

Issues on Assessment using CEFR in the Region

J. A. Foley

The Graduate School of Human Sciences, Assumption University, Thailand
jfoley@au.edu

Abstract

This paper first discusses the Common European Framework of Reference Languages: Learning, teaching and assessment CEFR 2001 and the revisions to this framework in CEFR 2018. It then looks at the Frameworks of Reference for English Language Education in Thailand, Japan, Vietnam, Malaysia and China which are based on CEFR. The main purpose of this study is to highlight the potential issues that need to be addressed including international forms of assessment such as TOEFL, IELTS, TOEIC and how the rating in CEFR can be compared to the scores obtained in these high-stake standardized tests of language proficiency. The indications are that using the same proficiency scales as the basis for rating scale criteria may lead to perceived equivalence but does not necessarily lead to a greater comparability of shared criteria. The implications from a number of studies are that different tests use similar criteria that are based on the same descriptors, but the comparability is only assumed.

Keywords: CEFR 2001, 2018, ASEAN and East-Asia, Assessment

Introduction

There has been a major concern to establish standards for the user/learner of English within the ASEAN region, and countries in East Asia such as Japan, and China orchestrated in part by the general trend towards globalization. Mainly, focusing on countries within the ASEAN region as well as Japan and China that have implemented versions of CEFR in their education systems, this article will outline the Common European Framework of Reference as applied to users/learners of English in both the 2001 and 2018 versions. There have been a number of modifications made over the intervening years after critical comments were made concerning the 2001 version. Changes have been brought about in the 2018 version particularly in relation to the concept of 'native-speakerism', the importance of plurilingual and pluricultural repertoires of users/learners as well as a 'can do' approach to language competence. A number of these changes are reflected in the adaptation used in the ASEAN region as well as Japan and China, although there are still issues that need to be addressed in particular with reference to assessment both local and international.

The historical background of the Common European Framework of Reference

The history of the Common European Framework of Reference for Languages (CEFR) is linked with the history of the Council of Europe (CoE). Founded after the Second World War, the CoE aims to protect human rights, the rule of law, and parliamentary democracy in its 47 Member States. The CoE strives to advance mutual understanding between nation states, language education, communication and multilingualism which are central to its mission. In 1959, the CoE launched an initiative to support communicative language teaching in Europe. This project was to investigate the possibility of developing a pan-European unit-credit system that would allow language learners to document their foreign language

qualifications in a modular way. These efforts led to the Threshold Level (van Ek 1975); a description of the day-to-day linguistic challenges' migrants faced when living in a foreign country (Deygers, 2019). Since the Threshold level was so popular, the authors were asked to develop additional levels. Initially, they were reluctant to do this, since they did not want to apply a compartmentalized, level-based logic to language learning (Trim 2012). Nevertheless, in order to advance communicative approaches to language learning, *Vantage* and *Waystage* were published as part of CEFR in 2001, incorporating previously established linguistic descriptions, such as *Threshold* (B1), *Vantage* (B2), and *Waystage* (A2), supplemented with newer levels and descriptors, in a framework with a vertical dimension to the levels by mapping them onto a common scale (Deygers, 2019). Additionally, the CEFR continued to promote the mobility of people and lifelong learning, while aspiring to create a common metalanguage to talk about language proficiency across educational systems and international borders (Council of Europe, 2001). The Common European Framework of Reference for Languages (CEFR) describing users as Basic (A1, A2), Independent (B1, B2) and Proficient (C1, C2) has impacted language teaching and assessment in Europe (Figueras, 2012; Barni, 2015) and across the globe (Byram & Parmenter, 2012). CEFR became the most widely used language proficiency framework worldwide. It has impacted language policies, language curricula and language tests (Figueras, 2012) and has also attracted scrutiny and criticism. The criticism typically focused either on the CEFR's use or on its scientific foundations. Usage-based criticism has highlighted that the CEFR allows policymakers to easily use language proficiency levels as gatekeepers without a thorough needs analysis (Barni, 2015). Scientific critique has focused on the development and validation of the level descriptors (Alderson, 2007; Fulcher, 2004), on theoretical gaps in the CEFR's foundation (Hulstijn, 2007), or on the wording of the level descriptors (Alderson 2007). Still others have questioned why multilingualism received comparatively little attention in the scales (Krumm, 2007) and why the CEFR (2001) appeared to uphold a native speaker norm (Barni, 2015; McNamara, 2014). Consequently, the purpose of the recently published CEFR (2018) was to expand, clarify, and update the earlier version. The CEFR (2018) provides new scales for language activities that were not covered in the CEFR (2001) (online communication) and presents more elaborately defined plus levels, pre-A1 levels, and C levels. It also focuses on plurilingualism and foregrounds mediation and new descriptors for sign language users and young learners.

One of the main purposes of CEFR is the promotion of the formulation of educational aims and outcomes at all levels. Its 'can do' aspects of proficiency are intended to provide a shared road-map for learning and a more flexible instrument to gauge progress than a focus on scores in tests and examinations. The principle is based on the CEFR view of language as vehicle for opportunity and success in social, educational and professional domains. This presents the language learner/user as a social agent, acting in the social world and exerting agency in the learning process (CEFR, 2018). The CEFR action-oriented approach represents a move away from syllabuses based on linear progression through language structures, or a pre-determined set of notions and functions. The goal is a communication's perspective guided by what someone 'can do' in terms of the descriptors rather than a deficiency perspective focusing on what the learners have not yet acquired. Fundamentally, the CEFR, as originally devised is a tool to assist the planning of curricula. Courses and examinations can be based on what the users/learners need to be able to do in the target language in their own context. To further promote and facilitate cooperation, the CEFR provides common reference levels A1-C2 defined by illustrative descriptors. However, CEFR is proposed as a tool to facilitate educational reform projects, not a standardizing tool. One of the major issues is whether any adaptations of CEFR might well lead to an over emphasis on testing as a standardized tool of language proficiency. As the CEFR document points out:

One thing should be made clear right away. We have NOT set out to tell practitioners what to do, or how to do it. We are raising questions and not answering them. It is not the function of the European Common Framework to lay down the objectives that users should pursue or the methods they should employ (CEFR: Notes to the User 2018:26).

The message from CEFR (2018) is that language learning should be directed towards enabling learners to act in real-life situations, expressing themselves and accomplishing tasks of different natures. The action-oriented approach puts the co-construction of meaning (through interaction) at the center of the learning and teaching process. The construction of meaning may take place across languages and draws upon users/learners' plurilingual and pluricultural repertoires (translanguaging). CEFR (2018) distinguishes between multilingualism (the co-existence of different languages at the social or individual level) and plurilingualism (the developing linguistic repertoire of an individual user/learner). The fundamental point is that plurilinguals have a single, inter-related, repertoire that they combine with their general competencies to accomplish tasks. Such tasks might require moving from one language to another or giving an explanation in another language to make sense of what is said or written (CEFR, 2018).

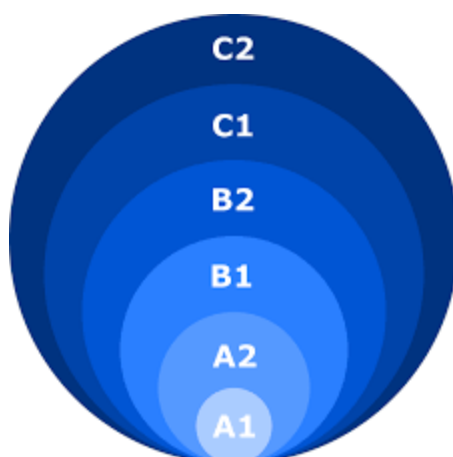


Figure 1: The six Common Reference Levels

CEFR (2018) has two axes: a horizontal axis of categories for describing different activities and aspects of competence, and a vertical axis representing progress in proficiency in those categories. To facilitate the organization of courses and to describe progress, the CEFR (2018) presents the same six Common Reference Levels providing a roadmap that allows user/learners to engage with relevant aspects of the descriptive scheme in a progressive way. However, the six levels are not intended to be absolute (CEFR, 2018).

There has been a tendency for some educational bodies and testing organizations to use these categories without the flexibility intended, but in fact, all categories in language testing are conventional and socially constructed concepts. Like the colors of the rainbow, language competence is a continuum, both vertical and horizontal. As with the rainbow, despite the fuzziness of the boundaries between colors, we tend to see some colors more than others (CEFR, 2018). CEFR was never considered to be a completed or standalone document, indeed supporting work on CEFR scales had started in 2005 with the English Profile Programme (EPP) (Green, 2012). Cambridge University has been developing reference level descriptions (RLD's) of English that provides language specific guidance for each level of CEFR. Komorowska (2004) had found that teachers and teacher trainees did not like the CEFR's lack of guidance for choosing curriculum options, nor the non-evaluative approach

to teaching methods. Costa (2007) also expressed doubts about the empirical and statistical validation outside the original Swiss context where it was being used. Hulstijn (2007) indicated that the empirical foundations of the CEFR scales were based on the judgements of teachers and experts and not on Second Language Processes or research. Poszytek (2012) warned publishers not to use CEFR's global scale or 'can do' concept to sell their textbooks as they were often misaligned with the CEFR scales and had limited theoretical background.

However, English Profile Project and the British Council- EAQUALS Core Inventory for General English were developed to provide language support with more finely tuned contextually, discrete language points in both global and illustrative scales (North et al. 2010). Equally important was the vertical and horizontal dimension of language development reflecting the fact that users develop their overall communicative language competence by improving the quality of their language (vertical development) and expanding the breath of communicative activities that they are engage in (horizontal development). Indeed, the CEFR's concept of partial competence can help in appreciating that language development does not solely have to be about moving up the vertical scale of complex language use. Broadening performance ability in communicative activities and strategies across domains is seen as equally important.

Proficiency scales of CEFR

In order to avoid inconsistencies between the CEFR (2001) and the CEFR (2018) the CEFR (2001) scales were included in the CEFR (2018) in their original form. Today, some descriptors read as outdated ('watching TV news' rather than on a 'smartphone'), or ('propose a toast' [at A1]). Moreover, their focus on target language use contexts, such as leisure, travel, and especially academia ('Can present a topic in a short report or poster' [at B1]) may diminish their applicability in a global community of language learners, where less than 10 per cent have a university degree (Barro & Lee, 2013). One of the most noticeable changes relates to the use of the term 'native speaker', which has been replaced with speakers of the target language (Deygers, 2019). Even though this change corresponds to current thinking in applied linguistics (Houghton et al. 2018) it does present a problem. The term 'native speaker', as used in the CEFR (2001), implies a competent, fluent language user who is able to convey and comprehend nuanced and idiomatic language use. The CEFR (2018) does not specify the proficiency level of speakers of the target language. The idea of uneven proficiency profiles is referred to as 'partial competence' which is significant because it recognizes that a language user's proficiency is fundamentally uneven. No two users share the same language profile, as even the most proficient language user is unlikely to have the same proficiency across all CEFR scales. Another major change in CEFR (2018) is the focus on 'mediation'—an activity whereby 'the user/learner acts as a social agent who creates bridges and helps to construct or convey meaning, sometimes within the same language, sometimes from one language to another' (Council of Europe, 2018 p. 103). Mediation was discussed in the CEFR (2001), but in the CEFR (2018) it has become such a central concept that it is listed as one of the four primary communicative language activities and strategies. Theoretically, mediation is loosely defined in CEFR (2018). It is stated that subdividing communication into reception, production, interaction, and mediation is 'closer to real-life language use, which is grounded in interaction where meaning is co-constructed' (Council of Europe, 2018 p. 31). When we use a language, several activities are involved; mediation combines reception production and interaction. Also, when we use language it is not just to communicate a message, but rather to develop an idea through 'linguaging' (using more than one language), for example, articulating our thoughts to facilitate understanding and communication. However, operationalizing mediation as a rating criterion might well present

a challenge for test developers (Deygers, 2019). They will need to consider how to reliably and validly score what as times can be a vague construct from other constructs.

In spite of these objections, mediation does fit the CEFR's communicative approach, links in with the CoE's values, and can lead to more communicative language teaching and testing. In fact, language testing organizations have already started developing integrated, communicative language tasks in response to the mediation scales. Even so, it remains to be seen if and how mediation will be scored and operationalized. Mediation has been subdivided into more than 20 subscales, and the CEFR (2018) includes over 30 new or redeveloped scales. These proficiency scales are introduced by a brief definition. In itself, that is a useful addition, but sometimes the dense writing style makes these introductions hard to grasp. Users who were troubled by the style of the CEFR (Martyniuk & Noijons, 2007) will find the CEFR (2018) a challenging read as well. The sometimes-vague language (Alderson, 2007) that was a problem for language testers working with the CEFR (2001) is still there in the CEFR (2018). On a more strategic level, the attention devoted to proficiency scales in the CEFR (2018) does mark a shift. CEFR proponents have often stressed that the scales are illustrative only and have become more important than originally intended (Trim, 2012; North, 2014). Consistency with this argument would have dictated that the CEFR (2018) focus less on the scales, not more. However, the writers of the CEFR (2018) have embraced the idea that for most users, the scales are the framework (Martyniuk & Noijons, 2007). The centrality of the scales in the CEFR (2018) warrants a closer look at their methodological foundation. The authors have provided ample documentation (North & Piccardo, 2016) to explain how they were designed and validated. These reports, however, do not justify why the validation and revision relied on language professionals judging and mapping descriptors in a contextual vacuum. It would seem that no learner performances were involved in the construction or calibration of the new scales. (Deygers, 2019). Moreover, even though the authors signal the danger of circularity in validation processes, it is unclear how circularity was avoided when participants had to demonstrate their knowledge of the CEFR before taking part in the validation of new descriptors (Deygers, 2019). It is difficult to see how CEFR (2018) remains true to the CEFR's approach to scale development when this methodology has received such fundamental criticism (Alderson, 2007; Fulcher, 2004). Studies have repeatedly shown that the original CEFR scales are largely unsuitable to reliably compare performances across or within educational systems (Deygers et al., 2018). It is still unclear how the CEFR (2018) will bring us closer to achieving this goal, when they seem to replicate a disputed methodology.

Summary of the major modifications/additions in CEFR 2001-2018:

- developing the illustrative descriptors of second/foreign language proficiency
- to produce versions of CEFR for young learners (7-10\11-15) and for sign language
- to develop more detailed coverage in the descriptors for A1 and the C levels
- complement the original illustrative scales with descriptors for mediating a text mediating concepts, mediating communication
- the provision of descriptors for plurilingual/pluricultural competence
- the removal of any reference to 'native speaker' being replaced with speakers of the target language
- the proficiency level of speakers of the target language is not specified and uses the term 'partial' competence, arguing that language users' are fundamentally uneven in different contexts

Adapting the Framework of Reference for English Language Education in the region

The CEFR was originally envisaged as a planning tool to provide a common language for communication between various participants in educational contexts across the member states of the Council of Europe and has now spread to other parts of the world including USA, South America and Asia and the Asia-Pacific. This involved a development of a ‘tool-kit’ of materials to provide guidance in making the best use of the CEFR for different purposes, such as elaboration of curricula and syllabuses for producing textbooks and developing assessment systems (Saville, 2016).

Recently frameworks based on CEFR have been adopted, with modifications as a proficiency benchmark for both English teachers and students in Thailand, Japan, Vietnam, Malaysia, and China. The CEFR version in Thailand (FRELE-TH) is one of the most recent attempts to do this and will be discussed in some detail to indicate the main global and illustrative scales used in CEFR in terms of English. Japan, Vietnam and Malaysia will be discussed briefly in relation to when they were first fully implemented and the common issues that arose. However, it should be pointed out that CEFR frameworks outlined here are all based on CEFR 2001 albeit modified for local requirements. It might also be noted that CEFR can be applied to any language, not just English as it has been in Indonesia (in French) for university students.

Thailand (FRELE-TH)

Thailand is ranked 53rd among 80 non-native speaking countries in Education First Standard English Test (2017) with a score of 49.78 which is classified as low proficiency. According to the Thai Minister of Education, 40,000 Thai English teachers were tested using Cambridge English standards. Only 6 scored at C level, indicating fluency, 350 score at B level or intermediate, while the majority was at advanced beginners’ level (*UNDP Report*, 2015; Mala, 2016, *Bangkok Post*, 9 August 2018). English, however, plays an increasingly important role in international communication for people in the region. This has seen an even greater emphasis with the ASEAN Economic Community Integration (AEC). With a view to enhancing the English abilities of Thai people to cope with and perform effectively in this changing context, in April 2014, the English Language Institute (ELI), a branch of the Ministry of Education (MOE) announced a policy of basing all aspects of English language curriculum reform on the CEFR framework. A local version of Common European Framework of References for Languages-Thailand, FRELE-TH (2018) was published including Evaluation and Accreditation of Quality Language Services (EAQUALS). The FRELE-TH has two scale types to describe the English proficiency levels: a global scale (overall descriptors) and illustrative scales, (communicative activities, communication strategies, and communicative language competence).

The context of introducing CEFR can be very different as the example of Thailand illustrates: the poor levels of English, poorly-trained teachers, poorly-motivated students and rare opportunities for students to have exposure to English outside the classroom (Dhanasobhon, 2006). The rationale behind the development of FRELE-TH lies in the principle of CEFR’s inception that CEFR does not offer ready-made solutions but must be adapted to the requirements of particular contexts. In order to meet these objectives a 10-level reference framework was developed as an adaptation of CEFR to make it relevant to English use in local and international communication in Thailand (Hirburana et al., 2018). English is one of the working languages in the ASEAN Economic Community (AEC), not only in education but for job applications and work promotion (Pitsuwan, 2014). However, many people do not have satisfactory proficiency in English. This is despite the 9-12 years that Thai

students spend in learning English in formal education. Consequently, the English language reform policy in Thailand (Ministry of Education, 2014) announced the use of CEFR in the design of language curricula, learning goals, testing and assessment as well as the development of the teaching (Hiranburana et al, 2018). It was hoped that the FRELE-TH global scale could be used for the design of specifications on the high-stakes standardized tests of English proficiency, the results of which can be benchmarked with those of international standards. In this way, in principle, students and users' performance and progress can be measured and tracked to be calibrated with other international standards for educational and professional purposes (Hiranburana et al, 2018).

Table 1: FRELE-TH Equivalency

FRELE-TH	CEFR Proficiency level	FRELE-TH level	
A1	A1	1	Basic User
A1+		2	
A2	A2	3	
A2+		4	
B1	B1	5	Independent
B1+		6	
B2	B2	7	
B2+		8	
C1	C1	9	Proficient
C2	C2	10	

Adapted from FRELE-TH (2018) based on CEFR (2001)

The FRELE-TH used the plus (+) levels to make sure that Levels A (Basic User) and B (Independent User) were not too high for Thai learners to achieve these levels of performance. Outlining a more detailed breakdown of levels makes sense given that the users are EFL learners (North, 2004, p. 48) as it shows that the FRELE-TH framework following CEFR is flexible, allowing levels and categories to merge and sub-divide as appropriate. FRELE-TH, descriptors were written to be more comprehensible and relevant to Thai learners and users of English. For example, A1 level has taken into consideration the fact that users of English begin with words, phrases and simple expressions with 'Can Do' statements on familiar topics and immediate surroundings.

Examples of 'Can Do' descriptors

A1 learners

Can recognize simply vocabulary and basic expressions concerning themselves or their family. Can understand and reply to simply expressions spoken very clearly and slowly

A2 learners

Can use basic sentence patterns and groups of phrases to communicate and describe personal information, routine activities and requests

B1 learners

Can understand the main points of clear speech on familiar topics. Can work out the main points they want to communicate in a range of contexts.

B2 learners

Can understand the main ideas of complex speech on concrete and abstract topics, including

technical discussions in their field of specialization.

C1 learners

Have a good command of vocabulary including some idiomatic expressions and speaks fluently

C2 learners

Have no difficulty in understanding any kind of spoken language, whether live or broadcast. Can express thoughts fluently and effectively. Can use a variety of cohesive devices in written language to produce a coherent and cohesive text.

FRELE-TH adopted components from EAQUALS (North, 2008), the Threshold Level (Trim & Trim, 1980; van Ek & Trim, 1990), the Core Inventory of General English (North, Ortega, & Sheehan, 2010), the English Profile Program (Salamoura & Saville, 2010) and the Word Family Framework (West, 2015). The FRELE-TH used the plus (+) levels from the Swiss Project (Goullier, 2007) to make sure that Levels A (Basic User) and B (Independent User) were not too high for Thai learners to achieve these levels of performance (Hiranburana et al., 2018). Similar practice can be seen in the CEFR-J for use in Japan (Negishi et al., 2013, p.156-163) and in China with three stages divided into nine levels (CSE, 2018). The rationale behind the development of FRELE-TH lies in the principle of CEFR's inception that CEFR does not offer ready-made solutions but must be adapted to the requirements of particular contexts. Recently a consortium of 12 Thai universities was proposed to help implement CEFR and expand changes in the local education system at all levels of education.

The establishment of the ASEAN Economic Community (AEC) in 2015 with its internal labor market and English as its sole working language raised concerns about the nation's economic competitiveness. The adoption of CEFR and the contracting of the British Council to deliver a CLT-based training program for Thai English language teachers were presented as a possible solution to Thailand's English language problems (Mala, 2016). In fact, Thailand was rather late in joining a global trend of countries embracing CEFR to reform their English language curriculums and assessment mechanisms. Japan in 2012 developed CEFR-J to suit the Japanese EFL context by re-mapping the 'can do' statements and subdividing the lower proficiency levels and by adding sub-levels to allow for more differentiation at the levels relevant to the majority of Japanese learners ((Tono, 2012). In 2008 Vietnam ratified 'Project 2020' to improve English language proficiency by basing the reform efforts around a CEFR framework to facilitate the teaching of English under Vietnamese conditions (Chung, 2014).

Japan (CEFR-J)

Japan used a modified version CEFR-J to ensure that the framework reflected its local standards in teaching and learning, curriculum development as well as assessment (Bucar et al. 2014). Part of the impetus for change came from the need to transition from a knowledge-based English curriculum to a competency-based language one. Stakeholders' consent for a new skill-based language curriculum was more in favor of curriculum objectives that aimed at marketable results on reputable language proficiency tests (Moser 2015). However, it was also realized that the proficiency level in English of students enrolling in tertiary education was too low to achieve the proficiency test results required. It was suggested that CEFR's globally recognized 'can do' scales could be used as these scales identified language gains at the lowest levels of language proficiency. The CEFR-J the Japanese Ministry of Education, Culture, Sports Science and Technology (MEXT, 2011) published a report encouraging the use of 'can do' lists in junior and senior high schools.

The ‘can do’ lists, in addition to using the CEFR descriptors, triangulation was used with banks of descriptors for EAQUALS/ALTE, ELP as well as textbooks influenced by CEFR such as Longman’s *Total*, and Cambridge University Press’ *English Unlimited*. Negishi et al, (2013) survey of Japanese EFL users indicated that 80 per cent were between A1 and A2. CEFR-J unlike CEFR introduced scales using a branching approach with narrower levels of A1+ and A2+ B1+ and B2+ to make CEFR more useable in the Japanese context (Negishi, Takada and Tono, 2013). It was felt that this increase in levels allowed teachers to better fine-tune student assessment, which meant being able to create more separation between students within a band. This use of CEFR-J scales allowed students of near A2 or A2 students who did not see their progress improve on the vertical scales in the initial stages of the program because of the time needed to acquire skills to be considered as A2+ or B1. As North (2007b) pointed out a branching approach with its narrow levels would allow teachers and students to see more progress, which especially at the earlier levels is critical for developing motivation. A drawback of this narrower scaling was distinguishing these sublevels became more nuanced and created a little more variability in teacher assessment.

Vietnam: CEFR-V

The Vietnamese Ministry of Education and Training (MOET) in 2008 officially began to use CEFR to define English language exit benchmarks for students ranging from primary through to tertiary levels of education. The national project Teaching and Learning Foreign Languages in the National Education System 2008-2020 (Hung 2013) expected all university graduates not majoring in languages to reach B1 English. MOET also adopted CEFR levels A1 (beginners) A2 and B1 as the required standards for students leaving Primary, Junior and Secondary High schools (Nguyen & Hamid, 2015). However, in a meeting organized by the education ministry, university and government representatives it was reported that the government’s targets for language proficiency were too ambitious (Nguyen and Hamid, 2015). According to a survey, only one in five students achieved that level in 2015. The consequence was that institutions had to lower the requirement to A2. The reasons given for not reaching the targets were the teachers’ poor English, lack of resources and outdated teaching methods with a heavy focus on traditional grammar. The government has reportedly moved some of the objectives of the language learning and teaching plan to 2025. A new approach was to be undertaken, creating CEFR-V, a Vietnamese version, similar to CEFR-J. The Management board for the National Foreign Language Teaching Program indicated that the original framework would be adjusted to make it more suitable for Vietnamese studying foreign languages. However, because it was felt that it would take a long time to fulfill the English teaching program, with MOET now focusing on training teachers of English, it is expected that Vietnam would need 100,000 teachers to fulfill the program’s objectives (Viet, 2015).

Malaysia: CEFR-M

The proposed implementation of CEFR in Malaysia started with the establishment of English Language Standards and Quality Council (ELSQC) 2013. Indeed, the alignment of the education system against CEFR has become an important element in the Malaysia Education Blueprint (MEB) with the aim to boost the level of education to international standards (Azman, 2016). The implementation of MEB brought about an additional impact on English Language education especially in primary schools such as the inclusion of English literacy in the Literacy and Numeracy Screening or the LINUS program. However, in 2018, the government introduced CEFR as well. Students’ proficiency was to be graded using CEFR

descriptors in order to ensure that the students' grades were recognized at international levels. LINUS 2.0 was seen to give more emphasis on English language literary skills together with numeracy, consequently, assessment was shared between LINUS 2.0 and CEFR. Only the first three levels of CEFR descriptors (A1, A2, B1) were to be used with LINUS 2.0 because of the low proficiency of the students and the fact that students might progress at a slower pace. However, it was felt that the integration of CEFR into existing programs had to take into account the reality of the Malaysian education landscape as well as whether the CEFR-LINUS screening program assessment is really measuring what it is intended to measure (Ishak and Mohamad, 2018).

CEFR has been integrated as part of the Malaysian roadmap with an overall plan covering a long-term goal from 2013 to 2015 having the aim of providing the best language education starting from pre-school up to tertiary education. The roadmap consists of three phases. Phase 1 (2013-2015) focused on raising the level of English proficiency of teachers. Phase 2 (2016), in the first part appropriate CEFR levels were to be matched against educational levels starting from pre-school to teacher education. While the second part of Phase 2, School Based Assessment (SBA), syllabus and curricula were also to be aligned with CEFR descriptors (National Education Blueprint, 2013). Finally, Phase 3 was for ELSQC to evaluate, review and revise the implementation of CEFR.

China's Standards of English (CSE)

China's Standards of English Language Ability (CSE) (2018) has been developed by the National Education Examinations Authority (NEEA) to develop a national framework of reference for English language education. The management structure of education in China had different governmental departments taking charge of education at different stages. One of the issues arising from such management structure was the inconsistent learning objectives specified in the curricula for learners of English at each educational stage (Yan et al., 2017). Another issue was reflected in the proficiency levels of national assessment aligned to the curriculum at each stage. National tests were developed and administered by different testing organizations. The introduction of a common English proficiency scale it was hoped would facilitate test construction and score interpretation. Added to this is the challenge of globalization by making the education system more transparent to the outside world. China has developed a nine-level scale similar to CEFR (2018) with plus levels, so that the standards of English language education can be aligned to international frameworks and thus prepares Chinese people to become global citizens. The descriptive framework has subdivisions of organizational knowledge (grammatical and textual); pragmatic knowledge (functional and sociolinguistic) and interpreting and translation following the genres outlined in sociolinguistic knowledge. In general, this seems to reflect a much more 'functional' approach to language knowledge than in the original CEFR (2001) document. For example, sociolinguistic knowledge is subdivided into genres, dialects/varieties, registers, idiomatic expressions and cultural figures of speech. As mediating activities, interpretation and translation occupy an important place in the linguistic functions of Chinese society and are taught as a language skill at tertiary level of education. Issues have also been identified particularly with the use of CEFR for developing examinations. Papageorgiou (2010) identified problems with some of the descriptors when used for setting cut-off scores, as CEFR was not designed specifically for test specifications. More importantly, in the Chinese context, the CEFR (2001) 'can do' descriptors were too narrowly focused to be useful for teachers to reflect on teaching and constructing a teaching syllabus. A key difference between CEFR (2001) and CSE is in the target users. CSE is intended for Chinese learners of English

at all educational stages, whereas, CEFR was developed principally, to aid foreign language learning in an adult context in Europe.

The implications of CEFR approaches in the region

Competency in a language is a multi-dimensional system that accounts for the situations, the functions, the linguistic elements needed in communicative competencies. However, measures of language competency can be arbitrary. North (2000) pointed out that CEFR as originally designed was a common measure for recording language competence and that the motivation for a common framework was more pragmatic (thus the ‘can do’) rather than academic. However, as we have seen, there were some inherent limitations in the original version of CEFR (2001) which did affect its applicability, not only in Europe but also in other parts of the world (Fulcher, 2004). There was in the CEFR (2001) a lack of empirical evidence between the products and the research to underpin the descriptions and reference levels of CEFR (2001) in its early stages. Consequently, as pointed out, examination providers, textbook publishers and curriculum developers made claims about the relationship between their products and CEFR (2001) but little hard evidence was produced to back up such claims (Alderson, 2007).

Creating a language competency framework for Thailand, Japan, Vietnam, Malaysia, and China has involved decisions which are more than simply transferring CEFR to other educational contexts. The various implementations of CEFR in this part of the world have been based on CEFR (2001) as the revised version of (2018) came later. In general, teachers and government officials’ saw the potential for the implementation of CEFR to help to raise the level of competence in English within the contexts of their educational system. However, the way CEFR was introduced has led many teachers to associate CEFR with the framework’s proficiency scale with too much emphasis on testing. This supports what Freeman (2017) called a ‘deficit view’ of for teachers and their teaching abilities. Wider forms of self-assessment advocated by the developers of CEFR seem to have been missed. For example, in Thailand, the feedback on the 2015 online placement test using either Cambridge / Oxford exam board was considered by the teachers as being more suitable for a European context. The teachers did not object to being tested as they wanted to improve their English proficiency as they felt it needed to be higher than their students. But for those English teachers below B1 in the test, there was apparently little additional support from the Ministry of Education in terms of offering special assistance (as it was to those attaining B 1 and above) to help improve their English proficiency (Franz & Teo, 2018).

A summary of the various issues identified in the implementation of CEFR in the region shows a number of similarities:

- ambitious target levels for students and teachers
- centralized decision making
- the need to resort to external consultancies
- teachers having very limited knowledge and exposure to CEFR
- teachers’ level of English proficiency
- the traditional resistance to change
- the lack of local CEFR experts who were able to construct and produce local CEFR textbooks’
- the lack of adequate teacher training and the notion that it would be difficult to incorporate CEFR in their teaching
- seeing CEFR as simple a measure of language proficiency rather than a goal in terms of a ‘can-do’ approach

Competency in a language is a multi-dimensional system that accounts for the situations, the functions, the linguistic elements needed in communicative competencies. However, measures of language competency can be arbitrary. North (2000) pointed out that CEFR as originally designed was a common measure for recording language competence and that the motivation for a common framework was more pragmatic (thus the ‘can do’) rather than academic. However, there were some inherent limitations in the original version of CEFR (2001) which did affect its applicability, not only in Europe but also in other parts of the world (Fulcher, 2004). There was, as we have seen, a lack of empirical evidence between the products and the research to underpin the descriptions and reference levels of CEFR (2001) in its early stages. Examination providers, textbook publishers and curriculum developers made claims about the relationship between their products and CEFR (2001) but little hard evidence was produced to back up such claims (Alderson, 2007).

Assessing Competence: The CEFR Framework

Creating a language competency framework for Thailand, Japan, Malaysia, Vietnam, and China has involved decisions which are more than simply transferring CEFR to other educational contexts. The various implementations of CEFR in the ASEAN and East Asia regions have been based on CEFR (2001). In general teachers and government officials saw the potential for the implementation of CEFR as helping to raise the level of competence in English within the contexts of their educational system. However, the way CEFR was introduced has led many teachers to associate CEFR with a proficiency scale with possibly and too much emphasis on testing. This supports what Freeman (2017) called a ‘deficit view’ of for teachers and their teaching abilities. Wider forms of self-assessment advocated in CEFR seem to have been missed. In Thailand, the feedback on the 2015 online CEFR based placement test using either Cambridge / Oxford exam boards was considered by the teachers as being more suitable for a European context. The teachers did not object to being tested as they wanted to improve their English proficiency as they felt it needed to be higher than their students. But for those English teachers below B1 in the test, there appeared to be little additional support in terms of offering special assistance to help improve their English proficiency (Franz and Teo, 2018). Indeed, as Byrnes pointed out the dangers of the simple and inappropriate transfer of CEFR content to other educational contexts called for CEFR research to focus more based on ‘how a context-free, though by no means context-indifferent, framework like CEFR can and should be translated into context-relevant forms in diverse educational environments in order to be implemented’ (Byrnes, 2007, pp. 642-643). CEFR’s original intention has been that language learning should be directed towards enabling learners to act in real-life situations, expressing themselves and accomplishing tasks of different natures. It would seem that the whole point of the development of local varieties of CEFR is ‘...to define ‘can do’ statements of users and learners of English relevant to the context of local and regional and international communication (Preface: FRELE-TH, iii 2017).

CEFR as ‘common currency’

As previously indicated, CEFR has become widely used and adopted by test developers. It has become to be seen as a common currency in language performance levels (Figueras, 2012). CEFR is so influential in Europe that it has become necessary for tests to be link to it to gain recognition. Outside of Europe too, many scoring systems and performance standards have been mapped onto CEFR (Tannenbaum and Wylie, 2008) for TOEFL, iBT and (Zheng & De Jong, 2011) for PTE Academic. Even though the goals of CEFR are more descriptive rather than normative (North, 2014), achieving score comparability across tests was one of the primary goals of its earliest drafts (van Ek, 1975). Today, the CEFR level

descriptors are used more in a normative way, as performance standards, or as labels to facilitate score transparency (Fulcher, 2012). However, two tests could have the same CEFR level but very different specifications, and it would be wrong to consider them as equivalent simply because they share a CEFR label (Taylor, 2004). Also, because CEFR is context and language independent, test developers need to add specific details to the descriptors when using them in a rating context (Harsch & Martin, 2012). CEFR descriptors have been criticized for their vagueness and inconsistencies, both within and across levels (Alderson, 2007; Harsch & Rupp, 2011) leaving room for dissimilar interpretations. CEFR (2001) was not designed as a ready-to-use normative tool, and its descriptors are unsuitable for unaltered use in rating scales (Deygerss, et al., 2018). However, changing the CEFR descriptors to meet the needs of a test is common (North, 2014) and it is not unlikely for two tests that were aligned to the same CEFR level to differ substantially in terms of content or construct yet, for test-score-users these tests might be considered equivalent because they share the same CEFR level. Fulcher (2010) suggested that the global spread of CEFR has been facilitated by the reification of the framework, the illustrative descriptors turning into prescriptive targets (B in the hands of policy makers and consultants. There is also the issue of whether there is sufficient evidence based on such CEFR tests to suggest that teacher's proficiency will translate into an increase in students' proficiency (Tsang, 2017). Based on his research in Hong Kong teachers of English found that having a native-like or high-proficiency does not equate to successful teaching and that factors other than proficiency may play a more important role (2017 p. 122).

FRELE-TH and other local versions based on CEFR are directed to treating all resources working together to make meaning, defining communicative activity and capacity along practice-based approaches. This means treating activity as the starting point for the analysis of competence; language is shaped by other resources in developing competence. Competency, especially language competency that prioritizes mastery of grammatical knowledge, can also be and has been, exclusionary in favoring different people, different countries and cultures. It is language that can create barriers as much as it presents possibilities of improving lives. Many English international (IELTS, TOEFL, TOEIC) and local proficiency tests can create barriers rather being ladders to future progress. Language as envisaged by CEFR, and other variations in the region is an attempt to look at competence more from its position as 'can do' performative communication given the contexts of use. This perspective takes diversity as the norm and challenges the assumption that 'sharedness' and uniformity are required to measure communicative success. An inflexible approach to CEFR levels suggesting that everyone who achieves a certain 'cut-off' score on a test is at the level indicated by CEFR appears to be more to do with the power of numbers/letters. This lies in the perception of the public that they are objective and therefore represent some truth with a strong implication that they are not open for discussion and challenge. There are certain numbers or letters in the case of CEFR that have become iconic in the consciousness of score users to the extent that they cannot be altered, even if the actual test undergoes radical revisions. This is the case of the perceived meaning of 6.5/7 on the IELTS reporting scale which has not fundamentally changed since its very different forerunner the English Language Testing System (ELTS). Score users, administrators and the public have come to accept the truth that this is the point on a nine-band scale at which a student is capable of undertaking English medium studies in higher education. The value of any number depends upon the extent to which it reflects the performance that the test intends to assess in this case language proficiency in terms of ability to communicate. But proficiency and competence may not be exactly the same.

For example, Speexx Language Assessment is based on CEFR (2001) and claims to be the most widely used and recognized language proficiency scale in the world. Speexx

projects a CEFR- aligned score with results that are comparable to other major assessments: TOEIC, IELTS, TOEFL (iBT).

Table 2: Comparative table of different assessments

CEFR BAND	Speexx Placement test	TOEIC	IELTS	TOEFL (iBT)
C2	95-100	801-900	7.0-7.5	95-99
C1	81-95	701-800	6.0-6.5	78-94
		601-700	4.5-5.5	52-77
B2	65-80	501-600	4.0	40-51
B1	49-64	401-500	3.0-3.5	29-39
		251-400	2.5-2.0	18-28
A2	33-48	201-250	1.0-1.5	9-17
		101-200	0-1.0	
A1	17-32	10-100	0.1.0	0-8

However as was originally pointed out *In the Introduction to ELTS* (1987) ‘The appropriate level of proficiency required for a given course of study or training is ultimately something which institutions/faculties/course tutors must determine in the light of knowledge of their own courses and their experience of overseas students taking them’ (1987 p. 7)

Testing and CEFR

In testing the most commonly used standard documents are normally those of the Ministries of Education. The aim is to provide a framework for the development of assessments for particular purposes or the harmonization of educational systems. It represents an attempt to ensure that the outcomes of all tests can be quantified in the same ‘currency’, which is given its value by the institutions that organizes the tests. The effect has been the creation of an alignment industry which has come to view mapping test scores indicating competence to external standards as a validation process in its own right (Fulcher, 2012). However, the measure of such competence should be empirically derived, that is, what counts as real-life communication and based on data of what people say and do (performance) as proposed in CEFR. Indeed, it has been often pointed out that language assessment is a social and political activity (Roever and McNamara, 2006; Shohamy, 2006). As the values of a globalized economy preoccupy many aspects of social and political life, so have they become central to language testing and assessment. Standardized tests make us less able to respond to the fact that communication, ever increasingly is taking place in the globalized workplace using English as a lingua franca. In current criterion-referenced approaches, scores are interpreted in relation to pre-determined standards of knowledge and capacity in functional terms, reflecting the communicative tradition of language teaching. This was the case of CEFR (2001) and with an ordered set of statements about aspects of communicative ability. In CEFR (2018) the design of the assessment and the construction of the test instrument begins with a set of standards that are general because tests have always been sensitive to policy and market constraints, as indicated in the history of IELTS (Davis, 2008). The wording of the original frame work of CEFR (2001) reflects the policies of European governments and educational institutions. The Council of Europe in the early 1970s was the principle influence on the European version of communicative language teaching as part of European integration. FRELE-TH and other versions in the region can be seen as part of ASEAN Economic Community (AEC) with Education as central to economic development. As such the curriculum framework is meant to determine the shape of assessment to establish general

educational standards and for personal educational achievement. The potential overwhelming political authority of the CEFR means that increasingly, language assessments in many areas of education must be calibrated against it. This is a complex and technically demanding process as published tests need to be explicitly related to CEFR levels. What is C1 on IELTS a 6.5 or a 7? (McNamara, 2010) This has become a crucial level for admission decisions into higher education and for continued employment. For example, in Vietnam educational policy for ELT now requires a C2 for University teachers a C1 for teachers in Higher Education and a B2 for teaching in secondary education and a B1 for primary. However, one could argue that if the standard of CEFR [2018] for university reference is C1, any test organization that wishes its own test to be used for this purpose will produce a cut-off score that links the test to the relevant level. The danger is that once a test has been scaled and standards are set to external standards, it has been criterion referenced. The interpretation of test score meaning is framed only in terms of the link, in this case to the CEFR. A further problem is that the test construct is embedded in policy processes. It is only through policy processes that the construct can be challenged or updated because existing accountability frameworks such as CEFR are deeply embedded in existing administrative procedures. Consequently, there will be a natural reluctance to change the construct because of all the administrative and legislative work that is entailed (McNamara, 2011).

Conclusion

There are several issues stemming from CEFR and its implementation. For example, language assessment might be seen as simply serving the goals and policies supporting a view of education as primarily preparing learners for the globalized work force. Indeed, it is true that in language education, we need to help create a capacity for participation in a world outside the classroom and for individuals to attain their own personal goals. Globalization demands flexibility which is precisely in danger of being forgotten in universalizing outcome statements such as in the original CEFR (2001). However, in CEFR (2018) the changes made in CEFR descriptors indicate a move towards a degree of flexibility in terms of the levels to be attained. In FRELE-TH and other variations in the region, it is suggested, can be used for the design of local high stakes standardized tests of English proficiency, and the consequent results be benchmarked with those of international standards (Hiranburana, et al., 2018) In testing and assessment the main concern is the validity of the construction. If the test specifications are based on the descriptors/can-do statements and test the salient features of the levels of user/learner competence, the degree of success will depend on how the users/learners can perform the task according to the criteria set earlier (criteria-based testing and evaluation). The issue is, can we have an ecologically sensitive assessment where interpretation of the levels attained are context dependent for learner/users to be aware of their performance benchmarked with international standards. It is recognized that testing is part of social reform in the sense of giving, in principle, unbiased access to educational and employment opportunities (Zeng, 1999). Going back to the validity argument, how do we reconcile external international assessment with local contexts as in the use of English as a Lingua Franca rather than a UK or US based model. Perhaps, the major issue is the potential effect on the creation of an alignment industry which has come to view this mapping of test scores to external standards as a validation process in itself (Martyniuk, 2009). There are always two sides to an argument; the use of a common currency based on forms of CEFR makes sense if qualifications and certificates are to be accepted across national boundaries (Jones, 2013). For others, this common currency can be a feature of political control that has more the role of ‘gatekeeper’.

About the Author

Joseph A. Foley is a Senior Lecturer, in the Graduate School of Human Sciences (GSHS), Assumption University. He obtained his Ph.D. in Linguistics and Education from University College London, Institute of Education. He worked in Singapore for more than 25 years at the National University of Singapore and SEAMEO (RELC, Singapore). His main activity in Assumption University involves supervising PhD. ELT candidates and teaching MA. and Ph.D. courses in English Language Teaching. He has published very widely and his books and research articles are in areas in Applied Linguistics with a specialization in Systemic Functional Linguistics.

References

- Alderson, C. (2007). The CEFR and the need for more research. *The Modern Language Journal*, 91(4) 695-663. Retrieved May 14, 2019 from: https://onlinelibrary.wiley.com/doi/abs/10.1111/j.1540-4781.2007.00627_4.x
- Azman, H. (2016). Implementations and challenges of English language education reform in Malaysian primary schools. *3L: The Southeast Asian Journal of English Language Studies*. 22(3), 65-78.
- Barni, M. (2015). In the name of the CEFR: Individuals and standards. In B. Spolsky, O. Inbar-Lourie, and M. Tannenbaum (Eds), *Challenges of language education and policy. Making space for people* (pp.40-52). London, Routledge.
- Barro, R. & J. Lee (2013). A new data set of educational attainment in the world, 1950–2010. *Journal of Development Economics*, 104, 184–98.
- Byram, M. & L. Parmenter (eds). (2012). *The Common European Framework of Reference: The globalisation of language education policy*. Multilingual Matters.
- Byrnes, H. (2007). Perspectives. *The Modern Language Journal*, 91(4), 641-645.
- Bucar, C.S., Ryu, H., Skof, N.M. & Sangawa, K.H. (2014). The CEFR and Teaching Japanese as a foreign language. *Linguistica*, 54(1), 455-469. Retrieved May 25, 2019 from: <https://revije.ff.uni-lj.si/linguistica/article/view/2619> doi: 10.4312/linguistica.
- China's standards of English language ability* (2018). Ministry of Education of the People's Republic of China.
- Chung, V. (2014). Vietnam considers using CEFR-V standard for English teaching. Viet/Nam/Net, Hanoi, Vietnam 15th November. Retrieved May 25, 2019 from: <http://english.vietnamnet.vn/fms/education/116476/vietnam-considers-using-cefr-v-standard-for-english-teaching.html>.
- Council of Europe (2001). *Common European Framework of Reference for Languages: Learning, Teaching and Assessment*. Council of Europe, Language Policy Unit: Strasbourg. Retrieved April 17, 2019 from: www.coe.int/lang-cefr.
- Council of Europe (2018). *Common European Framework of Reference for Languages: Learning, teaching and assessment. Companion volume with new descriptors*. Council of Europe, Language Policy Programme. Strasbourg. Retrieved APRIL 17, 2019 FROM: www.coe.int/lang-cefr.
- Costa, D. (2007). Contextualizing uses of the common European framework of reference for languages, In *Council of Europe, The common European framework of references for languages (CEFR) and the development of language policies: Challenge and responsibilities*. Strasbourg: Council of Europe. (40-49). Retrieved May 10, 2019 from: www.coe.int/t/dg4/linguistics/Publications_en.asp.
- Davis, A. (2008) *Assessing academic English: Testing English proficiency 1950-89 – The ETS solution*. Cambridge: Cambridge University Press.

- Deygers, B. (2019). *The CEFR companion volume: Between research-based policy and policy-based research*. Retrieved June 26, 2019 from: <https://academic.oup.com/applij/advance-article-abstract/doi/10.1093/applin/amz024/5487749>.
- Deygers, B., C. Carlsen, N. Saville, & K. Van Gorp (2018). The use of the CEFR in higher education: A brief introduction to this special issue. *Language Assessment Quarterly* 15(1), 1–2.
- Dhanasobhon, S. (2006). *English language teaching dilemma in Thailand*. Retrieved May 30, 2019 from: <http://www.curriculumandinstruction.org/index.php?lay=show&ac=article&Id=539134523&Ntype=7>.
- Education First Standard English (2017). *Test proficiency index*. Retrieved May 18, 2019 from: www.ef.cf.com/epi.
- English Language Testing Service (1987). *An introduction to the English language testing service*. The British Council. University of Cambridge Local Examinations Syndicate.
- Figueras, N. (2012). The impact of the CEFR, *ELT Journal*, 66(4), 477–85.
- Fulcher, G. (2004). Deluded by artifices? The Common European Framework and harmonization. *Language Assessment Quarterly*, 1(4) 23-266. Retrieved May 2, 2019 from: <http://citeseerx.ist.psu.edu/viewdoc/download?doi=10.1.1.460.2317&rep=rep1&typepdf>.
- Fulcher, G. (2010). The reification of the Common European Framework of Reference (CEFR) and effect-driven testing. In A. Psytaltou-Joycey and M. Mattaioudakis (Eds), *Advances in research in language acquisition and teaching*. GALA: Thessaloniki.
- Fulcher, G. (2012). Scoring Performance Tests. In Fulcher, G and Davidson, (Eds), *The Routledge Handbook of Language Testing* (pp.378-392). Abington: Routledge.
- Franz, J., & Teo, A. (2018). A2 is normal? - Thai secondary school English teachers' Encounters with CEFR. *RELC Journal*, 49(3), 322-338.
- Freeman, D. (2017). The case for teachers' classroom proficiency. *RELC Journal*, 48(1), 31-52.
- Garcia, O., and Li Wei (2014). *Translanguaging: Language, bilingualism and education*. New York: Palgrave Macmillan.
- Green, A. (2012). Language functions revisited: Theoretical and empirical bases for language construct definition across the ability range. *English Profile Studies* vol. 2. Cambridge: Cambridge University Press.
- Goullier, F. (2007). Impact of the Common framework of Reference for Languages and the Council of Europe's work on the New European educational area, in Council of Europe (2007) *The Common European Framework of References (CEFR) and the Development of Language Policies: Challenges and responsibilities*, Strasbourg: Council of Europe. 29-37.
- Harsch, C. & Martin, G. (2012). Adapting CEF-descriptors for rating purposes: Validation by a combined rater training and scale revision approach. *Assessing Writing*, 17(4), 228-250. Retrieved April 15, 2019 from: <https://www.researchgate.net/publication/257483759> doi:10.1016/j.asw.2012.06.003.
- Harsch, C., & Rupp, A. (2011). Designing and scaling level specific writing tasks in alignment with CEFR: A test-centered approach. *Language Assessment Quarterly*, 8(1), 1-33. doi:10.1080/15434303.2010.535575.
- Hiranbuana, H. Subphadoongchone, P. Tangkiengsirisin, S. Phoochaoensil, S. Gainey, J. Thongsongri, J. Sumonsriworakun, P. Somphong, M. Sappapan. P, & Taylor, P.

- (2017). *Framework of reference for English Language Education in Thailand (FRELE-TH)* based on the CEFR.
- Hiranbuana, H. Subphadoongchone, P. Tangkiengsirisin, S. Phoochaoensil, S. Gainey, J. Thongsongri, J. Sumonsriworakun, P. Somphong, M. Sappapan. P, & Taylor, P. (2017). A Framework of reference for English Language education in Thailand (FRELE-TH) – based on the CEFR, The Thai experience. *LEARN Journal*, 10(2), 90-119.
- Hiranbuana, H. Subphadoongchone, P. Tangkiengsirisin, S. Phoochaoensil, S. Gainey, J. Thongsongri, J. Sumonsriworakun, P. Somphong, M. Sappapan. P, & Taylor, P. (2018). Framework of Reference for English Language Education in Thailand – (FRELE-TH). Based on CEFR: Revisited in the English Educational Reform. *Pasaa Paritat Journal*, 33, 51-91.
- Houghton, S. A., D. J. Rivers, & K. Hashimoto (2018). *Beyond Native-speakerism. Current explorations and future visions*. Abington: Routledge.
- Hulstijn, J. (2007). The shaky ground beneath the CEFR quantitative and qualitative dimensions of language proficiency. *The Modern Language Journal*, 91(4), 663-667.
- Hulstijn, J. (2015). *Language proficiency in native and non-native speakers: Theory and research*. Amsterdam: John Benjamins.
- Hung, N.N. (2013). Vietnam's national foreign language 2020 project: Challenges, opportunities, and solutions. Retrieved May 15, 2019 from: <http://bruneiusprogramme.org/wp-content/uploads/2013-Forum-Publications-Complete.63-65.pdf>.
- Ishak, W. & Mohamad, M. (2018). The implementation of Common European Framework of Reference (CEFR): What are the effects towards LINUS students' achievements? *Creative Education*, 9, 2714-2731. doi: 10.4236/ce.2018.916205.
- Jones, N. (2012). Reliability and dependability. In Fulcher, G. and Davidson, F (Eds), *Routledge Handbook of Language Testing* (pp. 350-360). London: Routledge.
- Krumm, H. J. (2007). Profiles instead of levels: The CEFR and its (ab)uses in the context of migration, *The Modern Language Journal* 91/4: 667–9.
- Komorowska, H. (2014). The CEF in pre- and in-service teacher education. In K, Morrow (Ed), *Insights from the common European framework* (pp. 55-64). Oxford: Oxford University Press.
- Laufer, B. (2003). Vocabulary acquisition in a second language: Do learners really acquire most vocabulary by reading? Some empirical evidence, *Canadian Modern Language Review* 59/4: 567–87.
- Lo, Y.Y. (2018). English teachers' concern on Common European Framework of References for Languages (CEFR). An application of BAM. *Jurnal Hurikulu, dan Pengajaran Asia Pasifik* 6 (1). 46-58.
- Mala, D. (2016). English in the news: govt launches app, teacher's 'boot camp' Bangkok Post 9 March. <http://www.bangkokpost.com/learning/advanced/891472/english-in-the-news-govt-launches-app-teachers-boot-camp>.
- Martyniuk, M. (2009). *Alignment Tests with CEFR*. Cambridge: Cambridge University Press.
- Martyniuk, W. & Noijons, J. (2007). *Executive summary of results of a survey on the use of the CEFR at national level in the Council of Europe Member States*. Council of Europe Language Policy Division.
- McNamara, T. (2011). Managing Learning: Authority and Language Assessment. *Language Teaching* 44(4), 500-515.
- McNamara, T. (2014). 30 years on—Evolution or revolution? *Language Assessment Quarterly*, 11(2), 226–32.

- MEXT (2011). Five proposals and specific measures for developing proficiency in English for international communication. Retrieved May 18, 2019 from: www.mext.go.jp/english/elsec/1319701.htm.
- Ministry of Education Malaysia (2012). Malaysia Education Blueprint 2013-2025 Putrajaya: Ministry of Education Malaysia. Retrieved April 15, 2019 from: http://www.moe.gov.my/userfiles/fil/PPP/Preliminary-Blue_print-Eng.pdf.
- Moser, J. (2015). From a knowledge-based language curriculum to a competency-based one: The CEFR in action in Asia. *Asian EFL Journal*, 88, 1-29.
- Negishi, M. Takada & Tono, Y. (2013). A progress report on the development of the CEFR-J In Galaczi, E.D. & Weir. C.J. (Eds.), *Exploring Language Frameworks: Proceedings of the ALTE Kraków Conference July 2011*, *Studies in Language Testing*, 36, 135-163. Cambridge: UCLES/Cambridge University Press.
- Negishi, M. & Tono, Y. (2014). *An update on CEFR-J project and its impact on English language education in Japan*. Paris ALTE.
- Nguyen, V.H. & Hamid. M.O. (2015). Educational Policy borrowing in a Globalized World: A Case Study of Common European Framework of reference for Languages in a Vietnamese University. *Journal of English Teaching, Practice and Critique*. 14(1), 60-74. doi: 10.1108/ETPC-02-2015-0014.
- North, B. (2000). *The development of a common framework scale of language proficiency*. New York, Peter Lang.
- North, B. (2004). Europe's framework promotes language discussion, not directives. Retrieved April 15, 2019 from: <http://education.guardian.co.uk/tefl/story/0,1191130,00.html>.
- North, B. (2007). The CEFR illustrative scales. *The Modern Language Journal* 94 (4) 656-59 doi:10.1111/modl2007.91.issue-4.
- North, B. (2008). EAQUALS bank of descriptors – as levels. EAQUALS/ALTE portfolio descriptor revision project. *The European Association of Quality Language Services*. Retrieved May 22, 2019 from: www.eaquals.org.
- North, B., Oretaga, A, & Sheehan, S. (2010). *A core inventory for general English*. British Council: EAQUALS.
- North, B. (2014). Putting the Common European Framework of Reference to good use. *Language Teaching*, 47(2), 228–49.
- North, B. & E. Piccardo. (2016). Developing illustrative descriptors of aspects of mediation for the Common European Framework of Reference (CEFR): *A Council of Europe project*. *Language Teaching*, 49(3), 455–9.
- Nurul, F. M. U. & Sallehuddin, A. A. M. (2017). CEFR in Malaysia: Current issues and challenges in the implementation of the framework. The 3rd International Conference on Language Testing and Assessment and 5th British Council New Directions in Language Assessment Conference 2-3 December 2017. Shanghai, China.
- Papageorgiou, S. (2010). Investigating the decision-making process of standard setting participants. *Language Testing*, 27(2), 261-282.
- Pitsuwan, S. (2014). Teachers of English to speakers of other languages. *TESOL International Convention 2014*. Portland, Oregon, USA.
- Poszytek, P. (2012). Policy perspectives from Poland. In M. Byram & L. Parmenter (Eds.), *The common European framework of reference: The globalization of language education policy* (pp. 97-103). Bristol: Multilingual Matters.
- RELC Journal (April 2018). *Special Issue: Teaching English as an International language (TEIL): Realistic or Idealistic*, 49(1), SAGE Publications Ltd.
- Roever, C. & McNamara, T. (2006). Language testing: The social dimension. *International Journal of Applied Linguistics*, 16(2), 242-258. doi:10.1111/ijal.2006.16.issue-2.

- Salamoura, A. & Saville, N. (2010). Exemplifying the CEFR: criterial features of written learner English from the English Profile Programme. In I. Bartning and M. Marten (Eds.), *Communicative Proficiency and Linguistic Development: Intersection between SLA and Language Testing Research: Eurosla Monograph Series 1* European Second Language Association (pp 102-132).
- Saville, N. (2016) *Foreword* Framework of Reference for English Language Education in Thailand (FRELE-TH) based on CEFR 2017.
- Seidlhofer, B. (2007). Common Property: English as a lingua franca in Europe. In Cummins J., Davison C (Eds.), *International Handbook of Language Teaching*. London: Springer, 137-50.
- Shohamy, E. (2006). *Language policy: Hidden agendas and new approaches*. Abington: Routledge.
- Speexx Language Assessment* (2018) Retrieved April 28, 2019 from: <https://www.speexx.co.th/en>.
- Menon, S. (2019). English proficiency needs time. *The Star newspaper* (4 May 2019).
- Tannenbaum, R.J., and Wylie, C.R. (2008). *Linking English language test scores onto the Common European Framework of Reference: An application of standards setting methodology*. Princeton, NJ: ETS.
- Taylor, L. (2004). Issues of test comparability. *Cambridge Research Notes*, 15, 2-5.
- Trim, J. (2012). The Common European Framework of Reference for Languages and its background: A case study of cultural politics and educational influences. In M. Byram and L. Parmenter (Eds), *The Common European Framework of Reference: The Globalisation of Language Education Policy. Multilingual Matters* (pp. 14–36).
- Trim, J.L. & Trim, J.L.M. (1980). *Systems development in Adult language learning: A European unit/credit system for modern language learning by adults/ prepared for the Council of Europe*. Council of Europe: Strasbourg.
- Tono, Y. (2012). What is CEFR-J? CEFR-based framework for ELT in Japan, research project. Retrieved April 12, 2019 from: <http://www.tufs.ac.jp/ts/personal/tonolab/cefr-j/english/whatis.html>.
- Tono, Y. & Negishi, M. (2012). The CEFR-J: Adapting the CEFR for English language teaching in Japan. *Framework & Language Portfolio (FLP). SIG Newsletter*, 8 September 2012.
- Tsang K.S. (2017). *Conceptions of learning among secondary four students in Hong Kong. university of Hong Kong*. Pokfulam: Hong Kong.
- The Bangkok Post* August 9, 2018.
- UNDP (2015). *The Millennium Development Report*. Retrieved April 18, 2019 from: www.th.undp.org/content/thailand/en/home/library/mdg/the_millennium_development_report_2015.html.
- van Ek, J. (1975). *Systems Development in Adult Language Learning: The Threshold Level in a European-Unit/Credit System for Modern Language Learning by Adults*. Council of Europe.
- Van Ek, J.A. & Trim J.L.M. (1990). *Threshold 1990*. Cambridge: Cambridge University Press.
- Viet, H. P.. (2015). The CEFR in Vietnam from adoption to adaption. JALT Conference.
- Vantage* (2001). <http://www.cambridgeenglish.org/exams-and-tests/cefr>.
- Waystage*. (1990). Retrieved April 18, 2019 from: http://www.coe.int/t/dg4/linguistics/Waystage_CUP.pdf.
- West, R. (2015). Keeping it in the family. *English Language Teaching Professional*, 97.

- Yan Jin, Zunmin Wu, C. Alderson & Weiwei Song (2017). Common European Framework of reference for Languages (CEFR) for English language assessment in China *Language Testing in China* (Special Issue) Springer, Open.
- Zeng, K. (1999). *Dragon gate: Competitive examinations and their consequences*. London: Cassell.
- Zheng, Y. & Jong, J. (2011). Research note: Establishing construct and concurrent validity of Pearson test of English academic. Pearson Education Ltd. Retrieved April 15, 2019 from: http://pearsonpte.com/wp-content/uploads/2014/00000007/RN_EstablishingConstructAndConcurrentValidityOfPTEAcademic_2011.pdf.