



KollektiveIntelligenz

Kollektive Intelligenz – Übersetzungsmaschinen und Literatur

Workflow 6:

CAT-Tool mit Glossar und DeepL-Plug-in

bearbeitet von

Margarita Ruppel

mit einer Review von Daniel Landes

Inhalt

Beschreibung des Workflows	1
Reflexion: Ablenkende Tools: Oft eintönig und teils völlig am Thema vorbei	2
Review	6

gefördert von

**Deutscher
Übersetzerfonds**

**NEU
START
KULTUR**



Kooperationspartner

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

Beschreibung des Workflows

In Workflow 6 setzen die Teilnehmer*innen auf eine Kombination verschiedener Technologien. Computergestützte Programme à la Trados Studio und memoQ (CAT-Tools) kommen in Kombination mit einem DeepL-Plug-in zum Einsatz. Ob es hilft, in diesen Mix noch ein Glossar hineinzuzwerfen?

Arbeitsanweisungen:

1. Terminologie aus Ausgangstext extrahieren und Entsprechungen für Zieltext festlegen
2. Import der Termdatenbank in CAT-Tool
3. Segmente im CAT-Tool mit MÜ-Plugin übersetzen
4. In Word exportieren
5. Zieltext auf Stimmigkeit und Literarizität prüfen
6. Kontrolldurchlauf
7. Reflexion des Arbeitsprozesses verfassen

Reflexion: Ablenkende Tools: Oft eintönig und teils völlig am Thema vorbei

Margarita Ruppel experimentiert mit dem CAT-Tool memoQ und verwendet das DeepL-Plug-in. Durch den Einsatz der Tools kann sie sich nicht mehr so recht mit den Ergebnissen ihrer Arbeit identifizieren.

Vorbereitung: Die Tools

Da ich für gewöhnlich nicht mit einem CAT-Tool arbeite und nur Grundwissen aus früheren Erfahrungen mitbrachte, stand für mich zu Beginn des Experiments die Einarbeitung ins Übersetzungs-Tool memoQ. Die Grundfunktionen wie das Anlegen eines neuen Projektes, Einbinden eines Glossars oder des DeepL-Plug-ins sowie die Arbeitsoberfläche mit dem in einzelne Segmente (Sätze) aufgeteilten Ausgangstext hatte ich schnell verstanden, allerdings kann ich natürlich nicht beurteilen, ob ein erfahrener Nutzer mit dem Tool noch effektiver umgehen würde als ich.

Die Erstellung der Terminologieliste habe ich manuell vorgenommen, so erschien es mir am einfachsten. Ich habe dafür nur 9 Begriffe aus dem Originaltext, einer Passage aus Aubrey Gordons *What We Don't Talk About When We Talk About Fat* herausgesucht. Das Vorgehen an sich erscheint mir nicht sehr sinnvoll. Man muss sich vor dem Übersetzen fragen, welche Begriffe eine feste Entsprechung haben oder haben sollten, diese recherchieren und festlegen. Das mache ich lieber nach und nach während der Textarbeit, denn so kann ich die Begriffe stärker in ihrem Kontext betrachten. Es ist natürlich sinnvoll, wichtige und häufiger vorkommende Begriffe in einer Liste festzuhalten, damit die Übersetzung einheitlich ist, aber sie sollte im Laufe der Textarbeit noch angepasst und um weitere Entsprechungen ergänzt werden können, denn durchgängige 1:1-Lösungen sind meines Erachtens (auch in Sachbüchern) selten und sollten nicht stumpf übernommen werden. Ich kann mir sehr gut vorstellen, dass diese Funktion bei Fachübersetzungen mit

festen Vorgaben des Auftraggebers und gegebenenfalls mehreren Übersetzern hilfreich ist, aber in der Literatur aus meiner Sicht nicht.

Problemfelder: Bezug, Semantik, Idiomatik

Zur Arbeit mit den Vorschlägen von DeepL sei zunächst gesagt: Von 53 Segmenten habe ich bloß 4 ohne Veränderungen übernommen. Sobald Sätze etwas komplizierter oder bildlicher wurden, unterliefen der Maschinenübersetzung (MÜ) sehr eindeutige Fehler. Gravierende Bezugs- oder Bedeutungsprobleme tauchten zum Beispiel in folgenden Sätzen auf:

1) Ausgangstext (AT): »Which dates will mock my body?«

MÜ: »Welche Daten werden meinen Körper verhöhnen?«

2) AT: »But when I disclose the abuse I have faced as a fat person, I am frequently met with a steely refusal to believe it.«

MÜ: »Aber wenn ich von dem Missbrauch erzähle, dem ich als dicke Person ausgesetzt war, weigere ich mich oft, dies zu glauben.«

Auch die aufgezählten Synonyme für das Wort »dick« zu Beginn des Textes gingen teilweise völlig am Thema vorbei: »flauschig«, »buschig«.

Ein weiterer häufiger Grund für Korrekturen waren die fehlende lexikalische Vielfalt (z. B. haben/sein als Verben), Wiederholungen einzelner Wörter und vor allem auch von Satzanfängen und Satzstrukturen, die den Text sehr einförmig klingen lassen.

Die Betonung durch Kursivschreibung in einem Satz hat DeepL falsch gesetzt.

Besonders unidiomatisch war auch die Übertragung mündlicher Sprache:

AT: GET OUT AND TRY WALKING FOR A CHANGE.

MÜ: RAUSGEHEN UND ZUR ABWECHSLUNG MAL LAUFEN.

Erst beim abschließenden Lesedurchgang ist mir aufgefallen, dass auch indirekte Rede an einer Stelle nicht als solche gekennzeichnet war.

Anstatt darauf aufmerksam zu machen, dass eine Maßeinheit wie das Gewicht oder die Kleidergröße für die deutsche Leserschaft umgerechnet werden sollte, hat memoQ mich auf einen möglichen Fehler hingewiesen, weil in der Übersetzung durch meine Umrechnung andere Zahlen als im Original standen. Ich möchte nicht ausschließen, dass dies von den Einstellungen abhängt und sogar eine automatische Umrechnung möglich wäre, und hierin würde ich auch einen potenziellen Nutzen eines solchen Tools sehen. DeepL hat die Maßangaben jedenfalls nicht umgerechnet.

Zu guter Letzt sind mir beim abschließenden Lesen des fertigen, aus memoQ exportierten Texts noch einige Stellen aufgefallen, die man idiomatischer ausdrücken oder aber verdichten konnte (überflüssige Nebensätze). Das ist beim zweiten Durchgang nicht ungewöhnlich, war aber hier meinem Eindruck nach häufiger als bei meinen sonstigen Übersetzungen, da ich normalerweise direkt im ersten Durchgang stärker auf dichtere bzw. idiomatische Formulierungen achte und mich dazu auch schon stärker vom Original löse. Ein positiver Aspekt des Einsatzes von memoQ zeigte sich darin, dass im zweiten Durchgang die Kontrolle ausgelassener Sätze wegfällt. Dafür lässt sich der fertige Text dann intensiver auf Flüssigkeit und Sound prüfen, was aber wiederum auch dringend nötig ist.

Der Flow stellt sich nicht ein

Insgesamt lässt sich innerhalb dieses Versuchs schwer sagen, ob ich durch die eingesetzten Tools Zeit gespart oder zusätzlich aufgewendet habe – oder weder noch. Der gesamte Prozess fühlt sich jedoch zäher an. Ich muss zunächst den Ausgangssatz lesen und verstehen, dann den MÜ-Vorschlag lesen und überlegen, was dort eventuell richtig oder falsch ist. Manchmal ist das nicht auf den ersten Blick ersichtlich und erfordert einiges an Denkarbeit. Diese Auseinandersetzung mit einem ersten Übersetzungsergebnis (oder eher -vorschlag) kostet Zeit, die ich auch direkt in die Formulierung meiner eigenen Übersetzung hätte investieren können. So stellt sich außerdem kein »Flow« beim Übersetzen ein, der eine

angenehme Arbeitsweise schafft, Kreativität fördert und dadurch auch gute Ergebnisse hervorbringt.

Texttreue versus Kreativität

Mir ist besonders aufgefallen, dass die Auseinandersetzung mit der MÜ mein Dilemma zwischen Freiheit und Treue zum Original verstärkt. Wenn ich einen Satz umformulieren möchte, stelle ich mir mit der MÜ-Lösung vor Augen viel eher die Frage, ob diese Freiheit berechtigt ist. Ich weiß zwar, dass gerade diese Freiheit einen Text lebendig macht und die Treue zum Original in der Literatur genauso gut auf der Ebene des Gesamtwerks und seiner Wirkung gewahrt werden kann. Bei der Arbeit mit DeepL rückt dieses Wissen bei mir jedoch in den Hintergrund.

Womöglich trägt auch die Arbeit mit memoQ dazu bei, da hier der Text in einzelne Sätze segmentiert wird, was dazu verleitet, ihn weniger in seiner Gesamtheit zu betrachten. Das lässt sich natürlich nach dem Export des Ergebnisses in einem zweiten Durchgang überarbeiten. Trotzdem hat sich mir insgesamt kein Vorteil des Einsatzes der Tools gezeigt. Für mich besteht kein Zweifel daran, dass ich den Übersetzungsprozess in diesem Experiment als anstrengender und weniger schöpferisch empfinde und mich mit dem Ergebnis weniger identifiziere.

Review

von Daniel Landes

Einleitung

Im Rahmen des MÜ-Workshops von Kollektive Intelligenz setze ich mich mit der Übersetzung eines Textes aus Aubrey Gordons Buch *What We Don't Talk About When We Talk About Fat* auseinander. Die Übersetzung stammt aus der Feder von Margarita Ruppel, die mit dem CAT-Tool memoQ unter Zuhilfenahme des DeepL-Plug-ins gearbeitet hat. Für meine Analyse stehen mir der Originaltext, der finale deutsche Text sowie ein Word-Dokument mit den Änderungen, die nach dem Export aus der Übersetzungssoftware vorgenommen wurden, zur Verfügung.

Gedanken zum Text

Auf den ersten Blick liest sich der fertige Text wie eine professionelle Übersetzung, die ihren Weg so auch in den Buchladen finden könnte. Es deutet erst einmal nichts auf die zugrunde liegende Kollaboration von Mensch und Maschine hin. Diese Tatsache ist in meinen Augen auf Margaritas fähiges Selbstlektorat zurückzuführen. Nach eigenen Aussagen hat die Übersetzerin nämlich auch beim zweiten Durchlesen noch einige Verbesserungen vorgenommen, und zwar deutlich mehr als bei einer Übersetzung ohne KI-Unterstützung.

Vergleich mit einem gesondert erstellten Roh-Output

Um den ersten Schritt der Postedition (Bearbeitung eines maschinell übersetzten Textes durch eine professionelle Sprachexpertin) und die übersetzerische Leistung meiner Kollegin besser beurteilen zu können, vergleiche ich ihren Text zunächst mit dem »rohen« Output, den DeepL liefert, wenn man die Textpassage dort eingibt. Da die zugrunde liegenden Sprachmodelle nicht immer dieselben Ergebnisse liefern, hat Margarita nicht mit exakt demselben Output gearbeitet. Dennoch sollte die generelle Qualität vergleichbar und eine aussagekräftige Gegenüberstellung demnach möglich sein.

Um mir vor Augen zu führen, wie tiefgreifend die Arbeit des Posteditierens bei einem solchen Text ist, erstelle ich eine Excel-Tabelle mit verschiedenen Kategorien. In der ersten Spalte liste ich die Textstellen, die die MÜ gut gelöst hat (4 Zitate). Lösungen, die mir bei der MÜ besser gefallen haben als bei der posteditierten Version, kommen in die zweite Spalte (2 Zitate). In Spalte drei werden die Stellen aufgeführt, die der Mensch durch Umschreiben besser gelöst hat (42 Zitate).

Die Zahlen sprechen meiner Meinung nach eine deutliche Sprache: Ist die linguistische Qualität eines Textes wichtig (kommt es also auf Nuancen, Stimmung und Humor an), so ist es unabdingbar, die Übersetzung von einem Menschen vornehmen zu lassen, der all die Aspekte souverän übertragen kann, an welchen die Algorithmen regelmäßig scheitern.

Das lässt sich gleich an den ersten beiden Sätzen aufzeigen: Aus »I have always been fat. Not chubby or fluffy or husky or curvy—fat« macht DeepL »Ich war schon immer dick. Nicht mollig oder flauschig oder buschig oder kurvig-fett«. Die KI übersetzt hier wörtlich, ohne den Worten im Kontext eine angemessene Bedeutung zu geben. Margaritas Lösung geht über die Wörtlichkeit hinaus und sorgt dafür, dass die deutschsprachige Leserschaft nicht ratlos zurückbleibt: »Ich war schon immer dick. Nicht mollig, weich, stämmig oder kurvig – sondern dick.«

Hinzu kommt, dass die Maschine mit Umrechnungen von angloamerikanischen Maßeinheiten ins metrische System Schwierigkeiten hat. Ich könnte mir jedoch vorstellen, dass sich diese Schwäche in Zukunft programmiertechnisch beheben lässt.

Der Kontext dürfte in meinen Augen allerdings weiterhin ein Problem bleiben. Auch wenn große Sprachmodelle uns menschliche Intelligenz vorgaukeln, basieren die Prozesse dahinter auf Wahrscheinlichkeitsrechnung. Die sogenannte künstliche Intelligenz versteht nicht, was sie von sich gibt. Wer eine gute Übersetzung liefern will, muss jedoch verstehen, was er oder sie liest, um die Inhalte und den Subtext adäquat in der Zielsprache wiederzugeben.

Vergleich mit memoQ-Export

Wie Margarita in ihrer Reflexion schreibt, zeigen ihre Änderungen während des zweiten Durchlesens, dass sie den Text vor allem verdichtet und Formulierungen verändert, die in der DeepL-Fassung zu wörtlich sind.

Ein gutes Beispiel dafür ist die Stelle »[...] stoße ich immer wieder auf die felsenfeste Weigerung, es zu glauben«, welche zu »[...] stoße ich immer wieder auf eine unbeirrbar Ungläubigkeit« wird.

Ein letzter Feinschliff in der Wortwahl (»fett« zu »dick«, »Ernährungsratschläge« zu »Ernährungstipps«) springt ebenfalls ins Auge, doch meiner Erfahrung nach ist das auch bei Übersetzungen ohne maschinelle Assistenz nicht ungewöhnlich.

Es gibt zudem Stellen (11 Zitate in meiner Exceltabelle), bei denen ich mich frage, ob die MÜ den Menschen nicht aufs Glatteis geführt hat. Die Passage »Wer wird mir Dinge hinterherrufen? Welche Ärzte werden sich weigern, mich zu behandeln? [...] Welche Fremden werden mich fotografieren und ein Meme aus mir machen?« erinnert mich beispielsweise sehr an das englische *will*, das ich im Deutschen oft als wenig idiomatisch empfinde. Hätte meine Kollegin hier auch das Futur gewählt, wenn die KI ihr keinen fertigen Satz präsentiert hätte?

Gedanken zur Reflexion

Ich stimme Margarita in dem Punkt zu, dass die Glossarerstellung vor Beginn der Übersetzung nicht sehr sinnvoll ist. Wie sie bereits schreibt, ergibt sich die beste Begrifflichkeit oft aus dem Kontext. Außerdem ist der Textauszug zu kurz, um Glossarbegriffe verlässlich identifizieren zu können.

Bei meiner sonstigen Arbeit mit CAT-Tools habe ich allerdings die Erfahrung gemacht, dass es helfen kann, Glossarbegriffe vor Übersetzungsbeginn zu identifizieren, ohne sie gleich zu definieren. Das Glossar lässt sich auch später befüllen, wenn man bei der Übersetzung zum ersten Mal auf den fraglichen Begriff stößt.

Ich teile außerdem Margaritas Einschätzung, dass die MÜ Schwierigkeiten bekommt, sobald Sätze komplizierter oder bildlicher werden und eine wörtliche Übersetzung nicht ausreicht. Gerade bei der Arbeit mit dem MÜ-Plug-in in einem CAT-Tool sind auch einförmige Satzanfänge und Satzstrukturen ein Problem, da durch die Segmentierung sowohl Mensch als auch Maschine dazu verleitet werden, Sätze in Isolation zu betrachten. Hier ist ein zweites Durchlesen unabdingbar.

Dass die Kontrolle ausgelassener Sätze wegfällt, erleichtert mir die Arbeit ebenfalls enorm, doch auch ich habe die Erfahrung gemacht, dass der Arbeitsprozess zähflüssiger ist und sich kein Flow einstellt (was jedoch an der konzentrationsintensiven Arbeit des Posteditierens liegt und nicht am CAT-Tool an sich).

Anders als Margarita stelle ich mir allerdings nicht die Frage, wie viel Freiheit bei der Übersetzung berechtigt ist. Ich habe vielmehr das Gefühl, dass ich beim Einsatz von DeepL freier übersetze als sonst, da ich die enorme Wörtlichkeit der MÜ kompensieren möchte.

Abschließende Gedanken

Durch die Auflistung von konkreten Textbeispielen in meiner Tabelle konnte ich deutlich sehen, wie viel die Übersetzerin trotz MÜ geleistet hat, um einen qualitativ hochwertigen Text zu produzieren. In Zeiten des unkritischen KI-Hypes war das für mich persönlich eine ermutigende Erkenntnis: Vielleicht ist das Ende des menschlichen kreativen Schaffensprozesses doch noch nicht erreicht.

Mich würde es nun reizen, einen direkten Vergleich zwischen dem posteditierten Text sowie der Übersetzung desselben Textes ohne Zuhilfenahme technischer Mittel zu sehen. Inwieweit würde sich die Stimme der Übersetzerin deutlicher manifestieren? Welche Merkmale würden schwächer, welche stärker ausgeprägt auftreten? Und wie sieht es mit dem zeitlichen Aufwand aus, um eine ähnliche Qualität zu erreichen? Vielleicht können weitere Experimente diese Fragen in der Zukunft beantworten.