

Masterarbeit über das Thema

## **Post-Editese als verschärfte Form der Translationese?**

**Eine Korpusanalyse zu *Simplification* und *Interference* in  
posteditierten Texten**

Marlena Conze

## Inhaltsverzeichnis

1	Einleitung.....	5
1.1	Herleitung.....	5
1.2	Ziel der Arbeit.....	7
1.3	Aufbau der Arbeit.....	8
2	Hintergrund.....	9
2.1	Post-Editing.....	9
2.1.1	Definition.....	9
2.1.2	Arten von Post-Editing.....	11
2.1.3	Produktivitäts- und Kostenvorteile.....	14
2.1.4	Qualitätsanalysen.....	16
2.1.5	Einflussfaktoren auf die Qualität des MÜ-Outputs.....	18
2.1.6	Präferenz- und Akzeptanzstudien.....	19
2.1.7	Einstellungen von Übersetzern, Sprachdienstleistern und Kunden.....	22
2.1.8	Weitere Forschungsrichtungen.....	25
2.2	Translationese.....	26
2.2.1	Definition.....	27
2.2.2	Übersetzungsuniversalien nach Baker.....	27
2.2.3	Interference nach Toury bzw. Teich.....	30
2.2.4	Forschungsstand.....	31
2.2.5	Neuere Forschungsrichtungen.....	35
2.3	Post-Editese.....	37
2.3.1	Definition.....	37
2.3.2	Forschungsstand.....	38
2.4	Zwischenfazit und Forschungsdesiderat.....	42
3	Daten und Methoden.....	44
3.1	Korpus.....	44
3.1.1	Beschreibung der Korpustexte.....	44
3.1.2	Erstellung des Korpus.....	46
3.1.3	Begründung der Auswahl des Korpus.....	47
3.1.4	Aufbereitung der Korpustexte für die maschinelle Auswertung.....	48
3.2	Analysemethoden.....	49
3.2.1	Simplification.....	50

3.2.2	Interference .....	56
4	Ergebnisse der eigenen Korpusanalyse und Einordnung der Ergebnisse.....	66
4.1	Simplification.....	66
4.1.1	Lexikalische Diversität.....	66
4.1.2	Lexikalische Dichte.....	69
4.1.3	Einordnung der Ergebnisse zu Simplification .....	72
4.2	Interference .....	80
4.2.1	Verhältnis der Verhältnisse von nominalen und verbalen Wortarten zwischen AT und ZT.....	80
4.2.2	Verhältnis der Satzlängen zwischen AT und ZT .....	83
4.2.3	Einordnung der Ergebnisse zu Interference.....	85
5	Diskussion.....	94
6	Fazit und Ausblick .....	97
7	Literaturverzeichnis.....	100

## Abbildungsverzeichnis

Abbildung 1: Korpusaufbau. ....	45
Abbildung 2: Permutationsschema für die Datenerhebung (Čulo et al. 2017, 189).....	46
Abbildung 3: Berechnungsmethode der Durchschnitte.....	49
Abbildung 4: Sketch Engine Benutzeroberfläche, Ansicht Corpus details .....	52
Abbildung 5: Sketch Engine Benutzeroberfläche, Wordlist-Funktion (advanced) .....	53
Abbildung 6: Interpretationshilfe zum Verhältnis der Nominal-Verbal-Verhältnisse .....	58
Abbildung 7: Verschobene Satzalignierung nach Satzaufspaltung .....	62
Abbildung 8: TTR nach Übersetzungsmodus für alle Texte.....	66
Abbildung 9: TTR nach Übersetzungsmodus für alle med. Texte. ....	67
Abbildung 10: TTR nach Übersetzungsmodus für alle techn. Texte .....	67
Abbildung 11: TTR nach Übersetzungsmodus für alle Texte, aufgeschlüsselt.....	68
Abbildung 12: Lexikalische Dichte nach Übersetzungsmodus für alle Texte .....	69
Abbildung 13: Lexikalische Dichte nach Übersetzungsmodus für alle med. Texte.....	70
Abbildung 14: Lexikalische Dichte nach Übersetzungsmodus für alle techn. Texte.....	70
Abbildung 15: Lexikalische Dichte nach Übersetzungsmodus für alle Texte, aufgeschlüsselt.....	71
Abbildung 16: Verhältnis der Nominal-Verbal-Verhältnisse nach Übersetzungsmodus für alle Texte.....	80
Abbildung 17: Verhältnis der Nominal-Verbal-Verhältnisse nach Übersetzungsmodus für alle med. Texte.....	81
Abbildung 18: Verhältnis der Nominal-Verbal-Verhältnisse nach Übersetzungsmodus für alle techn. Texte .....	82
Abbildung 19: Verhältnis der Nominal-Verbal-Verhältnisse nach Übersetzungsmodus für alle Texte, aufgeschlüsselt .....	82
Abbildung 20: Satzlängenverhältnis nach Übersetzungsmodus für alle Texte .....	83
Abbildung 21: Satzlängenverhältnis nach Übersetzungsmodus für alle med. Texte .....	84
Abbildung 22: Satzlängenverhältnis nach Übersetzungsmodus für alle techn. Texte.....	84
Abbildung 23: Satzlängenverhältnis nach Übersetzungsmodus für alle Texte, aufgeschlüsselt.....	85
Abbildung 24: Satzlängenverhältnis nach Satz für das Korpus MED2_P23_FPE.....	90

## 1 Einleitung

Die vorliegende Arbeit beschäftigt sich mit der Beschreibung und dem Nachweis des Phänomens Post-Editese in einem Korpus aus technischen und medizinischen Texten. In diesem Zusammenhang werden auch Aspekte der angrenzenden Forschungsfelder Post-Editing und Translationese thematisiert. Im Rahmen einer Korpusanalyse wird untersucht, ob die Übersetzungsuniversalien *Simplification* und *Interference*, wie in früheren Beiträgen konstatiert, besonders stark in posteditierten Texten vorzufinden sind. Nach einer Herleitung des Themas (Abschnitt 1.1) wird das Ziel der Arbeit formuliert (Abschnitt 1.2), sowie der Aufbau der Arbeit geschildert (Abschnitt 1.3).

### 1.1 Herleitung

Sowohl Übersetzern<sup>1</sup> als auch Laien dürfte bekannt sein, dass die Maschinelle Übersetzung („MÜ“) in den letzten Jahren zunehmend an Relevanz in der Berufs- und Lebenswelt gewonnen hat: Nahezu jeder Internetnutzer kennt *Google Translate* oder *DeepL* und auch in der Übersetzungsbranche wird seit Jahren mehr und mehr auf MÜ gesetzt (Daems et al. 2017, 89).

Die maschinelle Übersetzung bietet eine Chance, dem durch die Globalisierung angefeuerten, wachsenden Bedürfnis nach Informationen und globaler Kommunikation – sprich dem steigenden Übersetzungsbedarf, der heute durch menschliche Übersetzer allein nicht mehr zu bewältigen ist – gerecht zu werden (Poulis/Kolovratnik 2012, 1; Carl et al. 2015, 146).

„Neuronale Maschinelle Übersetzung“ („NMÜ“) ist dabei das Wort der Stunde: Dieses neueste der Übersetzungsparadigmen, das auch den maschinellen Übersetzungssystemen von *Google Translate* und *DeepL* (Le/Schuster 2016; DeepL 2021) zugrunde liegt, hat in den letzten Jahren signifikante Qualitäts- und Produktivitätssteigerungen möglich gemacht (Doğru et al. 2018; SDL 2020) (siehe auch Abschnitte 2.1.3 und 2.1.4).

---

<sup>1</sup> Im Interesse einer besseren Lesbarkeit wird in dieser Arbeit nicht ausdrücklich in geschlechtsspezifischen Personenbezeichnungen differenziert. Die gewählte männliche Form schließt Personen aller Geschlechter gleichberechtigt ein.

Wenn maschinell erstellte Übersetzungen vor der Veröffentlichung von Menschen korrekturgelesen werden, ist von „Post-Editing“ bzw. „Posteditieren“ („PE“) die Rede (ISO:17100:2015). In den meisten Fällen und insbesondere bei Texten, die veröffentlicht werden sollen, ist ein Post-Editing trotz der genannten Qualitätssteigerungen weiterhin notwendig, um die geforderte Qualität sicherzustellen (Koponen 2016, 132). Post-Editing gibt es seit es leistungsfähige maschinelle Übersetzungssysteme (auch „MÜ-Engines“) gibt (Koehn 2010, 537). Es ist also keineswegs ein neues Phänomen, gewiss aber eines, das in den letzten Jahren zunehmend an Bedeutung gewonnen hat (Koponen 2016, 133). Heute ist Post-Editing ein integraler Bestandteil der Übersetzungslandschaft (SDL 2020).

Während bereits zahlreiche Studien zur Qualität von humanübersetzten und posteditierten Texten, sowie unzählige Produktivitätsstudien durchgeführt wurden, hat die Frage, inwiefern sich posteditierte Texte in Bezug auf verschiedene linguistische Eigenschaften von humanübersetzten Texten unterscheiden, und inwiefern diese Eigenschaften auf die maschinelle Übersetzung zurückzuführen sind, bisher wenig Aufmerksamkeit erfahren (Čulo et al. 2017, 183/184). Die Erforschung der Unterschiede zwischen humanübersetzten und posteditierten Texten ist Gegenstand der Post-Editese-Forschung. Der Begriff Post-Editese (nach Daems et al. 2017) wurde in Anlehnung an den Begriff Translationese (nach Gellerstam 1986) geprägt. Die Translationese-Forschung beschäftigt sich mit den Unterschieden zwischen übersetzten Texten und Originaltexten, weshalb Translationese auch als „language of translation“ verstanden werden kann (Lapshinova-Koltunski 2013, 78). Anders als das Phänomen Post-Editese ist Translationese bereits recht gut erforscht. So hat man etwa beobachtet, dass es sogenannte Übersetzungsuniversalien (Baker 1995; Teich 2003; Toury 2012) gibt, die – teilweise unabhängig vom Sprachpaar – in übersetzten Texten nachgewiesen werden konnten. Übersetzungsuniversalien sind also gewissermaßen Merkmale der Translationese.

Ein entscheidender Beitrag zur Post-Editese-Forschung stammt von Antonio Toral (2019) und trägt den Titel „Post-Editese: An Exacerbated Translationese“. Darin macht er sich die Erkenntnisse der Translationese-Forschung zu Nutze, indem er versucht, sie auf das angrenzende Forschungsfeld der Post-Editese zu übertragen. Er prüft im Rahmen einer umfassenden Korpusanalyse, ob sich *Simplification*,

*Interference* und *Normalization*, drei bekannte Übersetzungsuniversalien bzw. Merkmale der Translationese, auch in posteditierten Texten nachweisen lassen. Damit geht er der Frage nach, ob und inwiefern sich Translationese von Post-Editese unterscheidet.

Sein Korpus umfasst 2213 Satzpaare, verschiedene Übersetzungsmodi und fünf verschiedene Sprachpaare. Bei circa einem Viertel der Satzpaare handelt es sich um Untertitel und deren Übersetzungen, der Rest stammt aus dem Themengebiet Nachrichten. Anhand seiner Korpusanalyse stellt er fest, dass posteditierte Texte die oben genannten typischen Übersetzungsuniversalien tatsächlich ebenfalls aufweisen, und zwar sogar in höherem Maße als humanübersetzte Texte. Toral schlussfolgert also, Post-Editese sei eine verschärfte Form der Translationese. Die Erforschung dieser Post-Editese hat jenseits von Torals Beitrag bisher wenig Aufmerksamkeit erfahren.

Prognosen und aktuelle Entwicklungen lassen vermuten, dass Post-Editing auch künftig eine immer größere Rolle spielen und die klassische Humanübersetzung („HÜ“) zumindest in einigen Bereichen (weiter) verdrängen wird. Folglich wird ein immer größerer Anteil der Übersetzungsprodukte posteditiert und nicht humanübersetzt sein (Čulo et al. 2017, 198). Allein deshalb lohnt sich eine genauere Untersuchung dieser Art von Texten hinsichtlich möglicher Unterschiede zu Humanübersetzungen.

## **1.2 Ziel der Arbeit**

Die vorliegende Arbeit knüpft an Torals Studie an und ist in der Post-Editese-Forschung zu verorten. Ihr Ziel ist es zu prüfen, ob sich Torals (2019) Hypothese, Post-Editese sei eine verschärfte Form der Translationese, anhand der Ergebnisse eines weiteren Datensatzes – einem Korpus aus Fachtexten aus dem medizinischen und technischen Bereich – bestätigen lässt. Dazu wird eine Korpusanalyse durchgeführt, die sich im Wesentlichen an der von Toral (2019) und in geringerem Maße an der von Lapshinova-Koltunski (2015) verwendeten Methodik orientiert: Der Nachweis von Simplification soll über die lexikalische Diversität (Type-Token-Ratio) und die lexikalische Dichte erfolgen. Das Verhältnis der Verhältnisse von nominalen und verbalen Wortarten in Ausgangs- und Zieltext (orientiert an

Lapshinova-Koltunski 2015) sowie das Verhältnis der Satzlängen in Ausgangs- und Zieltext wird zum Nachweis von Interference herangezogen. Sollte es mithilfe dieser Methoden gelingen, eine stärkere Ausprägung von Simplification und Interference in posteditierten Texten im Vergleich zu Humanübersetzungen nachzuweisen, würde dies Torals Hypothese bestätigen, dass das Phänomen Post-Editese als verschärfte Form der Translationese begriffen werden kann.

### **1.3 Aufbau der Arbeit**

Der Aufbau der Arbeit orientiert sich an dem linguistischen Sechsschritt. Im nächsten Abschnitt wird zunächst in die Thematik eingeführt, beginnend mit verschiedenen Aspekten von Post-Editing (Abschnitt 2.1). Danach werden die Forschungszweige Translationese (Abschnitt 2.2) und Post-Editese (Abschnitt 2.3) nacheinander beleuchtet, da beide, bzw. ihre Schnittmenge, den theoretischen Rahmen für die Korpusanalyse bilden. In diesem Zusammenhang werden auch die für diese Arbeit relevanten Übersetzungsuniversalien Simplification und Interference (Abschnitt 2.2.2 und 2.2.3) näher erläutert. In einem Zwischenfazit werden die gewonnenen Erkenntnisse zusammengefasst (Abschnitt 2.4).

Im dritten Abschnitt wird das verwendete Korpus (Abschnitt 3.1) und die Methodik der eigenen Korpusanalyse (Abschnitt 3.2) vorgestellt. Der vierte Abschnitt befasst sich mit der Auswertung und Interpretation der Ergebnisse der eigenen Analyse und ist zweigeteilt: Zunächst werden die Ergebnisse der Untersuchungsaspekte, die Simplification nachweisen sollen, vorgestellt (Abschnitte 4.1.1 und 4.1.2) und anschließend eingeordnet (Abschnitt 4.1.3). Danach folgt der gleiche Ablauf für die Ergebnisse der Untersuchungsaspekte, die zum Nachweis von Interference gewählt wurden (Abschnitt 4.2). Im Rahmen der Einordnung der Ergebnisse werden mögliche Schwachstellen diskutiert und Besonderheiten in den Ergebnissen hervorgehoben. Teilweise werden Erklärungsversuche geliefert und ergänzende Forschungsrichtungen aufgezeigt. Im fünften Abschnitt werden allgemeine Probleme im Aufbau der eigenen Analyse oder bei der Auswertung der Daten, d. h. solche, die für alle Untersuchungsaspekte relevant sind, angesprochen. Schließlich werden im letzten Abschnitt die Erkenntnisse der Arbeit noch einmal zusammengefasst und Perspektiven für weitere Untersuchungen resümiert.



## 2 Hintergrund

Der Theorieteil dieser Arbeit gliedert sich in vier Teile. Im ersten Teil werden verschiedene Aspekte des Themenkomplexes Post-Editing, wie Produktivität, Qualität und Akzeptanz, thematisiert, bevor es im zweiten Teil um das Phänomen Translationese geht. In diesem Zusammenhang werden auch die Übersetzungsuniversalien, insbesondere Simplification und Interference im Detail erläutert. Der dritte Teil ist dem Phänomen Post-Editese, das im Rahmen der vorliegenden Arbeit untersucht wird, gewidmet. Schließlich folgt ein Zwischenfazit, in dem auch das Forschungsdesiderat formuliert wird.

### 2.1 Post-Editing

Neben der Definition (Abschnitt 2.1.1) und den verschiedenen Arten von Post-Editing (Abschnitt 2.1.2) wird in diesem Abschnitt der aktuelle Forschungsstand zur Produktivität (Abschnitt 2.1.3) von Post-Editing sowie zur erreichbaren Qualität (Abschnitt 2.1.4) von posteditierten Texten im Vergleich zu humanübersetzten Texten vorgestellt. Nach einem Exkurs zu den Faktoren, die die Qualität des MÜ-Outputs beeinflussen (Abschnitt 2.1.5), folgt die Vorstellung einiger Präferenz- und Akzeptanzstudien, die mit Zieltextadressaten durchgeführt wurden (Abschnitt 2.1.6). Anschließend werden die Einstellungen von Übersetzern, Sprachdienstleistern und Kunden zu Post-Editing dargelegt (Abschnitt 2.1.7). Zum Schluss werden weitere Forschungsrichtungen aus dem Forschungsfeld Post-Editing aufgezeigt (Abschnitt 2.1.8).

#### 2.1.1 Definition

Für den Begriff Post-Editing („PE“, auch „PEMT“) wurden über die Jahre eine Vielzahl verschiedener Definitionen vorgeschlagen. Wagner definiert Post-Editing 1985 als „[the] correction of a *pre-translated text* rather than translation 'from scratch'“ (Wagner 1985, 1, eigene Hervorhebung). Veale und Way (1997) hingegen verstehen darunter „the correction of *machine translation output* by human linguists/editors“ (Veale/Way 1997 zitiert nach O'Brien 2010a, 1, eigene Hervorhebung) und engen somit „pre-translated“ (Wagner 1985, 1) auf maschinell

übersetzte Texte ein. Der Begriff Post-Editing ist demnach insofern vom Korrekturlesen abzugrenzen, als mit letzterem das Korrekturlesen von nicht näher definierten Texten gemeint ist (ISO 17100:2015). Post-Editing ist gemäß dieser Definition also eine Sonderform des Korrekturlesens.

Der Begriff „correction“ bzw. ‚Korrekturlesen‘ ist in diesem Zusammenhang allerdings nicht unumstritten. Laurian konstatiert etwa:

„Post-editing is not revision, nor correction, nor rewriting. It is a new way of considering a text, a new way of working on it, for a new aim” (Laurian 1984, 237).

Modernere Definitionen, wie die von Carl et al., eröffnen neue Blickwinkel auf diese Art der Textbearbeitung:

„Post-editing is the correction of machine-generated translation output to ensure it meets a level of quality negotiated in advance between client and post editor” (Carl et al. 2015, 146).

Carl et al. scheuen sich also nicht, von ‚Korrektur‘ zu sprechen und verweisen zusätzlich auf die angestrebte Qualität – wie auch die Definition von Vashee (2011), in der allerdings der Begriff „correction“ vermieden wird: „PEMT [is] the effort and linguistic work done to raise raw MT to target quality levels”.

Ein weiterer Blickwinkel wird durch die Definition des TAUS-Report (*Translation Automation User Society*) von 2010 aufgezeigt, die zusätzlich eine Information über den Arbeitsaufwand, der zur Erreichung der gewünschten Qualität notwendig ist, enthält: „post-editing is the process of improving a machine-generated translation with a minimum of manual labor“ (TAUS 2010 zitiert nach O’Brien 2010a, 1).

O’Brien (2011, 197) zeigt mit ihrer Definition eine weitere Perspektive auf, indem sie nicht nur auf bestimmte geltende Qualitätskriterien und Richtlinien hinweist, sondern spezifiziert, dass das Post-Editing von einem „human translator“, also einem Humanübersetzer durchgeführt werden sollte:

„Post-editing is the correction of raw machine translated output by a human translator according to specific guidelines and quality criteria”.

Die wohl ‚offiziellste‘ Definition liefert die Internationale Norm ISO 17100:2015, in der die Anforderungen an Übersetzungsdienstleistungen festgehalten sind.

Post-Editing ist hier schlicht als „[the] edit[ing] and correct[ion] [of] machine translation output“ definiert (ebd., 2).

Neben den Abweichungen hinsichtlich der Art des zugrundeliegenden Textes („pre-translated text“ oder „machine translated output“) und der Frage, ob es sich beim Post-Editing um eine „Korrektur“ handelt, setzen die verschiedenen Definitionsversuche also auch unterschiedliche Schwerpunkte bzw. nähern sich dem Begriff aus unterschiedlichen Perspektiven, z. B. über die angestrebte Qualität oder den notwendigen Arbeitsaufwand, an. Dies sind Aspekte, die in den folgenden Abschnitten näher beleuchtet werden. Da es sich bei der Definition der ISO 17100:2015 um die ‚offiziellste‘ und ‚neutralste‘ handelt, ist diese Definition für die vorliegende Arbeit maßgebend.

### **2.1.2 Arten von Post-Editing**

Generell unterscheidet man zwei Hauptarten bzw. Grade von Post-Editing. Das sogenannte *Light Post-Editing* („LPE“) einerseits und *Full Post-Editing* („FPE“) andererseits. Zur Unterscheidung wurden zahlreiche Richtlinienkataloge von verschiedenen Autoren erarbeitet.<sup>2</sup> Darin werden – je nach Post-Editing-Grad – Handlungsempfehlungen für den Umgang mit verschiedenen Fehlerkategorien dargelegt. Im TAUS-Report von 2016 wird stattdessen eine Unterscheidung anhand der erwarteten Qualität des Endproduktes vorgeschlagen (Cadwell/Hu 2016, 347; TAUS 2016, 16). Aber auch hier werden Handlungsempfehlungen und Richtlinien präsentiert, um die jeweils gewünschte Qualität zu erreichen. Die folgende Darstellung stützt sich im Wesentlichen auf den TAUS-Report (2016, 16ff.).

Beim *Light Post-Editing* (auch „fast“, „rapid“, „minimal“ oder „gist post-editing“ (Laurian 1984, 237; Allen 2003, 304; Carl et al. 2015, 147)) wird ein Text erwartet, dessen Qualität „good enough“, d. h. ‚ausreichend‘ ist. In diesem Zusammenhang ist auch regelmäßig von einer ‚zweckgemäßen‘ („fit for purpose“) Qualität die Rede. Es werden nur essenzielle Änderungen vorgenommen, sodass das Endprodukt „comprehensible“ und „accurate“ ist und keine groben Fehler enthält. Jedoch ist zu erwarten, dass die Syntax untypisch, die Terminologie nicht immer

---

<sup>2</sup> Für einen Vergleich der verschiedenen Definitionen und Richtlinien siehe Cadwell/Hu 2016.

korrekt und der Text stilistisch nicht ausgefeilt ist oder klingt, als sei er von einem Computer generiert worden. Senez (1998, 5) bezeichnet dies auch als „awkwardness in style“. Die Devise für Posteditoren beim Light Post-Editing bzw. zur Erreichung einer „good enough quality“ lautet also, nur Änderungen vorzunehmen, die zwingend nötig sind, wie die Korrektur von Rechtschreibfehlern, Auslassungen, inhaltlichen Fehlern und unverständlichen Textpassagen. Ansonsten soll das MÜ-Output so weit wie möglich unverändert gelassen werden. Stilistische Änderungen sind explizit nicht vorzunehmen.

Ziel des Full Post-Editing (auch „maximal“, „conventional“ oder „heavy post-editing“ (Laurian 1984, 237; Allen 2003, 303; Cadwell/Hu 2016, 345) hingegen ist ein Text von „publishable“, d. h. von ‚veröffentlichungsreifer‘ Qualität. Das Endprodukt hat also den Anspruch, verständlich, präzise und stilistisch in Ordnung<sup>3</sup>, sowie im Bereich der Syntax unauffällig zu sein. Darüber hinaus soll auch auf eine korrekte Terminologie, Zeichensetzung und Formatierung geachtet werden. Auch bei dieser Form des Post-Editing gilt es, das unbearbeitete MÜ-Output wann immer möglich unverändert zu lassen – der Text soll nicht im großen Stil (etwa nach den Präferenzen des jeweiligen Posteditors) umformuliert werden. Ein „over-editing“ (TAUS 2016, 6) soll also – wie auch beim Light Post-Editing – vermieden werden. Aufgrund des größeren Umfangs der erlaubten bzw. empfohlenen Korrekturen erfordert FPE in der Regel mehr Zeit als LPE, jedoch weniger als klassisches Humanübersetzen (Wagner 1985, 1) (siehe auch Abschnitt 2.1.3).

Mitunter herrscht in der Fachliteratur Uneinigkeit darüber, welches Qualitätsniveau bei FPE angestrebt wird. Laut dem TAUS-Report von 2016 ist das Ziel eine Qualität die lediglich „*similar to* ‚high-quality human translation and revision‘“ (TAUS 2016, 16, eigene Hervorhebung) ist. Im TAUS-Report von 2013 heißt es dazu auch: „hinsichtlich des Stils [kann] die Qualität eines muttersprachigen Humanübersetzers nicht erreicht werden“. Eine abweichende Meinung vertreten Wagner und DePalma: Ziel beim FPE sei ein Text, der „*indistinguishable from what a good human translator can produce*“ ist (DePalma 2013 eigene Hervorhebung) und dessen Qualität „*identical to that of a conventional translation*“ (Wagner 1985, 1) ist.

---

<sup>3</sup> Je nach Definition sind stilistische Probleme mitunter auch beim FPE zu ignorieren, siehe dazu Cadwell & Hu 2016.

Dieser Arbeit liegen die Definitionen des TAUS-Report zugrunde. Dementsprechend wird bei FPE-Texten eine ‚vergleichbare‘ Qualität erwartet.

„Partial Post-Editing“ deckt das Spektrum zwischen FPE und LPE ab. Je nach Qualitätsanspruch des Kunden, zeitlichen Aspekten und anderen Faktoren (s. u.) können für einen bestimmten Auftrag individuelle Absprachen getroffen werden (Aziz et al. 2012 zitiert nach Carl et al. 2015, 147).

Eine weitere Sonderform des Post-Editing ist das sogenannte „monolingual post-editing“ (ebd.). In diesem Fall arbeitet der Posteditor ausschließlich mit dem Zieltext („ZT“) – entweder weil er die Ausgangssprache nicht beherrscht (Koehn 2010, 537) oder ihm der Ausgangstext („AT“) zum Abgleich nicht zur Verfügung steht (Carl et al. 2015, 147).

Entscheidend für die Wahl zwischen den verschiedenen Formen von Post-Editing und dem klassischen Humanübersetzen sind neben zeitlichen Aspekten und dem Qualitätsanspruch auch der Verwendungszweck, also ob der Text publiziert werden soll oder nur einen groben Eindruck („the gist“) über den Inhalt vermitteln soll. Darüber hinaus spielt die Kurz- bzw. Langlebigkeit des Zieltexts, sowie preisliche und datenschutzrechtliche Überlegungen,<sup>4</sup> der Textumfang, und letztlich der Wunsch des Kunden eine Rolle (ebd., 147/148; Allen 2003, 301; Specia 2011b zitiert nach Čulo et al. 2017, 186). Bei einem hohen Qualitätsanspruch und bei einer vorgesehenen Veröffentlichung des Textes beispielsweise sollte ein Full Post-Editing von einem qualifizierten Übersetzer mit Kenntnis beider Sprachen durchgeführt werden (Koehn 2009, 1), während für einen kurzlebigen Text, der nur unternehmensintern verwendet, d. h. nicht publiziert wird, und nur einen groben Überblick über den Inhalt geben soll, Light Post-Editing oder sogar eine unbearbeitete maschinelle Übersetzung ausreichend sein kann (Aymerich 2005, 7; Fiederer/O'Brien 2009, 52/53). Laurian fasst hierzu zusammen:

„Raw translations being usable, as they are quite often for rapid information needs, it is important to draw the limits between a style adequate for rapid

---

<sup>4</sup> Für weitere Informationen zum Thema Datenschutz bei Maschinellem Übersetzen siehe Kamocki/O'Regan 2016.

information, and an elegant, high quality style such as required for information large dissemination [sic!]“ (Laurian 1984, 236).

Allerdings soll an dieser Stelle nicht unerwähnt bleiben, dass die Toleranz für Texte geringerer Qualität, etwa stilistisch weniger gelungene Texte, in der Geschäftswelt – in Abwägung mit anderen Faktoren – gestiegen zu sein scheint:

„An [sic!] MT program can generate the gist of an email or other message and allow a rapid reply with a reasonable degree of accuracy. This has become more acceptable in business, where the tolerance for low-quality texts has risen [...] and users are more willing to concede some quality to achieve expediency in communication“ (Allied Business Intelligence 2002 zitiert nach Bowker 2009, 144).

### **2.1.3 Produktivitäts- und Kostenvorteile**

Der am häufigsten genannte Vorteil von Post-Editing im Vergleich zum klassischen Übersetzen ist sicherlich die höhere Produktivität und damit die geringeren Kosten, die mit diesem Übersetzungsmodus erzielt werden können (Guerberof 2009, 8; Carl et al. 2015, 146; Koponen 2016, 142; Čulo et al. 2017, 183 u.v.m.).

Konkret zeigen Studien, dass die schnellste und billigste Produktionsart die unbearbeitete MÜ ist, gefolgt von LPE, FPE und schließlich HÜ (Bowker 2009; Bowker/Buitrago Ciro 2015).

Der Zeitgewinn beim FPE gegenüber der HÜ kann laut Studien bei 30 bis 40 Prozent liegen (Aziz et al. 2012, 3982; Bowker/Buitrago Ciro 2015, 27), aber auch deutlich höher: Aymerich (2005, 2) zum Beispiel beziffert den Produktivitätsgewinn auf 50 bis 100 Prozent und vereinzelt konnten sogar noch höhere Werte erzielt werden, wie bei Vasconcellos und Léon (1985), die feststellten, dass zur Produktion eines veröffentlichungsreifen Textes mittels Post-Editing nur die Hälfte bis ein Drittel der Zeit im Vergleich zum klassischen Humanübersetzen benötigt wird. In diesen Studien beliefen sich die aus dem Produktivitätsgewinn resultierenden Kostenersparnisse auf 30 bis 55 Prozent (Aymerich 2005, 2; Bowker 2009, 136; Bowker/Buitrago Ciro 2015, 27).

Noch höhere Zeitgewinne und Kostenersparnisse wurden beim Vergleich von HÜ-Texten mit LPE-Texten festgestellt (ebd.) oder wenn das verwendete MÜ-

System proprietär ist oder speziell (etwa für ein bestimmtes Unternehmen) entwickelt wurde (vgl. Plitt/Masselot 2010).

Bei der Interpretation dieser Zahlen zur Produktivität ist jedoch zu bedenken, dass diese nur unter bestimmten Bedingungen zu erzielen sind (Garcia 2011, 228). Produktivitätsstudien dieser Art werden häufig unter Laborbedingungen durchgeführt und sind unter Umständen nicht ohne Weiteres auf die Praxis übertragbar (ebd.). Die Annahme, dass durch Post-Editing grundsätzlich Produktivitätssteigerungen zumindest möglich sind, ist zwar weitestgehend unstrittig, vereinzelt konnten in Studien aber auch keine oder keine signifikanten Produktivitätssteigerungen beobachtet werden. Post-Editing kann durchaus auch nur sehr geringfügig schneller und kostengünstiger oder sogar zeitaufwändiger und teurer sein als Humanübersetzen. Dies war beispielsweise in Studien von Carl et al. (2011) und Garcia (2010) der Fall, was laut Carl et al. mit der Tatsache zusammenhängen könnte, dass die Teilnehmer in beiden Studien angehende Übersetzer ohne Post-Editing-Erfahrung waren (Garcia 2010, 7; Carl et al. 2011, 137/138). Auch Guerberof (2009, 7) stellt fest, dass die Erfahrung der eingesetzten Übersetzer einen Einfluss auf ihre Produktivität beim Posteditieren hat (vgl. auch Kirchhoff et al. 2011, 478).

Zudem hat die Qualität des unbearbeiteten MÜ-Outputs einen entscheidenden Einfluss darauf, ob tatsächlich Produktivitätssteigerungen und Kostenersparnisse zu erzielen sind (Aziz et al. 2012, 3986; Čulo et al. 2017, 186), oder ob eine ‚translation from scratch‘ schneller und kostengünstiger zu einem vergleichbaren Ergebnis geführt hätte (Specia 2011a, 73; Gaspari et al. 2014, 70). Denn es ist in erster Linie die Qualität des MÜ-Outputs, die entscheidet, wie viel Post-Editing betrieben werden muss, um die gewünschte Qualität zu erreichen. Die Qualität des MÜ-Outputs wiederum hängt von einer Vielzahl von Faktoren ab, die in Abschnitt 2.1.5 näher erläutert werden.

Trotz der nicht ganz eindeutigen Datenlage lässt sich festhalten, dass durchaus leichte bis starke Produktivitätssteigerungen zu erwarten sind, wenn man statt auf Humanübersetzen auf Post-Editing setzt (Koponen 2016, 136). Wie hoch die Produktivitätszuwächse tatsächlich ausfallen hängt unter anderem entscheidend von bestimmten Bedingungen, wie der Qualität des MÜ-Outputs und der Erfahrung der Posteditoren, ab.

#### 2.1.4 Qualitätsanalysen

Bei allen Versprechungen von Produktivitätssteigerungen und Kosteneinsparungen darf der Aspekt der Qualität nicht in den Hintergrund geraten. Senez (1998, 6) formulierte das wie folgt: „The post-editing of machine translation seeks to strike the right balance between time, quality, and available capacity.“ Die Definition von Post-Editing nach Carl et al. (2015, 146) enthält – wie bereits erwähnt – sogar den Hinweis, dass durch das Post-Editing „a level of quality negotiated in advance between client and posteditor“ erreicht werden soll.

Die Qualität von posteditierten Texten gegenüber humanübersetzten Texten wurde in zahlreichen Studien untersucht. Wie bzw. mithilfe welcher Methoden die Qualität bewertet wird, unterscheidet sich von Studie zu Studie.

In einer Studie von Fiederer und O'Brien (2009) wurden die Texte von Sprachexperten nach drei Kriterien, nämlich „accuracy“, „clarity“ und „style“ bewertet. Sie stellten fest, dass die posteditierten Texte in der Kategorie „accuracy“ besser abschnitten als humanübersetzte Texte, aber schlechter in der Kategorie „style“. Für die Kategorie „clarity“ fielen die Ergebnisse in etwa gleich aus, sodass Fiederer und O'Brien auch insgesamt zu dem Urteil kommen, dass es zwischen posteditierten und humanübersetzten Texten keine signifikanten Qualitätsunterschiede gibt (2009, 62/63).

Auch Carl et al. (2011) baten Sprachexperten, blind je vier Übersetzungsvarianten des gleichen Ausgangstextes (zwei PE, zwei HÜ) in eine Rangfolge zu bringen und stellten fest, dass die posteditierten Texte als in etwa genauso gut eingestuft wurden wie die humanübersetzten.

Green et al. (2013) kamen in ihrer Studie zu dem gleichen Ergebnis. Allerdings war die Studie so konzipiert, dass nicht ganze Texte, sondern einzelne Sätze bewertet wurden, was auch bedeutet, dass Aspekte wie Kohärenz, Kohäsion und Logik bei der Bewertung nicht berücksichtigt wurde. An dieser Stelle sei darauf hingewiesen, dass eine Aneinanderreihung von ‚guten‘ Sätzen noch nicht zwangsläufig zu einem ‚guten‘, im Sinne von kohärenten, logischen Text führt.

In fehlerbasierten Qualitätsanalysen von Guerberof (2009) und O'Curran (2014) waren die Ergebnisse für posteditierte MÜ-Segmente ebenfalls deutlich besser als Segmente, die von Humanübersetzern oder unter Zuhilfenahme von *Translation*



*Memories* übersetzt wurden. In O'Curran Studie waren stilistische Fehler in HÜ-Segmenten überraschenderweise sogar drei Mal häufiger als in PE-Segmenten.

Im Gesamtbild ist also deutlich geworden, dass posteditierte Texte qualitativ in etwa genauso gut sind wie humanübersetzte Texte (Fiederer/O'Brien 2009; Garcia 2010; Carl et al. 2011; Daems et al. 2017, 92 u.v.m.) oder sogar etwas besser (Guerberof 2009; Green et al. 2013; O'Curran 2014). Anders sieht es jedoch aus, wenn man das unbearbeitete MÜ-Output mit humanübersetzten Texten vergleicht: die Qualität des unbearbeiteten MÜ-Outputs ist deutlich geringer als die der Humanübersetzungen und es taugt in der Regel nur zu groben Informationszwecken (Bowker 2009, 124; Kirchhoff et al. 2011, 475; Carl et al. 2015, 146 u.v.m.).

So wie die erzielbaren Produktivitätssteigerungen vom MÜ-Output abhängen (vgl. Abschnitt 2.1.3), hängt auch die Qualität von posteditierten Texten entscheidend vom MÜ-Output ab (ebd., 147; Koponen 2016, 142): Wenn die Qualität des MÜ-Outputs bereits relativ hoch ist, darf (bei vergleichsweise geringerem Aufwand) auch eine hohe Qualität des posteditierten Textes erwartet werden. Andersherum lässt sich ein MÜ-Output von sehr geringer Qualität auch durch Post-Editing unter Umständen nicht auf eine „publishable“ Qualität anheben (TAUS 2016, 16). Sofern dies durch besonders gründliches und damit zeitaufwändiges Post-Editing doch möglich sein sollte, sinken die Produktivitäts- und Kostenvorteile allerdings mitunter so stark, dass Post-Editing in der Gesamtabwägung nicht mehr die beste Lösung darstellt (Senez 1998, 3; Specia 2011a, 73). Die Frage, welche Faktoren die Qualität des MÜ-Outputs positiv oder negativ beeinflussen, wird in Abschnitt 2.1.5 beleuchtet.

Einschränkend sei hier noch darauf hingewiesen, dass Qualitätsanalysen dieser Art in der Regel auf Fehleranalysen oder Bewertungen von Sprachexperten basieren. Andere wichtige Aspekte, wie ‚instinktive‘ Meinungen und Präferenzen von Adressaten, die sich nicht unbedingt mit den Ergebnissen von Fehleranalysen decken, oder andere, weniger leicht quantifizierbare Eigenschaften eines Textes werden oftmals nicht in die Qualitätsbewertung einbezogen (Toral 2019, 274). Ein ‚guter‘ Text zeichnet sich aber nicht nur durch die Abwesenheit von Fehlern aus. Man könnte argumentieren, das Wichtigste sei, dass der Text für die Zieltextradressaten ‚funktioniert‘. Deshalb werden in Abschnitt 2.1.6 blind durchgeführte Akzeptanz- und Präferenzstudien mit Zieltextradressaten vorgestellt.

### 2.1.5 Einflussfaktoren auf die Qualität des MÜ-Outputs

Es wurde an vielen Stellen festgestellt, dass sich nicht alle Textsorten und -typen für Post-Editing bzw. für maschinelle Übersetzung eignen (Church/Hovy 1993; Povlsen et al. 1998, 31; Poulis/Kolovratnik 2012, 8). Grundsätzlich kann ein Text mit gewissen Eigenschaften dann als für Post-Editing geeignet gelten, wenn zu erwarten ist, dass das MÜ-Output qualitativ bereits recht ‚gut‘ ist, sodass der Umfang des nötigen Post-Editings in einem akzeptablen Rahmen bleibt (Povlsen et al. 1998, 27). Im Folgenden sollen ausgewählte Faktoren, die für die Qualität des MÜ-Outputs eine Rolle spielen, kurz thematisiert werden.

Für die Qualität des MÜ-Outputs spielt es zunächst einmal eine entscheidende Rolle, welches MÜ-Paradigma (regelbasiert, statistisch, neuronal)<sup>5</sup> und welche konkrete MÜ-Engine verwendet wird. Von besonderer Bedeutung ist dabei, wie ähnlich die Trainingsdaten des verwendeten MÜ-Systems dem Text sind, der maschinell übersetzt werden soll (Groves/Schmidtke 2009, 2) und natürlich die Qualität der Trainingsdaten (Doğru et al. 2018). Laut Groves und Schmidtke (2009) hängt die Tatsache, dass sich bestimmte Textsorten besser für MÜ eignen als andere, unter anderem damit zusammen, welche Textsorten zuvor zum Trainieren der MÜ-Engine genutzt wurden.

Zwei weitere wichtige Einflussfaktoren auf die Qualität des MÜ-Outputs sind das Sprachpaar und das Fachgebiet (O’Brien 2010a, 8; Poulis/Kolovratnik 2012, 8; Carl et al. 2015, 146). Auch hier ist zu erwarten, dass das MÜ-Output für diejenigen Sprachpaare und Fachgebiete ‚besser‘ ist, die schon in den Trainingsdaten mit großen Textvolumina vertreten waren. (ebd.). Da zum Trainieren einer MÜ-Engine große Datenmengen, d. h. konkret umfangreiche alignierte Parallelkorpora benötigt werden (Doğru et al. 2018, 1), sind Minderheitensprachen bzw. untypische Sprachpaare hier oftmals im Nachteil (Poulis/Kolovratnik 2012, 8; Doğru et al. 2018, 12). Ausschlaggebend ist nicht nur für welches Fachgebiet eine MÜ-Engine trainiert wurde, sondern, dass sie speziell für ein (oder wenige) Fachgebiet(e) trainiert wurde: Laut Hutchins (2001, 9) funktioniert eine MÜ am besten in „domain-specific and controlled environments“ und auch Doğru et al. (2018:13) konstatieren: „corpus based machine translation systems function better with more specific data.“

---

<sup>5</sup> siehe dazu Stein 2009 und Krüger 2017.

Als gut geeignet für MÜ gelten außerdem Texte und Textsorten mit einem „considerable amount of redundancies, universals or controlled language“<sup>6</sup> (Fiederer/O'Brien 2009, 68ff; O'Brien 2010b; Temnikova 2010), sowie Texte mit einem „formal writing style“ (Groves/Schmidtke 2009, 2). In diese Kategorien fallen beispielsweise technische Dokumentationen, Bedienungsanleitungen und – wie bereits erwähnt – Fachtexte (Aymerich 2005, 4; Kirchhoff et al. 2011). Bowker und Buitrago-Ciro (2015) haben Hinweise darauf gefunden, dass auch die Akzeptanz von Zieltextadressaten für PE bzw. MÜ bei „instructional texts“ wie Bedienungsanleitungen und technischer Dokumentation höher ist als bei anderen Texttypen<sup>7</sup> (ebd., 27).

Darüber hinaus wurden textinterne, mikrostrukturelle Eigenschaften identifiziert, die positive oder negative Auswirkungen auf die Qualität des MÜ-Outputs haben. Dazu zählen beispielsweise die Satzlänge (Groves/Schmidtke 2009, 3-5) sowie die Struktur, die Absatzlänge, die Komplexität eines Textes und die verwendete Fachterminologie (Laurian 1984, 236/237; Aymerich 2005, 4). Senez (1998, 3) weist zudem darauf hin, dass qualitative Mängel im Ausgangstext, wie Rechtschreibfehler, eine ambige Syntax oder eine fehlerhafte Interpunktion, ebenfalls oft zu einem schlechteren MÜ-Output führen.

Bei der Entscheidung, ob ein Text maschinell übersetzt und dann posteditiert oder lieber humanübersetzt werden soll, sind also auch all diese Faktoren zu berücksichtigen.

### **2.1.6 Präferenz- und Akzeptanzstudien**

Wie in Abschnitt 2.1.4 bereits angerissen wurde, basieren die meisten existierenden Qualitätsstudien auf Fehleranalysen oder Einschätzungen von Sprachexperten und die Frage, ob Post-Editing eine echte Alternative zum Humanübersetzen ist, wurde meist nur auf Grundlage von rigiden Qualitäts- oder Produktivitätsanalysen beantwortet.

In diesem Abschnitt sollen stattdessen die Präferenzen und Bedürfnisse der intendierten Zieltextadressaten im Mittelpunkt stehen, da diese Laut Löffler-Laurian

---

<sup>6</sup> Für nähere Ausführungen zu Controlled Language siehe auch Adriaens et al. 2000.

<sup>7</sup> Definition nach Reiß 1986.

am besten dazu in der Lage sind, einzuschätzen, ob ein übersetzter Text ihren Bedürfnissen gerecht wird. Hutchins hinterfragt sogar, ob der Aspekt Qualität als Bewertungskriterium überhaupt geeignet bzw. sinnvoll ist:

„It is more appropriate to judge the output in terms of its usefulness for a particular purpose – not how ‘good’ the ‘translation’ is, but whether the essential message has been conveyed” (Hutchins 2001, 16).

Auch Senez kritisiert die übliche Herangehensweise, bei der es Sprachexperten sind, die Urteile fällen: „Although professional linguists may well find fault with post-edited texts, the aim is to achieve standards which are acceptable to the users” (Senez 1998, 6).

In einer Studie von Bowker und Buitrago Ciro (2015) sollten spanischsprachige Immigranten in Kanada je vier Übersetzungsvarianten (HÜ, FPE, LPE, MÜ) von drei kurzen Texten (EN → ES) nach ihrer Präferenz sortieren, und zwar zunächst ohne zu wissen, welche Übersetzungsvariante wie produziert wurde. 49 Prozent der Befragten präferierten die PE-Varianten (24 % FPE und 25 % LPE), 42 Prozent präferierten die HÜ und nur 9 Prozent wählten das unbearbeitete MÜ-Output (ebd.). Hier zeichnet sich also zwar eine leichte, aber keine klare Präferenz für die Produkte des PE oder der HÜ ab. Dies bestätigt gewissermaßen die Aussage, dass humanübersetzte und posteditierte Texte qualitativ in etwa ‚gleich gut‘ sind (Abschnitt 2.1.4).

In einer nächsten Phase wurden den Teilnehmern dann die Metadaten zur Verfügung gestellt, d. h. sie erfuhren den Übersetzungsmodus und erhielten Informationen über die Kosten und den Zeitaufwand. Im Übersetzungsmodus FPE produzierte Texte waren in diesem Fall 67 Prozent günstiger als Humanübersetzungen und erforderten nur ein Drittel der Arbeitszeit. Unter Berücksichtigung dieser Daten sollten die Teilnehmer dann angeben, welche Übersetzung ihren Bedürfnisse am besten gerecht wird (ebd.).

Während die beiden ‚billigen‘, ‚schnelleren‘ Übersetzungsmodi nun besser bewertet wurden, schnitten die ‚teureren‘ Übersetzungsmodi schlechter ab: Unter Berücksichtigung der Metadaten wurde das unbearbeitete MÜ-Output nun in 23 % und die LPE-Texte in 54 % der Fälle als Favorit gewählt. Die FPE-Texte wurden nur noch in 17 % der Fälle präferiert, die HÜ-Texte sogar nur in 6 % der Fälle. (ebd., 25).

Es lässt sich also ableiten, dass das Wissen um Kosten- und Zeiteinsparungen durchaus einen Einfluss auf die Bewertung zu haben scheint, und die Teilnehmer Abwägungen treffen, ob beispielsweise eine nur geringfügig ‚bessere‘ Qualität einen deutlich höheren Preis rechtfertigt. In diesem konkreten Fall hielten die Teilnehmer unter Berücksichtigung der Metadaten die LPE-Texte, für die typisch ist, dass sie verständlich, korrekt, frei von groben Fehlern aber mitunter stilistisch nicht ausgefeilt sind (siehe Abschnitt 2.1.2), als am angemessensten für ihre Bedürfnisse (ebd.). Dies könnte auch damit zusammenhängen, dass viele der Teilnehmer angaben, die Übersetzung zu nutzen, um ihr Verständnis des Ausgangstextes zu bestätigen, oder weil sie Informationen in ihrer Muttersprache schneller verarbeiten können (ebd., 25, 27). Eine veröffentlichungsreife Qualität mit ‚perfektem‘ Stil ist für diese Zwecke nicht unbedingt notwendig.

Bowker führte 2009 eine ähnliche Studie mit Menschen, die französisch- und englischsprachigen Minderheiten in Kanada<sup>8</sup> angehörten, durch. Sie waren von Anfang an darüber informiert, welcher Text wie produziert wurde und wussten etwa um die geringeren Kosten von posteditierten Texten (30 bis 55 % geringer für FPE gegenüber HÜ). Sie wurden explizit darum gebeten, diese Metadaten bei ihrer Entscheidung zu berücksichtigen.

Interessanterweise bevorzugten die Teilnehmer der englischsprachigen Minderheit eindeutig die PE-Texte (88 %), während die französischsprachige Minderheit die HÜ den PE-Texten deutlich vorzog (71 %). Analog zu der oben genannten Studie würde man vermuten, dass das Wissen um die Metadaten auch hier in beiden Gruppen dazu führt, dass die posteditierten Texten häufiger als Favorit gewählt werden. Als mögliche Erklärung für die Diskrepanz zwischen den beiden Gruppen verweist Bowker darauf, dass unter den Teilnehmern aus der Gruppe der französischsprachigen Minderheit etwa die Hälfte Sprachexperten waren, was die Ergebnisse in Richtung der HÜ verfälschen könnte. Tatsächlich gaben 88 Prozent der teilnehmenden Sprachexperten in einem zuvor ausgefüllten Fragebogen an, dass für sie HÜ der einzig akzeptable Übersetzungsmodus ist. Ohne die Sprachexperten fiel

---

<sup>8</sup> Bei den „Official Language Minority Communities“ handelte es sich um die Fransaskois, eine französischsprachigen Gemeinschaft, die in der mehrheitlich englischsprachigen Provinz Saskatchewan lebt, und um eine englischsprachige Gemeinschaft, die in der mehrheitlich französischsprachigen Region Outaouais im Westen von Quebec lebt.

das Ergebnis ausgeglichener aus: 56 Prozent präferierten die HÜ und 44 Prozent die PE-Texte. Die Sprachexperten präferierten also häufiger die HÜ als der ‚Durchschnitts-Zieltextrezipient‘. Es ist daher durchaus denkbar, dass hier negative Einstellungen und Vorurteile von Sprachexperten gegenüber dem PE bzw. der MÜ in ihre Bewertung einfließen. Diese These wird in Abschnitt 2.1.7 näher erläutert und untermauert.

Die Ergebnisse dieser Studien können – auch wenn sie gemeinschaftsspezifisch sind und sich nicht ohne Weiteres verallgemeinern oder auf andere Personengruppen übertragen lassen – für die Praxis durchaus Konsequenzen haben. Denn wenn die jeweiligen Zieltextadressaten posteditierte Texte trotz womöglich stilistisch ungeschickter Formulierungen akzeptieren, und in der Folge des geringeren Preises womöglich auch noch größere Textvolumina übersetzt werden können (Löffler-Laurian 1996, Rn 21ff), ist es schlicht nicht logisch, Post-Editing trotzdem abzulehnen.

### **2.1.7 Einstellungen von Übersetzern, Sprachdienstleistern und Kunden**

Schließlich soll die Einstellung von Übersetzern, Sprachdienstleistern und Kunden, insbesondere von Unternehmenskunden, zu Post-Editing genauer beleuchtet werden. Denn wenn Sprachdienstleister und Übersetzer sich weigern, Post-Editing in ihren Leistungskatalog zu integrieren, kommen auch die Vorteile, die in Qualitäts- und Produktivitätsanalysen festgestellt wurden (Abschnitt 2.1.3 und 2.1.4), nicht zum Tragen. Das gleiche gilt, wenn das Angebot, MÜ zu nutzen, von Unternehmenskunden aufgrund von negativen Einstellungen und Vorurteilen aus Prinzip ausgeschlagen wird.

Unter professionellen Übersetzern reichen die Einstellungen gegenüber Post-Editing von strikter Ablehnung über Skepsis bis hin zu Enthusiasmus (Carl et al. 2015, 150). Ein viel diskutierter Punkt innerhalb dieser Debatte ist die Frage, wie sehr Übersetzer bereit sind, zuzulassen und zu akzeptieren, dass Maschinen den Workflow und das Berufsbild des Übersetzers verändern (Koponen 2016, 131). Ähnliche Diskussionen wurden im Übrigen bereits in den 1990er Jahren geführt, als Translation Memories bzw. CAT-Systeme („Computer-Assisted Translation“) zunehmend Fuß fassten (Carl et al. 2015, 150). Durch den wachsenden Einfluss von

MÜ wird die Selbstwahrnehmung des Übersetzers (erneut) in Frage gestellt (ebd.). Auch Hutchins (2001, 5), Yuste-Rodrigo (2001, 1) und Arnold et al. (1994, 8) stellen fest, dass sich einige Übersetzer durch die MÜ in ihrer Existenz bedroht sehen. Zudem befürchten sie, dass sie weniger zufrieden in ihrem Job sein werden, wenn sich dieser zum Posteditor wandelt (Church/Hovy 1993, 249), da sie Post-Editing als „tedious and dreary“ und als „frustrating“ wahrnehmen (Specia 2011a, 73). Hinzu kommt noch, dass sie sich in ihrer Freiheit, ihre Worte frei zu wählen, limitiert sehen (Wagner 1985, 1). Außerdem missfällt ihnen das Korrigieren von den immer gleichen Fehlern, die Menschen so nie machen würden und sie befürchten, dass sich die eigenen Sprachfähigkeiten verschlechtern, wenn sie sich ständig mit ‚schlechtem‘ MÜ-Output konfrontiert sehen (O’Brien 2002, 100). Laut Church und Hovy (1993, 249) sorgen sich einige Übersetzer auch um den Ruf ihres Berufs generell, da der Einsatz von MÜ ihrer Ansicht nach zur Produktion von Übersetzungen geringerer Qualität führt. Wagner (1985, 2) stellt in Bezug auf LPE fest: „translators must temporarily abandon their high standards in order to produce a 'quick and dirty' version of the source text“, was vielen Übersetzern nicht nur schwerfallen sondern auch missfallen dürfte (vgl. auch Gerber 2009). In einer Umfrage des Common Sense Advisory (2016, 20) bewerteten freiberufliche Übersetzer das Post-Editing von MÜ-Output auf einer Skala von 1 („I hate it) bis 10 („I love it) mit einem Wert von nur 3,7. Etwa 30 Prozent der Befragten vergaben die schlechteste Bewertung und nur 17 Prozent eine eindeutig positive.

Die oben genannten Einstellungen, Sorgen und Erfahrungen von Übersetzern haben zumindest zeitweise das Potenzial, bewusst oder unbewusst ihre Bewertung des MÜ-Outputs zu beeinflussen (Hutchins/Somers 1992, 173; Roturier 2006, 81), wie in der oben genannten Präferenzstudie von Bowker (2009) bereits angerissen wurde.

Laut O’Brien (2002, 100) sind es vor allem Übersetzer mit wenig Erfahrung und wenig Kenntnissen über Post-Editing, die diesem Modus „hostile“ gegenüberstehen. Senez beobachtet eine ähnliche Tendenz, offeriert aber zusätzlich einen Ausblick, in dem sie versöhnliche Worte findet:

„The experienced post-editor no longer feels threatened by the machine, but has learned to reap as much benefit as possible from what the computer gives him” (Senez 1998, 5).

Trotz der negativen Einstellung einiger Übersetzer hat Post-Editing über die Jahre an Bedeutung gewonnen und ist heute aus der Übersetzungsbranche nicht mehr wegzudenken: In einer Befragung von 438 „stakeholder in the translation and localisation field“ durch Gaspari et al. (2015 zitiert nach Koponen 2016, 134) gaben 30 Prozent der Befragten an, dass sie MÜ-Systeme bereits einsetzen. Weitere 21 Prozent hielten es für gesichert oder zumindest wahrscheinlich, dass sie sie künftig einsetzen würden. Unter denen, die MÜ bereits nutzen, setzten 70 Prozent – in unterschiedlichem Umfang – auch Post-Editing ein (ebd.). Laut der *Translation Technology Insights* Umfrage der RWS Group (2020, 15) nutzten 2016 bereits 50 Prozent der Sprachdienstleister MÜ – bis 2020 ist dieser Wert sogar auf 65 Prozent gestiegen. Unter freiberuflichen Übersetzern waren es 2020 demnach circa 40 Prozent, die MÜ einsetzen. Als Grund für diesen vergleichsweise geringen Wert werden auch hier negative Einstellungen der Übersetzer und die grundsätzliche Ablehnung dieser Technologie genannt. Erwartet wird allerdings, dass auch freiberufliche Übersetzer nach und nach ihre Skepsis abbauen und sich der Trend weiter zugunsten von MÜ entwickelt (ebd.).

Schließlich ist auch die Einstellung von Kunden, die Übersetzungen bei Sprachdienstleistern in Auftrag geben, zu berücksichtigen. In einer AMTRA/SDL-Umfrage unter Geschäftsführern von globalen Unternehmen im Jahr 2010 wurde festgestellt, dass 28 Prozent der Befragten MÜ bereits nutzen oder dies vorhaben, und weitere 57 Prozent positiv gegenüber PE eingestellt sind und in der Kombination aus durch Menschen sichergestellter Qualität und Produktivität der Maschine eine Win-Win-Situation sehen (Carl et al. 2015, 146). Auch der Allied Business Intelligence Report von 2002 sah MÜ als einen Bereich mit exponentiellem Wachstumspotential (Allied Business Intelligence 2002, 5.21 zitiert nach Bowker 2009, 124).

Heute haben große Unternehmen wie Intel, Adobe oder Continental Airlines MÜ längst in ihre Übersetzungs- und Lokalisierungsabläufe integriert, ebenso wie die Pan-American Health Organization und das Canadian-UN Global Public Health Intelligence Network (Kirchhoff et al. 2011, 475). Laut Common Sense Advisory (2016, 16) stammt ein Großteil des MÜ-Bedarfs von großen Unternehmen, wobei aber prognostiziert wird, dass auch kleine und mittelständische Unternehmen ihren Übersetzungsbedarf in den nächsten Jahren zunehmend mit MÜ decken werden (ebd.)



Insgesamt lässt sich also festhalten, dass MÜ und PE auch in der professionellen Praxis und trotz Vorurteilen und Skepsis seitens einiger Übersetzer immer nachgefragter und relevanter werden, sodass davon auszugehen ist, dass sich der Trend zu MÜ und PE unbeirrt fortsetzt (ebd.; Koponen 2016, 134; SDL 2020).

### **2.1.8 Weitere Forschungsrichtungen**

Darüber hinaus haben sich im Bereich Post-Editing weitere Forschungsrichtungen entwickelt, die an dieser Stelle kurz angesprochen werden sollen.

In der jüngsten Forschung wurden immer häufiger sogenannte Eyetracking- und Keylogging-Technologien genutzt, um den Prozess des Übersetzens oder Post-Editings besser verstehen zu lernen. O'Brien (2006) beispielsweise nutzte Eyetracking zur Erfassung der kognitiven Belastung („cognitive load“) und stellte – abgeleitet über Werte wie Pupillenweite – fest, dass diese beim Humanübersetzen höher als bei Post-Editing-Aufgaben ist (zitiert nach Green et al. 2013, 2). Die Annahme, dass Augenbewegungen, Fixierdauer und Sakkadenlängen Rückschlüsse über gleichzeitig stattfindende kognitive Prozesse zulassen (Doherty/O'Brien 2009, 1), nennt man auch „Eye-Mind-Hypothese“ (Ball et al. 2006 zitiert nach Doherty/O'Brien 2009, 1). Aus ihren Ergebnissen schließt sie, dass es sich beim Humanübersetzen um eine kognitiv anstrengendere Aufgabe handelt und untermauert insofern die Einschätzung von Übersetzern, Post-Editing sei „tedious and dreary“ (Church/Hovy 1993, 249) (siehe Abschnitt 2.1.7).

In einer Studie von Čulo et al. (2014) wurde ein Einfluss von der MÜ auf posteditierte Texte auf lexikalischer Ebene festgestellt: So wurden beispielsweise unidiomatische Formulierungen wie die Übersetzung von „in a gesture“ mit „in einer Geste“, die von der MÜ-Engine vorgeschlagen wurde, auch noch in vielen vollständig posteditierten Übersetzungen beobachtet (ebd., 205). Ebenso wurden die auf das MÜ-Output zurückzuführenden lexikalischen Inkonsistenzen wie die Bezeichnungen „Krankenpfleger“ und „Krankenschwester“ für „nurse“, die sich auf die gleiche Person beziehen, in der Hälfte der posteditierten Texte nicht vereinheitlicht (ebd., 206/207). Daraus lässt sich ein gewisser Einfluss von MÜ auf posteditierte Texte ableiten, der von Čulo et al. (2014, 212) mit dem Interferenz-Begriff („interference effects“) in Verbindung gebracht wird.

Als alternative Herangehensweise zur Bewertung der Qualität eines maschinell übersetzten Textes hat sich die Auswertung der für das Post-Editing benötigten Zeit und der sogenannten *Edit distance* (vorgenommene Korrekturen, gemessen in Zeichen) etabliert. Dabei gilt: Je weniger Zeit für das Post-Editing aufgewendet werden muss bzw. je geringer die *Edit distance*, desto besser das MÜ-Output (Specia 2011a, 73/74). Eine weitere alternative Qualitätsbewertung wurde in einer Studie von Carl et al. (2015) vorgestellt: Augenbewegungen und Fixierdauern sollen demnach Rückschlüsse auf den „cognitive effort“ und damit auf die Qualität des jeweiligen MÜ-Segments zulassen (vgl. auch Doherty et al. 2010; Čulo et al. 2014).

Im Hinblick auf die Ausbildung von Übersetzern bzw. Posteditoren weisen unter anderem O'Brien (2002) und Krings (2001, 12 zitiert nach O'Brien 2002, 99) auf die Notwendigkeit hin, Übersetzer auch speziell für Post-Editing auszubilden. Prinzipiell seien Übersetzer für diese Aufgabe zwar durchaus bereits geeignet, da sie über Fachkenntnisse in beiden Sprachen sowie über linguistische Expertise verfügen und mit Adressatenerwartungen vertraut sind. Jedoch erfordere Post-Editing auch zusätzliche Kenntnisse und Kompetenzen, die idealerweise schon während der Ausbildung vermittelt werden sollten (ebd., 100 ff.). Durch entsprechende Weiterbildungen bzw. durch ein Einbinden dieses Themas in die Ausbildung bzw. das Studium könne außerdem die Produktivität beim Post-Editieren erhöht und Skepsis gegenüber MÜ abgebaut werden (ebd., 100).

## **2.2 Translationese**

Im folgenden Abschnitt werden die oben erwähnten Übersetzungsuniversalien, insbesondere Simplification (nach Baker 1995) (Abschnitt 2.2.2) und Interference (nach Teich 2003) (Abschnitt 2.2.3), die grundlegend für die vorliegende Arbeit sind näher erläutert. Außerdem wird der aktuelle Forschungsstand der Translationese-Forschung zusammengefasst (Abschnitt 2.2.4). Schließlich werden neuere Tendenzen und Entwicklungen der Translationese-Forschung thematisiert (Abschnitt 2.2.5).

### 2.2.1 Definition

Gellerstam definiert Translationese als „any type of difference between original text and translated text“ (1986, 6 zitiert nach Daems et al. 2017, 90). Er beobachtet deutliche Unterschiede zwischen schwedischen Originaltexten und schwedischen Übersetzungen aus dem Englischen, wobei er nicht alle auf den Einfluss der Ausgangssprache zurückführt (Gellerstam 1986 zitiert nach Lapshinova-Koltunski 2015, 94). Translationese ist also eine Art „translational language“ bzw. ein Dialekt der jeweiligen Sprache (Volansky et al. 2014, 1). Die Charakteristika von Translationese unterscheiden sich sowohl von denen der Ausgangs- als auch der Zielsprache (Lapshinova-Koltunski 2015, 94). Frawley (1984) bezeichnet dieses Phänomen auch als „third code“ (zitiert nach Lapshinova-Koltunski 2015, 94/95). Dass es Unterschiede zwischen Originaltexten und Übersetzungen geben muss, begründet Baker folgendermaßen:

„[...] given that all language is patterned, and that this patterning is influenced by the purpose for which language is used and the context in which it is used, the patterning of translated text must be different from that of original text production; the nature and pressures of the translation process must leave traces in the language that translators produce“ (Baker 1996, 177).

Ungeachtet des teilweise unterstellten „pejorative ring“ des Begriffes Translationese (Tirkkonen-Condit 2002, 207), der auf seine Assoziation mit Begriffen wie *journalese*, *officialese* und *legalese* zurückzuführen ist (ebd.), verwende ich den Begriff in dieser Arbeit – wie auch von Gellerstam vorgesehen – konnotationsfrei und rein deskriptiv.

### 2.2.2 Übersetzungsuniversalien nach Baker

Einen der wichtigsten Beiträge zur Translationese-Forschung leistete Mona Baker (1993, 1995, 1996): Ausgehend von den in der Literatur bis dato beobachteten Eigenschaften von übersetzten Texten stellte sie vier Hypothesen zu sogenannten „translation universals“ („Übersetzungsuniversalien“) vor. Sie geht davon aus, dass diese vollkommen unabhängig von der Ausgangssprache und der Sprachrichtung in übersetzten Texten vorzufinden seien (ebd., 175). Empirische Studien, die diese

Hypothesen belegen oder dies zumindest versuchten, wurden nach und nach von ihr und anderen ‚nachgeliefert‘ (siehe Abschnitt 2.2.4). Inwiefern diese Übersetzungsuniversalien tatsächlich unabhängig von Ausgangssprache und Sprachrichtung sind, wurde in der Vergangenheit mehrfach in Frage gestellt (z. B. Puurtinen 2003, 401ff.; Bernardini/Zanettin 2004; Xiao 2010, 189; Volansky et al. 2014) oder gänzlich verneint (Tymoczko 1998, 657). Dementsprechend ist auch der Begriff „Universalien“ bzw. „universals“ ist nicht unumstritten. Lapshinova-Koltunski (2015, 94/95) präferiert die Begriffe „features“ oder „phenomena“, während bei Toury (2012, 92) und Chesterman (2004, 33) von „laws“ oder „regularities“ die Rede ist. Da „Übersetzungsuniversalien“ trotz aller Kritik der in der Literatur nach wie vor am häufigsten verwendete Begriff ist, wird er auch in dieser Arbeit genutzt.

Die vier Übersetzungsuniversalien nach Baker sind *Simplification*, *Explicitation*, *Normalization* und *Levelling Out*.

Unter *Simplification* versteht Baker (1996, 176) die Tendenz, dass in Übersetzungen einfachere Sprache verwendet wird – und zwar sowohl im Vergleich zum Ausgangstext als auch im Vergleich zu zielsprachlichen Originaltexten. *Simplification* kann auf lexikalischer, syntaktischer und stilistischer Ebene vorliegen (ebd., 181; Laviosa-Braithwaite 1998, 306ff.). Auf syntaktischer Ebene etwa kann ein Text durch das Aufspalten von langen, komplexen Sätzen in mehrere kurze Sätze vereinfacht werden (Baker 1996, 181; Lapshinova-Koltunski 2013, 96/97). Ein Beispiel für *Simplification* auf lexikalischer Ebene ist das Ersetzen spezieller durch allgemeine Begriffe.

Mit *Explicitation* meint Baker (1996, 180) die Tendenz, im AT implizite Informationen in der Übersetzung ‚auszubuchstabieren‘. Auf grammatischer Ebene geschehe dies z. B. durch das Auflösen von Ambiguitäten im Ausgangstext (Olohan 2001, 424) oder das Explizieren logischer Bezüge mithilfe von *function words*, d. h. z. B. durch Konjunktionen (Baker 1996, 181; Lapshinova-Koltunski 2015, 96/97): Ein impliziter modaler Bezug wird in der Übersetzung beispielsweise durch eine entsprechende modale Konjunktion explizit gemacht. Auch die Tatsache, dass Übersetzungen in der Regel länger sind als ihre Ausgangstexte (Baker 1996, 180/181), könne auf *Explicitation* zurückzuführen sein. Denn wenn implizite

Informationen des AT im ZT expliziter formuliert und unter Umständen sogar durch Hintergrundinformationen ergänzt werden, wird der Zieltext länger als der Ausgangstext (ebd.).

Lapshinova-Koltunski (2015, 96/97) sieht einen engen und gewissermaßen widersprüchlichen Zusammenhang zwischen Simplification und Explicitation: „Simplification and explicitation features correlate and may be just the opposite of each other.“ Beispielsweise sei das Ersetzen allgemeiner Terme durch spezielle ein Merkmal von Explicitation, während das Ersetzen von spezifischen Termen durch allgemeinere auf Simplification hindeuten könne (ebd.) Insofern wären Simplification und Explicitation „the opposite of each other“ (ebd.). Ein Beispiel für das Korrelieren von Simplification und Explicitation, bzw. für einen gewissen „overlap“ liefert Baker (1996, 182): Das Aufspalten eines Satzes auf zwei Sätze mache diesen nicht nur simpler, sondern gehe häufig mit einer Disambiguierung des Inhalts einher, d. h. mit Explicitation.

Normalization (auch *Conservatism* (ebd., 183) oder *Homogeneisation* (Toral 2019, 273), beschreibt „a tendency to exaggerate features of the target language and to conform to its typical patterns“ (Baker 1996, 183). In der Folge enthalten Übersetzungen gewisse typische Charakteristika der Zielsprache in höherem Maße als zielsprachliche Originaltexte dies tun (ebd.). Normalization stelle also einen Versuch dar, sich den Normen der Zielsprache anzupassen (Lapshinova-Koltunski 2015, 95). Vanderauwera führt aus:

„translations overrepresent features of their host environment in order to make up for the fact that they were not originally meant to function in that environment“ (Vanderauwera 1985, 11 zitiert nach Baker 1993, 245).

Beispiel für Normalization sind das ‚Grammatikalisieren‘ von ungrammatischen Strukturen im Ausgangstext und das Vermeiden unüblicher Interpunktion in Übersetzungen (Baker 1996, 184). Laut Baker (1996, 183) hängt Normalization zudem von dem Status des Ausgangstexts bzw. der Ausgangssprache ab: Je höher der Status der Ausgangssprache, desto geringer „the tendency to normalise“ (ebd.).

Levelling Out, auch *Convergence* (Laviosa 2002), beschreibt schließlich die Annahme, dass sich übersetzte Texte untereinander ähnlicher sind als Originaltexte

untereinander (Baker 1996, 177). Diese Ähnlichkeit kann sich auf verschiedene Aspekte beziehen und kann nur textübergreifend bewertet werden. Baker (1996, 186) begründet dieses Phänomen, indem sie darauf abstellt, dass Übersetzungen die Tendenz haben „to gravitate towards the centre of a continuum“ bzw. „to move away from extremes“. Ein Beispiel für Levelling Out sei, dass sich bei bestimmten Größen wie der lexikalischen Dichte, dem Type-Token-Ratio und der durchschnittlichen Satzlänge in Übersetzungen weniger Varianz zeigt als in Originaltexten (ebd., 184).

### 2.2.3 Interference nach Toury bzw. Teich

Gideon Toury, ein Vertreter der *Descriptive Translation Studies*, schlug 1985 zwei sogenannte „laws of translational behaviour“ (2012, 303) vor: Das *law of growing standardization* (vgl. Levelling Out nach Baker (1996), Abschnitt 2.2.2) und das *law of interference* (ebd.) (kurz auch *Interference* bzw. *Interferenz*). Im Folgenden werde ich nur auf letzteres näher eingehen.

Toury (2012, 310) definiert bzw. begründet die Existenz des Phänomens Interference wie folgt: „phenomena pertaining to the make-up of the source text tend to be transferred to the target text“. Dieses Phänomen ist also nicht sprachpaarunabhängig wie die Übersetzungsuniversalien von Baker, sondern ist den Einfluss der Ausgangssprache zurückzuführen (Koppel/Ordan 2011, 1318; Toury 2012, 310). Dennoch kann auch Interference als eine Art Übersetzungsuniversalie verstanden werden (Mauranen 2004, 79; Toral 2019, 273/274).

In Anlehnung an Tourys Interferenzbegriff prägte Teich 2003 den Begriff *shining through*:

„In a translation into a given target language (TL), the translation may be oriented more towards the source language (SL), i.e. the SL shines through“ (Teich 2003, 145).

Ein Beispiel für dieses shining through ist demnach eine überdurchschnittlich hohe Anzahl an Passiva in deutschen Übersetzungen englischer Texte im Vergleich zu deutschen Originaltexten (ebd., 187, 211; Lapshinova-Koltunski 2013, 78). Dies sei auf den Einfluss des Englischen zurückzuführen, da das Passiv im Englischen deutlich häufiger verwendet wird als im Deutschen (Teich 2003, 181; Čulo et al. 2017, 197). Ein weiteres Beispiel für shining through wäre ein untypisch starker

Nominalstil in englischen Übersetzungen deutscher Texte im Vergleich zu englischen Originaltexten. Dieser untypisch stark ausgeprägte Nominalstil wiederum könnte auf die Ausgangssprache Deutsch zurückgeführt werden, da der Nominalstil sehr typisch für das Deutsche aber eher untypisch im Englischen ist (Steiner 2012, 79; Lapshinova-Koltunski 2015, 96) (mehr dazu in Abschnitt 3.2.2.1).

Teichs „shining through“ ist weniger vage als Tourys „law of interference“. Wurm beschreibt das wie folgt:

„Teich operationalisiert dieses Phänomen [shining through, M.C.] jetzt jedoch für die korpuslinguistische Arbeit mit Übersetzungen, indem sie es auch für Gebrauchsfrequenzen verwendet“ (Wurm 2008, 44).

Da es sich bei dieser Arbeit um eine Korpusanalyse handelt, in der Gebrauchsfrequenzen untersucht werden, soll hier also der Interferenzbegriff nach Teich (2003) gelten, d. h. wann immer von Interference die Rede ist, ist damit der Interferenzbegriff im Sinne von shining through gemeint.

Lapshinova-Koltunski (2015, 96/97) weist auf die Beziehung von shining through und Normalization (vgl. Abschnitt 2.2.2) hin:

„normalisation implies the exaggerated use of the patterns typical for the target languages, whereas ‘shining through‘ involves the patterns typical for the source language (but not specific for the target language).“

Analog zu Bakers Annahme (1996: 183), dass Normalization mit geringerer Wahrscheinlichkeit auftritt, wenn der Status der Ausgangssprache hoch ist (siehe Abschnitt 2.2.2), konstatiert Lapshinova-Koltunski, dass shining through tendenziell auftritt, wenn die Ausgangssprache einen höheren Status als die Zielsprache genießt.

#### **2.2.4 Forschungsstand**

Seit den ersten Beschreibungen von Übersetzungsuniversalien haben sich viele Translatologen und Linguisten mit deren Untersuchung und Nachweis befasst. Die Übersetzungsuniversalien nach Baker (1993, 1995) sowie Tourys (2012) *law of interference* und *law of growing standardization* haben dabei besondere

Aufmerksamkeit erfahren. Eine Auswahl dieser Beiträge soll nun vorgestellt werden, um einen Überblick über den aktuellen Forschungsstand zu schaffen.

Laut Baker (1996, 175) dauerte es lange, bis elektronische Korpora in der Translationsforschung überhaupt genutzt wurden, da ihr Nutzen für die Untersuchung von linguistischen Eigenschaften und Mustern in (übersetzten) Texten lange verkannt wurde. Mitte der 90er Jahre begannen Baker und Laviosa-Braithwaite mit der Erstellung eines Korpus aus englischen Übersetzungen, das explizit für den oben genannten Zweck vorgesehen war und einen empirischen Nachweis der Übersetzungsuniversalien ermöglichen sollte (ebd., 178). Das *Translation English Corpus* („TEC“), das in Zusammenarbeit mit dem *Institute of Science and Technology* der *University of Manchester* entwickelt wurde, besteht aus englischen Übersetzungen verschiedener Ausgangssprachen und Textsorten: Etwa 82 Prozent der Texte sind dem Bereich Fiktion zuzuordnen, der Rest setzt sich aus Zeitungsartikeln, Biographien und Bordmagazinen zusammen (Olohan 2001, 423; The University of Manchester 2022).<sup>9</sup>

Viele der zum Thema Übersetzungsuniversalien durchgeführten Studien nutzten das TEC (Laviosa-Braithwaite 1997; Laviosa 1998; z. B. Baker/Olohan 2000; Kenny 2014) und zogen ein Subkorpus des *British National Corpus* („BNC“) als Vergleichskorpus heran: Ziel war es vor allem Unterschiede auf syntaktischer und lexikalischer Ebene zwischen englischen Übersetzungen und englischen Originaltexten festzustellen (Olohan 2001, 423).

Sara Laviosa (1998, Laviosa-Braithwaite 1997) untersuchte im Rahmen zweier Studien ausgewählte Subkorpora des *English Comparable Corpus* (ECC)<sup>10</sup> auf Bakers Übersetzungsuniversalien. In der ersten Studie untersuchte sie übersetzte Zeitungsartikel aus den britischen Zeitungen *The Guardian* und *The European*, in der zweiten Romanübersetzungen verschiedener Ausgangssprachen. In beiden Studien beobachtete sie, dass die Übersetzungen im Vergleich zu den Originaltexten des Vergleichskorpus eine geringe lexikalische Dichte aufwiesen (vgl. Abschnitt 3.2.1.2), was auf Simplification hindeutet. Die durchschnittliche Satzlänge, ebenfalls eine

---

<sup>9</sup> Stand 2001 nach Olohan 2001. Inwiefern das Korpus seitdem erweitert und seine Zusammensetzung verändert wurde ließ sich nicht abschließend klären.

<sup>10</sup> Das *English Comparable Corpus* besteht aus dem *Translation English Corpus* sowie einem Vergleichskorpus.



Methode zum Nachweis von Simplification, war nur bei den Zeitungsartikeln für die übersetzten Texte geringer – in den Romanübersetzungen war sie überraschenderweise höher. Auch die lexikalische Diversität, die sie über das Verhältnis der frequentesten Wörter zu weniger frequenten Wörter in einem Text untersucht, ist in Übersetzungen geringer als in Originaltexten, was ebenfalls auf Simplification hindeutet.

Al-Shabab (1996) untersuchte die lexikalische Diversität über eine andere Methode als Laviosa-Braithwaite, nämlich mithilfe des sogenannten Type-Token-Ratio (mehr dazu in Abschnitt 3.2.1.1). Er stellte fest, dass dieser in Übersetzungen geringer ist als in Originaltexten und wies somit ebenfalls Simplification nach (zitiert nach Volansky et al. 2014, 4).

Auch die Übersetzungsuniversalie Explicitation wurde vielfach untersucht. Olohan and Baker (2000 zitiert nach Olohan 2001, 423) nutzten das TEC und beobachteten, dass das optionale ‚that‘, das teilweise in der indirekten Rede im Englischen verwendet wird, in englischen Übersetzungen deutlich häufiger verwendet wird als in englischen Originaltexten. Dies bringen sie mit einer Form der Disambiguierung und somit mit dem Phänomen Explicitation in Verbindung (ebd.).

In Anlehnung daran versuchte Olohan (2001) mit einem anderen Subkorpus des TEC, Explicitation anhand von weiteren optionalen, syntaktischen Features nachzuweisen. Tatsächlich stellt sie fest, dass auch andere optionale Features wie z. B. die (in gewissen Relativsätzen) optionalen Relativpronomen ‚that‘ und ‚which‘ ebenso wie das optionale ‚in order‘ in der Phrase ‚in order to‘ in den englischen Übersetzungen häufiger als in englischen Originaltexten verwendet werden (ebd., 427). Sie bestätigte also die Ergebnisse von Olohan und Baker und wies ebenfalls Explicitation nach.

Teich (2003) prägte in ihrer 2003 erschienen Monographie nicht nur den Begriff *shining through*, sondern wies das Phänomen auch anhand eines Korpus aus populärwissenschaftlichen Texten am Sprachpaar Deutsch-Englisch nach. Gleichzeitig bestätigte sie aber auch Bakers Hypothese zu Normalization: Übersetzungen sind stärker an den Normen der Zielsprache orientiert, aber die Ausgangssprache bzw. der Ausgangstext ‚scheint‘ dennoch durch (ebd., 207).

Normalization und Interference (shining through) schließen sich also nicht gegenseitig aus, sondern können koexistieren.

Auch Neumann (2013, 258ff.) fand anhand des CroCro-Korpus<sup>11</sup> aus deutsch- und englischsprachigen Originaltexten und Übersetzungen und durch die Analyse verschiedener Dimensionen des Registers sowohl Hinweise auf shining through als auch auf Normalization.

Levelling Out konnte unter anderem in der oben genannten Studie von Laviosa (1997) nachgewiesen werden: in Bezug auf ihre lexikalische Dichte sind Übersetzungen homogener bzw. standardisierter als Originaltexte.

Korpusanalysen zur Untersuchung der Übersetzungsuniversalien beschränkten sich aber nicht ausschließlich auf englische Übersetzungen, auch wenn dort sicherlich der Fokus lag und nach wie vor liegt (Mauranen/Kujamäki 2004, 3). Puurtinen (2003) gelang es anhand eines Korpus aus finnischen Übersetzungen von Kinderbüchern universelle Features von „translated finnish“ zu identifizieren, allerdings deckten sich diese nur teilweise mit Bakers Übersetzungsuniversalien – teilweise widerlegten sie sie auch. Sie verweist darauf, dass Bakers Übersetzungsuniversalien auf indo-europäische Sprachen ausgerichtet seien und für eine agglutinierende Sprache wie Finnisch, die nicht Teil der indo-europäischen Sprachfamilie ist, nicht die gleiche Gültigkeit haben (ebd.). Kritik an der Universalität von Bakers Übersetzungsuniversalien kam, wie bereits in Abschnitt 2.2.2 erwähnt, von vielen Seiten. So wies auch Xiao (2010, 179, 189) darauf hin, dass sie sich in chinesischen Übersetzungen nicht nachweisen ließen und möglicherweise nicht universell sind, sondern lediglich charakterisierend für Übersetzungen aus indo-europäischen Sprachen ins Englische.

Wie dieser Überblick zeigt, ist die Studienlage in Bezug auf Übersetzungsuniversalien umfangreich. Zwar gibt es Zweifel daran, wie universell diese Universalien tatsächlich sind und ob sie tatsächlich existieren. Die Befundlage suggeriert dennoch, dass zumindest Übersetzungen bestimmter Textsorten und Sprachpaare im Vergleich zu Originaltexten simpler, expliziter und normalisierter

---

<sup>11</sup> Die Zusammensetzung des CroCro-Korpus wird in Neumann (2013, 83ff) beschrieben.

sind als Originaltexte, und dass Charakteristika der Ausgangssprache in der Übersetzung oftmals ‚durchscheinen‘ (Toral 2019, 273/274).

### 2.2.5 Neuere Forschungsrichtungen

Mit der Zeit entwickelte sich die Translationese-Forschung dahingehend weiter, dass nicht mehr nur humanübersetzte Texte Gegenstand der Forschung waren. Lapshinova-Koltunski (2013, 93) war eine der ersten, die bemängelte, dass bei korpusbasierten Analysen lediglich Originaltexte mit humanübersetzten Texten verglichen wurden, andere Übersetzungsmodi aber weitestgehend ignoriert wurden. Vereinzelt wurden humanübersetzte Texte mit maschinell übersetzten Texten verglichen (siehe White et al. 1994; Papineni et al. 2002; Babych/Hartley 2004). Ziel dieser Arbeiten war aber weniger die Erforschung linguistischer Charakteristika der aus den verschiedenen Übersetzungsmodi entstandenen Texte, sondern vielmehr die automatische Bewertung der Qualität von MÜ-Systemen (Lapshinova-Koltunski 2015, 94).

Lapshinova-Koltunski erstellte also das VARTRA-Korpus (VARiation in TRAnslation) (2013), das neben Originaltexten und Humanübersetzungen (mit und ohne Hilfe von CAT-Tools) auch maschinell übersetzte Texte verschiedener MÜ-Paradigmen (regelbasiert, statistisch) umfasst – posteditierte Texte werden aber auch hier nicht berücksichtigt. Zudem umfasst das VARTRA-Korpus verschiedene Sprachrichtungen, Textsorten und Register (ebd., 80). Mittels einer umfangreichen Korpusanalyse stellte Lapshinova-Koltunski fest, dass sich die Produkte dieser verschiedenen Übersetzungsmodi tatsächlich voneinander unterscheiden: In Bezug auf Simplification konnte sie nachweisen, dass maschinell übersetzte Texte eine geringere lexikalische Dichte und Diversität als humanübersetzte Texte haben. Darüber hinaus fand sie Hinweise auf Levelling Out in der Gesamtschau aller Übersetzungen, Schließlich wies sie über die Verhältnisse von nominalen und verbalen Wortarten in AT und ZT (vgl. Abschnitt 3.2.2.1) shining through nach, und zwar insbesondere bei den mithilfe eines CAT-Tools erstellen Humanübersetzungen (ebd., 83).

In einer anderen Studie (Lapshinova-Koltunski 2015), in der ein Subkorpus von VARTRA (VARTRA-SMALL, EN↔DE) genutzt wurde, konnte sie allerdings weder über die lexikalische Dichte noch über den Type-Token-Ratio zweifelsfrei

nachweisen, dass sich der Grad an Simplification vom Übersetzungsmodus beeinflusst ist (ebd.: 109). Auch der Nachweis von Explicitation anhand von verschiedenen Untersuchungsaspekten gelang nicht. Der Anteil an Konjunktionen in den human- und maschinell übersetzten Texten war nicht, wie erwartet, höher als in Originaltexten, sondern niedriger (ebd., 102/103). Nur in Bezug auf Convergence (Levelling Out) bestätigte sich ihre Hypothese: Die untersuchten Übersetzungen waren einander ähnlicher als die untersuchten Originaltexte (ebd., 109).

Trotz der Behauptung vieler Sprachexperten und Laien, sie erkennen einen übersetzten Text daran, dass er sich ‚irgendwie übersetzt‘ anhört, konnten Studien bisher nicht nachweisen, dass Leser in der Lage sind, Originaltexte von Übersetzungen zu unterscheiden (Tirkkonen-Condit 2002; Baroni/Bernardini 2006).

Weitere Studien aus dem Feld der Translationese beschäftigten sich mit der Frage, ob man einen Computer so trainieren kann, dass er (human)übersetzte Texte von Originaltexten unterscheiden kann. Baroni und Bernadini gelang dies (2006, 95/96) ebenso wie Rabinovich und Wintner (2015), Ilisei et al. (2010), Koppel und Ordan (2011) und Volansky et al. (2014). Es ist auch gelungen, Computer so zu trainieren, dass sie maschinell übersetzte Texte von humanübersetzten Texten unterscheiden können (siehe Aharoni et al. 2014). Anwendung können diese Erkenntnisse etwa bei der Bereinigung von Parallelkorpora aus dem Internet (z. B. wenn (maschinell) übersetzte Texte aus einem Korpus rausgefiltert werden sollen) oder bei der Verbesserung von Sprachmodellen für MÜ-Systemen finden (Lapshinova-Koltunski 2013, 78/79). Außerdem kann eine automatisierte Translationese-Erkennung als Bewertungstool, als webbasiertes Extraktionstool für Parallelkorpora und bei der multilingualen Plagiatserkennung hilfreich sein (Daems et al. 2017, 90).

Wie aus diesen Ausführungen deutlich geworden sein dürfte, hat sich die Translationese-Forschung mit der Zeit weiterentwickelt. In den 1990er Jahren, als sich der Forschungszweig Translationese entwickelte, wurden ausschließlich humanübersetzte Texte und ihre Ausgangstexte untersucht, da maschinelle Übersetzung noch keine große Rolle spielte. Mit der zunehmenden Bedeutung von MÜ (vgl. Abschnitt 1.1) wurden aber auch Stimmen lauter, die eine Berücksichtigung von maschinell-übersetzten Texten forderten bzw. maschinell-

übersetzte Texte humanübersetzten Texten gegenüberstellten (Lapshinova-Koltunski 2013; Aharoni et al. 2014). Zusammen mit maschineller Übersetzung hat aber auch Post-Editing an Bedeutung gewonnen – nicht zuletzt dank der Neuronalen Maschinellen Übersetzung, die entscheidend dazu beigetragen hat, dass der Einsatz von MÜ sich auch wirtschaftlich immer mehr lohnt und den Qualitätsansprüchen gerecht wird. Die Entstehung der Post-Editese-Forschung aus diesen Entwicklungen heraus ist in dieser Hinsicht ein logischer Schritt, um den Anforderungen der Zeit gerecht zu werden.

## **2.3 Post-Editese**

Der folgende Abschnitt ist der Post-Editese-Forschung gewidmet. Nach der Klärung des Begriffs (Abschnitt 2.3.1) folgt eine Vorstellung der bisherigen Studien zum Phänomen Post-Editese (Abschnitt 2.3.2). Die Ergebnisse dieser Studien, insbesondere von Toral (2019), beeinflussen die Hypothesen, die im dritten Abschnitt für die eigene Korpusanalyse aufgestellt werden.

### **2.3.1 Definition**

Der Begriff Post-Editese wurde 2017 von Daems et al. in Anlehnung an den von Gellerstam geprägten Begriff Translationese vorgeschlagen. Sie definieren Post-Editese als „the expected unique characteristics of a PE text that set it apart from a translated text“ (Daems et al. 2017, 90). Es ist also ein Sammelbegriff für Charakteristika eines posteditierten Textes in Abgrenzung vor allem zu humanübersetzten Texten. Die Annahme, dass es Unterschiede zwischen humanübersetzten und posteditierten Texten gibt, begründet Toral wie folgt:

„In theory, one would claim that HTs and PE translations are clearly different, since, in the translation workflow of the latter, the translator is primed by the output of an MT system“ (Toral 2019, 273).

Die logische Konsequenz daraus wiederum ist, dass die finale posteditierte Übersetzung zu einem gewissen Grad den „footprint“ des MÜ-Outputs trägt (ebd.).

### 2.3.2 Forschungsstand

Wie bereits erwähnt ist die Erforschung der Post-Editese ein recht junger Forschungszweig und hat – vor allem im Vergleich zur Translationese – bisher wenig Aufmerksamkeit erfahren. Die bisherigen einschlägigen Beiträge zum Thema werden in diesem Abschnitt vorgestellt.

Čulo et al. (2017) untersuchten die Terminologie in humanübersetzten und posteditierten Texten anhand eines Parallelkorpus, das verschiedene Übersetzungsmodi (HÜ, FPE, LPE, MÜ) umfasste. Damit waren sie die ersten, die die linguistischen Eigenschaften von Texten dieser Übersetzungsmodi gezielt verglichen (Daems et al. 2017, 91). Sie stellten anhand von Untersuchungen auf Grundlage des *perplexity coefficient* fest, dass die Terminologie in posteditierten Texten der MÜ-Terminologie ähnlicher ist als der Terminologie, die in Humanübersetzungen gewählt wurde und führen dies auf ein shining through des MÜ-Outputs in posteditierten Texten zurück (Čulo et al. 2017, 183).

Auch wenn der Begriff Post-Editese zum Zeitpunkt dieser beiden Studien noch nicht geprägt war, können sie doch als erste Beiträge zur Post-Editese-Forschung verstanden werden.

Erst 2017 wurde dann der Begriff Post-Editese von Daems et al. (2017) geprägt und wie oben definiert. Ihr Beitrag ist noch heute bedeutend für die Post-Editese-Forschung und gleichzeitig einer der umfangreichsten Beiträge zu diesem Thema. Sie weisen zunächst darauf hin, dass der mögliche Nachweis von Post-Editese-Merkmalen keineswegs direkt mit einer geringeren Übersetzungsqualität in Verbindung bringen lasse, denn wie in Abschnitt 2.1.4 dargelegt, sind posteditierte Texte laut Studien mindestens genauso gut wie humanübersetzte Texte (ebd., 90). Vielmehr könnte die Post-Editese-Forschung dabei helfen „more elusive aspects of translation quality“ zu verstehen (ebd.) (vgl. Abschnitt 2.1.6).

Daems et al. untersuchten, inwiefern Menschen zwischen posteditierten und humanübersetzten Texten unterscheiden können und ob Computer so trainiert werden können, dass sie eine richtige Zuordnung zu einem dieser beiden Übersetzungsmodi vornehmen können. Ziel der Studie ist also, das Phänomen Post-Editese nachzuweisen (ebd.). Aufgrund der bisherigen Forschungsergebnisse erwarten sie, dass das MÜ-Output einen Einfluss auf posteditierte Texte hat, da Posteditoren

zwangsläufig vom MÜ Output geprintet werden (vgl. Green et al. 2013): „Where we expect there to be source text (ST) interference in Translationese, we expect there to be MT interference in Post-editedese“ (Daems et al. 2017, 90).

Das Korpus von Daems et al. (2017, 91) besteht aus 87 Übersetzungen und 87 posteditierten (EN→NL) Texten, die sich gleichmäßig auf die acht Ausgangstexte aufteilen. Bei den Texten handelt es sich um Zeitungsartikel.

Im ersten Teil der Studie, der „reader perception study“ wurden den Teilnehmern (195 Translationsstudenten und Fakultätsmitarbeiter der Universität Gent) paarweise humanübersetzte und posteditierte Texte vorgelegt. Sie erhielten die Aufgabe zu entscheiden, ob es sich bei einem, beiden oder keinem der Texte um einen posteditierten Text handelt (ebd., 93/94). Bei der Auswertung stellten Daems et al. fest, dass die Leser nicht in der Lage sind, Humanübersetzungen von posteditierten Texten zu unterscheiden, da sie die Texte nur in etwa 30 Prozent der Fälle dem richtigen Übersetzungsmodus zuordneten (ebd., 95).

Der zweite Teil der Studie, die computergestützte Analyse, sollte klären, ob Computer – anders als Menschen – in der Lage sind, zwischen humanübersetzten und posteditierten Texten zu unterscheiden. Dazu wurden aus den Korpus-texten 55 Features ausgewählt, deren Werte Rückschlüsse auf den Übersetzungsmodus zulassen sollen. Darunter waren traditionelle (z. B. durchschnittliche Wort-/Satzlänge), lexikalische (Type-Token-Ratio), semantische (Häufigkeit von verschiedenen Konnektoren) und syntaktische Features (PoS-Tagging) (ebd., 96). Nach der Trainingsphase und einigen Anpassungen wurde der finale Test mit bis dahin ungesehenen Daten durchgeführt. Im Ergebnis zeigte sich, dass das Programm nicht erfolgreich darauf trainiert werden konnte, humanübersetzte Texte von posteditierten zu unterscheiden: Eine richtige Zuordnung erfolgte nur in 43,75 Prozent der Fälle (ebd., 100).

Somit fassen Daems et al. zusammen, dass sie im Rahmen dieser beiden Studien keine Beweise für die Existenz von Post-Editedese finden konnten (ebd., 101). Sie weisen allerdings auch auf zahlreiche mögliche Probleme im Aufbau ihrer Studien hin, wie etwa die geringe Erfahrung der Übersetzer bzw. Posteditoren und Überanpassung im Fall der computergestützten Analyse (ebd.).

Ein weiterer Beitrag zur Post-Editedese-Forschung kam von Farrell. Zwar wählt er einen etwas anderen Ansatz, doch ist auch diese Studie in der Post-Editedese-

Forschung zu verorten. Er untersucht posteditierte Texte auf sogenannte *MT Markers* (2018, 50/51) und meint damit nicht ‚globale‘ Features wie Daems et al., sondern konkrete sprachliche Formulierungen, die stark darauf hindeuten, dass es sich bei einem vorliegenden Text um einen posteditierten Text handelt. Dazu hat er kurze Wikipediaartikel übersetzen bzw. posteditieren lassen (EN → IT) und anschließend analysiert, wie einzelne Wörter und Phrasen – sogenannte *n-grams* – von den Übersetzern und Posteditoren übersetzt wurden (ebd., 51). Dabei fand er heraus, dass es in posteditierten Texten pro n-gram weniger verschiedene Übersetzungsvarianten als bei Humanübersetzungen gibt (ebd., 53/54), was auf eine Normalization hindeuten könnte (ebd., 58). Dies liegt ihm zufolge vermutlich daran, dass MÜ-Engines nur eine Untermenge der typischerweise in Humanübersetzungen vorkommenden Übersetzungsvarianten vorschlagen (ebd.). Folglich sind meist auch nur diese Übersetzungsvarianten in posteditierten Texten zu finden, und zwar jeweils „with a higher than natural frequency“ (ebd.). Es sind diese Übersetzungsvarianten, die Farrell *MT Markers* nennt. Mitgrund für die erhöhte Frequenz dürfte die Tatsache sein, dass Post-Editoren, wenn sie in einem MÜ-Text eine akzeptable Lösung vorfinden, diese i. d. R. auch dann übernehmen, wenn es nicht ihre präferierte Lösungsvariante ist (ebd., 54).

Auf dieser Basis ist es Farrell auch gelungen, einen Test für sein Versuchskorpus zu designen, mit dem sehr zuverlässig eine richtige Zuordnung getroffen werden konnte. Dieser Test basierte nur auf einem *MT Marker*, nämlich „ci sono“, der Vorschlag der MÜ-Engine. Farrell hatte festgestellt, dass keiner der Humanübersetzer „there are“ mit „ci sono“ übersetzt hatte, aber jeder Posteditor als Übersetzung für „there are“ mindestens einmal „ci sono“ unverändert gelassen hatte (ebd., 57). Wenn „ci sono“ in einem Text vorkommt, ist das also ein eindeutiger Hinweis darauf, dass es sich bei dem vorliegenden Text um einen posteditierten Text handelt. Zwar lässt sich der von Farrell entwickelte Test nicht ohne Weiteres auf andere Korpora übertragen, aber er stellt dennoch einen interessanten Beitrag zur Post-Editese-Forschung dar.

Wie bereits erwähnt leistete Antonio Toral (2019) mit dem Artikel „Post-Editese: An Exacerbated Translationese“ einen weiteren wertvollen Beitrag zur Post-Editese-Forschung. Auf Grundlage bisheriger Forschungsergebnisse zu Übersetzungsuniversalien untersucht Toral, ob diese typischen Eigenschaften von



humanübersetzten Texten auch in posteditierten Texten zu finden sind (ebd., 273/274). Sein Ziel ist es also, zu untersuchen, ob und inwiefern sich Translationese von Post-Editese unterscheidet. Dazu leitet er vier konkrete Untersuchungsmerkmale aus den Übersetzungsuniversalien Simplification, Normalization und Interference ab. Er nutzt den Type-Token-Ratio (TTR) zur Bestimmung der lexikalischen Diversität der Korpus-texte und vergleicht das Verhältnis zwischen sogenannten content words und der Gesamtwortzahl, um Aussagen über die lexikalische Dichte zu treffen. Ist die lexikalische Diversität und Dichte in posteditierten Texten ähnlich gering wie in humanübersetzten Texten, würde dies darauf hindeuten, dass das Translationese-Phänomen Simplification auch ein Post-Editese-Phänomen ist. Außerdem untersucht er mithilfe des *perplexity coefficient* Abweichungen beim PoS-Tagging zwischen AT und ZT, um eine mögliche Interference des AT in posteditierten Texten nachzuweisen. Schließlich prüft er, wie ähnlich die Satz-längen zwischen AT und ZT bei den verschiedenen Übersetzungsmodi sind, um Interference und Normalization nachzuweisen.

Die von Toral verwendeten Korpora bestehen aus Humanübersetzungen und posteditierten Texten (insgesamt 2213 Satzpaare) und umfassen fünf verschiedene Sprachrichtungen, sowie die beiden Arten von Post-Editing (LPE und FPE) und eine unbearbeitete MÜ-Fassung. Bei etwa einem Viertel der Korpus-texte handelte es sich um Untertitel, der Rest ist im Bereich Nachrichten zu verorten (ebd., 275/276). Die genaue Methodik, die in großen Teilen auch Grundlage der vorliegenden Arbeit ist, wird in Abschnitt 3.2 ausführlich erläutert.

Tatsächlich ist es Toral gelungen, die von ihm gewählten Translationese-Merkmale auch in posteditierten Texten nachzuweisen – und zwar meist in höherem Maße als bei humanübersetzten Texten. Posteditierte Texte haben eine geringere lexikalische Diversität und eine geringere lexikalische Dichte als Humanübersetzungen, was mit dem Phänomen Simplification in Verbindung gebracht wird. Das Translationese-Merkmal Simplification ist also in posteditierten Texten noch stärker ausgeprägt als in Humanübersetzungen. Außerdem stellt er fest, dass die Satz-längen und das PoS-Tagging bei posteditierten Texten zwischen dem AT und dem ZT ähnlicher sind als das für Humanübersetzungen der Fall ist, was auf eine besonders ausgeprägte Interference und Normalization in posteditierten Texten hinweist (ebd., 279). Somit beweist Toral anhand seines Korpus die Existenz von

Post-Editese, die als verschärfte Form der Translationese verstanden werden kann (ebd.).

Zudem weist er darauf hin, dass viele dieser Merkmale in den unbearbeiteten MÜ-Outputs noch stärker ausgeprägt sind als in posteditierten Texten. Beispielsweise ist die lexikalische Diversität beim MÜ-Output am geringsten, etwas höher bei posteditierten Texten und wiederum höher bei humanübersetzten Texten. Dies interpretiert Toral wie folgt: der Posteditor kann die lexikalische Diversität zwar erhöhen, aber durch das Priming durch die MÜ bleibt auch nach dem Posteditieren ein „footprint“ des MÜ-Outputs im posteditierten Text vorhanden (ebd.).

Forschungsrichtungen wie die Entwicklung einer Erkennungssoftware für posteditierte Texte, wie sie es analog in der Translationese-Forschung gab (vgl. Abschnitt 2.2.5), gab es in der Post-Editese-Forschung bisher nicht. Wenn das Phänomen Post-Editese umfangreicher erforscht ist, könnte in einem nächsten Schritt eine solche Software entwickelt werden (vgl. Daems et al. 2017, 90). Ein Anwendungsbeispiel für eine solche Software wäre die Bereinigung von Parallelkorpora – etwa, wenn posteditierte Texte nicht Teil des Parallelkorpus sein sollen.

#### **2.4 Zwischenfazit und Forschungsdesiderat**

In den vorangegangenen Abschnitten wurden die verschiedenen Forschungsrichtungen und -schwerpunkte der Themenkomplexe Post-Editing, Translationese und Post-Editese thematisiert. Viele der Forschungsfelder, wie Qualität und Produktivität beim Post-Editing, sind äußerst umfangreich erforscht worden. Andere wiederum, wie die Translationese-Forschung, haben in den vergangenen Jahrzehnten zwar viel Aufmerksamkeit erfahren, aber dabei unterrepräsentierte Sprachen und verschiedene Übersetzungsmodi nicht ausreichend beachtet (Mauranen/Kujamäki 2004, 3; Lapshinova-Koltunski 2013).

Aus dieser Forschungslücke heraus hat sich die Post-Editese-Forschung entwickelt. Wie die Zusammenfassung des Forschungsstands zu diesem Forschungsfeld gezeigt hat, ist dieses Phänomen bisher noch zu wenig empirisch beforscht worden, um generalisierte Aussagen über die Existenz oder die Eigenschaften von Post-Editese treffen zu können. In den wenigen Fällen, in denen

mögliche Eigenschaften von Post-Editese umfassend untersucht wurden, geschah dies meist – wie bei Toral (2019) und Daems et al. (2017) – anhand von Zeitungsartikeln oder sogenannter *general language* (Čulo et al. 2017, 184). Fachtexte und Texte aus dem Bereich Language für Specific Purposes (LSP) wurden in der Post-Editese-Forschung bisher weitestgehend ignoriert (außer bei Čulo et al. 2017), und das obwohl ein Großteil des Übersetzungsbedarfs in diesem Bereich zu verorten ist (ebd., 184). In der im Rahmen dieser Arbeit durchgeführten Korpusanalyse werden daher Fachtexte aus dem Bereich LSP untersucht, um einen ersten Schritt zur Schließung dieser Forschungslücke zu wagen.

Die übergeordnete Forschungsfrage lautet, ob es sich bei Post-Editese, wie von Toral (2019) behauptet, um eine verschärfte Form der Translationese handelt. Ein erster Beleg dafür wäre geliefert, wenn sich nachweisen ließe, dass Simplification und Interference, zwei bekannte Translationese-Merkmale, in posteditierten Texten stärker ausgeprägt sind als in humanübersetzten Texten. Wie bereits erläutert, werden für beide Phänomene zwei Untersuchungsmethoden herangezogen. Die Hypothesen dazu werden in den Abschnitten 3.2.1 und 3.2.2 anhand des Forschungsstands hergeleitet.

### **3 Daten und Methoden**

In diesem Abschnitt soll zunächst das in der eigenen Korpusanalyse verwendete Korpus vorgestellt werden (Abschnitt 3.1). Anschließend werden die verwendeten Analysemethoden erläutert und Hypothesen formuliert (Abschnitt 3.2). In diesem Zusammenhang wird die Eignung der Methoden für diese Korpusanalyse begründet, aber auch auf mögliche Limitationen hingewiesen.

#### **3.1 Korpus**

In dem nachfolgenden Abschnitt werden zunächst die untersuchten Korpustexte vorgestellt (Abschnitt 3.1.1) und erläutert, wie diese erstellt wurden (Abschnitt 3.1.2). Die Auswahl des Korpus für die vorliegende Korpusanalyse wird in Abschnitt 3.1.3 kurz begründet. Schließlich wird dargelegt, wie die Korpustexte für die maschinelle Auswertung aufbereitet wurden (Abschnitt 3.1.4).

##### **3.1.1 Beschreibung der Korpustexte**

Das der Analyse zugrundeliegende Korpus wurde mir freundlicherweise von Herrn Prof. Dr. Czulo zur Verfügung gestellt. Es ist das gleiche Korpus, das in „Contrasting terminological variation in post-editing and human translation of texts from the technical and medical domain“ von Čulo et al. (2017) verwendet wurde.

Der Aufbau des Korpus ist in Abbildung 1 visualisiert.

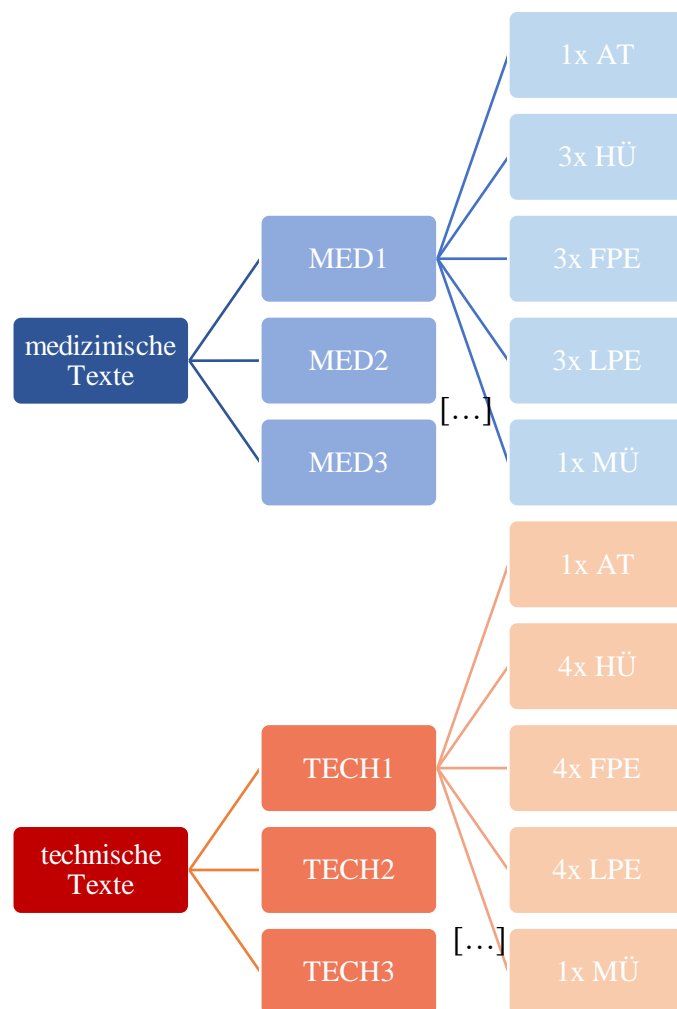


Abbildung 1: Korpusaufbau.

Die englischen Ausgangstexte stammen aus dem medizinischen und technischen Bereich und sind dem Bereich Language for Specific Purposes zuzuordnen (ebd., 188). Als solche sind sie geeignet, zur Schließung der in Abschnitt 2.4 aufgezeigten Forschungslücke innerhalb der Post-Editese-Forschung beizutragen.

Bei den Texten aus dem medizinischen Bereich handelt es sich um Ausschnitte aus Beipackzetteln zu einem Masernimpfstoff (MED1), einem Insulinmedikament (MED2) und einem Krebsmedikament (MED3). Alle drei technischen Texte (TECH1, TECH2, TECH3) sind Ausschnitte aus einer Bedienungsanleitung für einen Geschirrspüler. Bei allen sechs Ausgangstexten handelt es sich um Textausschnitte von je etwa 150 Wörtern mit einem „low formality level“, die zum Zeitpunkt der Korpuserstellung frei im Internet verfügbar waren (ebd., 188/189).

Für jeden medizinischen Text liegen – neben dem Ausgangstext (AT) – drei deutsche Humanübersetzungen (HÜ), drei vollständig posteditierte Texte (FPE) und

drei leicht posteditierte Texte (LPE) vor – bei den technischen Texten sind es jeweils vier. Zudem existiert zu jedem Ausgangstext ein unbearbeitetes MÜ-Output (MÜ). Dies wurde mithilfe von *Google Translate* generiert und ist auch Grundlage der posteditierten Texte (ebd.). Das Korpus besteht also insgesamt aus 75 Texten, darunter sechs Ausgangstexte, sechs maschinelle Übersetzungen, 21 Humanübersetzungen, 21 vollständig posteditierte Texte und 21 leicht posteditierte Texte.

### 3.1.2 Erstellung des Korpus

Bei der Erstellung des Korpus handelte es sich um eine „generic collection“, d. h. das Korpus wurde nicht zu einem bestimmten Zweck erstellt, sondern zur Gegenüberstellung von Humanübersetzung und Post-Editing (-prozessen) im Allgemeinen (Čulo et al. 2017, 188). Die sechs Ausgangstexte wurden von Studenten, die zum Zeitpunkt der Korpuserstellung in einem translationswissenschaftlichen Studiengang eingeschrieben waren, übersetzt bzw. posteditiert (ebd.).

Dadurch, dass zu jedem Ausgangstext jeweils drei bzw. vier Übersetzungsvarianten des gleichen Übersetzungsmodus vorliegen (Ausnahme: MÜ), kann zu einem gewissen Grad ausgeschlossen werden, dass eventuell festgestellte Unterschiede zwischen den Übersetzungsvarianten idiosynkratisch sind, also z. B. allein auf den Stil des jeweiligen Übersetzers bzw. Posteditors zurückzuführen sind. Das bei der Korpuserstellung angewendete Permutationsschema (siehe Abbildung 2) stellt sicher, dass kein Teilnehmer mehrere Aufgaben (also z. B. HÜ und LPE) für den gleichen Text erledigt, da es die Ergebnisse verfälschen würde, wenn ein Teilnehmer bereits mit einem Text vertraut ist.

**Table 1:** Permutation scheme for text translation, full, and light post-editing

Participant	Text 1	Text 2	Text 3
P01	HT	LPE	FPE
P02	FPE	HT	LPE
(...)			

Abbildung 2: Permutationsschema für die Datenerhebung (Čulo et al. 2017, 189)

Die Teilnehmergruppe für die technischen Texte bestand aus zwölf Translationsstudenten höherer Semester, die Teilnehmergruppe für die medizinischen Texte aus neun Teilnehmern (ebd.). Alle Teilnehmer befanden sich im mindestens dritten Jahr ihres Studiums und hatten mindestens einen Übersetzungskurs in dem relevanten Fachbereich bestanden (ebd.). Da Post-Editing kein obligatorischer Studieninhalt für die Teilnehmer war, hatten nur wenige von ihnen Post-Editing-Erfahrung, woraus Čulo et al. (2017, 188) schließen, dass alle Teilnehmer für die Übersetzungsaufgabe besser qualifiziert waren als für die Post-Editing-Aufgaben.

Allerdings erhielten die Teilnehmer Hinweise zum Post-Editing: Beim Full Post-Editing wurden sie unter anderem instruiert, auf die terminologische Konsistenz, d. h. die einheitliche Verwendung von Termini, zu achten. Für das Humanübersetzen gab es keine näheren Anweisungen, da vorausgesetzt wurde, dass die Teilnehmer mit gängigen Normen, darunter auch der Herstellung von terminologischer Konsistenz in den zielsprachlichen Fachtexten, vertraut waren (ebd., 189).

Im Rahmen der Korpuserstellung wurden auch Eyetracking- und Keylogging-Daten gesammelt. Diese finden allerdings in der vorliegenden Korpusanalyse keine Anwendung.

### **3.1.3 Begründung der Auswahl des Korpus**

Das Korpus erfüllt durch seinen Aufbau den Zweck, eine Untersuchung von Unterschieden zwischen humanübersetzten und posteditierten Texten zu ermöglichen: Das Vorhandensein mehrerer Übersetzungsvarianten des gleichen Übersetzungsmodus ist von entscheidender Bedeutung, da sonst keine generalisierbaren Aussagen über eventuell festgestellte Unterschiede zwischen den Übersetzungsmodi getroffen werden könnten. Das Korpus wurde für diese Analyse ausgewählt, um Torals und Lapshinova-Koltunskis Methoden an einer Textsorte zu überprüfen, die sich nicht nur von den von Toral (2019) und Daems (2017) untersuchten Textsorten unterscheidet, sondern auch in der übersetzerischen Praxis eine größere Relevanz hat (vgl. Abschnitt 2.4).

### 3.1.4 **Aufbereitung der Korpus­texte für die maschinelle Auswertung**

Für die vorliegende Analyse wurde für jeden der 75 Korpus­texte mithilfe der Sprachanalyse-Software Sketch Engine<sup>12</sup> ein Korpus erstellt. Bei der Erstellung eines Korpus mit Sketch Engine werden automatisch allgemeine Informationen wie die Anzahl der Wörter, Sätze und Types extrahiert. Außerdem wird ein automatisches PoS-Tagging (Sketch Engine 2018) vorgenommen. Für die deutschen Subkorpora wurde das *German RFTagger tagset* verwendet, das von Helmut Schmid und Florian Laws entwickelt wurde (Sketch Engine 2017b). Für die englischen Subkorpora wurde eine von Sketch Engine modifizierte Version des *English TreeTagger part-of-speech tagset* verwendet, welches ebenfalls von Helmut Schmid entwickelt wurde (Sketch Engine 2017a). Die Präzision der Tagger liegt laut Sketch Engine zwischen 95 und 98 Prozent (Sketch Engine 2018, 2021). Die Ergebnisse des PoS-Taggings werden für die Berechnung der lexikalischen Dichte (siehe Abschnitt 3.2.1.2) und zum Vergleich der Verhältnisse von nominalen und verbalen Wortarten (siehe Abschnitt 3.2.2.1) benötigt.

Die Korpusbezeichnung entspricht dem folgenden Muster:

TEXTBEZEICHNUNG\_(TEILNEHMERNUMMER\_)ÜBERSETZUNGSMODUS

Bei dem Dokument bzw. Korpus mit der Bezeichnung MED1\_P01\_HUE beispielsweise handelt es sich um eine Humanübersetzung des ersten medizinischen Textes (MED1), die von Teilnehmer Nr. 01 angefertigt wurde. Bei den Originaltexten und maschinellen Übersetzungen entfällt die Teilnehmernummer.

---

<sup>12</sup> [www.sketchengine.eu](http://www.sketchengine.eu).



### 3.2 Analysemethoden

Wie bereits erwähnt sind die in dieser Arbeit genutzten Methoden in großen Teilen identisch zu denen von Toral (2019). Lediglich der dritte Untersuchungsaspekt (siehe Abschnitt 3.2.2.1) orientiert sich an einer Studie von Lapshinova-Koltunski (2013, 82ff.). Die Ergebnisse von Toral und Lapshinova-Koltunski werden beim Aufstellen der Hypothesen für diese Korpusanalyse berücksichtigt und in diesem Zusammenhang noch einmal kurz wiederholt.

Zur Berechnung der relevanten Werte wurde eine Excel-Arbeitsmappe erstellt. Alle Werte werden für jeden der 75 Korpus Texte bzw. Subkorpora ermittelt und auf vier Nachkommastellen gerundet. Aus den ermittelten Werten werden jeweils verschiedene Durchschnitte (nach Übersetzungsmodus, Textgruppe, Text) berechnet (siehe Abbildung 3).

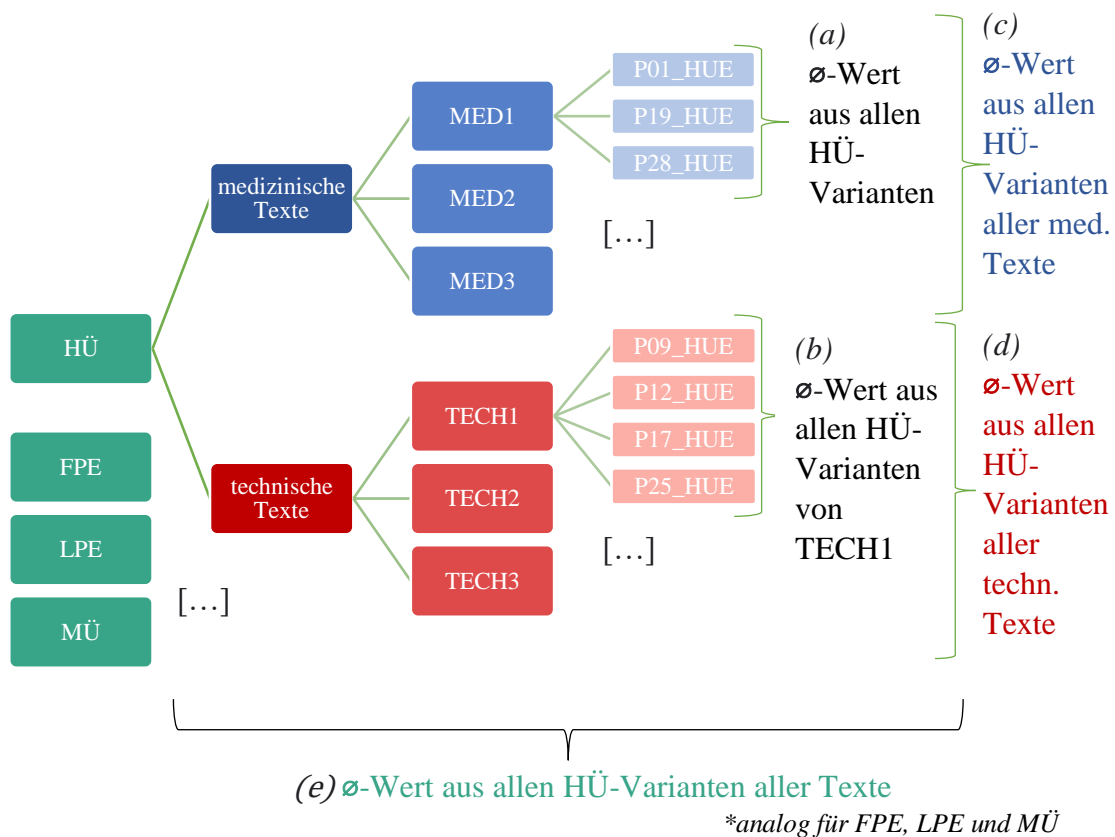


Abbildung 3: Berechnungsmethode der Durchschnitte

Für alle Übersetzungen des gleichen Ausgangstextes (z. B. MED1 oder TECH1) und des gleichen Übersetzungsmodus (z. B. HÜ) wird ein Durchschnittswert berechnet ((a) und (b)). Die Ergebnisse aller medizinischer (und analog aller technischer) Texte wiederum werden in einem Gesamtergebnis zusammengefasst, d. h. es wird ein Durchschnittswert aus allen Humanübersetzungen aller medizinischer (c) (und analog aller technischer (d)) Texte ermittelt. Diese Durchschnittswerte werden auch analog für die anderen Übersetzungsmodi (FPE, LPE, MÜ) berechnet. Schließlich wird der Durchschnitt für jeden Übersetzungsmodus unter Berücksichtigung aller Texte (medizinisch und technisch) errechnet (e). Die aufgeschlüsselte Betrachtung der Ergebnisse nach Ausgangstext sowie nach Fachgebiet (medizinisch oder technisch) dient der differenzierteren Betrachtung. Denkbar ist zum Beispiel, dass sich eine Hypothese nur für medizinische oder nur für technische Texte bestätigt oder, dass einzelne Texte ‚aus dem Rahmen‘ fallen, während die anderen die gleichen Tendenzen aufweisen.

Die Darlegung der Ergebnisse erfolgt für jeden Untersuchungsaspekt von den Gesamtergebnissen hin zu den aufgeschlüsselten Ergebnissen.

### **3.2.1 Simplification**

Die Untersuchung der ersten beiden Untersuchungsaspekte, d. h. der lexikalischen Diversität und der lexikalischen Dichte, dienen wie bei Toral (2019) dem Nachweis von Simplification. Die zugehörigen Analysemethoden werden im Folgenden vorgestellt. Mögliche Limitationen oder Schwachstellen werden angerissen.

#### **3.2.1.1 Lexikalische Diversität**

Zur Bestimmung der lexikalischen Diversität in den vorliegenden Übersetzungen nutzen wir, wie Toral (2019, 276), den sogenannten Type-Token Ratio (TTR) (auch Diversifikationsquotient):

$$\text{TTR} = \frac{\text{Anzahl der Types}}{\text{Anzahl der Tokens}}$$

*Formel 1: Lexikalische Diversität (Type-Token-Ratio) (Toral 2019, 277)*

*Types* sind dabei definiert als „number of unique words used in a text“ oder anders gesagt, die Anzahl der verschiedenen Wörter eines Textes (Oakes 2012, 133). Unter *Tokens* versteht man „the length of the text in words, regardless of any repetition“, also die Gesamtanzahl der Wörter (ebd.). Der TTR gibt also das Verhältnis zwischen der Anzahl verschiedener Wörter und der Gesamtanzahl der Wörter eines Textes an. Ein hoher TTR signalisiert „lexical variation“ während ein geringer TTR auf „lexical repetition“ und damit Simplification hindeutet (Xia/Li 2010, 65).

Der TTR ist eine bewährte Größe zur Untersuchung von „vocabulary variation“ in Korpora (Baker 1996, 183; Lapshinova-Koltunski 2013, 81) und wurde in verschiedenen Translationese-Studien bereits erfolgreich zum Nachweis von Simplification genutzt (Hansen 2003, 13; Lapshinova-Koltunski 2015, 96/97) (vgl. auch Abschnitt 2.3.2).

Bei der Verwendung des TTR gilt es zu beachten, dass dieser, und damit seine Brauchbarkeit („usefulness“) stark von der Textlänge abhängig ist (Olohan 2004, 81; Oakes 2012, 134). Je länger der untersuchte Text, desto geringer wird der TTR sein, da die Wahrscheinlichkeit, dass noch neue Wörter auftauchen, stetig sinkt (Olohan 2004, 81). Um diesem Problem zu begegnen, wurde u. a. der *Standardized Type-Token-Ratio* (STTR) entwickelt. Hier werden aus einem Korpus Blöcke von je 1000 Wörtern gebildet, für die dann jeweils ein TTR berechnet wird. Daraus wird dann der Durchschnitt berechnet, was zu einem aussagekräftigerem Ergebnis führt (ebd., 80/81; Lewandowska-Tomaszczyk 2012, 7; Oakes 2012, 137). Da die Texte des vorliegenden Korpus aus nur je etwa 144 Wörtern bestehen ( $\emptyset$  143,71 Wörter), ist die Verwendung des STTR in diesem Fall nicht notwendig.

Steiner (2004, 164) weist außerdem darauf hin, dass die Verwendung des TTR auch bei kurzen Texten problematisch ist: „type-token measurements of texts of such short length are not valid, for obvious reasons“. Ob die vorliegenden Korpus-texte in Steiners Bezeichnung von Texten von „such short length“ inbegriffen sind und die daraus resultierenden Probleme auch in dieser Analyse relevant sind, wird in Abschnitt 4.1.3.1 diskutiert.

Eine weitere Grundvoraussetzung bei der Verwendung des TTR ist, dass die zu vergleichenden Texte gleich lang oder zumindest ungefähr gleich lang sind, damit auch die entsprechenden TTRs vergleichbar sind (Olohan 2004, 80/81; Xiao 2010, 195). Olohan und Xiao machen keine näheren Angaben dazu, wie ähnlich die

Textlängen sein müssen, damit der TTR vergleichbar ist. Der mögliche Einfluss der unterschiedlichen Textlängen auf die Ergebnisse des TTR wird ebenfalls in Abschnitt 4.1.3.1 genauer thematisiert.

Die Anzahl der Types und Tokens wird mithilfe von Sketch Engine ermittelt. Unter *Corpus details* erscheint die folgende Ansicht:



The screenshot shows the 'Corpus details' view for a corpus named 'MED1\_P01\_HUE'. The interface is divided into several sections:

- GENERAL INFO:** Language: German; Tagset: LIST TAGS; Word sketch grammar: SHOW; Term grammar: SHOW.
- COUNTS:** Tokens: 168; Words: 137; Sentences: 12; Documents: 1.
- LEXICON SIZES:** word?: 99; ...: 74.
- COMMON TAGS:** adjective: ADJ.\*; preposition: ADPP.\*

Abbildung 4: Sketch Engine Benutzeroberfläche, Ansicht *Corpus details*

Die in Abbildung 4 im Bereich „Counts“ unter „Words“ angegebene Zahl entspricht der Gesamtanzahl der Wörter des Korpus, d. h. die Größe, die in Formel 1 „Tokens“ heißt. Die Zahl, die Sketch Engine für „Tokens“ angibt, ist höher als die Gesamtanzahl der Wörter. Dies liegt daran, dass „Tokens“ bei Sketch Engine standardmäßig neben Wörtern auch sogenannte „non-words“ einschließen (Sketch Engine 2016b). Diese definiert Sketch Engine als „tokens which do not start with a letter of the alphabet“ und nennt als Beispiele „numbers, punctuation but also tokens such as 25-hour, 16-year-old, !mportant, 3D“ (Sketch Engine 2016a). Für die Analysezwecke dieser Arbeit ist die Berücksichtigung von non-words nicht gewünscht.<sup>13</sup> Über die *Wordlist advanced* Funktion können non-words ausgeschlossen werden. Der Wert „items“ in der Kopfzeile (siehe Abbildung 5) entspricht nun den „Types“ und „total frequency“ entspricht den „Tokens“ gemäß den oben genannten Definitionen.

<sup>13</sup> Bei den non-words in den vorliegenden Korpus-texten handelt es sich ausschließlich um Satzzeichen, die allesamt nicht in der Wortzahl berücksichtigt werden.

WORDLIST   

word (89 items | 137 total frequency)

Word	Frequency ? ↓	Word	Frequency ? ↓
1 proquad	9 ...	14 dem	2 ...
2 ist	5 ...	15 eine	2 ...
3 die	5 ...	16 kindern	2 ...
4 und	4 ...	17 alter	2 ...
5 windpocken	3 ...	18 abgeschwächte	2 ...
6 mumme	3 ...	19 verabreicht	2 ...

Abbildung 5: Sketch Engine Benutzeroberfläche, Wordlist-Funktion (advanced)

Variationen wie „Filöter“ statt „Filter“ oder „deifach“ statt „dreifach“ (siehe TECH3\_P16\_HUE) oder „Herrsteller“ statt „Hersteller“ (TECH1\_P08\_HUE), bei denen es sich offensichtlich um Tippfehler handelt, werden – anders als bei der automatischen Zählweise von Sketch Engine – in dieser Korpusanalyse als nur ein Type gezählt. So wird verhindert, dass Tippfehler als (beabsichtigte) lexikalische Diversität interpretiert werden.

Torals Hypothese in Bezug auf die lexikalische Diversität war, dass diese in posteditierten Texten geringer sein wird als in Humanübersetzungen. Dies begründet er wie folgt: MÜ-Systeme produzieren tendenziell nur „a subset of all the possible translations in the target language“ (Farrell 2018, 57 zitiert nach Toral 2019, 276), nämlich diejenigen Übersetzungsvarianten, die in den Trainingsdaten am häufigsten vorkamen. Da Posteditoren durch das MÜ-Output geprint sind, vermutete er, dass auch posteditierte Texte nur ein „subset“ aller Übersetzungsmöglichkeiten enthalten und somit lexikalisch weniger divers sind. Seine Hypothese bestätigte sich. Er beobachtete sogar, dass der TTR für maschinell übersetzte Texte noch etwas geringer ist als für posteditierte Texte. Dies begründete er damit, dass der Posteditor zwar durch das MÜ-Output geprint sei, aber durch das Post-Editing noch Änderungsspielraum habe, im Rahmen dessen er noch mehr lexikalische Diversität einbringen könne (ebd.).

Ausgehend von diesen bisherigen Ergebnissen stelle ich in Bezug auf die lexikalische Diversität ebenfalls die Hypothese auf, dass der TTR bei posteditierten Texten geringer ist als bei humanübersetzten Texten. Ferner erwarte ich, dass der TTR für das unbearbeitete MÜ-Output am geringsten ist. Die erwartete Progression für die Ergebnisse ist also MÜ<LPE<FPE<HÜ. Sie korreliert mit dem Umfang der

menschlichen Beteiligung an den verschiedenen Übersetzungsmodi: Der menschliche Übersetzer ist am meisten bei der HÜ involviert, weniger beim FPE, wiederum weniger beim LPE und am wenigsten in der MÜ.

### 3.2.1.2 Lexikalische Dichte

Mit der lexikalischen Dichte wird „the amount of information present in a text“ (ebd., 276/277), d. h. die Informationsdichte, berechnet, indem das Verhältnis zwischen der Gesamtanzahl der sogenannten *content words* und der Gesamtanzahl der Wörter des Textes berechnet wird.

$$\text{Lexikalische Dichte} = \frac{\text{Anzahl der content words}}{\text{Gesamtanzahl der Wörter}}$$

*Formel 2: Lexikalische Dichte (Toral 2019, 277)*

Zu content words (auch *lexical words*) gehören dabei Adverbien, Adjektive, Nomen und (Voll-)Verben. Das Gegenstück zu content words sind *function words* (auch *grammatical words*), worunter Pronomen, Präpositionen, Konjunktionen, Hilfsverben, Modalverben und Artikel fallen (Olohan 2004, 81). Eine geringe lexikalische Dichte, d. h. ein verhältnismäßig geringer Anteil an content words an der Gesamtwortanzahl, signalisiert demnach Simplification. Stubbs (1996, 72/73) nennt für die lexikalische Dichte einen Referenzwert von mindestens 0,4, wobei dieser je nach Genre variiert: Wissenschaftliche Artikel etwa haben tendenziell eine höhere lexikalische Dichte als Prosatexte.

Ebenso wie der TTR ist auch die lexikalische Dichte eine bewährte und häufig verwendete Größe zum Nachweis von Simplification in übersetzten Texten (z. B. Laviosa 1998; Ilisei et al. 2010; Lapshinova-Koltunski 2013, 81, 2015, 96/97; Rabinovich/Wintner 2015) (vgl. Abschnitt 2.3.2). Schon Baker (1996, 183) hat die Berechnung der lexikalischen Dichte vorgeschlagen, um Simplification nachzuweisen und begründet dies wie folgt: „using more grammatical and fewer lexical words is a way of building in more redundancy and making a text easier to process“ (ebd.).

Auch die Untersuchungsmethode lexikalische Dichte ist nicht frei von Kritik. Steiner (2004, 134) kritisiert etwa, dass „any statistics based on numbers of words will be inevitably misleading“. Grund dafür seien unter anderem die systemischen Unterschiede bei der Kompositabildung (ebd.). Diese Kritik gilt also nicht nur für die Berechnung der lexikalischen Dichte, sondern auch für den TTR und insbesondere die Gegenüberstellung der Nominal-Verbal-Verhältnisse des AT und ZT, da diese ebenfalls wortbasiert sind. Er fährt fort:

„This should be common knowledge – or should be – for nominal compounds (e.g. *taxi driver* vs. *Taxifahrer*), but the same applies to the spelling of phrasal, prepositional and phrasal-prepositional verbs in English vs. the spelling of many of their counterparts in German (e.g. *sit down* vs. *hinsetzen*)“ (Steiner 2004, 134).

Die Ergebnisse zu diesen Untersuchungsaspekten sind daher mit Vorsicht zu interpretieren. In Abschnitt 4.1.3.2 wird diese Problematik genauer beleuchtet.

Der bereits in Abschnitt 2.2.2 erwähnte Zusammenhang zwischen Simplification und Explicitation wirkt sich auch auf die lexikalische Dichte und die Interpretation der Ergebnisse aus: Vielfach wurde angenommen, dass eine geringe lexikalische Dichte ein Hinweis auf Simplification ist (s. o.). Ebenso wurde aber angenommen, dass eine vermehrte Nutzung von function words (z. B. Konjunktionen) ein Hinweis für „increased logical explicitness“ (Steiner 2012, 82) und damit für Explicitation ist (Lapshinova-Koltunski 2013, 83). Der Zusammenhang besteht hier also darin, dass eine vermehrte Nutzung von Konjunktionen, d. h. von function words, einen Text einfacher zu verstehen und expliziter macht. Folglich kann eine geringe lexikalische Dichte gleichzeitig zum Nachweis von Simplification und Explicitation dienen. Dieser Zusammenhang ist bei der Interpretation der Ergebnisse zu berücksichtigen (siehe Abschnitt 4.1.3.2).

An dieser Stelle sei daran erinnert, dass Toral in seiner Studie Simplification, Interference und Normalization untersuchte – Explicitation war nicht Teil seiner Studie. Folglich thematisierte er auch nicht den Zusammenhang von Simplification und Explicitation und seinen Einfluss auf die Interpretation der lexikalischen Dichte.

Die Ergebnisse des automatisch durchgeführten PoS-Taggings, die für die Berechnung der lexikalischen Dichte benötigt werden, können über die *Wordlist-*

Funktion von Sketch Engine abgerufen werden, indem als Filter die gewünschte Wortart (adverb, adjective, noun, verb) eingestellt wird. Sketch Engine generiert dann eine Ansicht, in der nur die Wörter der entsprechenden Wortart mit ihren Häufigkeiten aufgelistet werden. Eine Unterkategorisierung von Verben in Vollverben, Modalverben und Hilfsverben ist mit dem PoS-Tagger von Sketch Engine nicht möglich, für meine Zwecke aber notwendig, da Vollverben zu content words, Modal- und Hilfsverben aber zu function words gezählt werden. Diese Art der Ungenauigkeit sowie festgestellte Tag-Fehler werden manuell berichtigt.

Innerhalb der Post-Editese-Forschung hat bisher nur Toral (2019) nachgewiesen, dass posteditierte Texte eine geringere lexikalische Dichte als humanübersetzte Texte haben. Zwischen maschinell übersetzten Texten und posteditierten Texten konnte er keinen signifikanten Unterschied feststellen (ebd.: 276/277).

In Anlehnung an Torals Ergebnisse zur lexikalischen Dichte, wird die Hypothese aufgestellt, dass die lexikalische Dichte in posteditierten und maschinell übersetzten Texten geringer ist als in humanübersetzten Texten. Da Toral bei der lexikalischen Dichte keine signifikanten Unterschiede zwischen maschinell übersetzten und posteditierten Texten festgestellt hat, wird in diese Richtung auch keine Hypothese formuliert.

### **3.2.2 Interference**

Auch zum Nachweis von Interference werden zwei Untersuchungsaspekte herangezogen. Zum einen wird für jeden Text das Verhältnis der Verhältnisse von nominalen zu verbalen Wortarten zwischen dem AT und dem ZT berechnet. Zum anderen wird jeweils das Verhältnis der Satzlängen zwischen dem AT und dem ZT ermittelt. In den folgenden Abschnitten werden die Berechnungsmethoden im Detail erläutert und mögliche Kritikpunkte angeführt.



### 3.2.2.1 Verhältnis der Verhältnisse von nominalen und verbalen Wortarten zwischen AT und ZT

Grundlage für diesen Untersuchungsaspekt ist die Methodik aus Lapshinova-Koltunski (2013, 82/83). Sie errechnet das Verhältnis der Wörter aus nominalen und verbalen Wortarten für jeden Korpus text.

$$\text{Nominal-Verbal-Verhältnis} = \frac{\text{Anzahl der Wörter aus nominalen Wortarten}}{\text{Anzahl der Wörter aus verbalen Wortarten}}$$

*Formel 3: Verhältnis der nominalen und verbalen Wortarten (Lapshinova-Koltunski 2013, 82)*

Zu den nominalen Wortarten („nominal word classes“) zählen dabei Nomen, Pronomen, Adjektive und Adpositionen (ebd.). Der Begriff Adposition ist ein Sammelbegriff für Präpositionen und Postpositionen (Glossary of Linguistic Terms 2015). Zu den verbalen Wortarten („verbal word classes“) zählen Verben (Vollverben, Modalverben und Hilfsverben), sowie Adverbien und Konjunktionen. Das anhand von Formel 3 berechnete Verhältnis gibt Auskunft darüber, ob die nominalen oder verbalen Wortarten in einem Text dominieren. Ein vergleichsweise hoher Wert spricht für eine hohe Nominalität, während ein niedrigerer Wert für eine hohe Verbalität spricht (Lapshinova-Koltunski 2013, 83).

Laut Steiner (2012, 79) handelt es sich bei Deutsch im Vergleich zu Englisch um eine weniger ‚verbale‘, bzw. um eine ‚nominalere‘ Sprache. Im Rahmen der Translationese-Forschung ist aber festgestellt worden, dass deutsche Übersetzungen verbaler sind als deutsche Originaltexte (ebd., 83). Lapshinova-Koltunski (2013, 83) schließt daraus, dass eine überdurchschnittlich hohe Verbalität in deutschen Übersetzungen englischer Texte als Indiz für shining through gewertet werden kann. Dies müsste aber anhand von Vergleichs- und Parallelkorpora geprüft werden.

Für diese Korpusanalyse wird Lapshinova-Koltunskis Formel (Formel 3) abgewandelt bzw. weiterentwickelt. Zunächst wird das Verhältnis von nominalen zu verbalen Wortarten zusätzlich zu den Ziltexten auch für die jeweiligen englischen Ausgangstexte berechnet, damit es einen Vergleichswert gibt. Da Interference, also ein Einfluss des Ausgangstexts auf den Ziltext nachgewiesen werden soll, werden

anschließend die errechneten Verhältnisse von den Ziltexten und den jeweiligen Ausgangstexten ins Verhältnis zueinander gesetzt. Die abgewandelte Formel lautet dann wie folgt:

$$\text{Verhältnis der Nominal-Verbal-Verhältnisse} = \frac{\text{Nominal-Verbal-Verhältnis des ZT}}{\text{Nominal-Verbal-Verhältnis des AT}}$$

Formel 4: abgewandelte Formel zur Bestimmung des Verhältnisses der Nominal-Verbal-Verhältnisse von AT und ZT

Zur Interpretation der Ergebnisse folgen nun nähere Ausführungen, die in Abbildung 6 visualisiert sind.

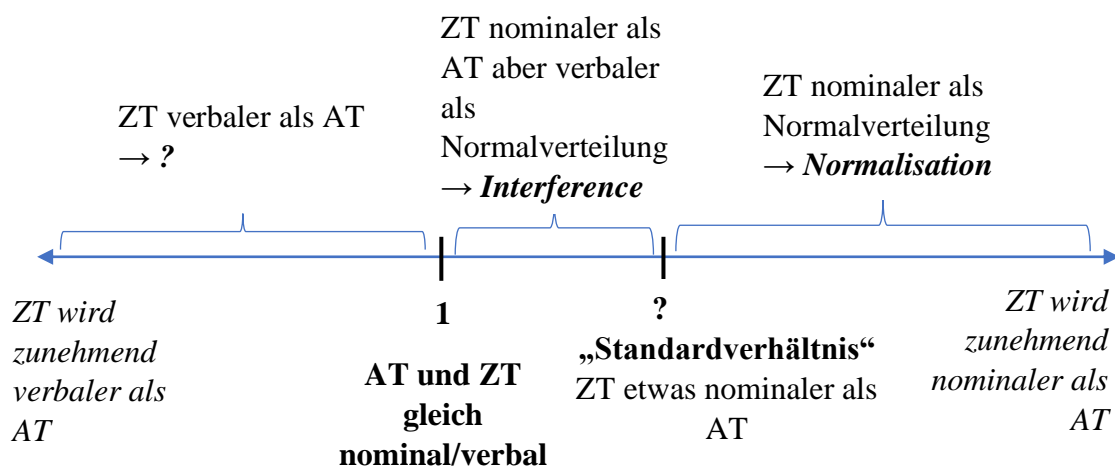


Abbildung 6: Interpretationshilfe zum Verhältnis der Nominal-Verbal-Verhältnisse

Entspricht das Nominal-Verbal-Verhältnis des ZT exakt dem Nominal-Verbal-Verhältnis des AT, d. h. wären der AT und der ZT exakt gleich nominal bzw. verbal, entspräche das nach Formel 4 einem Wert von 1. Ausgehend von den Erkenntnissen der kontrastiven Linguistik (z. B. Steiner 2012, 79ff.) wäre aber ‚normalerweise‘ (d. h. bei der Untersuchung von Originaltexten) zu erwarten, dass deutsche Texte nominaler als englische Texte sind. Daraus folgt, dass ein Wert von 1 bereits auf Interference hinweist.

Ein Wert von  $> 1$  ergibt sich, wenn der ZT einen größeren Wert für das Nominal-Verbal-Verhältnis hat als der AT, d. h., wenn der ZT nominaler ist als der AT (bzw. der AT verbaler als der ZT). Je weiter der Wert über 1 steigt, desto stärker die ‚nominale Dominanz‘ des ZT gegenüber dem AT. Der Wert, der ein ‚normales‘ Verhältnis von Nominalität bzw. Verbalität in deutschen Texten gegenüber

englischen Texten signalisiert (d. h. eine leichte nominale Dominanz) liegt an einem unbekanntem Punkt oberhalb von 1. An diesem Punkt („Standardverhältnis“, siehe Abbildung 6) sind deutsche Texte im Vergleich zu englischen genau in dem Maß nominaler, das laut Referenzkorpora ‚normal‘ ist (ebd., 79-83). Oberhalb des Standardverhältnisses allerdings wären die deutschen ZT im Vergleich zu ihren AT zu nominal (gemessen am Standardverhältnis). In diesem Fall läge die Vermutung nahe, dass es sich um einen Normalization-Effekt<sup>14</sup> handelt: In einem Versuch, den ZT an die deutschen Normen anzupassen, ‚schießt‘ der Übersetzer über das Ziel hinaus und produziert einen Text, der nominaler als deutsche Originaltexten ist.

Ein Wert von  $< 1$  ergibt sich, wenn der AT einen größeren Wert für das Nominal-Verbal-Verhältnis als der ZT hat, d. h., wenn der ZT verbaler ist als der AT. Dies wäre nach Steiners (2012, 83) Erkenntnissen für die Sprachrichtung EN  $\rightarrow$  DE eher ungewöhnlich und könnte daher nicht unmittelbar, bzw. zumindest nicht allein mit dem Phänomen Interference erklärt werden.

Aus diesen Ausführungen folgt, dass Interference bei Werten zwischen 1 und dem Standardverhältnis nachgewiesen wäre: In diesem Bereich ist der ZT verbaler als deutsche Originaltexte es normalerweise im Vergleich zu englischen Originaltexten sind (Wert  $\geq 1$ ), aber noch nicht so verbal, dass er verbaler als der AT ist (Wert  $<$  Standardverhältnis). Diese ungewöhnlich hohe Verbalität des deutschen ZT könnte dementsprechend auf den Einfluss des AT zurückgeführt werden.

Die von Steiner (2004, 133f) angesprochenen Sprachsystemunterschiede in Bezug auf die Kompositabildung (vgl. Abschnitt 3.2.1.2) dürften auch für diesen Untersuchungsaspekt eine Rolle spielen, da auch diese Berechnungen (Formel 5) auf der Anzahl von Wörtern bestimmter Wortarten basieren. Wie genau diese Unterschiede die Ergebnisse beeinflussen, wird in Abschnitt 4.2.3.1 näher beleuchtet.

Ebenso könnte die Unerfahrenheit der Übersetzer (vgl. Abschnitt 3.1.2) einen Einfluss auf die Ergebnisse haben. Laut Kujamäki (2004, 198) bleiben unerfahrene Übersetzer oft nah an der „source language sentence structure“, wodurch dem Zieltext „a taste of translationese“ (Tirkkonen-Condit 2002, 218) verliehen wird.

---

<sup>14</sup> Dass es einen Zusammenhang zwischen Interference und Normalization gibt, wurde bereits in Abschnitt 2.2.3 festgestellt.

Die Ergebnisse des PoS-Taggings werden wieder über Sketch Engines *Wordlist*-Funktion unter Verwendung der entsprechenden Filter<sup>15</sup> abgerufen. Festgestellte Tag-Fehler werden manuell korrigiert.

Lapshinova-Koltunski (2015, 106/107) stellte in ihrer Korpusanalyse fest, dass MÜ-Engines in der Regel keine Wortartwechsel vornehmen: „they [source verbal phrases, M. C.] are often left as verbal phrases in machine translation“. Daraus folgt, dass zwischen den Nominal-Verbal-Verhältnissen von den maschinell übersetzten Zieltexten und ihren AT mit keinen großen Abweichungen zu rechnen ist. Auf der anderen Seite beobachtet Lapshinova-Koltunski (ebd.), dass Verbalphrasen bei Humanübersetzungen vom Englischen ins Deutsche oft als Nominalphrasen übersetzt werden (ebd.).

Die Hypothese für diesen Untersuchungsaspekt lautet daher – in Anlehnung an Lapshinova-Koltunski (2013) und ungeachtet der Beobachtung von Kujamäki und Tirkkonen-Condit bezüglich der Unerfahrenheit von Übersetzern und deren Einfluss auf die Übersetzungen – wie folgt: Den geringsten Wert (nach Formel 4) und damit am meisten Interference haben MÜ-Texte, da MÜ-Engines in der Regel keine Wortartenwechsel oder ähnliches vornehmen. Die Werte für die posteditierten Texte dürften etwas höher liegen, da Posteditoren eventuell vom MÜ-Engine produzierte Interference noch abschwächen können und den Text hinsichtlich seiner Nominalität bzw. Verbalität etwas Richtung Standardverhältnis bewegen können. Für Humanübersetzungen wird der höchste Wert und damit am wenigsten Interference erwartet, da Humanübersetzer am meisten Handlungsspielraum haben, um bei der Übersetzung Wortklassenwechsel vorzunehmen (Lapshinova-Koltunski 2015, 106/107) und die Übersetzung in Bezug auf Nominalität bzw. Verbalität ebenfalls an das Standardverhältnis anzunähern. Die erwartete Progression lautet also MÜ<LPE<FPE<HÜ.

Denkbar ist auch, dass der Punkt des Standardverhältnisses überschritten wird, was bedeuten würde, dass die deutschen ZT nicht nur nominaler als die zugehörigen AT, sondern auch nominaler als vergleichbare deutsche Originaltexte sind. Aus der

---

<sup>15</sup> Bei der *Wordlist*-Funktion gibt es statt eines Filters für Adpositionen nur einen Filter für Präpositionen. Daher wurde über den Filter „tags“ zusätzlich zu Präpositionen nach Postpositionen gesucht.

Tatsache, dass die Ausgangssprache dieses Korpus Englisch ist – eine Sprache mit sehr hohem Status – folgt aber laut Lapshinova-Koltunski (2015: 97) , dass eher Interference (shining through) als Normalization zu erwarten ist (vgl. Abschnitt 2.2.3). Andererseits konstatieren Parthasarathi et al. (2021, 2), dass NMÜ-Systeme zu „normative language“ neigen und zwar „whether it is spelling, word order or choice of vocabulary“. Schließlich ist als dritte Option denkbar, dass sowohl Interference und Normalization in einem Text vorkommen können (vgl. Teich 2003: 207).

Die Interpretation der Ergebnisse zu diesem Untersuchungsaspekt (siehe Abschnitt 4.2.1 und 4.2.3.1) ist stark dadurch limitiert, dass der Wert des Standardverhältnisses nicht bekannt ist. Es lässt sich also nicht mit Sicherheit sagen, ob ein Wert ‚noch Interference‘ oder ‚schon Normalization‘ signalisiert. Ich beschränke mich in der Interpretation der Ergebnisse also vor allem auf die Progressionen der Werte der verschiedenen Übersetzungsmodi.

### 3.2.2.2 Verhältnis der Satzlängen zwischen AT und ZT

Analog zu Torals Studie (2019, 277) wird das Verhältnis der Satzlängen zwischen dem AT und dem ZT über den Betrag der absoluten Differenz der Satzlängen (gemessen in Zeichen) und normalisiert durch die Satzlänge des Ausgangstextes berechnet.

$$\text{Verhältnis der Satzlängen} = \frac{|\text{Satzlänge}_{\text{AT}} - \text{Satzlänge}_{\text{ZT}}|}{\text{Satzlänge}_{\text{AT}}}$$

*Formel 5: normalisiertes Satzlängenverhältnis (Toral 2019, 277)*

Die Berechnung erfolgt Satz für Satz. Aus den ermittelten Verhältnissen wird dann pro Text ein Durchschnitt berechnet.

Wie die anhand von Formel 5 errechneten Werte zu interpretieren sind, soll anhand der folgenden Beispiele erläutert werden. Wenn in einem Satzpaar die Sätze des AT und des ZT genau gleich viele Zeichen haben, ergibt sich für das

Satzlängenverhältnis ein Wert von 0.<sup>16</sup> Dies ist der geringstmögliche Wert für dieses Verhältnis, da negative Zahlen durch die Betragsstriche im Zähler ausgeschlossen sind. Daher gilt: Je näher der Wert an 0 liegt, d. h. je geringer der Wert für das Satzlängenverhältnis, desto ähnlicher sind sich der AT und der ZT in Bezug auf die jeweiligen Satzlängen. Eine starke Ähnlichkeit der Satzlängen wiederum, d. h. eine strenge Wahrung der AT-Satzgrenzen, ist ein Hinweis auf Interference. Auf der Kehrseite deutet ein hoher Wert darauf hin, dass sich etwa der Übersetzer in Bezug auf die Satzgrenzen gewisse Freiheiten genommen hat, d. h. zum Beispiel Sätze kombiniert und aufgespaltet hat.

Wenn ein Satz des AT im ZT in zwei Sätze aufgespaltet wird (oder zwei AT-Sätze zu einem ZT-Satz kombiniert werden), hat das direkte Konsequenzen für das Satzlängenverhältnis dieses Satzes sowie aller folgender Sätze, da sich die Satz-zu-Satz-Alignierung sozusagen um eine Stelle verschiebt. Diese verschobene Satzalignierung durch Aufspalten eines Satzes ist in Abbildung 6 (a) vereinfacht visuell dargestellt.

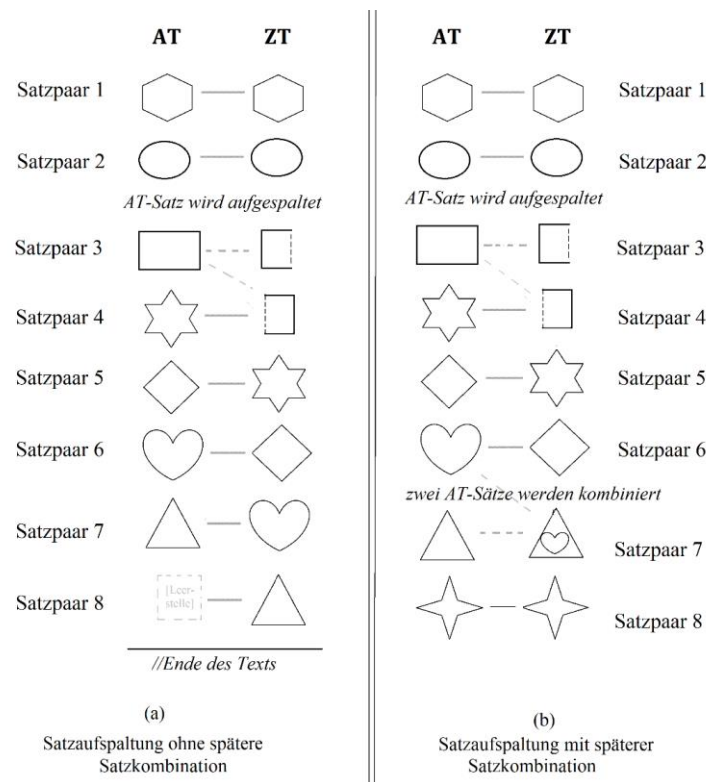


Abbildung 7: Vershobene Satzalignierung nach Satzaufspaltung

<sup>16</sup> denn wenn  $\text{Satzlänge}_{\text{AT}} = \text{Satzlänge}_{\text{ZT}} = x$ , dann ist  $\frac{|\text{Satzlänge}_{\text{AT}} - \text{Satzlänge}_{\text{ZT}}|}{\text{Satzlänge}_{\text{AT}}} = \frac{|x-x|}{x} = \frac{0}{x} = 0$ .

Ein konkretes Beispiel zur Erläuterung: Ein langer AT-Satz (Abbildung 6 (a), Satzpaar 3) wird im ZT in zwei kürzere Sätze aufgespalten. Der lange Satz des AT steht nun statt dem zugehörigen langen ZT-Satz einem kürzeren Satz gegenüber (Abbildung 6 (a), Satzpaar 3). Die daraus resultierende Diskrepanz in der Zeichenanzahl schlägt sich in einem vergleichsweise hohen Wert für das Satzlängenverhältnis für diesen Satz nieder, welcher wiederum den Durchschnittswert für den gesamten Text erhöht. Auch der nächste Satz des AT (Abbildung 6 (a) Satzpaar 4) steht durch das Aufspalten des vorangegangenen Satzes nicht mehr seinem ‚eigentlichen‘ ZT-Satz, d. h. der entsprechenden ZT-Sinneinheit (visualisiert durch gleiche Formen) gegenüber, sondern dem ‚Rest‘ des zuvor aufgespaltenen Satzes. Vor allem bei stark variierenden Satzlängen innerhalb des AT dürfte diese Diskrepanz zu stark erhöhten Werten für das jeweilige und die folgenden Satzlängenverhältnisse führen. Diese Diskrepanz zieht sich durch den gesamten Text (Abbildung 6 (a)) bzw. so lange, bis die Alignierung durch das Kombinieren zweier AT-Sätze zu einem ZT-Satz (d. h. der umgekehrten Operation) wiederhergestellt wird (Abbildung 6 (b), Satzpaar 7).

Für den Fall, dass entweder der AT oder der ZT in der Summe tatsächlich einen Satz weniger hat, wird an der letzten Position (vgl. Abbildung 6 (a), Satzpaar 8) für den ‚fehlenden‘ Satz eine Zeichenanzahl von 0 gesetzt.

Wenn der ZT einen Satz weniger hat, ergibt sich für das Verhältnis zwischen dem ‚fehlenden‘ Satz des ZT und dem zugehörigen Satz des AT zwangsläufig der Wert 1,

denn aus  $\text{Satzlänge}_{\text{ZT}} = 0$  folgt: 
$$\frac{|\text{Satzlänge}_{\text{AT}} - 0|}{\text{Satzlänge}_{\text{AT}}} = 1.$$

Wenn allerdings der AT einen Satz weniger hat als der ZT und dementsprechend  $\text{Satzlänge}_{\text{AT}} = 0$  ist, ergibt sich daraus für die Formel: 
$$\frac{|0 - \text{Satzlänge}_{\text{ZT}}|}{0}.$$

Eine Division mit einem Divisor von 0 ist nicht möglich. Toral beschreibt nicht, wie er in einem solchen Fall vorgeht. Um dieses Problem zu umgehen, führe ich in dieser Korpusanalyse eine Mehrfachalignierung durch. Wenn der AT einen Satz weniger hat,

werden die letzten beiden Sätze des ZT für die Berechnung des Satzlängenverhältnisses zu einem Satz zusammengefasst.

Ein weiterer Sonderfall liegt vor, wenn ein Satz ausgelassen wird, d. h. die Sinneinheit als solche fehlt. In diesem Fall wird nicht erst an der letzten Satzposition (s. o.), sondern unmittelbar für den ausgelassenen Satz im ZT eine Zeichenzahl von 0 gesetzt. Für das Satzlängenverhältnis dieses ‚Satzpaares‘ ergibt sich dadurch der Wert 1 (s. o.). Ein ausgelassener Satz hat also auf das Satzlängenverhältnis die nahezu gleiche Wirkung, wie wenn der ZT einen Satz weniger hat, ohne dass eine Sinneinheit ausgelassen wurde.

Die jeweiligen Satzlängen in Zeichen werden mithilfe der Zeichenzählfunktion (inkl. Leerzeichen) von Microsoft Word ermittelt und mithilfe von Microsoft Excel berechnet.

Toral (2019, 277/278) war bisher der Einzige, der das Satzlängenverhältnis zwischen AT und ZT als Nachweis für Interference nutzte – ähnliche Studien in der Translationese-Forschung untersuchten oftmals die durchschnittliche Satzlänge in Übersetzungen und Originaltexten und versuchten damit Simplification nachzuweisen (Baker 1996, 184) (siehe Abschnitt 2.2.2). Toral beobachtete, entsprechend seiner Hypothese, dass das Verhältnis der Satzlängen zwischen dem AT und dem ZT in posteditierten Texten meist geringer ausfiel als in Humanübersetzungen. Die AT und ZT sind sich also in Bezug auf ihre Satzlängen im Übersetzungsmodus Post-Editing ähnlicher als im Übersetzungsmodus Humanübersetzen, was auf mehr Interference in posteditierten Texten hindeutet. Diese Beobachtung erklärt er damit, dass MÜ-Engines in der Regel keine Sätze verbinden oder trennen, sondern den Satzgrenzen ‚treu‘ bleiben, während menschliche Übersetzer, die „from scratch“ übersetzen sich diese Freiheit eher nehmen können (Torale 2019, 277/278). Posteditoren auf der anderen Seite sind das MÜ-Output geprint (ebd.), und werden bzw. wurden zusätzlich explizit instruiert, nur notwendige Änderungen vorzunehmen. Das Verbinden oder Trennen von Sätzen gehört in den meisten Fällen wohl nicht zu den notwendigen Änderungen (siehe Abschnitt 2.1.2). Es ist also zu erwarten, dass Posteditoren die durch das MÜ-Output



eingebraachte Interference, hier in Form von strenger Beibehaltung der Satzgrenzen, nicht zwangsläufig ändern, sondern in großen Teilen übernehmen.

Einschränkend muss an dieser Stelle darauf hingewiesen werden, dass Toral (2019, 277/278) seine Hypothese nicht für alle Datensets bestätigen konnte. In einem von drei Datensätzen war das Satzlängenverhältnis in den Humanübersetzungen höher als in posteditierten Texten, d. h. für die Humanübersetzungen zeigte sich mehr Interference. Als möglichen Grund dafür führt er die mangelnde Erfahrung der Humanübersetzer an: Die humanübersetzten Texte wurden von Textproduzenten mit wenig Übersetzungserfahrung angefertigt, die posteditierten Texte allerdings von professionellen Posteditoren (ebd., 275). Er verweist darauf, dass nicht-professionelle Übersetzer dazu neigen, „more literal translations“ zu produzieren, bei denen es folglich zu erwarten ist, dass die Satzlängen im ZT nah am AT orientiert sind (ebd.). Ähnlich argumentiert auch Kujamäki (2004, 198) (vgl. Abschnitt 3.2.2.1). Da auch die Textproduzenten des vorliegenden Korpus vergleichsweise wenig Erfahrung haben, ist dies ein Faktor, der in der Diskussion (Abschnitt 4.2.3.2) aufgegriffen werden soll.

In Anlehnung an Torals These, dass Post-Editese eine verschärfte Form der Translationese ist, stelle ich als Hypothese auf, dass das nach Formel 5 berechnete Satzlängenverhältnis vom AT zum ZT bei posteditierten Texten geringer ist als bei humanübersetzten Texten. Torals Vermutung, dass unerfahrene Übersetzer dazu neigen, ‚wörtlichere‘ Übersetzungen zu produzieren, kommt in dieser Hypothese nicht zum Tragen, wird aber in Abschnitt 4.2.3.2 wieder aufgegriffen. Aufgrund der Funktionsweise der MÜ-Engines weisen die maschinell übersetzten Texte vermutlich einen ähnlich geringen oder geringeren Wert als die posteditierten Texte auf. Dementsprechend lautet die erwartete Progression MÜ<LPE<FPE<HÜ.

## 4 Ergebnisse der eigenen Korpusanalyse und Einordnung der Ergebnisse

### 4.1 Simplification

In diesem Abschnitt sollen die Ergebnisse der ersten beiden Untersuchungsaspekte, die dem Nachweis von Simplification dienen, vorgestellt werden. Eine Diskussion und Einordnung der Ergebnisse zu Simplification folgen direkt im Anschluss im Abschnitt 4.1.3.

#### 4.1.1 Lexikalische Diversität

Die durchschnittlichen Type-Token-Ratios für die verschiedenen Übersetzungsmodi unter Berücksichtigung aller Texte können Abbildung 8 entnommen werden.

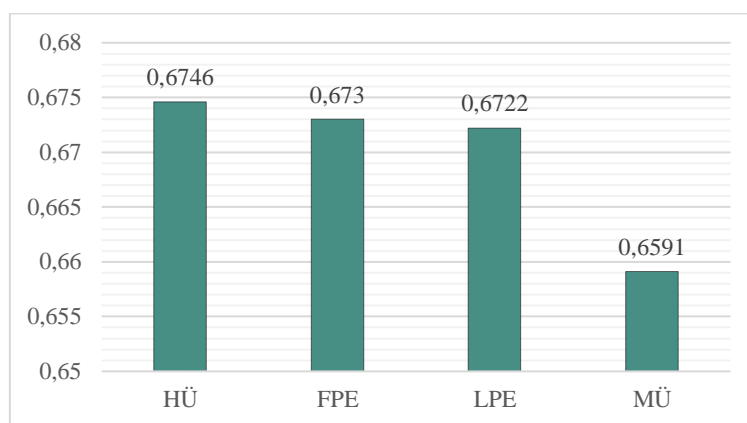


Abbildung 8: TTR nach Übersetzungsmodus für alle Texte

Die Ergebnisse scheinen die aufgestellte Hypothese zu bestätigen, dass der TTR in posteditierten Texten geringer ist als in Humanübersetzungen. Auch die Erwartung, dass der TTR in maschinell übersetzten Texten am geringsten sein würde, wurde erfüllt. Den höchsten TTR erzielt der Übersetzungsmodus HÜ mit 0,6746, gefolgt von FPE (0,6730), LPE (0,6722) und MÜ (0,6591). Die erwartete Progression MÜ<LPE<FPE<HÜ lässt sich also tatsächlich beobachten.

Das durchschnittliche Ergebnis nach Übersetzungsmodus für alle medizinischen Texte ist in Abbildung 9 dargestellt.

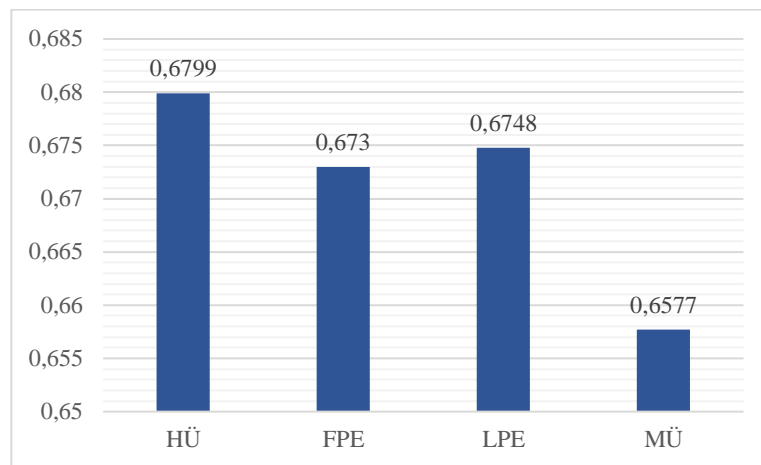


Abbildung 9: TTR nach Übersetzungsmodus für alle med. Texte.

Auch hier ist der TTR für humanübersetzte Texte mit 0,6799 am höchsten und für maschinell übersetzte Texte mit 0,6577 am niedrigsten. Der TTR von posteditierten Texten liegt erwartungsgemäß dazwischen (0,6730 bei FPE; 0,6748 bei LPE). Unerwartet ist an dieser Stelle nur, dass LPE-Texte einen (geringfügig) höheren TTR als FPE-Texte haben.

In Abbildung 10 kann das Gesamtergebnis für die technischen Texte abgelesen werden.

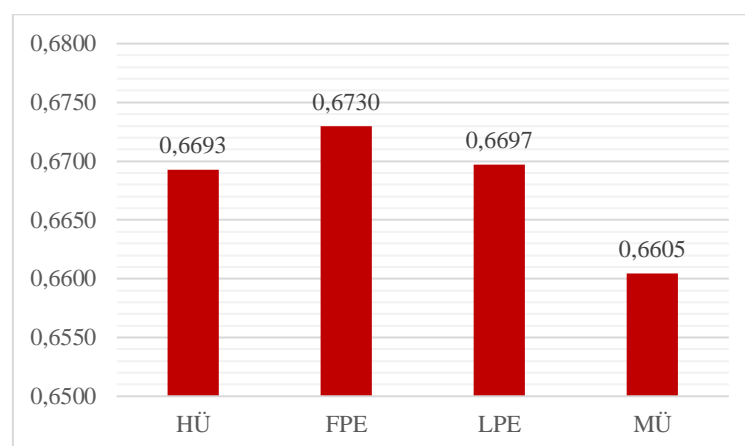


Abbildung 10: TTR nach Übersetzungsmodus für alle techn. Texte

Die Ergebnisse der technischen Texte bestätigen die Hypothese nur teilweise. Zwar ist der TTR für posteditierte Texte mit 0,6730 (FPE) bzw. 0,6667 (LPE) höher als für maschinell übersetzte Texte (0,6605), aber HÜ ist mit 0,6693 entgegen den Erwartungen nicht der Übersetzungsmodus mit dem höchsten TTR, sondern nur auf Rang 3.

Um genauer nachzuvollziehen zu können, wie die Ergebnisse zustande kommen, werden die Ergebnisse nun aufgeschlüsselt nach Ausgangstext betrachtet (Abbildung 11).

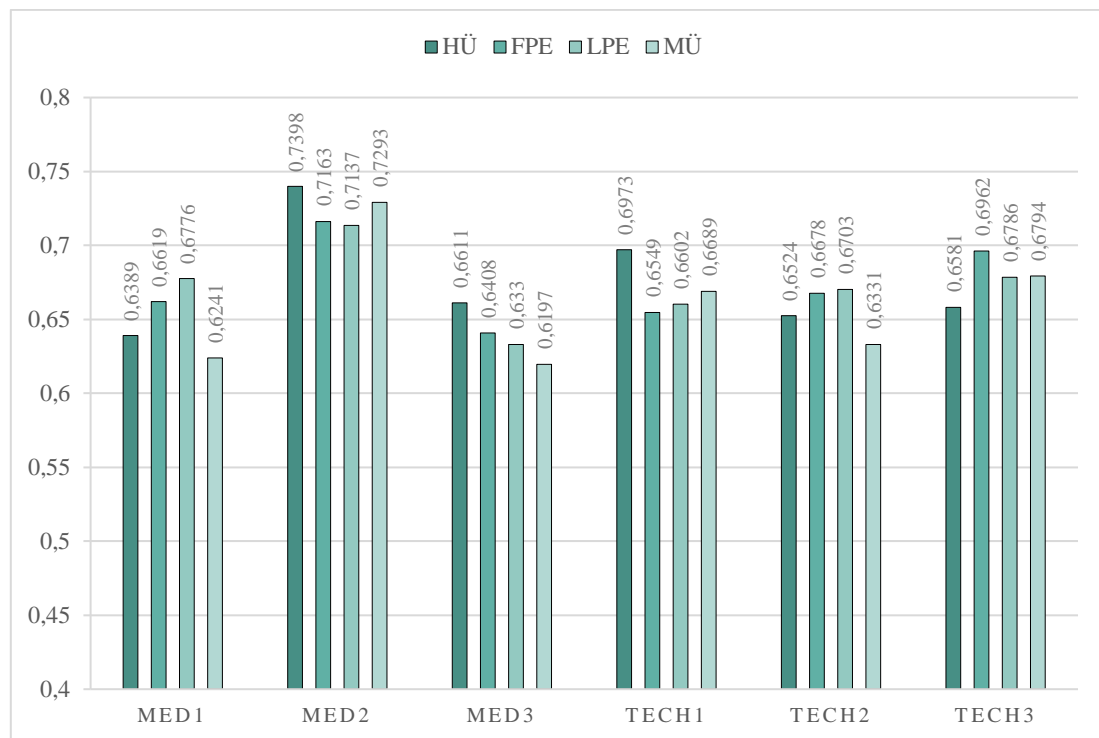


Abbildung 11: TTR nach Übersetzungsmodus für alle Texte, aufgeschlüsselt

In dieser Aufschlüsselung ist kein eindeutiges Muster zu erkennen. Die Hypothese, dass die lexikalische Diversität in humanübersetzten Texten am höchsten, geringer in posteditierten Texten und wiederum geringer in maschinell übersetzten Texten ist, wird nur durch MED3 vollständig bestätigt: Hier haben die HÜ-Texte den durchschnittlich höchsten TTR, gefolgt von FPE-, LPE- und MÜ-Texten, d. h. die erwartete Progression ist exakt eingetroffen.

Der erste Teil der Hypothese, d. h. die Annahme, humanübersetzte Texte haben einen höheren TTR als posteditierte Texte, bestätigt sich in dieser Einzelansicht insgesamt nicht: HÜ ist zwar in drei Fällen der Übersetzungsmodus mit dem höchsten TTR (MED2, MED3, TECH1), in den anderen drei (MED1, TECH2,

TECH3) aber der mit dem niedrigsten oder zweitniedrigsten. Für den zweiten Teil der Hypothese, die besagt, MÜ-Texte habe den geringsten TTR, sieht es ähnlich aus: MÜ ist zwar in drei Fällen (MED1, MED3, TECH2) der Übersetzungsmodus mit dem niedrigsten TTR, in den anderen drei (MED2, TECH1, TECH3) aber der mit dem höchsten oder zweihöchsten. Eine nähere Interpretation dieser Ergebnisse und eine Diskussion möglicher Gründe für dieses Ergebnis folgt in Abschnitt 4.1.3.1.

#### 4.1.2 Lexikalische Dichte

Die Gesamtergebnisse für die durchschnittliche lexikalische Dichte nach Übersetzungsmodus können Abbildung 12 entnommen werden.

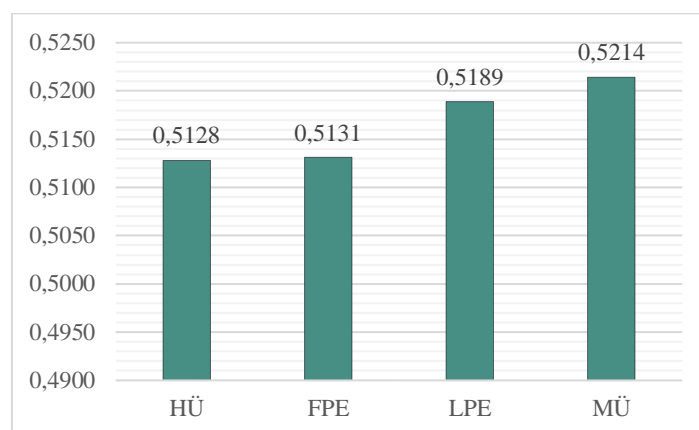


Abbildung 12: Lexikalische Dichte nach Übersetzungsmodus für alle Texte

Die Hypothese, dass die lexikalische Dichte von maschinell übersetzten und posteditierten Texten geringer ist als von humanübersetzten Texten, wird durch die vorliegenden Ergebnisse nicht bestätigt. Statt der erwarteten Progression PE<HÜ und MÜ<HÜ kann die Progression HÜ<FPE<LPE<MÜ beobachtet werden: Die HÜ-Texte haben mit einem Wert von 0,5128 nicht wie erwartet eine höhere lexikalische Dichte als posteditierte und maschinell übersetzte Texte, sondern eine niedrigere (FPE = 0,5131, LPE = 0,5189, MÜ = 0,5214).

Es folgt die Betrachtung der Ergebnisse aufgeschlüsselt in medizinische und technische Texte (Abbildung 13 und 14).

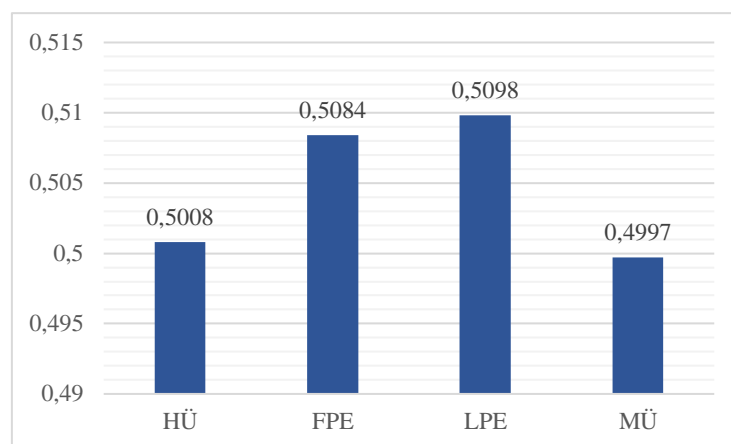


Abbildung 13: Lexikalische Dichte nach Übersetzungsmodus für alle med. Texte

Betrachtet man ausschließlich die medizinischen Texte, bestätigt sich die Hypothese ebenfalls nicht. Zwar hat das unbearbeitete MÜ-Output mit 0,4997 wie erwartet eine geringere lexikalische Dichte als die HÜ-Texte (0,5008), aber die lexikalische Dichte der PE-Texte liegt mit 0,5084 (FPE) und 0,5098 (LPE) unerwartet über der lexikalischen Dichte der HÜ-Texte.

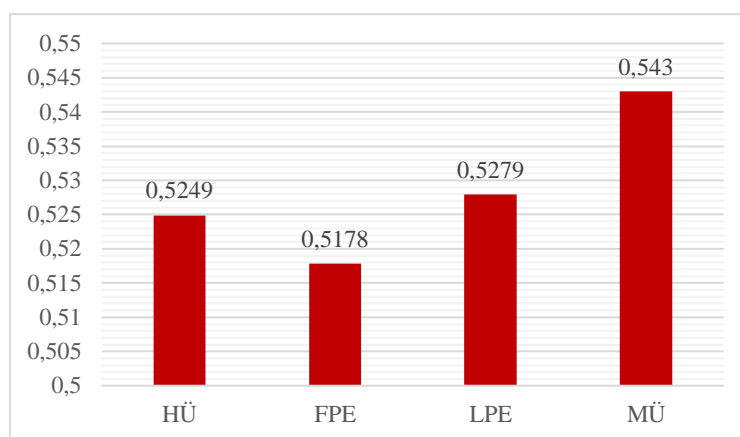


Abbildung 14: Lexikalische Dichte nach Übersetzungsmodus für alle techn. Texte

Auch bei den technischen Texten (Abbildung 14) bestätigt sich die Hypothese nicht. Die lexikalische Dichte der HÜ-Texte liegt mit 0,5249 zwar höher als die der FPE-Texte (0,5178), aber unterhalb der Werte von LPE (0,5279) und MÜ (0,5430). Mögliche Gründe für diese unerwarteten Ergebnisse und Progressionen werden in Abschnitt 4.1.3.2 erläutert.

Die aufgeschlüsselten Ergebnisse nach Ausgangstext können Abbildung 15 entnommen werden.



Abbildung 15: Lexikalische Dichte nach Übersetzungsmodus für alle Texte, aufgeschlüsselt

In der Einzelansicht nach Ausgangstext ergibt sich kein klares Muster. In MED3 fällt auf, dass die Werte der Übersetzungsmodi sehr nah beieinander liegen, sodass kaum eine Progression erkennbar ist. Die Hypothese, dass humanübersetzte Texte eine höhere lexikalische Dichte haben als posteditierte und maschinell übersetzte Texte wird nur von TECH1 (und teilweise MED3) bestätigt. In den meisten anderen Fällen haben die PE- und MÜ- Texte eine höhere lexikalische Dichte als die HÜ-Texte. Zwischen den PE-Übersetzungsmodi und der MÜ lässt sich – wie bei Toral – kein klares Muster erkennen: Das MÜ-Output hat in zwei Fällen (MED1 und MED2) eine geringere lexikalische Dichte als die Texte beider PE-Modi – in den anderen vier Fällen ist aber die lexikalische Dichte bei beiden PE-Modi geringer als die der HÜ-Texte. In MED3, TECH2 und TECH3 erzielt das MÜ-Output völlig entgegen den Erwartungen sogar die höchste lexikalische Dichte.

Auffällig ist außerdem, dass sich in den technischen Texten (insb. in TECH2 und TECH3) die Progression MÜ>LPE>FPE>HÜ zeigt, d. h. die geringste lexikalische Dichte haben HÜ-Texte, gefolgt von FPE-, LPE- und MÜ. Diese Progression war auch bei den Gesamtergebnissen (siehe Abbildung 12) zu erkennen. Aufbauend auf der eigentlichen Hypothese hätte ich eher eine entgegengesetzte Progression

(MÜ<LPE<FPE<HÜ) erwartet, d. h. je mehr der Mensch involviert ist, desto höher die lexikalische Dichte.

Es lässt sich also nicht feststellen, dass die lexikalische Dichte in Humanübersetzungen höher ist als in posteditierten und maschinell übersetzten Texten. Die möglichen Gründe hierfür werden im folgenden Abschnitt 4.1.3 näher erläutert.

### **4.1.3 Einordnung der Ergebnisse zu Simplification**

Es folgt eine Einordnung der bis hier hin vorgestellten Ergebnisse. Mögliche Gründe und Limitationen werden diskutiert und Perspektiven für künftige Studien eröffnet, bevor in einem Fazit resümiert wird, ob Simplification in den posteditierten Texten dieses Korpus in höherem Maße als in den humanübersetzten Texten nachgewiesen werden konnte.

#### **4.1.3.1 Lexikalische Diversität**

Zunächst einmal scheinen die Gesamtergebnisse die Hypothese zu bestätigen: Die geringste lexikalische Diversität wurde beim Übersetzungsmodus MÜ festgestellt, gefolgt von LPE, FPE und HÜ. Gemäß Toral lassen sich diese Ergebnisse wie folgt interpretieren: Die geringste lexikalische Diversität hat der Übersetzungsmodus MÜ, da MÜ-Engines in der Regel die Übersetzungsvarianten wählen, die am häufigsten in den Trainingsdaten vertreten waren (Farrell 2018, 57) (vgl. Abschnitt 3.2.1.1). Dadurch wird der Text lexikalisch vergleichsweise wenig divers. Durch leichtes und vollständiges Post-Editing kann diese geringe lexikalische Diversität erhöht werden, allerdings nicht so stark, dass das Niveau von Humanübersetzungen erreicht wird (Toral 2019, 276). Dementsprechend war die Erwartung, dass die Progression der TTR für die verschiedenen Übersetzungsmodi mit dem Einfluss des Menschen beim jeweiligen Übersetzungsmodus korrelieren würde, d. h. je mehr der Mensch involviert ist, desto höher die lexikalische Diversität und somit der TTR. Diese erwartete Progression (MÜ<LPE<FPE<HÜ) wurde in den Gesamtergebnissen beobachtet.

Allerdings lassen sich bei der Betrachtung der aufgeschlüsselten Ergebnisse keine eindeutigen Muster oder Progressionen erkennen. Fraglich ist, ob die kumulierten



Gesamtergebnisse trotzdem Gültigkeit haben. Die vorgestellten Ergebnisse zum Type-Token-Ratio (siehe Abschnitt 4.1.1) lassen an dieser Stelle keine eindeutigen Schlüsse zu, sondern müssen durch weitere Untersuchungen überprüft werden.

Wie in Abschnitt 3.2.1.1 bereits erwähnt wurde, hat die Verwendung des TTR zum Nachweis von Simplification durchaus Limitationen. So sieht Steiner (2004, 164) ein Problem bei der Anwendung des TTR auf kurze Texte. Bei den Texten, über die er die Aussage „type-token measurements of texts of such short length are not valid, for obvious reasons“ getätigt hat, handelte es sich um Texte von ein bis zwei Sätzen (ebd., 150). Deren Ergebnisse haben „no meaning“, da die Texte vermutlich nicht repräsentativ für ein größeres Korpus seien (ebd.). In einem nächsten Schritt untersuchte Steiner einen englischen Ausgangstext und seine Übersetzung (ebd., 171/172). Diese Texte umfassen je etwa 200 Wörter und sind somit ca. 39 % länger als die vorliegenden Korpustexte. Steiners Zurückhaltung bei der Formulierung von Schlussfolgerungen auf Grundlage dieser Ergebnisse scheint sich nun auf die Korpusgröße (ein AT, ein ZT) und nicht auf die Längen der Korpustexte zu beziehen: „Once again, these absolute figures do not mean anything here, because the ‚corpus‘ is too small“ (ebd., 172). Es lässt sich also nicht final beantworten, ob das vorliegende Korpus (6 x AT, 6 x MÜ, 21 x HÜ, 21 x FPE, 21 x LPE) groß genug und die Korpustexte (Ø 144 Wörter) lang genug sind, um repräsentative Ergebnisse zu erzielen. Die Tatsache, dass beim STTR Blöcke von 1000 Wörtern gebildet werden legt nahe, dass 144 Wörter pro Text nicht ausreichend sein könnten.

Ob und in welchem Maße die geringe Länge der Korpustexte die Belastbarkeit der Ergebnisse dieser Korpusanalyse beeinflusst, muss letztlich in künftigen Untersuchungen mit größeren Korpora geklärt werden.

Auch die Grundvoraussetzung, dass Texte ähnlich lang sein müssen, wurde bereits in Abschnitt 3.2.1.1 erwähnt. Genaue Angaben über den Grad der notwendigen Ähnlichkeit gibt es nicht. Die durchschnittlichen Textlängen nach Übersetzungsmodus sind im vorliegenden Korpus wie folgt: AT: Ø 148,83 Wörter; MÜ: Ø 139 Wörter; HÜ: Ø 144,57 Wörter; FPE: Ø 144,69 Wörter; LPE: Ø 141,43 Wörter. Wenn man die durchschnittliche Wortzahl der verschiedenen Übersetzungsmodi vergleicht, fällt auf, dass die MÜ-Texte durchschnittlich 2,4 bis 5,2

Wörter weniger haben als die Texte der anderen Übersetzungsmodi. Das mag nach einer geringen Differenz klingen, aber kann sich bei einem so kleinen Korpus durchaus stark auf den TTR auswirken.

Um herauszufinden, warum die MÜ-Varianten durchschnittlich weniger Wörter haben als alle anderen Texte, schauen wir uns beispielhaft die maschinellen Übersetzungen zweier Ausgangstexte (TECH1\_MUE und MED2\_MUE) genauer an: Beim Abgleich von MED2\_MUE und MED2\_AT fällt auf, dass immer wieder Wörter ausgelassen, d. h. nicht übersetzt wurden, z. B. ‚is‘ in Satz 1 ‚use‘ in Satz 6, ‚not‘ in Satz 7, ‚exercising‘ in Satz 11 und ‚(carry) on‘ in Satz 12. Es ist bekannt, dass MÜ-Engines dazu neigen, Wörter auszulassen (Castilho et al. 2017, 114; Yang et al. 2019, 6191). Zur Erinnerung: Die Ergebnisse von MED2 bestätigen die aufgestellte Hypothese nur teilweise. Die Progression LPE<FPE<HÜ entspricht der Hypothese, lediglich der TTR des MÜ-Outputs ist ‚zu hoch‘. Wenn man bei MED2\_MUE nun die fünf ausgelassenen Wörter zur Gesamtzahl der Wörter (Tokens) addiert und die Anzahl der Types entsprechend anpasst, so sinkt der TTR des MÜ-Outputs deutlich, und zwar beinahe unter den TTR-Wert der beiden PE-Modi, sodass das MÜ-Output beinahe den geringsten TTR hätte und die Hypothese rundum bestätigt wäre. An diesem Beispiel zeigt sich, dass bei einem Korpus von so geringer Größe bzw. bei Korpustexten von so geringer Länge schon wenige Wörter mehr oder weniger (sowohl bei den Types als auch den Tokens) stark ins Gewicht fallen.

Außerdem ist denkbar, dass die Textsorte (hier: medizinische und technische Fachtexte) einen Einfluss auf die Ergebnisse für die lexikalische Diversität haben. Es lässt sich in Frage stellen, ob der TTR für Fachtexte überhaupt in dieser Form geeignet ist. Farrell sagt dazu:

„Variety and inventiveness are not always desirable features in every kind of text. For example, excessive lexical variation might make a smartphone user’s guide more difficult to follow“ (Farrell 2018, 58).

Fachtexte wie Packungsbeilagen und Bedienungsanleitungen fallen sicherlich auch in die Kategorie der Textsorten, in denen „excessive lexical variation“ nicht gewünscht ist, sondern Einheitlichkeit. Auch Toral (2019, 280) gesteht ein, dass eine Gleichsetzung von hoher lexikalischer Diversität mit hoher Qualität, so wie auch er impliziert hat, nicht immer sinnvoll ist: „we acknowledge that in translation there is a

tension between diversity and consistency, especially in technical translation". Diese terminologische Konsistenz war, wie bereits in Abschnitt 3.1.2 erwähnt, sogar Teil der Arbeitsanweisung bei der Erstellung der FPE-Texte dieses Korpus. Umso überraschender ist es, dass das Wort „dishwasher“ in keiner der FPE- und LPE-Varianten der drei technischen Texte einheitlich übersetzt wurde, sondern verschiedene Terme genutzt wurden. Die auch im MÜ-Output vorhandenen Varianten „Geschirrspüler“, „Geschirrspülmaschine“ und „Spülmaschine“ wurden also von keinem Posteditor vereinheitlicht. Eine ähnliche Beobachtung machten auch Čulo et al. (2014) in ihrem von Translationsstudenten erstellten Korpus. In den Humanübersetzungen des vorliegenden Korpus gab es diese Inkonsistenz in 10 von 12 Fällen nicht, sondern ein Begriff wurde einheitlich verwendet. Für den TTR der posteditierten Texte bedeutet diese Unaufmerksamkeit, dass jeder technische FPE- und LPE-Text ein bis zwei Types weniger haben müsste, wodurch auch der TTR für die posteditierten Texte sinken würden. Bei Korpus-texten von so geringer Länge kann schon das Wahren oder Nicht-Wahren bzw. das Herstellen oder Nicht-Herstellen der terminologischen Konsistenz Auswirkungen auf den TTR und damit die lexikalische Diversität haben. Bei längeren Korpus-texten würden solche unerwünschten lexikalischen Inkonsistenzen nicht so stark ins Gewicht fallen

Tatsächlich könnte man argumentieren, die Tatsache, dass die lexikalische Diversität in den vorliegenden posteditierten FPE-Texten mitunter höher ist als erwartet (und insb. in zwei Fällen höher als in den humanübersetzten Texten), treffe in diesem Fall keine Aussage über die Existenz von Simplification, sondern deute lediglich darauf hin, dass die Posteditoren ihre Aufgabe in diesem Fall nicht entsprechend der Aufgabenstellung bzw. nicht gewissenhaft erfüllt haben (siehe Abschnitt 3.1). Insofern könnte also auch die Unerfahrenheit der Korpus-textproduzenten die Ergebnisse beeinflusst haben. Čulo et al. betonen in diesem Kontext die Relevanz des Erlernens von Post-Editing-Fähigkeiten (vgl. Abschnitt 2.1.8):

„Correcting MT-specific errors in single segments and ensuring consistency over the stretch of a text may thus be a task translators need to be trained for specifically, as they require heightened levels of effort“ (Čulo et al. 2017, 197).

Außerdem mutmaßen sie, dass Posteditoren das MÜ-Output – trotz der angeblichen Skepsis gegenüber MÜ (vgl. Abschnitt 2.1.7) – nicht kritisch genug rezipieren,

sondern sich zu sehr darauf verlassen (ebd.). Um der Frage nachzugehen, ob sich Übersetzungen weniger erfahrener Übersetzer tatsächlich signifikant von Übersetzungen erfahrener Übersetzer unterscheiden, bedarf es weiterer Korpusanalysen, in denen auch die Texte dieser beiden Personengruppen verglichen werden.

Ein abschließendes Urteil darüber, ob Simplification anhand vom TTR für dieses Korpus nachgewiesen werden konnte, ist an dieser Stelle nicht möglich. Für künftige Studien empfiehlt es sich, größere Korpora mit längeren und möglichst gleich langen (Fach-)Texten zu untersuchen. Idealerweise würden auch Texte von Textproduzenten mit unterschiedlich viel Erfahrung gegenübergestellt werden. So könnten die genannten offen gebliebenen Fragen mit größerer Sicherheit beantwortet werden.

#### **4.1.3.2 Lexikalische Dichte**

Zunächst einmal muss festgestellt werden, dass alle Werte für die lexikalische Dichte deutlich über 0,4, d. h. über dem Minimal-Referenzwert von Stubbs (1996, 72/73) liegen. Grund dafür ist vermutlich der recht hohe Fachlichkeitsgrad der Korpustexte, für den eine höhere lexikalische Dichte (d. h. Informationsdichte) üblich ist (ebd.).

Wie bereits in Abschnitt 4.1.2 dargelegt, hat sich die Hypothese, dass posteditierte Texte eine geringere lexikalische Dichte als humanübersetzte Texte haben, anhand der Ergebnisse nicht bestätigt. Bei der Betrachtung der Gesamtergebnisse konnte die Progression  $HÜ < FPE < LPE < MÜ$  festgestellt werden, die der aufgestellten Hypothese ( $PE < HÜ$ ,  $MÜ < HÜ$ ) klar widerspricht. Die Progression bedeutet, dass humanübersetzte Texte die geringste lexikalische Dichte haben, gefolgt von den beiden PE-Modi und der maschinellen Übersetzung. Fraglich ist jedoch, ob es sich bei dieser Progression, die sich neben den Gesamtergebnissen auch in TECH2 und TECH3 gezeigt hat, schon um ein Muster handelt und ob die kumulierten Ergebnisse Gültigkeit haben, obwohl sich die gleiche Progression nur teilweise in den aufgeschlüsselten Ergebnissen erkennen lässt.

Die Hypothese, dass posteditierte und maschinell übersetzte Texte eine geringere lexikalische Dichte als humanübersetzte Texte aufweisen würde, fußte auf der

Hypothese, Post-Editese sei eine verschärfte Form der Translationese und würde demnach die Translationese-Merkmale in besonderem Maße vorweisen.

An dieser Stelle möchte ich noch einmal auf Bakers (1996, 183) Begründung für die Sinnhaftigkeit der lexikalischen Dichte als Größe zum Nachweis von Simplification zurückkommen (siehe Abschnitt 3.2.1.2): „using more grammatical and fewer lexical words is a way of building in more redundancy and making a text easier to process“. Eine MÜ-Engine fügt keine Konjunktionen hinzu und baut auch keine Redundanzen ein, um einen Text besser verständlich oder leichter verarbeitbar zu machen. MÜ-Engines sind nicht in der Lage, Faktoren wie Verständlichkeit oder bessere Lesbarkeit zu berücksichtigen. Vielmehr ist bekannt, dass sie sich streng am Ausgangstext orientieren und in der Regel keine Wortartwechsel<sup>17</sup> vornehmen (Lapshinova-Koltunski 2015, 106/107). Insofern war Torals Hypothese, maschinell übersetzte und posteditierte Texte würden mehr Simplification (d. h. eine geringere lexikalische Dichte) als humanübersetzte Texte aufweisen, nicht vollkommen schlüssig. In der Konsequenz war es überraschend, dass sich seine Hypothese bestätigte. Denn auf Grundlage von Bakers Zitat wäre es eigentlich eher zu erwarten gewesen, dass HÜ-Texte eine geringere lexikalische Dichte als PE- und MÜ-Texte haben, weil Humanübersetzer am ehesten dazu in der Lage sind, solche bearbeitenden Eingriffe wie das Hinzufügen von Konjunktionen oder Redundanzen, die die lexikalische Dichte verringern, vorzunehmen. Dieser Art der Änderungen beruhen auf durch und durch menschlichen Eigenschaften, da dafür tatsächliches Verständnis notwendig ist. Für das Post-Editing bieten sich Änderungen dieser Art eigentlich schon aufgrund der Devise, dass nur notwendige Änderungen vorgenommen werden sollen, nicht an (vgl. Abschnitt 2.1.2). Dieser Logik folgend müssten HÜ-Texte die geringste Dichte haben, gefolgt von posteditierten und maschinell übersetzten Texten. Wenn die Hypothese so gelautet hätte, wäre sie von den Ergebnissen aus TECH2 und TECH3 sowie von den Gesamtergebnissen bestätigt worden: Hier haben die HÜ-Texte die geringste lexikalische Dichte und die MÜ-Outputs die höchste. Die lexikalischen Dichten der posteditierten Texte liegen dazwischen. In der Konsequenz würde das bedeuten, dass Simplification – zumindest wenn sie über die lexikalische Dichte nachgewiesen werden soll – kein Merkmal der

---

<sup>17</sup> Wortartwechsel könnten zu einer abweichenden lexikalischen Dichte führen, da die Formel für die lexikalische Dichte auf Wortartenzugehörigkeit (content words) basiert.

Post-Editese ist, bzw. zumindest kein stark ausgeprägtes. Es sei hier darauf hingewiesen, dass es sich bei diesen Überlegungen um Vermutungen handelt, die empirisch dringend belegt werden müssten.

Ein Kritikpunkt an der Methode zur Untersuchung der lexikalischen Dichte ist bereits in Abschnitt 3.2.1.2 angeklungen: Laut Steiner (2004, 133f) sind wortbasierte Untersuchungsmethoden, wie auch die lexikalische Dichte eine ist, generell kritisch zu sehen. Für das Sprachpaar Englisch-Deutsch stellen die sprachsystembezogenen Unterschiede bei der Kompositabildung ein Problem dar. In jedem der vorliegenden Korpustexte gibt es mehrere Beispiele für diese Problematik, insbesondere im Bereich der Nominalkomposita <sup>18</sup> : „vaccination program“ – „Impfprogramm“ (MED1), „immune system“ – „Immunsystem“ (MED1), „long-acting insulin“ – „Langzeitinsulin“ (MED2), „side effects“ – „Nebenwirkungen“ (MED3), „safety instructions“ – „Sicherheitshinweise“ (TECH1), „water sensor“ – „Wassersensor“ (TECH2), „foreign object“ – „Fremdkörper“ (TECH3), u. v. m. Da englische Nominalkomposita in der Regel nicht zusammen geschrieben werden, zählen sie in der Berechnung von Sketch Engine als zwei Nomen – und damit als zwei content words. In den deutschen Zieltexten wird jedes Nominalkompositum als nur ein Nomen und damit als ein content word gezählt. Da es in jedem Text zahlreiche Beispiele für dieses Phänomen gibt, muss man davon ausgehen, dass die Ergebnisse gewissermaßen verzerrt sind. Konkret ist die Anzahl der Nomen – die bekanntermaßen zu den content words gezählt werden (vgl. Abschnitt 3.2.1.2) – in den englischen Ausgangstexten nach oben verzerrt. Diese scheinbar höhere Nominalität der englischen Ausgangstexte im Vergleich zu den deutschen Zieltexten nenne ich im Folgenden ‚Pseudonominalität‘. Die Frage, wie stark diese Pseudonominalität die Ergebnisse verfälscht, kann in dieser Arbeit nicht beantwortet werden. Die geringe Länge der Korpustexte sorgt mutmaßlich dafür, dass diese durch die unterschiedlichen Sprachsysteme ausgelösten Diskrepanzen besonders stark ins Gewicht fallen, da schon ein Unterschied von wenigen content words mehr oder weniger einen starken Einfluss auf die lexikalische Dichte haben kann.

---

<sup>18</sup> Im Rahmen dieser Arbeit beschränke ich mich auf die Untersuchung von Nominalkomposita, auch wenn es sicherlich auch Beispiele für andere Arten von Komposita gibt.

Außerdem sei noch einmal darauf hingewiesen, dass auf Grundlage der Daten eigentlich nicht eindeutig konkludiert werden kann, ob die lexikalische Dichte in den untersuchten Texten Simplification oder Explicitation nachweist (vgl. Abschnitt 3.2.1.2). Dafür müsste etwa die Anzahl der Konjunktionen zwischen den Texten verglichen und die alignierten Texte auch qualitativ untersucht werden. Dies ist in dieser Arbeit nicht leistbar.

Auch für die lexikalische Dichte lässt sich also kein abschließendes Urteil fällen, da die Ergebnisse vermutlich verzerrt sind. Zudem sind sie nicht homogen genug, um eindeutige Muster erkennen zu lassen. In künftigen Studien müssen also längere Korpus-texte und größere Korpora verwendet und die Methoden ggf. so angepasst werden, dass sie den Unterschieden bei der Kompositabildung Rechnung tragen.

#### **4.1.3.3 Fazit zu Simplification**

Wie die Einordnung der Ergebnisse zur lexikalischen Diversität und lexikalischen Dichte gezeigt hat, konnte anhand dieses Korpus nicht zweifelsfrei nachgewiesen werden, dass das Translationese-Merkmal Simplification auch ein Merkmal der Post-Editese ist. Folglich kann auch die Frage, ob es sich bei Post-Editese um eine verschärfte Form der Translationese handelt, hier noch nicht beantwortet werden.

Bei der Untersuchung der lexikalischen Diversität über den TTR ließ sich nicht abschließend klären, ob die kumulierten Ergebnisse, die die Hypothese zu bestätigen schienen, auch Gültigkeit haben, obwohl die gleichen Muster und Progressionen in den nach Texten aufgeschlüsselten Ergebnissen nicht erkennbar waren. Vor allem die Länge der Korpus-texte und die Größe des Korpus erwiesen sich als limitierende Faktoren in der Interpretation der Ergebnisse, da bereits wenige Worte Unterschied bei den Types oder Tokens einen recht starken Einfluss auf die errechneten Werte haben. Würde es gelingen, diese Ergebnisse anhand eines größeren Korpus mit längeren Texten zu reproduzieren, würde dies dafür sprechen, dass Simplification auch ein Merkmal der Post-Editese ist und dass Post-Editese als verschärfte Form der Translationese verstanden werden kann.

Anhand der lexikalischen Dichte ist es ebenfalls nicht gelungen, Simplification in posteditierten Texten nachzuweisen. Auch hier bedarf es weiterer Untersuchungen, um zu prüfen, ob die in den Gesamtergebnissen festgestellte Progression

MÜ>LPE>FPE>HÜ, die entgegen der ursprünglichen Hypothese ist, als Muster interpretiert werden kann, obwohl sie nicht in allen Texten zu finden ist. Wäre dies der Fall, wäre das ein Widerspruch zu Torals Hypothese, dass Simplification auch ein Merkmal der Post-Editese ist. Unterschiede in den Sprachsystemen bei der Kompositabildung wurden als Faktor identifiziert, der die Ergebnisse mutmaßlich verzerrt.

Insgesamt wird deutlich, dass die Ergebnisse der lexikalischen Diversität und Dichte zu widersprüchlich sind, als dass sie Simplification in posteditierten Texten nachweisen könnten. Zur Kontrolle der Ergebnisse und zur Klärung offengebliebener Fragen bedarf es weiterer Studien mit längeren Korpus-texten und größeren Korpora.

## 4.2 Interference

In diesem Abschnitt sollen die Ergebnisse zu den beiden Untersuchungsaspekten, die Interference nachweisen sollen, vorgestellt werden. Eine Diskussion und Einordnung dieser Ergebnisse erfolgen in Abschnitt 4.2.3.

### 4.2.1 Verhältnis der Verhältnisse von nominalen und verbalen Wortarten zwischen AT und ZT

Die Gesamtergebnisse für das Verhältnis der Nominal-Verbal-Verhältnisse von AT und ZT können Abbildung 16 entnommen werden.

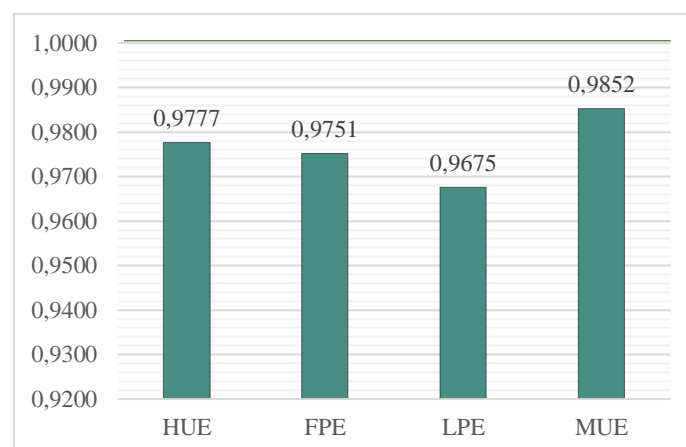


Abbildung 16: Verhältnis der Nominal-Verbal-Verhältnisse nach Übersetzungsmodus für alle Texte



Zunächst einmal fällt auf, dass alle Durchschnittswerte unter 1 liegen, was überraschend ist. Erklärungen dafür werden in Abschnitt 4.2.3.1 diskutiert.

Die Hypothese war, dass die Progression  $M\ddot{U} < LPE < FPE < H\ddot{U}$  zu beobachten sein würde, wobei auch angenommen wurde, dass alle Werte über 1 liegen würden. Die erwartete Progression ist nicht eingetreten. Der Wert für die MÜ ist mit 0,9852 nicht der niedrigste, sondern der höchste. Außerdem ist HÜ nicht wie erwartet der Übersetzungsmodus mit dem geringsten Wert, sondern der Wert (0,9777) liegt nur knapp hinter dem des MÜ-Outputs. Die Werte für die posteditierten Texte liegen dazwischen ( $FPE = 0,9751$  und  $LPE = 0,9675$ ). Gründe für diese überraschende Progression werden ebenfalls in Kapitel 4.2.3.1 thematisiert.

Die Ergebnisse für die medizinischen und technischen Texte können Abbildung 17 und 18 entnommen werden.

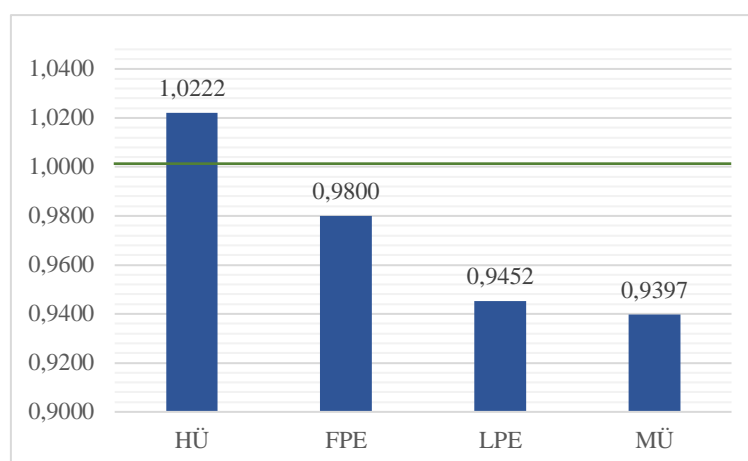


Abbildung 17: Verhältnis der Nominal-Verbal-Verhältnisse nach Übersetzungsmodus für alle med. Texte

Im Gesamtergebnis für die medizinischen Texte scheint sich die aufgestellte Hypothese auf den ersten Blick zu bestätigen, da die erwartete Progression  $M\ddot{U} < LPE < FPE < H\ddot{U}$  eingetroffen ist. Allerdings war die Annahme, dass alle Werte über 1 liegen – hier liegt nur der Wert der HÜ-Texte über 1. Konkret wurde für den Übersetzungsmodus HÜ ein Wert von 1,0222 errechnet, gefolgt von 0,9800 für FPE, 0,9452 für LPE und 0,9397 für MÜ.

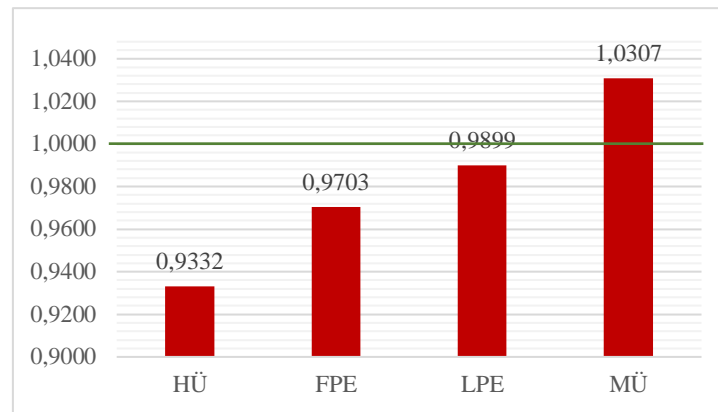


Abbildung 18: Verhältnis der Nominal-Verbal-Verhältnisse nach Übersetzungsmodus für alle techn. Texte

In der Gesamtansicht der technischen Texte zeigt sich ein völlig anderes Bild als bei den medizinischen Texten. Auch in diesem Fall liegen alle Werte – mit Ausnahme des Werts des MÜ-Outputs – anders als erwartet – unterhalb von 1. Die Progression ist allerdings genau andersherum und entgegen der Hypothese  $MÜ < LPE < FPE < HÜ$ . HÜ ist mit einem Wert von 0,9332 der Übersetzungsmodus mit dem geringsten Wert für das Verhältnis der Nominal-Verbal-Verhältnisse, gefolgt von FPE (0,9703), LPE (0,9899) und MÜ (1,0307).

Schließlich folgt die Betrachtung der aufgeschlüsselten Ergebnisse nach Ausgangstext (Abbildung 19).

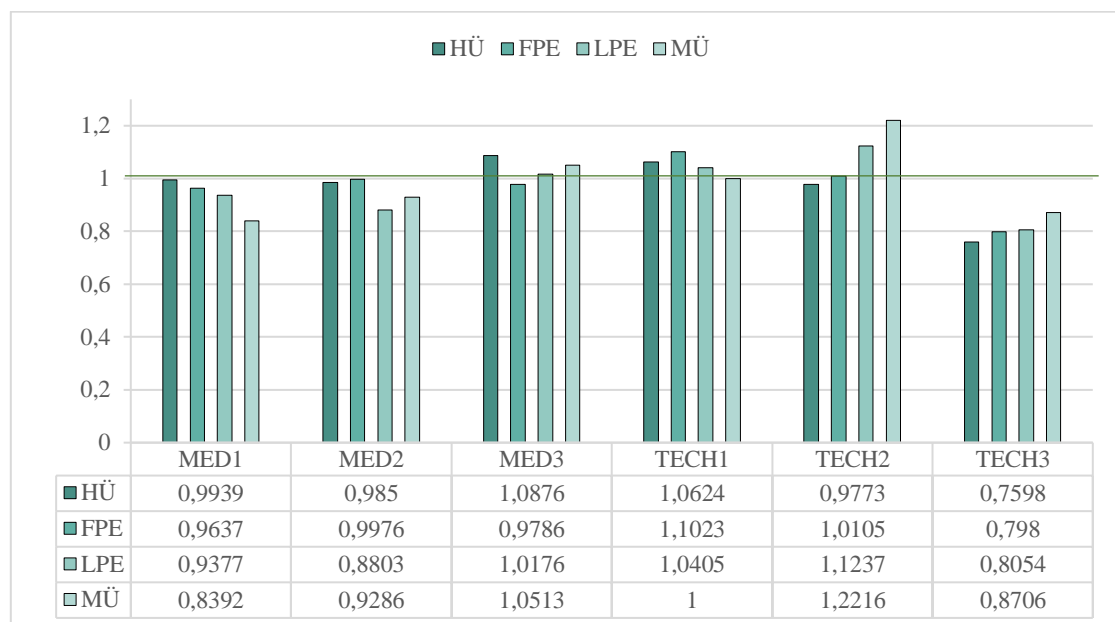


Abbildung 19: Verhältnis der Nominal-Verbal-Verhältnisse nach Übersetzungsmodus für alle Texte, aufgeschlüsselt

Auch in den aufgeschlüsselten Ergebnissen zeigt sich keine eindeutige Progression. Viele Werte liegen unter 1 – in MED1, MED2 und TECH3 liegen sogar alle Werte unter 1. Für HÜ-Texte wurde nur in zwei Fällen (MED1, MED3) tatsächlich der höchste Wert festgestellt. Die MÜ war ebenfalls nur in zwei Fällen der Übersetzungsmodus mit dem niedrigsten Wert (MED1, TECH1). Die erwartete Progression zeigt sich nur in MED1. In TECH2 und TECH3 ist sogar eine im Vergleich zur Hypothese entgegengesetzte Progression zu beobachten. Eine nähere Interpretation der Ergebnisse erfolgt in Abschnitt 4.2.3.1.

#### 4.2.2 Verhältnis der Satzlängen zwischen AT und ZT

Das Gesamtergebnis für die durchschnittlichen Satzlängenverhältnisse nach Übersetzungsmodus kann Abbildung 20 entnommen werden.

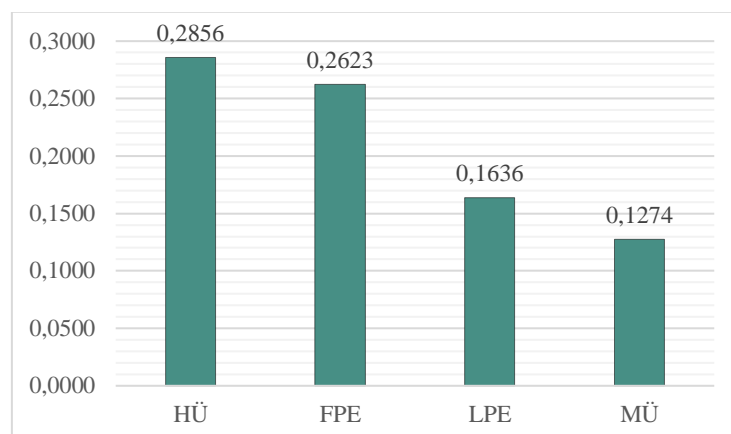


Abbildung 20: Satzlängenverhältnis nach Übersetzungsmodus für alle Texte

Die Hypothese lautete, dass posteditierte Texte eine geringere lexikalische Dichte haben als Humanübersetzungen, und maschinell übersetzte Texte wiederum eine geringe lexikalische Dichte als posteditierte Texte ( $MÜ < LPE < FPE < HÜ$ ). Diese Hypothese wird durch die Ergebnisse bestätigt. Das Satzlängenverhältnis ist tatsächlich bei maschinell übersetzten Texten am geringsten (0,1274), etwas höher bei posteditierten Texten (LPE: 0,1636; FPE: 0,2623) und am höchsten bei Humanübersetzungen (0,2856).

Die Ergebnisse für die medizinischen und technischen Texte können Abbildungen 21 und 22 entnommen werden.

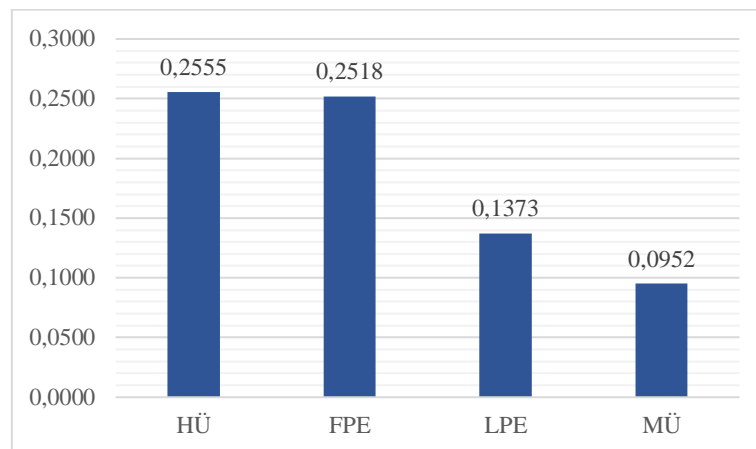


Abbildung 21: Satzlängenverhältnis nach Übersetzungsmodus für alle med. Texte

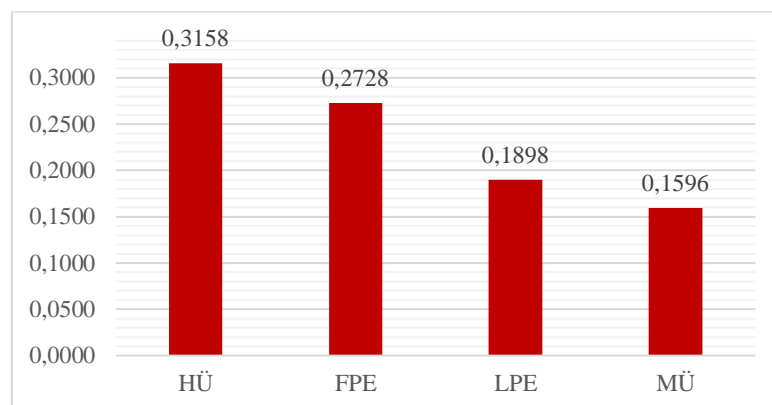


Abbildung 22: Satzlängenverhältnis nach Übersetzungsmodus für alle techn. Texte

Auch bei der differenzierten Betrachtung von technischen und medizinischen Texten bestätigt sich die Hypothese. Bei beiden Textgruppen ist das Satzlängenverhältnis für die MÜ-Outputs mit Werten von 0,0952 bzw. 0,1596 am geringsten, während die HÜ der Übersetzungsmodus mit dem höchsten Wert für das Satzlängenverhältnis ist (0,2555 bzw. 0,3158). Die Werte für die Übersetzungsmodi FPE (0,2518 bzw. 0,2728) und LPE (0,1373 bzw. 0,1898) liegen wie erwartet dazwischen.

Schließlich folgt die Aufschlüsselung nach Ausgangstext (Abbildung 23).

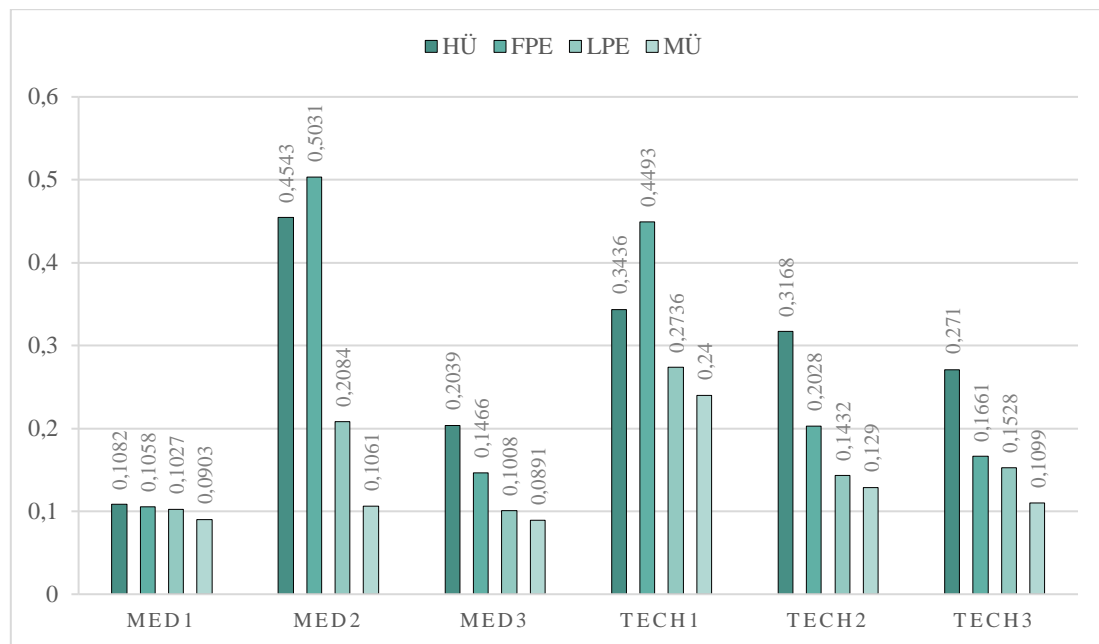


Abbildung 23: Satzlängenverhältnis nach Übersetzungsmodus für alle Texte, aufgeschlüsselt

In der Detailansicht ist die erwartete Progression nicht für jeden Text erkennbar. In MED1, MED3, TECH2 und TECH3 bestätigt sich die Hypothese nach wie vor in vollem Umfang. Auffallend ist bei MED1, dass alle Werte sehr niedrig und sehr nah beieinander sind – auch die Werte von MED3 sind verhältnismäßig niedrig. Die Hypothese bestätigt sich für MED2 und TECH1 nur teilweise: Dort ist das Satzlängenverhältnis in den MÜ-Outputs zwar am geringsten, jedoch ist es bei den FPE-Texten höher als bei den HÜ-Texten. Mögliche Gründe dafür werden in Abschnitt 4.2.3.2 diskutiert.

### 4.2.3 Einordnung der Ergebnisse zu Interference

Im folgenden Abschnitt werden die Ergebnisse aus den Abschnitten 4.2.1 und 4.2.2 eingeordnet und neue Forschungsperspektiven aufgezeigt. Am Ende des Abschnitts steht ein Fazit, in dem die Ergebnisse dieser Korpusanalyse zum Phänomen Interference resümiert werden.

### 4.2.3.1 Verhältnis der Verhältnisse von nominalen und verbalen Wortarten zwischen AT und ZT

Wie bereits in Abschnitt 3.2.2.1 erwähnt wurde, ist der Wert des Standardverhältnisses nicht bekannt. Folglich kann hier nicht eindeutig eingeordnet werden, ob die Werte, die oberhalb von 1 liegen, ‚noch‘ auf Interference oder ‚schon‘ auf Normalization hindeuten.

Der höchste Wert aus allen Texten ist 1,2216 (TECH2\_MUE), die meisten Werte liegen aber unterhalb von 1,1. Man könnte spekulieren, dass das Standardverhältnis vermutlich oberhalb von 1,1 oder sogar 1,2216 liegt und somit (fast) alle Werte, die über 1 liegen tatsächlich Interference und nicht Normalization nachweisen. Künftige Studien müssten diese Spekulation empirisch untermauern.

Als konkretes Beispiel soll TECH1 dienen, der einzige Text, bei dem alle Werte  $\geq 1$  sind. Im Übersetzungsmodus MÜ ist die Nominal-Verbal-Verteilung zwischen dem AT und dem ZT mit einem Wert von 1 exakt gleich, d. h. AT und ZT sind gleich nominal bzw. verbal. Dies entspricht der Erwartung, dass MÜ-Engines ‚nah‘ am AT bleiben und beispielsweise in der Regel keinen Wortklassenwechsel vornehmen (Lapshinova-Koltunski 2015, 106/107). Geht man davon aus, dass alle Werte noch unterhalb des Standardverhältnisses liegen, würde die Progression MÜ<LPE<HÜ<FPE bedeuten, dass das MÜ-Output am meisten Interference aufweist, gefolgt von den LPE- (1,0405), HÜ- (1,0624) und FPE-Texten (1,1023). Alle Zieltexte wären also nominaler als der AT (Wert  $\geq 1$ ), aber nicht so nominal wie deutsche Originaltexte normalerweise (Wert  $<$  Standardverhältnis). Die Verbalität des englischen Ausgangstextes würde folglich ‚durschneien‘. Dies entspräche der aufgestellten Hypothese nur teilweise, da die erwartete Progression MÜ<LPE<FPE<HÜ nicht ganz zu beobachten ist: Für die HÜ-Texte wurde der höchste Wert und damit die geringste Interference erwartet, stattdessen hat die HÜ in TECH1 den zweitniedrigsten Wert. Möglich ist natürlich auch, dass einer oder mehrere der Werte  $> 1$  bereits oberhalb des Standardverhältnisses liegen und somit Normalization signalisieren.

Solange der Wert des Standardverhältnisses nicht bekannt ist, kann anhand der vorgelegten Ergebnisse auch nicht geklärt werden, ob ‚noch Interference‘ oder ‚schon Normalization‘ vorliegt. Eine eventuelle Koexistenz der beiden Phänomene

(vgl. Teich 2003: 207) könnte nur mithilfe von qualitativen Studien nachgewiesen werden.

Überraschend war, wie bereits erwähnt, dass viele der Werte unerwartet niedrig ausgefallen sind. In MED1, MED2 und TECH3 lagen sogar alle Werte unterhalb von 0, was zeigt, dass hier die Zieltexte aller Übersetzungsmodi verbaler sind als die englischen Ausgangstexte. Denkbar ist, dass Interference-Effekte dazu beigetragen haben, aber allein dadurch lässt sich die hohe Verbalität im Vergleich zu den AT nicht erklären. Eine mögliche Erklärung hängt mit den von Steiner (2004, 133f) angesprochenen Unterschieden bei der Kompositabildung zusammen (vgl. Abschnitt 4.1.3.2): Ein Nominalkompositum wird von Sketch Engine in den englischen Texten als zwei Nomen gezählt, in den deutschen Texten lediglich als eins. Beispiele wurden ebenfalls bereits in Abschnitt 4.1.3.2 genannt, z. B. „vaccination program – „Impfprogramm“, „safety instructions“ – „Sicherheitshinweise“, „side effects“ – „Nebenwirkungen“, usw. Konkret führt dies dazu, dass die Anzahl der Nomen in den englischen Ausgangstexten nach oben verzerrt ist, bzw. die englischen Ausgangstexte nominaler wirken als sie ‚eigentlich sind‘. Diese Pseudonominalität könnte auch hier der Grund für die verzerrten Ergebnisse sein – d. h. konkret dafür, dass so viele Werte unterhalb von 0 liegen.

Ein oberflächlicher Vergleich von TECH3 mit den anderen Texten scheint diese These zu bestätigen: TECH3 hat insgesamt sehr niedrige Werte, d. h. die Zieltexte sind deutlich verbaler als ihre Ausgangstexte. Im Vergleich zu den anderen Texten gibt es in den Zieltexten zu TECH3 auch deutlich mehr Nominalkomposita: Während in TECH3 16 Beispiele für Nominalkomposita gezählt wurden, waren es in den anderen Texten zwischen fünf und neun. Tatsächlich steigen die berechneten Werte für das Verhältnis der Nominal-Verbal-Verhältnisse auf knapp über 1, wenn man bei der Berechnung bei jedem der Zieltexte von TECH3 16 Nomen addiert.<sup>19</sup> Die besonders stark ausgeprägte Pseudonominalität in TECH3 könnte also ein Grund für die besonders niedrigen Werte sein.

Einschränkend muss aber darauf hingewiesen werden, dass die Hypothese nicht nur aufgrund der insgesamt zu niedrigen Werte nicht bestätigt werden konnte, sondern auch, weil sich die erwartete Progression (MÜ<LPE<FPE<HÜ) nur in

---

<sup>19</sup> Es handelt sich hierbei um einen oberflächlichen Test. Eigentlich müsste die Zahl der fälschlicherweise ‚doppelt‘ gezählten Nomen natürlich für jeden Zieltext einzeln bestimmt werden.

MED1 und in den Gesamtergebnissen der medizinischen Texte zeigt. In TECH2 und TECH3 ist die Progression sogar entgegengesetzt zur Hypothese. In MED3 fällt im direkten Vergleich von den FPE- und LPE-Texten mit dem unbearbeiteten MÜ-Output auf, dass die Posteditoren die Texte durch ihre Korrekturen noch verbaler gemacht haben. Um Erklärungen dafür zu finden, wären tiefgreifende qualitative Analysen notwendig. Eine finale Interpretation der Ergebnisse unter 0 ist an dieser Stelle nicht möglich.

Neben den angesprochenen Faktoren gibt es weitere, die die Ergebnisse beeinflusst haben könnten. Dazu gehören zum Beispiel die Auslassung von Wörtern durch MÜ-Engines (insbesondere in MED2 und TECH1) (vgl. Abschnitt 4.1.3.1).

Außerdem wäre es interessant zu prüfen, inwieweit die Art der Anrede eine Auswirkung auf die Ergebnisse hat. Bei TECH1 handelt es sich um einen sehr appellbetonten Text: In beinahe jedem Satz werden die Adressaten aufgefordert, etwas zu tun. Im englischen Ausgangstext geschieht dies im Imperativ („Read all instructions before using the dishwasher“), in den deutschen Zieltexten fast einheitlich und entsprechend den deutschen Textsortenkonventionen über die Höflichkeitsform („Lesen Sie die Gebrauchsanweisung vor Benutzung des Geschirrspülers.“). Dadurch enthalten die deutschen Zieltexte allesamt deutlich mehr Pronomen (Teil der nominalen Wortarten) als der Ausgangstext, was logischerweise auch einen Einfluss auf das Verhältnis der Nominal-Verbal-Verhältnisse zwischen AT und ZT hat. Eine genauere Betrachtung des Einflusses dieser beiden Faktoren übersteigt den Rahmen dieser Arbeit und bleibt künftigen Forschungsanstrengungen überlassen.

Da Interference mit der gewählten Methode nicht nachgewiesen werden konnte, lassen sich auch keine Aussagen dazu treffen, inwiefern die Unerfahrenheit der Übersetzer bzw. Posteditoren einen Einfluss darauf hatte (vgl. Abschnitt 3.2.2.1).

An dieser Stelle kann also nicht final beantwortet werden, ob Interference für dieses Korpus anhand der Verhältnisse der Nominal-Verbal-Verhältnisse nachgewiesen werden konnte. Zur sinnvollen Interpretation der Ergebnisse müsste zunächst das Standardverhältnis der Nominalität bzw. Verbalität in deutschen Originaltexten beziffert werden. In künftigen Studien sollte untersucht werden, ob sich eine Betrachtung der Nominalität bzw. Verbalität von Texten beim Sprachpaar Englisch-



Deutsch überhaupt zum Nachweis von Interference eignet oder ob die Methodik so angepasst werden kann, dass die genannten Sprachsystemunterschiede nicht mehr so stark ins Gewicht fallen. In jedem Fall sollten dafür längere Korpus-texte und ein größeres Korpus genutzt werden.

#### 4.2.3.2 Verhältnis der Satzlängen

Wie bereits in Abschnitt 4.2.2 dargelegt, entsprechen die ermittelten Werte den Erwartungen und bestätigen die Hypothese. Maschinell übersetzte Texte haben die geringste lexikalische Dichte, gefolgt von posteditierten Texten und Humanübersetzungen. Dabei gilt: Je geringer der Wert, desto höher die Ähnlichkeit zwischen dem AT und dem ZT in Bezug auf die Satzlänge, was wiederum ein Zeichen von Interference ist. Die in fast allen Texten festgestellte Progression  $MÜ < LPE < FPE < HÜ$  entspricht der Erwartung und stimmt mit dem Grad des menschlichen Einflusses überein. Je stärker ein Mensch an dem jeweiligen Übersetzungsmodus beteiligt ist, desto mehr Abweichungen vom AT hinsichtlich der Satzlängen können festgestellt werden.

Die erwartete Progression konnte lediglich in MED2 und TECH1 nicht exakt wie erwartet festgestellt werden. Die Satzlängenverhältnisse für FPE-Texte sind dort höher als für HÜ-Texte und damit höher als erwartet. Ein Blick in die Daten und Texte liefert eine mögliche Erklärung: In MED2 setzt sich das unerwartet hohe durchschnittliche Satzlängenverhältnis (0,5031) aus den folgenden drei Werten der verschiedenen Übersetzungsvarianten zusammen:  $P18\_FPE = 0,3143$ ,  $P27\_FPE = 0,1882$  und  $P23\_FPE = 1,0068$ . Der von Teilnehmer P23 vollständig posteditierte Text hat also ein deutlich höheres Satzlängenverhältnis als die andere Varianten und stellt somit einen ‚Ausreißer‘ dar. Ein genauerer Blick in die Daten, d. h. in die Satz-für-Satz-Berechnung des Satzlängenverhältnisses und in die Korpus-texte selbst, zeigt, dass Satz 2 und 3 des AT zu nur einem Satz im ZT verbunden wurden (Abbildung 24). Folglich hat der ZT einen Satz weniger als der AT. An der letzten Satzposition wird für den ‚fehlenden‘ Satz der Wert 0 gesetzt.

MED2_P23_FPE				
Satzpaar Nr.	Satzlänge AT	Satzlänge ZT	Satzlängenverhältnis	
1	48	52		0,0833
2	50	91		0,8200
3	39	140		2,5897
4	147	86		0,4150
5	78	36		0,5385
6	31	33		0,0645
7	24	164		5,8333
8	142	108		0,2394
9	104	122		0,1731
10	105	111		0,0571
11	86	63		0,2674
12	51	0		1,0000
				1,0068

Abbildung 24: Satzlängenverhältnis nach Satz für das Korpus MED2\_P23\_FPE

Beginnend mit Satzpaar 2 entstehen für alle folgenden Satzpaare mitunter stark erhöhte Satzlängenverhältnisse, die sich auf den Durchschnitt auswirken. Die grauen Linien in Abbildung 24 verdeutlichen, welche Sätze ‚eigentlich‘ zusammengehört hätten. Die verschobene Satzalignierung (vgl. Abschnitt 3.2.2.2, Abbildung 7) führt insbesondere bei Satzpaar 7 zu einem enorm hohen Satzlängenverhältnis, da dort der kurze Satz des AT (24 Zeichen) auf einen sehr langen Satz im ZT (164 Zeichen) trifft. Dass Teilnehmer P23 überhaupt zwei Sätze zu einem kombiniert hat, ist für ein FPE eher unüblich, da nur notwendige Änderungen vorgenommen werden sollen (vgl. Abschnitt 2.1.2). Werfen wir einen Blick in den AT (1a), das unbearbeitete MÜ-Output (1b), sowie in die FPE-Variante von Teilnehmer P23 (1c) (jeweils Satz 2 und 3), um zu beurteilen, ob das Verbinden der Sätze als ‚notwendige Änderung‘ zu werten ist.

(1)a. *Protaphane is human insulin to treat diabetes.*

*Protaphane is a long-acting insulin.*

b. *Protaphane ist ein Humaninsulin zur Behandlung von Diabetes.*

*Protaphane ist ein lang wirkendes Insulin.*

c. *Protaphane ist ein Humaninsulin zur Behandlung von Diabetes mit langanhaltender Wirkung.*

Die im AT (1a) vorhandene Anapher wurde vom MÜ-Engine (1b) analog nachgebildet. Teilnehmer P23 hingegen hat sich beim Full Post-Editing (1c) dazu entschieden, die Anapher durch Kombinieren der beiden Sätze aufzulösen. Ob dies

etwa aus stilistischen Gründen nötig war, ist fraglich. In allen anderen FPE- und LPE-Texten von MED2 wurde die im MÜ-Output vorgeschlagene Anapher beibehalten und das MÜ-Output wurde in den meisten Fällen gänzlich unverändert übernommen. Auch in den HÜ-Texten wurden Satz 2 und 3 nur von einem einzigen Teilnehmer (P02) zu einem Satz verbunden.

Schauen wir nun zurück auf das durchschnittliche Satzlängenverhältnis aller FPE-Varianten von MED2. Zur Erinnerung: Es war höher als erwartet und setzte sich aus den folgenden Werten zusammen:  $P18\_FPE = 0,3143$ ,  $P27\_FPE = 0,1882$  und  $P23\_FPE = 1,0068$ . Wäre der Wert für  $P23\_FPE$  nur etwas geringer, etwa 0,8 (was immer noch deutlich über den Werten von  $P18\_FPE$  und  $P27\_FPE$  wäre), würde dies genügen, um das FPE-Satzlängenverhältnis für MED2 so weit zu senken, dass es unterhalb des HÜ-Satzlängenverhältnisses läge. Dann wäre die Hypothese vollständig bestätigt.

Dieses Beispiel zeigt, dass nur eine ‚unübliche‘ Entscheidung eines Posteditors, die eventuell auch über den eigentlichen Zweck vom Post-Editing hinaus geht, eine starke Auswirkung auf das Satzlängenverhältnis für diesen Text haben kann. Im Gesamtergebnis aller Texte fällt dieser ‚Ausreißer‘ zwar nicht mehr so stark ins Gewicht, in den nach Ausgangstexten aufgeschlüsselten Ergebnissen allerdings schon. Grund dafür ist die geringe Länge der Korpustexte.

Die von der Hypothese abweichenden Ergebnisse von TECH1 lassen sich auf ähnliche Weise erklären. Insgesamt kann die Hypothese daher als bestätigt gelten.

Torals (2019, 277/278) Vermutung, unerfahrene (Human-)Übersetzer würden „more literal translations“ produzieren und sich streng an Satzgrenzen halten (vgl. auch Kujamäki 2004, 198), hat sich im vorliegenden Datensatz nicht bestätigt. In diesem Korpus zeigen sich in den Humanübersetzungen durchaus Abweichungen von den Satzgrenzen: Ein Aufspalten oder Verbinden von Sätzen des AT hat in 6 von 21 HÜ-Texten, 4 von 21 FPE-Texten, und kein Mal in LPE- oder MÜ-Texten stattgefunden. Überraschend war allerdings, dass in der FPE-Aufgabe so oft Sätze kombiniert oder getrennt wurden, auch wenn das durch den Post-Editing-Auftrag nicht unbedingt angezeigt war. Diese Beobachtung wiederum könnte man mit der Unerfahrenheit der Textproduzenten in Bezug auf Post-Editing in Verbindung bringen. Čulo et al. (2017) haben auch selbst angemerkt, dass die Textproduzenten so gut wie keine Post-Editing-Erfahrung hatten.

Interessant ist auch, dass acht der zehn Satzaufspaltungen bzw. -kombinationen in den technischen Texten stattgefunden haben, während es in den medizinischen Texten nur zwei waren. Dies mag auch daran liegen, dass das Korpus aus 42 technischen und nur 33 medizinischen Texten besteht, aber das allein erklärt dieses Ungleichgewicht nicht. Weitere Studien müssten dies näher ergründen.

Ein möglicher Kritikpunkt an der Berechnung des Satzlängenverhältnisses nach Formel 5 (siehe Abschnitt 3.2.2.2) ist, dass die oben erläuterte, verschobene Satzalignierung (Abbildung 7) unterschiedlich starke Auswirkungen auf das durchschnittliche Satzlängenverhältnis des Textes hat, in Abhängigkeit davon, wie stark die Satzlängen innerhalb des AT variieren. Wechseln sich beispielsweise ständig sehr kurze und sehr lange Sätze ab, dürfte die Verschiebung stärker ins Gewicht fallen, als wenn es keine große Varianz zwischen den Satzlängen gibt (vgl. Abschnitt 3.2.2.2).

Zu kritisieren ist außerdem, dass es nicht unerheblich ist, wann ein Satz aufgespaltet oder zwei Sätze kombiniert werden: Passiert dies früh im Text, dürfte der Einfluss auf das durchschnittliche Satzlängenverhältnis für den ganzen Text größer sein als wenn es erst gegen Ende des Textes passiert, da sich die durch das Aufspalten oder Kombinieren erzeugte Diskrepanz in Form von erhöhten Satzlängenverhältnissen durch den ganzen Text zieht (sofern die eigentliche Satzalignierung nicht durch die umgekehrte Operation wieder hergestellt wird). Folglich hat auch die Textlänge (in Sätzen) einen Einfluss auf das Satzlängenverhältnis.

Darüber hinaus berücksichtigt die Formel nicht, dass auch Sprachsystemunterschiede das Ergebnis für die Satzlängenverhältnisse verfälschen können. Beispielsweise könnte man durch eine Konstruktion, die es nur in einem der beiden Sprachsysteme gibt (z. B. das *gerund* im Englischen) zu einer expansiven Übersetzung des Satzes ‚gezwungen‘ werden.

Inwiefern diese Faktoren die Satzlängenverhältnisse tatsächlich beeinflussen, muss ebenfalls in künftigen Analysen genauer untersucht werden.

Insgesamt liefert die Untersuchung der Satzlängenverhältnisse deutliche Hinweise darauf, dass Interference in posteditierten Texten in höherem Maße nachgewiesen werden kann als in humanübersetzten Texten. Noch mehr Interference weisen

maschinell übersetzte Texte auf. Diese Ergebnisse stimmen mit Torals Ergebnissen überein und bestätigen seine These, dass Interference (nachgewiesen anhand der Satzlängenverhältnisse) ein Merkmal der Post-Editese ist. Da Interference in posteditierten Texten in höherem Maße nachgewiesen werden konnte als in humanübersetzten Texten, kann Post-Editese in dieser Hinsicht tatsächlich als verschärfte Form der Translationese verstanden werden.

#### **4.2.3.3 Fazit zu Interference**

In der Gesamtschau beider Untersuchungsaspekte lässt sich kein eindeutiges Fazit dazu ziehen, ob Interference im Rahmen dieser Korpusanalyse besonders stark in posteditierten Texten nachgewiesen werden konnte.

In Bezug auf die Verhältnisse der Nominal-Verbal-Verhältnisse des AT und des ZT bedarf es weiterer Untersuchungen und Korpusanalysen, um die hier vorgestellten Ergebnisse sinnvoll zu interpretieren. In diesem Rahmen sollte nicht nur angestrebt werden, das Standardverhältnis zu beziffern, sondern auch, den Einfluss von sprachsystembezogenen Unterschieden wie der Kompositabildung zu bestimmen und die Methoden gegebenenfalls so anzupassen, dass die Ergebnisse weniger anfällig für diese Art der Verzerrung sind. Dies sollte unbedingt anhand von größeren Korpora mit längeren Texten geschehen.

Die Ergebnisse zu den Satzlängenverhältnissen zwischen dem AT und dem ZT hingegen haben die Hypothese rundum bestätigt. Sie weisen darauf hin, dass Interference (gemessen an den Satzlängenverhältnissen) stärker in posteditierten Texten als in humanübersetzten Texten vorkommt. Demnach wäre Interference ein Phänomen, was nicht nur ein Merkmal der Translationese, sondern auch der Post-Editese ist und dort besonders stark in Erscheinung tritt. Idealerweise sollten auch diese Ergebnisse anhand weiterer Studien mit größeren Korpora überprüft werden.

## 5 Diskussion

In diesem Abschnitt sollen weitere Probleme und Limitationen der vorliegenden Korpusanalyse thematisiert werden.

Neben den bereits angesprochenen Limitationen Größe des Textkorpus (siehe Abschnitte 4.1.3 und 4.2.3.1), Erfahrungsgrad der Textproduzenten (siehe Abschnitte 4.1.3.1 und 4.2.3), und wortbasierte Untersuchungsmethoden (Abschnitte 4.1.3.2 und 4.2.3.1), sollen an dieser Stelle noch weitere Limitationen, die für alle Untersuchungsaspekte gelten, aufgeführt werden.

Zunächst einmal sei darauf hingewiesen, dass es sich sowohl bei der lexikalischen Diversität (TTR) und der lexikalischen Dichte, als auch bei den Verhältnissen der Nominal-Verbal-Verhältnisse und der Satzlängenverhältnisse von AT und ZT um „shallow feature[s]“ (Teich 2003, 145; Xiao 2010, 165) bzw. um „superficial linguistic features“ handelt (Toral 2019, 280). Laut Steiner (2004, 132) sind all diese „low-level criteria“ problematisch, da ihr Zusammenhang mit „higher-level theoretical concepts such as ‚interference, normalisation, creativity, explicitation, levelling-out, simplification, understanding““ zweifelhaft ist.

Olohan (2004, 86) lehnt den Einsatz solcher „shallow features“ nicht grundsätzlich ab und verweist auf ihren Nutzen: „[...] the simple measure of frequency can produce interesting findings when comparing to comparable corpora.“ Allerdings sagt sie auch, solche oberflächlichen, rein quantitativen Features können nur ein erster Schritt sein, auf den tiefgreifendere, auch qualitative, Untersuchungen folgen müssen, damit Generalisierungen möglich sind:

„More often than not, quantitative data form the basis from which qualitative analysis is deemed interesting, useful or indispensable. Thus a combination of quantitative and qualitative analysis is, in most cases, desirable, particularly if fuller descriptions of linguistic and translational phenomena are to be given and reasons suggested for their occurrence“ (Olohan 2004, 86).

Qualitative Untersuchungen, wie Olohan sie fordert, konnten in dieser Analyse aufgrund des begrenzten Rahmens nur beispielhaft durchgeführt oder angerissen werden (siehe Abschnitte 4.1.3 und 4.2.3). In diesem Zusammenhang wurden auch konkrete Forschungsdesiderate aufgezeigt. Auf diese vornehmlich quantitativ ausgerichtete Arbeit sollten also sowohl weitere quantitative als auch qualitative

Analysen folgen, um die Ergebnisse zu sichern und Erklärungen für die verschiedenen Phänomene zu finden.

Neben diesen bereits angerissenen Faktoren, gibt es weitere, die die Ergebnisse beeinflusst haben könnten. Dazu gehören die genutzte MÜ-Engine (vgl. Čulo et al. 2017, 196), die Art der Trainingsdaten und das Sprachpaar (Daems et al. 2017, 101) (vgl. auch Abschnitt 2.1.5). Daems et al. weisen darauf hin, dass verschiedene Sprachgemeinschaften möglicherweise verschiedene Ansichten über MÜ und Post-Editing haben und die Aufgabe folglich unterschiedlich interpretieren (ebd.). In der Konsequenz können die Ergebnisse je nach Sprachpaar abweichen. Tatsächlich haben sich die meisten der bisherigen Post-Editese-Studien mit dem Sprachpaar Englisch-Deutsch (vgl. Toral 2019, Čulo et al. 2017) bzw. Englisch-Niederländisch (Daems et al. 2017) befasst.<sup>20</sup> Andere Sprachpaare und -richtungen – auch solche an denen Englisch nicht beteiligt ist – sollten künftig dringend mitberücksichtigt werden.

Denkbar ist außerdem, dass die Ergebnisse bei der Verwendung einer anderen MÜ-Engine (z. B. DeepL oder eine nicht-öffentliche bzw. proprietäre MÜ-Engine) erheblich abgewichen wären. Von zentraler Bedeutung für die Performance einer MÜ-Engine ist, wie bereits erwähnt (siehe Abschnitt 2.1.5), die Art der Daten, die bei deren Training verwendet wurde: Hätte man eine auf medizinische oder technische Fachtexte spezialisierte MÜ-Engine genutzt, wären möglicherweise abweichende Ergebnisse erzielt worden. Auch die Tatsache, dass jeweils nur ein MÜ-Output vorlag, limitiert die Aussagekraft der Ergebnisse. Zudem spielt der Zeitpunkt, zu dem das MÜ-Output generiert wurde eine Rolle, da MÜ-Engines stetig weiterentwickelt werden. Künftige Studien sollten also verschiedene MÜ-Engines gegenüberstellen, um zu erforschen, ob dies zu abweichenden Ergebnissen führt.

Womöglich haben die erzielten Ergebnisse auch mit dem persönlichen Stil der Teilnehmer zu tun. Die Entscheidung etwa, wie sehr man Redundanzen verhindert oder sogar bewusst einarbeitet, ob implizite Bezüge explizit gemacht werden, ob Verbalphrasen verbal oder nominal übersetzt werden, und ob Sätze aufgespaltet oder kombiniert werden, ist letztlich auch eine Frage des subjektiven Stils. Theoretisch ist denkbar, dass nur ein oder wenige Teilnehmer mit einem ‚Hang‘ zu Simplification

---

<sup>20</sup> Bei Farrell (2018) (EN → IT) handelt es sich nicht um eine Post-Editese-Studie im engeren Sinn.

oder Interference die Ergebnisse der vorliegenden Korpusanalyse verzerren – Grund dafür ist wiederum die begrenzte Größe des Korpus. Idealerweise liegen in künftigen Korpusanalysen mehr als drei oder vier Übersetzungsvarianten pro Übersetzungsmodus vor, um die Gefahr, dass subjektive Schreibstile die Ergebnisse verfälschen, weiter zu reduzieren. Konkretere Untersuchungen in diese Richtung, wie beispielsweise Stilometrie-Studien, könnten ebenfalls interessante Ergebnisse liefern.

Nicht zuletzt können subjektive Faktoren wie Konzentration, Arbeitsbedingungen und die Einstellung zu Maschinellem Übersetzung bzw. Post-Editing (vgl. Abschnitt 2.1.7) relevant sein. Auch der Kontext, in dem das Korpus erstellt wurde, kann einen Einfluss gehabt haben: Möglicherweise hat das Wissen darum, dass die Teilnehmer an einer Studie teilnehmen, sowie das Wissen, dass Eyetracking- und Keylogging-Daten gesammelt werden, die Art und den Umfang der Übersetzungs- bzw. Korrekturentscheidungen der Teilnehmer beeinflusst.

Zuletzt bleibt darauf hinzuweisen, dass im Rahmen dieser Korpusanalyse nicht geprüft werden konnte, ob die festgestellten Differenzen zwischen den errechneten Werten überhaupt signifikant sind. Auch deshalb sollten diese Ergebnisse anhand eines größeren Korpus überprüft werden. Die Ergebnisse dieser Arbeit sind demnach noch nicht als gesichert zu verstehen, sondern bieten einen ersten Anhaltspunkt für künftige Studien im Bereich der Post-Editese-Forschung.

Die festgestellten methodischen Schwächen dieser Arbeit können als Chance für künftige Forschungsanstrengungen begriffen werden. Durch sie eröffnen sich neue interessante Perspektiven für genauere Untersuchungen.



## 6 Fazit und Ausblick

Im Rahmen der Einführung in die Themenbereiche Post-Editing, Translationese und Post-Editese und durch die Vorstellung der Forschungsstände (Abschnitte 2.1, 2.2 und 2.3) hat sich eine Forschungslücke im Bereich der Post-Editese-Forschung herauskristallisiert (Abschnitt 2.4). Obwohl ein Großteil des Übersetzungsbedarfs im Bereich der Language for Specific Purposes zu verorten ist, wurde dieser Bereich bei der Erforschung der Unterschiede zwischen posteditierten und humanübersetzten Texten bisher beinahe vollständig ignoriert.

Ziel der vorliegenden Korpusanalyse war es, einen Beitrag zur Schließung dieser Forschungslücke zu leisten, indem die von Toral (2019) und Lapshinova-Koltunski (2013) genutzten Methoden in adaptierter Form auf ein Korpus aus medizinischen und technischen Fachtexten angewendet werden. Die übergeordnete Hypothese war, analog zu Toral (2019), dass die Phänomene Simplification und Interference in posteditierten Texten in höherem Maße als in humanübersetzten Texten vorzufinden seien. Folglich wären sie nicht nur Merkmale der Translationese, sondern auch der Post-Editese. Post-Editese könne demnach als verschärfte Form der Translationese verstanden werden.

Zum Nachweis von Simplification wurden die lexikalische Diversität und die lexikalische Dichte der Korpus Texte untersucht (Abschnitt 3.2.1). Das Verhältnis der Nominal-Verbal-Verhältnisse und das Verhältnis der Satzlängen zwischen AT und ZT diente dem Nachweis von Interference (Abschnitt 3.2.2).

In Bezug auf Simplification (Abschnitt 4.1) konnte die Hypothese anhand der Ergebnisse der eigenen Korpusanalyse nicht bestätigt werden. Die kumulierten Ergebnisse zur lexikalischen Diversität (TTR) zeigten zwar, wie erwartet, dass posteditierte Texte lexikalisch weniger divers als humanübersetzte Texte sind, es bedarf aber weiterer Studien mit größeren Korpora, um diese Ergebnisse zu sichern. Die kumulierten Ergebnisse zur lexikalischen Dichte wiederum deuten darauf hin, dass posteditierte Texte, entgegen den Erwartungen, eine höhere lexikalische Dichte als humanübersetzte Texte haben. Allerdings dürften die Ergebnisse aufgrund von Sprachsystemunterschieden verzerrt sein, weshalb grundsätzlich kritisch betrachtet werden müssen. Sollten sie durch weitere Studien bestätigt werden, wäre dies ein Widerspruch zur Hypothese, Simplification sei ein Merkmal der Post-Editese.

Folglich kann die Forschungsfrage, ob Simplification in posteditierten Texten tatsächlich in höherem Maße als in humanübersetzten Texten auftritt, an dieser Stelle nicht mit ausreichender Sicherheit beantwortet werden.

Auch die Ergebnisse zu Interference (Abschnitt 4.2) ließen keine eindeutigen Schlüsse dazu zu, ob Interference in posteditierten Texten in einem höheren Maße als in humanübersetzten Texten auftritt. Die Untersuchung der Satzlängenverhältnisse zwischen AT und ZT bestätigte die Hypothese in vollem Maße und zeigte somit, dass posteditierte Texte in Bezug auf die Satzgrenzen und -längen ihren Ausgangstexten ähnlicher sind als humanübersetzte Texte ihren Ausgangstexten. Jedoch konnten die Ergebnisse zu den Verhältnissen der Nominal-Verbal-Verhältnisse mangels Daten aus Vergleichskorpora nicht sinnvoll interpretiert werden. Außerdem sorgten die oben genannten Sprachsystemunterschiede auch hier vermutlich für Verzerrungen.

Insgesamt kann also festgehalten werden, dass trotz einiger Hinweise in diese Richtung weder Simplification noch Interference in besonders starkem Maße in posteditierten Texten nachgewiesen werden konnten. Die Frage, ob diese Translationese-Merkmale auch Merkmale der Post-Editese sind, ließ sich ebenso wenig endgültig beantworten wie die Frage, ob es sich bei Post-Editese um eine verschärfte Form der Translationese handelt.

Vor dem Hintergrund der festgestellten methodischen Schwächen (Abschnitte 4.1.3, 4.2.3 und 5) sollten sich künftige Forschungsanstrengungen auf größere Korpora mit längeren Korpustexten stützen und – aufbauend auf den quantitativen Ergebnissen – quantitative mit qualitativen Methoden verbinden. Im Idealfall beschränken sich diese Korpusanalysen nicht auf „shallow features“, sondern nutzen auch tiefergreifende Methoden, die so angepasst sind, dass sie eine Verzerrung der Ergebnisse aufgrund von Sprachsystemunterschieden vermeiden. Darüber hinaus sollten sie verschiedene MÜ-Paradigmen, Sprachpaare, Erfahrungsgrade der Textproduzenten sowie andere Textsorten aus dem Bereich Language for Specific Purposes berücksichtigen. Dies ist notwendig, um die in dieser Arbeit festgestellten Ergebnisse zu bestätigen oder zu widerlegen und sie einzuordnen, sowie um die offen gebliebenen Fragen zu beantworten.

Zusammenfassend ist es in dieser Korpusanalyse gelungen, einige Hinweise auf das verstärkte Vorkommen von Simplification und Interference in posteditierten Texten gegenüber humanübersetzten Texten zu liefern, auch wenn es für ein gesichertes Fazit weiterer Studien bedarf. Die festgestellten methodischen Schwächen und offengebliebenen Fragen eröffnen zahlreiche neue Perspektiven und interessante Forschungsrichtungen innerhalb der Post-Editese-Forschung.

## 7 Literaturverzeichnis

- Adriaens, G.; Allen, J.; Bernth, A.; Godden, K.; Mitamura, T.; Nyberg, E.; Wocjik, R.; Zajac, R. (2000) *CLAW 2000. Proceedings of the third International Workshop on Controlled Language Applications*. Seattle, Washington.
- Aharoni, R.; Koppel, M.; Goldberg, Y. (2014) Automatic Detection of Machine Translated Text and Translation Quality Estimation. In: *Proceedings of the 52nd Annual Meeting of the Association for Computational Linguistics. (Volume 2: Short Papers)*. Baltimore, Maryland: Association for Computational Linguistics: 289–295.
- Allen, J. (2003) Post-editing. In: Somers, H. (Hrsg.) *Computers and Translation. A translator's guide*. Philadelphia: John Benjamins Publishing Company: 297–317.
- Allied Business Intelligence (2002) *Language translation, localization and globalization: World market forecasts, industry drivers and eSolutions*. Oyster Bay, NJ: Allied Business Intelligence, Inc. (5.3).
- Al-Shabab, O. (1996) *Interpretation and the Language of Translation - Creativity and Convention in Translation*. Janus Publishing Company Limited.
- Arnold, D.; Balkan, L.; Meijer, S.; Humphreys, R.L.; Sadler, L. (1994) *Machine translation: an introductory guide*. Manchester: NCC Blackwell.
- Aymerich, J. (2005) Using Machine Translation for fast, inexpensive, and accurate health information assimilation and dissemination. Experiences at the Pan American Health Organization.  
<http://www.icml9.org/program/track3/public/documents/Julia%20Aymerich-174934.doc> (31.08.2022).
- Aziz, W.; De Sousa, Sheila C M; Specia, L. (2012) PET: a Tool for Post-editing and Assessing Machine Translation. In: *Proceedings of the Eighth International Conference on Language Resources and Evaluation (LREC'12)*: 3982–3987.  
[https://www.researchgate.net/publication/264886897\\_PET\\_a\\_Tool\\_for\\_Post-editing\\_and\\_Assessing\\_Machine\\_Translation](https://www.researchgate.net/publication/264886897_PET_a_Tool_for_Post-editing_and_Assessing_Machine_Translation) (31.08.2022).
- Babych, B.; Hartley, T. (2004) Extending the BLEU MT Evaluation Method with Frequency Weightings. In: *Proceedings of the 42nd Annual Meeting of the Association for Computational Linguistics (ACL-04)*: 621–628.  
<https://aclanthology.org/P04-1079/> (31.08.2022).
- Baker, M. (1993) Corpus Linguistics and Translation Studies — Implications and Applications. In: Baker, M.; Francis, G.; Tognini-Bonelli, E.; Sinclair, J. (Hrsg.) *Text and technology. In honour of John Sinclair*. Philadelphia: J. Benjamins Pub. Co: 233–250.
- Baker, M. (1995) Corpora in Translation Studies. An Overview and Some Suggestions for Future Research. In: *Target. International Journal of Translation Studies* 7(2): 223–243.

- Baker, M. (1996) Corpus-based translation studies: The challenges that lie ahead. In: Somers, H. (Hrsg.) *Terminology, LSP and Translation. Studies in language engineering in honour of Juan C. Sager*. Amsterdam and Philadelphia: Benjamins: 175–186.
- Baker, M.; Olohan, M. (2000) Reporting That in Translated English: Evidence for Subconscious Processes of Explication? In: *Across Languages and Cultures* 1(2): 141–158.
- Ball, L.; Eger, N.; Stevens, R.; Dodd, J. (2006) Applying the post-experience eye-tracked protocol (PEEP) method in usability testing. In: *Interfaces* (67): 15–19.
- Baroni, M.; Bernardini, S. (2006) A New Approach to the Study of Translationese: Machine-learning the Difference between Original and Translated Text. In: *Literary and Linguistic Computing* 21(3).  
[https://www.researchgate.net/publication/31201942\\_A\\_New\\_Approach\\_to\\_the\\_Study\\_of\\_Translationese\\_Machine-learning\\_the\\_Difference\\_between\\_Original\\_and\\_Translated\\_Text](https://www.researchgate.net/publication/31201942_A_New_Approach_to_the_Study_of_Translationese_Machine-learning_the_Difference_between_Original_and_Translated_Text) (31.08.2022).
- Bernardini, S.; Zanettin, F. (2004) When is a universal not a universal? Some limits of current corpus-based methodologies for the investigation of translation universals. In: Mauranen, A.; Kujamäki, P. (Hrsg.) *Translation universals. Do they exist?* Benjamins: 51–62.
- Bowker, L. (2009) Can Machine Translation meet the needs of official language minority communities in Canada? A recipient evaluation. In: *Linguistica Antverpiensia* 8: 123–155.  
[https://www.researchgate.net/publication/299060995\\_Can\\_Machine\\_Translation\\_meet\\_the\\_needs\\_of\\_official\\_language\\_minority\\_communities\\_in\\_Canada\\_A\\_recipient\\_evaluation](https://www.researchgate.net/publication/299060995_Can_Machine_Translation_meet_the_needs_of_official_language_minority_communities_in_Canada_A_recipient_evaluation) (31.08.2022).
- Bowker, L.; Buitrago-Ciro, J. (2015) Investigating the usefulness of machine translation for newcomers at the public library. In: *Translation and Interpreting Studies* 10(2): 165–186.  
[https://www.researchgate.net/publication/290210038\\_Investigating\\_the\\_usefulness\\_of\\_machine\\_translation\\_for\\_newcomers\\_at\\_the\\_public\\_library](https://www.researchgate.net/publication/290210038_Investigating_the_usefulness_of_machine_translation_for_newcomers_at_the_public_library) (31.08.2022).
- Cadwell, P.; Hu, K. (2016) A Comparative Study of Post-editing Guidelines. In: Carl, M.; Dragsted, B.; Elming, J.; Hardt, D.; Jakobsen, A. (2011) *The Process of Post-Editing: A Pilot Study*. In: *Copenhagen Studies in Language* (41): 131–142.
- Carl, M.; Gutermuth, S.; Hansen-Schirra, S. (2015) Post-Editing Machine Translation: Efficiency, Strategies, and Revision Processes in Professional Translation Settings. In: *Psycholinguistic and Cognitive Inquiries into Translation and Interpreting*: 145–174.
- Castilho, S.; Moorkens, J.; Gaspari, F.; Calixto, I.; Tinsley, J.; Way, A. (2017) Is neural machine translation the new state of the art? In: *The Prague Bulletin of Mathematical Linguistics* (108): 109–120. <https://doras.dcu.ie/23072/> (31.08.2022).

- Chesterman, A. (2004) Beyond the particular. In: Mauranen, A. (Hrsg.) *Translation Universals. Do they exist?* Philadelphia: John Benjamins Publishing Company: 33–49.
- Church, K.; Hovy, E. (1993) Good applications for crummy machine translation. In: *Machine Translation* 8(4): 239–258. <https://www.jstor.org/stable/40007972> (06.08.2022).
- Common Sense Advisory (2016) Europe’s Leading Role in Machine Translation. How Europe is Driving the Shift to MT. [http://cracker-project.eu/wp-content/uploads/Europes\\_Leading\\_Role\\_in\\_MT.pdf](http://cracker-project.eu/wp-content/uploads/Europes_Leading_Role_in_MT.pdf) (07.08.2022).
- Čulo, O.; Gutermuth, S.; Hansen-Schirra, S.; Nitzke, J. (2014) The Influence of Post-Editing on Translation Strategies. In: O’Brien, S.; Winther Balling, L.; Carl, M.; Simard, M.; Specia, L. (Hrsg.) *Post-editing of Machine Translation. Processes and Applications*. Newcastle upon Tyne: Cambridge Scholars Publishing: 200–218.
- Čulo, O.; Hansen-Schirra, S.; Nitzke, J. (2017) Contrasting terminological variation in post-editing and human translation of texts from the technical and medical domain: New Methodological and Theoretical Traditions. In: Sutter, G. (Hrsg.) *Empirical Translation Studies. New Methodological and Theoretical Traditions*. Berlin/Boston: De Gruyter Inc.
- Daems, J.; Clercq, O. de; Macken, L. (2017) Translationese and Post-editeese: How comparable is comparable quality? In: *Linguistica Antverpiensia, New Series – Themes in Translation Studies* 16. <https://lans-tts.uantwerpen.be/index.php/LANS-TTS/article/view/434> (31.08.2022).
- DeepL (2021) How does DeepL work? <https://www.deepl.com/en/blog/how-does-deepl-work> (08.08.2022).
- DePalma, D. (2013) Post-editing in practice. <https://www.tcworld.info/e-magazine/translation-and-localization/post-editing-in-practice-459/> (20.03.2022).
- Doğru, G.; Martín-Mor, A.; Aguilar-Amat, A. (2018) Parallel Corpora Preparation for Machine Translation of Low-Resource Languages: Turkish to English Cardiology Corpora. In: *Proceedings of the LREC 2018 Workshop*.
- Doherty, S.; O’Brien, S.; Carl, M. (2010) Eye tracking as an MT evaluation technique. In: *Machine Translation* 24(1): 1–13.
- Doherty, S.; O’Brien, S. (2009) Can MT Output be Evaluated Through Eye Tracking? In: *MT Summit XII*. Ottawa, Canada.
- Farrell, M. (2018) Machine Translation Markers in Post-Edited Machine Translation Output. In: Asling: International Society for Advancement in Language Technology (Hrsg.) *Translating and the Computer 40: proceedings.*: 50–59.
- Fiederer, R.; O’Brien, S. (2009) Quality and Machine Translation: A realistic objective? In: *The Journal of Specialised Translation*: 52–74 (02.02.2022).

- Frawley, W. (1984) Prolegomenon to a theory of translation. In: *Translation: Literary, Linguistic and Philosophical Perspectives*. Associated University Press: 159–175.
- Garcia, I. (2010) Is machine translation ready yet? In: *Target. International Journal of Translation Studies* (22 (1)): 7–21.  
<https://benjamins.com/online/target/articles/target.22.1.02gar> (31.08.2022).
- Garcia, I. (2011) Translating by post-editing: is it the way forward? In: *Machine Translation* 25(3): 217–237.
- Gaspari, F.; Almaghout, H.; Doherty, S. (2015) A survey of machine translation competences: Insights for translation technology educators and practitioners. In: *Perspectives* 23(3): 333–358.
- Gaspari, F.; Toral, A.; Naskar, S.; Groves, D.; Way, A. (2014) Perception vs. reality: measuring machine translation post-editing productivity. In: *Proceedings of the 11th Conference of the Association for Machine Translation in the Americas*: 60–72. <https://aclanthology.org/2014.amta-wptp.5/> (31.08.2022).
- Gellerstam, M. (1986) Translationese in Swedish novels translated from English. Vorgelegt beim Scandinavian Symposium on Translation Theory, Lund.
- Gerber, L. (2009) Machine Translation: Ingredients for Productive and Stable MT deployments - Part 1.  
<https://www.translationdirectory.com/articles/article1941.php> (04.08.2022).
- Glossary of Linguistic Terms (2015) Adposition.  
<https://glossary.sil.org/term/adposition> (01.08.2022).
- Green, S.; Heer, J.; Manning, C. (2013) The efficacy of human post-editing for language translation. In: Mackay, W. (Hrsg.) *Proceedings of the SIGCHI Conference on Human Factors in Computing Systems*. New York, NY: ACM: 439–448.
- Groves, D.; Schmidtke, D. (2009) Identification and Analysis of Post-Editing Patterns for MT. In: *Proceedings of Machine Translation Summit XII: Commercial MT User Program*.  
[https://www.researchgate.net/publication/228789664\\_Identification\\_and\\_Analysis\\_of\\_Post-Editing\\_Patterns\\_for\\_MT](https://www.researchgate.net/publication/228789664_Identification_and_Analysis_of_Post-Editing_Patterns_for_MT) (31.08.2022).
- Guerberof, A. (2009) Productivity and quality in MT post-editing.  
<https://core.ac.uk/download/pdf/228161879.pdf> (31.08.2022).
- Hansen, S. (2003) The Nature of translated text : an interdisciplinary methodology for the investigation of the specific properties of translations.  
[https://lib.ugent.be/en/catalog/rug01:002029480?i=14&q=Taaltechnologie&search\\_field=subject](https://lib.ugent.be/en/catalog/rug01:002029480?i=14&q=Taaltechnologie&search_field=subject) (31.08.2022).

- Hutchins, J. (2001) Machine translation and human translation - in completion or in complementation? In: *International Journal of Translation* (13(1-2)): 5–21. <https://aclanthology.org/www.mt-archive.info/00/IJT-2001-Hutchins.pdf> (04.08.2022).
- Hutchins, W.; Somers, H. (1992) *An introduction to machine translation*. London: Academic Press.
- Ilisei, I.; Inkpen, D.; Corpas Pastor, G.; Mitkov, R. (2010) Identification of translationese: a machine learning approach. In: 0302-9743. <https://wlv.openrepository.com/handle/2436/622559> (31.08.2022).
- ISO ISO 17100:2015: Translation services – Requirements for translation services. [http://www.iso.org/iso/catalogue\\_detail.htm?csnumber=59149](http://www.iso.org/iso/catalogue_detail.htm?csnumber=59149) (31.08.2022).
- Kamocki, P.; O'Regan, J. (2016) *Privacy issues in online machine translation services – European perspective*. European Language Resources Association (ELRA).
- Kenny, D. (2014) Lexical Hide-and-Seek: Looking for Creativity in a Parallel Corpus. In: Olohan, M. (Hrsg.) *Intercultural faultlines. Research models in translation studies*. London: Routledge: 93–104.
- Kirchhoff, K.; Turner, A.; Axelrod, A.; Saavedra, F. (2011) Application of statistical machine translation to public health information: a feasibility study. In: *Journal of the American Medical Informatics Association : JAMIA* 18(4): 473–478.
- Koehn, P. (2009) A Web-Based Interactive Computer Aided Translation Tool. In: *Proceedings of the ACL-IJCNLP 2009 Software Demonstrations*: 17–20. <https://aclanthology.org/P09-4005/> (31.08.2022).
- Koehn, P. (2010) Enabling Monolingual Translators: Post-Editing vs. Options. In: *Human Language Technologies: Conference of the North American Chapter of the Association of Computational Linguistics, Proceedings, June 2-4, 2010, Los Angeles, California, USA*: 537–545. <https://www.research.ed.ac.uk/en/publications/enabling-monolingual-translators-post-editing-vs-options> (31.08.2022).
- Koponen, M. (2016) Is Machine Translation Post-editing Worth the Effort?: A Survey of Research into Post-editing and Effort. In: *The Journal of Specialised Translation* (25): 131–148. <https://researchportal.helsinki.fi/en/publications/is-machine-translation-post-editing-worth-the-effort-a-survey-of-> (31.08.2022).
- Koppel, M.; Ordan, N. (2011) Translationese and Its Dialects. In: *Proceedings of the 49th Annual Meeting of the Association for Computational Linguistics: Human Language Technologies*: 1318–1326. <https://aclanthology.org/P11-1132/> (31.08.2022).
- Krings, H. (2001) *Repairing Texts: Empirical Investigations of Machine Translation Post-Editing Processes* (Geoffrey S. Koby, ed.), The Kent State University Press, Kent, Ohio & London, 558 p.



- Krüger, R. (2017) Neuronale Maschinelle Übersetzung - Von Netzen und Vektoren. In: *MDÜ* (63(1)): 38–44.
- Kujamäki, P. (2004) What happens to "unique items" in learners' translations? "Theories" and "concepts" as a challenge for novices' views on "good translation". In: Mauranen, A.; Kujamäki, P. (Hrsg.) *Translation universals. Do they exist?* Benjamins: 187–204.
- Lapshinova-Koltunski, E. (2013) VARTRA: A Comparable Corpus for Analysis of Translation Variation. In: *6th Workshop on Building and Using Comparable Corpora*. Sofia, Bulgarien: 77–86.
- Lapshinova-Koltunski, E. (2015) Variation in Translation: Evidence from Corpora. In: Fantinuoli, C.; Zanettin, F. (Hrsg.) *New directions in corpus-based translation studies. (Translation and Multilingual Natural Language Processing 1)*. Berlin: Language Science Press: 93–113.
- Laurian, A.-M. (1984) Machine Translation : What Type of Post-Editing on What Type of Documents for What Type of Users. In: *10th International Conference on Computational Linguistics and 22nd Annual Meeting of the Association for Computational Linguistics*: 236–238. <https://aclanthology.org/P84-1051/> (31.08.2022).
- Laviosa, S. (1998) Core Patterns of Lexical Use in a Comparable Corpus of English Narrative Prose. In: *Meta* 43(4). [https://www.researchgate.net/publication/264870606\\_Core\\_Patterns\\_of\\_Lexical\\_Use\\_in\\_a\\_Comparable\\_Corpus\\_of\\_English\\_Narrative\\_Prose](https://www.researchgate.net/publication/264870606_Core_Patterns_of_Lexical_Use_in_a_Comparable_Corpus_of_English_Narrative_Prose) (31.08.2022).
- Laviosa, S. (2002) *Corpus-based translation studies: theory, findings, applications*. Amsterdam; New York NY: Rodopi.
- Laviosa-Braithwaite, S. (1997) Investigating Simplification in an English Comparable Corpus of Newspaper Articles. In: Klaudy, K. (Hrsg.) *Transfere necesse est. proceedings of the 2nd International Conference on Current Trends in Studies of Translation and Interpreting*. Budapest: Scholastica: 531.540.
- Laviosa-Braithwaite, S. (1998) Universals of Translation. In: Baker, M. (Hrsg.) *Routledge encyclopedia of translation studies*. London: Routledge: 306–310.
- Le, Q.; Schuster, M. (2016) A Neural Network for Machine Translation, at Production Scale. Google AI Blog. <https://ai.googleblog.com/2016/09/a-neural-network-for-machine.html> (08.08.2022).
- Lewandowska-Tomaszczyk, B. (2012) Explicit and tacit: An interplay o the quantitative and qualitative approaches to translation. In: Oakes, M. (Hrsg.) *Quantitative methods in corpus-based translation studies. A practical guide to descriptive translation research*. Amsterdam/Philadelphia: John Benjamins Publishing Company: 3–34.
- Loffler-Laurian, A.-M. (1996) *La traduction automatique*. Villeneuve d'Ascq: Presses Universitaires.

- Mauranen, A. (2004) Corpora, universals and interference. In: Mauranen, A.; Kujamäki, P. (Hrsg.) *Translation universals. Do they exist?* Benjamins: 65–82.
- Mauranen, A.; Kujamäki, P. (2004) Introduction. In: Mauranen, A.; Kujamäki, P. (Hrsg.) *Translation universals. Do they exist?* Benjamins: 1–11.
- Mauranen, A.; Kujamäki, P. (Hrsg.) (2004) *Translation universals. Do they exist?* Benjamins.
- Mitkov, R. (Hrsg.) (1997) *International Conference, Recent Advances in Natural Language Processing*. Philadelphia: J. Benjamins.
- Neumann, S. (2013) *Contrastive Register Variation. A quantitative approach to the comparison of English and German*. Berlin, Boston: De Gruyter Mouton.
- O'Brien, S. (2002) Teaching post-editing: a proposal for course content. In: *Proceedings of the 6th EAMT Workshop: Teaching Machine Translation*. <https://aclanthology.org/2002.eamt-1.11/> (31.08.2022).
- O'Brien, S. (2010a) Introduction to Post-Editing: Who, What, How and Where to Next? In: *AMTA*. <https://www.semanticscholar.org/paper/Introduction-to-Post-Editing%3A-Who%2C-What%2C-How-and-to-O%E2%80%99Brien/54cf88bb96bdf9863120b73d95e7bdd245e6e1cc> (31.08.2022).
- O'Brien, S. (2011) Towards predicting post-editing productivity. In: *Machine Translation* 25(3): 197–215.
- Oakes, M. (2012) Describing a translational corpus. In: Oakes, M. (Hrsg.) *Quantitative methods in corpus-based translation studies. A practical guide to descriptive translation research*. Amsterdam/Philadelphia: John Benjamins Publishing Company: 115–148.
- Oakes, M. (Hrsg.) (2012) *Quantitative methods in corpus-based translation studies. A practical guide to descriptive translation research*. Amsterdam/Philadelphia: John Benjamins Publishing Company.
- O'Brien, S. (2006) Eye-tracking and translation memory matches. In: *Perspectives* 14(3): 185–205.
- O'Brien, S. (2010b) Controlled language and readability. In: Shreve, G.; Angelone, E. (Hrsg.) *Translation and cognition*. Amsterdam; Philadelphia: John Benjamins Pub. Co: 143–165.
- O'Curran, E. (2014) Translation quality in post-edited versus human-translated segments: A case study. In: *AMTA 2014 3rd Workshop on Post-editing Technology and Practice (WPTP-3)*: 113–118.
- Olohan, M. (2001) Spelling Out the Optionals In Translation: A Corpus Study. In: *UCREL Technical Papers*: 423–432. [https://www.academia.edu/277842/Spelling\\_Out\\_the\\_Optionals\\_In\\_Translation\\_A\\_Corpus\\_Study](https://www.academia.edu/277842/Spelling_Out_the_Optionals_In_Translation_A_Corpus_Study) (31.08.2022).

- Olohan, M. (2004) *Introducing corpora in translation studies*. London: Routledge.
- Papineni, K.; Roukos, S.; Ward, T.; Zhu, W.-J. (2002) Bleu: a Method for Automatic Evaluation of Machine Translation. In: *Proceedings of the 40th Annual Meeting of the Association for Computational Linguistics*: 311–318.  
<https://aclanthology.org/P02-1040/> (31.08.2022).
- Parthasarathi, P.; Sinha, K.; Pineau, J.; Williams, A. (2021) Sometimes We Want Translationese. <https://openreview.net/forum?id=51yU9Ht3r3z> (31.08.2022).
- Plitt, M.; Masselot, F. (2010) A Productivity Test of Statistical Machine Translation Post-Editing in a Typical Localisation Context. In: *Prague Bulletin of Mathematical Linguistics* 93(-1): 7–16.  
[https://www.researchgate.net/publication/228849469\\_A\\_Productivity\\_Test\\_of\\_Statistical\\_Machine\\_Translation\\_Post-Editing\\_in\\_a\\_Typical\\_Localisation\\_Context](https://www.researchgate.net/publication/228849469_A_Productivity_Test_of_Statistical_Machine_Translation_Post-Editing_in_a_Typical_Localisation_Context) (31.08.2022).
- Poulis, A.; Kolovratnik, D. (2012) To post-edit or not to post-edit? Estimating the benefits of MT post-editing for a European organization. In: *Workshop on Post-Editing Technology and Practice*. <https://aclanthology.org/2012.amta-wptp.7/> (31.08.2022).
- Povlsen, C.; Underwood, N.; Music, B.; Neville, Anne (1998) Evaluating Text-type Suitability for Machine Translation a Case Study on an English-Danish MT System. In: Rubio, A.; Gallardo, N.; Castro, R.; Tejada, A. (Hrsg.) *Proceedings of the First International Conference on Language Resources and Evaluation*. Granada, Spain: 27–31.
- Puurtinen, T. (2003) Genre-specific Features of Translationese? Linguistic Differences between Translated and Non-translated Finnish Children's Literature. In: *Literary and Linguistic Computing* (18(4)).
- Rabinovich, E.; Wintner, S. (2015) Unsupervised Identification of Translationese. In: *Transactions of the Association for Computational Linguistics* 3: 419–432.  
<https://aclanthology.org/Q15-1030/> (31.08.2022).
- Reiß, K. (1986) *Möglichkeiten und Grenzen der Übersetzungskritik. Kategorien und Kriterien für eine sachgerechte Beurteilung von Übersetzungen*. München: Hueber.
- Roturier, J. (2006) An investigation into the impact of controlled English rules on the comprehensibility, usefulness and acceptability of machine-translated technical documentation for French and German users. <https://doras.dcu.ie/18190/> (31.08.2022).
- RWS (2020) Translation Technology Insights 2020. Coping with the rise in pressure: Why humanizing technology is key to translation success. [http://cracker-project.eu/wp-content/uploads/Europes\\_Leading\\_Role\\_in\\_MT.pdf](http://cracker-project.eu/wp-content/uploads/Europes_Leading_Role_in_MT.pdf) (07.08.2022).
- SDL (2020) Maschinelle Übersetzungstechnologien sind immer mehr verbreitet – was bedeutet das für freiberufliche Übersetzer? (07.08.2022).

- Senez, D. (1998) The machine translation help desk and the post-editing service.  
<https://aclanthology.org/www.mt-archive.info/T&T-1998-Senez.pdf> (02.02.2022).
- Sketch Engine (2016a) non-word | Sketch Engine.  
[https://www.sketchengine.eu/my\\_keywords/non-word/](https://www.sketchengine.eu/my_keywords/non-word/) (01.08.2022).
- Sketch Engine (2016b) token | Sketch Engine.  
[https://www.sketchengine.eu/my\\_keywords/token/](https://www.sketchengine.eu/my_keywords/token/) (24.08.2022).
- Sketch Engine (2017a) English Penn Treebank tagset with modifications | Sketch Engine.  
<https://www.sketchengine.eu/english-treetagger-pipeline-2/> (01.08.2022).
- Sketch Engine (2017b) German RFTagger part-of-speech tagset | Sketch Engine.  
<https://www.sketchengine.eu/german-rftagger-part-of-speech-tagset/>  
 (23.06.2022).
- Sketch Engine (2018) POS tags | Sketch Engine.  
<https://www.sketchengine.eu/blog/pos-tags/> (23.06.2022).
- Sketch Engine (2021) Why are some words in the corpus tagged incorrectly?  
<https://support.sketchengine.eu/help/en-us/5-tags-lemmas-taggers-vertical-file/60-why-are-some-words-in-the-corpus-tagged-incorrectly> (13.08.2022).
- Specia, L. (2011a) Exploiting Objective Annotations for Measuring Translation Post-editing Effort. In: *Machine Translation*.  
[https://www.researchgate.net/publication/270878317\\_Exploiting\\_Objective\\_Annotations\\_for\\_Measuring\\_Translation\\_Post-editing\\_Effort](https://www.researchgate.net/publication/270878317_Exploiting_Objective_Annotations_for_Measuring_Translation_Post-editing_Effort) (31.08.2022).
- Specia, L. (2011b) Quality Estimation of Machine Translation. Abstract to Video talk on the Homepage of DCU, School of Computing, Dublin City University.
- Stein, D. (2009) Maschinelle Übersetzung - ein Überblick. In: *JLCL* (Vol. 24 (3)): 5–18.  
[https://www.researchgate.net/publication/220046884\\_Maschinelle\\_Ubersetzung\\_-\\_ein\\_Uberblick](https://www.researchgate.net/publication/220046884_Maschinelle_Ubersetzung_-_ein_Uberblick) (31.08.2022).
- Steiner, E. (2004) *Translated Texts: Properties, Variants, Evaluations*. Frankfurt: Europäischer Verlag der Wissenschaften.
- Steiner, E. (2012) A characterization of the resource based on shallow statistics. In: Hansen-Schirra, S. (Hrsg.) *Cross-Linguistic Corpora for the Study of Translations. Insights from the Language Pair English-German*. Berlin/Boston: De Gruyter Inc: 71–90.
- Stubbs, M. (1996) *Text and corpus analysis. Computer-assisted studies of language and culture*. Oxford: Blackwell.
- TAUS (2010) Post-editing in Practice. A TAUS Report. <https://www.taus.net>.
- TAUS (2016) Post-Editing Guidelines.  
<https://www.commonspaces.eu/document/39/download/>.

- Teich, E. (2003) *Cross-Linguistic Variation in System and Text. A Methodology for the Investigation of Translations and Comparable Texts*. Berlin, New York: Mouton de Gruyter.
- Temnikova, I. (2010) Cognitive Evaluation Approach for a Controlled Language Post-Editing Experiment.  
[https://www.researchgate.net/publication/220746821\\_Cognitive\\_Evaluation\\_Approach\\_for\\_a\\_Controlled\\_Language\\_Post-Editing\\_Experiment](https://www.researchgate.net/publication/220746821_Cognitive_Evaluation_Approach_for_a_Controlled_Language_Post-Editing_Experiment) (31.08.2022).
- The University of Manchester (2022) Translational English Corpus (TEC).  
<https://www.alc.manchester.ac.uk/translation-and-intercultural-studies/research/projects/translational-english-corpus-tec/> (09.08.2022).
- Tirkkonen-Condit, S. (2002) Translationese – a myth or an empirical fact? In: *Target. International Journal of Translation Studies* 14(2): 207–220.  
[https://www.researchgate.net/publication/233651302\\_Translationese\\_-\\_A\\_myth\\_or\\_an\\_empirical\\_fact\\_A\\_study\\_into\\_the\\_linguistic\\_identifiability\\_of\\_translated\\_language](https://www.researchgate.net/publication/233651302_Translationese_-_A_myth_or_an_empirical_fact_A_study_into_the_linguistic_identifiability_of_translated_language) (31.08.2022).
- Toral, A. (2019) Post-editeese: an Exacerbated Translationese. In: *Proceedings of Machine Translation Summit XVII: Research Track*: 273–281.  
<https://aclanthology.org/W19-6627/> (31.08.2022).
- Toury, G. (2012) *Descriptive translation studies - and beyond*. Amsterdam; Philadelphia: John Benjamins Pub. Co.
- Tymoczko, M. (1998) Computerized Corpora and the Future of Translation Studies. In: *Meta : journal des traducteurs / Meta: Translators' Journal* 43(4): 652–660.  
<https://www.erudit.org/fr/revues/meta/1998-v43-n4-meta169/004515ar/> (31.08.2022).
- Vanderauwera, R. (1985) *Dutch novels translated into English: the transformation of a minority literature*. Amsterdam: Rodopi.
- Vasconcellos, M.; León, M. (1985) SPANAM and ENGSPAN: Machine Translation at the Pan American Health Organization. In: *Computational Linguistics* (11).  
<https://www.semanticscholar.org/paper/SPANAM-and-ENGSPAN%3A-Machine-Translation-at-the-Pan-Vasconcellos-Le%C3%B3n/e05769cf33a509a61589c28aa72ca760792e2e98> (31.08.2022).
- Vashee, K. (2011) An Exploration of Post-Editing MT – Part I. <http://kv-emptypages.blogspot.com/2011/02/exploration-of-post-editing-mt-part-i.html> (09.02.2022).
- Veale, T.; Way, A. (1997) Gaijin : A Bootstrapping, Template-Driven Approach to Example-Based MT. In: Mitkov, R. (Hrsg.) *International Conference, Recent Advances in Natural Language Processing*. Philadelphia: J. Benjamins: 239–244.
- Volansky, V.; Ordan, N.; Wintner, S. (2014) On the features of translationese. In: *Digital Scholarship in the Humanities* 30(1).  
<http://cs.haifa.ac.il/~shuly/publications/vered.pdf> (31.08.2022).

- Wagner, E. (1985) Post-editing Systran - a Challenge for Commission translators. In: European Commission (Hrsg.) *Terminologie et Traduction*.
- White, J.; O'Connell, T.; O'Mara, F. (1994) The ARPA MT Evaluation Methodologies: Evolution, Lessons, and Future Approaches. In: *Proceedings of the First Conference of the Association for Machine Translation in the Americas*. <https://aclanthology.org/1994.amta-1.25/> (31.08.2022).
- Wurm, A. (2008) *Translatorische Wirkung. Ein Beitrag zum Verständnis von Übersetzungsgeschichte als Kulturgeschichte am Beispiel deutscher Übersetzungen französischer Kochbücher in der Frühen Neuzeit*. Frankfurt am Main: Lang.
- Xia, Y.; Li, D. (2010) Specialized Comparable Corpora in Translation Evaluation: A Case Study of English Translations of Chinese Law Firm Adverts. In: Xiao, R. (Hrsg.) *Using Corpora in Contrastive and Translation Studies*. Newcastle-upon-Tyne: Cambridge Scholars Publisher: 62–78.
- Xiao, R. (Hrsg.) (2010) *Using Corpora in Contrastive and Translation Studies*. Newcastle-upon-Tyne: Cambridge Scholars Publisher.
- Yang, Z.; Cheng, Y.; Liu, Y.; Sun, M. (2019) Reducing Word Omission Errors in Neural Machine Translation: A Contrastive Learning Approach. In: *Meeting of the Association for Computational Linguistics*: 6191–6196. <https://typeset.io/papers/reducing-word-omission-errors-in-neural-machine-translation-1fm0ayjluw> (30.08.2022).
- Yuste-Rodrigo, E. (2001) Making MT commonplace in translation training curricula - too many misconceptions, so much potential! In: *Workshop on Teaching Machine Translation*. <https://aclanthology.org/2001.mtsummit-teach.9/> (31.08.2022).