



**NILE**

PART OF THE **INTO** GROUP

**Dragging TEA into the 21<sup>st</sup> C**

**‘Unchain Your Minds’**

**Bolzano, 14 Nov 2014**


**Dave Allan**  
**Director, NILE**



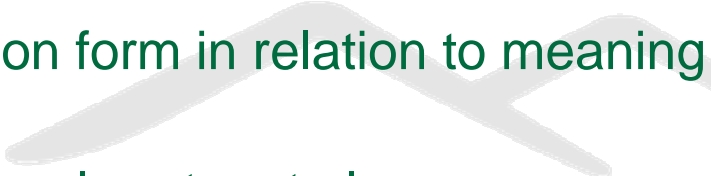
[www.nile-elt.com](http://www.nile-elt.com)


# Taking TEA in(to) the 21<sup>st</sup> Century

**Key changes (and progress?) in  
Testing/TEA in the last half  
century or so**

A faint, light gray graphic of a mountain range, consisting of several peaks and valleys, positioned below the main text.


## **Key changes and trends since 1960 in Testing and in TEA both worldwide and (?) wherever you personally work**

- From word and sentence level discrete-item testing to testing language in context
  - From a G'n'T approach to testing as well as to teaching to a wider range of test types (but N.B. 'Mediation')
  - From a focus on form to a focus on form in relation to meaning
  - From a focus on language as formal system to language as a communicative medium
- 
- A faint, light grey graphic of a mountain range is positioned in the lower right area of the slide, partially overlapping the text of the fourth bullet point.

- From testing the rules of form to testing the rules of use (e.g. 'tags')
  - From most weight and marks being given to knowledge of language system to focus on knowledge and skills
  - **From just testing and examining to testing, assessing, examining and evaluating: the extended assessment repertoire**
  - From test objectives being purely summative, to a balance of formative and summative (but n.b. exams)
- 
- A faint, light grey graphic of a mountain range is positioned in the lower right area of the slide, partially overlapping the text of the fourth bullet point.

- From testing the memorisable to testing and assessing the memorable
- From testing the predictable/preparable to the assessment of spontaneity and flexibility
- From testing written grammar to (also) testing spoken English and discourse in action
- From norm-referenced testing to criterion-referenced testing

- From marks for 'content' in tests called 'language tests' to tests marked for language only (exc. CLIL)
- From inward-looking individualistic criteria to externally recognised, transparent criteria
- From tests designed by isolated individuals to assessment as a collaborative activity
- From tests for the benefit of the teacher and the institution to tests benefiting the learner

- From tests informed only by local perspectives to assessment based on national standards
  - From mutual incomprehensibility over 'levels' to commonly interpreted frameworks (we hope!)
  - From single scores/grades to skills profiles
  - From under-use and misuse of new technology to increasingly effective use of IT
- 
- A faint, light gray graphic of a mountain range is positioned in the lower right area of the slide, partially overlapping the text of the fourth bullet point.

## Some key dates in the last half century of developments in T(TEA)

**1960** Publication of Robert Lado's 'Language Testing'

**1973** Specification of 'Threshold Level'

**1975** Publication of John Munby's 'Communicative Syllabus Design'

**1977** Publication of Keith Morrow's 'Techniques of Evaluation for a Notional Syllabus'



- 1983** Publication of Andrew Harrison's 'A Language Testing Handbook'
- 1987** Publication of Nick Underhill's 'Testing Spoken Language'
- 1989** IATEFL Testing SIG Conference 'Language Testing in the 90s – the Communicative Legacy'
- 1991** Publication of Charles Alderson (ed.) 'Language testing in the 90s – the Communicative Legacy'

- 1994** Publication of Harris and McCann's 'Assessment'
- 1997** Publication of O'Malley and Valdez Pierce's 'Authentic Assessment for English Language Learners'
- 2000** Publication of the first book in the CUP Cambridge Language Assessment Series – Charles Alderson's 'Assessing Reading'
- 2001** Publication of the Council of Europe's 'Common European Framework of Reference: Learning, teaching, assessment'

- 2001** Appearance of the first big ‘CAT’, the Cambridge ESOL/OUP QPT
- 2004** Publication of Sara Luoma’s ‘Assessing Speaking’
- 2004** Publication of the Oxford Placement Tests (3<sup>rd</sup> edition) linked to the CEFR
- 2009** The ‘Watershed’ year for placement testing to make proper use of the new technology:

**OOPT**

**AOC**

**CEPT**

**IUPT**

**PTE**

**‘Password’**

**BCPT**



# 2011 'Examining Speaking' Cambridge Assessment/CUP Lynda Taylor (ed.)

## EVP Project ('English Profile')

<http://vocabulary.englishprofile.org>

## Are we now in a post-methods era in TEA?

- A historical perspective on teaching and TEA
- What did we/do we mean by 'method' in TEA?
- A 'smorgasbord of ideas? (Oller 1993)
- Prescription for practice (this is how you have to do it, e.g. The Berlitz Method, Direct Method)
- Organising principles likely to lead to more effective practice?

**If there have been ‘methods’ in syllabus design and teaching pedagogy, have they been matched in testing and assessment?**

**What were/are the counterparts to ‘Grammar-Translation’, structural syllabuses, Direct Method, AV approaches, The Silent Way, CLL, Suggestopedia, TBL, LCL, CBI, CBT, CLIL and CALL/TALL et al?**

# The key changes and trends in testing and assessment since the mid-20<sup>th</sup> C

1950s G & T for teaching and testing

1960s Psychometric approaches (Lado et al)

1970s Communicative language teaching, but...

1980s More of the same, and still just 'testing'

1990s The start of the moves towards assessment:  
alternative assessment, authentic  
assessment, continuous assessment, peer and self-  
assessment; need for standardising frameworks &  
the development of the CEFR

# And this century.....?

- Testing **and** Assessment
- The CEFR
- Item banking
- Insights from test analysis and item analysis
- Insights from corpora, especially about the nature of spoken English
- Digital delivery including CATs
- Computerised marking (of productive skills??)
- EVP
- LOLA




# So are there ‘best methods’ for testing and assessing or is it just.....it depends?

- It depends why you’re testing and assessing
- It depends who you’re testing and assessing
- It depends where you’re testing and assessing

We are in a post-methods era of T & A, but we have been for at least 20 years.....

and if we want to do it better, then in the 3<sup>rd</sup> sense of ‘methods’, there are some really important ‘organising principles’

# The really important principles

- Good T & A is both formative and summative
  - We need to ensure appropriate levels of validity and reliability
  - We can't rely on computers to assess the 'productive skills', speaking and writing
  - We can't assess fairly without using suitable assessment criteria
  - We can't assess reliably without clear, explicit criteria which properly define the target domains and are open to common interpretation
- 
- A faint, light grey graphic of a mountain range is positioned in the lower right area of the slide, partially overlapping the text of the fifth bullet point.

## **So how do we get it right when we want to assess PLA (productive language ability)?**

- Learn to assess the positive – don't just count errors
- Identify what really matters in speaking and writing
- Use the CEFR and 'Can-Do' statements as guidelines and starting points, but remember that 'Can-Do' statements can't work as assessment criteria.
- Develop assessment criteria designed for context and purpose but consistent with widely accepted standards and provide training in their use

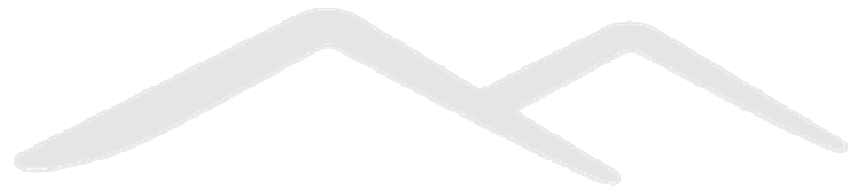
# How can the CEFR help us?

It's as old as this century (begun over 2 decades ago) and was produced by human beings, so is in no way a perfect solution, but it does:

- Provide us with widely recognised language levels and descriptions of them
- Offer us a common language for assessment
- Offer key links for good assessment practice
- Provide a basis for more detailed development
- Provide a framework for benchmarking

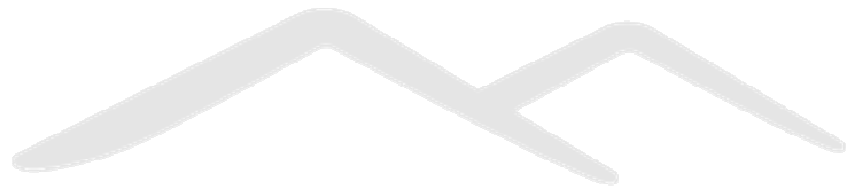
## How does the CEFR work?

- A specification of what users of a language *can do*
- A specification of the competenc(i)es they use
  - *general*
  - *communicative*
  - *linguistic*



## Range

- The range of competenc(i)es learners need to acquire
- The range of areas of language use they may want to become familiar with
- The horizontal dimension

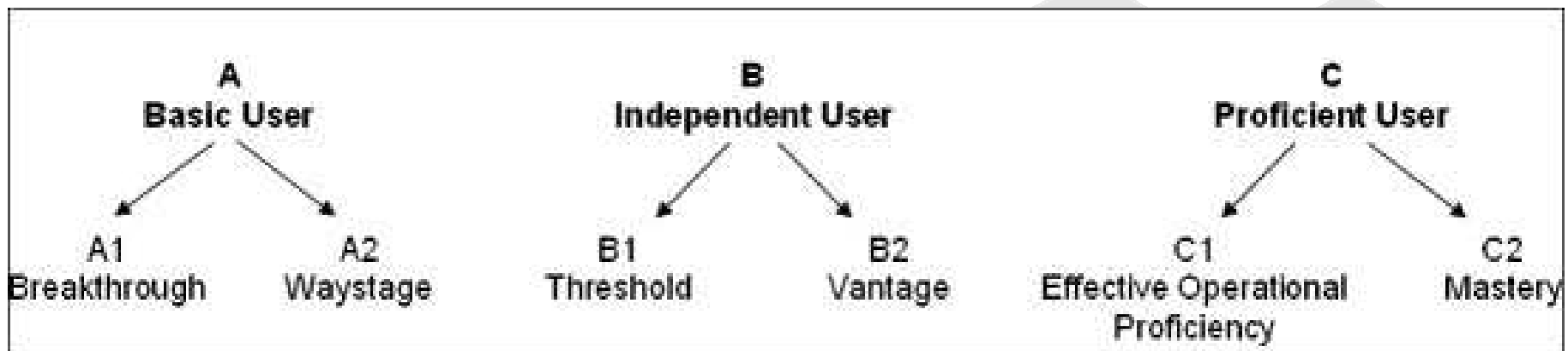


# Levels

Six levels of performance

– A1 A2; B1 B2; C1 C2

- Global scales; specific scales
- All scales are positive
- All scales are hierarchical
- The vertical dimension





PART OF THE INTO GROUP

IUPT Score Range %	CEFR Level	IELTS equivalent bandscore	Cambridge Main Suite Exams	TOEFL iBT	TOEFL CBT	TOEFL PBT	Previous OPT Score Range and Level
82 - 100	Close to bilingual	8.5+		Max: 120	Max: 300	Max: 677	Highly Proficient Expert User
78 - 81	C2 +	8.0 - 8.5	CPE 'B' & 'A'		Above 263	Above 625	Above 190
73-77	C2 'Mastery'	7.5 - 8.0	CPE 'Pass'		263	625	180 - 189
69-72	C1+ to C2	7.0 - 7.5	CAE 'B' & 'A'	120	250	600	170 - 179
64-68	C1 'Effective Operational Proficiency'	6.5 - 7.0	CAE 'Pass'	110 - 120	232	575	160 - 169
59-63	B2+	6.0 - 6.5	FCE 'A' to CAE 'Pass'	101 - 109	213	550	150 -159
54-58	B2 'Vantage - Independent User '	5.5 - 6.0	FCE 'B' & 'A'	94 - 101	195	525	142 - 149
50-53	B2-	5.0 - 5.5	FCE 'Pass'	87 - 94	173	500	135 - 141
45 - 49	B1+	4.5 - 5.0	PET 'B' & 'A'	77 - 86	152	475	128 - 134
40 - 44	B1 'Threshold Level'	4.0 - 4.5	PET 'Pass'	67 - 76	133	450	120 - 127
35 - 39	B1-	3.5 - 4.0	KET 'B' & 'A'	57 - 66	113	425	113 - 119
30 - 34	A2+ 'Waystage +'	3.0 - 3.5	KET 'Pass'		97	400	105 - 112
Below 30	A2, A1 and below	Below 3.0	Below 'Main Suite'		Below 97	Below 400	Verv Limited User



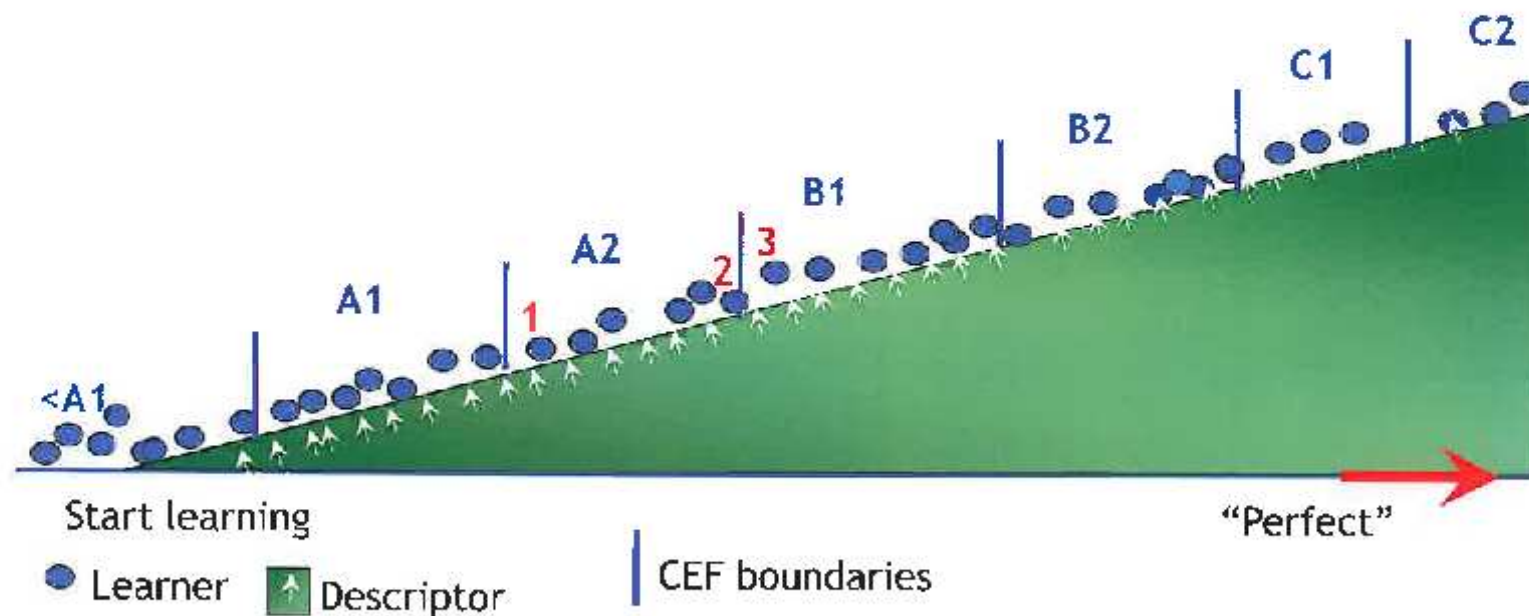
## Criterion referencing and the CEFR

### Intentions and applications

- Despite the previous image, the CEFR is not a set of 6 points, nor a series of plateaus with vertical climbs at level boundaries.



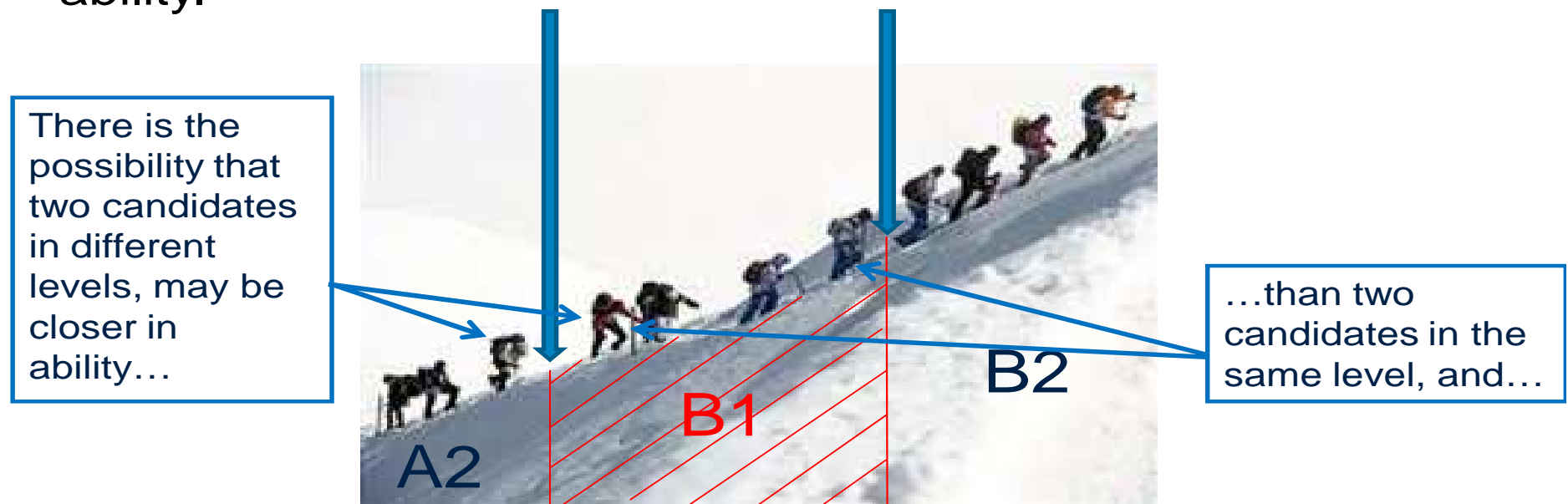
# Increasing language ability & CEF Descriptors

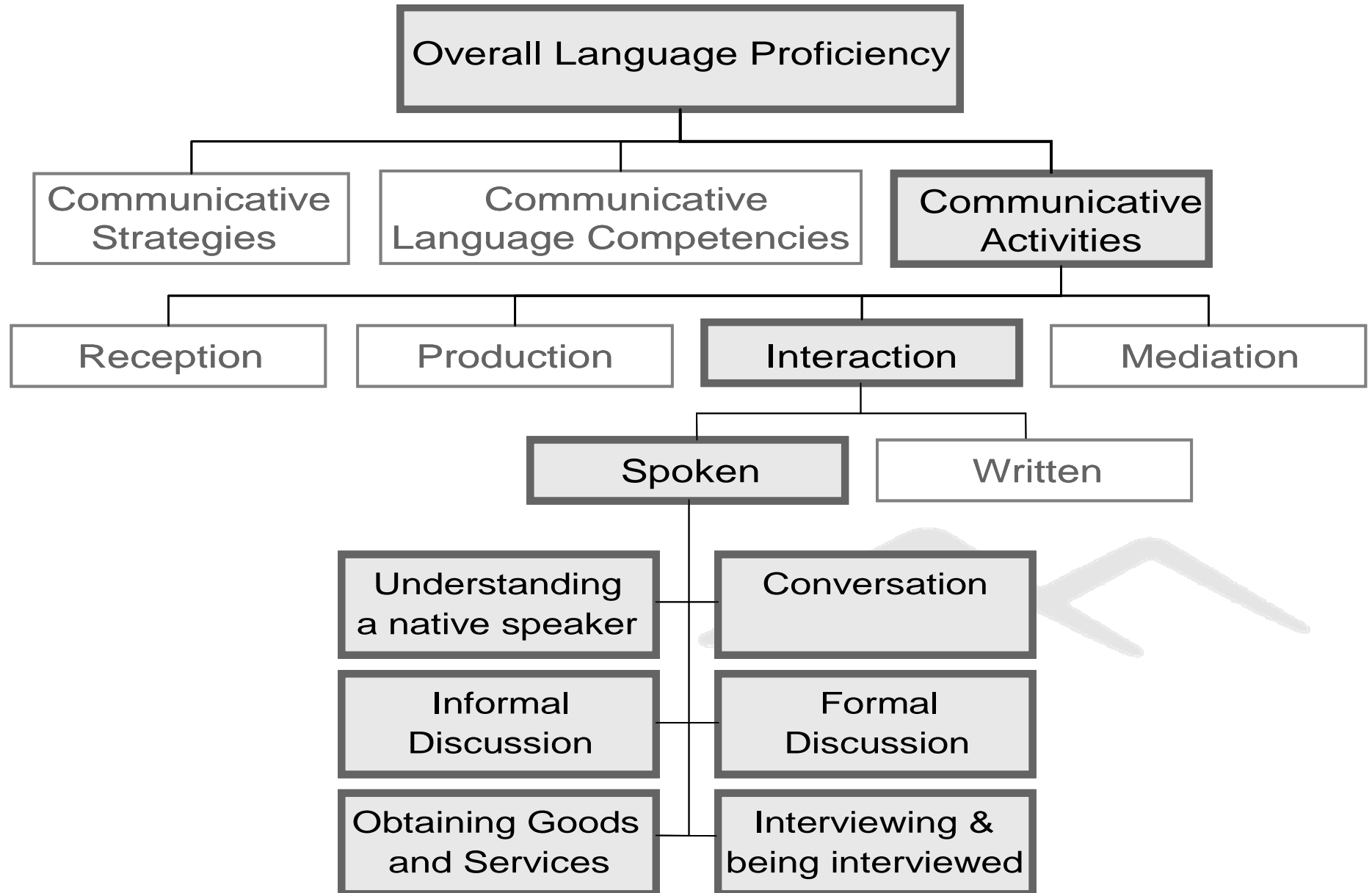


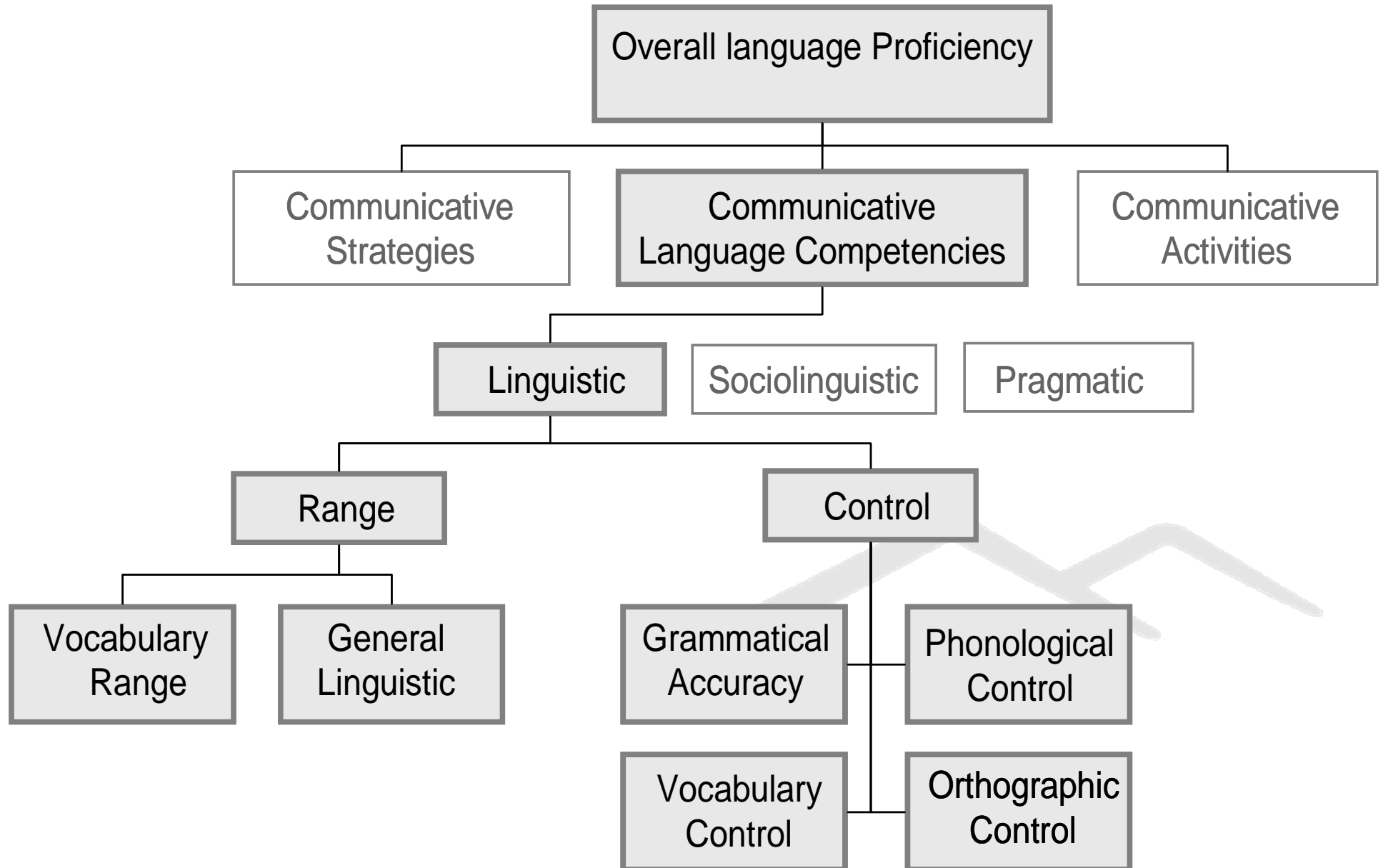
# Criterion referencing and the CEFR

## Intentions and applications

- The CEFR is a continuous scale of progress in language ability.









PART OF THE INTO GROUP

# Common Reference Levels



# Common Reference Levels

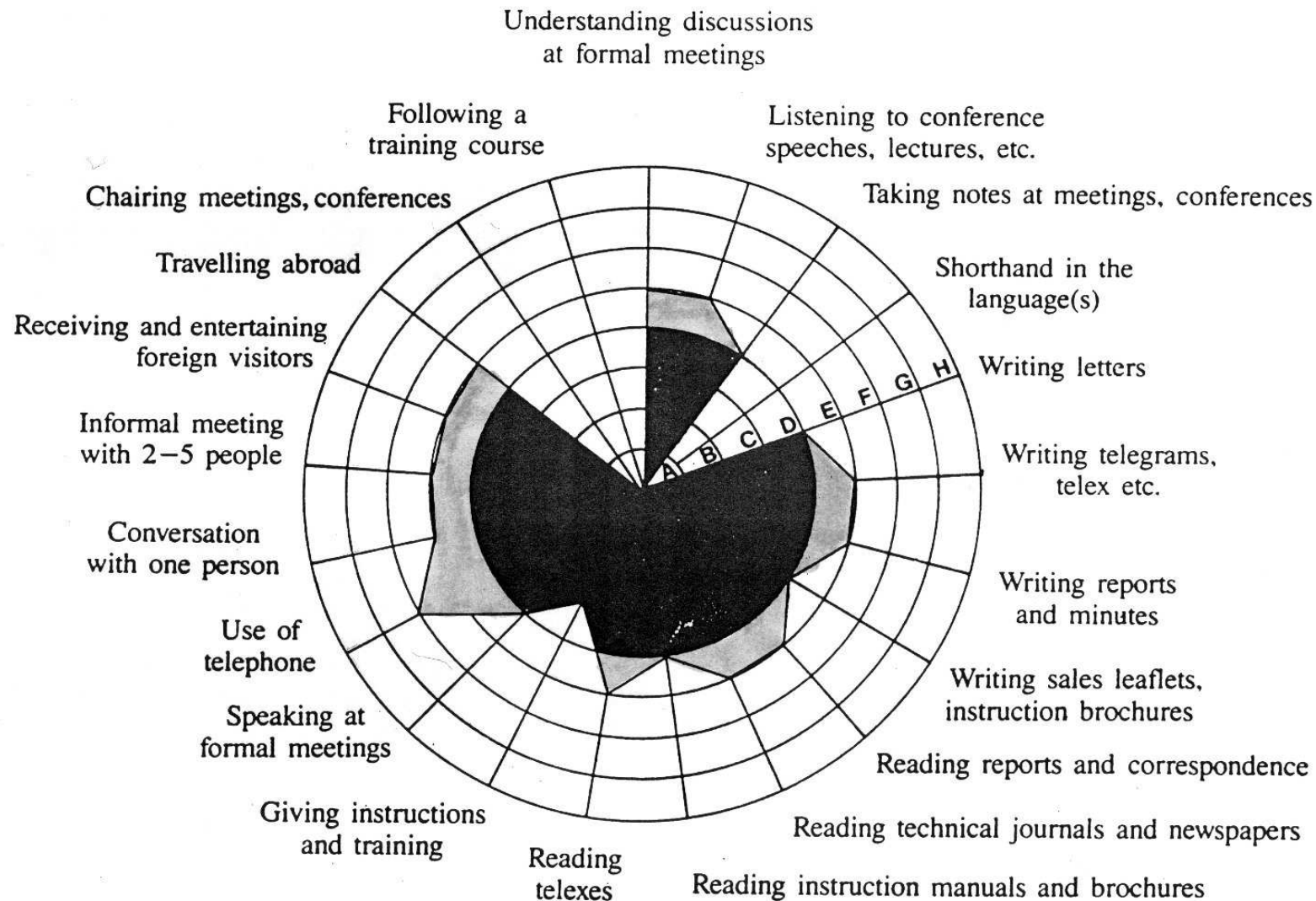


# Common Reference Levels

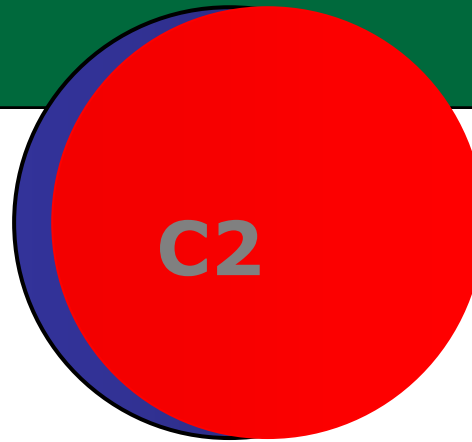
- **20 year process of convergence**
- 1976 “The Threshold Level” – “Waystage”
- 1978 Ludwighafen (Wilkins proposal for levels)
- 1991 Rüschtikon (Hargreaves “natural levels”)
- 1993-6 Swiss research project (North & Schneider)
  
- **Levels and descriptors validated**
  - Qualitative validation (32 workshops)
  - Mathematically scaled (2500 learners, 250 teachers)
  - Post-validation projects (ALTE, Finland, DIALANG, Pearson, English Profile, CEFR-J)



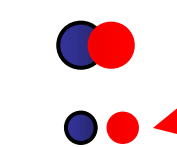
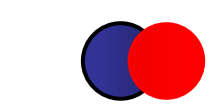
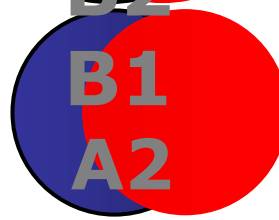
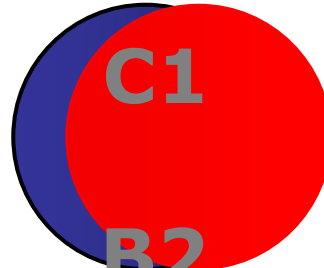
# Profiling – Professional needs



...there is no guarantee that two learners' ability profiles are exactly alike at a given level.



Note that the area of non-overlap is relatively small, but larger than the whole area of A1



< A1: Ability so limited that zero likelihood of overlap

Graphic: J. de Jong, (Prague 2012)

<b>Proficient User</b>	<b>C2</b>	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.
	<b>C1</b>	Can understand a wide range of demanding, longer texts, and recognise implicit meaning. Can express him/herself fluently and spontaneously without much obvious searching for expressions. Can use language flexibly and effectively for social, academic and professional purposes. Can produce clear, well-structured, detailed text on complex subjects, showing controlled use of organisational patterns, connectors and cohesive devices.
<b>Independent User</b>	<b>B2</b>	Can understand the main ideas of complex text on both concrete and abstract topics, including technical discussions in his/her field of specialisation. Can interact with a degree of fluency and spontaneity that makes regular interaction with native speakers quite possible without strain for either party. Can produce clear, detailed text on a wide range of subjects and explain a viewpoint on a topical issue giving the advantages and disadvantages of various options.
	<b>B1</b>	Can understand the main points of clear standard input on familiar matters regularly encountered in work, school, leisure, etc. Can deal with most situations likely to arise whilst travelling in an area where the language is spoken. Can produce simple connected text on topics, which are familiar, or of personal interest. Can describe experiences and events, dreams, hopes & ambitions and briefly give reasons and explanations for opinions and plans.
<b>Basic User</b>	<b>A2</b>	Can understand sentences and frequently used expressions related to areas of most immediate relevance (e.g. very basic personal and family information, shopping, local geography, employment). Can communicate in simple and routine tasks requiring a simple and direct exchange of information on familiar and routine matters. Can describe in simple terms aspects of his/her background, immediate environment and matters in areas of immediate need.
	<b>A1</b>	Can understand and use familiar everyday expressions and very basic phrases aimed at the satisfaction of needs of a concrete type. Can introduce him/herself and others and can ask and answer questions about personal details such as where he/she lives, people he/she knows and things he/she has. Can interact in a simple way provided the other person talks slowly and clearly and is prepared to help.

**B2**

Can understand the main ideas of complex text on both concrete and abstract topics, including technical discussions in his/her field of specialisation. Can interact with a degree of fluency and spontaneity that makes regular interaction with native speakers quite possible without strain for either party. Can produce clear, detailed text on a wide range of subjects and explain a viewpoint on a topical issue giving the advantages and disadvantages of various options.



PART OF THE INTO GROUP

# Qualitative aspects of spoken language use

	<b>RANGE</b>	<b>ACCURACY</b>	<b>FLUENCY</b>	<b>INTERACTION</b>	<b>COHERENCE</b>
<b>C2</b>	Shows great flexibility reformulating ideas in differing linguistic forms to convey finer shades of meaning precisely, to give emphasis, to differentiate and to eliminate ambiguity. Also has a good command of idiomatic expressions and colloquialisms.	Maintains consistent grammatical control of complex language, even while attention is otherwise engaged (e.g. in forward planning, in monitoring others' reactions).	Can express him/herself spontaneously at length with a natural colloquial flow, avoiding or backtracking around any difficulty so smoothly that the interlocutor is hardly aware of it.	Can interact with ease and skill, picking up and using non-verbal and intonational cues apparently effortlessly. Can interweave his/her contribution into the joint discourse with fully natural turntaking, referencing, allusion making etc.	Can create coherent and cohesive discourse making full and appropriate use of a variety of organisational patterns and a wide range of connectors and other cohesive devices.
<b>C1</b>	Has a good command of a broad range of language allowing him/her to select a formulation to express him/ herself clearly in an appropriate style on a wide range of general, academic, professional or leisure topics without having to restrict what he/she wants to say.	Consistently maintains a high degree of grammatical accuracy; errors are rare, difficult to spot and generally corrected when they do occur.	Can express him/herself fluently and spontaneously, almost effortlessly. Only a conceptually difficult subject can hinder a natural, smooth flow of language.	Can select a suitable phrase from a readily available range of discourse functions to preface his remarks in order to get or to keep the floor and to relate his/her own contributions skilfully to those of other speakers.	Can produce clear, smoothly flowing, well-structured speech, showing controlled use of organisational patterns, connectors and cohesive devices.
<b>B2</b>	Has a sufficient range of language to be able to give clear descriptions, express viewpoints on most general topics, without much con-spicious searching for words, using some complex sentence forms to do so.	Shows a relatively high degree of grammatical control. Does not make errors which cause misunderstanding, and can correct most of his/her mistakes.	Can produce stretches of language with a fairly even tempo; although he/she can be hesitant as he or she searches for patterns and expressions, there are few noticeably long pauses.	Can initiate discourse, take his/her turn when appropriate and end conversation when he / she needs to, though he /she may not always do this elegantly. Can help the discussion along on familiar ground confirming comprehension, inviting others in, etc.	Can use a limited number of cohesive devices to link his/her utterances into clear, coherent discourse, though there may be some "jumpiness" in a long contribution.
<b>B1</b>	Has enough language to get by, with sufficient vocabulary to express him/herself with some hesitation and circum-locutions on topics such as family, hobbies and interests, work, travel, and current events.	Uses reasonably accurately a repertoire of frequently used "routines" and patterns associated with more predictable situations.	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 initiate, maintain and close simple face-to-face conversation on topics that are familiar or of personal interest. Can repeat back part of what someone has said to confirm mutual understanding.	Can link a series of shorter, discrete simple elements into a connected, linear sequence of points.
<b>A2</b>	Uses basic sentence patterns with memorised phrases, groups of a few words and formulae in order to communicate limited information in simple everyday situations.	Uses some simple structures correctly, but still systematically makes basic mistakes.	Can make him/herself understood in very short utterances, even though pauses, false starts and reformulation are very evident.	Can answer questions and respond to simple statements. Can indicate when he/she is following but is rarely able to understand enough to keep conversation going of his/her own accord.	Can link groups of words with simple connectors like "and", "but" and "because".
<b>A1</b>	Has a very basic repertoire of words and simple phrases related to personal details and particular concrete situations.	Shows only limited control of a few simple grammatical structures and sentence patterns in a memorised repertoire.	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.	Can ask and answer questions about personal details. Can interact in a simple way but communication is totally dependent on repetition, rephrasing and repair.	Can link words or groups of words with very basic linear connectors like "and" or "then".

	<b>RANGE</b>	<b>ACCURACY</b>
<b>B2</b>	Has a sufficient range of language to be able to give clear descriptions, express viewpoints on most general topics, without much conspicuous searching for words, using some complex sentence forms to do so.	Shows a relatively high degree of grammatical control. Does not make errors which cause misunderstanding, and can correct most of his/her mistakes.
<b>B1</b>	Has enough language to get by, with sufficient vocabulary to express him/herself with some hesitation and circum-locutions on topics such as family, hobbies and interests, work, travel, and current events.	Uses reasonably accurately a repertoire of frequently used "routines" and patterns associated with more predictable situations.

	<b>Pronunciation: comprehensibility and effective communication of meaning, including stress and intonation</b>	<b>Lexical and structural range, flexibility and appropriacy: adequacy of repertoire in relation to tasks and topics</b>	<b>Accuracy: lexical and grammatical - frequency and communicative significance of mistakes and errors</b>	<b>Interaction: initiation, response, independence, participation, negotiation of meaning, turntaking &amp; accommodation</b>	<b>Communicative effectiveness: relevance, coherence, organisation of ideas, sequencing &amp; task achievement</b>
<b>10</b> <b>9</b>	Individual sounds and intonation almost always clear enough to communicate meaning, including in some longer, more complex utterances. Can produce stretches of language of fairly even tempo. Some pausing for grammatical and/or lexical planning and repair will be evident, especially in longer stretches of free production, but not so as to seriously impede comprehension	Has the range to be able to communicate with some flexibility and appropriacy, giving effective descriptions and expressing viewpoints on most familiar topics, without too much conspicuous searching for words, even when using some complex forms to do so.	Can use and adapt a repertoire of frequently used and familiar routines and patterns associated with more predictable situations with reasonable accuracy. While some errors are noticeable, there are few occasions when communication is seriously impaired.	Can initiate, maintain and close simple face-to-face conversation with little hesitation on topics that are familiar or of personal interest. Can repeat back part of what someone has said to confirm mutual understanding.	Can communicate relevantly and effectively in a mainly clear and comprehensible way by interacting coherently and/or linking a series of shorter, discrete simple elements into a connected linear sequence of points, when appropriate in longer turns or on prepared topics .
<b>8</b> <b>7</b>	Individual sounds and intonation generally comprehensible including in some longer utterances. Can mostly keep going without help, despite some problems with formulation resulting in pauses and 'cul-de-sacs'.	Noticeably limited range, but enough to get by, with sufficient vocabulary for largely effective and appropriate expression on familiar topics (family, hobbies, interests, travel) though with some hesitation and circumlocution. Limited evidence of awareness of appropriacy.	Can use some simple structures and expressions correctly but still makes quite frequent errors, even in dealing with familiar topics. Most of these errors, while noticeable, do not prevent communication of the basic message.	Can initiate, maintain and close some of the time in interactions of a familiar kind on simple familiar topics, though with some hesitation which does not significantly impede the interaction.	Can communicate largely relevantly and coherently when interacting or speaking on prepared topics of a fairly simple kind or on personal matters. Can use a wider range of linkers in mostly appropriate ways to produce some generally effective connected sequences.
<b>6</b> <b>5</b>	Individual sounds and intonation are generally clear enough to communicate broad meaning in short utterances. Pauses, false starts and reformulation are very evident. Largely comprehensible.	Has enough range to combine basic sentence patterns with memorised phrases, groups of few words and a few formulaic utterances to communicate limited information in simple everyday situations. Very limited awareness of appropriacy, e.g. some politeness markers	Can use some simple structures and expressions correctly but still makes frequent systematic errors while dealing with simple and familiar topics. While most of these errors do not prevent communication some of them make comprehension difficult.	Can interact in a limited way on simple and/or prepared topics. Can ask and answer questions, respond to simple statements and express simple opinions.. Can indicate when able to follow the interaction but is often unable to initiate, maintain or close the interaction.	Can communicate simple messages on personal and/or concrete matters in a simple way staying relevant to the topic/task and achieving some basic coherence and cohesion using simple linkers like 'and', 'but' and 'because'.
<b>4</b> <b>3</b>	Individual sounds and intonation are not always comprehensible. Communication is noticeably impaired by phonological errors, hesitations and false starts, when attempting to adapt rehearsed expressions to particular situations, even in short routine exchanges.	Has a basic repertoire of words and simple phrases related to personal details and particular concrete situations where the topics are familiar and the required vocabulary has been rehearsed. No evidence of awareness of appropriacy.	The frequency and significance of errors, even when using a limited range of language to handle simple and familiar topics, makes comprehension difficult a lot of the time and sometimes completely prevents communication.	Can manage some simple interactions of the question and answer type, but often fails to interact appropriately even in familiar situations dealing with simple topics. Often requires assistance to keep going.	Can only link words or groups of words with the most basic linear connectors and is unable to manage much relevant effective communication, even on simple, personal and prepared topics.
<b>2</b> <b>1</b>	Frequent and significant phonological errors seriously impede comprehension and prevent clear communication of even basic messages. Can only manage short, isolated and mainly pre-packaged	Extremely limited range of structure and vocabulary. Can only manage a very few words at a time from a very basic repertoire of known and partly-known phrases relating to a restricted set of simple utterances on a few	Even with very simple and familiar topics, lexical, grammatical and phonological errors are so frequent, systematic and serious that most of what is said is incomprehensible or	Can only ask and answer questions about a very limited range of personal details. Can manage minimal interaction on simple topics but communication is totally dependent on	Is unable to manage any level of effective, relevant communication, even on the simplest of topics. Does not have the linguistic repertoire to manage either adequate coherence or any meaningful linking of

<b>Pronunciation: comprehensibility and effective communication of meaning, including stress and intonation</b>	<b>Lexical &amp; structural range, flexibility and appropriacy: adequacy of repertoire in relation to tasks and topics</b>	<b>Accuracy: lexical and grammatical - frequency and communicative significance of mistakes and errors</b>	<b>Interaction: initiation, response, independence, participation, negotiation of meaning, turntaking &amp; accommodation</b>	<b>Communicative effectiveness: relevance, coherence, organisation of ideas, sequencing &amp; task achievement</b>
<p>Generally clear and natural pronunciation with few phonological errors and some evidence of ability to vary stress and intonation to express shades of meaning in contexts requiring detailed communication of fact or attitude</p>	<p>An extended repertoire allowing for the expression of depth of thinking and complex ideas; gaps are easily compensated for without much loss of fluency by effective paraphrase. Register almost always appropriate.</p>	<p>Accuracy is consistently high with very few serious errors; hardly any inaccuracies which seriously impede communication even of detailed meaning, whether fact, opinion or attitude</p>	<p>Highly effective participation in the development of the interaction; contributions skilfully related to those of other speaker(s)</p>	<p>Tasks dealt with almost always fully and effectively; clear, well-structured speech to present and relate quite complex ideas with elaborated use of cohesive devices as and when required</p>
<p>A few examples of incorrect pronunciation may occur but these only rarely affect comprehensibility including with respect to detailed as well as broad meaning</p>	<p>A good range allowing for the expression of substantial ideas with some complexity of supporting argument; generally good use of paraphrase to compensate for lexical gaps and more limited structural range. Register largely appropriate to the task/topic.</p>	<p>Largely accurate, though some lexical errors and minor grammatical inaccuracies occur; few of these prevent or significantly impede communication other than in a few instances where expression of detailed meaning is required</p>	<p>Active participation in the development of the interaction: contributions well related to those of other speaker(s)</p>	<p>Tasks dealt with generally effectively with clear presentation of the relationship between ideas; mostly well chosen use of cohesion devices as and when required</p>
<p>While an L1 accent may be evident and a number of specific, noticeable phonological errors occur, these do not often impede and hardly ever prevent communication of both the import and the detail of the message</p>	<p>Sufficient range to allow for the expression of ideas broadly adequate for requirements; use of paraphrase mostly successful for communication of general meaning but may produce the odd moment of unnatural hesitation</p>	<p>Quite a number of inaccuracies including some noticeable systemic errors, but the majority of these do not lead to any significant misunderstanding and are often repaired by the context and/or by self-correction</p>	<p>Reasonably active participation in the development of the interaction; contributions for the most part adequately related to those of other speaker(s)</p>	<p>Tasks dealt with mostly adequately, though sometimes without the clarity to provide consistent ease of understanding for the listener; use of a more limited range of cohesive devices for the most part adequate as and when required</p>
<p>Communication is impaired by frequent phonological errors, some of which significantly impede and in some cases prevent the effective communication of meaning</p>	<p>Limited range causing noticeable constraints on the communication of ideas relevant to the task, especially in respect of detailed meaning; limited ability to paraphrase effectively when facing new topics or changes of required register</p>	<p>Frequent inaccuracies and errors of a kind which are not only noticeable and intrusive but which lead to a significant incidence of distortion of meaning and the odd moment of communication breakdown</p>	<p>Mainly responsive role in participating in the interaction; contributions often inadequately related to those of other speaker(s)</p>	<p>Tasks often not adequately dealt with; noticeably limited range of cohesion devices, these often being inappropriately applied or missing, when required or desirable to aid effective communication</p>
<p>Frequent and severe phonological errors seriously impede communication of meaning, making important aspects of the message difficult or impossible to understand</p>	<p>Considerable limitations of range make it very difficult to deal with the tasks and communicate meaning effectively other than in the most general ways on familiar topics; very limited ability to paraphrase causes evident lack of fluency</p>	<p>Very frequent errors in attempting to cope with tasks clearly beyond the available range of language; frequently unable to communicate even broad meaning effectively with some instances of communication breakdown</p>	<p>Noticeable lack of attempts to initiate; contributions generally faltering and bearing little relationship to those of other speaker(s)</p>	<p>Frequently unable to deal with the tasks effectively; often lacking in both coherence and cohesion as a result of the inadequate use of, or absence of, even the most basic cohesion devices required</p>
<p>Sample of language inadequate to gain even the lowest mark</p>	<p>Sample of language inadequate to gain even the lowest mark</p>	<p>Sample of language inadequate to gain even the lowest mark</p>	<p>Sample of language inadequate to gain even the lowest mark</p>	<p>Sample of language inadequate to gain even the lowest mark</p>



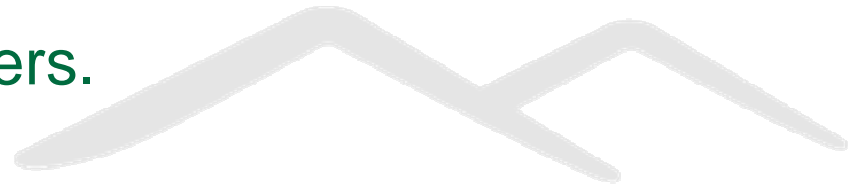
	<b>Pronunciation: comprehensibility and effective communication of meaning, including stress and intonation</b>	<b>Lexical &amp; structural range, flexibility and appropriacy: adequacy of repertoire in relation to tasks and topics</b>	<b>Accuracy: lexical and grammatical - frequency and communicative significance of mistakes and errors</b>	<b>Interaction: initiation, response, independence, participation, negotiation of meaning, turntaking &amp; accommodation</b>	<b>Communicative effectiveness: relevance, coherence, organisation of ideas, sequencing &amp; task achievement</b>
<b>10</b>	Clear, natural pronunciation with almost no phonological errors; able to vary stress and intonation consistently and appropriately to express a full range of shades of meaning in contexts requiring detailed communication of facts, opinions and attitudes.	A comprehensive repertoire allowing for the expression of depth of thinking and complex ideas; hardly any evidence of lexical deficit. Register consistently appropriate.	Hardly any errors. Any occasional mistakes do not impede communication.	Highly effective participation in the development of the interaction, including evident sensitivity to other speakers' discourse and non-verbal features; contributions very skilfully related to those of other speaker(s)	Tasks dealt with consistently well. Clear, well-sequenced and organised articulation of ideas. Able to present complex concepts coherently with consistently appropriate and effective use of cohesive devices.
<b>9</b>	Generally clear and natural pronunciation with few phonological errors and some evidence of ability to vary stress and intonation to express shades of meaning in contexts requiring detailed communication of facts, opinions and attitudes.	An extended repertoire allowing for the expression of depth of thinking and complex ideas; gaps are easily compensated for without much loss of fluency by effective paraphrase. Register almost always appropriate.	Accuracy is consistently high with very few serious errors; hardly any inaccuracies which seriously impede communication even of detailed meaning, whether fact, opinion or attitude	Effective participation in the development of the interaction; contributions skilfully related to those of other speaker(s)	Tasks dealt with almost always fully and effectively; clear, well-structured speech to present and relate quite complex ideas with elaborated use of cohesive devices as and when required
<b>8</b>	A number of examples of incorrect pronunciation may occur along with frequent evidence of L1 influence, but these only rarely affect comprehensibility, including with respect to most aspects of detailed as well as broad meaning	A good range allowing for the expression of substantial ideas with some complexity of supporting argument; generally good use of paraphrase to compensate for lexical gaps and more limited structural range. Register largely appropriate to the task/topic.	Largely accurate, though some lexical errors and minor grammatical inaccuracies occur; few of these prevent or significantly impede communication other than in a few instances where expression of detailed meaning is required	Active participation in the development of the interaction: contributions well related to those of other speaker(s)	Tasks dealt with generally effectively with clear presentation of the relationship between ideas; mostly well chosen use of cohesion devices as and when required
<b>7</b>	While an L1 accent may be consistently evident and quite frequent specific, noticeable phonological errors occur, these do not often impede and hardly ever prevent communication of both the broad meaning and the detail of the message	Sufficient range to allow for the expression of ideas broadly adequate for requirements; use of paraphrase mostly successful for communication of general meaning but may produce the odd moment of unnatural hesitation	Quite a number of inaccuracies including some noticeable systemic errors, but the majority of these do not lead to any significant misunderstanding and are often repaired by the context and/or by self-correction	Reasonably active participation in the development of the interaction; contributions for the most part adequately related to those of other speaker(s)	Tasks dealt with mostly adequately, though sometimes without the clarity to provide consistent ease of understanding for the listener; use of a more limited range of cohesive devices for the most part adequate as and when required
<b>6</b>	Communication is impaired by frequent phonological errors, some of which significantly impede and in some cases prevent the effective communication of meaning	Limited range causing noticeable constraints on the communication of ideas relevant to the task, especially in respect of detailed meaning; limited ability to paraphrase effectively when facing new topics or changes of required register	Frequent inaccuracies and errors of a kind which are not only noticeable and intrusive but which lead to a significant incidence of distortion of meaning and the odd moment of communication breakdown	Mainly responsive role in participating in the interaction; contributions often inadequately related to those of other speaker(s)	Tasks often not adequately dealt with; noticeably limited range of cohesion devices, these often being inappropriately applied or missing, when required or desirable to aid effective communication
<b>5</b>					
<b>4</b>					
<b>3</b>					
<b>2</b>					
<b>1</b>					

<b>Mark</b>	<b>Vocabulary: range, appropriacy and accuracy (including spelling)</b>	<b>Grammar: accuracy and range; communicative significance of mistakes and errors</b>	<b>Organisation: organisational structure; cohesion &amp; coherence; layout and punctuation</b>	<b>Task Achievement: relevance, variety and appropriacy of ideas and arguments; suitability of text type and length for topic and task</b>
<b>5</b>	Very wide range of vocabulary Almost entirely appropriate to topic and task	Almost entirely accurate - very few errors and almost none which prevent or significantly impair communication of meaning Range and complexity/simplicity of structures used almost entirely appropriate to topic and task	Organisational structure almost entirely appropriate to topic and task Consistently coherent development of ideas and highly appropriate use of suitable cohesion devices Formal conventions of layout and punctuation consistently observed	Task(s) almost fully achieved Range of ideas and arguments almost entirely appropriate, with text type and text length highly suited to topic and task Highly effective communication of both broad and detailed meaning
<b>4</b>	Wide range of vocabulary Very largely appropriate to topic and task	Mostly accurate - few errors and very few which prevent or significantly impair communication of meaning Range and complexity/simplicity of structures used very largely appropriate to topic and task	Organisational structure very largely appropriate to topic and task Mostly coherent development of ideas and very largely appropriate use of suitable cohesion devices Formal conventions of layout and punctuation very largely observed	Task(s) very largely achieved Range of ideas and arguments mostly appropriate, with text type and text length well suited to topic and task Largely effective communication of both broad and detailed meaning
<b>3</b>	Fairly wide range of vocabulary Generally appropriate to topic and task	Generally accurate - some errors, few of which prevent but a few of which may significantly impair communication of meaning Range and complexity/simplicity of structures used largely appropriate to topic and task	Organisational structure generally appropriate to topic and task Largely coherent development of ideas and generally appropriate use of suitable cohesion devices Formal conventions of layout and punctuation generally observed	Task(s) generally achieved Range of ideas and arguments mainly appropriate, with text type and text length for the most part suited to topic and task Largely effective communication of broad meaning with some evidence of ability to communicate more detailed meaning
<b>2</b>	Limited range of vocabulary Often inappropriate to topic and task	Often inaccurate - frequent errors, some of which prevent and many of which significantly impair communication of meaning Range and complexity/simplicity of structures used often inappropriate to topic and task	Organisational structure often inappropriate to topic and task Development of ideas often lacks coherence, with limited and/or inappropriate use of cohesion devices Formal conventions of layout and punctuation often not observed	Task(s) not adequately achieved Range of ideas and arguments often inappropriate to topic and task, with text type/length for the most part unsuited to requirements. Often ineffective communication of broad meaning, with very little evidence of any ability to communicate more detailed meaning
<b>1</b>	Very limited range of vocabulary Very frequently inappropriate to topic and task	Very often inaccurate - very frequent errors, many of which prevent or significantly impair communication of meaning Range and complexity/simplicity of	Organisational structure almost entirely inappropriate to topic and task Development of ideas very frequently lacks coherence, with almost no appropriate use of cohesion devices	Task(s) clearly far from being achieved Range of ideas and arguments almost entirely inappropriate to topic and task, with text type/length almost completely unsuited to

Organisational structure almost entirely  
inappropriate to topic and task  
Development of ideas very frequently  
lacks coherence, with almost no  
appropriate use of cohesion devices  
Formal conventions of layout and  
punctuation hardly ever observed

# 3 key principles for assessing

- It has to be criterion-referenced assessment – without good criteria you simply can't achieve the 2 key 'F-words' of testing and assessment.
- Rater reliability really matters.  
NB 'Intra-' and 'Inter-'
- Scoring validity really matters.



## Some ‘tips’ from 2 decades of experience

- You can’t reliably assess more than 5 ‘macrocategories’
- Equal weighting is easier to handle, so cluster correctly
- Avoid ‘double jeopardy’
- Always provide criteria to identify levels above and below your target range
- Always start at the top
- Always start by looking for the positive features of PLA
- Make the marks meaningful – equal steps on the ladder
- Evaluate your assessment grids with scissors!

**So that's it for this plenary session**

**Thank you for being here!**

**If you have any questions,**

**or want to get in touch,**

**it's [dave@nile-elt.com](mailto:dave@nile-elt.com)**

**Enjoy the rest of the conference!**