

# **DESCRIPTORS – A WAY FORWARD FOR TRANSLATOR AND INTERPRETER TEST ASSESSMENT IN AUSTRALIA**

*Paper presented to the University of Western Sydney Interpreting and Translating Research Symposium, 26 September 2008.*

Barry Turner  
barry.turner@rmit.edu.au

## **ABSTRACT**

This paper is substantially based on one given to the 18<sup>th</sup> World Congress of the International Federation of Translators (Turner, 2008). It examines some assessment methodologies used for large-scale translating and interpreting accreditation testing and particularly different approaches to the use of descriptors or negative marking systems in such assessment. A report by RMIT University (Turner and Ozolins, 2007) showed that the UK Institute of Linguists and the American Translators Association (ATA) have moved or are moving towards using descriptors or combining negative marking and descriptors. The paper will explore the advantages and disadvantages of using negative marking and descriptors in translator and interpreter accreditation examinations. It will also present the rationale behind the author's conclusions that NAATI should replace its current error analysis / deduction system for assessing translation tests and introduce descriptor-based testing for its translation and interpreting tests.

## **1. INTRODUCTION**

NAATI has been in existence for 30 years, during which time successive editions of the NAATI Examiners' Manual have provided a reasonably detailed error analysis / deduction system for marking translation tests. However, much less guidance has been provided for marking interpreting tests where examiners have lacked either a holistic (appraisal of overall quality guided by examples and / or criteria) or an error analysis / deduction system.

The results of NAATI translating and interpreting tests have always been expressed in percentage terms. Candidates are required to achieve a minimum overall mark of 70% to achieve accreditation.

I was a co-developer of the 2005 edition of the NAATI Manual, which provided some more holistic guidance on assessing both translating and interpreting tests. The Manual continued to provide for translating and interpreting test results to be expressed in percentage terms.

The error analysis / deduction method of marking translation tests was retained and this section of the manual was strengthened with much more detailed guidance on error analysis / deduction.

The interpreting section of the Manual continued to give much less guidance than the translation section. Some suggestions were offered on how marks might be deducted in the dialogue interpreting component of the interpreting test but relatively little guidance was provided on assessing the consecutive interpreting (interpretation of short speeches ) component.

In the NAATI-funded study referred to earlier (Turner and Ozolins, 2007), our research team gathered information on 24 translating and interpreting accreditation bodies around the world. We found that there are a number of accreditation systems in use, including:

- error analysis / deduction systems (like NAATI's system for assessing translation tests);
- a combination of criterion-referencing (the use of scales of descriptors to describe performance in tests) and error analysis / deduction; and
- criterion referencing (descriptors) with no system of error analysis / deduction.

In this paper I shall concentrate on the assessment of translation tests, although I will also touch upon the assessment of interpreting tests. Descriptive or theoretical themes, such as establishing the nature

of a good translation, the nature of translation errors, assessing quality on the basis of text linguistic analysis (Waddington 2001: 311, 312) are outside the scope of this paper. Rather, my emphasis will be on the assessment of translation tests in public accreditation systems, and I shall include the accreditation testing conducted by my University in this public accreditation category. RMIT, like a number of other universities in Australia, conducts NAATI approved translating and interpreting programs and our testing is carried out in accordance with NAATI guidelines.

Public accreditation systems, such as NAATI's, tend to recruit examiners in a large number of languages from a variety of professional backgrounds. Source texts used in such systems tend to be professional, rather than literary in nature and examiners often lack a detailed appreciation of translation theory. Public accreditation bodies throughout the world tend to employ relatively straightforward methods to identify and categorise errors in translation tests.

In this paper I shall identify what I believe to be the best aspects of the UK Diploma of Public Service Interpreting (DPSI) and the American Translators Association (ATA) tests. I shall outline my preference for a combination of descriptor-based testing with a system for identifying, categorising and taking into account errors in translated texts.

## **2. METHODS OF ASSESSING TRANSLATED TEXTS**

In an attempt to ascertain the validity of methods of examining student translations, Christopher Waddington surveyed teachers from 20 European and Canadian universities, of whom 52 responded. Of these respondents, 36.5% used a method based on error analysis, 38.5% used a holistic method, and 23% combined error analysis with a holistic method. (Waddington, 2001: 313)

Waddington then used five examiners to assess 64 Spanish-into-English second-year University translation paper (a newspaper editorial). After undergoing training, they applied each of the following four methods to assess the translations which they completed in different order and with a month's interval between each method (Waddington 2001: 316):

### **Method A**

This type of error analysis involves the categorisation of mistakes under the following headings:

- inappropriate renderings which affect the understanding of the source text;
- inappropriate renderings which affect expression in the target language;
- inadequate renderings which affect the transmission of either the main function or secondary functions of the source text;
- distinctions are made between serious errors (-2 points) and minor errors (-1point), and there is provision for the awarding of points for good (+1 point) or exceptionally good (+2 points) solutions to translation problems. (Waddington, 2001: 313)

### **Method B**

This second type of error analysis takes into account the negative effect of errors on the overall quality of translations. Examiners have to determine if errors are translation or language mistakes. Language mistakes are penalised with -1 point. Where translation errors are identified, the examiner has to identify to what extent the error affects the text and deduct points accordingly. For example, a mistake which has a negative effect on 1-5 words in the target text would be penalised -2 points and one that affects 81-100 words would incur a penalty of -7 points. (Waddington, 2001: 314)

### **Method C**

Waddington designed his own "holistic method" because the few sent to him by respondents were based on the requirements of professional translation and were of little use for judging the translations of students into a foreign language. He created scales of descriptors at five levels and

divided vertically between accuracy of transfer of source text content and quality of expression in the target language. For example, the level 4 descriptor for accuracy of transfer of ST content was:

Almost complete transfer: there may be one or two insignificant inaccuracies; requires certain amount of revision to reach professional standard. (Waddington 2001: 315)

Each level attracted one or two points and the total points that could be achieved was 10 (out of 10).

#### **Method D**

Method D combined error analysis Method B and holistic Method C in a proportion of 70/30, with Method B accounting for 70% of the total mark and Method C for the remaining 30%. (Waddington 2001: 315)

Interestingly, Waddington found that there were no significant differences in the grades given to students and that “all four methods have proved to be equally valid in spite of the considerable differences that exist between them.” (Waddington, 2001: 322)

### **3. WADDINGTON’S METHODS AND NAATI TRANSLATION TEST ASSESSMENT METHODS**

Although Waddington’s research involved translations by students who were not translating into their “A” language and their translations were not at the professional level, the methods he describes have wide currency. The closest of Waddington’s methods to the system of error analysis / deduction in the NAATI Examiners’ Manual is Method A.

The interpreting section of the Manual provides guidance on standards expected of interpreters at the Paraprofessional Interpreter and Interpreter levels but does not provide descriptors or a system for error analysis / deduction. The translation section provides a detailed system of error analysis / deduction and more detailed descriptions of what examiners should expect from translation test candidates. It provides examples of errors of different severity, and guidance on the points (from ½ a mark to 5 points or more) that should be deducted in each case. It categorises errors in terms of accuracy, quality of language and technique.

NAATI translation test passages at the Translator level are approximately 250 words in length. Candidates must select two passages and each is accorded 45 points (ten points are reserved for questions on ethics of the profession).

Errors of factual accuracy (distortions, omissions and additions) incur more penalty points than the other categories. A tendency to make a number of errors in these categories is viewed seriously. Language quality considerations relate to appropriate register, grammar, idiom, collocations and spelling. Significant and prolonged inappropriacies in these areas incur significant penalties.

The Manual advises that -.5 to -2 points may be deducted for errors involving incorrect spelling and punctuation, -1 to -3 points may be deducted for errors involving distortions that do not extend beyond the phrase, clause or sentence level, and -5 or more points may be deducted for more serious distortions that affect meaning in other parts of the text. There is a cut-off average percentage of 70% for the two translation passages that each candidate must attempt, below which test candidates are deemed to have failed.

To achieve 70% candidates must obtain a score of 63 points (out of 90) for both passages. A system of symbols is used to indicate whether mistakes affect accuracy or quality of language. For example, words deemed inaccurate are enclosed in a rectangle by the examiner, punctuation and spelling errors are circled, and serious extensive errors are wavy-underlined. Errors affecting accuracy are annotated in solid lines and those relating to quality of language are in dotted lines. This enables examiners to tally results later and provides an instant picture of how many errors have been identified and the number and type of errors throughout the translation.

If the same number of points per line of text is deducted as in the example below, it would not take long for a candidate to be penalised more than 27 points for the two passages and fail to achieve accreditation standard. In the table below, the title of a source text is shown in the first row. The points deducted, in order of the appearance of errors in the text, are shown to the left of the candidate's response in the second row:

<b>Source Text</b>	<i>Pollution Advocacy: Demanding the Public's Right to a Clean Environment</i>
2 1 ½ ½	<i>Pollution Lawyers: Suing for the Public's Right for a Clean (Living) Environment</i>

Accordingly, pass rates for NAATI Translator tests are generally low with only 12.89% achieving accreditation at the Translator (into English) level. Interestingly, despite the relatively low level of guidance in the interpreting section, pass rates are similar with only 15.22% of interpreting test candidates passing in the 2005-06 annual testing program (Turner and Ozolins, 2007: 18). These rates appear to be similar to those of similar overseas testing organisations (Stejskal, 2003: 30, 37).

#### 4. DESCRIPTORS

The use of descriptors in the language field has been pioneered by organisations that assess general language competence. One of the best known descriptor-based systems for assessing language competence is the International English Language Testing System (IELTS).

The introduction of holistic systems such as the IELTS has resulted in greater sophistication in language assessment. Unlike previous error analysis or intuitive and non-standardised means of assessing language, IELTS and similar descriptors describe the way test candidates use language beyond the level of a particular word or grammatical category which can be assessed as right or wrong. Importantly, they describe what a candidate *can* do as well as what they *cannot* do (e.g. uses less common and idiomatic vocabulary skillfully with occasional inaccuracies') and also what in this case the listener will hear: a high degree of idiomatic utterances, with occasional inaccuracies, leading the assessor to look at *both* skilful use and inaccuracies.

A major difference between descriptor systems designed to holistically assess language competence and those employed in assessing translation competence is that in translating and interpreting there is an absolute need to ensure that the outcome is a faithful rendition of the original source text that achieves a substantial degree of pragmatic and semantic equivalence, and that the quality of expression is accurate and appropriate. Of course, this requires constant references back to the source text, both by the candidate in producing the translation, and by the examiner in assessing it. The use of descriptors that fail to take sufficiently into account these particular aspects of translation and the assessment of translations could result in examiners failing to take sufficient note of errors in the target text.

#### 5. SUPPORT FOR THE CURRENT NAATI TRANSLATION TEST ASSESSMENT SYSTEM

The error analysis / deduction method currently employed in assessing NAATI translation texts enjoys wide support from the Authority's examiners. This was apparent at a series of workshops held throughout Australia in late 2005 and in the research we conducted for NAATI in 2007. Indeed, discussions at NAATI in-service workshops that I facilitated to introduce the new Manual and responses from some Panel chairs to RMIT survey questions indicate they would like similar error analysis / deduction guidance in assessing interpreting tests. (Turner and Ozolins, 2007: 28)

#### 6. ATTEMPTS TO DEVELOP A HOLISTIC SYSTEM FOR ASSESSING NAATI INTERPRETER TESTS

After the release of the 2005 NAATI Examiners' Manual, David Deck (a co-developer of the 2005 NAATI Manual) and I turned our attention towards providing a more systematic and rigorous

approach to marking interpreter tests that would incorporate both a holistic approach and a method of deducting points for errors involving accuracy. We concluded that a 'holistic' impression of a student's performance was most suited to evaluating 'performance' aspects, such as fluency and naturalness of delivery and appropriate choice of register, and also to 'quality of language' aspects, such as grammar, pronunciation, idiom, etc. On the other hand, we argued that errors involving accuracy needed to be assessed according to some sort of objective scale, as in marking translating tests. We also proposed that points for performance / quality of language be awarded holistically on the basis of a set of descriptors.

In a sense, we were proposing something similar to Waddington's Method D, where Method B and Method C are combined with Method B accounting for 70% of the total mark and Method C for the remaining 30% (Waddington, 2001: 315).

The examiners who attended subsequent NAATI workshops found it difficult to decide what percentage of points should be allocated to the holistic component and to mistakes affecting accuracy. For example, some examiners felt that 70% of the total could be allocated to the error analysis component and 30% to the holistic component, others argued that a more appropriate allocation was 50% to each component, while some felt that both should be allocated 100 points and that a final mark be determined by combining (merging) the two.

## **7. OTHER ACCREDITATION ORGANISATIONS THAT EMPLOY ERROR ANALYSIS / DEDUCTION**

In our research for NAATI we found that the American Translators Association (ATA) was using negative (error) deductions for its translation tests, but was moving towards a "holistic (functional) approach with the help of a rubric" that is "still very much a work in progress". The ATA was using "a point-based marking system for translation tests in which the grader assigned 1,2,4,8 or 16 points for each error." A passage with eighteen or more points deducted received a grade of fail (Stejskal, 2003: 96).

ATA examiners use a flow chart (Appendix A), a significant feature of which is its attempt to relate errors to how a prospective monolingual copy editor or typical target reader might be able to understand them.

The flow chart guides examiners in making error point decisions. It distinguishes between "mechanical errors" including "grammar, syntax, punctuation, spelling, diacritical marks/accents, upper/lower case, word form, and usage" and "errors that involve transfer of meaning" including "addition, omission, terminology/word choice, register, too freely translated, too literal, false cognate and ambiguity". (Turner and Ozolins, 2007: 44)

The ATA's system seems to be more punitive than NAATI's. To pass the ATA exam candidates have to score 82% or higher compared with NAATI's 70% and while NAATI's manual recommends deductions range from half a mark to five or more, the ATA recommends deductions of up to 16 points. Like NAATI's system of deducting errors, which was developed from opinions expressed at in-service workshops for examiners, the ATA scale:

... reflects experienced graders' judgements about the relationships among different types of errors, about what sorts of errors might be allowed in a translation that meets ATA standards. Criteria include whether a target-language reader would certainly recognise the error, and how serious the consequences would be in the context of the passage. One of the main advantages of the new point system now reflects five different levels of seriousness for each error, rather than the former system with two levels.

The ATA further advised:

By convention, we cap mechanical errors at four points (two points for spelling errors). Other conventions set a limit of eight points for reversal of meaning at the word level (left/right, open/closed, black/white) and errors of magnitude in numbers and units of

measurement (million/billion, dl/ml). The 16-point error applies to a mistranslation that seriously impairs the usefulness of the text as a whole. (Turner and Ozolins, 2007: 44)

We found that a number of other organisations that conduct translation exams only tolerate certain numbers of errors in specified categories. (Turner and Ozolins, 2007: 44, 45)

## **8. DESCRIPTOR-BASED SYSTEMS - THE DIPLOMA IN PUBLIC SERVICE INTERPRETING (DPSI) SYSTEM**

Like NAATI, the United Kingdom Diploma in Public Service Interpreting examination system has adopted different categories for its interpreting, sight interpreting, and translation tests. It divides its grading for interpreting tasks into 'accuracy', 'delivery', and 'language use'; categories for translation tasks are 'accuracy / appropriateness of translated text', 'cohesion, coherence and genre conventions', and 'effectiveness of communication'; and categories for sight translation are 'completeness', 'accuracy/appropriateness', and 'fluency/pronunciation' (UK Institute of Linguists, 2004).

However, unlike NAATI, DPSI candidates are assessed according to these categories mentioned earlier and within four bands of descriptors (A, B, C and D). Like the IELTS system for assessing language competence, these bands describe what test candidates can and cannot do. For example, a candidate who performed exceptionally well in "accuracy / appropriacy of translated text" in a DPSI written translation task would fulfill the following criteria (B and A):

The candidate

- conveys information with complete accuracy
- conveys all information without omissions, additions or distortions
- shows excellent use of vocabulary throughout
- uses excellent grammatical/syntactical constructions
- displays faultless spelling, accentuation/use of diacritics, faultless punctuation
- has excellent knowledge of specialised terminology with minimum paraphrasing

On the other hand, a candidate whose performance was barely satisfactory might fulfill the criteria set out in Band C:

The candidate

- displays some inaccuracies but no misunderstandings
- adequately conveys content
- makes minor omissions/distortions but not leading to misunderstandings
- makes adequate use of vocabulary
- uses adequate grammatical/syntactical constructions
- displays no major faults in spelling, accentuation, diacritics, punctuation
- has adequate knowledge of specialist terminology with paraphrasing kept to an acceptable level

In addition to these criteria, the DPSI system also uses numerical scores. Each band level is given a mark range. The mark range for Band D is 1-3, Band C is 4-6, Band B is 7-9, and Band A is 10-12.

A candidate who is assessed as fulfilling the criteria of Band C in the three categories (accuracy / appropriacy of translated text; cohesion, coherence and genre conventions; and effectiveness

of communication) might score 5 for the first category, 4 for the second, and 6 for the third, giving a total score of 15 for the test. The DPSI system requires candidates to achieve a minimum of 12 points with no fewer than 4 points per category to pass this translation test, meaning that this candidate would achieve a Pass grade.

## **9. A DESCRIPTOR SYSTEM SUPPLEMENTED WITH RIGOROUS ERROR IDENTIFICATION**

The RMIT report to NAATI suggested that the Authority pay particular attention to the Norwegian system whereby both descriptors and scales of deduction are used in interpreting tests and the UK Diploma of Public Service Interpreting tests where examiners rely exclusively on descriptors. We recommended that NAATI follow up ATA initiatives to introduce a descriptor-based system for its translation tests, and explore ways in which descriptors might be introduced into translator test marking. We also recommended that if NAATI introduced descriptor-based assessment, it should ensure that examiners are thoroughly trained to rate candidates against these descriptors.

I have since given more consideration to this and I now suggest the following approaches to testing are worthy of consideration in designing an improved way forward for translator and interpreter test assessment in Australia.

### **9.1. Holistic Guidance**

As discussed earlier, the 2005 edition of the NAATI Examiners' Manual states that while relatively minor transgressions involving factual accuracy might be tolerated examiners should view more seriously a tendency to make a number of errors. This edition of the Manual also advises that while some "language quality" errors involving inappropriate choices of grammar, idiom, collocations and spelling might be acceptable, significant and prolonged divergence should incur significant penalties. I believe that this guidance is impoverished and incomplete if compared with that given to examiners in descriptor-based assessment systems, such as the United Kingdom Diploma in Public Service Interpreting (DPSI) examinations, where examiners can choose from a scaled set of descriptors when grading test candidates.

I have sometimes felt the need for such descriptors at RMIT, where I have concluded that some examiners become too fixated on classifying errors and deducting points in accordance with the NAATI Manual, to the extent that it impairs their ability to arrive at a holistic appreciation of the translation they are assessing. In some cases, a tally of the points deducted has led to students being assessed at accreditation level, whereas it is obvious to a reader who takes a more holistic view that the test is in fact below the required level. On the other hand, I have also seen cases where too many points appear to have been deducted in a mechanistic way and the student was assessed to have failed while the general standard of the translation was actually at the accreditation standard.

### **9.2. Detailed and Consistent Feedback to Test Candidates**

From an educator's point of view, one of the greatest problems with the current NAATI system is that NAATI translation test candidates do not have descriptors to refer to in order to identify the proficiency they have demonstrated and the flaws in their performance in translating or interpreting tests. Where two assessors of the same test disagree markedly, candidates or students tend to focus on the differences in percentage terms, rather than on differing perceptions of examiners regarding their proficiency.

Percentage scores are supported by notations on the students' test papers by examiners, but these are sometimes not adequately or clearly expressed. Moreover, the use of percentage scores in this way implies that a mechanistic process, rather than a qualitative one, has been employed by the examiners.

I have often felt that if NAATI examiners had descriptors to refer to, they would be able to give more appropriate and comprehensible reasons for their decisions. In particular, they would be able to comment in a standardised and comprehensible manner on the proficiencies that had or had not been adequately demonstrated by the candidate or student.

### **9.3 Research at RMIT**

While writing the paper for the International Federation of Translators (FIT) Congress (Turner, 2008) I asked two staff members at RMIT to trial the UK descriptor system in assessing NAATI-style translation tests involving the translation of two 250-word passages. One was a second marker of ten English-into-Chinese translation tests while the second was a third marker of ten English-into-Japanese translation tests. The other markers used the NAATI error deduction system.

The English into Chinese marker is a very experienced user of the NAATI system. Seven of her results agreed with the first marker (five failed and two passed). Of the remaining three, she failed two that the first examiner had passed and passed one that the first examiner had failed. She found the descriptors very useful, but employed the NAATI error annotation symbols to provide a “picture” or “map” of her assessment. She would have liked more detailed guidance on numbers and types of errors that might be tolerated at each level.

The English into Japanese marker is a relatively new staff member and less experienced in using the NAATI system. Five out of ten of her results were very close to those of the other two examiners (less than five points over or under the other results), three results were close for one passage and more than ten points different for the other, and there was a more than 10 point difference between her and the other examiners for both passages in two of the exams. While the other teachers “accredited” four out of ten students, she “accredited” two (two of the four “accredited” by the other two teachers). Importantly, while the other teachers gave the two “fail” students a bare pass, she barely failed them. This indicates that while similar results were achieved, the teacher using the descriptors failed two students on “holistic” grounds.

Like the English-into-Chinese examiner, this examiner used the NAATI error analysis symbols (to complement the descriptors) and found them very useful in reaching conclusions. She commented that she had trouble with the working of criterion statements such as “good” and “adequate” at different band levels, a difficulty that might be overcome with appropriate training in their use.

While this research was at a very preliminary level, I find it interesting that both examiners found it useful to employ a standardised set of symbols to identify and categorise errors in addition to the DPSI descriptors. They were concerned that they might not be able to examine the renderings of test candidates in detail unless they identified different classes and types of errors in this way. Of course, both were no doubt influenced by prior exposure to the NAATI system of error analysis/deduction.

Perhaps the most interesting aspect of this preliminary research is that, like Waddington, our assessors did not arrive at markedly different results regardless of which system they were using. The research at RMIT is ongoing.

### **9.4 Support for a Descriptor System**

Bearing in mind the possibility that there will be little difference in outcomes if a descriptor system is introduced by NAATI, I nevertheless support the introduction of descriptors by NAATI. A particular benefit is that it would bring much greater clarity to the assessment of interpreting tests, where relatively little guidance is now available, and enable examiners who lack a theoretical knowledge of translating and interpreting to provide constructive, coherent

and rational descriptions to interpreting and translation test candidates of what they have and have not achieved.

### **9.5. A Rigorous Descriptor System**

I also believe that the particular nature of assessing translation and interpreting tests described earlier (the need to ensure that the outcome is a faithful rendition of the original source text and the requirement for examiners to constantly refer back to the source text) I have concluded that an optimum system for assessing translation tests would incorporate both holistic guidance and a rigorous approach to identifying and categorizing errors.

I admire the American Translators Association's reader-centred approach to its translation test marking. I believe that questions such as whether the error would be apparent to a monolingual copy editor or a typical target reader, and whether the error would be disruptive to a target reader are very useful in reminding assessors that they have to certify whether or not a candidate will actually be able to operate proficiently in the real world of translation practice. I also admire the ATA's emphasis on the reader in determining the seriousness of errors involving accuracy. I support the provision by the ATA and some other organisations to tolerate only a certain numbers of errors in specified categories, rather than NAATI's system of deducting points each time an error occurs.

I am also cognisant of the support of NAATI examiners for the current error analysis system in examining translation tests, the similar pass rates in NAATI exams to those of other public accreditation testing organizations, and the research by Waddington that all assessment methods might produce similar results.

### **9.6. A Combined System**

My preference is for a translation and interpreting test assessment system that provides for:

- retention of error analysis but without provision for examiners to negatively deduct points for each error;
- a set of holistic descriptors that –
  - like the DPSI's descriptors, describe what a test candidate can and cannot do at each level of proficiency,
  - incorporate the guidance currently offered by the NAATI system for identifying, categorising and annotating (with standardized symbols) different types of errors, and
  - like the ATA and similar systems, provide for a degree of error analysis and guidance on how many errors of different types may be tolerated at each level;
- derivation of the final grades entirely from the scales of descriptors (obviating concerns about the percentage of points to be allocated to the holistic component and to mistakes affecting accuracy) with no provision for percentage scores.

## **10. NAATI ASSESSMENT IN THE FUTURE**

When Uldis Ozolins and I briefed the NAATI Board about the RMIT report in mid-2007, the Board undertook to give serious study to all of its recommendations. Obviously, moving a large public examination system away from practices that have been in place for some 30 years is not something that NAATI will undertake lightly and I do not expect immediate changes from the Authority.

However, I believe that the work that our team did in researching international organizations points to the need for more reflection and research into the assessment of translating and interpreting tests in

Australia so that we can benefit from the best translating and interpreting assessment systems throughout the world and develop a more mature and effective system for Australia.

## REFERENCES:

- Stejskal, J. (2003) *International Certification Study*, American Translators Association.
- Turner, B. *Descriptors and Negative Marking Systems: An International Comparison of Translator and Interpreter Certification Testing*, paper presented at 18<sup>th</sup> World Congress of the International Federation of Translators (FIT), Shanghai, August 2008.
- Turner, B. and Ozolins, U. (2007) *The Standards of Linguistic Competence in English and LOTE among NAATI accredited Interpreters and Translators*, RMIT University, Melbourne.
- UK Institute of Linguists The Diploma of Public Service Interpreting: Assessment Criterion Statements, [www.iol.org.uk](http://www.iol.org.uk), accessed March 2007
- Waddington, C. "Different Methods of Evaluating Student Translations: The Question of Validity", *Meta*, XLVI, 2, 2001

