

『理研母子会話コーパス (R-JMICC)』構築の試みと研究成果
— 対乳児自発音声における
日本語特有の韻律的・分節的特徴の解明を目指して —

西海枝洋子 (理化学研究所 脳科学総合研究センター 言語発達研究チーム) †

渡辺和希 (筑波大学大学院)

小西隆之 (理化学研究所 脳科学総合研究センター 言語発達研究チーム)

伊藤直子 (筑波大学大学院)

金礪愛 (早稲田大学人間科学学術院)

五十嵐陽介 (広島大学文学研究科文学部)

宮澤幸希 (理化学研究所 脳科学総合研究センター 言語発達研究チーム)

西川賢哉 (理化学研究所 脳科学総合研究センター 言語発達研究チーム)

馬塚れい子 (理化学研究所 脳科学総合研究センター 言語発達研究チーム)

Riken Japanese Mother Infant Conversation Corpus (R-JMICC)
— **Compilation and Recent Findings of Japanese-Specific Prosodic and Segmental Characteristics in Infant-Directed Speech** —

Yoko Saikachi (Lab. for Language Development, RIKEN Brain Science Institute)

Kazuki Watanabe (Graduate school of University of Tsukuba)

Takayuki Konishi (Lab. for Language Development, RIKEN Brain Science Institute)

Naoko Ito (Graduate school of University of Tsukuba)

Ai Kanato (Waseda University)

Yosuke Igarashi (Hiroshima University)

Koki Miyazawa (Lab. for Language Development, RIKEN Brain Science Institute)

Ken'ya Nishikawa (Lab. for Language Development, RIKEN Brain Science Institute)

Reiko Mazuka (Lab. for Language Development, RIKEN Brain Science Institute)

1. はじめに

大人は乳児に語りかける際、韻律的・分節的特徴を強調した独特な話し方をする (対乳児音声、Infant-Directed Speech; IDS)。全体的に声が高く、ゆっくりとして、ピッチの変動幅が大きいなどの IDS の特徴は、多言語において報告されており (Fernald et al. 1989, Soderstrom 2007)、その普遍性と言語獲得における重要性が示唆されてきた。しかし、日本語は他の言語とは異なる独特な音韻体系、そして韻律構造を持つ。その特異性を考慮した上で、対乳児音声の音響的特徴を定量的に分析し、明らかにした研究はほとんど行われてこなかった。

そこで本研究チームでは、ここ数十年の間に開発が進められてきた音声コーパスの構築技術を応用し、『理研母子会話コーパス (R-JMICC)』 (Mazuka, Igarashi, and Nishikawa 2006, 五十嵐, 馬塚 2006) の構築を進めている。このコーパスは、2005 年に収録を行った 18-24 ヶ月の乳児を持つ母親 22 名による自発的な会話音声データ (対乳児音声および対成人音声)

† ysaikachi@brain.riken.jp

と、その4年後の2009年に収録を行った同じ母親による音声データ（対乳児音声、対成人音声および読み上げ音声）で構成されており、『日本語話し言葉コーパス（CSJ）』（前川2004, 2006）にほぼ準拠した、形態論情報、分節音情報、および韻律情報が付与されている。これらの付加情報を活用することにより、韻律句の構成やピッチアクセント、句末音調といった日本語の韻律特性を考慮しながら、自発音声における韻律的・分節的特徴を分析することが可能となった。

本稿では、まずコーパスの概要を説明した後に（2節）、コーパス分析に基づく最近の研究成果の具体例として、日本語対乳児音声における1)ピッチレンジの拡大という韻律的強調の局所性（Igarashi and Mazuka 2008, Igarashi et al. 2009）（3節）と、2)長短母音の持続時間長の分布特性（Bion et al. in press）（4節）を紹介する。

2. データ

2.1. 『理研母子会話コーパス』

『理研母子会話コーパス（R-JMICC）』は、2005年に収録されたデータと、2009年に収録されたデータで構成される（表1参照）。2005年収録のデータは、会話形式の自発音声（以下05セット）、2009年収録のデータは、会話形式の自発音声（以下09セット）と、文章の読み上げ音声（以下ATRセット）から構成される。

表1 『理研母子会話コーパス』の概要

収録年 (参加者数)	発話スタイル (名称)	内容 (収録時間/文の総数)	総時間	語数
2005年 (22人)	自発音声 (05セット)	IDS 絵本を見ながらの会話 (15分)	11時間	49340
		玩具で遊びながらの会話(15分)		
	ADS 育児に関する会話 (10分)	3時間	25012	
2009年 (20人)	自発音声 (09セット)	IDS 絵本を見ながらの会話 (10分)	7時間	33280
		玩具で遊びながらの会話(10分)		
	ADS 育児に関する会話 (10分)	3時間	25901	
	読み上げ音声 (ATRセット)	ATR 音素バランス 503 文 A セット(50 文)	4時間	14538

2.1.1 参加者

05セットの参加者は、母親22名（25-43歳、平均年齢33.0, SD±3.6）とその子供である。母親は全て関東地方（東京・神奈川・埼玉・千葉）出身であり、標準日本語を話す。子供の月齢は18-24カ月（平均20.4, SD±2.7カ月）であった。2009年には、2005年の収録に参加した親子22組中20組が再度参加した。

2.1.2 音声の収録環境

理化学研究所言語発達研究チーム内の防音室で、母親にヘッドセット型コンデンサマイク(CROWN, CM-312A)を装着してもらい、一組ずつ録音を行った。また、直接の分析対象とはしていないが、2005年にはコンデンサマイク(Behringer, B-5)をテーブル上に配置し、2009年には子どもにもヘッドセット型ダイナミックマイク(SHURE, SM10A)を装着してもらい、子供の発話も収録した。音声は、DAT(TASCAM, DA-P1)を用いて収録した(44.1 kHz、16ビット)。

2.1.3 収録内容

05 セットの内容は3種類である。まず、IDSとして1)絵本を見ながら母親が子供に話しかける音声と、2)玩具で遊びながら子供に話しかける母親の音声を収録した。次に、対成人発話(Adult-Directed Speech; ADS)として、3)同じ母親による成人の実験者(同年代かつ子育て中の女性)との会話音声も収録した。ADSは、育児に関する内容が多かった。2009年には、05セットと同一のタスクを用いた自発音声(09セット)に加え、母親による読み上げ文(ATRセット)の収録を行った。

2.2. 研究用付加情報

R-JMICCには、書き起こしテキスト(転記テキスト)、形態論情報(単語境界や、品詞、および活用形についての情報)、分節音情報、韻律情報といった様々な研究用付加情報が付与されている(図1参照)(Mazuka, Igarashi, and Nishikawa 2006)¹。付加情報は、概ねCSJ(国立国語研究所(編)2006)に準拠している。

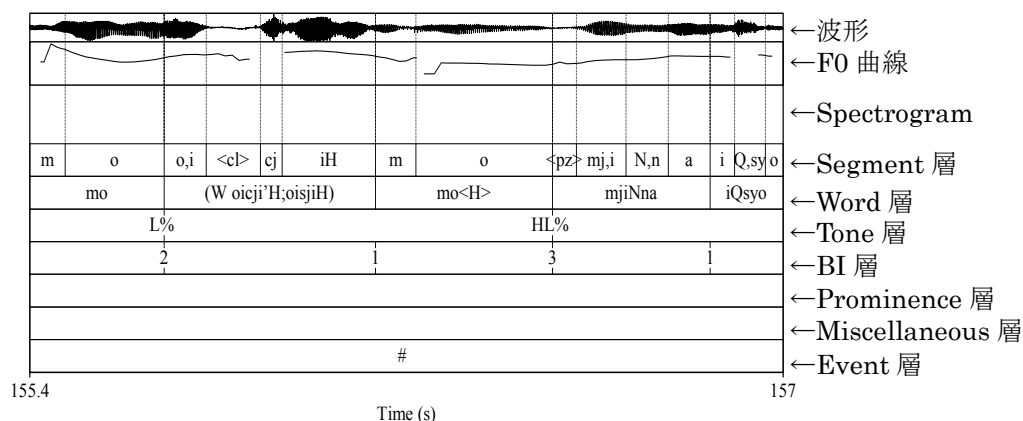


図1 研究用付加情報の一例

- (a) 音切り・書き起こし：発話を200msec以上のポーズで区切り(音切り)、区切られた単位(転記基本単位; IPU)ごとに、聞き取れる範囲で忠実に発話を記す。談話・音声現象(フィラー、語断片、歌声、対成人音声等)を記述するためのタグもこの段階で付与する。
- (b) 形態論情報付与：短単位(辞書の見出し語に相当)および長単位(複合語を一つと扱う)という二種類の形態論的単位を認定し、それぞれの単位に品詞などの付加情報を付与する。
- (c) 分節音情報および韻律情報付与：分節音情報として、子音・母音の種類とその境界位置に関する情報を、韻律情報として、アクセント・イントネーション(韻律句境界のレベル、句末音調の種類等)に関する情報を記す。韻律ラベリングスキームは、CSJで採用されているX-JToBI方式の体系を一部改訂したものである(五十嵐、馬塚 2006)。

3. 日本語対乳児自発音声における韻律的特徴：ピッチレンジ拡大の局所性

3.1 研究の背景および目的

¹ 05セットは、一連のアノテーション作業が終了しており、3節および4節で紹介する研究成果は、05セットの解析結果に基づくものである。ATRセットも一通りアノテーション作業が終了している。09セットは、(a)-(b)の作業が終了しており、現在(c)の分節音情報および韻律情報の付与を進めている。

これまで多くの研究により、ピッチレンジの拡大という韻律的強調が IDS の主要な特徴とされてきたが、日本語にはそのような強調が存在しないのではないかと示唆されてきた (Fernald et al. 1989)。五十嵐らは、日本語特有の韻律的強調の実態を明らかにするために、アクセント、ダウンステップ、句末音調などの日本語に特徴的な韻律構造による影響を考慮した上で、R-JMICC(05 セット)におけるピッチレンジの分析を行った (Igarashi and Mazuka 2008, Igarashi et al. 2009)。

3.2 日本語の韻律構造

3.2.1 韻律句の特徴

R-JMICC のラベリングスキームである X-JoBI では、アクセント句 (Accentual Phrase, 以下 AP) とイントネーション句(Intonational Phrase, 以下 IP) の2種類の韻律句が仮定されている。AP は、句頭で上昇し、その後句末にかけて段々と下がり、低く終わるといようなピッチ曲線で特徴づけられる (Venditti 2005)。AP がアクセントを含む有核句の場合、アクセントによる急激なピッチの下降があるため、無核句と比較すると、AP の最低ピッチが引き下げられ、ピッチレンジが拡大する。

AP より階層的に上位に位置づけられる IP は、アクセント核が、後続する AP の最大ピッチを反復的に低下させるダウンステップという音韻現象が生じる領域と定義され、IP 境界でピッチリセット (前の文脈とは独立した新たなピッチレンジの設定) が生じる。IP 内のアクセントの数が増えると、ダウンステップの影響により一番目の AP の最大ピッチが高くなり、そのため IP のピッチレンジが拡大すると考えられる。

3.2.2 句末複合境界音調 (Boundary Pitch Movement; BPM)

AP の終端には、下降調だけではなく、上昇調や上昇下降調などの局所的音調が生じる。句末複合境界音調 (Boundary Pitch Movement, 以下 BPM) と呼ばれるこの音調は、質問、強調、継続など、語用論的な意味あるいは発話意図の伝達に重要な役割を果たしており、R-JMICC では、主に H% (上昇調 1)、LH% (上昇調 2)、HL% (上昇下降調) によって表現されている (Igarashi and Mazuka 2008, Igarashi et al. 2009) (図 2)。

韻律句の BPM 以外の場所 (以下、主要部と呼ぶ) のピッチパターンは、アクセントやダウンステップ等、語彙情報によって規定されているため、韻律的強調による変化の度合いは限定的と考えられる。一方、韻律句末に生じる BPM はそのような制約がなく、発話の意図などを表現するためにピッチの特徴を強調しやすい場所であると考えられる。

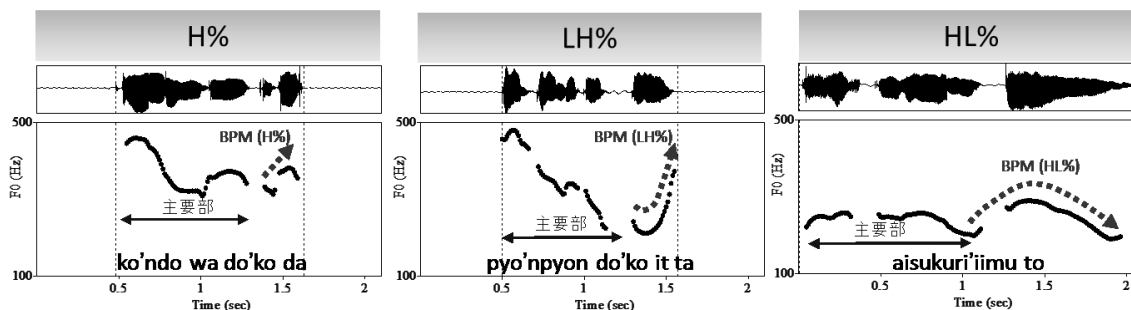


図 2 BPM および主要部のピッチ曲線

3.3 コーパス分析²

² IDS データは、絵本を読みながらの会話音声のみ分析対象としている。また、22 名中 1 名の母親のデータは声質に問題があるため分析対象外としている。

3.3.1 Utterance (発話)³全体の特徴

五十嵐らはまず、Utterance 全体を対象とした分析を行った(図3)。その結果、ピッチの最大値・平均値・最小値は、ADS と比較して IDS では有意に高くなってしたが、IDS におけるピッチレンジの拡大は観察されず、先行研究 (Fernald et al. 1989) と同様の結果が得られた。

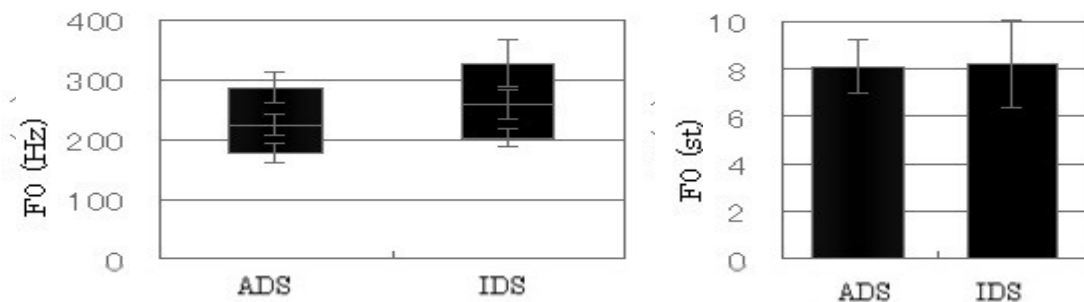


図3 Utterance 全体のピッチの特徴。左図：ピッチの最大値（図の上限值）・平均値（図の中央値）・最小値（図の下限值）。右図：ピッチレンジ。エラーバーは標準偏差を示す。

3.3.2 BPM および主要部の特徴

ピッチレンジの拡大を Utterance 全体で測定するのではなく、主要部と BPM に分けて分析を行った。まず、BPM の相対頻度を分析したところ、H%と LH%は、IDS でより頻繁に現れ、逆に、HL%は、ADS でより頻繁に観察された(図4)。

次に、BPM 毎にピッチの特徴量を分析したところ、IDS では、全ての BPM の種類で、ピッチの最大値および平均値が ADS よりも有意に高く、ピッチレンジが有意に拡大していた(図5、図6)。ピッチの最小値については、H%と LH%では、IDS のほうが有意に高かったが、HL%では発話者間に有意差はなかった。

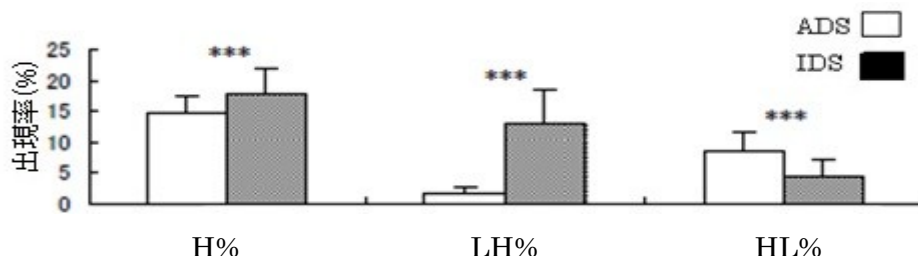


図4 BPM の出現率

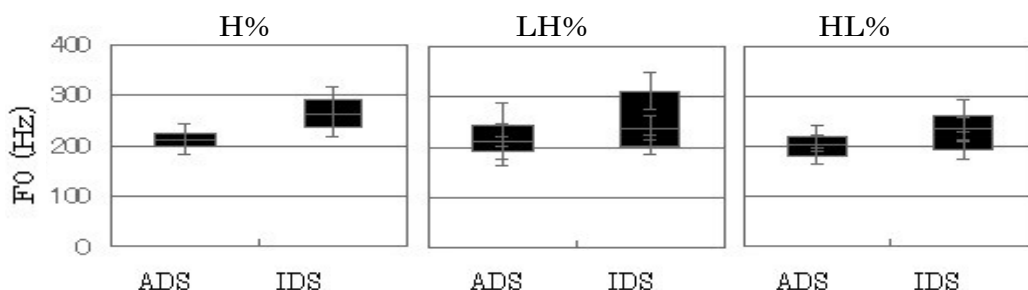


図5 BPM のピッチの特徴：最大値・平均値・最小値

³ ここでは、Utterance (発話) を、200msec 以上のポーズが後続する IP 境界で区切られる単位と定義している。

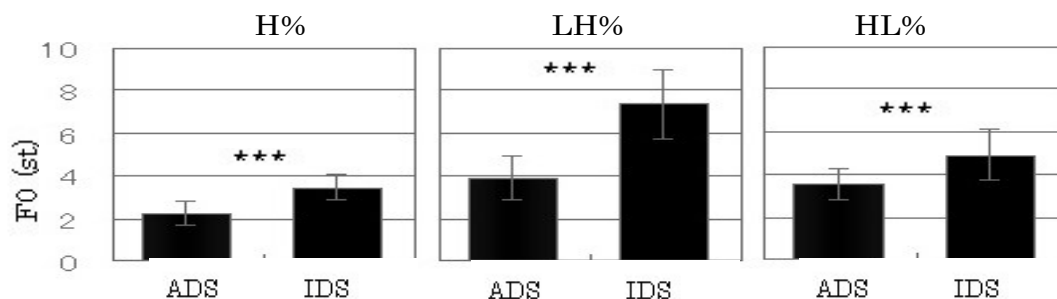


図6 BPMにおけるピッチレンジ

次に、主要部のピッチの特徴量を比較したところ、ピッチの最大値・平均値・最小値は、ADSと比較してIDSでは有意に高かったが、ピッチレンジはADSがIDSと比較して有意に大きかった(図7)。

つまり、Utteranceを主要部とBPMに分けて分析すると、IDSにおけるピッチレンジの拡大は韻律句末のBPMでのみ顕著に現れ、主要部ではIDSよりもADSのほうが、ピッチレンジが大きいことが明らかになった。

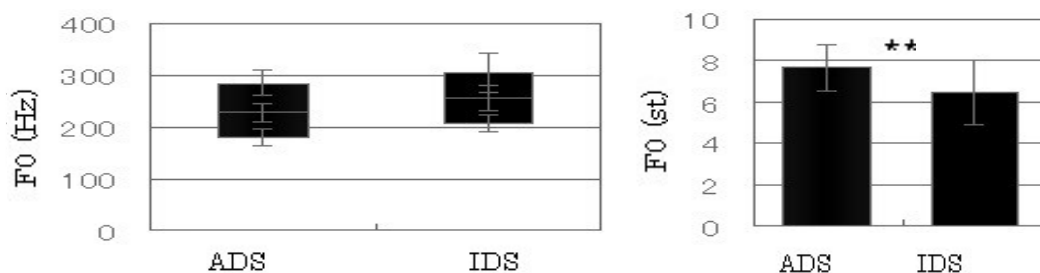


図7 主要部のピッチの特徴：最大値・平均値・最小値（左図）、ピッチレンジ（右図）

3.3.3 IPの長さを考慮した場合の主要部におけるピッチレンジ

最後に、主要部におけるピッチレンジを、IPの長さを考慮して分析した結果を紹介する。まず、持続時間長(msec)、モーラ数、短単位数、IPの数、APの数、アクセントの数といった項目について計測したところ、全てにおいて、ADSのほうがIDSより長かった。

次に、IPの長さ(アクセント数)を統一してADSとIDSのピッチレンジを比較した(図8)。その結果、アクセント数が1、2、3個では、IDSはADSと比べて、ピッチレンジが有意に大きかったが、0個および4個の場合はレジスター間に有意な差は認められなかった。

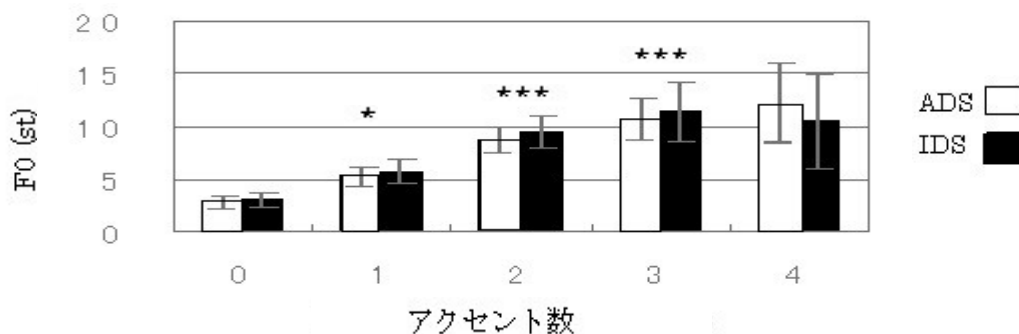


図8 IPの長さ(アクセント数)を統一した際のピッチレンジ

つまり、IPの主要部のピッチレンジは主にアクセントの数によって決められていること、そして主要部でADSの平均ピッチレンジが大きいのは、ADSでは、より多くのアクセントが含まれているためであり、このことがIDSにおけるピッチレンジの拡大を分かりづらくしている理由であることが明らかになった。

3.4 ピッチの特徴に関するまとめ

五十嵐らは、韻律情報が付与されているR-JMICCを使用して、日本語特有の韻律構造を考慮した分析を行うことにより、日本語の対乳児自発音声では、ピッチレンジの拡大という韻律的特徴の強調が、韻律句末のBPMに局所的に現れることを明らかにした。五十嵐らの研究の大きな意義は、ここ数十年で研究が積み重ねられてきた韻律の音韻構造(Ladd 1996, Pierrerhumbert and Beckman 1988)に基づく分析を行うことによって、対乳児音声における韻律的特徴の強調が、全ての言語で同じように現れているのではなく、個別言語独自の韻律構造の枠内で表出するものであることを示したことである。今後、多様な言語で解析が進むことにより、個々の言語特有の対乳児音声の特徴だけではなく、構造的な違いが大きい言語間における共通した特徴を明らかにすることが出来れば、音声言語獲得における言語普遍性の解明への一歩となることが期待される。

4. 日本語のIDSにおける長短母音の出現頻度分布特性

本節では、長短母音という日本語特有の音韻対立に焦点をあてて、対乳児自発音声における音韻の出現頻度分布特性を分析したBionらによる研究(Bion et al. in press)を紹介する。

4.1 研究の背景

日本語では、「床」/toko/と「渡航」/tokoo/のように、母音の長短が語の意味を区別するが、英語のような言語では母音の長短のみで語の意味を区別することはない。では、乳児は、一体どのようにして母語特有の音韻的な対立構造を習得しているのであろうか。乳児にとっての主要な入力音声である対乳児音声には、何か手がかりとなるような情報が含まれているのだろうか。

Werkerらは、日本語を母語とする母親による読み上げ音声の分析に基づき、日本語では長短母音間に母音長の確かな差が存在し、音韻的な長さの指標として使用することが出来ることを指摘している(Werker et al. 2007)。しかし、この結果は、無意味単語対を用いた、発話内容を予め統制している読み上げ音声の収録に基づく結果であり、多様な実在語を豊富に含む自発的な日本語対乳児音声においても、同様の結果が得られるのかどうかは明らかになっていない。

また、乳児のように、長母音、短母音というカテゴリーが存在すること自体をまだ認識していない場合、入力音声における音韻カテゴリーの出現頻度分布特性が、音韻の獲得に重要な役割を果たしていることが、統計学習モデルを用いた研究(Vallabha et al. 2007)や、乳児を対象とした実験的研究(Maye et al. 2002)によって指摘されている。このような研究では、入力音声の特徴として、1)異なる音韻カテゴリーが入力音声に同頻度で存在すること、そして2)出現頻度分布が二つの山を持つような双峰性であること、という二つの条件を前提にしているが、日本語の対乳児自発音声においても、前提とされている頻度分布特性が存在しているのかどうかは確かめられていない。

4.2 研究の目的

Bionらの研究の目的は、母語の音韻構造を習得するために重要とされている特徴が、実際の入力音声において存在しているのかどうか、R-JMICC(05 セット)における日本語の長

短母音の出現頻度分布特性を分析し検証することである。

4.3 コーパス分析

4.3.1 平均値の比較

Bion らはまず、自発的な日本語対乳児発話音声においても、長短母音間に確実な音響的な違いが存在するのかどうかを検討するために、日本語の各母音(a,e,i,o,u)について、短母音と長母音の長さの平均を比較した。図 9 に示すように、母音の種類に関係なく長母音は短母音よりも長く、その差は統計的に有意であった。

次に、Werker et al.(2007)に従って、母音のカテゴリが従属因子、母音長が独立因子、そして話者がランダム効果とするロジスティック回帰分析を行った。母音の持続時間長に基づいて母音のカテゴリ（短母音 vs. 長母音）を予想するモデルは有意であった。22 名それぞれの話者においても、同じような結果が得られた。

このように、大量の母音を含む自然発話の収録においても、Werker et al. (2007)と同様に、日本語の長短母音の長さには、確かな差があることが確認された。

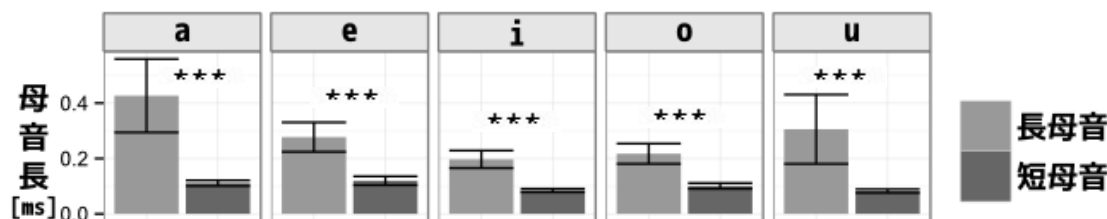


図 9 長短母音の持続時間長：被験者間平均および標準偏差

4.3.2 出現頻度分布特性の分析

次に、単純な出現頻度分布モデルによって、日本語における音韻的な長さの習得を説明することが可能なのかどうかを探るために、対乳児発話データにおける母音の頻度分布を分析した（図 10）。その結果、コーパスに現れている母音のほとんどが短母音であること（全体の 94%）、そしてそれぞれの母音について、入力音声における長短母音の分布が完全に重複していることが分かった。22 名それぞれの話者においても、同様の結果が確認された。

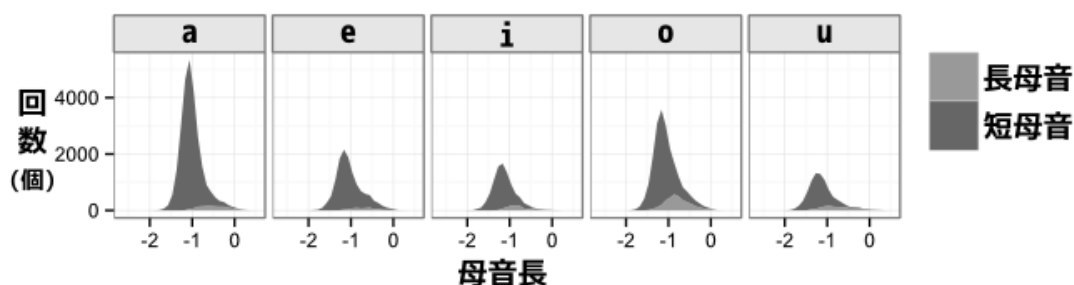


図 10 長短母音の出現頻度分布

つまり、実際の日本語の入力音声では、ほとんどの母音が短母音であり、もともとの出現頻度の偏りがあるため、長短母音を合わせた母音長の出現頻度分布は、音韻カテゴリ習得モデルが前提としたような双峰性ではなく単峰性であることが確認された。

4.4 長短母音の分析に関するまとめ

Bion らは、初期音韻獲得モデルの前提条件とされてきた入力音声における出現頻度分布の特性を検討するために、日本語の対乳児自発音声における長短母音の分析を行った。その結果、日本語の長短母音には、持続時間長に顕著な差があるが、出現頻度の偏りがあるため、出現頻度分布による音韻習得モデルのアルゴリズムのみでは習得は困難なことが示唆された。そのため、頻度分布以外にも韻律情報や語彙情報など、他の情報を考慮に入れなければ、日本語では母音長が語の意味の違いを区別しているということを、習得するのは難しいと考えられる。実際、日本語を母語として学んでいる乳児は、音韻的な母音長に対する感度を示すのに時間がかかることが明らかになっている(Mugitani et al. 2009, Sato, Sogabe, and Mazuka 2010)。最近では、音声・音韻知覚の発達過程を語彙獲得との関係で捉えようとするモデルも提唱されており (Feldman, Griffiths, and Morgan 2009)、今後の研究として、異なる習得モデルによる仮説を、入力音声における実際の語彙や音響的指標の分布によって検証していくことが重要である。

5. おわりに

本稿では、対乳児自発音声コーパス構築の取り組みと、最近の研究成果について紹介した。本コーパスのように、統制の加えられていない対乳児自発音声を、ある程度の量、収録し、詳細なアノテーションを行うことにより、日本語特有の音韻・韻律構造を考慮しながら、音声的特徴を分析することが可能になった。本稿で紹介したように、R-JMICC(05セット)の解析によって、今までの研究では見落とされていた日本語特有の韻律的強調の実態の解明 (Igarashi and Mazuka 2008, Igarashi et al. 2009)、そして母語固有の音韻獲得モデルとして重要視されてきた音韻の出現頻度分布特性の検証 (Bion et al. in press) に貢献してきた。

このような研究成果は、主に 18-24 ヶ月の乳幼児に語りかけた母親の発話データ (05 セット) の分析に基づくものである。対乳児発話音声は、語りかける子供の年齢の変化や言語発達とともに、その音響的特徴もさまざまな次元において次第に変化していくことが知られている (Kitamura et al. 2001)。今後、09 セットのラベリングを進め、05 セットと 09 セットの解析結果と組み合わせることで、言語発達に伴って対乳児音声がどのように変化していくのかを明らかにしていくことが可能になる。また、ATR セットの音声ラベルを合わせて解析し、読み上げ音声との比較分析を行うことで、対乳児発話および対成人発話という異なる自発音声スタイルの特徴を、より相対的に捉える事が可能になると考えられる。

謝辞

現在進行中のコーパス構築作業において韻律ラベリングに関する貴重な助言を頂いている宇都木昭氏、そして原稿準備にご協力いただいた浅井拓也氏に感謝いたします。本研究は、文部科学省科学研究費補助金・基盤研究(C)21610028 (研究代表者：馬塚れい子) および若手研究(B)23720223 (研究代表者：西海枝洋子) の補助を得ています。

参考文献

- Bion, Ricardo, Koki Miyazawa, Hideaki Kikuchi, and Reiko Mazuka (in press) “Learning phonemic vowel length from naturalistic recordings of Japanese infant-directed speech”, *PLoS ONE*.
- 五十嵐陽介、菊池英明、前川喜久雄 (2006) 「第 7 章韻律情報」『日本語話し言葉コーパスの構築法』(国立国語研究所報告書 124), pp.347-453.
- 五十嵐陽介、馬塚れい子(2006)「母親特有の話し方(マザリーズ)は大人の日本語とどう違うか-理研日本語母子会話コーパス-」電子情報通信学会技術研究報告,106:43, pp.31-35.
- Igarashi, Yosuke and Reiko Mazuka (2008) “Exaggerated prosody in infant-directed speech?”

- Intonational phonological analysis of Japanese infant-directed speech”, *BUCLD 32: Proceedings of the 32nd annual Boston University Conference on Language Development*, pp.177-188.
- Igarashi, Yosuke, Ken'ya Nishikawa, Kuniyoshi Tanaka and Reiko Mazuka (2009) “Pitch range expansion in Japanese infant-directed speech”, *The 4th International Workshop on the Interface between Prosody and Information Structure* (Jan. 11, 2009, Otsu, Shiga, Japan).
- Feldman, Naomi, Thomas Griffiths, and James Morgan (2009) “Learning phonetic categories by learning a lexicon”, *Proceedings of the 31st Annual Conference of the Cognitive Science Society*, Amsterdam, pp.2208-2213.
- Fernald, Anne, Traute Taeschner, Judy Dunn et al.(1989) “A cross-language study of prosodic modifications in mothers' and fathers' speech to preverbal infants”, *Journal of Child Language*, 16:3, pp.477-501.
- Kitamura, Christine, Chayada Thanavishuth, Dennis Burnham, and Sudaporn Luksaneeyanawin (2001) “Universality and specificity in infant-directed speech: Pitch modifications as a function of infant age and sex in a tonal and non-tonal language”, *Infant Behavior and Development*, 24:4, pp.372-392.
- 国立国語研究所編(2006)『日本語話し言葉コーパスの構築法』(国立国語研究所報告書 124).
- 前川喜久雄 (2004) 「『日本語話し言葉コーパス』の概要」『日本語科学』, 15, pp.111-133.
- 前川喜久雄 (2006) 「第 1 章概説」『日本語話し言葉コーパスの構築法』(国立国語研究所報告書 124), pp.1-22.
- Ladd, Robert (1996) *Intonational Phonology*, Cambridge: Cambridge Univ. Press.
- Maye, Jessica, Janet Werker, and LouAnn Gerken (2002) “Infant sensitivity to distributional information can affect phonetic discrimination”, *Cognition*, 82:3, pp.B101-B111.
- Mazuka, Reiko, Yosuke Igarashi, and Ken'ya Nishikawa (2006) “Input for learning Japanese: RIKEN Japanese Mother-Infant Conversation Corpus”, 電子情報通信学会技術研究報告, 106:165, pp.11-15.
- Mugitani, Ryoko, Ferran Pons, Laurel Fais et al. (2009) “Perception of vowel length by Japanese- and English-learning infants”, *Developmental Psychology*, 45:1, pp.236-247.
- Pierrehumbert, Janet and Mary Beckman (1988) *Japanese Tone Structure*, Cambridge: MIT Press.
- Sato, Yutaka, Yuko Sogabe, and Reiko Mazuka. (2010) “Discrimination of phonemic vowel length by Japanese infants”, *Developmental Psychology*, 46:1, pp.106-119.
- Soderstrom, Melanie (2007) “Beyond babytalk: Re-evaluating the nature and content of speech input to preverbal infants”, *Developmental Review*, 27, pp. 501-532.
- Vallabha, Gautam, James McClelland, Ferran Pons et al. (2007) “Unsupervised learning of vowel categories from infant-directed speech”, *Proceedings of the National Academy of Sciences*, 104:33, pp.13273-13278.
- Venditti, Jennifer (2005) “The J_ToBI model of Japanese intonation”, Jun, S. -A. (ed.) *Prosodic Typology: The Phonology of Intonation and Phrasing*, New York: Oxford Univ. Press, pp.172-200.
- Werker, Janet, Ferran Pons, Christiane Dietrich et al. (2007) “Infant-directed speech supports phonetic category learning in English and Japanese”, *Cognition*, 103:1, pp.147-162.