

日本語学習者の発話量と言語テストの得点の関連性について

李在鎬（筑波大学），村田裕美子（ミュンヘン大学）
小林典子（元筑波大学），酒井たか子（筑波大学）

1. 研究背景

学習者コーパスとは，言語学習者の産出データを格納したデータベースのことである。一般的には，学習言語の熟達度（proficiency）の差が言語使用にどのようなバイアスを与えるかを調査する目的で使用する。そのため，学習者コーパスの開発者は何らかの方法で学習者の（学習言語に対する）熟達度を判断し，アノテーション情報として公開している。

熟達度を判断する方法としては，2つの方法が利用される。1つ目は，産出データそのものに対して，直接的に熟達度を判断する方法，2つ目は，言語テストを使用し，産出データとは独立して熟達度を判断する方法である。1つ目の方法については，ACTFL OPI（oral proficiency interview；以下 OPI）の枠組みがよく利用される。そして，2つ目の方法については，SPOT（Simple Performance-Oriented Test；以下 SPOT，詳細は小林 2014，2015 参照）がよく利用される。1つ目の方法を利用したコーパスとしては，鎌田修氏と山内博之氏による「KY コーパス」が広く知られている。2つ目の方法を利用したコーパスとしては，伊集院郁子氏が構築した「日本・韓国・台湾の大学生による日本語意見文データベース」および金澤（2014）の「YNU 書き言葉コーパス」があげられる。なお，本研究が利用する「ドイツ話者日本語学習者話し言葉コーパス（Spoken Corpus of German Learners of Japanese；以下，GLJ コーパス）」は OPI と SPOT をともに利用しているコーパスである点で，ハイブリットのデータベースと言える（OPI と SPOT の関連については岩崎 2002 参照）。

本研究では，「GLJ コーパス」を用いて，言語テスト「SPOT」が日本語学習者の発話データをどの程度，説明できるかを調査する。具体的には重回帰分析と分散分析を行い，SPOT の得点と発話量の関連を考察する。

2. データと分析方法

2.1. データ 1：GLJ コーパス

GLJ コーパスとは，村田・李（2015）によって開発されている学習者コーパスで，ドイツ語母語話者 45 名の発話データを収録した学習者コーパスである。コーパスの基本設計において，KY コーパスと同様に，OPI を用いて熟達度を判断している。

コーパスの中には，テスターと学習者による 2 者の対話データが文字化されているが，OPI の判定ルールに基づいて，初級学習者，中級学習者，上級学習者にカテゴリー化されている。各集団の学習者数およびコーパスサイズを表 1 に示す。

GLJ コーパスの特徴として，すべての学習者はテスターと対話を行ったあとに，インターネット日本語テストである「SPOT」（<http://ttbj.jp/>）を受けており，コーパスデータ（話し言葉データ）と言語テストの得点が比較できるように構成されている。

表 1. GLJ コーパスのサイズ

熟達度区分	学習者数	延べ語数*
初級学習者	15 名	67,751
中級学習者	15 名	83,107
上級学習者	15 名	95,837
総計	45 名	246,695

*延べ語数は、UniDic+MeCab の解析結果に基づいて計算

2.2. データ 2 : SPOT

本研究では、GLJ コーパスの発話量と言語テストの関連を捉えるため、SPOT の得点データを利用する。ここでは、SPOT について述べる。SPOT は、日本語の客観テストの一つであり、以下の特徴を持っている（小林 2015）。

- (1) 言語運用能力を間接的・客観的に測定するテストである。
- (2) 短時間で実施できるテストである。
- (3) 能力差が比較的大きな集団を 2～4 段階程度の能力別グループに分けるテストである。

SPOT は自然な速度で読み上げられる文を 1 文ずつ聞きながら、1 カ所の空欄に平仮名 1 文字を挿入するという形式の問題である。今回は WEB 版 SPOT を使っており、図 1 のテスト項目が音声とともにディスプレイ上に表示される。

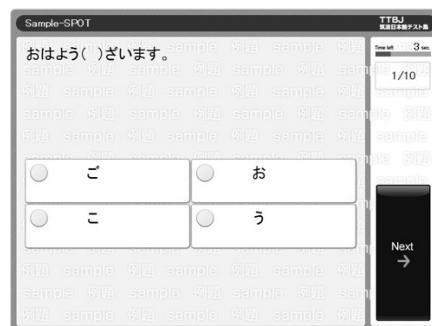


図 1. 「SPOT90」の問題提示画面

SPOT にはいくつかのバージョンがあるが、本研究では「SPOT90」を利用した。「SPOT90」は 30 問ずつの「SPOT90-1」「SPOT90-2」「SPOT90-3」で構成されており、それぞれのテストセットの難易度は次のように設定されている。まず、「SPOT90-1」は初級向けで、日本語能力試験の N4-N5 レベルの学習者を対象にしている。「SPOT90-1」の音声は声優による明瞭な

ものを使用している。次に、「SPOT90-2」は初級後半から上級前半向けで、日本語能力試験の N2-N4 レベルの学習者を対象にしている。「SPOT90-2」の音声は日本語教師のやや不明瞭なものを使用している。最後に、「SPOT90-3」は上級向けで、日本語能力試験の N1-N2 レベルの学習者を対象にしている。

2.3. 分析方法

本研究では、言語テストの成績と学習者の発話量の関連を明らかにする目的で、以下の分析を行った。

分析 1) GLJ コーパスの学習者の発話特徴量を独立変数、SPOT の得点を従属変数にして、重回帰分析を行った。

分析 2) GLJ コーパスにおける熟達度による集団間で、SPOT の得点および回答時間にどのような差があるかを分散分析で検討した。

分析 1) では SPOT の得点の違いが発話量にどのような関連を持っているか調査する。分析 2) では SPOT の得点と日本語の熟達度がどのような関連を持つか調査する。なお、学習者の発話特徴量を抽出するため、全文字データを形態素解析エンジンの MeCab 0.996 と解析辞書の UniDic 2.2.0 で解析し、李ほか (2014) で有効とされた変数を利用し、発話特徴量を抽出した。具体的には、語種別の使用頻度 (和語、漢語、外来語、混種語)、品詞の比率 (助詞率、名詞率、動詞率)、平均文長である。

3. 結果

3.1 重回帰分析の結果

重回帰分析では、「SPOT90」の合計得点を従属変数、発話特徴量を独立変数にして、ステップワイズ法で分析をしてみた。分析の結果、助詞率と平均文長による回帰モデルが得られ、高い予測力を持つことが明らかになった ($R^2=0.807$)。「SPOT90 の得点 $=-164.791+$ 助詞率 $*114.050+$ 平均文長 $*63.498$ 」の回帰式が得られた。この結果を受け、平均文長と助詞率の散布図を作成してみた (図 2)。

図 2 の数値は SPOT の得点であるが、左下に 30~40 点前後の学習者が分布しており、右上に進むにつれ、50 点、60 点、さらには 70 点前後の学習者が分布している。これにより、助詞率と平均文長が言語テストの得点に対して強い予測力を持つことが明らかになった。このことを踏まえ、2 つの変量は、言語能力の差を反映する変数である可能性があると考え、より詳細に分析してみた。

図 3 の横軸は、GLJ コーパスに付与されている学習者の熟達度であり、縦軸は変数を表す。左の図は平均文長であり、右の図は助詞率である。平均文長の場合、初級では、16 語程度の長さで構成されているが、中級では、23 語、上級では 29 語で構成されている。次に助詞率で言えば、初級は 14%、中級では 17%、上級では 20%程度が助詞である。

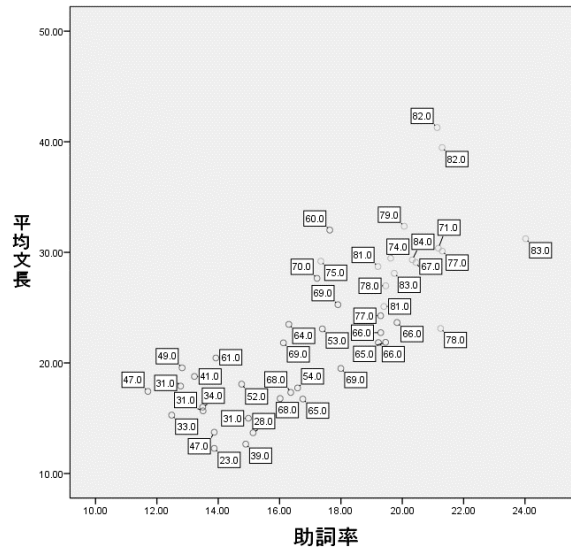


図 2. 平均文長と助詞率の散布図

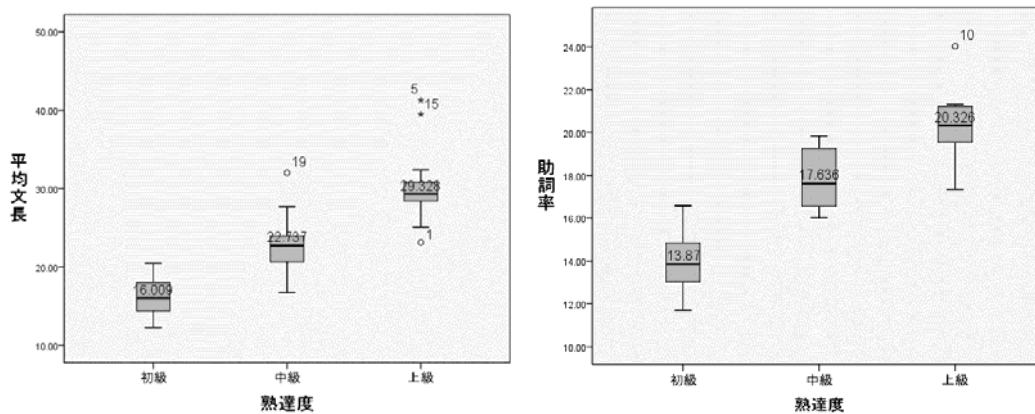


図3. 平均文長と助詞率の箱ひげ図

3.2 分散分析の結果

2つの方法で分散分析を行った。

方法1) SPOT90の得点を独立変数, GLJコーパスの熟達度を従属変数にして平均値に差があるか調査した。

方法2) SPOT90の回答時間を独立変数, GLJコーパスの熟達度を従属変数にして平均値に差があるか調査した。

方法1) による調査結果として, SPOT90-1, SPOT90-2, SPOT90-3のいずれにおいても, 有意差があった (SPOT90-1は $F(2,42)=51.422, p<.001$, SPOT90-2は $F(2,42)=74.004, p<.001$, SPOT90-3は $F(2,42)=68.505, p<.001$, SPOT90全体は $F(2,42)=99.080, p<.001$) (得点分布は図4参照)。

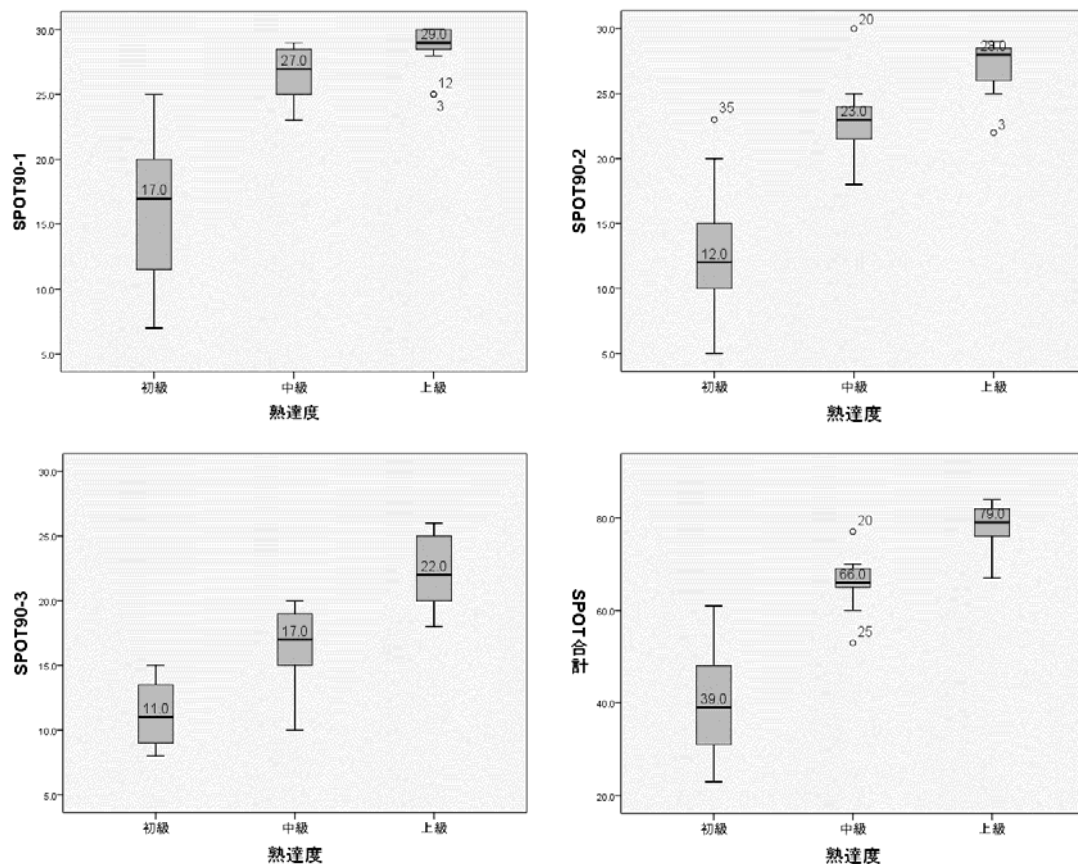


図 4. SPOT90 の箱ひげ図

初級向け SPOT である「SPOT90-1」の場合、初級においてはかなり幅があるが、中央値は 17 点であった。中級は 27 点であるが、上級は 29 点となり、中上級に関しては初級に比べ、差が小さい。中級向け SPOT である「SPOT90-2」では、初級の幅が大きいことは、SPOT90-1 と同じであったが、中級と上級の間で中央値で言えば、5 点ほど差がついた。上級向け SPOT である「SPOT90-3」では上級と中級と初級でほぼ等間隔の差がついており、能力の差を適確に捉えられていると考えられる。「SPOT90 合計」では初級の幅が非常に大きいものの、3 つの熟達度で十分な得点差が見られる。

SPOT は、即時応答的な問題形式を使って、言語運用力を測るという目的で作られたテストである。このことを検討すべく、方法 2) として、回答時間と OPI レベルの関連を調べてみた。方法 2) による調査の結果、SPOT90-1 と SPOT90-2 で OPI レベルによる回答時間の差に統計的有意が認められる結果となった (図 5)。SPOT90-3 については、有意である可能性が見られた ($F(2,42)=2.948, p=0.063$)。SPOT90-3 は SPOT 全体においてもっとも難しいものであり、非母語話者全体において時間を要している可能性がある。これについては矢崎 (2006) の考察を参考に、母語話者と非母語話者の差という視点から、今後、さらに調査する必要がある。

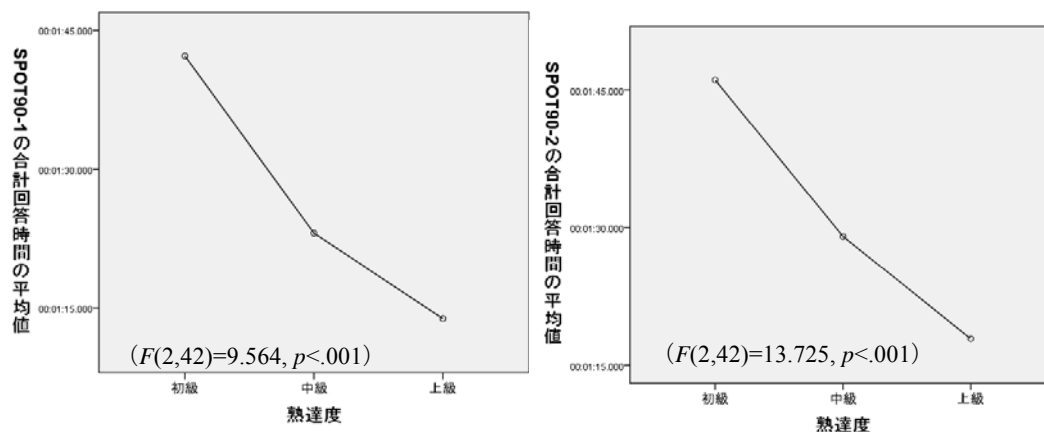


図 5. OPI レベルによる SPOT90-1 と 90-2 の回答時間

4. まとめ

発話データの産出量で SPOT の得点を予測させると 80%以上、妥当な結果が得られるため、SPOT は会話の産出力を測定するテストであると言える。また、産出能力を測った OPI レベルと言語テストの能力を示す SPOT の得点の関連を分析した結果、OPI レベルによって、すべてのテストセットの得点に統計的な有意差が観察された。そして、SPOT の回答時間においては、初級向けの SPOT90-1 と中級向けの SPOT90-2 において、 $p<.001$ 水準で差があることが明らかになった。以上の分析結果により、GLJ コーパスの収録データにおける言語能力の差を SPOT は正確に捉えていると結論づけられる。

【参考資料】

- 矢崎 彩, 李 海南, ウォンサミン スリーラット, 高橋 美野梨, 酒井 たか子, 小林 典子 (2006) 「上級日本語学習者向け SPOT(Simple Performance-Oriented Test)の開発」『日本語教育方法研究会誌』13(1), 16-17.
- 岩崎典子 (2002) 「日本語能力試験 (SPOT) の得点と ACTFL 口頭能力評価 (OPI) のレベルの関係について」『日本語教育』114 号, 100-105.
- 小林典子・フォード順子・山元啓史 (1996) 「日本語能力の新しい測定法『SPOT』」『世界の日本語教育』6 号, 201-236.
- 小林典子 (2014) 「SPOT について」『月刊日本語学』2014 年 10 月号, 42-51.
- 小林典子 (2015) 「SPOT」『日本語教育のための言語テストガイドブック』110-126, くろしお出版
- 村田裕美子・李在鎬 (2015) 「ドイツ語母語話者の話し言葉コーパスの開発」(Japanologentag 2015 - LMU München)
- 李在鎬・宮岡弥生・林炫情(2013) 「学習者コーパスと言語テストー 言語テストの得点と作文のテキスト情報量の関連性」『言語教育評価研究 (AELE)』2 号, 22-31.