

## 大東文化大学 博士学位論文審査報告書

氏 名	三浦 邦彦
学 位	博士 (英語学)
学 位 記 番 号	乙第37号
学位授与年月日	平成30年3月22日
審 査 研 究 科	外国語学研究科
論 文 題 目	A Japanese Lower Secondary School English Learner Corpus and its Applications
論 文 審 査 委 員	(主査) 大東文化大学教授 山崎 俊次 (副査) 大東文化大学教授 北林 光 (副査) 大東文化大学准教授 ロバート J シグレー (副査) 中央大学教授 新井 洋一

#### 博士論文 審査報告

この部分に掲載されている内容については、論文の内容の要旨及び論文審査の結果の要旨に関する箇所では無い為、加工がされておりますので、ご了承願います。

## 2. 論文の要旨およびその特色

本論文は、東京大学教育学部附属中等教育学校の中等部生徒の3年間に及ぶ品詞、特に形容詞、動詞、前置詞の理解と発達についての研究論文である。英語の授業を Teacher Talk の機能別発話分析として意思・感情等に関わる言語機能を4項目、小項目も含めて30種類の言語機能を設定して研究したものである。また Learner-centered activities としてペアワーク・グループワークを通して英語で授業を行い、また年度末に生徒の書いた英文の Learner Corpus を構築して、生徒の形容詞、動詞、前置詞の理解と発達を定量分析 (quantitative analysis)、定性分析 (qualitative analysis)、内容分析 (content analysis) などの複合分析 (multiple analysis) を用いて統計解析した研究論文である。論文全体は9章で構成され、巻末には生徒との会話の分析、詳細な品詞 tag sets、そして主要文献があげられている。

本論文の構成は以下の通りである。

1. Introduction
2. Literature review
  - 2.1. Introduction
  - 2.2. What is a learner corpus?
  - 2.3. Important elements of learner corpora design
  - 2.4. Trend of learner corpus research
  - 2.5. Previous research of learner corpus
3. Research purposes
4. Methodology
  - 4.1. Data collection
  - 4.2. How to build a learner corpus
  - 4.3. How to analyze a learner corpus
5. Results: First Year
  - 5.1. Learner corpus size
  - 5.2. Word frequency list
  - 5.3. Analysis of POS usage
    - 5.3.1. Verbs
    - 5.3.2. Nouns
    - 5.3.3. Adjectives
    - 5.3.4. Adverbs
    - 5.3.5. Prepositions
    - 5.3.6. Conjunctions
  - 5.4. Analysis of learners' collocation patterns
  - 5.5. Chi-squared test
  - 5.6. Data analysis of verb in detail
    - 5.6.1. Frequency list of verbs
    - 5.6.2. Data analysis: be verb; *am, is, are*
    - 5.6.3. Data analysis: auxiliary verb; *do, does, did*

- 5.6.4. Data analysis: modal auxiliary verb; *can, will*
- 5.7. Focus on collocation patterns of past tense of be verbs and some lexical verbs
  - 5.7.1. Past tense of be verbs: *was, were*
  - 5.7.2. Lexical verb: *playing*
  - 5.7.3. Lexical verb: *have*
  - 5.7.4. Lexical verb: *get*
- 6. Results: Second Year
  - 6.1. Introducing second year corpus research
  - 6.2. How was the second year learner corpus built?
  - 6.3. Second year learner corpus size
  - 6.4. Wordlist of the second year learner corpus
    - 6.4.1. How to make a wordlist of the second year learner corpus
    - 6.4.2. Getting to know more about keywords
  - 6.5. Overview the vocabulary usage in different part of speech
    - 6.5.1. Comparing noun word lists
    - 6.5.2. Comparing verb word lists
    - 6.5.3. Comparing preposition word lists
    - 6.5.4. Comparing conjunction word lists
    - 6.5.5. Comparing adjective word lists
    - 6.5.6. Comparing adverb word lists
  - 6.6. Analysis of adjectives and adverbs to use a new tag set
  - 6.7. Wordlist of verbs in the second year learner corpus
    - 6.7.1. How to make a wordlist of verbs
    - 6.7.2. A word list of verbs
    - 6.7.3. Overviewing 100 most frequent verbs
    - 6.7.4. Comparing 10 most frequent verbs in the 1<sup>st</sup> year learner corpus and the second year learner corpus
    - 6.7.5. 10 most frequent verbs in the second year learner corpus
  - 6.8. Examining developmental verb usages
    - 6.8.1. As the usage of that-clause (Ranking 25<sup>th</sup>: *think*)
    - 6.8.2. As the usage of present & past passive voice (Ranking 26<sup>th</sup>: *spoken*, 53<sup>rd</sup>: *seen*, 543<sup>rd</sup>: *used*, 89<sup>th</sup>: *made*)
    - 6.8.3. As the usage of S+V+O+C (Ranking 28<sup>th</sup>: *tell*, 29<sup>th</sup>: *call*)
    - 6.8.4. As the usage of -ing participle (Ranking 35<sup>th</sup>: *playing*, 43<sup>rd</sup>: *watching*, 46<sup>th</sup>: *listening*, 67<sup>th</sup>: *reading*, 79<sup>th</sup>: *working*, 82<sup>nd</sup>: *eating*)
    - 6.8.5. As the usage of lexical verb *want* (Ranking 22<sup>nd</sup>)
    - 6.8.6. As the usage of a past participle *been* (Ranking 56<sup>th</sup>)
  - 6.9. Preposition use in the second year learner corpus (Ranking 56<sup>th</sup>)
    - 6.9.1. Preposition frequency list by CLAWS 7
    - 6.9.2. Preposition frequency list by a new tag set
    - 6.9.3. Analysis of prepositions through N-gram analysis, concordance analysis and a new tag set analysis
  - 6.10. Error analysis of most frequent 10 prepositions
- 7. Results: Third Year
  - 7.1. Introducing the third year learner corpus research
  - 7.2. How was the third year learner corpus built?
  - 7.3. Third year learner corpus size
  - 7.4. How to make a wordlist of the third year learner corpus
    - 7.4.1. Wordlist of the third year learner corpus
    - 7.4.2. Getting to know more about keywords
  - 7.5. Overview the vocabulary usage in different part of speech
    - 7.5.1. Comparing noun wordlists
    - 7.5.2. Comparing verb wordlists
    - 7.5.3. Comparing preposition wordlists
    - 7.5.4. Comparing conjunction wordlists
    - 7.5.5. Comparing adjective wordlists
    - 7.5.6. Comparing adverb wordlists
  - 7.6. Analysis of adjectives and adverbs to use a new tag set
    - 7.6.1. Analysis of adjective uses based on a new tag set
    - 7.6.2. Analysis of adverb uses based on a new tag set

- 7.7. Wordlist of verbs in third year learner's corpus
  - 7.7.1. How to make a wordlist of verbs
  - 7.7.2. A wordlist of verbs
  - 7.7.3. Overviewing 100 most frequent verbs
  - 7.7.4. Focusing on 10 most frequent verbs
  - 7.7.5. Comparing 10 most frequent verbs in the second year learner corpus and the third year learner corpus
  - 7.7.6. 10 most frequent verbs in the third year learner corpus
- 7.8. Characteristic verb usages in the third year learner corpus
  - 7.8.1. Clarifying learner's developmental language use from the usage of to-infinitive as a key collocation
  - 7.8.2. Clarifying learner's developmental language use from the usage of present participle
  - 7.8.3. Clarifying learner's developmental language use from the usage of past participle
  - 7.8.4. Clarifying learner's developmental language use from the usage of relative pronouns
- 7.9. Content analysis focused on the usage of verbs and nouns
- 7.10. Preposition use in the third year learner corpus
  - 7.10.1. Preposition frequency list by CLAWS 7
  - 7.10.2. Preposition frequency list by a new tag set
  - 7.10.3. Analysis of prepositions through doing N-gram analysis, concordance analysis and a new tag set analysis
- 7.11. Error analysis of most frequent 10 prepositions
- 8. Applications
  - 8.1. Corpus-based grammar teaching
  - 8.2. Introducing two C-R activities
  - 8.3. Corpus based vocabulary teaching
  - 8.4. Textbook page sample and related issues
  - 8.5. Pair work and related issues
  - 8.6. Teacher talk
  - 8.7. On increasing student talk use
  - 8.8. On using the Sinclair and Coulthard Model
- 9. Conclusion
  - 9.1. The present situation
  - 9.2. Future English teacher training issues
  - 9.3. Research limitations
  - 9.4. Research implications
- Appendix I. Teacher Talk & Student Talk Analysis
- Appendix II. New Tag Set
- Selected Bibliography and References

### 3. 論文の審査内容および評価

本論文は、日本の英語教育において中学生がいかにかに効果的に英語を学ぶことができるか、またそれを助長するために取り得るべき方法を模索した論文である。東京大学教育学部附属中等教育学校の中等部生徒の3年間に渡る品詞の理解・発達、特に形容詞、動詞、前置詞の理解・発達について corpus-based approach に基づき生徒の書いた英文の Learner Corpus を統計分析した研究論文である。

本論文で取られた授業形式として、英語の授業を英語で行ない、Learner-centered activities を促進するためにまず教員が行なう Teacher Talk の機能別発話分析として意思・感情等に関わる言語機能の4項目を設定している。

1. 意思を伴う機能 (Expressing and finding out intellectual attitudes) (小項目: 8種類)
2. 感情を伴う機能 (Expressing and finding out emotional attitudes) (小項目: 5種類)
3. 相手の行為に影響を与える機能 (Getting things done) (小項目: 5種類)
4. 挨拶等の社交上の機能 (Socializing) (小項目: 6種類)

さらに小項目も加えて30種類の言語機能 (language functions) を設定している。そのように英語で授業をすることにより、また Teacher Talk の言語機能によって、生徒との相互作用、新出語彙の紹介などに大いに役立つことが判明した。

日本の英語教育は、演繹的、帰納的英語教育の観点からその方法論が多く模索されてきたが、伝統的に grammar translation method が一般的に用いられて来ていてその功罪について多くが述べられている (Thornbury, 1999)。またそれに対する教授法として communicative language teaching (CLT)、direct method、audio-lingual method などの言語教育法が提唱されてきた (一部には、行き過ぎた CLT により英文

の読み書き能力の欠如の結果、基本に戻るべき (back to the basics) であるという意見もある)。そこで、本論文では演繹法と帰納法の統合として意識高揚法 (conscious-raising activities, C-R) を挙げている (Alwright, 1984)。本論文では、C-R Approach を次のように表現している。'C-R emphasizes the process of discovery learning. C-R might, therefore, be profitably considered as an approach which would allow the integrating of some elements of grammar focused lessons with communicative activities'.

本論文の根幹をなす言語研究の方法としてコーパスを利用した言語研究について詳細に論述している。コーパスは様々な言語研究の目的のために集積された電子化された大量 (100 万語から数億語) の生のデータ (言語資料) であり、それを使って客観的な語彙リストを作成し、使用頻度によりその目的に適合した基本語彙の作成が可能である (例えば Paul Nation の基本語彙 2000 語)。実際の例として英米の新聞記事の社説の比較語彙リスト、年齢差、性差、社会階層差等における比較語彙リストなど多方面の語彙比較が可能である。また電子化されたコーパスを利用する大きなメリットとして語彙の共起に伴うコロケーション研究が客観的、科学的に遂行できる特徴がある。連続した語彙の強さの繋がりや勿論、離れた位置における語彙の共起コロケーションの研究が統計処理することで数値として証明される (Log-likelihood など)。

本論文の大きな特徴として、生徒の理解・発達を図る尺度として年度末に 3 年間生徒に英語で書かせた Learner Corpus の構築にある。生徒の品詞、特に形容詞、動詞、前置詞の理解・発達を客観的に分析するために、そのコーパスに英国のランカスター大学で構築された CLAWS 7 を用いて品詞タグを付与した。しかし、生徒の書く英文には正誤含めて多種多様になり、95% から 97% の正確さを持ち 135 の品詞タグを付与する CLAWS7 では対応しきれない部分があるために、独自のタグを付与するために考案した所に特徴がある。例えば、can に対して CLAWS7 では '\_VM(modal auxiliary; can, will, would, etc.)' と一種類しか付与していない法助動詞に present, plural, 1<sup>st</sup> person など細かいタグを付与し、助動詞だけで 63 種類のタグを付与し、動詞全体で 326 種類のタグが設定されている (その他は形容詞・副詞: 45 種類、前置詞: 58 種類)。勿論 CLAWS7 は文法タグであり、本論文のタグは意味タグも包含している点は指摘する必要がある。

また、語彙共起のコロケーション研究では、N-gram を使って分析を行い、生徒の英文の誤文を探す誤差解析 (Error analysis) も行っている。対応分析 (correspondence analysis) を使って、形容詞+名詞のコロケーション分析をし、3 年生の形容詞+単数名詞+普通名詞の tri-gram における発達の分析を試みている。生徒のコーパスを利用して Concordance lines を分析することによって生徒の例文に散見される誤文が歴然となる。例えば、日本語が影響すると思われる \*I can't tennis (私はテニスができない)、\*Can you tennis? などが挙げられている。またこれらの英文において、統計的解析を行ない、定量分析 (quantitative analysis)、定性分析 (qualitative analysis)、内容分析 (content analysis) などの複合分析 (multiple analysis) を使って生徒の理解・発達を分析している日本の中等学校の英語教育に大きく寄与する研究論文である。

#### 4. 結論

以上の審査内容、評価に基づき、本論文を審査対象とする学位論文審査委員会は、全員一致をもって、本論文は博士 (英語学) の学位を授与するに値するものと判断し、ここに報告する。

#### 審査委員・副査の評価

##### ■北林光

This dissertation breaks new ground in learner corpus research because it deals with junior high school student English. Following the progress of a group of Japanese children over a three-year period, it describes in detail the second language acquisition process. To accomplish this, the dissertation makes use of a new tag set that supplements the CLAWS 7 tag set. It is also unique in that it makes extensive use of error analysis, thus showing stage by stage the learning process. Even if it does not always fully succeed in what it attempts to accomplish, this dissertation is highly original in terms of concept and opens a path for other researchers to follow and explore.

##### ■Robert Sigley

Miura has constructed a longitudinal learner corpus representing written English output from 120 students at a Japanese junior high school in three consecutive years. This will certainly be a valuable resource for future research into development of English vocabulary and grammar among Japanese learners of English. The present research uses the output of an automatic part-of-speech tagger (CLAWS7), and conducts a broad-spectrum survey of vocabulary and grammar in each year.

In terms of methodology, the results are limited by the accuracy of the automatic tagger output, and by the use of restricted-length word sequences (n-grams) to retrieve constructions. Nevertheless, the findings, firstly, serve to identify some important areas where the tagger output needs to be augmented to deal with the language produced by Japanese learners of English (essentially, performing error analysis); and

secondly, confirm that these students do in fact show considerable development of their English system over their course of study. Finally, the results suggest some promising directions for future use of corpora to analyze development of vocabulary and grammar among Japanese learners of English, and as a way of consciousness-raising in the classroom.

■新井洋一

本論文は CLAWS7 に加えて新しいタグを付与して分析している意欲的な論文である。文法タグである CLAWS7 に意味範疇タグまで踏み込んで多くのタグを付与しているが対応する具体例を示し、多少重複するタグの整理が必要である。またスペルミスや語法の修正が必要な個所があるので注意が必要である。全体的に英語教員として英語教育の向上を意図した興味ある研究論文である。

以上