

翻译能力评分量表的设计与开发^{*}

吕晓轩

(黑龙江大学 哈尔滨 150080)

提 要: 本文从测试学视角出发,以翻译测试设计原则为出发点,阐释现有翻译测试评分方法的分类以及评价量表的重要性。基于对翻译能力的回顾,并借鉴 Angelelli 的翻译能力构念框架,构建翻译能力评分量表,以期能够更好地测量译者的翻译能力。

关键词: 翻译测试; 评分量表; 构念; 翻译能力

中图分类号: H315.9

文献标识码: A

文章编号: 1000-0100(2016)03-0137-4

DOI 编码: 10.16263/j.cnki.23-1071/h.2016.03.028

Designing and Developing Scoring Rubrics for Translation Competence

Lv Xiao-xuan

(Heilongjiang University, Harbin 150080, China)

This paper focuses on language testing, discusses the principles of translation assessment and elaborates the classification of scoring methods and the significance of scoring rubrics. With reference to Angelelli's construct framework of translation competence, it provides scoring rubrics for translation competence in order to better assess translators' competence.

Key words: translation assessment; scoring rubrics; construct; translation competence

1 引言

随着我国“一带一路”战略思想的提出,翻译在推动我国政治、经济、文化和科技走向世界的进程中发挥越来越重要的作用。我国经济的快速发展对翻译人才培养提出更高的要求,并引发对翻译能力培养和翻译测试等问题的深入探讨。翻译测试是用来检验被测对象翻译实践能力情况的一种手段,涉及评判答题内容,因此与翻译评估或评价、语言测试等有关。尽管“语言测试的发展在近三十年取得巨大的成就,但很少触及翻译测试”(肖维青 2010: 171)。究其原因,翻译测试具有较强的主观性,如何保障评分信度成为其发展的阻碍。因此,只有选择恰当的评分标准,并且由经过训练的评分员持续性地使用相关评分量表是保证行为测试信度和效度的关键因素。

2 翻译测试设计原则

翻译测试属于行为测试的一种,测试设计者要遵循

以下 3 个原则:一致性、信度和效度。一致性是指测试设计程序上的标准化,即设计、管理和评分要求测试设计者和评分员能够遵循相同的规范和标准,并且对测试有相同的预期。同时,一致性也体现在评分员对评分量表的解读方面。

“翻译测试信度的影响因素包括试题的质量、考试实施和阅卷三个方面。”(李筱菊 2001: 35) 保证翻译测试的信度,须要对上述 3 个要素实施严格管理和规范。由于翻译测试的主观性特征,其信度更大程度上取决于评分是否客观,因此评分标准的规范性和评分员的素质非常关键。

测试效度是“针对经验证据和理论论证对测试分数或其他评估结果所做出推断采取行动的充分性与合适性的支持程度的综合评价性判断”(Messick 1989: 13)。翻译测试的效度是对翻译测试进行效验,根据已有效度理论框架“对翻译测试结果的使用、解释以及根据该结果所

^{*} 本文系国家社科基金项目“基于体裁的商务英语话语研究”(12BYY128)、黑龙江省社科基金项目“翻译测试评分体系构建”(WY2015108-C)和黑龙江大学青年基金项目“多模态视阈下商务英语读写能力评价体系构建”(QW201321)的阶段性成果。

可能做的推断或决策提供一些可供参考的理论论证和经验证据,进而促进翻译测试的改进,并且提供有价值的反馈信息”(邹申 2005:403)。

3 翻译评分方法与评价量表

翻译测试的评分方法包括机械法、整体评分法和分项评分法 3 种类型。其中,机械法又称计算错误法,也称为倒扣法,即按照译文中的错误数目扣分。有时,评分员还会根据误译的严重性或者误译的性质扣分。整体评分法也称印象评分法,指评分员根据对译文的总体印象打分的方法。其优点是评分效度高,评分迅速。但评分主观性大,影响评阅的因素比较多。分析评分法把译文分解成各种构成要素,并根据各要素量表对每个要素打分,最后合成一个总分。其中,整体评分法和分项评分法的实施都是评分员参照评分量表进行评阅。

“评分量表是一种描述性教学评价方法,是一种评分工具,能够针对学习成果表现列出相应的标准,并详细描述从优到差的每一个标准的水平。”(Goodrich 1997: 14-17)“评分量表是由教师或评价者开发出的一种描述性的评分量表,其目的是为分析学生学习结果,包括学习作品和学习过程。”(Moskal 2000: 71-81)因为,评分量表不但能够像教学法一样能明确地指导学生有效学习,而且能提供评价标准去评估学生学习程度好坏及掌握知识的多少。借鉴评分量表,学生就能明确翻译教学的要求,并以此作为翻译学习和自评的工具;当学生把评分量表内化为翻译学习的标准和目标时,就能更好地促进和完善翻译学习,从而促进翻译能力的提高。

4 翻译测试构念

4.1 构念的界定

构念是对可观测的行为进行有意义地阐释(Messick 1989: 13)。21 世纪以来的测试观并未将构念看成通过测试任务考查某种能力的潜在特征。构念属于多个要素的结合,包括特征、完成任务的语境、评价任务表现行为的标准以及对所得分数的解释(Weir 2013: 3)。因此,如果测试翻译能力,首先应该明确界定测量的内容。不仅要明确被测量的能力、知识或行为,还要将其分解为具体的要素。因此,为测量译者的专业能力,须要首先界定构成该专业能力的具体技能和次技能。翻译能力构念的确立对翻译考试内容、考试方式和分数意义的解释息息相关,是效度研究的基础,是翻译考试设计和翻译考试评价的重要方面。

4.2 翻译能力研究综述

学者们分别从不同视角对翻译能力进行过诠释(Wills 1976; Toury 1986; Hatim, Mason 1997; Schäffner 2000; Neubert 2000; Colina 2003)。一些学者倾向于认为

翻译能力由一系列子能力组成,但是对这些子能力的界定又存在巨大差异。即便是子能力的名称相同,表达的意义也千差万别。

Wills(1976)指出翻译能力由接受能力、生产能力和超能力组成,即在源语、目标语的语言和语篇系统间转换信息的能力,这是最重要的构成成分。翻译中的转换(transfer)不可或缺,它随翻译能力的发展而发生变化(Toury 1986: 72-94)。Bell(1991)认为翻译能力是多种能力的总和,即包括目标语知识、文本类型知识、源语知识、客观世界知识、对比知识以及解码和编码技能的一种交际能力。Cao(1996)指出翻译的语言能力包括语法能力、篇章能力、言外能力和社会语言学能力 4 个子能力。就原文和译文的文本转换而言,译者能力包括原文处理能力、转换能力和译文处理能力,其重心在于协调源语语篇和译语语篇的交际特征(Hatim, Mason 1997: 205)。也有学者将转换能力作为一种与语言能力、文化能力、语域能力和研究能力并存的能力(Schäffner 2000: 147)。Neubert(2000)认为,转换能力建立在语言能力、语篇能力、主题能力和文化能力之上,具有统一协调的功能。Colina(2003)将翻译视为一种交际活动,强调交际的互动性。因此,除能够译出符合译文语境和任务要求的译文外,与原文及其语境、翻译任务要求、翻译过程的参与者互动的能力成为重要的子能力之一,此外还包括母语、外语的交际能力以及跨语言和跨文化交际能力。尽管学者对翻译能力的诠释不尽相同,但是在翻译能力可以被分解这一观点上具有一致性。但这些翻译能力的界定并不是从测试的角度出发,所以如何应用到翻译测试中还有待商榷。

4.3 翻译能力构念的界定

翻译测试的构念首先要回答什么是翻译能力,而翻译能力如何定义和分解一直是翻译研究者关注的焦点之一。构念的界定需要具备可操作性,“能力的界定必须具体、清晰,从而指导翻译能力测试和开发”(Fulcher 2003: 44)。因此,提出一个全面、系统、科学的翻译能力构念对于翻译测试十分必要,但是也极为复杂。翻译能力构念框架的提出旨在为翻译测试研究和发展提供导向功能。结合交际能力要素以及众多翻译能力研究的成果,Angelelli(2009)将翻译能力分解为语言能力(linguistic-level competence)、文本能力(textual competence)、语用能力(pragmatic competence)及策略能力(strategic competence) 4 个子能力。从狭义的角度而言,“语言层面能力特指语法能力,包括译者对词汇、形态、句法、字形方面知识的掌握和运用”(Cao 1996: 325-340)。文本能力指译者对修辞和衔接连贯的规则和约定的理解和运用能力。语用能力包括言外能力和社会语言能力,指译者为实现翻译目标有效使用语言的能力和在语境中理解语言的能力。策略能力指译者在翻译过程中有效识别和解决翻译问题的

能力。能力构念的提出为进一步制定评分标准提供理论依据。

5 评分方法与评分量表设计

评分方法确保评分员趋向一致,既包括评分结果,也指他们自身判断的一致性,那么,制定恰当的评分量表以及基于评分量表对评分者的定标显得尤为重要(Weir 2005: 17)。“为了客观准确地评估翻译能力,拟采用整体评分法和分析评分法相结合的评分方法,由两位评分员分别根据整体印象评分量表和分项评分量表评分,然后将二者的分数合成一个总分数。”(严明 2012: 27-28)

“翻译评分量表包括 3 个要素:一系列评价标准,即决定表现任务质量的各个指标;等级水平,即描述评价标准在质量上从差到好的序列;详细说明,即对各个标准在不同水平上的表现进行详细描述。”(Goodrich 1997: 14-17)

整体评分量表关注“原文的理解能力”和“译文的表达能力”,评分量表的等级水平设定为 5 个级别。5 级代表翻译质量最高,即最高级,而 1 级代表翻译质量最差,也是最低级。从 5 级至 1 级的详细描述如下:(1) 译文充分体现对原文的主旨和次要主题的详尽理解和细微之处的辨析,并再现原文的表达方式。译文娴熟地再现原文的意思;(2) 译文体现对原文的重要主旨和次要主题的充分理解,并再现原文的表达方式。译文熟练地再现原文的意思;(3) 译文能体现对原文的重要主旨和次要主题的基本理解,并再现原文的表达方式。译文恰当地再现原文的整体意思,但是偶尔存在一些错误;(4) 译文体现出对原文的重要主旨信息或一些次要主题或表达方式的理理解存在偏差,译文多次出现对原文的错误理解;(5) 译文体现出对原文意义的严重误解。(Angelelli 2006: 41)

通过分析文本能力、语用能力、语言能力和策略能力的典型特征,进而确立“风格与衔接”、“情境适切性”、“语法与技巧”和“翻译技巧”4 个维度作为分项评分量表的 4 个指标(Angelelli 2009: 40-42)。其中,“风格与衔接”是通过译文观测考生对体裁知识的掌握程度及衔接连贯知识的运用能力。“情境适切性”是考察考生对目标语的功能和社会文化方面掌握和运用的能力。“语法与技巧”是考察考生对目标语书写和语法方面掌握的水平。“翻译技巧”主要考察考生运用翻译策略或翻译资源解决翻译问题的能力。分项评分量表的等级水平与整体评分量表一致。4 个维度在 5 个级别上的详细表述如下:

风格与衔接:(1) 译文能够充分实现源语文本与目标语文本章节或段落组织的一致性。译文风格娴熟。译文流畅无瑕疵,浑然一体。(2) 译文能够实现源语文本与目标语文本章节或段落组织的一致性。译文具有一定的风格。译文语篇流畅、结构连贯。(3) 译文能够基本实现源语文本与目标语文本章节或段落组织的一致性。译文风

格可能不太一致。偶尔会出现不符合目标语习惯的结构或内容。(4) 译文呈现出源语文本与目标语文本章节或段落组织方面的一致性。译文风格拙劣。译文不流畅,频繁出现不符合目标语习惯的结构或内容。(5) 译文组织混乱。译文缺少风格。译文极不流畅,且内容不符合目标语表达习惯。句子和思想不相关。

情境适切性:(1) 译文能够充分考虑目标语受众,并在目标语中实现翻译预期的目的。词汇选择具有技巧且贴切。文化偏好、话语及语域特征符合目标语的文本特征及目标语受众。(2) 译文能够考虑目标语受众,并在目标语中实现翻译预期的目的。词汇选择恰当。文化偏好、话语及语域特征符合目标语的文本特征及目标语受众。(3) 译文基本能够考虑目标语受众,并在目标语中实现翻译预期的目的。词汇选择恰当。文化偏好、话语及语域特征基本符合目标语领域。但是对于目标语、文本类型和目标语受众而言,遣词造句过于正式或口语化。(4) 译文不能考虑目标语受众,也很难在目标语中实现翻译预期的目的。词汇选择恰当。文化偏好、话语及语域特征有时不符合目标语领域。对于目标语、文本类型和目标语受众而言,大量的遣词造句太过于正式或口语化。(5) 译文不能考虑目标语受众,也未能在目标语中实现翻译预期的目的。词汇选择恰当。文化偏好、话语及语域特征始终不符合目标语领域。对于目标语、文本类型和目标语受众而言,几乎全部的遣词造句都太过于正式或口语化。

语法与技巧:(1) 译文充分体现译者精通目标语语法、拼写和标点,几乎没有错误。(2) 译文体现译者熟悉目标语语法、拼写和标点,偶尔有小错误。(3) 译文体现译者不擅于使用目标语语法、拼写和标点,频繁出现小错误。(4) 译文体现译者基本不会使用目标语语法、拼写和标点,大量的错误对译文产生影响。(5) 译文体现译者不能使用目标语语法、拼写和标点,频繁出现严重的错误。

翻译技巧:(1) 译文体现出译者创造性解决翻译问题的能力,以及熟练运用资源和材料的能力。(2) 译文体现出译者解决翻译问题的能力,没有显著错误且很少有小错误出现,运用资源和材料时未出现明显错误。(3) 译文基本体现出译者识别和解决翻译问题的能力。但显著错误或累积过多的小错误影响译文整体质量。未能恰当使用资源或材料。(4) 译文体现出译者在识别和解决翻译问题方面存在困难。几处显著错误或大量的小错误严重影响译文整体质量。未能恰当使用资源或材料。(5) 译文基本体现出译者不能识别和解决翻译问题。大量的显著错误或小错误导致译文存在严重问题。始终未能恰当使用资源或材料。(Angelelli 2006: 41-67)

将整体评分法与分析评分法相结合的方法称为综合法,它集合两种评分方法的优点,具体表现为:第一,既能

将翻译能力作为整体进行评价,同时也能够判断考生翻译能力的各个子能力的具体情况。第二,分析法还能够指导评分者的评分过程(王克非 杨志红 2010: 58-59)。整体评分法和分项评分量表规范化的设计是实现客观评分的第一步,翻译测试的信度还取决于评分员能否正确使用量表,以及评分的一致性。

本文借鉴 Angelelli(2006)的翻译能力构念模式构建评分量表,并制定评分方法,是一项探索性研究成果,而这种评分方式在实践中的可行性、量表的改进与完善等问题还须通过进一步的实证研究来验证。

参考文献

- 李筱菊. 语言测试科学与艺术[M]. 长沙: 湖南教育出版社, 2001.
- 王克非 杨志红. 翻译测试中的理论与实践问题[J]. 外国语, 2010(6).
- 肖维青. 翻译测试的信度研究[J]. 广西民族大学学报, 2010(2).
- 严明. 商务英语能力构念界定和测试[J]. 当代外语研究, 2012(2).
- 邹申. 语言测试[M]. 上海: 上海外语教育出版社, 2005.
- Angelelli, C. V. Minding the Gaps: New Directions in Interpreting Studies[J]. *Translation and Interpreting Studies*, 2006(1).
- Angelelli, C. V. Using a Rubric to Assess Translation Ability: Defining the Construct[A]. In: Angelelli, C. V. (Ed.) *Testing and Assessment in Translation and Interpreting Studies: A Call for Dialogue between Research and Practice*[C]. Amsterdam: John Benjamins, 2009.
- Bell, R. *Translation and Translating: Theory and Practice*[M]. London: Longman, 1991.
- Cao, D. A Model of Translation Proficiency[J]. *Target*, 1996(2).
- Colina, S. *Translation Teaching: From Research to the Classroom: A Handbook for Teachers*[M]. Boston: McGraw-Hill, 2003.
- Fulcher, G. *Testing Second Language Speaking*[M]. Harlow: Pearson Education Limited, 2003.
- Goodrich, H. Understanding Rubrics[J]. *Educational Leadership*, 1997(4).
- Hatim, B., Mason, I. *The Translator as Communicator*[M]. London: Routledge, 1997.
- McNamara, T. F. *Measuring Second Language Performance*[M]. London: Longman Limited, 1996.
- Messick, S. Validity[A]. In: Linn, R. L. (Ed.) *Educational Measurement*[C]. New York: Macmillan, 1989.
- Moskal, B. M. Scoring Rubrics: What, When and How?[J]. *Practical Assessment Research & Evaluation*, 2000(3).
- Neubert, A. Competence in Language, in Languages, and in Translation[A]. In: Schäffner, C., Adab, B. (Eds.) *Developing Translation Competence*[C]. Amsterdam: John Benjamins, 2000.
- Schäffner, C. Running before Walking? Designing a Translation Programme at Undergraduate Level[A]. In: Schäffner, C., Adab, B. (Eds.) *Developing Translation Competence*[C]. Amsterdam: John Benjamins, 2000.
- Toury, G. Monitoring Discourse Transfer: A Test-case for Developmental Model of Translation[A]. In: House, J., Blum-Kulka, S. (Eds.) *Interlingual and Intercultural Communication: Discourse and Cognition in Translation and Second Language Acquisition Studies*[C]. Tübingen: Gunter Narr Verlag, 1986.
- Weir, C. J. *Language Testing and Validation: An Evidence-based Approach*[M]. New York: Palgrave Macmillan, 2005.
- Weir, C. J. An Overview of the Influences on English Language Testing in the United Kingdom 1913-2012[A]. In: Weir, C. J., Vidakovic, I., Galaczi E. (Eds.) *Measured Constructs: A History of the Constructs Underlying Cambridge English Language (ESOL) Examinations 1913-2012*[C]. Cambridge: Cambridge University Press, 2013.
- Wills, W. Perspectives and Limitations of a Didactic Framework for the Teaching of Translation[A]. In: Brislin, R. W. (Ed.) *Translation*[C]. New York: Gardener, 1976.

定稿日期: 2016-03-01

【责任编辑 陈庆斌】