

A Review of Translation Quality Assessment: Achievements, Problems and Methods

YANG Ningwei

Henan Normal University, China

Received: April 15, 2023

Accepted: May 5, 2023

Published: September 30, 2023

To cite this article: YANG Ningwei. (2023). A Review of Translation Quality Assessment: Achievements, Problems and Methods. *Asia-Pacific Journal of Humanities and Social Sciences*, 3(3), 083–089, DOI: [10.53789/j.1653-0465.2023.0303.011](https://doi.org/10.53789/j.1653-0465.2023.0303.011). p

To link to this article: <https://doi.org/10.53789/j.1653-0465.2023.0303.011>. p

The research is supported by the Teaching Reform Foundation of Xinlian College, HNU (No. 22A002).

Abstract: As an essential part of the application field of translation, translation quality assessment (TQA) is greatly useful in standardizing the translation market, improving translation quality, and promoting translation teaching. This paper explores different kinds of TQA models chronologically from the qualitative and quantitative perspectives, respectively, and points out some problems in current studies on TQA in China. The author, based on the new trends in translation theory and application, proposes three new directions in TQA model research and underlines the positive influence of some language ability scales on the construction and improvement of TQA models.

Keywords: translation quality assessment; qualitative and quantitative perspectives; literature review; language ability scales

Notes on the contributor: YANG Ningwei holds a master's degree in Translation Studies. His major research interest lies in translation criticism and translation teaching. His email address is yangningwei6@163.com.

翻譯品質評估研究綜述：成就、困境與出路

楊寧偉

河南師範大學

摘要：翻譯品質評估是翻譯應用領域的重要組成部分，對規範翻譯市場，提升翻譯品質，促進翻譯教學都有著極其重要的意義。本文從定性和定量兩個方面對國內外翻譯品質評估模式進行了歷時考察，指出我國翻譯品質評估研究存在起步晚、發展慢、後勁不足等問題。針對我國翻譯品質評估研究的困境，筆者結合當前翻譯研究和語言評測等領域的發展趨勢和最新成果，提出了翻譯品質評估研究的三個新方向，並強調了評測量表對於構建和完善翻譯品質評估模式的積極意義。

關鍵詞：翻譯品質評估；定性；定量；研究綜述；量表

基金專案：河南師範大學新聯學院教學團隊計畫專案「基於筆譯量表的大學英語翻譯教學改革研究」(22A002)。

引言

新世紀以來，隨著對外開放不斷深入，中國與世界的關係日益緊密，交流日趨頻繁。翻譯是不同種族間溝通的橋樑，資訊傳遞的準確性是實現高質量對話的根本。自上世紀70年代以來，作為翻譯應用領域的重要組成部分，翻譯品質評估在學術界的持續關注下，不斷迭代更新，日趨完善。在新時代加強國際傳播能力建設要求下，我國語言服務市場呈現出一片繁榮景象，但譯文品質參差不齊，甚至是粗製濫造，溝通失效的現象層出不窮。我們當前面臨的最大問題是缺乏客觀、科學、可操作性強並適用於評估各種類型譯文品質的方法或模式。加強翻譯品質評估模式的完善及其相關研究，提升譯文與原文在微觀和宏觀兩個層面的評估效度和信度，對規範翻譯市場，提升翻譯品質，促進翻譯教學都有著重要的現實意義。本文將從定性和定量兩個方面，對國內外翻譯品質評估模式研究成果進行歷時考察，並結合近幾年翻譯理論和翻譯教學等方面的新進展，嘗試提出翻譯品質評估研究的新方向。

一、國內外翻譯品質評估研究綜述

首先，我們需要明確翻譯品質評估模式的分類，西方大體分為原則參照模式和參數參照模式，國內大體分為言外評價模式和言內評價模式。儘管名稱不同，但分類的依據大致相同，前者較多關注如讀者反應和語篇類型等宏觀層面因素，而後者側重評價語言層面是否等效和「忠信」。根據不同的評估主體和評估目的，楊志紅將翻譯品質評估模式分為定性評估和定量評估，定性評估是根據既定的框架或模式來評價譯文

品質,但並不給出一個確切的分數,而定量評估則以分數的形式呈現評估結果^①。下文將從定性評估和定量評估兩個方面考察不同時期具有代表性的翻譯品質評估模式。

(一) 翻譯品質定性評估

翻譯品質定性評估都是基於某種語言學理論或框架進行的,所以不同的定性評估模式的區別在於不同的理論基礎。定性評估模式主要包括基於讀者反應的評估模式,基於文本類型與功能的評估模式、基於系統功能語言學的評估模式和基於論辯理論的評估模式。

基於讀者反應的評估模式的理論基礎為上世紀 60 年代 Nida 等人所提出的讀者反應理論,即好的譯文應該使目標語讀者的反應與源語讀者的反應相一致。這一重要論斷打破了以往只關注譯文與原文在字詞層面對等與否的傳統模式,從更加宏觀的層面來考察譯文品質。但 Munday 認為,「反應」是一個較為主觀的概念,如果沒有嚴謹的實證研究,是很難判斷目標語讀者的反應是否與源語讀者反應相一致^②。所以,儘管讀者反應論的提出在當時具有重大歷史意義,但不可否認,該評估模式的主觀成分居多,在操作層面上的可行性並不高。

語篇類型學為基於文本類型與功能的評估模式提供了理論基礎。20 世紀 70 年代,Reiss 將文本分為資訊型文本、表情型文本、操作型文本和視聽媒體文本四類,並指出不同類型和功能的文本應該使用不同的品質評估標準。評估資訊型文本譯文品質時主要看原文資訊是否得到了完整、準確的傳達;評估表情型文本時主要考察譯文是否能實現與原文相一致的美學功能;評估操作型文本時主要考察譯文是否與原文具有同等感染力;評估視聽媒體文本主要看譯文能否體現媒體文本的特徵,能否使譯文目標讀者作出與原文目標讀者相一致的反應。基於文本類型與功能的評估模式綜合考慮了文本類型學、語言學、語用學和美學等知識,明確地將譯文的功能、使用者的需求和期待等因素納入了譯文品質評估中,主客觀相結合,是一個系統性和操作性都比較強的翻譯品質評估模式。

上世紀 70 年代末,House 建立了基於系統功能語言學的評估模式,其代表作《翻譯品質評估模式》(1977)和《翻譯品質評估:修訂模式》(1997)在翻譯品質評估研究領域享有盛名。該模式從單個文本的功能、語域、體裁、語言/文化情境等四個層面評估譯文品質^③。此外,House 結合文本性質,提出了「顯性翻譯」和「隱形翻譯」的概念,並區分出了不同的翻譯方法。與基於文本類型與功能的評估模式相比,House 模式更注重微觀的語言層面分析,她強調基於文本對比的語言學途徑能夠確保翻譯品質評估的客觀性,翻譯批評的主要任務應該是對原文和譯文的語言—文本分析和對比,對其他任何社會因素的考量都處於次要地位^④。

基於論辯的評估模式可分為兩類,創立者分別為 Williams 和 Bittner,兩個評估模式儘管名稱相似,但所依據的理論基礎卻相去甚遠,不可混為一談。Williams 認為不同種類的語篇中存在著不同的論辯結構,文本的論辯結構可以為翻譯品質評估提供依據,可以通過考察譯文是否準確地反映原文的論辯結構來評價譯文品質^⑤。基於論辯理論的評估模式主要從論辯宏觀結構和修辭類型兩個方面對原文與譯文進行對比。前者主要涉及語篇的宏觀資訊,後者則用於考察語篇如何通過辭彙、句法等方面的微觀語言手段來實現宏觀論辯結構。武光軍對該模式給予了積極的評價,認為它涵蓋了整個語篇的所有宏觀成分和微觀成分,也體現了兩者之間的互動關係,但該模式在評估賦值上的主觀隨意性降低了其客觀性^⑥。

2020 年,Bittner(2020)在《評價評估者:翻譯品質評估新視角》一書中以教師關於學生評論性翻譯論文



的評估報告為語料,提出了以論辯為基礎的翻譯品質評估新理論。這裏的論辯指的是語篇外評估者為捍衛或否定某一翻譯結果而進行的說理或論證。Bittner 通過「譯者的水仙花」這一圖式,盡可能多地涵蓋影響翻譯品質的因素(如原文本、文本形式、客戶、譯者、文化、政治等),組成論辯內容,然後借鑒 Gregor Betz 的辯證結構理論,提出了客觀全面的論辯評估方法。她認為此模式以論辯為基礎,是獨立於翻譯之外的,所以適用於任何翻譯類型^⑦。Bittner 還指出翻譯品質評估中的主觀性無法規避,但可通過質評標準的系統篩選和運用、實證研究和評估結果的量化達到弱化的目的。該模式「充分考慮到實際的翻譯品質評估中可能會遇到的多元化難題,具有強大的思辨性、開放性以及實操性」^⑧。

(二) 翻譯品質的定量評估

定性評估模式使我們可以從不同維度對譯文品質作出評價,大大提高了翻譯品質評估的全面性、科學性和有效性。但開展定性評估耗時費力,且需要評價者具有一定的語言學理論基礎,並不適用於翻譯水準測試。特別是在大規模翻譯測試中,評分員需要在有限時間內用一個分數(或等級)對譯文品質做出評估,複雜的定性評估模式顯然無法應用,這就需要定量評估模式的介入。開展翻譯評估定量模式研究對翻譯教學和翻譯測試均具有重要意義。

國內學者穆雷提出了錯誤扣分法、綜合法和分析法三種定量評估方法^⑨。錯誤扣分法是指根據譯文中的錯誤數量和性質扣分;綜合法的實質是印象評分法,該方法具有較強的主觀性;分析法是指先制定若干評分項,依據一定的評分標準給每個評分項打分,最後將各個專案的得分相加得到總分。分析法較為複雜繁瑣,所以國內教師很少使用。肖維青提出了錯誤扣分法、印象評分法、整體評分法、分析評分法和綜合評分法^⑩。值得注意的是,她提出的綜合評分法是將錯誤扣分法與整體評分法或印象評分法相結合,抑或請多位評分員評分的一種方法。楊志紅將定量評估方法分為(1)錯誤扣分法;(2)標準參照模式,即參照一定的標準對譯文進行評分。該模式又可細分為整體性評分法和分析性評分法兩類。兩者的區別是,整體性評分法只需得出一個總分,而分析性評分法需要提供各個評分項的分數,最後通過一定比例計算出總分。分析性評分法又可分为多標準模式和簡化標準模式;(3)綜合性評分法,在此模式下,針對同一個翻譯任務採用不同的評分方法,然後對不同的評分方法所得到的結果進行交叉驗證。如劍橋商務英語考試口語的評分就是由兩位考官完成,考官甲進行整體性評分,只需給出一個總分數;考官乙進行分析性評分,對考生在各個維度上的能力表現打分,在兩位考官打分的基礎上產生考生的最終得分^⑪。

錯誤扣分法操作簡單,經常應用於翻譯測試,尤其是大規模測試中。但它容易使評分員糾結於微觀的語言錯誤,而忽視譯文傳達和整體風格等宏觀層面的問題。過多地使用錯誤扣分法對考生也會產生誤導,不利於語篇翻譯觀的培養。整體性評分法操作相對簡單,評分員發揮主觀能動性的空間較大。分析性評分法耗費的精力較多,但更能清晰地指導評分員的評分過程,也可幫助考生瞭解譯文各個方面的得分情況,從而有針對性地完善自己的譯文。分析性評分法中,簡化標準模式操作性更強,更適用於較大規模的翻譯評估。綜合性評分法更能客觀準確地評估考生的能力,但前提是要證明擬運用的不同評分方法之間存在高度的相關性,且都能有效測試考生的翻譯能力。

二、翻譯品質評估的發展趨勢

經過學者們近半個世紀的努力,翻譯品質評估模式研究取得了長足的發展,其科學性和可操作性越來越強。很明顯,質評模式的完善與翻譯學科的拓展是緊密相連的,甚至是「同呼吸,共命運」的。當前翻譯(批評)研究已經進入到了一個新階段,我們預測翻譯品質評估模式發展將呈現以下四個趨勢:(1)譯文在特定情境中的功能將日益得到重視,譯文是否實現了交際功能成為評價翻譯品質的重要標準;(2)更多的大型翻譯測試中會使用分析性評分法。儘管評分員比較認可錯誤扣分法,但其缺點已顯露無疑。2019年,NAATI考試就發佈最新版評分模式,改為使用分析性評分法。而事實證明,標準參照模式下的分析性評分法能夠幫助評分員脫離懲罰式的錯誤扣分模式,鼓勵他們關注影響翻譯品質的宏觀層面因素。這與第一個趨勢是相對應的;(3)定性評估和定量評估的結合。未來合理的翻譯品質評估模式應包含兩個層次上的評估:整體上的模糊評估與局部上的精確評估^⑩。模糊評估反映出了翻譯品質評估的主體性。精確評估反映著翻譯品質評估的客觀性。模糊評估即定性評估,精確評估即定量評估;(4)從單一模式走向多種模式的結合,電腦輔助的品質評估模式將受到更多關注。多種模式的結合必然會增強質評的效度和信度,但同時也需要更多的人力和物力的投入,而電腦技術和人工智慧的發展可以很好解決這個問題。

三、國內翻譯品質評估模式研究的困境

相比於西方,我國的翻譯品質評估模式研究呈現出以下特點:(1)起步晚,發展慢。直到2007年,在系統功能語言學框架下,司顯柱(2007)才建構了國內第一個翻譯品質評估模式,晚於豪斯模式近三十年。在中國知網以篇名「翻譯品質評估」檢索,截止到2023年5月,去除無關文獻後,僅有不足400篇;(2)學科獨立性不強。定性評估研究常以翻譯批評的名義進行,而定量研究常作為翻譯教學或翻譯測試研究的一部分進行,真正冠以翻譯品質評估之名的研究少之又少;(3)代表性成果少,缺少領軍人物。截至目前,僅有司顯柱(2007)在系統功能語言學框架下建構的翻譯品質評估模式和何三寧(2015)從關聯理論出發開展的翻譯品質評估研究產生了一定的影響;(4)用力分散,後勁不足。絕大多數研究集中於探討西方的翻譯品質評估模式在實踐中的應用,對國內相關研究關注嚴重不足。囿於各個模式的自身局限性和翻譯實踐的不斷發展,上述模式在翻譯學界的闡釋、應用與發展逐漸式微甚至銷聲匿跡。

造成這種情況最根本的原因,就是對翻譯品質評估的定位不准,主要關涉跟翻譯批評的關係。張美芳、司顯柱認為兩者是從屬關係,即翻譯品質評估是翻譯批評的一個下義位概念^{⑪⑫}。劉雲虹主編的《批評之批評:翻譯批評理論建構與反思》(2020)及其與許鈞主編的《翻譯批評研究之路:理論、方法與途徑》(2015)皆收錄了近年來國內有關翻譯品質評估的研究成果,這說明兩者在一定程度上都將翻譯品質評估置於翻譯批評研究視域之下。翻譯品質評估的從屬地位極大影響了社會關注度和科研資源的分配,不利於學術共同體的構建。所以,要改變翻譯品質評估模式研究的現狀,必須重新討論並確定其在翻譯研究中的地位。

四、國內翻譯品質評估模式研究的出路

(一) 厘清與翻譯批評之間的關係,強化翻譯品質評估的學科地位

在新形勢下,我們需要重新審視翻譯品質評估和翻譯批評之間的關係。何三寧認為,翻譯品質評估和翻譯批評是既交叉互動又是辯證平等的關係,兩者存在相互利用和借鑒的內容,但它們在研究內容與標準、方法與手段、目的與作用等方面存在明顯差異。他形象地把翻譯批評描述為「打假」環節,把翻譯品質評估描述為「質檢」環節,前者注重產品的真與假、優與劣,而後者則依據一定的標準和參數判斷和檢驗譯文的品質^⑤。在學術層面,我們可以積極借鑒翻譯批評的最新理論成果,來充實翻譯品質評估模式的研究。在社會層面,我們可以通過校際合作,建立學術共同體,爭取學術資源,推動翻譯品質評估的學科建設。確立並提升翻譯品質評估研究地位對譯學具有巨大的應用前景和理論意義。

(二) 加強定量研究,推動定性和定量一體化研究

定性與定量研究適用範圍不同,各有特點,但在翻譯品質評估過程中,兩者同等重要,不可偏廢一方,特別是在社科類譯文品質評估中,沒有定性指導的翻譯品質評估往往會誤入歧途。定性評估主觀性較強,缺乏具體的操作流程,降低了品質評估模式的準確性和可行性,而定量研究正好可以彌補定性研究的缺點。目前翻譯品質評估模式的薄弱點是定量研究,尤其是人工翻譯品質評估的定量研究。當然,定量模式絕不局限於完全形式化的「定量」,在量化過程中,勢必會有定性的判斷。定量與定性相結合,突出定量模式的優勢,在定量模式的基礎上得出定性的結論,才能建立起有效的評估模式。另外,利用電腦技術和人工智慧建立起高效的量化和自動化評估模式也是推動翻譯品質評估的一個重要途徑。

(三) 引進語言測試量表,提升翻譯品質評估模式的適用性和科學性

為改進現有評量工具,讓翻譯測驗更為可靠和有效,臺灣師範大學賴慈芸對錯誤扣分法和三種量表(級分制)進行了對比研究^⑥。其中三種量表指的是劉敏華等人提出的五/五量表(「忠實」五級分/「通順」五級分,兩組獨立評分)、修正後的五/五量表(「正確」五級分/「表達」五級分,兩組合並評分)和六/四量表(「訊息準確」六級分/「表達風格」四級分)。經實證性研究得出:(1)使用錯誤扣分法,評分人對翻譯的主觀看法較易影響結果,適用於需要排名次,如翻譯比賽或入學考,或一個評分人即可完成的評量,如課堂成績或專業考;(2)量表比錯誤扣分法更適用於大型翻譯能力測試,其中六/四量表更為穩定可靠;(3)由一位評分人給兩項分數的方法,可以提高信度。結合賴慈芸的研究,我們認為科學、客觀、有效的翻譯能力評測量表可為建立和完善翻譯品質評估模式提供參照和理論支撐,增強翻譯品質評估的適用性和實用性。《中國英語能力等級量表》(以下簡稱《量表》)作為國家語委語言文字規範自2018年正式實施。教育部考試中心負責人在答記者問時指出,《量表》就像一把能力尺規,對提升考試品質、推動國內外考試對接、促進多元評價、回饋教學等方面都將發揮積極作用,有利於實現「量同衡」。筆譯力量表作為《量表》的重要組成部分,是世界首個全面系統開發的翻譯能力系列量表。它對筆譯能力進行了科學合理的界定,並運用實證研究方法對

筆譯能力進行了科學的分級描述,將在制定合理的筆譯學習目標、學習結果評測、課程設置、編寫教學大綱和考試大綱、編寫教材、制定教育政策法規、評定各類人員的筆譯能力各方面工作中發揮重要指導作用^⑰。結合中國英語能力筆譯量表開展翻譯品質評估研究對翻譯測試和翻譯教學無疑具有重要意義,也是一個大有可為研究方向。

結語

經濟的快速發展和國際政治地位的提升促使中國漸漸走向世界舞臺的中心,參與到更多的國際事務中。中國共產黨二十大報告中明確提出要增強中華文明傳播力、影響力,講好中國故事、傳播好中國聲音,推動中華文化更好走向世界的要求。經過多年的發展,翻譯品質評估研究取得了階段性的成果,但就目前情況而言,還未受到應有的重視,在理論、模式和標準等層面進行的研究還不足以滿足具體的應用要求,大多數的翻譯研究活動還停留在理論分析和學術談論中,應用研究缺少前瞻性和系統性,翻譯實踐中應用的實例還不充實。因此,今後的翻譯品質評估研究應該結合中國語境下具體的翻譯要求,充分利用電腦和人工智慧領域的成果,加強翻譯品質評估理論與實踐的並行發展。

注釋

- ① 楊志紅:《翻譯測試與評估研究》,外語教學與研究出版社2019年版,頁149。
- ② Munday, J. (2001). *Introducing translation studies: Theories and applications*. Routledge, 43.
- ③ House, J. (1997). *Translation quality assessment: A model revisited*. Gunter Narr Verlag, 105–108.
- ④ House, J. (2001). Translation quality assessment: Linguistic description versus social evaluation. *Meta*, 46(2), 255.
- ⑤ Williams, M. (2001). The application of argumentation theory to translation quality assessment. *Meta*, 46(2), 326–344.
- ⑥ 武光軍:《翻譯品質評估:論辯理論模式評介》,《外語研究》,2006年第4期,頁72–74。
- ⑦ Bittner, H. (2020). *Evaluating the evaluator: A novel perspective on translation quality assessment*. Routledge, 276.
- ⑧ 李夢利、張娟:《翻譯品質評估中的「論辯」——「評價評估者:翻譯品質評估新視角」述評》,《外文研究》,2023年第1期,頁101。
- ⑨ 穆雷:《翻譯測試及其評分問題》,《外語教學與研究》2006年第6期,頁466–471。
- ⑩ 肖維青:《本科翻譯專業測試研究》,人民出版社2012年版。
- ⑪ 楊志紅:《翻譯測試與評估研究》,外語教學與研究出版社2019年版。
- ⑫ 武光軍:《當代中西翻譯品質評估模式的進展、元評估及發展方向》,《外語研究》,2007年第4期,頁78。
- ⑬ 張美芳:《翻譯研究的功能途徑》,上海外語教育出版社2005年版,頁9。
- ⑭ 司顯柱:《翻譯研究:理論·方法·評估》,中國文史出版社2005年版,頁182。
- ⑮ 何三寧:《並列還是從屬:翻譯品質評估與翻譯批評之關係》,《南昌大學學報》,2009年第4期,頁130。
- ⑯ 賴慈芸:《四種翻譯評量工具的比較》,《編譯論壇》2008年第1期,頁71–92。
- ⑰ 白玲、馮莉、嚴明:《中國英語筆譯能力等級量表的構念與原則》,《現代外語》,2018年第1期,頁108。