

Proseminar SS 06: Einführung in die Kognitionswissenschaft – 10. [9.] Sitzung [04.07.06]:
Die subjektphilosophische Kritik des symbolisch-informationsverarbeitenden Paradigmas der Kognitionswissenschaft

Thesenübersicht zum Bezugstext: Searle, J. R. (1992) *The Rediscovery of the Mind*, Cambridge/MA [Dt: *Die Wiederentdeckung des Geistes*, Frankfurt 1993].

- (1) „Most mainstream cognitive scientists simply repeated the worst mistake of the behaviorists: They insisted on studying only objectively observable phenomena, thus ignoring the essential features of the mind“ – „The philosophy of mind of the past fifty years ... is a [collective] compulsive neurotic ... repeating the same pattern of behavior over and over“ (XII, 31)
- (2) Die moderne Physik zeigt, dass die Materie im traditionellen Sinn (*res extensa*) nicht existiert, sondern es viele unterschiedliche ontologische Ebenen der natürlichen, physischen Realität gibt mit einem je eigenen „way of existing“ (25).
- (3) Searles Position ist ein biologischer Naturalismus: „Consciousness is a higher-level or emergent property of the brain“ (14) wie die emergenten Eigenschaften des Flüssigseins (Wasserzustand) oder der Festigkeit (Eiszustand) bei H₂O-Molekülen.: „Consciousness qua conscious, qua mental, qua subjective, qua qualitative is physical“ (15).
- (4) „The ontology of the mental is an irreducibly first-person ontology [of conscioius subjectivity]“ (20, 95): „Es gibt immer eine 'Erste Person', ein 'Ich', das die mentalen Zustände hat“ (20)
- (5) Das materialistische Paradigma in der Philosophie des Geistes zielt auf eine reduktive Erklärung des Geistes in materiellen, physikalischen Begriffen, die regelmäßig an den nicht materialistisch reduzierbaren „essential features of the mind“ scheitern: „Consciousness“ / Bewusstsein – „Qualia“ / phänomenale Wahrnehmungsqualitäten – „Semantic content“ / Begriffsbedeutungen (30).
- (6) Abriss der immer neu scheiternden Theorienabfolge im materialistischen Paradigma (33–46). Searles Darstellung entspricht etwa folgender komprimierten Übersicht:

„28.3.4.1 Reduktionistischer und nichtreduktionistischer Physikalismus

Popper/Eccles (1991), Churchland (1992), Beckermann (1992) und (2000), Chalmers (1996) und Pauen (2000) bieten hier konzentrierte Überblicke über den einschlägigen Hintergrund. Dieser besteht v.a. in der sachlogischen Abfolge und Weiterentwicklung der Theorien zum Physikalismus (Materialismus) und Reduktionismus des Mentalen – speziell in der analytischen Philosophie seit dem logischen Positivismus des Wiener Kreises. Die Diskussion begann mit dem logischen Physikalismus oder Behaviorismus (Hempel, Carnap) der 30er Jahre des 20. Jh., der eine in allen Möglichen Welten geltende analytische Definierbarkeit psychologischer Terme durch physikalische annahm: psychologische Prädikate sind nichts anderes als äußere Verhaltensdispositionen (vgl. aktuell hierzu z.B. Kutschera (1993, 1–6)).

Da sich diese sehr starke These als nicht haltbar herausstellte, wurde sie zum psychologischen Behaviorismus (Feigl) der 50er Jahre abgeschwächt, der nur noch eine in der aktuellen Welt geltende nomologische Definierbarkeit psychischer Terme durch physikalische annahm. Diese Position wurde meist als generische oder Typenidentitätsthese vertreten: jedem psychologischen Prädikatentyp ist nomologisch ein physikalischer Prädikatentyp zugeordnet (vgl. Kutschera 1993, 6–7, 10–23). Aus der nomologischen Korrelation wurde in logisch problematischer Weise auf Identität geschlossen. Dieses *Non-sequitur* wurde von der Kritik (Nagel, Kripke, vgl. neuerdings auch Kutschera (1993), Metzinger (1993) und Chalmers (1996)) beanstandet: Mentale Prädikate (phänomenale *Qualia* und intensionale Bedeutungen) sind prinzipiell nicht identisch mit physikalischen Zuständen, auch wenn eine theoretische Identifikation vorgenommen werden kann, wie: Wasser = H₂O. Denn bei phänomenalen Qualitäten und intensionalem Sinn ist das mentale Sein das primäre und wesentliche Sein (kantisch: das Ding in der Erscheinung *ist* das Ding an sich). Das zugeordnete objektive (physikalische oder neuronale) Korrelat ist sekundär.

Durch diese Kritik wurde der Physikalismus veranlasst, sich weiter abzuschwächen in Richtung einer Vorkommnis-Identitätsthese (Davidson) und Supervenienztheorie, die nur noch behauptet, dass jedem psychologischen Unterschied ein physikalischer Unterschied korrespondiert. Diese Supervenienz kann dreifach interpretiert werden. Einmal als *schwache* Supervenienz: Jeder psychologischen Eigenschaft entspricht eine physikalische Eigenschaft, die in verschiedenen möglichen Welten verschieden sein kann. Zweitens als *starke* Supervenienz: Jeder psychologischen Eigenschaft entspricht in allen möglichen Welten ein und dieselbe physikalische Eigenschaft, die aber wegen ihrer Komplexität und aufgrund ihrer definitonischen Abkünftigkeit von der psychologischen Eigenschaft nicht als deren nomologische Definition auftreten kann. Drittens die *globale* Supervenienz: Einem Unterschied in der Totalität der psychologischen Eigenschaften einer Welt entspricht stets irgendein Unterschied in der Totalität der physikalischen Eigenschaften dieser Welt (vgl. Kutschera (1993, 23–33)). Die Vorkommnis-Identitätsthese und Supervenienztheorie sind Formen des heute vorherrschenden nichtreduktionistischen Physikalismus, der das Mentale als irreduzible, holistisch emergente Eigenschaft betrachtet, dessen Natur und Leistung meist funktionalistisch interpretiert wird, also durch die essentiell *kausale* Rolle zwischen sensorischer

Eingabe und motorischer Ausgabe im Organismus. Schwierig, wenn nicht sogar logisch fragwürdig ist dabei die Spannung zwischen einerseits der Behauptung der Epiphänomenalität des Mentalen, und andererseits der Behauptung der Kausalität (Abwärtsverursachung) des Mentalen (vgl. Kim 1992).

Neben den genannten Theorien existiert seit den 80er Jahren des 20. Jh. eine Position, die als eliminativer Materialismus bezeichnet wird (Feyerabend, Rorty, Churchland). Er entwickelt nichts sachlich Neues, sondern ist eher als gegen den nichtreduktionistischen Physikalismus gerichtete Utopie und Spekulation zu fassen, die postuliert, dass die kulturelle Evolution und der wissenschaftliche Fortschritt trotz allem im zukünftigen Denken und Erleben der Menschen zu einer Metamorphose des Mentalen im Sinne einer physikalischen Rekonstruktion führen wird. Dass wir später also nicht mehr etwa „Wärme“ sagen werden, sondern „molekulare kinetische Energie“. [Natterer, Paul: *Systematischer Kommentar zur Kritik der reinen Vernunft. Interdisziplinäre Bilanz der Kantforschung seit 1945*, Berlin/New York 2003, 634–635]

- (7) „Die Basisannahme hinter dem Materialismus ist im Wesentlichen genau die Cartesische Annahme, dass Materialismus notwendig Antimaterialismus einschließt und Mentalismus notwendig Antimaterialismus einschließt“ (54).
- (8) Theorieunabhängige Existenz / Ontologie mentaler Zustände: „We do not postulate beliefs and desires to account for anything. We simply experience conscious beliefs and desires“ (59). Die mentalen Realitäten sind nicht vorwissenschaftliche empirische Hypothesen oder mythische Konstrukte, sondern harte existenzielle Fakten, die das Gegenstandsgebiet der Philosophie des Geistes definieren. Und: Selbst theorie- oder kulturabhängige Entitäten wie z.B. Zinssätze, Literaturgattungen und Fußballspiele sind in ihrer Existenz nicht abhängig von einer gelingenden Reduktion auf Physik oder Neurobiologie (59).
- (9) Die Analyse des Geistes muss drei Faktoren berücksichtigen: (a) Biologisches *Gehirn* – (b) Mentales *Bewusstsein* – (c) Äußeres intelligentes *Verhalten*. Der entscheidende und zentrale Faktor ist (b) Mentales Bewusstsein, da rein logisch im Blick auf den Faktor (a) statt eines biologischen Gehirns (auf Kohlenwasserstoffbasis) ein Gehirn z. B. auf Siliziumbasis vorstellbar ist, und bei (c) trotz intelligentem Verhalten das Fehlen mentalen Bewusstseins vorstellbar ist (Zombies, Knowbots, Cyborgs) (66).
- (10) Zwei Ecksteine des naturwissenschaftlichen Weltbildes der Gegenwart: (a) Atomtheorie (Felder und Teilchen) der Materie und (b) Evolutionstheorie der Biologie: „Although, like any other theory, they might be refuted by further investigation [...] this world-view is not an option“ (86, 90). Searle setzt die physikalische Methode absolut und verbindet mit ihr im Gegensatz etwa zu Heisenberg, Dürr und Hawking, ein atheistisches philosophisches Weltbild: „The existence of God ... and the hypothesis of an afterlife ... we cannot take ... seriously“ (90)
- (11) Bewusstsein, Intelligenz, Sprache, rationales Denken sind – so Searle – biologische Phänomene wie Photosynthese, Zellteilung und Verdauung: Bewusstsein ist eine emergente Eigenschaft komplexer, physikalisch-chemischer und biologischer Systeme (90). Dies bedeutet aber andererseits auch: „The real world ... described by physics and chemistry and biology, contains an ineliminably subjective element“ (95)
- (12) Das subjektive Bewusstsein (= Beobachter) ist Bedingung der objektiven Beobachtung und Weltansicht (98).
- (13) Bewusstsein und bewusstes Verhalten gibt uns eine neue Dimension von Flexibilität – Kreativität – Sensitivität gegenüber unbewusstem Verhalten, das auf Routine – Gelerntes – Gespeichertes eingeschränkt ist (107).
- (14) Arten von Reduktionismus (111–117):
 - (a) Ontologische Substanz [*type-type*] Reduktion: Stuhl → Molekülsammlung; Gen → DNS-Molekül.
 - (b) Ontologische Eigenschafts-Reduktion: Wärme → mittlere kinetische Energie der Molekülbewegung; Licht → elektromagnetische Strahlung; Farbe → Lichtreflektanz bzw. Photonenemission.
 - (c) Theoretische Reduktion: Ableitung einer untergeordneten Theorie aus einer übergeordneten Theorie; z. B. Klassische Gasgesetze → Statistische Thermodynamik.
 - (d) Logische Reduktion: Definitorenische Reduktion, z.B. Zahlen → Mengen von Mengen.
 - (e) Kausale Reduktion: Reduktion der Existenz und der kausalen Kraft einer untergeordneten Ebene auf kausale Kräfte einer übergeordneten Ebene: Solide Festkörper und ihre kausalen Kräfte (Undurchdringlichkeit, Widerstand gegen Druck) → kausale Kräfte von molekularen Vibrationsbewegungen in Kristallgittern.Searle: Bewusstsein ist ein irreduzibles emergentes Merkmal der Realität, das nicht in Form der

- Reduktionen (a) und (b) eliminiert werden kann, sondern nur eine kausale Reduktion (e) auf die Ebene der Physik ist möglich. Nagels, Kripkes, Jacksons Argumente gegen die Reduzierbarkeit von Bewusstsein gelten nicht nur epistemologisch, sondern auch ontologisch: „Consciousness consists in the appearances themselves. Where appearance is concerned we cannot make the appearance-reality distinction because the appearance is the reality“ (122).
- (15) Strukturelle Merkmale des Bewusstseins (128–141):
- (a) Finite Modalitäten: Äußere und innere Sinne + Denken + Emotionen + Triebe + Lust / Unlust
 - (b) Horizontale, diachrone, zeitliche Einheit (Synthesis) und vertikale, synchrone, räumliche Einheit (Bindung)
 - (c) Mehrheitlich intentional + perspektivisch: I am conscious **of** x [z.B. Trier] **as** y [z.B. spätromische Hauptstadt]
 - (d) Subjektives Empfinden
 - (e) Verknüpfung von Bewusstseinsinhalten
 - (f) Figur (Fokus)-Grund (Hintergrund)-Struktur
 - (g) Vertrautheit und Erwartung der Bewusstseinsinhalte entsprechend apriorischen angeborenen Instinkten und Kategorien
 - (h) Überschuss (*overflow*): Der implizite Inhalt ist reicher und weiter als der explizite Gedanke
 - (i) Zentrum (Aufmerksamkeit)-Umfeld (Begleitbewusstsein)-Charakteristik
 - (j) Situiertheit: raumzeitliche und sozioemotionale Randbedingungen
 - (k) Stimmung
 - (l) Lust / Unlust
- (16) Klärung dreier häufiger Fehler einer Theorie des Bewusstseins: (a) Selbstbewusstsein ist nicht aktuelles, sondern nur potenzielles intentionales (inhaltliches) Bewusstsein = Reflexion [auf eigenen Körper und] auf vorgängiges intentionales objektives Bewusstsein; (b) Introspektion ist nicht eine spezielle intentionale Erkenntnisfähigkeit, sondern das reflexive Denken an eigene mentale Zustände; (c) Eine Irrtumslosigkeit von 1. Person-Berichten gibt es nur von intentionalem objektivem Bewusstsein, nicht von reflexiv-subjektivem Bewusstsein (141–145).
- (17) *Connection principle* zwischen Bewusstem und Unbewusstem: „The notion of an unconscious mental state implies [potential] accessibility to consciousness“ (152), auch wenn viele mentale Einzelzustände zeitweise oder für immer nicht bewusst sind (z.B. bei Schock, Amnesie, Verdrängung, Gehirntrauma). Bewusstes und Unbewusstes muss „aus demselben Stoff“ und „von derselben Art“ sein (160).
- (18) Radikal Unbewusstes ist nur und genau die neurophysiologische Basis bewusster Zustände (167–168)
- (19) Freud / Psychoanalyse: Problematische mentale Ontologie des Unbewussten: Alles Mentale ist an sich unbewusst. Bewusstsein ist nur ein spezieller Wahrnehmungsmodus des Mentalen, der diesem äußerlich und akzidentell ist wie der von außen kommende Lichtkegel einer Taschenlampe auf den Gegenständen eines dunklen Speichers. Tatsächlich gilt aufgrund des *Connection principle* (siehe 17) der Primat und die Priorität des mental Bewussten (169).
- (20) Die radikale Unterbestimmtheit (*vagueness*) von Wörtern und die Unvollständigkeit (*incompleteness*) von Sätzen (und inneren mentalen Zuständen) wird einerseits durch holistisches intentionales Allgemeinwissen (*The Network*) ergänzt, andererseits durch den Hintergrund (*The Background*) von apriorischen und erworbenen, automatisierten hermeneutisch-linguistischen Kompetenzen der menschlichen Natur und Kultur (175–192).
- (21) „The Critique of Cognitive Reason“: Grundannahmen des traditionellen symbolisch-informations-verarbeitenden Paradigmas der KW und KI sind:
- (a) Kognitive Prozesse sind weitgehend unbewusst (Chomsky, Marr, Fodor)
 - (b) Das Gehirn bzw. die Neurobiologie sind nebensächlich
 - (c) Bewusstsein ist nebensächlich
 - (d) Kognitive Prozesse sind formale Rechenprozesse (Symbolmanipulation) (197–198)
- (22) Die Kernthese der traditionellen KW und KI ist somit: „The mind is to the brain as the program is to the hardware“ (200). Diese These impliziert, so Searle, drei unterschiedliche Fragen:
- (a) Sind Gehirnprozesse Berechnungen? Ist das *Gehirn* ein digitaler Computer (= These des Funktionalismus)? Antwort: Mentale Zustände sind zumindest *auch* syntaktische Rechenzustände.
 - (b) Ist der *Geist* ein Computerprogramm (= Starke KI). Antwort: Nein, ein Programm ist *per*

definitionem nur formal und syntaktisch. Der Geist enthält jedoch intrinsische mentale Gehalte (Bedeutungen), und „semantics is not intrinsic to syntax“ (210) wie Searles eigenes berühmtestes Gedankenexperiment mit dem Chinesischen Zimmer (*Chinese room*) zeigt.

(c) Können die Operationen des *Gehirns* auf einem digitalen Computer simuliert werden (= Schwache KI)? Antwort: Ja, denn dies besagt die Church-Turing-These: Jeder Algorithmus kann auf einer Turing-Maschine (= digitaler binärer Rechner) implementiert werden (200).

(23) Problematisch ist also v.a. die Frage bzw. Interpretation (b) Der *Geist* ist ein Computerprogramm (= Starke KI). Diese Problematik erhellt aus folgenden vier Einwänden:

(a) Der binäre Maschinenkode des Computers (Turing-Maschine) ist in unterschiedlichen physischen Medien realisierbar (*multiple realizability*) und ist in sich nur die Produktion physikalischer Wirkungen, deren syntaktische Zuordnung und Interpretation „depend on an interpretation from outside“ (209). Syntax ist kein physikalisches Merkmal: „syntax is not intrinsic to physics“ (210).

(b) Echte, reale Computer benötigen einen intelligenten Benutzer oder Beobachter. Ohne ein kognitives bewusstes Subjekt als Nutzer sind sie unverstandene elektrische Schaltkreise und nur *potenzielle Denkwerkzeuge*. Wenn der Geist ein Computerprogramm ist, muss ein bewusster Beobachter / Nutzer im Gehirn („*Homunculus*“) postuliert werden (210).

(c) Syntax hat keine ursächlichen Kräfte. Das Computerprogramm-Modell der Kognition bietet keine kausale Erklärung der Kognition: Die Simulation der Kognition ist keine kausale Erklärung oder Theorie der Kognition (210).

(d) Das Gehirn macht keine Informationsverarbeitung nach Art eines Computers, wo ein äußerer Agent Information in einer syntaktischen Form enkodiert, die von den Schaltkreisen des Rechners mechanisch implementiert und verarbeitet werden kann – in Schritten und mit abstrakten Ausgabesymbolen, die ein Nutzer / Beobachter von außen syntaktisch und semantisch interpretieren kann, obwohl der Rechner (*hardware*) keine intrinsische Syntax oder Semantik hat: „It is all in the eye of the beholder“ (223). Dabei ist das physikalische Medium uninteressant.

Die neurobiologischen Prozesse des Gehirns sind dagegen nicht beobachterabhängig und verkörpern unmittelbar – in intrinsischer Intentionalität – konkrete spezifische *bewusste* Lebensphänomene und Welterfahrungen, nicht nur mittelbar – in extrinsischer Intentionalität – abstrakte Worte oder Symbole: „The study of the mind is the study of consciousness“ (227). Dabei ist das physische Medium wichtig: Geist und Kognition sind ontologisch ein biologischer Lebensprozess (227).

(24) Informationsverarbeitung im Sinne der traditionellen KW „is at much too high a level of abstraction to capture the concrete biological reality of intrinsic intentionality“ (224).

(25) Bei vorbewusster Informationsrepräsentation und -verarbeitung (z.B. im Fall der linguistischen Tiefengrammatik oder vorbewussten Rechenleistungen der visuellen Wahrnehmung) sollte man nicht von unbewussten Repräsentationen und Regeln sprechen, sondern sagen: „There is simply a [neurobiological] hardware mechanism that functions in a certain way and not in others“ (245). Eine zu begrüßende Korrektur des problematischen symbolisch-informationsverarbeitenden Paradigmas der KW und KI findet sich im Ansatz der parallel-verteilten Informationsverarbeitung (Konnektivismus, PDP), welcher der tatsächlichen Natur der Kognition näher kommt (246–247).