



# Kollektive Intelligenz

Kollektive Intelligenz – Übersetzungsmaschinen und Literatur

## Workflow 3: Post-Editing in Word

bearbeitet von

**Nina Restemeier**

mit einer Review von Katharina Meyer

### Inhalt

Beschreibung des Workflows .....	1
Reflexion: Der Inhalt erschließt sich, doch es bleibt viel zu tun .....	2
Review .....	7

gefördert von

**Deutscher  
Übersetzerfonds**



Kooperationspartner

**Kulturwerk Berliner  
Schriftstellerinnen &  
Schriftsteller e. V.**

## Beschreibung des Workflows

»Diesen Text hat eine Kollegin *in house* übersetzt. Sie müssen nur noch drübersehen, ein Übersetzungslektorat anfertigen.« Die fragliche Kollegin heißt in diesem Fall DeepL, und unsere Teilnehmerinnen haben den maschinenübersetzten Text bearbeitet, gewissenhaft mit dem Original verglichen und ihre Beobachtungen aufgeschrieben.

Arbeitsanweisungen:

1. Text vollständig übersetzen lassen (maschinelle Übersetzung)
2. Satz-für-Satz-Vergleich: Ausgangstext vs. Zieltext
3. Zieltext auf Stimmigkeit und Literarizität prüfen
4. Kontrolldurchlauf
5. Reflexion des Arbeitsprozesses verfassen

Reflexion: Der Inhalt erschließt sich, doch es bleibt viel zu tun

Nina Restemeier lädt eine DeepL-Übersetzung in ihr Textverarbeitungsprogramm und widmet sich dem Post-Editing. Welche Fallstricke diese Vorgehensweise für sie bereithielt, schildert sie in ihrem Arbeitsbericht.

### Die Maschinenübersetzung: Weit entfernt von »gelungen«

Motiviert und neugierig mache ich mich an die Arbeit, überfliege das Original, ein Ausschnitt aus Aubrey Gordons *What We Don't Talk About When We Talk About Fat*, wenn auch nur oberflächlich, um den ersten Entwurf der Maschinenübersetzung (MÜ) unvoreingenommen auf Verständlichkeit zu überprüfen. Dann geht es ans Übersetzen(lassen), und schon stellt DeepL mich vor die erste Entscheidung: Ich kann den Text auf der Website in die Eingabemaske kopieren oder das ganze Word-Dokument hochladen. Ich probiere spaßeshalber beides und vergleiche die Versionen mit der »Dokumente vergleichen«-Funktion von Word. Das Ergebnis überrascht: die Übersetzungen unterscheiden sich geringfügig, sie sind sogar minimal unterschiedlich lang.

Originaltext: 895 Wörter

Gesamte Datei übersetzt: 933 Wörter

Webübersetzung: 922 Wörter

Finale Version: 856 Wörter

Beide Versionen übersetzen *fat* manchmal mit »fett«, manchmal mit »dick«, allerdings an unterschiedlichen Stellen. Die Webversion gewinnt in der Kategorie Idiomatik, wenn sie zum Beispiel »freely removing items they don't think I should eat« mit »nehmen einfach Dinge heraus« übersetzt, während in der übersetzten Datei »nehmen aus freien Stücken Artikel heraus« steht. Auch »Körper wie meiner« ist der etwas sperrigen Formulierung »Körper wie der meine« vorzuziehen. Weniger treffsicher ist die Webversion dagegen, wenn sie das wiederkehrende *abuse* allzu

wörtlich mit »Missbrauch« wiedergibt. Die Übersetzung des kompletten Dokuments benutzt dagegen (zumindest an zwei Stellen) das in diesem Kontext eher treffende »Beschimpfungen«. Doch dazu später mehr.

Für die weitere Arbeit verwende ich schließlich die komplett übersetzte Datei, weil darin Formatierungen wie Kursivschrift erhalten bleiben, die in der Webversion verloren gehen. Ich lese die von DeepL ausgespuckte Übersetzung gründlich und stelle fest: Für ein reines Erfassen des Inhaltes ist sie vollkommen ausreichend. Von einer gelungenen Übersetzung kann dennoch keine Rede sein, darüber kann nicht einmal eine idiomatische Wendung wie »ein gefundenes Fressen« hinwegtäuschen. Folgende Schwierigkeiten möchte ich exemplarisch herausgreifen:

**Wörtlichkeit:** »Ich war schon immer dick.

Nicht *mollig* oder *flauschig* oder *buschig* oder *kurvig-fett*.« Oder: »Mein Body-Mass-Index (BMI) beschreibt meinen Körper als ›super krankhaft fettleibig‹.« Anstatt deutsche Euphemismen für Übergewicht oder die im Deutschen im Zusammenhang mit dem BMI üblichen Formulierungen zu verwenden, übersetzt DeepL, was da steht, mit unfreiwillig komischem Ergebnis.

**Syntax:** An einigen Stellen übernimmt die DeepL-Übersetzung die Satzstellung des Originals, was zu inhaltlichen Verzerrungen führen kann: »Als weiße, queere Frau mag es schwierig sein, die Herausforderungen zu beschreiben, denen ich durch Frauenfeindlichkeit und Homophobie ausgesetzt bin ...« Hier geht es allerdings nicht darum, was der Autorin als weißer, queerer Frau schwerfällt, sondern welche Misogynie und Homophobie einer weißen, queeren Frau entgegengebracht werden.

**Lokalisierung:** In der Maschinenübersetzung werden Gewichtsangaben in Pfund und Kleidergrößen nach dem amerikanischen System wiedergegeben, was bei einer deutschen Leserschaft das Verständnis erschweren kann. Kilogramm und deutsche Konfektionsgrößen rufen hingegen konkretere Vorstellungen hervor.

**Register:** In der automatischen Übersetzung wird ausnahmslos gesiezt, was dem Kontext nach unwahrscheinlich ist. »STEIGEN SIE AUS UND GEHEN SIE ZUR

ABWECHSLUNG MAL ZU FUSS.« Wer fremde Menschen auf der Straße beleidigt, legt sicher keinen Wert auf Höflichkeit. »Haben Sie etwas getan, was sie verärgert hat?« In diesem Satz geht es um der Autorin nahestehende Menschen, die versuchen, das diskriminierende Verhalten anderer zu relativieren.

**Tempus:** Mitten in einem durchgehend im Präsens verfassten Text stehen plötzlich Sätze im Präteritum: »Die meisten fettfeindlichen Haltungen, mit denen ich konfrontiert wurde, kamen von dünnen Menschen, die fest daran glaubten, das Richtige und Gute zu tun, ermutigt durch eine Kultur, die ihnen von ganzem Herzen zustimmte.« Im Original steht der Satz im Perfekt, impliziert also, dass der Verfasserin fortwährend bis in die Gegenwart Fettfeindlichkeit entgegenschlägt. Die DeepL-Übersetzung stellt diese Tatsache dagegen als abgeschlossen dar.

**Realia:** »In diesen Tagen lasse ich meine Fenster meist hochgekurbelt.« Hier ist schön zu erkennen, dass die Maschinenübersetzung mit Wahrscheinlichkeiten arbeitet. Jahrzehntelang wurden Autofenster manuell hochgekurbelt, so dass sich dieses Wort sicher in unzähligen Textbeispielen findet. In Zeiten von automatischen Fensterhebern ist diese Wortwahl dennoch veraltet.

**Wiederholungen:** Die Autorin verwendet mehrmals kurz hintereinander *face/faced with* und *met with*, doch da es sich um kurze Wörter handelt, ist diese Häufung im Original unauffälliger als »ausgesetzt sein« und »auf etwas stoßen« in der DeepL-Übersetzung: »Als weiße, queere Frau mag es schwierig sein, die Herausforderungen zu beschreiben, denen ich durch Frauenfeindlichkeit und Homophobie ausgesetzt bin, aber ich stoße zunehmend auf ein gewisses Maß an Sympathie. Wenn ich jedoch von dem Missbrauch erzähle, dem ich als dicke Person ausgesetzt war, stoße ich häufig auf eine eiserne Ablehnung.«

Zugleich produziert DeepL auch Wiederholungen, wo im Original keine sind: »Ich vermeide den Blickkontakt (*eye contact*) mit Fremden, denn ich weiß, dass unser Blickkontakt (*our locking eyes*) allzu oft als Aufforderung missverstanden wird ...«

Und dann sind da noch die gewollten Wiederholungen. Entsprechend den Bestrebungen der Autorin, *fat* nicht als herabwürdigendes, sondern als ein

neutrales Wort anzusehen, sollte es auch in der Übersetzung durchgehend als »fett« übersetzt werden. DeepL dagegen benutzt mal »dick«, mal »fett«, ohne ein konkretes System erkennen zu lassen.

### **Das Post-Editing: Auf falschem Gleis**

Im ersten Schritt redigiere ich Formales: Anführungszeichen werden händisch an deutsche Konventionen angepasst, Pfundangaben in Kilogramm umgerechnet, US-Kleidergrößen durch europäische ersetzt. Ich suche nach BMI-Tabellen, um die in diesem Zusammenhang üblichen Formulierungen zu verwenden. Ich besuche Reiseblogs und Websites von Airlines, um mich zu informieren, inwiefern übergewichtige Passagiere benachteiligt werden. Dabei stoße ich auch auf den Blog der Autorin und ihren Artikel »Fat isn't a bad word«. Damit ist die Entscheidung gefallen, auch in der Übersetzung durchgehend das Wort »fett« zu benutzen, so befremdlich das für deutsche Leser\*innen zunächst auch wirken mag.

Bei der Überarbeitung gehe ich so vor, wie ich auch bei meinen eigenen Übersetzungen vorgehen würde: Ich versehe unklare Formulierungen mit Fragezeichen und Kommentaren, schreibe viele Varianten und Alternativen in den Text. Dafür gleiche ich anfangs jeden Satz akribisch mit dem Original ab, doch irgendwann fällt mir auf, dass ich mich zunehmend auf den deutschen Text verlasse und ihn zu schleifen versuche, ohne konkrete Formulierungen im Original gegenzuchecken. Dadurch laufe ich Gefahr, Missverständnissen aufzusitzen oder mich zu lange an einer Formulierung aufzuhalten, obwohl ein Blick ins Original sofort Klarheit schaffen könnte.

In der zweiten Überarbeitungsrunde entscheide ich mich zwischen meinen Varianten – oder freue mich, wenn ich auf weitere, noch treffendere Lösungen komme. So fällt mir die Übersetzung »hadern« für *struggle* erst über den Umweg »zu kämpfen haben« ein, da mich DeepL mit seinem Vorschlag »Ich kämpfe nicht mit meinem Selbstwertgefühl« zunächst auf ein falsches Gleis gesetzt hat.

Da ich bei der Übersetzung ohne Glossar gearbeitet habe, fällt mir im Verlauf der Überarbeitung auf, dass ich im ersten Durchgang doch ein paar Sachen übersehen

habe: So haben sich ein »dick«, ein »Pfund« und ein »Missbrauch« der Angleichung entzogen.

Den finalen Korrekturdurchgang erledige ich wie auch bei meinen eigenen Übersetzungen in Papyrus Autor, da dieses Programm über eine hervorragende Synonymdatenbank verfügt. So ist »Demütigung« für *abuse* ein Resultat dieser letzten Überarbeitungsrunde (frühere Varianten waren Beleidigungen, Beschimpfungen, Schmähungen, Kränkungen, die mir aber allesamt zu verbal waren.)

### **Fazit: Keinerlei Arbeitersparnis durch DeepL**

Auch wenn ich nicht mit allen Lösungen hundertprozentig zufrieden bin, stelle ich die Arbeit am Text nach etwa 8 Stunden verteilt auf 4 Arbeitstage bewusst ein. Außerhalb eines Experiments wäre ein solcher Zeitaufwand für 4 Normseiten einfach nicht wirtschaftlich, ganz egal, wie hoch das Honorar ausfällt. Zudem würde im echten Berufsleben auf die Übersetzung noch ein Lektorat folgen, bei dem eventuell verbliebene Zweifelsfälle geklärt werden können.

Ein letztes Mal lasse ich die ursprüngliche Maschinenübersetzung mit meiner finalen Version vergleichen und stelle fest: Es gibt genau 3 Sätze im Text, an denen ich nichts geändert habe. Das Arbeiten mit DeepL ist also zumindest im Workflow »komplette Maschinenübersetzung, händische Überarbeitung« keine Arbeitersparnis.

### **Kleiner Nachtrag**

Mittlerweile gibt es auch DeepL write, das für »bessere Texte im Handumdrehen« wirbt. Aus Interesse gebe ich meine redigierte finale Version dort ein und erhalte eine ganze Reihe weiterer Verbesserungsvorschläge. Auf einige wäre ich tatsächlich selbst nicht gekommen, andere hatte ich bei meiner Überarbeitung bewusst verworfen. Aber das zu analysieren wäre Stoff für ein eigenes Experiment.

## Review

von Katharina Meyer

Als ich die Dokumente für die Peer Review in den Händen halte, würde ich am liebsten mit dem Reflexionstext beginnen, weil ich allein schon die übersichtliche Struktur so ansprechend finde.

Zunächst lese ich aber den Zieltext in seiner finalen Version nach der DeepL-Übersetzung und dem Post-Editing. Die Übersetzung liest sich flüssig und gut und lässt einen deutlichen Ton erkennen.

Zum Vergleich lese ich als nächstes den Ausgangstext und kann den Tonfall aus der Übersetzung wiedererkennen.

Im nächsten Schritt lese ich die Variante, in der die finale Version der von DeepL »ausgespuckten« Variante gegenübergestellt wird. Dabei wird schnell deutlich, dass zwar der Inhalt des Ausgangstexts grob erfasst wurde, aber keine gelungene dem Original entsprechende Übersetzung entsteht. Durch die DeepL-Übersetzung werden zusätzliche Fehlerquellen eingebaut, da beim Post-Editing zwischen Übersetzung und Original noch ein weiterer Text, die KI-Variante, zwischengeschaltet ist. Nina schreibt: »Anfangs gleiche ich akribisch jeden Satz einzeln mit dem Original ab, doch irgendwann fällt mir auf, dass ich mich zunehmend auf den deutschen Text verlasse und ihn zu schleifen versuche, ohne konkrete Formulierungen im Original gegenzuchecken. Dadurch laufe ich Gefahr, Missverständnissen aufzusitzen oder mich zu lange an einer Formulierung aufzuhalten, obwohl ein Blick ins Original sofort Klarheit schaffen könnte.«

Semantische Äquivalenz, lexikalische Adäquatheit und stilistische Korrektheit sind häufig nicht gegeben. Nina führt exemplarisch eine Reihe von Schwierigkeiten an, die sie mit gut nachvollziehbaren Beispielen untermauert, wie die Wörtlichkeit, die zu unfreiwillig komischen Ergebnissen führe, die 1:1-Übertragung der Syntax, die inhaltliche Verzerrungen mit sich brächte, nicht umgerechnete Gewichtsangaben oder Kleidergrößen, die das Verständnis erschweren, Tempusfehler,

Wiederholungen, die nicht im Original vorkommen, oder auch gewollte Wiederholungen, die in der MÜ verloren gehen, oder Registerfehler, die den Text in eine ganz andere Richtung führen.

Die Arbeit der Übersetzerin besteht aus mehr als dem reinen Übertragen von Wörtern, wie Nina mit einem anderen Beispiel deutlich macht: Bei ihrer Recherche suchte sie nach BMI-Tabellen, um die in diesem Zusammenhang üblichen Formulierungen zu verwenden, besuchte Reiseblogs und Websites von Airlines, um sich darüber zu informieren, inwiefern übergewichtige Passagiere benachteiligt werden, und stieß schließlich auch auf den Blog der Autorin und einen ihrer Artikel mit dem Titel »Fat isn't a bad word«. Daraufhin entschied sich Nina dafür, das Wort »fett« zu benutzen. DeepL benutzt trotz der gewollten Wiederholung verschiedene Synonyme, und so geht die Intention der Autorin verloren, »fat« nicht als herabwürdigendes, sondern als neutrales Wort zu gebrauchen. An anderer Stelle weist die Autorin auf die Scheinheiligkeit mancher Sprecher hin, die zwar Euphemismen wie »chubby, fluffy, husky or curvy« anstatt »fat« verwenden, durch ihr Verhalten und ihre Haltung aber eine ganz andere Sprache sprechen. Die MÜ überträgt diese Begriffe wortwörtlich mit *mollig*, *flauschig*, *buschig* oder *kurvig*, wodurch die Intention verloren geht und der Satz unfreiwillig komisch wirkt. Insgesamt entspricht der Ton der MÜ nicht dem der selbstbewussten, mit Witz ausgestatteten Sprecherin des Originals.

Die post-editierte Fassung von Nina halte ich für sehr gelungen, denn sie bringt den Text wieder auf Spur, gibt ihm den richtigen Tonfall, behebt die von ihr aufgezeigten Schwierigkeiten, passt das Register an und findet richtig gute Lösungen. Doch ich gehe stark davon aus, dass sie dieses Ergebnis auch ohne die zwischengeschaltete MÜ erzielt hätte. Wie sie selbst schreibt, bedeutet die Arbeit mit DeepL im Workflow »komplette MÜ, händische Überarbeitung« keine Arbeitersparnis, sondern der Zeitaufwand sei nicht wirtschaftlich.

Diese Erkenntnisse decken sich mit meinen eigenen Erfahrungen: Der Zeitaufwand ist deutlich erhöht, und die Arbeit mit einem weiteren Text bietet zusätzliche Fehlerquellen. Der Priming-Effekt bestimmt den Ton, und man läuft Gefahr, Missverständnissen aufzusitzen. Die Anforderungen an die Post-Editorin sind daher sehr hoch.