

Another critique of CEFR is that, although it was based on extensive L2 testing research and consultation with L2 teachers, it has not really been validated by parallel second language acquisition developmental data, for example monitoring how students progress from one level to another, if indeed that is how they progress. The levels make great sense intuitively but a stronger interface between testing research and second language acquisition research would further strength them. Alderson (2007) therefore suggests that the test data need to be verified with test corpus data. Alderson and Little (2007) point out that the CEFR has to date had more impact on the field of testing such as the Association of Language Testers of Europe (ALTE), and especially private companies' testing interests, than on official high school matriculation testing, curriculum design, materials, and pedagogy.

Other limitations of the CEFR are the following:

- (1) It has been used primarily with young adults. With the introduction of foreign language teaching (and assessment) at earlier grade levels CEFR tasks or competencies likely need to be adapted somewhat.
- (2) For content-specific learning (called "language of schooling" in Europe) rather than general-proficiency language teaching and learning, additional modifications might be necessary.
- (3) Although it accounts for second-language pragmatics (appropriateness of language use), CEFR doesn't directly and explicitly take into account cultural or literary knowledge.

V. Other Issues Related to Assessment and Standards

1. Assessing language learners across APEC economies

The previous section highlighted the strengths and limitations of CEFR for potential adaptation in and across APEC economies. Certainly, it has numerous strengths. In considering the matter of adopting or adapting such instruments in APEC, a tension must be acknowledged between the desire to establish comparisons in learning outcomes (or standards) across economies/languages by using well-field-tested instruments, on the one hand, and the need for local autonomy, responsiveness to local contexts, and a sense of agency and ownership of policy/standards/practices on the part of local experts/teachers, on the other hand. Furthermore, borrowing curriculum or assessment instruments developed in a very different educational and geopolitical context does require a full understanding of how and why particular instruments were developed in the first place and how best to use or adapt them.

Within APEC economies presently, according to the 2007 EDNET survey, there are many approaches to testing: from local classroom-based and national standardized instruments to international standardized tests such as those developed by the University of Cambridge, UK. In general, it appears that most APEC language tests are locally developed, but ensuring that tests reflect curriculum contexts/levels and objectives well has been an ongoing concern.

One advantage of using an internationally standardized examination system is that it facilitates comparisons of results across contexts and helps establish the readiness of learners to study abroad or in second-language immersion programs, for example. However, again the

suitability of the assessment tool in the local curricular context must be established. Also, testers and policy-makers must decide whether they wish to assess students' achievement, based on the learning they have done in their coursework (favouring criterion-based assessment), or whether more global proficiency measures, independent of coursework, are sought. The latter would include such standardized tests as the U.S. Educational Testing Services' Test of English for International Communication (TOEIC) and the Test of English as a Foreign Language (TOEFL), and the UK/Australian administered International English Language Testing System (IELTS). Both TOEFL and IELTS are widely used standardized tests for academic English, for international or English-medium education, but their target audience is not school-aged learners but rather students aspiring to study in an English-medium university. That said, these large-scale international tests do provide interesting comparison data across countries/economies and reveal progress toward international English norms, especially referenced to postsecondary education. Of course, such comparisons must be interpreted cautiously, with full recognition first of all that different APEC economies have completely different histories—colonial, postcolonial, or other—with English as a second, foreign, international or lingua franca language; and second, that international standardized test-takers reflect just a fraction of English language learners in language, possibly skewing or inflating scores (based on “the cream of the crop”) or reflecting differences in test-preparedness. The following table provides some data from the new Internet-based TOEFL for all test-takers from September 2005 to December 2006.

TOEFL Internet Based Test Results Sept 05-Dec. 06:

Section score (scaled) means by selected geographic regions and by native country
(Source: ETS, 2007, pp. 10-11)⁴

Native Economy (per APEC)	Number of Examinees	Reading/30	Listening/30	Speaking/30	Writing/30	Total/120
Chile	830	23	24	21	21	89
PR China	20,450	20	19	18	20	76
Hong Kong, China	2,763	18	21	19	22	80
Indonesia	1,875	19	21	19	21	80
Japan	17,957	15	17	15	17	65
Korea	31,991	17	19	17	19	72
Malaysia	920	22	23	20	24	89
Peru	1437	20	22	20	20	82
Philippines	5,882	20	22	22	21	85
Russian Federation	2,922	20	23	22	21	85
Singapore	144	25	25	24	26	100
Chinese Taipei	10,022	16	18	17	19	71
Thailand	3,886	17	19	17	18	72
Viet Nam	2,320	17	17	17	19	71

⁴ ETS (Educational Testing Service). (2007). TOEFL® Test and Score Data Summary for TOEFL Internet-Based Test: September 2005-December 2006 Test Data. Retrieved Jan. 7/08 from <http://www.ets.org/Media/Research/pdf/TOEFL-SUM-0506-iBT.pdf>. The nomenclature of the column on the left has been edited for APEC purposes.

Comparable data have been collected and are freely available online for IELTS and TOEIC as well, and from approximately the same time period (2006). Below the Academic IELTS data are reproduced rather than IELTS General test data, since the majority of test takers (80%) take the Academic version and it's similar in objectives to TOEFL. Selected APEC economies included in this table are highlighted.

IELTS Mean band score by most frequent countries or regions of origin (2006)

International English Language Testing System (IELTS): 9 Bands⁵

	Listening	Reading	Writing	Speaking	Overall
Bangladesh	5.58	5.38	5.38	5.62	5.55
China	5.47	5.80	5.23	5.39	5.53
Germany	7.44	7.23	6.75	7.26	7.23
Hong Kong, China	6.70	6.75	5.91	6.06	6.42
India	6.30	5.82	5.79	6.10	6.07
Indonesia	6.10	6.27	5.43	5.83	5.97
Iran	6.04	5.96	5.81	6.31	6.09
Japan	5.87	5.86	5.33	5.80	5.78
Korea	5.87	5.87	5.36	5.72	5.77
Malaysia	6.93	6.85	6.13	6.41	6.64
Nepal	6.34	5.79	5.71	5.88	5.99
Nigeria	5.65	5.84	6.22	6.93	6.22
Pakistan	5.83	5.58	5.49	5.86	5.75
Philippines	6.68	6.27	6.18	6.74	6.53
Russia	6.49	6.48	5.98	6.68	6.47
Sri Lanka	6.27	5.97	5.93	6.39	6.21
Chinese Taipei	5.52	5.81	5.23	5.66	5.62
Thailand	5.82	5.89	5.28	5.70	5.74
United Arab Emirates	4.99	5.10	4.86	5.43	5.16
Viet Nam	5.59	6.01	5.56	5.70	5.78

Again, these data only capture the mean scores of some of the highest-achieving students in those economies, specifically those who seek opportunities for further study (typically graduate study) abroad. They do not indicate the levels of typical school leavers.

In the teaching of Chinese, the standardized HSK Proficiency Test developed in Beijing and loosely modeled on an older version of TOEFL, is becoming more widely used both inside and outside of Chinese regions for learners of Mandarin. However, there has been insufficient research on its reliability and validity with heritage-language learners in North America, many of whom take it to demonstrate that they satisfy additional-language requirements. A variety of other standardized tests also reviewed by Chen et al. (2008) indicate the range of choices available for test takers who seek international validation of their L2 proficiency.

Whatever tests are used, it can be helpful to try to equate local tests with standardized ones or to map them onto instruments such as CEFR (e.g., Chen et al., 2008) to assist with interpreting results. Many European-language tests have already done so (e.g., French DELF, German TestDAF) and the Council of Europe publishes an online manual⁶ to assist with this kind of equating or referencing to CEFR specifically. For example, some IELTS and

⁵ IELTS is managed by University of Cambridge Local Examinations Syndicate (UCLES), with British Council and IDP Australia. Retrieved Jan. 8/08 from <http://www.ielts.org/teachersandresearchers/analysisoftestdata/article382.aspx>. Some "country/region" names have been edited to reflect APEC economies.

⁶ <http://www.coe.int/T/DG4/Portfolio/documents/CEF%20ref%20supp%20%20intro%20sep.pdf>

Cambridge English examinations are compared or equated below, with the column in the far right representing the CEF(R) levels. This kind of comparison would be very helpful across economies using data from their own standardized tests, and particularly for those economies that encourage mobility for educational and work purposes.

Comparisons of University of Cambridge Examinations and CEF(R)⁷

NQF=National Qualification Framework;
 CELS=Certificates in English Language Skills;
 BEC=Business English Certificates

IELTS	Main Suite	BEC	CELS	NQF	CEF
9.0					
8.0					
7.0	CPE			3	C2
6.0	CAE	BEC H	CELS H	2	C1
5.0	FCE	BEC V	CELS V	1	B2
4.0	PET	BEC P	CELS P	Entry 3	B1
3.0	KET			Entry 2	A2
				Entry 1	A1

Chen et al. (2008) include a table from Educational Testing Service mapping the new TOEFL Internet-based Test (iBT) onto CEFR. The CEFR level B-2, for example, corresponds roughly to the iBT TOEFL total score of 87-109, whereas C-1 is in the 110-120 range. Such mappings are obviously very helpful for nonspecialists who must try to interpret scores across contexts.

Turning to local or national standardized testing within economies, Chen et al. (2008) report on some very impressive, rigorous test development taking place in APEC economies, such as in Korea. For example, the G-TELF (General Tests of English Language Proficiency) in Korea is a criterion-referenced, task-based, diagnostic instrument, based on communicative competence that is suitable for EFL contexts. The test is relevant for general, academic and business settings.

2. Standards for teachers

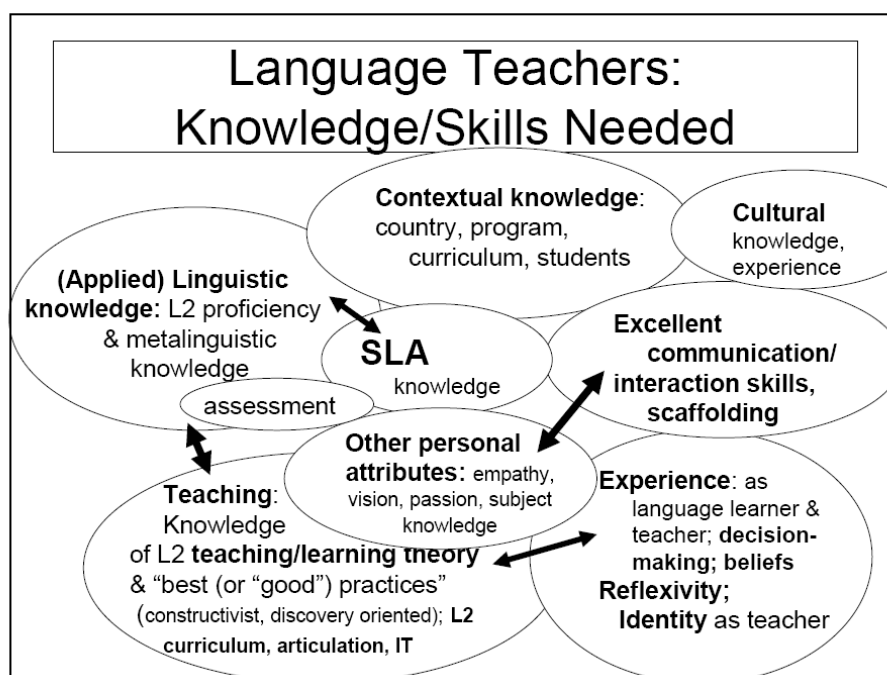
2.1 21st century professional standards and competencies

Up to this point, we have discussed ways of assessing students' competencies in additional languages, especially high-stakes international languages such as English. Here we turn to a discussion of standards for teachers, first in terms of pedagogical competencies for the 21st

⁷ Retrieved Jan. 7/08 from <http://www.ielts.org/teachersandresearchers/commonEuropeanframework>

century, and then in terms of linguistic competency, which is naturally a subset of their overall competency as educators. Most APEC economies have their own standards and procedures for accreditation, assessment, and for professional development. They specify various criterion L2 proficiency levels and many professional knowledge parameters. Again, the question is, are there any widely agreed-upon standards or assessment tools that might facilitate mobility among teachers and also more parsimonious teacher development?

Based on my own work in language teaching and language teacher education for nearly three decades, I suspect that most people would agree that the constellation of knowledge and skills shown in the following figure are needed by language teachers:



They must have considerable knowledge of the curricular context in which they are teaching, must have a high degree of proficiency in, and metalinguistic knowledge of, their own language and of the language they are teaching (if different), they must understand the processes of second language acquisition and principles of assessment, and they must understand culture (e.g., as processes, ideologies, and behaviours shared by groups of learners). In addition, they must have a strong foundation in pedagogy: of best (or sound) teaching practices reflecting 21st century priorities, and they must be effective communicators who know how to organize classroom learning interactionally and in relation to course and curricular objectives. They must also have a number of personal attributes, such as self-discipline, empathy, vision, passion, and subject-matter knowledge; not just knowledge of language and language teaching but also knowledge of the kinds of subjects that students will learn about *through* language. And, ideally, teachers will learn to reflect on their own experiences as teachers in such a way that they improve in their online decision-making, in their planning and assessment, and that they develop identities as teaching professionals.

Some organizations are currently advocating for greater enforcement of, and compliance with, such standards in English language or other modern language programs. In Australia, Ingram (2007) outlines standards for teacher accreditation in use, such as those put forward by the Australian Federation of Modern Language Teachers Associations (AFMLTA, 2005,

Professional Standards for Accomplished Teaching of Languages and Cultures). It stresses and elaborates on the following dimensions:

- Educational theory & practice
- Language & culture
- Language pedagogy
- Ethics and responsibility
- Professional relationships
- Active engagement with wider context
- Advocacy
- Personal characteristics (AFMLTA, 2005, cited in Ingram, pp. 13-14)

Ingram (2007) reports that International Second Language Proficiency Ratings (referred to earlier in Section IV) are used to assess teachers' functional proficiency. The Australia Council of TESOL Associations lists 27 standards for teaching ESL, such as:

- Dispositions toward TESOL
- Understandings about TESOL
- Skills in TESOL

2.2 Teachers' L2 metalinguistic knowledge and proficiency

With respect to assessing language teachers' knowledge of their L2 (if not native-like), it is important to determine threshold levels required for different grade levels, and then determining ways of assessing these fairly and realistically. Proficiency assessment/standards for L2 teachers are also highly relevant for English-dominant economies in which trained immigrant teachers wish to be (re)certified to teach English locally. Again, having international standards or instruments for teachers' L2 proficiency can assist with mobility and also with cross-national/economy research.

In the United States, the powerful international association of Teachers of English to Speakers of Other Languages (TESOL), has recently provided some leadership on teacher standards, both within the United States and elsewhere, such as in China with Chinese teachers of English (China English as a Foreign Language Project).⁸ Four recent publications in a series called *Integrating ESL Standards into Chinese Classroom Settings* reflect this trend: a Teachers' handbook on *Portfolio-based Teacher Development and Appraisal with Teacher Performance Standards* and books focusing on primary to senior levels. Similar standards for the teaching of Chinese as an international language, based on those developed for English, are also currently being developed (Jun Liu, TESOL Past President, personal communication, Nov. 2007).

Within the United States, TESOL has produced an elaborate description of the competencies English language teachers—and mainstream teachers of English-language learners in English-medium mainstream classes—should have,⁹ and that teacher education programs should also focus on (TESOL, 2003)¹⁰. TESOL's model, officially endorsed by a powerful

⁸ http://www.tesol.org/s_tesol/seccss.asp?CID=366&DID=1983

⁹ http://www.tesol.org/s_tesol/seccss.asp?CID=219&DID=1689

¹⁰ TESOL (2003). TESOL / NCATE program standards. Standards for the accreditation of initial programs in P-12 ESL teacher education. Alexandria, VA: TESOL, Retrieved Jan. 1, 2008 at http://www.tesol.org/s_tesol/bin.asp?CID=219&DID=2135&DOC=FILE.PDF.

national accrediting body known as the National Council for the Accreditation of Teacher Education (NCATE), highlights four intersecting circles--*Language, Culture, Instruction, and Assessment*—with a fifth circle, *Professionalism*, at the centre. That model stresses foundations, knowing about language and culture, for example, and applications: planning for and implementing ESL and content instruction, and understanding principles and practices in language proficiency assessment, inside and outside of classrooms.

VI. Conclusion

There is considerable momentum at present toward establishing useful standards for language learning, language teaching, language programs and language teacher education programs, especially for English as L2, but also for other widely taught additional languages. At present, information is being shared across APEC economies vis-à-vis language learning, through ED-NET surveys and syntheses of that material. Additional information that would assist with standards-setting, standards comparisons and cross-referencing, and assessment would likely be beneficial to all stakeholders. Many sources point to the potential for using the European CEFR in particular as a reference point for APEC language teaching and learning standards, for teacher education, and for assessment. Other compatible standards documents for the accreditation of teachers and teacher education programs also identify key areas in which teachers need preparation, in addition to language proficiency.

Possibilities exist for increased communication and sharing of strategies for improving teaching and assessment across APEC economies by the demonstration and annotation of best practices using new technologies, in the manner that has been successfully done with mathematics education lesson studies. Although ongoing attention must be paid to L2 teachers' language proficiency standards and assessment across all economies, English-dominant-economies in particular must continue to find ways to motivate learners—and *teachers*--to study other languages, one way being through better instruction and the use of engaging online and other multimedia 21st century resources and subject matter. Furthermore, more study-abroad programs and student and teacher exchanges, co-op programs, service learning opportunities and better modeling of teaching by language teachers will serve students well, transcending their current circumstances to enable many future possibilities.

References

- AFMLTA, 2005. Professional standards for accomplished teaching of languages and cultures. Canberra: Australian Government Department of Education, Science and Training. (<http://www.afmlta.asn.au>).
- Alderson, J. C. (2007). The CEFR and the need for more research. *Modern Language Journal*, 91, 659-663.
- Brinton, D., Snow, M.A., & Wesche, M. (1989). *Content-based second language instruction*. New York: Heinle and Heinle.
- Byrnes, H. (Ed.). (2007). Perspectives. *Modern Language Journal*, 91, 641-685.
- Crandall, J. 1987. *ESL through content-area instruction*. New Jersey: Prentice-Hall Regents.
- Byrnes, H. (Ed.). (1998). *Learning foreign and second languages: Perspectives in research and scholarship*. New York: Modern Language Association of America.
- Buck, G. (2007). The Common European Framework of Reference for Languages: Learning teaching, assessment. A developing international standard. Paper presented at the APEC EDNET Project Seminar on Language Standards and their Assessment. Taipei, Dec. 3, 2007.