研究論文

変分オートエンコーダーを用いた乳幼児期の語彙発達過程の探索

萩原 広道 [1,2,*]　水谷 天智 [3]　山本 寛樹 [2,4,5]　阪上 雅昭 [3]

[1] 東京大学　[2] 日本学術振興会　[3] 京都大学　[4] 北陸先端科学技術大学院大学　[5] 大阪大学

Exploring young children’s vocabulary development trajectories using variational autoencoders

Hiromichi Hagihara [1,2,*]　Tenchi Mizutani [3]　Hiroki Yamamoto [2,4,5]　Masa-aki Sakagami [3]

[1] *The University of Tokyo*　[2] *Japan Society for the Promotion of Science*　[3] *Kyoto University*
[4] *Japan Advanced Institute of Science and Technology*　[5] *Osaka University*

This study aimed to explore young children’s vocabulary development using a machine learning technique, variational autoencoder (VAE). The VAE is an unsupervised neural network that maps high-dimensional input data onto a dimension-reduced latent space and then regenerates the data. The complex input features could be visualized in a low-dimensional latent space while maintaining its interpretability. We used parent-reported questionnaire data extracted from a publicly available database, involving American young children (N = 5,520) and applied VAE. The two-dimensional latent space in the adopted model demonstrated that vocabulary development had a quasi-one-dimensional structure shaped by an arc. Its rotation and radial directions represented changes in total vocabulary size and individual differences, respectively. We found that some categories in the questionnaire (e.g., Sounds, Animals) were more likely to develop earlier in the outer path of the arc, whereas others (e.g., Action words, Pronouns) tended to develop predominantly in the inner path of the arc. Furthermore, a simulation case study using longitudinal data suggested that some specific lexical items were crucial in characterizing the universality and diversity of different developmental trajectories in the latent space. Our approach will contribute to quantitatively depicting the development of children’s vocabulary in a more fine-grained and nuanced manner, providing a synergetic bridge between machine learning and developmental science.

Keywords: language development（言語発達）, productive vocabulary（表出語彙）, individual differences（個人差）, case study（事例分析）, machine learning（機械学習）, explainable artificial intelligence（説明可能な人工知能）

Received 1 March 2023; Accepted 2023; Available online at J-STAGE 2023

1. はじめに

言語発達は，言語学や心理学，認知科学，情報学といった多様な分野において，重要な探究の対象であり続けてきた．言語は，音韻・語彙・文法・語用論等の諸側面から構成される (萩原, 2021)．そのなかでも，乳幼児期の語彙発達は，後の言語スキル (Borovsky, 2022; Marchman & Fernald, 2008) や学力 (Bleses et al., 2016; Morgan et al., 2015) の予測因子となったり，非定型発達児において特異的な発達パターンが観察されたりする (Jiménez et al., 2021; Jiménez & Hills, 2022a, 2022b) ことから，理論的観点のみならず，教育・臨床といった応用的観点からも広く研究されてきた．子どもは，“banana” や “bottle” といった生活中で身近に触れる単語であれば，生後 6〜9 ヶ月頃から理解の兆しを見せ始めるが (Bergelson & Swingley, 2012), 実際に単語を表出するようになるのは生後 12 ヶ月頃である (Moore et al., 2019)． 1 歳前半の語彙発達は緩

萩原広道の現所属は大阪大学，東京大学．山本寛樹の現所属は Indiana University，大阪大学．
ORCID iD : 萩原広道（https://orcid.org/0000-0003-3316-600X），山本寛樹（https://orcid.org/0000-0003-1051-6599）．
本研究の一部は日本物理学会 2023 年春季大会にて報告した (阪上他，2023)．本論文ではさらに，VAE における潜在空間について，個々の意味カテゴリーや語彙項目ごとに詳細に検討し考察を深めた．
* 責任著者．E-mail: hiromichi.h@gmail.com

やかに進むのに対し，1 歳半頃になると，表出語彙の急激な増加が見られる (小林他, 2013)．英語やフランス語等の多くの言語圏において，初期の語彙には名詞が多く含まれており，それに対して動詞等他の品詞の発達は遅れることが知られている (Frank et al., 2021; Gentner & Boroditsky, 2001)．日本語も例外ではなく，初期の語彙は名詞優位となることが報告されている (小椋, 2007; 小椋他, 2016)．なお，中国語や韓国語等の一部の言語では，初期の語彙は名詞よりも動詞優位になることが指摘されている (Choi & Gopnik, 1995; Tardif, 1996)．

語彙の発達を調べる手法には，(a) 親子会話等の自然場面の観察法，(b) 選好注視法や馴化－脱馴化法を用いた実験心理学的手法，(c) 語彙チェックリスト法等，様々な種類がある (Ambridge & Rowland, 2013)．そのうち，予め用意した語彙チェックリストを用いて養育者からの報告をもとにデータを収集する手法は，比較的簡便に実施でき，かつ，既に多言語で翻訳・標準化された質問紙が存在する (Fenson et al., 1993) ことから，過去数十年にわたり多くの研究で採用されてきた．国際的に最も広く使用されている語彙チェックリストに，マッカーサー・ベイツ乳幼児言語発達質問紙 (MacArthur–Bates Communicative Development Inventories: MCDI; Fenson et al., 1993) がある．MCDI を用いて収集された語彙データの一部は，Wordbank と呼ばれるオープンデータベース (Frank et al., 2016) にて公開されており，公開データに対する二次利用研究も実施されている (例えば，Frank et al., 2021; Jiménez & Hills, 2022a)．2023 年 7 月現在，Wordbank には 38 の言語からなる 94,451 件分の MCDI データが集積され公開されており，個々の研究者が協力を申し出て Wordbank にデータを提供している．

MCDI 等の語彙チェックリストデータに対して解析を行う場合，およそ 600～700 の語彙項目について，表出語彙であれば子どもが「言う」か「言わない」かを記録したものがデータとなる．語彙チェックリストは多次元データとなるため，解釈可能な解析を実施するためには一般に何らかの集約が必要になる．例えば，各月齢における平均的な表出語彙数や，表出語彙数の個人差を産出する分析 (Frank et al., 2021) は，定型的な言語発達の幅を知る上で有用である．これに特定の閾値を設けて，語彙発達に遅れのみられるレイト・トーカー (late talker) を特定し，その語彙発達の特徴を検討した研究も存在する (Beckage et al., 2011; Jiménez & Hills, 2022a; 奥村・小林, 2019)．また，いくつかの意味または文法カテゴリーごとに語彙数を集計し，各カテゴリーにおける語彙数の発達的変化を記述・比較する研究も多くなされてきた (例えば, Frank et al., 2021)．

このように，語彙発達データを用いた従来の研究では，月齢群や意味カテゴリー等，ある特定のア・プリオリな基準によってデータを要約・集計し，それらを記述・比較することが中心だった．これらの既存手法に対し，本研究では，語彙発達データをより詳細かつデータ駆動的に探索する新たな方策として，Variational Autoencoder (VAE; Kingma & Welling, 2013; Rezende et al., 2014) という機械学習手法を用いた解析アプローチを提案する．VAE とは，高次元の入力データを低次元の潜在空間上に縮減した後，その潜在空間の情報からデータを再度復元するようにニューラルネットワークを訓練する教師なし学習の一手法である．VAE を MCDI の語彙データに適用することには，以下の利点があると考えられる．第一に，何百もの語彙項目からなる MCDI の多次元データがもつ特徴量を，低次元空間に縮約して表現・可視化することが可能になる．第二に，潜在空間から高次元データを復元するという VAE の特質を活用することにより，手持ちのデータには存在しない新しい MCDI データを仮想的に生成できる．第三に，潜在空間へのデータの縮減と，潜在空間からのデータの生成というこれら 2 つの利点を組み合わせることで，ある特定の語彙項目の表出確率が，潜在空間上でどのように分布しているかを詳細かつ容易に把握し，潜在空間の解釈を進めることができる．第四に，その分布をもとに各語彙項目を少数のパラメータで特徴づけることによって，「名詞」「動詞」といったカテゴリーをア・プリオリに仮定することなく，データ駆動的に各語を分類し，語彙発達過程を議論する新たな視点および定量指標を提供することができる．このように，VAE は，MCDI における語彙発達の様相をデータ駆動的に，かつ解釈可能性を担保しながら探索できる解析アプローチとして優れた手法である．

本研究では，MCDI データを VAE に適用する端緒であるという位置づけから，以下の 2 点の目的を掲げる．第一に，MCDI データを用いて学習された VAE について，その潜在空間の解釈を試みる．具体的には，表出語彙の発達が潜在空間上でどのようなパターンとして表現・可視化されるかを調べるとともに，

MCDI を構成する各意味カテゴリーの表出確率が, 潜在空間上でどのように分布するかを調べる．上述の通り，VAE では個々の語彙項目ごとに表出確率の分布を調べることが可能ではあるが，本研究ではまず潜在空間の大まかな解釈を行う目的から，カテゴリーごとの表出確率の分布を分析対象とすることにした. 第二に, 実在する MCDI の縦断データを用いて，その発達の経路が潜在空間上でどのように表現されるかを調べるとともに，もし実際とは異なる経路を辿った場合にどのような語彙発達の過程が想定されうるかを探索する．これにより, 実際に想定される語彙発達の経路に照らし合わせながら，潜在空間のより詳細な解釈を目指す.

2. 方法

2.1 対象とするデータセット

本研究では，乳幼児期の語彙発達を調査する際に国際的に広く用いられている MCDI (Fenson et al., 1993) の回答データを用いた．MCDI は，日本語を含む多くの言語に翻訳されており (例えば, 小椋・綿巻, 2004), 予め用意された語彙チェックリストに対して，子どもが言えるようになったものを養育者が回答することによってデータを得る．MCDI には，8〜18 ヶ月用の「語と身振り」版と，16〜30 ヶ月用の「語と文法」版の 2 種類がある［脚注 1］．前者は，最初期の理解語彙・表出語彙を対象としているのに対して，後者は 3 歳頃までの表出語彙をより広く扱う内容となっている. 質問紙は Animals (動物の名前), Food and drink (食べ物・飲み物)，Action words (動作語) 等の意味カテゴリーごとに並んでおり，養育者はそれぞれの語彙項目について，「語と身振り」版であれば「わからない」「わかるが言わない」「わかる＋言う」から当てはまるものを選び,「語と文法」版であれば「言う」か「言わないか」を回答する.

本研究では，潤沢かつ一部に縦断データを含むアメリカ英語の「語と文法」版を対象に，Wordbank (Frank et al., 2016) にて公開されているデータセットから MCDI データを抽出した (抽出日：2022 年 2 月 23 日)．同様の手法は Jiménez and Hills (2022a) にも見られる．このデータセットには，16〜30 ヶ月の乳幼児 5,520 名が含まれており，各児について，22 の意味カテゴリーに分類される 680 の語彙項目についての表出語彙データが含まれている. このうち, 1,349 名のデータは，異なる時期に同じ子どもから複数回にわたりデータを収集した縦断データとなっている. データ収集の回数およびデータ収集時の月齢は子どもによって異なるが, 計 3,223 回分のデータが縦断データと見なされる.

2.2 Variational Autoencoder (VAE)

VAE は，高次元の入力データを低次元の潜在空間に縮減した後，再度データを復元するようにネットワークを訓練する教師なし学習である (Kingma & Welling, 2013; Rezende et al., 2014)．本研究で用いた VAE の構成を**図 1** に示す.

［図 1 を挿入］［脚注 2］

高次元データを縮減する簡便でよく知られた類似手法として, 主成分分析が挙げられる. 実際に, MCDI を使った既存研究でも主成分分析を用いた報告が存在する (Frank et al., 2021)．しかし，主成分分析は変数を線形変換し低次元空間に射影することによってデータを縮減するため，表現力や低次元空間の解釈可能性といった点で限界もある. 例えば, 非線形な変換ができないことや，低次元空間において実データがない位置が元の高次元空間のどこに対応するかが分からないといった問題点が挙げられる.

これらの限界を克服する手法として，Autoencoder (AE; Hinton & Salakhutdinov, 2006) が挙げられる．AE は，高次元の入力ベクトル$\boldsymbol{x}$を低次元の潜在変数ベクトル$\boldsymbol{z}$に変換する encoder と，$\boldsymbol{z}$を高次元ベクトル$\hat{\boldsymbol{x}}$に復元する decoder の 2 つのニューラルネットワークから構成される．複雑な関数を表現するために，encoder と decoder のそれぞれに中間層を置き，活性化関数により非線形な変換を行うことで，モデルの表現力を高めることができる．このため，AE は非線形化された主成分分析として位置づけられる．さらに，decoder を用いることで，潜在空間上の任意の座標から未観測データを仮想的に生成できることも AE の利点である.

VAE ではさらに, 入力ベクトル$\boldsymbol{x}$から潜在空間ベクトル$\boldsymbol{z}$を直接学習するのではなく，$\boldsymbol{z}$はある分布からのサンプリングによって得られるものとし，分布のパラメータを入力データから学習するようにモデル化する点に特徴がある．性質がよく分かっている正規分布$N(\boldsymbol{\mu}, \boldsymbol{\Sigma})$を採用しモデルの自由度に制限をかけることで, 潜在空間の解釈可能性を高めている. 例えば，同じ入力$\boldsymbol{x}$に対して，試行の度に異なる潜在変数$\boldsymbol{z}$がサンプリングされるが，それらの異なる$\boldsymbol{z}$に対し

ても出力は元の$\boldsymbol{x}$と同じになるように学習が進む．その結果，潜在空間において距離の近い点同士は入力・出力空間においても距離が近くなるため，潜在空間の解釈が容易になる．また，AE 同様に decoder から潜在空間上の任意座標について仮想的なデータを生成できる．VAE を構成する encoder と decoder を行き来することで，実データの有無に関わらず潜在空間と元の高次元データとの対応関係を把握できることから，VAE の潜在空間には高い解釈可能性が期待される．なお，正規分布からのサンプリングという拘束によってモデルの表現力が低下することが危惧されるが，encoder と decoder を構成するニューラルネットワーク自体の表現力が高いので，VAE は依然として高い表現力が維持される．

このように，VAE では正規分布からのサンプリングによって入力データと潜在空間との対応関係を構築する．潜在空間の事後分布は，多変量標準正規分布$N(\mathbf{0}, \mathbf{I})$を事前分布として，変分ベイズ法によって推定される．VAE の学習手続きについては，**付録**の概説を参照されたい．

2.3 モデルの学習および分析方針

VAE の学習に用いた入力データは，680 語×5,520 名であった．したがって，モデルの入力および出力次元は 680 次元となる．入力データの成分は，「言う」(＝1)，「言わない」(＝0) の 2 値である．モデルの中間層はそれぞれ 512 次元とし［脚注 3］，活性化関数には ReLU を用いた．本研究では，潜在変数の可視化を容易にしモデルの解釈可能性を高めるために，潜在空間の次元は 2 次元とした．潜在空間の次元数は自由に設定することが可能ではあるが，解釈可能性を優先して次元数を 2 次元に留める方法は，主成分分析や AE を用いた他の発達研究においても用いられている (例えば，Westermann & Mareschal, 2014)．出力層の活性化関数には sigmoid を用いた．潜在空間へのサンプリングを除いて，隣接する層同士は全結合するようにモデルを実装した．

VAE の学習は以下の手続きによって進めた．まず，誤差関数の最適化手法には Adam を用いた．初期値をランダムに変えながら 1,000 通りの学習を行い，データを 7 分割したクロスバリデーションによる検証誤差が最も小さい結果を最終的なモデルとして採用した．その際，検証誤差が相対的に小さい複数の結果において潜在空間の特徴を目視にて確認したところ，いずれも類似の構造を有していた (潜在空間の具体的な特徴については 3.1 節を参照)．したがって，採用されたモデルのみが特異な構造をもつわけではないことは確かめられている．

採用されたモデルの encoder と decoder を駆使して，以下の 2 つの観点から潜在空間の解釈を試みた．第 1 の観点は，潜在空間上で，表出語彙数の発達がどのようなパターンとして表現・可視化されるかを調べることである．潜在空間には，個人ごとの実データ点を投射できる．これらのデータ点に加えて，MCDI 上で語彙数が最小 (＝全成分が 0) および最大 (＝全成分が 1) となる仮想データを生成し，同様に潜在空間上に投射することで，潜在空間における発達の方向性を特定した．さらに，decoder を用いて，潜在空間上の任意の点から仮想的な MCDI データを生成し，その結果をもとに総表出語彙数の期待値および各語の表出確率を算出した．なお，decoder によって出力される 680 次元のデータの各成分は，本来の MCDI 入力データ (2 値の離散値) とは異なり，0 から 1 の範囲における連続値で与えられる．この値は，各語がどのくらいの確率で「表出する」と回答されるかの表出確率を示している．そして，各成分の表出確率の総和をとったものは，潜在空間上のその座標における総表出語彙数の期待値と見なされる．以降，この期待値を decoder から得られる総表出語彙数と表記する．これら decoder から生成されたデータをもとに，潜在空間上で総表出語彙数や各語の表出確率の等高線を描出することで，潜在空間における発達の向きを可視化し，潜在空間を解釈した．

第 2 の観点は，MCDI の縦断データを用いて，ある特定の個人が潜在空間上でどのような発達経路を辿りながら表出語彙を増やしていくかを探索することである．事例分析として，異なる月齢で複数のデータ点をもつ参加児を 1 名選び，表出語彙の変化を潜在空間上に図示した．さらに，もし実際とは異なる経路を辿った場合に，どのような表出語彙発達の過程が生じるかを，個々の語彙項目レベルで詳細に調べた．

なお，本研究における潜在空間の図示・解釈には，$\boldsymbol{z}$の事後分布$N(\boldsymbol{\mu}, \boldsymbol{\Sigma})$のうち，代表値となる$\boldsymbol{\mu}$のみを用いることとした．

3. 結果

3.1 潜在空間の特徴とその解釈

本研究で採用したモデルの学習結果について，2 次元潜在空間を可視化したものを**図 2** に示す．図の各点は，MCDI の実際のデータ点 (N＝5,520) を encoder

によって潜在空間上にプロットしたものである.

［図 2 を挿入］

まず，表出語彙数の発達的始点および終点を特定することで, MCDI における発達経路が全体として潜在空間上のどこからどこへ進むのかを調べた．MCDI の語彙項目 680 語を全て表出しない場合 (全成分が 0) と，全て表出する場合 (全成分が 1) の仮想的なデータを生成し，これらを encoder によって潜在空間上に投射した．本研究では，これらの点をそれぞれ birth point，end point と呼ぶ．その結果，図 **2-a** の通り，表出語彙数の発達は birth point から end point へと向かって，厚みのある弧のような擬一次元構造として表現されることが分かった．実際に，MCDI の縦断データを月齢の向きに沿って潜在空間上にプロットすると，birth point から end point に向かって時計回りに語彙の発達が進む様子が明確に表現された (図 **2-b**).

次に, MCDI における総表出語彙数と潜在空間上の座標との対応関係を調べることで，擬一次元構造のさらに詳細な解釈を試みた．具体的には，潜在空間上のプロット範囲において 30×30 の格子点を作り，その各点から decoder を用いて仮想的な MCDI データを生成した．これは，潜在空間上に一様に分布する 900 人分の MCDI データを新たに生成することに対応する．生成されたデータから総表出語彙数を算出し，そのデータをもとに潜在空間上に総表出語彙数の等高線を描いた (図 **2-c**)．この等高線からも，語彙数が birth point から end point に向かって，定型的には時計回りに発達が進む (表出語彙数が大きくなる) ことが見て取れる．すなわち，擬一次元構造を弧として見た場合，回転方向は総語彙数の変化を示している．一方，動径方向に注目すると，弧の厚みに対して中央を通る経路は最も定型的な発達経路だと見なされる．実際に，MCDI データを総表出語彙数が小さい順に 300 名ごとの bin に区切り，各 bin に含まれるデータに対して潜在空間における平均座標を算出し描写したところ (図 **2-c**，緑の点)，それらの平均点は動径方向に対して中央を通るように布置された．本研究では，この経路を潜在空間における平均発達曲線と呼ぶこととする．平均発達曲線は，総語彙数の等高線に対しておよそ垂線となっていた．したがって，動径方向におけるデータの分布は，総語彙数が同程度の場合の個人差・発達の多様性を表現していると想定される．

総表出語彙数の等高線で見たときに同程度となる点同士には，どのような違いがあるだろうか．これについて検討するために，MCDI を構成する 22 の意味カテゴリーについて，潜在空間における表出確率の分布を調べた．上述の手続きと同様，decoder を用いて生成された仮想的な MCDI データから，各語の表出確率をカテゴリーごとに平均し，潜在空間上に等高線を描いた (図 **3**)．同時に，総表出語彙数が全体の約 50% (＝330～350 語) となるデータ点を図示した．各意味カテゴリーの表出確率が 50%となる等高線と，総表出語彙数が 50%となる帯とを見比べることで，総語彙数が同程度である場合に，どのような語彙発達の多様性が見られるかを探索した．具体的には，各意味カテゴリーの等高線を参照しながら，擬一次元構造の外側と内側とを比較した．

［図 3 を挿入］

例えば, Sounds カテゴリーは“moo”や“woof woof”等の擬音語に対応するカテゴリーであり，早期に習得されやすい語彙群だとされる (Laing, 2019a, 2019b). 緑色の帯の外側に注目すると，Sounds の等高線は既に表出確率 75%を越えているが，内側に注目すると，この等高線はまだ 50～75%の間に留まる．すなわち，総表出語彙数が同程度の場合，Sounds カテゴリーは擬一次元構造の外側において発達が早く，内側において発達が遅いという特徴をもつ. Sounds 以外にも，いわゆる具体名詞に分類される意味カテゴリー (Animals から People まで) の多くは，同様に擬一次元構造の外側の方が，内側よりも表出確率が高くなる傾向を有していた．

これに対して，“me” や “you” 等の Pronouns (代名詞) カテゴリーは, 内側と外側の関係性が逆転することが分かった．すなわち，総表出語彙数が 50%となる緑色の帯の外側に注目すると，Pronouns の等高線は表出確率 25%を下回っている一方で，内側の方では 25%の等高線を上回っていた．したがって，総表出語彙数が同程度の場合，Pronouns カテゴリーは擬一次元構造の内側において発達的に優位であるといえる. 同様の傾向は, 文法的には述部 (動詞や形容詞) に分類される Action words や Descriptive words，閉じた語に分類されるカテゴリー群 (Time words，

Quantifiers 等) にも見られた．

別の観点として，25%の等高線から 75%の等高線に至るまでの区間の広さを意味カテゴリーごとに比較することもできる．例えば，**図 2** で見た平均発達曲線に沿って，**図 3** の等高線の区間の広さを見てみると, Action words では 25%から 75%に至るまでの区間は相対的に狭くなる．これは, 総表出語彙数でいうところの約 250〜460 語までに対応しており，この小さい幅の総語彙数の変化の間に，Action words 全体の表出語彙が一気に増加することを示している．一方で，People では等高線が 25%から 75%に至るまでの間に, 平均発達曲線に沿った総表出語彙数は約 150〜550 語に増える．つまり，Action words と比較して，People のカテゴリー全体の語彙発達は緩徐に進むことが示された．

以上のように，VAE の encoder と decoder を用いて表出語彙の発達の向き, 擬一次元構造における内側・外側の比較を行うことにより，潜在空間が解釈可能となることが明らかとなった．

3.2 縦断データを使った事例分析

潜在空間についてさらに詳細な解釈を進めるため, MCDI の縦断データを用いた事例分析を実施した. 単一事例の分析であるため，得られた知見の一般化可能性には限界を伴うが，個々の子どもの語彙発達の過程を定量的かつ詳細に捉え可視化するアプローチとしての VAE の活用法を提示するという点で，本節の分析には方法論的な意義も認められると考える．

本事例の MCDI データは，16 ヶ月〜29 ヶ月までの間に計 9 回収集された．実データ点を潜在空間上で月齢順に繋いだものを**図 4** の実線で示す．この発達軌跡を見ると，始点から O まで，および P は，平均発達曲線に近い位置を通っている．それに対して，O→C→P の区間は，動径方向において一度大きく右側に逸脱しており，平均発達曲線に対して内側を回っている. この内回りの迂回は, 平均的な発達経路と比べてどのような語彙発達の特徴をもつだろうか．また, もし当該事例が内回りではなく, 反対に大きく外側を迂回するような経路を辿っていたら, MCDI データはどのような様相を呈しただろうか. 以下では，こうした仮想的な発達経路を探索することで，異なる発達経路間の特徴を比較し，擬一次元構造の動径方向, すなわち語彙発達の個人差・多様性についての理解を深める．さらに，decoder を用いて個々の語の単位で表出確率を計算することにより，発達経路ごとの特徴を顕著に示す語彙項目を特定し，潜在空間のより詳細な解釈を目指す.

［図 4 を挿入］

実データと比較する仮想的な発達経路は，次の手順によって定めた. まず, 実データにおいて内側に迂回する直前の点 O，および再び平均的な発達経路に戻ってくる点 P を定義した．簡便のため，定型的な発達経路は OP を結ぶ直線とし，この直線上で，decoder で生成される総表出語彙数が実データ点 C と等しくなる点 B を定めた．これは，各経路における発達の違いを探索するには，総表出語彙数が同じになる点同士で比較するのが適切だと考えたためである．さらに，点 B がちょうど中点になるように直線 A’C を定め，直線 OA’上で総表出語彙数が点 C と等しくなるように点 A を定めた．点 C に対応する実データの総表出語彙数点は 29 語であり，decoder で求めた点 A，B，C の総表出語彙数は 25.4 語であった. こうして定められた外回り (O→A)，中央 (O→B)，内回り (O→C) の 3 つの発達経路に対して, 680 語すべての語彙項目について decoder から求められる各語の表出確率の増分を算出した．この値を比較することで，発達経路によって表出されやすい語彙項目の差異を調べたり，逆に発達経路によらない普遍的な特徴を抽出したりできると考えた.

表 1 は，3 つの発達経路において，表出確率の増分が大きかった語彙項目上位 10 語を抽出したものである．外回りと中央の経路では，ともに “ball” の表出確率の増分が突出して大きかったのに対し，内回り経路では“go”の表出確率の増分が最大となった. “go” は表中で唯一 Action words (動作語) に属する語彙項目であり，内回りでは他にも “me” や “mine” といった Pronouns が見られる点に特徴があった．一方で，外回りで顕著なのは“moo”や“woof woof”等の Sounds が上位に多く現れた点である. これらの知見は, 擬一次元構造では外側で擬音語や具体名詞が，内側で動詞・形容詞や閉じた語が発達的に優位になるという上述の知見と合致する．ただし，“go” は実際には “Let’s go!” や “Where’s it gone?” といった定型句に埋め込まれる形で現れやすいという報告もあるため (深田, 2018; Theakston et al., 2002)，必ずしも特定の行為を示す語とは限らない点には注意を要する.

［表 1 を挿入］

3 つの発達経路すべてに当てはまる共通点も存在した．例えば，“shoe”はどの発達経路でも表出確率が顕著に増加していた．また，どの発達経路でも**表 1** のリストには Games and Routines (日課とあいさつ) が複数含まれていた．ただし，意味カテゴリーは同じでも，リストに含まれる語彙項目それ自体は経路ごとに異なっており，意味カテゴリーとは別に，個々の語も発達的に多様な特徴をもつことが分かった．

実際に，外回り・中央で表出確率の増分が最大となる“ball”と，内回りで表出確率の増分が最大となる“go”の 2 つの語彙項目について，表出確率の等高線を描いたところ (**図 5**)，個々の語彙項目がもつ発達的特徴の多様性はより顕著に確認できた．例えば，“ball”の場合，確かに外回り O→A の経路を辿ると等高線に対して垂直に進むことになるので，表出確率が増えやすくなる．しかし，この語彙項目は点 O の時点で既に表出確率 25%を上回っているため，内回り O→C であっても表出確率は 50%近傍に迫っていた．これに対して，“go”は点 O の時点では表出確率は非常に低く，点 C に至ってようやく 25%を上回る程度であった．このように，発達経路と各語の等高線とを重ね合わせることで，外回りと内回りのどちらで表出確率が増えやすいかという点のみならず，表出確率でみたときにどのような発達フェーズにあるかについても同時にかつ詳細に探索できる可能性が示唆された．

［図 5 を挿入］

さらに，外回り経路と内回り経路との間で表出確率の差分を取ってみると (**図 6**)，ある特定の語彙項目が，特定の発達経路に対して強い感受性をもつことが確認された．これまで見てきたように，外回りにおいて特徴的なのは全体として擬音語や具体名詞であり，内回りの場合はそれらが動作語や代名詞等のカテゴリーに推移する．しかし，個々の語彙項目に注目すると，外回りの場合では“ball”等が経路をよく特徴づける重要な語であることがうかがえる．同様に，内回りの場合は“go”等の語が当該経路において代表的な語であることが分かる．

［図 6 を挿入］

このように，VAE を用いることで，実際の縦断データが辿った発達軌跡と，それとは異なる別の軌跡とを仮想的に比較することが可能である．以上の探索によって，潜在空間のより詳細な解釈が可能となることが示唆されただけでなく，語彙発達の多様性を個人や語といったよりミクロなレベルで調べられる可能性が示された．

4. 考察

本研究では，VAE という機械学習の手法を MCDI データに適用し，潜在空間において語彙発達がどのように表現・可視化されるかを調べた．その結果，2 次元潜在空間において，語彙発達は擬一次元構造をもつこと，その構造を弧として見た場合に，回転方向は総語彙数の変化を，動径方向は総語彙数が同程度の場合の個人差または発達の多様性を，それぞれ表現していることが明らかとなった．また，擬一次元構造の動径方向における多様性について，弧の外側の経路で表出されやすい意味カテゴリー (Sounds 等) と，内側の経路で表出されやすい意味カテゴリー (Pronouns 等) とが存在することが示された．さらに，仮想的に生成された縦断データを用いて事例を分析した結果，異なる発達経路によって表出されやすくなる語彙項目に相違点や共通点があることを明らかにした．それらの発達経路を特徴づけるのは意味カテゴリーだけでなく，特定の語彙項目が重要な役割を担う可能性が示された．

4.1 語彙発達過程を詳細に調べるツールとしての VAE の潜在空間

2 次元潜在空間上に 680 次元の MCDI データを落とし込むことにより，定型的な語彙発達の向き (回転方向) と，それに付随する語彙発達の多様性 (弧の厚み:動径方向) とが一目で確認できる形で可視化された．MCDI データを VAE に適用することで，encoder を用いて特定の事例における MCDI データを潜在空間上に投射したり，decoder を用いて潜在空間上の任意の座標と MCDI データとを対応づけたりする等，総表出語彙数，意味カテゴリー，語彙項目のそれぞれのレベルにおいて，語彙発達の過程を詳細に検討できることが示された．特に，擬一次元構造の外側の経路は，Sounds をはじめ，Animals や Toys などの具体名詞がより早期に表出されやすい経路である一方で，内側の経路は，動詞や形容詞などの述部，閉じた語がより早期に表出されやすい経路であることが可視化

された．このうち，名詞と動詞との対比関係は，言語発達研究において広く関心が寄せられてきたトピックであり (Gentner, 1982; Gentner & Boroditsky, 2001)，名詞語彙数や動詞語彙数がそれぞれどのような別の認知機能と関連するか等が調べられている (例えば，Hagihara et al., 2022a, 2022b; Zuniga-Montanez et al., 2021)．単なるカテゴリーの語彙数を超えて，潜在空間上での位置と認知発達との関連を探索することで，発達の個人差を新たに類型化したり，語彙発達と認知発達との新たな関係性を見出したりすることが今後の研究として期待される．古典的な研究では，早期の語彙発達には，モノの名前を中心に語彙を構成する「参照型」(referential) と，対人的な語を中心に語彙を構成する「表出型」 (expressive) に大別されるという見方 (Nelson, 1973) や，個々の単語の習得から始める「分析型」(analytic) と，決まり文句から習得して後から単語に分解する「全体型」(gestalt/holistic) とに大別されるという見方 (Peters, 1977) 等が提案されてきた．VAE を用いることで，これらの古典的な分類の妥当性を定量的に吟味したり，データ駆動的な分類によってこれまで想定されていなかった新たな語彙発達のスタイルが見つかったりする可能性が考えられる．

また，潜在空間から得られた特徴量から，各語彙項目がどのような発達的特質をもつかを，小数のパラメータによって特徴づけることが可能かもしれない．例えば，総表出語彙数が 50%となる等高線に対して，その語の表出確率が 50%になる等高線が相対的にどのような位置にあるかを定量化することで，その語の習得が早いか遅いかの情報を抽出することができる．また，25%と 75%の等高線の幅の広さを定量化することで，その語が総表出語彙数に対してどの程度急峻または緩徐に習得されるかの情報が得られる．従来，各語の発達的特徴は，その表出が 50%を越える月齢 (＝Age of Acquisition: AoA) として定量的に評価され，発達研究に活用されてきた (例えば，Barrow et al., 2019)．しかし，本研究で探索したように，語彙項目ごとの発達的特徴は，単なる習得の早い／遅いだけには還元されない．潜在空間上にプロットした等高線の情報から，各語がもつ発達的多様性を定量的な特徴量として抽出することで，語彙発達のモデル研究を一層推し進めることが可能になると期待される．

事例分析で確認されたように，異なる発達経路に対して表出確率が大きく変化する語彙項目には違いがあった．このことは，本事例のように総表出語彙数が少ない場合，ある特定の語を表出するかどうかで潜在空間上の座標が変化しやすいことを示唆する．複数の仮想的発達経路を比較し，それぞれの発達経路の特殊性を示す少数の語彙項目を特定できれば，それらの語は語彙発達の多様性の向きを調べる重要なキーと見なせるだろう．これらの少ない語彙項目についてデータを収集することで，ある子どもがどのような発達経路を辿りながら語彙を発達させていくかを予測できる可能性がある．MCDI には，実施の簡便化のためにいくつかの短縮版が存在し，その中には短縮版の得点から本来の MCDI で測られる総表出語彙数を推定できるものもある (Kachergis et al., 2022; Mayor & Mani, 2019)．この方法では単に総表出語彙数しか予測できないが，VAE の潜在空間をもとに特定の発達経路に特に感受性の高い語彙をリスト化することで，語彙発達の個人差の度合いを考慮した短縮版 MCDI を開発できるかもしれない．こうして開発した MCDI を用いることで，子どもの語彙発達のタイプを特定し，後の語彙発達の経路を予測したり，語彙発達のタイプに適した言語入力の方法について提案したりする等の応用可能性が見込まれる．

加えて，異なる発達経路のいずれにおいても表出確率が増えやすい語彙項目が存在した点も興味深い．このことは，語彙発達には多様性だけでなく普遍性も備わっていることを示唆する．例えば，本研究で確認された“shoe”は，名詞優位となる外回り経路でも，動詞優位となる内回り経路でも，表出確率の増分が上位となる語のリストに含まれていた．“shoe”は明らかに「靴」という具体名詞だが，靴には「履く」という特定の関連行為も存在する．このような語の場合，特に語彙発達の初期には意味的に〈モノ〉と〈行為〉とが未分化になることが指摘されている (萩原・阪上, 2019; Hagihara et al., 2022b)．“shoe”には名詞的な側面と動詞的な側面の両面が意味論として含まれていたために，外回り・内回りの両経路において表出しやすかった可能性が考えられる．VAE を用いた仮想的発達経路の生成によって，こうした語彙発達の普遍的性質についてもさらなる探究が可能になるかもしれない．

4.2 今後の展望

本研究は，VAE を用いて MCDI データを分析するという，強い探索的性質をもつ．手法およびデータの

特質から，当然限界も存在する．まず，本研究では，モデルの解釈可能性を高めるために潜在空間を 2 次元としたが，より高次元のモデルの方が適合性が高かった可能性がある．ただし，本研究で用いた MCDI データを予備的に主成分分析した場合，2〜3 次元で十分な説明力があったため，VAE においても低次元の潜在空間に十分に特徴が埋め込まれていると期待される．また，潜在空間の図示・解釈には，事後分布の平均$\boldsymbol{\mu}$のみを使用したが，分散$\boldsymbol{\Sigma}^2$も考慮する方がより適切である．特に，潜在空間上の異なる座標同士では，対応する分散の大きさも異なっているため，例えば潜在空間における 2 つの座標の距離を解析する場合などには，分散を考慮した定式化が求められる．

次に，MCDI は子どもの発話を直接調べたものではなく，あくまでも養育者による報告に基づいた間接データである．そのため，本研究で得られた潜在空間上の発達経路が，子どもの表出語彙発達を本当に反映しているのか，それとも単に養育者にとっての観察しやすさを反映しているのか区別することは容易ではない．実際に，MCDI は他の言語指標とよく相関するという報告もある一方で (Ambridge & Rowland, 2013)，特に動詞については子どもの実際の語彙が適切に反映されていないという問題点も指摘されている (Arunachalam et al., 2022)．近年は，子どもにウェアラブルデバイスを装着して，直接発話データを収集するという取り組みも盛んに行われているため (例えば，Mitsven et al., 2022)，MCDI による間接データとそのような直接データとの統合が将来的には期待される．

最後に，VAE を MCDI データに当てはめる際の限界として，end point に近づくほど，語彙発達の個人差が潜在空間上に適切に反映されていない可能性が挙げられる．MCDI の語彙リストには上限 (680 語) があるため，表出語彙数がリストの上限値に近づくほど，ほとんどの語彙が「言う」として計上される．そのため，リストに含まれていない語彙項目を考慮すれば，end point 近傍の語彙発達は本来であればより高い多様性をもっていると考えられる．

本研究は，VAE を MCDI データに当てはめ，データ駆動的に語彙発達の様相を議論するための基盤を提供するものである．本研究で得られた知見から，今後どのような展望が期待されるだろうか．ここでは 4 点ほど挙げる．まず，各語彙項目の表出確率が潜在空間上でどのように分布するかを調べることによって，各語がもつ特徴量を小数のパラメータで規定し，それらのパラメータによって乳幼児期の語彙をデータ駆動的に再分類することが挙げられる．「名詞」「動詞」といった文法カテゴリー，または「生き物」「小さな家具」といった意味カテゴリーによって各語彙をトップダウン式に集約してしまうのではなく，潜在空間上における特徴量からボトムアップ式に，各語彙を平等に扱いながら再分類することによって，発達的観点から見た場合の語彙の分類法を新たに確立できるかもしれない．実際に，Roy et al. (2015) は，語の文法カテゴリーに依らず，空間的・時間的・言語文脈的に特異性の高い語ほど表出語彙として早期に習得されやすいことを指摘しており，発達的な観点においては，ア・プリオリな語彙の集約方法が必ずしも適切ではないことを示唆している．

次に，潜在空間から得られた各語の発達を特徴づけるパラメータが，語のどのような特性と関係するかを探索することが挙げられる．例えば，語のイメージのしやすさ (Ma et al., 2009) や具体性 (Brysbaert et al., 2014) といった語彙それ自体に内包された特性が，あるパラメータと関連している可能性が考えられる．また，養育者の言語入力における特徴，具体的には多様な語との共起しやすさ (Hills, 2012; Rosa et al., 2017) なども，潜在空間上の特定のパラメータと相関する可能性が期待される．

第三に，本研究ではアメリカ英語の MCDI データを用いたが，他の言語データに本研究の手法を適用することも今後の展望として挙げられる．特に，多くの言語圏と異なり，名詞に対する動詞の優位性が報告されている中国語や韓国語などの言語 (Choi & Gopnik, 1995; Tardif, 1996) の場合，語彙発達が潜在空間上でどのように表現されるかは非常に興味深い．多言語間で潜在空間の特徴を比較することで，語彙発達の言語普遍性と個別性をより深く理解できる可能性がある．なお，Wordbank 上に日本語データは現状存在しないが，今後日本語データも公開されることを強く期待する．特に，日本語の早期語彙は名詞優位である一方で，日本語には主語や目的語が省略可能であったり，行為動詞の意味範囲が英語に比べて狭いものが多かったりする等，動詞優位な側面も併存する (Hagihara et al., 2022b; Imai et al., 2005)．日本語の MCDI データが公開されることで，こうした特殊性を VAE の解析アプローチから議論できるようになるかもしれない．

第四に，特定のポピュレーションにおける語彙データが，VAE の潜在空間上でどのように表現されるかを調べることも今後の課題である．例えば，MCDI を用いて自閉スペクトラム (Jiménez et al., 2021) やレイト・トーカー (Jiménez & Hills, 2022a, 2022b; 奥村・小林, 2019) の語彙発達の特徴を議論した研究があるが，本研究で構築した VAE の潜在空間上にそのような MCDI データを布置することで，定型発達児と比較した際の発達の質的な違いや想定される発達経路などを詳細に議論できる可能性がある．すなわち，本研究で得られた知見は，語彙発達の評価や予測，介入のためのツールとしても応用可能性が期待される．

5. 結語

本研究では, VAE を MCDI データに適用することにより，乳幼児期の語彙発達の普遍性と個別性をより明確に, かつ定量的に描出できることを示した．本研究で提案した手法をより発展させることで，言語発達過程をより詳細に議論するのみならず，言語発達教育・支援への貢献も可能になると期待する．さらに，機械学習モデルを単なる回帰や分類の道具として使うのではなく，生成モデルとしての VAE の特質を最大限活用することで潜在空間の解釈性を追及した点も，本研究の意義であると考える．近年，人工知能をいかにブラックボックスに留めずに，データの解釈・説明のための道具として使うかという議論が盛んに行われている (Tjoa & Guan, 2021). 本研究は，こうした議論をさらに推し進め，認知科学をより発展させることに寄与するだろう．

謝　辞

本研究は JSPS 科研費 (JP20H05002，JP21J00750, JP22K13664, JP23K11786) の助成を受けた．本研究の実施にあたり有益な指摘をくださった浅井怜斗氏，深田智氏に感謝申し上げます．

文　献

Ambridge, B., & Rowland, C. F. (2013). Experimental methods in studying child language acquisition. *WIREs Cognitive Science Interdisciplinary Reviews*, *4*(2), 149–168. https://doi.org/10.1002/wcs.1215

Arunachalam, S., Avtushka, V., Luyster, R. J., & Guthrie, W. (2022). Consistency and inconsistency in caregiver reporting of vocabulary. *Language Learning and Development*, *18*(1), 81–96. https://doi.org/10.1080/15475441.2021.1931233

Barrow, C. W., Björkenstam, K. N., & Strömbergsson, S. (2019). Subjective ratings of age-of-acquisition: exploring issues of validity and rater reliability. *Journal of Child Language*, *46*(2), 199–213. https://doi.org/10.1017/S0305000918000363

Beckage, N., Smith, L., & Hills, T. (2011). Small worlds and semantic network growth in typical and late talkers. *PloS One*, *6*(5), e19348. https://doi.org/10.1371/journal.pone.0019348

Bergelson, E., & Swingley, D. (2012). At 6–9 months, human infants know the meanings of many common nouns. *Proceedings of the National Academy of Sciences*, *109*(9), 3253–3258. https://doi.org/10.1073/pnas.1113380109

Bleses, D., Makransky, G., Dale, P. S., Højen, A., & Ari, B. A. (2016). Early productive vocabulary predicts academic achievement 10 years later. *Applied Psycholinguistics*, *37*(6), 1461–1476. https://doi.org/10.1017/S0142716416000060

Borovsky, A. (2022). Lexico-semantic structure in vocabulary and its links to lexical processing in toddlerhood and language outcomes at age three. *Developmental Psychology*, *58*(4), 607–630. https://doi.org/10.1037/dev0001291

Brysbaert, M., Warriner, A. B., & Kuperman, V. (2014). Concreteness ratings for 40 thousand generally known English word lemmas. *Behavior Research Methods*, *46*, 904–911. https://doi.org/10.3758/s13428-013-0403-5

Choi, S., & Gopnik, A. (1995). Early acquisition of verbs in Korean: A cross-linguistic study. *Journal of Child Language*, *22*, 497–529. https://doi.org/10.1017/S0305000900009934

Deng, L. (2012). The MNIST database of handwritten digit images for machine learning research. *IEEE Signal Processing Magazine*, *29*(6), 141–142. https://doi.org/10.1109/MSP.2012.2211477

Fenson, L., Dale, P. S., Reznick, J. S., Thai, D., Bates, E., Hartung, J. P., Pethick, S., Relly, J. S. (1993). MacArthur communicative development inventories: User's guide and technical manual. Singular Press.

Frank, M. C., Braginsky, M., Yurovsky, D., & Marchman, V. A. (2016). Wordbank: An open repository for developmental vocabulary data. *Journal of Child Language*, *44*(3), 677–694. https://doi.org/10.1017/S0305000916000209

Frank, M. C., Braginsky, M., Yurovsky, D., & Marchman, V. A. (2021). Variability and consistency in early language learning: The Wordbank project. MIT Press.

深田 智 (2018). “Let's + 移動動詞”表現と子どもの運動能力及び社会性の発達．日本認知言語学会論文集, *18*, 555–560.

Gentner, D. (1982). Why nouns are learned before verbs: Linguistic relativity versus natural partitioning. In S. A. Kuczaj II (Ed.), *Language development, Vol. 2: Language, thought, and culture* (pp. 301–334). Lawrence Erlbaum.

Gentner, D., & Boroditsky, L. (2001). Individuation, relativity, and early word learning. In M. Bowerman & S. C. Levinson

(Eds.), *Language acquisition and conceptual development* (pp. 215–256). Cambridge University Press.

萩原 広道 (2021). 言語の発達 浅野大喜 (編) 人間発達学 (pp.140–175) メジカルビュー社

萩原 広道・阪上 雅昭 (2019). 初期言語発達における意味の全体性と可塑的変化：子どものことばに品詞構造はあるか？. ベビーサイエンス, *18*, 14–24.

Hagihara, H., Ishibashi, M., Moriguchi, Y., & Shinya, Y. (2022a). Object labeling activates young children's scale errors at an early stage of verb vocabulary growth. *Journal of Experimental Child Psychology*, *222*, 105471. https://doi.org/10.1016/j.jecp.2022.105471

Hagihara, H., Yamamoto, H., Moriguchi, Y., & Sakagami, M. (2022b). When "shoe" becomes free from "putting on": The link between early meanings of object words and object-specific actions. *Cognition*, *226*, 105177. https://doi.org/10.1016/j.cognition.2022.105177

Hills, T. (2012). The company that words keep: Comparing the statistical structure of child- versus adult-directed language. *Journal of Child Language*, *40*(3), 586–604. https://doi.org/10.1017/S0305000912000165

Hinton, G. E., & Salakhutdinov, R. R. (2006). Reducing the dimensionality of data with neural networks. *Science*, *313*(5786), 504–507. https://doi.org/10.1126/science.1127647

Imai, M., Haryu, E., & Okada, H. (2005). Mapping novel nouns and verbs onto dynamic action events: Are verb meanings easier to learn than noun meanings for Japanese children? *Child Development*, *76*(2), 340–355. https://doi.org/10.1111/j.14678624.2005.00849_a.x

Jiménez, E., Haebig, E., & Hills, T. T. (2021). Identifying Areas of Overlap and Distinction in Early Lexical Profiles of Children with Autism Spectrum Disorder, Late Talkers, and Typical Talkers. *Journal of Autism and Developmental Disorders*, *51*(9), 3109–3125. https://doi.org/10.1007/s10803-020-04772-1

Jiménez, E., & Hills, T. T. (2022a). Semantic maturation during the comprehension‐expression gap in late and typical talkers. *Child Development*, *93*(6), 1727–1743. https://doi.org/10.1111/cdev.13815

Jiménez, E., & Hills, T. T. (2022b). Differences in the semantic structure of the speech experienced by late talkers, late bloomers, and typical talkers. *Developmental Psychology*, *59*(1), 141–160. https://doi.org/10.1037/dev0001446

Kachergis, G., Marchman, V. A., Dale, P. S., Mankewitz, J., & Frank, M. C. (2022). Online computerized adaptive tests of children's vocabulary development in English and Mexican Spanish. *Journal of Speech, Language, and Hearing Research*, *65*(6), 2288–2308. https://doi.org/10.1044/2022_JSLHR-21-00372

Kingma, D. P., & Welling, M. (2013). Auto-encoding variational Bayes. *arXiv*. https://doi.org/10.48550/arXiv.1312.6114

小林 哲生・南 泰浩・杉山 弘晃 (2013). 語彙爆発の新しい視点：日本語学習児の初期語彙発達に関する縦断データ解析. ベビーサイエンス, *12*, 40–64.

Laing, C. (2019a). A role for onomatopoeia in early language: Evidence from phonological development. *Language and Cognition*, *11*(2), 173–187. https://doi.org/10.1017/langcog.2018.23

Laing, C. E. (2019b). Phonological motivation for the acquisition of onomatopoeia: An analysis of early words. *Language Learning and Development*, *15*(2), 177–197. https://doi.org/10.1080/15475441.2019.1577138

Ma, W., Golinkoff, R. M., Hirsh-Pasek, K., McDonough, C., & Tardif, T. (2009). Imageability predicts the age of acquisition of verbs in Chinese children. *Journal of Child Language*, *36*(2), 405-423. https://doi.org/10.1017/S0305000908009008

Marchman, V. A., & Fernald, A. (2008). Speed of word recognition and vocabulary knowledge in infancy predict cognitive and language outcomes in later childhood. *Developmental Science*, *11*(3), F9–F16. https://doi.org/10.1111/j.1467-7687.2008.00671.x

Mayor, J., & Mani, N. (2019). A short version of the MacArthur–Bates Communicative Development Inventories with high validity. *Behavior Research Methods*, *51*, 2248–2255. https://doi.org/10.3758/s13428-018-1146-0

Mitsven, S. G., Perry, L. K., Tao, Y., Elbaum, B. E., Johnson, N. F., & Messinger, D. S. (2022). Objectively measured teacher and preschooler vocalizations: Phonemic diversity is associated with language abilities. *Developmental Science*, *25*(2), e13177. https://doi.org/10.1111/desc.13177

Moore, C., Dailey, S., Garrison, H., Amatuni, A., & Bergelson, E. (2019). Point, walk, talk: Links between three early milestones, from observation and parental report. *Developmental Psychology*, *55*(8), 1579–1593. https://doi.org/10.1037/dev0000738

Morgan, P. L., Farkas, G., Hillemeier, M. M., Hammer, C. S., & Maczuga, S. (2015). 24‐month‐old children with larger oral vocabularies display greater academic and behavioral functioning at kindergarten entry. *Child Development*, *86*(5), 1351‐1370. https://doi.org/10.1111/cdev.12398

Nelson, K. (1973). Structure and strategy in learning to talk. *Monographs of the Society for Research in Child Development*, *38*(1-2, Serial No 149), 136. https://doi.org/10.2307/1165788

小椋 たみ子 (2007). 日本の子どもの初期の語彙発達. 言語研究, *132*, 29–53. https://doi.org/10.11435/gengo.132.0_29

小椋 たみ子・綿巻 徹 (2004). 日本語マッカーサー乳幼児言語発達質問紙「語と文法」手引 京都国際社会福祉センター

小椋 たみ子・綿巻 徹・稲葉 太一 (2016). 日本語マッカーサー乳幼児言語発達質問紙の開発と研究 ナカニシヤ出版

奥村 優子・小林 哲生 (2019). 日本語レイト・トーカーにおける表出語彙のカテゴリ構成の検討. 音声言語医学, *60*(3), 205–213. https://doi.org/10.5112/jjlp.60.205

Peters, A. M. (1977). Language learning strategies: Does the whole equal the sum of the parts? *Language*, *53*(3), 560–573. https://doi.org/10.2307/413177

Rezende, D. J., Mohamed, S., & Wierstra, D. (2014). Stochastic backpropagation and approximate inference in deep generative models. *Proceedings of the 31st International Conference on Machine Learning*, *32*(1), 1278–1286.

Rosa, E., Tapia, J. L., & Perea, M. (2017). Contextual diversity facilitates learning new words in the classroom. *PLoS One*, *12*, e 01790 04. https://doi.org/10.1371/journal.pone.0179004

Roy, B. C., Frank, M. C., DeCamp, P., Miller, M., & Roy, D. (2015). Predicting the birth of a spoken word. *Proceedings of the National Academy of Sciences*, *112*(41), 12663–12668. https://doi.org/10.1073/pnas.1419773112

阪上 雅昭・水谷 天智・萩原 広道・山本寛樹 (2023, March 24). 変分オートエンコーダーによる乳幼児の語彙発達の解析. 日本物理学会 2023 年春季大会　オンライン

Tardif, T. (1996). Nouns are not always learned before verbs: Evidence from Mandarin speakers' early vocabularies. *Developmental Psychology*, *32*(3), 492–504. https://doi.org/10.1037/0012-1649.32.3.492

Theakston, A. L., Lieven, E. V. M., Pine, J. M., & Rowland, C. F. (2002). Going, going, gone: The acquisition of the verb 'go.' *Journal of Child Language*, *29*(4), 783–811. https://doi.org/10.1017/S030500090200538X

Tjoa, E., & Guan, C. (2021). A survey on explainable artificial intelligence (XAI): Toward medical XAI. *IEEE Transactions on Neural Networks and Learning Systems*, *32*(11), 4793-4813. https://doi.org/10.1109/TNNLS.2020.3027314

Westermann, G., & Mareschal, D. (2014). From perceptual to language-mediated categorization. *Philosophical Transactions of the Royal Society B: Biological Sciences*, *369*(1634), 20120391. https://doi.org/10.1098/rstb.2012.0391

Zuniga-Montanez, C., Kita, S., Aussems, S., & Krott, A. (2021). Beyond the shape of things: Infants can be taught to generalize nouns by objects' functions. *Psychological Science*, *32*(7), 1073–1085. https://doi.org/10.1177/0956797621993107

脚注

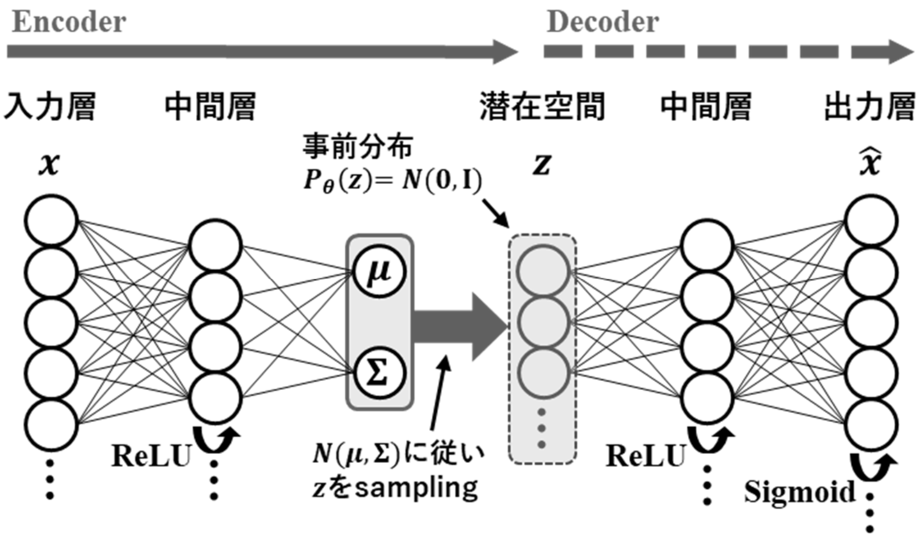
1. 脚注 1　なお，日本語版 MCDI の「語と文法」版の対象年齢は 16〜36 ヶ月である．
2. 脚注 2　図 1 では，encoder と decoder の間は正規分布からのサンプリングという過程を経ていわば断絶されており，バックプロパゲーションに必要な微分ができない．そこで，reparameterization trick (Kingma & Welling, 2013) と呼ばれる方法を採る．これは，$\boldsymbol{z}$を正規分布$N(\boldsymbol{\mu}, \boldsymbol{\Sigma})$から直接ランダムサンプリングする代わりに，ランダムノイズ $\boldsymbol{\varepsilon} \sim N(\mathbf{0}, \mathbf{I})$ を別に生成し，これを分散$\boldsymbol{\Sigma}^2$に掛け，$z_i = \mu_i + \Sigma_i \times \varepsilon_i$ (iは潜在空間の次元数) としてモデルに潜り込ませる操作を指す．この操作により，$\boldsymbol{z}$のサンプリング過程を模しつつ，$\boldsymbol{z}$が$\boldsymbol{\mu}$と$\boldsymbol{\Sigma}$に対して確定的となり微分できるようになる．
3. 脚注 3　機械学習のデータセットとして広く用いられている MNIST データ (Deng, 2012) の入力次元は 784 次元であり，MCDI データと同程度と見なせる．このデータを使った VAE のモデル (http://mbd.riken.jp/VAE/lecture.html) を参考に本研究における中間層の次元を定めた．

表 1

	外回り経路 (O→A)			中央経路 (O→B)			内回り経路 (O→C)		
順位	カテゴリー	語彙項目	表出確率の増分	カテゴリー	語彙項目	表出確率の増分	カテゴリー	語彙項目	表出確率の増分
1	Toys	ball	0.526	Toys	ball	0.391	Action Words	go	0.225
2	Games & Routines	bye	0.359	Clothing	shoe	0.280	Games & Routines	yes	0.204
3	Clothing	shoe	0.339	Games & Routines	no	0.265	Games & Routines	thank you	0.177
4	Sounds	moo	0.328	People	baby	0.254	Sounds	ouch	0.170
5	Games & Routines	no	0.307	Games & Routines	bye	0.254	Pronouns	me	0.167
6	Sounds	woof woof	0.299	Games & Routines	hi	0.200	People	baby	0.164
7	People	baby	0.296	Food & Drink	banana	0.199	Pronouns	mine	0.155
8	Toys	book	0.280	Toys	book	0.199	Clothing	shoe	0.154
9	Sounds	uh oh	0.279	People	grandma*	0.186	Games & Routines	shh/shush/hush	0.151
10	Food & Drink	banana	0.266	Descriptive Words	hot	0.185	Question Words	what	0.149

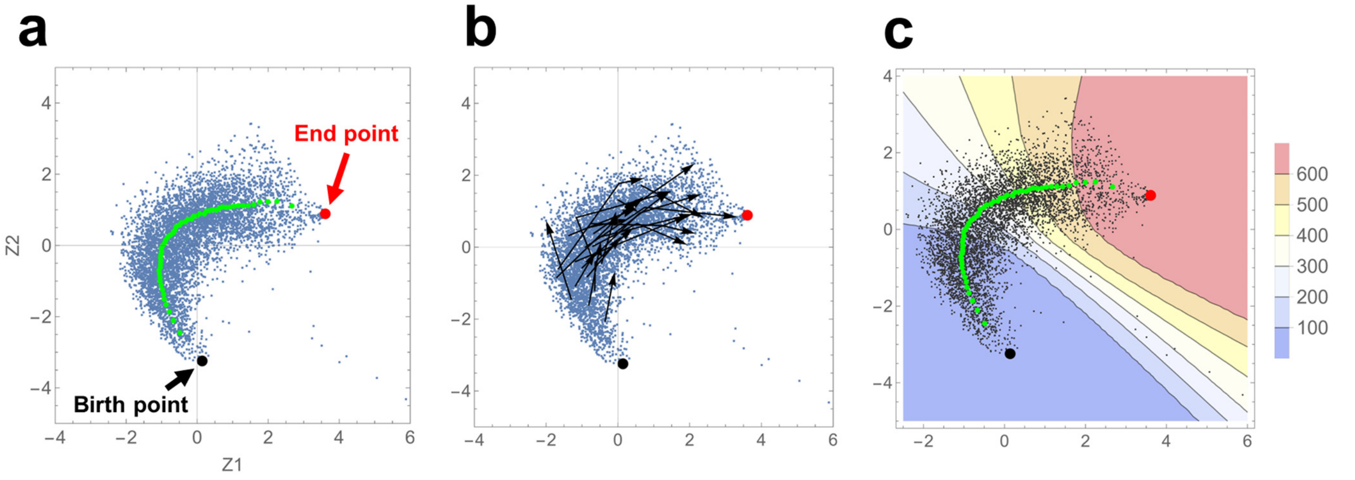
［表タイトル］各発達経路における表出確率の増分が大きかった語彙項目 (上位 10 語)

図 1



［図タイトル］本研究で用いた VAE の概略図［脚注 1］

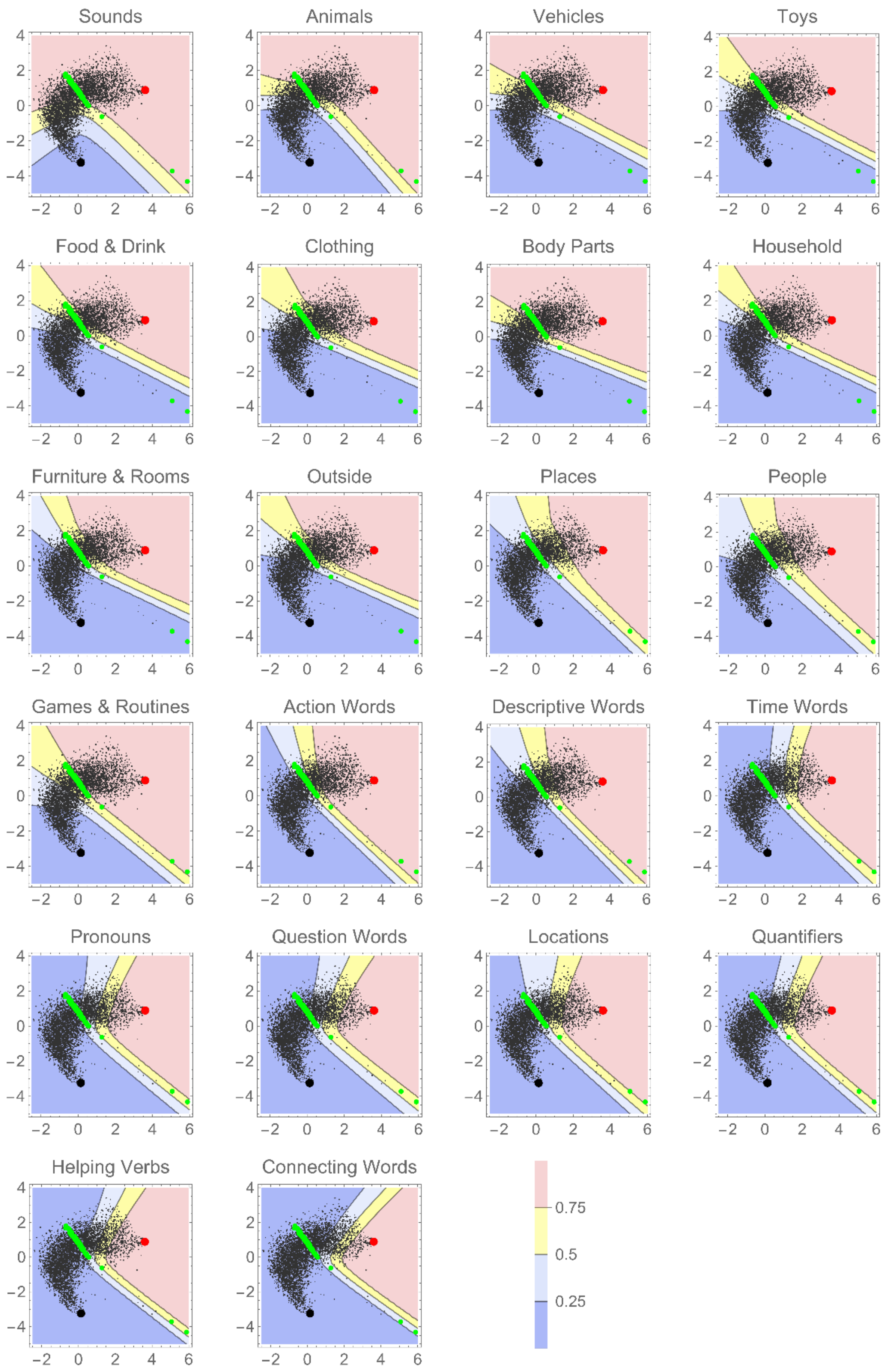
図 2



［図タイトル］本研究で採用した VAE モデルの潜在空間

［キャプション］表出語彙発達の擬一次元構造．(a) 表出語彙数の始点 (birth point) および終点 (end point) を encoder によって投射し，その他の各点は実際の子どものデータ点 (N = 5,520) を投射した．(b) MCDI の縦断データの一部を，月齢の向きにそって矢印で繋いだ．(c) decoder から求めた総表出語彙数について，潜在空間上に等高線を重ね描きした．緑の点は平均的な発達経路 (平均発達曲線) を示す．

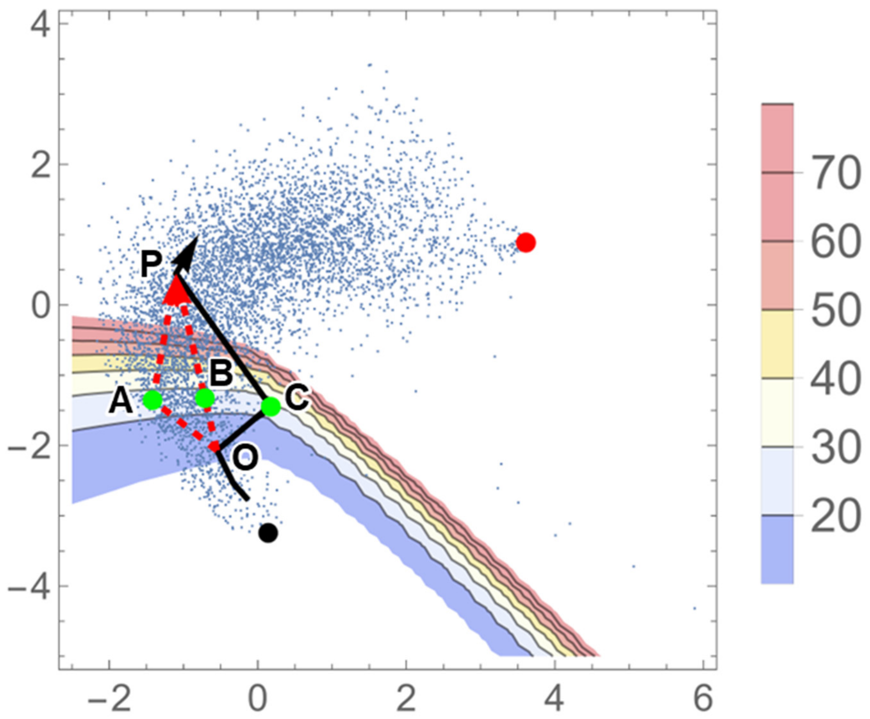
図 3



［図タイトル］各意味カテゴリーにおける潜在空間上の表出確率の分布

［キャプション］アメリカ英語版 MCDI を構成する 22 の意味カテゴリーについて，表出確率の平均が 25，50，75% の場合の等高線を図示した (birth point，end point，および各データ点は図 2 と同様)．緑で示した帯状のデータ点は，総表出語彙数が全体の約 50% (＝330〜350 語) となる点である．意味カテゴリーの並びは MCDI のリスト順に準拠するが (Sounds, Animals, …の順に，左上から右下へ)，概ね擬一次元構造の外側で発達が早い (＝表出確率が高い) ものから順に並んでおり，リストの後半になるほど内側で発達が早いものに推移している．

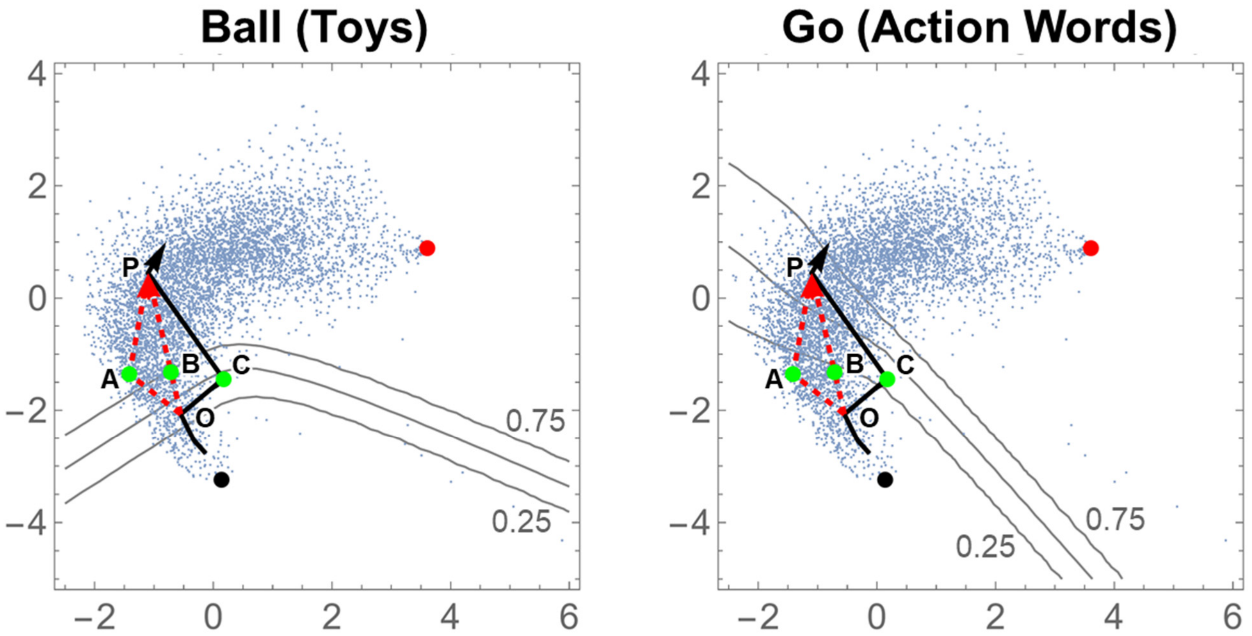
図 4



［図タイトル］表出語彙発達の事例分析

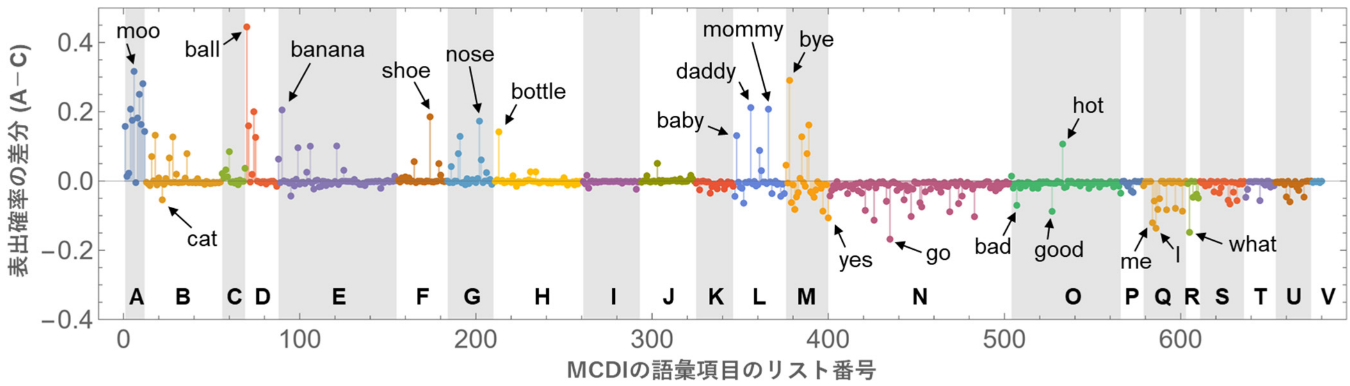
［キャプション］実線は実データの発達経路を，点線は仮想的な発達経路を示す．点 A，B は，decoder で求めた総表出語彙数が点 C と一致するように定めた．等高線は総表出語彙数から描出した．

図 5



［図タイトル］語彙項目“ball”と“go”の表出確率の分布
［キャプション］各語の表出確率を等高線で示した．その他の凡例は図 4 と同じ．

図 6



［図タイトル］MCDI を構成する各語彙項目の表出確率の差分

［キャプション］外回り (O→A) の表出確率から, 内回り (O→C) の表出確率を差し引いた. A) Sounds, B) Animals, C) Vehicles, D) Toys, E) Food & Drink, F) Clothing, G) Body Parts, H) Household, I) Furniture & Rooms, J) Outside, K) Places, L) People, M) Games & Routines, N) Action Words, O) Descriptive Words, P) Time Words, Q) Pronouns, R) Question Words, S) Locations, T) Quantifiers, U) Helping Verbs, V) Connecting Words.

付録：VAE の学習手続き

VAE は元来，潜在空間の事後分布を変分ベイズ法で推定するという文脈で提案されたものであり，その過程で AE の構造が自ずと現れる．以下，この文脈に沿ってネットワークの訓練過程について概説する．

VAE では，潜在変数$\boldsymbol{z}$はあるパラメータθに従う確率分布$p_\theta(\boldsymbol{z})$から得られたと仮定し，さらに観測データ$\boldsymbol{x}$は, 条件付き確率$p_\theta(\boldsymbol{x}|\boldsymbol{z})$から生成されたと考える．高次元ベクトルの入力データセットを$\boldsymbol{X} = \{\boldsymbol{x}_i | i = 1, \dots, N\}$とすると，$\boldsymbol{x}$の周辺確率

$$p_\theta(\boldsymbol{x}) = \int p_\theta(\boldsymbol{x}|\boldsymbol{z})\, p_\theta(\boldsymbol{z}) d\boldsymbol{z} \tag{1}$$

から得られる対数尤度

$$\log p_\theta(\boldsymbol{x}_1, \dots, \boldsymbol{x}_N) = \sum_{i=1}^{N} \log p_\theta(\boldsymbol{x}_i) \tag{2}$$

を最大化することで，潜在変数$\boldsymbol{z}$が従うパラメータθや確率分布$p_\theta(\boldsymbol{x})$を求めることがここでの目的となる．しかし，(1) の積分を直接計算することは難しいため，変分ベイズ法を用いて推定を行う．以下，VAE における推定の手続きについて述べる．

まず，潜在変数の事前分布には多変量標準正規分布$p_\theta(\boldsymbol{z}) = N(\boldsymbol{0}, \mathbf{I})$を置く．さらに，真の事後分布を近似する確率分布として，$q_\phi(\boldsymbol{z}|\boldsymbol{x})$を導入する．ここで$\boldsymbol{\phi}$は変分パラメータを表す．次に，潜在変数の事後確率を$p_\theta(\boldsymbol{z}|\boldsymbol{x}) = q_\phi(\boldsymbol{z}|\boldsymbol{x}) = N(\boldsymbol{\mu}, \boldsymbol{\Sigma})$という正規分布でモデル化する．ニューラルネットワークでモデル化された事後分布$q_\phi(\boldsymbol{z}|\boldsymbol{x})$は図 1 の encoder に対応しており, その出力は多変量正規分布の平均$\boldsymbol{\mu}$と分散$\boldsymbol{\Sigma}^2$になっている．他方, 条件付き確率$p_\theta(\boldsymbol{x}|\boldsymbol{z})$は decoder に対応している．したがって, $\boldsymbol{\phi}$とθはそれぞれ encoder, decoder がもつ多次元パラメータとなる．

変分ベイズ法では，対数尤度$\log p_\theta(\boldsymbol{x}_i)$に

$$\begin{aligned} \log p_\theta(\boldsymbol{x}_i) &> L(\theta, \phi; \boldsymbol{x}_i) \\ &= -D_{KL}\left(q_\phi(\boldsymbol{z}|\boldsymbol{x}_i) \parallel p_\theta(\boldsymbol{z})\right) \\ &\quad + E_{q_\phi(\boldsymbol{z}|\boldsymbol{x}_i)}[\log p_\theta(\boldsymbol{x}_i|\boldsymbol{z})] \end{aligned} \tag{3}$$

という変分下限Lが存在することを利用し，このLを最大化するようにパラメータ$\boldsymbol{\phi}$，θの学習を行う．ここで，第 1 項D_{KL}はカルバック・ライブラー情報量であり，事前分布$p_\theta(\boldsymbol{z})$に近い事後分布$q_\phi(\boldsymbol{z}|\boldsymbol{x}_i)$を選ぶための条件となっている．すなわち，encoder から生成される$\boldsymbol{z}$の確率分布$q_\phi(\boldsymbol{z}|\boldsymbol{x}) = N(\boldsymbol{\mu}, \boldsymbol{\Sigma})$を，$p_\theta(\boldsymbol{z}) = N(\boldsymbol{0}, \mathbf{I})$と比較し，これらが近くなるように学習する．なお，第 1 項は$q_\phi(\boldsymbol{z}|\boldsymbol{x}) = p_\theta(\boldsymbol{z})$のときに最小値 0 をとる非負値であり，両者の分布が近いほど値は小さくなる．第 2 項Eは$\log p_\theta(\boldsymbol{x}_i|\boldsymbol{z})$の期待値であり，encoder および decoder で再構成された生成データ$\hat{\boldsymbol{x}}_i$が，元の入力データ$\boldsymbol{x}_i$にどれだけ近いかを示す対数尤度である．したがって，変分下限Lは，潜在変数$\boldsymbol{z}$の値の取り方が正規分布に近くなるように制約をかける部分 (第 1 項) と，入出力データの差分を測る部分 (第 2 項) から構成されるということになる．