博士論文報告

胡　晓琳

フ シャオリン

HU XIAOLIN

複雑性、正確性、流暢性の観点からの
第2言語としての英語基準特性の研究

Criterial Features in L2 English, Based on Complexity, Accuracy and Fluency

Table 1. CEFR levels

Criterial Features of CEFR

Proficient Users	C2	Mastery
	C1	Effective Operational Proficiency
Independent Users	B2	Vantage
	B1	Threshold
Basic Users	A2	Waystage
	A1	Breakthrough

Council of Europe's 2001 document *Common European Framework of Reference for language learning, teaching, and assessment*

The list of Can Do descriptors

ALTE Level	Council of Europe Levels	Listening/Speaking
ALTE Level 5	C2	CAN advise on or talk about complex or sensitive issues, understanding colloquial references and dealing confidently with hostile questions.
ALTE Level 4	C1	CAN contribute effectively to meetings and seminars within own area of work or keep up a casual conversation with a good degree of fluency, coping with abstract expressions.
ALTE Level 3	B2	CAN follow or give a talk on a familiar topic or keep up a conversation on a fairly wide range of topics.
ALTE Level 2	B1	CAN express opinions on abstract/ cultural matters in a limited way or offer advice within a known area, and understand instructions or public announcements.
ALTE Level 1	A2	CAN express simple opinions or requirements in a familiar context.
ALTE Level Break-through Level	A1	CAN understand basic instructions or take part in a basic factual conversation on a predictable topic.

Table 1. CEFR levels

CAF measures:

Complexity: the extent to which the language produced in performing a task is elaborate and varied' (Ellis 2003: 340)

Accuracy: the ability to produce error-free speech (Lennon 1990: 390)

Fluency: the ability to process the L2 with 'native-like rapidity' (Lennon 1990: 390) or 'the extent to which the language produced in performing a task manifests pausing, hesitation, or reformulation' (Ellis2003: 342)

Example of CAF Features

Criteria	Sub-measures		Descriptions
Fluency	Syllables per second		The mean number of syllables produced per second
	Mean length of utterance		The average number of syllables produced between pauses of 0.1 secs or above
	Phonation time ratio		The percentage of time spent speaking as a proportion of the total time
	Mean length of silent pauses		The total length of silent pause time divided by the number of silent pauses
	Mean length of filled pauses		The total length of filled pauses divided by the number of filled pauses (filled pause: e.g., uh, um, well, like, you know)
Accuracy	Grammatical Accuracy	Error free per t-unit	The ratio of the number of error-free t-units
		Specific types of errors	Articles, preposition, singular/plural, subject-verb agreement, and tense
Complexity	Grammatical Complexity	T-unit complexity ratio	The total number of clauses per T-unit
	Lexical Complexity	Lexical diversity-	Type/token ratio
		Lexical richness	Proportion of k1, k2, awl words
		Word length	Number of letters per word

Purpose of Research:

1. to explore whether and to what extent different CAF features are associated with CEFR proficiency levels in speaking performance;

2. to examine what CAF features are best predictors to distinguish L2 learners' oral performance at different CEFR levels;

3. by observing the CAF components' correlational effects on each other, to explore whether trade-off effects are inevitable at different CEFR levels.

Method:

Data Base:

NICT JLE Corpus:

L2 Learners: CEFR Level A1, A2, B1 (SST level 2-8)

20 native speakers

CEFR-J	SST
PreA1	1
A1.1	2/3
A1.2	3
A1.3	4
A2.1	4
A2.2	5
B1.1	6/7
B1.2	8
B2.1-C2	9

CEFR-J	SST	
PreA1	1	
A1.1	2	A1
A1.2	3	
A1.3	3	
A2.1	4	A2
A2.2	5	
B1.1	6/7	B1
B1.2	8	
B2.1-C2	9	

（学習者コーパスによる英語 CEFR レベル基準特性の 特定と活用に関する総合的研究，投野 由紀夫，2016）

A Validation Study of the Accuracy of Lexical Diversity Tools

Research Question

1. To what extent can tools accurately measure the lexical diversity features?

2. What factors might affect the outputs of different tools?

Method:

Data set:

24 transcriptions from NICT JLE Corpus

CEFR Levels	A1	A2	B1	Native Speaker(USA)
Number	6	6	6	6

6 tools:
- Text Inspector (2020)
- VocabProfile Program (Cobb, 2002)
- Coh-Metrix (Graesser, 2004)
- Lexical Complexity Analyzer (Lu, 2012)
- CLAN (MacWhinney, 2000)
- TAALED (Kyle & Crossley, 2015).

Gold Standard:
- TTR (type/token ratio)
- Guiraud' s index

To divide the tools into 2 Groups:

Group 1 :
- Text Inspector (2020)
- Coh-Metrix (Graesser, 2004)
- VocabProfile Program (Cobb, 2002)
- CLAN0 (MacWhinney, 2000)

No lemmatization
Lemma 0

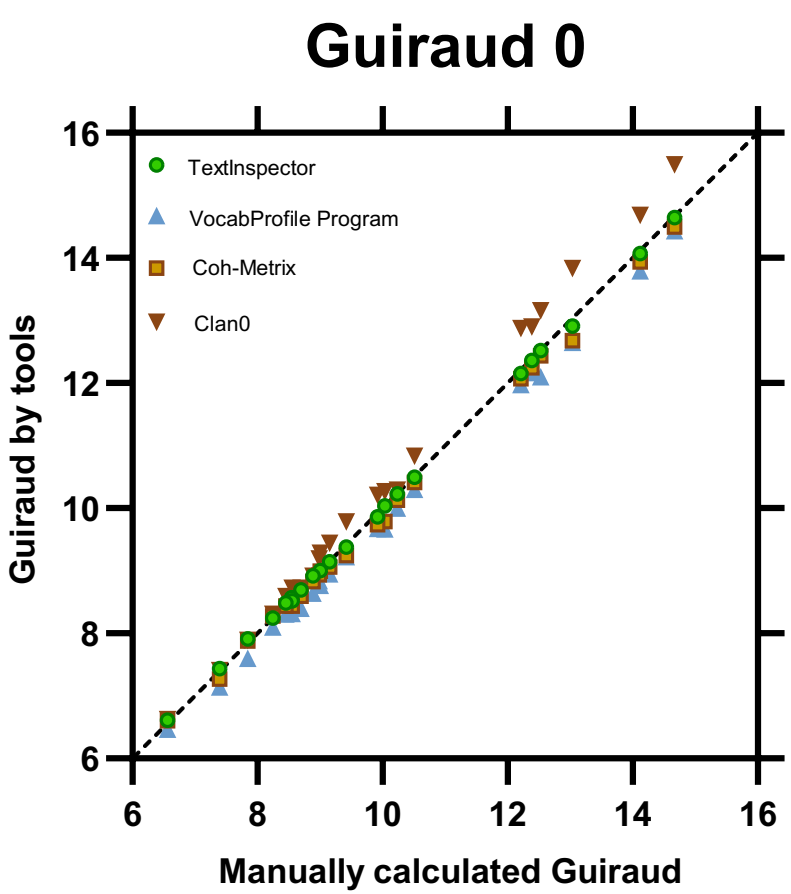
Group 2 :
- CLAN 1 (MacWhinney, 2000)
- Lexical Complexity Analyzer (Lu, 2012)
- TAALED (Kyle & Crossley, 2015).

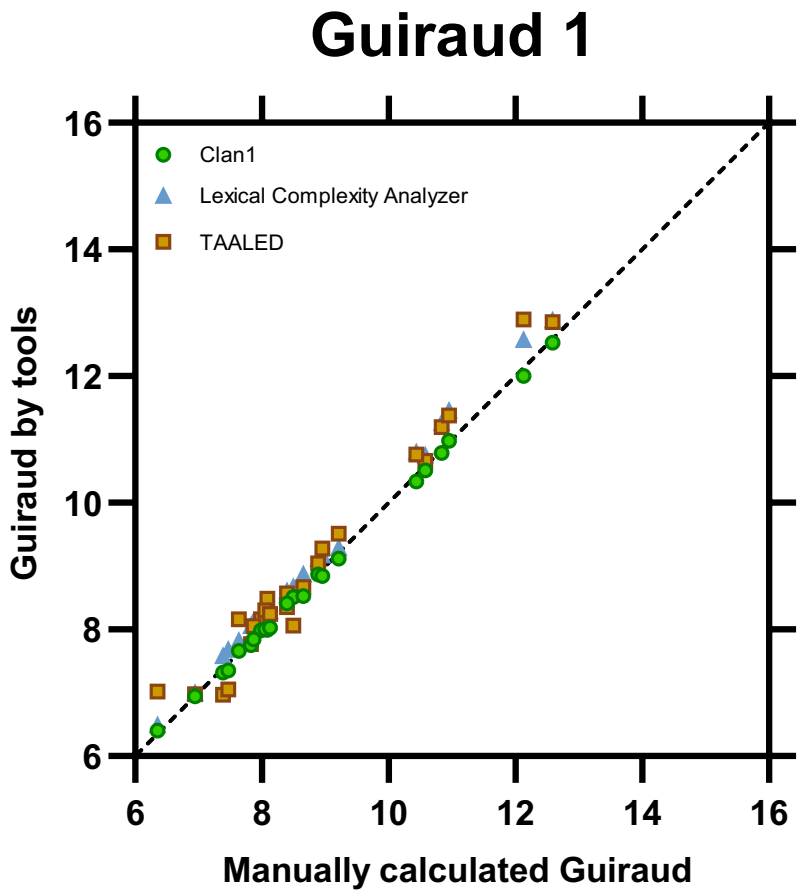
With lemmatization
Lemma 1

The recognition and calculation of tokens in different tools

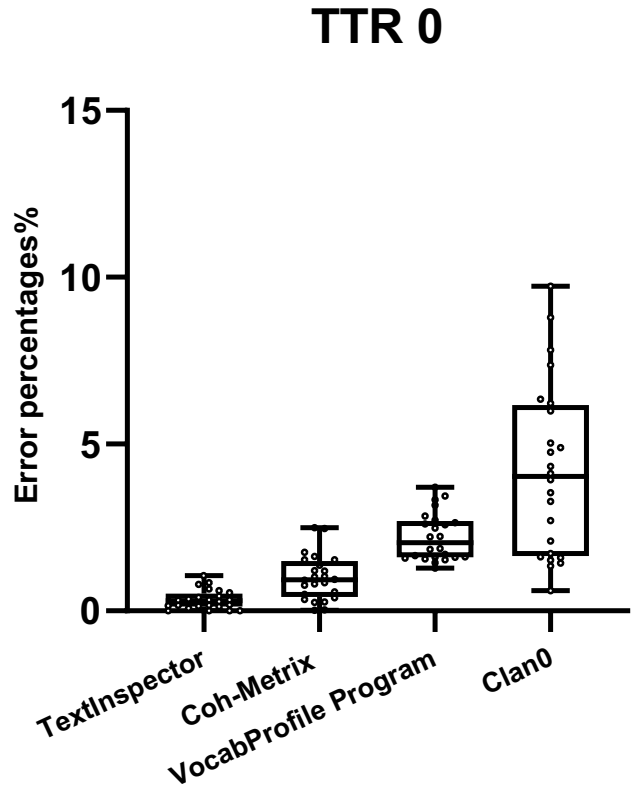
		Manual Counting	Text inspector	Coh-Metrix	VocabProfile Program	CLAN		TAALED	LCA	Manual Counting
Lemmatization principle		Lemma 0	Lemma 0	(Lemma 0)	Lemma 0	Lemma 0	Lemma 1	Lemma 1	Lemma 1	Lemma 1
don't		2 (do/n't)	2 (do/n't)	2 (do/n't)	2 (do/not)	1 (don't)	2 (do/not)	1 (do)	2 (do/not)	2 (do/not)
I'm		2 (I/ 'm)	2 (I/ 'm)	2 (I/ 'm)	2 (I/ am)	1 (I'm)	2 (I/ be)	1 (I)	2 (I/ be)	2 (I/ be)
's	he's	2 (he/ 's1)	2 (he/ 's)	2 (he/ 's)	2 (he/ is)	1 (he's)	2 (he/be)	1 (he)	2 (he/be)	2 (he/ be)
	mom's(+n)	2 (mom/ 's2)	2 (mom/ 's)	2 (mom/ 's)	1 (mom)	1 (mom's)	1 (mom)	1 (mom)	1 (mom)	2 (mom/ 's)
	let's	2 (let/ 's3)	2 (let/ 's)	2 (let/ 's)	1 (let)	1 (let's)	2 (let/us)	1 (let)	2 (let/us)	2 (let/ us)
	it's been	3 (it/ 's4/ been)	3 (it/ 's/ been)	3 (it/ 's/ been)	3 (it/is/been)	2 (it's/been)	2 (it/ be)	2 (it/ be)	2 (it/ be)	3 (it/have/ be)
twenty-four		2	2	1	2	1	1	2	1	2
o'clock		1	1	2	2	1	1	1	2	1
fragments	c	1	1	1	0	1	1	0	1	1
	co	1	1	1	1	1	1	0	1	1
big/ biggest		2	2	2	2	2	1	1	2	1
XXX01/ XXX02		2	2	2	1 (XXXNumber)	2	2	0	2	2

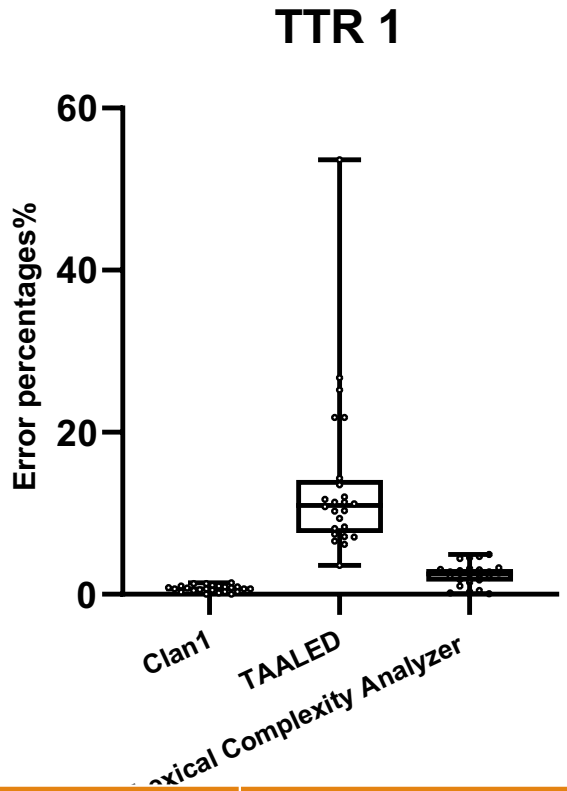
Result:


Guiraud 0
TextInspector
VocabProfile Program
Coh-Metrix
Clan0
Guiraud by tools
Manually calculated Guiraud
6
8
10
12
14
16


Guiraud 1
Clan1
Lexical Complexity Analyzer
TAALED
Guiraud by tools
Manually calculated Guiraud
6
8
10
12
14
16

Result:

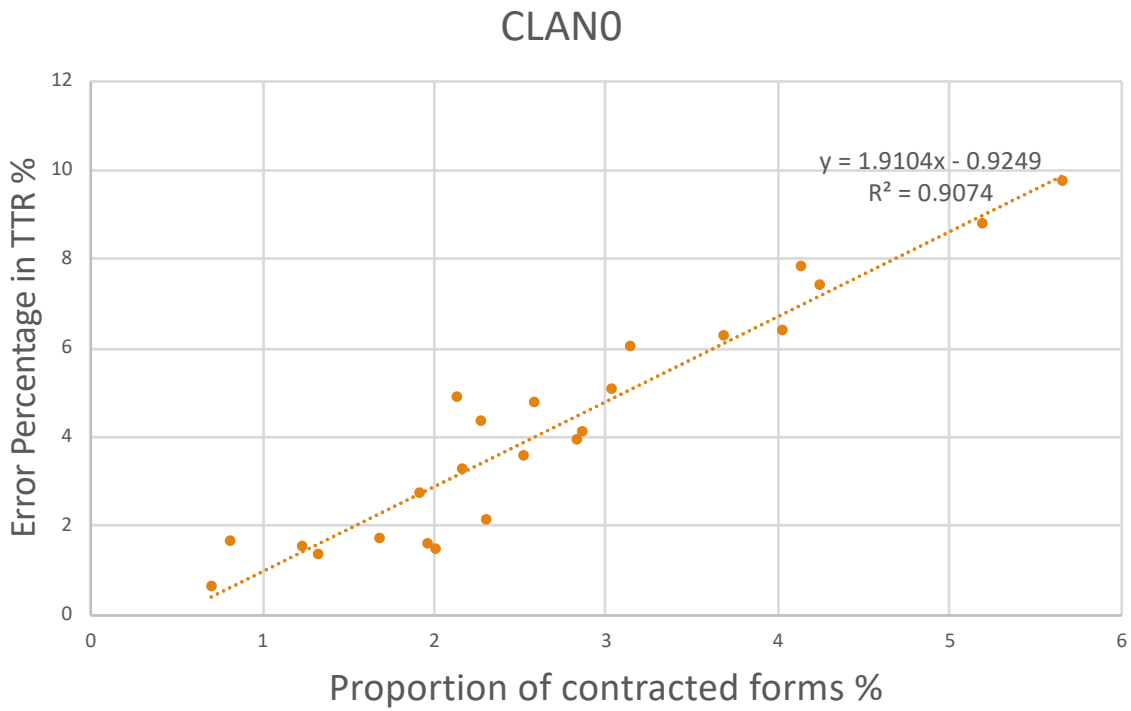




	Group 1				Group 2		
Error percentage of TTR	Text Inspector	Coh-Metrix	VocabProfile	Clan0	Clan1	TAALED	LCA
Mean (SD)	0.33% (0.29%)	1.00%(0.68%)	2.22% (0.72%)	4.20% (2.57%)	0.70% (0.36%)	13.74% (13.83%)	2.43% (0.97%)

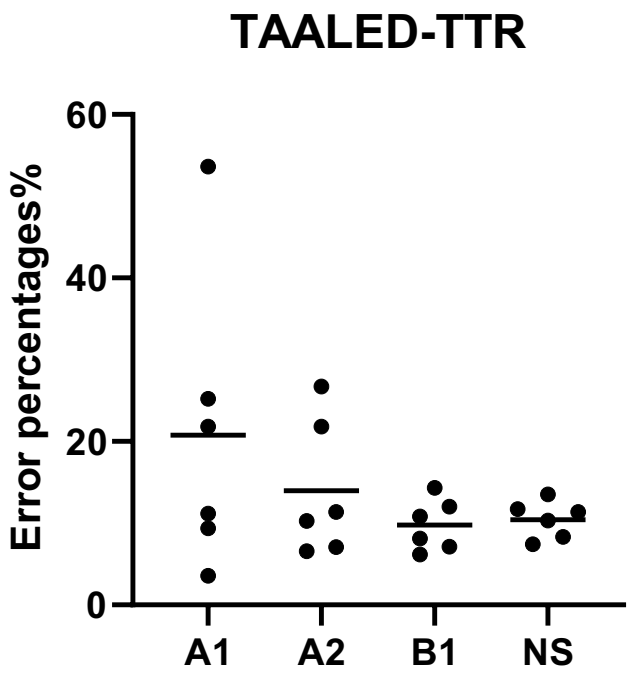
The Correlation between

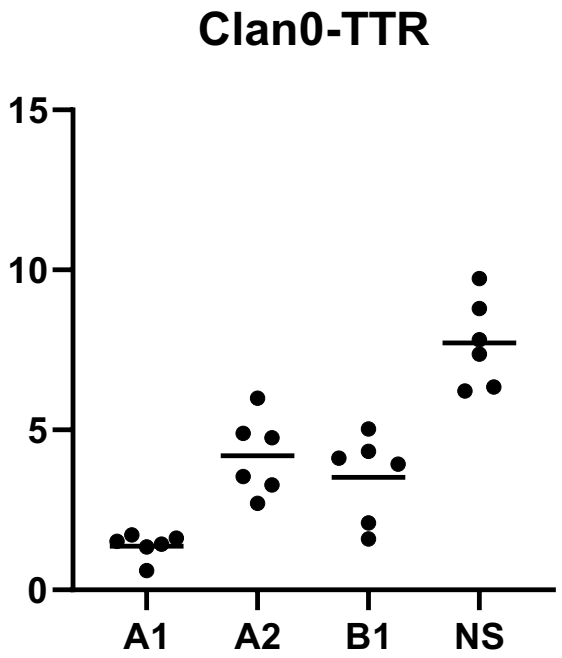
the Proportion of Contracted form in texts and Error Percentage of TTR:



The Error percentage of TTR

in different CRFR Levels





Conclusion

Tools for lexical diversity should be selected carefully according to different data types and research aims.

Error percentage in TTR < 2.5% :

No lemmatization
- Text Inspector (2020)
- Coh-Metrix (Graesser, 2004)
- VocabProfile Program (Cobb, 2002)

With lemmatization
- CLAN 1 (MacWhinney, 2000)
- Lexical Complexity Analyzer (Lu, 2012)

Thank you !