

インタラク션을介した音韻意識獲得に関する認知モデルの検討 A Study of Cognitive Model about Phonological Awareness Acquisition through Interaction

西川 純平[†], 森田 純哉[†]

Jumpei Nishikawa, Junya Morita

[†] 静岡大学

Faculty of Informatics, Shizuoka University

bi16050@s.inf.shizuoka.ac.jp

Abstract

Humans acquire language through social work, that is, interaction between themselves and others. Especially in the process of mastering spoken language, it is important to be able to be aware of phonemes from continuous sounds. In this research, we aim to understand the process of acquisition of phonological awareness through interaction by modeling. For that purpose, we constructed a cognitive model that interacts using the cognitive architecture ACT-R, and performed simulations using that model. From the results of the simulation, factors inhibiting the phonological errors and a method to model underdeveloped phonological awareness has been suggested.

キーワード：音韻意識，認知アーキテクチャ，認知モデル

1. はじめに

言語は，社会的な営み，すなわち当人と他者とのインタラクシヨンのなかで習得される．典型的な言語習得プロセスにおいては，乳幼児は養育者のふるまいを観察し，その模倣を行うことで言語を習得していく [6]．このプロセスのなかで，乳幼児は，音の分節化のパターン，記号と対象の対応関係など，膨大なパラメータの値を，共同注視や役割交代を伴う模倣などの生得的に埋め込まれた社会認知的機能に誘導されながら推定している [2]．とくに音声言語習得の過程においては，音韻に関する意識が重要な役割を果たす．連続した音声のなかで音のくぎりに注目したり，それぞれの音を聞き分けたりする音韻の知覚機構は，生後約1年で母語の音韻体系に沿うかたちが形成され，4, 5歳までかけて確立されていくということが知られている．このような能力の発達を，一連の流れの中で，個人差を含めて理解し，発達支援につながる要素を検討するためには，計算機上にインタラクシヨンの一

連のプロセスをモデル化し，シミュレーションを通して観察することが有効であると考えられる．

これらを踏まえて，本研究では，インタラクシヨンを介した音韻意識獲得のプロセスをモデル化することを目指す．音韻に着目するインタラクシヨンの例として言葉遊びを扱い，モデル化のためのフレームワークに認知アーキテクチャACT-R(Adaptive Control of Thought-Rational)[1]を用いる．

本稿の構成は次の通りである．まず，2節にて本研究と関連する研究をレビューする．そののちに，本研究で実装を進めているプロトタイプモデルとモデルを用いたシミュレーションを示す．最後に現状のまとめと今後の課題を示す．

2. 関連研究

2.1 言葉遊びの利用

本研究では，音韻意識に関わるインタラクシヨンの例に言葉遊び，とくにしりとりを扱う．しりとりと学習に関する研究として，学習者の語彙モデルを構築する学習支援システムのなかで，タスクの一つとしてしりとりを課す山本と柏原によるもの [9] が存在する．また失語症の治療，あるいは自閉症の療育など，言語聴覚療法においても，しりとりは頻繁に用いられている．いくつかの論文の中で，療育中の自閉症児の検査にしりとりが利用されている [11, 13]．

高橋による研究 [8] では，幼児がしりとりを可能とする条件を定型発達の幼児を対象とした横断的な発達心理学実験を通して調査している．この調査は，しりとりを遂行するためには，音を音韻に分割する音韻意識や音韻による索引が付与された心的な語彙辞書が必要であること，音韻による語彙への索引付けに，かな文字の獲得が有効であることを示している．さらに，しりとりの遂行に必要な音韻意識を持たない子供であっても，ヒント提示などの大人の援助により，遊び

の活動に参加可能であることも示されている。このことから高橋は「子ども達はことば遊びの活動に最初は周辺的に参加して行く中で音韻意識が高まって行き、それを支えとして文字の読みを習得する、といった過程をたどる」と考察している。

本研究においても、しりどりのなかで、単語にあたる連続音から語尾音を切り出す能力や、語頭音によって単語を検索する能力としての音韻意識に着目してモデル化を行う。

2.2 認知アーキテクチャの利用

認知アーキテクチャは、個別の課題において生起する認知プロセスをモデル化する基盤である。認知アーキテクチャを利用したモデルにより、課題の達成に関わる要因を切り分けたモデルを構築できる。様々な認知アーキテクチャが開発されるなかで、本研究ではACT-R[1]を利用する。

ACT-Rは複数のモジュールを持つプロダクションシステムである。外界とのインタラクションを受け持つモジュールや、知識の検索、イメージの役割を果たすモジュールを持ち、反応時間の予測が可能である。また、モジュールの動作を規定する様々なパラメータが存在し、個人のモデル化を容易にしている。

ACT-Rを用いた言語の獲得に関する研究は多く行われている。英語の学習における不規則動詞の獲得に関わるモデル[5]、幼児による名詞の学習などのモデル[7]が構築されている。脳機能障害に関わる検討もなされており、失語症の文理解において生じるエラーをACT-Rのパラメータによって説明した研究も存在する[4]。

ACT-Rによるしりどりのモデルとして、著者ら[10]は、知識の活性値と音韻意識を対応づけるシミュレーションを行なった。その結果、音韻意識の高まりによってしりどりの継続数が増加すること、しりどりの遂行に伴って音韻意識が増強されることを示した。この結果は、先述の高橋による調査と整合的なものといえる。

3. モデル化とシミュレーション

3.1 モデルの構成

しりどりのプロセスを検討するため、モデルを構築した。図1にモデルの概観を示す。モデルの中で、同じ内部構造を持つ2体のエージェントが交互に単語を回答することでしりどりが進行していく。エージェント中のボックスはACT-Rの各モジュールに対応する。

以下に、ACT-Rのモジュール構造によって、しりどりのプロセスがどのように実現されるかを示す。

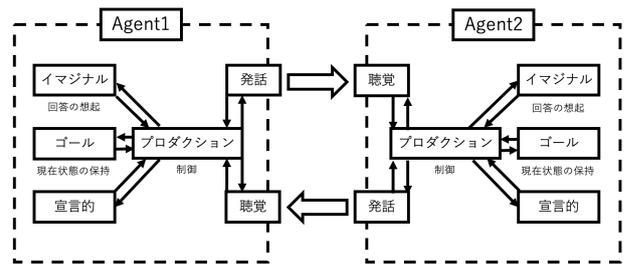


図1 モデル概観

3.1.1 宣言的モジュール

ACT-Rの宣言的モジュールを用いて、しりどりの遂行に必要な知識をモデル化する。ACT-Rの宣言的モジュールにおいて、知識はチャンクと呼ばれる構成要素からなる。本研究のモデルが保持するチャンクには、単語の知識（語彙）に関するものと、音韻の知識に関するものがある。前者のチャンクとして、単語の音情報を表す words、単語の語頭音の知識である word-heads、単語の語尾音の知識である word-tails というタイプを用意した。以下にそれぞれのタイプに含まれるチャンクの例を示す。

```
(ringo ISA words sound "ringo")
```

```
(word-head-ringo ISA word-heads
  meaning ringo
  head-char "ri")
```

```
(word-tail-ringo ISA word-tails
  meaning ringo
  tail-char "go")
```

各チャンクは、先頭にチャンク名が示され、その後スロット名と値の組が続く。ISA スロットの値によってチャンクの種類（チャンクタイプ）が示され、以降にチャンクタイプごとに共通のスロットを持つ。words をタイプとするチャンクは、sound という名前のスロットを持ち、単語の音の情報（“ringo”）を保持する。word-heads タイプは、meaning スロットに単語の知識、head-char スロットに語頭音の情報を持つ。word-tail タイプも同様に、単語の知識と語尾音を組み合わせた情報を保持している。ここで、words タイプのチャンクは word-heads および word-tails タイプの構成要素となっている。

本研究における ACT-R モデルは、この他に、音韻に関わる知識をチャンクとして持つ。以下にその例を示す。

(a ISA kana mean a sound "a")
 (ka ISA kana mean ka sound "ka")
 (sa ISA kana mean sa sound "sa")

(n ISA kana mean n sound "n")

このチャンクは、先に示した単語に関わるチャンクの構成要素となる。つまり、本研究のモデルにおいて、しりとりにおける単語から文字を切り出す音韻意識は、word-heads あるいは word-tails タイプのチャンクに含まれる語頭、語尾音の知識をもとにして kana タイプのチャンクを検索するということに対応すると言える。

また、本モデルはしり通りのゲーム中で既に回答された単語に関する知識を保持するためのタイプ past を持つ。このタイプは単語の知識 past-word と単語の音知識 past-sound からなる。past タイプのチャンクは、はじめは宣言的モジュール内には存在せず、しり通りの進行に従って生成され、格納されてゆく。

3.1.2 ゴールモジュール

ゴールモジュールは、課題の状態を保持する。本モデルでは、ゴールモジュールに保持される短期記憶は回答単語を表すスロット (a-word)、回答単語の語頭音を表すスロット (a-head)、回答単語の語尾音を表すスロット (a-tail) とモデルの状態を表すスロット (state) によって構成される。これらのスロットの値はプロダクションモジュールによって逐次的に挿入、更新される。

3.1.3 イマジナルモジュール

イマジナルモジュールは短期記憶に関するモジュールであり、そこに保持された情報をチャンクとして生成する機能を持つ。本モデルの中では、相手の回答を受け取ったとき、または単語を回答するときに、回答済みの単語を表すチャンクを新たに生成し、宣言的モジュールへ格納する役割を果たす。

3.1.4 聴覚モジュール

聴覚モジュールは、耳で聞きとった音について、位置を把握し、内容を理解することに要する時間をシ

ミュレートする。本モデルでは、聴覚的な情報の入力があるまで待機し、入力があった際には、その情報をバッファに保持したのち、宣言的知識として格納する。

3.1.5 発話モジュール

発話モジュールは、口から言葉を発することや、頭の中で言葉を思い浮かべることによって要する時間をシミュレートする。本モデルでは、相手の回答した単語を思い浮かべること、自分の回答を発話することに利用される。

3.1.6 プロダクションモジュール

プロダクションモジュールは、他のモジュールが保持する情報や状態を利用しながら、ルールを選択、適用し、モジュールを操作する様々な処理を行う。本研究のモデルに含まれるエージェントは、相手の回答として単語の情報を受け取ると、しり通りのルールに即して単語を検索し回答する。

モデルには、相手の回答を聞き取り自分の回答を発話する一連のプロセスの他に、自分の回答をチェックするプロセスが存在する。回答候補の想起プロセスにおいて、まず聞き取った単語の情報が聴覚モジュールから受け渡され、ゴールモジュールの a-word スロットに配置される。その後、a-word に配置された知識をもとに宣言的モジュール内の単語と語尾を結ぶ知識 (word-tails タイプのチャンク) が呼び出される。また、a-word の知識はイマジナルモジュールを利用して既に回答した単語の知識 (past チャンクタイプ) として記憶される。その後、語尾音に注目して、ゴールモジュールの a-tail スロットに格納し、a-tail の音を語頭に持つ単語 (word-heads タイプのチャンク) を検索する。単語が検索されると、単語知識を次のゴール (a-word スロット) にセットする。

この後、モデルは想起された回答候補が語尾に「ん」を持たないことをチェックする。想起した回答候補の語尾が「ん」でなければ、その単語を回答とする。回答候補の単語の語尾が「ん」であるとき、回答候補単語の語頭文字との関連記憶を検索し、語頭文字に着目し、語頭文字の知識を利用することで、再度音韻の知識から回答候補となる単語の検索を行う。

「ん」のチェックの後、現在ゴールに保持されている単語が、自分もしくは他者によってすでに回答されていないか判定する。モデルの宣言的モジュールは、過去に想起された単語を経験として保持している (past

タイプのチャンク)。回答の前に past タイプのチャンクを検索し、失敗した場合（過去にその単語を想起した経験がない場合）、その単語を回答とする。回答候補を手掛かりとした宣言的モジュールの検索によって、過去に想起した経験が思い出された場合（回答候補の単語が既回答だった場合）、回答候補となる単語を再検索する。

現在ゴールに保持されている単語が、既に回答された単語ではなく、語尾の文字が「ん」でもないとき、その単語を回答済みの単語（past チャンクタイプ）として記憶した上で、発話モジュールを用いて単語を回答する。以上が本モデルにおける回答の流れである。しりとりは、2体のエージェントによってこのプロセスが連鎖することで遂行される。

3.1.7 モデルのパラメータ

ACT-R のモジュールの動作は、数値的なパラメータによって調整される。本研究では、エージェントの音韻意識を、単語知識をもとにして語尾音の知識を検索することで表現する。このことから、知識の検索されやすさを表す ACT-R パラメータ（活性値）の設定が重要となる。活性値は学習や忘却、文脈などに対応する複数の項の加算として定義され、算出された値がひとつひとつの知識（チャンク）に付与される。このうち、本研究において、学習と忘却の効果を表すベースレベルに注目する。

$$B_i = \ln \left(\sum_{j=1}^n t_j^{-d} \right) + \beta_i \quad (1)$$

上記の式 1 において、 B_i はあるチャンク i のベースレベルを表す。ベースレベルは、そのチャンクが参照された回数 n および ACT-R 内部でパラメータ (:bll) によって指定される減衰率 d と j 回目にチャンクが参照されてからの経過時間 t_j 、パラメータ (:blc) によって指定されるオフセット β_i から算出される。

ベースレベルを項として算出される活性値は、そのチャンクの想起に要する時間に影響する（活性値の高いチャンクほど素早く想起される）。さらに、活性値が一定（パラメータ:rt によって指定される閾値）よりも低いチャンクは、長時間の検索の試行が行われた後に失敗する。つまり、シミュレーションの進行にともなう活性値の変化を観察することで、ACT-R 内部での知識の使われやすさの変化を推測することができる。

活性値は繰り返しの効果によって強化されるため、通常の ACT-R において自由連想の課題をモデル化する

ると、検索が一部の知識に偏り、病的な検索になるということが指摘されている [3]。これに対し、知識の検索条件に制限をかける方法が存在する。プロダクションモジュールに含まれる記憶検索のルールにコマンド (Recently-Retrieved) を設定することにより、直近に検索されたことのある知識のみを指定した検索や、直近の記憶を除外した検索が可能となる。このコマンドを設定し、直近に検索された知識を除外することで、特定の事柄に対するこだわりを抑制することができるような個人を表現することが可能となる。

3.2 しりとりを失敗するモデル

未発達な音韻意識を想定したシミュレーションを行うために、3.1.1 節で示した音韻の知識について、誤った知識を作成した。誤った音韻の知識を表すチャンクの例を以下に示す。

```
(ka-error ISA kana mean ka sound "a")
(ki-error ISA kana mean ku sound "i")
(ku-error ISA kana mean ku sound "u")
```

この知識では、ka という知識が、誤ってその母音のみを表す “a” という知識と紐つけられている。競合する音韻の知識によって、エージェントはしり通りのルールに沿わない誤った回答（「りんご」に対して「おかし」と答えるなど）をするようになる。

しりとりを通して言語の習得に至るためには、このような誤った回答に対するフィードバックや補助が必要となる。モデルに含まれる 2 体のエージェントのうち、一方にのみ誤った音韻の知識を搭載することで、非対称なインタラクションとするとともに、他方のエージェントには、誤った回答を受け取ったとき、先の回答をもう一度提示するというプロセス（「りんご」に続く「おかし」という誤った回答に対しては、もう一度「りんご」を提示する）を組み込んだ。これは、同じ単語に対する回答を繰り返し考えさせるという補助に対応すると考えられる。

3.3 シミュレーション

モデルを用いて、シミュレーションを行い、結果から考察を行った。このモデルの各エージェントには天野、小林による『基本語データベース』[12]に掲載される単語の中から、20,544 単語を知識として搭載した。本研究では、処理の容易さとしり通りのルールから、名詞以外の品詞、同音異義語などの重複、一音のみで構成される単語（“木”や“輪”など）を除外して

モデルの知識としている。また、単語末尾の長音記号はすべて削除した。モデルの持つ音韻の知識には、単語の知識を音韻の単位で分割し、出現しうる音韻をすべて与えた。さらに、音韻の知識の検索において、Recently-Retrieved の設定により直近に検索された知識を抑制する設定（抑制 ON）をしたエージェントを含むモデルと、指定をしない（抑制 OFF）モデルの2パターンを用意した。

シミュレーションにおいて、モデルは 10,000 秒間（約 2.8 時間）のしりとりを 250 回実行した。それぞれ試行の中で、一方のエージェントの回答から他方のエージェントが別の単語を回答することを 1 チェーンと数え、しりとりが繋がった回数をチェーン数として記録している。

以下では、モデルに組み込んだしり通りの失敗と補助の効果を検討する。図 2 は、各試行における達成チェーン数をヒストグラムにまとめたものである。このヒストグラムから、2 モデルでともに達成チェーン数の小さな試行が多く、長くしりとりを続けられた試行は極端に少ないことが見て取れる。とくに達成チェーン数が少なくなる理由は、単語の知識に語頭音として含まれる音韻に偏りがあることによると考えられる。ある音を語頭音にもつ単語が極端に少ない場合、数回のチェーンによって次の回答候補がなくなるといった状況が考えられる。達成チェーン数が少ない階級において、抑制 ON モデルは抑制 OFF モデルよりも度数が小さい。これは直近に検索された知識を抑制する意識により、しりとりが継続する可能性を示唆する。

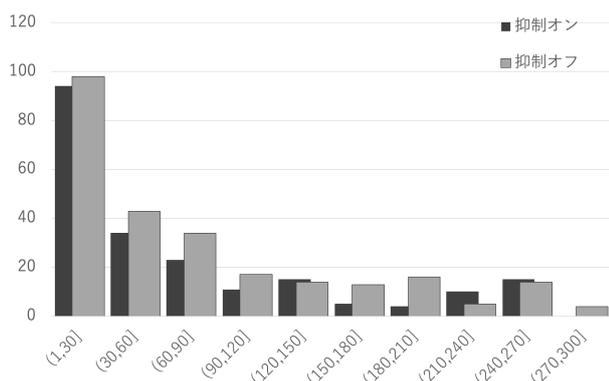


図 2 達成チェーン数の分布

図 3 は 1 チェーンに含まれる誤答の数の変化を表す。図 3 の横軸はチェーン数を示す。誤った知識を持つエージェントは誤った回答をするため、1 チェーンの中には複数の誤答が含まれる場合がある。図の縦軸はそのチェーンに含まれる誤答数の平均値を示す。

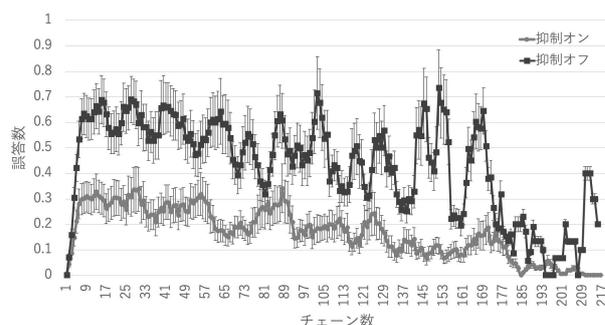


図 3 誤答数の推移，エラーバーは $\pm \frac{SD}{15}$

このグラフから、2 モデルでともに初期は誤答数が比較的多く、しり通りの進行に伴って誤答数が減少傾向にあることが読み取れる。これは、同じ単語への回答を繰り返し考えることで、正しい音韻の知識の活性化が相対的に高くなり、誤った知識が検索されることが少なくなったためであると考えられる。また、抑制 ON モデルでは、抑制 OFF モデルに比べて一貫して誤答数が少ない。これは直近の記憶を抑制する設定により、誤答したあとに同じ誤答を繰り返すことが少なくなるためであると考えられる。

4. まとめと今後

本研究では、インタラクションを介した音韻意識獲得のプロセスをモデル化することを目的とした。これを達成するために、ACT-R を用いて対話的に言葉遊びを行うモデルを構築し、シミュレーションを行った。このシミュレーションから、未発達な音韻意識を、誤った音韻の知識と正しい音韻の知識が検索の際に競合することに対応づけてモデル化できることが示唆された。また、同じ単語への回答を繰り返し考えさせる補助により正しい音韻が想起されやすくなることや、同じ間違いを繰り返さないという意識によって誤った音韻の知識による誤答が減少することが示唆された。さらに、直近の知識への抑制がないモデルでは、相対的に誤答が多く、しり通りのチェーン数も少ない傾向にあった。これは、間違えた経験など特定の事柄へのこだわりによって課題遂行に失敗している状態であると考えられ、自閉症による言語発達遅滞者などといった個人に対応づけられるモデルであることが示唆された。

本研究に対して、多くの課題が残されている。まず、シミュレーション結果の精査と、条件の異なるシミュレーションに対する結果の比較を含むさらなる分析が必要である。本研究のシミュレーションでは、約 2.8 時間のしりとりによって、誤答数に改善の傾向が見ら

れた。これは実際の音韻意識の発達過程にくらべて非常に短い期間である。モデルの振る舞いを分析し、このような変化の要因をさらに詳細に調べる必要がある。また、モデルの発展が必要である。達成チェーン数が小さくなるのは、特定の語頭音を持つ単語が少ないことによると考えられる。単語数の少ない語頭音として、たとえば、“シェ”や“ヴァ”などと言った音は、日本語の音韻に無く、外国語に由来するものである。日本語の音韻意識獲得に関するシミュレーションにおいて、このような音をどのように扱うか（“ヴァ”は“バ”に置換するなど）検討の上でモデルに与える知識を更新する必要がある。

文献

- [1] Anderson, J. R., (2007) “How can the human mind occur in the physical universe?”.
- [2] Baron-Cohen, S., (1997) “Mindblindness: An essay on autism and theory of mind”: MIT press.
- [3] Lebiere, C. and B. J. Best, (2009) “Balancing long-term reinforcement and short-term inhibition,” in “Proceedings of the 31st annual conference of the cognitive science society”, pp. 2378–2383, Cognitive Science Society Austin, TX.
- [4] Mätzig, P., S. Vasishth, F. Engelmann, D. Caplan, and F. Burchert, (2018) “A computational investigation of sources of variability in sentence comprehension difficulty in aphasia,” *Topics in cognitive science*, Vol. 10, No. 1, pp. 161–174.
- [5] Taatgen, N. A. and J. R. Anderson, (2002) “Why do children learn to say “broke”? A model of learning the past tense without feedback”, *Cognition*, Vol. 86, No. 2, pp.123–155.
- [6] Tomasello, M., (1999) “The Cultural Origins of Human Cognition.”
- [7] Van Rij, J., H. Van Rijn, and P. Hendriks, (2010) “Cognitive architectures and language acquisition: A case study in pronoun comprehension,” *Journal of Child Language*, Vol. 37, No. 3, pp. 731–766.
- [8] 高橋登, (1997) “幼児のことば遊びの発達: “しりとり”を可能にする条件の分析”, *発達心理学研究*, Vol. 8, No. 1, pp.42–52.
- [9] 山本米雄・柏原昭博, (1989) “知識定着を目的とした開放型 CAI のモデル化”, *電子情報通信学会論文誌 D*, Vol. 72, No. 9, pp.1459–1471.
- [10] 西川純平・森田純哉, (2018) “認知アーキテクチャを利用したことば遊びにおける音韻意識のモデル化”, 日本認知科学会第 35 回大会.
- [11] 大石敬子, (1994) “学習障害における言語の問題”, *聴能言語学研究*, Vol. 11, No. 2, pp.57–63.
- [12] 天野成昭・小林哲生, (2008) “基本語データベース: 語義別単語親密度”.
- [13] 日山美子, (2012) “自閉症児の読み書きの発達を促すための指導: 認知特性を踏まえた音節分解の指導”, *筑波大学特別支援教育研究*, Vol. 6, pp.31–36.