamoto WH (panel A) and YES-NO (panel B) questions



音韻論的モデルの妥当性は記号レベルのみにとどまらず実験的・定量的にも検討されることがのぞましい。イントネーションの場合、一定の限界内では物理的関連量を基本周波数だけにかぎることができるので、LPC分析/合成の手法を利用することによって、比較的簡単に実験的検討をおこなうことができる。今回は分節音に関与するスペクトル構造や持続時間を一定にたもったまま、基本周波数パターンだけを上記モデルにもとずいて変化させた合成音を作成し、その自然性に関する知覚実験をおこなった。合成の素材はできるだけ単純な韻律構造をもつことがのぞましいので、発話全体がただひとつのIpから構成されるとかんがえられる疑問詞疑問文「ナンノミュ」(=何が見える?)をえらんだ。イントネーションモデルの構築段階で中心的な資料として利用した20代男性話者1名の発話の実測値を参考にして、各トーンの実現値を以下のように設定した。

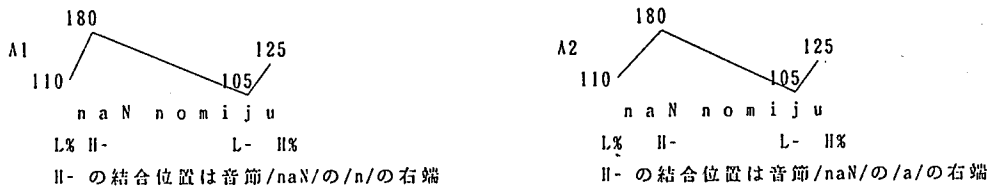
Ut左の L%=110Hz
 H-=180Hz
 L-=105Hz
 H%=230Hz
 Ut右の L%=70Hz

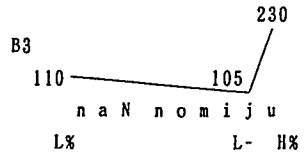
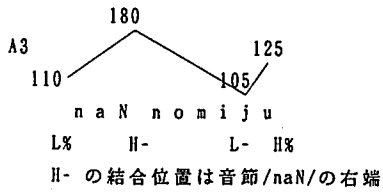
各トーンがテキストのどの位置に結合されるかはトーン結合規則で決定される。熊本方言のトーン結合規則を以下に示す。

- 1) Ut 左の L% は発話冒頭の音節に結合される。
- 2) Ut 右の L% ないし H% は発話の最終音節に結合される。
- 3) Ip 左の H- は所属する Ip 内部の任意の音節に結合される。
- 4) Ip 右の L- は所属する Ip の末尾音節に結合される。

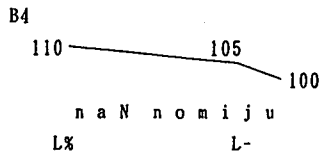
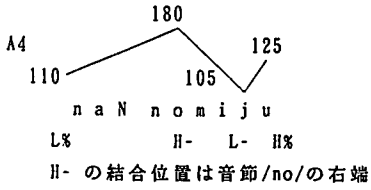
2 イントネーションの合成

Fig. 3 知覚実験用イントネーションの構造 (図中の数字は周波数値 [Hz])

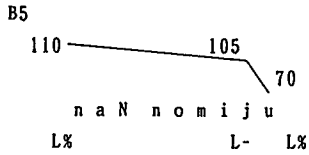
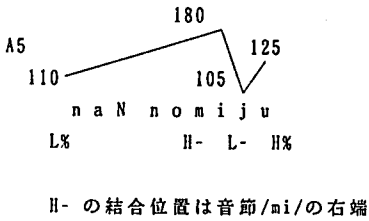




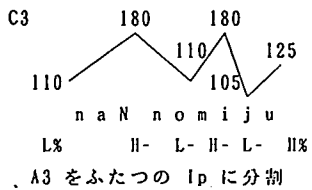
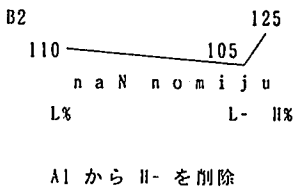
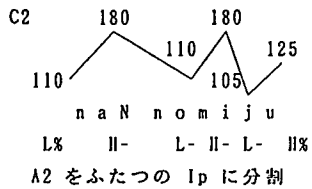
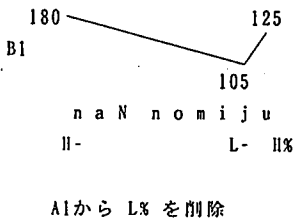
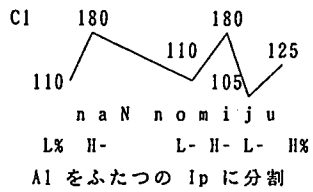
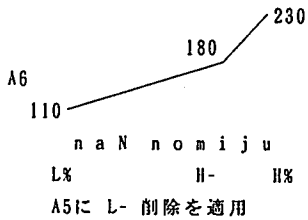
B2 の H% を 230Hz に変更



A1 から H- を削除し Ut 右端の H% も削除



A1 から H- を削除し Ut 右端に L% を結合



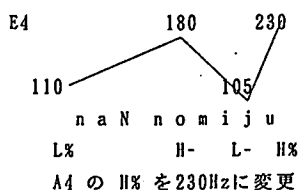
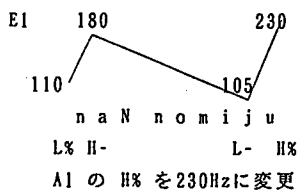
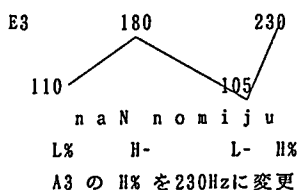
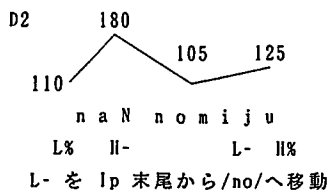
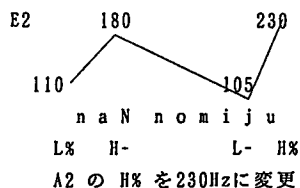
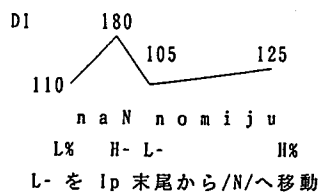


Fig. 3 にしめした 20 種類のイントネーションのうち、A1～A6 だけがモデルによって生成される「自然」なイントネーションである。A1～A5 では H- の位置が浮動しているが、これはトーン結合規則によって許容される変異である。A6 では A5 に L- 削除規則が適用されている。カテゴリ B～E のイントネーションにはいずれも何らかの点で「自然」ではない。カテゴリ B では削除規則の対象となりえないトーンが削除されている。カテゴリ C のイントネーション構造そのものはモデルによって生成されるものである。しかし、さきにのべたように今回のテキストは統語上の特徴から全体がひとつの lp をなす必要があるので、統語構造と韻律構造間の写像に異常がある。カテゴリ D では L- が lp 末尾音節に結合されていない。最後にカテゴリ E では H-L- のうしろにある H% が低下していない。これは音韻表示ではなく音声実現規則レベルでの異常である。

合成音の作成にあたっては今川博・桐谷滋両氏の開発された PANASYS プログラムを利用させていただいた。合成の基礎となった音声は概略 A4 のイントネーションで発話されたものであった。

3 知覚実験

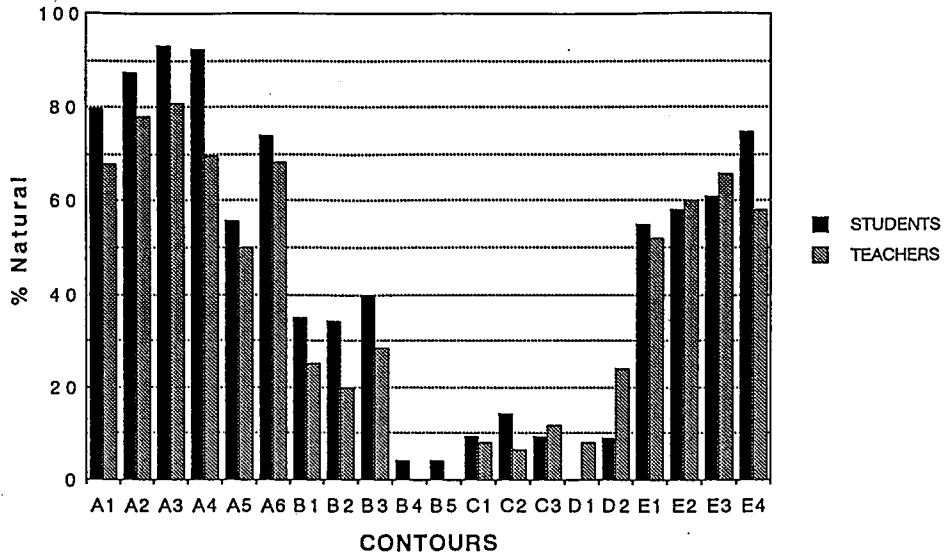
知覚実験では 28 種類の合成音が 3 秒間隔で

ランダムに 5～11 回録音されたテープを再生し、被験者に各イントネーションが、被験者が日常使用している熊本方言の疑問詞疑問文として「自然」か「不自然」かの強制判定を要求した。被験者となったのは熊本市および同市周辺の無アクセント地域出身の高校生 15 名・短大生 4 名（以下ではこのグループを「学生」とよぶ）と 50 代を中心とする高校教員 10 名（以下「教師」）である。

3.1 全体的傾向

Fig. 4 に知覚実験の結果をしめす。学生の反応と教師の反応をそれぞれプールして集計した。縦軸は各イントネーションが「自然」と判定された比率を百分率で表示している。Fig. 4 ではカテゴリ A とそれ以外のカテゴリのあいだにはあきらかな自然性の差が存在することがわかる。（ただし A5, A6 の自然性は、同カテゴリの他のメンバーにくらべるとひくい。3.3.3.4 参照。）これは想定したイントネーションモデルの基本的な妥当性をしめす結果といってよい。また学生と教師の反応とは全体的によく一致している（学生の反応を y, 教師の反応を x とし計算した回帰直線は $y = 4.578 + 0.890x$ となる）。

Fig.4 Result of Perception Test: Naturality Judgment by Students and Teachers



以前発話データの分析結果を口答発表した際に、筆者らの分析しているイントネーションは若年の個人に特有のもので、熊本方言全般の特徴とはみとめられないとの意見があったが、今回のデータにみられる学生と教師の一致は、この意見に対する反証となるものである。

A以外の「不自然」なカテゴリを検討すると、E > B > C > Dの順に自然性が低下しており、「不自然」さにもさまざまな水準が存在することをしめしている。Eの自然性が相対的にたかいことは、音声規則レベルの異常が音韻表示の異常ほどは自然性を低下させないことを示唆している。不適切なトーン削除がおこなわれているカテゴリBのなかでは、発話末にH%が存在するB1~B3の自然性がかなりたかい。これは異常な音韻表示ではあっても、発話末の上昇が疑問文らしさに貢献するためであろう。カテゴリC, Dはともに韻律構造のフレージングに異常が存在するケースであるが、この種の異常は自然性を最低のレベルにまで低下させる。なかでも学生によるD1の自然性判定は0%になっているが、実験後数名の被験者がもらした感想によればD1は「何ノ見ユ？」よりも「ナンノミヨル？」(=「何飲ミヨル？」)にちかくきこえたため、確実に「不自然」と判定

することができたとのことである。

3.2 トーンの実現位置

A1~3とE1~3は発話末のH%の実現値以外は同一のイントネーションであり、音節/naN/内でのH-の実現位置のみが相違している。Fig.4ではH-の実現位置によって両カテゴリに共通した変動が生じている。つまりH-が音節/naN/の右端に結合された場合(A3, E3)に自然性が最高値をとる。これは熊本方言の音声実現規則におけるトーン実現の正準的位置が音節の右端にあることを示唆している(この現象は今回報告の対象からのぞいたイントネーションにおいても確認することができた)。その意味では、A1, 2も完全に「自然」なイントネーションではありえないわけであるが、トーン実現位置の決定は音声実現規則レベルの問題とかがえられるので、規則違反による自然性の低下もすくないものとおもわれる。

3.3 A6の問題

適格なイントネーションであることが予想されるカテゴリAのなかで、A5とA6だけは、相対的に自然性が低下している。これらのイントネーションに対しては反応の分散がおおきい。つま

り、これらを一貫して「自然」と判定する被験者と反対に一貫して「不自然」と判定する被験者とが同時に存在している。

Table 1 には両イントネーションへの「自然」反応の分布を83%以上(6回中5回)以上、17%以下(6回中1回以下)とその中間にわけて年代差とともにしめた。各セルの数字は被験者の人数をしめしている。A6の場合、被験者の6割以上は83%以上「自然」と判定しているのだが、小数の被験者が平均値の「足をひっぱる」結果となっていることがわかる。この現象のくわしい解釈は今後の分析にかかっているが、ひとつの可能性として、このイントネーションの生成に關与するL- 削除規則をもつ個人とそうでない個人とが存在することがかんがえられる。

Table 1 Distribution of "Natural" judgment for Contours A5 and A6

		>83% 中間 17%>		
合計 Total	A 5	7	1 7	5
	A 6	1 8	7	4
学生 Stu- dents	A 5	5	1 1	3
	A 6	1 1	6	2
教師 Teach- ers	A 5	2	6	2
	A 6	7	1	2

3.4 A5の問題

一方、A5に対する反応では被験者の半数程度は「中間」の反応をしめしており、分布がA6とはことなっている。またA5の場合、このイントネーションだけに關与する規則は存在しないので、個人による規則の有無を論じることはできない。それでは何故、現実の発話中には実際に観察されるイントネーションであるA5の自然性がひくく判定される傾向が存在するのだろうか。その原因としては相互に關連するふたつの可能性がかんがえられる。

まず第一にかんがえられるのが、フォーカス

との關係である。熊本イントネーションではH-の位置は音韻論だけの情報では決定することができない。これは前川(1991a)の段階では未解明の問題とされていたのだが、その後出版された吉岡・都築(1991)では、H-の位置が発話のフォーカスに対応する例が報告されている。動詞「ミュ」の冒頭にH-が結合されたイントネーションであるA5が文の述語にフォーカスがおかれているとの解釈をうけるならば、今回の実験のように特に文脈が指定されていない実験ではA5を疑問詞疑問文イントネーションとして不適格と判断する被験者がいても不思議ではない。

第二の可能性は、音声実現規則レベルの問題である。今回の合成音ではテキスト中の位置によらずH-の値を180Hzに固定しているが、A5ではH-が発話の末尾ちかくに位置するために、知覚における自然下降の補償を考慮するとH-がA1~A3等にくらべて知覚上きわだかって知覚された可能性がある。(その場合H-のきわだちがフォーカスの存在を知覚させる結果となったであろうと想像される。)

A5の自然性のひくさはさきに提案したトーン結合規則3)の妥当性に直接關係するので、今後の検討が必要である。(なお、カテゴリA・Eともに音節/no/の右端にH-が結合されたイントネーション(A4, E4)は、/naN/右端のイントネーションと同程度のたかい自然性を有する。これはトーン結合規則3)の妥当性をしめしていることに注意。)

4 被験者の分類

ここまでの議論は実験に供したイントネーション個々の性質に關するものであった。しかしA5・A6両イントネーションの分析は、今回の被験者がかならずしも均一な反応をしめしていないことをあきらかにした。以下では被験者相互間の相関關係を簡単に分析する。

Fig. 5はイントネーションA1~A6に対する被験者29名の相関行列をもとに主成分分析を実行し、その第1・第2主成分得点によって被験者を分類した結果である。ふたつの主成分の累積寄与率は約68%、第3以降の主成分の固有値は1に達しない。第1主成分得点に対応する横軸は、主成分分析の結果がしばしばそうなるように、6種類のイントネーション全体に対する自然性判定の強度を反映している。すなわち軸の右側にはイ

ントネーションに対する自然性判定の平均値がたかい被験者が位置し、左側には平均値のひくい被験者が位置している。Fig. 6にはA1～A6に対する反応の平均値と第1主成分得点の関係を個人ごとにプロットした。Fig. 5の第1軸で興味ぶかいは被験者33を例外として、軸の左側に位置する被験者がすべて教員(7,13,14,15)であることである。さらにこれら4名の教員全員が熊本市以外(鹿本郡ないし菊池市)の出身であることも注目される。Fig. 5第1軸上の分布は熊本イントネーションの地域差の存在を示唆しているのかもしれない。(ただしすべての鹿本郡/菊池市出身者が軸の左に位置するわけではないことに注意。)

Fig. 5の第2主成分得点(縦軸)は、被験者がA5・A6の両イントネーションをともに自然と判定するかどうかに関係していると解釈される。すなわち軸の上方に位置する被験者はA5・A6をともに不自然と判定しており、反対に軸の下方に位置する被験者はA5・A6をともに自然と判定している。この軸においても教員が両極をしめ

ていることが注目される。上方に位置する2,9,下方に位置する3,15はすべて教員である。Fig. 7は第2主成分得点とA5・A6への反応の平均値の関係を教員に関してだけプロットしたものである。

Fig. 5は全体としてFig. 4からはよみとれなかった学生と教員の反応の相違をあきらかにしている。学生被験者は反応の均一性がたかく、カテゴリAのイントネーションには総じてたかい自然性判定をあたえている。それに対して教員被験者のなかには、適格なイントネーションであるカテゴリAのメンバに対しても自然性をひくめに判定するものがあり、またA5・A6に関しては両イントネーションをともに自然と判定するものと、ともに不自然と判定するものとが比較的明瞭にわかれて存在する。このような教員(50代)被験者の特異性が何を意味しているのかについては、十分な考察ができていない。今後、社会言語学的な観点からの分析が必要である。

Fig.5 Subjects Classified by Factor Scores of Principal Component Analysis

Outlined digits stand for teachers. Underlined digits stand for the subjects from the surrounding area of the city of Kumamoto: Kamoto district and the city of Kikuchi.

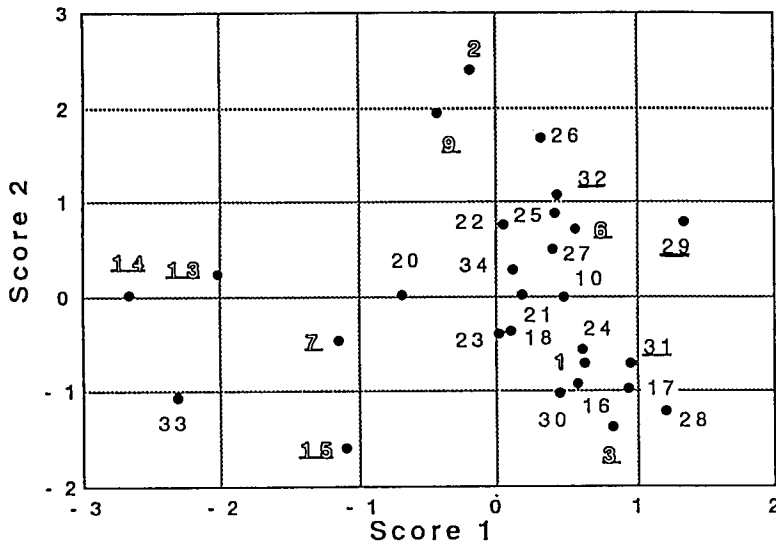


Fig.6 Correlation between Score 1 and mean naturality of A1-A6 (All subjects' data)

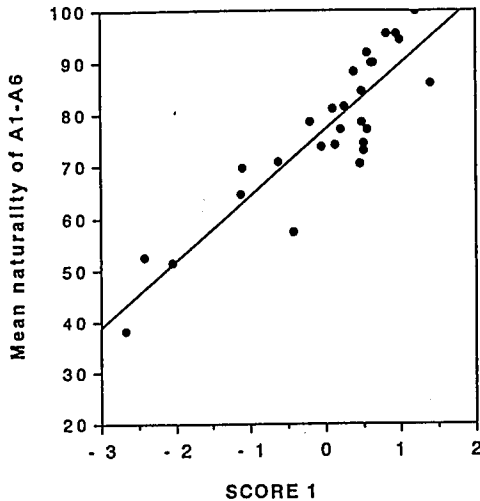
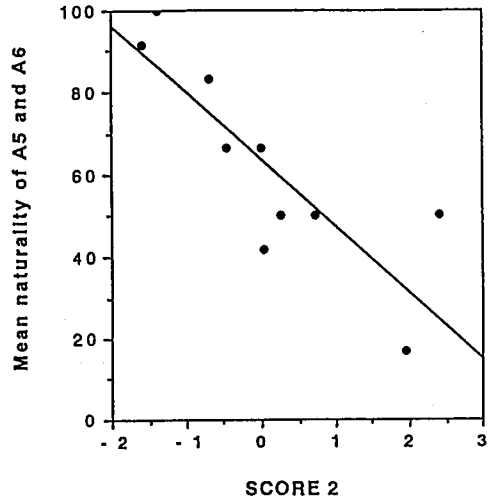


Fig.7 Correlation between Score 2 and mean naturality of A5 and A6 (Teachers' data)



5 まとめ

無アクセント方言である熊本方言を対象に構築された音韻論的イントネーションモデルの妥当性を10代および50代の方言話者を被験者とした合成音の知覚実験によって検討した。モデルによって許容されるイントネーション群と許容されないイントネーション群の間には自然性の判定において明瞭な差が観察され、モデルの基本的な妥当性があきらかになった。しかし一部のイントネーションに関しては、被験者の反応が均一でないこともわかった。主成分分析法によって被験者間の相関を分析してみると、上述の不均一はもっぱら50代被験者によってうみだされていることが判明した。

謝辞：本研究は科研費重点領域研究「日本語音声」（課題番号03208114）の補助をえて遂行されたものです。杉藤代表はじめ研究のさまざまな段階で助言をいただいた方々に感謝いたします。また発話資料の収録と知覚実験に御協力いただいた熊本の皆様は心より御礼申し上げます。

参考文献

- [1] 前川喜久雄「無アクセント方言のイントネーション(試験)」, 音声言語IV, 1990.
- [2] 前川喜久雄「無アクセント方言の音調」, 重点領域「日本語音声」シンポジウム「アクセントからイントネーションへ」配付資料, 1991.
- [3] 前川喜久雄「東京方言疑問文のイントネーション」, 平成3年度日本音声学会全国大会発表論集, 1991.
- [4] 前川喜久雄「熊本無アクセント方言のイントネーション」, 月刊言語, 21-9, 1992.
- [5] 吉岡泰夫・都築直也「無型アクセント地域におけるイントネーション」, 「日本語音声」研究成果中間報告書, 1991.
- [6] Pierrehumbert, J. B. and M. Beckman, *Japanese Tone Structure*, MIT Press, 1988.
- [7] Kiritani, S., H. Hirose, K. Maekawa & T. Sato, "Electromyographic studies on the production of pitch contours in accentless dialects in Japanese," *Annual Bulletin RILP*, U. Tokyo, 25, 1991.
- [8] Maekawa, K., "Perception of intonational characteristics of WH and NON-WH questions in Tokyo Japanese," in *Proc. of 12th ICPhS*, Aix-en-Provence, France, 1991.