

中国人日本語学習者の文法能力は作文の特性にどう影響するか

楊艷萍（中南林業科技大学）・張婧禱・玉岡賀津雄（名古屋大学）

1. 研究の目的

インターネットの普及によって、文字によるコミュニケーションが頻繁に行われるようになった。村田（2018）は、ドイツ人日本語学習者の日本語の習熟度が、作文の文の長さ、誤り数などに影響することを報告している。本研究では、中国語人日本語学習者の文法能力が、作文の言語的な特性にどう影響するかを検討した。

2. リーダビリティと作文特性の関係—ピアソンの積率相関係数

中国の大学で日本語を専攻する 2 年から 4 年生までの合計 26 名に、「中国の若者が見つけた日本の新しい魅力」というテーマで 26 名が作文を書いた。年齢は、平均が 20 歳 10 カ月、標準偏差が 1 歳 2 カ月であった。3 カ月以上日本に滞在した経験はない。本研究では、彼らの作文について、リーダビリティ（読解可能指標）を「日本語文章難易度判別システム」（<https://jreadability.net/>）で調べた。リーダビリティは、 $Y = -0.056 * \text{平均文長} - 0.126 * \text{漢語率} - 0.042 * \text{和語率} - 0.145 * \text{動詞率} - 0.044 * \text{助詞率} + 11.724$ で計算される。26 名の作文の平均は 3.38 で、標準偏差は 0.28 であった。26 名の学習者は、中級レベルの作文を書いた。総文数は、35 文から 61 文で、平均の総文数は 44 文、標準偏差が 6 文であった。総延べ文字数は、1,326 文字から 1,686 文字で、平均は 1,552 文字、標準偏差は 64 文字であった。1 文の平均語数は、17 語から 30 語で、平均は 23 語、標準偏差は 3 語であった。作文の総文数と 1 文の平均語数 ($r = -.92$, $p < .001$) の相関は、極めて高く、逆相関であった。つまり、文数が多い作文は、1 文当たりの語数が少ない。文数の多い作文は、語彙数の少ないシンプルな文で書かれている。リーダビリティと総文数 ($r = .66$, $p < .001$) は高い正の相関であるが、リーダビリティと 1 文あたりの語数 ($r = -.68$, $p < .001$) は高い負の相関であった。文数が多い作文はリーダビリティの得点が高く、難易度の低い。1 文当たりの語数が多い作文はリーダビリティの得点が低く、難易度の高い作文である傾向があることが窺える。総文数が多く、語数が少ないシンプルな文で書かれているほうが、リーダビリティが高く、分かり易い作文になっていることが分かる。

3. 文法能力—ステップワイズ法による重回帰分析

26 名の学習者の文法テスト（早川・玉岡 2015）の得点は、最も得点の低い学習者は 11 点で、最も得点の高い学習者は 29 点であった。26 名の学習者の平均は 19.85 点

で、標準偏差が 4.72 点であった。11 種類の品詞については、総語彙数で各品詞の頻度を割った頻度比率を計算した。そして、文法テストの (1) 形態素変化、(2) 局所依存、(3) 構造の複雑性の 3 つの下位カテゴリーで各品詞比率を予測する重回帰分析を行った。その結果、動詞 [$R^2=0.20$, $\beta=0.45$, $t=2.48$, $p<.05$]、副語 [$R^2=0.30$, $\beta=0.54$, $t=3.17$, $p<.01$]、助動詞 [$R^2=0.31$, $\beta=0.55$, $t=3.26$, $p<.01$] の頻度比率について、構造の複雑性のみが有意な予測変数となった。文全体の統語規則の理解が、述部 (predicate) を構成する主要な動詞、助動詞および副詞の頻度比率を予測した。次に、普通名詞 [$R^2=0.23$, $\beta=-0.47$, $t=-2.64$, $p<.05$]、形容詞 [$R^2=0.23$, $\beta=0.48$, $t=2.65$, $p<.05$] の頻度比率には、局所依存のみが有意な予測変数となった。語の並び方の規則である局所依存の理解が、普通名詞とそれを修飾する形容詞の頻度比率を予測した。文法能力と品詞の頻度比率の関係が密であることを示した。

4. 日本語学習者 26 名の作文特性についての検討—クラスタ分析と判別分析

重回帰分析で文法能力が有意な予測変数となった動詞、副詞、助動詞、普通名詞、形容詞の 5 つの頻度比率を使って、階層的クラスタ分析を行った。クラスタ間の距離はウォード法、日本語学習者間の距離は平方ユークリッド距離を使った。その結果、25 ポイント中の 8 ポイントで 4 つのクラスタが得られた。これら 4 つのクラスタの分類の精度を判別分析で調べた結果、本データのグループ化について 100% が正しく判別された。また、交差妥当化の結果、96.2% の判別率であった。2 つの判別関数が有意に 4 つのグループを予測した。第 1 正順判別関数 (#1) は、64.4% を説明し、有意であった [Wilk's $\lambda=0.03$, $\chi^2(15)=73.72$, $p<.001$]。第 2 正順判別関数 (#2) は、32.9% (累積は 97.3%) を説明し、やはり有意であった [Wilk's $\lambda=0.19$, $\chi^2(8)=33.62$, $p<.001$]。構造行列から各変数と判別関数間の最大絶対相関が有意であった変数をみると、4 つのグループ化について普通名詞が最も強い負の方向の原因帰属 (#1=-0.75) を示し、次に動詞 (#1=0.30)、副詞 (#1=0.29)、助動詞 (#2=0.74)。つまり、2 つの正順判別関数の係数を使って 2 次元空間で 4 つのグループの成員である日本語学習者 26 名をプロットして、作文の特性を検討した (研究会で紹介)。

引用文献

- 早川杏子・玉岡賀津雄 (2015) 「改訂版・構造分類による日本語文法知識テストの開発—中国人日本語学習者のデータによるテスト評価—」『ことばの科学』29、5-24
- 村田裕美子 (2018) 「ドイツ語を母語とする日本語学習者の作文に見られる言語的特徴：習熟度の差は産出にどう現れるのか」『第二言語としての日本語の習得研究』21、61-76
- 日本語文章難易度判別システム <<https://jreadability.net/>> (2019 年 5 月 4 日 アクセス)