

## 科学研究費助成事業 研究成果報告書

平成 28 年 6 月 13 日現在

機関番号：11301

研究種目：若手研究(B)

研究期間：2013～2015

課題番号：25770140

研究課題名(和文)韓国語テンス・アスペクトの第一言語習得過程及び習得データのコーパス構築

研究課題名(英文)The L1 Acquisition of Korean Tense-Aspect Markers and the Construction of L1 Korean Corpus

研究代表者

柳 朱燕 (RYU, Juyeon)

東北大学・国際文化研究科・GSICSフェロー

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、「韓国語のテンス・アスペクトの第一言語習得過程」を明らかにすることと共に、「韓国語の言語習得データをコーパス化する」ことを実施してきた。本研究によって、幼児による韓国語の過去形「-ess-」と進行相「-ko iss-」の習得過程が明らかになった。本研究で集めた韓国語の言語習得データは、1歳半から3歳半までの3人の幼児による自然発話データで、全部で177回81時間の膨大なデータである。そのデータを国際的な幼児言語データ交換システムであるCHILDES(Child Language Data Exchange System)にアップロードし、コーパスとして一般公開されるようにした。

研究成果の概要(英文)：This study investigates the development of tense-aspect markers in Korean, -ess (past-perfective) and -ko iss (progressive), and examines the relation between child-directed speech and children's speech. The results indicate that in the early stages of development, the three Korean children used past marker -ess- predominantly with telic verbs, following the Aspect Hypothesis. However, the acquisition of progressive marker -ko iss did not follow the prediction of the Aspect Hypothesis in that the children were not observed to associate progressive marking with activity verbs. In this study, the Korean L1 data was collected from three Korean children (Jong 1;3-3;5, Joo 1;9-3;10, Yun 2;3-3;9). The data consists of longitudinal video-recorded speech samples of adult-child interaction. In total, about 81hr 15min of speech data was collected. This data was uploaded on CHILDES (Child Language Data Exchange System) named as Ryu corpus. Therefore, the Korean L1 data became commonly available.

研究分野：言語学

キーワード：第一言語習得研究 韓国語のテンス・アスペクト 言語習得資料のコーパス構築 アスペクト仮説 CHI LDES インプットの影響

## 1. 研究開始当初の背景

言語習得研究において、テンス・アスペクトの習得は動詞の内在アスペクト(inherent aspect)と強い相互作用があることが広く知られている。Andersen と Shirai は、テンス・アスペクト形式の習得過程には普遍的なパターンが存在するという「アスペクト仮説 (the Aspect Hypothesis)」を提案した(Andersen & Shirai, 1994)。アジア言語においては、日本語や中国語の習得に関して様々な研究が行われ、多くの研究成果が報告されている。しかし、韓国語のテンス・アスペクトの習得を対象にした研究はまだ非常に少ないのが現状である。そこで、本研究では「韓国語のテンス・アスペクトの習得過程」を明らかにすることと共に、韓国語の言語習得データをコーパス化することを目的としている。

## 2. 研究の目的

### (1) 韓国語テンス・アスペクトの第一言語習得過程の解明

テンス・アスペクトの習得は、動詞が持っている内在アスペクト(inherent aspect)と強い相関関係があることがよく知られており、動詞の内在アスペクトによってテンス・アスペクトの習得過程が予測されるという「アスペクト仮説 (the Aspect Hypothesis, Andersen & Shirai 1994)」が様々な言語において検証されてきた。「アスペクト仮説」では、学習者又は幼児は過去形態素の使用を「到達動詞」、「達成動詞」から「活動動詞」、「状態動詞」へ、そして、進行形態素の使用を「活動動詞」、「達成動詞」、「到達動詞」へと発展させていくと主張している。

博士課程での研究では、韓国語の未完了アスペクト形式「-ko iss / a iss-」を対象とし、第一言語における未完了アスペクトの習得過程を明らかにした。研究の結果、韓国語の第一言語習得研究では「アスペクト仮説」に従うとは断定しがたい結果が得られた。これは「アスペクト仮説」における「動作の進行」形態素の習得過程を検証した研究であるが、「アスペクト仮説」では動作の進行の形態素のみならず、「過去形」の習得に関しても言及しているため、「アスペクト仮説」を検証するに当たり、過去形の形態素の習得順序についても検証する必要が生じる。

テンス・アスペクトの習得がなぜこのような過程をたどるかについて、「アスペクト仮説」では、親のインプットの影響を中心に説明している。Shirai & Andersen(1995)では英語児を対象とし、子どもに対する親の発話を分析した結果、「過去形」は「到達動詞」と、「進行相」は「活動動詞」と結び付いた使用頻度が高く、テンス・アスペクトにおいて親の偏った使用パターンが子どもの習得に影響を及ぼす(Distributional Bias Hypothesis)と出張

している。

一方、Bickerton(1981, 1984)は、生得的能力である PNPD(the Punctual-Non Punctual Distinction)によって子どもがテンス・アスペクトを習得していくという「Bioprogram 仮説」を主張している。そこでは、子どもがテンス・アスペクトを習得する際、[限界性(telicity)]と[瞬間性(punctuality)]による認知的区別に基づいて習得していくとされ、「過去形」は[限界性(telicity)]と[瞬間性(punctuality)]を持つ動詞から、「進行相」は[限界性(telicity)]と[瞬間性(punctuality)]を持たない動詞から習得を始めるとされている。問題は、親のインプットが「過去形」は「到達動詞」に、「進行相」は「活動動詞」に偏った場合、「アスペクト仮説」にしても「Bioprogram 仮説」にしても、同じ予測が立てられ、根本的な原因について解明することができない。英語やインド・ヨーロッパ言語がその例に当てはまる。しかし、博士課程での韓国語の場合、親のインプットの偏りが西欧言語と異なる傾向を見せており、「アスペクト仮説」の予測と「Bioprogram 仮説」の予測が異なるパターンを示している。韓国語を対象にテンス・アスペクトの習得研究を行うことによって、今まで検証し難かった二つの理論の妥当性を語るができると考えられる。また、「進行相」の習得と「過去形」の習得が対立関係にありながら、どちらか一つを欠いては習得理論が成り立たないという関係にあることを考えると、「過去形」の習得過程を分析し、より広い観点から「アスペクト仮説」及び「Bioprogram 仮説」を検証する研究は必要である。

一方、本研究が対象とする韓国語の「過去/完了」形態素の習得過程に関しては、まだ研究が十分に行われていない。韓国語の過去形の習得に関する研究として、Seda & Lee (2002)があるが、子どもの言語使用を養育者が筆記により記録したものに基いた研究であり、データの質量に関して問題が残る。博士課程での研究で収集したデータは、4人の子どもの急激な言語発達の時期である1歳半から3歳半までの2年をかけて、ビデオで収録したデータである。このデータには子どもに話しかける養育者の発話も含まれているため、親のインプットを分析することができる。そこで、本研究では、韓国語の「過去形」の習得過程を明らかにした上で、韓国語のテンス・アスペクトの習得過程を明らかにし、認知言語学を基盤にする「アスペクト仮説」と言語の生得性を基盤にする「Bioprogram 仮説」の検証を試みる。

### (2) 韓国語の幼児言語習得データのコーパス化

表1のように博士課程では韓国語を母語として習得している子どもの自然発話データを収集した。これは量的及び質的に L1 習得

研究を行うのに十分なデータである。公開された韓国語 L1 習得のデータがない現実を勘案すると収集したデータを整備した上でコーパス化し、公開することは極めて学問的に価値があると考えられる。

表 1 収集された幼児言語データの詳細

	年齢	収集期間	収録時間
Jong	1;3~3;5	2年2カ月	31hr37min
An	1;7~3;1	1年6カ月	16hr15min
Joo	1;9~3;10	2年1カ月	29hr2min
Yun	2;3~4;3	2年0カ月	26hr36min
合計			103hr30min

韓国語の第一言語習得研究の分野では、まだ公開された幼児言語コーパスがないという非常に深刻な問題を抱えている。言語習得研究に関しては、杉浦・中・宮田・大嶋(1997:80)が述べているように、「データの収集と文字化そして分析処理に膨大な時間がかかるため、従来、個々の研究者による少数のケースに基づく研究しかできなかった」という問題が今の韓国の現実である。しかし、現在では、コンピューターを利用したデータの共有化と分析ツールの提供を行う CHILDES(Child Language Data Exchange System)プロジェクト(MacWhinney 2000)により、大量の自然発話データの分析が可能となっており、欧米の L1 習得研究の分野では共有化されたコーパスを利用して多くの研究が行われている。日本語を母語とする子どものデータについても、日本語版 CHILDES である JCHAT プロジェクト(Oshima-Takane & MacWhinney 1998)により、公開されたデータを使用した大規模な分析が可能になっており、L1 習得研究の分野では近年盛んに利用されている。韓国語の L1 習得コーパスも、欧米又は日本のように公開され、誰でも大量の自然発話データを分析することが可能となることが望ましい。このため、本研究では、博士課程の研究で集めたデータをコーパス化して公開(共有化)することを目的としている。

### 3. 研究の方法

第一言語における韓国語のテンス・アスペクトの習得過程を明らかにするために、本研究では博士課程で収集した幼児言語データを用い、動詞の内在アスペクトによるテンス・アスペクトの使用パターンを分析する。その際、子どもの発話だけではなく、養育者の発話も分析の対象にし、インプットの影響など、原因について考察を行う。

韓国語の幼児言語習得のコーパスを構築するためには、韓国語で文字化されているデータをローマ字表記にする作業、形態素の分析ができるように形態素のタグ付けをする作業、文種類や状況説明などコメントを付ける作業を順次に行わなければならない。最終的には CHILDES に掲載することを

目指している。

前述した研究目的を達成するために、本研究では以下の二つの研究課題を立てる。

課題 1. 韓国語の第一言語習得過程において、テンス・アスペクトの習得はどのような発達過程をたどるか。

課題 2. 韓国語の幼児言語データをコーパス化し、一般公開できるウェブページ CHILDES にアップロードする。

課題(1)に関しては、韓国語の幼児言語習得データを分析し、過去形の習得過程を明らかにする。その際、養育者のデータも分析し、親のインプットの影響を検証する。海外又は国内の学会で研究内容を発表し、*First Language, Applied Psycholinguistics* 等での原著論文発表を計画している。実際、日本語のテンス・アスペクトの習得に関しては Shirai (1993, 1998)が、中国語の習得に関しては Chen & Shirai (2010)で論文が発表されており、韓国語のテンス・アスペクトの習得に関する論文に関心が高まっている状況である。

課題(2)に関しては、博士課程で収集した 100 時間を超えるデータを基にして、コーパスとして利用できるように整備をする膨大な作業が行われる計画である。コーパス化することを説明するために、現在 CHILDES に公開されてある日本語のデータを例として見せる(図 1)。

```

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

現在、韓国語の幼児言語のデータはすべて韓国語で文字化されており、コーパス化するためには、韓国語で文字化された文をローマ字表記にする作業と、形態素の分析ができるように形態素のタグ付けをする作業、文種類や状況説明などコメントを付ける作業を行わなければならない。平成 25 年度には、ローマ字表記を中心にして作業を行う予定で、その際、MLU(平均発話長)の計算が可能になるように、一定の基準を作って既存のデータを再整理するつもりである。ローマ字表記がある程度できたら、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きたい。最終的には CHILDES に掲載することを目指している。

韓国語のローマ字表記と形態素分析に関しては、韓国ソウル大学 Chungmin Lee 教授の協力を得て韓国の国立国語研究院の「ローマ字変換プログラム」と「形態素分析基準」

を参考して行う予定である。また、JCHAT プロジェクトに関わっていた宮田 Susanne 教授（愛知淑徳大）、CHILDES の責任者である MacWhinney 教授（カーネギーメロン大教授）には、韓国語のデータのタグ付けを含め分析が可能になるコーパスを構築し、CHILDES にアップロードすることを既に報告しており、協力を得る計画である。

本研究では、今まで皆無であった韓国語 L1 習得データのコーパスを構築するという目標を立てている。大量の自然発話データが共有化されることによって、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が期待される。日本語は韓国語との構造的類似性の最も高い言語であると言われており、語順の一致、助詞の用法、活用、漢語語彙、敬語など他の外国語にはない多くの共通性を持っている。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、研究を進めていきたい。

また、既存の言語学の理論は主にインド・ヨーロッパ言語に基づいて検証されてきているが、日本語を対象にして検証した研究では、これまで言語学の研究とは異なる結果が得られた研究が少なくない（大関 2008 など）。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。

#### 4. 研究成果

##### (1) 韓国語テンス・アスペクトの第一言語習得過程の解明

本研究では、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を用いた。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「過去形-ess-」について、子どもが産出した 1,491 文と養育者が産出した 12,556 文を、「進行形-ko iss-」についても、子どもが産出した 92 文と養育者が産出した 1,386 文を分析対象とし、各子どもと養育者がすべての調査期間で発話した「過去形-ess-」と「進行形-ko iss-」の使用状況を表 2 でまとめた。分析の結果から、韓国語のテンス・アスペクトの習得過程に関しては、次の 3 点が分かった。

子どもの過去形「-ess-」の使用  
過去形の習得過程はデータの全期間（2 年）と初期段階（stage 1）の使用パターンを総合的に分析してみると、3 人とも「到達動詞」の使用が最も多く、次は「活動動詞」と「達成動詞」の順で使っているが、初期ではあま

り使っておらず、また「状態動詞」の使用を徐々に広げていることが分かった。韓国語の「過去形-ess-」の習得過程は「アスペクト仮説」に従う結果になった。

子どもの進行形「-ko iss-」の使用  
全期間での「進行形-ko iss-」の使用と初期段階の使用を見ると、Joo は「活動動詞」の方を初期段階でも全期間でも多く使っているが、Jong と Yun は「到達動詞」の使用が初期段階で最も多くなり、次の段階では「到達動詞」の使用が減り、「活動動詞」の使用が増えていくことが分かる。従って、「進行形-ko iss-」の習得は、「アスペクト仮説」に沿う子どもが一人、沿わない子どもが二人という、子どもによって個人差が見られる結果が得られた。

養育者のインプットの影響  
インプットの影響を調べるため、子どもと養育者の使用パターンに相関関係があるかを文法形式ごとに比べた。図 1 で示したように、「過去形-ess-」に関しては相関関係が高く（ $r=0.94$ ）、養育者のインプットは子どもの過去形使用パターンに大きな影響を及ぼすことが分かる。しかし、「進行形-ko iss-」に関しては相関関係が低く（ $r=0.58$ ）、養育者のインプットが子どもの進行形使用パターンに影響を及ぼすとは言いにくい。

なぜ「過去形-ess-」は「アスペクトの仮説」にも従うし、インプットの影響も大きく見えるのに対し、「進行形-ko iss-」ははっきりした結果が得られなかったのかについては、「過去形」と「進行形」の使用頻度の偏り、複雑な韓国語の非完結相の体系(Shirai, 2004)、子どもの文法形式習得における個人差(Budwig, 1996)が挙げられる。

##### (2) 韓国語の幼児言語習得データのコーパス化

幼児言語収集に参加した 4 人の子どもの内 An は自閉症スペクトラム（以下、autism spectrum disorder の略称である“ASD”と表記する）を持つ子どもである。実際、データを収録していた 1;7~3;1 の時期では ASD であることを知らず、他の 3 人の健常児と同様に自然発話データを取っていたが、データの収集が終わった 2,3 カ月後、An の養育者から An は ASD だと医療機関から診断されたという連絡を受けた。そこで、An のデータは 3 人の健常児と区別し、本研究の分析対象から排除することと、コーパスとして一般公開しないことに決めた。

「韓国語の言語習得データのコーパス化」に関しては、研究実施計画の通りに CHILDES(Child Language Data Exchange System, <http://chilDES.psy.cmu.edu/>)にアップロードが完了し、韓国語の言語習得データが一般公開されている。CHILDES の database に入

り「Browsable database → East Asian → Korean」を押すと Ryu corpus が出てくる。Ryu corpus の中には3人の幼児 jong, Joo, Yun のコーパスが入っており、jong のコーパスには動画も公開されて、動画を見ながら同時に script を見ることができるようになっている。Jong は 62 回 31 時間 37 分、Joo は 77 回 29 時間 2 分、Yun は 38 回 20 時間 36 分のデータが、分析可能なコーパスとして備えている。Ryu corpus に関しては「Database Manuals → East Asian Language → Korean」に入ると、database の紹介や謝辞など詳しい説明がマニュアルとして書いてある。本研究は、今まで皆無であった韓国語の言語習得データのコーパスを構築するという大きな成果を出している。大量の自然発話データが共有化されることによって、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が期待される。

しかし、研究方法に記載していた Tag 付け（形態素分析）に関しては、現在検討中であり、今後の研究課題として後続研究を行う予定である。Tag 付け（形態素分析）に関して本研究では、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）で本データを処理してみた。その結果、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おり、CLAN を通した形態素分析には不適切なところが沢山出ている。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）と多角度で検討しており、MacWhinney 教授は日本語版 CHILDES である JCHAT に習って韓国語の形態素分析を行うことを勧めている。今後、JCHAT プロジェクトに関わっていた宮田 Susanne 教授（愛知淑徳大）に研究協力を得て、日本語で行われた形態素分析基準を韓国語でどのように適用するか助言をもらうつもりである。

## 5. 主な発表論文等

〔雑誌論文〕(計3件)

Ryu, Juyeon, Horie, Kaoru and Shirai, Yasuhiro (2015). Acquisition of the Korean Imperfective Aspect Markers *-ko iss-* and *-a iss-* by Japanese Learners: A Multipul-Factor Account, *Language Learning*, Volume 65, Issue 4, 791-823. DOI:10.1111/lang.12132 査読有

Ryu, Juyeon and Shirai, Yasuhiro (2014). The First Language Acquisition of the Korean Imperfective Aspect markers *-ko iss-/-a iss-*, In S. Nam, H. Ko, & J. Jun (Eds.), *Japanese/Korean Linguistic*, 21 (pp. 265-279). Stanford,

CA: CSLI. 査読有

柳朱燕(2013)「韓国語未完了アスペクトの第一言語習得過程における誤用の分析」『韓国語教育研究』査読有,第3号,68-87.

〔学会発表〕(計6件)

柳朱燕「韓国語におけるテンス・アスペクトの第一言語習得過程」日本韓国語教育学会(Japan Korean Education Conference).札幌市教育文化会館 2015年10月31日-11月1日

柳朱燕, 白井恭弘「韓国語の第一言語習得におけるテンス・アスペクト形式の習得過程 - アスペクト仮説の検証とインプットの影響を中心に - 」日本言語科学会(Japanese Society for Language Sciences). 別府国際コンベンションセンター 2015年7月18-19日

柳朱燕「韓国語未完了アスペクトの第一言語習得過程における語用の分析」日本韓国語教育学会(Japan Korean Education Conference).九州大学 2014年12月6日

Ryu Ju-yeon and Shirai Yasuhiro. *The L1 Acquisition of Tense-Aspect Marking in Korean*. Poster presented at the 13<sup>th</sup> International Congress for the Study of Child Language. University of Amsterdam. 14-18 July 2014.

Ryu Ju-yeon and Shirai Yasuhiro. *The Role of Input on the L1 Acquisition of Tense-Aspect Marking in Korean*. Paper presented at the Georgetown University Round Table on Languages and Linguistics. Georgetown University. 14-16 March 2014.

Ryu Ju-yeon. *The L1 Acquisition of the Past Marking -ess- in Korean*. Paper presented at New Zealand Linguistic Society Conference. University of Canterbury. 20-22 Nomenber 2013.

〔図書〕(計1件)

柳朱燕、堀江薫(印刷中)「韓国語の未完了アスペクトの習得と教育 - 日本人学習者を対象に - 」稲垣俊史・庵功雄編『日本語のテンス・アスペクトの習得と教育』ひつじ書房

〔その他〕

Ryu コーパス公開サイト CHILDES

データ

<http://childes.psy.cmu.edu/browser/index.php?url=EastAsian/Korean/>

コーパスのマニュアル

<http://childes.psy.cmu.edu/manuals/10eastasian.pdf>

6. 研究組織

(1) 研究代表者

柳 朱燕 (RYU, Juyeon)

東北大学・国際文化研究科・GSICS フェロー  
研究者番号: 40647682