

乳幼児の歌唱様音声の韻律的特徴

坂井 康子・志村 洋子¹⁾・山根 直人²⁾・岡林 典子³⁾

Prosodic Feature of Song-like Voices of Infants

SAKAI Yasuko, SHIMURA Yoko, YAMANE Naoto and OKABAYASHI Noriko

Abstract : We have been studying the prosodic feature of infant voices in order to clarify the way Japanese children express singing elements in their voices. In our previous thesis (Sakai et al. 2013) we carried out a listening experiment by presenting 88 types of three-tone voices to the adults, then asked them to assign a score for each of the three-tone voices on a perceived musicality scale that we have provided. The data obtained from this experiment pointed to three general tendencies. First, the adults tended to assign higher scores to the three-tone voices with longer sounds. Secondly, if the last sound in the three-tone voices were the highest it was perceived as more musical. Thirdly, although the last tendency was not as clearly visible as the first two, the musicality score also tended to depend on the overall pitch – the higher it was, the more points it received. Our current thesis elaborates on our previous findings in order to confirm our results by refining the listening experiment. By selecting 5 of the previous 88 types of the three-tone voices and synthesizing them into various types of lengths, pitches and intonation. The results confirmed three of the tendencies observed in our initial experiment. Longer sounds scored considerably higher than shorter ones. Overall pitch influences the score as well ; higher three-tone voices received better scores. Finally, the three-tone voices where the last sound was the highest also scored higher. Our previous experiment showed only a general tendency, however this current experiment gave us much clearer and more detailed results.

要旨：これまで筆者らは、日本語環境にある乳幼児が歌唱様の音声をどのように表出しているかを明らかにするために乳幼児音声の韻律的特徴について検討してきた。前稿（坂井他 2013）において我々は、歌唱様音声の韻律的特徴を明らかにするために、乳幼児の 88 種の 3 音等時音声について歌っていると知覚される度合い（以下「うた度」と呼ぶ）に関する聴取テストをおこない、うた度が高いと評価された音声の特徴として、「音声長が比較的長い」「音声の末尾が上昇しているものが比較的多い」傾向を検出した。また、有意差は認められなかったが、うた度の高い音声はピッチの最高値が高い傾向がみられた。そこで本稿では、前稿での結果についてさらに検討を加えるために、前稿で用いた音声のうち 5 パターンの音声について「長さ」「高さ」「抑揚」の三つの要素を変化させた加工音声を作成し、聴取実験をおこなった。その結果、短い音声より長い音声、また低い音声よりは高い音声、そして音声の末尾が上昇している音声、歌らしく聞き取られており、これらは前稿の結果を裏付ける結果であった。

1. 本稿までの経緯

日本語の獲得過程にある 1 歳前後の子どもの喃語や日本語の音声表現には、3 音ひとまとまりの音声特徴的にみられることが多くの研究者によって指摘されている(伊藤 1978/永田 1981/志村 1991/南 1991/岡林・坂井 2007/坂井・岡林・佐野 2008)。また、幼児期の発話において、3 モーラの音声が多多いということも明らかにされている(市島 2004, 益子 2004)。これらのことから、筆者らは乳幼児音声の 3 音ひとまとまりの音声について分類をおこない、乳幼児の音声のリズムや抑揚に関する研究をおこなってきた(坂井他 2012, 2013)。これまで筆者らがおこなってきた喃語を中心とした乳幼児音声の研究において、以下の点が検出されている。

1.1. 等時音声の特徴

8 ヶ月・12 ヶ月・17 ヶ月齢児の、3 音が比較的同じ長さ(以下等時とする)のひとまとまりの音声⁴⁾を比較した結果、短い 3 音等時音声は 17 ヶ月齢において増加する傾向がみられた。このことにより、短い 3 音等時音声は言語とつながっているという可能性が示唆された。次に、1 音ずつに特殊拍を伴う 3 音等時音声(重音節)の実態について明らかにするために、特殊拍部分の分類をおこなった。その結果、特殊拍の出現頻度としては、促音の 3 連続(例: たったった)、長音の 3 連続(例: たーたーたー)が高く、出現上位は促音と長音という一定の特徴がみられた。口形の変化を必要としない促音と長音による延音の頻度が高いことは、生理学的に自然であると思われる。

これまでの分析結果から筆者らは、喃語の音声長の変化について「まず喃語期初期において乳児は、重音節様の音声(各音に特殊拍を伴う)を発声しやすかったが、しだいに軽音節様の音を発声することができるようになり、その延長上において様々な単語を発話できるようになる」ととらえた。3 音が重音節の音声については、ことばとなる(例: アンパンマン)音声のほか、音声を引き延ばすことによって「歌っている」音声も含まれる可能性がある。しかし、喃語の音声を「ことば」と「うた」に分類するためには、何をもって「うた」ととらえるかということを明らかにする必要がある。

1.2. うた度の高い音声の特徴

喃語期の音節を「ことば」か「うた」かに分類するための試みとして、「歌っているように聞こえる音声」はどのような韻律的特徴を持つかを調べることにした。88 パターンの乳幼児の 3 音等時音声⁴⁾(音声の状態が良好な 8 ヶ月・12 ヶ月・17 ヶ月齢児の 3 音等時音声)を用いて「うた度」を測るための聴取テストをおこなった。うた度の高い上位 10 位までの音声とうた度の低い下位 10 位の音声を比較した結果、うた度の高い音声では「音の長さが比較的長い」「音声の末尾が上昇しているものが比較的多い」ことが明らかになり、また前者はピッチが高めでピッチレンジが広い傾向が認められた。

「坂井(2008)」でおこなった聴取調査においても、音声が定常であること、引き延ばされていることに加え、最後の音の上昇が「うた」と聞き取られる特徴として上げられており、これらの点についてさらに検証することによって、「うた」と聞き取られる音声の韻律的特徴が明らかにされると考えられる。

2. 加工音声による聴取実験

2.1. 作成した加工音声

本稿では、前稿での聴取テストの結果についてさらに検討を加えるため、前稿で用いた 88 パターンの音声のうち、うた度評定が 1 位であった「ヘーヤーアー」、2 位の「ギャッタッタ」、8 位の「シーシーオー」、11 位の「コークーオー」23 位の「シーイーエー」の 5 パターンの音声について、韻律の要素(実験条件)を変化させた加工音声を作成した。加工音声作成の方法の詳細は以下のとおりである。

今回音声を加工するために「CeVIO」を用いた。他にも「VOCALOID」など人間の声を加工できるものはいくつかあるが、「VOCALOID」は歌を歌うために作られており「歌っているように聞こえる音声」かどうかを判断する本実験ではあまり適していないと判断した。「CeVIO」は音程・長さ・速さ・発音などを変更することが可能であり、設定次第では人間が話しているときにかかなり近い加工音声を作成できるため、本実験で用いることとした。

加工音声の作成においては、まず乳幼児の 3 音等時音声 5 パターンの音声を「CeVIO」を用いて再現した(この音声をオリジナル加工音声と呼ぶ)。なおオリジナル加工音声は 3 音等時音声でなくならないように注意し、また音の長さはオリジナル加工音声を

作る前の元の音声（生音声と呼ぶ）における間隙を変化させないように注意して作成した。生音声の有声区間が滑らかにつながっている場合は加工後も滑らかにつながるように加工を行い、逆に生音声において間隙が顕著な音声については、その間隙ができるだけ保存されるようにした。こうして作成したオリジナル加工音声を基に、音の長さ（音の速さ）、全体の音の高さ、抑揚（音声の末尾の高低）の三つの要素を実験条件として変化させた。

まず音の長さ（音の速さ）を変化させ、オリジナル加工音声と、それより速いもの、遅いものの3水準を作成した。変化の度合いは、オリジナル加工音声を1とすると、速いものは音の長さが約0.8倍と短く、遅いものは音の長さが約1.2倍と長くなるように設定した。

次に全体の音の高さは3音全ての高さを同程度変化させるために、「CeVIO」によって作成したオリジナル加工音声を「聞々ハヤえもん」によって変化させた。5パターン全ての音声において印象が異なる音声とするために、音の長さを維持したまま音声全体の高さを3半音程度低くすることによってオリジナル加工音声と低めの音声の2水準を作成した。

最後に抑揚について、3音等時音声の2音目を基準とした3音目の音高が高い場合、同じ場合、低い場合の3水準を作成した。まず2音目と3音目が同じ音高の場合を作成し、その3音目とオリジナル加工音声の3音目の音程の幅（高くなる場合と低くなる場合がある）が同じになるように3水準目の音声を作成する。つまり、3音目の音高が高い場合と低い場合の音声は2音目の音高から同じ幅で変化するように設定した。

以上により合計90の加工音声を作成した。実験条件と水準を表1にまとめる。

5パターン×「長さ3水準×高さ2水準×抑揚3水準」の計90音声为本稿の聴取実験用音声である。論文末尾に1例（「シーシーオー」）、生音声（図2）と加工音声（図3）のピッチデータを掲載する。

2.2. 聴取実験の方法

作成した90音声を用いておこなった聴取実験の被験者は、関東出身者23名と、関西出身者22名、計45名である。両者の実験結果が近似していたため、両者の評定結果を統合して提示し、考察をおこなう。

まず実験においては、回答用紙の初めに記述した以

下の文章を読み上げた。

回答用紙の番号にあわせ、その声が「赤ちゃんが歌っているよう」に聞こえたかどうか、1「まったく聞こえない」から6「とても聞こえる」までの、6段階のうち一つの+に必ず○を付けてください。

45名の被験者に、ランダム配列した90音声を3回ずつ聴取させ、それぞれうた度を6段階で記入させた。

また実験終了後に「歌っているように聞こえた音声の特徴を書いてください」、および「聴取実験の印象を書いてください」と依頼した。

3. 聴取実験の結果と考察

3.1. 各パターンの結果と考察

聴取実験の結果と考察をパターンごとに記す（表2）。

3.1.1. 「ヘーヤーアー」

「ヘーヤーアー」は今回の評定で5パターン中うた度が最も高かった音声（1位音声の評定値4.8）である。「ヘーヤーアー」は前稿のうた度評定でも1位（前回実験の評定値5.4）であった。オリジナル加工音声は末尾の音が高い抑揚の音声で、今回の実験では18位中4位である。今回の評定で1位の音声はオリジナル加工音声と長さが長いことだけが異なっており、2位は長さが長いことと、高さが低いことが異なっていた。3位は長さが短いことだけが異なっていた。これらから、長さが長いこと、高さが高いこと、末尾の音が高いことはうた度が高いことと関係がある可能性が認められる。

3.1.2. 「ギャッタッタ」

「ギャッタッタ」（1位音声の評定値3.7）は前稿のうた度評定で2位（前回実験の評定値5.2）だった音声である。オリジナル加工音声は今回の実験で1位となっており、続く2位は長さが長いだけが異なるもの、3位は長さが短いだけが異なるものである。これらから、長さが長いこと、高さが高いこと、末尾が高いことはうた度が高いことと関係がある可能性が認められる。

3.1.3. 「シーシーオー」

「シーシーオー」（1位音声の評定値4.0）は前稿のうた度評定で8位（前回実験の評定値4.9）だった音声である。この音声はオリジナル加工音声は今回3位で、同率1位がもとの音声と比べて長さだけ長いもの、同じく短いもの、同率3位は長さが長く、高さが低いものであった。「シーシーオー」の場合、うた度上

表 1 加工音声の実験条件と水準 (網掛けはオリジナル加工音声の水準)

長さ	短い (0.8 倍) : 0.8	オリジナル加工音声の長さ : 1	長い (1.2 倍) : 1.2
高さ	オリジナル加工音声の高さ : a		3 半音程度低い高さ : b
抑揚	末尾の音が高い : 高	2 音目と同じ高さ : 中	末尾の音が低い : 低

表 2 5 パターンの「長さ 3 水準, 高さ 2 水準, 抑揚 3 水準」の加工音声のうた度評定値と順位, およびオリジナル加工音声が生音声時の評定値と順位

ヘーヤーアー			加工音声		生音声	
長さ	高さ	抑揚	順位	評定値	順位	評定値
1.2	a	高	1	4.8		
1.2	b	高	2	4.7		
0.8	a	高	3	4.3		
1.0	a	高	4	4.2	1	5.4
1.2	a	低	4	4.2		
1.0	b	高	6	4.1		
1.0	a	低	7	3.8		
1.0	b	低	7	3.8		
0.8	b	高	7	3.8		
1.2	b	低	10	3.7		
1.2	a	中	11	3.4		
0.8	a	低	11	3.4		
0.8	b	低	13	3.3		
1.0	a	中	14	3.2		
1.2	b	中	15	3.0		
0.8	a	中	15	3.0		
1.0	b	中	17	2.9		
0.8	b	中	18	2.7		

ギャットタ			加工音声		生音声	
長さ	高さ	抑揚	順位	評定値	順位	評定値
1.0	a	高	1	3.7	2	5.2
1.2	a	高	2	3.6		
0.8	a	高	3	3.5		
1.0	b	高	4	3.3		
1.2	a	低	5	3.2		
0.8	b	高	6	3.1		
1.0	a	中	6	3.1		
1.0	a	低	6	3.1		
0.8	a	低	6	3.1		
1.0	b	中	10	3.0		
0.8	a	中	11	2.9		
1.2	b	低	11	2.9		
1.2	a	中	13	2.8		
0.8	b	低	13	2.8		
1.2	b	高	15	2.7		
1.0	b	低	15	2.7		
0.8	b	中	15	2.7		
1.2	b	中	18	2.5		

シーシーオー			加工音声		生音声	
長さ	高さ	抑揚	順位	評定値	順位	評定値
1.2	a	高	1	4.0		
0.8	a	高	1	4.0		
1.2	b	高	3	3.9		
1.0	a	高	3	3.9	8	4.9
0.8	b	高	5	3.7		
1.2	a	中	6	3.6		
1.0	a	中	6	3.6		
0.8	a	中	8	3.4		
1.0	b	高	9	3.3		
1.2	b	中	10	3.2		
1.0	b	中	10	3.2		
0.8	b	中	12	3.1		
1.0	a	低	13	3.0		
0.8	a	低	13	3.0		
1.2	a	低	15	2.8		
1.0	b	低	15	2.8		
1.2	b	低	17	2.7		
0.8	b	低	18	2.6		

コーケーオー			加工音声		生音声	
長さ	高さ	抑揚	順位	評定値	順位	評定値
1.2	a	高	1	4.1		
1.0	a	低	2	4.0		
1.2	a	低	2	4.0		
0.8	a	高	4	3.9		
1.0	a	高	4	3.9	11	4.7
1.0	b	高	4	3.9		
0.8	a	低	7	3.8		
1.2	b	中	8	3.7		
1.2	b	高	9	3.6		
1.0	a	中	10	3.5		
0.8	a	中	10	3.5		
0.8	b	高	10	3.5		
1.2	a	中	13	3.4		
1.2	b	低	13	3.4		
0.8	b	低	13	3.4		
1.0	b	低	13	3.4		
1.0	b	中	17	3.3		
0.8	b	中	18	3.2		

ソーイーエー			加工音声		生音声	
長さ	高さ	抑揚	順位	評定値	順位	評定値
1.0	a	低	1	4.2	23	4.1
1.2	a	低	2	3.8		
1.2	a	高	3	3.6		
1.0	b	低	3	3.6		
1.0	a	高	3	3.6		
1.2	b	低	6	3.5		
1.2	a	中	6	3.5		
1.2	b	高	8	3.4		
0.8	a	高	9	3.3		
1.0	b	高	10	3.2		
1.2	b	中	11	3.1		
1.0	a	中	11	3.1		
0.8	a	中	13	3.0		
0.8	b	低	14	2.9		
0.8	a	低	14	2.9		
0.8	b	高	16	2.8		
1.0	b	中	17	2.7		
0.8	b	中	18	2.6		

※「生音声」は加工する前の音声
 ※加工音声欄の順位は 18 水準内での順位であり, 生音声欄の順位は前稿での聴取実験 88 パターン内での順位である。

位5位まで末尾が高く加工された音声であり、末尾の音が高いことはうた度の高さと結びついているとみられる。

3.1.4. 「コークーオー」

「コークーオー」(1位音声の評定値4.1)は前稿のうた度評定で11位(前回実験の評定値4.7)だった音声である。今回はオリジナル加工音声は4位であった。今回1位の音声はオリジナル加工音声より長さのみ長いもの、同率2位は末尾のみ異なるもの、また長さが長く、末尾が異なるものであった。「コークーオー」はうた度上位で末尾が低い抑揚のものもあり、抑揚については前稿での傾向が認められないが、高さに関しては上位5位まで高い音声を選ばれており、高い音声がうた度の高いことと関係する傾向がみられる。

3.1.5. 「シーイーエー」

「シーイーエー」(1位音声の評定値4.2)は前稿のうた度評定で23位(前回実験の評定値4.1)だった音声である。5パターンの音声中唯一、元の音声は末尾の低い音声である。そのことが関係しているのか偶然か、この音声のみ末尾の低いオリジナル加工音声は1位で、2位は末尾の音が低い長さが長いものであった。同率3位は長さが長く、全体が高い、かつ末尾が高い音声と、高さのみ異なる音声と、末尾が異なる音声であった。この音声に関してはなぜ末尾の低いオリジナル加工音声が最も歌っている様に聴き取られたかについて疑問が残る。この3音音声は、何らかの理由で下がるのが歌として自然であったのではないか。

3.2. 全体的な考察

今回の実験では、うた度が前稿の実験より相対的に

下がっており、この点については実験用音声加工加工音声であることが影響していると考えられる。加工音声を用いて実験をおこなったことでうた度が下がり、本実験が前稿を裏付けるために最適の方法であったとは言えない。しかし「3.1」にまとめたように、音声長が長い、全体の音高が高い、末尾が高いことが「うた」として聴き取られるという傾向がみられた。

特に抑揚(末尾が高い)に関しては図1に見られるように、末尾が高い「高抑揚」の平均値が高く、また平坦な「中抑揚」の平均値が低いという顕著な傾向がみられる。

また「3.1」ではうた度上位についてのみ触れたが、下位の特徴を見ると「ヘーヤーアー」においては下位に「中抑揚(末尾音が2音目と同じ高さ)」が集中している。また、「ギャッタッタ」においては下位に高さの低い音声が集まっている。「シーシーオー」においては下位に末尾の低くなっている音声が続いている。5パターンすべての音声においてうた度評定の下位に末尾の高い音声がほとんど見られないことから、やはり末尾が高いことは「うた」と結びつく要素であると考えられる。これに対し、高さとは抑揚ほど顕著な傾向はみられなかった。この点について、高さに関してはオリジナル加工音声aが高い方の音声であったことが影響していると考えている。一般的な赤ちゃんの音声の高さとそれより低い音声が比べられたため、低い音声に違和感が生まれた可能性がある。また長さに関しては、0.8, 1, 1.2の3音声の違いが顕著に聞き分けられる範囲ではなかった可能性がある。この点については再度実験を試みたい。

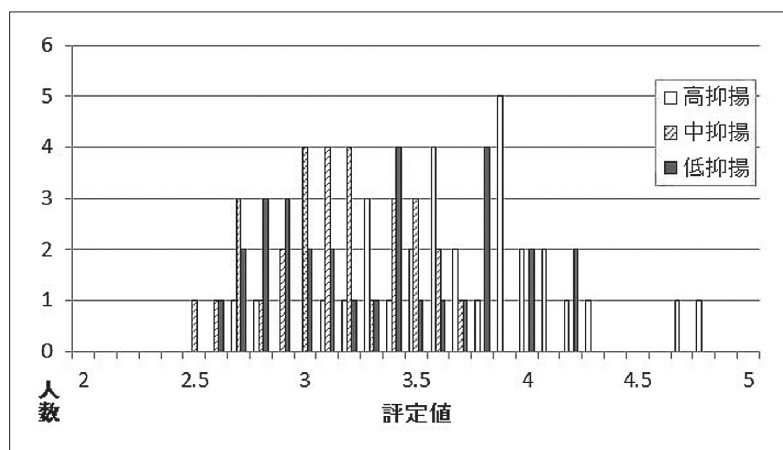


図1 90 音声を用いた聴取実験における抑揚別の評定値と各抑揚のうた度平均値

	低抑揚	中抑揚	高抑揚
平均値	3.25	3.11	3.64

4. ま と め

実験後のアンケートでは、「赤ちゃんの声というよりも機械音に近かった気がします」、「同じような声がたくさんあったように感じました」という記述が数例あり、この点については今後の課題である。しかし、「同じ種類の音でも最後の音が高く上がるか下がるかで聞こえ方が変わっていてふしぎでした」など聴取した感覚を率直に述べている記述や「下から上に声のトーンが高くなっていくときの方が下がっていくよりも歌っている様に聞こえました」と、歌っているように聞こえる音声の特徴についての自身の判断を分析している記述も見られ、被験者は無意識にうた度を記入しているだけではないということがわかった。

今回の実験では、前回の実験に比べ全体的に「うた度」が下がったこと、終了後のアンケートにおいて音声の不自然さを複数名が記述していることから、詳細な統計的分析をおこなわなかった。今回の実験でうたと聴き取られる要素を特定することはできなかったが、うたと聴き取られる音声の韻律的特徴として「長い>短い」、「高い>低い」、「末尾が高い>平坦・下降」という前稿の結果を裏付けることができたと考えている。

しかし「ンーイーエー」において末尾が下降しているオリジナル加工音声が高いうた度を得ていることから、おそらく本稿で取り上げた3要素の実験条件以外に「うた」を特徴づける要素があると考えられる。今後、加工することによって失われたと考えられる音声の抑揚の動態や赤ちゃんらしい声質の音響特性等についても検討しなければならない。

本稿で用いた聴取実験用音声の加工にあたりご指導いただいた埼玉大学島村徹也教授にお礼を申し上げます。

本稿の執筆は坂井が、実験に関しては志村、山根、岡林が、全体の校閲は志村がおこなった。

本研究は、JSPS (課題番号: 25381104, 25381279) の研究費助成を受けている。

注

- 1) 埼玉大学非常勤講師
- 2) 理化学研究所
- 3) 京都女子大学
- 4) NTT 乳幼児音声データベース (音声資源コンソーシアム) から抽出。発話音声ファイルのフォーマットは、量子化ビット数 16 ビット、量子化周波数 16kHz、モノラルの wav 形式である。

文 献

- 市島民子 2004「日本語における初期音韻獲得」上智大学博士論文 (言語学)
- 伊藤勝志 1978「幼児初期の歌唱行動について」『北海道教育大学紀要 第一部 C 教育科学編』第 28 巻第 2 号, pp.157-170
- 岡林典子・坂井康子 2007「母子コミュニケーション場面にみられる創造的なことばのやりとり—日本語のリズム感に注目して」『表現文化研究』第 7 巻第 1 号, pp.11-26
- 坂井康子 2008「幼児の音声表現における歌唱様発声」『甲南女子大学研究紀要』44, pp.29-36
- 坂井康子・岡林典子・佐野仁美 2008「日本語の韻律の獲得—母子間で交わされた 3 拍の唱えことばの抑揚」『表現文化研究』第 8 巻第 2 号, pp.85-97
- 坂井康子・岡林典子・山根直人・志村洋子 2012「喃語のリズムの変化—生後 8 ヶ月, 12 ヶ月, 17 ヶ月の音声の比較から—」, 『甲南女子大学研究紀要』48, pp.43-52
- 坂井康子・岡林典子・山根直人・志村洋子 2013「乳幼児の音声表現のリズムと抑揚」『甲南女子大学研究紀要』49, pp.41-48
- 志村洋子 1991「一歳児の歌—歌唱様発声の音響分析的研究」『音楽教育学の展望Ⅱ下』日本音楽教育学会編, 音楽之友社, pp.152-165
- 永田栄一 1981「子どもの音楽表現の形成と学習 (1)」『季刊音楽教育研究』No.26, pp.160-167
- 益子幸江 2004「日本語の言語獲得過程における音声表出に関する研究」『東京外国語大学論集』No.67, pp.15-37
- 南曜子 1991「言語習得期の音楽的表現『即興うた』の旋律性」『音楽教育学の展望Ⅱ下』日本音楽教育学会編, 音楽之友社, pp.166-175

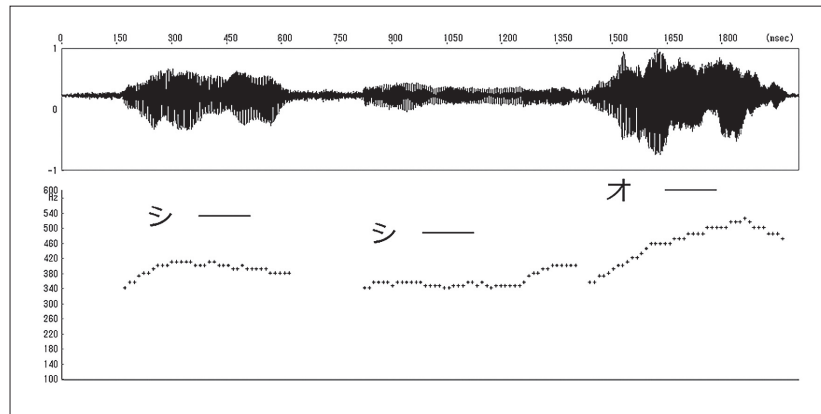
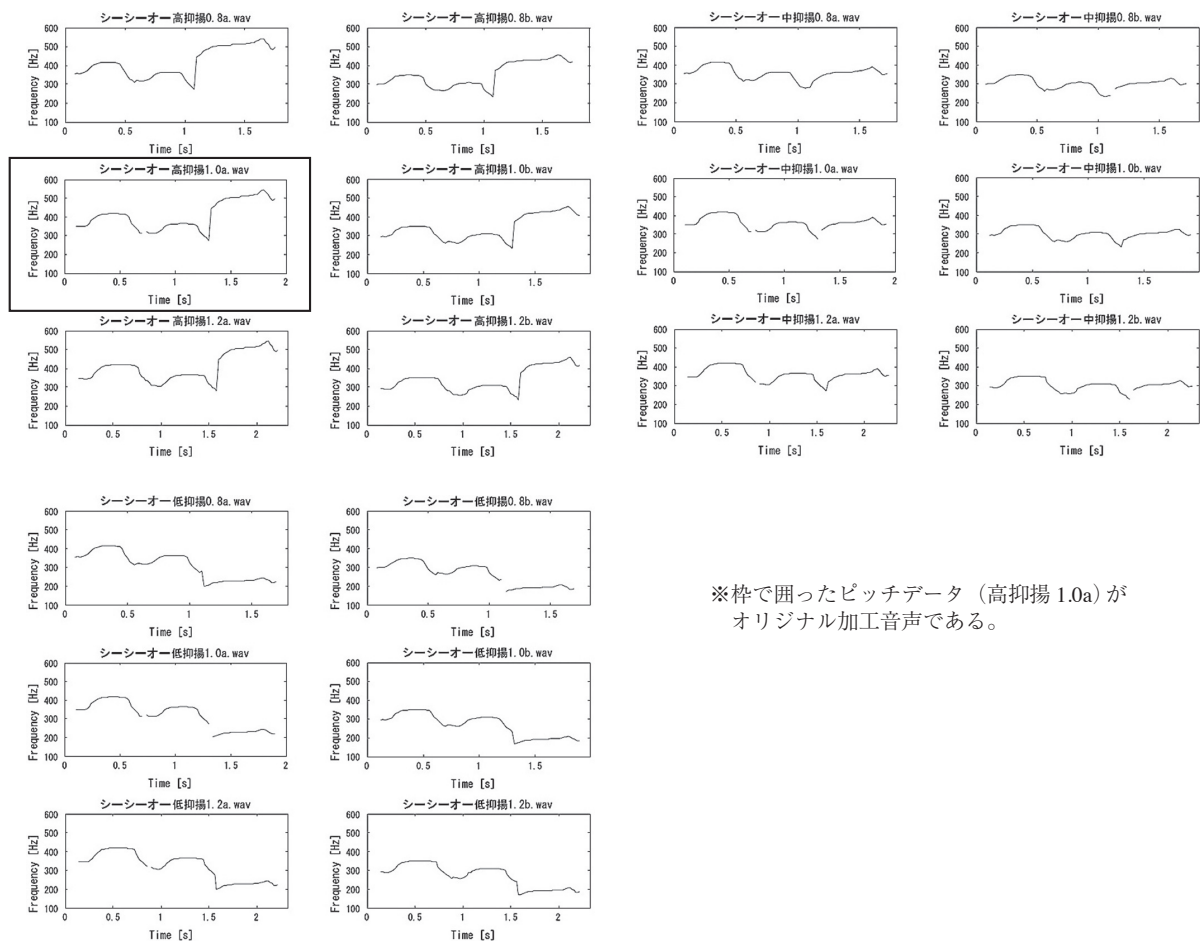


図2 生音声「シーシーオー」のピッチデータ



※枠で囲ったピッチデータ（高抑揚 1.0a）が
オリジナル加工音声である。

図3 加工音声「シーシーオー」のピッチデータ