

目的

- ・ 学習者コーパスICNALEを用いて、複数の言語的特徴のなかからCEFRレベル基準特性となる指標を抽出する
- ・ 抽出された指標とCEFRレベルとの相関関係を明らかにする

CEFR

Common European Framework of Reference for Languages

欧州評議会によって、外国語の学習、教育、評価に共通の基盤を提供するために制定された枠組み

学習者がどのような言語知識と言語運用能力を備えていなければならないのかを、6つの習熟度レベル(A1、A2、B1、B2、C1、C2)で記述

基準特性

CEFRの各習熟度レベルを特徴付け、下位レベルから上位レベルを区別する学習者言語の特性

ICNALE

International Corpus Network of Asian Learner of English

英語母語話者およびアジア10か国(地域)の学生の、英作文(200~300語指定)と英語発話(60秒間)を大規模に収集したコーパス

総語数：書き言葉 約130万語 話し言葉 約50万語

学習者のCEFRレベルは、標準テストのスコアに基づき、A2_0(A2以下)、B1_1(B1下級)、B1_2(B1上級)、B2_0(B2以上)に分けられる

テストスコアとCEFRレベルの対応表

Levels	TOEIC	TOEFL PBT	TOEFL IBT	IELTS
A2_0	-545	-486	-56	3+
B1_1	550+	487+	57+	4+
B1_2	670+	527+	72+	4+
B2_0	785+	567+	87+	5 (5.5)+

データ数

	A2_0	B1_1	B1_2	B2_0
書き言葉	960	1904	1872	464
話し言葉	400	884	1876	640

研究手法 1

- ① 研究対象とする言語的指標として、流暢さの指標3つ(総語数、文の数、平均文長)、複雑さの指標1つ(従属節の数)、語彙の豊富さの指標1つ(GI)を選択
- ② ICNALEの書き言葉データ、話し言葉データをダウンロードし、各データの総語数、異なり語数、文の数、従属節の数を求める
 - ・ 総語数、異なり語数、文の数… WordSmithのWord Listを使用
 - ・ 従属節の数… Stanford Parserを用いて構文解析情報を付与 → AntConcを用いて“(SBAR)”の数をカウント
- ③ 各指標のCEFRレベルごとの平均値を求める
 - ・ 総語数：1作文/発話当たりの総語数の平均
 - ・ 文の数：1作文/発話当たりの文の数の平均
 - ・ 平均文長：1作文/発話における平均文長(総語数÷文の数)の平均
 - ・ 従属節の数：1作文/発話当たりの従属節の数
 - ・ GI：1作文/発話におけるGI(総語数÷√異なり語数)の平均

研究手法 2

- ① テストスコアとCEFRレベルの対応表に基づき、各CEFRレベルをTOEICのスコアに換算

CEFRレベル	A2_0	B1_1	B1_2	B2_0
TOEICスコア	387.5	610.0	727.5	865.0
- ② 各指標と換算スコアとの相関係数を求める

考察

- ・ 書き言葉
 - 基準特性として有効な指標は、総語数、平均文長、GI
 - 換算スコアとの相関係数が最も高い指標はGI
 - 上級学習者になるほど文の数が減少(増加すると予想)
 - ←データ収集の際に語数の制限があったためか
- ・ 話し言葉
 - 基準特性として有効な指標は、平均文長、従属節の数、GI
 - 換算スコアとの相関係数が最も高い指標は従属節の数

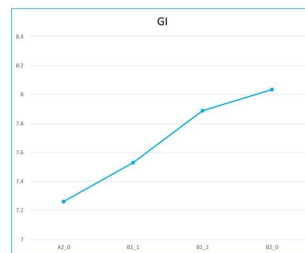
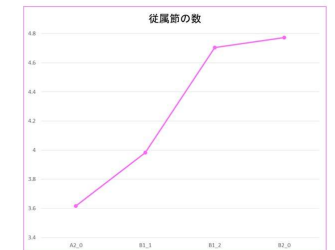
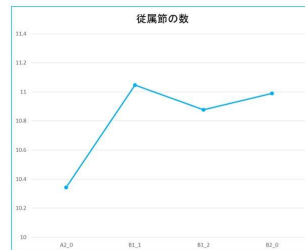
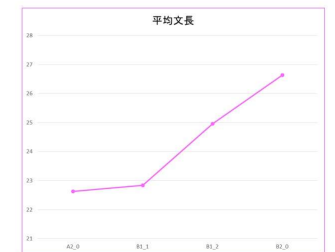
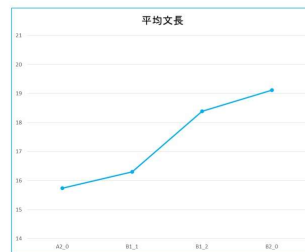
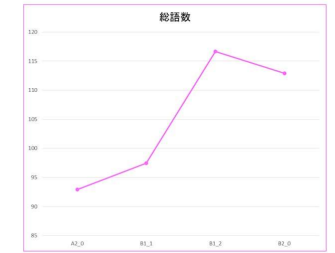
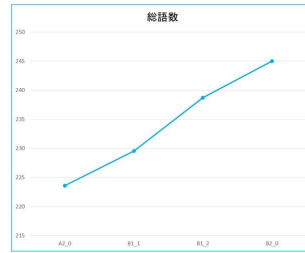
結果 1

書き言葉

	A2_0	B1_1	B1_2	B2_0
総語数	223.62	229.57	238.69	245.01
文の数	15.19	15.04	13.91	13.63
平均文長	15.74	16.30	18.39	19.11
従属節の数	10.34	11.05	10.88	10.99
GI	7.26	7.53	7.89	8.04

話し言葉

	A2_0	B1_1	B1_2	B2_0
総語数	92.92	97.43	116.67	112.93
文の数	4.26	4.46	4.95	4.31
平均文長	22.63	22.84	24.96	26.64
従属節の数	3.62	3.98	4.70	4.77
GI	5.14	5.21	5.79	5.85



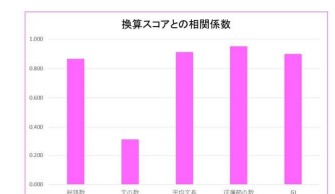
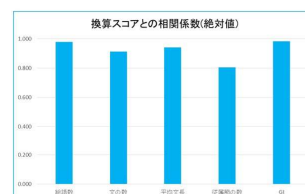
結果 2

書き言葉

指標	総語数	文の数	平均文長	従属節の数	GI
相関係数	0.978	-0.913	0.941	0.804	0.982

話し言葉

指標	総語数	文の数	平均文長	従属節の数	GI
相関係数	0.866	0.307	0.912	0.950	0.897



参考文献

- 投野由紀夫編 (2015) 『英語コーパス研究シリーズ 第2巻 コーパスと英語教育』 ひつじ書房
- 投野由紀夫・金子朝子・杉浦正利・和泉絵美編著 (2013) 『英語学習者コーパス活用ハンドブック』 大修館書店