Lang. Teach. (2011), **44.4**, 500–515 © Cambridge University Press 2011
doi:10.1017/S0261444811000073 First published online 19 April 2011

Plenary Speeches

Managing learning: Authority and language assessment

Tim McNamara School of Languages and Linguistics, The University of Melbourne, Australia
tfmcna@unimelb.edu.au

A feature of language assessment internationally is its role in the enforcement of language policies established by governments and other educational and cultural agencies. This trend has led to the near-universal adoption of curriculum and assessment frameworks, the clearest example of which is the Common European Framework of Reference for Languages (CEFR). What does this movement represent? What is the source of the authority of frameworks such as the CEFR, and why are they so appealing to governments? The paper argues that the determination of test constructs within policy-related frameworks leads to inflexibility, and considers the case of the testing of competence in English as a lingua franca (ELF) communication, in the context of international civil aviation communication.

1. Introduction

Language assessment, as we are increasingly realizing, is first and foremost a social and political activity (Spolsky 1995; Shohamy 2001, 2006; McNamara & Roever 2006). It is a site where social values are expressed and contested. As the values of a globalized economy have come to occupy and preoccupy many spheres of social and political life, so have they become central to current developments and challenges in language testing and assessment. I will try to argue that the values of globalization are embedded in test constructs through the development of powerful outcome statements and frameworks. I will then argue that this leads to an inflexible situation, where language education, controlled by outcomes assessment in the form of universal standards, is erasing the historical and cultural complexity and specificity of language learning in particular settings, and the meaning of language learning in the lives of individuals. I will also argue that, ironically, standards-based language learning, as currently formulated, makes us less able to respond to another result of globalization, the fact that communication in the globalized workplace takes place using English as a lingua franca. I will illustrate this last point with reference to the role of language testing in the management of the international aviation workplace, whose business depends on successful English as a lingua franca for communication. I will argue that current constructs as articulated in

Revised version of a paper given as a plenary address at the Third International CELC Symposium, 'Global Perspectives, Local Initiatives: Reflections and Practices in English Language Teaching', National University of Singapore, May 2010.

https://doi.org/10.1017/S0261444811000073 Published online by Cambridge University Press

powerful outcomes statements are inadequate to deal with the needs of this increasingly common situation.

2. Test constructs and policy

Test scores tend today to be interpreted in terms of real world abilities and skills. In earlier norm-referenced approaches to assessment, an estimate of a candidate's ability, in the form of a test score, was interpreted with reference to the performances of other test takers, that is, with reference to the scores of the population of test takers for that test. Thus, test score meaning was always relative. In current criterion-referenced approaches, scores are interpreted in relation to pre-determined standards of knowledge and capacity. Such standards are typically couched in practical, functionalist terms, reflecting the functionalist, communicative tradition of language teaching. For example, the widely used and very influential Common European Framework of Reference for Languages (CEFR: Council of Europe 2001) consists of an ordered set of statements about aspects of communicative ability, grouped into six levels (from lowest to highest: A1, A2, B1, B2, C1, C2) according to the complexity of the skills concerned. It is claimed, for example, that a speaker assessed as meeting the standard for level A2

> Can understand sentences and frequently used expressions related to areas of most immediate relevance (e.g. very basic personal and family information, shopping, local geography, employment). Can communicate in simple and routine tasks requiring a simple and direct exchange of information on familiar and routine matters. Can describe in simple terms aspects of his/her background, immediate environment and matters in areas of immediate need.

while a speaker at B2

> Can understand the main ideas of complex text on both concrete and abstract topics, including technical discussions in his/her field of specialisation. Can interact with a degree of fluency and spontaneity that makes regular interaction with native speakers quite possible without strain for either party. Can produce clear, detailed text on a wide range of subjects and explain a viewpoint on a topical issue giving the advantages and disadvantages of various options. (Council of Europe 2001, p. 24)

The wording of these standards is important, as they represent the test CONSTRUCT, the assumed view of language proficiency which is assessed in the test and which is the target of measurement in any individual case. (The reference to the native speaker in the definition of level B2 is crucial to the argument of this paper and will be considered below; cf. Seidlhofer 2003.)

Increasingly, the design of an assessment and the construction of a test instrument begin with such a set of standards. These may be rather general, as in the case of the CEFR, or more specific, as in standards relevant to a particular workplace or educational setting; we will consider English language standards for the aviation workplace later in the paper.

Policy initiatives expressed in the wording of standards are increasingly central to language testing. The fact the standards set by policy act as the basis for the test has changed the nature of language testing. Previously, test developers, informed by the latest theories of language proficiency, of communicative interaction, and of the way in which language

https://doi.org/10.1017/S0261444811000073 Published online by Cambridge University Press

knowledge intersects with and interacts with other aspects of professional oreducational competence, defined the test construct themselves in a stage preparatory to the development of a procedure which would provide evidence relevant to that construct. This 'thinking stage' of test development has now been removed. Instead, given that the test construct, and its wording, are determined by complex policy procedures, evolving discussions within applied linguistics on the nature of language proficiency and communicative interaction are far less likely to inform test constructs once the standards as initially formulated have been adopted – they are cast in stone. While tests have always been sensitive to policy and market considerations (see Spolsky 1995 on the history of TOEFL and Davies 2008 on the history of IELTS), these pressures are now directly and explicitly expressed within the very wording of the test construct, from which everything in a test flows.

The CEFR is a case in point. The wording of the framework inevitably reflects the policies and values of its original sponsor, the Council of Europe, and its current users, governments and educational institutions throughout the world. The CEFR has its origins in the work of the Council of Europe in the early 1970s. The Council was looking for a way of thinking about achievement in language study which was independent of the structural particularities of particular languages, in the interests of labour mobility and readily transportable and interpretable credentials. The policies of the Council at that time were critical in the development of the notional/functional syllabus, and the definition of levels of syllabus in such functionalist terms; the Council was the principal influence on the European version of communicative language teaching. It can therefore be argued that the functionalist orientation of communicative language teaching and the curriculum and assessment frameworks that go with it are a reflection of the values of the Council – broadly, European integration. In other words, the terms in which the outcomes of language education are specified, and which are the construct of language assessment, are expressions of policy. And the formulations of such frameworks, including the way in which they lend themselves to a reduction to simple numbers – A1, A2, etc. – are designed to meet the needs of those responsible for the accountability of educational systems.

The credit transfer ('unit credit scheme': Trim 1980) and labour mobility motivation of the Council of Europe is explicable in terms of its policy goals, as described on its website in the period in which the CEFR was adopted (they have since been revised to focus more on the promotion of human rights and of democratic governance systems):

> The Council of Europe is the continent's oldest political organisation, founded in 1949… [It] was set up to … develop continent-wide agreements to standardise member countries' social and legal practices [and] promote awareness of a European identity based on shared values and cutting across different cultures.
>
> Within the area of education, its mission involves carrying out major projects on education policy, content and methods in and out of school… Special importance is attached to … the mutual recognition of competences to improve mobility and employment prospects, and lifelong learning for participation in an international society. (www.coe.int/T/e/Com/about_coe/, accessed 10 January 2006)

The cultural context of the goal of an integrated European economy, a regional form of globalization, is clear, and is strongly reminiscent of the mission of the Organization of Economic Cooperation and Development (OECD), an organization more explicitly committed to furthering the processes of globalization. It may be worth comparing the

https://doi.org/10.1017/S0261444811000073 Published online by Cambridge University Press

actual wording of the goals of the two organizations (the goals that seem shared are shown in bold):

OECD	**COUNCIL OF EUROPE**
The Organization's mission is essentially to help governments and society reap the full benefits of **globalization**,	. . . the mutual recognition of competences to improve **mobility and employment prospects**, and lifelong learning for **participation in an international society**
while tackling the economic, **social and governance challenges** that can accompany it.	[It] was set up to . . . develop continent-wide agreements to standardise member countries' **social and legal practices**. . .

The OECD specifically focuses on economic development, and the Council of Europe on education, but in a deeper sense the goals of the two organizations in the period leading up to the formulation and adoption of the CEFR reflect the same mission. Education is seen as centrally concerned with economic development, and the curriculum frameworks determining the shape of the accompanying assessment regimes for educational achievement, including that for languages, are central to the fulfilment of the goals of each organization. The OECD also has a testing regime, the Program for International Student Assessment or PISA, which is having profound impacts on the educational systems of member countries (McNamara 2011) – it tests the literacy of 15 year olds in the language or languages used as the medium of education in the member states, as well as their knowledge of mathematics and science.

This linking of language policy to educational goals through processes of examination and certification is, of course, not new. In a discussion of the linguistic unification of France in the eighteenth century, the French sociologist Pierre Bourdieu writes:

> But it was doubtless the dialectical relation between the school system and the labour market – or, more precisely, between the unification of the educational (and linguistic) market, linked to the introduction of educational qualifications nation-wide, independent (at least officially) of the social or regional characteristics of their bearers, and the unification of the labour market (including the development of the state administration and the civil service) – which played the most decisive role in devaluing dialects and establishing the new hierarchy of linguistic practices. (Bourdieu 1991: 49)

3. Elaborating the construct

Increasingly, then, test constructs, which are the heart of tests, are dictated to us by policies governing teaching, and hence assessment practices, in educational systems. The wording of the statements of test constructs in standards documents tends to be brief and directed at the audience of those using and interpreting test scores, rather than those responsible for constructing and conducting tests: in other words, the wording is user-oriented rather than constructor- or assessor-oriented (Alderson 1991). Further elaboration is often provided to help teachers, syllabus developers and test developers interpret the meaning of the standards

https://doi.org/10.1017/S0261444811000073 Published online by Cambridge University Press

and to guide them in the development of syllabuses and the operational tests that accompany them.

The CEFR has considerable elaboration of the descriptors set out in the global scale, as described above, in the form of a manual designed to assist in the interpretation of the levels, and the application of the scale in different contexts is set out in a book-length publication (Council of Europe 2001). More broadly, a great deal of current work in language testing is designed to provide interpretative commentary on the meaning of standards such as the CEFR, and more and more detailed advice to test developers and teachers about relating the necessarily abstract and general statements of standards to the day-to-day realities of test design and classroom assessment. Language testers thus play a mediating role between policy statements and the needs of teachers and learners, though it should be recognized that in many respects the needs of administrators and those of teachers and learners are at odds, placing an impossible burden on language assessment (not to mention on the teachers who have to do the assessing) (McNamara 2001).

Now the impact of the CEFR will in some respects be positive. The spread of policies favouring communicative methods of teaching for a wide range of languages throughout Europe can be attributed to the influence of the CEFR. A shared framework across languages means that shared professional development of teachers of different languages is possible, at least so far as familiarizing them with the approach of the CEFR is concerned. The articulation of the goals of language education and the course of development of the learner's competence leads to valuable discussion among teachers. There is already some documentation of these benefits (Morrow 2004).

The policy goals of the CEFR are now dominating language education policy at every level in Europe, in a striking example of policy-driven assessment. It is worth noting in passing that the appeal of this management function of the CEFR to other governmental systems, and of course the power and influence of Europe, mean that it is being used as a crucial reference point – in fact the only reference point authorized by policy – well beyond Europe: for example in North and South America, Australia and Asia. And the impact of the CEFR extends to areas beyond language education. The levels of the CEFR are enshrined in law, in policies incorporating the use of language tests in granting immigrants rights to entry, residency and ultimately citizenship (Van Avermaet 2009).

4. Imposing a common currency: Calibrating tests against the CEFR

The overwhelming political authority of the CEFR means that increasingly, language assessments in many, indeed most, areas of education must be calibrated against it. This is a complex and technically demanding process. Frameworks like the CEFR are like the Euro, or like the European Committee for Standardization (CEN). They make exchange across boundaries easier, but reduce local variation, and render unintelligible other accounting systems, or sets of cultural values, or formulations of the goals of language education, which cannot be directly translated into the language of the CEFR. Moreover, if the construct of a curriculum or an assessment cannot be calibrated against the CEFR, it has no currency.

https://doi.org/10.1017/S0261444811000073 Published online by Cambridge University Press

It is becoming imperativefor published tests to be explicitly related to the CEFR. There is a clear commercial advantage for those which can do so. Cambridge ESOL, for example, has linked the levels of its main suite of examinations to the levels of the CEFR: in its own words:

> Cambridge ESOL's exam system has developed in interaction with the Council of Europe language policy initiatives which have culminated in the publication of the CEFR – a process described by North 2006 and Taylor & Jones 2006. This history of interaction gives Cambridge ESOL a very strong, one might say organic, relationship to the CEFR. (www.cambridgeesol.org/what-we-do/research/cefr/common-scale-levels.html)

This leads to mapping exercises such as that shown on Cambridge ESOL's website at www.cambridgeesol.org/assets/img/exams/cefr-diagram-large.jpg; this is also included as supplementary material on this journal's website (http://journals.cambridge.org/lta).

But linking is no simple matter. The question of the calibration of individual tests to the CEFR is very demanding and very complex, even for an organization like Cambridge. Recent changes to immigration law in Britain tie admission under certain visa categories to the applicant's demonstration of a certain level of proficiency on the CEFR (www.ukba.homeoffice.gov.uk/policyandlaw/immigrationlaw/immigrationrules/appendixb/). This means that the calibration of levels on a widely available test such as IELTS becomes hugely consequential. Cambridge itself is unclear about the correct calibration. On the one hand, it states

> Our recent work to align the Cambridge tests within the CEFR suggests that a CAE pass and IELTS Bands 6.5/7.0 are located at Level C1 of the Framework.

but on the other,

> In 1998 and 1999 a research project examined the relationship between IELTS and the Cambridge Main Suite Examinations, specifically CAE (C1 level, Common European Framework) and FCE (B2 level). Under examination conditions, candidates took both IELTS and tests containing CAE or FCE items. Although the study was limited in scope, the results indicated that a candidate who achieves an overall Band 6 or 6.5 in IELTS would be likely to pass CAE.
>
> Further research to link IELTS and the Cambridge Main Suite Examinations/ALTE levels was conducted in 2000 as part of the ALTE Can Do Project. Can-do responses by IELTS candidates were collected over the year and matched to grades, enabling Can-do self-ratings of IELTS and Main Suite candidates to be compared. The results, in terms of mean 'can-do self-ratings', supported the findings from the earlier work.
>
> Both these studies support the placing of IELTS Band 6 to 6.5 at NQF Level 2 along with CAE. (www.cambridgeesol.org/assets/pdf/ielts-research-faqs.pdf)

So what is C1 on IELTS? 6? 6.5? 7? 8? These differences are very important for implementing an immigration policy where C1 is a crucial level for admission decisions. There is evidence from other carefully conducted linkage studies of competitor tests to IELTS such as TOEFL iBT and the new Pearson Academic test that there may be significant problems with the IELTS calibration. For as long as it lasts, Cambridge has a significant commercial advantage over its competitors, as it is far easier to meet the C1 standard in

https://doi.org/10.1017/S0261444811000073 Published online by Cambridge University Press

terms of the claimed IELTS equivalence than it is to meet the same C1 standard by using TOEFL iBT or Pearson Academic. This discrepancy needs to be resolved urgently.

5. The costs of a 'common currency'

Even if it were possible to impose a common currency for standards in language education, such a development would come at enormous cost, the discussion of which has barely begun (Fulcher 2004). The quote from Bourdieu reminds us of what this cost might be. The cost of unification is the devaluing of the local interpretation of the goals of education. One issue is that the limiting of the goals and meaning of language learning to functional, communicative objectives ignores the role of language learning in the subjective experience of the learner as an individual with a history, both personal and cultural.

Research in applied linguistics is beginning to address the meaning of language learning to individuals. The theme of identity and second language learning is currently being explored in a number of major publications (for a survey, see Block 2007). In particular, Pavlenko's studies of literary autobiographies (Pavlenko 2007) have shown how language learning is deeply embedded in the life trajectories of the authors; functionalist accounts are inadequate to deal with the complexities revealed in these accounts. Her work on language learning and emotions (Pavlenko 2005) has extended the scope of this research into other autobiographical accounts of language learning, language use and language loss. Kramsch's discussion of subjectivity in language learning (Kramsch 2009) interprets narratives of bilinguals and other learners within contemporary social theory, and shows how 'individual' experience is constructed within the operation of larger social forces. Judith Butler's work on desire in the construction of subjectivity (Butler 1997) points to a new framework for understanding such tired concepts as attitudes and motivation in language learning, which have lacked an appropriate social theory to guide research in this field.

Although recent work on language learning and individual identity is increasingly located within contemporary social theory, and adopts a social perspective, it is still focused on interpreting individual experiences of language learning, albeit in the light of social forces. We have yet to give an account of how language learning in different societies and cultures will have specific social meanings, and hence potential meanings for individuals, within the history of contact, cultural and political, of those societies with the societies and communities in which the target language is spoken. These historical conditions will mediate the encounters of individuals from different societies with the same language, so that the learning of a particular language may have different meanings according to the specific social and cultural background of the learner.

The imposition of a single set of cultural meanings and social and political values for language education, for each setting in which the CEFR is adopted, eviscerates the traditions of language teaching which are incompatible with the CEFR. In cultural and historical terms, learning English is simply not the same for a Singaporean, an Indonesian, a Vietnamese, a French person, a Dutch person, or a Hungarian. And different languages are indeed different, and vary in socio-political role, range and purpose of use, and carry very different histories of contact, often violent, between speakers of the speech communities involved. The CEFR

https://doi.org/10.1017/S0261444811000073 Published online by Cambridge University Press

puts all foreign languages into one and the same category, thereby erasing the fact that they are 'foreign' in very different ways. Learners from those backgrounds may be aware of the cultural and political significance of the act of learning the language of a cultural group with a particular historical relationship to their own, unless they are invited, in an act of collective historical amnesia, to wipe the slate clean and re-identify themselves as citizens of the new globalized world.

While the CEFR, in other words, is conceived of as something like the Euro, the currency of language is not simply convertible. Managers love the CEFR and other standards-based frameworks all over the world because they make available a certain kind of quantitative accountability in a way that was not possible before. But this all-too-ready surface translatability denies the untranslatable, the part of language and language learning which gives the enterprise value beyond labour force mobility. And being capable of 'participation in an international society' is not best achieved through historical amnesia, tempting though that may be given the horrors of the previous century.

A further problem with the imposition of universal language test constructs by managerial systems is that the constructs cannot easily be changed or challenged, because researchers and theorists of language testing are no longer responsible for determining the test construct. The evolution of notions of language proficiency will only slowly be reflected in test constructs, for two reasons. Firstly, because the test construct is embedded in policy processes, it is only through policy processes that the construct can be challenged or updated. Academic researchers are not necessarily very good at influencing policy debates. Secondly, because existing accountability frameworks such as the CEFR are deeply embedded in existing administrative procedures, there will be a natural reluctance to change them because of all the administrative and even legislative work that is entailed.

6. The need for change: English as a lingua franca

It seems to me that there is a great need for re-thinking the construct involved in language assessment. I think we are at a moment of very significant change, the sort of change that only comes along once in a generation or longer – the challenge that is emerging in our developing understanding of what is involved in ELF communication (Jenkins 2000, 2006, 2007; House 2003; Seidlhofer 2005, 2006; Seidlhofer, Breiteneder & Pizl 2006; Canagarajah 2007; Mauranen & Ranta 2009; Seidlhofer 2009; Pitzl 2010). I feel that this will be as big a change for language testing as the communicative revolution was for language teaching and for language testing 35 years ago.

Let me explain what I mean. Hyejeong Kim is currently completing a Ph.D. at Melbourne on the policy of the International Civil Aviation Organization (ICAO) governing English language testing, which has recently introduced English proficiency standards to be met by pilots and air traffic controllers working in international aviation settings (Kim & Elder 2009). Kim's study investigates attitudes to this policy within the aviation industry in Korea, which has in various ways resisted and sought to weaken the policy, through, for example, controversially revealing material from the English test used in Korea on the web well in

https://doi.org/10.1017/S0261444811000073 Published online by Cambridge University Press

advance of test dates, to assist those facing the test to pass. She has gained access to recordings of interaction between Korean air traffic controllers at airports in Korea (non-native speakers of English), and the pilots with whom they are communicating (a mix of non-native and native speakers of English). In some of these interactions, difficulties of mutual understanding emerged, which had potentially dangerous implications. Among the situations she studied are the following: a conversation with the Russian pilot of a Russian passenger plane who wanted to change course because of a technical problem, and who needed permission to do so; a conversation with an Australian pilot, a native speaker of English, whose plane had suffered a mechanical failure after landing and who had no power to leave a busy runway, and needed his plane to be towed to safety; a conversation with a French pilot who misheard an instruction and had therefore made a runway incursion, disrupting the landing path of incoming planes. I think we can agree that in all these situations, the communication breakdowns which occurred, but which were ultimately resolved, were potentially significant for airline and passenger safety.

What is the nature of the miscommunication in these incidents? Kim played recordings of the incidents and showed transcripts of the recordings to two groups of informants, Korean air traffic controllers, and Korean pilots, and using a form of think-aloud methodology, elicited their perspective on these incidents. The informants were asked to articulate their understanding of what was going on, and the nature and sources of the miscommunication, as the situations unfolded. These comments form the main data for the thesis. Kim then coded the data to see what they focused on. (The subject of study is thus informants' attitudes to examples of lingua franca communication in this setting.)

The main conclusion from the analysis is that while the language proficiency of some of the Korean air traffic controllers is an issue in some of the situations, the reasons for the communicative breakdown go far beyond the language proficiency of one of the participants. A number of factors come into play. First, the professional competence of the pilot and the air traffic controller is a crucial factor in avoiding or successfully resolving miscommunication. Professional competence is an important factor in shaping discourse, to the point that it is difficult to separate them. Second, and most relevant here, the responsibility for communication in English as a lingua franca is a joint responsibility, and failure is rarely the fault of the second-language speaker alone. The rules of communication in the aviation industry set out clear procedures for routine communication, involving the required use of set phraseology for routine requests, situation reports and explanations, and required procedures for reading back what has been heard, and confirmation that what has been read back is in fact what was originally said. For less predictable situations, spontaneously formulated 'plain English', that is, the ordinary code, is permitted. Kim's informants commented on two aspects of ELF communication. First, the native speaking or fluent non-native speaking pilots did not observe the rules about using routine phraseology, and unnecessarily used ordinary conversational English with the Korean air traffic controllers in entirely routine and predictable situations, which sometimes resulted in miscommunication. Second, the native speaker pilots did not know how to accommodate to their ELF interlocutors in terms of accent features, lexical choice, speed of delivery, and so on. In turn, air traffic controllers did not always acknowledge their lack of understanding, which subsequently complicated the misunderstanding.

https://doi.org/10.1017/S0261444811000073 Published online by Cambridge University Press

We can see then three features of ELF communication in this example (and these are amply borne out in the extensive existing literature on the subject, a sample of which was cited above):

(1) The interlocutor of the second-language speaker of English will not necessarily be a native English speaker (though he/she may be), but a speaker of another second language.
(2) Factors other than simple language proficiency – for example, professional competence – are an essential component of the communication.
(3) The responsibility for successful communication is a joint responsibility between the participants in the communication – it does not lie entirely with the non-native speaker.

Let us now consider the ICAO English proficiency standards (ICAO 2004). As you might expect, the wording of standards reflects to a certain extent the demands of the workplace setting. The following is the definition of the operational proficiency required to gain a professional licence:

> Proficient speakers shall:
>
> a) communicate effectively in voice-only (telephone/radiotelephone) and in face-to-face situations;
> b) communicate on common, concrete and work-related topics with accuracy and clarity;
> c) use appropriate communicative strategies to exchange messages and to recognize and resolve misunderstandings (e.g. to check, confirm, or clarify information) in a general or work-related context;
> d) handle successfully and with relative ease the linguistic challenges presented by a complication or unexpected turn of events that occurs within the context of a routine work situation or communicative task with which they are otherwise familiar; and
> e) use a dialect or accent which is intelligible to the aeronautical community. (ICAO 2004, A-2)

Although the ICAO proficiency standards do not refer explicitly to the fact that the relevant work setting involves ELF communication, the mention of the need to 'recognize and resolve misunderstandings' is potentially relevant to that construct. In order to further elaborate the meaning of the standard, six levels of proficiency are defined in terms of the degree to which they meet the standard. Air traffic controllers and pilots must reach at least Level 4 in order to get a certificate allowing them to continue working in their chosen profession. In defining the levels, various aspects of performance are identified, and descriptions of performance are provided at the relevant level for each aspect. The following dimensions of proficiency are identified: *Pronunciation*, *Structure*, *Vocabulary*, *Fluency*, *Comprehension* and *Interactions*. Apart from *Interactions*, this is quite a traditional set of rating categories for the assessment of speaking, and the descriptions within each are informed by understandings of language learning and language pedagogy that are quite general and not necessarily specific to the particular work setting. Here are the definitions of two of these dimensions, *Comprehension* and *Interactions*, at Level 4, the Operational Level:

COMPREHENSION

> Comprehension is mostly accurate on common, concrete, and work-related topics when the accent or variety used is sufficiently intelligible for an international community of users. When the speaker is

https://doi.org/10.1017/S0261444811000073 Published online by Cambridge University Press

confronted with a linguistic or situational complication or an unexpected turn of events, comprehension may be slower or require clarification strategies.

INTERACTIONS

Responses are usually immediate, appropriate, and informative. Initiates and maintains exchanges even when dealing with an unexpected turn of events. Deals adequately with apparent misunderstandings by checking, confirming, or clarifying. (ICAO 2004, A-8)

Distinctive features of the work setting are acknowledged at various points: note the mention of a 'situational complication or an unexpected turn of events', and there is an acknowledgement that communication will be among 'an international community of users', that is, that the target of the test is ELF communication as much as, or more than, communication with native speakers of English. The phrase 'deals adequately with apparent misunderstandings by checking, confirming, or clarifying' again acknowledges the mutual nature of the communication.

Nevertheless, it is interesting to note that as a matter of policy only one side is subjected to the examination – the native speakers are exempt, although Kim's study shows their contribution to miscommunication. Moreover, the scale of proficiency defines subsequent higher levels in terms of increasing control of the code, not increasing ability to deal with increasingly complex situations of miscommunication.

For example, the distinction between subsequent levels (5 and 6) does not appear to be specifically related to the work setting. The criterion of comprehension at Level 5 refers to ability 'to comprehend a range of speech varieties (accent and/or dialect) or registers' whereas Level 6 states 'Comprehension is consistently accurate in nearly all contexts and includes comprehension of linguistic and cultural subtleties'. One might have here expected some further reference to the conditions of communication in terms of greater demand – unfamiliar accent, styles of delivery, and so on. Interestingly, in their actual working life, the pilots themselves are oriented to the demands of lingua franca communication. Kim (personal communication) reports that when flying to a new destination, the pilots listen to tapes of communication in English by speakers of ELF from that region, in order to familiarize themselves with the accent of the air traffic controllers in that destination. They do not listen to more tapes of native speakers in order to improve their comprehension of native speaker speech. Native speaking pilots might be recommended to listen to tapes of local air traffic controllers, but there is no ICAO policy to enforce that – in fact it is never discussed, because of the presupposed privileging of native speaker norms.

There have been several developments in response to the ICAO policy. ICAO's framework and the definition of operational competence are not tied to any particular test; each national system can choose to meet the standard in whatever way it wants. Many national civil aviation authorities have endorsed certain existing tests, usually general proficiency tests, as suitable instruments for eliciting evidence of the standing of individual pilots and air traffic controllers in relation to the standard. These tests are often associated with general proficiency scales that are in turn linked to the CEFR. There are two problematic aspects of this development. The first is that general proficiency scales do not articulate the relevant construct, which

https://doi.org/10.1017/S0261444811000073 Published online by Cambridge University Press

is ELF communication. Secondly, they locate the issue in the second language proficiency, narrowly defined, of the non-native speaker. The International Language Testing Association is currently attempting to establish itself as an accrediting agency for the tests being adopted by different national civil aviation authorities. While this has many positive aspects, in that it will deal with aspects of the FAIRNESS of the test, it will not deal with the lack of JUSTICE embodied in the construct (McNamara & Ryan 2011). And because to change the construct involves a complex political process, it is unlikely to change in a hurry. In response to this impasse, the Korean civil aviation industry have taken the controversial position of publishing the test online in advance so that pilots and air traffic controllers can more easily pass it. They are indicating in this way their own disbelief in the construct of the test; it is not that they want to conduct civil aviation in Korea recklessly or dangerously – in fact Korea has an excellent civil aviation record – but they resent the injustice of very experienced and competent air traffic controllers losing their jobs needlessly.

Turning to the CEFR, to what extent is the construct of ELF communication adequately represented? As Seidlhofer (2003) pointed out, it is absent from the framework – not surprising, given its intellectual antecedents in the European version of the communicative movement of the 1970s. Here are the definitions of the two most advanced levels of the CEFR, Level C, Proficient User, divided into two levels, a higher (C2) and a lower (C1).

Proficient User

C2 – Overall

Can understand with ease virtually everything heard or read. Can summarize information from different spoken and written sources, reconstructing arguments and accounts in a coherent presentation. Can express him/herself spontaneously, very fluently and precisely, differentiating finer shades of meaning even in more complex situations.

C1 – Overall

Can understand a wide range of demanding, longer texts, and recognize implicit meaning. Can express him/herself fluently and spontaneously without much obvious searching for expressions. Can use language flexibly and effectively for social, academic and professional purposes. Can produce clear, well-structured, detailed text on complex subjects, showing controlled use of organizational patterns, connectors and cohesive devices.

For a more detailed understanding of how the levels are defined for the skills of listening and speaking, we can turn to the descriptors in the self-assessment grid:

Listening

C2

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

https://doi.org/10.1017/S0261444811000073 Published online by Cambridge University Press

C1

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

Spoken interaction

C2

I can take part effortlessly in any conversation or discussion and have a good familiarity with idiomatic expressions and colloquialisms. I can express myself fluently and convey finer shades of meaning precisely. If I do have a problem I can backtrack and restructure around the difficulty so smoothly that other people are hardly aware of it.

C1

I can express myself fluently and spontaneously without much obvious searching for expressions. I can use language flexibly and effectively for social and professional purposes. I can formulate ideas and opinions with precision and relate my contribution skilfully to those of other speakers.

It is clear that the concept of ELF communication does not inform these descriptors. In terms of COMPREHENSION,

(1) The interlocutors are assumed to be native speakers.
While the wording 'provided I have some time to get familiar with the accent' might lead us at first to think that this means non-native speaker accents, we realize that this cannot be right, as the references to 'at fast native speed' and 'a good familiarity with idiomatic expressions and colloquialisms' make clear.
(2) The role of other competences as an essential basis for comprehension is not considered.
(3) The responsibility for successful communication is held to lie entirely with the non-native speaker.

In terms of PRODUCTION, the references in the definition of Spoken Interaction are revealing: 'I can express myself fluently and convey finer shades of meaning precisely. If I do have a problem I can backtrack and restructure around the difficulty so smoothly that other people are hardly aware of it'. The problem is presumably 'mine' alone, not a problem caused by the competence of my interlocutor. What is required in such a situation is not the ability to 'convey finer shades of meaning'. Again, the ever more demanding native speaker interlocutor is the presumed target of the communication efforts.

My point here is not only the need for change, the need to challenge the construct, but what it would take to move the mountain of the CEFR in the direction of change. The arguments of applied linguists are unlikely to be persuasive in themselves – policy statements need policy action to alter them.

https://doi.org/10.1017/S0261444811000073 Published online by Cambridge University Press

7. Conclusion

Let me make a very brief conclusion. In this paper I have argued that language assessment, like language education more generally, is increasingly serving the goals of policy, and specifically of policies supporting a view of education as primarily preparing learners for participation in a globalized workforce. This view of education, and of assessment within it, seems to me to constitute a serious impoverishment. Moreover, the impact of globalization has had a profound impact on the role of English. As the extensive literature on ELF communication makes clear, the emergence of English as a lingua franca as a key feature of a globalized world presents a powerful challenge to assumptions about the authority of the native speaker, an authority which is enshrined in test constructs; tests are the enforcers of native speaker privilege. Claims to the ownership of English and the privileging of native speaker identity inhibit the appropriation of English to the actual demands of communication in a globalized world. I think we have here a powerful example of the theme of the conference, 'Global Perspectives, Local Initiatives'. In language education, a core cultural practice, we need to base our capacity for participation in such a world on local histories of the meaning of learning a language such as English. The tension between the global and the local is felt here as in other areas of contemporary life. Moreover, what globalization itself demands is flexibility, and flexibility is precisely what is lacking in universalizing outcome statements such as the CEFR, which are the lever of control of education systems and social policies involving language throughout the world. It is time to engage with the serious re-thinking that has been going on for a decade or more now – but this is not something we can expect to be sponsored within organizations such as the Council of Europe, or testing organizations such as Cambridge, whose interests lie elsewhere, and who we can expect to resist the fundamental reorientation that I believe is required. The levers of change we must seek are initially conceptual, but given the politicization of test constructs that I have outlined, that is inadequate in itself. Here the resistance of the aviation industry in Korea to the demands of the ICAO policy is an example of the complex political work that will be required to achieve change. Applied linguists may not be at the forefront of this kind of work, but we can at least supply the arguments and the understanding to support the activity of those who are in a position to force the policy change that is clearly needed.

Acknowledgements

The initial spark for the ideas in the latter half of this paper was a conversation initiated by Barbara Seidlhofer of the University of Vienna at the annual meeting of the British Association of Applied Linguistics at Bristol in 2005, about the implications for testing of work on English as a lingua franca. I am grateful to Barbara and to Henry Widdowson for the ongoing stimulus and opportunity to think through and discuss this complex question, and to Henry Widdowson for detailed commentary on an earlier version of this paper. This paper draws in part on material published as 'The use of language tests in the service of policy: Issues of validity', which appeared in *Revue Française de Linguistique Appliquée* XV, 1: 7–23 (2010) (Special issue: Evaluer les compétences langagières), accessible at www.cairn.info/revue-francaise-de-linguistique-appliquee.htm, and 'A two-edged sword: Comparative assessment

https://doi.org/10.1017/S0261444811000073 Published online by Cambridge University Press

frameworks and language education', which appeared in *Toegepaste Taalwetenschap in Artikelen* 82, 2: 9–19 (2009).

Note

Supplementary material accompanies this paper on the journal's website (http://journals.cambridge.org/lta).

References

Alderson, J. C. (1991). *Bands and scores*. In J. C. Alderson & B. North (eds.) *Language testing in the 1990s*. London: Modern English Publications & The British Council, 71–86.

Block, D. (2007). *Second language identities*. London: Continuum.

Bourdieu, P. (1991). *Language and symbolic power*. London: Polity Press.

Butler, J. (1997). *The psychic life of power: Theories in subjection*. Stanford, CA: Stanford University Press.

Canagarajah, S. (2007). Lingua franca English, multilingual communities, and language acquisition. *The Modern Language Journal* 91, Focus Issue, 923–939.

Council of Europe (2001). *Common European framework of reference for languages: Learning, teaching and assessment*. Cambridge: Cambridge University Press.

Davies, A. (2008). *Assessing academic English: Testing English proficiency 1950–89: The IELTS solution*. Cambridge: Cambridge University Press.

Fulcher, G. (2004). Deluded by artifices? The Common European Framework and harmonization. *Language Assessment Quarterly* 1.4, 253–266.

House, J. (2003). English as a lingua franca: A threat to multilingualism? *Journal of Sociolinguistics* 7, 556–578.

International Civil Aviation Organization (ICAO) (2004). *Manual on the implementation of ICAO language proficiency requirements*. Montréal: International Civil Aviation Organization.

Jenkins, J. (2000). *The phonology of English as an International Language*. Oxford: Oxford University Press,

Jenkins, J. (2006). Current perspectives on teaching World Englishes and English as a lingua franca. *TESOL Quarterly* 40, 157–181.

Jenkins, J. (2007). *English as a lingua franca: Attitude and identity*. Oxford: Oxford University Press.

Kim, H. & C. Elder (2009). Understanding aviation English as a lingua franca: Perceptions of Korean aviation personnel. *Australian Review of Applied Linguistics* 32.3, 23.1–23.17.

Kramsch, C. (2009). *The multilingual subject*. Oxford: Oxford University Press.

Mauranen, A. & E. Ranta (eds.) (2009). *English as a lingua franca: Studies and findings*. Newcastle upon Tyne: Cambridge Scholars Publishing.

McNamara, T. (2001). Language assessment as social practice: Challenges for research. *Language Testing* 18.4, 333–349.

McNamara, T. (2011). Measuring deficit. In C. N. Candlin & J. Crichton (eds.), *Discourses of deficit*. London: Palgrave Macmillan, 311–326.

McNamara, T. & C. Roever (2006). *Language testing: The social dimension*. Malden, MA & Oxford: Blackwell.

McNamara, T. & K. Ryan (2011). Fairness vs. justice in language testing: The place of English literacy in the Australian citizenship test. *Language Assessment Quarterly* 8.2.

Morrow K. (ed.) (2004). *Insights from the Common European Framework*. Oxford: Oxford University Press.

North, B. (2006). The Common European Framework of Reference: Development, theoretical and practical issues. Paper presented at the symposium 'A New Direction in Foreign Language Education: The Potential of the Common European Framework of Reference for Languages', Osaka University of Foreign Studies, Japan, March.

Pavlenko, A. (2005). *Emotions and multilingualism*. Cambridge: Cambridge University Press.

Pavlenko, A. (2007). Autobiographic narratives as data in applied linguistics. *Applied Linguistics* 28.2, 163–188.

https://doi.org/10.1017/S0261444811000073 Published online by Cambridge University Press

Pitzl, M.-L. (2010). *English as a lingua franca in international business. Resolving miscommunication and reaching shared understanding*. Saarbrücken: VDM-Verlag Müller.
Seidlhofer, B. (2003). *A concept of international English and related issues: From 'real English' to 'realistic English'?* Strasbourg: Council of Europe, Language Policy Division DG IV – Directorate of School, Out-of-School and Higher Education.
Seidlhofer, B. (2005). Standard future or half-baked quackery? Descriptive and pedagogic bearings on the globalisation of English. In C. Gnutzmann & F. Intemann (eds.), *Globalisation and the English language classroom*. Tübingen: Narr, 155–169.
Seidlhofer, B. (2006). English as a lingua franca in the expanding circle: What it isn't. In R. Rubdy & M. Saraceni (eds.), *English in the World: Global rules, global roles*. London: Continuum, 40–50.
Seidlhofer, B. (2009). Common ground and different realities: World Englishes and English as a lingua franca. *World Englishes* 28.2, 236–245.
Seidlhofer, B., A. Breiteneder & M-L. Pitzl (2006). English as a lingua franca in Europe. *Annual Review of Applied Linguistics* 26, 1–34.
Shohamy, E. (2001). *The power of tests: A critical perspective on the uses of language tests*. London: Pearson.
Shohamy, E. (2006). *Language policy: Hidden agendas and new approaches*. Abingdon & New York: Routledge.
Spolsky, B. (1995). *Measured words*. Oxford: Oxford University Press.
Taylor, L. & N. Jones (2006). Cambridge ESOL exams and the Common European Framework of Reference. *University of Cambridge ESOL Examinations Research Notes* 24: 2–5.
Trim, J. L. M. (1980). *Developing a unit/credit scheme of adult language learning*. Oxford & New York: Pergamon Press on behalf of the Council of Europe.
Van Avermaet, P. (2009). Fortress Europe? Language policy regimes for immigration and citizenship. In G. Hogan-Brun, C. Mar-Molinero & P. Stevenson (eds.), *Discourses on language and integration: Critical perspectives on language testing regimes in Europe*. Amsterdam: John Benjamins, 15–43.

TIM MCNAMARA is Professor in the School of Languages and Linguistics at The University of Melbourne, where he teaches and supervises at graduate level in Applied Linguistics. His publications include *Measuring second language performance* (Longman, 1996), *Language testing* (OUP, 2000) and *Language testing: The social dimension* (with Carsten Roever, Blackwell, 2006), which won the Sage/ILTA Award in 2009. His main areas of research are in language testing, including its social and political aspects, and poststructuralist approaches to language.

https://doi.org/10.1017/S0261444811000073 Published online by Cambridge University Press