特別講演

親子間相互作用が結ぶ言の葉のはじめ —認知発達ロボティクスからのアプローチ—

浅田 稔

大阪大学大学院工学研究科

Beginning of words emerges from caregiver-infant interactions —An approach from cognitive developmental robotics—

Minoru Asada

Osaka University

Key words: cognitive developmental robotics, physical embodiment, sociality, mother-infant speech interaction, sensori-motor magnet bias, auto-mirroring bias

Abstract: This article summarizes the special talk entitled “Beginning of words emerges from caregiver-infant interactions” at 55th Scientific Meeting of the Japanese Society of Stomatognathic Function. First, cognitive developmental robotics (hereafter, CDR) which the speakers have been advocating for more than 15 years was introduced. The key ideas of CDR are “physical embodiment,” and “sociality” which can be seamlessly connected through “development.” Symbol emergence also shares these ideas, especially, in case of speech perception and vocalization since the process from random cooing and babbling to vowel utterances of infants can be regarded as a process of symbol emergence through the interaction with their mothers. Next, computational approaches are shown, especially focusing on mother-infant speech interaction. Not only the learning algorithms on infant (robot) side, but also how mothers interact with their infants, which could be a sort of entrainment, is discussed. Simulations based on a sensori-motor magnet bias and an auto-mirroring one show how these biases affect the process of vowel acquisition of infant. Finally, future issues are shown.

抄録 本稿は、日本顎口腔機能学会第55回学術大会において、『親子間相互作用が結ぶ言の葉のはじめ』と題した特別講演の内容をとりまとめたものである。最初に講演者らが提唱し、推進してきた認知発達ロボティクス^{1,2)}の概要を説明した。認知発達ロボティクスのキーアイデアは「身体性」と「社会性」であり、これらをシームレスで結びつける基盤が発達である。

講演では、身体性、社会性、発達をキーにした認知発達の凝縮課題の一つが記号創発であると捉え、その研究分野の一例として、「音声の知覚と発声の発達」課題をとりあげた。乳幼児のランダムなクーイング期（およそ4ヶ月頃）やバブリング期（およそ5～6ヶ月頃）から母音発声に至る過程は、母親とのさまざまな相互作用を通じた記号創発過程と見なせる。最初に、認知発達ロボティクスの観点から、母子間

相互作用による言語獲得過程に焦点をおいた音声の知覚と発声に関する基本課題を示した。次に、認知発達ロボティクスで行われてきた関連研究を紹介した。母親が持っているバイアスとして、知覚範疇化や構音機構の拘束によるマグネット効果に加え、自身を模倣してくれるという期待から生じる解釈の偏り（自己鏡映バイアス）のモデルを紹介し、いかに身体性や社会性が密に関連しているかを示し、最後に、今後の課題を議論した。

キーワード：認知発達ロボティクス，身体性，社会性，母子間相互作用，マグネット効果，自己鏡映バイアス

I. 母子間相互作用による言語獲得過程の課題

1. 音声模倣による共通音韻獲得課題

乳幼児の音声コミュニケーション発達過程において、乳児（あるいはロボット）と養育者は、互いに異なる身体で声を発する。異なる身体から発せられる音声は、当然異なる音響特性をもつ。また、自身が発した音声は、自身が聞くようには、他者は聞かない。理由は以下である。

人の音声は、声帯に呼吸が送り込まれることにより生じた空気の振動が、声道を通過する際に共鳴することによって産出されるため、その音響特性を表すフォルマント（共鳴周波数）は声道がどのように形作られているかによって決まる。フォルマントは、母音識別に有効な特徴で、音声スペクトルのピークを指し、低い周波数から順に、第一フォルマント、第二フォルマントと呼ばれる。霊長類のコミュニケーションにも利用され、そのため生得的に埋め込まれているようである⁵⁾。従って、身体の違いのため、両者が発することができる声の音響特徴の範囲は異なり、養育者の声と物理的に同じ音響特性を再現しようとしても、それが産出可能範囲内のものであるとは限らない。

人は体内にある鼓膜の振動として音を聴取するため、相手の声についてはその気導音のみが知覚されるのに対し、自身の声については気導音と体内を伝わってきた骨導音とが合わさった形で知覚される。従って、自分自身にとっては相手と物理的に同じに知覚される声を産出できたとしても、それから骨導音が除かれた気導音を聴取する他者にとっては、同じと知覚される音になっているとは限らない。従って、物理的な音響特徴そのものを模倣の指標とするのではなく、乳児は何が対応するのかをどのように発見可能かが重要な問いとなる。

これらのことから、乳幼児（あるいはロボット）と養

育者の間での音声相互模倣においては、音響的な整合性では、模倣が成立せず、先に述べた二つの課題、すなわち、自身の音声知覚と発声の対応付け、及び養育者の音声と自身の音声の対応付け（音韻の共有化）を同時に解決する必要がある、その意味で、養育者の関わり方は重要である。

表1に、乳児の構音、知覚、母親とのインターアクションの発達の変化をまとめた。6ヶ月頃までに知覚範疇（マグネット効果）獲得準備が行われ、特に3～4ヶ月頃の相互模倣が後の相互作用にも効果を現していると察せられる。

II. 音声の知覚と発声の発達における身体性と社会性

認知発達ロボティクスでは、構成的手法により、人間の認知発達過程を構成論的に理解するアプローチである。特に、人間の発達の起点と想定される赤ちゃんを対象として、養育者との相互作用を通じた社会的行動獲得のメカニズムを明らかにすることが狙いの一つである。この裏には、自他認知発達の基本課題にアタックする流れがある。図1に、その考え方を示す⁷⁾。最初に自らの身体と環境との物理的な相互作用により、主に自己身体に関する表象が獲得され（自己の萌芽）、次に養育者との相互作用により、自他を同一視する機構が発達する（ミラーニューロンシステムの基盤）。そして、最後には、自他を分離し、他者との相互作用を制御可能な構造に発展すると考えられる。本稿で扱う「音声の知覚と発声の発達」課題では、自分の声と自分の声を産出するための構音運動を関連付けることが第一段階に対応し、生後の周りの大人とのやりとりの中から、親の声と自分の声の対応を見出すこと（音韻の共有化）が第二段階に対応する。ここでは、主に後者の状況を想定するが、両方含めて、身体性、社会性の意味について考えよう。

表1 乳児の構音, 知覚, 母親とのインターアクションの発達の変化 (文献6の表から適応)

	～一ヶ月	～二ヶ月	～三ヶ月	～四ヶ月	～五ヶ月	～六ヶ月	六ヶ月～
構音の発達			クーイング		過渡期の喃語 聴覚障害児の過渡期の喃語が消滅		
知覚の発達	母親と他の女性の声の区別	/a/と/i/, /u/と/u/の弁別	/b/と/p/, /d/と/g/, /i/と/i/の弁別	母親と異なる周波数で母親を模倣			声道長の正規化, マグネット効果
インターアクション	マザーリーズで語りかけ	乳児がマザーリーズを好む	親が乳児を高頻度で模倣	親の随伴的応答を促す. 母親が母音様の音声応答. 乳児の母音様のクーイングが母親の模倣を促進	母親の模倣が乳児のクーイングを促進	マザーリーズに回答音声返す.	

(I)生態学的自己

自己の萌芽

身体と環境の同調

(II)対人的自己

自他の同一視
(MNSの基盤)

養育者による同調

(III)社会的自己

自他の分離

養育者からの脱同調

図1 身体性から社会性に至る過程における自他認知の発達

1. 身体性

身体性に関しては、すでにいくつかの定義や議論がなされているが、浅田、國吉らは、「行動体と環境との相互作用を身体が規定すること、およびその内容。環境相互作用に構造を与え、認知や行動を形成する基盤となる。」と規定している⁸⁾。身体を感覚・運動・認知を支える物理的基盤と考えると、身体の物理的構造による拘束(形態)だけでなく、感覚器、運動器、認知の機能が相互にダイナミックに関連しながら、情報を生成し行動を規定していく。

音声の知覚と発声に限ると、ハードとしては、聴覚器官(マイクroフォン)、構音器官(人工声帯、声道形状制御機構)、これらの入出力データを処理し、保持するストレージなどが考えられる。これらは、予め準備され、備えられており、それらの配置は、当然、処理の形態に影響を与える。環境内の静的な物体を対象とした相互作用では、図1(I)のように、主に自己身体、すなわち、ここでは、自身の発声と音声の知覚の関連づけの学習に留まる。構音機構による発声範囲と聴覚機構による感覚空間によって、拘束される写像である。この段階では、環境情報の取得には限界がある。例えば、(叫び)声をあげて、その残響特性などから、環境の構造などを推定することを考える。その場合、首を振ったり、さらには、移動して壁にぶつかったり、触れることができる他の身体部位がないことには、環境の構造の意味づけが困難である。そもそも、そのような動機付けを仮定する正当性も充分でない。ここに社会性導入の理由がある。

2. 社会性

図1(II)に示す、養育者との相互作用では、養育者がさまざまな意味合いで、学習者(赤ちゃん)の発達の多くの可能性の中から、ある経路に絞り込んで行く過程

が存在する。その最重要基盤として、相互模倣がある。乳児の音韻発達において養育者に自身の発話を模倣されることが重要な役割を果たしているとし唆されているが、その具体的な機能については明らかでない。これに対し、筆者らは養育者による模倣に二つの役割があるとの仮説を提案している。

- (1) 乳児の行動に対応する養育者の行動、すなわち身体構造の違いを吸収する対応づけを正しく学習するための正解例を乳児に例示する働き
- (2) 養育者が乳児の発声を再現する時に身体構造の違いのため完全には再現できず、無意識のうちに自身の普段よくする行動、つまり母音の発話に置き換えた模倣をすることで、乳児の発話カテゴリを母音に誘導する働き

である。前者は教示行動の表れであり、他の霊長類にはあまり見られない明示的な社会的行動である。これに対し、後者は、身体性に拘束された中で、音声不一致が引き出す、養育者の引き込み行動である。もし、完全に音声的な整合(同じ音声が発声可能)がとれていれば、引き込みの必要もない代わりに、教示行動にもなり得ない。両者に共通するのは、親の愛着による明示的、非明示的な社会的行動であり、それにより、学習者の発達経路の複数の可能性を絞り込んでいく過程でもある(例えば、英語か日本語かなど)。このように、身体性に裏打ちされた自己の表象が、養育者の身体との相互作用により形作られる構造の性質や構造構築そのものの可能性をここでは、社会性と呼ぶ。

3. 認知発達

先の身体性で、ハードが予め準備され、備わっていると説明したが、厳密には、発達過程を考慮すると、感覚情報の表象、運動の制御性能、記憶容量などの能力は、

図2 音声の知覚と発声の発達

未熟な状態から始まる。むしろ未熟であることが、発達を促すとの知見もあり、その代表が範疇知覚であろう。6ヶ月以前は、連続母音をそのまま知覚可能であるが故に、世界のあらゆる母音に対応可能だが、6ヶ月以降は母国語の影響により離散的な母音識別に変化する⁹⁾。さらに発声自体もその影響を受けるとして、知覚運動マグネット効果と呼ばれている。また、社会性のところで説明した親の期待や思い込みを計算モデルとして具現化した自己鏡映バイアスも提案され、計算機シミュレーションにより、それらの有効性も検証されている¹⁰⁾。図2は、音声の知覚と発声の発達に関連する項目を図示した。これらを含め計算モデルを以下で紹介する。

Ⅲ. 身体構造の異なる他者との母音の対応学習

これまで概観してきた計算モデルは、母子間にみられる身体構造の異なるエージェント間の相互作用を対象としていない。ここでは、筆者らが、認知発達ロボティクスのアプローチとして、身体構造が異なる場合を想定した母子間相互作用における母音対応課題を扱った例を示す。

1. オウム返し教示による母音獲得学習

最初に行なった母音獲得学習¹¹⁾は、母子間相互作用モデルとして図3に示す構造を考えた。先行研究¹²⁾を参考に試作したロボットはランダムクーイングし、養育者の応答が入力となり、フォルマント抽出器により聴覚層にフォルマントベクトルが保持される。

構音層は音源からの音声をシリコンチューブを5つのモーターで変形することで発声するための構音ベクトルを保持する。聴覚層、構音層、それぞれ自己組織化によ

図3 母子間相互作用モデルによる音声模倣ロボットシステム

(a) 通常のヘブ学習による結果 (b) 修正されたヘブ学習による結果

図4 構音ベクトルの主要二成分の分布(詳細は本文参照)

りクラスタリングされると同時に二つの層間でヘブ学習によりクラスターを連結する。養育者は、ロボットが発声する音声がいずれかの母音に聞こえたならば、その母音を返すというオウム返し行動をとる。最初、母音の範疇も発声の仕方も知らなかったロボットが相互作用による学習を通じて、母音の範疇獲得、ならびに発声の仕方も同時に獲得する。

音声発声は通常、声道の形状で決まるフィルター関数によって音源が変調された結果だとみなされ、音声生成の「ソース-フィルター理論」と呼ばれている。ここでは、バイブレータを音源とし、シリコンチューブを5つのモーターで変形可能な声道に見立てた。

図4に実験結果を示す。通常のヘブ学習で獲得された構音ベクトルを(a)に示す。養育者がロボットに合わせて、自身の母音を様々に変えて応答した結果、ばらつきが多い。それに対し、(b)では、より少ない力やシリコンチューブのより少ない変形を好む修正ヘブ学習の結果を示す。ばらつきが少なくなっている様子がうかがえる。オウム返しという養育者からの社会的応答に加え、疲労が少ないなどの主観的基準に基づく学習過程が、声道というロボットの身体性拘束を反映した結果となっている。

図5 発話ロボット：Burpy

2. 相互模倣による音韻共有化モデル

前節の研究では、模倣は養育者側だけであり、しかもロボットは、養育者の応答音声しか聞かず、自身の聴覚フィードバックは含まれておらず、不自然であった。そこで、Miura et al.⁴⁾ は、口唇形状の視覚模倣も考慮した発話ロボット（図5）を試作し、その発話を養育者役の実験者に模倣させることを通じて、人が同じと思える母音の発話をロボットに獲得させる問題に取り組んだ。ロボットは、先のバイブレータによる音源ではなく、人工声帯に圧縮空気を流入することで基本周波数の音源を生成し、シリコン製の人工声道でこれを共鳴させることで音声を産出し、人工声道の喉・顎・舌にあたる筒の部分と唇にあたる開口部を電動モータにより変形させることで、様々なフォルマントの音を調音することができる。

ロボットと人それぞれが生成可能な音声のフォルマントの分布範囲をおおよそ対応づける一次変換、および人の口の形を大まかに再現する視覚的な模倣メカニズムにより、ロボットは観測した人の母音発話を、人によっては母音と聞き取れる程度に模倣することができる。しかし、どのような変形により生成される音が、より母音様であると人に思わせるものであるかは不明であった。Miura et al.¹³⁾ は、これをフォルマント空間上の発話プロトタイプを適応させる問題と捉え、人の発話音声とその人の典型的な発話音声とどのように異なるかを基に、発話プロトタイプを移動させる学習メカニズムを提案した。ロボットの発話が人に模倣されることを仮定したインタラクション実験を実施し、実際にロボットの発話する母音をより明瞭なものへと誘導できたことが報告されている。

実際の環境では、親の模倣率はGros-Louis et al.¹⁴⁾ によれば、2割の低頻度と報告されている。このような環境でも、母音獲得が可能なモデルとして、Miura et al.¹³⁾

図6 低頻度模倣環境での母音獲得学習モデル

図7 低頻度模倣環境での母音獲得学習過程

は、声域の異なる親（養育者）と乳児ロボット（学習者）の間で、親の母音候補に対して多数のプリミティブ用意し、ランダムに配置して（図6）、それらのうち幾つかがより自然な母音の方向に収束するように学習した。より正確には、乳児に養育者の模倣を引き出しやすい音を積極的に発話するモチベーションを与えることで、インタラクションを繰り返すごとに乳児の発話が母音様へと近づいていくことが示された。図7のその結果を示す。当初はプリミティブが親の5つのフォルマントベクトルの重心近くに集中していたが、後期では、プリミティブがほぼ親の自然な母音のフォルマントベクトルに到達している様子が分かる。

3. 養育者の模倣メカニズム

Miura et al.⁴⁾ の研究では、本来未知であるはずの身体構造の対応付け、すなわち人とロボットの声域をおおまかに対応付ける写像が与えられていた。また、人の模倣が誘導的な働きを持つことが議論されていたが、どのようなしくみで人がそのような模倣を示すのかについての

図8 養育者の知覚運動バイアスと自己鏡映バイアス

議論は十分ではなかった。模倣一般に言われているように、人は何かを模倣する際、感覚器からの信号をそのまま模写し、行動するわけではなく、なんらかの知覚運動バイアスがかかると考えられる。Ishihara et al.¹⁰⁾ はさらに相互作用がフィードバックを増幅する自己鏡映バイアスという概念を提案し、計算機シミュレーションによって、二つのバイアス（図8）が乳児の、身体構造の対応付けの学習と並行した音韻共有化過程に及ぼす効果を検討している。

前者の知覚運動バイアスは、乳児の発話をより母国語母音らしい音として解釈するバイアスであり、これは母国語母音の典型例（/a//i//u//e//o/ のような母音プロトタイプ）を中心とした知覚バイアスである知覚のマグネット効果⁹⁾である。ただし、音声を構音する際にも自身の普段発声し慣れた母音へのバイアスが生じると考えられるため、これらを模倣におけるプロトタイプへのバイアスとしてまとめて扱い、以後知覚構音バイアスと呼ぶ。

もう一つのバイアスは、乳児の応答をより精度のよい模倣として解釈するというものである。これは、相互模倣の最中に親は乳児に模倣されることを期待しており、その期待に沿うように乳児の応答を解釈するという仮定に基づくものである。以下ではこれを自己鏡映バイアスと呼ぶ。これは、プラシーボ効果に通じるもので、被験者実験を通じて検証した。コンピュータで生成した合成音に対し、被験者に単純に模倣してもらうグループと、模倣に加え、「コンピュータも被験者を模倣する場合がある」と説明を受けたグループ間で、被験者の応答の時間的変化を計測した。後者のグループでは、相手が自分の模倣をしているという思い込みから、最初のある時刻で被験者が発声した音声を次の時刻でコンピュータが模倣したと錯覚し、その次の時刻で被験者がコンピュータ

図9 二つのバイアスを組み込んだ模倣システム（養育者側）¹⁰⁾

の模倣をすると、被験者の最初の時刻とその次の発声の変化が、バイアスがない場合に比較して、少なく、明らかなバイアス効果が観測された¹⁵⁾。

図9に、二つのバイアスを組み込んだ養育者側の模倣システムを示す。養育者は乳児（学習ロボット）の時刻 t の音声 $s(t)$ を聴き、その音声を構音 $a(t)$ で模倣する。この模倣過程は三つのモジュールの機能からなる。一つは、入力音声にバイアス进行自己鏡映バイアスモジュール (b)、二つ目は、マグネット効果を持つ知覚構音バイアスモジュール (f)、そして、三つ目は、自己鏡映による期待を計算する期待モジュール (g) である。先に述べたように、「自己鏡映による期待」は、他者の音声自身が自身の音声の模倣であるとの期待から、他者の音声自身が自身の音声に似ているという期待にバイアスされた知覚と定義される。

他者の音声 $s(t)$ は、自己鏡映による期待 $s^g(t-1)$ にバイアスされ、 $s^b(t)$ に変換される。この過程は以下で与えられる。

$$s^b(t) = b(s(t), s^g(t-1); \eta) \quad (1)$$

$$= s(t) + \eta (s^g(t-1) - s(t)) \quad (2)$$

ここで、 η ($0.0 \leq \eta \leq 1.0$) は、自己鏡映バイアスの強度を表すパラメータであり、 η が 0 に近いと、出力 $s^b(t)$ は、自己鏡映バイアスが弱く、元の入力 $s(t)$ に近づく。逆に η が 1 に近いと自己鏡映バイアスが強く、出力 $s^b(t)$ は、自己鏡映による期待音声 $s^g(t-1)$ に近づく。

図7に示すように、養育者と乳児のそれぞれの発声可能な音域は異なるので、養育者によって知覚された乳児の音声は、養育者自身の発声可能な構音パラメータに変換されなければならない。ここでは、正規化ガウス関数ネットワーク (Normalized Gaussian Network, 以

図10 ただ一つの回帰行列を持つ ($M = 1$) NGnet による知覚構音マグネットの表現 (文献10より)

図11 知覚運動バイアスと自己鏡映バイアスの違いによる相互模倣発話分布の違い (本文参照)

下, NGnet と略称する) を用いて, 乳児の発声可能領域から養育者の発声可能領域へ写像する. NGnet は, M 個の線形写像の組み合わせで非線形関数を近似するものであり, 線形写像の組み合わせの重みを決定するための M 個のガウス関数 $g_i (i = 1, \dots, nM)$ が入力空間に配置されている. ガウス関数と線形写像は 1 対 1 の対応関係を持っており, 入力と各々のガウス関数の中心との距離によって, そのガウス関数に対応する線形写像の重ね合わせる際の重みの大きさが決定される. すなわち, 親の写像 f の NGnet の場合, 入力である知覚 $s^b(t)$ は,

$$a(t) = f(s^b(t); p_i; \lambda) \quad (3)$$

として, 構音 $a(t)$ に変換される. p_i は養育者の i 番目のプロトタイプであり, λ は, 知覚構音バイアスの強度を示している. 図10は, $M = 1$ の NGnet により知覚構音マグネットがどのようにモデル化され, また λ の設定

によってバイアスの強さがどのように調節されるかを示している. 横軸が 1 次元の入力空間であり, いくつかの入力が NGnet のガウス関数の中心ベクトル周りに正規分布状に与えられ, それぞれの入力が線形回帰行列 $\tilde{w}_1$ で 1 次元の出力空間に写像されるとする. この時, 出力データの分布は線形回帰行列の固有値 (この場合は線形写像の傾き) の大きさによって異なった結果となる. すなわち, 固有値 $(1-\lambda)$ が小さいほど, 出力データはガウス関数中心の像を中心に小さく収斂する. すなわち λ が大きいほど, 知覚構音バイアスが強いことを示している. 乳児の方の写像も NGnet で記述され, そのパラメータは EM アルゴリズムで推定される.

図11に, 二つのバイアスの強さを変えて相互作用させたときの, 養育者と乳児が発声した発話音声のフォルマントの分布を示す. 赤の分布が乳児の模倣音声であり, 5つの黒点は相互作用終了時の乳児の5つの母音プロトタイプの位置を示す. 青の分布は養育者の模倣音声

である．図11 (a) は，両方のバイアスが存在し，相互模倣を通じて，互いに正しい母音位置に収束している様子を示す．同図 (b) では，知覚運動バイアスにより収束しているものの，正しい位置ではない．同図 (c) と (d) では，自己模倣バイアスのみが作用している場合と二つのバイアスがともに無い状態で，ばらけた音を互いに適当に模倣しあっている．

IV. 終わりに

言語は，二重の意味でシンボリックである．アイコンニックからインデキシャル，そしてシンボリックに至る言語特有の参照形式の意味合い¹⁷⁾ と，言語という人間に特異的と思える認知機能の象徴的な意味合いである．本稿では，後者の意味で，記号創発ロボティクスの象徴的課題として，音声の知覚と発声の発達過程を取り上げた．認知発達ロボティクスのアプローチのキーワードである身体性，社会性，そして発達を，特に母子間相互作用に焦点を置いて再考した．音韻の共有化の発達過程の構成的理解を目指して，イメージング研究を含んだ発達心理の知見を紹介し，この過程のメカニズムを解明すべく，認知発達ロボティクスのアプローチを含む計算モデルを紹介した．

音声の知覚と発声の発達過程の実装は，身体性とその発達が鍵である．現状では，母音を中心としたフォルマント空間を用いた手法がメインであるが，今後，子音を含めた語彙獲得への拡張を考えると，子音の発声機構の実装，発声器官の発達の模倣が焦点となる．早稲田大学で開発されている WT シリーズは主に大人の機構であるので，乳幼児のサイズ及び構造の発声機構の開発が急務である．身体発達は，筆者の過去のプロジェクト (jeap.jp) でも，課題であった．そこでは，月齢に応じた赤ちゃんロボットのプラットフォームを複数開発し，月齢に応じた研究項目を実施した．それに倣えば，音声発達の大きなハードの変化，すなわち咽頭蓋の沈下前後のプラットフォームを開発し，実験することが考えられる．

語彙獲得は次のステップであるが，ここで紹介したいいくつかの方法も，実際の母子間相互作用のリアリティはまだまだ低く，相互模倣の低頻度だけではなく，実際の状況から想定される，さまざまなコンテキスト情報の利用は欠かせない．乳児だけでなく養育者，あるいは模倣だけでなく語彙に関するマッピングというように，乳児の発達を支える複数の要素の間の相互作用のモデル化を進め，実際の乳児に対しては倫理的には許されない操作

的な統制を施した計算機上の実験を実施することで，発達過程の理解に対して洞察を得，更に詳細なモデルを構築することが期待される．特に，視覚・聴覚だけでなく，触覚などのマルチモダルな情報の統合，物体操作の運動スキルの発達による物体モデルの深化やアフォーダンスの表象の発達も考慮する必要があるだろう（主に身体性の詳細化）．これらの研究を統合し，音韻と語彙の共発達における養育者の誘導的役割をモデル化することも重要である（主に社会性の詳細化）．これらの研究を遂行する上で，人との相互作用を通じて音声模倣および語彙を獲得するロボットの設計・製作は，認知発達ロボティクスのアプローチの観点からも重要であり，発声機構のための素材の選択，アクチュエータの開発なども急務である．これらの実装により，仮説検証，モデル再構築を繰り返し，初期言語発達の構成的モデルが構築されると期待される．

謝 辞

本稿をまとめるにあたり，吉川雄一郎博士（大阪大学准教授），三浦勝司君（現，(株)富士通），石原尚君（大阪大学特任研究員），笹本勇輝君（大阪大学大学院生）の協力を得た．ここに，深く感謝する．また，一部，科学研究費補助金特別推進研究（24000012）の補助を受けた．

参考文献

- 1) Minoru Asada, Koh Hosoda, Yasuo Kuniyoshi, Hiroshi Ishiguro, Toshio Inui, Yuichiro Yoshikawa, Masaki Ogino, and Chisato Yoshida. Cognitive developmental robotics: a survey. IEEE Transactions on Autonomous Mental Development, 1(1): 12-34, 2009.
- 2) 浅田稔. ロボットという思想—脳と知能の謎に挑む—. NHK ブックス (1158), 2010.
- 3) P. K. Kuhl and A. N. Meltzoff. Infant vocalizations in response to speech: Vocal imitation and developmental change. Journal of Acoustic Society of America, 100: 2415-2438, 1996.
- 4) Katsushi Miura, Yuichiro Yoshikawa, and Minoru Asada. Unconscious anchoring in maternal imitation that helps finding the correspondence of caregiver's vowel categories. Advanced Robotics, 21: 1583-1600, 2007.

- 5) Marc D. Hauser, Noam Chomsky, and W. Tecumseh Fitch. The faculty of language: What is it, who has it, and how did it evolve? *Science*, pages 1569-1579, 2002.
- 6) 三浦勝司. 乳児の明瞭な母音獲得を助ける自身の学習バイアスと養育者の認知バイアスに関する構成的研究. PhD thesis, 大阪大学大学院工学研究科, 2011.
- 7) 浅田稔. 共創知能を超えてー認知発達ロボティクスよる構成的発達科学の提唱ー. *人工知能学会誌*, 27(1): 2-9, January 2012.
- 8) 浅田稔 and 國吉康夫. *ロボットインテリジェンス*. 岩波書店, 2006.
- 9) P. K. Kuhl. Human adults and human infants show a “perceptual magnet effect” for the prototypes of speech categories, monkeys do not. *Perception & Psychophysics*, 50: 93-107, 1991.
- 10) H. Ishihara, Y. Yoshikawa, K. Miura, and M. Asada. How caregiver's anticipation shapes infant's vowel through mutual imitation. *IEEE Transactions on Autonomous Mental Development*, 1(4): 217-225, 2009.
- 11) Yuichiro Yoshikawa, Minoru Asada, Koh Hosoda, and Junpei Koga. A constructivist approach to infants' vowel acquisition through mother-infant interaction. *Connection Science*, 15(4): 245-258, Dec 2003.
- 12) T. Higashimoto and H. Sawada. Speech production by a mechanical model construction of a vocal tract and its control by neural network. In *Proc. of the 2002 IEEE Intl. Conf. on Robotics & Automation*, pages 3858-3863, 2002.
- 13) Katsushi Miura, Yuichiro Yoshikawa, and Minoru Asada. Vowel acquisition based on auto-mirroring bias with a less imitative caregiver. *Advanced Robotics*, 26: 23-44, 2012.
- 14) Julie Gros-Louis, Meredith J. West, Michael H. Goldstein, and Andrew P. King. Mothers provide differential feedback to infants' prelinguistic sounds. *International Journal of Behavioral Development*, 30(6): 509-516, 2006.
- 15) 石原尚, 若狭みゆき, 吉川雄一郎 and 浅田稔. 乳児母音発達を誘導する自己鏡映的親行動の構成論的検討. *認知科学*, 18(1): 100-113, 2011.
- 16) P. Kuhl, J. Andruski, I. Chistovich, L. Chistovich, E. Kozhevnikova, V. Ryskina, E. Stolyarova, U. Sundberg, and F. Lacerda. Cross-language analysis of phonetic units in language addressed to infants. *Science*, 277: 684-686, 1997.
- 17) Terrence W. Deacon. *The Symbolic Species: The co-evolution of language and the brain*. W. W. Norton & Company, New York, London, 1998.