From job analysis to task design: different approaches to simulating teacher language behaviour

Kathryn Hill
The NLLIA Language Testing Research Centre
The University of Melbourne

Abstract

Three language teacher proficiency tests (Italian, Indonesian, Japanese) were developed. Although based on the same job analysis and theoretical underpinnings, it was necessary to operationalise them differently. The factors that motivated the differences among the tests are explored as are the implications for setting common benchmarks for language teacher accreditation.

1. Background

In Australia, a target has been set for 25% of all final year high school students to be studying a foreign language by the year 2006 (Rudd 1994: 106).

Demand for language teachers created by government policy and community pressure for programmes in certain languages has resulted in "an increasing number of unqualified teachers... being pressed into teaching these languages. This is especially the case in those states which lack a proficiency standard for employment" (Rudd 1994: 92). Variable levels of language proficiency amongst LOTE teachers had already been identified as a problem in an earlier report on the employment and supply of teachers (Nicholas et al 1993).

In line with these concerns the Rudd Report calls for national standardised teacher proficiency tests (1994: 140). However, putting aside the issue of whether individual Australian states could ever agree on a national system of language teacher certification, there is a question of whether uniform testing procedures and a common minimum standard for language teachers are possible across all languages.

This paper will consider this question in relation to the development of three LOTE teacher proficiency tests: Italian (Elder 1993) , Japanese (Elder, Iwashita & Brown 1995) and Indonesian (Hill & Arnost, 1996). The three tests share the same conceptual basis and many common features. However, there a number of differences between the tests which, we suggest, arise from factors which are specific to each language.

2. Specificity versus uniformity

The tests are 'specific purpose' not only in terms of being specific to the profession of language teaching but also specific to the language concerned (i.e. Indonesian, Italian or Japanese). The idea of a common minimum standard, however, implies a need for uniformity in testing across the three languages both in terms of test design (format and content), and standard (i.e. overall difficulty). The question raised here then is, is there a tension between specificity and uniformity?

2.1 Uniformity

Test purpose

On the side of uniformity, the three tests have the same purpose, sample from the same language domain and share many common features in their design. All three tests are intended for use in the screening, diagnosis and/or certification of language teachers.

Screening

It is assumed, not always correctly, that students have already attained an acceptable level of proficiency by the end of their first degree. Teacher education courses in Australia, therefore, do not typically provide further language tuition. The tests provide a means for identifying applicants for teacher training courses who have inadequate language skills.

Diagnosis

The tests make the language abilities expected of the language teacher explicit so that individuals' strengths and weaknesses can be measured in relation to this.

Certification

Particularly in the case where applicants do not have formal training or qualifications in the language (e.g. mother tongue and background speakers, or those who have spent time 'in country'), the tests provide an indication of whether an applicant's language proficiency is sufficient to carry out the role of language teacher.

Test domain

Another point in common is that all three tests are based on a content analysis conducted by Elder for the Italian teacher proficiency test (see Elder 1994). In her article, Elder describes the process undergone during the test design stage. This involved three stages: a review of the literature on second language acquisition research, a job analysis and a review of the curricula for Italian in schools.

On the basis of her literature review, Elder concluded that teachers need to demonstrate an ability to:

(i)   use the target language as medium of instruction

(ii)   modify the target language input in such a way as to render it comprehensible to learners

(iii) produce well-formed input for learners

(iv)   draw learners' attention to the formal features of the target language

Elder (1994)

The job analysis drew on expert opinion, interviews with 'excellent' teachers and classroom observation to determine both what language teachers say they do as well as what experts believe they should be able to do.

This analysis resulted in an inventory of tasks which were classified into interactions oriented towards 'core' pedagogic goals, those which serve to create a framework within which teaching can take place (both based on Ellis', 1984 goal-based framework of classroom interaction) and those involving 'extra classroom use' of the language. An example of core pedagogic goals is modelling the target language. An example of creating a framework for teaching is issuing homework, and examples of extra classroom use include

preparing for the classroom or interacting with the relevant language community.

The original content analysis was used as the basis for the Japanese and the Indonesian tests.

Test design

The third area of similarity between the tests is test design. The Rudd Report (1994) calls for tests which measure

"-what <u>candidates can do</u> with the language rather than what they know about the language

- what skills, tasks and content areas are <u>relevant to the teaching of the language;</u>

-what <u>reading, writing, listening, speaking</u> tasks and types of topic a teacher would be expected to encounter." (Rudd, 1994: 140)

In line with these recommendations, the three tests are all performance based, i.e. they include tasks which resemble the types of tasks required of candidates in their role as teachers. However, whilst many of the tasks bear an obvious relation to teaching, they are not intended to be a test of methodology, i.e. it is the abilities underlying task performance which are of interest.

Also in line with these recommendations, each test covers the four macro skills.

2.2. Specificity

However, although the tests share very many features in common, they nevertheless differ in a number of respects. These differences can be attributed to differences in:

- the nature of the L2;
- the nature of the target population;
- the content of teaching;
- the culture of teaching, and
- the political context.

Nature of the language

Many of the differences were driven by the discrepancy in the inherent difficulty of the three languages for English speakers. It has been suggested that Chinese and Japanese take a considerably longer time for English speakers to learn than European languages, with Indonesian somewhere in between (Kirkpatrick 1995). Evidence for these differences is the well attested discrepancy between the exit proficiencies of students of European languages and Asian languages in Australian universities.

The difficulty of the language dictated what was possible on a given test. For example, for the read aloud/retelling task task, candidates are required to read a text aloud and then retell it as if to a class of learners. For Italian and Indonesian, the text was a short story and a short newspaper article respectively. However, for Japanese it was deemed necessary to use a much simpler text, a post card, with a limited number of kanji.

Again due to the difficulty factor, it was considered necessary to reduce the number of written prompts in the Japanese test, which relies more heavily on picture prompts than the other tests. Topics in the test are also linked thematically in an attempt to reduce the cognitive load for candidates.

Finally, whilst candidates were permitted to use dictionaries for all sections of the Japanese test, it was not considered appropriate for the Italian and Indonesian tests. Dictionaries were considered necessary for Japanese as teachers who have been away from formal study for a period tend to have a character attrition problem[1].

Nature of the target population

Italian is a community (heritage) language in Australia. One is therefore more likely to encounter mother tongue or background speakers teaching Italian than is the case for Indonesian and Japanese. Italian teachers are also more likely to have significant numbers of mother tongue or background speakers in their classes.

---

[1] Access to dictionaries is also the convention in final year high school examinations.

Whilst they may have difficulty with reading and writing, spoken fluency tends to be less of a problem for Italian teachers than for teachers of Indonesian and Japanese. Hence, during the Speaking section of the Italian test candidates are required to set up (i.e. explain) as well as participate in a role play task. However, this requirement was not considered reasonable for the other languages.

For the same reason, all instructions for the Italian test are in the target language. However, for Indonesian and Japanese it was felt that having instructions in the L2 might confound the measurement of the skill of interest, i.e. there was concern that a candidate might get the item wrong through failing to understand the prompt (though, for Indonesian there was some dissent on this issue).

Content of teaching

Going back once again to the Rudd Report (1994), the test designers were encouraged to consider

"- what skills, tasks and content areas are <u>relevant to the teaching</u> of the language;

-what reading, writing, listening, speaking <u>tasks and types of topic a teacher would be expected to encounter</u>." (Rudd 1994) emphasis added.

Any language teacher proficiency test, therefore, needs to take the demands of the syllabus into account, i.e. the test developer needs to have in mind the materials and activities the teacher will have to be able to deal with in the classroom.

The issue of which variety of the target language should be represented in the test did not arise in the case of the Italian and Japanese tests. However, a large amount of work has recently been done in Australia to develop teaching materials which include authentic texts containing non-standard Indonesian. It was therefore considered important that the test include reading and listening texts with this type of language (though teachers are not expected to be able to produce this language themselves).

Culture of Teaching

In Italian classrooms, using the L2 to explain grammatical points is standard practice. Hence, the test includes a task where the

candidate is required to use Italian to explain errors in a written text as if to a student.

For Indonesian, it was deemed important that teachers use Indonesian for this type of task. However, as this is not part of what teachers already do in the Indonesian classroom, there was a lot of negative feedback from teachers about this task. Some claimed that their *students* wouldn't be able to cope if they used Indonesian whilst others insisted that this type of task was more properly carried out in English. As they were being asked to perform an unfamiliar task, candidates were given 15 minutes before the interview (with access to a dictionary) to prepare for this task.

For Japanese, however, this task was felt to be far beyond the level of competence of even the most proficient teachers.

	Italian	Japanese	Indonesian
the target language			
read aloud / retell task	story	post card	short article
access to dictionary	✗	✔	✗
the target population			
assigning a role play	✔	✗	✗
instructions in L2	✔	✗	✗
content of teaching			
language variety	standard	standard	standard + non-standard
culture of teaching			
use L2 metalanguage	✔	✗	✗

Table 1. Summary of differences between the three teacher proficiency tests

Political concerns

It is important to note that not all differences were driven by linguistic and cultural concerns but were also influenced by external

factors. Each test was developed at a different time, under different funding schemes and in different political climates.

For example, the Indonesian test was required to articulate with the recently developed specific purpose version of ASLPR for Second Language Teachers. One consequence of this was that the separate test of formal language knowledge, included on the other tests, was not included on the grounds that there is no facility in the scales for reporting on this.

3. Conclusion

Whilst I have dwelt on differences in this paper, I should stress again that three tests share a large amount in common. The tests are all concerned with language ability as relevant to teaching and all were carefully designed, trialled and validated.

However, we believe that this study indicates that, if tests are to be truly specific in terms of language and context of use, they will need to differ, at least in some respects, across languages. Our experience also suggests that a uniform minimum standard for language teachers is not enforceable in the short term.

These differences have implications for equivalence of qualifications for teachers of different languages. Whilst format differences (e.g. video/audio) should not be seen as a barrier, the different levels of difficulty across the three tests appear more problematic.

References

Elder, C. (1993) The Proficiency Test for Language Teachers: Italian. Volume 1: Final Report on the Test Development Process. National Project Report for DEET, Canberra. 34 pp.

Elder, C., N. Iwashita & A. Brown (1995) The Proficiency Test for Language Teachers: Japanese. National ILOTES Project Report for DEET, Canberra. Final report on the test development process. National ILOTES Project Report for DEET, Canberra. NLLIA Language Testing Research Centre, University of Melbourne. 30pp.

Elder, C. (1994) Performance Testing as a benchmark for LOTE Teacher Education, <u>Melbourne Papers in Language Testing.</u> Vol. 3, No. 1., 1994: 1-25

Ellis, R. (1984) <u>Classroom Second Language Development</u> Prentice Hall.

Hill, K. and V. Arnost (1996) <u>The Asian Languages Teacher Proficiency Project - Indonesian</u> Canberra: DEET

Kirkpatrick, A. (1995) Learning Asian Languages in Australia: Which languages and when? <u>Babel: Journal of the Australian Federation of Modern Language Teachers' Association</u> 30, 1: 4-11 and 26-29.

Nicholas, H. , H. Moore, M. Clyne and A. Pauwels (1993) <u>Languages at the Crossroads: The Report of the National Enquiry into Employment and Supply of Teachers of Languages Other Than English</u>. Canberra: National Languages and Literacy Institute of Australia.

Rudd, K.M. (1994) <u>Asian Languages and Australia's Economic Future: A report prepared for the Council of Australian Governments on a proposed national Asian languages/studies strategy for Australian Schools</u>. Queensland Government Printer.