

CEFRに基づく英語リスニング・レベルの基準特性を探る

Identifying Criterial Features of English Texts Based on the CEFR Listening Levels

根岸雅史

Masashi NEGISHI

東京外国語大学

Tokyo University of Foreign Studies

Abstract

The Common European Framework of Reference for Languages (CEFR) is spreading its influence outside Europe. However, this framework is not language-specific, and therefore it is necessary to identify so-called “criterial features” for each language, when it is actually used for the design of learning, teaching, and assessment of the language. This article focuses on the CEFR listening descriptors, and investigates whether we can predict the CEFR listening levels from textual features of the texts. The present study analysed authentic texts which were related to the CEFR levels by experts. All the texts were transcribed, and they were analysed statistically. The results of multiple regression analyses indicate that it is possible to make fairly accurate predictions of the CEFR listening levels from textual features, if they are combined with subjective judgements as to “complexity” and “degree of specialisation” of the texts. The results of the principal component analysis suggest that criterial features of the CEFR English listening levels consist of such components as “linguistic difficulty”, “here-and-nowness”, and “diversity of topics”. This study has implications for the selection of authentic listening texts and for the validation of existing teaching materials based on the Common European Framework.

Keywords

CEFR, Listening, Criterial Features

1. CEFRのリスニング記述文における基準特性

Common European Framework of Reference for Languages (CEFR)は、外国語教育の参照枠として、ヨーロッパの枠を超え、世界にインパクトを及ぼし始めている。しかしながら、その枠組みは、外国語教育の共通枠組みとなることを目指しているために、個別言語に関する記述はなく、実際の運用にあたっては、解決しなければならない課題が出てきているのも事実である。

CEFRでは、コミュニケーション活動をListening, Reading, Spoken Interaction, Spoken Production, Writingの5技能に分類しているが、本研究ではこのうちリスニング(Listening)に焦点を当てて研究を行うこととする。CEFR Common Reference Levels: self-assessment grid (Council of Europe, 2001, pp.26-27)におけるリスニングの記述文は、次の通りである。

A1

I can recognise familiar words and very basic phrases concerning myself, my family and immediate concrete surroundings when people speak slowly and clearly.

A2

I can understand phrases and the highest frequency vocabulary related to areas of most immediate personal relevance (e.g. very basic personal and family information, shopping, local area, employment). I can catch the main point in short, clear, simple messages and announcements.

B1

I can understand the main points of clear standard speech on familiar matters regularly encountered in work, school, leisure, etc. I can understand the main point of many radio or TV programmes on current affairs or topics of personal or professional interest when the delivery is relatively slow and clear.

B2

I can understand extended speech and lectures and follow even complex lines of argument provided the topic is reasonably familiar. I can understand most TV news and current affairs programmes. I can understand the majority of films in standard dialect.

C1

I can understand extended speech even when it is not clearly structured and when relationships are only implied and not signalled explicitly. I can understand television programmes and films without too much effort.

C2

I have no difficulty in understanding any kind of spoken language, whether live or broadcast, even when delivered at fast native speed, provided I have some time to get familiar with the accent.

この記述文には、「言語的困難度(e.g. A1: familiar words and very basic phrases)」「トピックの馴染みの度合い(e.g. B2: provided the topic is reasonably familiar)」「興味の度合い(e.g. B1: current affairs or topics of personal or professional interest)」「理解の詳細さ(e.g. B1: understand the main points of clear standard speech)」「長さ(e.g. B2: understand extended speech)」「議論の複雑さ(e.g. C1: when it is not clearly structured and when relationships are only implied and not signalled explicitly)」「発話速度(e.g. C2: when delivered at fast native speed)」「なまり(e.g.

C2: provided I have some time to get familiar with the accent)」「明瞭さ(e.g. A1: when people speak ... clearly)」「ノイズ(e.g. A2: clear, ... announcements)」といった要素が含まれている。これらの要素は、いずれもテキストの困難度に関係していると考えられていることがわかる。たとえば、馴染みのない語彙は馴染みのある語彙より困難度が高いし、議論が複雑であればあるほど理解は困難になる。しかしながら、問題は、現実のテキストではそれぞれの困難度が手に手を取り合って増していくわけではないということである。つまり、「議論は複雑ではない」が「なまりが強い」という組み合わせがあったり、「長さは短い」が「早口である」という組み合わせがあったりするということだ。また、「理解の詳細さ」と「議論の複雑さ」のように、相互依存する可能性のあるものもある。こうした現実には、CEFRのような枠組みの適用にあたり、困難を引き起こすものと思われる。

リスニングの記述文に含まれる要素はすべてのレベルで言及されているわけではなく、いくつかのレベルにおいてのみ言及されているのが一般的である。たとえば、「なまり(accent)」に関してはC2で言及されているが、それより下のレベルでは、B2で「標準方言(standard dialect)」に関する言及があるだけである。つまり、そのレベルを特徴づけるところにおいてのみ言及される傾向がある。この点に関して、Council of Europe(2001)には、次のような記述がある。

Not every element or aspect in a descriptor is repeated at the following level. That is to say that entries at each level describe selectively what is seen as salient or new at that level. They do not systematically repeat all the elements mentioned at the level below with a minor change of formulation to indicate increased difficulty. (Council of Europe, 2001, p.37)

こうした問題は、具体的な教材やテストの開発において、より鮮明になる。とりわけオーセンティック・テキストを用いた場合は、様々な要素が調整された人為的なテキストとは異なり、異なったレベルの要素が組み合わされる可能性が高い。これまでの研究(根岸 2008; 2009a; 2009b)からも、記述文には困難度を定めるいくつかの要素が含まれているが、CEFRのレベル判断に決定的な影響を与えるものは、いくつかの「中心的な要素」と推測される。そこで、本研究では、リスニング・レベルを決定するこれらの「基準特性」の存在を実証的に探ることとする。研究設問は次の通りである。

1. CEFR に基づく英語リスニング・レベルはテキスト情報から推定可能か。
2. CEFR に基づく英語リスニング・レベルの基準特性は何か。

2. 研究1

2.1 調査に用いたスクリプト

現実には CEFR に関連付けられたとされる英語の音声テキストは、教材やテストにはあるが、オーセンティックな音声テキストはまだまだ数少ない。Horváth and Pižorn(2005)の *Listening* は、そうした音声テキストを含んだ数少ない文献のひとつである。ここに含まれる

音声は、「現実の生活から録った本物のテキスト(authentic texts recorded from real life)」であり、CD により実際の音声を提供されている。また、英語も英米の母語話者による英語のみならず非英語母語話者の英語も含まれている。ここでもっとも特徴的なのは、それぞれの録音に対して、CEFR の分類に基づくレベル分けがなされている点である。この判断は、英語教育とアセスメントの専門家によってなされている。

この音声 CD の文字起こしは、言語学修士号を持つアメリカ人英語母語話者に依頼して行った(問題構成の概要は表1を参照)。ただし、問題 42, 43, 44, 48 は、テキストが複数の会話から成っており、他のテキストと構成が異なるために、分析対象としなかった。

表 1 Contents of the CDs (Horváth and Pižorn: 2005, Appendix 2)

テキスト	タイトル	CEFR レベル	長さ(分:秒)
1	The National Palace Museum - Sample	B2	2:51
2	Corfu	A2	2:48
3	Game Boy and music	B1	2:12
4	Celebrating the New York	B1	2:11
5	A townie turned to farming	B2	3:29
6	Albanian farming	B2	3:30
7	Scottish kilts	B2	2:35
8	Casey	B2	2:57
9	El Niño	C1	1:52
10	Climbing	C1	3:31
11	Rainbow, the billy goat	C1	2:14
12	Shuttle launch - Sample	C1	3:51
13	Sleepy Hollow	A2	1:54
14	Nicole Kidman	A2	3:05
15	The jazz singer	B2	2:29
16	Vision reined	B2	4:07
17	Home of the future	C2	2:02
18	Peter's day - Sample	B1	2:48
19	Harry Potter	B1	2:04
20	Midwest people	B1	4:06
21	A female astronaut	B2	3:14
22	Jumble at the sale	B2	1:45
23	Empty seas	C1	2:27
24	Air chair	C1	2:33
25	Courmayeur - Sample	A2	2:29
26	Pension calculator	A2	1:50
27	Hotel Sao Joao	B1	2:07

表 1(続き) Contents of the CDs (Horváth and Pižorn; 2005, Appendix 2)

テキスト	タイトル	CEFR レベル	長さ(分:秒)
28	Gerald's family	B1	2:59
29	Cycling	B1	2:34
30	Earthquake	B1	2:13
31	Ken Follet	B2	3:31
32	Cheddar in the cave	B2	1:52
33	Nyasa animals	B2	2:07
34	Haunting monk	B2	3:28
35	Coffee	B2	2:47
36	Healthy urban living	B2	1:50
37	Paris skaters	B2	3:33
38	Weather forecast	C1	2:35
39	Tulips	C1	3:07
40	English courses - Sample	B1	2:57
41	My room	A1	1:15
42	Dialogues	A1	2:35
43	So embarrassing!	A2	3:01
44	Radio	B1	3:15
45	Eileen's objects	B1	3:29
46	Attitudes to parks	B2	2:20
47	Listeners' views	C1	2:47
48	Intuition and logic	C1	2:32

2.2 予測変数と分析方法

主な分析方法は、テキストの客観的指標として考えられるすべてを数量化し、CEFRレベルの推定が可能かどうかを統計的にみるというものである。具体的には、テキストの客観的指標として33の言語項目を用いて、表1のテキストを対象に重回帰分析を行った。目的変数は、CEFRのA1, A2, B1, B2, C1, C2 をそれぞれ1, 2, 3, 4, 5, 6 と置き換えた。

独立変数としては、次の1~33を設定した。1. 時間 2. 発話スピード 3. 総語数 4. 異なり語数 5. センテンス数 6. センテンス内の平均語数 7. フレッシュ・リーダビリティ・イーゼ (Flesch Readability Ease) 8. フレッシュ・キンケイド・グレード・レベル (Flesch-Kincaid Grade Level) 9. 内容語 10. 語彙密度 11. レクサイル指数といった変数の他に、言語的要素に関する変数として、以下のようなBiber(1988)の多次元分析(MD分析)の次元スコアを用いた(なお、タイプ・トークン比については、これに代わり、ギロー指標を用いた)。12. 人称代名詞 13. it+指示代名詞 14. 前置詞 15. 名詞化 16. 助動詞 17. 場所を表す副詞類 18. 時を表す副詞類 19. 他の副詞的従属接続詞 20. 否定 21. 省略 22. 句をつなぐ等位接続詞 23. 独立節をつなぐ等位接続詞 24. ギロー

指標 25. 主動詞としてのbe 26. 限定形容詞 27. 強調語句 28. to不定詞 29. 過去形
30. 名詞 31. 副詞 32. 平均語長 33. 受動態

2.3 結果

重回帰分析の結果は、表2の通りである。自由度調整済重相関係数の二乗が0.31と低く、33項目の変数からリスニング教材のCEFRレベルを推定できるとは言い難い。

表2 重回帰分析の結果(研究1)

重相関係数	重相関係数の二乗	自由度調整済重相関係数の二乗
0.90	0.81	0.31

研究1ではCEFRリスニングのレベルをテキストの客観的な情報から推定しようとしたが、客観的指標だけを用いての精度の高い推定は困難であった。そこで、研究2では、主観的な判断を伴う説明変数を追加して、分析を行う。

3. 研究2

3.1 分析方法

研究2では、研究1で用いたテキストの客観的指標(33の言語項目)に主観的指標を追加した。この主観的指標は、CEFRの記述文、および、これまでの分析から「複雑さ」と「専門性」という指標を立て、筆者自身が実際のCDの音声を聞いて4段階で2度評価し(採点者内信頼性は、複雑さが0.86、専門性が0.84であった)、判断の異なったものは再度CDを聞いて最終判断を下した。

CEFRレベルの推定が可能かどうかを統計的に分析したが、その手順は、主成分分析から主成分得点を導き、それを重回帰分析にかけるというものであった。因子分析ではなく主成分分析を用いたのは、後者がデータ数が少ない時に頑健であるためである。因子分析を行う際は、サンプル数(データ数)が変数の数の3倍から5倍、あるいは、研究者によっては10倍必要だという主張もあり、今回のように変数の数とサンプル数にあまり差がない場合、因子分析は適さないとされる。

3.2 結果

主成分得点による重回帰分析の結果は、表3の通りである。この結果から、CEFRに基づく英語リスニング・レベルはテキスト情報に主観的指標を加えることで、およその推定は可能であると言えるだろう。

表3 重回帰分析の結果(研究2)

重相関係数	重相関係数の二乗	自由度調整済重相関係数の二乗
0.73	0.54	0.46

表4 主成分分析の結果

	偏回帰 係数	標準誤差	<i>t</i> 値	<i>p</i> 値	標準化偏 回帰係数
PC.1	-0.11	0.04	-2.61	1.30E-02	-0.29
PC.2	0.20	0.05	4.28	1.28E-04	0.48
PC.3	-0.01	0.06	-0.21	8.37E-01	-0.02
PC.4	-0.05	0.07	-0.71	4.81E-01	-0.08
PC.5	-0.30	0.07	-4.14	1.94E-04	-0.46
PC.6	0.00	0.08	0.05	9.59E-01	0.01
定数項	3.68	0.11	32.32	1.01E-28	NA

表4は主成分分析の結果である。*p* 値から、第3、第4、第6主成分は説明変数の予測に役立っていない($p > .05$)と考えられるために、第1、第2、第5主成分のみの解釈を試みた。これらの主成分の負荷の高いテキストを主成分ごとに実際に聞き直し、その特性を推定した。第1主成分は、テキストの中でどれだけ複数のトピックが展開されているかに関わっていると思われる。この主成分は、複数のトピックが展開されていればテキストの困難度は上がり、1つのトピックを中心に展開されていれば理解は容易となるため、「トピックの複合性」とも呼べるであろう。これは、CEFR リスニングの記述文で言えば、「議論の複雑さ」や「理解の詳細さ」に関わる成分と考えられる。第2主成分は、語彙の困難度を中心とした「言語的困難度」に関わるものと思われる。第5主成分は、第1主成分や第2主成分とは困難度の方向が逆で、こちらの度合いが高くなるとテキストの困難度は低くなる。第5主成分の解釈は容易ではなかったが、空間的・時間的に目の前で起こっているものほど、この主成分の得点が高いことから、最終的に「目の前度 (here-and-nowness)」のように解釈するのが適当であろうとの結論に達した。この観点からすると、目の前で起こっている事柄を扱ったテキストは、この主成分の得点が高く、過去の出来事や抽象度の高い出来事を扱ったテキストは、得点が低くなる。これらをまとめると、CEFR に基づく英語リスニング・レベルの基準特性については、「言語的困難度」「目の前度 (here-and-nowness)」「トピックの複合性」と解釈されるような主成分を抽出したと解釈できるだろう。

上の分析を元に、次のような回帰式を得ることができた。

$$\text{CEFR 推定レベル} = -0.11 \cdot \text{PC.1} + 0.2 \cdot \text{PC.2} - 0.01 \cdot \text{PC.3} - 0.05 \cdot \text{PC.4} - 0.3 \cdot \text{PC.5} + 0 \cdot \text{PC.6} + 3.68$$

この回帰式から算出したテキストの困難度とそれを四捨五入した値(推定レベル)、専門家の判断した CEFR レベルを示したものが、表5である。ここからわかるとおり、44 テキスト中 22 テキストを正しく推定し、残りの 22 テキストは±1の差となった。

表5 回帰式による推定レベルとの差

テキスト	CEFR レベル (5段階)	推定レベル	推定レベル (四捨五入)	推定レベル との差
1	4	4.10	4	0.10
2	2	3.46	3	1.46
3	3	2.71	3	-0.29
4	3	2.69	3	-0.31
5	4	3.58	4	-0.42
6	4	4.62	5	0.62
7	4	3.69	4	-0.31
8	4	2.62	3	-1.38
9	5	3.68	4	-1.32
10	5	4.81	5	-0.19
11	5	4.47	4	-0.53
12	5	4.40	4	-0.60
13	2	2.54	3	0.54
14	2	3.25	3	1.25
15	4	3.50	4	-0.50
16	4	4.50	5	0.50
17	5	3.88	4	-1.12
18	3	2.93	3	-0.07
19	3	3.21	3	0.21
20	3	4.10	4	1.10
21	4	4.32	4	0.32
22	4	3.22	3	-0.78
23	5	4.27	4	-0.73
24	5	4.26	4	-0.74
25	2	3.08	3	1.08
26	2	3.02	3	1.02
27	3	2.73	3	-0.27
28	3	2.88	3	-0.12
29	3	3.09	3	0.09
30	3	4.44	4	1.44
31	4	3.90	4	-0.10
32	4	3.52	4	-0.48
33	4	4.38	4	0.38
34	4	4.62	5	0.62
35	4	3.80	4	-0.20

表5(続き) 回帰式による推定レベルとの差

テキスト	CEFR レベル (5段階)	推定レベル	推定レベル (四捨五入)	推定レベル との差
36	4	3.68	4	-0.32
37	4	4.51	5	0.51
38	5	4.92	5	-0.08
39	5	4.37	4	-0.63
40	3	2.61	3	-0.39
41	1	1.68	2	0.68
45	3	3.45	3	0.45
46	4	4.11	4	0.11
47	5	4.38	4	-0.62

注:「推定レベル(四捨五入)」で網掛けが付いているのは、
 専門家の判断による CEFR レベルとの差が±1のもの。

4. 結論

CEFR に基づく英語リスニング・レベル推定は、テキスト情報に関する客観的な指標だけでは、可能とは言い難いが、これらに「複雑さ」と「専門性」といった主観的指標を加えることで、およその推定は可能となることがわかった。

CEFR のリスニングの記述文は、様々な観点から記述されているわけであるが、これらはすべてが常に効いているわけではない。本研究により明らかになった英語リスニング・レベルの基準特性は、「言語的困難度」「目の前度 (here-and-nowness)」「トピックの複合性」となった。「言語的困難度」は既述した通り、記述文に様々な形であらわれており、これは当然の結果と言えるだろう。また、「トピックの複合性」は、記述文の中にある「議論の複雑さ」や「理解の細かさ」と深く関わっている。トピックが1つに収斂していれば、議論は単純なものになる傾向があり、それゆえ、求められる理解の質も高度なものではないだろう。反対に、トピックが複数にまたがり、いくつものトピックからなる場合は、それぞれの関係は複雑なものになり、それゆえ求められる理解の質も高度なものとなる。

ただ、「目の前度 (here-and-nowness)」という基準特性は、やや意外なものであった。これが意味することは、目の前で起こっていることを理解するのは比較的容易で、空間的・時間的に離れた事柄を理解するのは、抽象度が上がるためか、難しいということだ。この基準特性に直接該当するような記述文は存在しないと言っていいだろう。これが現れたのは、たまたま今回扱った音声テキストの影響なのかどうかは、今後の研究結果を待たなければならない。

課題としては、今回の CEFR のレベル判断は完全に Horváth and Pižorn (2005) の *Listening* に依拠しているが、この判断の妥当性に問題はなかったかという点がある。ただし、CEFR のレベルは、誰かが権威的に判断を下すことができるというようなものではなく、難しい問題である。とはいえ、本研究の結果については、他の専門家が同様の判断を下すかなど、妥当性の検証を今後行うべきかもしれない。

統計分析にあたっては、重回帰分析の目的変数である CEFR の A1, A2, B1, B2, C1, C2 をそれぞれ 1, 2, 3, 4, 5, 6 と置き換えたが、これらが厳密な意味での間隔尺度である根拠はない。今回は他の妥当性の高い方法がみつからなかったため、この点も今後の課題と言えるかもしれない。

この研究の成果をもとに、英語の自然言語コーパスなどから採った教材をある程度の精度で CEFR レベルに分類することが可能となった。英語は、音声に関してもオーセンティックな素材にあふれていると言ってもよい。このような素材をレベル分けして学習者に提供することができれば、学習者は自分のレベルに合った音声素材を自分の興味に合わせて選択し、学習することができるようになるだろう。

また、逆に、今回の手法は、今日使われている様々な英語リスニング学習教材の CEFR レベルを判断する助けとなるだろう。たとえば、日本の学習指導要領に基づいて作成された中学校や高等学校の検定教科書のリスニング教材などを分析すれば、それぞれのおよその CEFR レベルがわかることになる。

参考文献

- Biber, D. (1988). *Variation across speech and writing*. New York: Cambridge University Press.
- Council of Europe. (2001). *Common European Framework of Reference for Languages: Learning, teaching, assessment*. Cambridge: Cambridge University Press.
- Horváth, K. F., and Pižorn, K. (2005). *Into Europe : Listening*. Budapest: British Council.
- 根岸雅史 2008. 「CEFRリスニングレベルの決定要因を探る」『現場型リサーチと実践へのアプローチ 英語教育・英語学習研究 金谷憲教授還暦記念論文集』桐原書店.
- 根岸雅史 2009a. 「オーセンティック・リスニング・テキストのCEFRリスニングのレベル判断における諸問題」『ARCLE REVIEW』No. 3, pp.100-109.
- 根岸雅史 2009b. 「自然言語コーパスに基づく学習教材作成のための基礎的研究:英語リスニング・テキストのCEFRレベルの決定要因とそれに基づくレベル推定の可能性」『コーパスに基づく言語学教育研究報告3 フィールド調査, 言語コーパス, 言語情報学』pp.195-210.