

Zur Psychobiologie des Sprechens als zeitstrukturiertes Verhalten und ihre verhaltenstherapeutische Aussagekraft

H.-P. Krüger

Einleitung

In der Sprache der Verhaltensforscher katalogisiert das Ethogramm das gesamte Verhalten einer Spezies (Tembrock, 1980, 37). Dabei sind drei Dimensionen dieses Ethogramms zu unterscheiden:

- Jedes Verhalten hat eine Funktion, hat meist Bedeutung, ist zielgerichtet (wie etwa die Aggression, das Bindungs- oder Dominanzverhalten). Das *Aktogramm* beschreibt diese Ebene.
- Ein großer Teil des Verhaltens führt aber auch zu einer Veränderung der räumlichen Umwelt: eine Person bewegt sich, eine Geste geht auf den Partner zu. Andererseits führen besondere räumliche Konstellationen zu bestimmtem Verhalten: man denke etwa an die Folgen zu großer physischer Nähe («crowding»). Beschrieben werden diese räumlichen Charakteristika des Verhaltens im *Topogramm*, zu dem Begrifflichkeiten gehören wie Territorialität, Distanz, persönlicher Raum.
- Viel weniger auffällig, aber dennoch in gleicher Konsequenz hat alles Verhalten eine zeitliche Dimension, hat eine Dauer, geschieht in einer Folge. Ein Teil unserer Kommunikationskanäle kann nur sequentiell funktionieren – beim Dialog müssen Sprechen und Zuhören abwechseln. Mit dem *Chronogramm* soll diese zeitliche Dimension gefaßt werden. Einschlägige Begrifflichkeiten sind hier etwa Rhythmus, Dauer, Zyklen.

Die Sozialpsychologie hat sich wie die Humanethologie sehr lange nur mit dem Aktogramm beschäftigt (siehe dazu etwa die Darstellung von Eibl-Eibesfeldt, 1984). Vor allem Hall (zuletzt in 1988) hat mit schlagkräftigen Titeln wie «The silent language» (1959), «The hidden dimension» (1966) oder «The dance of life» (1983) die Raum- und Zeitbedingtheit des Verhaltens in die Wissenschaft eingebracht. Mit den Forschungen zur sozialen Bedeutung von Raum, den «proxemics», tat sich sehr schnell ein potenter Ansatz auf, der neben Grundlagenverständnis auch enorme Auswirkungen in der Anwendung fand, bis hin zur architektonischen Planung von Stadtteilen. Erst in den letzten Jahren richtete sich das Forschungsinteresse auf die zeitliche Dimension, wobei der thematische Bogen von unterschiedlichen zeitlichen Rhythmen in verschiedenen Kulturen (eine Zusammenfassung ist in Davis, 1982, zu finden) bis hin zu Mikrorhythmen physiologischer Reaktionen im Dialog gespannt wurde.

Wie die Motorik eine zentrale Rolle im Topogramm spielt, ist es das Sprechen für das Chronogramm. Verbale Kommunikation kann nur sequentiell ablaufen, aus einem einzelnen Zeitschnitt ist nichts zu verstehen. Erst die Folge von Lauten kreiert die Botschaft, erst die Folge von Äußerungen schafft das Gespräch. Um diesen Tatbestand zu kennzeichnen, führten wir den Begriff «speech chronemics» ein (Krüger, 1989), wobei Bezug genommen wurde auf Poyatos, der mit «chronemics» einen Forschungsgegenstand bezeichnen wollte: «That would deal with our conceptualization and handling of time as a biopsychological and cultural element» (1972, 61).

Zum Material und zur Messung

Die Zeitstruktur des Sprechens zeigt sich erst dann, wenn man aus der Sprache alle Qualitäten herausnimmt. So wird durch die alleinige Betrachtung des Energieverlaufs des Sprechens (d.h. der Mikrofonspannung) der Inhalt des Sprechens eliminiert. Reduziert man über eine Schwelle dieses physikalische Signal weiter in hoch- und niederenergetische Perioden, resultiert eine Sequenz von ON/OFFs, die nur noch die zeitliche Verteilung des Sprechverhaltens repräsentieren.

Bereits um 1901 begann Wallin Versuche zu machen, die zeitliche Struktur der Sprache zu erheben (eine historische Darstellung vor allem der psycholinguistischen Forschungstradition findet sich in Kowal, 1991). Bekannt wurde Chapple mit seinem Interaction Chronograph (beschrieben in Chapple 1948/49), der in der Klinischen Psychologie und Psychiatrie vor allem durch Matarazzo (1973) benutzt wurde. Ein Meilenstein war die Entwicklung des AVTA-Systems durch Cassotta, Feldstein & Jaffe (1964). Neuere Entwicklungen sind in Kowal (1992) beschrieben. Das neueste System ist wohl das von Craig (1991) entwickelte Listener System, das auf dem IBM Speech Viewer basiert. Alle diese Geräte setzen Sprachmaterial voraus, das entweder direkt im Labor gewonnen oder über Tonträger aufgezeichnet wurde. Demgegenüber ist unsere Entwicklung, das Logoport (Krüger, 1989), für die Datenaquisition und -vorverarbeitung im Feld (ambulant) konzipiert.

Das Logoport (10 × 12 × 3 cm) wird in einem Schulterhalter unter der Kleidung getragen. Am Hals wird ein Klebemikrofon angebracht, dessen Signal verstärkt, bandpaßgefiltert, gleichgerichtet und mit 1 kHz A/D-gewandelt wird. Diese Werte werden über eine einstellbare Schwelle in ON/OFF transformiert. Acht dieser Elemente (einstellbar) werden aufsummiert. Sind mehr als ein ON vorhanden (einstellbar), wird das gesamte 8-Millisekunden-Intervall als ON gewertet (digitale Schwelle). Damit liegt am Ende ein 125 Hz schnelles ON/OFF-Signal vor, das in einem 1 Megabyte-Speicher abgelegt wird. Integriert ist in das Logoport ein EKG via Brustwandableitung. Das Gerät wertet diese Signal on-line auf RS-Komplexe aus und legt die Interbeatintervalle in Millisekunden ebenfalls im Speicher ab. Zusätzlich gibt es einen Ereignismarker, dessen Betätigung ebenfalls registriert wird.

Zu Beginn einer Messung wird das Logoport mit einem PC verbunden. Über ein Initialisierungsprogramm werden alle Schwellen und Variablen für die zu untersuchende Person eingestellt. Die interne Uhr wird so programmiert, daß das Logoport zu einstellbaren Zeitpunkten losläuft und endet. Damit ist es auch möglich, mehrere Logoports zu synchronisieren (etwa für Gruppengespräche). Nach dieser Einstellung können sich die Untersuchungspersonen frei bewegen. Am Ende der Messung nimmt die Person selbst das Gerät ab und bringt es zur Auswertung ins Institut (Datenerhalt vor Auslesung über einen Monat).

Einen Ausschnitt aus einer solchen Messung zeigt Abbildung 1, die etwa zwei Minuten eines Gesprächs repräsentiert. Eine Zeile «dauert» etwa 4,5 Sekunden. Erkennlich

sind drei längere Sprechepisoden von etwa 15, 10 und 30 Sekunden Dauer und einer im oberen Drittel vorfindbarer Periode fast vollständigen Schweigens. Jede der drei Sprechepisoden ist selbst wieder unterteilt in zusammengehörige Stücke von einer bis drei Sekunden Dauer, wobei auch diese Stücke wieder weiter unterteilt sind: ein erster Hinweis auf die hierarchische Natur der Zeit.

Die zwei Perspektiven der Zeitstruktur

Prinzipiell gibt es zwei Möglichkeiten, an eine solche Aufzeichnung heranzugehen. Der Sprecher kann als «information processor» oder als «interlocutor» gesehen werden (Levelt, 1989). Im ersten Fall wird die Zeit als Kennzeichen oder indirektes Maß eines

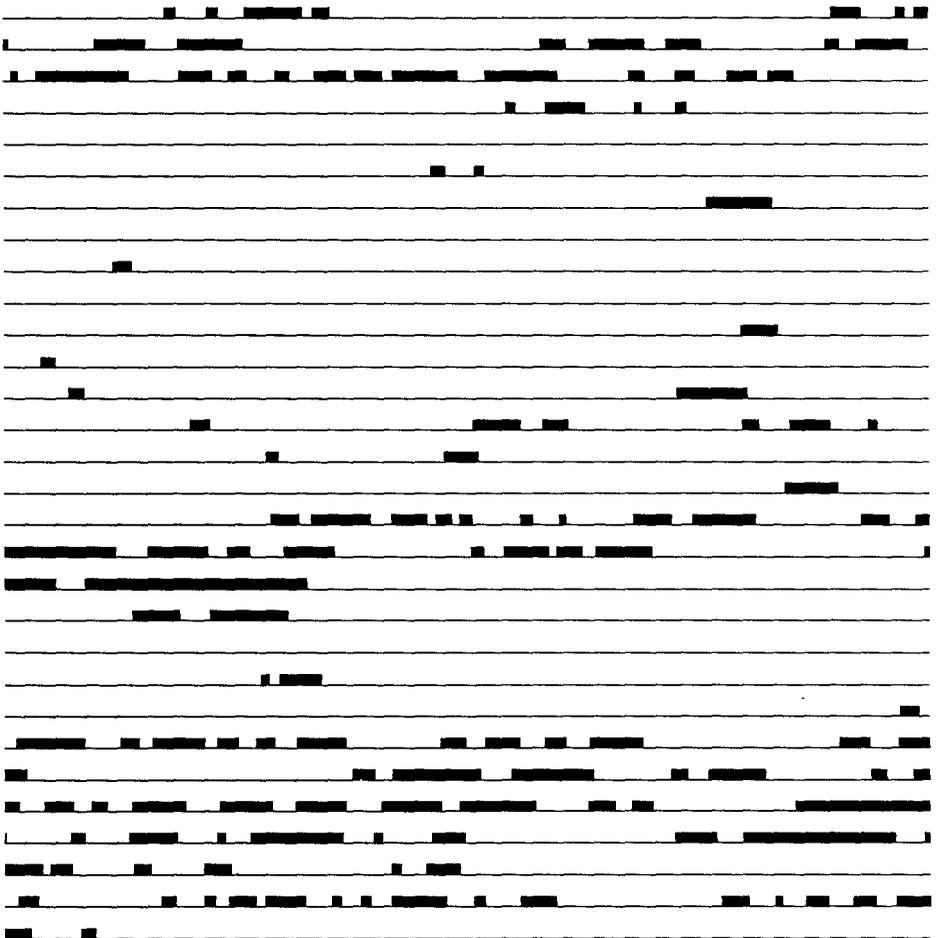


Abbildung 1: Muster von ONs und OFFs mit einer Auflösung von 8 msec. Nach rechts ist die Zeit aufgetragen, ein Balken ist ON, eine Linie OFF. Eine Zeile entspricht ungefähr 4,5 Sekunden.

komplexen Sprachgenerierungs-Prozesses benutzt. In Analogie: zwei unterschiedliche konstruierte Computer haben eine unterschiedliche Zeitstruktur in der Lösung der gleichen Aufgabe. Damit ist aber diese Zeitstruktur ein indirektes Maß für die Architektur der Geräte. In ähnlicher Weise wird etwa die Reaktionszeit in den Verhaltenswissenschaften als Maß für Komplexität gebraucht. Dieser Ansatz hat sich als sehr fruchtbar erwiesen (die meisten Maße der kognitiven Psychologie sind Zeitmaße), benutzt die Zeit aber nur als *akzidentelles*, nicht als konstituierendes Merkmal eines Verhaltens.

Demgegenüber betont der funktionale Zugang die biologische Bedeutung der Zeit. Ob die Zeitstruktur eine Bedeutungsträgerin ist, läßt sich einfach an folgendem Test entscheiden: ändert sich die Botschaft, wenn ihre zeitlichen Parameter (etwa Dauern, Artikulationsgeschwindigkeiten) verändert werden? Wenn ja, hat die Zeitstruktur selbst eine biologische Bedeutung.

Die zwei Domänen der Zeitstruktur

Speech chronemics betrachtet Zeit unter dem biologischen Aspekt der Ko-Evolution von Zeichen. Letzteres bekommt seinen «Sinn» daraus, daß der Sender es in seine einzigartige Form bringt und der Empfänger es «eindeutig» versteht. Dabei müssen offensichtlich im Bereich der Sprache zwei ko-evolutionäre Prozesse stattgefunden haben. Der erste betrifft die Dualität Sprecher/Hörer, der zweite die Dualität Dialogpartner/Dialogpartner.

Alle Möglichkeiten, sich sprachlich auszudrücken, sind nur in dem Maße «sinnvoll», in dem sie auch unser akustisches System und unsere daraus abgeleitete Wahrnehmung aufnehmen können. Es bedarf damit einer Isomorphie sprachlicher und akustischer Strukturen. Dies gilt auch für die zeitlichen Charakteristika der Sprachproduktion, die hier unter dem Begriff der «Prosodie» zusammengefaßt werden. Prosodische Sprachmerkmale sind damit alle zeitlichen Charakteristika des Sprechens, deren Variation durch den Sprecher eine Information für den Hörer bedeutet.

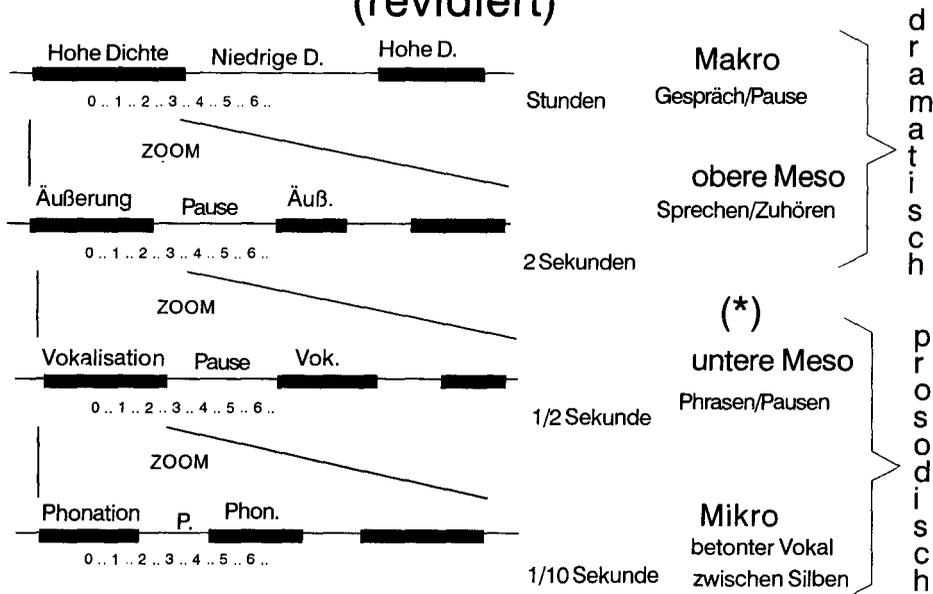
Der zweite Prozeß betrifft die Produktion des Dialogs. Da die Botschaft des Sprechers sich erst in einer (zumindest relativ) ungestörten zeitlichen Sequenz eines Sprechers entwickeln kann, braucht der Dialog eine Steuerung des Wechsels zwischen Hören und Sprechen, die grundsätzlich von der zeitlichen Struktur des monologischen Sprechens unabhängig ist. Diesen Bereich bezeichnen wir als die «Dramaturgie» des Sprechens.

Die hierarchische Natur des Sprechens

Ein Blick auf Abbildung 1 zeigt, daß die ON/OFFs weder gleichmäßig noch zufällig in der Zeit verteilt sind. Vielmehr deuten sich verschiedene Strukturierungsebenen an, die wir als die Jaffe-Hierarchie bezeichnen wollen. Jaffe hat als erster auf diese Tatsache hingewiesen (in Jaffe, Anderson & Stern, 1979), wobei wir seine methodische Definition der Ebenen unter Verwendung neuer Erkenntnisse der Psycholinguistik und der Ergebnisse aus Logoport-Messungen revidierten (Abb. 2).

Bei hoher zeitlicher Auflösung zeigt sich die Sprache (im übrigen in allen Sprachen der Welt nachgewiesen) als eine Folge «stimmhaft – stimmlos – stimmhaft...». Das heißt, hochenergetische Laute wie Vokale wechseln ab mit niederenergetischen wie

Die JAFFE-Hierarchie (revidiert)



(*) Intonationsebene kann hier eingefügt werden

Abbildung 2: Die Jaffe-Hierarchie.

Frikative, Plosive, andere Konsonanten (eine genauere Darstellung findet sich bei Studdert-Kennedy, 1975). Zu den niederenergetischen Lauten sind wohl auch die Artikulationspausen zu zählen, die bei der Bildung von Plosiven entstehen und bei der Spracherkennung eine entscheidende Bedeutung haben: Wie Liberman (zusammenfassend in 1978) gezeigt hat, ist die Dauer der Pause Bedeutungsträger! Levelt (1989) bezeichnet diese Ebene als die «skeletal structure» der Sprache, wir ordnen sie als Mikroebene ein. Um die Ereignisse dieser Ebene adäquat darstellen zu können, braucht es eine zeitliche Auflösung des Signals von etwa 100 Hz.

Die nächsthöhere zeitliche Integrationseinheit sind die Vokalisationen, die Levelt «phonological words», Liberman «syllabic bundles», Jaffe «phrases» nennt. Aus der ON/OFF-Sequenz der Mikroebene entstehen hier Wörter, größere sprachliche Einheiten, die etwas weniger als eine Sekunde lang sind und die als untere Mesoebene in unsere Hierarchie eingehen.

Wieder eine Stufe höher liegen die «intonational phrases» oder «phonemic clauses» (Boomer, 1965), sprachliche Einheiten, die unter bewußter Kontrolle des Sprechers stehen und ein wichtiges Ausdrucksmittel darstellen (Levelt, 1989, 385). Sie dauern zwischen einer und vier Sekunden mit einem Modus von etwa zwei Sekunden. Wenn die Sprechgeschwindigkeit erhöht wird, zerfällt diese Struktur als erstes. Da ihre Dauer andererseits mit den Ereignissen der unteren Meso-Ebene hoch korreliert ist, haben wir auf die Einführung einer eigenen Ebene verzichtet.

Dennoch ergibt sich aus diesen intonational phrases ein interessanter Untersuchungsaspekt: wie Studien von Schleidt (1988, 1989) zeigten, sind transkulturell die Längen motorischer Ausdruckseinheiten (Bewegungen, Gesten) sehr vergleichbar und verteilen sich um einen Mittelwert von etwa drei Sekunden. Pöppel (1985) unterstellt, daß mit diesen drei Sekunden eine generelle zeitliche Verarbeitungseinheit gefunden worden sei, die für viele Verhaltensbereiche, insbesondere auch für das Sprechen, gelte. Unsere Ergebnisse bestätigen dies, zeigen aber eine mittlere Länge dieser Einheiten von etwa 2,5 Sekunden auf. Insbesondere Kendon arbeitet an dieser strukturellen Gleichheit von Sprache und Gestik, wie seine Überschrift «How gestures can become like words» (1988) zeigt.

Die zwei bisherigen Zeitebenen (bzw. die drei Ereignis-Klassen) gehören alle zum Bereich der Sprachproduktion und stellen Organisationsformen der Prosodie dar, die insgesamt unter dem Regime möglichst optimaler Verständlichkeit steht. Anders die langen Pausen zwischen den Sprechereignissen der Abbildung 1. Was hier geschah, ist offensichtlich: der Partner redete, die aufgezeichnete Person hörte mit aller Wahrscheinlichkeit zu. Dabei zeigt sich, daß dieses Zuhören ein sehr aktiver Prozeß ist: während des Zuhörens werden immer wieder kurze sprachliche Einheiten eingeworfen, die aus Interjektionen wie «mmmh», «ja», «klar» usw. bestehen. Wie unsere Aufzeichnungen zeigen, ist ihre zeitliche Positionierung von erstaunlicher Exaktheit. Sie werden vom Hörer so präzise in kurze Sprechpausen des Sprechers gesetzt (Größenordnung etwa 500 Millisekunden), daß dies lediglich durch Antizipation des prospektiven Verlaufs des Sprachflusses möglich ist.

Damit sind die langen Pausen – wie die dazwischenliegenden Sprechstücke – nur aus Gestaltungsprinzipien des Dialogs verständlich, gehören mithin zum Bereich der eingangs definierten «dramatischen» Prinzipien der Zeitstruktur. Verständlich wird diese Struktur nur aus dem Dialog, der eine Fülle neuer Ereignisse bringt wie das «Unterbrechen», das «gemeinsame Schweigen», die Zeiten, die es für einen Sprecherwechsel braucht usw.

Aus monologischen Daten (wenn nur ein Sprecher aufgezeichnet ist), können diese dyadischen Strukturen nur approximativ erschlossen werden. Der dialogische Sprecherwechsel – und damit das Ende des Sprechens – ist nicht oder nur höchst selten bedingt aus strukturellen Eigenschaften des Sprechens selbst. Der Inhalt des Gesagten, aber auch dynamische Aspekte der Beziehung zwischen den Sprechern sind machtvollere Kriterien für diesen Wechsel. Damit ist die «Äußerung» (und damit auch die Pause zwischen Äußerungen) als dialogische Einheit nur bedingt aus der Betrachtung nur eines Sprechers zu erschließen. Dennoch konstituiert dieses Ereignis Äußerung die obere Mesebene des Sprechens.

Die oberste Zeitebene des Sprechens ist erreicht, wenn man das Sprechverhalten eines Menschen etwa über einen ganzen Tag hinweg verfolgt. Dort ergibt sich eine neue ON/OFF-Struktur von «Gesprächen» und Pausen dazwischen. Daß dies nicht so einfach ist, zeigt Abbildung 3.

Im oberen Teil der Abbildung ist der Sprechverlauf eines Ehepaares über den ganzen Tag hinweg dargestellt. Die Messung beginnt gegen 9 Uhr morgens und endet um Mitternacht. Die Säulen über der Abszisse stellen die Sprechhäufigkeiten des Mannes dar, die der Abszisse «hängenden» Säulen die Häufigkeiten der Frau. Je höher die Säule, desto mehr wurde gesprochen. Im unteren Teil der Abbildung ist die nach dem gleichen Schema aufgebaute Messung für ein zweites Paar dargestellt. Überraschend ist die Tatsache, daß es nur wenige wirklich stille Perioden über den Tag hinweg gibt. Die Schweigephasen sind leicht zu erklären: im oberen wie unteren Abbildungsteil hielten die Paare etwa zwischen 14 und 15 Uhr einen Mittagsschlaf, der im unteren Abbildungsteil gefolgt war von einer Arbeitsphase, bei der jeder an seinem Schreibtisch saß. Die lange Schweigephase der Frau am Vormittag erklärt sich daraus, daß sie hier allein zuhause war.

Im Gegensatz zur logischen Erwartung zeigt der Sprechverlauf über den Tag hinweg kein klares ON/OFF-Muster. Vielmehr gibt es Perioden dichten und geringen Sprechens – vergleichbar den Ereignissen auf den anderen Zeitebenen, wo ebenfalls nicht ein reines ON/OFF, sondern ein HI/LO (Abk. für high/low oder hoch/nieder) gefunden wurde. Offensichtlich bedeutet Zusammensein auch miteinander reden. Wer nicht redet, ist faktisch allein, auch in der Anwesenheit des Partners. Dies ist der Fall, wenn etwa am Schreibtisch gearbeitet wird, im Schlaf, aber auch, wenn die Partner sich streiten.

Damit bekommt der Begriff der «Stimmföhlung» aus der Ethologie hier eine ganz besondere Bedeutung. Unsere Ergebnisse zeigen, daß beim Zusammensein ein «Teppich» von kurzen Äußerungen ausgelegt wird, in denen – wie Befragungen der Partner ergaben – kaum Information (und wenn, dann mit hoher Redundanz) ausgetauscht wird. Das Reden über gänzlich Unwichtiges scheint eine eigene Funktion der sozialen Rückmeldung zu haben in dem Sinne: «Ich bin noch da – du bist noch da – alles ist o.k. – nichts ist passiert.» So unwichtig der Inhalt dieses Sprechens zu sein scheint, so auffällig ist, daß genau dieser Dialogtyp als erstes verschwindet, wenn im Zusammensein Ärger auftaucht.

Ein weiterer Beleg für die grundsätzliche Unterscheidung zwischen «Stimmföhlung» und «Gespräch» ergibt sich aus der (ex post in der Nachbefragung erhobenen) Antwort der Probanden, wann sie im Verlauf des Meßtages ein «Gespräch» miteinander gehabt hätten. Für den oberen Abbildungsteil wurde lediglich ein Gespräch zwischen 19 und 20 Uhr angegeben, für den unteren Teil zwei Gespräche (eines gegen 14 Uhr, das zweite gegen 20 Uhr). Keines dieser so bezeichneten «Gespräche» ist in auffälliger Form aus dem Verlauf des Sprechverhaltens «herauszusehen». Wie hier, konnten wir auch bei unseren anderen Messungen nicht finden, daß die im Urteil als «Gespräch» qualifizierten Situationen sich im Sprechverhalten besonders ausgezeichnet hätten: aus der Inspektion von Sprechhäufigkeiten ist damit nur schwerlich ein Kriterium für das Vor-

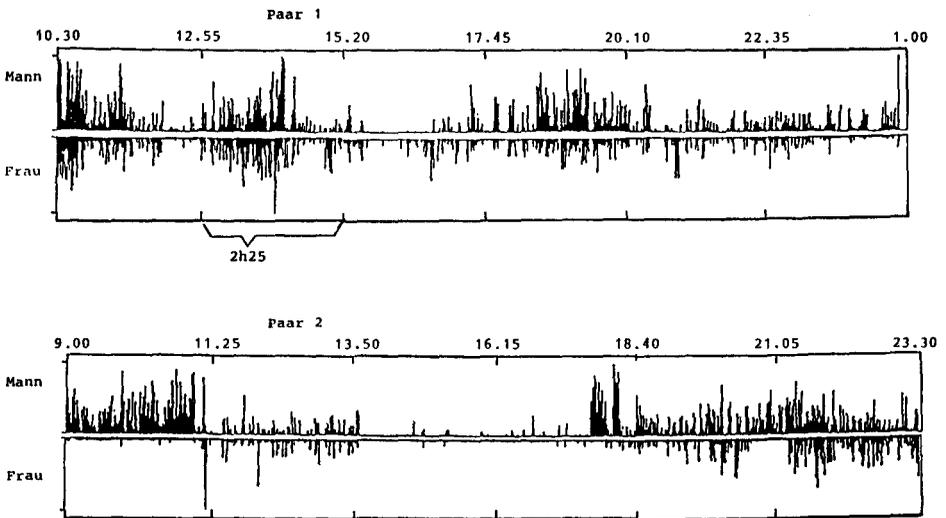


Abbildung 3: Zeitverlauf des Sprechverhaltens zweier Paare über den Tag hinweg. Erklärung im Text.

Tabelle 1: Mittelwert («Mittel»), Standardabweichung («Std») und Median der Parameter.

Ebene	Mittel	Std. der ONs	Median	Mittel	Std. der OFFs	Median
Macro (Minuten)	55,6	39,7	42,5	29,1	19,8	21,4
Obere Meso (Sekunden)	4,9	2,0	4,4	11,7	4,3	12,2
Untere Meso (msek.)	797	231	808	744	182	724
Micro (msek.)	180	48	183	94	11	94

liegen von «Gesprächen» zu gewinnen. Mithin scheint dieser Begriff eine besondere Kombination aus (erlebtem) Informationsaustausch und Kommunikationsintention zu sein, aber kein objektivierbarer Verhaltensbegriff.

Alle vier Zeitebenen sind in der gleichen Messung der ON/OFF-Strukturen zu finden und müssen durch eine entsprechende Auswertung isoliert werden. Diese bedient sich einer bottom-up/top-down-Prozedur. Es werden auf einer Ebene E1 Sprechereinheiten (etwa Phoneme) zu Sprechereinheiten der Ebene E2 (etwa Wörter) zusammengefaßt (bottom-up). Die Pausen zwischen E1-Ereignissen (Artikulationspausen) werden dann in einem zweiten Schritt nur innerhalb eines E2-Sprechereignisses (Wort) ausgewertet (top-down). Aus dieser Prozedur ergeben sich für die vier Zeitebenen die in Tabelle 1 dargestellten deskriptiven Statistiken aus dem uns vorliegenden, einen weiten Anwendungsbereich abdeckenden Material.

Die Validität der Zeitebenen

Da die vier Zeitebenen nicht zwingend in die gleiche Richtung variieren (lange Gespräche bestehen nicht unbedingt aus langen Äußerungen, jene nicht aus langen Artikulationen usw.), ergibt sich die Frage nach Sensitivität und Spezifität der Parameter der einzelnen Ebenen. Jaffe et al. (1979, 407) vermuten, daß die schnellen Zeitebenen (Mikro-, untere Mesoebene) Unterschiede in psychischen Zuständen abbilden, während die oberen Zeitebenen eher sensitiv sind für Veränderungen der sozialen Situation. Unsere Ergebnisse stützen und verfeinern diese Annahmen. Danach ist die aktuelle Ausprägung der Zeithierarchie das Ergebnis aus dem Zusammenwirken von aktivationalen, emotionalen, motivationalen und kognitiven Prozessen (Abbildung 4).

Dabei repräsentiert die Mikroebene am ehesten die motorischen Prozesse der Sprachproduktion und ist von daher besonders sensitiv für Veränderungen des basalen Erregens (arousal), aber auch der Schwierigkeit des Sprechens (etwa durch Steigern der Sprechgeschwindigkeit). Vollrath (1992) konnte an Studien mit psychotropen Substanzen (vor allem Alkohol, Coffein) aufzeigen, daß diese Ebene massiv durch Erregungsänderungen beeinflusst wird. Der Blutspiegel des Alkohols wie der des Coffeins bildet sich im Verlauf der Dauer von Artikulationspausen ab. Damit bietet sich mit der Mikroebene ein äußerst sensitives und nonreaktiv zu erhebendes Erregungsmaß an.

Validität der Zeitebenen

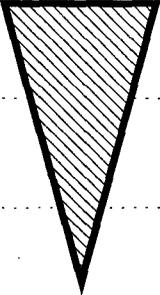
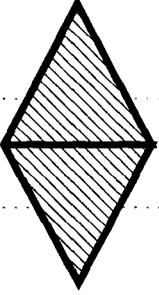
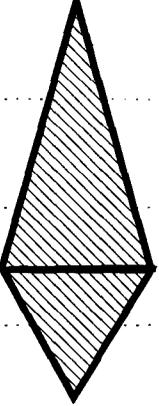
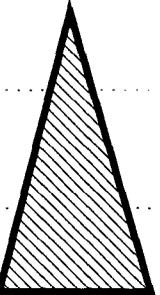
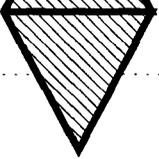
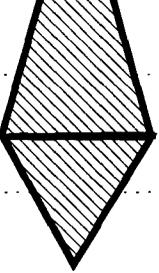
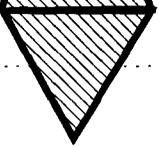
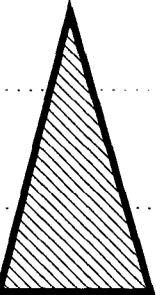
Zeit- ebene	Aktivierung Emotion	Kognitive Prozesse	Sozio- emotional	Sozio- personal
Mikro				
untere Meso				
obere Meso				
Makro				
z.B.	<i>Müdigkeit</i> <i>Erregung</i> <i>Drogen</i>	<i>Schwierigk.</i> <i>linguistische</i> <i>Faktoren</i>	<i>Flirt</i> <i>Streit</i> <i>Neutral</i>	<i>Introversion</i> <i>soziale</i> <i>Angst</i>

Abbildung 4: Validität der verschiedenen Zeitebenen.

Auf der unteren Mesoebene wirken kognitive und emotionale Prozesse zusammen. Die Variabilität der taktischen Prozesse bei der Auswahl von semantischen und grammatikalischen Einheiten zur Generierung der verbalen Botschaft bildet sich am ehesten auf dieser Ebene ab. Zusätzlich bekommt die Botschaft auf dieser Ebene noch ihre emotionale Tönung.

Die obere Mesoebene ist der Bereich der strategischen Sprechplanung und -ausführung, die sowohl kognitive wie sozioemotionale Komponenten enthält. Auf dieser Ebene wird die intendierte Äußerung zusammengestellt und in den situativen Kontext des Sprechens eingepaßt. Zusätzlich werden auf dieser Ebene die unterschiedlichen Ausdruckskanäle (etwa Mimik, Gestik, Sprache, Blickkontakt) integriert. Auf der anderen Seite bilden sich hier sozioemotionale Einflüsse aus dem Gesprächstyp (freundlich, aggressiv usw.) und aus dem sozialen Verhältnis der Sprecher zueinander ab.

Die Ereignisse der Makroebene sind am stärksten beeinflusst durch Sprechercharakteristika der sozialen Motivation und sozialen Intelligenz. Diese Ebene ist auch am besten geeignet, um psychopathologische Zustände wie etwa Depression oder soziale Ängstlichkeit und deren Veränderung unter Therapie abzubilden.

Damit ergibt sich: je langsamer die Ereignisse sind, um so mehr reflektieren sie Unterschiede im trait-Bereich («Eigenschaft»), je schneller, desto eher liegen die Sensitivitäten im state-bereich («Zustände»). Ein methodisches Hauptproblem ist die Selektion der adäquaten Sprechsituation zur Erhebung des Sprechmaterials. Geht es nur um die Abbildung motorischer Prozesse, hat sich das Lesen von standardisierten Texten bewährt. Sollen kognitive und/oder emotional-aktivationale Veränderungen abgebildet werden, bedarf es eines entsprechenden situativen Kontexts, der etwa durch Interviews oder durch Rollenspiele hergestellt werden kann. Geht es um Persönlichkeitsmerkmale genereller Art, ist die Messung des Sprechverhaltens im freien Feld unter natürlichen Kommunikationsbedingungen die Methode der Wahl.

Damit erweisen sich die «speech chronemics» als potente Methodologie – vor allem im Bereich der Psychopathologie – wenn folgende Prinzipien eingehalten werden:

- Der Meßgegenstand muß sein eine Veränderung der Geschwindigkeit des Verhaltens und/oder eine Veränderung von dessen zeitlicher Strukturierung,
- die richtige Zeitebene muß gefunden werden, die sensitiv für diese Veränderungen ist und schließlich,
- die richtige Sprechsituation muß realisiert werden, in der diese Veränderung beobachtet werden soll.

Literatur

- Boomer, D. S.: Hesitation and grammatical encoding. *Language and Speech*, 8, 1965, 148–158.
- Cassotta, L.; Feldstein, S.; Jaffee, J.: AVTA: A device for automatic vocal transaction analysis. *Journal of Anal. Behavior*, 7, 1964, 99–104.
- Chapple, E. D.: The interaction chronograph: Its evolution and present application. *Personnel*, 25, 1948/49, 295–307.
- Craig, R.: The user's guide to Listener Speech Signal Processor. Version 2.0. Menlon Park, CA: Apian Software, 1992.
- Davis, M.: (Ed.) *Interaction Rhythms*. Human Sciences Press, New York, 1982.
- Eibl-Eibesfeldt, I.v.: *Die Biologie des menschlichen Verhaltens*. Piper, München/Zürich, 1984.
- Hall, E. T.: *The silent language*. Doubleday, New York, 1959.
- Hall, E. T.: *The hidden dimension*. Doubleday, New York, 1966.
- Hall, E. T.: *The dance of life*. Doubleday, New York, 1983.
- Hall, E. T.: The hidden dimensions of time and space in today's world. In Poyatos, F. (Ed.). *Cross-cultural perspectives in nonverbal communication*. Hogrefe, Toronto, 1988, 145–152.
- Jaffe, J.; Anderson, S. W.; Stern, D. N.: Conversational rhythms. In: Aaronson, D.; Rieber, R. W. (Eds.), *Psycholinguistic research: implications and applications*, 393–431. Lawrence Erlbaum, Hillsdale, 1979.
- Kendon, A.: How gestures can become like words. In Poyatos, F. (Ed.). *Cross-cultural perspectives in nonverbal communication*. Hogrefe, Toronto, 1988, 131–141.
- Kowal, S.: *Über die zeitliche Organisation des Sprechens in der Öffentlichkeit*. Huber, Bern/Stuttgart, 1991.
- Kowal, S.: Comparison of real-time analytic methods. Paper presented at the Annual Meeting of the AAAS in Chicago, 1992.
- Krüger, H. P.: Speech chronemics – a hidden dimension of speech. *Psychopharmacology*, 22 (1), 1989, 6–12.
- Levelt, W. J. M.: *Speaking. From intention to articulation*. MIT Press, Cambridge, 1989.
- Liberman, A. M.; Studdert-Kennedy, M.: Phonetic Perception. In Held R., Leibowitz, H. & Teuber H.-L. (Eds.). *Handbook of sensory physiology*, Vol. VIII, Perception (S. 143–178). Springer, Heidelberg, 1978.
- Matarazzo, J. D.: A speech interaction system. In: Kiesler, D. J. (Ed.), *The process of Psychotherapy*, Aldine, Chicago, 1973, 138–146.

- Pöppel, E.: Grenzen des Bewußtseins. DVA, Stuttgart, 1985.
- Poyatos, F.: The communication system of the speaker-actor and his culture. A preliminary investigation. *Linguistics*, 83, 1972, 64-86.
- Schleidt, M.: A universal time constant operating in human short-term behavior repetitions. *Ethology*, 77, 1988, 67-75.
- Schleidt, M.; Feldhütter, I.: Universal time constant in human shortterm behavior. *Naturwissenschaften*, 76, 1989, 127-128.
- Studdert-Kennedy, M.: From continuous signal to discrete message: syllable to phoneme. In Kavanagh, J. F.; Cutting, J. E. (Eds.). *The role of speech in language*. MIT Press, Cambridge, Mass., 1975, pp. 113-125.
- Tembrock, G.: *Grundriß der Verhaltenswissenschaften*. Fischer, Stuttgart/New York, 1980.
- Vollrath, M.: *Apparative Messung und psychologische Bedeutung von Mikropausen beim Sprechen*. Inauguraldissertation Würzburg (in Vorbereitung).