

## [招待講演] 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—

五十嵐 陽介<sup>†</sup> 馬塚 れい子<sup>‡</sup>

<sup>†</sup> 日本学術振興会, 国立国語研究所, 理化学研究所 〒190-8561 東京都立川市緑町 3591-2

<sup>‡</sup> 理化学研究所, Duke University 〒351-0198 埼玉県和光市広沢 2-1

E-mail: <sup>†</sup> yosuke.igarashi@kokken.go.jp <sup>‡</sup> mazuka@brain.riken.jp

あらまし 乳幼児は母親や周囲の人間の話し言葉を聞くことにより言語を獲得する。言語獲得の入力となるこのような話し言葉は、「マザリーズ」(motherese)あるいは“Infant-Directed Speech”(IDS)と呼ばれる。IDS は大人に向けて話される言葉と異なることが知られているが、言語発達の研究のためには、IDS の性質を理解することが不可欠となる。本講演では、我々が言語発達の研究目的で構築した『理研日本語母子会話コーパス』を紹介する。この音声データベースは、母親 22 人による自分の子供(17~24 ヶ月)に向けた発話、および大人に向けた発話を格納している。このコーパスは、約 14 時間の音声信号とともに、転記テキスト、分節音情報、形態論情報、韻律情報を検索可能な形で与えている。講演の最後では、コーパス使用法の例として IDS の発話速度の分析結果を示す。

キーワード 音声コーパス、マザリーズ、infant-directed speech、発話速度

## How Does Japanese Motherese Differ from Adults' Speech? —RIKEN Japanese Mother-Infant Conversation Corpus—

Yosuke IGARASHI<sup>†</sup> Reiko MAZUKA<sup>‡</sup>

<sup>†</sup> JSPS, NIJLA, RIKEN BSI 3591-2 Midori-cho, Tachikawa-shi, Tokyo, 190-8661 Japan

<sup>‡</sup> RIKEN BSI, Duke University 2-1 Hirosawa, Wako-shi, Saitama, 351-0198 Japan

E-mail: <sup>†</sup> yosuke.igarashi@kokken.go.jp <sup>‡</sup> mazuka@brain.riken.jp

**Abstract** Infants cannot acquire language without hearing the speech that mothers and other adults produce around them. The speech input from which infants learn language is called “motherese” or “Infant-Directed Speech” (IDS). It has been well documented that IDS has a number of characteristics that are clearly distinct from adults' speech when they talk to other adults (“Adult-Directed Speech” (ADS)). In order to investigate language acquisition, it is necessary to understand the nature of IDS. In this talk, we will introduce the RIKEN Japanese Mother-Infant Conversation Corpus, which contains 22 mothers' conversations with their 17-24 month old children as well as with adults. This speech database consists of approximately 14 hours speech signals along with the various annotations (transcriptional texts and segmental, morphological and prosodic information), all of which are linked to each other in a searchable way. For purposes of illustration, we will present a comparison of speech rates between IDS and ADS.

**Keywords** speech corpus, motherese, infant-directed speech, speech rate

### 1. 導入

母親や周囲の人間の話し言葉を聞くことによって乳幼児は言葉を学ぶ。言語獲得について生得説の立場をとろうと経験説の立場をとろうと、言語獲得の入力となるこのような言葉が与えられなければ、乳幼児は言語を獲得できないという事実は揺るがない。したがって言語発達の研究のためには、言語獲得のために乳幼児が利用する入力に関して正確な理解を持つことが重要となる。

乳幼児に対する話し方は「マザリーズ」(motherese)、あるいは“Infant-Directed Speech”(IDS)と呼ばれるこ

とがあり、それは大人に対する話し方、すなわち“Adult-Directed Speech”(ADS)とは様々なレベルで異なっていることが知られている[10]。IDS の特徴としてよく知られているのは、全体的にピッチが高くなることやピッチ幅(pitch range)が拡大することである[2, 3]。IDS はまた、ポーズがより長く、より頻繁に生じることや、発話がより短く単純であることが知られている[3, 4]。発話速度が遅いことも IDS の特徴であると主張されている[11]。このように言語発達研究者は、IDS は ADS と明らかに異なる特徴を有することを示しており、言語発達理解には IDS の分析が不可欠であ

ることを認識している。

日本語の言語発達を研究するためには、日本語 IDS などのような音声的・言語的特徴を持つかを分析することが不可欠である。もし、音声信号と種々の言語学的情報を格納する日本語 IDS の音声データベースが利用できれば、非常に有益であることは間違いない。しかしながら、日本語 IDS の特性を様々な側面から分析するという目的に適い、かつ質・量ともに十分なデータベースは今までなかった。

もちろん、日本語 IDS のデータが存在しないというわけではない。例えば野路[7]や大久保[9]は乳幼児と両親との会話の膨大な転記(transcription)データを提供している。日本語の CHILDES プロジェクト[8]は乳幼児と大人との自然な会話の優れたソースである。

しかし、既存のデータベースが提供するの文字に書かれた転記のみであるのが普通であり、これを利用して IDS の音声的特徴を分析することはできない。CHILDES のデータベースの一部は音声信号をも提供しているが、種々の言語学的情報を検索可能な形で提供していない。

我々は日本語の言語発達の研究を目的として『理研日本語母子会話コーパス』(以後『母子会話コーパス』)を構築している。『母子会話コーパス』は、母親による IDS と ADS の発話を格納した音声データベースであり、音声信号とともに、転記テキスト、分節音情報、形態論情報、韻律情報という研究に必要な情報が検索可能な形で与えられている。第 2 節では『母子会話コーパス』の概要を述べる。

## 2. 『理研日本語母子会話コーパス』

### 2.1. 録音

録音に参加したのは、いわゆる標準語圏(東京・神奈川・千葉・埼玉)出身の母親 22 人とその子供である。子供の月齢は 17~24 ヶ月とした。この月齢の子供はなんらかの言葉を発することができるので、(母親によるモノログではなく)子供との会話が得られることになる。また、この月齢の子供の発話は限定されているので、母親の発話が子供によって中断されることが少なくなるという利点もある。

録音は理化学研究所の防音実験室で行われた。IDS の録音は、実験室に母親とその子供のペアのみが入った状態で行ったが、ADS の録音は母親、子供、そして母親の話し相手となる実験者の計 3 人が入った状態で行った。

母親の発話は、ヘッドセット式ダイナミックマイクローフオンを用いてモノラルで録音した。同時に母親の話し相手(IDS 録音の場合は子供、ADS 録音の場合は

実験者)の発話を納めるためにコンデンサマイクローフオンをテーブルの上に置いてモノラルで録音した。発話は DAT に収録されたが、母親の発話は左チャンネルに、話し相手の発話は右チャンネルに納められるようにしてある。加えて、天井に設置されたカメラを用いて動画撮影も行った。動画は mini DV に収録された。

IDS の録音のために 2 種類の緩やかなタスクを課した。ひとつは絵本を使って遊んでもらうタスクであり、もうひとつは用意されたおもちゃを使って遊んでもらうタスクである。双方のタスクとも所要時間は約 15 分である。はじめに実験室にある本を用いて遊んでもらい、15 分後に実験者が本を片付け、代わりにおもちゃを入れて、再び遊んでもらうという手順で行った。

ADS の録音は、上記 2 種類のタスクの終了後、同年齢の子供を持つ女性の実験者ひとりが実験室に入り、母親と話をすることで行った(子供は実験室に残ったままである)。所要時間は約 10 分である。会話の内容は特に指定しなかったが、主として育児に関する会話が行われた。

録音された発話は 44kHz、16 ビットでデジタル化された。

ひとりの母親あたり約 45 分、全体で約 14 時間の録音が行われた。IDS は約 11 時間(約 5 万語)、ADS は約 3 時間(約 3 万語)である(ここで「語」とは後述する「短単位」を指す)。

### 2.2. 研究用付加情報

『母子会話コーパス』には以上の方法で録音した音声信号とともに、言語研究に必要な種々の研究用付加情報(アノテーション)が与えられている。

研究用付加情報の付与は、『日本語話し言葉コーパス』(以後 CSJ)構築のために開発された手法に準拠している[6]。ただし与えられる情報は、転記テキスト、形態論情報、分節音情報、韻律情報のみである。CSJ が提供する節単位情報、係り受け構造情報、要約・重要文情報、談話境界情報、印象評定データは与えられない。情報付与の体系は必要に応じて(特に母子会話という特性に応じて)縮小・拡張している。なお子供の発話には形態論情報、分節音情報、韻律情報は与えられない。

#### 転記テキスト

転記テキストとは会話全体を文字で書き起こしたものである。まず音声信号を 200ms 以上の物理的なポーズで区切られた区間で分割し、これを「転記基本単位」とした。この単位内の発話を一貫した基準で書き起こした。転記テキストは片仮名による表記を行っており、CSJ のような平仮名、片仮名、漢字を使用した

表記は行っていない。

文字化された情報には、フィラー(filled pause)、いい誤り、いい淀みなどによって生じた語断片などの非流暢性(disfluency)を表すタグなども付与されている。転記基本単位とは独立に、笑い、咳、息などの非言語的なイベントなどの情報も付与されている。『母子会話コーパス』独自のものとして、母親が歌っていることを示すタグなども付与されている。なお、子供の発話にはおおよその文字化が行われているだけであり、その他の詳細な情報は付与されない。

転記テキストの体系の詳細は[6, chap.2]参照。

### 形態論情報

形態論情報の付与のためには語に相当する単位の同定が重要である。当該の単位として CSJ と同様に長短 2 種類の単位、すなわち「短単位」と「長単位」を同定した。短単位は単一あるいはふたつの「最小単位」(およそ形態素に相当)の結合からなる単位である。長単位はおおよそ複合語に相当する単位である。例えば、「アンパンマン」は、短単位としては「アンパン」と「マン」というふたつの単位として、長単位としては「アンパンマン」という単一の単位として分析される。

すべての短単位および長単位には品詞、活用の種類、活用形などの情報が付与されている。品詞情報はいわゆる学校文法に依拠している。

『母子会話コーパス』では、IDS で多用されるとされる擬音語・擬態語などのオノマトペ、「あんよ」「ねんね」などのいわゆる育児語(nursery word)も検索可能のように形態論情報のラベリングを拡張した。

形態論情報の体系の詳細は[6, chap.3]参照。

### 分節音情報

コーパスには子音や母音に関する分節音情報も与えられている。分節音ラベルは CSJ の体系に従っている。この体系では基本的に音素記号が用いられているが、母音の無声化や/i/の前の硬口蓋化(palatalization)、閉鎖音・破擦音の閉鎖区間などの音声レベルの情報も与えられている。

長母音(第二要素)、促音、撥音といういわゆる特殊モーラは、それぞれ“H”, “Q”, “N”というラベルで表記される。例えば「学校」には“gaQkoH”というラベルの連鎖が認定される。なお、非語彙的な母音・子音の伸張は長音・促音とは区別される。非語彙的な伸張は転記テキストのタグにより表記され、分節音ラベルでは表記されない。例えば「行くー?」のような発話末の母音の引き伸ばしは、転記テキストでは“iku<ー>”のように表記されるが、分節音ラベルでは“iku”というラベルの連鎖として表記される。

個々の分節音ラベルは CSJ の基準に従って、音声信号と同期される。

分節音情報の体系の詳細は[6, chap.6]参照。

### 韻律情報

ここで韻律情報とは、アクセント核(語彙的ピッチアクセント, lexical pitch accent)とイントネーションに関する情報のことである。日本語の韻律ラベリング体系としては、J\_ToBI (Japanese Tone and Break Indices)がよく知られている[12]。CSJ では、J\_ToBI の問題点を解消しさらに自発音声に適用可能なように拡張した X-JToBI というモデルを用いて韻律情報が与えられている。『母子会話コーパス』でも X-JToBI を用いて韻律情報を付与した。

X-JToBI は、イントネーションパターンを記述する「トーン層」、発話を構成する語を記述する「単語層」、韻律的な境界の深さを記述する「BI 層」と、その他の情報を記述する層という複数の層からなるラベルで韻律情報を記述するモデルである。単語層を構成する語としては、CSJ と同様に形態論情報における短単位を用いている。アクセント核の位置は単語層に記述される。

韻律情報を利用することによって、例えばトーン層からはどのような句末境界音調(韻律的フレーズの終端に生じるピッチ運動)が生じているのかなどを知ることができる。同様に、単語層からはどのようなアクセント型が用いられているかなどを、BI 層からは発話がどのように韻律的に分割されているかなどを知ることができる。

『母子会話コーパス』では作業時間上の制限から、X-JToBI を以下のように単純化した。第 1 に、トーン層のラベルは音声基本周波数曲線上における物理的実現位置と対応させない(これは旧 J\_ToBI[12]の体系に近い)。第 2 に、単語層・BI 層から自動的に派生可能なトーン層のラベルは付与しない。

韻律情報の体系の詳細は[6, chap.7]参照。

## 2.3. コーパスの構造

『母子会話コーパス』におけるすべての研究用付加情報と音声信号は、時刻情報によって関連付けられるため、種類の異なる情報に跨ったデータ検索が可能である。例えば韻律情報と形態論情報を参照して、「上昇調の句末境界音調が生じている内容語」のみを取り出し、分節音情報を参照してその語の最終音節の母音の生起時刻を抽出し、音声信号と関連付けることによって、「内容語に生じた上昇音調の句末境界音調」の音響特徴が計測可能となる。

現時点では、形態論情報は Microsoft Access に格納し、転記テキスト、分節音情報、韻律情報は Praat [1]

の TextGrid ファイル形式で格納している（転記テキストはアルファベットに変換してある）。コーパス完成時に情報をどのような形式で格納するかは未定である。Access に格納された形態論情報の検索や修正は、表計算ソフト Microsoft Excel で行っている。このための仕組みは Excel のマクロで実現している。各情報の検索・修正の際、必要に応じて音声・動画が参照できるよう、指定された範囲の音声・動画を再生する機能や、Praat を立ち上げ韻律情報を表示する機能などを持っている。

## 2.4. 現状

コーパスはまだ構築作業中である。転記テキスト、形態論情報、韻律情報の付与はおおよそ終了しているが、個々の分節音を音声信号に正確に同期させる作業が進行中である。完成までには各情報間の整合性を確認する作業を行う必要がある。

## 2.5. 『理研日本語母子会話コーパス』の限界

『母子会話コーパス』にはいくつかの限界がある。第 1 に規模が限定されている。例えば CSJ の「コア」と呼ばれる一部（詳細な情報が付与されている部分）が約 50 万語であるのに対して、『母子会話コーパス』は IDS, ADS あわせて約 8 万語である。言語発達の研究目的に作られた他のコーパスと比較すれば、十分に大規模なものといえるとはいえ、コーパスの分析結果の解釈には十分な注意が必要である。

第 2 に本コーパスには（狭義の）統語論情報を与えていないため、統語分析を行いたい場合は十分な情報を得ることができない。例えば、日本語では格助詞が省略されることが知られているが、格助詞省略に関して IDS と ADS にどのような差異が観察されるかは言語発達の研究者の主要な関心事のひとつであろう。本コーパスを用いることによって、格助詞を伴わない名詞がどれくらい出現するかを計測することは原理的に可能であるが、省略された格助詞が何であるかを知ることができない。例えば「クマさん見た」という発話では、省略された格助詞が「を」なのか（「クマさんを見た」）それとも「が」なのか（「クマさんが見た」）を判断することは不可能である。これは（格助詞によって明示的に表示されていない）格関係という統語論情報を本コーパスは与えていないからである。

## 3. 『理研日本語母子会話コーパス』の使用例： 日本語 IDS, ADS の発話速度の比較

### 3.1. IDS の発話速度

IDS の発話速度が ADS と比較して遅いということが、IDS の有する普遍的特徴であると主張されている [11]。日本語 IDS の発話速度も同様に ADS よりも遅いのであろうか。日本語はいわゆるモーラの等時性 (mora isochrony) を有する言語であるので、IDS における継続長の制御が、今まで発話速度の報告されている言語（例えば英語）などと異なっていなくてもなんら不思議はない。類型論的に異なる日本語のような言語を分析することで、普遍性に関する主張の妥当性を検証することができる。我々は『母子会話コーパス』を用いて日本語 IDS の発話速度を分析した [5]。

### 3.2. コーパス分析

発話速度を、発話の継続長、会話全体に対するポーズの割合、およびモーラの継続長という尺度を用いて計測した。まず発話の継続長を計測した。これは 200ms 以上のポーズで区切られた単位（転記基本単位）の継続長を計測して求めた。結果は、発話の継続長の平均値は IDS よりも ADS の方が有意に長いことを示した。次にポーズの割合であるが、IDS のポーズは会話全体の 60% 以上を占めるのに対して、ADS は 50% 以下にとどまることがわかった。以上の結果から、IDS は ADS と比較して、発話はより短く、ポーズはより長くより頻繁であるということになる。発話が短いこと、ポーズが長く頻繁に生じることが発話速度を示す妥当な尺度であるならば、日本語の IDS も、報告されている他の言語と同様に、より遅い発話速度に特徴付けられるということになる。

次にモーラ数の継続長であるが、これはおおよそ調音の速さを表す尺度となるものである。分節音情報の付与がまた終了していないので、モーラの平均継続長は、語（短単位）の継続長を、語を構成するモーラ数で割ることによって間接的に求めた（フィラーや語断片は分析の対象外となる）。平均値を比較すると、IDS のモーラの継続長は ADS よりも有意に長いことがわかった。この結果は一見、調音の速さという観点から見た日本語 IDS の発話速度もまた、他の言語と同様に遅いことを示すものと思われる。しかしこの結果は誤解を招くものである。

上に論じたように IDS の個々の発話は ADS よりも短い。このことはすなわち、IDS では発話終端に位置するモーラの数 ADS と比べて多いことを意味する。発話終端に限らず何らかの韻律的フレーズ (prosodic phrase) の終端のモーラないし音節には、いわゆる final lengthening が生じることがある。Final-lengthening をコントロールせずにモーラの継続長を計測することは、IDS のモーラの平均継続長を長めに見積もってしまう

ことになる。

そこで我々は、韻律的なフレーズ内部のモーラの位置を考慮に入れてモーラ継続長を計測した。『母子会話コーパス』には X-JToBI を用いた韻律情報が付与されているので、この情報を利用することができる。X-JToBI では韻律的フレーズとして「アクセント句」(accentual phrase)とそれより階層的に上位の「イントネーション句」(intonation phrase)が認められている。この韻律情報を利用して、フレーズ終端以外、アクセント句終端、イントネーション句終端にそれぞれ位置するモーラの継続長を別々に計測した。

IDS, ADS とともに、かなりの程度の final lengthening が観察された。イントネーション句終端に位置するモーラ継続長の平均値は約 210ms であったが、それは終端以外に位置するモーラ (約 110ms) およびアクセント句終端に位置するモーラ (約 130ms) より有意に長いものであった。一方、IDS と ADS のモーラ継続長を比較すると、それらいずれの位置においても、有意差が認められなかった。

『母子会話コーパス』の分析結果は、調音の速さ (どれくらいの速さでモーラが調音されるか) という尺度で計測する限り、日本語 IDS の発話速度は ADS よりも決して遅いわけではないことを示した。この結果は以下の 2 つの点で重要である。

第 1 に、この結果によって、IDS は普遍的に遅い発話速度を有するとする主張の妥当性が疑われることになる。確かに日本語 IDS の発話は短く、ポーズが長く頻繁に生じるという事実は、IDS の発話速度が遅いという印象を与えるであろう。しかし、ポーズを除いた発話部分における調音の速さを計測してみると、日本語の IDS の発話速度は、ADS より遅いわけではないことがわかる。

第 2 に、発話速度の結果のボタンが、イントネーション句の終端とそれ以外の位置のモーラを別々に計測した場合にのみ明らかになったことは重要である。もし (IDS が遅い発話速度を有するとする) 過去の報告が、IDS に final lengthening が頻繁に生じる事実を勘案していないのならば、IDS の音節ないしモーラの継続長を長めに見積もりすぎている可能性がある。このような韻律的なフレーズ (prosodic phrasing) をコントロールした分析は、『母子会話コーパス』に付与されているような韻律情報などの種々の研究用付加情報を利用して初めて可能となる。

## 文 献

- [1] P. Boersma and D. Weenink, Praat: doing phonetics by computer (version 4.4.24) [Computer program] <http://www.praat.org/>, 2006.
- [2] A. Fernald and T. Simon, "Expanded intonation contours in mothers' speech to newborns," *Developmental Psychology*, 20, pp.104-113, 1984.
- [3] A. Fernald, T. Taeschner, J. Dunn, M. Papoušek, B. de Boysson-Bardies, and I. Fukui, "A cross-language study of prosodic modifications in mothers' and fathers' speech to preverbal infants," *Journal of Child Language*, 16, pp.477-501, 1989.
- [4] D. Grieser, and P. Kuhl, "Maternal speech to infants in a tonal language: Support for universal prosodic features in motherese," *Developmental Psychology*, 24, pp. 14-20, 1988.
- [5] Y. Igarashi and R. Mazuka, "Speech rate of infant-directed speech is NOT slower than adult-directed speech, Poster presented at International Conference of Infant Studies, Kyoto, Japan, June. 2006.
- [6] 国語研, 国立国語研究所報告 124: 日本語話し言葉コーパスの構築法, 独立行政法人国立国語研究所(編), 東京, 2006.
- [7] 野地潤家, 幼児期の言語生活の実態, 文化評論出版, 広島, 1974-1977.
- [8] 宮田 Susanne, 今日から使える発話データベース CHILDES 入門, ひつじ書房, 東京, 2004.
- [9] 大久保愛, 幼児のこことば資料, 秀英出版, 東京, 1981.
- [10] C. Snow and C. Ferguson, *Talking to Children: Language Input & Acquisition*, Cambridge University Press, London, 1977.
- [11] L. Trainer, C. Austin and R. Desjardins, "Is infant-directed speech prosody a result of the vocal expression of emotion?," *Psychological Science*, 11, pp.188-195, 2000.
- [12] J. Venditti, "The J ToBI model of Japanese intonation," *Prosodic Typology: The Phonology of Intonation and Phrasing*, ed. S.A. Jun, pp. 172-200, Oxford University Press, New York, 2005.