

[原著論文]

L1・L2における英語関係節の習得プロセス －CHILDESコーパスを用いた分析－

田島 弥生

山陽小野田市立山口東京理科大学 共通教育センター

Acquisition Process of English Relative Clauses in L1 and L2: An Analysis with the use of CHILDES

Yayoi TAJIMA

Center for Liberal Arts and Sciences, Sanyo-Onoda City University

Abstract

It has long been debated whether the process of L2 learning is similar to that of L1. Some hypotheses assert that L2 learners as well as L1 speakers have a sort of innate apparatus for language acquisition, like LAD posited by Chomsky in the field of L1 studies. Others suggest that the process of L2 learning should be divergent depending on the L1 background. This study is intended to explore the differences between L1 and L2 in the acquisition process of English relative clauses. Two experiments were conducted—one with English-speaking children aged 7 ($n = 24$) and another with Japanese L2 learners of English ($n = 35$). The Child Language Data Exchange System (CHILDES) corpus was used to investigate the frequency of relative clauses in the natural utterances of children (L1 English), while forced-choice, fill-in-the-blank questions were designed to measure how accurately Japanese L2 learners acquire the form and function of relative clauses. As a result, there was no difference found between the frequency of object-extracted relative clauses (ORCs) and that of subject-extracted relative clauses (SRCs) in natural utterances of English-speaking children. In contrast, Japanese L2 learners of English were clearly found to have much more difficulty in mastering ORCs than SRCs ($p < .01$). These findings are taken to indicate that there exists a process of L2 acquisition, which is different from that of L1.

Key words: CHILDES, corpus, relative clauses, L2, acquisition order

キーワード: CHILDES, コーパス, 関係節, L2, 習得順序

1. はじめに

第二言語習得論のアプローチには、大きく分けて次の三つの段階がある。まず一つ目は、1950年代から1960年代にかけてアメリカ言語学の主流であった構造言語学の分析手法に基づいて、二つの言語間の構造上の差異を体系的に記述しようとした対称分析である。二つの言語はお互いに似ているほど習得しやすく、相違点が多いほど習得しにくいとされ、その相違点は学習者にとっての干渉、すなわち学習に何らかの障害を生じさせる困難点 (trouble spot) になるであろうと想定された。当時、行動主義 (behaviorism) の影響を受けて、言語習得とは刺激に対する反応の繰り返しによって獲得される一種の習慣形成であるとの見方が広まっていた米国では、そのような困難点を集中的に訓練・矯正すれば、飛躍的に学習が促進されると考えられた。

Lado (1957) に代表されるこのような対称分析仮説 (Contrastive Analysis Hypothesis) は英語教育に貢献するものと当時大いに期待されたが、実際には、二言語間の類似点は習得しやすく相違点は習得が難しいという仮説が十分に証明されることはなかった (Gass 1996)。また、言語習得は一種の習慣形成であるという言語観においては、第二言語習得とは詰まるところ、L1の習慣をL2の習慣に置き換えることに他ならない。それ故、L2におけるすべてのエラーはL1による母語干渉の結果であり、L1とL2の構造上の違いさえ特定できれば学習者のエラーは予測できるはずと考えられていたが、それも十分に証明されることはなかった。

例えば、英語とフランス語はどちらもSVOの語順を取るが、目的語が代名詞の場合のみ、フランス語には英語と異なる次のような語順が現れる (馬場・新多 2016)。

- (1) The dog eats *the cookie*. The dog eats *it*.
 Le chien mange *le biscuit*. Le chien *le* mange.
 (犬がクッキーを食べる。犬はそれを食べる。)

つまり、英語は目的語が名詞 (*the cookie*) であろうと代名詞 (*it*) であろうとSVOの語順に変わりはないが、フランス語は目的語が代名詞の場合のみ (二文目の *le* が英語の *it* に相当)、SOVの語順を取るのである。この語順の違いに着目し、対称分析仮説に従って、L2におけるすべてのエラーはL1に起因する母語干渉であると仮定すると、次の二通りのエラーが予想される。

- (2) The dog *it* eats.
 (3) Le chien mange *le*.

しかし実際には、英語母語話者による(3)のようなエラーは頻繁に観察されるが、フランス語母語話者による(2)のようなエラーは観察されない (馬場・新多 2016)。このように、二言語間の構造の違いから第二言語習得プロセスの解明を目指した対称分析仮説ではあったが、実際の教育現場で起きている現象を十分に説明することはできなかった。

構造主義的な言語観の限界、および当時の言語学に多大なる影響を与えていた行動主義の衰退に伴い、第二言語習得研究は対称分析から次の時代、すなわち学習者内部の認知能力に焦点を当てた認知的アプローチの時代へと移っていく。その大きな契機となったのが、Corder (1967) による誤用分析 (Error Analysis) である。言語習得を「刺激→反応」の繰り返しによる習慣形成と位置付けていたそれまでの対称分析においては、誤用は学習者の単なるミスであり、L2の模倣に失敗した結果にすぎなかった。しかし、Corderは、無意識に繰り返される誤用は、学習者自身が自らL2に関する仮説を立て、それを検証している証拠であり、学習者の発達段階を示す重要な手がかりになると指摘した。

例えば、フランス語には英語の動名詞に相当する文法範疇がないため、フランス語母語話者のL2発話には、*We just enjoyed to move* といった誤用が現れる。これはL1とL2の違いに由来する誤用、すなわちL1からL2への言語転移 (母語干渉) であると考えられるため、それまでの対称分析の枠組みで解釈できる。しかし一方、初期の学習者には、L1には起因しない *He comed yesterday* といった誤用も頻繁に現れる。このような誤用はL1とL2の構造上の違いから生じたわけではないため、対称分析の枠組みでは説明できない。むしろ、動詞の過去形には *ed* を付けるというL2の文法的特徴を過剰般化した例であり、学習者内部でL2の規則に関する仮説形成と仮説検証の試行錯誤が行われている証拠である。

このように、対称分析から誤用分析へと関心が移った第二言語習得理論であったが、その認知的アプローチはSelinkerによって、さらに次の段階へと発展していく。Corder(1967)の誤用分析は、学習者の産出言語に現れる誤用のみを分析対象としていたため、例えば、L2学習者の自信の欠如や不十分な理解により産出されていなかった回避エラー等については何も考察されていない。このような不備を補い、L2学習者の言語能力全体を包括的に把握するアプローチとして提唱されたのが、Selinker(1972)による中間言語分析である。Selinkerは、学習者はL2のシステムについて仮説形成と仮説検証を絶えず繰り返しながら、独自の体系的な学習者言語を構築すると提唱し、それを中間言語(interlanguage)と名付けた。それは、L1ともL2とも異なる学習者独自の自律した言語体系であり、絶えずL2へと近づいていく成長の連続体である。中間言語の構築に作用するプロセスとしては、具体的に、言語転移、過剰般化、回避、化石化(fossilization)等、今日でも第二言語習得理論を研究する上で欠かせない概念が指摘されている。

以上、第二言語習得研究の初期のアプローチとして、対照分析、誤用分析、中間言語分析の三つを概観してきたが、このように構造主義的な言語観から学習者内部の創造的な仮説検証へと関心が移った背景には、心理主義の台頭とそれに伴うL1習得研究の発展があった。1960年代後半から1970年代前半にはChomskyの変形生成文法がL1習得研究において注目を集め、またBrown(1973)は心理主義の立場からL1習得過程を創造的構築過程であると表現している。第二言語習得モデルとしての認知的アプローチとは、L2習得過程は基本的にL1習得過程と同じであるという前提に立った主張であり、Chomskyの唱えた言語獲得装置(LAD)によく似た言語習得のための生得的能力がL2習得においても存在すると提唱されていた。

その主な根拠となったのが、Dulay & Burt(1974)の形態素の習得順序に関する研究である。認知的アプローチ以前の言語観では、第二言語習得とは訓練によってL1の習慣をL2の習慣に置き換えることであったため、L2の習得プロセスは学習者のL1によって当然異なるはずであると想定されていた。しかし、Dulay & Burtは、中国語を母語とする子どもの英語学習者グループと、スペイン語を母語とする子どもの英語学習者グループを対象に文法的形態素の習得過程を観察し、L2学習者はL1が何であるかに関わらず、ほぼ同じ習得順序を辿ることを明らかにしている。また、L2学習者の誤りの多くが、L1習得過程の子どもの誤り、つまり発達上の誤りと似通っていることや、L1が何であるかに関わらず、初期の学習者には似たようなエラーが共通して観察されることも分かった(Kellerman 1984)。

このように、L2習得はL1習得と同じような創造的構築過程を辿るのか、それとも学習者のL1によって異なる習得プロセスを辿るのかについては、これまで様々な議論されてきた。そこで本研究では、子どもの発話コーパスであるCHILDES(MacWhinney 2000)を用いて、英語を母語とする子どもの関係節の出現頻度を分析し、その結果を日本語を母語とする英語学習者の関係節習得率と照らし合わせることによって、関係節におけるL1習得プロセスとL2習得プロセスのそれぞれの特徴を探索的に調査する。

2. 関係節の出現頻度(L1)

2.1. 分析手順

CHILDESとは、Child Language Data Exchange Systemの略称であり、子どもの言語発達過程を明らかにすることを目的に作成された発話データベースである(中谷 2019)。本稿では、CHILDESのBrowseable Databaseを使用し、Fletcherコーパスにおける24人の子どもの発話データを元に分析した。発話者は英語を母語とする6歳から7歳の子ども(女児13名、男児11名)で、同じく英語を母語とする調査者との一対一の会話が記録されている。データ採取場所は英国のLiverpool、Berkshire、Yorkの三か所で、小学校の教室や図書館の一室を利用し、Farm Gameというボードゲームで遊びながら自由に会話が展開されている。24名の子どもによる発話は合計6727、一人当たりの発話数平均は280.3、形態素数平均は1438.5であった。本研究では、CHILDESのfreq機能を用いて関係節の出現頻度を関係代名詞の種類別に算出し、それぞれ統計分析を行い検証した。

2.2. 結果

まずは、Fletcherコーパスに観察された関係節の出現頻度を、5つの関係代名詞の種類(who / whose / whom / which / that)別に算出した。その結果、24名の子どもによる合計6727の発話のうち、whoが30発話、whichが23発話、thatが510発話確認された。そのうち、先行詞が明示されており、大人の文法研究において議論されている意味機能と質的に同じ意味機能で使用されている発話のみを関係節として特定した結果、主格のwhoが12発話、主格のwhichが6発話、目的格のwhichが

3発話、主格のthatが10発話、目的格のthatが26発話確認された。それ以外は、疑問詞、指示代名詞、指示形容詞、接続詞等として使用されていた。なお、関係代名詞としての使用か否かに関わらず、whoseとwhomはどちらも出現しなかった(図1)。

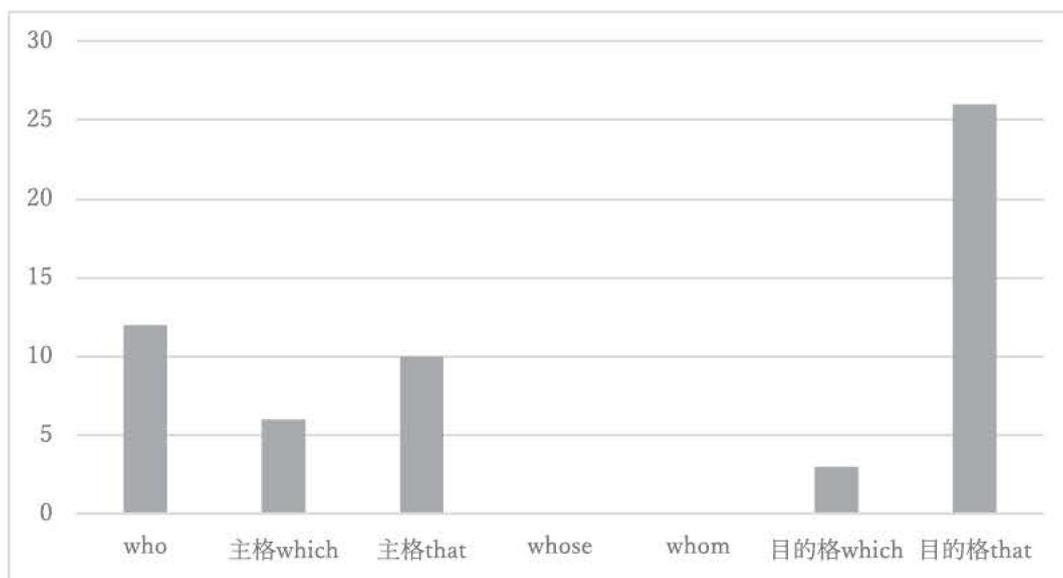


図1 Fletcherコーパスに出現した関係代名詞の発話数

一元配置の分散分析により、英語を母語とする子どもの自然発話コーパスにおける関係節の出現率には、関係代名詞の種類による有意差が確認された($F(4, 119) = 2.504, p = 0.046$)。さらにTukey-KramerのHSD検定により、目的格関係代名詞whichと目的格関係代名詞thatの出現率の間には有意差が認められた($p = 0.036, \eta^2 = 0.678$)。

また、関係節の使用に関しては、子どもによる個体差も観察された。種別に関わらず、上記いずれかの関係代名詞を使用した複文を発話した子どもは24名中17名であり、残りの7名は関係節自体を使用していなかった。最も数多く使用した子どもの関係節発話数は10であり、被験者間の個体差には一標本のt検定による有意差が確認された($M = 2.375, SD = 2.871, p = 0.0002$)。種類別に見ると、主格whoは7名、主格whichは3名、目的格whichは2名、主格thatは5名、目的格thatは9名の子どもによる使用がそれぞれ確認された。なお、24名の子どものMLUm値と関係節の出現頻度には正の相関が確認され($r = .47, p = .018$)、発話数の多い子どもほど関係節を使用していることも分かった。^{*1}

一方、関係代名詞の種類に関わらず、主格関係節全体と目的格関係節全体の出現頻度には違いは確認されなかった。24名の子どもの主格関係節(who / which / that)全体の発話数は28、目的格関係節(which / that)全体の発話数は29のほぼ同数であり、一人あたりの平均発話数も主格関係節1.167($SD = 1.401$)、目的格関係節1.208($SD = 2.105$)であった($F(1, 47) = 0.007, p = 0.936$)。

3. 日本語母語話者の関係節習得率

3.1. 実験手順

次に、日本語母語話者を対象とした関係節の習得率について調査した。実験参加者は山陽小野田市立山口東京理科大学に在籍する18歳～22歳の大学生35名(男子学生27名、女子学生8名)である。いずれも第二言語として大学のカリキュラムに沿って英語を学んでいる学生であり、長期間の海外滞在経験はない。英語運用能力はTOEIC Listening and Reading Testのスコア300-600間に位置するPre-Intermediateレベルである。

【注】

¹ MLUm (Mean Length of Utterances in morphemes) とは、英語を母語とする子どもの発達段階を示すことを目的にBrown (1973) によって考案された指標で、発話の長さの平均値を指す。具体的には発話された形態素数を発話数で割ったもので、MLUm値が高いほど大人に近い文法を獲得していると推定される。

実験参加者には、関係節の理解を問う24問の選択式空所補充問題に解答してもらった。内訳は、主格関係代名詞whoが正解となる問題(4a)が6問、目的格関係代名詞whomが正解となる問題(4b)が6問、所有格関係代名詞whoseが正解となる問題(4c)が6問、そして目的格関係代名詞が省略されている問題(4d)が6問である。

- (4) a. I have a friend () helps me when I am in trouble.
 ① who ② whose ③ whom
- b. He is a soccer player () everyone knows.
 ① whose ② whom ③ which
- c. I have a friend () father is a police man.
 ① whom ② who ③ whose
- d. This is a book () bought yesterday.
 ① who ② I ③ which

3.2. 結果

上記の実験の結果、関係代名詞の種類によって正解率に大きな差があることが判明した。最も正解率が高かったのは、79.5%の主格whoであった($M = 4.77, SD = 1.29$)。次に、所有格whoseが68.5%($M = 4.11, SD = 1.29$)、目的格関係代名詞の省略が62.3%($M = 3.74, SD = 1.77$)と続き、最も正解率が低かったのは57.7%の目的格whom($M = 3.46, SD = 1.75$)であった(図2)。一元配置の分散分析により、関係代名詞の種類による正解率の違いには有意差が確認され($F(3, 139) = 3.86, p = .011$)、さらにTukey-KramerのHSD検定により、whoとwhomの正解率には明らかな有意差が認められた($p = .009, \eta^2 = 0.852$)。

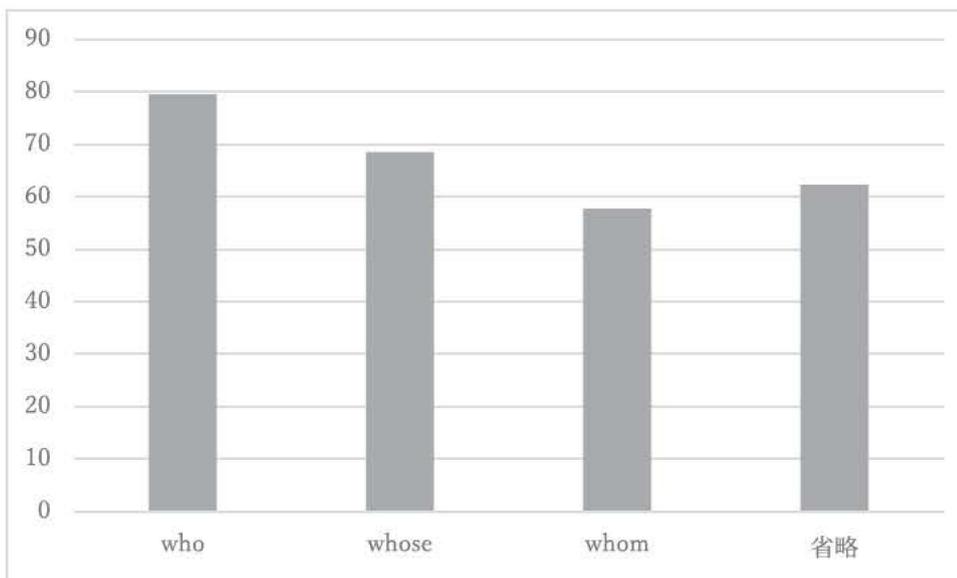


図2 日本語母語話者による関係節空所補充課題の正解率

4. 考察

本研究では、関係節のL1習得プロセスとL2習得プロセスについて探索的に調査することを目的に、まずは英語を母語とする24人の子どもの自然な発話を記録したFletcherコーパスを用いて、そこに現れる関係節の出現頻度を分析した。その結果、最も出現率が高かったのは、全体の45.6%を占める目的格関係代名詞のthatであった。同じ目的格であってもwhichの使用はわずか5.3%に留まり、whomに至っては一例も確認されなかった。一方、主格の関係代名詞に関しては、which(10.5%)よりもwho(21.1%)やthat(17.5%)の使用率の方が高かった。つまり、英語を母語とする子どもの自然な発話にお

いては、目的格関係節には主にthatを使用する一方、主格関係節にはwhoとthatを併用する傾向があることが分かった。

特に顕著な有意差が現れたのは、目的格関係代名詞のthatとwhichの出現率である。目的格thatを使用した子どもの数は9名に上り、合計26の発話が確認されたが、目的格whichを使用した子どもの数は2名のみであり、発話数もわずか3に留まった($p = 0.036, \eta^2 = 0.678$)。また、目的格whichを使用した子どもは同時に目的格thatも併用していたが、目的格thatを使用せずに目的格whichのみを使用した子どもはいなかった。もちろん横断的調査における出現頻度をそのまま子どものL1獲得順序の指標にすることは出来ないが、これらの結果から、英語を母語とする子どもは目的格whichよりも先に目的格thatを習得する傾向があると推測できる。

一方、主格関係節に関しては、その出現頻度に関係代名詞の種別による統計的有意差は確認されなかったが、その併用パターンについては目的格関係節とは異なる傾向が観察された。主格関係代名詞を使用した子どもの人数はそれぞれ7名(who)、5名(that)、3名(which)であったが、whoとwhichを併用した子どもも、thatとwhichを併用した子どももいなかった。また、whoを使用した7名のうち、thatを併用した子どもは2名確認されたが、whoを使用せずにthatのみを使用する子どもも観察されたため、目的格関係節のように、特定の主格関係代名詞が先に習得されているという仮説を立てることはできなかった。むしろ主格に関しては、子どもが自らの言語環境において、おそらくはそのインプットに従い、いずれか一つの種類の主格関係代名詞をそれぞれ習得し始めると推論すべきであろう。

そして最も特筆すべきは、英語を母語とする子どもの主格関係節全体の発話数と目的格関係節全体の発話数の間には、何の違いも見出されなかったことである。24人の子どもの発話数は、主格関係節が28、目的格関係節が29のほぼ同数であり、一人あたりの平均発話数も主格関係節と目的格関係節とでさほど変わらなかった($F(1, 47) = 0.007, p = 0.936$)。一方、日本語母語話者を対象とした関係代名詞の選択式空所補充課題の結果では、主格関係節の正解率(79.5%)が目的格関係節の正解率(57.7%)を明らかに上回り、両者の正解数の平均値には顕著な有意差が認められた($p = .009, \eta^2 = 0.852$)。また、有意差までは出ていないものの、主格関係節の平均正解数は、関係代名詞が省略された目的格関係節の平均正解数も上回っていた($p = .062, \eta^2 = 0.067$)。

本研究で分析したFletcherコーパスは英語を母語とする子どもの自然発話データであり、whoseとwhomを含む発話は一例も確認されていない。一方、日本語母語話者を対象とした実験はwhoseとwhomを含む選択式の空所補充問題であるため、この両者の分析結果を単純に比較することはできないが、全体的な傾向として、英語を母語とする子どものL1習得過程においては、主格関係節と目的格関係節の習得順序にさほどの違いはないが、日本語母語話者のL2習得過程においては、主格関係節よりも目的格関係節の方が明らかに習得に時間が掛かるということは言えるだろう。さらに、日本語母語話者のwhomの正解率はwhoseの正解率すら下回っており、日本語を母語とするL2学習者にとって目的格関係節がいかに習得の難しいtrouble spotであるかがわかる。

Keenan & Comrie (1977) は、世界の様々な言語の関係節を調査し、関係節化が可能となる位置には普遍的な階層が存在するとして、以下のようなNoun Phrase Accessibility Hierarchy (NPAH) を提唱している。これは、特定の言語において、ある階層に位置する語が関係詞化できれば、それよりも上位にあるすべての階層において関係詞化が可能であることを示している。

Noun Phrase Accessibility Hierarchy (NPAH)

SU (subject)	the man who bought the book
DO (direct object)	the book which he bought
LO (indirect object)	the boy whom he gave the book
OBL (oblique object)	the boy whom the man gave the book to
GEN (genitive)	the man whose book was sold out
OCOMP (object of comparison)	the woman whom he is taller than

NPAHは元々、関係節化の普遍的階層性を示すものではあったが、この階層と関係節のL2習得における難易度との関連性についてはこれまでに何度も指摘されており、下位の階層ほど習得が難しいと考えられている(Gass 1979, Kawauchi 1988, Kubota 1993)。本研究の実験結果においても、日本語を母語とするL2学習者は主格関係節よりも目的格関係節の習得に困難を抱えていることが示され、NPAHをL2習得の難易度を示す指標として応用する妥当性が一部示されたと言える。

しかし、この関係節のL2習得過程は、本研究で観察された英語を母語とする子どものL1習得過程とは一致していない。まだ完全に大人と同じ文法が獲得されているとは言えない6歳～7歳の子どものグループで、そのうちほぼ3割の子どもの発話にはそもそも関係節自体が出現していないというこの段階で、主格関係節と目的格関係節の発話数に差がないという結果は、L2習得とは異なるL1習得プロセスの存在を示唆していると言えるだろう。

構造主義的な対称分析から始まり、誤用分析、中間言語分析を経て、認知的アプローチに関心が移った第二言語習得理論ではあったが、前述した通り、その創造的構築仮説には、L2習得過程は基本的にL1習得過程と同じであり、言語獲得装置(LAD)によく似た言語習得のための生得的能力がL2習得においても存在するという前提があった。しかし、本研究の結果はこれにはそぐわない。L2習得プロセスは必ずしもL1習得プロセスと一致するものではなく、やはりそこにはL1の影響も含めたL2独自の習得プロセスが存在していると解釈すべきであろう。

5. 結論

本稿では、関係節におけるL1習得プロセスとL2習得プロセスの違いを探索的に調査することを目的に、英語を母語とする子どもの自然発話コーパスの分析と、日本語母語話者を対象とした選択式空所補充課題とを実施した。その結果、英語を母語とする子どもの発話には、主格関係節と目的格関係節が同程度の頻度で産出されていたが、日本語を母語とするL2学習者は主格関係節よりも目的格関係節の習得に時間が掛かることが明らかに示された。これにより、L2習得プロセスとは異なるL1習得プロセスの存在が示された。

しかし、横断的調査によって採取された特定の文法要素に関するデータは、その時点における当該要素の獲得を示唆するものではあるが、何をもちて獲得と判断すべきかについては様々な議論がある(Stromswold 1996, Snyder 2007)。より精密に獲得順を判断するためには、横断的調査だけではなく、縦断的調査も今後併せて行う必要があるだろう。

また、英語を母語とする子どもの発話データからは関係代名詞の種類による使用率の違いが観測されたが、本研究のデータはコーパスによる口語データのみであったため、口語体と文語体による関係代名詞の使用率の違いについては何も検討されていない。L1習得における文法要素の獲得順をより正確に判断するためには、産出された文語体データも含めて評価していく必要がある。さらに、日本語母語話者のデータに関しても、今回実施した選択式空所補充課題だけではなく、自由英作文課題や文法性判断テスト等も併せて行うことによって、より包括的にL2習得プロセスを解明し、効果的な学習法の開発に寄与することができるだろう。

【引用・参考文献】

- Brown, R. (1973). *A first language: The early stages*. Cambridge, MA: Harvard Univ. Press.
- Corder, S. P. (1967). The significance of learner's errors. *International Review of Applied Linguistics in Language Teaching*, 5, 161-170.
- Dulay, H., & Burt, M. (1974). Natural sequences in child second language acquisition. *Language Learning*, 24, 37-53.
- Gass, S. (1979). Language transfer and universal grammatical relations. *Language Learning*, 19, 327-344.
- Gass, S. (1996). Second language acquisition and linguistic theory: The role of language transfer. In W. Ritchie & T. Bhatia (eds.) *Handbook of Second Language Acquisition* (pp. 317-345). New York: Academic Press.
- Kellerman, E. (1984). The empirical evidence for the influence of the L1 in interlanguage. In A. Davies, C. Criper & A. Howatt (eds.) *Interlanguage* (pp. 98-122). Edinburgh: Edinburgh University Press.
- Kawauchi, C. (1988). Universal processing of relative clauses by adult learners of English. *JACET Bulletin*, 19, 19-36.
- Keenan, E., & Comrie, B. (1977). Noun phrase accessibility and universal grammar. *Linguistic Inquiry*, 8, 63-99.
- Kubota, M. (1993). Accuracy order and frequency order of relative clauses as used by Japanese senior high school students of EFL. *Institute for Research in Language Teaching Bulletin*, 29, 367-374.
- Lado, R. (1957). *Linguistics across cultures*. Ann Arbor, MI: University of Michigan Press.
- MacWhinney, B. (2000). *The CHILDES Project: Tools for analyzing talk*. 3rd edition, Vol. 2. *The Database*. Mahwah: LEA.
- Selinker, L. (1972). Interlanguage. *International Review of Applied Linguistics in Language Teaching*, 10, 209-231.
- Snyder, W. (2007). *Child language: The parametric approach*. Oxford: Oxford University Press.

- Stromswold, K. (1996). Analyzing children's spontaneous speech. In D. McDaniel, C. McKee & H. S. Cairns (eds.) *Methods for assessing children's syntax* (pp. 23-53). Cambridge: MIT Press.
- 馬場今日子・新多了 (2016).『はじめての第二言語習得論講義』大修館書店.
- 中谷健太郎 (2019).『ことばの実験研究の方法－容認性調査、読文・産出事件からコーパスまで』ひつじ書房.