

# Transparenz bei der Evaluierung von Übersetzungen

*Blanka Jenčíková*

Die Beurteilung von Übersetzungen zusammen mit der Revision und der Qualität der translatologischen Tätigkeit gewinnen in den letzten Jahrzehnten an Bedeutung und rücken immer öfter in den Fokus der translatologischen Untersuchungen. Trotzdem spürt man heute in der Praxis immer noch das Fehlen einer systematischen Konzeption in diesem Bereich und es stellt sich auch die Frage, ob es überhaupt möglich ist, für die Bewertung und auch die Revision von Übersetzungen – also für zwei höchst subjektive Tätigkeiten – eine Methodologie festzulegen, die zu objektiven und transparenten Ergebnissen führen würde. In diesem Artikel werden zuerst sowohl didaktische als auch praxisorientierte Konzepte von Bewertungskriterien dargelegt, im weiteren Teil folgt ein Vergleich des Evaluierungsprozesses im Übersetzungsunterricht und in der übersetzerischen Praxis mit Fokus auf ihre Transparenz. Abschließend wird die Anwendung einer Software zur effizienteren analytischen Leistungsbewertung vorgestellt.

## 1 Evaluierung von Übersetzungen

Im Rahmen der Translatologie wird die Evaluierung von Übersetzungen dem Bereich der angewandten Translatologie zugeordnet. Die traditionelle Bezeichnung des sich mit der Beurteilung der Übersetzungsqualität befassenden Fachs ist Übersetzungskritik, mit deren Aufgaben und Zielen sich die Übersetzungswissenschaft erst intensiver in den 70er Jahren des 20. Jahrhunderts zu beschäftigen begann, und deren Gegenstand ursprünglich hauptsächlich literarische Übersetzungen bildeten. Mit der allmählichen Erweiterung des Interesses an den pragmatischen Texten und der damit zusammenhängenden Steigerung der Anforderungen an die Qualität der übersetzerischen Tätigkeit etablierte sich gegen Ende des 20. Jahrhunderts der Begriff Translation Quality Assessment (TQA), also die Beurteilung der Übersetzungsqualität. Zurzeit werden beide Begriffe – also die Übersetzungskritik und die Beurteilung der Übersetzungsqualität – von einigen Autoren synonym verwendet, andere Autoren halten dagegen die Beurteilung der Übersetzungsqualität für einen allgemeineren Begriff als die Übersetzungskritik (vgl. Zehnlová 2015: 43-44). Allerdings wird heute in der Praxis von den Übersetzungsdienstleistern und Übersetzern im Zusammenhang mit dem Qualitätsmanagement und der Evaluierung von Fachübersetzungen vorwiegend der Begriff Translation Quality Assessment angewendet, wobei unter Übersetzungskritik eher die Beurteilung der Qualität von literarischen Übersetzungen verstanden wird.

## 2 Bewertungskriterien und Übersetzungsfehler

Jede Evaluierung von Übersetzungsleistungen – in der Praxis sowie im didaktischen Prozess – erfordert ein möglichst hohes Maß an intersubjektiver Zuverlässigkeit, Konsistenz und Effizienz der Bewertung. Die intersubjektive Zuverlässigkeit führt dazu, dass unterschiedliche Bewerter bei der Evaluierung derselben Übersetzung unter Nutzung von gegebenen Methoden und Kriterien zu einem ungefähr gleichen Ergebnis kommen sollten. Unter Konsistenz wird die Beständigkeit von Bewertungsmethoden und -kriterien verstanden und die Effizienz hängt mit dem Zeitaufwand der Beurteilung zusammen (vgl. Kubánek, Molnár 2015: 148-149). Um dies zu erzielen, sind plausible und effiziente Bewertungskriterien notwendig. „Nur solche Kriterien

gestatten es, einerseits Übersetzungen sachgerecht zu kritisieren, andererseits mit Kritik richtig umzugehen, d. h. berechtigte Kritik zu akzeptieren, unberechtigte Kritik aber mit entsprechender Begründung auch zurückweisen zu können“ (Kautz 2002: 278).

Bei der Ermittlung der Qualität einer Übersetzung wird üblicherweise aufgrund von festgelegten Kriterien nach „defekten“ Stellen, also nach Fehlern gesucht. Somit wird der Übersetzungsfehler zum Hauptkriterium der Bewertung. Mit dem Problem, was ein Fehler und was keiner ist und wie schwerwiegend er sich auf die Gesamtqualität des Zieltextes auswirkt, werden sowohl Übersetzer in der Praxis als auch Lehrkräfte ständig konfrontiert. Im Fremdsprachenunterricht werden Fehler allgemein als Abweichung von einem Normen- oder Regelsystem definiert. In der Translationswissenschaft kam es in den letzten Jahrzehnten zu einer Entwicklung in der Auffassung des Übersetzungsfehlers von den linguistisch orientierten Ansätzen bis zum funktionalen Ansatz, der die Translatfunktion als Hauptkriterium in den Focus zieht: „Ein Übersetzungsfehler ist also eine „Nicht-Erfüllung“ des Übersetzungsauftrags in Bezug auf bestimmte funktionale Aspekte. Das bedeutet: Jede Übersetzungsleistung kann nur in Bezug auf ein vorgegebenes (funktionales) Übersetzungsziel beurteilt werden. Dieses Übersetzungsziel muss dem Übersetzer bzw. der Übersetzerin bekannt sein.“ (Nord 2010: 179).

## 2.1 Didaktisch orientierte Modelle von Bewertungskriterien

Die Erstellung von einheitlichen oder zumindest maßgebenden Kriterien im Bereich der Evaluierung von Übersetzungen ist eine anspruchsvolle Aufgabe. In diesem Abschnitt werden drei Modelle vorgestellt, die konkrete didaktisch anwendbare Bewertungskriterien anbieten.

Laut R. Stolze (1999: 240ff.) sollte bei der Frage nach Bewertungskriterien unterschieden werden, ob die Übersetzung als Produkt, das eingekauft wird, im Prozess ihrer Erstellung oder als Objekt der Analyse betrachtet werden soll. In der Praxis prüfen die von den Auftraggebern bestimmten Überprüfer prospektiv, ob die Übersetzung – also das Produkt – die angesetzten Qualitätsanforderungen erfüllt (Layout, Kohärenz, Terminologie, Textsorte, Stil). Die Perspektive des Translators im Übersetzungsprozess ist wiederum eher dynamisch, und zwar bei der Anwendung von einigen ganzheitlichen Begründungsmaßstäben hinsichtlich der Klarheit, Genauigkeit und Wirkung des Zieltextes. Im Übersetzungsunterricht erfolgt die Evaluation von Übungsleistungen in einer statischen Fehleranalyse anhand der festgelegten Fehlerklassifikation. R. Stolze versucht diese unterschiedlichen Perspektiven auf Qualität und deren Bewertung bei Auftraggebern, Übersetzern und Übersetzungslehrern durch eine Systematik der translatorischen Kategorien Thematik, Lexik, Pragmatik und Stilistik zu lösen, in welchen alle drei Sichtweisen – prospektive, dynamische und statische – vereinbart werden können. Solche translatorischen Kategorien als Bewertungskriterien für Übersetzungen können ihrer Meinung nach mit dem linguistischen Instrumentarium dargestellt werden, denn beim Übersetzen gehe es um ein Handeln in und mit Sprache. Diese Kategorien können vielleicht auch helfen, den viel strapazierten Begriff der „Qualität der Übersetzung“ besser zu fundieren. Unter der Kategorie Thematik versteht R. Stolze die Kohärenz (Textgestalt, logischer Aufbau), die Ausgangstextanalyse (textexterne und textinterne Faktoren, Wortbedeutung, Zeichenfunktion) und die Semantik (thematische Struktur, Gliederungssignale, Textstatus). Die Lexik bezieht sich auf die Aspekte der Fachterminologie, also auf ihre Konsistenz und Exaktheit. Die Kategorie Pragmatik verweist auf die Adressaten in der Zielkultur. Als Unterkategorien führt sie hier die Funktionsadäquatheit, den Umgang mit Kulturunterschied und die Textfunktion in Bezug auf den Empfänger an.

In der didaktisch orientierten Klassifizierung von Übersetzungsfehlern von C. Nord in ihrem Buch *Fertigkeit Übersetzen*, das 2010 erschien, führt die Autorin bei der Kategorisierung der Übersetzungsfehler diese drei Kategorien an (Nord 2010: 181ff.):

1. Pragmatische Übersetzungsfehler (PÜF): also Verstöße gegen die Funktionshierarchie und Verstöße gegen die Situationseinbettung, wie z. B. Inkohärenz, falsche Information, unverständliche Realiendifferenz, falsche Deixis;
2. Kulturelle Übersetzungsfehler (KÜF): Nichteinhaltung der durch den Übersetzungsauftrag geforderten Konventionen, wie z. B. falsches Maßsystem, Nichtanpassung an Textsortenkonventionen, Nichteinhaltung der stilistischen Konventionen;
3. Sprachliche Übersetzungsfehler (SÜF): Nichteinhaltung der zielsprachlichen Gebrauchsnormen infolge von Interferenzen mit ausgangssprachlichen Strukturen, wie z. B. lexikalische Interferenz, falscher Bezug, Übernahme von AS-Fokusstrukturen.

Zusätzlich zu diesen drei Hauptkategorien nennt sie auch den zielsprachlichen Normverstoß, also z. B. falsche Verbform, falsche Wortbildung, Kommafehler usw., wobei sie meint, dass diese Art der Fehler in der Übersetzer Ausbildung bei der Übersetzung in die Fremdsprache zumindest in der Anfangsphase aus der Bewertung ausgeklammert werden sollte.

G. Hansen (2008b: 114ff.), die Autorin des in den 80er Jahren entwickelten CBS-Modells für das Management der Übersetzungsqualität, führt folgende Fehlklassifikation an, die für das Sprachpaar Dänisch – Deutsch zusammengestellt wurde:

1. Fehlertypen sortiert nach den betroffenen Stufen und Ebenen der Sprachbeschreibung:
  - Pragmatische Fehler: falsche Einschätzung des Auftrags und/oder der Situation (Auftrag missverstanden, Übersetzungstyp falsch, Text nicht der Zielgruppe und dem Zweck angepasst, Informationen fehlen oder sind zu umfassend, Deixis falsch oder nicht beachtet, Normen und Konventionen missachtet, direkte Übertragung von Abkürzungen)
  - Textlinguistische Fehler: Verstöße gegen die semantische, logische, stilistische Kohärenz (semantisch-logische Fehler, Referenz unklar/falsch, Zeitverhältnisse unklar, Kategorie falsch, Modalität falsch, Informationsstruktur falsch, Stilbruch)
  - Semantische Fehler: falsche Bedeutung auf Wort- oder Satzebene
  - Idiomatiche Fehler: Wörter und Wendungen, die zwar semantisch deckend sind, aber in einem bestimmten Zusammenhang oder in einer bestimmten Situation in der ZS nicht benutzt werden - Stilistische Fehler: Falsche Stilebene, falsche Stilfärbung, falscher Ton
  - Morphosyntaktische Fehler: Falsche Wortstruktur im Verhältnis zur Funktion im Satz, falsche Beugungen der Wörter
  - Syntaktische Fehler: Falsche Satzstruktur, Wortstellung
  - Fakten falsch: Fehler im Hinblick auf Zahlen, Daten, Eigennamen usw.
  - Auslassungen: Fehlende Einheiten im Verhältnis zum AT
2. Fehlertypen sortiert nach der Ursache „Interferenz“:
  - Interferenz auf Wortstufe
  - Syntaktische Interferenz
  - Textsemantische Interferenz
  - Kulturelle Interferenz

G. Hansen (2008a: 320ff.) präsentiert ihre Klassifikation als ein offenes System, das um weitere Kriterien ergänzt werden kann. Im Rahmen der Bewertung empfiehlt sie, als Motivationsfaktor gute Übersetzungslösungen mit Pluspunkten zu belohnen. Die logische Struktur dieser Kriterien unterstützt ihre Übersichtlichkeit und ermöglicht eine einfache Anwendung, allerdings wird hier das Gewicht der Fehler nicht in Betracht gezogen.

Die Auswirkung der (Übersetzungs-)Fehler auf den Zieltext und damit auf die Qualität der Übersetzung ist nicht gleich. Ein Tippfehler oder ein fehlendes Komma, solange sie den Sinn des Zieltextes nicht verändern, sind in der Übersetzung eines Studenten oder in der Praxis in einer nicht zur Veröffentlichung bestimmten Übersetzung, die nachher vielleicht von zwei oder drei Empfängern gelesen wird, von keiner großen Bedeutung. Völlig anders würden diese

Fehler in einer druckreifen Übersetzung ausfallen. Bei der Bewertung von Übersetzungen ist es also notwendig, die Relevanz der Fehler in Bezug auf den durch den Übersetzungsauftrag definierten Verwendungszweck zu differenzieren. Eine dichotomische und rein qualitative Unterscheidung zwischen „richtig“ und „falsch“ wäre in diesem Fall nicht ausreichend, da die Folgen eines pragmatischen Fehlers bestimmt wesentlich schwerwiegender sind als die eines Tippfehlers. Die Fehlergewichtung stellt hier eine Lösung dar, die aber nicht einheitlich und global anzuwenden ist.

Laut C. Nord (2011: 276ff.) ergibt sich aus der funktionalen Orientierung, dass pragmatische Aspekte, und unter diesen vor allem die Textfunktion, Vorrang vor anderen Aspekten haben und die Vermittlung ihrer Bedeutung sollte eine Priorität der ersten Phasen des Übersetzungsunterrichts sein. Das Gewicht der kultur- und sprachpaarspezifischen Fehler ergibt sich aus der Übersetzungssituation.

## 2.2 Bewertungskriterien in der Praxis

Der wachsende Druck auf die Qualität der Übersetzungsdienstleistungen führte in den letzten Jahrzehnten zur Einführung von entsprechenden Standards und Normen zur Qualitätssicherung. Seit 1987 etablierte sich in der Industrie und im Dienstleistungsbereich die ISO 9000er-Reihe als bewährtes Instrument des Qualitätsmanagements. Die Übersetzungsdienstleister lassen sich nach ISO 9001 – der weltweit anerkannten und meistverbreiteten Norm für Qualitätsmanagement – zertifizieren. 2006 wurde die erste internationale Norm für Übersetzungsdienstleistungen – die EN 15038 – veröffentlicht, die anschließend im Mai 2016 durch die ISO 17100 „*Übersetzungsdienstleistungen – Anforderungen an Übersetzungsdienstleistungen*“ abgelöst wurde. ISO 17100 enthält die „Vorschriften für Übersetzungsdienstleister (ÜDL) zur Abwicklung der Kernprozesse, zu den Mindestanforderungen an die Qualifikation, die Verfügbarkeit und den Einsatz der Ressourcen sowie zu den weiteren für die Bereitstellung einer qualitativ hochwertigen Übersetzungsdienstleistung erforderlichen Tätigkeiten“ (www1).

In Bezug auf die Zertifizierung sind Übersetzungsdienstleister verpflichtet, für einen angemessenen Qualitätsstandard ihrer Dienstleistungen zu sorgen, wodurch sie eigene Verfahren zur Revision und Evaluierung von Übersetzungen zwecks Überprüfung sowohl der Qualität von Translaten als auch der translatorischen Kompetenzen der Übersetzer entwickeln müssen. Dem Übersetzungsmarkt stehen im Bereich der Bewertungskriterien mehrere Normen und Modelle zur Verfügung, die ursprünglich überwiegend für Softwareanwendungen oder für die maschinelle Übersetzung entwickelt worden sind, wie SDL TMS Classic Model, Lisa, MQM oder SAE J2450.

Das SDL TMS Classic Model ist ein sehr einfaches Bewertungsmodell, das eine einfache Anwendung ermöglicht. Es wird im Translation-Memory-System SDL Trados Studio als Grundlage für die Bewertung von Übersetzungen angewendet. Das Modell umfasst sieben linguistische Fehlerkategorien, die ohne Unterscheidung zwischen verschiedenen Schweregraden oder Gewichtungen jeweils mit einem Fehlerpunkt bewertet werden: Gesamtbewertung, Genauigkeit, Grammatik, Rechtschreibung, Stil, Terminologie, fachliche Genauigkeit. Das LISA QA Model wurde von der Localization Industry Standards Association, kurz LISA, ursprünglich für die Bewertung von Übersetzungen von Software-Oberflächen entwickelt. Es werden drei Schweregrade von Fehlern unterschieden: leicht, schwer, kritisch (minor, major, critical). Neben sprachlichen Fehlertypen werden weitere Typen berücksichtigt, wie Fehler im Layout und solche, die die Funktionalität der Software beeinträchtigen, zum Beispiel Links. Das Modell umfasst 25 Kategorien für das Software User Interface und 20 für die Dokumentation. Das LISA-Modell ist deutlich komplexer und differenzierter als das von SDL, vor allem weil es Gewichtungen zulässt. Allerdings ist es stark auf die Bedürfnisse der Lokalisierungsindustrie

zugeschnitten, für die es ursprünglich vorgesehen war. Die Vielzahl von Kategorien erfordert eine längere Einarbeitungsphase (vgl. Ottmann 2017: 36).

Das Modell Multidimensional Quality Metrics, kurz MQM, wurde zunächst für die Beurteilung maschineller Übersetzungen entwickelt und inzwischen wird es auch für die Bewertung von Humanübersetzungen genutzt. Gemessen wird nach folgenden Kriterien: Angemessenheit, Flüssigkeit, Originaltreue, Internationalisierung, Gebietsschema, Stil, Terminologie, Design, Sonstige. Die neun Hauptkategorien werden in insgesamt über 100 sprachliche und nichtsprachliche Unterkategorien verästelt. Das Modell unterscheidet „potenzielle Probleme“ und Fehler. Das MQM-Modell kann mit einer hohen Flexibilität und einem umfassenden Kategorienkatalog punkten, der an die jeweiligen Anforderungen des Übersetzungs-Workflows angepasst werden kann. Die hohe Komplexität erfordert jedoch eine längere Einarbeitungsphase und zudem ist bei diesem Modell keine Fehlergewichtung vorgesehen (vgl. Ottmann 2017: 37-38).

Das Modell auf Grundlage der Norm SAE J2450 wurde von der Society of Automotive Engineers (SAE) speziell für die Bedürfnisse der Automobilindustrie entwickelt und wird heute auch in anderen Bereichen verwendet. Das Modell hat sieben primäre Fehlerkategorien (Terminologiefehler, inhaltlicher Fehler, Auslassung, Strukturfehler, Rechtschreibfehler, Interpunktionsfehler, sonstiger Fehler) und zwei sekundäre Kategorien für jede primäre Kategorie (schwere und leichte Fehler). Außerdem ist für jede primäre und sekundäre Kategorie eine Gewichtung festgelegt. Die Kategorisierung der Fehler erleichtern vier Metaregeln, nach welchen ein Fehler im Zweifel zum Beispiel als schwerwiegend zu betrachten ist (www2).

### 3 Transparenz des Evaluierungsprozesses in Bezug auf seine Bestandteile

Die Evaluierung in der Praxis und die Evaluierung im Übersetzungsunterricht sind zwei unterschiedliche Verfahren, da sie verschiedene Ziele verfolgen und wesentliche Differenzen bei den Beziehungen zwischen den einzelnen Prozessbeteiligten aufweisen.

#### 3.1 Evaluierungsprozess im Übersetzungsunterricht

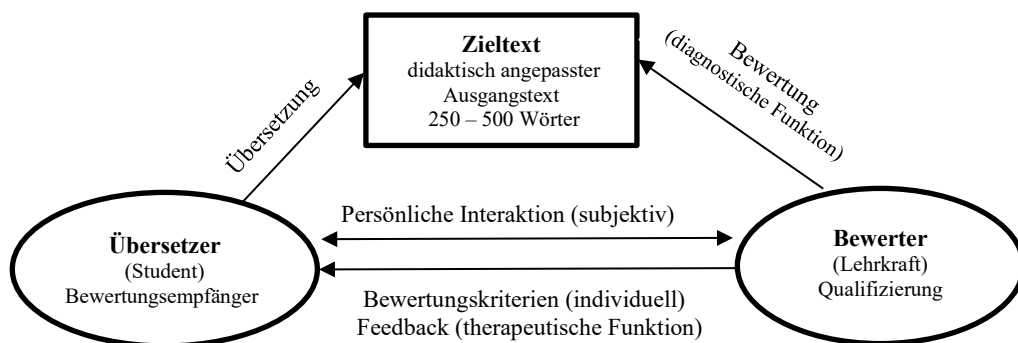


Abb. 1: Evaluierungsprozess im Übersetzungsunterricht

Im Übersetzungsunterricht stellt das Objekt der didaktischen Evaluierung ein anhand eines didaktischen Übersetzungsauftrags erstellter Zieltext dar, dessen Länge in der Regel 1-2 Normseiten (250 – 500 Wörter) beträgt. Als Beteiligte treten in den Evaluierungsprozess die Lehrkraft als Bewerter und der Student als Übersetzer und Bewertungsempfänger ein, die üblicherweise in persönlichem Kontakt stehen, wodurch eine gegenseitige Interaktion ermöglicht wird. Diese Tatsache trägt bedeutend zur Erhöhung der Transparenz bei der Bewertung

bei, indem den Studenten die angewendeten Bewertungskriterien bekannt sind. Von großer Bedeutung ist ebenfalls das mit der Bewertung erhaltene Feedback, das der Verstärkung der therapeutischen Funktion der Evaluierung dient. Allerdings bringt der persönliche Kontakt die Gefahr der Subjektivität mit sich.

Was die Qualifizierung des Bewerbers betrifft, wird davon ausgegangen, dass eine Hochschullehrkraft über ausreichende Kenntnisse und Erfahrung verfügt. Doch kann die Objektivität der Bewertung durch Probleme kognitiver Art gestört werden, und zwar infolge der fehlenden Konsequenz. Besonders bei größeren Studiengruppen, wenn zum Beispiel mehr als 20 Übersetzungen desselben Ausgangstextes zu bewerten sind, kann es passieren, dass der Bewerter identische Fehler in allen Zieltexten nicht bemerkt bzw. nicht gleich bewertet. Diese Inkonsequenz kann auch mit der Ermüdung und dem Nachlassen der Aufmerksamkeit zu tun haben, oder auch mit der Tatsache, dass der Bewerter sich an die Fehler gewöhnt und toleranter wird. Die Gefahr der Inkonsequenz und Ermüdung wächst proportional zur Anzahl der zu bewertenden Zieltexte (vgl. Hansen 2007: 122-123).

Die Bewertungskriterien legt üblicherweise die Lehrkraft (eventuell aufgrund von Vorgaben der Ausbildungseinrichtung) fest, wobei sie in Abhängigkeit von dem verfolgten didaktischen Ziel bei jeder Leistungsbewertung die Schwerpunkte und auch die Fehlergewichtung individuell anpassen kann. In dem Sinne steht der Lehrkraft zum Beispiel die Möglichkeit zur Verfügung, bei einer Übersetzung in die Fremdsprache rein sprachliche Fehler aus der Bewertung auszuklammern. Die Bewertungskriterien sollten verständlich, plausibel und effizient sein, im Rahmen des Übersetzungsunterrichts konsistent angewendet werden und vor allem den Studenten bekannt sein, damit die Transparenz der Bewertung sichergestellt wird.

Ferner kann die Lehrkraft selbst entscheiden, ob ein holistisches oder ein analytisches Bewertungsverfahren angewendet wird. Bei der holistischen Bewertung erfolgt eine ganzheitliche Beurteilung aller relevanten Kriterien. In Form einer wörtlichen Bewertung wird der Gesamteindruck beurteilt und die Bewertung erfolgt oft schneller und einfacher als bei der analytischen Bewertung. Die Gewichtung der einzelnen Kriterien kann von Fall zu Fall unterschiedlich ausfallen und der Bewerter hat eine größere Bewertungsfreiheit, indem er leichter gute Übersetzungslösungen berücksichtigen kann. Der Erfahrungsschatz und die Intuition des Beurteilers haben bei dieser Bewertungsmethode größeren Einfluss auf das Ergebnis. Bei der analytischen Bewertung werden dagegen einzelne Kriterien festgelegt, die alle zur Gesamtbewertung beitragen. Jedes Kriterium wird einzeln beurteilt und die Einzelbewertungen werden schließlich zu einem Gesamtergebnis zusammengefasst. Hierbei ist es möglich, die einzelnen Kriterien unterschiedlich zu gewichten, um so ihren Anteil an der Gesamtbewertung anzupassen. Die Gesamtleistung der Studenten kann so objektiver, reliabler und valider bewertet werden (vgl. Kreuzer 2018: 3).

Um den therapeutischen Ansatz der Evaluierung zum Tragen zu bringen, ist im Übersetzungsunterricht die Wahl der geeigneten Form der Präsentation von Evaluierungsergebnissen von besonderer Wichtigkeit. Bei einer summativen Bewertung werden die Ergebnisse den einzelnen Studenten individuell zugänglich gemacht. Der Bewerter muss fähig sein, seine Bewertung zu begründen und eventuell mögliche Lösungen von aufgetretenen Problemen anzubieten. Ein anderer Weg wäre die formative Bewertung, bei welcher es sich um eine kollektive Präsentation im Unterricht handelt und die zur Diskussion über einzelne Probleme, ihre Ursachen und Lösungen anregen soll (Kubánek/Molnár 2015: 151). Der Vorteil der formativen Bewertung aus der Sicht der Fehlertherapie beruht in der Möglichkeit aus Fehlern von anderen Studienkollegen zu lernen und über die Fehlerursachen und die zulässigen Lösungen zu diskutieren.

Zusammenfassend kann festgestellt werden, dass die Evaluierung im Übersetzungsunterricht eine sowohl diagnostische als auch therapeutische Funktion erfüllt. Die Transparenz des gesamten Prozesses gegenüber dem Übersetzer (dem Studenten) liegt in den Händen der

Lehrkraft als Bewerter. Und nicht zuletzt ist zu betonen, dass eine transparente Leistungsbeurteilung unter anderem eine bildende Funktion erfüllt. Die ständig wachsenden Anforderungen an das Qualitätsmanagement hat zur Folge, dass es in den letzten Jahren zur Neuprofilierung des Übersetzerberufs kommt, konkret geht es um die Erweiterung um die Revisions- und Evaluierungstätigkeiten. Aus dieser Sicht erweist sich eine praxisnah gestaltete und transparente Evaluierung im Rahmen des Übersetzungsunterrichts als gutes Instrument für die breit gefächerte Vorbereitung der künftigen Übersetzer auf die Praxis.

### 3.2 Evaluierungsprozess in der Praxis

Die Übersetzungen werden in der Praxis meistens in zwei Situationen bewertet. Erstens prüfen die Übersetzungsdienstleister (ÜDL) bei den neuen Bewerbern anhand von kurzen Testübersetzungen aus verschiedenen Fachbereichen oder anhand von kurzen realen Aufträgen ihre translatorischen Kompetenzen und hauptsächlich ihre „Tauglichkeit“ für die künftige Zusammenarbeit. Zweitens bewerten überwiegend die zertifizierten ÜDL im Rahmen des eingeführten Qualitätsmanagementsystems die erstellten Übersetzungen verschiedener Länge oft im Zuge einer bilingualen Revision, um die Qualität ihrer Dienstleistungen Auftraggebern (Kunden) und Auditoren gegenüber nachweisen zu können. An dem Evaluierungsprozess beteiligen sich demnach der ÜDL, der Bewerter, der Übersetzer und indirekt auch der Kunde (Auftraggeber), der bei der Übernahme des fertiggestellten Auftrags eine dichotomische Beurteilung von gut oder schlecht vornimmt. Erachtet er die Qualität der bestellten Dienstleistung als nicht genügend, folgt eine Reklamation als negatives Ergebnis seiner Qualitätsprüfung. Als Bewerter treten in der Praxis oft erfahrene Übersetzer auf. Hier bietet sich die Frage an, in wieweit diese Personen für die Evaluierung qualifiziert sind. Dies erfordert nämlich sowohl linguistische als auch Fachkenntnisse aus dem betreffenden Fachgebiet, um zum Beispiel die Richtigkeit der Terminologie prüfen zu können. Auf der anderen Seite trägt die Tatsache, dass der gesamte Prozess anonym verläuft, also der Bewerter und der Bewertete sich nicht kennen, zur Erhöhung der Objektivität bei.

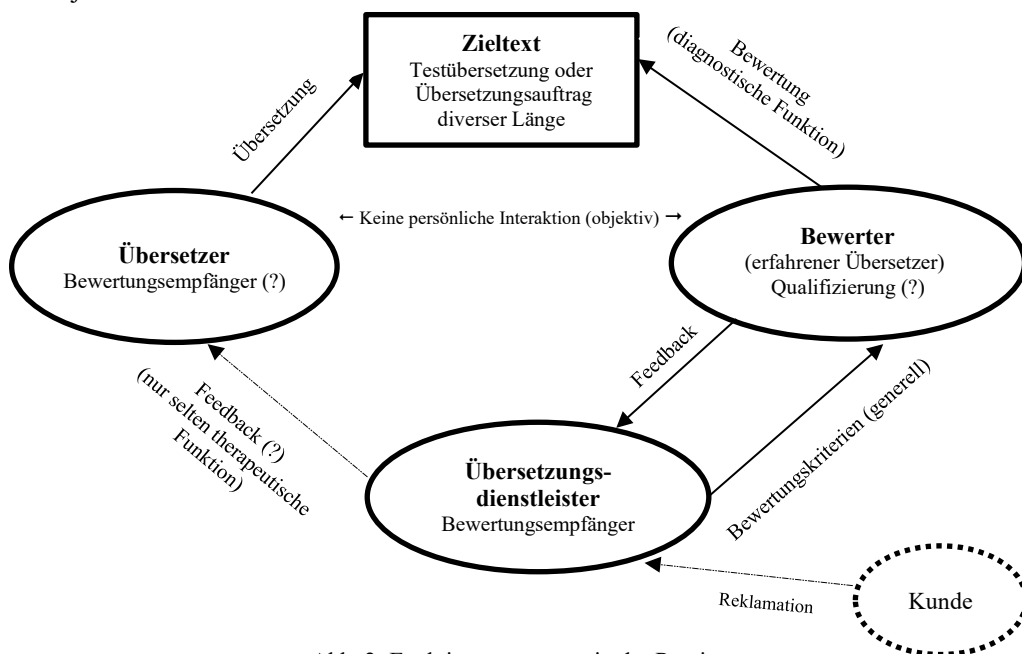


Abb. 2: Evaluierungsprozess in der Praxis

Der primäre Bewertungs- und somit auch Feedbackempfänger ist hier im Gegensatz zur Evaluierung im Übersetzungsunterricht der ÜDL, der unter Umständen die Bewertung und das Feedback an den Übersetzer weiterleiten sollte. In der Realität wird den neuen Bewerbern um die Zusammenarbeit mit einem ÜDL lediglich das Ergebnis mitgeteilt, nämlich ob sie als Übersetzer akzeptiert oder abgelehnt werden, und die Ergebnisse von laufenden internen Qualitätsprüfungen erfahren die Übersetzer höchstens bei Feststellung mangelnder Kompetenzen eines Übersetzers oder natürlich im Falle einer Kundenreklamation. Es ist wirklich schade, dass hier bis auf einige wenige Ausnahmen die therapeutische Funktion der Evaluierung verloren geht. Die Evaluierung in der Praxis verfolgt also in erster Reihe diagnostische Zwecke, nämlich die Beurteilung, ob die Translatqualität den gestellten Anforderungen entspricht.

Die Bewertungskriterien und das Bewertungsverfahren, d. h. ob die Evaluierung holistisch oder analytisch erfolgen soll, bestimmt ebenfalls der ÜDL. In der Regel werden bei einem ÜDL generell gleiche Bewertungskriterien angewendet, die den Übersetzern nicht bekannt sind, solange sie bei demselben ÜDL nicht als Revisoren oder Bewerter tätig sind. Um ein möglichst hohes Maß an intersubjektiver Zuverlässigkeit, Konsistenz und Effizienz der Bewertung sicherzustellen, sind die Bewertungskriterien eindeutig und nachvollziehbar festzulegen. Die Fehlerkategorien und die Gewichtung von Fehlern müssen verständlich und logisch aufgebaut werden und vor allem eine einfache und effiziente Handhabung ermöglichen. Erhält ein Bewerter nur eine Liste von Fehlerkategorien ohne nähere Erklärung, ist aufgrund der unklaren und nicht eindeutigen Definition die intersubjektive Zuverlässigkeit gefährdet, da dieselben Fehlerkategorien bei verschiedenen Bewertern unterschiedlich verstanden werden können. Wenn aber ein Bewerter andererseits zusammen mit einem Bewertungsauftrag ein Manual von 10 Seiten erhält, leidet wieder die Effizienz des gesamten Bewertungsverfahrens.

Zusammenfassend ist festzustellen, dass der Evaluierungsprozess in der Praxis wesentliche Abweichungen von dem didaktischen Evaluierungsprozess aufweist. Die wichtigsten Faktoren stellen vielleicht eine andere Zielsetzung und die Anonymität dar, infolge welcher der therapeutische Ansatz der Bewertung stark unterdrückt wird und die Forderung der Transparenz sich auf die Beziehung zwischen dem Bewerter und dem ÜDL verschiebt.

### 3 Transparenz durch die softwaregestützten Bewertungsinstrumente

Zur Evaluierung von Übersetzungen in der Übersetzerausbildung stehen heutzutage viele elektronische Möglichkeiten zur Verfügung. Jedes angewendete Bewertungsinstrument sollte, wie bereits oben erwähnt, intersubjektiv zuverlässig, konsistent und effizient sein. Von großem Vorteil wäre auch seine Benutzerfreundlichkeit und Flexibilität, also eine rasche Einarbeitungsphase, klare und verständliche Handhabung sowie die Möglichkeit der Einstellung und Änderung von Parametern, ihrer Auswertung und die Möglichkeit der flexiblen Nutzung auch für andere Forschungsvorhaben und für den Unterricht. Als Instrument, das diese Anforderung erfüllt – hauptsächlich bezüglich der automatischen Auswertungsoption, die zum Beispiel bei MS Word oder Moodle nicht vorhanden ist – wurde MAXQDA gewählt (www3). Diese Software wird für die qualitative Inhaltsanalyse mit anschließender quantitativer Auswertung (Mixed Methods) genutzt. Das zentrale Analysewerkzeug stellen hier die Codes dar, welche die einzelnen Fehlerkategorien repräsentieren und den markierten Teilen des Materials zugeordnet werden. Jedem Code werden bei seiner Erstellung ein Name, eine Farbe und eine Beschreibung zugeordnet, wobei die auf diese Weise entstandene Codeliste zu jeder Zeit umorganisiert oder geändert werden kann. Im Zuge der Markierung und Codierung von Fehlern kann zu jedem Fehler ein Kommentar mit der richtigen Lösung oder sonstigen Hinweisen hinzugefügt werden.



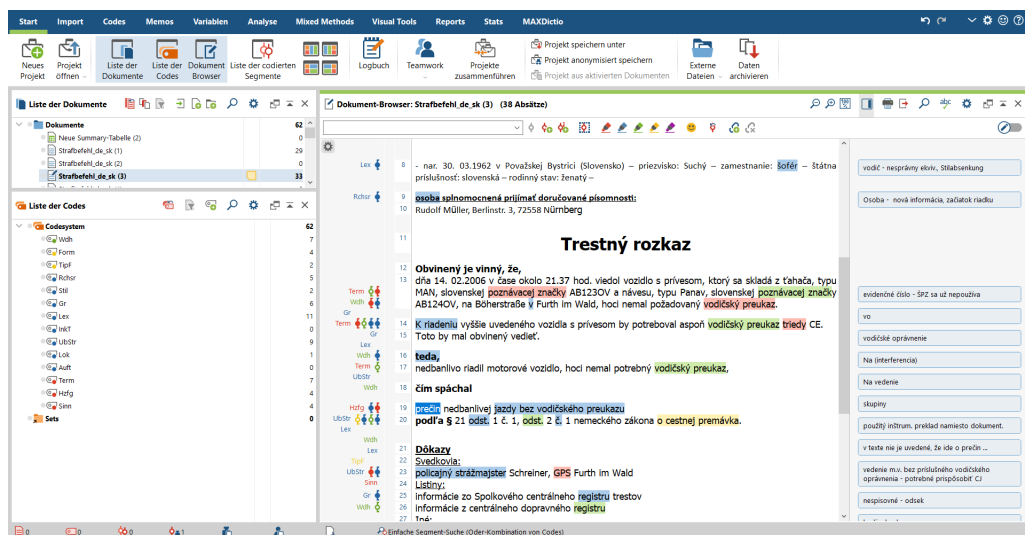


Abb. 3: Codieren von Fehlern mit Kommentar

Wie in Abbildung 3 zu sehen ist, werden während der Korrektur links die Codes angezeigt und rechts vom Text können die Kommentare eingblendet werden. Nach der Fertigstellung der Korrektur kann an den jeweiligen Studenten als Feedback eine PDF-Datei mit farbig markierten mangelhaften Textstellen und dazu eine Excel-Datei mit Kommentaren zu einzelnen codierten Fehlern (siehe Abb. 2) gesendet werden, oder die Studenten können den kostenfreien MAXQDA Reader nutzen.

Dokumentname	Codierte S...	Wdh	Form	TypF	Rchr	Stil	Gr	Lex	InkT	UbStr	Lok	Auft.	Term	HfG	Si
Strafbehf_de_sk (1)	29	1	3	1	3	1	2	6	0	5	0	0	4	1	
Strafbehf_de_sk (3)	33	6	1	1	2	1	4	5	0	4	1	0	3	3	

Abb. 4: Übersichtstabelle mit der Anzahl der codierten Fehler

Diese Software ermöglicht zudem die Gewichtung von Fehlern. Da sie aber lediglich die Zuordnung eines gemeinsamen Standardgewichts für alle Codes zulässt und die Änderung dieses Gewichts erst beim Codieren bei einem konkreten Code, und nicht bei allen gleich codierten Textstellen möglich ist, scheint es effektiver zu sein, die Gewichtung erst in der abschließenden Analyse vorzunehmen. Die automatische Auswertung und Bewertung von Übersetzungen in Form von Tabellen, wo auf eine sehr intuitive und benutzerfreundliche Weise neue Variablen und Gewichte hinzugefügt werden können, stellt eine wesentliche Erleichterung des gesamten Bewertungsverfahrens dar. In einer übersichtlichen Tabelle wird bei jedem Dokument die Anzahl der codierten Fehler gegliedert nach einzelnen Fehlerkategorien angezeigt (Abb. 4). Durch Definieren von Berechnungsformeln, in denen auch die Fehlergewichtung sehr einfach zu berücksichtigen ist, werden auf transparente Weise automatisch Bewertungsergebnisse für Lernergruppen jeder Größe gewonnen. Obwohl für die Anwendung dieser Software eine aufwendigere Vorbereitungsphase und gewisse Vorkenntnisse im Bereich der Excel-Berechnungsformeln erforderlich sind, erleichtern Instrumente dieser Art das

analytische Bewertungsverfahren im Übersetzungsunterricht wesentlich. Außerdem finden sie Anwendung auch bei der Sammlung von Materialien und Daten für linguistische und translatorische Untersuchungen.

### 3 Zum Schluss

Die Frage, ob Transparenz in der Evaluierung erreicht werden kann, ist bei den vielen Sprachkombinationen und den vielfältigen Übersetzungsaufträgen überaus schwer zu beantworten. Allgemein geltende Evaluierungskriterien sind anscheinend eine Utopie, doch in jedem Fall sind in den letzten Jahrzehnten große Bemühungen im Bereich der Entwicklung und Umsetzung von praktischen Qualitäts- und Evaluierungskriterien sowohl in Bildungseinrichtungen als auch in der translatorischen Praxis zu verzeichnen. In dem Artikel wurden die Konzepte der im Übersetzungsunterricht anwendbaren Bewertungskriterien sowie Modelle, die als Vorlage für die in der Übersetzungspraxis angewendete Kriterien dienen, dargestellt. Anschließend wurde festgestellt, dass der didaktische Evaluierungsprozess und die in der Praxis realisierte Evaluierung wesentliche Unterschiede besonders in Bezug auf ihre Ziele, Funktionen und Beteiligten aufweisen. Im Übersetzungsunterricht ist die Transparenz der Evaluierung von den Entscheidungen des Bewerbers abhängig, in der Übersetzungspraxis ist es der Übersetzungsdienstleister, der bezüglich der Transparenz das entscheidende Wort hat. Die ständige Entwicklung im Software-Bereich eröffnet zudem allerdings neue Möglichkeiten, die nicht nur auf die Zeitersparnis bei der Arbeit der Bewerter, sondern auch auf die Erhöhung der Bewertungstransparenz Einfluss nehmen.

### Literaturverzeichnis

- Hansen, Gyde (2007): *Ein Fehler ist ein Fehler ... oder? Der Bewertungsprozess in der Übersetzungsforschung*. In: Gerd Wotjak (Hg.): *Quo vadis Translatologie? Ein halbes Jahrhundert universitäre Ausbildung von Dolmetschern und Übersetzern in Leipzig*. 115-131. Berlin: Frank & Timme.
- Hansen, Gyde (2008a): *A classification of errors in translation and revision*. In: Martin Forstner, Hannelore Lee-Jahnke, & Peter A. Schmitt (Hgg.): *CIUTI-Forum 2008 Enhancing translation quality. Ways, means, methods*. 313-326. Berlin: Lang
- Hansen, Gyde (2008b): *Triangulierung im Unterricht: Übersetzungsprozesse, Eigenkorrektur und Fremdkorrektur*. *Informatologia*, 41(2), 105-115.
- Kautz, Ulrich (2002; 2000): *Handbuch Didaktik des Übersetzens und Dolmetschens*. München: IUDICIUM Verlag GmbH
- Kreuzer, Pia (2018): *Handreichung der Prüfungswerkstat: Kriterienraster*. [Abgerufen am 12.02.2021]. Verfügbar unter: [https://www.zq.uni-mainz.de/files/2018/08/6\\_Kriterienraster-erstellen.pdf](https://www.zq.uni-mainz.de/files/2018/08/6_Kriterienraster-erstellen.pdf).
- Kubánek, Michal; Molnár, Ondřej (2015): *Využití nástroje pro tvorbu korpusů studentských překladů*. In: Jitka Zehnalová et al. (Hgg.): *Kvalita hodnocení překladu: Modely a aplikace*. 137-165. Olomouc: Univerzita Palackého v Olomouci.
- Nord, Christiane (2010): *Fertigkeit Übersetzen. Ein Kurs zum Übersetzenlehren und -lernen*. Berlin: BDÜ Verlag.
- Nord, Christiane (2011): *Funktionsgerechtigkeit und Loayalität. Theorie und Didaktik des funktionalen Übersetzens*. Berlin: Frank & Timme Verlag.
- Ottmann, Angelika; Canfora, Carmen (2017): *Übersetzungen richtig bewerten. technische kommunikation*, 06/17, 35-40. [Abgerufen am 12.02.2021]. Verfügbar unter: [https://risikoscouts.de/wp-content/uploads/2017/11/tk617\\_Artikel-Ottmann-Canfora\\_%C3%9Cbersetzungsevaluierung.pdf](https://risikoscouts.de/wp-content/uploads/2017/11/tk617_Artikel-Ottmann-Canfora_%C3%9Cbersetzungsevaluierung.pdf).
- Stolze, Rade Gundis (1999). *Die Fachübersetzung. Eine Einführung*. Tübingen: Gunter Narr Verlag.
- Zehnalová, Jitka et al. (2015): *Kvalita hodnocení překladu: Modely a aplikace*. Olomouc: Univerzita Palackého v Olomouci, 2015.

www1: ISO 17100. Anforderungen an Übersetzungsdienstleistungen. [Abgerufen am 12.02.2021]. Verfügbar unter: <<http://qualitätsstandard.iso17100.com>>.

www2: Evaluierung von Übersetzungen nach SAE-J2450. [Abgerufen am 12.02.2021]. Verfügbar unter: <http://www.atesman.info/wp-content/uploads/2015/10/SAE-Qualit%C3%A4tsmetrik.pdf>.

www3. Maxqda. [Abgerufen am 12.02.2021]. Verfügbar unter: <https://www.maxqda.de/>.

## Annotation

### **Transparency in the evaluation of translations**

*Blanka Jenčíková*

The issue of transparency when assessing translation quality is often discussed, both in practice and in the context of university translator training. However, there is a lack of systematic concepts in this area and the question arises as to whether a methodology can be established for this naturally subjective activity that would lead to objective and transparent results. In this article both didactic and practice-oriented concepts of evaluation criteria are presented, in the further part a comparison of the evaluation process in translation lessons and in translation practice with a focus on their transparency follows. Finally, the application of software for more efficient analytical performance evaluation is presented.

*Keywords:* transparency, performance evaluation, translation quality, translation errors, translation criteria

Mgr. Blanka Jenčíková  
Prešovská univerzita v Prešove  
Ul. 17. novembra 15  
SK-080 01 Prešov  
[blanka.jencikova@unipo.sk](mailto:blanka.jencikova@unipo.sk)