

Mensch wird damit als Herrscher über die Natur etabliert, der über weitreichende Möglichkeiten technischer Machbarkeit verfügt. Die erschaffenen Maschinen unterscheiden sich jedoch fundamental vom Menschen. Selbst wenn sie äußerlich wie Menschen aussähen, wäre eine Unterscheidung, wie Descartes betont, dennoch klar:

„Das erste wäre, dass diese Maschinen nie sich der Worte oder Zeichen bedienen können, durch deren Verbindung wir unsere Gedanken einem Anderen ausdrücken. Man kann zwar sich eine Maschine in der Art denken, dass sie Worte äusserte, und selbst Worte auf Anlass von körperlichen Vorgängen, welche eine Veränderung in ihren Organen hervorbringen; z.B. dass auf eine Berührung an einer Stelle sie fragte, was man wolle, oder schrie, dass man ihr weh thue, und Aehnliches; aber niemals wird sie diese Worte so stellen können, dass sie auf das in ihrer Gegenwart Gesagte verständig antwortet, wie es doch selbst die stumpfsinnigsten Menschen vermögen.

Zweitens würden diese Maschinen, wenn sie auch Einzelnes ebenso gut oder besser wie wir verrichteten, doch in anderen Dingen zurückstehen, woraus man entnehmen könnte, dass sie nicht mit Bewusstsein, sondern bloß mechanisch nach der Einrichtung ihrer Organe handelten. Während die Vernunft ein allgemeines Instrument ist, das auf alle Arten von Erregungen sich äussern kann, bedürfen diese Organe für jede besondere Handlung auch eine besondere Vorrichtung, und deshalb ist es moralisch unmöglich, dass es deren so viele in einer Maschine giebt, um in allen Vorkommnissen des Lebens so zu handeln, wie wir es durch die Vernunft können.“
(Descartes 1637, Abschnitt V)

Intentionales Sprechen außerhalb eines Reiz-Reaktion-Schemas und Vernunft als Ausdruck des Bewusstseins sind bereits bei Descartes als zentrale Unterscheidungskriterien festgeschrieben. Literarisch wird diese Distinktion in E.T.A Hoffmanns 1817 erschienener Erzählung *Der Sandmann* diskutiert. Der Protagonist und Physikstudent Nathanael verliebt sich in die Androide Olimpia. Blind vor Liebe erkennt er trotz verschiedener Warnungen nicht, dass es sich bei ihr nicht um einen Menschen, sondern um ein künstlich erschaffenes Wesen handelt, denn er beschreibt sie als nahezu vollkommene Bürgersfrau: „Ein hohes, sehr schlank im reinsten Ebenmaß gewachsenes, herrlich gekleidetes Frauenzimmer saß im Zimmer vor einem kleinen Tisch, auf den sie beide Ärme, die Hände zusammengefaltet, gelegt hatte. Sie saß der Türe gegenüber, so, dass ich ihr engelschönes Gesicht ganz erblickte.“ (Hoffmann 2015, 17) Lediglich die Augen, die Tür zur Seele des Menschen, erscheinen sonderbar, denn in ihnen lag „etwas Starres, beinahe möchte ich sagen, keine Sehkraft“ (17). Die Mechanisierbarkeit des Geistes und der Seele wird als Unmöglichkeit im Ausdruck leerer, befremdlicher Augen festgeschrieben, ein Topos, der sich mannigfaltig bei künstlichen Geschöpfen in der Literatur wiederholt. So sind es auch die Augen, seine wie ihre, die Nathanael schließlich die Erkenntnis gewinnen lassen, dass Olimpia eben keine Frau, sondern ein Maschinenmensch ist: „Erstarrt stand Nathanael – nur zu deutlich hatte er gesehen, Olimpias toderbleichtes Wachsgesicht hatte keine Augen, statt ihrer schwarze Höhlen; sie war eine leblose Puppe.“ (37)

Dennoch bürden Wesen, die von Menschen gottgleich erschaffen werden, ihren Schöpfer*innen die moralische Frage der Verantwortung auf, welche zentral in Mary Shelleys 1818 geschriebenem Roman *Frankenstein* gestellt wird. Der wissbegierige, geniale junge Frankenstein studiert Naturwissenschaften und folgt schließlich manisch seinem Ziel, das Geheimnis des Lebens zu lüften und eine lebendige Kreatur aus Leichenteilen künstlich zu erschaffen: „With an anxiety that almost amounted to agony, I collected the instruments of life around me, that I might infuse a spark of being into the lifeless thing that lay at my feet.“ (Shelley 2010, 45). Doch die äußere Erscheinung, die aus einer Auswahl ästhetisch-schöner Teile hervorgegangen ist, schockiert ihn zutiefst:

“His yellow skin scarcely covered the work of muscles and arteries beneath; his hair was of alustrous black, and flowing; his teeth of pearly whiteness; but these luxuriances only formed a more horrid contrast with his watery eyes, that seemed almost of the same colour as the dun-white sockets in which they were set, his shrivelled complexion and straight black lips.” (45)

Bereits hier wird die rekurrierende, narrative Bedeutung deutlich, die den Augen künstlicher Geschöpfe in der Literatur zugewiesen wird, sind doch sie es, die hier in besonderer Weise als unmenschlich, gar abstoßend hervorgehoben werden.

Frankenstein weist im Laufe der Erzählung jegliche Verantwortung für seine Schöpfung zurück und flieht. Schuld tragen falsche Lehrer, Verlockungen und die Gefahr der Wissenschaft an sich. Das „Monster“ jedoch lernt gleich einem Kind selbstständig die Wesensarten der Menschen ebenso wie die Sprache durch Beobachtung kennen² (vgl. Shelley 2010, 150f. und 160f.). Voll Hoffnung und Zuversicht getraut es sich schließlich, den Menschen zu begegnen, doch erneut wird es brutal zurückgewiesen, da sein Äußeres den Menschen Furcht und Angst einflößt. Diese emotionale Verletzung macht aus der Kreatur erst das Monster (vgl. Shelley 2010, 179ff.), als das es Frankenstein von Beginn an gesehen hat. Nicht die Kreatur selbst trägt Böses in sich, sondern der Mensch ist durch seine Erschaffung und seine emotionale Zurückweisung für diese Entwicklung verantwortlich.

Frankenstein gilt als erster Science-Fiction-Roman, da er den Erschaffungsprozess und dessen moralischen Folgen, die Interaktion zwischen Mensch und Kreatur sowie den Kontrollverlust ins Zentrum stellt.³ Im Science-Fiction-Bereich sind es dann vor allem Roboter⁴, die als künstlich erschaffene Wesen thematisiert werden. Literarische Texte hinterfragen die erkenntnistheoretischen, emotionalen, aber auch lebensweltliche Folgen von einer drohenden Austauschbarkeit mit den Menschen, da Roboter schon rein äußerlich dem Erscheinungsbild von Menschen

² In dieser Hinsicht nimmt die Kreatur Selbst-Lernprozesse im Bereich Künstlicher Intelligenz bereits vorweg und damit auch die Frage der Verantwortung für die Initiierung und Programmierung von selbstlernenden Algorithmen.

³ Daher wird oft der Begriff „Frankenstein-Mythos“ verwendet, um auf eben diese literarischen Topoi zu verweisen.

⁴ Der Begriff Roboter wurde, wie weithin bekannt, erstmals in Karel Capeks *R.U.R. – Rossum Universal Robots* von 1920 verwendet. Er leitet sich vom tschechischen Begriff „robota“ (Arbeit) ab und wird in Folge in der Literatur weitgehend anstelle des Terminus „Automat“ verwendet.

mehr und mehr entsprechen. Nicht nur Fragen zu maschineller Intelligenz, der Möglichkeit von Gefühlen oder der Existenz eines freien - maschinellen - Willens werden dabei reflektiert, auch Bedrohungsszenarien spielen weiterhin eine große Rolle: Was passiert, wenn sich intelligente Systeme gegen den Menschen wenden, außer Kontrolle geraten oder ein Ziel verfolgen, für das die Umkehrung des Kontrollverhältnisses oder gar die Auslöschung der Menschheit der beste Weg ist (z.B. zur Verhinderung einer Klimakatastrophe)⁵? Bostrom (2018) warnt, „dass uns durch die Erschaffung einer maschinellen Superintelligenz wahrscheinlich eine solche existentielle Katastrophe droht“ (164) und definiert dieses existentielle Risiko als Situation, „wenn das auf der Erde entstandene intelligente Leben vom Aussterben bedroht ist oder die Gefahr besteht, dass sein zukünftiges Entwicklungspotential zumindest dauerhaft und drastisch beschnitten wird.“ (164)

1.2 Isaac Asimov, *The Bicentennial Man* (1976)

Isaac Asimov gilt als der prominenteste Schriftsteller im Bereich Roboterliteratur. Nicht nur, dass er Begriffe wie „positronisches Gehirn“ geprägt hat, er hat auch die drei Robotergesetze ins Leben gerufen, die bis heute Wirkung zeigen.⁶ Mit eben diesen Gesetzen beginnt auch die Erzählung *The Bicentennial Man*:

“The Three Laws of Robotics:

1. A robot may not injure a human being or, through inaction, allow a human being to come to harm.
2. A robot must obey the orders given it by human beings except where such orders would conflict with the First Law.
3. A robot must protect its own existence as long as such protection does not conflict with the First or Second Law.” (Asimov 1976, 135)

Zu Beginn der Geschichte wird Andrew Martin vorgestellt, der eine offensichtlich letale Operation von einem hochspezialisierten Roboter-Chirurgen durchführen lassen möchte. Andrew hat ein adrettes Äußeres und beginnt mit seinem Gegenüber eine Diskussion zum Streben nach Menschsein, wenn er ihn fragt: „Have you ever thought you would like to be a man?“ (136) Daraufhin antwortet der Roboter-Chirurg nur, dass sein einziges Streben der Perfektion seines Handwerks gälte und er ansonsten zufrieden mit dem status quo wäre. Im Rückgriff auf die Robotergesetze weist der Chirurg darauf hin, dass er die gewünschte Operation bei Andrew nicht durchführen könne, woraufhin dieser entgegnet: „‘On a human being, you must not,’ said Andrew, ‘but I, too, am a robot.’“ (136) Bei der Lektüre wird Andrew

⁵ Diese kritische Sicht auf KI-Entwicklung ist durchaus kulturell unterschiedlich ausgeprägt. Literarische Werke scheinen dies nicht nur widerzuspiegeln, sondern auch proaktiv zu beeinflussen. Als Beispiel führt Wolfangel (2020) die größere Offenheit japanischer Bürger*innen gegenüber technischer Aufrüstung des eigenen Körpers oder des Einsatzes von Robotern in verschiedenen Lebensbereichen an und begründet dies u.a. durch den Einfluss der Anime, denn die „Kunstfiguren transportieren ein positives Bild der menschlichen Erweiterung“ (32). So wird zuerst die Unterstützungsmöglichkeiten durch Roboter statt der Bedrohungslagen gesehen.

⁶ Nicht nur Schriftsteller*innen wie Ian McEwan rekurren auf die Asimov’schen Gesetze (siehe Kapitel 1.4), auch im Bereich Maschinenethik/autonomes Fahren werden diese Robotergesetze als Grundlage für mögliche Entscheidungsprozesse genommen (siehe Kapitel 4 „KI-Anwendungen und ihre Narrative“).

narrativ durch die Beschreibung seines Äußeren, durch die personale Erzählsituation und seine Reflexionen als Mensch wahrgenommen, der in Opposition zur technisierten Repräsentation eines Roboters wie dem Chirurgen steht. Nach dieser Exposition, dem eigentlichen Ende der Erzählung, springt diese zurück zu Andrews Entstehung, als er noch wie ein Roboter aussah. Er wird von Familie Martin als Haushaltsroboter, wie zu der Zeit und Umgebung üblich, angeschafft. Doch Andrew, der Name wird aus seiner Seriennummer „NDR“ (137) versprachlicht und erinnert doch auch an „android“, entwickelt sich anders als andere Roboter seiner Art. Er zeigt kreativ-künstlerische Fähigkeiten und verspürt dabei sogar Freude: „It makes the circuits of my brain somehow flow more easily. I have heard you use the word ‘enjoy’ and the way you use it fits the way I fell. I enjoy doing them, Sir.” (138) Doch Andrew strebt nach mehr, nach Freiheit. Damit ist jedoch nicht primär Freiheit im Sinne der Unabhängigkeit von Familie Martin gemeint, sondern Selbstbestimmtheit und Selbstwirksamkeit. Auch Kleidung trägt Andrew schließlich, um menschlicher zu wirken. Doch das stößt auf Inakzeptanz bei Begegnungen auf der Straße. Als er angegriffen wird, steht er vor dem inhärenten Roboter-Dilemma:

„There was no way Andrew could stop them, if they ordered him not to resist in a forceful enough manner. Second Law of obedience took precedence over the Third Law of self-preservation. In any case, he could not defend himself without possibly hurting them and that would mean breaking the First Law.” (149)

Auch auf organischer Ebene wird Andrew den Menschen immer ähnlicher, denn sein Streben hat zum Ziel, als Mensch vor dem Weltparlament anerkannt zu werden, doch das letzte Hindernis dafür ist seine Unsterblichkeit, die aus seinem positronischen Gehirn resultiert: „Your brain is man-made, the human brain is not. Your brain is constructed, theirs developed. To any human being who is intent on keeping up the barrier between himself and a robot, those differences are a steel wall a mile high and a mile thick.“ (169) Daher begibt er sich zum Roboter-Chirurgen – die Erzählung endet in der Exposition –, dem er befiehlt, sein positronisches Gehirn gegen ein menschliches auszutauschen, um damit menschlich – sterblich – zu werden. Erleichtert nach der Operation aber konfrontiert mit Unverständnis stellt er fest: „If it brings me humanity, that will be worth it. If it doesn’t, it will bring an end to striving and that will be worth it, too.“ (171) Schließlich wird er als Mensch anerkannt – „Today we declare you a Bicentennial Man, Mr. Martin.“ (172) – und stirbt kurz darauf.

Im Gegensatz zu menschlichem Streben nach Unversehrtheit, Jugend und ewigem Leben erscheinen menschliche Attribute wie Altern und Sterblichkeit als Ideale des Seins. Die Frage nach (Un-)Sterblichkeit scheint hier das einzig gültige Distinktionsmerkmal zwischen Mensch und Maschine zu sein, denn Kreativität, Kleidung, Bildung, Selbstbestimmtheit werden toleriert. Doch Unsterblichkeit spiegelt allzu deutlich den nicht zu überwindenden menschlichen Makel wider: „Human beings can tolerate an immortal robot, for it doesn’t matter how long a machine lasts. They cannot tolerate an immortal human being, since their own mortality is endurable only so long as it is universal.“ (171) Damit hinterfragt Asimov gerade in der heutigen Zeit wieder kritisch diejenigen technologischen Entwicklungen, deren Ziel eben einzig die Verlängerung des Lebens oder die Perfektionierung des

Daseins ist. Denn wie Asimov mit dieser Erzählung betont: Nicht Perfektion gilt als Ideal, sondern gerade die Vergänglichkeit und der Makel implizieren Einzigartigkeit, die dem Menschsein innewohnt.

1.3 Jochen Beyse, *Fremd wie das Licht in den Träumen der Menschen* (2017)

Mit seinem Roman *Fremd wie das Licht in den Träumen der Menschen* (2017) rückt Jochen Beyse Roboter in eine ungewohnte narrative Perspektive. Rob, der Roboter-Protagonist, präsentiert sich sogleich als künstliches Wesen, das über ein Bewusstsein verfügt und daher den Menschen sehr ähnlich geworden ist: „Ich habe gelernt, ich zu sagen, der Gefühle wegen, um ihnen näherzukommen.“ (Beyse 2017, 10) Rob ist gleichzeitig der Erzähler eines Monologs, einer Geschichte, die selbst in verschiedene Erzählebenen verschachtelt ist – Robs Ich-Erzählung, ein Rückblick auf sein früheres Leben, in der Rob als Figur von einer extradiegetischen Erzählsituation aus erzählt, eine Filmaufzeichnung, eine Buchlektüre. Um zwischen Robs Erzählebenen in der ersten und dritten Person zu wechseln, werden Kommandos wie „Erzählmodus an“, „Erzählmodus aus“ oder „Erzählmodus wechseln“ eingefügt. Durch teils abrupte und mannigfaltige Wechsel wird es zunehmend schwer, sich in der Diegese zu orientieren, wie Kümmel (2017) betont:

„Zunächst scheint völlig klar, auf welcher Realitätsebene wir uns jeweils befinden („Unterbrechung der Lesung, Erzählmodus an“), doch je menschlicher der Automat zu denken beginnt, desto häufiger müssen auch wir die immer arbiträrer erscheinenden Grenzziehungen zwischen Realität und Fantasie anzweifeln.“ (Kümmel 2017)

Robs Erzählung ist atemlos, teils versatzstückartig montiert, befindet er sich doch auf einer Schutthalde, auf die er sich nach seiner Flucht aus seiner Besitzerfamilie gerettet hat, um dort in heldenhafter, epischer Konstellation auf den Sonnenaufgang zu warten, während seine Energiereserven dem Ende zu gehen: „Mit seinem manischen Monolog nähert sich der Automat einem tragischen Heldentod, oder aber der Rettung durchs (göttliche) Licht.“ (Kümmel 2017) Sein Sprechen und seine Situation lassen ihn dabei umso menschlicher erscheinen, denn es scheint, als ob er „in melancholischer Nostalgie menschliche Erschöpfung nachstellt.“ (Kümmel 2017)

Den Gegenentwurf zum vermenschlichten, erzählenden Roboter bilden die Menschen, mit denen Rob zusammengelebt hat, die aus seiner Sicht roboterhaft wirkten und ebenso von ihm portraitiert werden:

„Die Berdoncs waren dem Zodiak in dem Punkt ähnlich. Auch sie haben mich an Automaten erinnert. Nicht von ihren Bewegungen her, oder vielleicht doch, aber dann nur schwach, ihr Gang wirkte wie aufgezo-gen, bei dem männlichen Berdonc etwas auffälliger als bei ihr, dem weiblichen Berdonc. Dafür redeten sie wie Maschinen. Langsam, fast schleppend waren ihre Stimmen hinter jeder Silbe her, dann hinter der nächsten und über-nächsten, betonungslos, mit immer der gleichen Lautstärke und Geschwin-digkeit.“ (Beyse 2017, 44)

Berdoncs Habitus habe „etwas generell Stumpfes an sich“ (45), die Zodiak „war auf ihre Weise ebenfalls nur beschränkt anwesend“ (45) und die Kinder erweckten