

Übungsblatt 1 – Musterlösung

1.1 Maschinelle Übersetzung

(a)

Original	Übersetzung
<p>Your letter reached me just a few days ago. I want to thank you for the deep and loving trust it(1) revealed. I can do no more(2). I cannot comment on the style of your(3) verses; critical intent is too far removed from my nature. [...] You ask whether your poems are good. You send them to publishers; you compare them with other poems; you are disturbed(4) when certain publishers reject your attempts. Well now, since you have given me permission to advise you, I suggest that you give all that up. You are looking outward and, above all else, that you must not do now(5). No one can advise and help you, no one.</p>	<p>Dein Brief hat mich erst vor wenigen Tagen erreicht. Ich möchte dir für das tiefe und liebevolle Vertrauen danken, das es(1) offenbart hat. Ich kann nicht mehr(2). Über den Stil Ihrer(3) Verse kann ich nichts sagen; kritische Absicht ist zu weit von meiner Natur entfernt. [...] Du fragst, ob deine Gedichte gut sind. Sie senden sie an Verlage; Sie vergleichen sie mit anderen Gedichten; Sie werden gestört(4), wenn bestimmte Verlage Ihre Versuche ablehnen. Nun, da Sie mir die Erlaubnis erteilt haben, Sie zu beraten, schlage ich vor, dass Sie das alles aufgeben. Sie schauen nach außen und vor allem, was Sie jetzt nicht tun dürfen(5). Niemand kann dich beraten und dir helfen, niemand.</p>

Ein Problem, das sich durch diesen Übersetzungsvorschlag zieht, ist das der Pronomenwahl: Mehrmals wird zwischen *du* und *Sie* gewechselt. Die **anaphorische Referenz**, also Rückbezug, zwischen den Pronomen wird nicht berücksichtigt. Um solche Bezüge richtig herzustellen, muss man den Gesamtkontext betrachten, in dem ein Satz geäußert wird, was für NLP-Systeme wie maschinelle Übersetzer noch immer eine große Herausforderung darstellt. Ähnlich verhält es sich mit Fehler (1), bei dem die Kongruenz zwischen dem Pronomen und dem vorangegangenen Bezugsnomen nicht gegeben ist.

Des Weiteren problematisch sind ungewöhnlich verwendete Ausdrücke wie *disturbed* bei Fehler (4), das man wohl in den meisten Fällen tatsächlich am ehesten als *gestört* übersetzen würde. Ein beinahe lyrisch anmutender Sprachstil wie in diesem Textabschnitt fördert solche Ausdrucksweisen natürlich. Genauso führt der ungewöhnliche Satzbau bei *that you must not do now* zu Fehler (5). Fehler (2) wäre leicht behoben, indem man das Wort (*für Sie*) *tun* einfügt. Hier entsteht durch diese Auslassung eine Formulierung (*Ich kann nicht mehr.*), die im Deutschen eine starke Eigenbedeutung hat. Diese Nuance konnte die maschinelle Übersetzung hier nicht fassen.

(b)

#	Original	Fehler	Korrektur
1	es	Kongruenz	er
2	kann nicht mehr	Sinnfehler	kann nicht mehr tun
3	Ihrer	Köhäsion	deiner
4	gestört	Sinn/Lexik	verärgert
5	was Sie jetzt...	Sinn/Satzbau	das dürfen Sie nicht mehr tun

Damit der Text im Deutschen wirklich natürlich klingt, müsste man ihn vielleicht noch freier umgestalten.

(c) Text 1: News (http://www3.nhk.or.jp/nhkworld/en/news/20161121_29/)

Original	Übersetzung
Death toll in Indian train derailment rises to 145	Todesgebühr in der indischen Zugentgleisung steigt auf 145
The death toll following the derailment of a train in northern India has risen to 145, with 170 others injured .	Die Todesopfer nach der Entgleisung eines Zuges im Norden Indiens ist auf 145 gestiegen, mit 170 anderen verletzt .
The 14-car express train ran off the rails near Kanpur in the state of Uttar Pradesh early on Sunday.	Die 14-Wagen-Express-Zug lief die Schienen in der Nähe von Kanpur in den Zustand von Uttar Pradesh früh am Sonntag.
Rescuers worked overnight at the site using lights and heavy machinery.	Rettungskräfte arbeiteten über Nacht am Ort mit Licht und schweren Maschinen.
The Railway Ministry said on Monday that poor maintenance of railway tracks may have led to the accident.	Das Eisenbahnministerium sagte am Montag, dass schlechte Wartung der Eisenbahnschienen zum Unfall geführt haben kann.
India has one of the world's largest railway networks.	Indien hat eines der weltweit größten Eisenbahnnetzwerke.

Ähnlich wie bei der Übersetzung des Twain-Textes macht Google Translate hier hauptsächlich Ausdrucksfehler bzw. übersetzt einzelne Wörter falsch. Allerdings gibt es auch Fehler im Satzbau (Die Todesopfer [...] ist gestiegen) sowie Genusfehler (Die 14-Wagen-Express-Zug). Besonders interessant ist hier die unterschiedliche Übersetzung von "death toll": In der Überschrift wird dies als "Todesgebühr" übersetzt, während es im nächsten Satz zu "Todesopfer" wird.

Text 2: Gedicht (<http://www.gandjlawrence.co.uk/Werdz/buy.htm>)

Original	Übersetzung
Beware of heard, a dreadful word That looks like beard and sounds like bird.	Vorsicht vor Gehör, ein schreckliches Wort Das sieht aus wie Bart und klingt wie Vogel.
And dead: it's said like bed not bead - For goodness sake don't call it "deed".	Und tot: es heißt wie Bett nicht Perle - Um Himmels willen nicht nennen es "Tat".

Wie erwartet gehen in der Übersetzung Reime und Versschemata verloren. Dadurch wird bei diesem Gedicht allerdings der Sinn nicht mehr erkennbar (Gehör [...] sieht aus wie Bart und klingt wie Vogel). Auch hier finden sich wieder Ausdrucks- ("it's said like" wird zu "es heißt wie") und Syntaxfehler (nicht nennen es "Tat").

1.2 Informationssuche

- (a) Sucht man direkt auf Englisch nach der Frage in Google, bekommt man zunächst das Geburtsdatum von Barack Obama angezeigt. Darunter befinden sich Seiten, die sich zwar mit Obama befassen, allerdings nicht das gewünschte Ergebnis liefern. Sucht man auf Deutsch, bekommt man eine Liste der US-Präsidenten sowie den Wikipedia-Artikel über Obama ausgeworfen.

Mit den Ergebnisseiten der deutschen Suchanfrage kann man sowohl das Geburtsdatum als auch den amtierenden US-Präsidenten zu dieser Zeit (Kennedy) finden. Da man mit den englischen Ergebnissen nur Obamas Geburtsdatum bekommt, ist hier eine zweite Anfrage (z.B. "us president 1961") nötig, um die Frage zu beantworten.

Sucht man auf `google.com/ncr` nach der etwas modifizierten Frage "Who was president when Obama was born", werden direkt die US-Präsidenten im Jahr 1961 angezeigt.

Sucht man in der semantischen Suchmaschine "Wolfram Alpha" nach der ursprünglichen Fragestellung, wird diese als `President(United States, date_of_birth(Barack Obama))` interpretiert und sofort das gewünschte Ergebnis ausgegeben.

Die Frage lässt sich nur schwierig beantworten, da Google in erster Linie keine semantische Suchmaschine ist, sondern nach Seiten sucht, die die gewünschten Suchbegriffe enthalten. Da es (fast) keine Seiten gibt, welche die Fragestellung und Antwort enthalten, führt die direkte Suche nach der Frage nicht zum Erfolg.

- (b) Viele Fragen, die Sonderzeichen beinhalten können nicht mit einer einzigen Suchanfrage beantwortet werden, da Google viele dieser Sonderzeichen ignoriert, wenn diese mit Text zusammen stehen. So sind zum Beispiel mindestens zwei Suchanfragen nötig, um die Bedeutung der Tilde unter Unix herauszufinden ohne deren Namen zu kennen, da "`~ unix`" als "`unix`" interpretiert wird. "`~`" liefert allerdings Informationen zur Tilde und "`unix tilde`" dann die gewünschte Information zur Verwendung.

Soeben beschriebener Spezialfall lässt sich auf umschreibende Ausdrücke verallgemeinern. Ist der genaue Begriff des Gesuchten nicht bekannt, führt die Verwendung einer Umschreibung nur dann zum Erfolg, wenn bereits Andere die gleiche Beschreibung genutzt haben, um die gleiche Frage zu stellen. Generell ergibt sich daraus das Problem, dass Suchen mit steigender Wortzahl ungenau werden dann die gewünschte Information zur Verwendung.

1.3 Morphologie

- (a) Annäherung (Derivation)

An	näher	ung
Prefix	Stamm	Suffix

- (b) ausverkauft (Flexion & Derivation)

Aus	ver	kauf	t
Präfix	Präfix	Stamm	Flexionssuffix

Aus und *ver* können wir genauer als Derivationspräfixe bezeichnen.

- (c) Handelskontor (Komposition)

Handel	s	Kontor
Bestimmungswort	Fugenelement	Grundwort

1.4 Ambiguität

- (a)
- Ich habe Ente für sie zubereitet.
 - Ich habe ihre Ente zubereitet.
 - Ich habe ihre Ente hergestellt.
 - Ich habe sie dazu veranlasst, sich zu ducken.
 - Ich habe sie in "Ente" verwandelt.

- (b) *Er geht zur Bank.*

- Er geht zum Geldinstitut
- Er geht zur Sitzgelegenheit

Die Ambiguität in diesem Satz resultiert aus der Ambiguität des Substantives *Bank*. Dieses hat mehrere Bedeutungen, die normalerweise aus dem Kontext des Satzes deutlich werden. Da es hier aber keinen Kontext gibt, ist der Satz ambig.

1.5 Coli-Wiki und E-Mail

- (a) Die ersten drei Themen im User-Wiki (die obersten Verweise auf andere "Unterwikis" ausgenommen) sind:
- Allegro
 - Bachelorpool
 - cl4lrl
- (b) Der Coli-Account wurde erfolgreich auf meinem Mail-Clienten eingerichtet.