What Was So Hard About That?
Test Errors and Source Passage Challenges

DIANE HOWARD

Universitat Rovira i Virgili, Spain

Descriptors of foreign-language reading ability, such as the U.S. Inter-agency Language Roundtable (ILR) reading levels and the Common European Framework of References for Languages (CEFR), are sometimes uses to select source passages for official and professional translation tests. However, little research has been done on whether these descriptors correspond to translation difficulties. This paper examines errors made in a small set of Japanese-to-English tests from the American Translators Association Certification Examination to determine whether the errors could be predicted from the reading level descriptors.

Keywords: CEFR, ILR, reading level, translation difficulty, translator errors, American Translators Association

Introduction

The short-passage translation test is a standard way of screening translators in the professional world. However, beyond the requirement that the material used for testing be authentic, little consensus exists on how to choose passages for such tests. An approach taken by some official and professional organizations is to use one of several existing systems for classifying foreign-language reading level difficulty (e.g. the Interagency Language Roundtable or the Common European Framework of Reference for Languages systems) as a way of selecting test passages. The assumption is that reading-level difficulty can be equated to translation difficulty, but this has yet to be demonstrated with analyses based on test data. In this paper, I will compare errors appearing in a set of Japanese-to-English tests from the 2004 American Translators Association (ATA) Certification Examination to determine whether the errors the candidates made were those that could be predicted from the reading level of the test passage. In other words, those taking the test will be indirectly asked: What was so hard about that passage?

For the analysis, I use the general (mandatory) passage from Japanese-to-English 2004 ATA certification examination (see the Appendix for the text and a possible translation) and six candidate translations. All translations were produced under approximately the same circumstances: In a proctored examination setting, candidates were required to produce handwritten translations of two passages of 325 to 400 Japanese characters (about 250

English words) within 3 hours. Paper references were allowed, but electronic resources and Internet access were prohibited. Candidates could not share references or consult other people (i.e., all test papers were the result of individual work). Information about the ATA Certification Examination can be found on the ATA website and will not be repeated here (ATA 2008). The passage will be evaluated in terms first of Interagency Language Roundtable (ILR) Skill Level Descriptions: Reading (ILR 2006), then in light of the Council of Europe Common European Framework of Reference for Languages reading descriptions (CEFR 2001).

Both the ILR and CEFR systems were designed to assess second-language proficiency and not the translation difficulties of a given text. However, ILR reading level is currently used to select texts for translation tests given by various agencies of the U.S. government and as been adopted as part of the passage selection guidelines for the ATA certification examination (ATA Graders 2008). Thus an assumption is being made by the examiners that reading-level difficulty somehow correlates with translation difficulty, an assumption that also underlies the recently developed ILR *Skill Level Descriptions for Translation Performance* (ILR 2006). To avoid confusion between reading difficulty as described in the ILR and CEFR systems and what might constitute a "translation difficulty", in the following discussion I will use the term "challenge" rather than "difficulty" when referring to translation. Difficulty is a notoriously relative concept. A challenge, such as jumping a one-meter fence, is there to be met or not. Challenges can also be said to separate those who can meet them from those who cannot, a concept that fits the screening purpose of professional examinations.

ILR Language Skill Level Descriptions: Reading

The ILR language skill level scales grew out of efforts by the U.S. State Department to develop tests to measure the language abilities of foreign-service officers beginning in the 1950s. The system measures the four language skills—speaking, reading, writing, and listening—on a scale (including plus values) of 0 to 5. Descriptions of translation skill levels were added in 2006 (ILR 2006).

Although the system was developed to measure language proficiency, several U.S. government agencies demand that entry-level candidates for positions requiring a second language pass a reading level 2 (minimum working proficiency) translation test and later an additional test at ILR reading levels 3 (general professional proficiency). Thus, despite the system having originally been designed to measure language proficiency, it has been used for a number of years to measure language performance (Child, Clifford, Lowe 1993 and Child 1998).

The description of reading levels has been further refined through the addition of text typologies (Child 1981) corresponding to the levels: orientation mode (level 1, with a one-to-one correspondence of language and content), instructive mode (level 2, straightforward information about the real world), evaluative mode (level 3, analysis and evaluation against "a backdrop of shared information"), and projective mode (level 4, extensive author input and shaping).

Superficially, these modes resemble the text types described by Reiss (2000) and the text functions set out by Nord (1997), but one should be careful about comparisons because the initial aim of a system to describe reading proficiency is obviously different from typologies created for translation or descriptive pragmatics. However, with that caveat in mind, the systems of Reiss and Nord can be used to critique Child's modes.

The "easiest" of Child's classification modes, "orientation," corresponds to Reiss's content-focused texts and Nord's referential function, but also to Nord's phatic function, which often demands a fair amount of transfer skill. Child (1981:100) considers phatic utterances easy to understand and/or produce because of their recurrent, *pro forma* nature, but instrumental translations of such phrases frequently require replacement with a phrase used in the same manner but with a different literal meaning or even omission when the target culture has no equivalent phatic category or the target situation does not require it.

In terms of author shaping and use of language, the evaluative and projective modes generally match Reiss's form-focused text, although Reiss only includes literary texts in her category. They are also fit well with Nord's expressive function, which insists on the sender orientation of such texts.

The mode of the general test passage under consideration is evaluative, with a point of view working toward a conclusion. The development of the topic is linear and organized in the form of problem–cause–result with supporting evidence. The sentence structure is reasonably complex, with insertion and qualification. The vocabulary is generally concrete, with several linking lexical chains. Viewed sociolinguistically, the passage does not create a wide cultural gap for this language pair, but some cultural background is helpful. The style is not highly individual or idiosyncratic, and the passage is well written.

Moving from mode to specific reading level, the passage is ILR reading level 3 (general professional proficiency). The text includes "hypothesis, argumentation and supported opinions" and requires some ability to "relate ideas and 'read between the lines' " (ILR 2006). While the structure of two of the sentences is somewhat complicated, both remain within the bounds of normal Japanese usage and do not reach the level 3+ requirement of "intentionally complex structures" (ILR 2006). One somewhat unusual term (駆り立ててきた) appears and two clichés (一挙に and 言うまでもない),

but no idioms and certainly none of the "low frequency idioms" of level 3+. Some cultural knowledge would be useful in analyzing the passage, but no deep knowledge is required.

If we transfer these descriptions of reading text level into predicted translation challenges, we can assume that candidates will need to follow the following argument:

- Up until the time of Masao Maruyama, the level of research on Japanese politics by Japanese scholars had not been high.
- Maruyama changed that situation, but not completely.
- Only recently have Japanese scholars produced genuine studies of contemporary Japanese politics.

The markers that candidates have reproduced the argument could include the following:

- correct use of English tenses to indicate the time sequence;
- demonstration that the second sentence is the explanation for the situation set out in the first sentence;
- possibly replacement of これまで with "up to the time of Masao Murayama" or some other indication of a definite time.

Candidates might need some cultural background about Japanese views concerning foreign scholarship and the emperor system during the mid-twentieth century to sort out the second sentence, although I suspect that good grammatical and syntactic knowledge could compensate for lacunae here.

The clichés and one unusual phrase would appear to present not challenges but more of an opportunity to show transfer and target-language writing skills.

CEFR

Although I do not know of any organization currently using the CEFR for selection of translation test passages, it is a logical system to consider both because of the range of languages it covers and because of the extensive research being done to validate the system. While the ILR scale was developed through intuition and experience, the CEFR was created using more statistically rigorous methods (2001: 217-225), suggesting that it might provide better descriptors of difficulty than the ILR system or at least better validated ones.

In the levels set for Overall Reading Comprehension (2001: 69), level B1 essentially corresponds to ILR level 2: "Can read straightforward factual texts on subjects related to his/her field and interest with a satisfactory [not

defined] level of comprehension." Levels C1 and C2 are also compatible with, although not completely identical to, ILR levels 3 and 4.

The CEFR descriptors are further broken down into Reading Correspondence, Reading for Information and Argument, and Reading Instructions (2001: 69-71), a more nuanced approach than that of the ILR (Reading for Translation would be a welcome addition). If we consider the Japanese-to-English test passage in terms of both Overall Reading Comprehension and Reading for Information and Argument, it falls somewhere between levels B1 and C1. Because this system is entirely proficiency based (the CEFR descriptors apply to the language learner and not the text) and not a proficiency/performance hybrid like the ILR descriptors, it is difficult to extract the amount of detail about a text that one can achieve with the ILR levels, but a text at the B1/C1 level should have the following characteristics: complexity (not defined), possibility unfamiliar subject matter, a limited number of low frequency idioms, and both stated and implied opinion. These elements are consistent enough with the ILR system that we can say they predict the same translation challenges. Also, one can argue that the two systems validate each other to some extent, so that overlapping categories of description should be reasonably reliable.

Candidate Performance

The only information available about the candidates is that they met the ATA eligibility requirements for taking the certification examination. These include certification from another member of the Fédération Internationale des Traducteurs, a degree or certificate in translation and/or interpreting, high school or college graduation with a specified amount of translation or interpreting experience, or an advanced degree in any field with no translating experience required (ATA 2008). Information about whether people were working into their A or B language is not available.

As scored by the ATA grading system, one exam passed (fewer than 17 error points) (A), three were in the 18 to 25 point range (B-D), and two failed by more than 45 error points (E, F).

Responses to predicted translation challenges

A primary feature of ILR reading level 3 is the ability to follow a supported argument. All of the candidates except F appear to have done this. However, three candidates (A, B, and C) did not use English tenses correctly to indicate the time sequence in the passage. All three used the present perfect extensively, and C switched to the present when introducing the information about Maruyama's achievement. Unexpectedly, the two candidates who failed by the largest margin did the best with temporal cohesion. Candidates C and F did not present the second sentence as the reason for the situation in

the first sentence. Candidate D used the dummy subject "it was" as an inappropriate transition between the two sentences ("It was because excellent political scientists…").

All of the candidates omitted これまで, the time marker in the first sentence, which may have accounted for the subsequent tense problems. In 2004 the ATA examination presented no information about the source text or context for the translation. Happily, this situation has been corrected, but one wonders whether having Maruyama's dates (1914-1996) might have helped people produce more cohesive translations.

None of the candidates had trouble with the reference to the emperor system, and A, B, and C made the target language appropriate addition of "of Japan" or "Japanese." The phrase 学問の源泉が外国にあるという古来の学問的伝統 (the long-standing scholarly tradition that the fountainhead of scholarship was overseas) was more of a problem. The translation B gave was "the restraint placed on this [no antecedent] scholarly tradition, which has foreign origins, by the Japanese Emperor system". The text given in D read, "the old academic tradition that found sources of academic studies in foreign countries". E had "the traditional belief that any scholarly source should be found in foreign countries".

All of the candidates rendered cliché for cliché, which was certainly adequate for the purpose at hand. 駆り立ててきた invoked five different translations, all correct, and one omission (B). Campbell and Hale have suggested that different translations of a source text item is an indicator of textual difficulty (Campbell 1999; Campbell and Hale 1999), but later conclude that, "if subjects are faced with multiple choices this does not necessarily mean that the item in question is difficult" (Hale and Campbell 2002: 29). That the test takers were presented with a term which did not have an immediately obvious English equivalent does count as a challenge in our sense of an obstacle to be overcome—finding a suitable target term may have slowed them down and required additional cognitive processing—but the fact that this challenge was successfully met confirms Hale and Campbell's finding that a term without an immediately obvious target equivalent cannot be universally defined as a translation difficulty.

Responses to unpredicted translation challenges

In discussing why the test passage was at ILR reading level 3 and not 3+, we noted that the text contained two somewhat complicated sentences (sentences 2 and 4) that we did not rate as "intentionally complex". However, three candidates (B, E, and F) had difficulty sorting out the connections in sentence 2 and two candidates (E and F), the links in sentence 4. This could indicate that the sentences were more complex than we thought or that sentences with the degree of modification and insertion seen in the test

passage are appropriate for a general profession level test. Based on this tiny sample, as well as other experience, I am inclined toward the latter explanation. The candidates who had trouble with both sentences were also those who failed by the most points and so the sentences with this level of complexity can be seen as proper screening challenges. Therefore, in passage selection for screening test, fairly complicated complex and compound sentences should be noted as translation challenges and some effort should be made to find texts that include such sentences.

The term 思想史 (history of ideas/thought) was an unexpected challenge. Candidate C translated it as "history of ideology" while E opted for "philosophical history". Ideally, terminology as a test challenge should be linked to research skills, an area of professional competence that most organizations would want to evaluate. On an examination restricted, as this one was, to paper sources only, recognition of terminology becomes more a test of general (or perhaps specific) knowledge, which may not be relevant to evaluating a translator's professional skills.

Conclusion

While no real conclusions can be drawn from such a small sample, this very preliminary study suggests that systems for classifying reading difficulty of texts for foreign language learners, and the ILR system in particular, are useful in identifying challenges in passages for translation tests. The challenges identified by the criteria for ILR reading level 3 do seem to be the items that the test takers found difficult, in that they generally failed to follow the argument of the passage. Beyond the obvious advantages of providing criteria to select passages of equivalent difficulty across languages and over time, the salient point is that the system provides a means of identifying the challenges a passage presents. Such identification should guide not only passage selection, but also passage evaluation. Error marking in the absence of predetermined challenges often becomes a line-by-line search for grammatical and equivalence mistakes. Starting with passage-level challenges, such as identifying links in an argument with their associated cohesion patterns, should help focus evaluators' attention on the translated passage as a whole and to determine what they are marking before they pick up their red pens.

Another advantage of connecting a reading level system to selection of translation passages is that the process of matching the descriptors to the text forces one to read in a different manner, which could be the start of "reading for translation". This, as Juan Sager has noted, is a different process than other types of reading (1994: 111-113). Using reading levels to focus student attention on potential translation challenges would give identification of reading levels pedagogical, as well as testing, utility.

Appendix: Japanese-to-English General Passage from the 2004 ATA Certification Examination with Possible Translation

日本人研究者による日本政治研究の水準は、これまで、少数の優れた研究があったとはいえ、全体として高くはなかった。学問の源泉が外国にあるという古来の学問的伝統に、天皇制による束縛、さらには、政治学者の近代化・民主化への関心が重なって、優れた政治学者を政治的先進国の政治の研究に駆り立ててきたからである。日本政治の研究はこれら学者の余技か、政治評論・文明批評として行われることが多かった。丸山真男の業績は、日本政治の研究水準を一挙に高めるとともに、才能のある政治学者の目を日本に向けることに貢献したことは言うまでもないが、いわゆる丸山学派の人たちの分析対象は主として思想史であって、現代日本政治の科学的研究に及ぶことが少なかった。日本政治の研究のレベルが一般的に高まり、日本政治専攻者の数が増えるようになったのは、ごく最近のことである。最近話題になった政治学書の多くは、現代日本政治を対象とした本格的研究である。

Apart from a few excellent studies, up to the time of Masao Maruyama [これまで "this time" literally] the general level of research on Japanese politics by Japanese scholars was rather low. This was because restrictions resulting from the emperor system and the interest of political scientists in modernization and democracy, coupled with the long-standing scholarly tradition that the fountainhead of scholarship was overseas spurred [駆り立ててきた] the best scholars to study the politics of politically advanced countries. For the most part, these scholars studied Japanese politics as a hobby or in the form of political or cultural commentary. Needless to say, the achievement of Maruyama was that he raised the level of research on Japanese politics virtually overnight and also helped direct the attention of talented political scientists to Japan. However, the members of the so-called Maruyama school primarily focused on analysis of the history of ideas [思想史] and rarely produced scientific studies of contemporary Japanese politics. The overall improvement in the level of research on Japanese politics and the increased number of scholars specializing in that field is an extremely recent development. Many of the writings on political science that are current topics of discussion are genuine studies of contemporary Japanese politics.

References

American Translators Association. 2008. *ATA Certification Program*. http://www.atanet.org/certification/index.php/. Entry point for information about the certification examination. Visited February 2008.

ATA Graders. *ATA Graders*. Visited February 2008. http://tech.groups.yahoo.com/group/ATA_Graders/. Password protected Web site.

Campbell, Stuart. "A cognitive approach to source text difficulty in translation". *Target* 11(1): 33-63.

Campbell, Stuart and Sandra Hale. 1999. "What Makes a Text Difficult to Translate?" *Proceedings of the 1998 ALAA Congress*. http://www.latrobe.edu.au/alaa/proceed/camphale.html/. Visited February 2008.

Child, James. 1981. "Language Proficiency Levels and the Typology of Texts". Unpublished paper, given at a pre-conference workshop on language testing sponsored by the Georgetown University Language Roundtable on Languages and Linguistics.

Child, James R. 1998. "Language Skill Levels, Textual Modes, and the Rating Process". *Foreign Language Annuals* 31(3): 381-391.

Child, James R., Ray T. Clifford, Pardee Lowe, Jr. 1993. "Proficiency and Performance in Language Testing". *Applied Language Learning* 4(1/2): 19-54.

Council of Europe. 2001. *Common European Framework of Reference for Languages*. Cambridge: Cambridge University Press. http://www.coe.int/T/DG4/ Linguistic/Source/Framework_EN.pdf/. Visited February 2008.

Hale, Sandra and Stuart Campbell. 2002. "The interaction between text difficulty and translation accuracy". *Babel* 48(1): 14-33.

Interagency Language Roundtable. 2006. *Skill Level Descriptions. Introduction*. http://www.govtilr.org/ILRscale1.htm/. Entry point for information about the ILR skills levels. Visited November 2008.

Nord, Christiane. 1997. *Translating as a Purposeful Activity: Functionalist Approaches Explained*. Manchester: St. Jerome.

Reiss, Katharina. 2000 (1971/1984). *Translation Criticism—The Potentials & Limitations: Categories and Criteria for Translation Quality Assessment*. Manchester: St. Jerome.

Sager, Juan C. 1994. *Language Engineering and Translation. Consequences of Automation*. Amsterdam and Philadelphia: Benjamins.