解 説

幼児は親の発話をどのように分節化するのか： 比較・照合アルゴリズムに基づく音素系列の分節化と再認のモデル

牧 岡 省 吾

大阪府立大学人間社会学部人間科学科*

How Do Children Learn to Segment Continuous Speech?: Models of Lexical Segmentation and Recognition

Shogo Makioka

Department of Human Sciences, Osaka Prefecture University

Two models are presented for generating a representation of words from the input phoneme sequences. They use an unsupervised learning algorithm that compares the input with its internal representation and generates a new representation of each subsequence. Simulation using child-oriented utterances in the CHILDES database as the training stimuli showed that the model performs lexical segmentation better than SRN and that it has fairly good generalization ability.

1. は じ め に

これまでに提案された言語に関するニューラルネットワークモデルの多くは、幼児の言語獲得過程あるいは成人の言語処理過程の、ある特定の側面のみを扱っている。たとえば動詞の過去形の獲得に関するモデルは、語幹と過去形のマッピングの学習のみを扱っている¹⁻³⁾。それらのモデルは、幼児が英語の過去形を学習する過程でみられる不規則動詞の過剰規則化（たとえば“go”→“goed”）の出現と消滅など、いくつかの心理現象を再現することに成功した。しかし実際の幼児のモデルとして考えた場合、語幹と過去形の対応のみを学習するという手続きは、不自然であると言わざるを得ない。幼児が母国語を学習する際、語幹と過去形が同時に対呈される機会は減多にないからである。幼児は、周囲で交わされる会話や自分に向けられた発話、さらに自分の身の回りで生じる出来事を分析し、発話のどの部分が動作を指し、どの部分が過去という意味を担っているのかを発見しなければならない。語幹と過去形の対応関係の学習を可能にするためには、様々な能動的処理がその背後で働いている必要がある。

以上のような問題は、単語の文法的カテゴリや意味を学習するモデルについても同様に当てはまる⁴⁻⁶⁾。多くのモデルは、あらかじめ分節化された単語とそれに

対応する意味特徴とのマッピング過程を扱っている。しかし、現実の幼児は、発話の中のどこからどこまでを1つの単語とみなすべきか、その単語と対応させるべき意味属性は何か、といったことを能動的に学習しなければならない。

本研究は、自然な状況下における幼児の語彙獲得過程を、できるだけ包括的にモデル化する試みの一環である。最終的には、音素系列の単語への分節化、単語の再認、語形変化、単語と意味の対応付けといった過程をすべて扱うことができるようなモデルの構築を目指している。本論文では、その最初の段階として、音素系列を単語へと分節化する過程を扱う。

子どもが母語を獲得する過程において、親が単語を1つずつ区切って発話する場合はまれである。子どもは、切れ目のない音素系列として発話される文を適切に分節化することを学習しなければならない。分節化の手がかりとしては、(1) 音素系列の統計的特性、(2) プロソディやアクセントなどの音声の手がかりの2つが考えられる。本研究では、(1) を用いて分節化を行う方法について検討する。

以下に、統計的アプローチによる研究においてしばしば使用される Korman のコーパス⁷⁾に収録された発話の最初の部分を抜粋する。これは、生後6-16週の乳児に対する親の発話を収録したものである。

Hello hello.

* 〒 599-8531 大阪府堺市学園町 1-1

Oh dear.
 Oh come on.
 Are you a sleepy head?
 Are you?
 A sleepy head?
 Oh dear.
 Oh dear.

このように、ことばの話せない乳児に対する話しかけでは、同一の単語や句、文が繰り返されることが多い。これは日本語においても同様である。いわゆる母親語 (motherese) がこのような統計的特性を持つことは、子どもが分節化の学習を行う際の有効な手がかりになっていると思われる。

先行研究について解説する前に、分節化の成功率を定義しておくことにする。本論文では、分節化の成功率の尺度として recall 率を用いる。以下に述べるモデルの多くでは、訓練刺激として、Korman のコーパスのような CHILDES⁹⁾ に登録されたコーパスを使用している。CHILDES のコーパスには、自然な状況下で採取された親と子の発話が一定の様式で記録されている。それぞれの発話は、通常の英語の文章と同様、単語間に空白を挿入した形で記録されており、この空白の位置を正しい分節化の位置とみなす。モデルが単語の前後を正しい位置で分節化できた回数 (コーパス全体における延べ回数) を、コーパス中の全単語の延べ出現回数で割ったものが recall 率である。

さて、統計的アプローチによる研究として代表的なものに、Brent らの記号表現を用いたモデルがある^{9, 10)}。彼らは、入力文の可能な分節化パターンに対して DR function と呼ばれる評価関数をあてはめることにより、教師なしで分節化の学習が可能であることを示した。Bernstein-Ratner のコーパス¹¹⁾に対する recall 率は、DR function のみを使用した場合に 47.3%、単語の始点と終点の発音可能性に関する手がかりを外部的から与えた場合には 69.2% だった。しかし彼らのモデルはコーパス全体を一度に入力することによって最適化を行っており、実際の幼児のように 1 文ずつ学習を行うわけではない。

一方、Yang は、音節 (syllable) の表現を獲得済みのモデルにおいて、音節の遷移確率に関する統計的学習と、音声的手がかりに関する生得的制約条件を組み合わせることによって分節化が可能であることを示した¹²⁾。彼らは、音節の遷移確率だけに基づく分節化は局所解に落ち込みやすいことを指摘し、「単語は必ず第一強勢 (primary stress) をもつ」という制約条件を用いれば局所解への落ち込みが避けられることを示し

た。しかし彼らのモデルには、冠詞や前置詞など、強勢が置かれられない機能語をうまく分節化することができないという問題点がある。

ニューラルネットワークを用いた研究としては、Christiansen らのモデルがある¹³⁾。彼らは、単純帰帰型ネットワーク (SRN)¹⁴⁾に、入力文の終端の予測を学習させた。Korman のコーパスに収録された 1 万文程度を学習させた結果、文の終端を予測するユニットが単語の終端に対しても反応するようになり、このユニットの活性度にもとづいてある程度分節化が可能であることが分かった (Recall 率は 40.4%)。しかし後述のように、彼らのモデルには、訓練刺激を増やしても分節化の成功率が改善されないという問題点がある。また、彼らのモデルは単語間の区切りを検出するだけで、単語の内部表現を獲得する訳ではない。

分節化の学習を語彙獲得過程と結びつけたモデルとしては、Davis のモデルがある¹⁵⁾。このモデルも SRN によって構成され、逐次的に入力される音素系列と、その文に含まれる単語の意味の局所表現との対応付けを学習する。文の最初の音素が入力されたときから文全体の意味表現を教師として与え、学習させることにより、各単語の音素系列が、それぞれ別々の意味表現へと対応づけられる。このモデルは意味との対応付けが分節化の手がかりとして役立つことを示したという点で興味深い。訓練刺激には機能語を含まない人工言語が用いられており、自然な発話に対して適用できるかどうかは未知数である。

2. モデル

2.1 モデルのアーキテクチャ

上述のように、これまでに提案された分節化のモデルの多くでは、音素または音節の遷移確率を分節化の手がかりとして用いている (遷移確率の低い場所で分節化が起こる)。これに対して、本研究では、入力された音素系列を比較・照合することによって音素系列から単語を直接抽出するモデルを構築している。このようなアプローチには、分節化と語彙獲得という 2 つの過程を、比較・照合による学習という単一の原理で説明することができるという利点がある。

提案するモデルは、言語処理に特化したアーキテクチャを備えている。これは、言語を扱う脳内の神経回路の構造が、遺伝的に強く制約されていることを仮定することに相当する。しかし言語に関する生得的知識を仮定する訳ではなく、モデルは Hebb¹⁶⁾型の教師なし学習規則によって音素系列の分節化を学習していく。訓練刺激としては、先行研究と同様、CHILDES に登

録されたコーパスを用いる。

モデルは音韻層と項目層の2層から構成される (Fig. 1)。音韻層ではスロット型の入力表現を用いており、それぞれのスロットに、各音素に対応するユニットが配置されている。各音素は、スロット内の1つのユニットが活動することによって表現される (局所表現)。最初に入力された音素は左端のスロットを活動させ、次に入力された音素は左から2番目のユニットを活動させる。このように、音素系列は入力層における空間的パターンとして表現される。

上記のような入力表現には、有限の種類の音素の組み合わせによって単語が構成されるという言語の特性を忠実に反映できるという利点があり、後に説明する比較・照合アルゴリズムを適用するためのワーキングメモリとしての役割を果たすのに適している。一方、SRNでは、隠れ層と文脈層がワーキングメモリとして働く。SRNではこれらの層におけるユニット群全体が時系列パターンを保持し、しかもユニットの入出力関数が非線形なので、パターンを下位の要素に分解することが難しい。これに対して、本モデルの入力表現ではスロット位置によって簡単にパターンを分解することができる。このような表現がヒトの脳内に存在するという証拠は今のところない。しかし視覚系においては、一次視覚野のコラム構造のような、網膜位置ごとに同じ神経回路網が繰り返されるという構造が普遍的にみられる。時系列パターンの処理において、系列位置ごとに同じ構造が繰り返すような回路網を想定することは不自然ではないであろう。また、本研究のアプローチとはやや異なるが、O'Reillyらは、音韻ループ (言語的ワーキングメモリ) のモデルにおいて、スロット型の内部表現を学習によって生成可能であることを確かめている¹⁷⁾。

Fig. 1のように、入力層には12個のスロットが配置されている。項目層には30,000個のユニットが存在し、入力層のユニットと双方向に結合している。結合

Fig. 1 Model architecture (simulation 1).

強度の初期値はランダムである。

2.2 学習アルゴリズム

牧岡^{18,19)}は、Grossberg²⁰⁾のART (Adaptive Resonance Theory) を、須賀・久野ほか^{21,22)}が提案した「差異・共通性抽出原理」を参考に拡張することにより、切れ目のない音素系列から単語の内部表現を自己組織的に生成するアルゴリズムを提案した。このアルゴリズムは、学習済み項目と新奇項目を比較・照合することによって内部表現を生成する。須賀・久野らが提案した言語獲得のモデル MLAS は LISP ベースの記号処理モデルであるが、その学習は入力パターン同士の比較に基づいており、ニューラルネットワークとの親和性は高い。

モデルの学習過程を以下に示す (Fig. 2)。ここでは、次の3つの文の学習過程について例示する。

/@UdI@r/ (Oh dear.)
 /@U/ (Oh.)
 /@UkVmOn/ (Oh come on.)
 (※音素記号は CELEX²³⁾ に準ずる)

1. 最初の文 (/@UdI@r/) が呈示されると項目層のユニットが活動し、競合が始まる。
2. 学習の初期においては、入力された音素系列に対応する内部表現が存在しない。その場合、入力パターン全体を表現する内部表現が生成される。
3. 次の文 (/@U/) が入力されると、共通部分をもつ /@UdI@r/ のユニットが最も強い入力を受けて活動する。
4. 活動したユニットは、下向きの結合を通して入力パターンとの照合を行う。
5. 照合が完全でない場合、項目層から音韻層へ向かう活動パターンと入力パターンとの共通部分 (/@U/) を表現するユニットがつくられる。
6. 次の文 (/@UkVmOn/) が入力されたとき、/@U/ を表現する項目ユニットからの照合は完全である。完全な照合が持続すると、その部分の入力が消失し、入力パターンが左方向に詰められる。完全な照合が持続した項目層のユニットの閾値は低下する。
7. 入力層に残された音素系列 (/kVmOn/) はこれまでに生成された項目ユニットと共通部分をもたないので、新たなユニットが割り当てられる。

以上のような比較・照合過程を繰り返すことにより、

Fig. 2 Learning process of proposed model. Active units are denoted by black circles.

モデルは音素系列を徐々に分節化していく。このアルゴリズムは、ある特定の音素系列が異なる文において異なる音素系列に隣接しているときに、両方の文で共通している音素系列を抽出し、項目層のユニットを割り当てる。そのため、ある単語が異なる文脈で使われるほど、その単語を分節化しやすくなる。さらに、項目層では局所表現が用いられているために、学習の干渉が生じにくい。これらの点から、訓練刺激が増えるとともに分節化の成功率が上昇することが期待される。

新たな内部表現を生成する際には、牧岡²⁴⁾が提案した忘却付き即時学習規則を用いる。この規則は Hebb 型の即時学習と素早い忘却を組み合わせたものである。下のように、ユニット間の結合強度は slow weight と fast weight の 2 つの成分からなる。

$$w = w^s + w^f \quad (1)$$

これら 2 種類の結合強度は、以下の式によって更新される (a は項目層のユニットの活性度、 m は入力層のユニットの活性度)。

$$\Delta w^s = (1 - \varepsilon)w^s + \varepsilon am \quad (2)$$

$$\Delta w^f = am - w^s \quad (3)$$

slow weight は 1 回の学習で学習係数 ε に応じた量だけ変化する。一方、fast weight は slow weight に対して相補的に働き、両者の総和は常に am に等しくなる。すなわち、

$$w = am \quad (4)$$

であり、1 回で完全な学習が成立する。

fast weight は急速に減衰し、slow weight は減衰しない。すなわち、

$$w^f(t+1) = \delta w^f(t) \quad (\delta < 1) \quad (5)$$

$$w^s(t+1) = w^s(t) \quad (6)$$

となる。

このように急速な学習と忘却を組み合わせることには、次のような利点がある。前述のように、自然な状況下における親の発話では、近接する文において同じ単語や句が繰り返されることが多い。入力文に対する学習が即時に行われることにより、近接する文同士の間接性を抽出し、新たな内部表現を生成することができる。そのように生成された内部表現が妥当なものであるとは限らないが、妥当でない内部表現に対応する音素系列が入力される頻度は小さいため、そのような表現は忘却によって消滅することが期待できる。

3. シミュレーション 1

3.1 訓練刺激

訓練刺激は、CHILDES データベースに登録された以下の 3 つの英語のコーパスから、親から子に向けた発話のみを抽出して作成した。

Korman⁷⁾ (子どもの年齢: 6-16 週, 12,419 文)

Bernstein-Ratner¹¹⁾ (子どもの年齢: 1 歳 9 カ月-1 歳 11 カ月, 8,700 文)

Theakston ら²⁵⁾ (子どもの年齢: 1 歳 9 カ月-3 歳 0 カ月, 20,000 文)

これらのコーパスには、分かち書きされた文として発話内容が記録されている。まず、コーパス中の文を辞書データベース CELEX²³⁾に収録された発音にもとづいて音素表記に変換した。固有名詞等はこの辞書に収録されていないことが多いため、別途リーダーズ英和辞典²⁶⁾で検索し、音素表記に変換した。次に、単語

Fig. 3 Correct segmentation rate (recall rate) for the three corpora.

間の空白を消去し、音素表記を入力ボタンに変換した。入力ボタンは 37 次元で、各音素が 1 個のユニットに対応する局所表現だった。

3.2 学習方法

コーパス中に収録された順序で各文を 1 回ずつネットワークに呈示し、学習を行わせた。slow weight の学習係数 ϵ と fast weight の減衰係数 δ の値は、ともに 7 だった。訓練刺激を増やすことによる効果を検討するために、同一のネットワークに 3 つのコーパスを順次学習させた。各コーパスの学習終了後に、学習を停止させてもう一度最初からコーパス全体を呈示し、正答率を算出した。

さらに学習の般化について検討するために、Korman⁷⁾のコーパス全体をテスト刺激として用い、このコーパスだけを学習した後にテストした場合と、その他のコーパスを継続して学習した後にテストした場合の recall 率を比較した。これにより、新奇な文を学習することによってモデルの能力が向上するかどうか分かる。比較対象として、Christiansen ら¹³⁾が使用したものと同様な構成の SRN に、同じ刺激を用いて訓練を行い、同一の基準で recall 率を算出した。

3.3 結果と考察

各コーパスに対する本モデルの recall 率を Fig. 3 に、Korman のコーパスに対する recall 率の変化を Fig. 4 に示す。本モデルの正答率は、全体に SRN と比べて良好であり、訓練刺激を増やすことによって Korman のコーパスの recall 率が向上している。これは、後続するコーパスの学習による般化が生じたことを意味する。これに対して SRN では、訓練刺激を増やすことによる成績の向上が頭打ちになっている。学習係数が小さい場合も recall 率の向上がみられないので、SRN ではコーパス間での般化が生じていないと判断できる。

Fig. 4 Generalization performance of the model.

SRN は訓練刺激における音素系列の統計的パターンを学習するが、各コーパスに固有の統計的性質を学習してしまい、般化能力に乏しいと思われる。一方、本モデルは、それまでに学習した音素系列と新たに入力された音素系列の共通部分を抽出することによって学習を進める。そのため、単語同士の新奇な組み合わせに遭遇することで、学習がむしろ促進されることが考えられる。

4. モデルの改良

4.1 シミュレーション 1 の問題点

シミュレーション 1 では SRN と比べて良好な結果が得られたが、いくつかの問題点が残されていた。第 1 の問題点は、文中のどこかで完全な照合が成立せずに分節化に失敗すると、残りの系列すべての分節化も失敗することである。人間の場合は、文頭の語が聞き取れなくても残りの部分の単語を認識できることが多い。

第 2 の問題点は、語頭の音素系列が共通する単語の区別ができないことである。たとえば“cap”と“captain”はともに独立した単語であるが、シミュレーション 1 のモデルは、いずれか一方を学習してしまうと、もう一方の認識が不可能になる。

第 3 の問題点は、単一の音素から構成される単語を分節化できないことである。シミュレーション 1 では、単一の音素に項目ユニットを割り当てることを許容していなかった。これは、文を音素に分解してしまうことを避けるためである。ほとんどの場合、個々の音素の出現頻度は、それらが構成する単語より高い。そのため、単一音素に分節化することを許容すると、モデルは文を単語ではなく音素に分解してしまうようになる。そこで、シミュレーション 1 では、音素を 2 つ以上含む系列だけに項目ユニットを割り当てるような

Fig. 5 Input representation of the modified model.

制約を与えていた。しかし、英語の場合、冠詞“a”は多くの場合/@/と発音されるため、このような制約を課すことによって冠詞の分節化が不可能になってしまう。このような事態を避けるため、シミュレーション1では冠詞“a”の発音を/eI/と登録していた（この点はChristiansenらのシミュレーション¹³⁾も同様である）。これは自然な状況下での発話と一致しないため、/@/のような単一の母音を単語として学習可能でなければならない。

4.2 アーキテクチャと入力表現

以上のような問題点を解決するために、入力表現を改良したモデルを構築した。シミュレーション1のモデルは、文全体を一度に入力され、それを分節化することによって単語の内部表現を獲得する。しかし現実の単語認知過程においては、音素系列は逐次的に入力される。成人は、単語の終端を待たず、その単語を一意に特定できるだけの音素が入力された時点でその単語を同定できることが知られている²⁷⁾。たとえば、/k&pt@/ (captain) の場合であれば、終端の/n/を待たず、/k&pt@/が入力された時点で“captain”であることが確定し、認知が成立する。

このような聴覚的単語認知の時間的特性を再現できるように、Fig.5に示すような入力表現を用いたモデルを構築した²⁸⁾。この入力表現では、系列の最初の音素は左端のスロットに入り、順次右側のスロットに音素が入力されていく。Fig.5の例では、/@Ud/が入力された時点で、/@Ud/とトップダウン・パターンが一致する項目ユニットが存在しないことにより、その直前に完全な照合が成立していた/@U/が認識されることになる。このような入力表現を用いることにより、単語の分節化と再認の過程を統一的に説明することが可能になる。

Fig. 6 Architecture of the modified model.

改良したモデルの全体像をFig.6に示す。モデルは文比較モジュールと認知／産出モジュールから構成される。音素系列は2つのモジュールの音韻層に同時に入力される。文比較モジュールは文同士の共通部分を抽出し、認知／産出モジュールは抽出した共通部分の内部表現を生成し、それを用いて単語認知を行う。Fig.5の入力表現は認知／産出モジュールの音韻層で用いられている。一方、文比較モジュールの音韻層では、入力された音素系列が右から左に流れていく。これにより、文頭以外の部分での音素系列の比較が可能になる。シミュレーション1と同様、両モジュールの音韻層はスロット型の構造をもつ。両モジュールとも実際には12個のスロットを持つが、単純化のため図には6個のみを示す。さらに両モジュールは項目層をもち、シミュレーション1と同様、音韻層と項目層の間は双方向性の結合で結ばれている。文比較モジュールの項目層は5,000個、認知／産出モジュールの項目層は10,000個のユニットから構成されている。

文比較モジュールは、文全体が入力された時点、あるいは音韻層のすべてのスロットに入力があった時点における音韻層の活動パターンを、シミュレーション1と同一の即時学習則によって記録する（Fig.6）。さらに、入力された文をシミュレーション1と同様なトップダウン照合によって記録済みの文と比較し、一致する部分があれば、どのスロットが一致しているのかに関する情報を認知／産出モジュールに送る。認知／産出モジュールは、一致したスロット内の音素系列の表現を生成する（Fig.7）。一致部分は単語に対応していることが多く、単語に対応するユニットが徐々に生成されていく。シミュレーション1と同様、認知／産出モジュールで認知が成立したスロットの内容はキャンセルされ、右側のスロットの内容が左側にコピーされる。これにより、文中の単語を次々に認識していくことが可能になる。

Fig. 7 Learning process of the modified model.

しかし、以上のような仕組みだけでは、単一の子音や、子音の発音不能な組み合わせからなる系列が抽出されてしまうことがある。そこで、認知／産出モジュールにおいて「母音を含んでいるかどうか」「語頭に発音不能な子音の組み合わせを含んでいないかどうか」を確認し、これらの条件に該当しない系列の学習をキャンセルするようなメカニズムを付加した。これは音素系列の発音可能性に関する情報を外部から与えることに相当するが、幼児は構音運動を学習する過程で発音可能性に関する情報がある程度抽出できるのではないと思われる。

5. シミュレーション 2

5.1 訓練刺激と学習方法

シミュレーション 1 と同一の 3 つの CHILDES コーパスを使用した。自然な状況下における発音により近づけるため、個々の単語の発音について見直しを行った。前述の通り、冠詞 “a” の発音としては /@/ を採用した。さらに、“are” や “dear” の末尾の /r/ のように、発音されないことが多い子音は削除した。これらの見直しにより単語の長さは全体的に短くなった。これは系列を特定するための手がかりが減少したことを意味し、分節化はより困難になると予想される。学習の手続きや学習係数の値、結果の評価方法は、シミュレーション 1 と同一だった。

5.2 結果と考察

Korman のコーパスの冒頭部分に関する分節化の結果を以下に示す。

[正しい分節化の例]

/h@l@U/ /h@l@U/ (Hello hello.)
 /@U/ /dl@/ (Oh dear.)
 /@U/ /kVm/ /On/ (Oh come on.)

Fig. 8 Correct segmentation rate (recall rate) for the three corpora.

Fig. 9 Generalization performance of the modified model.

/A:/ /ju:/ /@/ /sli:pI/ /hEd/ (Are you a sleepy head?)

/A:/ /ju:/ (Are you?)

[モデルの出力]

/h@l@U/ /h@l@U/

/@U/ /dl@/

/@U/ /kVmOn/

/A:j/ /u:/ /@/ /sli:pI/ /hEd/

/A:j/ /u:/

シミュレーション 1 では不可能だった、単一の音素から構成される単語 “a” の抽出に成功していることが分かる。

さらに、各コーパス全体に対する recall 率を Fig. 8

に、Korman のコーパスに対する般化成績を Fig. 9 に示す。シミュレーション 2 の recall 率は、シミュレーション 1 と比べて若干低下している。発音可能性に関する手がかりを外から与えたにもかかわらず正答率が低下した理由としては、以下の 2 つが考えられる。第 1 に、より自然な発音から構成されるように訓練刺激を修正したことにより単語の長さが全体に短くなり、分節化が困難になったと考えられる。第 2 に、音素系列を逐次的に入力しながら単語の認識と内部表現の生成を同時に行うという学習方法を採用したため、いったん短い音素系列の認識が成立してしまうと、それを語頭に含むような単語が学習されにくくなっていることがある。たとえば /dI/ という系列の内部表現を生成してしまうと、/dIn@/ (dinner) という系列は抽出されにくくなる。シミュレーション 1 では文全体を 1 度に呈示していたため、このような問題は起こりにくかった。この問題を解決するためには、長い音素系列を積極的に抽出するように学習方法を改良する必要があると思われる。

6. まとめと展望

シミュレーション 1 では、学習済み知識と新たな入力との比較・照合に基づく学習アルゴリズムによって音素系列の分節化が可能であることを確かめることができた。シミュレーション 2 では、音素系列を分節化することにより単語の内部表現を獲得し、かつその内部表現に基づく単語認知が、人間の場合と同様な時間的特性をもつようにモデルを改良することを試みた。シミュレーション 2 における分節化の成功率 (recall 率) は先行研究と比べて低い水準にとどまったが、その原因は逐次的な入力に基づく学習を行ったことにあると考えられる。分節化の成功率を向上させるためには、入力された文全体を使用する学習と組み合わせる必要があると思われる。これは、直前に聞いた文をどのように分節化すべきかを幼児が再吟味する過程に相当する。

本研究で提案した学習アルゴリズムは、過去形などの語形変化の学習や、意味との対応付けを含む語彙獲得過程にも適用可能であると思われる。今後は、それらの過程を包含するモデルを構築していく予定である。

謝 辞

本研究は科学研究費補助金・基盤研究 (B) 「話し言葉の理解における複合語処理の認知神経機構の解明」(課題番号 17320064) の助成を受けて行われた。

参 考 文 献

- 1) Rumelhart, D.E., McClelland, J.L. (1986): On learning the past tenses of English verbs. In D.E. Rumelhart, J. McClelland (Eds.), *Parallel distributed processing Vol.2*. (pp.216–271) Cambridge, MIT Press
- 2) Plunkett, K., Marchman, V.A. (1991): U-shaped learning and frequency effects in a multi-layered perceptron: Implications for child language acquisition, *Cognition*, Vol.38, pp.43–102
- 3) Plunkett, K., Marchman, V.A. (1993): From rote learning to system building: Acquiring verb morphology in children and connectionist nets, *Cognition*, Vol.48, pp.21–69
- 4) Seidenberg, M.S., Gonnerman, L.M. (2000): Explaining derivational morphology as the convergence of codes, *Trends in Cognitive Sciences*, Vol.4, pp.353–361
- 5) Hinton, G.E., Shallice, T. (1991): Lesioning an attractor network: Investigation of acquired dyslexia, *Psychological Review*, Vol.98, pp.74–95
- 6) Li, P., Farkas, I., MacWhinney, B. (2004): Early lexical development in a self-organizing neural network, *Neural Networks*, Vol.17, pp.1345–1362
- 7) Korman, M. (1984): Adaptive aspects of maternal vocalizations in differing contexts at ten weeks, *First Language*, Vol.5, pp.44–45
- 8) MacWhinney, B. (2000): *The CHILDES project: Tools for analyzing talk*. Third Edition, Lawrence Erlbaum Associates, Mahwah, NJ
- 9) Brent, M.R., Cartwright, T.A. (1996): Distributional regularity and phonotactic constraints are useful for segmentation, *Cognition*, Vol.61, pp.93–125
- 10) Brent, M. (1999): Speech segmentation and word discovery: a computational perspective, *Trends in Cognitive Science*, Vol.3, pp.294–301
- 11) Bernstein-Ratner, N. (1984): Patterns of vowel modification in motherese, *Journal of Child Language*, Vol.11, pp.557–578
- 12) Yang, C.D. (2004): Universal grammar, statistics or both?, *Trends in Cognitive Science*, Vol.8, pp.451–456
- 13) Christiansen, M.H., Allen, J., Seidenberg, M.S. (1998): Learning to segment speech using multiple cues: a connectionist model, *Language and Cognitive Processes*, Vol.13, pp.221–268

- 14) Elman, J.L. (1990): Finding structure in time, *Cognitive Science*, Vol.14, pp.179–211
- 15) Davis, M.H. (2002): Connectionist modeling of lexical segmentation and vocabulary acquisition, In P. Quinlan (Ed.), *Connectionist approaches to developmental psychology*, Psychology Press, Hove, UK
- 16) Hebb, D.O. (1949): *The Organization of Behavior*, New York, Wiley
- 17) O'Reilly, R.C., Soto, R. (2002): A Model of the Phonological Loop: Generalization and Binding, In T.G. Dietterich, S. Becker, Z. Ghahramani, (Eds.), *Advances in Neural Information Processing Systems (NIPS) 14*. MIT Press, Cambridge, MA
- 18) 牧岡省吾 (2003): 語形変化のコネクショニストモデル～単語の分節化過程との関連を探索～, 日本認知科学会第 20 回大会発表論文集, pp.208–209
- 19) 牧岡省吾 (2004): 単語の分節化のコネクショニストモデル, 日本認知科学会第 21 回大会発表論文集, pp.20–21
- 20) Grossberg, S. (1978): Adaptive pattern classification and universal recoding: Part I. Parallel development and coding of neural feature detectors, *Biological Cybernetics*, Vol.23, pp.121–134
- 21) 須賀哲夫, 久野雅樹 (2000): ヴァーチャルインファント—言語獲得の謎を解く. 北大路書房, 京都
- 22) Suga, T. (2004): Solving the riddle of language acquisition: How to design a virtual Infant, *Studies in Language Sciences*, Vol.3, pp.231–246
- 23) CELEX Lexical Database, English version 2.5 (1995): Centre for Lexical Information, Max Planck Institute for Psycholinguistics, Nijmegen
- 24) 牧岡省吾 (2003): 音韻ループのコネクショニストモデル, 認知科学, Vol.10, pp.93–103
- 25) Theakston, A.L., Lieven, E.V.M., Pine, J.M., Rowland, C.F. (2001): The role of performance limitations in the acquisition of verb-argument structure: an alternative account, *Journal of Child Language*, Vol.28, pp.127–152
- 26) リーダーズ英和辞典 第 2 版, 松田徳一郎, 東 信行, 豊田昌倫, 原 英一, 高橋作太郎, 木村建夫, 山県宏光, 馬場 彰 (編) (1999): 研究社, 東京
- 27) Marslen-Wilson, W.D., Welsh, A. (1978): Processing interactions and lexical access during word recognition in continuous speech, *Cognitive Psychology*, Vol.10, pp.29–63
- 28) 牧岡省吾 (2005): 単語の分節化過程のコネクショニストモデル (2): ～単語認知の時間的特性に関する検討～, 日本認知科学会第 22 回大会発表論文集, pp.282–283