

# 日本語と英語の母語の発達過程の比較 — 発話数・形態素数・平均発話長の分析 —

稲葉 みどり

日本語教育講座

## Developmental Changes of MLU in Japanese- and English- Language Narratives

Midori INABA

Department of Teaching Japanese as a Foreign Language, Aichi University of Education, Kariya 448-8542, Japan

### 要 約

本研究では、第一言語としての日本語の発達過程を英語の発達過程と比較した。日本語を第一言語とする子ども (JL1) と英語を第一言語とする子ども (EL1) の物語文を同じ手法で分析し、結果を比較することにより、両者の類似点、相違点から発達の普遍的側面や個別言語に特有の側面等を探ることを目的としている。また、英語との比較により、日本語の発達の特徴をより明確に提示し、新たな知見を得ることをめざしている。研究では、JL1とEL1について、1) 発話数の年齢による推移、2) 形態素数の年齢による推移、3) 平均発話長から見た発達過程、及び、4) 個人差から見た発達過程等を比較分析した。結果から、発話数は、JL1では3~9歳間であまり増加しないが、EL1では4~9歳にかけてJL1よりも大きな増加が見られ、9歳児の発話数は両者で近い値を示した。形態素数はJL1の方がEL1よりも全体に高い値を示したが、年齢と共に純増する発達傾向は両言語で類似していた。日本語の平均発話長をMLUm、MLUwの2つの手法で算出した結果、MLUmは、MLUwよりかなり高い値となり、分析方法によってかなり数値に差が出ることが分かった。さらに、大人の値への到達度という観点から平均発話長を分析し、JL1とEL1において5歳頃までにほぼ大人の9割の値に到達するという類似の傾向が見られること発見した。個人差の分析では、JL1、EL1において、バラツキと安定を繰り返しながら次第に高い値になる様子が見られた。これらの結果は、本研究の対象となった言語資料のみから得られた結果であり、今後はより広い言語資料を用いた更なる分析により検証していく必要がある。

Keywords : 物語文、発話数、形態素数、平均発話長 (MLUw, MLUm)、個人差

### 1. 研究の目的

本研究では、日本語と英語の母語 (第一言語) の発達過程を比較する。対象となる発話資料は、日本語と英語の口頭による物語文である。発話資料は、日本語を第一言語とする子どもと大人から収集した日本語の物語文、及び、英語を第一言語とする子どもと大人から収集した英語の物語文である。

発達過程の分析は、発話数、形態素数、平均発話長を指標として、年齢による量的な推移を捉える。物語文の発達においては、発話量 (発話数、単語数を含む) の増加は語彙や文法等の発達に関わる重要な要因の一つ、形態素数や平均発話長の増加は発話の文法的複雑さの増加と関わっていると考えられるからである。

発話資料となる物語文は、文字のない絵本 *Frog, Where Are You?* (Mayer, 1956) を見て語った口頭作話で、同じ物語文を扱った研究では、FrogStory研究として知られ、様々な言語で発達の過程が研究されている。日本語のFrog Storyを対象とした先行研究で、筆者は、物語文の談話構成 (Inaba, 2002)、全体構造 (Inaba, 1999)、局所構造 (Inaba, 2001)、時制転換 (Inaba, 2003) 等に関する研究を行ってきたので、本研究では、物語文の発達を発話数、形態素数、平均発話長という言語知識の発達の観点から分析する。

分析では、日本語と英語の発達を比較し、両者の類似点と相違点を明らかにすることで、発達の普遍的側面や個別言語に特有の側面等を探る。

## 2. 研究の背景

### 2.1 物語文の発達指標

本研究は、第一言語としての日本語発達の過程を考察した稲葉 (2017a)、及び、第一言語としての日本語と第二言語としての日本語の発達過程を比較した稲葉 (2017b) との関連研究 (2.2参照) である。これらの研究では、言語知識の発達に焦点を絞り、物語文の発達過程を、発話数、単語数、形態素数、文法発達 (平均発話長) の観点から考察した。

本筋、前景描写や背景描写、出来事、評価、背景・設定等の要素が適宜含まれている整った物語文を構成するには、ある程度の発話数が必要であり、これらの要素を含んだ物語文への発達の過程では、語彙の増加、様々な機能を持つ形態素の獲得、それに伴う文法構造の複雑化等が起こることが予測される。

よって、これらを発達の指標と捉えるのは、発話数は物語文の構成面での発達の1つの指標、単語数は語彙の発達の1つの指標、形態素数は文法構造の発達の1つの指標、平均発話長 (MLUm・MLUw) は発話の文法的複雑さの1つの指標になるという考えに基づいている。勿論、発話数、単語数、形態素数、文法発達等の言語知識は、整った物語文かどうかを判定する究極の指標とは言えないが、物語談話の構成に最低限必要な要因である。

### 2.2 関連研究

稲葉 (2017a) では、日本語を第一言語とする子どもの物語文の発達を年齢による発話数、単語数、形態素数、形態素MLU (MLUm) と自立語MLU (MLUw) の推移の観点から分析した。その結果、発話数の変化は全年齢であまり大きく増加しないこと、単語数、形態素数は3~5歳で著しく増加すること、MLUmとMLUwは、3~5歳で増加が見られ、5歳で大人の値に近づくこと、MLUmの方がMLUwよりも文法発達の推移を敏感に捉えていること等が明らかになった。

稲葉 (2017b) では、日本語を第二言語として学成人の英語話者 (JL2) と日本語を第一言語とする子ども (JL1) の物語文の発達過程を比較した。稲葉 (2017a) と同様に、発話数、単語数、形態素数、MLUm、MLUwの比較分析を行った。結果から、発話数はJL2では、発達初期から比較的高い値で、レベルが上がると更に増加するが、JL1の場合は、発話数は、年齢が上がってもほぼ横這いに近い状態であること、単語数、形態素数もJL2はJL1に比べると、発達初期から比較的高い値であるが、JL2、JL1ともにレベル、年齢とともに増加すること、MLUm、MLUwは、JL2では、レベルIからIVにかけて純増するが、レベルVでは減少すること、一方、JL1では、3~5歳にかけて急増し、その後は横ばい状態であること等を報告した。

### 2.3 研究課題

本研究は、稲葉 (2017a) の続編で、第一言語としての日本語の発達の過程を英語の発達過程と比較する。日本語を第一言語とする子ども (JL1) と英語を第一言語とする子ども (EL1) の物語文を同じ手法で分析を行い、結果を比較することにより、両者の類似点、相違点から発達の普遍的側面や個別言語に特有の側面等を考察する。また、比較により、日本語の発達の特徴をより明確に提示することをめざしている。

研究は以下の手順で進める。JL1とEL1について、1) 発話数の年齢による推移、2) 形態素数の年齢による推移、3) 平均発話長から見た発達過程、4) 個人差から見た発達過程等の分析を行い、両者を比較する。MLUは、一般に第一言語の早期の発達の指標とされているが、あえて大人の言語資料にまで適用し、同じ指標で分析することにより、新たな知見を得ることを目的とする。

## 3. 研究の方法

### 3.1 発話資料と被験者

日本語の言語発達資料 (Frog Story) は、3、4、5、9歳の日本語を母語とする子ども (各年齢10人、合計40名)、及び、大人 (50人) から収集したものである。被験者の概要は【表1】に示した。

英語を第一言語とする子どもと大人のFrog Storyは、CHILDES database online の English-Slobin Corpus<sup>1</sup>よりダウンロードした。これは、Berman and Slobin (1994) が収集したFrogStoryである。被験者の概要は【表2】の通りである。

【表1】英語を第一言語とする子どもと大人 (JL1)

年齢	年齢レンジ	平均	人数
3歳児	3:2-3:11	3:7	10
4歳児	4:0-4:11	4:6	10
5歳児	5:0-5:10	5:4	10
9歳児	9:0-9:11	9:5	10
大人	-	-	50

【表2】英語を第一言語とする子どもと大人 (EL1)

age	age range	mean	N
3 yrs	3:1-3:11	3:7	12
4 yrs	4:0-4:11	4:7	12
5 yrs	5:1-5:11	5:7	12
9 yrs	9:1-9:11	9:7	12
大人	-	-	12

\*Berman and Slobin (1994: 31) を基に作成

### 3.2 コーディングと解析プログラム

本研究で収集した発話資料は、MacWinney (2000) による CHILDES (Child Language Data Exchange System) を用いて、CHAT (Codes for the Human Analysis of Transcripts) 形式でデータベース化した。日本語フォーマットは、宮田・森川・村木 (2004) を用いた。データには、宮田 (2012) の JMOR 形態素コードに沿って、形態素コード MOR (Miyata & Naka, 2010) を付与した。MLU<sub>w</sub>、及び、MLU<sub>m</sub> の計算は、宮田 (2012) の以下のガイドライン (3.3 参照) に沿って行った。データ解析には、CLAN (Computerized Language Analysis) プログラムを用いた。

### 3.3 平均発話長の解析方法

平均発話長 (Mean Length of Utterances, MLU) は、発話の長さを表し、1 発話あたりの語数の平均である。これは、Brown (1973) により英語を獲得する子どもの文法発達の指標として提案されたもので、Brown (1973) の研究では、第一言語の発達においては、月齢との相関があるとされている。しかし、この語数をどのように捉えるかで、数値が変わってくる。これは、もともと英語を対象として考案された方法なので、膠着語である日本語にはそのまま当てはめるのは難しい。文法発達の指標として用いる場合、日本語の形態素の特徴をうまく反映できるような調整が必要である。そのために開発されたのが、宮田 (2012) の提案する自立語 MLU<sub>w</sub> (以下、MLU<sub>w</sub>) と、形態素 MLU<sub>m</sub> (以下、MLU<sub>m</sub>) である。

本研究では、日本語の物語文の平均発話長の算出に、宮田 (2012) の提案する自立語 MLU<sub>w</sub> (以下、MLU<sub>w</sub>) と、形態素 MLU<sub>m</sub> (以下、MLU<sub>m</sub>) の両方の方法を用いることにする。また、英語の物語文は CLAN の mlu コマンドによる解析を行い、結果を MLU<sub>e</sub> と表記することにする。

### 3.4 平均発話長の算出方法

宮田 (2012) では、主に以下のような平均発話長の算出方法を提案している。MLU<sub>w</sub> は、自立語の数を基に MLU を算出する方法で、接頭辞、活用語尾、助詞などの付属語、判定詞は、個別にカウントされない方法である。主な方法は、1) 名詞および固有名詞は付属語を含めて 1 語として計算する。名字と名前は別の単語として分ける。2) 格助詞、終助詞、取立助詞、接続助詞はカウントされない。接続詞はカウントする。3) 動詞の活用語尾は個別にカウントしない。助動詞もカウントしない。判定詞 (だ、です等) もカウントしない。形容動詞の場合も、判定詞 (だ、な、で、に、です、だった等) の部分はカウントしない。形容詞の活用語尾も個別に数えない。4) する、なるは独立した動詞として数える。5) 複合語は分けない。数字は百、

千、万単位で分ける。助数詞は分けない等である。

MLU<sub>m</sub> は、「子ども + たち + が」語幹 + 語尾 + 格助詞や「食べ + られ + た」語幹 + 受け身 + 完了形 (両方 3 形態素) というように単語と付属語を分けて計算を行う方法である。具体的な方法は、1) 名詞の場合は語幹のほか語尾や接頭辞をそれぞれ別個に数える。2) 助詞のが、を、の、は、等の助詞は別個に数える。3) 動詞の語尾は、別個にカウントする。時制の語尾の前に受け身、使役などの語尾が入っている場合は、それらも別個に数える。形容詞も同じく活用語尾が個別に数えられる。判定詞 (コピュラ) は基本的に動詞と同様に扱う。4) 複合語は単語毎に分ける。助数詞 (例: ~ 枚、~ 冊) は一形態素として数える等である。

## 4. 結果と考察

### 4.1 発話数の比較

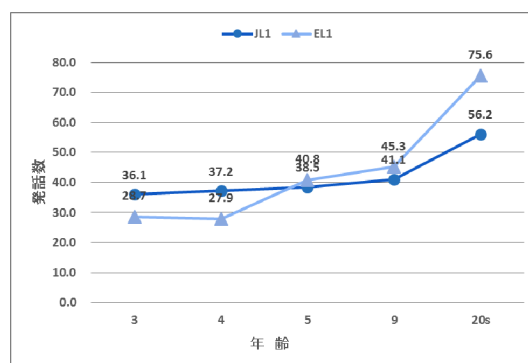
最初に、物語全体で用いられている発話数の年齢による推移を、日本語を母語とする子どもと大人 (JL1) と英語を母語とする子どもと大人 (EL1) で比較する。

【図1】は、JL1 と EL1 の 3、4、5、9 歳児、及び、大人 (グラフ中では「20s」と表記する) が 1 つの物語文全体で用いた発話数の平均値を年齢毎に表したものである。JL1 の推移は、3 歳児で 36.1、9 歳児で 41.1 で、3 歳児の 1.1 倍とあまり大きな増加はない。

一方、EL1 の推移を見ると、3 歳児で 28.7、4 歳児で 27.9 と日本語よりやや値は低いが、5 歳児で 40.8 となり、JL1 を幾分上回る。9 歳児では、45.3 で 3 歳児の 1.6 倍に増加する。

よって、発話数の変化は、JL1 では 3~9 歳間であまり大きく増加しないが、EL1 では、4~9 歳にかけて JL1 よりも増加が大きい。ただし、9 歳児の発話数は両者で近い値となっている。

大人では、JL1 は 56.2 であるが、EL1 は 75.6 と JL1 の 1.4 倍であり、この発話資料では、大人の発話数は、EL1 の方が JL1 よりも高い値を示している。

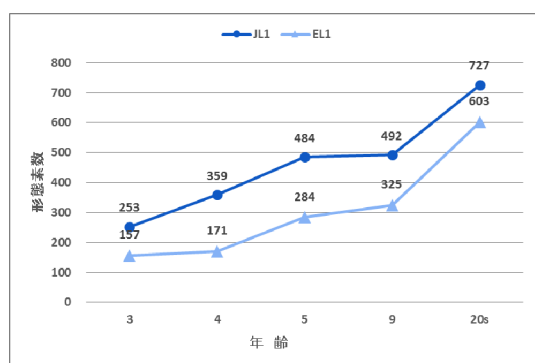


【図1】 JL1 と EL1 の発話数の推移

## 4.2 形態素数の比較

次に、形態素数の推移を比較する。【図2】は、JL1とEL1の形態素数の推移の比較である。JL1の形態素数は、3歳児では253、4歳児では359、5歳児で484で、年齢と共に増加している。9歳児では、492で、5歳から9歳にかけては横這いである。大人では、727と形態素数は9歳児の約1.5倍になる。すなわち、9歳頃までに大人の70%前後の形態素数に到達していることが分かる。

EL1の形態素数は、3歳児では157、4歳児では171、5歳児で284で、年齢と共に増加している。形態素数は全体にJL1よりも少ない。9歳児では、325で、5歳から9歳にかけての増加はJL1同様に横這いである。大人では、603と形態素数は9歳の約2倍になる。言い換えれば、9歳頃までに50%前後の形態素数になる。年齢と共に形態素数が純増する点はJL1の場合と類似している。よって、形態素数はJL1の方が全体に高いが、増加の過程はEL1の場合もJL1の場合も概ね類似した傾向にあると言える。



【図2】 JL1とEL1の形態素数の推移

## 4.3 平均発話長の比較

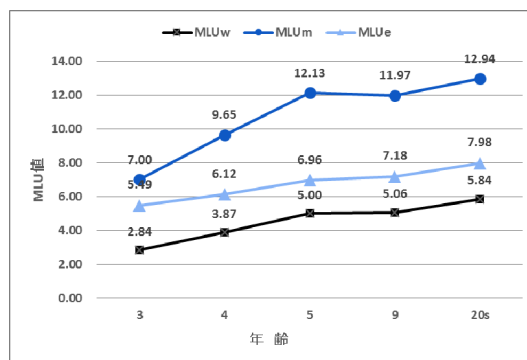
ここでは、JL1の平均発話長（MLUm・MLUw）をEL1と比較する。【図3】は、JL1のMLUmとMLUwとEL1のMLUeの推移を比較したものである。これを見ると、MLUmが全体に一番高い値を示している。また、MLUwは、一番低い値となっている。

まず、MLUmの推移を見る。3歳児では7.00、4歳児では9.65、5歳児では12.13と3～5歳の間で急激に増加している。3～4歳と4～5歳での区間の平均値の差の検定の結果、3～4歳間では $t=9.6476$  ( $p<.01$ )、4～5歳間では、 $t=8.5346$  ( $p<.01$ ) で、両区間において有意な増加が確認された（稲葉, 2017a）。よって、3歳から5歳にかけて文法発達が進み、発話の文構造の複雑さが著しく増していくと言える。

次に、MLUwの推移を見る。3歳児では2.84、4歳児では3.87、5歳では5.00と、この年齢の区間で次第に増加している。3～4歳と4～5歳の区間での平均値の差の検定の結果、3～4歳間では $t=12.2904$  ( $p<.01$ )、4～5歳

間では、 $t=11.8889$  ( $p<.01$ ) で、両区間でMLUw値の有意な増加が確認された（稲葉, 2017a）。

MLUeの推移は、MLUeは、MLUmとMLUwの間に位置している。3歳児で5.49、4歳児で6.12、5歳児で6.96と純増している。増加の割合は、MLUwとほぼ平行している。よって、JL1、EL1において、年齢と共に平均発話長の値が増加し、特に3～5歳の区間での増加が大きい点は類似の傾向である。



【図3】 MLUm・MLUw・MLUeの推移

## 4.4 到達度によるMLUの分析

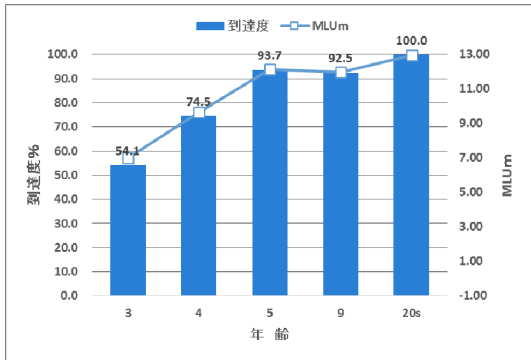
4.3では、日本語の平均発話長をMLUm、MLUwの両手法で算出したが、これらの両値は同じ発話を分析したにも関わらず、数値に大きな開きが出た。よって、分析方法によってかなり数値に差が出ることが分かった。同様に、MLUを通言語間で比較する場合、言語の構造的な特徴からその数値に幅が生じる可能性がある。よって、MLUの測定値をそのまま用いて発達の度合いを比較するのは適切でないと思われる。

例えば、本研究の言語資料では、JL1のMLUmの値は、EL1のMLUeを大きく上回っているから、日本語の方が発達が進んでいるとは言い難い。逆に、MLUw値は、MLUe値より低いことから、発達が遅れているというふうには判断するのは誤りである。JL1のMLUm値は、大人で12.94と非常に高いことから、これを到達点と考えると、数値だけで英語よりも発達が進んでいるとは言えない。反対に、大人のMLUw値は低いので、数値は低くても、発達は進んでいる可能性がある。したがって、数値の変化の過程を見ることはできるが、これらは絶対的な数値での比較には適さない。これは宮田（2012）の過評価や過少評価が起こる可能性があるという指摘を支持するものである。

そこで、本研究では、JL1とEL1の発達を大人の平均値を目標値（Target：TG値）とし、それにどれだけ近づいているかで発達の度合い（「到達度」と定義する）を比較することにする。以下、JL1のTG値をTGJ、EL1のTG値をTGEと表記する。

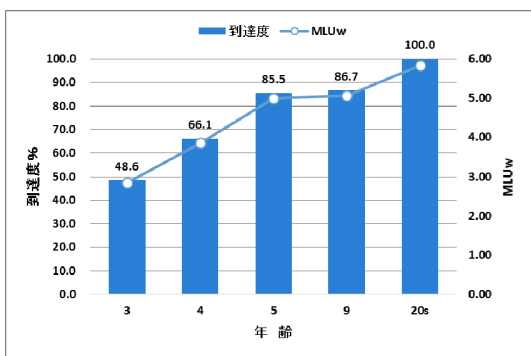
【図4】は、3～9歳児のMLUmをTGJ値を100%とした場合の割合で換算して示している。3歳児のMLUm

は7.00なので、TGJ12.94を100%と考えた場合、その到達度は54.1%となる。MLUmの到達度は、3歳児で54.1%、4歳児で74.5%、5歳児で93.7%となっており、5歳児では90%を超える。よって、発達を到達度の推移で捉えると、3~5歳で急激に上昇し、5歳で大人に近い値に到達することが分かる。



【図4】 MLUmの到達度の推移

次にMLUwの到達度の推移を見る。【図5】は、MLUwの3~11歳児の到達度を表している。MLUmと同様に、3歳児のMLUwは2.84、TGJは、5.84なので、その到達度は48.6%となる。MLUwの推移は、3歳児で48.6%、4歳児で66.1%、5歳児で85.5%となっており、MLUwでも5歳で達成度は8割を上回っており、MLUmと類似の傾向を示している。ここで、着目すべき点は、【図3】に示したMLUmとMLUwの値よりも開きが少ないことである。以上から、発達をMLUm、MLUwを指標として捉えた場合、5歳頃までにTGJの8~9割に到達すると言える。

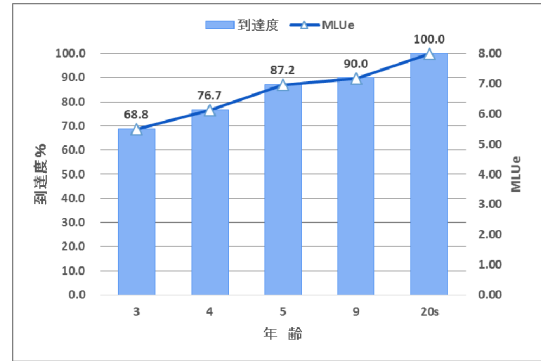


【図5】 MLUwの到達度の推移

次にMLUeの到達度の推移を見る。【図6】は、EL1の3、4、5、9歳児のMLUeの到達度の推移を表している。日本語の物語文と同様に、大人値(TGE)を100%として各年齢の値の割合を算出したものである。3歳児のMLUeは5.49で、到達度は68.8%、4歳児のMLUeは6.12で、到達度は76.7%、5歳児のMLUeは6.96で、到達度は87.2%である。よって、到達度は5歳までに

概ね9割に達する。

以上から、英語においてもMLUeは概ね5歳ごろまでに増加し、9歳頃までにはTGEの9割程度の達することが分かる。日本語の発達も5歳頃までにほぼTGJの8~9割程度に到達する言う点では、傾向が類似していると言える。



【図6】 MLUeの到達度の推移

## 5. ナラティブ萌芽期の個人差の分析

### 5.1 発話数の変化

ここでは、発達が著しいことが分かった発達の3~5歳児の個人の発話数、形態素数、平均発話長を見てみる。そして、これまでの各年齢の平均値だけでは分からない発達の特徴を探る。

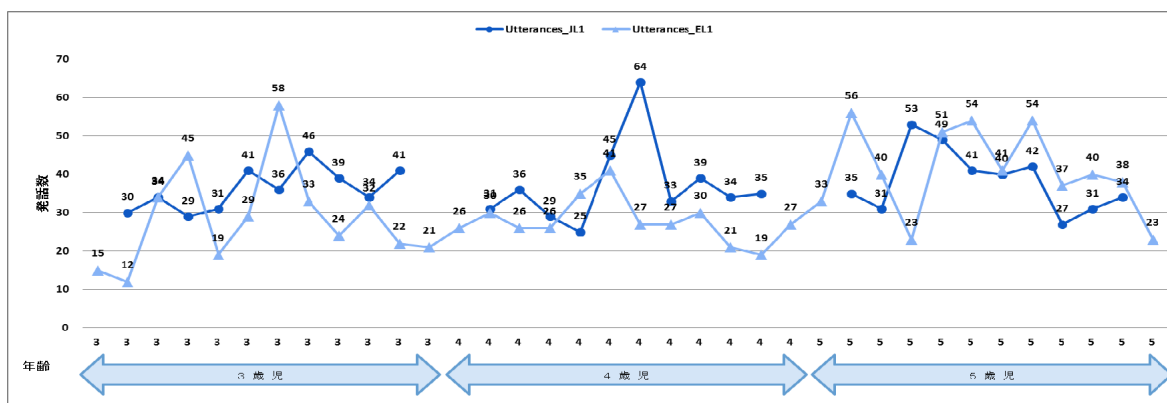
まず、各年齢でどのぐらい個人差があるのかを見てみる。【図7】は、JL1とEL1の3~5歳児の発話数を月齢の低い方から並べグラフに表したものである。EL1の被験者は各年齢12人であるが、JL1は各年齢10人であるので、グラフでは便宜上、JL1の各年齢の最初と最後の箇所を空欄として表示してある。

JL1の発話数は、3歳児では最小値29、最大値は46、レンジは17で、あまり増加は見られない。標準偏差(STD)【表3】は5.55で、個人による差はそれほど大きくない。4歳児では、最小値25、最大値は64、レンジは39で、月齢が上がってもあまり増加は見られない。標準偏差は10.99で、かなり個人差が現れてくる。5歳児では、最小値27、最大値は53、レンジは26で、同じく月齢が上がってもあまり増加は見られない。標準偏差は8.10で、個人差は4歳児よりは低い。

以上から、JL1の発話数は、3歳から5歳まであまり増加しないこと、3歳児では個人差は小さいが、4歳以降で顕著に現れることが分かった。しかし、3~4歳児、4~5歳児の発話数の分散分析(【表4】)の結果では、両区間のバラツキは有意には至らなかった。

EL1の発話数は、3歳児では最小値12、最大値は58、レンジは46、標準偏差【表5】13.03で、個人による開きが大きい。4歳児では、最小値19、最大値は45、レンジは26、標準偏差5.81で、個人による開きは小さい。





【図7】 JL1とEL1の3～5歳児の発話数の比較

【表3】 JL1の3～5歳児の発話数の標準偏差STD

年齢	3歳	4歳	5歳
発話数_STD	5.55	10.99	8.10

【表4】 JL1の3～5歳児の発話数の分散分析

区間	F値	自由度1	自由度2	P値	有意差
3-4歳	0.0798	1	13.3038	0.7819	-
4-5歳	0.0906	1	16.5479	0.7671	-

Welchの検定 [\* : P<0.05 \*\* : P<0.01]

【表5】 EL1の3～5歳児の発話数の標準偏差STD

年齢	3歳	4歳	5歳
発話数_STD	13.03	5.81	11.29

【表6】 EL1の3～5歳児の発話数の分散分析

区間	F値	自由度1	自由度2	P値	有意差
3-4歳	0.0332	1	15.2013	0.8579	-
4-5歳	12.4242	1	16.4408	0.0027	-

Welchの検定 [\* : P<0.05 \*\* : P<0.0]

5歳児では、最小値23、最大値56、レンジは33、標準偏差11.29で再び個人差が現れる。

以上から、EL1においては、発話数は、3歳と4歳では平均発話数にはあまり差が見られないが、3歳児では個人差が大きく、4歳児では、個人差が少なく、安定した状態であると言える。5歳児になると、平均発話数はそれ以前より増加し、再び個人差が現れることが分かった。しかし、の3～4歳児、4～5歳児の発話数の分散分析（【表6】）の結果では、両区間のバラツキは有意には至らなかった。

## 5.2 形態素数の変化

ここでは、JL1とEL1の3～5歳児ひとりひとりの形態素数を見てみる。【図8】3～5歳児の個人の形態素数を月齢の低い方から並べグラフに表したものである。

最初にJL1の形態素数を見る。3歳児では最小値115、最大値は430、レンジは315、標準偏差【表7】は116.09である。このデータでは、最大値の430以前のバラツキは少ないが、これ以降急激に形態素数が増加している。4歳児では、最小値199、最大値は793、レンジは594、標準偏差172.50で、全体に形態素数が増加し、中には突出して高い値もある。バラツキもさらに大きくなる。5歳児では、最小値242、最大値は636、レンジは394、標準偏差122.57で、形態素数はさらに増加するが、バラツキは4歳児よりは小さくなっている。

以上から、この物語文では、形態素数は3歳後半から増加し始め、4歳では乱高下を繰り返す、5歳にかけてさらに増加していくことが分かった。標準偏差は4歳児の値が一番大きい。しかし、の3～4歳児、4～5歳児の形態素数の分散分析（【表8】）では、両区間のバラツキは有意には至らなかった。

EL1の形態素数を見る。3歳児では最小値60、最大値は410、レンジは294、標準偏差【表9】は105.73である。このデータでは、最大値の410以前のバラツキが大きい、これ以降は安定している。4歳児では、最小値114、最大値は245、レンジは131、標準偏差36.80で、全体にバラツキは少ない。5歳児では、最小値138、最大値は399、レンジは261、標準偏差87.36で、形態素数は増加するが、バラツキは4歳児よりは大きい。

以上から、EL1のこの物語文では、発達初期の3歳

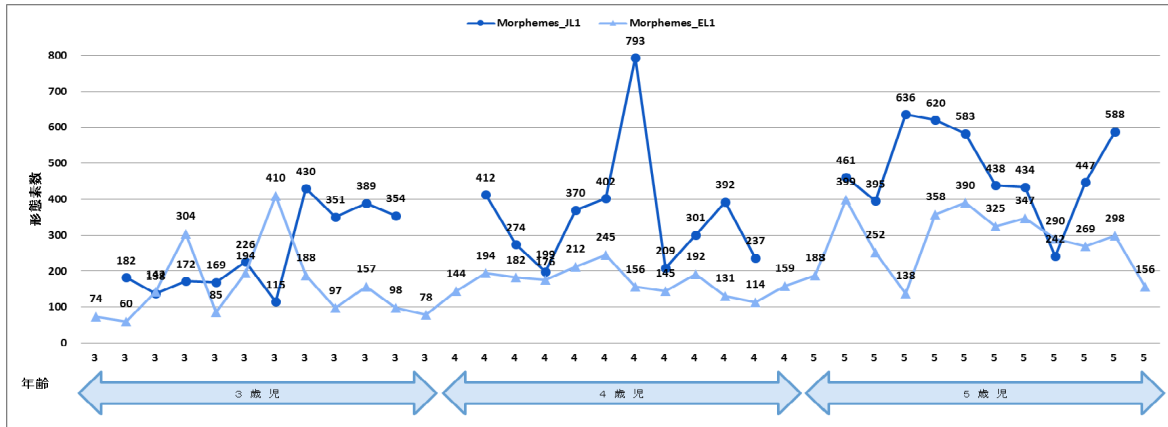
【表7】 JL1の3～5歳児の形態素数の標準偏差STD

年齢	3歳	4歳	5歳
形態素数_STD	116.09	172.50	122.57

【表8】 JL1の3～5歳児の形態素数の分散分析

区間	F値	自由度1	自由度2	P値	有意差
3-4歳	2.6138	1	15.7646	0.1258	-
4-5歳	3.5173	1	16.2422	0.0788	-

Welchの検定 [\* : P<0.05 \*\* : P<0.01]



【図8】 JL1とEL1の3～5歳児の形態素数の比較

【表9】 EL1の3～5歳児の形態素数の標準偏差 STD

年齢	3歳	4歳	5歳
形態素数_STD	105.73	36.80	87.36

【表10】 EL1の3～5歳児の形態素数の分散分析

区間	F値	自由度1	自由度2	P値	有意差
3-4歳	0.1766	1	13.6262	0.6808	-
4-5歳	17.1468	1	14.7830	P<0.001	**

Welchの検定 [\* : P<0.05 \*\* : P<0.01]

前半ではバラツキが見られ、その後、4歳にかけて安定する。形態素数は5歳はじめ頃から増加し始め、再びバラツキが見られるようになることが分かった。

標準偏差は5歳児の値が一番大きい。3～4歳児、4～5歳児の形態素数の分散分析（【表10】）では、4～5歳児のバラツキに有意差が見られた。

### 5.3 平均発話長の変化

ここでは、JL1のMLUwとMLUm、及び、EL1のMLUeのひとりひとりの値を見る。【図9】は、3～5歳児のMLUwとMLUm、及び、EL1のMLUeを月齢の低い方から並べグラフに表したものである。

まず、JL1の3～5歳児までのMLUwとMLUmの推移を比較すると、MLUmの方が全体に増減が大きく、MLUwよりも変化を敏感に捉えていると言える。

MLUmは、3歳児では最小値3.19、最大値は11.44、レンジは8.25、標準偏差【表11】は2.61である。3歳児の初期の発達ではバラツキはあまり大きくないが、発達の後期で急激に上昇する。

4歳児では、最小値6.33、最大値は14.80、レンジは8.47、標準偏差は3.10で、MLUm比較的高い値域で乱高下を繰り返し、バラツキ（個人差）が非常に大きい。5歳児では、最小値8.64、最大値は17.29、レンジは8.65、標準偏差は2.40である。5歳児の前半では、比較的高

い値で安定しているが、後半では、最小値までに一旦下がり、その後、上昇し最大値に至る。

一方、MLUwを見ると、3歳児では最小値1.58、最大値は5.18、レンジは3.6、標準偏差は1.04で、バラツキは小さい。4歳児では、最小値2.58、最大値は6.28、レンジは3.7、標準偏差は1.27で、3歳児よりは大きなバラツキが見られる。5歳児では、最小値3.07、最大値は6.18、レンジは3.18、標準偏差は0.88で、前半では、比較的高い値で安定しているが、後半では、一旦下降し、その後、上昇する。

MLUwとMLUmの標準偏差を比較すると、MLUmでは、3歳児で2.61、4歳児で3.10、5歳児で2.40で、4歳児が一番高い値となっている。MLUwについても、全体に数値はMLUwよりは小さいが、4歳児の標準偏差が1.27で一番高い。3～4歳児、4～5歳児のMLUmの分散分析（【表12】）では、両区間のバラツキには有意差が見られた。

次に、MLUeを見ると、3歳児では最小値7.07、最大値は7.93、レンジは0.86、標準偏差【表13】105.73で、多少のバラツキが見られる。4歳児では、最小値5.54、最大値7.00、レンジ1.56、標準偏差36.80で、3歳児同様バラツキは小さい。5歳児では、最小値5.70、最大値7.93、レンジ2.23、標準偏差87.38で、バラツキはあまり見られない。以上から、MLUeでは、発達初期の

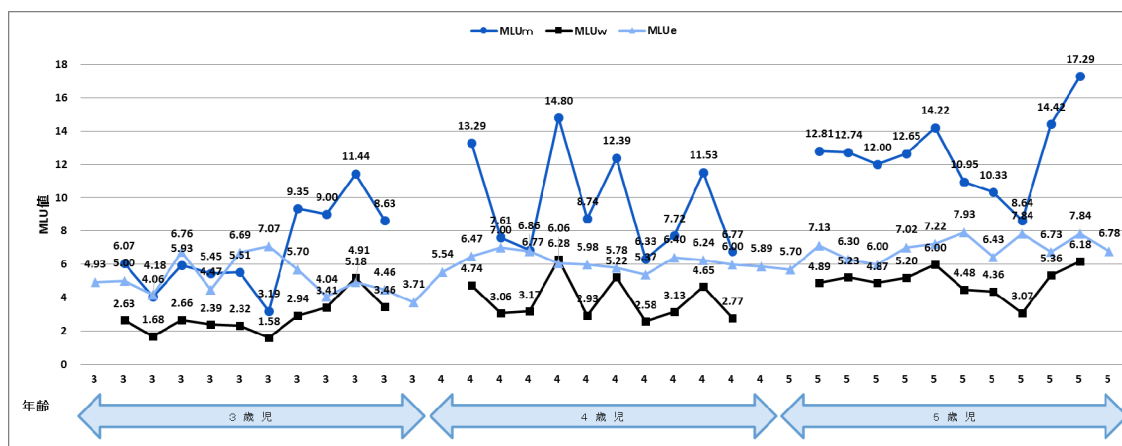
【表11】 JL1の3～5歳児の平均発話長の標準偏差 STD

年齢	3歳	4歳	5歳
MLUm_STD	2.61	3.10	2.40
MLUw_STD	1.04	1.27	0.88

【表12】 JL1の3～5歳児のMLUの分散分析

区間	F値	自由度1	自由度2	P値	有意差
3-4歳	4.5656	1	17.4971	0.0470	*
4-5歳	5.8550	1	16.9386	0.0271	*

Welchの検定 [\* : P<0.05 \*\* : P<0.01]



【図9】 JL1とEL1の3～5歳児のMLU値の比較

【表13】 EL1の3～5歳児の平均発話長の標準偏差 STD

年齢	3歳	4歳	5歳
形態素数_STD	105.73	36.80	87.38

【表14】 EL1の3～5歳児の平均発話長の分散分析

区間	F値	自由度1	自由度2	P値	有意差
3-4歳	7.3225	1	14.7736	0.0164	-
4-5歳	9.6957	1	18.9668	0.0057	-

Welchの検定 [\* : P<0.05 \*\* : P<0.01]

3歳ですこしバラツキが見られるが、その後は、あまりバラツキは見られなかった。3～4歳児、4～5歳児の発話数の分散分析（【表14】）の結果では、両区間のバラツキは有意には至らなかった。

これらの共通点は、発達の過程で、低い値から高い値に変化する前には、バラツキが大きくなるという点である。例えば、JL1の場合は、3歳児から5歳児間でみれば、3歳前半は比較的個人差が少ないが、後半から4歳にかけてバラツキが大きくなり、その過程を経て、5歳では高い値になる。5歳後半ではまたバラツキが見られ、さらに高い値に向かっていく。EL1においては、3歳児でバラツキが見られた後、5歳では比較的安定した状態になる。5歳では再びバラツキが現れた。その後は大人の値に向かって上昇していくのではないかと予測される。以上から、バラツキと安定を繰り返しながら、発達していくのではないかと仮説を立てた。この仮説の検証には、本研究のデータ以外の更なる分析が必要である

## 6. まとめと課題

本研究では、JL1、EL1の発話数、形態素数、平均発話長MLUm、MLUw、MLUeの分析結果から、以下のようなことが明らかになった。

発話数は、JL1では3～9歳間であまり増加しないが、EL1では4～9歳にかけてJL1よりも大きな増加が見られた。9歳児の発話数は両者で近い値を示した。24場面から成るこの物語文では、メインラインを語る発話数自体はそれほど変化しないのではないかと考えられる。大人の発話は、JL1もEL1も9歳児よりはるかに高い値を示したが、これは、前景や背景、状況説明などより多くの要素を物語に加えるからではないだろうか。これは、言語の違いを超えて物語文に共通する特徴である。

形態素数は、JL1の方がEL1よりも全体に高い値を示した。これは、日本語の方が英語よりも言語的な特徴からより多くの形態素を発話に含む構造であるからではないだろうか。しかし、年齢と共に純増する発達傾向は両言語で共通していた。この点を明らかにするには、さらに発話資料を精査する必要がある。

平均発話長は、JL1 (MLUm, MLUw) においても、EL1 (MLUe) においても、値の隔たりはあるが、年齢と共に平均発話長の値は増加すること、特に3～5歳の区間での増加が大きいことは類似した傾向である。

本研究では、日本語の平均発話長をMLUm、MLUwの2つの手法で算出したが、これらの両値は同じ発話文を分析したにも関わらず、数値に大きな開きが出た。MLUmは、MLUwよりかなり高い値となり、分析方法によってかなり数値に差が出ることが分かった。したがって、MLUを通言語間で比較する場合、言語の構造的な特徴からその数値に幅が生じる可能性がある。これらは、数値の変化の過程から、発達を読み取ることはできるが、絶対的な数値での比較には適さないとと言える。これは、宮田 (2012) で指摘されている分析の方法により、過評価や過少評価が起こる可能性があるという指摘を支持するものである。

そこで、本研究では、MLU (MLUm, MLUw, MLUe) の測定値をそのまま用いて発達の度合いを比較する代わりに、JL1とEL1それぞれの言語の大人の平均値を



目標値 (Target : TG 値) とし、それにどれだけ近づいているかで発達の度合い (「到達度」と定義) を比較した。その結果、JL1においても、EL1においても、3~5歳にかけて純増し、5歳までにほぼTG 値のほぼ9割に到達する点で、類似の傾向が見られた。

個人差の分析からは、JL1においても、EL1においても、バラツキと安定を繰り返しながら、次第に高い値になる様子が見られた。特に、低い値から高い値に変化する前には、大きなバラツキが見られた。JL1の場合で言えば、3歳前半は比較的個人差が少ないが、後半から4歳にかけてバラツキが大きくなり、その過程を経て、5歳では高い値になる。5歳後半ではまたバラツキが見られ、さらに高い値に向かっていく。EL1においては、3歳児でバラツキが見られた後、5歳では比較的安定した状態になる。5歳では再びバラツキが現れた。この後、大人の値に向かって上昇していくのではないかと予測される。言い換えれば、バラツキと安定を繰り返しながら、発達していくのではないかという仮説を得た。この仮説が成立するかどうかは、3歳以降の個人差や発達の過程を詳しく調べる必要がある。また、バラツキと安定を繰り返しながら、発達していくことが一般的な過程なのかどうかの検証には、本研究のデータ以外の更なる分析が必要である。

物語文においては、表現方法、レトリック、語りのスタイル等の多様性が考えられ、成熟した大人の物語文では、物語の長さ (発話量) 1つをとってもかなりの個人差が見られる。物語文の発達を考えると、子どもは様々な過程を経て、最終的には、大人の個人差の分布に近づいていくのではないかという仮説を筆者はもっている。ただし、その個人差は、言語発達の段階によって異なるのではないか。言語発達の初期においては、発話数 (発話量) や形態素数の差異、発達の中期以降においては、談話能力の発達と相まって変化するのではないだろうか。さらに、個別言語の特性によっても異なるのではないかと考えている。これらの点の解明は今後の課題である。

## 注

<sup>1</sup> <http://childes.psy.cmu.edu/access/Frogs/>

## 参考文献

- Berman R. & Slobin, D. I. (1994). *Relating events in narrative: A cross-linguistic developmental study*. Hillsdale, NJ: LEA Publishers.
- Brown, R. (1973). *A First Language*. Cambridge, Mass.: Harvard University Press.
- Inaba, M. (1999). Development of global structure in first-language narratives. *Bulletin of Aichi University of Education (Humanities and Social Sciences)*, 48, 103-112.
- Inaba, M. (2000). Narrative discourse processing in Second-Language

*Japanese. Bulletin of Aichi University of Education (Humanities and Social Sciences)*, 49, 127-136.

Inaba, M. (2001). Backtracking and reorganization in narrative. *Bulletin of Aichi University of Education (Humanities and Social Sciences)*, 50, 29-39.

Inaba, M. (2003). Anchor tense in Japanese Narrative. *Bulletin of Aichi University of Education (Humanities and Social Sciences)*, 52, 51-61.

稲葉みどり (2017a). 「日本語の物語文における言語知識の発達過程の考察—発話数・単語数・形態素数・平均発話長の解析—」『教科開発学論集』5, 23-32.

稲葉みどり (2017b). 「第二言語と第一言語の物語文における MLU (平均発話長) の変化」『愛知教育大学研究報告—人文科学編』66, 1-7.

MacWhinney, B. (2000). *The CHILDES Project: Tools for Analyzing Talk. Third Edition*. Mahwah, N.J.: Lawrence Erlbaum Associates.

Mayer, M. (1969). *Frog, where are you?* New York: Dial Press.

宮田 Susanne (2012). 日本語 MLU (平均発話長) のガイドライン: 自立語 MLU および形態素 MLU の計算法『健康医療科学』2, 1-15. <http://aska-r.aasa.ac.jp/dspace/bitstream/10638/5113/1/0039-002-201203-1-17.pdf> アクセス日: 2013年12月1日.

宮田 Susanne, 村木恭子, 森川尋美 (編) (2004). 『今日から使える発話データベース CHILDES 入門』ひつじ書房.

Miyata, S. & Naka, N. (2010). *JMOR05.1: The Japanese Morphological Analysis Program based on CLAN*. Retrieved December 1, 2010, from <http://childes.psy.cmu.edu/morgrams/Japanese>.

(2017年8月4日受理)