

学習者スピーキングデータにおける 基準特性の候補

—complexity, fluency, accuracyに着目して

金子 恵美子(会津大学)

和泉 絵美(京都外国語大学)

CEFR-J/SST対応表

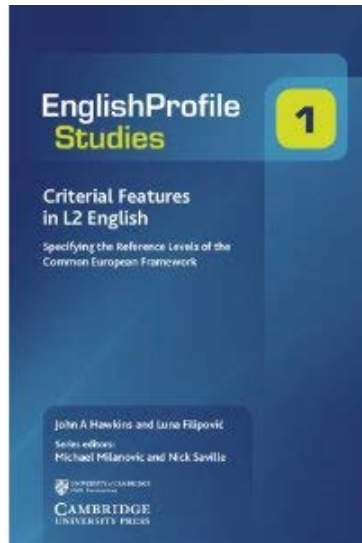
Table 1: CEFR-J/SST対応表 Ver. 1

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

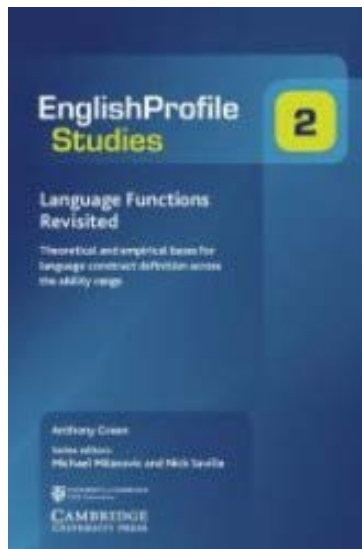
Ordering of the SST and CEFR-J levels

PreA1→**SST1**→**SST2**→A1.1→**SST3**/A1.2→A1.3→A2.1/**SST4**→A2.2/**SST5**→**SST6**→B1.1→**SST7**→B1.2→**SST8**→B2.1/**SST9**→B2.2→C1→C2

English Profile Programmeの報告書



Criterial Features in L2 English: Specifying the Reference Levels of the Common European Framework (English Profile Studies)



Language Functions Revisited : Theoretical and empirical bases for language construct definition across the ability of range (English Profile Studies Vol.2)

criterial features from the English Profile Project (syntax)

Table 2:

Level	frame	example
A2	NP-V <i>Ving</i>	his hair needs combing
	NP-V	he went
	NP-V (reciprocal sub)	they met
	NP-V-NP	he loved her
	NP-V-NP-PP (P= <i>for</i>)	she bought a book for him
	NP-V-NP-PP	she added the flowers to the bouquet
	NP-V-PP (P= <i>loc/to</i>)	they apologized to him
	NP-V-S	they thought that he ...
	NP-V-NP <i>inf</i> (sub control)	I wanted to come
	B1	NP-V-S (<i>Wh</i> -movement)
NP-V-VP <i>infin</i> (<i>Wh</i> -movement)		he explained how to do it
NP-V-NP-NP		she asked him his name
NP-V-NP-PP (P= <i>to</i>)		he gave a big kiss to his mother
NP-V-NP-(<i>to be</i>)-NP		I found him (<i>to be</i>) [a good doctor]
NP-V-NP-V <i>pastpart</i> (V= <i>passive</i>)		he wanted the children found
NP-V-P-Ving-NP		they failed in attempting the climb
NP-V-P-S		he thought about whether he wanted to go
NP-V-Part		she gave up

excerpt from p. 88

Vocabulary (numerical measures)

- Type Token Ratio (TTR)
 - Index of Guiraud (G)
 - Mean Segmental Type-Token Ratio (MSTTR)
 - Index of lexical diversity (D) $TTR = \frac{D}{N} \left[\left(1 + 2 \frac{N}{D} \right) - 1 \right]$
 - Measure of textual lexical diversity (MTLD)
- MTLD was least affected by text length, and that it should be used with texts of at least 100 tokens. (Koizumi, in print)
 - None of them makes distinctions between the lexical items. (Daller, van Hout, & Treffers-Daller, 2003)

Criterial features from the English Profile Project (vocabulary)

- 現時点では動詞と句動詞
- primary meaning → extended meaningの順で使われるようになる。
- extended meaningの中にも
 - simple/complex
 - figurative/metaphorical

のような、下位分類があるだろう。

Table 3:

	A1	A2	B1	B2	C1	C2
paint	[P] I like to <i>paint</i> (intr)	[P] <i>paint</i> my bedroom (tr)	same	[E] Her kindness is always <i>painted</i> at her wrinkled face	[F] <i>paint</i> the town red	[E] <i>paint</i> a sad picture of a future world

[P]: primary meaning

[E]: extended meaning

[F]: Fixed expression

excerpt from p. 136-142

Problems

- 書きことばコーパス (Cambridge Learner Corpus: 45 million words of learner data taken from Cambridge ESOL examination)に基づく
- A1が少ない
- A1からB2のSemantic criterial featureに関しては、English Profile Wordlists (Capel, 2010)を使用 (p. 49)
 - The Wordlists data take into account what is generally known at a given level regardless of how well a particular candidate may have performed and been graded
- SST4 (CEFR-J A2.1)以下はやっと文章が作れる程度 (SST3, 4の受験者で6割近くを占める)

スピーキングを細かく評定する

Complexity, Accuracy, Fluency (CAF)

Complexity

- unitの長さ (T-unit, AS-unitなど)
- 文法的複雑さ (clause/unit, non-finite verb/unit)

Accuracy

- 間違いのある・なし (global)
- Major (global) errorとminor (local) errorの数
- 具体的な間違いの数 (語彙選択、冠詞、三単現など)

Fluency

- temporal measures (speech rate, phonation-time ratio, mean length of utterance など)
- disfluency (repetition, self-correction, pauseの数、頻度)

CLC subcorpus

- Mean Length of Utterances (= words per sentence, fragmented utterances, NPs, formulaic expressions excluded)

Table 5: MLU figures from the manual corpus (p. 23)

A2	7.9
B1	10.8
B2	14.2
C1	17.3
C2	19

- The number of words per sentence may be a good indicator in writing.

Complexity of Japanese learners of spoken English

- Transcription of 120 interviews of the SST (20 speakers each from Levels 4 to 9) of the NICT JLE Corpus
- Monologue tasks (picture description and past tense narration tasks)

Measures

- clause per AS-unit (complexity)
- token per AS-unit (length)

AS-unit: “a single speaker’s utterance consisting of an independent clause, or sub-clausal unit, together with any subordinate clause(s) associated with either.”

Complexity

Table 6 : Means, Standard Deviation of the objective measures and the results of omnibus ANOVA

Level	4 (A1-2)	5 (A2)	6 (B1)	7 (B1)	8 (B1)	9 (B2~)	F-value	Sig.	Partial η^2
CL/AS (complexity)	1.23 (0.14)	1.33 (0.18)	1.43 (0.19)	1.56 (0.22)	1.52 (0.16)	1.68 (0.37)	10.52	<.0001	.316
TOKEN/AS (length)	7.64 (1.21)	8.81 (1.05)	9.50 (1.53)	9.47 (1.27)	9.88 (1.54)	10.57 (3.02)	6.66	.001	.226
CLC		7.9	10.8			14.2			

Table 7 : Multiple comparison

CL/AS (complexity)	9 = 7 = 8, 7 = 8 = 6 = 5, 6 = 5 = 4
TOKEN/AS (length)	9 = 8 = 6 = 7 = 5, 7 = 5 = 4

note. $\alpha = .0125$

- レベル7は、短く、複雑なセンテンスを作る。
 - データをみると、it seems, it looks like, I thinkを多用する傾向。つなぎ言葉？
- (レベル7が逸脱とすると)、大雑把な特徴として、
 - まず長いセンテンスを作る(レベル4, 5と5以上の弁別)
 - 複雑さは少しずつ増していく。

Complexityに関して考察

- vocabulary complexityは既存の数値的指標をいくつか試してみる？
 - 語彙は、数量的な指標だけでなく、使われ方、意味も考慮すべき？
 - *But closer inspection revealed that some of these frames were being used with only one or two verbs or adjectives and were not therefore fully productive at the relevant level. Grammatical constructional analysis must therefore be supplemented in these cases with searches for relevant and representative lexical items... (p. 111)*
- 文章の長さは、lower levelを弁別する指標として使えそう。
- 自発発話で従属節が使えないレベルでは、形容詞の数、前置詞句の数、不定詞、分詞の数も指標になるのでは？
 - discrete sentence > strings of sentences (coordinated sentences)
 - > subordinate sentences
- もっと下のレベルなら、名詞句で判定することも可能かも。

スピーキングを細かく評定する

Complexity, Accuracy, Fluency (CAF)

Complexity

- unitの長さ (T-unit, AS-unitなど)
- 文法的複雑さ (clause/unit, non-finite verb/unit)

Accuracy

- 間違いのある・なし (global)
- Major (global) errorとminor (local) errorの数
- 具体的な間違いの数 (語彙選択、冠詞、三単現など)

Fluency

- temporal measures (speech rate, phonation-time ratio, mean length of utterance など)
- disfluency (repetition, self-correction, pauseの数、頻度)

Spoken **Fluency** Illustrative Scale (Council of Europe 2001)

- B1** Can keep going comprehensibly, even though pausing for grammatical and lexical planning and repair is very evident, especially in longer stretches of free production. Can make him/herself understood in short contributions, even though pauses, false starts and reformulation are very evident.
- A2** Can construct phrases on familiar topics with sufficient ease to handle short exchanges, despite very noticeable hesitation and false starts.
- A1** Can manage very short, isolated, mainly pre-packaged utterances, with much pausing to search for expressions, to articulate less familiar words, and to repair communication.

Temporal Characteristics

- B1** Can keep going *comprehensibly*, even though pausing for grammatical and lexical planning and repair is very evident, especially *in longer stretches of free production*. Can make *him/herself understood in short contributions*, even though pauses, false starts and reformulation are very evident.
- A2** Can construct *phrases* on familiar topics with sufficient ease to handle *short exchanges*, despite very noticeable hesitation and false starts.
- A1** Can manage *very short, isolated, mainly pre-packaged utterances*, with much pausing to search for expressions, to articulate less familiar words, and to repair communication.

Disfluency

- B1** Can keep going comprehensibly, **even though pausing for grammatical and lexical planning and repair is very evident**, especially in longer stretches of free production. Can make him/herself understood in short contributions, **even though pauses, false starts and reformulation are very evident**.
- A2** Can construct phrases on familiar topics with sufficient ease to handle short exchanges, **despite very noticeable hesitation and false starts**.
- A1** Can manage very short, isolated, mainly pre-packaged utterances, **with much pausing to search for expressions, to articulate less familiar words, and to repair communication**.

Fluencyを手がかりに

NICT JLE Corpus では・・・

- Temporal Characteristics に関して
 - 時間情報(タイムスタンプ)が付与されていない
 - ポーズがマークアップされている
 - Token per AS-unit の算出(ある程度のレベルのみ)
- Disfluency に関して
 - フィラー、言いよどみ、ポーズがマークアップされている

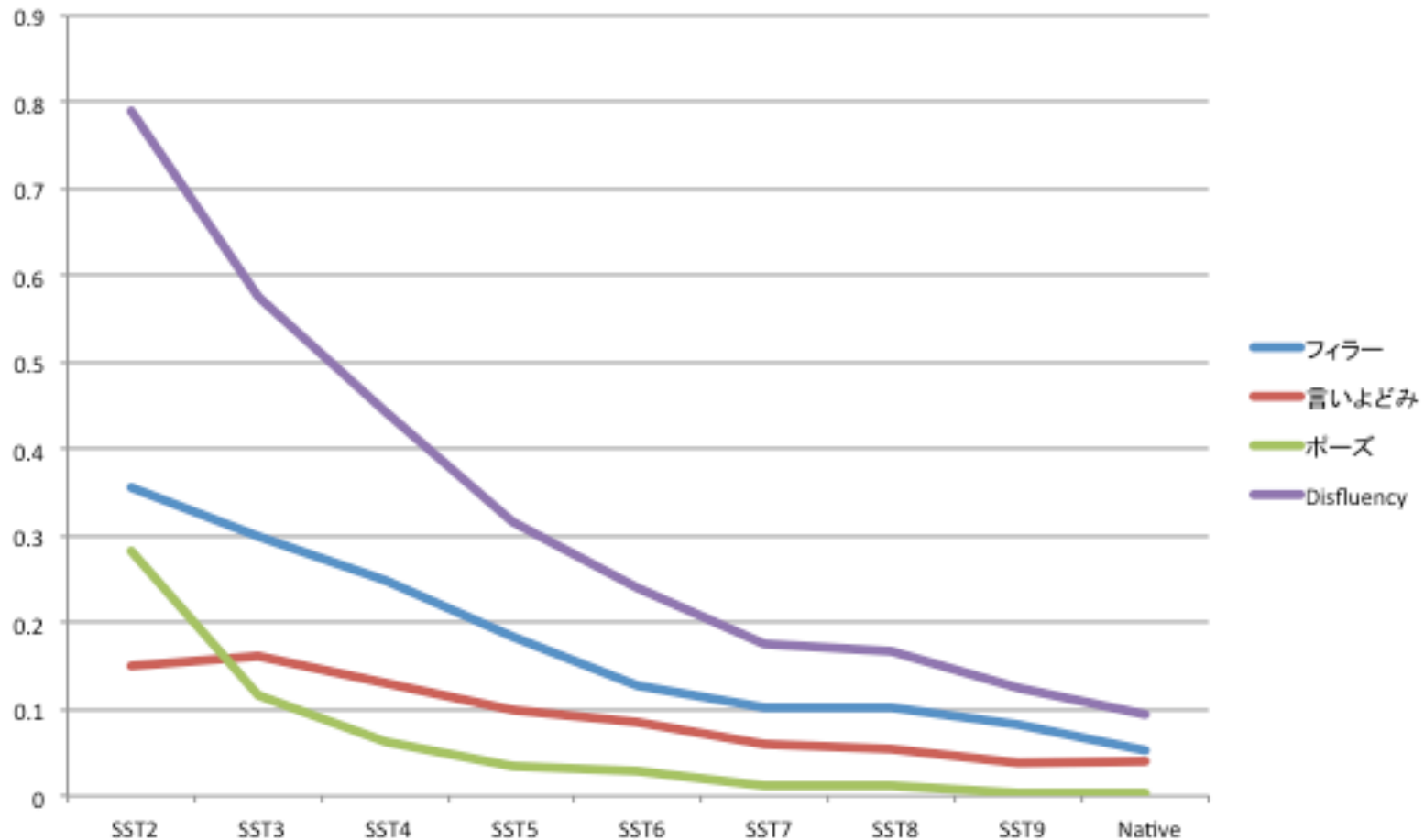
NICT JLE Corpusにおける Disfluency要素へのマークアップ(例)

Urm I have um I have mom and I have a dad and
I have ur younger and elder older sister.



<F>Urm</F> <R>I have</R> <F>um</F> I have
mom and I have a dad and I have <F>ur</F>
younger and <SC>elder</SC> <./> older sister.

Disfluency要素の 有効発話に対する比率



Disfluencyに関する考察

- フィラー、言いよどみ、ポーズなどの Disfluency 要素は、話し言葉に特有の基準特性の候補

Disfluencyに関する考察

- ただし、要素別の出現率は個人のクセに影響されやすいかもしれない
 - 各要素とも、発話内容や表現をプランニングするための空白時間を埋めるもの



学習者①

ほとんど言いよどまないが、
ポーズが多い



学習者②

言いよどみが多いが
ほとんどフィラーがない

→ 3つのDisfluency要素をセットで一つのDisfluency要素として捉える必要がある？

Disfluencyに関する考察

- 逆に、他の要素に比べ、あるレベルをより特徴付ける要素があるかもしれない



A2の学習者

フィラーや言いよどみの頻度の差はないが、
ポーズの頻度が顕著に異なる



B1の学習者

→ ポーズの頻度は他の要素に比べ、
よりA1を特徴付ける特性かもしれない

Disfluencyに関する考察

- フィラー、言いよどみ、ポーズなどの Disfluency 要素は、話し言葉に特有の基準特性の候補
 - ただし、分散なども参照して、要素単体でのレベルとの相関を注意深く検証する必要がある

スピーキングを細かく評定する

Complexity, Accuracy, Fluency (CAF)

Complexity

- unitの長さ (T-unit, AS-unitなど)
- 文法的複雑さ (clause/unit, non-finite verb/unit)

Accuracy

- 間違いのある・なし (global)
- Major (global) errorとminor (local) errorの数
- 具体的な間違いの数 (語彙選択、冠詞、三単現など)

Fluency

- temporal measures (speech rate, phonation-time ratio, mean length of utterance など)
- disfluency (repetition, self-correction, pauseの数、頻度)

Accuracyに基づいて

NICT JLE Corpus では・・・

- 文法・語彙誤りがマークアップされている
(エラータグ)

エラータグのリスト

名詞

活用の誤り<n_lxc>

単複の誤り<n_num>

格の誤り<n_cs>

可算・不可算の誤り<n_cnt>

補部の誤り<n_cmp>

語彙選択の誤り<n_lxc>

動詞

活用の誤り<v_inf>

主語・動詞の人称・数の不一致</v_agr>

形の選択の誤り<v_fml>

時制の誤り<v_tns>

相の誤り<v_asp>

態の誤り<v_vo>

定形・不定形の選択の誤り<v_fin>

否定形の誤り<v_ng>

疑問形の誤り<v_qst>

補部の誤り<v_cmp>

語彙選択の誤り<v_lxc>

形容詞

活用の誤り<aj_inf>

原級・比較級・最上級の用法の誤り<aj_us>

数の誤り<aj_num>

数量詞の誤り<aj_qnt>

数・性別の不一致<aj_agr>

補部の誤り<aj_cmp>

語彙選択の誤り<aj_lxc>

副詞

活用の誤り<av_inf>

原級・比較級・最上級の用法の誤り<av_us>

位置の誤り<av_pst>

語彙選択の誤り<av_lxc>

冠詞

選択の誤り<at>

助動詞

語彙選択の誤り<mo_lxc>

代名詞

活用の誤り<pn_inf>

数・性別の不一致<pn_agr>

格の誤り<pn_cs>

語彙選択の誤り<pn_lxc>

接続詞

語彙選択の誤り<con_lxc>

関係詞

格の誤り<rel_cs>

語彙選択の誤り<rel_lxc>

疑問詞

語彙選択の誤り<itr_lxc>

前置詞

補部の誤り<prp_cmp>

従属前置詞以外の誤り<prp_lxc1>

従属前置詞の誤り<prp_lxc2>

その他

和製英語<o_je>

語順の誤り<o_odr>

種類が特定できない誤り<o_uk>

発話意図が不明<o_uit>

連語表現の誤り<o_lxc>

全47種類

Illustrative Error Rates and Bandwidths (Hawkins and Buttery 2010)

- 決定詞の誤り ex) *Shes name was Anna.

CEFR Level	B1	B2	C1	C2	
Error Rate	0.263	0.171	0.097	0.053	
		0.217	0.134	0.075	

– B1 > B2 > C1 > C2

決定詞の誤りはB1～C2へレベルが上がるごとに
顕著に減少していく(Progressive learning
pattern)

Illustrative Error Rates and Bandwidths (Hawkins and Buttery 2010)

- 前置詞の脱落 ex) **I gave it John.*

CEFR Level	B1	B2	C1	C2
Error Rate	3.297	2.998	1.684	1.253
			2.341	

– [B1, B2] > [C1, C2]

前置詞の脱落は、B1からB2に上がる過程ではそれほど減少しないが、B2からC1に上がる際に顕著に減少する。C1からC2に上がる際もそれほど減少しない(Progressive learning pattern with different transitional features)

Illustrative Error Rates and Bandwidths (Hawkins and Buttery 2010)

- 名詞の可算性に関する誤り
ex) **I don't have any monies.*

CEFR Level	B2	C1	B1	C2	
Error Rate	0.249	0.183	0.114	0.105	
			0.149		

- [B2, C1] > [B1, C2]
名詞の可算性に関する誤りは、B1からB2に上がる際に増加し、C1に上がる際に再び減少するものの、B1よりも依然多い。C2で最小値になる
(**Inverted U**)

エラータグのリスト

名詞

活用の誤り<n_lxc>

単複の誤り<n_num>

格の誤り<n_cs>

可算・不可算の誤り<n_cnt>

補部の誤り<n_cmp>

語彙選択の誤り<n_lxc>

動詞

活用の誤り<v_inf>

主語・動詞の人称・数の不一致</v_agr>

形の選択の誤り<v_fml>

時制の誤り<v_tns>

相の誤り<v_asp>

態の誤り<v_vo>

定形・不定形の選択の誤り<v_fin>

否定形の誤り<v_ng>

疑問形の誤り<v_qst>

補部の誤り<v_cmp>

語彙選択の誤り<v_lxc>

形容詞

活用の誤り<aj_inf>

原級・比較級・最上級の用法の誤り<aj_us>

数の誤り<aj_num>

数量詞の誤り<aj_qnt>

数・性別の不一致<aj_agr>

補部の誤り<aj_cmp>

語彙選択の誤り<aj_lxc>

副詞

活用の誤り<av_inf>

原級・比較級・最上級の用法の誤り<av_us>

位置の誤り<av_pst>

語彙選択の誤り<av_lxc>

冠詞

選択の誤り<at>

助動詞

語彙選択の誤り<mo_lxc>

代名詞

活用の誤り<pn_inf>

数・性別の不一致<pn_agr>

格の誤り<pn_cs>

語彙選択の誤り<pn_lxc>

接続詞

語彙選択の誤り<con_lxc>

関係詞

格の誤り<rel_cs>

語彙選択の誤り<rel_lxc>

疑問詞

語彙選択の誤り<itr_lxc>

前置詞

補部の誤り<prp_cmp>

従属前置詞以外の誤り<prp_lxc1>

従属前置詞の誤り<prp_lxc2>

その他

和製英語<o_je>

語順の誤り<o_odr>

種類が特定できない誤り<o_uk>

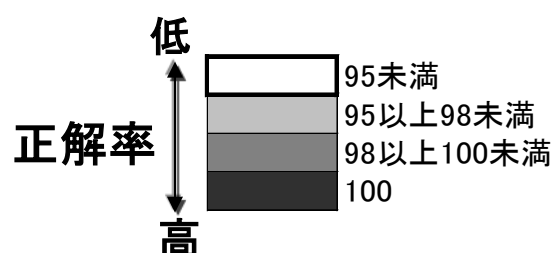
発話意図が不明<o_uit>

連語表現の誤り<o_lxc>

全47種類

NICT JLEコーパスに含まれる誤りの レベル内順位およびレベル間の推移

レベル2	レベル3	レベル4	レベル5	レベル6	レベル7	レベル8	レベル9								
冠詞	40.57	冠詞	52.75	冠詞	58.46	冠詞	65.68	冠詞	77.73	冠詞	77.38	冠詞	78.67	冠詞	84.43
前置詞の語彙選択	76.44	前置詞の語彙選択	80.58	動詞の補部	85.61	動詞の補部	86.36	動詞の補部	87.08	形容詞の活用	80.00	動詞の補部	88.42	動詞の補部	88.13
助動詞の語彙選択	81.07	動詞の補部	84.93	前置詞の語彙選択	85.83	前置詞の語彙選択	91.14	形容詞の活用	91.50	動詞の補部	87.26	前置詞の語彙選択	92.80	前置詞の語彙選択	94.23
動詞の補部	82.81	助動詞の語彙選択	90.01	形容詞の活用	92.81	動詞の時制	95.15	前置詞の語彙選択	92.39	前置詞の語彙選択	93.48	動詞の時制	96.77	動詞の時制	97.87
動詞の否定形	80.00	動詞の語彙選択	93.24	助動詞の語彙選択	94.29	形容詞の活用	95.55	助動詞の語彙選択	96.12	動詞の時制	96.09	動詞の語彙選択	97.47	主語・動詞の人称一致	98.23
副詞の語彙選択	90.63	動詞の時制	94.57	形容詞の語彙選択	94.66	助動詞の語彙選択	95.65	動詞の時制	96.50	助動詞の語彙選択	96.97	形容詞の語彙選択	97.70	助動詞の語彙選択	98.41
動詞の語彙選択	91.50	主語・動詞の人称一致	94.78	動詞の語彙選択	95.10	動詞の語彙選択	96.27	形容詞の語彙選択	96.96	動詞の語彙選択	97.72	名詞の数	98.54	名詞の数	98.71
主語・動詞の人称一致	95.22	動詞の否定形	95.71	動詞の時制	95.14	動詞の否定形	96.39	主語・動詞の人称一致	97.39	形容詞の語彙選択	98.19	副詞の語彙選択	98.69	動詞の語彙選択	98.95
形容詞の語彙選択	97.02	形容詞の語彙選択	96.38	主語・動詞の人称一致	95.17	形容詞の語彙選択	96.42	動詞の語彙選択	97.46	主語・動詞の人称一致	98.33	名詞の数語彙選択	98.81	接続詞の語彙選択	98.97
動詞の時制	97.20	副詞の語彙選択	96.97	動詞の否定形	97.40	主語・動詞の人称一致	97.21	名詞の数	98.43	名詞の数	98.54	主語・動詞の人称一致	98.82	形容詞の語彙選択	99.04
名詞の数	97.50	形容詞の活用形	97.12	名詞の数	97.45	名詞の数	97.54	接続詞の語彙選択	98.63	接続詞の語彙選択	98.58	助動詞の語彙選択	98.84	副詞の語彙選択	99.14
動詞の定形・不定形	97.62	名詞の数	97.52	動詞の活用	97.52	副詞の語彙選択	98.54	名詞の語彙選択	98.71	副詞の語彙選択	99.01	接続詞の語彙選択	99.31	名詞の語彙選択	99.29
副詞の位置	97.68	名詞の語彙選択	98.62	接続詞の語彙選択	98.11	接続詞の語彙選択	98.59	動詞の否定形	98.98	名詞の語彙選択	99.11	副詞の位置	99.50	前置詞の補部	99.69
前置詞の補部	98.98	接続詞の語彙選択	99.05	副詞の語彙選択	98.22	名詞の語彙選択	98.62	副詞の語彙選択	99.01	副詞の位置	99.31	動詞の態	99.61	副詞の位置	99.83
動詞の相	99.08	名詞の活用	99.32	名詞の語彙選択	98.48	副詞の位置	99.54	副詞の位置	99.58	動詞の否定形	99.65	動詞の否定形	99.74	形容詞の級	99.83
名詞の語彙選択	99.16	動詞の相	99.34	副詞の位置	99.36	前置詞の補部	99.65	動詞の態	99.68	前置詞の補部	99.65	動詞の相	99.90	名詞の可算・不可算性	99.87
形容詞の級	99.64	副詞の位置	99.47	動詞の相	99.67	名詞の活用	99.66	前置詞の補部	99.83	名詞の可算・不可算性	99.85	形容詞の級	99.93	形容詞の補部	99.89
名詞の可算・不可算性	99.90	動詞の活用	99.59	前置詞の補部	99.69	動詞の態	99.67	名詞の可算・不可算性	99.85	動詞の態	99.85	名詞の可算・不可算性	99.95	名詞の活用	100.00
名詞の活用	100.00	動詞の態	99.73	動詞の態	99.75	動詞の相	99.73	動詞の相	99.90	動詞の活用	99.88	名詞の活用	100.00	動詞の活用	100.00
接続詞の語彙選択	100.00	前置詞の補部	99.73	名詞の活用	99.75	動詞の活用	99.82	名詞の活用	99.93	形容詞の級	99.97	動詞の活用	100.00	動詞の相	100.00
動詞の態	100.00	名詞の可算・不可算性	99.86	形容詞の級	99.81	名詞の可算・不可算性	99.83	動詞の活用	99.97	動詞の相	99.98	形容詞の活用	100.00	動詞の態	100.00
形容詞の活用	100.00	形容詞の級	99.93	名詞の可算・不可算性	99.95	形容詞の級	99.90	形容詞の補部	99.97	形容詞の補部	99.98	形容詞の補部	100.00	動詞の否定形	100.00
形容詞の補部	100.00	形容詞の補部	100.00	形容詞の補部	99.98	形容詞の補部	100.00	形容詞の級	100.00	名詞の活用	100.00	前置詞の補部	100.00	形容詞の活用	100.00



Accuracyに関する考察

(English Profileの手法に関連して)

- エラータグを利用する
- English Profileではあまり検証されていない
Aレベルのエラー状況を細やかに見る
- 誤りの種類だけでなく、出現率情報付きの
基準特性とする？
 - 上級レベルになっても依然として残る誤りも多い
 - 減少率等から成長期を把握できる
- 教科書コーパスと連携し、学習ターゲットの
増加がエラーを誘発する影響を考察する？
 - 学習ターゲットが増えると、一旦減少していた誤りが
再度増加 (Inverted U) するケースがあり得る