

# CLD 生徒の作文における語彙と「外国人児童生徒のための JSL 対話型アセスメント DLA」〈話す〉のステージとの関係

山元一晃(金城学院大学)

## 1. はじめに

本発表では、CLD (Culturally and Linguistically Diverse) 生徒が、「外国人児童生徒のための JSL 対話型アセスメント DLA」(DLA, 文部科学省 2014)の一環として書いた作文について、語彙的な観点から考察を試みる。山元 (2022)は、DLA 〈話す〉のアセスメントによって判定されたステージと、異なり語数および語彙多様性とに一定の関係があることを確認しており、ステージが高いほど、異なり語数が多くなり、語彙多様性が高くなることを指摘している。しかし、話しことばで高い評価を得ていることが、書きことばの語彙の指標に反映されるのかは明らかになっていない。

そこで、本発表では、DLA〈話す〉のアセスメントの判定と、学習者が書いた作文の語彙の指標とに一定の関係があるかを検討し、〈話す〉ことについてのアセスメントが作文にも反映されているのかを明らかにしたい。

## 2. 先行研究

CLD 児童生徒などの産出語彙に関する研究としては、和語動詞の産出力を調査した西川・青木 (2018)、一二三 (1996)、生田 (2001)などが挙げられる。西川・青木 (2018)は、CLD 児童生徒<sup>1</sup>と日本語を母語とする子どもに対し、イラストと動詞の部分を示すマス目を提示し、適切な動詞を調査票に記入するという形式で調査を行っている。その結果、JSL の子どもは「学校生活ではあまり使用しない動詞・用法のアイテム」(西川・青木 2018: 121)を苦手とすること、「家庭内でよく経験し、家庭から言語インプットが多く与えられるであろうアイテム」(西川・青木 2018: 122)は学年が上がっても正答率があまり伸びないことを指摘している。

一二三 (1996)は、範疇語連想と語頭連想とによる調査を行っており、「在日期間が長いほど語彙数が増え、学習段階の高い語彙も習得される」(一二三 1996: 22)ことを指摘している。ブラジル人中学生の作文における語彙を調査した生田 (2001)は滞日年数が増えるにつれて語彙の多様性が増えるが、日本での滞在が 6~8 年になると、日本人生徒と差がなくなり、滞在 8~11 年には、むしろ日本人生徒よりも多様な語彙を使っていることを指摘している。一方、漢語使用率については、日本人生徒を上回ることはなかったという。

母語話者の児童生徒の語彙を調査したものには、宮城・今田 (2018)や笹島 (2019)などがあり、前者は漢字使用能力の発達過程を分析しており、後者は、語彙多様性といった観点から学年による語彙使用の違いを明らかにしている。

## 3. 使用するデータと分析方法

本発表において使用するデータは、茨城県教育委員会から研究受託を受けて行っている茨城県教育委員会グローバル・サポート事業「オンライン学習による日本語初期指導カリキュラム開発・検証に関する研究」(澤田 2021)において実施されたアセスメントの一部である。本事業では茨城県内の公立中学校に在籍する CLD 生徒に対する日本語支援を行っており、その一環として DLA を実施している。DLA には、「普段の様子や DLA の実践を通して観察された子どもの日本語能力を具現化して捉えるための枠組み」(櫻井 2022)である「JSL 評価参照枠」があり、それぞれ〈技能別〉と〈全体〉に分かれている。そのうち、本発表では DLA〈話す〉のアセスメントの結果と、DLA〈書く〉の一環として執筆した作文を使用した。本発表で用いるデータは、DLA〈話す〉のステージ 1 が 1 名、ステージ 2 が 6 名、ステージ 3 が 9 名、ステージ 4 が 1 名、ステージ 5 が 3 名、ステージ 6 が 1 名であった。作文についてはステージ 2 の 1 名を除き「学校」をテーマとして書かれた作文である。

<sup>1</sup> 西川・青木 (2018) などでは、「JSL (Japanese as a Second Language)の子ども」としており、厳密には、対象となる者が異なるが、用語の統一を図るため、本稿においては「CLD 児童生徒」とした。

DLA はそれぞれの生徒に対して複数回行っているが、同条件で比較するため、初回の1回を対象とした。手書きで書かれた作文を文字起こしし、形態素解析用辞書として「現代書きことばUniDic 3.1.1」を、形態素解析器としてMeCab 0.996を使用して短単位に形態素解析を行った。形態素解析結果から「記号」「補助記号」として解析された語を除いた上で、延べ語数、異なり語数、語彙多様性を算出した。語彙多様性については、延べ語数の影響を受けにくい修正 TTR (Carroll, 1967) を使用した。修正 TTR (CTTR) は、異なり語数を、延べ語数の2倍の平方根で除した値である。DLA<話す>のステージ判定との関係を把握するため相関比  $\eta$  を算出した。相関比  $\eta$  を用いると、1つの尺度が質的データ、もう一方の尺度が量的データである場合に、相関があるかを調べることができる(島田・野口 2017)。さらに、5名以上のデータがあるステージ(ステージ2、ステージ3)について、箱ひげ図を作成し、弁別性があるかを検証した。その際、異なるテーマで書かれた母語話者の中学生の作文のデータも参考として図に加えた。

上記の方法によって、ステージとの相関や弁別性が明確でない場合は、その他の指標についても検討した。具体的には、Lee & Hasebe (2020) で指標としているようなテキストに占める漢語の割合と平均文長について検討した。前者は、文の総延べ語数に対する漢語の割合、後者は、総延べ語数を句点(「.」)の数で除した値とした。なお、平均文長については句点の使用が極端に少ない(2以下)3名を除外している。

#### 4. 結果と考察

延べ語数、異なり語数、語彙多様性(CTTR)のいずれについても有意な結果は出なかった。延べ語数については、 $\eta = 0.39$  ( $F(5, 16) = 0.58, p = .71$  n. s.), 異なり語数については  $\eta = 0.52$  ( $F(5, 16), p = .33$  n. s.), 語彙多様性(修正 TTR)については  $\eta = 0.64$  ( $F(5, 16) = 2.27, p = .09$  n. s.)となった。山元(2022)では、DLA<話す>のステージとアセスメント場面の生徒の発話の異なり語数および語彙多様性とに強い相関があることが確認されたことと対照的である。このことから、話しことばによる評価が、作文には必ずしも反映されないことが明らかになった。話しことばから、書きことばの様相について捉えることは適切ではないといえる。分布については、図1、図2、図3に箱ひげ図を示す。延べ語数、異なり語数については、ステージが上がると増加する傾向にあるが、語彙の多様性については、その傾向が見られない。

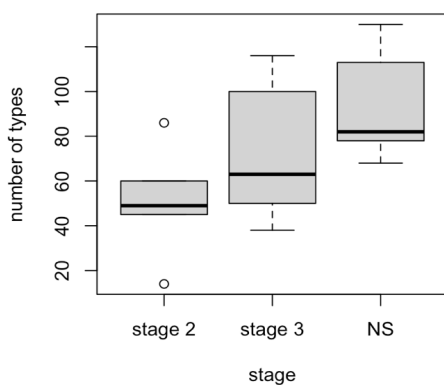


図1 DLA<話す>のステージと<書く>の延べ語数の関係

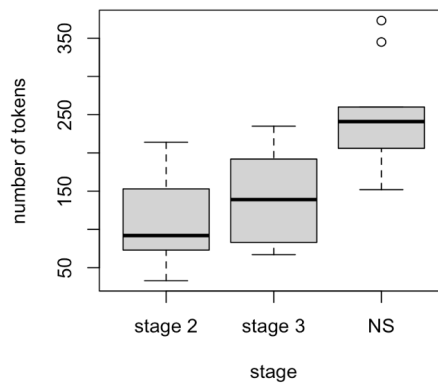


図2 DLA<話す>のステージと<書く>の異なり語数の関係

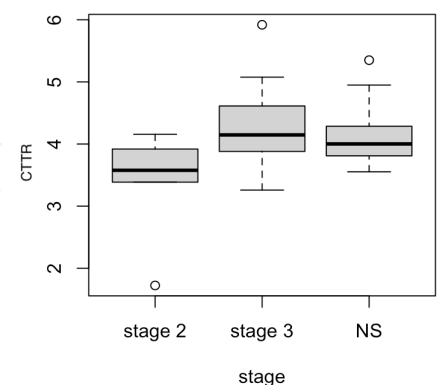


図3 DLA<話す>のステージと<書く>のCTTRの関係

図1~図3においては、母語話者(NS)のデータについても示している。延べ語数・異なり語数については母語話者で多くなっているが、語彙多様性(CTTR)が高くなることはないことが分かる。実際、母語話者を含めた場合の相関比を求めると、延べ語数については、 $\eta = 0.71$  ( $F(6, 24) = 3.50, p = .01$ ), 異なり語数については、 $\eta = 0.63$  ( $F(6, 24) = 2.64, p = .04$ )となり、いずれも相関があることが分かる。一方で、CTTRについては、 $\eta = .61$  ( $F(6, 24) = 2.33, p = .06$  n. s.)となり、母語話者を含めても有意とはならなかった。山元(2022)では、<話す>のステージが上がると、アセスメント場面におけるCTTRも高くなり、かつ、母語話者で最も高くなることを指摘しているが、そのこととは対照的な結果となった。

また、漢語率および平均文長についても検討した。漢語率については、形態素解析によって「漢語」とされた語の延べ語数を全体の延べ語数で除した値を使用した。それぞれについて、相関比を求めたところ、漢語率については、 $\eta = 0.56$  ( $F(5, 16) = 1.42, p = .27$  n. s.), 平均文長については、 $\eta = 0.56$  ( $F(4, 14) = 1.59, p = .23$  n. s.)となり、いずれについても有意な結果とはならなかった。なお、平均文長については、句点の使用が極端に少ない3名を除いている。図4および図

5に、箱ひげ図を示す。Lee & Hasebe (2020) では学習者のレベルが上がると漢語の使用率が上がり、平均文長も延びることが示されているが、いずれについても有意な結果は出ず、むしろステージ2とステージ3を比較すると、ステージ3の学習者の平均文長が下がっていた。ステージ3については、ステージ2の生徒に比較し、使用語彙が増え、一語で言えることが増えたこと、一方で、複文などは回避されたのではないかと考えられるが、この点については検証が必要である。図1～3と同様に、図4、図5にも母語話者(NS)のデータを示した。漢語率については、母語話者ではむしろ低くなるのが分かる。このことは、漢語使用率が日本人生徒と上回ることはなかったという生田(1996)の指摘と相反する結果である。なお、母語話者を含めた場合の相関比は、 $\eta = 0.60$  ( $F(6, 24) = 2.29, p = 0.07$  n. s.) であり、有意でなかった。一方で、平均文長については、母語話者で顕著に長くなっていることが分かる。実際、母語話者を含めて相関比を求めると、 $\eta = 0.71$  ( $F(5, 22) = 4.47, p = .005$ ) となり、強い相関があることが分かる。

DLAにおいては、〈話す〉と〈書く〉について、それぞれの別に評価が行われる。そのことの適切性を裏付ける結果になったといえる。

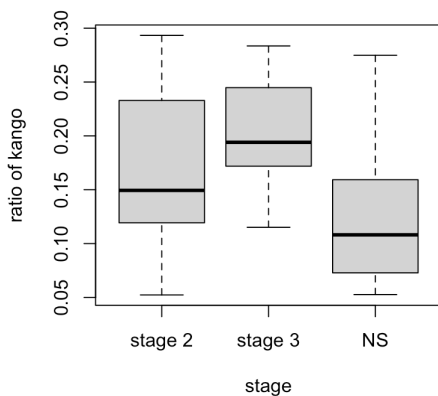


図4 DLA〈話す〉のステージと  
〈書く〉の漢語率の関係

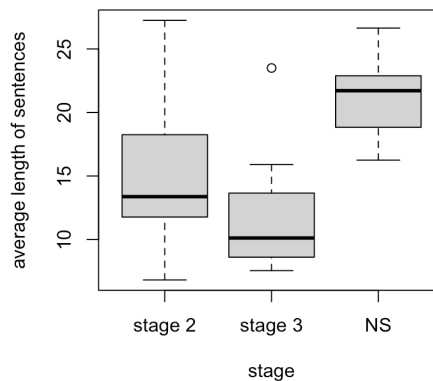


図5 DLA〈話す〉のステージと  
〈書く〉の平均文長の関係

## 5. まとめ

本発表で得られた知見をまとめると以下の通りとなる。

- 1) 延べ語数、異なり語数、CTTR のいずれの指標についても、DLA〈話す〉のステージとの相関について有意な結果は得られなかった。CTTRについては、ステージが上がってもほとんど変化がなかった。
- 2) 平均文長、漢語率についても、DLA〈話す〉のステージとの相関について有意な結果は得られなかった。
- 3) 母語話者を含めて相関を検証した場合、「延べ語数」「異なり語数」「平均文長」については、相関があることが分かった。

1)および2)について、DLA〈話す〉のステージ判定が、学習者の日本語能力と関連していると考えられるさまざまな作文に関する指標には、少なくとも本研究で対象とした学習者の産出したものについては、反映されないことが分かった。先述したようにDLAにおいては、作文についてもDLA〈書く〉として評価をしている。今後は、DLA〈書く〉のステージ判定についても作文に反映されているかを確認していきたい。また、〈話す〉のアセスメント場面において多様な語彙を使えることが、〈書く〉においても反映されているかを確認していきたい。

また、3)について、母語話者は、CLD 生徒と比較し、相対的に、作文を長く書くことができ、かつ、一文も長いことが分かった。その要因については、今後、それぞれの作文を具体的に調べていく必要がある。

西川・青木(2018)は、範疇語連想と語頭連想とによる調査を行った一二三(1996)や作文についての調査を行った生田(2001)について、「思いついた語の産出や与えられたテーマについて自由に書く作文であるため“知らない語”が回避された可能性がある」(西川・青木 2018: 17)ことを指摘している。そのため、動詞について検討した西川・青木(2018)やNishikawa(2022)ではコロケーションテストを作成して、来日からの年数や母語など、複数の観点から、相対的に苦手だと考えられる語を抽出している。西川・青木(2018)の指摘については、本研究にも当てはまっている可能性があり、本発表の限界であるともいえる。

謝辞 本研究は JSPS 科研費 JP22H00666 の助成を受けたものである。

## 参考文献

- Carrol, J. B. (1967). On Sampling from a Lognormal Model of Word Frequency Distribution. *Computational Analysis of Present-Day American English*, 406-424.
- Lee, J.-H., & Hasebe, Y. (2020). Quantitative Analysis of JFL Learners' Writing Abilities and the Development of a Computational System to Estimate Writing Proficiency. *Learner Corpus Studies in Asia and the World*, 5. <https://doi.org/10.24546/81012493>
- Nishikawa, T. (2022). Non-nativelike outcome of naturalistic child L2 acquisition of Japanese: The case of noun-verb collocations. *International Review of Applied Linguistics in Language Teaching*, 60(2), 287-314. <https://doi.org/10.1515/iral-2018-0292>
- 生田裕子 (2001). ブラジル人中学生の語彙の発達-作文のタスクを通して. *日本語教育*, 110, 120-129.
- 笹島眞実 (2019). 児童作文における使用語彙と文章様式からみたその量的特徴. *学芸国語国文学*, 51, 259-244. [https://doi.org/10.24672/gkokugokokubun.51.0\\_259](https://doi.org/10.24672/gkokugokokubun.51.0_259)
- 櫻井千穂 (2022). 子どもの日本語力を評価する. 西川朋美 (編著) 外国につながる子どもの日本語教育 くろしお出版 pp. 43-59.
- 澤田浩子 (2021). 茨城県教育委員会グローバル・サポート事業「オンライン学習による日本語初期指導カリキュラム開発・検証に関する研究」令和2年度実施のご報告. 筑波大学 人文社会系リサーチグループ「多文化的背景を持つ児童生徒教育のための研究グループ」.
- 島田めぐみ・野口裕之 (2017). 日本語教育のためのはじめての統計分析 ひつじ書房
- 西川朋美・青木由佳 (2018). 日本で生まれ育つ外国人の日本語力の盲点-簡単な和語動詞での隠れたつまづき ひつじ書房
- 一二三朋子 (1996). 年少者の語彙習得過程と言語使用状況に関する考察-在日ベトナム 人子弟の場合. *日本語教育*, 90, 13-24.
- 宮城信・今田水穂 (2018). 『児童・生徒作文コーパス』を用いた 漢字使用能力の発達過程の分析. *計量国語学*, 31(5), 352-369. [https://doi.org/10.24701/mathling.31.5\\_352](https://doi.org/10.24701/mathling.31.5_352)
- 文部科学省 (2014). 外国人児童生徒のための JSL 対話型アセスメント. 文部科学省初等中等教育局国際教育課.
- 山元一晃 (2022). CLD 生徒に対するアセスメント場面における生徒の発話の語彙について. 第66回計量国語学会大会予稿集, 49-53.