

Zur Adressierung des Unbelebten – Grenzen von pragmatischer Konzeption

Netaya Lotze

1. Künstliche Intelligenz im öffentlichen Diskurs

Künstliche Intelligenzen (KIs) nehmen derzeit rasant Einzug in unsere Alltagswelt – sei es in Form von Netzwerk-Algorithmen, KI-gesteuerten Industrierobotern, selbst-fahrenden Autos oder Dialogsystemen. So wurde das Jahr 2019 von der Bundesregierung sogar zum KI-Jahr ausgerufen, um Entwicklungsprozesse in diesem Bereich an der Schnittstelle zwischen Informatik, Technikwissenschaften, Kognitionswissenschaften und Linguistik noch weiter zu beschleunigen. Sich ein Abbild zu schaffen ist dabei ein uralter Traum der Menschheit. Die Metapher des künstlichen Menschen, der als Helfer*in oder zur Unterhaltung realer Menschen konstruiert wird, zieht sich auch als wiederkehrendes Motiv durch Technik- und Literaturgeschichte. Das menschenähnliche Wesen übt eine Technologie-historisch nicht zu unterschätzende Faszination aus. Barrat (2013) thematisiert aber auch die Gefahren einer solchen Entwicklung in seinem Buch *Our Final Invention* und Sykora gibt zu bedenken: »Als Kippfigur zwischen Ding und Mensch wohnt den anthropomorphen Kreaturen das Potenzial einer Emanzipation von ihren Schöpfern inne« (Sykora 2000: 118).

Die Konstruktion eines sogenannten virtuellen Assistenzsystems, das anthropomorph gestaltet und durch natürliche Sprache intuitiv nutzbar ist, erscheint vor diesem soziokulturellen Hintergrund betrachtet trotz aller Bedenken naheliegend, denn die Möglichkeit der Entwicklung künstlicher Menschen beschäftigt die Menschheit von alters her, als an eine technische Entwicklung von menschenähnlichen Automaten, Robotern, Androiden oder virtuellen anthropomorphen Systemen noch nicht zu denken war (vgl. Dotzler, Gendolla & Schäfer 1992). Während sich die Literat*innen in den ersten Jahrhunderten auf künstliche Körper konzentrierten (z.B. die Puppen des Dädalos bei Aischylos oder Galatea in Ovids Pygmalion), rückte später die Diskussion um eine mögliche Seele des Maschinenmenschen in den Vordergrund (z.B. Mary Shelly's Frankenstein, vgl. Krämer 2008: 23). Daraus wurden philosophische Diskurse abgeleitet, in denen bereits Argumentationen angelegt waren, die man heute aus der Diskussion um die Möglichkeit einer

menschenähnlichen KI und deren ethischer Implikationen kennt (detaillierter dazu Kap. 2).

Heute verschwimmen die Grenzen zwischen natürlicher und künstlicher Intelligenz im öffentlichen Diskurs bereits, wenn Vertreter*innen der Transhumanismus-Bewegung im Silicon Valley (u.a. Kurzweil 2005) die »*technologische Singularität*« für 2045 proklamieren: der Tag, an dem KIs intelligenter sein werden als Menschen. In dieser Argumentationslinie wird das Moore'sche Gesetz zum exponentiellen Anstieg von Rechenleistung auf die KI-Forschung übertragen. Wenn also nur quantitativ immer mehr Rechenprozesse in einer »lernfähigen« KI abliefern, würde diese von selbst »aufwachen« und zu einem Bewusstsein gelangen, das nicht nur menschenähnlich sei, sondern dem Menschen überlegen. Diese nahe Zukunftsperspektive lässt sich im nächsten Schritt gleichermaßen zur Utopie einer ethischen Vernunft Herrschaft oder Dystopie einer Roboter-Tyrannis verdichten. Beide Szenarien aber sind angesichts des derzeitigen Stands der Technik unwahrscheinlich und mythisch überhöht. Schon der Analogieschluss vom quantitativen Anstieg der Rechenleistung in einem binären System auf Künstliche Intelligenz, deren qualitative Parameter noch völlig unklar sind, ist angreifbar (vgl. Lotze 2018, Moore & Nikolao 2017).

Die Problematik des öffentlichen Diskurses um KI liegt also in einem hochkomplexen Spannungsfeld aus historisch gewachsener Überhöhung des Topos KI, die die Haltung der Rezipient*innen prägt, und strategischer Werberhetorik, die dieses Potenzial ausbeutet in Form metaphorischer Anthropomorphisierung der Systeme: Personennamen (*Siri, Alexa, Alice, Eliza*), pronominale Referenz (Apples *Siri* = »sie«, Coca Colas *Hank* = »er«) sowie Zuschreibungen von Kognition und Agentivität (»*Siri versteht Dich*«¹). Die Anthropomorphisierung und mythische Überhöhung von KI im öffentlichen Diskurs führt zu gleichermaßen überhöhten Erwartungshaltungen der User*innen an heutige Dialogsysteme. Denn diese sind – je nach Systemarchitektur – leider immer noch einerseits entweder reine *Request-Response-Systeme*, die nicht über das isoliert stehende Adjazenzpaar von Frage und Antwort hinaus operieren oder einem starren Handlungsplan aus geskripteten Sequenzen folgen, oder sie agieren andererseits auf der Basis von *Big-Data-Analysis* und *Machine Learning* flexibler, was aber zu völlig sprunghaften, thematisch inkohärenten Dialogen führt (vgl. Lotze 2016, 2018). Die Illusion des virtuellen Gegenübers funktioniert also nur äußerst eingeschränkt.

Beispiel HCI (Cleverbot):

1 Apple-Hilfsportal »machilfe«, Artikel »Damit Siri Dich noch besser versteht«, www.maclife.de/ratgeber/ultimate-siri-ratgeber-10073699.html?page=2

1. Cleverbot: Well, my wife left me today.
2. User: Feels bad for you.
3. Cleverbot: Not really, I have an amazing boyfriend :).

Cleverbot nutzt einen Big-Data-Ansatz und recyclet die Gesprächsbeiträge seiner früheren User*innen. Dies führt zu einer inkonsistenten System-Persona, die je nach recyceltem User*innen-Turn das Alter, den Beruf oder die Genderrolle etc. wechselt.

Die Illusion der anthropomorphen, intuitiv nutzbaren Schnittstelle ist dennoch der technologie-historische Kern der *Assistenz-Metapher*, welche die *Desktop-Metapher* in naher Zukunft ablösen soll. Wir bedienen also das Interface nicht mehr, indem wir metaphorische Gegenstände wie »Dokumente«, »Ordner« oder »Marker« manipulieren, sondern indem wir mit dem oder der virtuellen Assistent*in natürlich-sprachlich interagieren. Genau darin besteht die Pointe einer vermeintlich intuitiv, mobil und barrierefrei steuerbaren Assistenz-Metapher seitens der *Usability*-Forschung. Diese Form der reibungslosen Kommunikation zwischen Mensch und Maschine ist zum Idealtypus der heutigen Human-Computer-Interaction (HCI) stilisiert worden, sodass im Diskurs um innovative Sprachsysteme und neue Herausforderungen an die Entwickler*innen häufig von *Star-Trek-Kommunikation* als erklärtem Ziel gesprochen wird. Auch hier verschwimmen die Grenzen zwischen Science und Fiction – was ganz offensichtlich ein Hauptcharakteristikum des KI-Diskurses darstellt. Der Begriff *Star-Trek-Kommunikation* geht auf eine Aussage von Dix et al. zurück, in der die eingeschränkten Möglichkeiten heutiger Systeme mit den menschenähnlichen der *Star-Trek*-Systeme verglichen werden. »*In contrast to most human conversation, dialog with computers is relatively structured and constrained. It is only on Star Trek that one can chat freely to the computer and expect a response [...]!*« (Dix et al. 2004: 545).

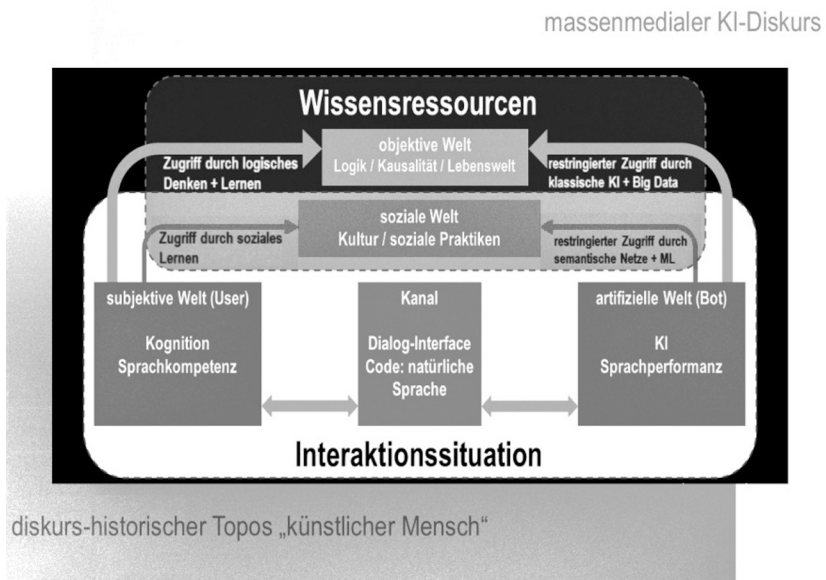
Doch wie reagieren User*innen abseits der Science Fiction, wenn sie mit dieser selten störungsfreien Illusion von Agentivität (im philosophischen Sinne) konfrontiert sind? Die Leittendenzen, die in Lotze (2016, 2018, 2019) nachgewiesen werden konnten, sind: Anpassung des Menschen an das System, Simplifizierung und Explizitheit im Sinne eines »Computer Talk« (Zoeppritz 1985) und eine steigende Akzeptanz gegenüber den vom System vorgegebenen Inhalten. D.h., User*innen passen sich in ihrem sprachlichen Dialogverhalten spontan, flexibel und interaktiv dem System ein Stück weit an und formulieren ihre Gesprächsbeiträge gegenüber der KI so, (wie sie denken,) dass diese sie am besten verarbeiten kann.

Werden also die Systeme menschlicher und die Menschen verhalten sich wie Maschinen? Ist Mensch-Maschine-Interaktion ein treffendes Beispiel für ein Verschwimmen der Grenzen zwischen Mensch und Objekt, wie es der Posthumanismus (vgl. Pennycook 2018) als differenzierteres, weniger extremes Dachkon-

zept zum Transhumanismus behauptet? Argumentiert man allein auf der Ebene der sprachlichen Performanz und radikal aus empirischer Perspektive, kann dieser Eindruck eventuell zunächst entstehen. Analysiert man HCI-Dialoge aber im Detail, wird man feststellen, dass menschliches Dialogverhalten nicht erklärbar ist, ohne die Ebene der Kognition miteinzubeziehen (vor allem in Form von *Partnermodellen* und Referenz auf *Common Ground*). Menschen haben – je nach User*innen-Typ und Technik-Expertise – eine bewusste Vorstellung von der Dialogfähigkeit des Systems, auf die sie sich einstellen und die sie bei Störungen flexibel variieren.

Die Interaktionssituation Mensch-Maschine wird also von einer Vielzahl von Faktoren beeinflusst: den technischen Grundprinzipien der KI in Form von Systemarchitektur und Dialog-Design, die sehr unterschiedlich sein können (Chatbot-Architektur, plan-basierte Systeme, Big-Data-Analysis und Machine Learning, vgl. Lotze 2018) sowie der anthropomorphen Gestaltung des Interfaces und auf der anderen Seite den Variablen der User*innen in Form von Alter, Technikexpertise und User*innen-Typ (Fischer 2006, 2010, 2016). Zusätzlich beeinflusst wird die Interaktionssituation durch den als Science Fiction überhöhten und durchkommerzialisierten Diskurs um KI, der a) innovative Design-Entscheidungen der Entwickler*innen prägt (vgl. Star-Treck-Kommunikation) und der b) Erwartungshaltungen und Rezeptionsverhalten der Nutzer*innen formt.

Abbildung 1: Erweitertes Interaktionsmodell der Mensch-Maschine-Interaktion



Bei der HCI handelt es sich um eine extrem heterogene neue Form der Dialogizität, die von unterschiedlichsten Variablen abhängig ist. An ihr werden die Unterschiede zwischen Mensch und Maschine gerade emergent, anstatt sie zu nivellieren.

Im Folgenden sollen grundsätzliche Restriktionen von KIs gegenüber Menschen definiert werden und die Relevanz der fraglichen Aspekte für Dialogfähigkeit am Beispiel der semantischen Kohärenz diskutiert werden. Warum gelingt es Menschen, semantisch kohärente Informationsstrukturen in Dialogen wechselseitig aufzubauen, während auch die avanciertesten Systeme noch an dieser Herausforderung scheitern?

Im zweiten Teil des Artikels soll am Beispiel von pronominaler Referenz, Anrede und sprachlicher Höflichkeit gezeigt werden, wie Menschen das virtuelle Gegenüber als »Kippfigur« im Sinne Sykoras (2000) im Dialog adressieren. Exemplarisch wird hier eine Studie zu medial-schriftlichen Chatbots aus mikro-diachroner Perspektive (avanciertere und weniger avancierte, ältere Systeme) vorgestellt. Abschließend werden die Ergebnisse der Studie vor dem Hintergrund der Transhumanismus-Debatte eingeordnet und diskutiert.

2. Minimalbedingungen für Dialogizität und Restriktionen von KI

Wodurch zeichnen sich Dialoge unter Menschen aus und woran scheitern auch modernste KIs noch immer? In aller Kürze und stark vereinfacht: Effizienz, Kohärenz, gegenseitiges Verstehen und Spontaneität im Dialog.

Effizienz Gerade in alltäglicher Face-to-Face-Kommunikation beobachten wir, dass Menschen in der Lage sind, äußerst effizient mit wenigen Worten komplexe Zusammenhänge zu kommunizieren. *Kommunikation* heißt hier, dass die Gesprächspartner*innen einander wirklich verstehen. Dies setzt voraus, dass beide Dialogpartner*innen jeweils über ein Bewusstsein verfügen, das die kommunizierten Inhalte denkt und sich selbst in Relation zu diesen Inhalten reflektiert im Sinne des »Self«-Begriffs nach Locke (1975).

Mensch-Mensch-Kommunikation (*Human-Human Communication*, HHC) ist deshalb so effizient, weil sie auf der Basis von geteiltem Wissen funktioniert. Menschen machen nicht jede Vorbedingung eines Gesprächsbeitrags und auch nicht jedes Detail in der Kommunikation untereinander explizit, sondern setzen den überwiegenden Teil der in einer Kommunikationssituation theoretisch verfügbaren Informationen als *Common Ground* (vgl. Stalnaker 2002) voraus. Sprachlich formuliert wird nur, was angesichts eines *Partnermodells* (vgl. Clark & Marshall 1981) des Gegenübers nicht als geteiltes Wissen angenommen und was außerdem als relevant für die Kommunikationssituation eingeschätzt wird

vor dem Hintergrund sozialen und kulturellen Wissens. Alles Weitere können die Kommunizierenden im Geiste ergänzen (*Etcetera-Prinzip*, Cicourel 1973). Die größte Herausforderung der Mensch-Mensch-Kommunikation ist es also das, was implizit bleibt, zu ergänzen. Menschen beherrschen dies virtuos auf der Basis ihres seit dem kindlichen Erstspracherwerb angesammelten *Weltwissens* im Sinne Habermas' (1993). Auf der Ebene von *Common Ground* und *Partnermodellen* sowie von *Weltwissen* fehlt es allen Dialogsystemen an den grundsätzlichen Wissensressourcen und der Fähigkeit, Vorannahmen (Präsuppositionen), die ihre User*innen nicht verbalisieren, zu inferieren. Also: was nicht verbalisiert wird, kann auch von keinem Parser erkannt werden. Leider machen die nicht verbalisierten Informationen das Gros der Informationsstruktur menschlicher Dialoge aus. Formuliert wird nur, was in der konkreten Dialogsituation als relevant bewertet wird. Diese Relevanzbewertung stellt dann im nächsten Schritt eine enorme Herausforderung für die Generierung von Gesprächsbeiträgen einer KI dar: Nach welchen Kriterien entscheiden Menschen, welcher Gesprächsbeitrag in einer konkreten Interaktionssituation gerade relevant ist? Selbst äußerst avancierte Systeme wie IBMs »WATSON« scheitern an dieser Aufgabe. Weltwissen ist nicht also nicht als Datenbank modellierbar (Big-Data-Ansatz), weil die Referenzrelationen fehlen, und auch nicht als semantisches Netzwerk künstlicher Assoziationsketten, weil diese unterspezifiziert und stereotyp sind.

Beispiel aus HCI:

1. System Max: Wo wohnst du? Wie wollen wir weitermachen?
2. User: in einem land hinter den sieben bergen
3. System Max: Du musst schon genauer erklären, was du meinst Hans.

(*Max-Korpus 1053-1080*)

Die Anspielung auf das Märchen »Schneewittchen« der Gebrüder Grimm kann vom Dialogsystem Max der Universität Bielefeld nicht geparkt werden, obwohl dieses in Europa als geteiltes kulturelles Wissen vorausgesetzt werden kann. Das hat zur Folge, dass auch die Ironie der User-Äußerung nicht als solche erkannt wird.

Kohärenz: Dialoge unter Menschen haben in der Regel einen »roten Faden«; d.h. die Gesprächsbeiträge bauen logisch kohärent aufeinander auf. Diese Turnübergreifende Referenz führt zu längeren Dialogsequenzen mit nachvollziehbarer logischer Progression. Menschliche Sprecher*innen beziehen sich aufeinander auf der inhaltlich-logischen Ebene. So entsteht ein engmaschiges Netz aus Referenzen: innerhalb des eigenen Gesprächsbeitrags, an den Vorgängerbeitrag anschließend, auf frühere Beiträge referierend und selbstverständlich auch auf *Common Ground*. Kohärenz manifestiert sich sprachlich in sog. Kohäsionsmarkern, die Aussagen in Relation zueinander setzen (*Topic-Comment*-Strukturen, Wiederaufnahmen, bestimmte Artikel, Anapher und Katapher usw.). Für dieses sprachimmanente

logische Kalkül sind KIs nicht sensibel. In der HCI finden Kohärenz und Kohäsion nur am konkreten Turn-Wechsel als ineinandergreifende Frage-Antwort-Sequenz statt.

Verstehen: Tatsächliche Kommunikation findet zwischen menschlichen Gesprächspartner*innen dann statt, wenn sie einander verstehen. Verstehen im kognitiven Sinne setzt aber voraus, dass da ein Bewusstsein ist, das sich selbst reflektiert, während es die Information denkt, sprachlich formuliert oder perzipiert. Searle hebt in seiner KI-Kritik am Beispiel des Gedankenexperiments »Chinese Room« vor allem die Referenzrelation zwischen dem menschlichen Bewusstsein und einer verstandenen Aussage als ausschlaggebend hervor und bezeichnet sie mit dem Terminus »*intentionality*«. Intentionalität versteht Searle (2006: 171ff.) also als die allgemeine Gerichtetheit von Aussagen. Dabei geht es primär nicht um eine Absicht, sondern um eine semantische Relation. Aussagen müssen auf etwas gerichtet sein. Das kann ein empirisches Objekt oder ein Handlungsziel gleichermaßen sein. Das menschliche Bewusstsein ist in der Lage, große Mengen intentionaler Bezüge zu koordinieren. Die Zielorientiertheit (»*goal directedness*«, Cassell et al. 2000) von KIs wie planbasierten Agenten-Systemen ist ohne Autonomie und ein Bewusstsein nicht als Intentionalität zu bezeichnen.

Spontaneität und Autonomie: Menschen verfolgen intentional Dialogziele, können diese aber spontan und bewusst ändern. Ein deterministisches System kann das nicht, denn agentive Handlungsplanung (im philosophischen Sinne) setzt Autonomie voraus. Agenten-Systeme, die planbasiert unter Berücksichtigung eines einprogrammierten Dialogziels »agieren«, können unter diesen Prämissen (Plan und Ziel) zwar Entscheidungen treffen, sind dabei aber nicht autonom, sondern allenfalls autopoietisch (vgl. Schatter 2011: 15).

Table 1: Minimalbedingungen von Dialogizität (vgl. Lotze 2016: 73)

Mensch	System
Autonomie	Autopoiesis
Intentionalität	Zielorientiertheit (Goal-Directedness)
Spontaneität	Determination
Identität der Erfahrung	Isolierte Verarbeitungsprozesse
Weltwissen	Begrenzte Datenbanken

Das System ist deterministisch, die Dialoge sind (mit einem unterschiedlichen Grad an Flexibilität) vorformuliert. Der Autonomiebegriff ist in der KI-Forschung folglich anders belegt als in der klassischen Philosophie, wenn er auf Agenten-

Systeme angewandt wird. Ein System gibt immer nur aus, was seine Entwickler*innen der Möglichkeit nach implementiert haben. Zwar sind moderne Systeme so komplex und gehen mit so vielen Wahrscheinlichkeiten um, dass es selbst für die Konstrukteur*innen manchmal nicht trivial ist, Systemreaktionen nachzuvollziehen. Es entsteht der Eindruck von Spontaneität, der sich in vermeintlich freien Entscheidungen manifestiert. Doch das scheinbar intelligente Verhalten ist nachvollziehbares Produkt einer zufälligen Kausalkette oder von probabilistischen Entscheidungen, auf die das System festgelegt ist. Kant begreift die autonome Person auf Grund ihrer Entscheidungsfreiheit als ethisches Subjekt. »Person ist dasjenige Subjekt, dessen Handlung einer Zurechnung fähig ist« (Kant 1999, VI 223). Der Mensch soll aus Einsicht in die Pflicht gegenüber der Menschheit als Totalität so handeln, dass er nicht dem Gesetz der Vernunft zuwider handelt (vgl. Kategorischer Imperativ). Diesem Gedanken liegt zu Grunde, dass Kant den Menschen als autonom begreift. Diese Autonomie besteht darin, dass ein Mensch über einen freien Willen verfügt und sich frei Ziele setzen kann. Ein Bot hat diese Möglichkeit nicht – eine einfache Erkenntnis, die jeder ethischen Diskussion um KI-Rechte vorausgesetzt sein muss.

Dass sich User*innen wie oben erwähnt in ihrem sprachlichen Dialogverhalten spontan, flexibel und interaktiv dem System ein Stück weit anpassen und ihre Gesprächsbeiträge gegenüber der KI so formulieren, (wie sie denken,) dass diese sie am besten verarbeiten kann, spricht also auf kognitiver Ebene betrachtet dafür, dass Menschen gerade im Kern grundsätzlich anders kommunizieren als KIs – nämlich autonom und bewusst.

3. Sprachliche User*innen-Strategien

Die internationale HCI-Forschung konzentriert sich im Bereich der Linguistik seit den 1990er Jahren vor allem auf die Frage, ob User*innen Dialogsystemen gegenüber eine simplifizierte Befehlssprache gebrauchen (»Computer-Talk« (CT), Zoepritz 1985) oder ein »natürlicheres« Verhalten aus der Mensch-Mensch-Kommunikation auf die HCI übertragen, wie z.B. sprachliche Höflichkeit oder Strategien der sozialen Aushandlung und Anpassung (Richards & Underwood 1984; Morel 1989; Amalberti et al. 1993; Johnstone et al. 1994; Hitzenberger & Womser-Hacker 1995; Porzel & Baudis 2004; Fischer 2006, 2010, 2016). Wie User*innen sprachlich auf **die Illusion eines agentiven Gesprächspartners (im philosophischen Sinne)** reagieren, wenn sie mit Störungen der logischen Kohärenz im Dialog konfrontiert sind, ist in der linguistischen KI-Forschung zwar hinlänglich untersucht, die Studien kommen aber zu unterschiedlichen Ergebnissen, je nachdem, ob strukturelle oder funktionale Aspekte der HCI in den Fokus genommen werden (vgl. Krause, Hitzenberger & Womser-Hacker 1998; Fischer 2006, 2010,

2016; Lotze 2016, 2018, 2019). Auch die Interpretation der User*innen-Strategien als vorbewusste Übertragung von Verhaltensweisen aus der HHC oder als bewusst simplifizierter »Computer-Talk« (CT, Zoeppritz 1985) variieren.

3.1 Computer-Talk (CT)

Treibende Kraft für die erste Welle der HCI-Forschung war ein ganz grundsätzliches informationswissenschaftliches Anliegen: Brauchen wir überhaupt intuitive, natürlich-sprachliche Interfaces? In älteren Studien werden strukturelle Marker für »Computer-Talk« (z. B. Imperative oder Infinitive) oder Übertragungen aus der Kommunikation unter Menschen (z. B. Höflichkeitsmarker wie *danke*, *bitte*, *Entschuldigung*) analysiert, der jüngere HCI-Diskurs fokussiert dagegen seit Fischer 2006 die konzeptionell-funktionale Ebene (HCI als »Simplified Register«).

3.2 Grounding und interaktives Alignment

Die Mensch-Mensch-Kommunikation kann mit Grounding definiert werden als »reaching the mutual belief that B has understood A well enough for current purposes« (Clark & Schäfer 1989: 265). Jönsson (1996) stellt fest, dass in der HCI von menschlichen Sprecher*innen weniger Bezug auf Kontextinformationen genommen, aber dennoch früher Erörtertes als gesichert wahrgenommen werde. Die erste, die die Relevanz von »Grounding« – also der Etablierung und Sicherung von geteiltem Wissen – in den Fokus ihrer Forschung stellt und als problematische Größe in der HCI benennt, ist Fischer (2006). Sie kann in ihren »*Verbomobil-Daten*« (2006) zeigen, dass Menschen sich dem System gegenüber auf vermeintlich geteiltes Wissen berufen wollen, über welches das System aber nicht verfügt. Dies führt zu Störungen im Dialog und zu Frustrationen seitens der User*innen.

3.3 Interaktives Alignment

Aus psycho-linguistischer Perspektive ist interessant, inwiefern Menschen zu virtuellen Dialogpartner*innen *alignen* – sich also lexikalisch und syntaktisch dem System durch (vorbewusste) Konstruktionsübernahmen anpassen (vgl. Pickering & Garrod 2004). Denn vorbewusstes Alignment könnte als Indikator für eine Übertragung von Konzepten aus der Mensch-Mensch-Kommunikation auf die Mensch-Maschine-Interaktion interpretiert werden sowie (unter bestimmten Bedingungen) als sprachlicher Indikator für die Attribuierung sozialer Eigenschaften an das System. User*innen-Alignment ist aber immer nur als *reaktiv* zu bezeichnen, weil tatsächliche Interaktivität mit dem System nicht erzielt werden kann. Bei Konstruktionsübernahmen des Systems aus dem vorangehenden Nutzer*innen-Gesprächsbeitrag handelt es sich immer um Artefakte

(*Pseudo-Alignment*, Lotze 2016), wenn ein System geparte Keywords aus dem vorangegangenen User*innen-Turn zur Erstellung des eigenen Folge-Turns nutzt. Dies sieht aus wie eine Konstruktionsübernahme im Sinne von interaktivem Alignment, ist aber ein rein deterministischer Prozess, der auf der Ebene der sprachlichen Performanz die Illusion von Alignment schafft. Alignment der Nutzer*innen bleibt also **einseitig** und erfolgt entweder **vorbewusst** im Sinne einer Übertragung von Verhaltensweisen aus der HHC oder **strategisch**, um sich dem Vokabular des Systems anzupassen und so Störungen zu vermeiden. Dies erschwert die linguistische Interpretation von Konstruktionsübernahmen in der HCI hinsichtlich der Motivation der User*innen als Übertragungsphänomen aus der HHC oder als Simplifizierungs-Strategie und damit CT. Computer-Talk als bewusste Simplifizierungs-Strategie und Übertragungsphänomene aus der HHC wie Grounding und Alignment schließen sich bei User*innen moderner Systeme jedoch keineswegs wechselseitig aus (Lotze 2016, 2018), da heutige KIs anwendungsspezifisch einerseits über längere Sequenzen störungsfreie, intuitive Nutzbarkeit zulassen und andererseits Bedienbarkeit durch Eingabe isolierter Keywords fördern (vor allem auffällig bei Social Bots, also solchen Bots, die auf Social-Media-Plattformen (Facebook-Messenger, Twitter etc.) eingesetzt werden; Lotze & Ohrndorf in Vorb.). Hier entsteht eine neue Form der Dialogizität zwischen CT und HHC. Je nach Typus (»Player*in« vs. »Non-Player*in« nach Fischer 2006) lassen User*innen sich mehr oder weniger auf die Illusion eines agentiven Gegenübers und damit auch auf eine intuitive, natürlich-sprachliche Interaktion ein.

3.4 Die Chatbot-Studie

Die korpuslinguistische sowie konversationsanalytische Studie thematisiert die Mensch–Maschine-Interaktion mit Chatbots. Dabei wird der Fokus auf die Evaluation des sprachlichen User*innen-Verhaltens gelegt gegenüber vier deutschsprachigen, medial-schriftlichen Dialogsystemen (*Twipsy*, *Karlobot*, *Elbot*, *Max*) auf unterschiedlichem Stand der technischen Entwicklung seit dem Jahr 2000 bis 2012 und ist damit (mikro-)diachron in zeitlichen Schnitten angelegt: *Twipsy*-Korpus = 2000, *Karlobot*-Korpus = 2003, *Max*-Korpus = 2005, *Elbot*-Korpus = 2008. Das für diesen Artikel weiterführend einbezogene Social-Bot-Korpus wurde im Jahr 2012 erhoben. Also nicht nur die älteren Systeme sind weniger avanciert, sondern auch ihre User*innen haben eine geringere Technikexpertise. Es geht nicht um die Weiterentwicklung bzw. Verbesserung der Systeme, sondern um die Entwicklung der Interaktionsform HCI seitens der Systeme und ihrer User*innen gleichermaßen. Ein Korpus zur Chat-Kommunikation unter Menschen stand als Vergleichskorpus zur Verfügung (Dortmunder Chat-Korpus Storrer/Beißwenger, Teilkorpus »Bibliotheksinformation«: Bib). Untersucht wurden 183 Dialoge, also rund 60.000

Wortformen nach qualitativen und quantitativen Kriterien. Es handelt sich um eine Erstbeschreibung dieser *neuen Form der Dialogizität*. Chatbots und Social Bots sind im Web in der virtuellen Kund*innen-Beratung bereits weit verbreitet und ihr Dialog-Design dient als Vorbild für moderne Assistenzsysteme. Da in Zukunft künstliche Intelligenzen eine immer größere Rolle spielen werden, legt diese Auseinandersetzung im Hinblick auf interaktives Alignment und Computer-Talk einen Grundstein zur linguistischen Erforschung von Dialogsystemen. Exemplarisch sollen hier Aspekte der Adressierung von Bots sowie Formen der Anrede (Duzen vs. Siezen) und allgemeine Aspekte der sprachlichen Höflichkeit in Abgrenzung von Computer-Talk diskutiert werden.

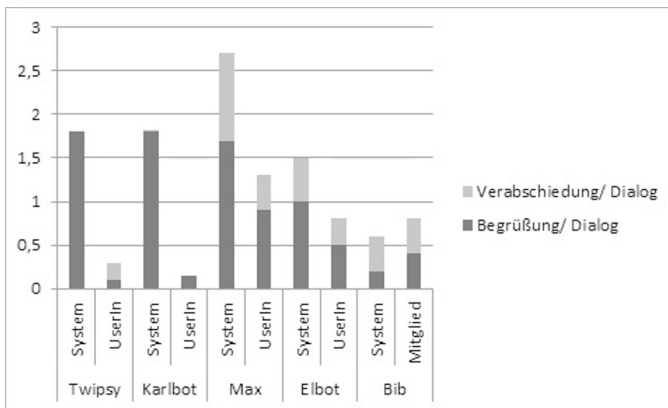
3.5 Sprachliche Höflichkeit in der HCI

Unter sprachlicher Höflichkeit wird ein komplexes kulturspezifisches Konzept verstanden, das unterschiedliche Dimensionen aufweist und das je nach Sprecher*innen-Gemeinschaft an unterschiedlichen sprachlichen Merkmalen festgemacht werden kann (Brown & Levinson 1987). Als Indikatoren für einen höflichen Stil in der HCI wurden in der vorliegenden Studie Begrüßungen und Verabschiedungen sowie Dank, Bitten und Entschuldigungen in jeglicher sprachlichen Realisierung qualitativ untersucht und für die quantitative Untersuchung nach SWBD-DAMSL-HCI (Lotze 2016) manuell annotiert (Adaptation des *Switchboard Dialog Markup in Several Layers* (Jurafsky, Schriberg & Biasca 1997) für die HCI). Clifford Nass und seine verschiedenen Koautor*innen gehen davon aus, dass User*innen Höflichkeitskonzepte aus der HHC auf die HCI übertragen und dass die entsprechenden sprachlichen Marker gleich häufig auftraten (vgl. Nass & Reeves 1996, Nass & Moon, 2000; Nass & Brave, 2005). Auch Fischer (2006) beobachtet höfliches Verhalten gegenüber dem System. Krause (1992) und Johnstone (1994) argumentieren gegen einen höflichen Stil in der HCI, da die User*innen sich darüber im Klaren seien, dass sie dem System gegenüber keine höflichen Formulierungen gebrauchen müssten. Im Gegenteil zeichne sich CT gerade durch die Abwesenheit von Höflichkeitsmarkern aus. Hinzu kämen spezielle Indikatoren für unhöfliches Verhalten, wie Beschimpfungen (*Flaming*), knappe Befehle und Gesprächsabbrüche, die typisch für die HCI seien und als strukturelle Indikatoren für einen CT gewertet werden könnten. Im Rahmen der vorliegenden Studie wurden diese Verhaltensweisen ebenfalls erhoben. Als funktionales Konzept ist Höflichkeit empirisch schwer greifbar, daher wurden sprachliche Indikatoren für die o.g. Aspekte per Hand annotiert und in ihrem Kontext interpretiert. So konnten im Bereich der sprachlichen Höflichkeit große Unterschiede zwischen User*in und System in den Dialogbeiträgen nachgewiesen werden. Damit weicht die chat-basierte HCI in Bezug auf Höflichkeit stark von der HHC ab (vgl. auch Salem/Ziadee/Sakr 2014, Burton & Gaskin 2019).

Bereits bei der manuellen Annotation der Korpora zeigte sich, dass User*innen (vor allem bei den neueren Systemen) zwar Grußfloskeln erwidern, komplexere Höflichkeitskonzepte aber in den Untersuchungskorpora keine Rolle spielen. Die Auswertung musste also aufgrund der defizitären Datenlage trotz qualitativer Analyse auf die eine einfache Auszählung von parallelen Grußfloskeln und einigen weiteren direkten Höflichkeitsindikatoren wie *bitte* und *danke* (s.u.) beschränkt werden.

Begrüßung und Verabschiedung: Die unterschiedlichen Anwendungsszenarios, in denen die untersuchten Bots eingesetzt werden, haben großen Einfluss auf das Grußverhalten der User*innen. Wechselseitige Grüße und Verabschiedungen sollten in der HCI das erklärte Ziel eines gelungenen Dialog-Designs sein, zumal solche stark routinisierten Sequenzen leicht zu implementieren sind.

Abbildung 2: Begrüßungen und Verabschiedungen (absolute Häufigkeiten pro Dialog (mikro-diachron in Schnitten))



Auch wenn im untersuchten HHC-Szenario an der chat-basierten Bibliotheksinformation in weniger als der Hälfte der Chats eine Begrüßung oder Verabschiedung formuliert wird, ist es für die HCI sicher einfacher, alignte Grußsequenzen zu implementieren als einen Algorithmus zur Erkennung der Angemessenheit eines Grußes innerhalb eines bestimmten Kontexts zu definieren.

Die Abbildung zeigt deutlich, dass die User*innen der neueren Systeme häufiger grüßen als die User*innen von Twipsy und Karlbot, die sich z.T. grundsätzlich nicht verabschieden. Dieses Verhalten kann als CT interpretiert werden, da das System gewissermaßen wie ein Such-Tool auf einer Website verwendet wird, das man nicht begrüßt und den Vorgang abbricht, wenn man fündig geworden ist

oder frustriert aufgibt. Durch ihr eigenes Grußverhalten fördern die alten Systeme wechselseitige Grußsequenzen auch nicht, da sie im Schnitt zwei Begrüßungen pro Dialog ausgeben, aber keine Verabschiedung. Dieses Design ist unnatürlich und wird von den User*innen auch so wahrgenommen. Auch Max gibt in der Regel zwei Begrüßungen pro Dialog aus, weil er durch die erste überhaupt erst die Aufmerksamkeit seiner User*innen erregt und durch die zweite diese dann persönlich begrüßt.

Beispiel:

- (1) Max: Hallo! Komm und spiel mit mir! [Default-Turn, um Aufmerksamkeit zu erregen, mit anderen in randomisierter Endlosschleife]
 - (2) UserIn: Hallo Max
 - (3) Max: Guten Tag und willkommen in der Ausstellung [...]
- (Max-Korpus 955-975)

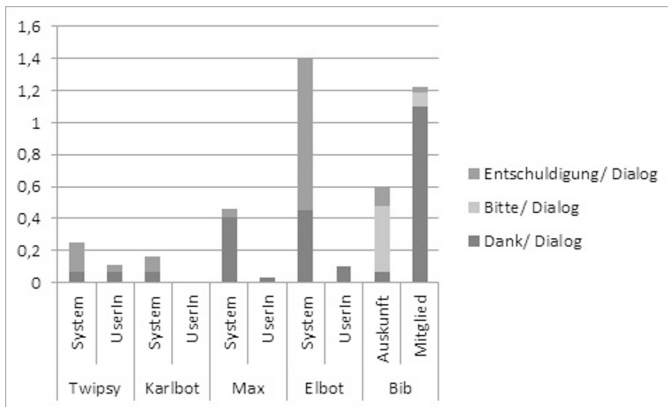
Das System verabschiedet sich aber zu 100 %. Die User*innen reagieren mit 100 % Begrüßungen als Reaktion auf den ersten Gruß, aber nur 50 % der Nutzer*innen verabschieden sich von Max. Eine qualitative Durchsicht der Dialoge legt nahe, dass User*innen sich dann verabschieden, wenn sie mit der Interaktion zufrieden waren. »User, die durch das Bot-Verhalten frustriert sind, werden höchstwahrscheinlich keine Höflichkeitsformen verwenden, sondern die Interaktion einfach abbrechen. Daher kann man sagen, dass eine Verabschiedung durch den User [...] ein Indiz für eine erfolgreich abgeschlossene Interaktion ist« (Schwarz 2013: 81).

Elbot begrüßt zu 100 %, verabschiedet sich aber auch nur in 50 % der Dialoge. Seine User*innen grüßen seltener als die von Max (50 %) und verabschieden sich nur in einem Drittel der Dialoge. Dieses Verhalten ist durch unterschiedlichen Bot-Personae bedingt. Eine deutliche Verbesserung des Grußverhaltens im Sinne einer Annäherung an HHC-Skripte kann aber bei den neueren Systemen verzeichnet werden. Hier kann ein Rückgang von CT zugunsten eines interaktiveren, dynamischeren und natürlicheren Verhaltens seitens der User*innen beobachtet werden. In medial-mündlichen Assistenzsystemen wie Siri oder Alexa wird die Adressierung seitens der User*innen als Aktivierungswort operationalisiert (»(Hey,) Siri!«) und damit technologisch erzwungen. Was in modernen Systemen also augenscheinlich dem Grußverhalten der HHC näher kommt, ist doch eigentlich als CT zu interpretieren, weil »(Hey,) Siri!« funktional eben nicht als Begrüßung fungiert, sondern das System startet (wie ein »Start«-Knopf oder die »Enter«-Taste).

Danke, Bitte und Entschuldigung: Während ein natürlicheres User*innen-Verhalten durch ein konsequentes Dialog-Design bei den Gesprächsrahmensequenzen offensichtlich leicht getriggert werden kann und so den neueren Systemen gegenüber Grußfloskeln formuliert werden, stellt es sich in den freieren Sequenzen in der Gesprächsmitte als schwieriger dar, durch geschickte Bot-Beiträge einen höflichen

Ton zu etablieren. Die Verteilung der Dankesbekundungen, Bitten und Entschuldigungen vermittelt einen aufschlussreichen ersten Eindruck.

Abbildung 3: Höflichkeit Gesprächsmitte (absolute Häufigkeiten der Wortformen pro Dialog (mikro-diachron in Schnitten))



In der HHC spielen gerade am virtuellen Help-Desk Bitten und Danken eine große Rolle und die Zahlen spiegeln eindeutig wider, dass vor allem die Auskunft Fragen und Aufforderungen als höfliche Bitten formuliert, während die Bibliotheksmitglieder, deren Anliegen bearbeitet werden, nicht vergessen, sich höflich zu bedanken. Entschuldigungen machen einen kleineren Anteil aus und finden sich in der HHC vor allem bei der Auskunft zum Einleiten von Reparaturen. In allen User*innen-Korpora dagegen spielen die Untersuchungsparameter eine untergeordnete Rolle. Für Bitten wurden gar keine Belege in den HCI-Korpora gefunden. Während die älteren Systeme auch wenig Dank und Entschuldigungen verwenden, enthalten bei den neueren Systemen 40 % der Dialoge Dankesbekundungen jeglicher Form. Bei Elbot wurde das Prinzip der Entschuldigung mit Rücksicht auf seine Persona, die vom Psychologen Fred Roberts zum Zweck der Unterhaltung durchgängig als Neurotiker gestaltet wurde, bewusst überstrapaziert. 95 % der Dialoge enthalten eine Entschuldigung.

Das Verhalten der Bots – sowohl der älteren als auch der neueren – ist zwar unterschiedlich, entspricht aber nie dem Vorbild der HHC. Vielleicht liegt darin der Grund, warum die User*innen auch bei den neueren Systemen kaum Höflichkeitsmarker in der Gesprächsmitte verwenden. Die Ergebnisse zeigen keine Entwicklung und legen eine Interpretation als CT nahe. Durch ein innovatives Dialog-Design könnte dem entgegengewirkt werden, denn die Gesprächsforschung hat

klare Muster für Sequenzen von Bitten und Danken definiert, die man in einfache Regeln für einen Dialog-Manager² übersetzen könnte.

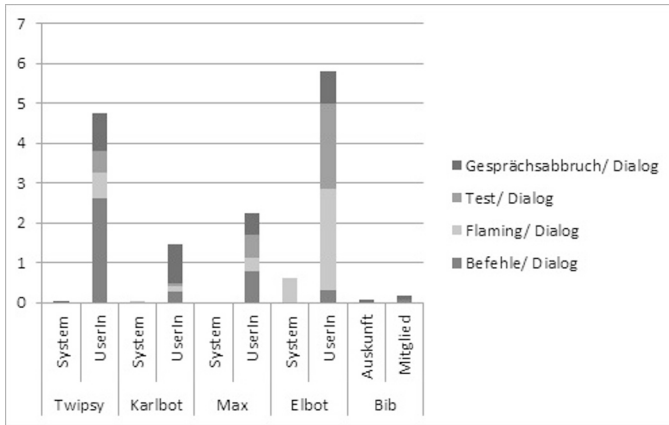
Indikatoren für unhöfliches Verhalten: Während Fischer (2006) auf der Grundlage ihrer Daten keine direkte Unhöflichkeit der User*innen gegenüber den Systemen nachweisen kann, finden sich in den Untersuchungskorpora der vorliegenden Studie zahlreiche Belege für grob unhöfliches Verhalten. In den untersuchten HCI-Dialogen sind Verhaltensweisen seitens der User*innen häufig, die in der HHC soziale Sanktionen nach sich ziehen würden. Am extremsten sind hier wohl *vulgäre Beschimpfungen (Flaming)*. Darüber hinaus können weitere Verhaltensweisen beobachtet werden, die in der HHC als grob unhöflich eingestuft würden. So werden z.B. *einfache Befehle* eingegeben (vgl. metaphorischer Sprachgebrauch, Krause 1992), *Gespräche ohne Verabschiedung* abgebrochen und die Systemfunktionen mit unterschiedlichen Mitteln (auch ohne vulgäre Sprache) getestet. Krause, Womser-Hacker und Hitzenberger (1992) sehen in diesen Verhaltensweisen Indikatoren für CT, da das System nicht als Gesprächspartner*in behandelt wird, sondern seine Funktionen als Maschine im Fokus der Interaktion stehen. User*innen, die solche Interaktionsstrategien verfolgen, sind sich darüber im Klaren, dass sie ein System bedienen. Dazu können sie mittels möglichst direkter Befehle navigieren und die Leistungsfähigkeit der Anwendung während des Dialogs testen, den sie einfach abbrechen, wenn sie fertig sind oder keine Lust mehr haben.

Die Verteilung zeigt, dass die fraglichen Verhaltensweisen in der HHC innerhalb eines Help-Desk-Szenarios nicht üblich sind. Auch die Systeme verwenden die o.g. Interaktionsstrategien nicht. Bei den User*innen spielen sie aber sowohl gegenüber den alten als auch gegenüber den neuen Systemen eine Rolle und stellen vielleicht die charakteristischsten Ausprägungen eines CTs dar. Wie häufig sich die User*innen dieser Strategien bedienen, scheint abhängig zu sein von Faktoren des Systems, denn die Ergebnisse sind sehr unterschiedlich. Interessant ist, dass einfache Befehle und Gesprächsabbrüche bei den älteren Systemen häufiger vorkommen, während bei den neueren Systemen Beschimpfungen und Tests dominieren. Dabei kann eine Koinzidenz von Beschimpfungen und Tests beobachtet werden. Diese Punkte sprechen für eine Veränderung der Wahrnehmung der Systeme durch die User*innen, je avancierter das System ist.

Flaming und System-Tests: Die bei den neueren Systemen häufigen vulgären Beschimpfungen können in der chat-basierten HCI als *Flaming* analysiert werden.

2 Der sog. Dialog-Manager ist eine Systemkomponente, die das Herzstück aller Dialogsysteme darstellt. Hier werden Nutzer*innen-Turns und System-Turns regelbasiert hinsichtlich der relevanten vom Parser erkannten Muster (Patterns) und der für die Antwortgenerierung verwendeten Schablonen (Templates) gematcht.

Abbildung 4: Indikatoren für Unhöflichkeit (absolute Häufigkeiten der Wortformen pro Dialog (mikro-diachron in Schnitten))



Flaming und Funktionstests sind besonders häufig bei Elbot zu finden. Dieser Bot schürt ein solches UserInnen-Verhalten geradezu, indem er seine Gesprächspartner*innen teilweise selbst beschimpft und allgemein provoziert. Er ist als Spielzeug für seine User*innen konzipiert und das Spiel der HCI liegt vor allem auch darin, die Funktionen des Bots zu erkunden und zu verstehen. Funktionstests und Tests durch Flaming sind probate Mittel, um Reaktionen zu evozieren, die die Grenzen des Dialog-Designs aufzeigen und so einen ersten Eindruck von den Funktionsweisen des Systems vermitteln. Sowohl bei Elbot als auch bei Max wurde an die Implementierung entsprechender Retourekutschen gedacht.

Beispiel:

- (1) UserIn: hi, du arsch
 - (2) Max: sehr kreativ, wirklich!
- (Max-Korpus 318-327)

Form der Anrede: Obwohl drei der vier Systeme ihre User*innen siezen, werden alle Systeme umgekehrt von User*innen durchgängig geduzt – egal, ob das System siezt oder duzt. Allein im Karlbot-Korpus konnten einige Belege für Siezen seitens der User*innen gefunden werden. Der Kontext *Kundengespräch mit dem virtuellen CEO Karlbot* kann als die Gesprächssituation mit der größten sozialen Distanz zwischen den Interagierenden interpretiert werden, was das vereinzelte Siezen erklärt.

Befehle: Häufig sind gerade bei den älteren Bots Aktionen, die an Navigation auf einer Website erinnern; z.B. mit einzelnen Nomina einen Menüpunkt auswählen (»Angebote«) und mit einzelnen Präpositionen navigieren (»weiter«, »zurück«). Bei Max finden sich vor allem Verben im Imperativ (»Tanz!«, »Sing!«). Dies ist einerseits der Interface-Animation zu verdanken, da der Avatar auf solche Befehle tatsächlich reagiert, andererseits manifestiert sich hier aber auch eine Entwicklungsstufe in der User*innen-Sprache: Der Bot wird nicht wie eine Navigation gesteuert, sondern ihm wird mittels einer konkreten Verbform befohlen. Der HHC ähnlich wäre erst eine Dialogsituation, in der der Bot höflich gebeten wird, etwas zu tun. Interessant ist, dass sowohl isolierte Keywords als auch einfache Befehle im Imperativ gegenüber modernen Social Bots (Keywords) und Assistenzsystemen (Imperative) derzeit ein Revival erleben, weil sich hier eine neue Kompetenz jüngerer User*innen entwickelt, die die robustesten Methoden zur Bedienung der Systeme erkannt haben.

Abbildung 5: Social Bot im Facebook-Messenger



4. Ausblick: Computer-Talk oder Attribuierung sozialer Eigenschaften an das System?

Dieser Auszug aus Ergebnissen der Teilstudie »Computer-Talk« aus der Gesamtstudie zu Chatbots (Lotze 2016) zeigt deutlich, dass User*innen teilweise Strategien aus der HHC auf die HCI übertragen (Duzen, Siezen, Adressierung durch Bot-Namen, Begrüßung, Verabschiedung, Danken, Bitten und Entschuldigen), dass die sprachlichen Marker für das funktionale Konzept »sprachliche Höflichkeit« aber

quantitativ bedeutend seltener sind als im Vergleichskorpus zur der HHC. Hinzu kommen eindeutige Parameter eines CTs wie einfache Befehle (Imperative), isolierte Keywords (Nomina, Präpositionen), Flaming, System-Test und Gesprächsabbrüche. Zusätzlich wird anhand der Korpus-Daten klar, dass die vom Bot vorgegebene Gesprächsatmosphäre (Anwendungskontext, soziale Nähe, Humor etc.) sprachliche Höflichkeit seitens der User*innen beeinflusst. User*innen-Sprache ist also im Spannungsfeld zwischen HHC und CT zu verorten und abhängig von einer Vielzahl von Variablen.

Tabelle 2: Dimensionen der dialogimmanenten Heterogenität der HCI (Lotze 2016)

Systemarchitektur	User*innen-Führung: geführt – frei – hybrid
Dialog-Design	Monolog – Dialog Schriftlichkeit – Mündlichkeit Distanz – Nähe Unterschiedliche Dialogphasen (Begrüßung – Mitte – Verabschiedung) Umgang mit Störungen: Inkohärenzen, Quasi-Kohärenzen, Default-Antworten oder Rückfragen
User*in	User*innen-Typ bewusst strategisch – vorbewusst bzw. routiniert CT – HHC

Bei HCI handelt es sich also um eine maximal heterogene Interaktionsform, die sich in naher Zukunft entweder in ihren Modalitäten und stilistischen Merkmalen ausdifferenzieren und festigen wird (Register-Hypothese) oder sich mit technologischem Fortschritt weiter der HHC annähern wird (Star-Trek-Kommunikation), wobei die neuralgischen Punkte der KI-Forschung (Bewusstsein, Intentionalität, Autonomie, Flexibilität) – vorerst und vor allem für binäre Systeme – eine natürliche Grenze dieser Entwicklung darstellen.

Literatur

- Amalberti, René; Carbonell, Noelle; Falzon, Pierre (1993): User representations of computer systems inhuman–computer speech interaction. In: *International Journal of Man–Machine Studies* (38), S. 547–566.
- Barrat, James (2013): *Our Final Invention: Artificial Intelligence and the End of the Human Era*. New York: Thomas Dunne Books.
- Brown, Penelope; Levinson, Stephan C. (1987): *Politeness. Some Universals in Language Usage*. 2. Aufl. Cambridge: Cambridge University Press.

- Cassell, Justine; Sullivan, Joseph; Prevost, Scott; Churchill, Elizabeth F. (Hg.) (2000): *Embodied conversational agents*. Cambridge: MIT Press.
- Cicourel, Aaron Victor (1973): Basisregeln und normative Regeln im Prozess des Aushandelns von Status und Rolle. In: Arbeitsgruppe Bielefelder Soziologen (Hg.): *Alltagswissen, Interaktion und gesellschaftliche Wirklichkeit*. Frankfurt a.M.: Fischer Taschenbuch Verlag (1), S. 147-188.
- Clark, Herbert H.; Marshall, Catherine R. (1981): Definite reference and mutual knowledge. In: Aravind K. Joshi, Ivan A. Sag und Webber Bonnie L (Hg.): *Elements of discourse understanding*. Cambridge: Cambridge University Press.
- Clark, Herbert H.; Schäfer, Edward F. (1989): Contributing to discourse. In: *Cognitive Science* (13), S. 259-294.
- Dix, Alan; Finlay, Janet; Abowd, Gregory; Beale, Russell (2004): *Human-Computer Interaction*. <http://myweb.tiscali.co.uk/lindasnape/CO2401/HCI/Dix,Finlay,Abowd,Beale.pdf>.
- Dotzler, Bernhard; Gendolla, Peter; Schäfer, Jörgen (1992): *MaschinenMensch. Eine Bibliographie*. Frankfurt a.M.: Lange.
- Fischer, Kerstin (2006): *What Computer Talk is and Isn't: Human-Computer Conversation as Intercultural Communication*. Saarbrücken: AQ.
- Fischer, Kerstin (2010): Why it is interesting to investigate how people talk to computers and robots. Introduction to the special issue. In: *Journal of Pragmatics* (42), S. 2349-2354.
- Fischer, Kerstin (2016): *Designing Speech for a Recipient. The Role of the Addressee in So-called Simplified Registers*. John Benjamins.
- Burton, Nathan; Gaskin, James (2019): »Thank you, Siri«: Politeness and Intelligent Digital Assistants. *AMCIS 2019 Proceedings*, <https://aisel.aisnet.org/cgi/viewcontent.cgi?article=1040&context=amcis2019>.
- Grice, H. Paul (1975): *Logic and Conversation*. In: Peter Cole und Jerry L. Morgan (Hg.): *Syntax and Semantics. Speech Acts*. London, New York, San Francisco: Academic Press (3), S. 41-58.
- Habermas, Jürgen (1993): *Theorie des kommunikativen Handelns*. 2 Bände. Berlin: Suhrkamp (2).
- Hitzenberger, Ludwig (1992): Modellbildung, Versuchsaufbau und Durchführung in DICOS. In: Jürgen Krause und Ludwig Hitzenberger (Hg.): *Computer Talk*. Hildesheim, Zürich, New York: Olms, S. 62-75.
- Hitzenberger, Ludwig; Womser-Hacker, Christa (1995): Experimentelle Untersuchungen zu multimodalen natürlich sprachigen Dialogen in der Mensch-Computer-Interaktion. In: *Sprache und Datenverarbeitung* (19 (1)), S. 51-61.
- Johnstone, Anne; Berry, Umsher; Ngyuen, Tina; Asper, Alan (1994): There was a long Pause: Influencing turn-taking behavior in human-human and human-computer spoken dialogues. In: *International Journal of Human-Computer Studies* (41), S. 383-411.

- Jönsson, Arne: A model for dialogue management for human-computer interaction. www.ida.liu.se/~arnjo/papers/issd-96.pdf.
- Jurafsky, Daniel; Schriberg, Elizabeth; Biasca, Debra (1997): Switchboard SWBD-DAMSL Shallow-Discourse-Function Annotation Coders Manual, Draft 13, University of Colorado, Boulder. Colorado: Institute of Cognitive Science Technical Report. www.stanford.edu/~jurafsky/ws97/manual.august1.html.
- Kant, Immanuel (1999): *Metaphysik der Sitten*. Hamburg: Meiner Verlag.
- Krämer, Nicole C. (2008): *Soziale Wirkungen virtueller Helfer. Gestaltung und Evaluation von Mensch-Computer-Interaktion*. Stuttgart: Kohlhammer.
- Kurzweil, Ray (2005): *The singularity is near: When Humans Transcend Biology*. New York: Viking Adult.
- Locke, John (1975): *An Essay concerning Human Understanding* Edited with an introduction, critical apparatus and glossary by Peter H. Nidditch. Oxford: Clarendon Press.
- Lotze, Netaya (2016): »Chatbots – Eine linguistische Analyse«. Peter Lang.
- Lotze, Netaya (2018): Zur sprachlichen Interaktion mit Chatbots – Eine linguistische Perspektive. In: Theo Hug, Günther Pallaver (Hg.). *Talk with the Bots – Gesprächsroboter und Social Bots im Diskurs*. Innsbruck University Press. S. 29-50.
- Lotze, Netaya (2019): Psycholinguistik der KI-Forschung – Beeinflussen Künstliche Intelligenzen unsere Art zu kommunizieren?. In: *Psychologische Aspekte von Automation und Robotik*. *Psychologie in Österreich* 4/2019. [PIOe_04-19_Lotze](https://doi.org/10.1026/0013-7285/a000419)
- Lotze, Netaya; Ohnrndorf Laura (in Vorb.): *Logische Kohärenz und Filterblaseneffekte in Social Bots*. Lang.
- Moore, Roger. K. ; Nikolao, Mauro. (2017): *Toward a Needs-Based Architecture for »Intelligent« Communicative Agents: Speaking with Intention*. In: *Frontiers in Robotics and A.I.*
- Morel, Marie-Annick (1989): *Computer-human communication*. In: Martin M. Taylor, Françoise Nèel und Don Bouwhuis (Hg.): *The Structure of Multimodal Communication*. Amsterdam: North-Holland Elsevier, S. 323-330.
- Nass, Clifford; Brave, Scott (2005): *Wired for Speech: How Voice Activates and Advances the Human-Computer Relationship*. Cambridge: MIT Press.
- Nass, Clifford; Moon, Youngme (2000): *Machines and mindlessness: Social responses to computers*. In: *Journal of Social Issues* (56 (1)), S. 81-103. <http://stanford.edu/~nass/vita.htm>.
- Nass, Clifford; Reeves, Byron (1996): *How people treat computers, television, and new media like real people and places*. CSLI Publications: Cambridge University Press.
- Pennycook, Alastair (2018). *Posthumanist Applied Linguistics*. Routledge.
- Pickering, Martin J.; Garrod, Simon (2004): *Towards a mechanistic psychology of dialogue*. In: *Behavioural and Brain Sciences* (27), S. 169-225.

- Porzel, Robert; Baudis, Manja (2004): The Tao of CHI: towards effective human-computer interaction. In: Susan Dumais, Daniel Marcu und Salim Roukos (Hg.): *Proceedings of HLT-NAACL 2004*, May 2-7. Boston: ACM Digital Library, S. 209-216.
- Richards, M. A.; Underwood, K. (1984): Talking to machines: how are people naturally inclined to speak? In: *Proceedings of the Ergonomics Society Annual Conference*, S. 62-67.
- Salem, Maha; Ziadee, Micheline Sakr, Majd (2014): »Marhaba, How May I Help You?«: Effects of Politeness and Culture on Robot Acceptance and Anthropomorphization. In: *Proceedings of the 2014 ACM/IEEE international conference on Human-robot interaction*. Bielefeld: ACM, 74-81.
- Schatter, Günther (2011): Affektive Agenten, Avatare, Apparate. Emotionale Empathie als Voraussetzung für überzeugende Charaktere künstlicher Subjekte. In: Klaus-Dieter Felsmann (Hg.): *Mein Avatar und ich. Die Interaktion von Realität und Virtualität: Erweiterte Dokumentation zu den 14. Buckower Mediengespräche 2010*. München: Kopaed Verlag, S. 13-24.
- Schwarz; Benjamin (2013): »Pseudo-Kohärenz in Chatbot-Dialogen«. Unveröffentlichte Masterarbeit.
- Searle, John R. (2006): *Geist: Eine Einführung*. Frankfurt a.M.: Suhrkamp.
- Stalnaker, Robert C. (2002): Common Ground. In: *Linguistics and Philosophy* (25), S. 701-721.
- Sykora, Katharina (2000): Der Index ist die Nabelschnur. Über Foto-Doubles und digitale Chimären. In: Rolf. Aurich, Wolfgang Jacobsen und Gabriele Jatho (Hg.): *Künstliche Menschen. Manische Maschinen. Kontrollierte Körper*. Berlin: Filmmuseum Berlin – Deutsche Kinemathek, S. 117-123.
- Zoeppritz, Magdalena (1985): *Computer talk? Technical Report TN 85.05*. Heidelberg: IBM Heidelberg Scientific Center.

