

令和 4 年 6 月 14 日現在

機関番号：33921

研究種目：若手研究

研究期間：2018～2021

課題番号：18K12384

研究課題名（和文）平均発話長MLUによる韓国語の初期文法発達と韓国語の幼児言語コーパスの形態素分析

研究課題名（英文）The Early Acquisition of Korean Morphology based on MLU and The Korean Morphological analysis

研究代表者

柳 朱燕 (RYU, JU-YEON)

愛知淑徳大学・教育部門・センター・講師

研究者番号：40647682

交付決定額（研究期間全体）：（直接経費） 2,500,000円

研究成果の概要（和文）：本研究では、CHILDES(Child Language Data Exchange System)で公開されているRyuコーパスを利用し、MLU（平均発話長）を基にして韓国語の基礎文法を中心に当たる格助詞の獲得順序を明らかにした。韓国語を母語とする子どもたちは主題「nun/un」、独立助詞（追加）「do」、主格「ka/i」、目的格「lul/ul」、所有格「ui」の習得過程を見せており、習得過程の基準としてのMLUの適合性を検証した。また、Ryuコーパスにおいて、形態素分析を行い、大量の自然発話データを効率的に分析できる言語処理方法を探ってみた。

研究成果の学術的意義や社会的意義

Slobin (2014) によると、韓国語の第一言語習得に関する論文は1974年から2013年までに10本が出ており、日本語の23本と、中国語のMandarin 27本とCantonese 11本に比べたら、大変遅れていることが分かる。特に、今まで行われた韓国語の研究は発話データがコーパス化されていないため、MLU(平均発話長)を基準にして幼児の言語発達を分析する研究は皆無であった。本研究では、研究代表者が構築したRyuコーパスを利用し、MLUを基にして韓国語の助詞の獲得順序を明らかにした。また、大量の自然発話データの分析が可能となるようにRyuコーパスに形態素のタグ付けの作業を行った。

研究成果の概要（英文）：In this study, using the Ryu corpus published in CHILDES (Child Language Data Exchange System), we clarified the order of acquisition of case particles, which is the center of the basic grammar of Korean, based on MLU (average speech length). The results indicate that in the early stages of development, the Korean children began to produce and productively use from nun/un (topic) to do (delimiting) to ka/i (nominative) to lul/ul (accusative) and finally, to ui (possessive), suggesting that the MLU can be the measure of the acquisition of L1 Korean. We also performed morphological analysis on the Ryu corpus to find a language processing method that can efficiently analyze a large amount of naturally spoken data.

研究分野：第一・第二言語習得研究

キーワード：第一言語習得研究 韓国語の幼児言語獲得過程 言語習得資料のコーパス構築 MLU（平均発話長） CHILDES

科研費による研究は、研究者の自覚と責任において実施するものです。そのため、研究の実施や研究成果の公表等については、国の要請等に基づくものではなく、その研究成果に関する見解や責任は、研究者個人に帰属します。

1. 研究開始当初の背景

(1) 第一言語習得過程では、文法の発達が進むと、文構造が複雑になり、子どもが様々な活用語尾や助詞を使うようになるため、発達が文の長さには反映されると言われている。そのため文法発達を平均発話長 (MLU: Mean length of utterances, 1 発話当たりの平均単語数) でとらえることができる。さらに、言語発達の速度に大きな個人差が見られるが、子どもを MLU で合わせると、類似した発達順序が観察される。しかし、韓国語の第一言語習得研究においては、Brown (1973) や小椋(1999)の研究のように発達初期段階で特定の文法項目の獲得を明らかにした研究がまだされていない。

Journal of Child Language のエディターを務めている Slobin (2014) によると、韓国語の第一言語習得に関する論文は 1974 年から 2013 年までに 10 本が出ており、日本語の 23 本と、中国語の Mandarin 27 本と Cantonese 11 本に比べたら、大変遅れていることが分かる。特に、今まで行われた研究は心理学的な実験によるものが多く、縦断的な発話データを使った研究においても、そのデータがコーパス化されていないため、MLU を基準にして幼児の言語発達を分析する研究は皆無であった。MLU は言語発達の尺度として用いられるが、1 発話当たりの平均単語数で計算するため、縦断的な発話資料をコーパス化しないと、簡単に計算することができない。本研究で使用する Ryu コーパスは、研究代表者が CHILDES(Child Language Data Exchange System) にアップロードしたコーパスであり、CHILDES で提供する言語分析ツール CLAN (Computerized Language Analysis) を適用すると、大量の自然発話データの分析が可能となり、MLU を基準にして言語発達過程を客観的に見ることが出来る。

既存の言語学の理論は主にインド・ヨーロッパ言語(SVO 言語)に基づいて検証されてきているが、日本語又は韓国語(SOV 言語)を対象として検証した研究では、これまでの言語学理論とは異なる結果が得られた研究が少なくない(Choi & Gopnik 1995, Shirai 1993, 大関 2008 など)。そこで、日本語だけではなく韓国語を対象に検証を行うことによって、双方の言語で得られた結果の妥当性を高めることができ、全般的な言語習得理論において「普遍性と個別性」を追究するためにも学術的に意義があると考えられる。また、MLU を基にした韓国語の文法習得過程を明らかにする本研究の成果を、言語発達の尺度として用いることにより、韓国語の第一言語習得研究分野に大きな示唆ができることを狙っている。

(2) Ryu コーパスが韓国語の幼児言語のデータとして公開されているものの、以下の図 1 の日本語コーパスのような形態素のタグ付けと文種類や状況説明などのコメントがまだ完成されていない。コーパスとして大量の自然発話データを効率よく分析するためには、形態素のタグ付けは必要不可欠である。本研究では、形態素のタグ付けとコメントを付ける作業に着手し、2~3 年に渡りコーパス化を進めていく。その際、MLU (平均発話長) の計算が可能となるように、形態素 MLU などといった一定の基準を作って既存のデータを再整理する。

```

895 *CHI: eki . ← 子どもの発話をローマ字で書いたもの
896 %xmor: n|eki=station . ← 形態素分析(タグ付け)
897 *AMO: doko ga ii ? ← お母さんの発話
898 %xmor: n:deic:wh|doko=where ptl:case|ga=NOM adj|i-PRES=good ?
899 %cod: $Q ← 文種類(疑問、否定)の表示 ← 形態素分析
900 *CHI: koko &tja:n .
901 %xmor: n:deic:dem|koko=here .
902 %gpx: pointing ← 状況説明コメント
    
```

<CHILDES に公開されている EastAsian/Japanese/Miyata/Aki24.CAT より>

図 1 日本語の幼児言語コーパスの例

2. 研究の目的

(1) 第一言語習得過程では、文法の発達が進むと、文構造が複雑になり、子どもが様々な活用語尾や助詞を使うようになるため、発達が文の長さには反映されると言われている。そのため文法発達を平均発話長 (MLU: Mean length of utterances, 1 発話当たりの平均単語数) でとらえることができる。しかし、韓国語の第一言語習得研究では、MLU を基準にして幼児の言語発達を分析する研究がまだ行われていない。本研究では、申請者が構築した Ryu コーパスを利用し、MLU を基にして韓国語の基礎文法の中心に当たる助詞及び動詞の活用形態素の獲得順序を明らかにすることを目的とする。

(2) CHILDES(Child Language Data Exchange System)で公開されている Ryu コーパスにおいて、形態素分析を行い、大量の自然発話データを効率的に分析できるようにシステムを構築することも本研究の目的である。そのため、形態素分析によるタグ付け又は、コーパス分析ツールの使用において、既に完成している日本語の L1 習得コーパスを見習い、研究を進めていく。本研究によって、韓国語の L1 習得コーパスが大量の自然発話データを効率的に分析できるようになり、韓国語の第一言語習得研究分野に大きな進歩をもたらすことを期待している。

3. 研究の方法

(1) 韓国語の第一言語習得における格助詞の習得過程

本研究で分析する格助詞を表1に示した。韓国語の場合、格助詞を付ける際、直前の単語が母音で終わるか子音で終わるかによって異なる助詞を使う。例えば、主格「ガ」に当たる韓国語の助詞は ka と i があり、直前の単語が母音で終わると ka を、直前の単語が子音で終わると i を使う。このようにペアーの格助詞が主格「ka/i」、主題「nun/un」、目的格「lul/ul」の三つがある。本研究ではその三つの格助詞に加え、所有格「ui」と独立助詞「do」を入れ、五つの助詞の習得過程を分析した。

表1 本研究で分析する格助詞

助詞	主格「ka/i」(母音+kaㄱ、子音+iㅇ)
	主題「nun/un」(母音+nunㅇ、子音+unㅇ)
	目的格「lul/ul」(母音+lulㅇ、子音+ulㅇ)
	所有格「ui」(ㅇ)
	独立助詞(追加)「do」(ㅇ)

本研究では、CHILDES(Child Language Data Exchange System)で公開されている Ryu コーパスを利用し、格助詞の使用過程を分析した。Ryu コーパスは、韓国語を母語として習得している3人の子ども(Jong, Joo, Yun)から、1歳半から3歳半までの2年をかけて養育者との自然発話を収録し、文字化したデータベースである(<https://chilides.talkbank.org/access/EastAsian/Korean/Ryu.html>)。データは養育者と子どもが絵本を読みながら対話する場面とし、1か月1時間のペースでビデオカメラで収録した。言語データは全部で81時間15分の自然発話であり、その中から研究の対象である五つの格助詞について、子どもが産出した3,723文を分析対象とした。

(2) 韓国語の幼児言語コーパスの形態素分析

形態素のタグ付けに関しては、韓国で広く使用されている形態素分析プログラムである UTagger (韓国蔚山大学校韓国語処理研究室 2011) で本データの処理を試みたが、形態素のタグ付けの正確率が70%にしか達していなかった。それは、言葉を学び始めた幼児のデータでは助詞の省略や統語的な情報の曖昧さが激しいため、形態素分析機が正確に判断し難いところがあるからである。また、UTagger では CHILDES での形態素分類基準と異なる基準で運営されているため、CLAN を通した形態素分析を行うと、適切な処理がなされない結果となる。特に、動詞の活用形の処理に関する分類基準の違い、同音異義語の処理の仕方、文法的に省略が多い幼児言語の処理に関する困難さなど、様々な問題が生じる。この問題点に関しては CHILDES の責任者である MacWhinney 教授(カーネギーメロン大教授)と JCHAT プロジェクトに関わっていた研究者の協力により、JCHAT で行われた形態素分析基準を韓国語へ適用することを検討した。

4. 研究成果

(1) 韓国語の第一言語習得における格助詞の習得過程

各子どもがすべての調査期間で発話した五つの格助詞の使用状況を表2でまとめた。本研究では子どもが、発話を収録した1時間の内10回以上格助詞を使い始めた時を習得初期段階として認め、それに基づいて習得過程を表3に示した。表2では、網掛けで習得初期段階を表した。表2と表3の結果から、韓国語格助詞の習得過程に関しては、次の2点があった。

表2 各子どもの格助詞の使用

JONG	Age																				SUM				
	1;7	1;8	1;9	1;10	1;11	2;0	2;1	2;2	2;3	2;4	2;5	2;6	2;7	2;8	2;9	2;10	2;11	3;0	3;1	3;2	3;3	3;4	3;5	token	(%)
主格ka/i	1	6	24	9	24	17	44	31	28	40	23	23	27	38	27	50	34	53	25	18	24	26	79	671	46
目的格lul/ul	0	0	3	0	5	2	12	2	1	2	4	9	1	5	1	8	9	8	11	10	11	5	4	113	8
主題nun/un	29	4	34	1	16	4	14	15	27	24	10	19	47	18	21	13	13	17	19	20	16	11	40	432	30
所有格ui	0	0	2	0	0	0	0	0	0	1	0	1	7	1	0	0	0	2	3	0	1	0	3	21	1
追加do	1	12	11	11	16	34	12	6	7	14	4	8	4	5	15	9	9	9	12	8	5	0	2	214	15
Sum	31	22	74	21	61	57	82	54	63	81	41	60	86	67	64	80	65	89	70	56	57	42	128	1451	100
MLU	2.4	2.6	3.1	2.8	2.8	2.9	3.2	3.3	2.8	2.8	3.0	2.7	3.1	2.8	2.9	2.8	2.6	3.4	3.4	2.7	3.1	3.0	3.8		

JOO	Age															SUM								
	2;2	2;3	2;4	2;5	2;6	2;7	2;8	2;9	2;10	2;11	3;0	3;1	3;2	3;3	3;4	3;5	3;6	3;7	3;8	3;9	3;10	token	(%)	
主格ka/i	0	0	0	0	0	0	0	2	1	3	11	17	12	16	44	99	82	73	42	7	14	423	33	
目的格lul/ul	0	0	0	0	0	0	0	1	1	2	0	1	2	25	26	16	38	13	10	4	1	140	11	
主題nun/un	0	0	0	0	0	1	1	2	10	5	18	34	65	69	54	60	48	17	31	16	32	463	37	
所有格ui	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	1	0	0	0	1	0	
追加do	1	2	1	0	0	2	0	4	5	16	9	23	43	12	13	25	16	10	23	9	23	237	19	
Sum	1	2	1	0	0	3	1	9	17	26	38	75	122	122	137	200	184	114	106	36	70	1264	100	
MLU	2.9	2.4	2.5	2.5	2.3	2.1	2.2	2.2	2.2	2.4	2.4	2.7	2.9	2.8	3.4	3.7	3.7	3.5	3.5	3.2	3.0			

YUN	Age																		SUM		
	2;3	2;4	2;5	2;6	2;7	2;8	2;9	2;10	2;11	3;0	3;1	3;2	3;3	3;4	3;5	3;6	3;7	3;8	3;9	token	(%)
主格ka/i	2	2	2	2	3	6	13	22	17	40	28	12	16	17	60	37	43	72	45	439	11
目的格lul/ul	0	0	3	1	0	1	2	3	13	8	6	5	2	7	15	8	6	18	14	112	11
主題nun/un	2	3	3	9	32	2	13	34	24	12	3	7	14	27	35	27	20	35	21	323	32
所有格ui	0	0	1	0	0	0	3	2	0	3	0	0	1	0	0	5	1	0	0	16	2
追加do	2	3	3	0	5	2	17	1	1	10	2	5	3	2	20	16	4	19	3	118	12
Sum	6	8	12	12	40	11	48	62	55	73	39	29	35	54	130	88	78	145	83	1008	100
MLU	2.2	2.4	2.9	2.5	2.6	3.5	3.4	3.0	3.3	2.9	2.7	3.2	3.0	3.2	3.6	3.3	3.2	4.1	3.8		

表3 格助詞の習得過程

子ども	習得過程					
JONG	主題	追加	主格	目的格	所有格	
	nun/un	do	ka/i	lul/ul	ui	
	Age	1;7	1;8	2;1	2;7	
	MLU	2.435	2.553	3.140	3.185	3.088
JOO	主題	追加	主格	目的格	所有格	
	nun/un	do	ka/i	lul/ul	ui	
	Age	2;10	2;11	3;0	3;7	
	MLU	2.238	2.352	2.350	2.752	3.473
YUN	主題	追加	主格	目的格	所有格	
	nun/un	do	ka/i	lul/ul	ui	
	Age	2;7	2;9	2;9	2;11	3;7
	MLU	2.603	3.412	3.412	3.282	3.234

格助詞の習得過程

表3を見ると、各子どもの格助詞習得パターンが一定した過程を辿ることが分かる。子どもたちは主題「nun/un」 独立助詞(追加)「do」 主格「ka/i」 目的格「lul/ul」 所有格「ui」の習得過程を見せており、三人の習得順序は一致している。しかし、これは韓国語の先行研究の結果(主格「ka/i」 独立助詞(追加)「do」 主題「nun/un」 目的格「lul/ul」)と異なる結果である。特に主題「nun/un」が本研究では最も早い段階で習得される結果になり、その原因に関しては現在も調査中である。また、本研究の結果を日本語の助詞の習得過程と比較してみると、小椋(1999)の研究から本研究の対象になる助詞だけ選別してみた結果『「ガ」「ノ」「モ」「ハ』』とは大きな相違点が見られたが、宮田・大伴・白井(2015)の結果『「ハ」・「ノ」 「モ」・「ガ」「ヲ』』からは「ノ」を除いて一致している。

習得過程における年齢と MLU

表3の結果から、格助詞の習得過程と年齢においては個人差が大きく見られることが分かった。一番目に習得された主題「nun/un」を見ると、JONGは1;7で習得したが、YUNは2;7、JOOは2;10で習得し、言語発達が早い子と遅い子の間に1年間の差ができ、年齢は習得過程の基準として使えないことが分かった。しかし、MLUは3人の子どもの個人差は少なく、習得過程の基準として使えることが分かった。例えば、習得初期の主題「nun/un」を習得する際のMLUが2.2~2.6の分布を持ち、最後の所有格「ui」を習得する際のMLUが3.1~3.5の分布をすることが本研究で明らかになった。本研究で用いるMLUは、小椋ら(1997)による四つ分類、接辞・活用形付属語MLU、活用形付属語MLU、自立語付属語MLU、自立語MLUの中、「自立語MLU」を使用している。今後の課題としてMLU産出方法による適合性判定をする予定である。

(2) 韓国語の幼児言語コーパスの形態素分析

JCHATプロジェクトの責任者の協力を得て、JCHATのプログラムを入手し、分析することができた。コーパスの形態素分析が可能となるためには文法分析に加え、語彙辞書(lexical dictionary)が必要である。Ryuコーパスの3人の幼児が発話した9,460tokenから抽出した単語をリストアップして名詞・動詞・形容詞・助詞・副詞・オノマトペの6種類に分類し、意味を英語に直して語彙辞書(lexical dictionary)を完成した。しかし、文法分析プログラミングが未完成で、韓国語の動詞活用形の処理という課題が残っている。これから更なる研究を続けて解決していきたい。

< 引用文献 >

小椋たみ子(1999) 初期言語発達と認知発達の関係。風間書房
 小椋、中、山下、村瀬、マユ(1997) 日本語獲得児の語彙と文法の発達 : Clanプログラムによる分析、神戸大学発達科学部研究紀要4(2):31-57
 宮田・大伴・白井(2015) 初期文法発達と平均発話長(MLU)段階、日本発達心理学会第25回大会ポスター発表
 Brown, Roger. 1973. A first language. Cambridge, MA: Harvard University Press
 MacWhinney, Brian. 2000. The CHILDES Project: Tools for Analyzing Talk. Third Edition. Mahwah, NJ: Lawrence Erlbaum Associates.

5. 主な発表論文等

〔雑誌論文〕 計6件（うち査読付論文 3件/うち国際共著 2件/うちオープンアクセス 4件）

1. 著者名 Ju-Yeon RYU and Yasuhiro SHIRAI	4. 巻 First View
2. 論文標題 L1 acquisition of the tense-aspect markers -ess (past-perfective) and -ko iss (imperfective) in Korean	5. 発行年 2022年
3. 雑誌名 Journal of Child Language	6. 最初と最後の頁 1-21
掲載論文のDOI（デジタルオブジェクト識別子） 10.1017/S0305000922000113	査読の有無 有
オープンアクセス オープンアクセスではない、又はオープンアクセスが困難	国際共著 該当する

1. 著者名 RYU Ju-Yeon	4. 巻 第11号
2. 論文標題 The First Language Acquisition processes of Case Markers in Korean-Focusing on the results of an analysis of the CHILDES Corpus-	5. 発行年 2021年
3. 雑誌名 韓国語教育研究	6. 最初と最後の頁 68-86
掲載論文のDOI（デジタルオブジェクト識別子） なし	査読の有無 有
オープンアクセス オープンアクセスではない、又はオープンアクセスが困難	国際共著 -

1. 著者名 柳朱燕	4. 巻 第11号
2. 論文標題 オンライン授業に関する実施状況と学生の評価 言語活用科目としての韓国・朝鮮語授業を中心に	5. 発行年 2021年
3. 雑誌名 愛知淑徳大学論集 交流文化学部篇	6. 最初と最後の頁 41-57
掲載論文のDOI（デジタルオブジェクト識別子） なし	査読の有無 無
オープンアクセス オープンアクセスとしている（また、その予定である）	国際共著 -

1. 著者名 Ryu, Ju-Yeon	4. 巻 第10号
2. 論文標題 The Ryu Corpus: Longitudinal Speech Data from Three Korean Children Added to the CHILDES Database	5. 発行年 2020年
3. 雑誌名 愛知淑徳大学論集 交流文化学部篇	6. 最初と最後の頁 29-48
掲載論文のDOI（デジタルオブジェクト識別子） なし	査読の有無 無
オープンアクセス オープンアクセスとしている（また、その予定である）	国際共著 -

1. 著者名 柳朱燕	4. 巻 第9号
2. 論文標題 韓国・朝鮮語プレゼンテーション授業におけるCLIL教授法の応用と評価	5. 発行年 2019年
3. 雑誌名 愛知淑徳大学論集 交流文化学部篇	6. 最初と最後の頁 67-83
掲載論文のDOI (デジタルオブジェクト識別子) なし	査読の有無 無
オープンアクセス オープンアクセスとしている(また、その予定である)	国際共著 -

〔学会発表〕 計6件(うち招待講演 0件/うち国際学会 3件)

1. 発表者名 Ryu Ju-yeon
2. 発表標題 The L1 Acquisition of Case Markers in Korean using CHILDES Corpus
3. 学会等名 the 15th International Congress for the Study of Child Language (国際学会)
4. 発表年 2021年

1. 発表者名 柳朱燕
2. 発表標題 韓国語科目の受講生を対象としたオンライン授業に関するアンケート調査結果 - 対面授業との比較を中心に -
3. 学会等名 日本韓国語教育学会(Japan Korean Education Conference)
4. 発表年 2021年

1. 発表者名 柳朱燕
2. 発表標題 韓国語の格助詞における第一言語習得過程 - 日本語の習得過程との比較を中心に -
3. 学会等名 日本韓国語教育学会(Japan Korean Education Conference)
4. 発表年 2020年

1. 発表者名 柳朱燕
2. 発表標題 韓国語の第一言語習得における格助詞の習得過程 - 主格ka/i、主題nun/un、目的格lul/ul、所有格ui、独立助詞(追加)doを中心に -
3. 学会等名 日本言語科学会(Japanese Society for Language Sciences)(国際学会)
4. 発表年 2019年

1. 発表者名 Ryu Ju-yeon and Shirai Yasuhiro
2. 発表標題 L1 Acquisition of Tense-Aspect Markers -ess (past-perfective) and -ko iss (imperfective) in Korean
3. 学会等名 the 43rd annual BUCLD (Boston University Conference on Language Development)(国際学会)
4. 発表年 2018年

1. 発表者名 柳朱燕
2. 発表標題 CHILDESに公開されている韓国語幼児言語コーパスの概要と応用
3. 学会等名 日本韓国語教育学会(Japan Korean Education Conference)
4. 発表年 2018年

〔図書〕 計0件

〔産業財産権〕

〔その他〕

-

6. 研究組織	氏名 (ローマ字氏名) (研究者番号)	所属研究機関・部局・職 (機関番号)	備考
---------	---------------------------	-----------------------	----

7. 科研費を使用して開催した国際研究集会

〔国際研究集会〕 計0件

8 . 本研究に関連して実施した国際共同研究の実施状況

共同研究相手国	相手方研究機関
---------	---------