

# 博士論文和文要旨

論文題目：フランス語学習者の発話における使用語彙分析

氏名：杉山香織

本論文の目的は、日本人フランス語学習者がフランス語を口頭産出する際に、母語話者の口頭産出と比較してどのような差異が見られるのかについて、使用語彙の観点から考察し、記述することである。先行研究より、体系的に構築された日本人フランス語学習者話し言葉コーパスは存在しないことが分かったことから、フランス語学習者コーパスの構築を行った。そしてこのコーパスを使用して、学習者の使用語彙について「語彙の豊かさ」、「特徴語」、「コロケーション」、「N-grams」という4つの観点から計量的そして質的な分析を行い、学習者の特徴の記述をした。より幅の広い学習者を対象とするため、フランス語初級学習者は口頭産出をより簡単に行うことができるタスク形式で会話を行い、フランス語で会話ができるレベルの学習者は、自由会話を行った。このため、初級学習者39名によって構成されている「タスクに基づくコーパス」と、初中級学習者から上級学習者までの計38名で構成されている「自由会話コーパス」の2種類のコーパスについて、以上の4つの観点から分析を行った。母語話者との比較を行うため、母語話者についても同様に「タスクに基づくコーパス」と「自由会話コーパス」を使用した。

本論文では、語彙は全ての言語的要素の根幹となるものであるという立場を取る。したがって、使用語彙研究を出発点として、統語レベルや談話レベルでの分析も行った。

本論文の構成は以下の通りである。

序章では、本論文の目的を示し、構成を説明している。

本研究ではコーパスに基づいて分析を行うため、第一章である「コーパス言語学」では、コーパス言語学とはどういう分野であるかを俯瞰している。「1.1. 言語学におけるコーパスとは何か」では、言語コーパスの特徴や、コーパスに必須の概念である代表性について説明を記している。「1.2. コーパス言語学の歴史」では、コーパス言語学がどのような状

況下で誕生をしたのか、どのように発展の過程を辿ったのかについて述べた上で、近年のコーパスについてまとめている。その後「1.3. コーパス分析ツール」では、コーパス分析を可能とする分析ツールについて紹介する。「1.4. 学習者コーパスとプロジェクト」では、中間言語の説明を行った後、コーパス分析によって中間言語分析にどのような変化がもたらされたのかを明らかにし、学習者コーパスとは何かについて定義を行っている。そして、代表的なフランス語学習者コーパスプロジェクトを概観している。「1.5. コーパスの利点と欠点」では、コーパス研究についての注意点を再考している。そして、「1.6. まとめ」で第一章のまとめを行う。

第二章は、語レベルの研究についてまとめられている。まず「2.1. 『一語』とは何か」で、様々な語の単位について整理している。「2.2. 語彙リスト開発の歴史」で、語彙研究が言語教育分野で発展する礎となった語彙リストの開発の歴史を俯瞰している。「2.3. 語彙と第二言語の能力」で、第二言語能力において語彙が中心的役割をすることを確認した後、「2.4. 語彙能力とは何か」では、受容語彙と産出語彙、語彙の深さと語彙の広さという特に2つの対立関係についての説明が行われている。「2.5. 伝統的な語彙の測定」では、これまで行われてきた語彙研究における測定対象と測定方法について、整理を行っている。コーパスの登場とともに注目を集めるようになった語彙の豊かさについて、「2.6. 語彙の豊かさとその測定法」で、これまで考察が行われてきた語彙の豊かさという概念の説明とその測定方法について整理を行い、「2.7. 語彙の豊かさの問題点」でその測定方法の欠点を指摘する。「2.8. 頻度情報を使用した語彙の測定法」では、英語における産出言語の語彙の測定で用いられてきた *Lexical Frequency Profile* についての説明を行い、そのフランス語版の *VocabProfil* の紹介とその問題点を提示している。「2.9. 過剰使用と過少使用」では、コーパスに基づいて行われてきた語レベルの研究の一つである過剰使用と過少使用研究についてまとめられている。そして、「2.10. 学習者の語彙研究」では、語彙レベルから行われてきた学習者の使用語彙に関する研究例と *VocabProfil* を使用した研究例を挙げている。最後に第二章のまとめを「2.11. まとめ」で行う。

第三章では、語レベルの研究よりも大きな単位である **Multi-Word Units** についてのまとめが行われている。「3.1. 単語以上の単位 **Multi-Word Units**」では、単語以上の単位として使用されてきた様々な名称について整理されている。「3.2. フレイジオロジーの歴史的背景」では、**Multi-Word Units** 研究であるフレイジオロジー研究について、その伝統的アプローチと頻度アプローチの特徴についてまとめられている。「3.3. MWUs の測定法」について、その主な分析であるコロケーション分析と **N-grams** 分析について測定方法がまとめられている。「3.4. 中間言語のフレイジオロジー」では、学習者が使用する **MWUs** について、その理論的背景と先行研究を概観している。そして、「3.5. まとめ」

で第三章のまとめを行う。

第四章では、研究方法について説明している。「4.1. 研究目的と研究背景」では、改めて本研究の目的について確認を行い、「4.2. コーパスデータ」では、研究対象であるコーパスデータについて詳細な説明を記している。「4.3. 使用語彙分析の構造」では、本研究で行う4つの分析の相互関係についてまとめている。「4.4. 研究背景とリサーチクエスチョン」では、各分析のリサーチクエスチョンを述べ、「4.5. 分析方法」で各分析の分析方法を説明している。

第五章では、語彙の豊かさの分析を行った。語彙の豊かさの分析ではまず、1位から1000位までの機能語 (K1 機能語)、1位から1000位までの内容語 (K1 内容語)、1001位から2000位までの語 (K2)、2001位から3000位までの語 (K3)、それ以下の頻度の語 (Off-list) の各頻度層におけるトークンとタイプの比率について、学習者コーパスと母語話者コーパスを比較した。次に、各頻度層におけるギロー値についても2つのコーパスの比較を行った。その結果、タスクに基づくコーパスも自由会話コーパスも、語彙の多様性、語彙密度、語彙の洗練さが組み合わされているギロー値を用いた語彙の豊かさの測定法が、最も信頼性の高い測定法であることが分かった。次に、学習者と母語話者それぞれ、ギロー値と頻度層を変数としてピアソンの積立相関係数を求めたところ、K1 機能語のギロー値の値のみが異なる性質を持つことが分かった。したがって、K1 機能語のギロー値が低いか、そして他の頻度層のギロー値が高いかで、ある程度学習者の言語能力を測れることが分かった。最後にギロー値を用いて主成分分析を行い、第1主成分得点をもとに学習者と母語話者の語彙の豊かさを点数化した。その結果、例外は見られたが総じて母語話者グループが学習者グループよりも得点が高いことが分かった。そして、自由会話コーパスでは、学習者グループの間には大きな得点差が見られたためグループ分けを行い、上位グループと下位グループの特徴を調べたところ、上位グループは留学経験を持ち修士以上に在籍する学習者であり、下位グループは留学経験がなく、学部生であった。また、言語能力が高い学習者は母語話者の語彙の豊かさに近いということが分かった。

第六章「特徴語分析」ではまず、対数尤度比を用いて過剰使用語は  $G^2 \geq 50$ 、過少使用語は  $G^2 \leq -50$  として特徴語を抽出した。特徴語の傾向として、タスクに基づくコーパスも自由会話コーパスも特徴語のほとんどが高頻度語 (K1 機能語か K1 内容語) であり、フランス語の頻度の上位1000位以内のものであることが分かった。特徴語を品詞別に分類したところ、過剰使用語は名詞や代名詞が多く見られ、過少使用語は副詞や多義語であることが分かった。

第七章「コロケーション分析」では、特徴語を中心語として調整頻度20以上の左右3語のコロケートをまとめた。そして、学習者と母語話者によって特徴語がどのように使用さ

れているかについて比較を行った。その結果の一例として、タスクに基づくコーパスの過剰使用語とのコロケーション分析から、人称代名詞主語の過剰使用が明らかになった。また、否定辞 *ne* を初級学習者は保持する傾向にあることも明らかになった。さらに、依頼表現において法の選択に違いが見られた。学習者は直接法を用いるのに対して母語話者はより丁寧度が上がる条件法を使用していた。このように学習者は形式的な語を使用するのに対して、母語話者は語の省略や俗語の使用といった非形式的な語を用いるなど、語のスタイルの選択について 2 グループ間に違いが見られた。また、意味・統語的選択についても違いが見られた。過少使用語とのコロケーション分析では、人称代名詞主語の選択の差が見られた。母語話者は *on* を使用するのに対して、学習者は *nous* を使用していた。また、ディスコースマーカの過少使用や、曖昧性付加句の過少使用も明らかになった。

自由会話コーパスにおける過剰使用語とのコロケーション分析では、学習者の発話は語や句の繰り返しに特徴付けられていることが分かった。そして、タスクに基づくコーパス分析同様、発話スタイルの違いが見られた。また、レベルの違いによる特徴の差も多く見られた。たとえば、初中級学習者は **EST-CE QUE** 型の疑問文を有意に多く使用するのに対して、上級学習者は母語話者と同様にイントネーション型の疑問文を使用する傾向にあった。また否定辞についても、初中級学習者は否定辞 *ne* を保持するが、上級者は *ne* を用いない傾向にあることが分かった。コミュニケーション方略としての **MWUs** の使用では、初級学習者は日本語や英語などのフランス語以外の語を使用していたが、上級学習者は単語がうまく出てこない場合や言葉につまった際に *comment dire* という **MWU** を使用して、沈黙を回避していた。過少使用語とのコロケーション分析では、関係代名詞を含む文や強調構文などの複文の使用がレベルを問わず学習者によって過少使用されていた。また、曖昧性付加句の過少使用やディスコースマーカの過少使用も学習者のレベルを問わず顕著であった。

第八章「**N-grams** 分析」では、まず学習者は一部の **N-grams** を繰り返し使用しているのかについて数量的分析を行った。その結果、タスクに基づくコーパス分析からも自由会話コーパス分析からも、学習者の発話は頻度の高い **N-grams** に依存していることが明らかになった。次に、過剰使用 **N-grams** と過少使用 **N-grams** といった特徴的 **N-grams** について、それらの構成する語の頻度層について分析を行ったところ、過剰使用語については高頻度語 (**K1** 機能語と **K1** 内容語) がほとんどを占めており、中頻度語 (**K2**) や低頻度語 (**K3**、**Off-list**) があつたとしても、学習に困難をきたすものではないことが分かった。過少使用語については、ほとんどの構成語が高頻度語であった。特徴的 **N-grams** の質的分析を行ったところ、過剰使用語や過少使用語の影響がコーパスの種類を問わず多く見られた。その他の点として、タスクに基づくコーパスでは、動詞の選択、疑問文における型の選択、否

定辞 *ne* の保持と脱落などに母語話者との違いが見られた。また、複文の過少使用、ディスコースマーカの過少使用、エラーも学習者の特徴であった。

自由会話コーパス分析から、語や句の繰り返しの多用、否定辞 *ne* の保持、疑問文 **EST-CE QUE** 型の使用は主に初中級学習者の特徴であることが分かった。一方、複文の過少使用、順接関係を示す語の過剰使用、方略的 **N-grams** の過剰使用、同意表現の過剰使用、そしてディスコースマーカの過少使用はレベルを問わず学習者全体の特徴であることが分かった。

第九章は結論の章であり、第五章から第八章までに行った分析や考察が結論としてまとめられている。そして、最後に今後の展望や課題などについて言及している。