

乳幼児の歌唱様音声の音響的特徴

坂井 康子・志村 洋子¹⁾

山根 直人²⁾・岡林 典子³⁾

Acoustic Features of Song-like Voices of Infants

SAKAI Yasuko, SHIMURA Yoko, YAMANE Naoto and OKABAYASHI Noriko

Abstract: The prosodic features of infant voices have been studied in order to clarify the way Japanese children express singing elements in their voices. We carried out a listening experiment by presenting 88 types of infant's 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 had 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 was the highest it was perceived as more musical. Thirdly, the musicality score also tended to depend on the overall pitch – the higher it was, the more points it received. In our current research, it was analyzed whether there were any differences of sound quality between the voices that scored high and the ones that scored low of the previous listening experiment. As a result, it was found that each kind of voices tended to have different power level in their high-frequency band.

要旨: 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、筆者らはこれまで乳幼児音声の韻律的特徴について検討してきた。乳幼児の88種の3音等時音声を対象として、歌っていると知覚される度合いに関する聴取実験をおこなった。この結果「うた度」が高いとされた音声の特徴として、①音声長が比較的長い、②音声の末尾が上昇しているものが比較的多い、③ピッチの最高音が高い傾向があることなどが見いだされた。さらに、うた度評定の高い音声とうた度評定の低い音声では音質に何らかの差異があるかどうかを音響的に分析した結果、両音声には高周波帯域のエネルギーに異なる傾向があることがわかった。

1. はじめに

乳幼児の音声は非常に多様で、特に喃語期には人の声かどうかを疑うような音声さえ表出される。しかしそれらの音声は確実に意味を聞き取られるものになっていく。この言語獲得の過程において、子どもは声を出して話そうとするだけでなく、自分の声を確かめた

り楽しんだりなど、様々に発声を試みていることが考えられる。筆者らは喃語期の様々な音声のうちの歌っているかのような音声（歌唱様音声）について研究してきた。喃語期の乳幼児に歌っているかを問うことは無意味であり、また乳幼児に歌うことの意識があるかどうかも分からない。そのため筆者らは喃語期の音声を聴く我々（日本人）はどのような音声を聴いて歌っていると感じているのかを知るために、歌っていると

1) 同志社大学赤ちゃん学研究センター

2) 理化学研究所

3) 京都女子大学

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、うた度評定の上位と下位の音声の韻律的特徴について分析をおこなってきた（坂井他 2012, 2013, 2015）。

本稿では、「2.」でこれまでに分析した韻律的特徴に関する結果を説明したのち「3.」で「2.」の分析を掘り下げ、「4.」で新たにおこなった音響的分析の結果を報告する。

2. これまでの分析結果

2.1. テスト音声とうた度聴取テスト

うた度を測るための聴取テストでは、喃語期の 8 カ月・12 カ月・17 カ月齢児の、3 音が比較的同じ長さ（以下等時とする）のひとまとまりの喃語音声を用いた。日本語として聴取できる音声では評価に偏りが出る可能性があるため喃語音声を採用し、顕著な月齢差があるかどうかを検討するため 3 種の月齢を選択している。また、音数が異なると比較が難しくなるため音数を統一し、各音の長さが異なることにより「うた」としての評価が生じる可能性があるため 1 音声内の各音の長さがほぼ同じ音声を選んでいる。乳幼児の音声において 3 音のまとまりを持った音声が増えること（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 4 音という音数が音声表出の単位としてのまとまりをもたらしやすいという日本語に特徴的な等時性を持った枠組み（坂野 1996）を考慮し、3 音（3 音と休止ととらえている）の音声を選択した。

テスト音声の抽出とうた度聴取テストの概要は以下のとおりである。

(1) NTT 乳幼児音声データベース⁴⁾の 4 児 (sk, mk, sa, ma) の 8 カ月 (mk は 8 カ月齢のデータが無い)・12 カ月・17 カ月齢の喃語 3 音等時音声のうち、それぞれの児の抽出数が一定になるように、音声の状態が比較的良好な音声を選択した（8 カ月齢は 3 児 4 音声ずつ、12 カ月と 17 カ月齢は各児 10 音声ずつ）。これ

ら 92 音声のうち雑音等が含まれた 4 音声を除き、88 音声をうた度聴取テスト用音声とした。

(2) 関東の学生 37 名と関西の学生 38 名に、続けて 3 回ずつテスト用音声を聴取してもらい、「乳幼児が歌っている」ように聴こえたかどうか、1. 「まったく感じない」から 6. 「とても感じる」の 6 段階評価を求めた。

上記方法で聴取テストをおこなった結果、関東と関西では上位下位に大きな差異がみられなかったことから、東西学生 75 名の評価を一括し、うた度の高いものから順に並べた。

2.2. うた度聴取テストの結果—うた度評定上位 10 音声と下位 10 音声の比較

うた度評定の上位 10 音声（1 位から 10 位）と下位 10 音声（88 位から 79 位）がどの児であったか、およびその音声発声時の月齢を表 1 に示す。上位 10 位に 8 カ月齢の音声は 1 例含まれる以外は上位下位とも 12, 17 カ月の音声を選択された。また 4 児は上位下位に偏りがあり、発声（発声におけるうた度）に特徴があることがうかがえた。

うた度評定上位 10 音声と下位 10 音声の各音声の音声長、抑揚（各音間の上昇下降）、最高音を比較した。その結果、まず上位 10 音声は下位 10 音声に比べ音声長の平均が 1 秒近く長く、上位 10 音声は有意に長かった。また 3 音の抑揚を比較した結果、下位 10 音声には末尾が上昇するものは無かったが、上位 10 音声では 10 音声中 6 音声の末尾が上昇していた。また、上位 10 音声の最高音の平均が高い結果であったがこれに有意差は無かった（坂井他 2013）。

88 の音声は生音声であるが故にゆれや音質の違いなどの様々な要素が含まれているため、要素を統一するために加工音声を作成し、聴取実験（生音声での聴取を「聴取テスト」、加工音声での聴取を「聴取実験」と呼んでいる）をおこなった。生音声の長さ、高さ、抑揚を合成音声ソフトで加工し（表 2）、長さとは高さとは抑揚の違いのみの相違を聴き取ることができるよう

表 1 上位下位 10 音声の児と月齢

順位	1	2	3	4	5	6	7	8	9	10	79	80	81	82	83	84	85	86	87	88
4 児	mk	sk	sk	sk	mk	mk	ma	mk	mk	sa	sk	sa	sa	ma	sa	sa	sa	sk	sa	sa
月齢	17	17	17	12	17	17	12	17	17	8	12	17	12	17	17	12	17	17	12	17

4) NTT 乳幼児音声データベース（音声資源コンソーシアム）収録の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 16 kHz、モノラルの wav 形式。

表2 加工音声の実験条件と水準（網掛けは生音声の韻律的特徴を踏襲した加工音声）

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

にした。聴取実験の結果、生音声でのテスト結果と同様に、うた度評定上位の音声は、音声長が長い、比較的末尾が上昇している、音高が高めであるという特徴を持っており、加工音声による聴取実験によって、生音声での結果を裏付けることができた（坂井他 2015）。

3. 音高とピッチレンジ—うた度評定 上位 20 音声と下位 20 音声の比較

うた度評定の上位 10 音声と下位 10 音声の音高の比較において、うた度上位の音声の音高がやや高い傾向がみられ、加工音声の場合でも同様な結果が確認された。しかし有意差がみられなかったことから、分析する音声の数を上位下位 10 ずつ増やし、上位 20 位までと下位 20 位までの音声の音高について、最高音、最低音、平均音高、ピッチレンジを測定した。平均音高は 3 音各音の平均を求めたのち、1 音声の平均を算出した。これは、各音声の特徴が多様であり、例として

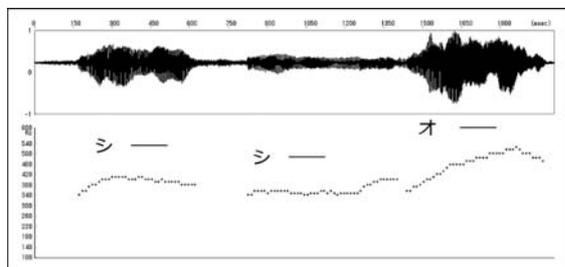


図1 うた度評定 8 位の音声の波形と周波数曲線

表3 最高音、平均音高およびピッチレンジの平均

最高音の平均		平均音高の平均		ピッチレンジの平均	
上位音声	下位音声	上位音声	下位音声	上位音声	下位音声
429.31	370.08	371.14	318.67	104.63	104.58

挙げる図 1⁵⁾のように 1 音が平坦な部分もあれば大きく上昇していることもあることから、最高音と最低音の間での平均では聴取の実態と異なる可能性があると考えたためである。3 音の最高音と最低音、そして 3 音それぞれの最高音、最低音から算出した平均音高をプロットする（図 2）。なお、今回最高音だけではなく一音一音を厳密に測定したため周波数が一部検出できず、上位 16 音声、下位 12 音声のみを示す。

表 3 には図 2 に示した上位下位の音声の最高音の平均、3 音の平均音高の平均、およびピッチレンジの平均を示す。

最高音の平均、平均音高の平均ともに上位音声が高く、これはこれまでの結果をさらに裏付けるものであった。ピッチレンジについては上位下位で差が無く、ピッチレンジはうたとして聴き取ることに影響がない、あるいは影響が少ないとみられる。

4. 音響的特徴の分析

次に、これまでの研究で用いた音声の中から 1 児の音声の音響的特徴（音質）について検討をおこなった。

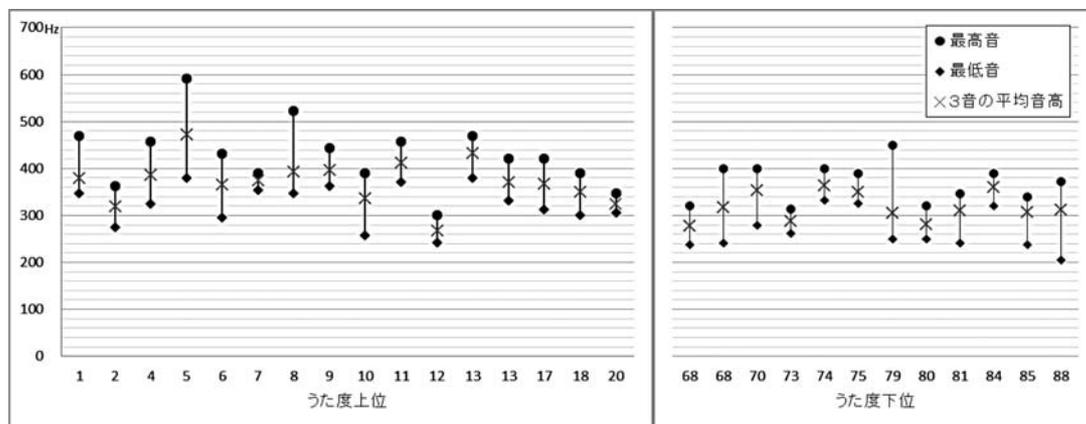


図2 音高とピッチレンジ

5) SUGI SpeechAnalyzer（杉藤美代子 2000 Animo）を使用。

4.1. 分析方法

うた度の高い音声と低い音声の音響的特徴を比較するにあたり、各児の発声の偏りによる影響をなくするため 1 児に限定して分析をおこなった。また分析する部分ができるだけ似通った音韻であることが必要であると考え、上下 20 位まで音声を確認し、mk の /i/ の母音を比較することとした。

以下に今回おこなった音響的分析の手順を示す。

①分析対象として、mk の上位 8 位と 9 位、下位

(下位 10 位以内には mk の音声が含まれていなかった) 17 位と 14 位の音声を選択した。

②スペクトルを観察するため、選択した 4 例の音声の最も /i/ らしい部分 (ジェは特に前方) において瞬時スペクトルを測定⁶⁾した (図 3 左)。

③4 例の高周波成分を明示するため、音質評価システムを用いて、9600 Hz までの周波数成分の測定をおこなった (図 3 右の上)

④音質の特徴を把握するために、音質評価システム

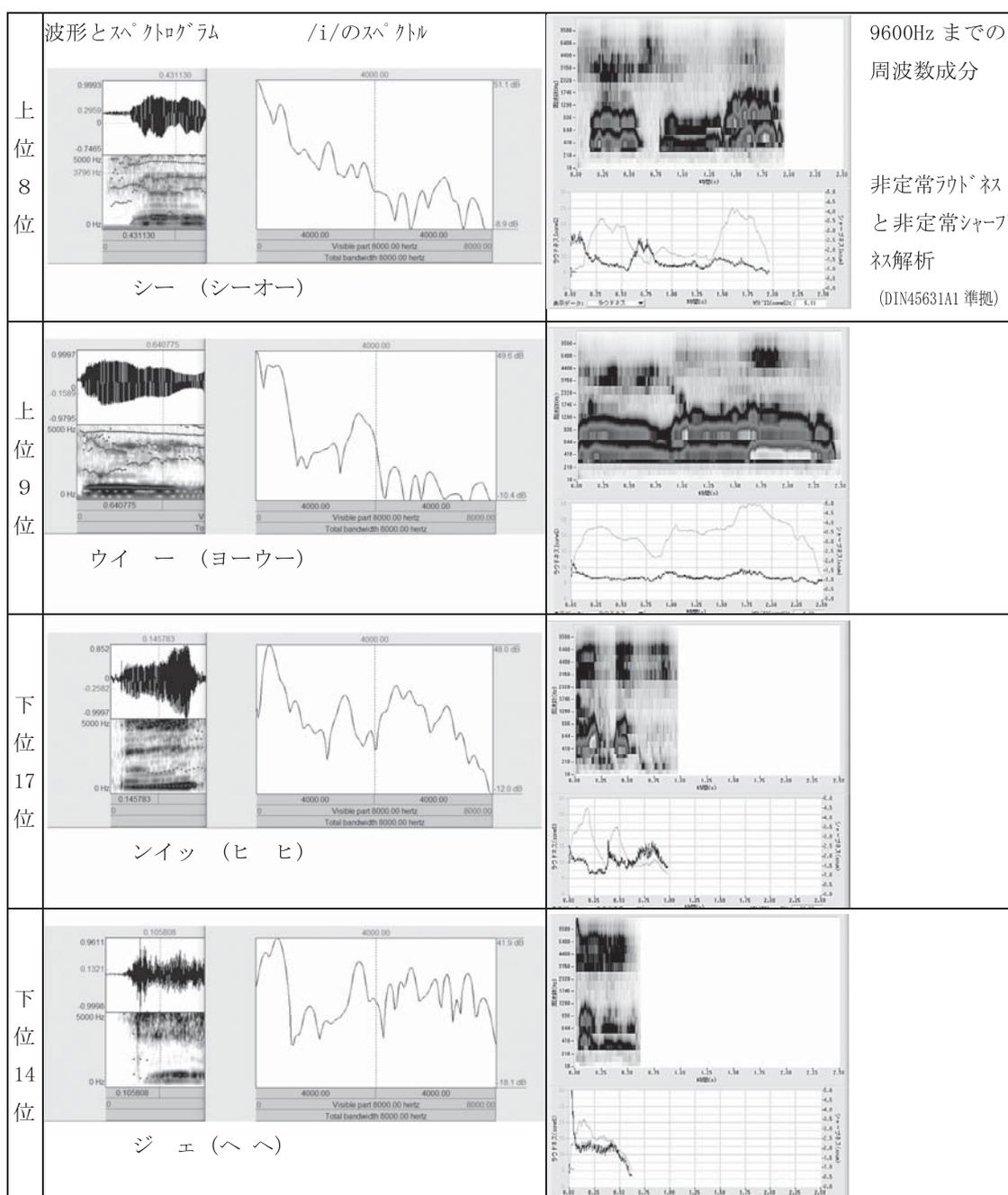


図 3 上位 8 位、上位 9 位、下位 17 位、下位 14 位の音声の音響的分析

6) praat (フリーソフト) を使用。

を用いて、非定常ラウドネスと非定常シャープネス解析 (DIN45631A1 準拠) をおこなった (図3 右の下)

4.2. 分析結果

瞬時スペクトルの測定 (図3の左図左部分「波形とスペクトログラム」上の縦点線の位置のスペクトルを左図右に展開させている) の結果、うた度の高い上位8位と上位9位では周波数の高い部分 (左図右の点線4000 Hz 以上) のエネルギーが弱く、うた度の低い下位17位と下位14位とでは4000 Hz 以上のエネルギーが強いという結果が得られた。これは4音声の高い周波数帯域までを分析した右図上 (横軸は時間 (s), 縦軸は周波数 (最大値 9500 Hz)) のうち、下位17位と下位14位の高周波部分に比較的強い成分がみられることと呼応する。右図下は音質を表す代表的なパラメータ (天津・長松 2009) の中から、非定常ラウドネス、非定常シャープネスの解析をおこなったものである。右図下の2本の線グラフのうち、概ね高い位置にある薄い色の線はラウドネス (sone: 目盛は左) を示し、概ね低い位置にある濃い線はシャープネス (acum: 目盛は右) を示している。ラウドネスが高くなっているときのシャープネスを見ることによって、音声の甲高さ (歌声とは異なる特徴と考えられる) を計測することができると考えられるが、上位8位と上位9位ではシャープネスの位置が相対的に低いことが見て取れた。

4.3. 機嫌の良い音声と悪い音声の比較との比較

「おかあさん」と発声している音声の機嫌のよし悪しと歌っていると聞こえるかどうかの聴取テストをおこない、上記分析対象の mk が発声した上位下位3種

の音声を音質評価システムで分析した (坂井・天津 2015)。ここでは4.2. の結果とこの3種の音声を比較する。

まず「おかあさん」音声の抽出方法や聴取テストについてまとめておく。

先述の NTT 乳幼児音声データベースに収録された音声のうち24ヵ月から36ヵ月齢の「おかあさん」と発音している音声347例 (助詞を伴っていない「おかあさん」のみ) をピックアップし、筆者を含む研究者2名が、まず機嫌の良い音声 (「機嫌良」とする) 20例と機嫌が悪い音声 (「機嫌悪」とする) 20例を選出した。この40例をランダム配列したテスト音声について、成人21名にそれぞれ「とても機嫌が良い」から「とても機嫌が悪い」までの6段階評価を求めた。また歌っているように聞こえる音声にチェックを入れることも併せて依頼した。

テスト音声のうち、分析対象 mk の一般的な機嫌良音声とうた度評定の高い機嫌良音声、および機嫌悪音声の3種の音声を取り上げ、これらの音声の音響的な差異の分析を試みた。図4上段では、薄色部分でラウドネスが大きく、濃色部分はラウドネスが小さい。機嫌良の2例 (左と中央) では、上段薄色部分の周波数が比較的一定に継続しているが、機嫌悪 (右) は周波数成分が一定でなく、また4000 Hz 以上の高い周波数の声のでている特徴が見て取れる。引き続き図4下段で、同児の同じ音声をラウドネスとシャープネスの時間変化で比較した。上方にある薄い線がラウドネス、下方の濃い線がシャープネスである。機嫌良の2例 (左と中央) では、ラウドネス値が大きい時のシャープネス値 (acum) (図中太い部分) が2.0 (点線) を上回ることはないが、機嫌悪 (右) では一部2.0を超える結果となった。

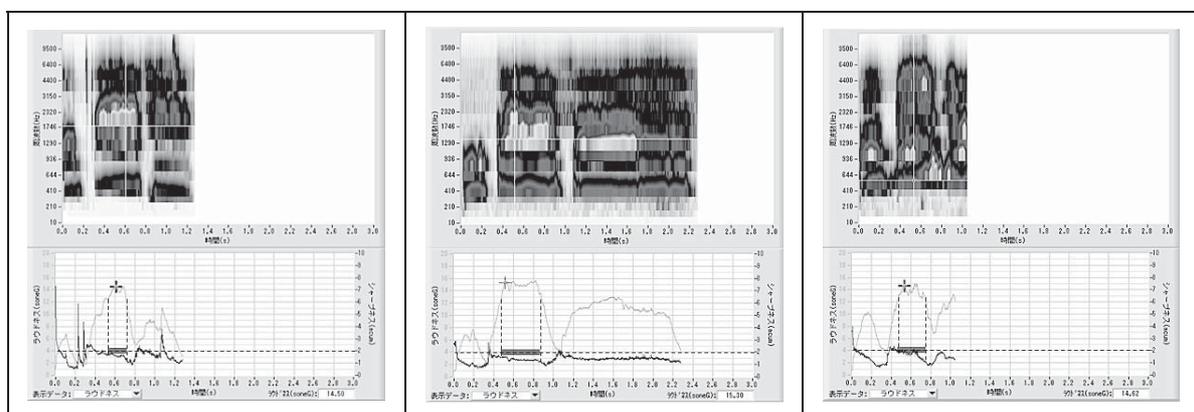


図4 左から機嫌良、機嫌良且つうた評価大、機嫌悪の音声の音質評価分析 (横軸は時間 (s), 縦軸は上段: 周波数 (最大値 9500 Hz), 下段左: ラウドネスの目盛 (sone), 右: シャープネスの目盛 (acum))

分析の結果から、機嫌悪の音声は機嫌良の音声と比較して、高い周波数成分が比較的多く含まれ、周波数成分が重層的に現れている、またシャープネスがラウドネス増大時に相対的に高いという特徴を持っていることが明らかになった。

「4.2.」と「4.3.」を比較すると、喃語 4 音声のうちのうた度の高い 2 例と「おかあさん」のうた評価の高い機嫌良音声とが、また喃語 4 音声のうちのうた度の低い 2 例と「おかあさん」の機嫌悪音声とがそれぞれ似通っており、各々の共通点からみて、これらの音響的分析による歌唱様音声の解析は有効であると考えられる。

5. ま と め

乳幼児の歌唱様音声について、韻律的特徴の分析を追加しておこない、また音質の差異を音響的に明らかにする分析をおこなった。

音高に関しては、分析件数を増やし 1 音ずつを詳細に測定した結果、うた度の高い音声の音高が高めであるという結果を得、これによりこれまでの分析結果をさらに裏付けることができた。ピッチレンジについてはほぼ上位下位が同じ値であり、ピッチレンジとうたとしての聞こえとの関係は認められなかった。

音質の特徴を探るために今回おこなった瞬時スペクトルの測定では、うた度の高い音声において比較的高い (4000 Hz 以上) 帯域のエネルギーが弱いという特徴が抽出された。また、音質評価システムによる非定常ラウドネス解析から、うた度の高い音声はシャープネスがラウドネス増大時に相対的に低い傾向が認められた。

大人の歌唱においては Singer's formant (3000 Hz 付近のエネルギーの集中) が存在する歌声もあるが、子どもの自発的な歌唱においては現状、Singer's formant を検出してはいない。しかしこの度おこなった喃語音声、および「おかあさん」音声の音響的分析において、うた度の高い音声の高い周波数帯域のエネルギーが弱いという結果がみられたが、このことは「共鳴が散在していない」ためなのではないかと考えている。

歌声と話し声にどのような差異があるかについてはこれまでほとんど研究が進んでいなかったが、ここ数年、様々な角度から研究が試みられ (阿曾他 2011, 山崎他 2014), 「歌うこと」に関する研究の進展が期待される。

筆者らの研究においても、今回これまでの韻律的研

究に加え音響的研究を試みる事ができたが、少数の音声の比較にとどまっており、今後歌唱様音声の特徴についてさらに検討をおこなう予定である。

本稿の執筆は坂井が、データ抽出は坂井・岡林、分析に関しては坂井・山根が、全体の校閲は志村がおこなった。

なお、図 3 左の音響分析は朱春躍先生 (神戸大学) に、図 3 右、図 4 は天津成美氏 (キャテック) にお願した。心より感謝申し上げます。

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

文 献

- 阿曾慎平, 齋藤毅, 後藤真孝, 糸山克寿, 高橋徹, 尾形哲也, 奥乃博 2011 「F0・音韻長・パワー制御による歌声らしさ・話声らしさの変化の評価」第 73 回全国大会講演論文集 (1), pp.255-256
- 天津成美・長松昭男 2009 「振動を測る - 振動と音響の計測に関する動向 -」, 機械の研究, 第 61 巻第 1 号, pp.40-48
- 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
- 坂井康子・志村洋子・山根直人・岡林典子 2015 「乳幼児の歌唱様音声の韻律的特徴」甲南女子大学研究紀要, 51, pp.67-73
- 坂井康子・天津成美 2015 「幼児の音声における『機嫌』の音響的特徴」日本音声学会第 29 回全国大会予稿集, pp.182-187
- 板野信彦 1996 『七五調の謎をとく』大修館書店
- 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
- 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
- 山崎健史, 池宮由楽, 糸山克寿, 奥乃博 2014 「歌声 - 話声変換における動的音響特徴量が話声らしさに及ぼす影響」第 76 回全国大会講演論文集 (1), pp.373-374