The CEFR Levels: Key Points and Key Problems

Brian North
www.eurocentres.com
Over 20 schools worldwide

- Languages in cultural context
- Educational foundation since 1960
- NGO to Council of Europe since 1968
- Language proficiency framework since 1989
- Development of CEFR descriptors
- Academic excellence
- Quality management
CEFR Levels: Key Points

- Origin of the CEFR levels and descriptors
- Salient characteristics of the levels
- Life beyond C2
- Validity claim of the illustrative descriptors
- Consistent interpretation of the levels
<table>
<thead>
<tr>
<th></th>
<th></th>
<th></th>
</tr>
</thead>
<tbody>
<tr>
<td>Ambilingual Proficiency</td>
<td>Proficiency</td>
<td>Mastery</td>
</tr>
<tr>
<td>Comprehensive Operational Proficiency</td>
<td>DALF / CAE</td>
<td>EOP</td>
</tr>
<tr>
<td>Adequate Operational Proficiency</td>
<td>FCE Vantage</td>
<td>Vantage</td>
</tr>
<tr>
<td>Limited Operational Proficiency</td>
<td>Threshold</td>
<td>Threshold</td>
</tr>
<tr>
<td>Basic Operational Proficiency (Threshold Level)</td>
<td>Threshold</td>
<td>Threshold</td>
</tr>
<tr>
<td>Survival Proficiency</td>
<td>Waystage</td>
<td>Waystage</td>
</tr>
<tr>
<td>Formulaic Proficiency</td>
<td>Waystage</td>
<td>Breakthrough</td>
</tr>
</tbody>
</table>
Intuitive Phase:
  > Creating a pool of classified, edited descriptors

Qualitative Phase:
  > Analysis of teachers discussing proficiency
  > 32 teacher workshops sorting descriptors

Quantitative Phase:
  > Teacher assessment of 2800 learners on descriptor-checklists (500 learners, 300 teachers)
  > Teacher assessment of videos of some learners

Interpretation Phase:
  > Setting “cut-points” for common reference levels
CEFR: Concertina-like Reference

A
Basic User

A1
A1.1  A1.2  A1.3
1      2      3

A2
A2.1  A2.2
4      5

B
Independent User

B1
6
CEFR Levels: Key Points

- Origin of the CEFR levels and descriptors
- Salient characteristics of the levels
- Life beyond C2
- Validity claim of the illustrative descriptors
- Consistent interpretation of the levels
Salient Characteristics A1

The point at which the learner can:

> interact in a simple way
> ask and answer simple questions about themselves
> respond to statements in areas of immediate need

rather than relying purely on a rehearsed repertoire of phrases
The majority of descriptors stating social functions:
> *greet people, ask how they are and react to news*
> *handle very short social exchanges*
> *discuss what to do, where to go and make arrangements*

Descriptors on getting out and about:
> *make simple transactions in shops, banks etc.*
> *get simple information about travel and services*
Salient Characteristics  B1

Maintain interaction and get across what you want to:

> give or seek personal views and opinions
> express the main point comprehensibly
> keep going comprehensibly, even though pausing evident, especially in longer stretches

Cope flexibly with problems in everyday life:

> deal with most situations likely to arise when travelling
> enter unprepared into conversations on familiar topics
Salient Characteristics  B2

Effective argument:
> account for and sustain opinions in discussion by providing relevant explanations and arguments
> explain a viewpoint on a topical issue giving the advantages and disadvantages of various options

Holding your own in social discourse:
> interact with a degree of fluency and spontaneity that makes regular interaction with native speakers possible
> adjust to changes of direction, style and emphasis

A new degree of language awareness:
> make a note of "favourite mistakes" and monitor speech for them
Fluent, well-structured language:

> good command of a broad lexical repertoire allowing gaps to be readily overcome with circumlocutions

> express self fluently and spontaneously, almost effortlessly

> produce clear, smoothly-flowing, well-structured speech, showing controlled use of organisational patterns, connectors and cohesive devices
Salient Characteristics

C2

Precision and ease with the language:

> convey finer shades of meaning precisely by using, with reasonable accuracy, a wide range of modification devices

> show great flexibility reformulating ideas in differing linguistic forms to give emphasis, to differentiate and to eliminate ambiguity
CEFR Levels: Key Points

- Origin of the CEFR levels and descriptors
- Salient characteristics of the levels

Life beyond C2

- Validity claim of the illustrative descriptors
- Consistent interpretation of the levels
<table>
<thead>
<tr>
<th>Level</th>
<th>Proficiency</th>
</tr>
</thead>
<tbody>
<tr>
<td>A1</td>
<td>Formulaic Proficiency</td>
</tr>
<tr>
<td>A2</td>
<td>Survival Proficiency</td>
</tr>
<tr>
<td>B1</td>
<td>Basic Operational Proficiency</td>
</tr>
<tr>
<td>B2</td>
<td>Limited Operational Proficiency</td>
</tr>
<tr>
<td>C1</td>
<td>Adequate / Effective Operational Proficiency</td>
</tr>
<tr>
<td>C2</td>
<td>Comprehensive Operational Proficiency</td>
</tr>
<tr>
<td>D1</td>
<td>Language professionals: Interpreters, translators, some university professors</td>
</tr>
<tr>
<td>D2</td>
<td>Genuine bilinguals (+ Beckett etc.)</td>
</tr>
<tr>
<td>E</td>
<td>WENS: Well-educated Native Speaker</td>
</tr>
</tbody>
</table>
Global Scale: C2

> Can understand with ease virtually everything heard or read.

> Can summarise information from different spoken and written sources, reconstructing arguments and accounts in a coherent presentation.

> Can express him/herself spontaneously, very fluently and precisely, differentiating finer shades of meaning even in more complex situations.
Salient Characteristics C2

Precision and ease with the language:

> convey finer shades of meaning precisely by using, with reasonable accuracy, a wide range of modification devices

> show great flexibility reformulating ideas in differing linguistic forms to give emphasis, to differentiate and to eliminate ambiguity
Salient Characteristics

Apparent ambilingualism:

> Convey, elaborate or translate to explicit expression the nuances and subtleties of their own and of others’ meaning by exploiting a comprehensive knowledge of the language to do so

> function in all situations to all intents and purposes exactly as the mother tongue; use the language in a sophisticated, natural, accurate manner apparently indistinguishable from the performance of a native speaker
CEFR Levels: Key Points

- Origin of the CEFR levels and descriptors
- Salient characteristics of the levels
- Life beyond C2
- Validity claim of the illustrative descriptors
- Consistent interpretation of the levels
Informal Discussion: Level B2:

> “Can take an active part in informal discussion in familiar contexts.”

> “Can with some effort catch much of what is said around him/her in discussion, but may find it difficult to participate effectively in discussion with several native speakers who do not modify their language in any way.”

> “Can account for and sustain his/her opinions in discussion by providing relevant explanations, arguments and comments.”
Validity: Scales before CEFR

> Wording tended to be relative. The descriptors were seldom stand-alone criteria one could rate “Yes” or “No”

> Wording often created semantic appearance of a scale, without actually describing anything

> Situation of descriptors at a particular level was arbitrary - following convention/cliché

> Lower levels tended to be worded negatively
Validity: Methodology

Developed scientifically:

- comprehensive documentation of existing descriptions
- relation to theory through descriptive scheme
- *what* learners can do *and how well* they do it
- positive, independent criterion-descriptors
- checking teachers could use categories & descriptors
- scaling on same scale as learners (video samples)
- data from real, end-of-year assessment
- four educational sectors in a multi-lingual environment
- three foreign languages (English, French, German)
- values replicated: ALTE 0.97; DIALANG: 0.92 / 0.96
## Validity: Content coherence

<table>
<thead>
<tr>
<th>SETTING</th>
<th>SPEECH</th>
<th>HELP</th>
</tr>
</thead>
<tbody>
<tr>
<td>C1</td>
<td></td>
<td></td>
</tr>
<tr>
<td>B2+</td>
<td>animated conversation between native speakers</td>
<td></td>
</tr>
<tr>
<td>B2</td>
<td>even noisy environments</td>
<td>standard spoken language</td>
</tr>
<tr>
<td>B1+</td>
<td>topics which are familiar</td>
<td></td>
</tr>
<tr>
<td>B1</td>
<td>extended everyday conversation</td>
<td>clearly articulated standard speech</td>
</tr>
<tr>
<td>A2+</td>
<td>simple, routine exchanges - familiar matters</td>
<td>clear, standard - directed at him/her</td>
</tr>
<tr>
<td>A2</td>
<td>simple everyday conversation</td>
<td>clear, slow, standard - directed at him/her</td>
</tr>
<tr>
<td>A1</td>
<td>needs of a concrete type - short, simple questions &amp; instructions</td>
<td>very clear, slow, carefully articulated repeated speech directed at him</td>
</tr>
</tbody>
</table>
# Validity: Content coherence

<table>
<thead>
<tr>
<th></th>
<th>A1</th>
<th>A2</th>
<th>A2+</th>
<th>B1</th>
<th>B2</th>
<th>C1</th>
</tr>
</thead>
<tbody>
<tr>
<td>Recognition not a native-speaker</td>
<td>✓</td>
<td>✓</td>
<td>✓</td>
<td>✓</td>
<td>✓</td>
<td>✓</td>
</tr>
<tr>
<td>Low background noise</td>
<td>✓</td>
<td>✓</td>
<td>✓</td>
<td>✓</td>
<td>✓</td>
<td></td>
</tr>
<tr>
<td>Familiar everyday topics</td>
<td>✓</td>
<td>✓</td>
<td>✓</td>
<td>✓</td>
<td>✓</td>
<td></td>
</tr>
<tr>
<td>Clear articulation</td>
<td>✓</td>
<td>✓</td>
<td>✓</td>
<td>✓</td>
<td>✓</td>
<td></td>
</tr>
<tr>
<td>Chance to get repetition</td>
<td>✓</td>
<td>✓</td>
<td>✓</td>
<td>✓</td>
<td>✓</td>
<td></td>
</tr>
<tr>
<td>Non standard, simplified</td>
<td>✓</td>
<td>✓</td>
<td>✓</td>
<td></td>
<td>✓</td>
<td></td>
</tr>
<tr>
<td>Directly to the user</td>
<td>✓</td>
<td>✓</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Overtly helpful interlocutor</td>
<td>✓</td>
<td>✓</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Slow</td>
<td>✓</td>
<td>✓</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Careful articulation with pauses</td>
<td>✓</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Very concrete, immediate topics</td>
<td>✓</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>
Validity: Replication/contradiction

Qualitative Analysis: Cambridge Writing scale

> Substantial independent confirmation of salient features of levels from Cambridge draft Common Writing Scale project

> Contradiction very limited and restricted to non-calibrated content elements (of socio-linguistic competence)
What is the purpose of the CEFR?
Where do the Common Reference Levels come from?
What claim to validity have the illustrative descriptors?
How can we ensure consistent interpretation of the levels?
Linking Assessment to the CEFR

Specification (of content in relation to CEF)
  > Description; Coverage: CEF categories, levels

Standardisation (of interpretation of levels)
  > Training with calibrated examples provided
  > Transfer to local examples (videos, scripts, items)

Empirical Validation (of test cut-scores to levels)
  > Internal (test characteristics)
  > External (linking to calibrated tests, descriptors)
External Validation

Correlation

> Is it worth trying to relate the two things. (0.75 = 50% shared variance)

Decision Power

> How many matching classifications are there?
# External Validation: Decision Table

Test under study (Eurocentres Itembank – German)

<table>
<thead>
<tr>
<th>Teachers</th>
<th>A1 (1)</th>
<th>A2 (2+3)</th>
<th>B1 (4+5)</th>
<th>B2 (6+7)</th>
<th>C1 8+9</th>
<th>Total</th>
</tr>
</thead>
<tbody>
<tr>
<td>A1</td>
<td>4</td>
<td>1</td>
<td></td>
<td></td>
<td></td>
<td>5</td>
</tr>
<tr>
<td>A2</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td>18</td>
</tr>
<tr>
<td>B1</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td>20</td>
</tr>
<tr>
<td>B2</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td>19</td>
</tr>
<tr>
<td>C1</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td>6</td>
</tr>
<tr>
<td>Tot</td>
<td>4</td>
<td>20</td>
<td>20</td>
<td>21</td>
<td>3</td>
<td>68</td>
</tr>
</tbody>
</table>
CEFR Levels: Key Problems

- Danger of differing interpretations for different languages
- Under-definition of C2, + some reversals of C1/C2 descriptors (ALTE, DIALANG, Catalonia)
- Weak definition of socio-linguistic competence (and some contradiction to Cambridge qualitative research)
- Unrealistic expectations in relation to receptive skills
Differing Interpretation of the Levels

- Translations, reference levels, samples produced independently, possibly importing problems from 1970s
  - **Lead language; Cross-linguistic benchmarking**
- Use of relative/normative terminology banned from English original (e.g. B2 = “avancé”)
  - **Remove from secondary docs & next printing**
- Use of criteria & samples for older frameworks rather than illustrative descriptors and samples calibrated to them (= indirect linking)
  - **Formally link older frameworks to CEFR; avoid borderline samples**
Under-definition of C2

> Mostly uncalibrated as very few C2 descriptors calibrated in CEFR/Swiss project

> Integrate suitable descriptors from ALTE, DIALANG, Catalonia, Portfolio bank

> Occasional C1/C2 reversals

> Investigate cases; Incorporate insights from qualitative analysis of samples (e.g. Cambridge)

> C1 descriptors tend to be more concrete, C2 descriptors less so – but try to avoid “native speaker” attributes

> Define Level D, at least in outline, to give upper boundary; Consult curriculum descriptors
Weak definition of Socio-cultural

> Mostly uncalibrated as very few C2 descriptors calibrated in CEFR/Swiss project; none in ALTE, DIALANG etc.

> ?? Project ??

> Some contradictions of uncalibrated with Cambridge Common Scale project

> Investigate cases; Incorporate insights from qualitative analysis of samples
Socio-cultural: contradictions  B2

CEFR
> Can express him or herself appropriately in situations and avoid crass errors of formulation. Can vary formulation of what she wants to say.

Cambridge
> Can only occasionally and quite often inappropriately match style of expression to the topic or use idioms correctly.
CEFR

> Can use language flexibly and effectively for social purposes, including emotional, allusive and joking usage.

> Can express him or herself confidently, clearly and politely in a formal or informal register, appropriate to the situation and person(s) concerned.

> Can adjust what he/she says and the means of expressing it to the situation and the recipient and adopt a level of formality appropriate to the circumstances.

Cambridge

> Can make a positive impact by effectively varying style of expression and sentence length for effect, and through the use of idiom and/or humour, though the use of the latter is not always completely appropriate.
Differing Interpretation of the Levels

- Translations, reference levels, samples produced independently, possibly importing problems from 1970s
- **Lead language; Cross-linguistic benchmarking**
- Use of relative/normative terminology banned from English original (e.g. B2 = “avancé”)
  - **Remove from secondary docs & next printing**
- Use of criteria & samples for older frameworks rather than illustrative descriptors and samples calibrated to them (= indirect linking)
  - **Formally link older frameworks to CEFR; avoid borderline samples**
 Danger of differing interpretations for different languages

 Under-definition of C2, + some reversals of C1/C2 descriptors (ALTE, DIALANG, Catalonia)

 Weak definition of socio-linguistic competence (and some contradiction to Cambridge qualitative research)

 Unrealistic expectations in relation to receptive skills
Common Framework of Reference:
“Learning, Teaching, Assessment”

NOT a harmonisation tool
“We have NOT set out to tell practitioners what to do or how to do it. We are raising questions not answering them.”

NOT a theory of language or skills development
Scales describe learning outcomes, learner behaviours, not the invisible processes involved.
CEFR “Dutch grid” – variables didn’t explain difficulty either

NOT a test specification
Scales and lists can be consulted when drawing up a task specification (Ch4) or defining assessment criteria (Ch5) but need reference to detailed specs for language & context
> **Samples**: Benchmark performance samples in international, cross-linguistic seminars (like Sevres for French)

> **Competences**: Define key *assessment criteria* and *salient features* in those categories at each level as seen in samples *across languages*:
  > Confirmation of existing illustrative descriptors
  > Enrichment of existing illustrative descriptors
  > Focus on CEFR weak points (socio-cultural etc)

> **Activities**: Examine other descriptors, esp. C1, C2
  > Calibrated: ALTE, DIALANG, Catalonia
  > Non-calibrated: Portfolio descriptor bank, EAQUALS workshop
The CEFR Levels: Key Points and Key Problems

> North, B. (forthcoming) The CEFR Levels and descriptor scales. 2nd ALTE International Conference, Berlin 19-21 May 2005
