

Peter Auer und Elizabeth Couper-Kuhlen

Rhythmus und Tempo konversationeller Alltagssprache

In: Zeitschrift für Literaturwissenschaft und Linguistik 96 (1994), 78- 106

Rhythmus und Tempo konversationeller Alltagssprache

0. Einleitung

Die Bedeutung der Prosodie für die Konstitution alltäglicher¹ Gesprächsinteraktionen wird heute zumindest theoretisch kaum mehr bestritten. Aus dem weiten Feld der Prosodie ist aber bisher – wenn überhaupt – vor allem die Intonation (Tonhöhenbewegung) in der Forschung zur Gesprächs- und Konversationsanalyse bzw. gesprochenen Sprache empirisch untersucht worden (vgl. für das Deutsche z. B. Selting, im Druck). Die mit dieser Vorliebe verbundene Vernachlässigung des Rhythmus und des Tempos mag unter anderem daran liegen, daß für diesen Bereich (in dem sich doch die Zeitlichkeit der mündlichen Sprache, also eines ihrer konstitutiven Merkmale, unmittelbar niederschlägt) das geeignete Untersuchungsinstrumentarium noch fehlt, der Phänomenbereich also »schwer faßbar« ist.

Der vorliegende Beitrag soll diesem Manko abhelfen und zugleich anhand einiger Beispiele aus englischen und deutschen Konversationsausschnitten zeigen, daß und wie Rhythmus und Tempo für das alltägliche Sprachhandeln als »Kontextualisierungshinweise«² relevant werden. Bei der Klärung der konzeptuellen und methodischen Probleme, die mit der Untersuchung der Rhythmen konversationeller Alltagssprache verbunden sind, werden wir auf Ergebnisse eines umfangreicheren Forschungsprojekts zum gleichen Thema zurückgreifen, die an anderer Stelle in extenso dargestellt sind und die wir hier nur andeuten.³

1 Der Begriff des »Alltags« im hier verwendeten Sinn bezieht sich auf alle Formen von Sprachverwendung, die in sog. natürlichen Interaktionszusammenhängen verwendet werden (vgl. Auer 1991); er schließt also lediglich diejenigen linguistischen Objekte als Untersuchungsgegenstand aus, die in manchen Strömungen der Linguistik besonders beliebt sind, nämlich kontextfreie Beispielsätze, außerdem experimentelle oder literarische Daten.

2 Vgl. Auer/Di Luzio (Hgg.) 1992.

3 Das Projekt wurde 1987–1994 von der DFG unter dem Titel »Kontextualisierung durch Rhythmus und Intonation« gefördert. Ergebnisse können Couper-Kuhlen (1993) und Auer (Hg.) (in Vorb.) sowie einzelnen Beiträgen in Couper-Kuhlen/Selting (im Druck) und Auer/di Luzio (eds.) 1992, außerdem Auer (1990) und Couper-Kuhlen/Auer (1991) entnommen werden. Vorabveröffentlichungen der zur Publikation vorgesehenen Beiträge zur interaktiven Funktion des Rhythmus sowie zu begleitenden sprachstrukturellen Untersuchungen sind in der Preprint-Serie »KontRI« (Kontextualisierung durch Rhythmus und Intonation) erschienen und über die Autoren erhältlich. Vgl. außerdem den Beitrag von F. Müller in diesem Band.

1. Die Entzeitlichung der Sprache: Gründe für die Vernachlässigung von Rhythmus und Tempo in der Sprachwissenschaft

Rhythmen – vorläufig in allgemeiner Form als in der Zeit rekurrente Ereignismuster definierbar – bestimmen unser Leben, vom Wechsel zwischen Tag und Nacht, den Jahreszeiten, Monaten und Jahren über die Körperrhythmen wie Schlafen/Wachen, Herzschlag, Atmung und die grundlegenden Bewegungsabläufe wie Gehen, Schwimmen oder Laufen bis zu den zentralen, vielleicht universalen menschlichen Kunstformen Musik, Tanz und Dichtung. Aber während z. B. eine Beschreibung des menschlichen Gehens ohne Bezug auf seine temporale Rekurrenzstruktur unvollständig, ja unsinnig wäre, kommen die traditionellen und gegenwärtigen Lehrbücher der Linguistik bei der Beschreibung der Sprache sehr wohl ohne einen solchen Verweis aus: Keines der großen und einflußreichen sprachwissenschaftlichen Lehr- oder Überblicksbücher der letzten 50 Jahre räumt dem Rhythmus der Sprache auch nur ein Nebenskapitel ein. Spielt Rhythmus also für die (Alltags-)Sprache keine Rolle?

Wir werden in diesem Beitrag zeigen, daß ein solcher Schluß völlig verfehlt wäre. Die Gründe für die Vernachlässigung zeitlicher Aspekte der Sprache (für deren »Entzeitlichung«) sind vielmehr, so meinen wir, Resultat und Indikator des europäischen Denkens über Sprache (also unserer »Sprachideologie« im Sinne von Silverstein 1979), die sich sowohl in der Laienkonzeption von Sprache als auch in strukturalistischen und nach-strukturalistischen (»generativen«) Konzeptualisierungen des Gegenstands der Sprachwissenschaft niederschlagen. Ein Blick in die sprachwissenschaftlichen Hauptwerke des ausgehenden 19. und beginnenden 20. Jahrhunderts (vgl. besonders: von der Gabelentz 1891, Wundt 1911–1912³ mit jeweils einem umfangreichen prosodischen Kapitel, sowie außerdem Saran 1907) beweist übrigens, daß sich dieses Denken in der Sprachwissenschaft erst relativ spät und allmählich im Sinne einer Reduzierung der Sprache auf ihre »Grammatik«, also ein System von zeit-neutralen Zeichen durchsetzen konnte. Auf wissenschaftsgeschichtlicher Ebene ließe sich diese Entzeitlichung als Prozeß rekonstruieren, für den der ambige und zwischen *langue* und *parole* schwankende Begriff des Syntagmas im *Cours de linguistique* als eine Schaltstelle zu betrachten wäre: denn während spätere strukturalistische Lehrbücher (z. B. Lyons 1968) Syntagma und Paradigma als die definierenden Dimensionen des linguistischen Systems – der *langue* – begreifen, ist für den *Cours* das Syntagma ein Element des *discours* (das Paradigma aber »en dehors du discours«, »dans la mémoire«, »dans le cerveau«), eine Relation »in praesentia« (das Paradigma aber »in absentia«), und Saussure ([1915] 1972, S. 173) weiß:

il faut reconnaître que dans le domaine du syntagme il n'y a pas de limite tranchée entre le fait de langage, marque de l'usage collectif, et le fait de parole, qui dépend de la liberté individuelle.

Das europäische Denken über die Sprache, das in diesem Sinn zu ihrer Entzeitlichung geführt hat, ist seinerseits Konsequenz einer langdauernden, dieses Denken prägenden Schriftlichkeit; die Schrift, folgenreichste aller menschlichen »Technologien« (Goody 1990), hat die zeitgebundene Eindimensionalität der mündlichen Sprache neutralisiert. Im wesentlichen lassen sich drei für unser Thema relevante Komponenten der europäischen Sprachideologie identifizieren. Die erste ist die Auffassung von der Sprache als Werkzeug (des Denkens, oder des Austausches von Gedanken), das von seinen Verwendungen abgelöst existiert, so wie die Werkzeuge eines Maurers von den Bauten, die damit errichtet werden, unabhängig sind. Die Sprache hebt sich damit aus dem kommunikativen Handeln (Sprechen) ab, ja sie steht ihm in Opposition gegenüber. Während dieses nicht aus seiner zeitlichen Erstreckung isoliert werden kann, ist sie selbst wesentlich über Situationen hinweg stabil.

Die zweite Komponente der europäischen Sprachideologie, die sich unmittelbar aus der ersten ergibt, ist ihre repräsentationale (oder referentielle) Voreingenommenheit. Sprache ist in der Hauptsache ein Instrument, um auf Dinge »draußen in der Welt« zu verweisen (d. h., sie darzustellen). Die Beziehung zu den Denotata ist wiederholbar und daher zeitneutral; d. h. die Verwendung derselben Zeichen ist imstande, die von diesen getrennt gedachten Referenten beliebig oft zu identifizieren bzw. zu repräsentieren. Nur die Beschränkung auf die Darstellungsfunktion von Sprache – also die Ausblendung aller »Symptom«- bzw. »Signal«-Funktionen – ermöglicht eine solche Konstanz der Beziehungen zwischen Zeichen und Denotata (oder, im Sinne dieses Arguments ersetzbar, zwischen *signifiants* und *signifiés*). Korrelat dieser ideologischen Präsupposition ist der Verweis auf die Lexik als Ansammlung stabiler Bezeichnungsmöglichkeiten, die quasi als Bausteine in sprachliche Äußerungen eingebaut werden. Erst die semantisch neutrale Wiederholbarkeit von repräsentationalen Akten ermöglicht es, *die Sprache* als Werkzeug gegen *das Sprechen* als dessen Verwendung auszuspielen.

Zum dritten ergibt sich aus dem Gesagten als weitere zentrale Komponente europäischer Sprachideologie die Dekontextualisierung der Sprache zur Grammatik. Dekontextualisierung meint hier nicht nur die Loslösung der Sprache von den Aspekten ihres unmittelbaren situativen oder sprachlichen (ko-textuellen) Umfelds, sondern auch die Ausblendung jenes Konglomerats von Wissensbeständen und Bewertung(smaßstäb)en, das man als »Kultur« bezeichnen kann. Kontexte in diesem umfassenden Sinn könnten nur als emergente, in der Zeit geschaffene Bedingungen für sprachliches Handeln rekonstruiert werden: Sie werden durch die sprachlichen und nicht-sprachlichen Handlungen, die in ihnen stattfinden (und durch sie erst erkennbar bedeutungsvoll werden) von einer Phase des Gesprächs zur nächsten neu indiziert und konstituiert.

Werkzeugcharakter der Sprache, Beschränkung auf ihre Darstellungsfunktion und Dekontextualisierung der Sprache zur Grammatik bilden den ideologischen Rahmen, in dem ihre Entzeitlichung möglich, ja folgerichtig erscheint. Die Ver-

nachlässigung von Rhythmus und anderen zeitgebundenen Aspekten der Sprache in strukturalistischen (einschl. »generativen«) Konzeptualisierungen ist ihr wissenschaftlicher Reflex. Er beginnt sich erst seit den 70er Jahren in dem Maß aufzulösen, in dem die genannten Einengungen des sprachwissenschaftlichen Forschungsgegenstands wegfallen; und je weiter diese Auflösung fortschreitet, werden diejenigen Aspekte von Sprache einer wissenschaftlichen Analyse zugänglich, die, wie der Rhythmus, weder losgelöst von ihren lokalen Funktionen beschreibbar sind (als »Werkzeug« also keinen Sinn unabhängig von ihrer Verwendung haben), noch Darstellungsfunktion haben, noch außerhalb ihres Kontexts in ihrer Bedeutung erschlossen werden können.

2. *Impulse für die Erforschung des Rhythmus aus den Nachbarwissenschaften der Linguistik*

Die linguistische Forschung zum Rhythmus hat sich entsprechend diesem theoretischen und ideologischen Umfeld bis vor kurzem auf die Analyse virtueller (abstrakter) Rhythmen beschränkt, wie sie z. B. im Rahmen der Metrischen Phonologie in die Grammatik integriert worden sind.⁴ Obwohl die theoretischen und deskriptiven Fortschritte der Metrischen Phonologie im Vergleich zu früheren, strukturalistischen Versuchen der Berücksichtigung prosodischer Phrasierung evident und nicht gering zu bewerten sind (wir verweisen als Beispiel lediglich auf die Fortschritte bei der Analyse des *stress shift*), beruht sie doch ausschließlich auf der Diskussion sog. wohlgeformter rhythmischer Strukturen, die allein durch die Kompetenz des Sprecher-Hörers (oder besser: der Autorin-Leserin) kontrolliert werden. Im Gegensatz zu manchen Bereichen der Syntax oder Morphologie steht unsere Intuition aber gerade im Falle des Rhythmus auf wackeligen Füßen. Es läßt sich vermuten, daß die Möglichkeiten des metasprachlichen Zugriffs auf sprachliches Wissen, wie es die Methodologie der »generativen« Grammatik mit der Arbeit an erfundenen Beispielsätzen erfordert, ihrerseits schriftsprachlich geprägt sind und deshalb gerade in jenen Bereichen der Sprache, die typischerweise dieser Schriftlichkeit entzogen sind (wie der Prosodie), an ihre Grenzen stoßen. In jedem Fall ist die Analyse abstrakter/virtueller Rhythmen, wie sie in der Metrischen Phonologie gepflegt wird, nur in sehr eingeschränktem Maß geeignet, unsere Kenntnis von Struktur und Funktion des Rhythmus der Alltagssprache zu erweitern.

Die wesentlichen Impulse für die Analyse des Rhythmus der konversationellen Alltagssprache kommen deshalb aus Nachbargebieten der Sprachwissenschaft. Auf die kontroverse und intensive *phonetische* Debatte um »Isochronie« als

4 Vgl. z. B. Selkirk 1984.

rhythmischem Parameter wollen wir hier (aus Platzgründen) lediglich verweisen; in der meist experimentellen phonetischen Forschung wird zwar keine Alltagssprache untersucht, dennoch sind ihre Ergebnisse auch für die Erfassung von Rhythmus im Diskurs wichtig.⁵

Aus der *psychologischen* Forschung kommen Anregungen zur Erforschung des Rhythmus aus drei unterschiedlichen Traditionen:

– Die Grundlagen der Interpretation und Verarbeitung rhythmischer Muster wurden schon um die Jahrhundertwende durch die Gestaltpsychologie erforscht.⁶ Der wichtigste theoretische Beitrag dieser Forschungstradition ist die Erkenntnis, daß rhythmische Strukturen nicht als solche im Signal vorhanden sind, sondern die Wahrnehmung rhythmischer Gestalten ein aktiver Prozeß der zeitlichen *Gestaltung* ist. Alle sprachlichen Rhythmen, die für die Interaktion relevant werden können, sind also »subjektive Rhythmen« (Fraisse 1974). Manche der Prinzipien, nach denen die menschliche Wahrnehmung physikalische Stimuli gruppiert, so daß sie als rhythmische Gestalten erkannt werden, sind gesichertes psychologisches Wissen, andere werden erst errahnt. Seit den Anfängen der Gestaltpsychologie ist z.B. bekannt, daß die subjektive Rhythmisierung auch Folgen von physikalisch identischen Stimuli gruppiert (meist in Zweier-, seltener in Dreiergruppen) und dabei einen Stimulus (meist den ersten) zum Akzent erhebt, also hervorhebt. Wir wissen außerdem, daß für subjektive Rhythmisierung bestimmte Toleranzgrenzen bezüglich des Abstands der Stimuli zu respektieren sind (die untere Schwelle scheint bei ca. 150–200 msec., die obere bei 1.5–2 sec. zu liegen) und daß es auch für die Dauer eines rhythmischen Intervalls optimale (ca. 400 msec.) und obere (4–5 sec.) Werte gibt (vgl. Fraisse 1974). Psychologische Experimente weisen außerdem darauf hin, daß bei der Wahrnehmung isochroner sprachlicher Muster (also von rhythmischen Intervallen gleicher Dauer) »objektive« Schwankungen von bis zu 30 % nicht notwendigerweise als rhythmusstörend empfunden werden;⁷ umgekehrt werden physikalisch isochrone sprachliche Stimuli in Abhängigkeit von ihrer phonologischen Struktur unter Umständen als unrhyth-

5 Ausführlicher dazu Auer/Uhmann 1988. Vgl. als Beispiel für die Relevanz phonetischer Rhythmus-Forschung die Diskussion um die P-Centers (Anm. 8). Es sollte allerdings berücksichtigt werden, daß in der Phonetik meist ein engerer Isochronie-Begriff verwendet wird als in unseren Untersuchungen.

6 Vgl. die Zusammenfassung von Fraisse (1974). Die Relevanz gestaltpsychologischer Prinzipien für das Hören von rhythmischen Strukturen wird in Couper-Kuhlen (1993, S. 66 ff.) detailliert dargestellt.

7 Bei Allen 1975, Lehiste 1977 und Couper-Kuhlen 1991 finden sich genauere Informationen.

misch interpretiert.⁸ Methodologisch muß aus diesen Erkenntnissen die Konsequenz gezogen werden, daß eine einfache und unproblematische meßphonetische Kontrolle von Rhythmizitäts- oder Isochronie-Urteilen nicht möglich ist; um die Beziehung zwischen wahrgenommenen und »objektiven« Rhythmen erfassen zu können, sind komplizierte und bis heute erst sehr unvollständig bekannte Theorien notwendig, was die Frage der Priorität zwischen gemessenen und gehörten Rhythmen zumindest offen erscheinen läßt (vgl. unten, Abschnitt 4). Gerade für eine interaktiv orientierte Untersuchung der Rhythmisierung (spontaner) Alltagssprache kommt dem auditiven Vorgehen deshalb nach wie vor eine bedeutende Rolle zu.

– Weniger ertragreich als die Gestaltpsychologie sind psychologische Untersuchungen zur sog. »interactional synchrony«, die – in der Tradition E.D. Chapples (1939) – vor allem Condon (1980, Condon/Ogston 1967) unternommen hat. Obwohl in dieser »chronographischen« Tradition mit interaktiver Alltagssprache gearbeitet wird, zielt das Forschungsinteresse in erster Linie auf Fragen der Synchronisierung von Bewegungsabläufen im Sprecher (z. B. zwischen phonologisch-segmentaler und grammatischer Struktur auf der einen, gestischer Struktur auf der anderen) oder zwischen Sprechern ab. Die von Condon unterstellten mikroskopischen Synchronisierungseffekte sind aber weder unumstritten,⁹ noch wird ihre interaktive Bedeutung ausreichend thematisiert. Akkomodation zwischen Sprechern in bezug auf Pausenstruktur, Länge von Redebeiträgen u. a. (vgl. Jaffe/Feldstein 1970), wie sie in dieser Tradition ebenfalls mit einigem Erfolg untersucht worden ist, betrifft zwar den Zeitcharakter von Sprache, nicht jedoch deren Rhythmus.

– Schließlich gibt es eine wesentliche Richtung in der Psycholinguistik, die Rhythmus in der Sprache unter dem Aspekt der Informationsverarbeitung und -speicherung betrachtet. Grundlegend ist dabei die Überlegung, daß der rhythmische Wechsel zwischen akzentuierten und nicht-akzentuierten Silben bestimmte Ereignisse zeitlich vorhersagbar macht (nämlich diejenigen, die auf den rhythmischen Schlag fallen) und daß phonetisch hervorgehobene Strukturen (z. B. Silben, Wörter, Fokuskonstituenten) wichtigere Information enthalten als nicht hervorgehobene. Die Verarbeitung sprachlicher Information wird dadurch erleichtert, daß solche rhythmischen Schläge mit einem Zustand

8 Zur Abhängigkeit isochroner Rhythmisierungen von der phonologischen Struktur der Stimuluswörter gibt es in der Phonetik eine umfangreiche jüngere Forschungstradition, die unter dem Stichwort »P-centers« vor allem den Einfluß des Silbenanstiegs auf die Lokalisierung des rhythmischen Referenzpunkts analysiert hat; vgl. etwa in jüngerer Zeit Pompino-Marschall 1989, Hoequist 1983.

9 Vgl. vor allem Rosenfeld 1981.

erhöhter perzeptorischer Zuwendung und Aufmerksamkeit koinzidieren, während die Zeit, die zwischen den Schlägen mit phonetisch weniger hervorgehobenen Silben gefüllt wird, genutzt werden kann, um jene zu verarbeiten (vgl. Allen 1975, Martin 1972, Allen/Hawkins 1980, Jones 1986). Es gibt überdies Evidenz dafür, daß rhythmisierte Information besser erinnert wird als nicht-rhythmisierte (Payne/Holtzman 1986).

Die dritte nicht-linguistische Forschungsrichtung neben Phonetik und Psychologie, aus der Impulse für die sprachwissenschaftliche Untersuchung des Rhythmus der Alltagssprache kommen, ist die *linguistische Anthropologie*. Wichtig sind hier weniger die anthropologischen Arbeiten über »exotische« Sprachkulturen, für die zwar die Bedeutung des Rhythmus vor allem im Rahmen von formellen, teils rituellen Gattungen erkannt und gelegentlich auch untersucht wurde,¹⁰ als vielmehr Untersuchungen zur amerikanischen Alltagssprache. Neben Gumperz' Analyse schwarzer Rhetorik (1980) und Scollon's Anmerkungen zum Rhythmus im weißen amerikanischen Englisch (Scollon 1982) sind hier vor allem die Arbeiten Fred Ericksons zu nennen, die unseren eigenen Ansatz wesentlich beeinflußt haben (Erickson 1982, 1992, Erickson/Shultz 1982).

Zwischen Ericksons Konzeption von Rhythmus und der hier vertretenen gibt es allerdings einige wesentliche Unterschiede. Erickson unterstellt die durchgängige isochrone Rhythmisierung konversationeller Daten, was seiner Auffassung entspricht, daß diese »constitutive for social meaning« (Erickson/Shultz 1982, S. 96), also eine »normal condition« (1982, S. 105) von Interaktion ist. Entsprechend stuft er Unterbrechungen dieser Rhythmizität als schwerwiegende Ereignisse ein; und in der Tat scheinen in seinen Daten (Konversationen und institutionelle Beratungen) solche Unterbrechungen auch nicht bzw. nur mit nachweisbar dramatischen Folgen (bzw. als Indikatoren für dramatische Problemsituationen) aufzutreten (»serious interactional trouble«, S. 103). Nach unserer eigenen Auffassung ist Rhythmizität (wie bei Erickson in Sinne von Isochronie verstanden) im Englischen – und noch viel mehr im Deutschen – keine durchgängige Erscheinung von Alltagsgesprächen. Zum Beispiel wird rhythmische Integration im Englischen eher beim Übergang von einem Redebeitrag zum nächsten beobachtet als innerhalb eines Redebeitrags; in der Regel findet man außerdem in beiden Sprachen häufige Tempoveränderungen sowie Phasen, in denen der Rhythmus leicht schwankt oder die Teilnehmer rhythmisch gesehen »aus dem Tritt« kommen. Diese rhythmischen Störungen (oder genauer: Übergänge von Phasen stärkerer und schwächerer Rhythmisierung) können zwar interaktiv durchaus bedeutungsvoll sein (wie aus den Beispielen des Abschnitts 4 ersichtlich werden wird), sie haben aber nicht den radikalen Effekt,

10 Ausnahmen sind verschiedene Arbeiten zur »oral poetry« (vgl. Finnegan 1992², S. 88 ff., Edwards/Sienkewicz 1990).

den Erickson ihnen unterstellt, nämlich die Grundlage der Interaktion aufzulösen.¹¹

Abgesehen von den genannten Autoren ist der Rhythmus der Alltagssprache bisher kaum untersucht worden. Besonders auffällig ist das völlige Fehlen rhythmischer Analysen in der Konversationsanalyse. Soweit temporale Parameter überhaupt berücksichtigt worden sind (wie bei Jefferson 1989 in ihrer Diskussion eines »Einsekundenmaßes für Pausen«) werden sie – in bemerkenswerter Dissonanz zu den phänomenologischen Wurzeln der Konversationsanalyse – in externer (objektiver) Zeit gemessen, nicht als wahrgenommene (subjektive) Zeitstruktur.

3. Der Rhythmus der Alltagssprache als wahrgenommene Gestaltstruktur

Wir wollen in diesem Abschnitt unser eigenes Konzept von Rhythmus vorstellen; auf dieser Grundlage werden wir im nächsten Abschnitt anhand exemplarischer Fälle auf einige seiner interaktiven Funktionen hinweisen.

Eine Analyse, die zu den interaktiven Funktionen von Rhythmus vorstoßen will, darf, wie schon erwähnt, diesen nicht als Eigenschaft physikalischer Signale (wie akustischer Wellenformen) auffassen. Wahrgenommene bzw. wahrnehmbare rhythmische Gestalten lassen sich aus den in diesen Signalen erkennbaren Grundfrequenzverläufen, Amplituden- und Dauerschwankungen nicht direkt ablesen. Sie sind vielmehr erst über zahlreiche interpretative Transformationen auf sie bezogen, die die Hörer zu leisten haben. Auf einer untersten Stufe müssen zunächst Veränderungen in Grundfrequenz, Amplitude, Dauer sowie evtl. Formantenstruktur als solche der Tonhöhe (Intonation), Lautheit, Länge (Quantität) und Segmentqualität verstanden werden. Auf einer zweiten, schon abstrakteren Ebene müssen diese Parameter in einem komplexen und nur ansatzweise bekannten Prozeß als *Akzente* interpretiert werden. Schließlich werden auf einer wiederum übergeordneten Ebene Sequenzen von Akzenten nach be-

11 Eine genauere Lektüre der Schriften Ericksons zeigt, daß das im »Counselor as Gatekeeper« verfolgte Konzept von Rhythmus dem in seinen übrigen Beiträgen zum Thema Rhythmus nicht genau entspricht. In diesen Beiträgen – die bezeichnenderweise mit einer musikalischen Rhythmusnotation arbeiten – bleibt die Beziehung zwischen phonetischen Hervorhebungen, Akzenten und Schlägen teilweise implizit und in jedem Fall sehr indirekt; das rhythmische Tempo ist langsam. In Erickson/Shultz (1982) wird rhythmische Integration hingegen annähernd in unserem Sinn definiert (vgl. den folgenden Abschnitt); wir vermuten, daß die Durchgängigkeit der Rhythmisierung in diesem Fall im Zusammenhang der spezifischen Aktivitätstypen gesehen werden muß, die diese institutionelle Beratungsinteraktion definieren, besonders im Zusammenhang der stereotypen Frage/Antwort-Sequenzen, denen die meisten Beispiele des Buchs entnommen sind.

stimmten Prinzipien zu rhythmischen Gestalten zusammengefaßt. Wir nennen Akzente, die in solchen Gestalten eine Rolle spielen, *Schläge*.

Angesichts dieses mehrfachen Interpretationsprozesses, dem das wahrnehmende Subjekt das akustische Material unterzieht, läßt sich Rhythmus mit Sonnenschein (1925, S. 26) vielleicht am sinnvollsten definieren als

[...] that property of a sequence of events in time which *produces in the mind of the observer the impression* of proportion between the durations of the several events or groups of events of which the sequence is composed (unsere Hervorhebung).

Die Definition erfaßt sowohl das Verhältnis zwischen langen und kurzen Silben (im Sinne klassischer metrischer Muster wie Anapäst oder Iambus), als auch das zwischen Gruppen von Silben (Füßen, Takten), z. B. deren zeitliche Gleichförmigkeit (Isochronie).

Betrachten wir die Wahrnehmung rhythmischer Gestalten im Einzelnen. Im einfachsten Fall kann das die Wiederkehr von Ereignissen in Abständen sein, die für den Hörer oder die Hörerin als annähernd gleich gelten können. Schematisch: wenn $x_1 \dots x_n$ solche Ereignisse sind und $|$ ein in der Wahrnehmung konstantes Zeitintervall, dann ist

$$|x_1 \quad |x_2 \quad |x_3 \quad \dots \quad |x_n$$

ein solcher einfacher Rhythmus.

Es sind also zumindest drei Ereignisse des Typs (x) – zwei Intervalle gleicher Dauer – notwendig, damit von einem rhythmischen Muster (Gestalt) die Rede sein kann. Ist ein solches Muster erst konstituiert, so liefert es eine Erwartungsfolie für die Platzierung des nächsten Ereignisses dieses Typs in der Zeit; da diese Folie die Wahrnehmung steuert, wird die Wahrscheinlichkeit, daß ein nächstes solches Ereignis tatsächlich als rhythmisch interpretiert wird, steigen. Allgemein gilt, daß etablierte und/oder prägnante rhythmische Gestalten die rhythmische Integration weiterer Ereignisse erleichtern, während »junge« und/oder wenig prägnante Gestalten sie erschweren.

Ein komplizierteres, für sprachliche Rhythmen aber ungleich wichtigeres Muster ergibt sich, sobald zwischen den Schlägen, die den Rhythmus definieren, weitere Ereignisse (Silben) stehen. In diesem Fall unterscheidet die Wahrnehmung zwischen starken Elementen (Schlägen) und schwachen. Isochrone Rhythmizität liegt dann vor, wenn die Schläge einer solchen Gestalt in subjektiv gleichen Intervallen auftreten, während die zeitliche Platzierung der schwachen Silben ohne Bedeutung ist. (Ihre individuelle zeitliche Ausdehnung ist im wesentlichen durch die Anzahl und den Aufbau der Silben im rhythmischen Intervall – dem *Fuß* – bestimmt.) Bezeichnen wir starke Ereignisse (Silben) mit s , und schwache mit w , so ergibt sich:

$$|s_1 \ w_{1,1} \dots w_{1,i} |s_2 \ w_{2,1} \dots w_{2,j} |s_3 \ w_{3,1} \dots w_{3,k} \dots |s_n \ w_{n,1} \dots w_{n,k}$$

Im Fall des Englischen oder Deutschen werden solche Gestalten meist so gehört, daß die starken Elemente den schwachen vorausgehen, nicht etwa umgekehrt.

Eine noch prägnantere rhythmische Gestalt entsteht, wenn die Anzahl der schwachen Silben von Fuß zu Fuß identisch ist. Auch hier ist der Rhythmus isochron, er zeigt aber zugleich noch eine konstante (iso-)metrische¹² Struktur:

$$|s_1 \ w_{1,1} \dots w_{1,i} | s_2 \ w_{2,1} \dots w_{2,j} | s_3 \ w_{3,1} \dots w_{3,k} \dots | s_n \ w_{n,1} \dots w_{n,k} \\ \text{(für } i = j = k = 1 \text{)}$$

Schließlich kommen (iso-)metrische Strukturen auch ohne Isochronie vor; wir können diesen Typ von Rhythmus, der sich als

$$s_1 \ w_{1,1} \dots w_{1,i} \quad s_2 \ w_{2,1} \dots w_{2,j} \dots \quad s_n \ w_{n,1} \dots w_{n,k} \quad \text{(wobei } i = j = k \text{)}$$

symbolisieren ließe, hier aber nicht weiter verfolgen.

Rhythmus im hier verstandenen Sinn läßt sich also in die beiden zentralen Komponenten (wahrgenommene) »Isochronie« und »Metrum« zerlegen. Im einen Fall geht es um die regelmäßige Abfolge der Schläge, im anderen um die Rekurrenz bestimmter Proportionen von schwachen zu starken Silben. Wir werden im folgenden vor allem den erstgenannten Aspekt weiter ausführen.

Eine zentrale Rolle spielt in unserer Herangehensweise an Rhythmus die Unterscheidung zwischen Akzent und Schlag. Nicht jede Akzentsilbe ist ein Schlag; vielmehr können zwischen rhythmischen Schlägen weitere Akzente stehen, die dann als schwache rhythmische Ereignisse eingestuft werden. Umgekehrt kann ein rhythmischer Schlag in einer schon etablierten Gestalt gelegentlich leer bleiben (d.h., auf der dafür vorgesehenen Stelle findet kein Ereignis statt). Wir sprechen dann (in Anlehnung an Abercrombie 1967, 1971, aber im Detail abweichend¹³) von einem »leeren Schlag« (silent beat). Obwohl die Ebenen der Identifizierung von Akzenten und die der Wahrnehmung von Schlägen streng geschieden werden müssen, beeinflussen sie sich doch gegenseitig. Diese reflexive Beziehung impliziert einerseits natürlich, daß Akzentsilben eher als Schläge gehört werden als nicht-akzentuierte; umgekehrt aber auch, daß rhythmische Gestalten die Wahrnehmung einer auf den nächsten Schlag fallenden Silbe als »akzentuiert« begünstigen, auch wenn Intonation oder Lautstärke eine solche Interpretation nicht unbedingt nahelegen. Der gesamte Wahrnehmungsprozeß sollte also eher wie in Abbildung 2 als wie in Abbildung 1 visualisiert werden:

12 Isometrische Strukturen spielen für besonders emphatische Rhythmisierungen eine Rolle (»Skandierungen«; vgl. F. Müller 1991), vielleicht sind sie für die Rhythmisierung sog. silbenzählender Sprachen (vgl. für das Italienische F. Müller 1989) auch wichtiger als für sog. akzentzählende.

13 Vgl. Couper-Kuhlen 1993, S. 72f.

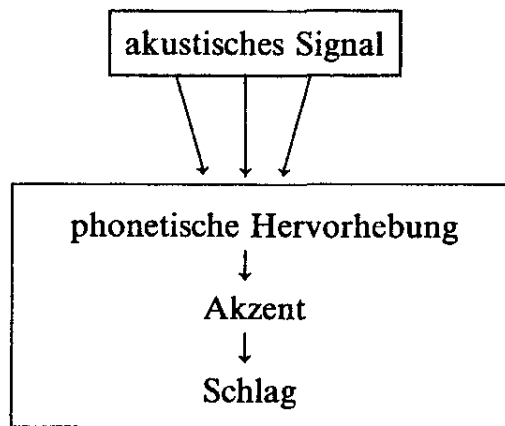


Abb. 1

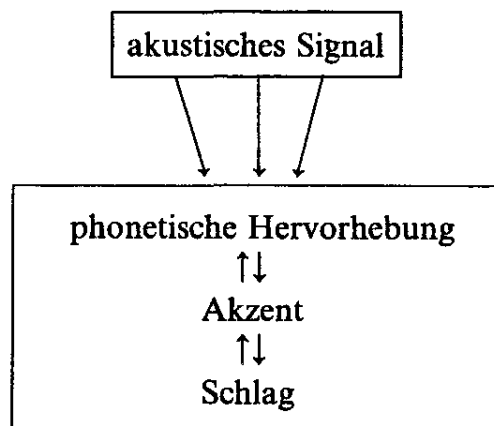


Abb. 2

Wenn wir von Rhythmus als wahrgenommenen Gestalten reden, so meinen wir damit auch (und zwar in Abweichung von der traditionellen phonetischen Forschung à la Abercrombie zum Thema Isochronie, aber auch in Abweichung von mikroethnographischen Studien im Sinne Ericksons), daß »bessere« und »schlechtere« Rhythmisierungen möglich sind (je nach Prägnanz der entstehenden rhythmischen Gestalt). Es wird also nicht nur zwischen rhythmischen und nicht-rhythmischen Strukturen unterschieden, sondern auch zwischen Graden von Rhythmisität. Zu diesem Zweck sind Präferenz-Prinzipien nötig, die die Prägnanz rhythmischer Gestalten bestimmen helfen. Ihnen zufolge ist eine »optimale« Rhythmisierung eine, die

- keine Akzente außerhalb des Schlags und keine leeren Schläge enthält (in der also das Verhältnis zwischen Akzenten zu Schlägen 1 : 1 ist);
- die Grenzen für gut erkennbare rhythmische Intervalle nicht unter- bzw. überschreitet (und sich möglichst zwischen 30 und 100 csec. bewegt);
- Akzente mithilfe von Tonhöhenbewegungen realisiert;¹⁴
- die Anzahl der schwachen Silben zwischen den Schlägen nicht zu stark variiert;
- prosodische Rhythmen durch segmentale Eigenschaften wie Wiederholungen und Parallelismen unterstützt.

Fehlen eines oder mehrere dieser Merkmale, so wird die Äußerung nicht un-rhythmisch; die so realisierte rhythmische Gestalt ist aber weniger prägnant als eine, die den genannten Prinzipien ausnahmslos gehorcht.

4. Rhythmus als Kontextualisierungshinweis

Während bisher konzeptuelle Aspekte der Erfassung alltagssprachlicher Rhythmen erläutert wurden, sollen in diesem Abschnitt einige ihrer interaktiven Funktionen skizziert werden. Dazu sind zwei Vorbemerkungen notwendig. Es ist

¹⁴ Dieses Prinzip gilt vor allem für das Englische.

zum einen der semiotische Rahmen zu skizzieren, innerhalb dessen rhythmisch-zeitliche Phänomene der Sprache interpretiert werden können, und es ist zum anderen die Praxis des Hörens und Notierens rhythmischer Strukturen zu erläutern.

Wie schon ausgeführt, sind Rhythmus und Tempo (a) kein Bestandteil der Grammatik im Sinn einer von der Sprachbenutzung abgetrennten Kompetenz oder eines »Werkzeugs« zur Vermittlung von »Gedanken« (Propositionen), (b) haben sie keine referentiell-darstellende Funktion (sind also, im Sinne Peirces, keine Symbole, sondern indexikalisch-ikonische Zeichen) und (c) läßt sich ihre Bedeutung – im Sinne dieser Indexikalität – nur im Kontext ihrer Verwendung rekonstruieren. Der theoretische Rahmen, der uns für die Analyse solcher Zeichen am geeignetsten erscheint, ist der der »Kontextualisierung« im Sinne von J. Gumperz.¹⁵ Bekanntlich werden darunter die Verfahren verstanden, mittels derer die Interaktionsteilnehmer durch sog. Kontextualisierungshinweise Kontexte schaffen bzw. indizieren (»aufrufen«), die die für die Interpretation der Äußerung notwendige Einbettung der Äußerung leisten. Kontextualisierungshinweise sind also materiale Hilfsmittel für das Verstehen von Äußerungen; sie helfen den Interaktionsteilnehmern, aus dem Gesagten Sinn zu machen, indem sie bestimmte Inferenzen nahelegen und andere ausschließen.

Das Konzept der Kontextualisierung impliziert, daß es wenig sinnvoll wäre, nach der »Bedeutung des Rhythmus« als solchem zu fragen. Ihrem Wesen als Kontextualisierungshinweis entsprechend, ist der Beitrag rhythmisch-zeitlicher Parameter für das Verstehen sprachlicher Strukturen nur im jeweiligen lokalen Umfeld ihrer Verwendung zu erfassen. Generalisierte Bedeutungszuschreibungen, die über solche Kontexte hinwegpauschalisieren – wie etwa, daß »kompaktere Informationen kompaktere Rhythmen« erfordern, daß rhythmische Integration stärkeres »conversational involvement« indiziert, oder daß der Anapäst einen »dramatischen«,¹⁶ der Iambos aber einen »komödiantischen« Rhythmus auszeichnet – sind nicht notwendigerweise falsch; sie sind dennoch problematisch, weil ihre Abstraktheit, also ihre Entfernung von den lokalen interaktiven Funktionen des Rhythmus im jeweiligen Einzelfall, dem Wesen eines ikonisch-indexikalischen Zeichens widerspricht. Der Ansatz der Kontextualisierungsanalyse erfordert mehr als die Diskussion solcher hochgradig generalisierter »Bedeutungspotentiale« die genaue Analyse der tatsächlichen Rolle, die bestimmte prosodische Parameter im sequenziellen Ablauf der Konversation für die Konstitution interaktiver Bedeutungen spielen.

Bevor wir anhand einiger englischer und deutscher Beispiele *typische* lokale Verwendungsweisen von Rhythmus erläutern, müssen wir darlegen, wie wir konversationelle Rhythmen in Alltagsgesprächen feststellen und notieren wol-

15 Vgl. ausführlicher dazu: Gumperz 1982, 1992; Auer 1986 und 1992 sowie die Beiträge in Auer/Di Luzio (Hgg.) 1992.

16 Vgl. K. Müller 1992.

len. Um zu einer rhythmischen Analyse einer verbalen Interaktion zu gelangen, müssen mehrere, aufeinander aufbauende Schritte durchgeführt werden. Ausgehend von einer orthographischen Transkription des verbalen Ablaufs einer auf Ton- bzw. Videoband festgehaltenen Interaktion werden (i) die prosodischen Hervorhebungen (d.h. die Akzente) und die Pausen festgestellt, (ii) die regelmäßigen zeitlichen Abstände, falls vorhanden, zwischen diesen Hervorhebungen bestimmt und so isochrone Rhythmen identifiziert, und (iii) leere Schläge gegebenenfalls in diese Rhythmen integriert. Die so ermittelten Rhythmen und Tempi der Interaktion werden dann (iv) in einer rhythmischen Notation festgehalten.

(i) *Feststellen der prosodischen Hervorhebungen und Pausen.* Unter prosodischen Hervorhebungen verstehen wir Silben, die aufgrund einer prototypischen Konstellation phonetischer Merkmale wie erhöhter Tonhöhe, größerer Lautstärke und/oder längerer Dauer aus der Umgebung der sie angrenzenden Silben gewissermaßen hervorstehen, bzw. so wahrgenommen werden. Da »Hervorhebung« ein relativer Begriff ist, kann die Bestimmung von hervorgehobenen Silben auf unterschiedlichen Feinheitsstufen stattfinden. Nach unserer Erfahrung ist es in der Regel ausreichend – zumindest für das Englische und das Deutsche –, beim ersten Durchgang ungefähr die Ebene der phonologischen Phrase anzu steuern. Es sollte aber in jedem Fall mindestens eine Hervorhebung pro intonatorische Phrasierungseinheit festgelegt werden. Im ersten Stadium genügt es, die hervorgehobenen Silben im orthographischen Transkript, etwa durch Unterstreichungen, festzuhalten. Um ein Beispiel aus einem englischen Gespräch zu nehmen:

he was talking about infected meat being sold for humans instead of being processed for pet foods

Während prosodische Hervorhebungen am besten auditiv zu identifizieren sind, ist es unserer Meinung nach sinnvoll, Pausen instrumentalphonetisch zu erfassen. Die rein auditive Wahrnehmung von Pausen in gesprochener Sprache ist sehr heterogenen Beeinflussungen ausgesetzt. So haben psycholinguistische Untersuchungen gezeigt, daß z. B. beim Sprechen manche Pausen generell überhört, andere, etwa nach Silbendehnungen, hingegen regelrecht »hineingehört« werden. Um dies zu vermeiden, ist eine instrumentelle Analyse hilfreich. Als Pause kann eine Unterbrechung im Phonationsstrom von ungefähr einer Zehntelsekunde (oder mehr) gelten (vgl. Butcher 1981).¹⁷ Ähnliches gilt für den Übergang von einem Sprecher zum nächsten. Für die Analyse von Rhythmus und Tempo reicht eine Angabe über Pausenlänge in Zehntelsekunden aus. Z. B.:

17 Die Phonationsunterbrechung während der Verschlußphase eines Plosivlautes bleibt natürlich unberücksichtigt.

he was talking: about (0.1) infected (0.6) meat being sold (0.2) for humans instead of being processed for pet foods

Die instrumentelle Messung von Pausen ist der üblichen konversationsanalytischen Methodik der Stoppuhr bzw. des langsamen Zählens (»one thousand and one, one thousand and two«) aus Gründen der größeren Zuverlässigkeit vorzuziehen. Der objektive Zeitwert wird in einer späteren Phase unserer Analyse in Verbindung zum (inter-)subjektiven Rhythmus gesetzt, so daß Pausen im Endeffekt als subjektive Zeit dargestellt werden.

(ii) *Bestimmung von regelmäßigen Abständen zwischen prosodischen Hervorhebungen.* Zur Bestimmung der Isochronie zwischen prosodischen Hervorhebungen ist es hilfreich, beim wiederholten Vorspielen kürzerer Abschnitte eine Körperbewegung (z. B. Klopfen oder Kopfnicken) mit den vermeintlichen Schlägen zu synchronisieren. In einem bestimmten Datenausschnitt sollten probeweise unterschiedliche Abfolgen von prosodischen Hervorhebungen auf diese Weise »getestet« werden. Stellt sich eine kinetische Regelmäßigkeit beim Synchronisieren ein, so kann von einer *isochronen* Sequenz ausgegangen werden. Oder umgekehrt, paßt keine kinetische Regelmäßigkeit auf die vorgesehenen prosodischen Hervorhebungen, so ist die Abfolge als *nicht isochron* einzustufen. Sowohl größere als auch kleinere Abschnitte aus den Daten sollten in dieser Weise abgehört werden, da die Größe des »Fensters« bekanntermaßen zur unterschiedlichen Wahrnehmung von Gestalten führen kann.

Es kann bei diesem Verfahren vorkommen, daß rhythmische Gestalten sich überlappen oder ineinander verschachteln. Insbesondere bei der Verschachtelung kann ein kleineres rhythmisches Muster sich innerhalb eines größeren so etablieren, daß die größeren Intervalle in gleichmäßige kleinere Intervalle aufgeteilt werden.

Isochrone Sequenzen können zunächst im orthographischen Transkript durch gebundene Linien festgehalten werden. Zum Beispiel:

he was talking: about (0.1) infected (0.6) meat being sold (0.2)
for humans instead of being processed for pet foods

Zur Überprüfung der zunächst auditiv festgestellten Isochronien können die Intervalle zwischen den gehörten rhythmischen Schlägen auch instrumentell gemessen werden. Unter der Voraussetzung, daß der Silbenkern der Ankerpunkt für unsere zeitliche Wahrnehmung von Silben ist, sollte entweder von Vokalmitte zu Vokalmitte oder von Vokalansatz zu Vokalansatz gemessen werden. Bei dieser Meßweise werden nach unserer Erfahrung Abweichungen in der Dauer der Intervalle bis zu ungefähr 20 %, bei vorhandener Intonationsgrenze bzw. Sprecherwechsel bis zu ungefähr 30 % in der Regel nicht als nach-

teilig für perzeptuelle Isochronie empfunden.¹⁸ Bei größeren Abweichungen empfiehlt es sich, das auditive Urteil nochmals zu überprüfen. Aufgrund des Gestaltcharakters der rhythmischen Gefüge in der gesprochenen Sprache sollte allerdings der auditive Eindruck immer den Ausschlag geben. Völlig verkehrt wäre es, zuerst die Intervalle zwischen hervorgehobenen Silben akustisch zu messen und darauf bei regelmäßigen Dauerverhältnissen die Abfolge als (perzeptuell) isochron zu bezeichnen. Eine Reihe von akustisch gleich langen Intervallen muß nicht als rhythmische Gestalt wahrgenommen werden: Das als Gestalt empfundene Ganze ist immer *mehr* als die Summe seiner Teile.¹⁹

Angaben (in Zehntelsekunden) über die durchschnittliche Dauer der Intervalle in einer rhythmischen Gestalt sind für die Feststellung des Tempos²⁰ nützlich.

(iii) *Identifizierung der leeren Schläge*. Der nächste Schritt besteht darin, zu überprüfen, ob die im Transkript festgehaltenen Pausen sich zusammen mit benachbarten prosodischen Hervorhebungen in die rhythmischen Strukturen eingliedern lassen. Nach unserer Erfahrung können sich (abhängig vom Tempo) problemlos bis zu vier leere Schläge zwischen »gefüllten« Schlägen so einfügen, daß die zeitliche Regelmäßigkeit des Taktes nicht gestört wird. In diesem Fall wird das Ganze zu einem übergeordneten rhythmischen Gefüge. In unserem Beispiel fügt sich ein leerer Schlag zwischen den Hervorhebungen auf *infected* und *meat* ein:

he was talking: about (0.1) infected (0.6) meat being sold

Bei der Festlegung von leeren Schlägen ist es wichtig, daß mindestens zwei rhythmische Schläge ein geeignetes Intervall *vor* der Pause etablieren und daß

18 Diese Prozentzahlen ergeben sich aus einer am sprachlichen Material durchgeführten Pilotstudie zur perzeptuellen Isochronie (Couper-Kuhlen 1993, S. 25 ff.) und haben sich in der auditiv-akustischen Untersuchung von Rhythmus in einem großen Korpus englischer Alltagsgespräche bewährt. Sie bilden sozusagen einen praxisbezogenen Toleranzbereich für die Feststellung von Rhythmus in gesprochener Sprache. Die verhältnismäßig große Ausdehnung dieses Bereichs ist vermutlich auch darauf zurückzuführen, daß der Ankerpunkt für die menschliche Wahrnehmung von zeitlicher Regelmäßigkeit in der Sprache, das sogenannte *Perceptual center*, noch unzureichend erforscht ist (vgl. Hoequist 1983, Buxton 1983 und die Diskussion in Couper-Kuhlen 1993).

19 Zu den Faktoren, die den Eindruck der perzeptuellen Isochronie verhindern können, gehören z. B. eine semantisch/syntaktische bzw. eine intonatorische Grenze oder prägnantere Gestalten in der Umgebung. Für eine ausführliche Diskussion siehe Couper-Kuhlen 1993, S. 60 ff.

20 Der von uns verwendete Tempo-Begriff ist also in Bezug auf Rhythmus definiert und diesem nachgeordnet. Eine traditionellere, nicht-rhythmische Herangehensweise an Tempo (Artikulations- und Sprechgeschwindigkeit) wird z. B. in Uhmann (1992) vertreten, wo auch weitere Literatur zum nicht-rhythmischen Tempo-Begriff diskutiert wird.

mindestens ein gefüllter Schlag *nach* der Pause vorhanden ist. Andernfalls gibt es keine rhythmische Struktur, in die sich die Pause integrieren kann.

(iv) *Zur rhythmischen Notation.* Es erweist sich als nützlich, die festgestellten Rhythmen und damit verbundenen Tempi in einer teilweise ikonischen Notation darzustellen, in der die rhythmische Konstanz bzw. der rhythmische Wechsel im Verhältnis zum Wortlaut besser zum Ausdruck kommt. Bei isochronen Rhythmisierungen im Gespräch setzen wir schräge oder gerade Striche vor die prosodisch hervorgehobenen Silben, um die Schläge zu kennzeichnen, und adjustieren diese Striche so, daß sie genau untereinander auf dem Blatt stehen. Rechte Striche werden hinzugefügt, um einen optischen Eindruck des Tempos zu geben. Arythmische Passagen werden ohne Striche und fortlaufend (wie Prosa) geschrieben. Die rhythmischen Strukturen in unserem Beispiel haben z.B. ein moderates Tempo (circa 0.6 Sek.), das so dargestellt werden würde:

he was	
/ tálkin g: a-	/
/ bóut (0.1) in-	/
/ fécted	/
/ ^ (0.6)	/
/ méat bein g sóld (0.2) for	
/ húmans, in-	/
/ stéad of being	/
/ prócessed for	/
/ pét foods.	/

Da die Hervorhebung auf *sold* zeitlich nicht mehr in die vorhergehende rhythmische Struktur hineinpaßt (die Sprecherin verlangsamt an dieser Stelle), muß mit *humans* eine neue Struktur beginnen. Bei einem schnelleren bzw. gemäßigteren Tempo würden die Striche etwas näher zusammenstehen bzw. weiter ausinanderrücken.

Zu welchen Zwecken werden nun Rhythmus und Tempo von Gesprächsteilnehmern in der Interaktion eingesetzt? In unseren Untersuchungen hat sich herausgestellt, daß sie auf mehreren Ebenen relevant sein können. Einmal auf der »mechanischen« Ebene des *Turntaking*: Hier fungiert ein isochroner Schlag im Englischen als Taktgeber für den Einsatz des nächsten Sprechers. Er trägt also zur zeitlichen Koordinierung der Redebeiträge im Gespräch bei. Zweitens werden Veränderungen von Rhythmus und Tempo als Kontextualisierungshinweise auf der *Aktivitätsebene* wichtig. Insbesondere werden Abweichungen vom zeitlich wohlplazierten Turneinsatz im Gespräch als »markiert« behandelt und so im jeweiligen lokalen Kontext für den Interpretationsprozeß relevant. Drittens werden Rhythmus und Tempo auf *größeren organisatorischen Ebenen* zur Einheitsmarkierung im Gespräch eingesetzt. Sie tragen z. B. zur Aushandlung von Themenbeendigungen bei und werden zur Markierung von Beendigungs-

phasen (»closings«) benutzt. In den nächsten Abschnitten wollen wir diese interaktiven Funktionen exemplarisch darstellen.

4.1 Rhythmus und Turntaking

Es ist eines der zentralen Probleme im Gespräch, darüber zu entscheiden, ob und wann der andere Gesprächsteilnehmer beim Reden zu einem möglichen Abschlußpunkt kommt. Im Zweiergespräch ist das der Punkt, an dem man selber ohne soziale Sanktionen das Wort ergreifen kann. Die Forschung über Sprecherwechsel in Alltagsgesprächen hat gezeigt, daß Gesprächsteilnehmer auf eine Reihe linguistischer und paralinguistischer Signale achten, um diesen Zeitpunkt zu erkennen. Syntax und Intonation leisten hier wichtige Hilfen, indem sie Konstituenten- bzw. Phrasierungsgrenzen erkennbar werden lassen. Darüber hinaus ermöglichen sie die *Projektion* dieser Grenzen dadurch, daß sie sozusagen einen syntaktischen bzw. melodischen Bogen in der Zeit spannen, in dessen Laufzeit die Vorhersagbarkeit der noch verbleibenden syntaktischen bzw. melodischen Komponenten ständig zunimmt (Auer, im Druck). So ist das (mögliche) Ende eines Redebeitrags vorauszusehen. Aber auch das Blickverhalten des Sprechenden kann darüber Auskunft geben, ob und wann ein Sprecherwechsel beabsichtigt ist: Der Sprecher richtet typischerweise kurz vor Ende des Redebeitrags den Blick auf den Rezipienten, um den beabsichtigten Wechsel anzukündigen.

Vieles deutet nun darauf hin, daß auch der Rhythmus zur Ankündigung eines intendierten Sprecherwechsels verwendet wird. Nehmen wir den Abschnitt, dem unsere Beispielaussage entnommen worden ist (es handelt sich um eine englische Rundfunksendung, bei der die Zuhörer im Studio anrufen):²¹

SA: *hello. Perm, (0.1) I'm ringing, (0.1) Perm about (0.5) a (0.1) a talk by a health inspector. on your programme a couple of days ago. (0.5) he was*

/ tálking: a- /
/ bóut (0.1) in- /
/ fécted /
/ ^ (0.6) /
/ méat being sóld (0.2) for

/ húmans, in- /
/ stéad of being /
/ prócessed for /
/ pét foods. /
/ óh yes. /

→ DH:

21 Die Teile des Transkripts, die keiner rhythmischen Analyse unterworfen worden sind, werden kursiv gesetzt.

Hier fällt auf, daß die Sprecherin nicht nur einen möglichen syntaktischen, thematischen und intonatorischen Abschlußpunkt nach »instead of being processed for pet foods« erreicht hat; dieser Punkt wird auch durch eine prägnante rhythmische Struktur angekündigt. Die vorhergehende Teiläußerung »talking about infected meat being sold« könnte zwar syntaktisch und thematisch ebenfalls als abgeschlossen gelten, wirkt aber intonatorisch unvollständig und ist rhythmisch nicht prägnant: Der vorhandene Rhythmus, obwohl isochron, hat einige unerwartete Längungen und Pausen und ist nicht sofort erkennbar. Hinzukommt, daß dieser Rhythmus aufgrund des verspäteten *sold* zusammenbricht; d. h., der Fluß des Redens gerät vorübergehend ins Stocken. Kennzeichnend für ein bald zu erwartendes Turnende ist aber, daß nach möglicherweise zeitweilig wechselnden Rhythmen eine sehr klare rhythmische Gestalt produziert wird, die mit den auslaufenden melodischen und syntaktischen Bögen zusammenfällt.

Ein prägnantes (oft im Vergleich zum Vor-Text etwas beschleunigtes) rhythmische Muster erfüllt an dieser Stelle eine doppelte Funktion. Zum einen wird ein mögliches Ende des jetzigen Turns indiziert. Zum anderen wird prospektiv ein Zeitraster aufgebaut, nach dem sich der Ansatz des nächsten Turns richten kann. Der Sprecher sorgt sozusagen nicht nur für den eigenen Abgang, sondern auch für den Übergang zum nächsten.²²

In der Tat ist der Zeitpunkt, an dem der nächste Sprecher das Wort übernimmt, nicht dem Zufall überlassen. In unserem Beispiel benutzt der Moderator das Raster der Anruferin, um den Beginn seiner Folgeäußerung in der Zeit zu plazieren: Die erste hervorgehobene Silbe seines Turns ist isochron mit den letzten Schlägen ihres Turns. Wir sagen, der zweite Turn und der erste sind *rhythmisch integriert*. Wären die letzten rhythmischen Schläge der Anruferin in einem schnelleren Tempo gekommen (mit z. B. 0.3 statt 0.6 Sek. Abstand), so wäre zu erwarten, daß der Moderator seinen Einsatz entsprechend früher gebracht hätte. Sprecher gehorchen also nicht einem rigiden Zeitgeber etwa von 1 Sekunde (Jefferson 1989) für die Übernahme des Turns, sondern passen sich dem Redefluß, d. h. dem Rhythmus und Tempo des Vorgängers, flexibel an. Die letzten Schläge des Vorgängerturns geben dem neuen Sprecher die Möglichkeit, sich im Rhythmus einzuüben, so daß es, wie Sacks/Schegloff/Jefferson (1974) bemerken, in der Mehrzahl der Fälle zu einem erstaunlich glatten zeitlichen Übergang kommt.

Gibt es unbetonte Auftaktsilben am Anfang eines neuen Turns, müssen sie so plaziert werden, daß die rhythmische Koordinierung gewährleistet wird. So kann es z. B. vorkommen, daß unbetonte Auftaktsilben des nächsten Sprechers sozusagen ins Territorium des vorigen Sprechers wandern müssen. Es kommt dann zu Überlappungen, wie im folgenden Ausschnitt aus einem englischen *face-to-face* Familiengespräch:

22 Näheres hierzu bei Couper-Kuhlen 1991 und 1993, S. 126ff.

E: /éven Conti-
 /néntal ones. /
 → P: I'll tell
 [you a- /
 /nóther thing ...

Umgekehrt können Lücken entstehen, wenn das Tempo relativ langsam ist und es keine oder nur wenige unbetonten Silben nach dem letzten Schlag des Vorgängers bzw. vor dem ersten Schlag des neuen Sprechers gibt. Solche »Pausen« sind aber für die Aushandlung von interaktiven Bedeutungen irrelevant, denn sie entstehen erzwungenermaßen aus rhythmischen Gründen. Vgl. aus dem oben erwähnten Telefongespräch:

SA: *if it's so bad*
 (0.2)
 DH: *yeah*
 (0.9)
 SA: *it (.) it* (0.2)
 /shóuldn't be sóld at /
 → DH: /áll! (0.2) /
 I /
 /quíte agréé

Eine solche Lücke ist für die Interaktionsteilnehmer weit weniger relevant als z. B. die folgende:

SA: *the fact that it was just taken for granted that* (0.3)
because this was so horrible it was in inverted commas (0.1)
 /ó n l y /
 /fít for /
 /pét food. /
 → DH: / ^ (0.5) /
 /yés. I – I'm not ...

In diesem Fall fällt die Unterbrechung des Redestroms mit einem leeren Schlag zusammen. Dadurch, daß der rhythmische Takt vorübergehend unterbrochen ist, bekommt eine solche Pause mehr Gewicht.²³ Laien hören Pausen, die mit leeren Schlägen zusammenfallen, eher als solche, die keine leeren Schläge bilden.²⁴ Für den Analytiker heißt das, daß erst der rhythmische Wert einer Unterbrechung im Redestrom darüber entscheiden kann, ob sie zu vernachlässigen oder zu interpretieren ist.

23 Natürlich neigen längere Pausen eher als kürzere Pausen dazu, mit rhythmischen Schlägen zusammenzufallen. Pausen im mittleren Dauerbereich (d.h. zwischen 0.4–0.6 Sek.) enthalten jedoch annähernd genauso häufig leere Schläge wie keine.

24 Dieses geht unter anderem aus einer informellen Befragung hervor, die in Couper-Kuhlen 1993, S. 52 geschildert wird.

4.2 Rhythmus und Aktivität

Die rhythmische Integration eines zweiten Turns mit dem ersten ist in unserem englischen Gesprächsmaterial eher die Regel als die Ausnahme. An den rhythmisch integrierten Übergängen gibt es auch keine Zeichen für eine Gesprächsstörung. Aus diesen Gründen nehmen wir an, daß es sich um den *unmarkierten* Fall, um eine implizite »Norm« handelt, und daß sich markierte Fälle, sozusagen »Abweichungen«, vor allem im Bezug zu diesem bemerkbar machen.

Bei Redebeiträgen, die als potentiell abgeschlossen gelten, sind zwei Arten von *markierten* zeitlichen Übergängen denkbar: Der Fall der zeitlichen *Verspätung* und der der zeitlichen *Verfrühung*. Diese zwei Fälle lassen sich jeweils in zwei Unterklassen aufteilen, je nachdem ob der verspätete oder der verfrühte Schlag nur eine vorübergehende Störung darstellt oder aber den Rhythmus völlig unterbricht. Sämtliche so markierten Fälle kommen in Alltagsgesprächen vor.

Nehmen wir das Beispiel der verspäteten Einsätze, so stellt sich heraus, daß die damit verbundene Unterbrechung des rhythmischen Musters einen wesentlichen Hinweis auf die interaktive Bedeutung des Redebeitrags geben kann. So z. B. im folgenden Ausschnitt aus dem selben Telefongespräch:

DH: *you're saying that if it- that you object to the idea that if it's
unfit for human consumption it is*
/ therefore as- /
/ súmed to be /
/ fit for /
/ pét's con- /
/ súmption.

→ SA: (0.6) y- yes ?uh ?uh (?0.1) u:n:fi:t (?0.4) e:rm (0.2)

Berücksichtigen wir nur den Wortlaut der Anruferin (»yes unfit«), so könnten wir zu der Behauptung verleitet werden, es handele sich hier um eine Zustimmung. Jedoch legen die Verzögerungspartikeln (»uh«, »erm«) und der Rhythmus dieses Beitrags eine andere Inferenz nahe, nämlich daß die Anruferin erhebliche Probleme hat, der Formulierung des Moderators zuzustimmen.²⁵ Um welche Probleme es sich handelt, wird erst aus einem späteren Redebeitrag klar: »I'm not talking about scraps and things like that ...«, »I'm not saying that other stuff shouldn't go into pet food, meat by-products and things like that«. Jedoch wird das Vorhandensein eines Problems schon an dieser Stelle rhythmisch angedeutet, indem der Takt, der vom Moderator am Ende seines Turns klar etabliert wird, von der Anruferin nicht aufrecht gehalten wird. Somit wird ein Bruch in der vom Moderator initiierten Aktivität des Formulierens regelrecht sichtbar, d. h. hörbar gemacht.

²⁵ Verzögerungspartikeln allein reichen nicht aus, um eine solche Inferenz auszulösen, denn wie wir an anderer Stelle gezeigt haben (Couper-Kuhlen/Auer 1991), können sie genau so gut benutzt werden, um den Rhythmus aufrechtzuhalten.

Es kann aber vorkommen, daß ein nur vorübergehend unterbrochener Rhythmus auf einer lokalen sequentiellen Ebene ein Problem andeutet, während auf einer übergeordneten Ebene der interaktive Zusammenhalt (das *ensemble* im Sinne von Scollon 1982) aufrecht erhalten wird. Dies ist der Fall, wenn die Verzögerung mit einem oder mehreren leeren Schlägen verbunden ist, z. B. im folgenden Ausschnitt aus einem englischen *face-to-face* Familiengespräch:

- C: what's that guy called now,
 Z: what works in Comet's.
 C: yeah!
 Z: Jack Barnes;
 C: Jack Barnes, he told me
 / Amstrad was /
 / rúbbish! /
 → Z: / ^ (1.4) well /
 / I don't know; /
 C: he /
 / sélls them.

Es handelt sich hier um eine Bewertungssequenz: Cs indirekt wiedergebene Bewertung »he told me Amstrad was rubbish« macht eine Zustimmung des Rezipienten erwartbar (Pomerantz 1984). Jedoch deutet schon die Verzögerung, mit der Zs Erwiderung einsetzt, darauf hin, daß sie nicht ohne weiteres zustimmen kann. Diese Inferenz wird selbstverständlich auch verbal durch die *hedges* »well«, »I don't know« zum Ausdruck gebracht. Sie wird durch die Reaktion von C im nächsten Turn bestätigt. Denn Cs Beitrag »he sells them« läßt sich als Unterstützung für die Jack Barnes zugeschriebene Bewertung verstehen. Diese Reaktion ist nur dadurch zu erklären, daß C den vorhergehenden Beitrag von Z als Nichtübereinstimmung aufgefaßt hat. Zu einem wesentlichen Teil schuldet C diese Inferenz der rhythmischen Verspätung von Zs Einsatz. Gleichzeitig muß aber festgestellt werden, daß im Vergleich etwa zum vorherigen Beispiel der Rhythmus durch die Verzögerung von Z nicht völlig zusammenbricht, sondern durch die Koordinierung des Akzents auf »I« mit dem übernächsten Schlag rückwirkend auf einer höheren Ebene bestätigt wird. Trotz der lokalen Störung bleibt also die globale rhythmische Struktur hier intakt. Damit liefern die Teilnehmer indexikalisch den Hinweis, daß sie eine Aktivität (die des Bewertens bei entgegengesetzten Meinungen) kollaborativ durchführen. Das Zusammentreffen von verspätetem Einsatz und dispräferiertem Turn wie im vorigen Beispiel darf allerdings nicht so verstanden werden, daß die rhythmische Verzögerung eine *conditio sine qua non* der Dispräferenz und somit als ein distinktives Merkmal des Präferenzstatus aufzufassen sei, obwohl dies öfters in der Literatur behauptet wird (vgl. z. B. Levinson 1983). Mehrere Gründe sprechen dagegen: Zum einen ist nicht jeder dispräferierte Turn (im Sinne von Präferenzen als Erwartungsstrukturen, die auf kulturellen Konventionen, z. B. Höflichkeit, beruhen) rhythmisch verspätet. Mancher rhythmisch integrierte

Turn »verstößt« gegen die Präferenz, etwa in Argumentationen oder bei Reparaturen. Darüber hinaus können dispräferierte Turns auch *zu früh* einsetzen: Hier kann die rhythmische Verfrühung je nach Aktivität Ungeduld, Übereifer, Attacke/Gegenwehr, usw. (im Gegensatz zu Muße, Nachdenklichkeit, Desinteresse, Schwierigkeiten, usw. bei rhythmischer Verzögerung) kontextualisieren. Drittens kann der rhythmisch verspätete Einsatz sogar das »präferiertere« Format sein. Dies ist z. B. bei »hot news«-Ankündigungen der Fall, wo eine deutlich spürbare Pause zwischen dem ersten und dem zweiten Teil der Sequenz den intendierten Überraschungseffekt nur noch erhöht, vgl. folgendes Beispiel aus der erwähnten englischen Rundfunksendung:

- S: *apparently; according to this chap; who's an English doctor; the chap in Spandau is*
 / n ó t; /
 / R ú d o l p h /
 / H é s s. /
 → H: / ^ (1.2) /
 / ^ /
 / y ó u /
 / w h á - ?
 S: it is /
 / n ó t R ú d o l p h /
 / H é s s ! /
 H: / s ó t h i s g u y /
 / s á y s;

Die rhythmische Verzögerung des Rezipienten kontextualisiert hier seine Überraschung, was durchaus im Sinne des Berichterstatters sein dürfte. Eine solche Reaktion ist beim Empfang von »hot news« (in unserer Kultur) erwartbar. In diesem Fall also scheinen die Präferenzformate in bezug auf Verzögerung gerade umgekehrt zum »Normalfall« zu sein.

Wir schließen daraus, daß der Rhythmus nicht als Reflex des Präferenzstatus eines Turns aufzufassen ist, sondern als zusätzliche Komponente, die in bestimmten Handlungssequenzen quer zu sonst kulturell erwartbaren Reaktionen eine dritte Dimension des *keying* einbringt. Wie der damit angedeutete *key* genau auszulegen ist, hängt vom jeweiligen sequentiellen Kontext und von dem mit der Handlung assoziierten, konventionalisierten Interpretationsrahmen ab.²⁶

4.3 Rhythmus und Gesprächsorganisation

Während sich die Funktionen, die wir bisher betrachtet haben, auf der eher lokalen Ebene des Turntaking bzw. der Aktivitätstypen ansiedeln, gibt es auch

26 Für eine ausführlichere Diskussion siehe Couper-Kuhlen 1993, S. 254ff.

zurückgekehrt werden kann, muß aber abgeklärt werden, ob die Teilnehmer damit einverstanden sind. Dies geschieht durch einen »turn pass«, der sich rhythmisch durch einen leeren Schlag bemerkbar macht (erster Pfeil). Nachdem keine weitere Frage über David gestellt wird, kann das Gespräch zum Thema der Weihnachtsfeier zurückkommen. Hier leitet nun T als Haupterzählerin durch den Vorlauf (preclosing) »so« einen Themenabschluß ein. Daraufhin haben die Gesprächspartner die Möglichkeit, weitere Aspekte des übergeordneten Themas nachzutragen. Da aber nochmals durch leere Schläge (zweiter Pfeil) erkennbar gemacht wird, daß das Wort zu diesem Thema nicht gewünscht wird, kann T mit einem zusammenfassenden »it was a lovely evening« einen Themenabschluß vorschlagen. Ihr Angebot wird dann durch den darauffolgenden leeren Schlag (dritter Pfeil) von den anderen Teilnehmern gebilligt. Damit hat B grünes Licht, ein neues Thema einzuführen.

Wichtig erscheint uns bei solchen Themenbeendigungssequenzen, daß ohne ein Zeitraster der Turnverzicht nicht erkennbar wäre. Ausgehend vom unmarkierten Fall der Turnübergabe in englischen Alltagsgesprächen scheint uns der Rhythmus dieses Raster zu liefern.

Die höchste Komplexitätsebene, auf der wir die kontextualisierende Funktion von Rhythmus und Tempo nachweisen konnten, ist die der Gesprächsbeendigung. Eine Untersuchung von deutschen Telefon-Beendigungssequenzen (Auer 1990 a, 1990 b) ergab die folgenden Regelmäßigkeiten:

- Beim Übergang in die Beendigungssequenz (beginnend mit dem ersten »Beendigungsvorlauf«/preclosing) oder innerhalb dieser Sequenz werden selbst solche Gesprächspassagen, die vorher nicht rhythmisiert (isochron) waren, normalerweise einem regelmäßigen isochronen Rhythmus unterworfen.
- Innerhalb dieses isochronen Musters verändert sich in der Regel – je nach Ausdehnung der Beendigungssequenz u.U. auch mehrmals – das Tempo; diese Tempoveränderungen korrespondieren mit der Bewegung der beiden Gesprächspartner auf das Gesprächsende zu (Verschnellerung) bzw. von diesem Ende weg (»moving out of closings« – Verlangsamung).
- Das »accelerando« erreicht seinen Höhepunkt mit dem abschließenden Austausch von Grüßen.

Ein typisches Beispiel – das Ende eines Telefongesprächs zwischen Vater und Tochter – mag das Gesagte demonstrieren:

M:	also f- ich wünsch dir viel Vergnüg	[	n	
F:			hn,	
M:	und s=is gút daß der Mammi was híngeschrieben has			
	da			
→	/ fréut se sich	/		
	/ séhr daß du kürz	/		
	/ hèimgekommen	[	bist	/
F:			já = à,	
	/ ^	/		
M:	/schréib noch drúnter			

→		/ Páppi hat mit mir /
		/ áuch gesprochen /
	F:	/ já gu:t
→	M:	/ dánke /
		/ tschúß /
	F:	/ tscháu /
	M:	/ sérvus /

Der Gesprächsausschnitt beginnt unrhythmisch, aber während M die Beendigung einleitet (vgl. den Beendigungsvorlauf »also«) verfällt er zunächst in einen langsamen Rhythmus (1. Pfeil). Typischerweise wird die beginnende Abschlußrhythmisierung von demjenigen Gesprächsteilnehmer vorangetrieben, der auch die Beendigung einleitet, also die Initiative zum Abschluß der Interaktion ergreift. Die Partnerin F synchronisiert zunächst ihre Äußerungen noch nicht mit dem von M aufgebauten rhythmischen Muster; ihr asynchrones »já = á« und der folgende leere Schlag bringen F zunächst »aus dem Takt«; auf inhaltlicher Ebene entspricht diesem Verlust an Rhythmisierung das Einfügen weiterer Schlußverabredungen (»schreib noch drunter ...«), die den Übergang in die terminale Grußsequenz zunächst aufschieben. M baut jedoch einen zweiten, etwas schnelleren Rhythmus auf (2. Pfeil), in den sich nun auch F integriert (»ja gu:t«). Gleichsam als ob diese rhythmische Ensemble-Bildung M den Schwung gegeben hätte, um nun die Beendigungssequenz weiterzutreiben, liefert er nun durch sein »danke tschúß« nicht nur das erste Glied der terminalen Grußsequenz, sondern auch die Bausteine für einen wesentlich schnelleren Rhythmus, der die letzte Phase der Beendigungssequenz einleitet (3. Pfeil). Es folgen zwei weitere Grüße seitens Fs und erneut Ms, womit das Gespräch beendet ist.²⁷

Erhöhte Rhythmisierung und damit einhergehende Verschnellerung des rhythmischen Tempos finden sich regelmäßig in Telefonbeendigungen, bei denen die Interaktion nicht einseitig und/oder übereilt abgebrochen wird. Je länger die Beendigungssequenz, um so perfekter ist normalerweise auch die Rhythmisierung. Es ist, als ob die Beteiligten noch einmal ein Stück gemeinsam tanzen wollten, bevor sie sich trennen. Tatsächlich läßt sich beobachten, daß die Weigerung des einen Partners, sich auf den Rhythmus des anderen einzulassen, zur Expansion des Gesprächs insgesamt führt: Der die Beendigung einleitende Gesprächsteilnehmer versucht dann immer wieder, mit neuem Anlauf und neuem Schwung den anderen zum »gemeinsamen Tanz« zu gewinnen.

27 Im Gegensatz zu den von Schegloff/Sacks 1973 gefundenen Ablaufschema enthalten – deutsche – Abschlußsequenzen sehr oft mehr als zwei abschließende Grüße; bei familiären bis intimen Gesprächen läßt sich die Grußsequenz fast beliebig ausdehnen.

5. Schlußbemerkung

Wir haben im letzten Abschnitt versucht, einen Eindruck von der Reichweite rhythmischer Erscheinungen in der Interaktion zu geben. Diese Rhythmen erfüllen vielfältige interaktive Funktionen auf unterschiedlichen Organisationsebenen des Gesprächs. Rhythmische Integration, wie wir sie an Turnübergängen feststellen konnten, ist ein *interaktives* Phänomen, weil sie erst im Zusammenspiel von mindestens zwei Interaktanten zustande kommt. Aber auch die »Bedeutungen« (Funktionen), die mit diesen rhythmischen Strukturen verbunden sind, müssen als interaktive Hervorbringung verstanden werden: um als solche zu »gelten«, bedürfen sie der gegenseitigen Aushandlung. Daß diese Bedeutungen so heterogen sind, liegt vor allem daran, daß die dazu verwendeten Zeichen indexikalischen Charakter haben. Sie werden kontextstiftend (kontextualisierend) eingesetzt, sind aber selber in ihrer Funktion und lokalen Bedeutung abhängig von ihrem jeweiligen Kontext.

Der rhythmische Ansatz, den wir hier vorgestellt haben, ist bislang hauptsächlich bei der Analyse von Turnübergängen eingesetzt worden, und zwar vorwiegend für das Englische und das Deutsche. Er ist aber genauso auf den Rhythmus monologischen Sprechens anwendbar. Auch eine Ausdehnung auf typologisch unterschiedliche Sprachen läßt sich gut vorstellen.²⁸ Die Untersuchung von Rhythmus und Tempo in konventioneller Alltagssprache könnte so dazu beitragen, die Zeit in die Sprache als Objekt der Linguistik zurückzuholen.

Literatur

- Abercrombie, David (1967): *Elements of General Phonetics*, Edinburgh: University Press.
- Abercrombie, David (1971): Some functions of silent stress, in: A.J. Aitken/Angus McIntosh/Hermann Pálsson (Hgg.), *Edinburgh Studies in English and Scots*, London: Longman, S. 147–156.
- Allen, George D. (1975): Speech rhythm: its relation to performance universals and articulatory timing, in: *Journal of Phonetics* 3, S. 75–86.
- Allen, George D./Hawkins, S. (1980): Phonological rhythm: Definition and development, in: O.H. Yeni-Komshian (Hg.), *Child Phonology*, New York, S. 227–256.
- Auer, Peter (1986): Kontextualisierung, in: *Studium Linguistik* 19, S. 22–47.
- Auer, Peter (1990a): Rhythmic integration in phone closings, in: B. Conein/M. de Fornel/L. Quéré (Hgg.), *Les formes de la conversation*, Paris: CNET, S. 203–250.
- Auer, Peter (1990b): Rhythm in telephone closings, in: *Human Studies* 13, S. 361–392.
- Auer, Peter (1991): *Phonologie der Alltagssprache*, Berlin: de Gruyter.
- Auer, Peter (1992): Introduction: John Gumperz' approach to contextualization, in: P. Auer/Aldo di Luzio (Hgg.), *The Contextualization of Language*, Amsterdam: Benjamins, S. 1–38.

28 Als Beispiel hierfür seien (aus unserem Projekt) die Arbeiten von F. Müller 1989, 1991 für das Italienische genannt.

- Auer, Peter (1993): Is a rhythm-based typology possible? KontRI Working Papers, Universität Konstanz, Fachgruppe Sprachwissenschaft.
- Auer, Peter (im Druck): On the prosody and syntax of turn continuations, in: Elizabeth Couper-Kuhlen/Margret Selting (Hgg.), *Prosody in Conversation: Interactional studies*, Cambridge: University Press.
- Auer, Peter (Hg.) (in Vorb.): *Language in Real Time: The rhythm and tempo of verbal interaction*, Cambridge: University Press.
- Auer, Peter/di Luzio, Aldo (Hgg.) (1992): *The Contextualization of Language*. Amsterdam: Benjamins.
- Auer, Peter/Uhmann, Susanne (1988): Silben- und akzentzählende Sprachen, in: *Zeitschrift für Sprachwissenschaft* 7, S. 214–259.
- Butcher, Andrew (1981): Aspects of the Speech Pause: Phonemic correlates and communicative functions, *Arbeitsberichte Nr. 15*, Institut für Phonetik, Universität Kiel.
- Buxton, Hilary (1983): Temporal predictability in the perception of English speech, in: A. Cutler/D. R. Ladd (Hgg.), *Prosody: Models and Measurements*, Berlin: Springer, S. 111–121.
- Chapple, E. D. (1939): Quantitative analysis of the interaction of individuals, in: *Proceedings of the National Academy of Academic Sciences* 25, S. 57–67.
- Condon, W. S. (1980): The relation of interactional synchrony to cognitive and emotional processes, in: M. R. Key (Hg.), *The Relationship of Verbal and Nonverbal Communication*, The Hague: Mouton, S. 49–65.
- Condon, W. S./Ogston, W. D. (1967): A segmentation of behaviour, in: *Journal of Psychiatric Research* 5, S. 221–235.
- Couper-Kuhlen, Elizabeth (1991): A rhythm-based metric for turntaking, in: *Proceedings of the 12th International Congress of Phonetic Sciences*, Vol. 1, Aix-en-Provence, Service des Publications, Université de Provence, S. 275–278.
- Couper-Kuhlen, Elizabeth (1993): *English Speech Rhythm. Form and function in everyday verbal interaction*, Amsterdam: Benjamins.
- Couper-Kuhlen, Elizabeth/Auer, Peter (1991): On the contextualizing function of speech rhythm in conversation: Question-answer sequences, in: Jef Verschueren (Hg.), *Levels of Linguistic Adaptation. Selected Papers of the 1987 International Pragmatics Conference*, Vol. II, Amsterdam: Benjamins, S. 1–18.
- Couper-Kuhlen, Elizabeth/Selting, Margret (Hgg.) (im Druck): *Prosody in Conversation: Interactional Studies*, Cambridge: University Press.
- Edwards, Viv/Sienkewicz, Thomas J. (1990): *Oral Cultures Past and Present: Rappin' and Homer*, Cambridge: Blackwell.
- Erickson, Frederick (1982): Money tree, lasagna bush, salt and pepper: Social construction of topical cohesion in a conversation among Italian Americans, in: Deborah Tannen (Hg.), *Analyzing Discourse: Text and talk*, Washington: GURT, S. 43–70.
- Erickson, Frederick (1992): They know all the lines: Rhythmic organization and contextualization in a conversational listing routine, in: Peter Auer/Aldo di Luzio (Hgg.), *The Contextualization of Language*, Amsterdam: Benjamins, S. 365–398.
- Erickson, Frederick/Shultz, Jeffrey (1982): *The Counselor as Gatekeeper*, New York: Academic Press.
- Finnegan, Ruth (1977, 1992²): *Oral Poetry*, Bloomington: Indiana University Press.
- Fraïsse, Paul (1974): *Psychologie du rythme*, Paris: Presses universitaires de France.
- Gabelentz, Georg von der (1891, 1901²): *Die Sprachwissenschaft, ihre Aufgaben, Methoden und bisherigen Ergebnisse*, Leipzig: Weigel.
- Goody, Jack (1990): Technologies of the intellect: writing and the written word, Working Paper Nr. 5, Projektgruppe Kognitive Anthropologie, MPI Nijmegen.
- Gumperz, John (1980): The role of dialect in urban communication, in: *Zeitschrift für Dialektologie und Linguistik (Sonderheft Dialektologie)*.

- Gumperz, John (1982): *Discourse Strategies*, Cambridge: University Press.
- Gumperz, John (1992): Contextualization and understanding, in: A. Duranti/Ch. Goodwin (Hgg.), *Rethinking Context*, Cambridge: University Press, S. 229–252.
- Hoequist, Charles E. (1983): The perceptual center and rhythm categories, in: *Language and Speech* 26, S. 367–376.
- Jaffe, Joseph/Feldstein, Stanley (1970): *Rhythms of Dialog*, New York: Academic Press.
- Jefferson, Gail (1989): Notes on a possible metric which provides for a 'standard maximum' silence of approximately one second in conversation, in: D. Roger/P. Bull (Hgg.), *Conversation: An interdisciplinary perspective*, Clevedon: Multilingual Matters, S. 166–196.
- Jones, Mari Riess (1986): Attentional rhythmicity in human perception, in: J. R. Evans/M. Clynes (Hgg.), *Rhythm in Psychological, Linguistic and Musical Processes*, Springfield, Ill.: Charles C. Thomas, S. 13–40.
- Lehiste, Ilse (1977): Isochrony reconsidered, in: *Journal of Phonetics* 5, S. 253–263.
- Levinson, Stephen C. (1983): *Pragmatics*, Cambridge: University Press.
- Martin, J. G. (1972): Rhythmic (hierarchical) versus serial structure in speech and other behavior, in: *Psychological Review* 79, S. 487–509.
- Müller, Frank (1989): *Rhythmische Alternation im Italienischen*, KontRI Working Paper Nr. 10, Universität Konstanz, Fachgruppe Sprachwissenschaft.
- Müller, Frank (1991): *Metrical emphasis: Rhythmic scansions in Italian conversation*, KontRI Working Paper Nr. 14, Universität Konstanz, Fachgruppe Sprachwissenschaft.
- Müller, Klaus (1992): Theatrical moments: On contextualizing funny and dramatic moods in the course of telling a story in conversation, in: Peter Auer/Aldo di Luzio (Hgg.), *The Contextualization of Language*, Amsterdam: Benjamins, S. 199–222.
- Payne, M. C./Holzman, Th. G. (1986): Rhythm as a factor in memory, in: J. R. Evans/M. Clynes (Hgg.), *Rhythm in Psychological, Linguistic and Musical Processes*, Springfield, Ill.: Charles C. Thomas, S. 41–54.
- Pomerantz, Anita (1984): Agreeing and disagreeing with assessments: Some features of preferred/dispreferred turn shapes, in: J. Maxwell Atkinson/John Heritage (Hgg.), *Structures of Social Action. Studies in conversation analysis*, Cambridge: University Press, S. 57–101.
- Pompino-Marschall, Bernd (1989): On the psychoacoustic nature of the P-center phenomenon, in: *Journal of Phonetics* 17, S. 175–192.
- Rosenfeld, Howard M. (1981): Whither interactional synchrony?, in: K. Bloom (Hg.), *Prospective Issues in Infant Research*, Hillsdale, N.J., S. 71–97.
- Sacks, Harvey/Schegloff, Emanuel A./Jefferson, Gail (1974): A simplest systematics for the organization of turn-taking for conversation, in: *Language* 50, S. 696–735.
- Saran, Franz (1907): *Deutsche Verslehre*, München.
- Saussure, Ferdinand de (1915, 1972): *Cours de linguistique générale*, Publié par Charles Bally und Albert Secheheye, Critical edition by Tullio de Mauro, Paris: Payot.
- Schegloff, Emanuel A./Sacks, Harvey (1973): Openings up closings, in: *Semiotica* 8, S. 289–327.
- Scollon, Ron (1982): The rhythmic integration of ordinary talk, in: Deborah Tannen (Hg.), *Analyzing Discourse: Text and talk*, Washington: GURT, S. 335–349.
- Selkirk, Elisabeth O. (1984): *Phonology and Syntax: The relation between sound and structure*, Cambridge, Mass.: MIT Press.
- Selting, Margret (im Druck): *Prosodie im Gespräch*, Tübingen: Niemeyer.
- Silverstein, Michael (1979): Language structure and linguistic ideology, in: P. R. Clyne/W. F. Hanks/C. L. Hofbauer (Hgg.), *The Elements: A parsession on linguistic units and levels*, Chicago Linguistic Society, S. 193–247.
- Sonnenschein, E. A. (1925): *What is rhythm?*, Oxford: Basil Blackwell.

Uhmann, Susanne (1992): Contextualizing relevance: on some forms and functions of speech rate changes in everyday conversation, in: Auer/Di Luzio (Hgg.) (1992), S. 297–336.

Wundt, Wilhelm (1911–1912³): Völkerpsychologie, Leipzig: Engelmann.

Rhythm and Speech in Conversational Everyday Speech

Summary

The authors show that the most important impulses in this field came from »Gestalttheorie« in psychology: rhythm is a perceptive »Gestalt« of speech. After some methodological discussion on the question how to identify this rhythmic »Gestalt« the authors discuss the interactive functions of rhythm in conversation: turn taking, theme change, preclosing.