

Ohne Redundanz keine Anschlusskommunikation

*Zum Verhältnis von Information und Kommunikation**

Rudolf Stöber

Kommunikation, Information und Redundanz bedingen sich wechselseitig. Da Information eine Zeitfunktion ist, liegt die Ursache kommunikativer Redundanz in der Notwendigkeit, mit „Anschlusskommunikation“ aus Verständlichkeitsgründen eine Brücke über die Zeit zu schlagen. In dem Aufsatz werden drei Informationsbegriffe unterschieden: der informationstheoretische, der sprachliche und der kommunikationswissenschaftliche. Der sprachliche Informationsbegriff wird gängiger Systematik entsprechend in einen semantischen, syntaktischen und pragmatischen Informationsbegriff geschieden. Den drei Haupt-Informationsbegriffen werden drei Redundanzen 1. bis 3. Ordnung zugeordnet. Es wird gezeigt, wie über die Informationsbegriffe und die Redundanzen hinweg Verständlichkeit erzeugt wird. Die verschiedenen Informationsbegriffe lassen sich als einander bedingende Teilmengen darstellen, die einerseits aus verschiedenen Regelwerken und Aushandlungen abzuleiten, andererseits in unterschiedlichem Umfang zu messen sind. Abschließend werden die unterschiedlichen Begriffsauffassungen von Information und Kommunikation, die sich in verschiedenen Wissenschaften herausgebildet haben, zueinander in Beziehung gesetzt.

Schlagwörter: Redundanz, Information, Kommunikation, Zeit, Verständigung, Sprachpragmatik, Handlungstheorie, Systemtheorie, Verhaltenspsychologie

Meinem Vater Wilhelm Stöber zum 80. Geburtstag

1. Problemaufriss

Sichtet man kommunikationswissenschaftliche Literatur, insbesondere Einführungen, Handbücher und Lexika, so stellt man rasch fest, dass zentrale Begriffe des Fachs unterschiedlich definiert werden. Insbesondere gilt dies für die Begriffe Information und Kommunikation. Je nachdem, ob systemtheoretisch, handlungstheoretisch oder aus anderer Perspektive argumentiert wird, lassen sich gar gegensätzliche Begriffsauffassungen finden. Wird Kommunikation systemtheoretisch verstanden, stößt man unweigerlich auf Niklas Luhmanns Diktum „Kommunikation ist unwahrscheinlich“ (Luhmann 1981: 26). Vielzitiert, nicht selten ohne den Kontext, findet sich das erste Axiom der verhaltenspsychologisch orientierten Palo-Alto-Schule: „Man kann nicht nicht kommunizieren“ (Watzlawick/Beavin/Jackson 1972: 51). Häufig erwähnt, zumeist als Gegenposition zur Kommunikationswissenschaft, wird der informationstheoretische Informationsbegriff nach Claude E. Shannon (1948).¹ Den kommunikationswissenschaftlichen Informationsbegriff definierte Harry Pross am prägnantesten: Information sei ein „Korrelat von Unkenntnis“ (Pross 1977: 23).

* Ich danke Helmut Glück, Gabriele Mehling und Gunda Stöber für ihre kritische Lektüre des Manuskripts und instruktive Hinweise.

1 Nicht selten wird er jedoch verkürzt dargestellt, oder man bezieht sich auf die Variante eines Werks aus der Koautorschaft mit Warren Weaver, die missverständlich suggeriert, die semantischen Aspekte von Information seien ebenso berechenbar wie die informationstheoretischen; vgl. Shannon/Weaver 1976: 35.

Die Spannweite der Begriffsauffassung liegt also je nach handlungstheoretischen, verhaltenspsychologischen, systemtheoretischen oder kommunikationswissenschaftlichen Perspektiven so weit auseinander, dass zwischen den Auffassungen kaum noch eine vermittelnde Position möglich scheint (Merten 1977). Dabei ist es lohnend, die unterschiedlichen Begriffsauffassungen von Information und Kommunikation so zueinander in Beziehung zu setzen, dass sie sich auseinander ableiten lassen und in einen kommunikationswissenschaftlichen Kommunikations- und Informationsbegriff münden. Das soll mit folgendem Beitrag geschehen. Er schenkt der Redundanz besondere Beachtung, die als konstitutives Element von Kommunikation bislang in der Kommunikationswissenschaft eher vernachlässigt wurde. Zugespitzt: Ohne Redundanzen (verschiedener Ordnung) wäre Kommunikation nicht möglich.

Sprache dient hier als Beispiel für diverse Formen der Humankommunikation, denn sie ist das mit Abstand flexibelste und komplexeste Instrument menschlicher Kommunikation. Dabei wird im Folgenden der Einfachheit halber mit einem Idealbild natürlicher Sprachen operiert. Sprache wird idealtypisch in die folgenden Erörterungen eingeführt, um sich nicht in die Untiefen sprachwissenschaftlicher Diskussionen begeben zu müssen. Es werden also bewusst Sonderformen und -fragen, z. B. Spezialaspekte restringierter Codes, konstruktionsgrammatische Fragen, Regelunsicherheiten, eingeschränkte Wortschätze bei den Kommunikationsteilnehmern etc., ausgeblendet. Es wird ausdrücklich die sprachliche Kommunikation behandelt, weil sich anhand visueller oder nonverbaler Kommunikation die fraglichen Probleme nicht so leicht veranschaulichen lassen. Die Schlussfolgerungen gelten aber *mutatis mutandis* auch für andere Ausdrucksformen menschlicher Kommunikation. Auch die Scheidung in Semantik, Syntaktik und Pragmatik hat analytische und darstellende Gründe, die soziale Kommunikation operiert mit ihnen in unauflöslicher Verbindung. Eine letzte Vorbemerkung: Die Gleichungen sind als abstrakte Symbolsprache zu verstehen, nicht im Sinne exakter mathematischer Terme.

2. Information informationstheoretisch

Sybillinisch formulierte der Mathematiker Norbert Wiener, „Information ist Information, weder Materie noch Energie“ (Wiener 1963: 192). Information sei eine Zeitfunktion: $\inf = f(t)$: Informationstheoretisch betrachtet ist der Informationsgehalt der spezifischen Information $t \rightarrow t'$ gleich derselben Information $t' \rightarrow t''$ – unabhängig vom Startzeitpunkt t oder t' (Wiener 1963: 111, 150; ähnlich Shannon 1948: 2). Darin ist schon implizit enthalten, dass sowohl die Informationstheorie Shannons als auch die Fortentwicklung zur Kybernetik durch Wiener axiomatisch von der Berechenbarkeit der Information ausgingen. Informationssysteme tendieren zum Informationsverlust; darum hat Shannon 1948 den Begriff der Entropie in Anlehnung an die Thermodynamik als Berechnungsmaß eingeführt. Damit hat er einige Verwirrung gestiftet. Die Unsicherheit resultiert daher, dass Entropie (= Unordnung) sowohl der Redundanz (= Wiederholung) als auch der Information gegenübergestellt werden kann. Die Informationstheorie definierte die Information als negative Entropie $[-H]$, die maximale Information(smenge) h , die ein System enthalten kann, entspricht jedoch zugleich dessen maximaler Entropie $[H_{\max}]$. Der bis zum Extremwert wichtige Vorzeichentausch wird mit dem Erreichen des Maximums hinfällig. Was bedeutet im Sinne der Informationstheorie die konkrete „Information“ als negative Entropie bzw. maximale Information als maximale (negative) Entropie: $h = -H$ oder $h_{\max} = +/ - H_{\max}$?

Zumeist beginnt die Erklärung der maximalen Informationsmenge h mit der Ableitung aus der als bekannt vorausgesetzten Zahl der Gleichwahrscheinlichkeiten n und

dem verfügbaren Primärzeichensatz. Das liest sich in Anlehnung an Shannon dann als $h = {}_2\log^n$. D. h. beispielsweise im dualen Zeichensatz der binären Computerwelt: Es gibt nur zwei Primärzeichen, die 0 und die 1. Das eine steht für „Strom fließt nicht“, das andere für „Strom fließt“. Ein Bit (binary digit), doppeldeutig im Englischen (= bisschen), kann zwei Werte annehmen. Wenn man einen natürlichen Zeichensatz von 256 Zeichen konstruieren will, hieße das $h = {}_2\log^{256}$ = acht Bit = ein Byte. Computerworte, Bytes, bestehen daher aus 8 Bit oder 8 primären Zeichen – sie entsprechen *grosso modo* einem Zeichen (Z) der natürlichen Sprachen. Doch ist diese Annäherung unnötig kompliziert, da in der digitalen Welt die maximale Information h vielfach als bekannt vorausgesetzt werden kann: Festplattengrößen werden derzeit üblicherweise in Giga- oder Terabyte angegeben, also als $8 \cdot 10^9$ bzw. $8 \cdot 10^{12}$ Bit. Ein Giga- oder Terabyte enthält jedoch nur dann die maximale Informationsmenge h , wenn zwei Bedingungen erfüllt sind: 1. Jeder Speicherplatz ist belegt, d. h. Leerstellen (0000 ...) gibt es nicht. 2. Redundanzen müssen ausgeschlossen sein. Paradoxerweise erfüllt nur eine Giga- oder Terabyte-lange, absolut zufällige Zeichenkette beide Voraussetzungen, z. B. die Zahlenfolge 1 4 1 5 9 2 6 5 3 5 8 9 7 9 3 auf ein Giga- oder Terabyte zufällig verlängert.² Dann wäre $h = -H_{max}$. Doch ist die maximale Information(sichte) als maximale Entropie für den Menschen „sinnlos“ und bedeutungsleer. Eine maximal befüllte Festplatte wäre ohne jede sinnvolle Information. Der informationstheoretische Ausdruck h soll im Folgenden durch $Inf_{(inf)}$ ersetzt werden, um deutlich zu machen, dass verschiedene Informationsbegriffe unterschieden werden müssen.

Spannender ist daher die umgekehrte Frage, wie wahrscheinlich ein spezifischer Informationszustand ist. Die Wahrscheinlichkeit, mit der einer der beiden möglichen Werte eines Bit vorkommt, beträgt $1:2 = \frac{1}{2} = 0,5 = 50$ Prozent. Ein Byte besitzt $2^8 = 256$ mögliche Werte. Die Wahrscheinlichkeit, mit der das „Computerwort“ den entsprechenden Wert annimmt, beträgt demnach $1/256$ oder ca. 0,39 Prozent. Die 256 Werte findet man in herkömmlichen PCs als Zeichensatz abgelegt, der, neben den Klein- und Großbuchstaben, Ziffern und Interpunktionszeichen auch diakritische Zeichen, mathematische Operatoren, Währungssymbole sowie weitere Sonderzeichen umfasst. Die meisten dieser Zeichen sind Primärzeichen, einige sind darüber hinaus zudem Sekundärzeichen: u. a. die Währungssymbole und die Ziffern von 0 bis 9, die zugleich den Zahlen von Null bis Neun entsprechen. Obwohl ein Zeichensatz von 256 möglichen Werten schon umfangreich ist und obwohl er mehr Zeichen enthält, als zur Codierung einer natürlichen Sprache notwendig sind, reicht er für alle Sprachen der Welt nicht hin. Daher gibt es eine Handvoll internationale Zeichensätze, weitere proprietäre Zeichensätze und zudem mehr als ein Dutzend nationale Varianten, um die binäre Primärzeichenwelt aus Nullen und Einsen der Primärzeichenwelt der natürlichen Sprachen anzupassen. Der Zeichensatz soll im Folgenden als

$$\sum_i^n Z_{(1. \text{ Ordnung})} (\cong Inf_{(inf)}) \text{ abgebildet werden.}^3$$

Bei einem „Text“ aus N Zeichen (Z) und mit einem Zeichensatz, der aus dem einfachen Alphabet mit seinen 26 Zeichen bestünde, gälte

2 Doch ist die Ziffernfolge nicht so zufällig, wie sie scheint, da sie die ersten 15 Nachkommaziffern der Zahl π listet.

3 Dabei besteht zwischen beiden Ausdrücken nur der Unterschied, dass mit der Summation natürlicher Zeichen auf den für natürliche Sprachen geeigneten Zeichensatz verwiesen werden soll.

$$N = \sum_i^n \log 26 \text{ bzw. } N = \sum_i^n Z(26).$$

Somit könnte der Text folglich 26^N gleichwahrscheinliche Lösungen annehmen. Bei einer Zeichenkette von vier Zeichen dieses Zeichensatzes hieße das $26^4 = 456.976$ Möglichkeiten.⁴ Die Gleichwahrscheinlichkeit, mit der eine dieser Möglichkeiten zutrifft, beträgt $26^{-4} = 1/456.976$. Diese Gesamtmenge aller möglichen Informationen, die ein System annehmen kann, ist mithin $= Z^N$, bei einem dualen Zeichensystem $= 2^N$, im 26-Zeichensystem $= 26^N$ usw. Die konkrete Information umfasst jedoch immer nur N Zeichen.⁵ Die Zahl der Primärzeichen eines Zeichensatzes bestimmt mithin die Basis, die Zeichenlänge eines Textes den Exponenten. Aus dem positiven Term (Z^N) ergibt sich die Zahl der gleichwahrscheinlichen Möglichkeiten. Ist der Exponent negativ (Z^{-N}), wird die Wahrscheinlichkeit eines spezifischen Zustands angezeigt.

Da das duale Zeichensystem für die natürlichen Sprachen ebenso wenig ausreicht wie 26 Buchstaben, soll im Folgenden zur Vereinfachung ein Byte als das Äquivalent zu den natürlichen Zeichen gelten. Bezogen auf diesen mit Hilfe eines PC geschriebenen Aufsatzes hieße das: Das mathematische Produkt der Möglichkeiten, die dieser Text mit der Länge von ca. 50.000 Zeichen annehmen könnte, beträgt $256^{50.000}$. Im dualen Zahlensystem ausgedrückt, also in Bit statt Byte, ist von $2^{(50.000 \cdot 8)}$ gleichwahrscheinlichen Möglichkeiten auszugehen. Beide Produkte entsprechen sich: $256^{50.000}$ Byte $= 2^{(400.000)}$ Bit. Das Ergebnis ist unvorstellbar groß, herkömmliche Tabellenkalkulationen können es nicht mehr potenzieren. Die Unwahrscheinlichkeit ist ungeheuer, so dass sie zunächst einmal veranschaulicht werden sollte:

In dem Möglichkeitsraum steht jeder sinnvolle wie sinnlose Text von der entsprechenden Zeichenlänge, der je geschrieben wurde, der jemals geschrieben werden wird oder der geschrieben werden könnte, den aber niemand je schreiben wird. Um (halbwegs) Sinnvolles zu erzeugen, gelten einige weitere Restriktionen, die noch erörtert werden. Aber selbst darin stecken potenziell die Schriftrollen der Bibliothek von Alexandria oder Babel (frei nach Jorge Luis Borges), Shakespeares Sonette, Auszüge aus seinen Komödien und Dramen, Goethes „Leiden des jungen Werther“, Einsteins allgemeine und seine spezielle Relativitätstheorie; ebenso stecken darin alle Untertitel für Hörgeschädigte zu jedem beliebigen Fernsehfilm, die Sprechblasentexte aus „Asterix“ oder „Donald Duck“, Beschreibungen von Manga Comics, Texte von Rosamunde Pilcher oder Karl May, Kontoauszugsdaten, Geheimdienstberichte oder was immer wir uns vorstellen können und selbstredend auch all das, wovon wir keine Vorstellung haben. Unter den Optionen befinden sich ebenso sinnvolle wie abstruse Kombinationen. Wir können uns einen Text abwechselnd mit Sätzen oder Absätzen aus Pierre Bourdieus „Die feinen Unterschiede“ und Niklas Luhmanns „Soziale Systeme“ ebenso vorstellen wie eine exakte Übersetzung dieses Textes ins Flämische, Finnische oder eine andere natürliche Sprache.⁶

Soweit, so gut, könnte man sagen, Luhmann ist zuzustimmen, bzw. er ist noch zu präzisieren: Kommunikation ist *extrem* unwahrscheinlich. Allerdings wird das Zwi-

4 Der Logarithmus von N zur Basis N ist bekanntlich immer 1: $\log_2 2 = 1$; $\log_3 3 = 1$; $\log_4 4 = 1$; ... $\log_n n = 1$. Bei 4-facher Aufsummierung des $\log_{26} 26 [= 1]$ im Exponenten zu $26 = 26^{1+1+1+1}$.

5 Das kann man zwar kompliziert als $N \cdot \log_2 2$ im dualen System oder $N \cdot \log_{26} 26$ im 26-Zeichensystem etc. schreiben, ist aber überflüssig.

6 Letzteres allerdings nur, sofern der Primärzeichensatz dies zulässt und ein Ausgleich für die Textlängendiskrepanz in den verschiedenen Sprachen gefunden wird. Die konkrete Information besteht jedoch jeweils aus 50.000 Zeichen oder 400.000 Bit.

schenergebnis weder Luhmann noch dem Problem gerecht. Luhmanns Unwahrscheinlichkeitstheorem ist nur partiell auf die Informationstheorie zurückzuführen; es hängt ebenso mit der Dreifachschachtelung der Kommunikation aus Information, Mitteilung und Verstehen zusammen. Die Informationstheorie hingegen reduziert den Informationsbegriff auf die Wahrscheinlichkeitsrechnung und schließt explizit alle „semantic aspects“ als irrelevant für das informationstechnische Problem der Datenübertragung aus: „The fundamental problem of communication is that of reproducing at one point either exactly or approximately a message selected at another point. Frequently the messages have meaning; that is they refer to or are correlated according to some system with certain physical or conceptual entities. These semantic aspects of communication are irrelevant to the engineering problem. The significant aspect is that the actual message is one selected from a set of possible messages. The system must be designed to operate for each possible selection, not just the one which will actually be chosen since this is unknown at the time of design. If the number of messages in the set is finite then this number or any monotonic function of this number can be regarded as a measure of the information produced when one message is chosen from the set, all choices being equally likely.“ (Shannon 1948: 1)

Die Gleichwahrscheinlichkeit der Myriaden von denkbaren Kombinationen, welche die Informationstheorie betont, hat also eine Doppel-Funktion: Zum einen, den technischen Vermittlungskanal so zu konstruieren, dass er unabhängig von den Bedeutungsinhalten für jede denkbare, gegenwärtige wie zukünftige Symbolübertragung zwischen Sender und Empfänger gleich geeignet ist. Zum anderen, das technische Speichermedium so zu entwerfen, dass es jede denkbare Information hier, heute und in Zukunft zum Gebrauch der Nutzer speichern kann. Bezogen auf den Speicherplatz eines Speichermediums bedeutet der informationstheoretische Informationsbegriff also ein Hohlmaß: Wenn man das Volumen eines Tanks berechnen kann, hat man damit noch keine Information über dessen Inhalt. Bezogen auf ein technisches Übertragungsmedium berechnet man die maximal mögliche Kanalkapazität je Zeiteinheit, nicht jedoch die spezifische Informationsübertragung. Damit wird aber erst die technische Grundlage menschlicher Kommunikation umschrieben. Daher ist anschließend in zwei Schritten die sprachliche Bedeutungskonstruktion zu reflektieren, um den Informationsbegriff weiter präzisieren zu können.

3. Information semantisch-syntaktisch

Auf die semantische Ebene begrenzt, werden die informationstheoretischen Gleichwahrscheinlichkeiten schon deutlich weniger: Jede natürliche Sprache besitzt, bezogen auf acht-zeichenlange Einheiten, eine Redundanz von ungefähr 50 Prozent.⁷ Weil es noch weitere Redundanzen gibt, soll sie als informationstheoretische *Redundanz*_(1. Ordnung) bezeichnet und als statistischer Ausdruck für die zur Sprachverständlichkeit notwendigen Wiederholungen verstanden werden.⁸ Semantische Informationen bestehen, so könnte man *näherungsweise* formulieren, aus der maximalen Infor-

7 Bei längeren Zeichenketten liegt sie höher (Shannon 1948: 14; Shannon/Weaver 1976: 23). Nach Krippendorff erreicht der Redundanzwert – ohne die Achtzeichenspezifikation – bis zu 70 %, vgl. Krippendorff 1986: 28. Krippendorffs Angabe liegt im Bereich der Beispiele in Fußnote 17.

8 Diese Redundanz scheint Sprache informationsineffizienter zu machen. Das stimmt jedoch nur bedingt, da sie Übermittlungs- und sonstige Störungen reduzieren hilft. Krippendorff scheint hinsichtlich der Korrekturfunktion der Redundanz gegenteiliger Auffassung zu sein. Es

mation (= negative Entropie = $h = Inf_{(inf)}$) abzüglich der *Redundanz*_(1. Ordnung), die natürlichen Sprachen eigen ist:

$$Inf_{(sem)} \cong Inf_{(inf)} - Redundanz_{(1. Ordnung)} (A)$$

Dabei ist $Inf_{(sem)}$ eine Teilmenge von $Inf_{(inf)}$ – d. h. $Inf_{(sem)} \subseteq Inf_{(inf)}$. Zugleich ist mit $Inf_{(sem)}$ analog zur $Inf_{(inf)}$ keine konkrete Information und nicht einmal ein spezifischer Wortschatz gemeint, sondern der maximal mögliche Wortschatz einer natürlichen Sprache. Allerdings ist der Term (A) unscharf, denn die Subtraktion der *Redundanz*_(1. Ordnung) beschreibt das Ergebnis nur messtechnisch und kann, muss aber nicht, semantisch richtige oder gar sinnvolle Information anzeigen.

Natürliche Sprachen verfügen nicht und benötigen auch keinen unendlich großen Thesaurus. Im Gegenteil, der aktuelle Wortschatz lebender Sprachen ist beschränkt. Das Englische wie das Deutsche umfassen, nimmt man das „Oxford Dictionary“ oder das „Deutsche Wörterbuch“ zum Maßstab, annähernd 500.000 Wörter. Allerdings sind die Sprachen flexibel und werden zudem durch die Übernahme von Fremdwörtern regelmäßig erweitert. Sollte es sich einmal als nötig erweisen, werden durch Neubildung oder Import die notwendigen Ergänzungen vorgenommen, zugleich geraten Wörter außer Gebrauch. Auch benötigt kein Deutscher oder Engländer 500.000 Wörter. Ein aktiver Wortschatz von ca. 10.000 Wörtern reicht vollständig. Cäsars Werk „De Bello Gallico“, immerhin als Werk der Weltliteratur anerkannt, kam mit einem Thesaurus von knapp 1.500 Wörtern hin. Charles K. Ogden's „Basic English“ von 1930 umfasst gar nur 850 Wörter und ermöglicht dennoch Verständigung.

Natürliche Sprachen sind Veränderungen unterworfen. Jede neue soziale Praxis verlangt über kurz oder lang nach einem neuen Begriff. So hätte man vor dreißig Jahren das Wort „simsen“ wohl nicht verstanden oder vielleicht als einen trivialen Verschreiber (vielleicht statt „bimsen“) interpretiert. Erst die Kommunikationspraxis, eine SMS zu schreiben,⁹ verlangte nach einem Verb. Die ungefilterte Übernahme des Kürzels SMS hätte – und wurde anfangs auch – in der Lautung holperig in das Wort „es-em-essen“ übertragen werden können. Doch ausspracheökonomisch bedingt verlangen natürliche Sprachen den Wechsel von Konsonanten und Vokalen. Das Verb zum Kürzel hätte durchaus auch „samsen“, „semsen“ oder „somsen“ lauten können. Die umfangreiche Literatur zu Wortbildung und -schöpfung muss hier nicht einmal ansatzweise aufgegriffen werden. Es reicht zum Zweck der Abgrenzung verschiedener Informationsbegriffe hin, auf die Existenz der Restriktionen und Mechanismen – vom Wechsel primärer Zeichen (Vokale und Konsonanten in den Silben) bis zu Suffixen, Präfixen und Zusammensetzungen – zu verweisen.

Analog zum informationstheoretisch begründeten Design des idealen Informationsspeichers bzw. Übertragungskanal kann man formulieren: Die Sprache muss so beschaffen sein, dass sie mit einem beschränkten Satz an primären Zeichen (Ziffern, Buchstaben etc.) und einem ebenso begrenzten Satz an Konstruktionsregeln alle Wörter, gegenwärtige wie zukünftige, ermöglicht. Analog zum Axiom der Generativen Transformationsgrammatik, dass sich mit einer beschränkten Zahl von Regeln eine

könnte sich jedoch um ein Versehen handeln, vgl.: „A nonredundant language would be a more efficient means of communication but totally insensitive to transmission errors of any kind“ (Krippendorff 1986: 28).

9 Die deutsche Bezeichnung ist genau besehen falsch. Mit „Short Message System“ wird das technische Kommunikationsmedium bezeichnet, die einzelne Nachricht müsste korrekt SM abgekürzt werden.

unendliche Zahl von verständlichen Aussagen erzeugen (generieren) lasse, ist mit einem konkreten Bestand von primären Zeichen und einer überschaubaren Zahl von Kombinationsmöglichkeiten eine hinreichend große (nur durch Raum und Zeit beschränkte) Zahl von Wörtern zu bilden.¹⁰ Wortschöpfung und -bildung vermehren also einerseits die Ausdrucksmöglichkeiten, schränken andererseits aber den informationstheoretisch unendlichen Möglichkeitsraum ein: Die Regeln schaffen Ordnung, indem sie zu dem Zeichenvorrat in Beziehung gesetzt werden.¹¹ Mithin kann der Wortschatz einer natürlichen Sprache in Abwandlung von Term A hergeleitet werden als:

$$Inf_{(sem)} = \sum_i^n Z_{(1. \text{ Ordnung})} / \text{semantische Regeln (A')}$$

Nun mag man einwenden, dass natürliche Sprachen nicht allein semantisch operieren, sondern die Syntax und pragmatische Aushandlungsprozesse involviert sind. Die Lexikalisierung schafft Ordnung und bildet aus den Zeichen 1. Ordnung die Symbole 2. Ordnung (Wörter, Zahlen), und ihre Verkettung erzeugt Sätze. Die Regeln einer natürlichen Sprache sind dabei nur eine Möglichkeit von vielen, Ordnung zu erzeugen. Jede natürliche Sprache erzeugt Redundanz, deren Quantität zwar immer ähnlich groß ist, deren Ordnungen sich hingegen unterscheiden. Auch sind Regelwerke denkbar, die Ordnung erzeugen, ohne (bislang) zu einer natürlichen Sprache zu gehören.

Zwar besitzen natürliche Sprachen mit ihren Wort- und Satzbildungsregeln Elemente, die die statistische Berechnung der Sprachstruktur erleichtern. Die Regelmäßigkeiten schaffen eine notwendige Voraussetzung für Verständlichkeit: Z. B. bestehen im Deutschen die Silben zumeist aus Konsonanten und Vokalen. Mehr als fünf Konsonanten folgen in einer Silbe nie aufeinander, mehr als drei Vokale auch nicht. Mit einer gewissen Wahrscheinlichkeit reiht sich somit an einen Konsonanten ein Vokal und umgekehrt. Aber welcher konkrete Buchstabe folgt, lässt sich nicht sicher prognostizieren: „Der Zwe?? ?ei?igt d?? ?ittel“ würde sicherlich von jedermann sofort zu „Der Zweck heiligt die Mittel“ ergänzt. Doch denkbar wäre auch: „Der Zwerg reinigt den Kittel“.¹² Das Argument gilt abgewandelt auch für ganze Sätze. Jede natürliche Sprache kennt sogenannte Normalsätze, bestehend aus Subjekt, Prädikat und Objekt. Somit darf man mit einiger Wahrscheinlichkeit, da nicht alle Sätze Normalsätze sind, nach Subjekt und Prädikat das Objekt erwarten.¹³

Auf der semantischen Ebene steigt mithin die Wahrscheinlichkeit der folgenden Buchstaben an, wenn einige voraussetzende primäre Zeichen schon bekannt sind; auf der syntaktischen Ebene gilt Gleiches, wenn die vorigen sekundären Zeichen feststehen. Beides sind Markov-Ketten: Eine zeit-räumlich vorher getroffene Entscheidung schränkt die Möglichkeit für die unmittelbar folgenden ein (Markov 2006 [1913]). Al-

10 Das könnte sogar eine unendlich große Zahl sein, allerdings nur, wenn man – wie Alan Turing in seinem Schlüsseltext „on computable numbers“ – mit dem Gedanken an einen in Raum und Zeit unendlichen Symbolmanipulator spielt. In der Praxis kann allerdings von „unendlich“ keine Rede sein, weil beides begrenzt ist; vgl. Turing 1936, 1937.

11 Dafür kommt weder Addition noch Subtraktion noch Multiplikation, sondern nur die Division in Betracht. Die Addition scheidet aus, weil die maximale Information nicht vermehrt werden kann. Außerdem beschreiben grammatische Regeln und (Informations-)Zeichen unterschiedliche Ebenen des Sprachproblems. Wegen des Ebenenproblems (Äpfel mit Birnen) scheidet auch die Subtraktion als Operation aus. Für die Multiplikation gilt wiederum der 1. Einwand gegen die Addition. Mithin verbleibt die Division, das in allen ähnlichen Fällen übliche Verfahren, Relationen anzuzeigen.

12 Den Scherz erzählte mir Hans-Peter Ecker.

13 Realiter ist es komplexer. Für diese Argumentation reichen die Vereinfachungen gleichwohl.

lerdings hilft aufgrund der „Gedächtnislosigkeit“ der Ketten die Kenntnis von weiter oben stehenden Zeichen nicht, sondern nur die der Zeichen in unmittelbar vorausgehender Stellung zu dem zu prognostizierenden Zeichen. Zudem bleibt das Problem des Wort- oder Satzauftakts, der Startwahrscheinlichkeit, unberührt. Statistische Berechenbarkeit von semantisch-syntaktischen Strukturen ist in Grenzen möglich, sie setzt aber immer Wissen über den Inhalt der Mitteilung voraus. Erwartbarkeitsberechnungen zu menschlicher Kommunikation, die diese Einschränkungen außer Acht lassen, sind daher zumindest missverständlich (Schulz 2009: 177-182).

Die Berechenbarkeit betrifft nur die Häufigkeit der primären und sekundären Zeichen. Deskriptiv lässt sich die Struktur von Mitteilungen erfassen und in den beschriebenen Grenzen vorhersagen. Aus der Struktur von Mitteilungen lässt sich auf vieles schließen, z. B. auf die Sprache eines verschlüsselten Textes. Mit semantisch-syntaktischen Strukturanalysen lassen sich sogar anonyme Autoren enttarnen, wie der Newsweek-Journalist Joe Klein erfahren musste, der zunächst als Anonymus einen Schlüsselroman über den 1. Präsidentschaftswahlkampf Bill Clintons geschrieben hatte (Anonymus; i. e. Klein 1996; vgl. Diederich et al. 2003). Doch hat die statistische Beschreibung und Analyse mehr oder weniger umfangreicher Korpora mit der Berechenbarkeit oder Erwartbarkeit von Bedeutung wenig zu tun. Seit Längerem wird zwar im Kontext von Semantic-Web- bis hin zu KI-Forschung an probabilistischen Verfahren gearbeitet, z. B. um über kontextorientierte Suchalgorithmen sog. Entitätenidentitäten zu ermitteln, aber selbst das gelingt erst mit Zuordnungswahrscheinlichkeiten von bis zu ca. 90 Prozent.¹⁴

Das semantisch-syntaktische Problem lässt sich mit Gleichung B veranschaulichen, in der $Inf_{(syn)}$ als Teilmenge in $Inf_{(sem)}$ enthalten ist – d. h. $Inf_{(syn)} \subseteq Inf_{(sem)}$. Sie berücksichtigt sowohl das semantische als auch das syntaktische Regelwerk. Wiederum soll analog zu den bisherigen Informationsbegriffen mit $Inf_{(syn)}$ die maximale Menge jener Sätze angezeigt werden, die dem Regelwerk entsprechen:

$$Inf_{(syn)} = \sum_i^n Z_{(1. \text{ Ordnung})} / \text{semantisch-syntaktisches Regelwerk (B)}$$

Doch diese Menge besteht nicht nur aus sinnvollen Sätzen – z. B. „Der Baum läuft keine Hunde“. Mithin ist die Konstruktion von Bedeutung erst teilweise umrissen; auch ihre weitergehende mathematische Rekonstruktion scheint noch nicht möglich.

4. Information pragmatisch

Das Bedeutungsproblem lässt sich weder allein semantisch noch semantisch-syntaktisch lösen, denn die aus theoretischen Überlegungen heraus notwendige strikte Trennung von Semantik, Syntaktik und Pragmatik läuft der Bedeutungskonstruktion in natürlichen Sprachen zuwider. In der Sprachwissenschaft gibt es darum sowohl Versuche, die Semantik zu pragmatisieren als auch umgekehrt die Pragmatik zu semantisieren, denn die Bedeutung von sprachlichen Ausdrücken wird ausgehandelt. Mithin bedarf die Grenzziehung zwischen Semantik und Pragmatik der Klärung im konkreten Einzelfall (Glück 2010: 601-603). Das Wort „Gerechtigkeit“ macht das Problem anschaulich. Es erscheint zunächst selbst ohne weiteren Kontext als ein allgemein ver-

14 Damit ist die Identifikation der konkreten Bedeutung sekundärer Symbole im jeweiligen Kontext gemeint, eine Aufgabe, die Kinder im Spiel „Teekesselchen“ erlernen: Wann ist z. B. mit dem Wort „Horst“ ein Vorname, wann ein Adlernes gemeint? Vgl. Reichartz et al. 2009.

ständliches, positiv belegtes Wort. Wie aber steht es um die Bedeutung von sozialer, politischer, juristischer, höherer oder einer sonstigen Gerechtigkeit? Das Symbol 2. Ordnung *Gerechtigkeit* ist schon ohne sein jeweiliges Attribut eine Projektionsfläche für soziale Konventionen, die sich ohne den (pragmatischen) Kontext nicht näher bestimmen lassen. Oder anders: Alle können den Begriff verwenden, ohne dass sie das Gleiche darunter verstehen müssen. Die Attribute scheinen das Wort *Gerechtigkeit* zwar einerseits zu präzisieren, erzeugen jedoch andererseits selbst Verständigungsbedarf und erhöhen damit den pragmatischen Aufwand. Unberührt davon unterstellen die Menschen den Anderen jedoch, dass diese die signifikanten Symbole – wie sie George Herbert Mead (1968) bezeichnete – jeweils so auffassen werden, wie man sie selbst auffassen würde.

Mit der Pragmatik wird das eigentliche Feld sozialer Bedeutungskonstruktion betreten. Dabei gilt wiederum eine Teilmengenbeziehung: $Inf_{(prag)}$ ist als Teilmenge in $Inf_{(syn)}$ enthalten, d. h. $Inf_{(prag)} \subseteq Inf_{(syn)}$. Analog zu den bisherigen formalen Beschreibungen können wir setzen:

$$Inf_{(prag)} = \frac{\sum_i^n Z(1. \text{ Ordnung})}{\text{semantische} * \text{syntaktische Regeln} * \text{pragmatische Aushandlungen}} \quad (C)$$

Unter $Inf_{(prag)}$ werden Mitteilungen gefasst, auf die sich die jeweils beteiligten Kommunikationspartner verständigen können, aber das müssen sie auch wollen. Damit ist – wiederum in Anlehnung an die bisherigen Informationsbegriffe – die Gesamtmenge aller möglichen verständlichen Mitteilungen bezeichnet. Dabei wird bewusst zwischen semantischen und syntaktischen Regeln einerseits und pragmatischen Aushandlungen andererseits unterschieden. Denn das Regelwerk auf dem Gebiet natürlicher Sprachen ist erlernbar. Es steht überdies vergleichsweise festgefügt und ist nur allmählichem Wandel unterworfen. Die grundlegenden Routinen des sozialen Handelns werden zwar mit der Sozialisation erworben. Die Sprachpragmatik, als Sonderfall sozialen Handelns, ist dennoch nur begrenzt zu erlernen, da sich die Aushandlungsprozesse je nach Kontext und Intention *ad hoc* verändern. Man kann die bisherigen Eingrenzungen des Informationsbegriffs als verschachtelte Teilmengen veranschaulichen:

$$\text{Mitteilung} = Inf_{(prag)} \subseteq Inf_{(syn)} \subseteq Inf_{(sem)} \subseteq Inf_{(inf)} = h \quad (D)$$

D. h., sinnvolle und verständliche Mitteilungen sind eine Teilmenge semantisch-syntaktischer Informationen, diese eine Teilmenge semantischer Informationen und diese wiederum eine Teilmenge informationstheoretischer Informationen. Mitteilungen oder $Inf_{(prag)}$ bestehen aus sinnvollen Sätzen, ein sinnvoller Satz aus der Menge der $Inf_{(syn)}$ besteht aus sinnvollen Kombinationen von Symbolen 2. Ordnung (= Wörtern) aus der Menge der $Inf_{(sem)}$, diese wiederum bestehen aus sprachpraktisch sinnvollen Kombinationen von Zeichen 1. Ordnung (= Primärzeichen). Es ist offensichtlich, dass das Problem umso weniger trivial ist, je stärker der linke Rand der Gleichung in den Blick genommen wird. Erst Mitteilungen = $Inf_{(prag)}$ ermöglichen Humankommunikation, die immer umfangreicheren Obermengen sind nur die jeweils notwendige Voraussetzung für die Informationen der nächsten Teilmenge. Keine der Eigenschaften der Obermengen ist für sich in der Lage, die Voraussetzungen von Humankommunikation hinreichend zu klären. Oder anders: Von rechts ($Inf_{(inf)}$) nach links ($Inf_{(prag)}$) wird Information immer mehr mit Sinn aufgeladen, und je „sinnvoller“ sie wird, desto stärker ist ihr Möglichkeitsraum eingegrenzt. Dabei verschleiert diese Teilmengenbe-

trachtung den Blick dafür, dass an zwei Scharnierstellen wesentliche Veränderungen im Abstraktions- und folglich Verständlichkeitsniveau auftreten:

Die 1. Niveauänderung tritt an der Schnittstelle $Inf_{(inf)}$ zu $Inf_{(sem)}$ und $Inf_{(syn)}$ auf. Schon das Berechenbarkeitsproblem zeigt es an. Zudem mangelt es den Primärsymbolen der Menge $Inf_{(inf)}$ an der doppelten Repräsentanz. Sekundärsymbole hingegen besitzen sie, oder anders: Es macht keinen Sinn, bei primären Zeichen zwischen Signifikant und Signifikat zu unterscheiden, bei sekundären Symbolen, insbesondere Wörtern, allerdings durchaus: Das Zeichen und sein Bezugspunkt in der Welt sind nicht identisch; die doppelte Repräsentation muss, wie im semiotischen Dreieck üblich, durch den Interpreten (entweder als Person oder als Abstraktum) hergestellt werden (Morris 1971: 21-54; vgl. Bühler 1934: 73-75). Die Niveauänderung lässt sich an einem Beispiel verdeutlichen: „Gmäß eneir Sutide eneir elgnihcesn Uvinisterät, ist es nchit witihcg in wlecehr Rneflogheie die Bstachuebn in eneim Wrot snid, das ezniige was wethiig ist, ist daß der estre und der letzte Bstachue an der rithicegn Pstioion snid. Der Rset knan ein ttoaelr Bsinöldn sien, tedztorm knan man ihn onhe Pemoblre lseen. Das ist so, wiel wir nicht jeedn Bstachuebn enzelin leesn, snderon das Wrot als geas-tems“ (Schnabel 2006; vgl. Rayner 2006).

Da Wort für Wort und nicht Buchstabe für Buchstabe gelesen wird und zudem das Regelwerk des Satzbaus erlernt wurde, kann auch der Sinn der Buchstabensuppe erfasst werden. Im Umkehrschluss heißt das, primäre Zeichen sind redundant in den Wort- und sekundäre in den Satzstrukturen abgelegt. Dabei sind syntaktische Regeln in Hochsprachen einerseits durch Konjunktionen, Präpositionen und andere Funktionswörter markiert, andererseits durch Konjugationen, Deklinationen und vergleichbare Beugungen der Wörter selbst. Semantik, Syntaktik und Pragmatik erzeugen also eine Redundanz höherer Ordnung, die wiederum der Verständlichkeit dient. Informationstheoretisch betrachtet liegt der neue Redundanzwert jedoch nur noch geringfügig höher als jene 50 Prozent, die hinsichtlich der Kettenlängen von bis zu acht Zeichen festgestellt wurden. Somit lassen sich die semantisch-syntaktisch-pragmatischen Regeln zwar *auch* als *Redundanz*_(1. Ordnung) berechnen. Doch die resultierende informationstheoretische Redundanz, entstanden als ein Produkt aus dem Prozess semantisch-syntaktischen Regelwerks in Kombination mit den pragmatischen Aushandlungen, ändert sich kaum, obwohl der Verständnissgewinn durch die *Redundanz*_(2. Ordnung) groß ist.¹⁵ Wenn also konsequenterweise gilt

semantisch-syntaktisch-pragmatische Regeln $Redundanz_{(1. Ordnung)} + Redundanz_{(2. Ordnung)}$,

dann kann die Gleichung C auch ausgedrückt werden als:

$$Inf_{(prag)} \cong Inf_{(inf)} - Redundanz_{(1. Ordnung)} - Redundanz_{(2. Ordnung)} (C')$$

Die 2. Niveauänderung ist der Schnittstelle von $Inf_{(syn)}$ zu Mitteilungen (= $Inf_{(prag)}$) zuzuordnen, denn die regelgeleitete, sinnvolle Kombination von Primärzeichen zu Wörtern und von diesen zu Sätzen bedarf einer intendierten sozialen Praxis, für die sich in der Sprechakththeorie, der Theorie des kommunikativen Handelns und anderen Theoriebildungen ein eigenes Wissensgebiet entwickelt hat. Die Sprechakththeorien in ihren Varianten von John L. Austin bis John R. Searle und die darauf basierende Theorie des kommunikativen Handelns von Jürgen Habermas zeigen gerade in ihrem Mangel an Einvernehmlichkeit, wie voraussetzungsreich menschliche Kommunikation letztlich ist.

15 Die Sprachwissenschaft spricht hier von grammatischer Redundanz (Glück 2010: 552).

5. Zur Kommunikationspraxis: Information stillt Neugier, Redundanz macht anschlussfähig

Nun gibt es aber mit den Wiederholungen von Altbekanntem noch eine weitere Redundanz, die als *Redundanz*_(3. Ordnung) eingeführt werden soll. Über die drei Redundanzen hinweg lässt sich eine abnehmende messtechnische und eine zunehmende kommunikative Offensichtlichkeit konstatieren: Die *Redundanz*_(1. Ordnung) ist nur messtechnisch zu erfassen, in der sprachlichen Kommunikation selbst fällt sie schon deshalb nicht auf, weil sie der Kommunikation inhärent ist. Die im komplexen Regel- und Handlungswerk abgelegte *Redundanz*_(2. Ordnung) erzeugt eine messtechnisch nur geringe Differenz zur *Redundanz*_(1. Ordnung), ist aber den Sprachbewussten durchaus deutlich. Bei unreflektiertem Sprachgebrauch (hinsichtlich der Sprache selbst) wird diese Redundanz wohl kaum bemerkt. Wiederholungen, Geschwätzigkeiten und andere *Redundanzen*_(3. Ordnung) können zwar messtechnisch enorme Ausmaße annehmen.¹⁶ Sie müssen aber keine messbare Differenz erzeugen, denn ihre Wiederholungen können auch Pleonasmen oder Tautologien sein: In der Aussage „ein weißer Schimmel“ ist zwar das Attribut überflüssig, informationstheoretisch ist es hingegen nicht redundant. Werden die *Redundanzen*_(3. Ordnung) jedoch übertrieben, so werden sie selbst von jenen bemerkt, die im jeweiligen Moment nicht über Sprache als solche nachdenken. Bei „normalen“ Texten lässt sich von geringerer oder größerer messtechnischer Redundanz nicht auf geringere oder größere Neuigkeits- bzw. Überraschungswerte schließen, wie sich an einigen exemplarisch durchgeführten Redundanzberechnungen leicht zeigen lässt. Man muss sogar eher mutmaßen, dass sprachmächtige, poetische Werke eine geringere messtechnische Redundanz aufweisen als wissenschaftliche und journalistische Texte.¹⁷

Mit der *Redundanz*_(3. Ordnung) wird – *ex negativo* – der Kern des kommunikationswissenschaftlichen Informationsbegriffs umschrieben, den Harry Pross als „Korrelat von Unkenntnis“ definierte. Die Unkenntnis kann nur über Rezipienten bzw. Empfänger von Informationen definiert werden, und zwar doppelt: Zum einen existiert sie deshalb, weil ein aktuelles Ereignis noch nicht bekannt ist. Dann beseitigt die erste Nachricht beim Empfänger dessen Nichtwissen über diesen einen Tatbestand – z. B. über das Ableben einer prominenten Person. Schon die zweite Nachrichtenerwähnung fügt jedoch als *Redundanz*_(3. Ordnung) keinen weiteren Informationsgewinn hinzu.

16 Die messtechnische Unauffälligkeit gilt nur solange, wie die Redundanz_(3. Ordnung) nicht bis zum Exzess endloser Wiederholung getrieben wird. Wenn z. B. das berühmte Gedicht von Gertrude Stein „A rose is a rose“ immer wieder aneinandergereiht wird, führt das zu fast maximaler informationstechnischer Entropie: Die Kompressionsrate des Gedichts – auf 1,5 Megabyte Umfang aufgeblasen – beträgt mit dem ZIP-Programm dieses Rechners schon 99,997 Prozent, die tatsächliche Redundanz hingegen beläuft sich auf 99,999994 Prozent, d. h. ein optimales Kompressionsprogramm müsste diesen Wert erzielen, da der Informationsinhalt nur noch 0,000006 Prozent der Textlänge beträgt.

17 Mit folgenden Maßnahmen lassen sich beliebige Texte auf ihre (ungefähre) messtechnische Redundanz prüfen: Sie müssen als Volltext ohne Grafiken und andere Abbildungen sowie ohne Formatierungsmerkmale abgespeichert werden. Den ASCII- oder ANSI-Text zippt man mit einem handelsüblichen Kompressionsprogramm. Die gezippten Texte belegen in der Regel weniger als die Hälfte des ursprünglichen Speicherplatzes (die entsprechend feine Clustergröße der Festplattenformatierung vorausgesetzt). Beispielsweise hat dieser Aufsatz eine Kompressionsrate von ca. 62 Prozent. Ausgewählte Texte weisen folgende Durchschnittsredundanzen (1. und 2. Ordnung) bzw. bezogen auf einzelne Kapitel Minimal- und Maximalredundanzen auf:

Messtechnisch allerdings bleibt die Redundanz der gestern aktuellen journalistischen Texte (z. B. in Fußnote 17) heute und morgen exakt die gleiche. Zum anderen konstituiert sich Unkenntnis bezogen auf einen grundsätzlich längst bekannten Tatbestand – z. B. dass Cäsar an den Iden des März 44 v. Chr. ermordet wurde. Wer das nicht wusste bzw. vergessen hat, kann es in den entsprechenden Informationsträgern (Lexika etc.) nachschlagen: Man liest nach, informiert sich und beseitigt die eigene Unkenntnis. Nichtwissen Dritter wird allerdings nicht gemindert; das geschieht nur durch individuelle Rezeption der entsprechenden Information und reicht so weit, bis man durch Vergesslichkeit wieder in den Status der Unkenntnis zurückgesetzt wird. Information, in diesem engen kommunikationswissenschaftlichen Sinn verstanden, stillt also Neugier.

Die *Redundanz*^(3. Ordnung) ist hierzu der Gegenbegriff. Sie besteht in der Wiederholung ganzer Aussagen bzw. ihrer Aussageteile. Man könnte *Redundanz*^(3. Ordnung) als überflüssig ignorieren, da sie keine neuen Informationen hinzufügt. Doch sie kann, wie jeder in der Lehre Tätige weiß, durchaus der Allgemeinverständlichkeit dienen. Sie hebt Wichtiges hervor, macht Kompliziertes verstehbarer, steuert Aufmerksamkeit und wirkt schleichendem Informationsverlust entgegen.¹⁸ Generell ist sie, wie die *Redundanz*^(1. Ordnung) und die *Redundanz*^(2. Ordnung), Voraussetzung der Verständlichkeit. Redundante Mitteilungen werden jedoch dann dysfunktional, wenn sie als Geschwätzigkeit die Kommunikation aufblähen, erschweren und den mit dem Geschwätz Traktierten dazu bewegen, die langweilig werdende Kommunikation ganz ab-

	Redundanzrate (bezogen auf den Gesamtkorpus)	Minimale Redundanz (bezogen auf Teilkorpora)	Maximale Redundanz (bezogen auf Teilkorpora)
<i>journalistische Texte</i>			
Nachrichten „Die Welt“ online (20.12.2010)	63	–	–
<i>eigene wissenschaftliche Texte</i>			
dieser Aufsatz	62	–	–
Auftakt Einführung	56	–	–
Einführung insgesamt	65	–	–
Habilitation	63	–	–
Dissertation	65	–	–
<i>Literatur</i>			
Goethe: Faust I	–	50	57
Goethe: Faust II	–	49	51
Goethe: Werther	–	51	60
Fontane: Stechlin	–	48	56
Fontane: Schach	60	–	–
Stein: A Rose 1,5MB	99,997	–	99,999994

Selbstironisch könnte ich mir zugutehalten, zwischen Dissertation und Habilitation immerhin einen kleinen Redundanzfortschritt erzielt zu haben und mit diesem Aufsatz einen weiteren.

- 18 Wichtiges kann selbstredend auch ohne Redundanzen markiert werden, z. B. über meta-kommunikative Hinweise wie: „Was nun folgt, ist wichtig!“

zuberechnen. Ob im positiven ersten Fall oder im schlechten zweiten: Die *Redundanz*_(3. Ordnung) liefert keinen Neuigkeitswert, beseitigt kein Unwissen und stillt keine Neugier; sie ist daher keine Information im kommunikationswissenschaftlichen Sinn. Information im kommunikationswissenschaftlichen Verständnis lässt sich ergo abbilden:

$$Inf_{(kowi)} = \frac{\sum_i^n Z(1. \text{ Ordnung})}{\text{semantische} * \text{syntaktische Regeln} * \text{pragmatische Aushandlungen}} - \text{Redundanz}_{(3. \text{ Ordnung})} \text{ (E)}$$

oder

$$Inf_{(kowi)} \cong Inf_{(inf)} - \text{Redundanz}_{(1. \text{ Ordnung})} - \text{Redundanz}_{(2. \text{ Ordnung})} - \text{Redundanz}_{(3. \text{ Ordnung})} \text{ (E')}$$

Information nach diesem Verständnis erzeugt also einerseits Ordnung und reduziert andererseits Unkenntnis oder Ungewissheit. Das klingt verdächtig nach Niklas Luhmann, der Kommunikation als Dreifach-Selektion von Information, Mitteilung und Verstehen definierte. Allerdings weist der Luhmannsche Informationsbegriff einige Unschärfen auf (Luhmann 1994: 102f., 203). Zunächst einmal scheinen sich Luhmanns und der informationstheoretische Informationsbegriff als jeweils größte Obermenge weitgehend zu entsprechen. Doch sprach Luhmann bei der Primärauswahl nicht nur von Information als Reduktion der Entropie; hätte er das getan und wäre er dabei geblieben, müsste man seine *Information* mit $Inf_{(inf)}$ übersetzen. Allerdings hat er dabei die *Redundanz*_(1. Ordnung), die natürlichen Sprachen auf semantischer Ebene zu eigen ist, nicht berücksichtigt. Zugleich machte er weitere Einschränkungen: Das Gegenteil von Entropie übersetzte er mit Sinn. Wäre er hierbei geblieben, müsste man seine *Information* mit $Inf_{(prag)}$ gleichsetzen. Hierbei jedoch hat er die *Redundanz*_(2. Ordnung), die im grammatischen Regelwerk abgelegt ist, ignoriert. Und um die Begriffsverwirrung komplett zu machen, forderte Luhmann von einer Information den Neuigkeitswert ein. Hätte er diese dritte Begriffsverwendung durchgehalten, wäre seine *Information* gleich $Inf_{(kowi)}$. Das berücksichtigt zwar die *Redundanz*_(3. Ordnung), aber die Einschränkung hat zwangsläufig zur Folge, dass Kommunikation nur Neuigkeiten prozessieren könnte, denn die folgenden Selektionen von Mitteilung und Verstehen wählen ja nur noch aus dem Pool der Neuigkeiten. Zudem können Kommunikatoren gar nicht wissen, was für die Rezipienten neu ist, sie vermuten es nur.

Weniger unscharf, aber dennoch doppeldeutig ist auch Luhmanns Kommunikationsbegriff. Aus seinem Diktum „Kommunikation ist unwahrscheinlich“ wird nach den bisherigen Erörterungen: Die konkrete Kommunikation ist unwahrscheinlich, nicht jedoch Kommunikation an sich. Luhmann setzt sogar hinzu: „obwohl wir sie jeden Tag erleben, praktizieren und ohne sie nicht leben würden“ (Luhmann 1981: 26). Somit kann er Kommunikation an sich als Grundvoraussetzung der Vergesellschaftung, als „kleinstmögliche Einheit eines sozialen Systems“ betrachten (Luhmann 1998: 82). Der wichtigste Unsicherheitsfaktor für gelungene Kommunikation ist hier wie da ähnlich: Luhmanns „Verstehen“ durch *Ego* bzw. die pragmatische Aushandlung zwischen den Kommunikationspartnern. Kommunikation an sich ist aber nicht nur nicht unwahrscheinlich, sondern sie ist ein konstituierendes Element der Gesellschaft.

Die Präzisierung, dass das spezifische Kommunikat und nicht Kommunikation als solche unwahrscheinlich ist, lässt sich mit Blick auf das – je nach Theorie – System oder Feld der Wissenschaft noch zuspitzen. Man könnte zunächst annehmen, dass originelle Kommunikation die unwahrscheinlichste Kommunikation darstellt. Luhmann hätte für das Wissenschaftssystem feststellen können: Originell ist unwahrscheinlich. In der Wissenschaft ist formale Originalität und inhaltlich Originelles dennoch ein Er-

fordernis, beides wird gewünscht und verinnerlicht. Selbst wenn die Handlungspraxis nur zu bemühter Originalität (\neq Originelles) führt, spätestens von Gutachterseite wird beides vorausgesetzt oder, wenn fehlend, angemahnt. D. h., das Bemühen um Originalität und Originelles konstituiert das wissenschaftliche Feld oder System. Aber vielleicht liegt ja hier der Schlüssel: Wer die Wissenschaft als System betrachtet, kann zu der Schlussfolgerung kommen, Originelles sei selten. Wird die Wissenschaft hingegen handlungstheoretisch als Feld aufgefasst, ist Originalität die das Feld konstituierende Praxis. Das Beispiel ließe sich mühelos auf andere Systeme/Felder, z. B. den Kunst- oder Literaturbetrieb, übertragen.

Verallgemeinert und zurückgeführt auf Kommunikation: Als *Kommunikation an sich* ist sie regeldominiert, intentional und als Aushandlungsprozess zugleich Grundlage der Vergesellschaftung. Kommunikation ist nicht Ausnahme, sondern die Regel. Aber wird damit das 1. Axiom der Palo-Alto-Schule automatisch richtig: „Man kann nicht nicht kommunizieren“? Betrachtet man genauer, was Watzlawick, Beavin und Jackson gemeint haben, wird klar, dass sie nicht von Kommunikation im bisher gebrauchten Sinn sprachen, vielmehr setzten sie Kommunikation und Verhalten gleich: „Man kann sich nicht nicht verhalten. Wenn man also akzeptiert, dass alles Verhalten in einer zwischenpersönlichen Situation Mitteilungsscharakter hat, d. h. Kommunikation ist, so folgt daraus, dass man, wie immer man es auch versuchen mag, nicht nicht kommunizieren kann. Handeln oder Nichthandeln, Worte oder Schweigen haben alle Mitteilungsscharakter: Sie beeinflussen andere, und diese anderen können ihrerseits nicht nicht auf diese Kommunikationen reagieren und kommunizieren damit selbst“ (Watzlawick et al. 1972: 51). An Max Webers Definition des sozialen Handelns anschließend könnte man Watzlawick, Beavin und Jackson in das bisherige Schema als weitere Obermenge integrieren und Kommunikation über eine Kaskade von Eingrenzungen bestimmen:

Kommunikation \subseteq *soziales Handeln* \subseteq *Handeln* \subseteq *Verhalten* (F)

Allerdings sollte Kommunikation nicht, wie Watzlawick, Beavin und Jackson postulierten, mit *Verhalten* gleichgesetzt werden, sondern Erstere macht nur als mehrfach eingeschränkter Spezialfall von Letzterem Sinn. Humankommunikation als *kommunikatives Handeln* kann als ein Spezialfall des sozialen Handelns betrachtet werden. *Soziales Handeln* ist eine Sonderform des Handelns analog zu Max Webers klassischer Definition: „Soziales Handeln aber soll ein solches Handeln heißen, welches seinem von dem oder den Handelnden gemeinten Sinn nach auf das Verhalten *anderer* bezogen wird und daran in seinem Ablauf orientiert ist“ (Weber 1980: 1). Soziales Handeln ist auf soziale Gemeinschaft gerichtet und mithin reflexiv; das soziale Handeln des Menschen ist intentional, eher kognitiv als vegetativ, doch gleichermaßen intellektuell wie emotional begründet. *Handeln* wiederum, als Spezialfall des Verhaltens, dient immer einem Zweck: Intentionales folgt dem Zweck bewusst, bei nichtintentionalem Handeln ist das dem Handelnden hingegen nicht bewusst. Aktives *Verhalten* schließlich ist dem Lebendigen vorbehalten.¹⁹

19 Zwar spricht man auch toter Materie Verhalten zu, doch nur in dem passiven Sinn, dass, wenn Material einer physikalischen oder chemischen Belastung ausgesetzt wird, es sich in bestimmter Weise verhält.

6. Schlussfolgerungen

Kommunikation ist demnach der Austausch von und die Verständigung über Bedeutungen, woran mindestens zwei Menschen beteiligt sind. Sie beruht auf der Verwendung signifikanter Symbole, bei der die Kommunikationspartner einen primären und sekundären Symbolvorrat teilen, das notwendige grammatische Regelwerk beherrschen, willens und imstande sind, Bedeutungsverschiebungen *ad hoc* auszuhandeln und gegebenenfalls zu korrigieren. Die Kommunikationspartner setzen voraus, dass ihre Symbole vom jeweils anderen in ähnlicher Weise wie von ihnen selbst benutzt werden. Kommunikation beinhaltet aktive wie reaktive, kognitive wie emotive Komponenten; sie ist ein selbst- und fremdreflexiver sozialer Prozess, Teil des sozialen Handelns und somit immer in Teilen intentional.²⁰ Kommunikation enthält nicht nur Informationen, sondern auch Redundanzen auf mehreren Ebenen, um Verständlichkeit zu erzeugen. Da sie jedoch auch *Redundanzen*_(3. Ordnung) umfasst, um die man den Kommunikationsprozess (theoretisch) entschlacken könnte, sind *Informationen*_(kowi) eine Sonderform von Mitteilungen, die beim Rezipienten Unkenntnis beseitigen und Neugier befriedigen. Als Teilmengen ließe sich das in zwei Dimensionen – der informationsbezogenen und der handlungstheoretischen – wie folgt abbilden:

$$\begin{aligned}
 Inf_{(kowi)} \subseteq Kommunikation \subseteq Mitteilung &= Inf_{(prag)} \subseteq Inf_{(syn)} \subseteq Inf_{(sem)} \subseteq Inf_{(inf)} = h \\
 &\subseteq \\
 &\text{soziales Handeln} \\
 &\subseteq \\
 &\text{Handeln} \\
 &\subseteq \\
 &\text{Verhalten}
 \end{aligned}$$

(G)

Zumindest die Waagerechte enthält (rekursiv) das ganze Thema: An Anfang wie Ende steht Information. Kommunikation kommt die Mittler-, Mittel- und letztlich die Schlüsselposition zu.

Allerdings ist auch das scheinbar Unwichtige wichtig, die Redundanz. Kommunikationspragmatisch kommt Kommunikation keineswegs ohne *Redundanz*_(3. Ordnung) aus. Denn würde sie nur auf die jeweilige Neuigkeit reduziert, würde sie rasch erliegen. Redundanzen ergeben sich aus der Richtung des Zeitpfeils, letztlich also aus den Naturgesetzen, auf denen unsere reale Welt aufbaut. In fundamentalem Sinne folgen sie aus Norbert Wieners (und Claude Shannons) Definition der Information als Zeitfunktion:

Redundanzen (aller drei Ordnungen) schlagen unablässig die Brücke über den Strom der Zeit. Sie verbinden das Soeben mit dem Jetzt und dem Gleich, sie verknüpfen gestern, heute und morgen. Sie sind nicht nur Gegenbegriffe zur jeweiligen Information, sondern als Komplementäre zur Information in graduell unterschiedlichem Maß für den Kommunikationsfluss notwendig. Die *Redundanz*_(1. Ordnung) ermöglicht Informationsaustausch und korrigiert Informationsstörungen, die *Redundanz*_(2. Ordnung) stabilisiert Sprache, die *Redundanz*_(3. Ordnung) hält Kommunikation in Gang, lenkt Aufmerksamkeit und gleicht allmählichen Informationsverlust aus.

20 Intentionalität als Voraussetzung der Kommunikation ist umstritten. Das erscheint nach dieser Argumentation jedoch wenig sinnvoll, da die Sprachpragmatik intentionales soziales Handeln voraussetzt. Und auch die Verstetigung der Kommunikation (Anschlusskommunikation) bedarf der Absicht.

Redundanzen sind mithin nicht überflüssig, sie erfüllen einen wichtigen Zweck: Aktualisierungen, symbolische Kommunikationen, Rituale, Habitualisierung der Nutzung, Routinen der Mediengestaltung, Erinnerung(en) und vieles andere wäre ohne Redundanz, ohne Wiederholungen zur Überbrückung der Zeit, undenkbar. Ohne sie gäbe es keine Anschlusskommunikation; ohne rückbezügliche Verweise, ohne rekursive Schleifen, nur beschränkt auf die Kommunikation der Information zur Reduktion von Unkenntnis, müsste man auf Bestätigung verzichten. Doch ausschließlich redundant lässt sich ebenfalls nicht kommunizieren. Information und Redundanz bedingen sich wechselseitig und ermöglichen die Kommunikation.

Literatur

- Anonymus; i. e. Klein, Joe (1996): *Mit aller Macht* (Primary Colors). Aus dem Englischen von Uda Strätling, München.
- Bühler, Karl (1934): *Sprachtheorie. Die Darstellungsfunktion der Sprache*. Jena.
- Diederich, Joachim; Kindermann, Jörg; Leopold, Edda; Paass, Gerhard (2003): Authorship Attribution with Support Vector Machines. In: *Applied Intelligence*, 19/2003, Nr. 1-2, S. 109-123.
- Glück, Helmut (Hrsg.) (2010): *Metzler Lexikon Sprache*, 4. Aufl. Stuttgart/Weimar.
- Krippendorff, Klaus (1986): *Information Theory. Structural Models for Qualitative Data*. Newbury Park/London/New Delhi (Quantitative Applications in the Social Sciences, Bd. 62).
- Luhmann, Niklas (1981): *Soziologische Aufklärung 3: Soziales System, Gesellschaft, Organisation*. Opladen.
- Luhmann, Niklas (1994): *Soziale Systeme. Grundriß einer allgemeinen Theorie*, 5. Aufl. Frankfurt a. M.
- Luhmann, Niklas (1998): *Die Gesellschaft der Gesellschaft*, 2 Bde., Frankfurt a. M.
- Markov, Andrej Andrejewitsch (2006 [1913]): An Example of Statistical Investigation of the Text Eugene Onegin Concerning the Connection of Samples in Chains. Classical Text in Translation. Lecture at the Physical-Mathematical Faculty, Royal Academy of Sciences, St. Petersburg, 23 January 1913. In: *Science in Context*, 19/2006, Nr. 4, S. 591-600.
- Mead, George H. (1968): *Geist, Identität und Gesellschaft aus der Sicht des Sozialbehaviorismus*. Mit einer Einleitung hrsg. v. Charles W. Morris. Frankfurt a. M.
- Merten, Klaus (1977): *Kommunikation. Eine Begriffs- und Prozeßanalyse*. Opladen (Studien zur Sozialwissenschaft, Bd. 35).
- Morris, Charles (1971): *Writings on the General Theory of Signs*. The Hague/Paris.
- Pross, Harry (1977): *Kommunikationspolitik und neue Medien*. In: Helga Reimann, Horst Reimann (Hrsg.), *Information*. München, S. 21-36.
- Rayner, Keith (2006): Raeding Wrods with jubmeld letters. There is a Cost. In: *Psychological Science*, 17/2006, Nr. 3, S. 192f.
- Reichartz, Frank; Korte, Hannes; Paass, Gerhard (2009): Machine Learning and Knowledge Discovery in Databases. Dependency Tree Kernels for Relation Extraction from Natural Language Text. In: *Lecture Notes in Computer Science*, 5782/2009, S. 270-285.
- Schnabel, Ulrich (2006): Bnuter Bhutsabensalat. In: *Die Zeit*, 9.2.2006, Nr. 7, S. 36, http://www.zeit.de/2006/07/S_36_Kleintext [09.06.2011].
- Schulz, Winfried (2009): *Kommunikationsprozess*. In: Elisabeth Noelle-Neumann, Winfried Schulz, Jürgen Wilke (Hrsg.), *Fischer Lexikon Publizistik Massenkommunikation*, akt., vollst. überarb. u. erg. Aufl. Frankfurt a. M., S. 169-199.
- Shannon, Claude E. (1948): A Mathematical Theory of Communication. In: *The Bell System Technical Journal*, 27/1948, Nr. 3, S. 379-423 und Nr. 4, S. 623-656. Zitiert nach dem korrigierten Reprint in: <http://cm.bell-labs.com/cm/ms/what/shannonday/shannon1948.pdf> [09.06.2011].
- Shannon, Claude E.; Weaver, Warren (1976): *Mathematische Grundlagen der Informationstheorie*. München/Wien (Scientia Nova).
- Turing, Alan M. (1936): On Computable Numbers – with an Application of the Entscheidungsproblem. In: *Proceedings of the London Mathematical Society*, 42/1936, S. 230-265.

- Turing, Alan M. (1937): On Computable Numbers – with an Application of the Entscheidungsproblem. A Correction. In: Proceedings of the London Mathematical Society, 43/1937, S. 544-546.
- Watzlawick, Paul; Beavin, Janet H.; Jackson, Don D. (1972): Menschliche Kommunikation. Formen, Störungen, Paradoxien, 3., unv. Aufl., Bern/Stuttgart/Wien.
- Weber, Max (1980): Wirtschaft und Gesellschaft. Grundriß der verstehenden Soziologie. Studienausgabe, hrsg. v. Johannes Winckelmann, 5. rev. Aufl. Tübingen.
- Wiener, Norbert (1963): Kybernetik. Regelung und Nachrichtenübertragung im Lebewesen und in der Maschine, 2., rev. und erg. Aufl. Düsseldorf/Wien.

Short Cuts | Cross Media



Populäre Musik, mediale Musik?

Transdisziplinäre Beiträge zu den Medien der populären Musik

Herausgegeben von Christofer Jost,
Daniel Klug, Axel Schmidt und
Klaus Neumann-Braun

2011, Band 3, 248 S., brosch., 29,— €

ISBN 978-3-8329-6719-2

Erscheint ca. August 2011

nomos-shop.de/13819

Der transdisziplinär ausgerichtete Sammelband legt den Fokus auf die technisch-medialen Rahmenbedingungen der populären Musik und untersucht den Einsatz

von Medien als Schnittstelle ästhetischer Transformationsprozesse und kultureller Sinnzuschreibungen. Ein Vorgehen, das als Desiderat einer zeitgemäßen Musikforschung erscheint.

Bitte bestellen Sie im Buchhandel oder
versandkostenfrei unter ► www.nomos-shop.de



Nomos