



Rater reliability in the context of translator assessment

Helen Slatyer & Kehui Luo, Macquarie
University

Catherine Elder, University of
Melbourne

EALTA, Sitges June 15-17, 2007

Translator and interpreter testing in Australia

- The National Accreditation Authority for Translators and Interpreters (NAATI)
 - Set up in 1977 to create and implement standards for translators & interpreters in Australia
 - Tests translators & interpreters as the basis for accreditation & entry to the profession
 - Tests in 57 languages

Translator and interpreter testing

- Little published research
- Tests often developed on the basis of educational models of assessment
- Construct definitions lack empirical basis
- Methodologies gradually being adopted from educational measurement and language testing
- Challenges relating to
 - the number of languages tested
 - complexity of constructs
 - practical difficulties finding and training examiners in rare languages

[Research project]

- Aims:
 - to identify extent of rater agreement within language panels and across language panels
 - to investigate intra-rater reliability
 - to develop and trial rater training package

[Translator test format]

- 3 hour exam
- 1 direction (e.g. Fr>Eng)
- Choice of 3 passages in source language
- 2 passages ~250 wds to be translated into target language + 2 ethics questions
- Hand written translations
- Paper & electronic dictionaries & glossaries allowed

[The role of language panels]

- Each language specific panel is headed by a panel chair
- Panels develop test tasks and rate scripts
- Reference translation for evaluation of the text for inclusion in the ATP & to guide rating
- Panel members are translators and/or educators
- Panel members either LOTE or English dominant

[Rating process]

- Double independent rating
- Scripts assigned to raters by NAATI
- Rating criteria derived from theoretical definitions of translator competence
- Scoring based on points deduction for errors
- Rater training recommended but non-compulsory & unpaid

[Rater training]

- 2 rater training sessions
- Operational data from ATP

Melbourne	Sydney
Amharic 1	Arabic 3
Filipino 2	Bangla 5
French 2	Czech 1
Greek 2	French 1
Hindi 4	Hindi 2
Indonesian 1	Indonesian 1
Japanese 2	Japanese 1
Chinese 2	Korean 4
Persian 1	Chinese 2
Sinhalese 3	Portuguese 2
Spanish 1	Romanian 1
Turkish 1	Tamil 1
Vietnamese 1	
23	27

Preliminary findings from qualitative data analysis

- Rater discussions during training audio recorded & transcribed
- Content analysis of discussions in language specific groups of raters where possible

[Issues raised]

- Industrial issues
 - Payment
 - Training
 - Apprenticeships/mentoring of newly accredited translators
- Difficulty of interpreting guidelines
 - Definition of error (accuracy vs readability, fidelity)
 - Relationship between error type and points deducted (5, 2, 1, 0.5)
 - Award of discretionary 5 points
- Text selection
 - Should language distance influence text difficulty?
 - Text length based on ST or TT?

[Issues cont'd]

- Lack of uniform standards for translation quality (no standards definitions, language panels define standards)
 - Confusion about whether the test is identifying entry to the profession or fully functional translators
 - Translation style
- Translation challenges relative to language pair (acronyms/punctuation/culturally specific concepts)
 - greater use of passive in English than French
 - domain specific terminology not reflected in all languages
 - lack of syntactic equivalence in certain language pairs
 - different educational systems across cultures
- Authenticity of the test
 - Lack of contextualisation (e.g. audience definition)
 - Problems of ambiguity in ST
 - Test conditions don't allow for professional level translation (handwritten/lack of resources for research/time)

[What's next]

- Symposium at AILA, 2008
- Questions/suggestions email:
Helen.Slatyer@ling.mq.edu.au