SCHRIFTEN ZUR SPRECHWISSENSCHAFT
UND PHONETIK

Susanne Voigt-Zimmermann/
Frank Lorenz/André Hüttner (Hg.)

Sprechwissenschaftliche Phonetik

Frank & Timme

Susanne Voigt-Zimmermann / Frank Lorenz / André Hüttner (Hg.)
Sprechwissenschaftliche Phonetik

Schriften zur Sprechwissenschaft und Phonetik,
herausgegeben von Ines Bose, Kati Hannken-Illjes,
Ursula Hirschfeld, Baldur Neuber und
Susanne Voigt-Zimmermann
Band 18

Susanne Voigt-Zimmermann / Frank Lorenz / André Hüttner (Hg.)

Sprechwissenschaftliche Phonetik

Frank & Timme
Verlag für wissenschaftliche Literatur

ISBN 978-3-7329-0575-1
ISBN E-Book 978-3-7329-9422-9
ISSN 2364-4494

© Frank & Timme GmbH Verlag für wissenschaftliche Literatur
Berlin 2019. Alle Rechte vorbehalten.

Das Werk einschließlich aller Teile ist urheberrechtlich geschützt.
Jede Verwertung außerhalb der engen Grenzen des Urheberrechts-
gesetzes ist ohne Zustimmung des Verlags unzulässig und strafbar.
Das gilt insbesondere für Vervielfältigungen, Übersetzungen,
Mikroverfilmungen und die Einspeicherung und Verarbeitung in
elektronischen Systemen.

Herstellung durch Frank & Timme GmbH,
Wittelsbacherstraße 27a, 10707 Berlin.
Printed in Germany.
Gedruckt auf säurefreiem, alterungsbeständigem Papier.

www.frank-timme.de

Inhaltsverzeichnis

Vorwort 7

Gottfried Meinhold
Kontextphonetische Zeitdaten gesprochener Sprache 9

Gottfried Meinhold
Phonostilistische Ebenen 101

Rodmonga Potapova, Vsevolod Potapov
Temporal Speech Indicators of the Emotional State “Fear”
(Forensic Aspects) 265

Rodmonga Potapova, Ludmila Veličkova, Eberhard Stock
Das Ökonomieprinzip in der Sprechkommunikation und die
Sprechwissenschaftliche Phonetik 277

Eberhard Stock
Internationale Aspekte der sprechwissenschaftlich-phonetischen
Lehre und Forschung an der Universität Halle 1949–1990 287

Ludmila Veličkova
Die Artikulationsbasis und das Phonemsystem 347

Verzeichnis der Autorinnen und Autoren 357

© Frank & Timme Verlag für wissenschaftliche Literatur

Vorwort

Mit vorliegendem Band werden verschiedene Arbeiten zur sprechwissenschaftlichen Phonetik vorgelegt. Sie stellt neben der Rhetorik, der Sprechkunst und der Klinischen Sprechwissenschaft (in Halle) eine Teildisziplin der Sprechwissenschaft dar, die auf der physiologischen Phonetik und der Orthoepie basiert.

Hervorzuheben ist, dass in diesem Band Autorinnen und Autoren versammelt sind, die die sprechwissenschaftliche Phonetik maßgeblich begründet und entwickelt haben. Sie alle gelten als ausgewiesene Experten mit dennoch jeweils individuellen Forschungsprofilen in den Bereichen Phonetik, Phonologie und Orthoepie. Über viele Jahrzehnte hinweg haben sie zudem in unterschiedlichen Arbeitsgruppen und -kontexten gemeinsam geforscht, gearbeitet und publiziert.

Durch ihre ausgewiesene Lehr- und Forschungstätigkeit haben sie zudem Generationen von Sprechwissenschaftlerinnen und Sprechwissenschaftlern für die theoretischen und praktischen Aspekte und Fragen der sprechwissenschaftlichen Phonetik vorbereitet und zu eigenen Forschungsfragen und -projekten angeregt und betreut.

Die Idee zu diesem Band entstand während der Planungen zur Tagung „Phonetik international“, die anlässlich der Emeritierung von Ursula Hirschfeld am 16. Juni 2018 in Halle stattfand. Da Ulla Hirschfeld nicht nur Schülerin von Eberhard Stock sondern auch langjährige Kollegin und Forschungspartnerin aller hier versammelten Autoren und Autorinnen ist, war es nur folgerichtig, diese zur Veröffentlichung in einem gemeinsamen Band einzuladen. Die Herausgeber schätzen sich glücklich, dass alle Autorinnen und Autoren bereitwillig zugesagt haben. Die Originalität, Diversität und Einzigartigkeit der einzelnen Beiträge spiegeln zudem die besondere Wertschätzung der langjährigen, intensiven und erfolgreichen Zusammenarbeit mit Ursula Hirschfeld wider.

Gottfried Meinhold stellt zuerst einen Beitrag zu kontextphonetischen Zeitdaten gesprochener Sprache zur Diskussion, der gleichsam von seiner diesbezüglichen Forschung seit den sechziger Jahren des vorigen Jahrhunderts auch im Zusammenhang mit anderen Themen, wie Phonostilistik, an der Universität Jena berichtet. In seinem zweiten Beitrag zum Inventar der phonetischen Formen mit

© Frank & Timme Verlag für wissenschaftliche Literatur

phonostilistischem Bezug wird in einer systematischen Übersicht die phonetische Formenvielfalt segmentaler phonostilistischer Varianten vorgestellt.
Rodmonga Potapova und Vsevolod Potapov gehen den Hypothesen nach, dass die zeitliche Strukturierung der Sprache sehr individuell und am stärksten abhängig vom phonetisch-phonologischen Material einer Sprache ist. Ihrer Meinung nach ist sie am wichtigsten für das Erkennen von Veränderungen im emotionalen Zustand eines Sprechers (insbesondere für Furcht und Angst).

Die Bedeutung des Ökonomieprinzips für Sprache und Sprechen, etwa im Fremdsprachenunterricht, diskutieren Rodmonga Potapova, Ludmila Veličkova und Eberhard Stock. Dabei gehen sie auch auf das kognitiv-semiotische Modell der sprechsprachlichen Tätigkeit in Bezug auf das Ökonomieprinzip sowie auf den Einfluss des Ökonomieprinzips im Hinblick auf den Wandel von Laut und Prosodie ein.

Eberhard Stock beschreibt in seinem Beitrag zu internationalen Aspekten der sprechwissenschaftlich-phonetischen Lehre und Forschung von 1949 bis 1990 die historischen Entwicklungen der sprechwissenschaftlichen Phonetik an der Martin-Luther-Universität Halle-Wittenberg.

Den Abschluss des Bandes bildet eine Darstellung des Zusammenhanges von Artikulationsbasis und Phonemsystem von Ludmila Veličkova.

Die vorliegende Publikation ist als Würdigung der sorgfältigen und beharrlichen Arbeit der Autorinnen und Autoren zu verstehen, die mit ihren Leistungen erheblich dazu beigetragen haben, die Phonetik als wichtigen Teilbereich der Sprechwissenschaft zu etablieren.

Halle, den 16. Mai 2019 Die Herausgeber

© Frank & Timme Verlag für wissenschaftliche Literatur

Kontextphonetische Zeitdaten gesprochener Sprache

Gottfried Meinhold, Jena

1 Vorbemerkung

Dieser Bericht stützt sich auf eine Reihe von Forschungen, die seit den sechziger Jahren des vorigen Jahrhunderts sowohl in Qualifikationsschriften – Dissertationen, Habilitationen, Diplomarbeiten – oder auch im Zusammenhang mit anderen Themen wie Phonostilistik an der Universität Jena durchgeführt wurden. Sie werden nur zum Teil im Einzelnen zitiert, doch seien die studentischen Forscherinnen und Forscher – unter ihnen vorwiegend ausländische Studierende – hier ausdrücklich erwähnt: Es handelt sich um die deutschen Studentinnen Andrea Hornbogen und Regina Hurt, um die russischen Studierenden Olga Iljina, Albina Kalinina, W. Loginow und Jewgeni Jemelin, die burjätische Studentin Sojelma Budajewa und die ukrainische Studentin Irina Friedrich. Die Erhebung von Zeitdaten an umfangreichen, großenteils durch Experimente zustande gekommenen Korpora der gesprochenen Sprache erbrachte so viel interessante Resultate, dass bereits in den achtziger Jahren der Plan entstand, die wichtigsten Ergebnisse zu publizieren, was im Fall von Iljina und Kalinina auch geschah. Erst als sich erwies, dass trotz neuerer Untersuchungen – zum Beispiel zur Sprechgeschwindigkeit von Hartmut R. Pfitzinger (2001) oder durch die komplexe Untersuchung zu Zeitparametern von Sabine Kowal (1991) – die älteren Resultate der Jenaer Forschung, selbst diejenigen der sechziger Jahre (Meinhold 1968 und später, Raabe 1979) nicht überholt waren, wurde der Entschluss in den neunziger Jahren erneut gefasst. Doch zog sich seine Verwirklichung noch eine Weile hin. Nunmehr also werden – in sehr verknappter Form – ausgewählte Daten aus einem größeren Bestand der Öffentlichkeit übergeben, in der Hoffnung, sie könnten nützlich sein.

© Frank & Timme Verlag für wissenschaftliche Literatur

2 Gesprochene Sprache im zeitlichen Ablauf. Allgemeines

Gesprochene Sprache – ob spontan produziert oder von einer Textvorlage reproduziert – ereignet sich in der Zeit, als Sequenz von akustischen Schallereignissen, die in strenger Nacheinanderordnung erscheinen, in einer linearen Abfolge nämlich, die keinerlei Vertauschung duldet. Die auditive Sprachwahrnehmung geht somit anders vonstatten als die optische. Bei ihr gibt es, trotz grundsätzlicher Aufrechterhaltung der zeitlichen Sukzession des Vorher und Nachher, einen höheren Grad von teilweiser Simultanität im Wahrnehmungsprozess: Graphische Wortgestalten (bzw. optische Zeichenkörper) werden großenteils, jedenfalls bei gut automatisiertem Leseprozess, als Ganzheiten erfasst, und zwar in kürzeren oder längeren Gruppen. Durch diese Besonderheit ergibt sich für das optische Aufnehmen von Texten gegenüber dem akustischen, hörenden Aufnehmen eine höhere Rezeptionsgeschwindigkeit, also ein zeitlicher Vorteil.

Bei der auditiven Sprachwahrnehmung kommt das Signalmaterial, die lautliche Sequenz, in eben der strengen zeitlichen Anordnung beim Rezipienten an, in der es artikulatorisch vom Produzenten hervorgebracht wird. Der Hörer muss somit das Zeitmaß, das der Sprecher seiner Äußerung aufprägt, als gegeben hinnehmen, er hat nicht wie der Leser die Möglichkeit, seine Rezeptionsgeschwindigkeit selbst zu bestimmen. Er ist gezwungen, sich der Produktionsgeschwindigkeit des Sprechers bzw. der Sprecherin anzupassen. Die Zeitstruktur der gesprochenen Sprache kann für den Erfolg der Kommunikation mittels gesprochener Sprache zu einer entscheidenden Bedingung werden. Doch nicht nur aus diesem Grunde lohnt es sich, damit zusammenhängende Probleme wissenschaftlich, zudem empirisch, zu betrachten.

2.1 Temporale Struktur – prosodische Struktur

Die temporale Struktur ist nämlich als Teilkomponente einer prosodischen Gesamtstruktur der gesprochenen Sprache einzuordnen; hinsichtlich der Variabilität und funktionellen Vielfalt ist sie ein ebenso vielgestaltiges Phänomen wie andere Komponenten der Prosodie, etwa die Melodisierung oder auch die Intensitätsverteilung und die daraus hervorgehenden Erscheinungen wie Akzent und Rhythmus. Sie ist ein prägender Faktor der Sprechweise und wie diese abhängig von individuellen, situativen, textlichen und anderen Gegebenheiten. Für die Charakterisierung von signalphonetischen Vorgängen und deren auditiver Verarbeitung und Wahrnehmung ist die zeitliche Dimension neben der Grundfrequenz, der spektralen Struktur und der Intensitätscharakteristik eine der bestimmenden Größen. Es handelt sich bei der temporalen Struktur um ein supra-

© Frank & Timme Verlag für wissenschaftliche Literatur

segmentales, d. h. ein die Segmentfolge überlagerndes und meistens erst durch die Beziehung mehrerer Segmente untereinander, also in der Lautsequenz deutlich erkennbares Phänomen. Damit soll nicht ignoriert werden, dass die zeitliche Ausdehnung auch für den engeren sog. Minimalsegment-Bereich bedeutsam ist, da ja jedes wahrnehmungsseitig identifizierbare Segment eine ihm zukommende Dauer besitzt, die an der für seine spezifische Lautklasse typischen Dauer („spezifische Lautdauer“) gemessen wird. Ob dem Hörer ein Laut besonders lang oder weniger lang bzw. hinsichtlich seiner Dauer nicht normgerecht erscheint, liegt daran, dass er als Sprachträger die Norm der spezifischen mittleren Lautdauer der betreffenden Lautklasse kennt bzw. die betreffenden Kenngrößen verinnerlicht hat.

Lautdauerphänomene spielen bekanntlich nicht nur suprasegmental eine Rolle, sondern auch segmental-phonologisch. Zudem kennen wir die Einwirkungen suprasegmentaler Erscheinungen auf die Dauereigenschaften der Segmente: Was wir als temporalen Akzent zu bezeichnen gewohnt sind, ist zumeist nichts anderes als eine Wirkung der melodisch-dynamischen Akzentmarkierung in der Zeitstruktur. Es kommt dann zur weiteren Dehnung von langen Akzentvokalen oder bei kurzen Akzentvokalen zur Dehnung der in derselben Silbe nachfolgenden Konsonanten, zumal sonorer.

Die zeitliche Variabilität der Minimalsegmentfolgen – das ist zunächst die Variabilität der Anzahl der Lautereignisse pro Zeiteinheit – beruht auf Dauerveränderungen identifizierbarer Einzelsegmente. Erhöhung der Sprechgeschwindigkeit, d. h. mehr Lautereignisse pro Zeiteinheit, ist stets mit Verringerung der mittleren Lautdauer verbunden, die ihrerseits eine Teilerscheinung des rhythmischen Gesamtgeschehens der gesprochenen Sprache ist. Darf man nun bei solchen Voraussetzungen überhaupt den Versuch wagen, textliche Zeitparameter ausfindig zu machen, zeitliche Kenngrößen also, die geeignet sind, textspezifische Struktureigenschaften des zeitlichen Verlaufs auszudrücken? Ein erster Gedanke, der diese Erwartung nahe legt, resultiert aus der Erkenntnis, dass es bei der Realisierung bestimmter Texte, z. B. literarischer, in ähnlicher Weise wie in der Musik, auf die Realisierung adäquater „Tempi“ ankommt. Mit anderen Worten, es gibt Tempi, die als fehlerhaft zu beurteilen sind, zu langsam oder zu schnell für den betreffenden Text, seiner Spezifik nicht „angemessen“. Dies ist eine Erfahrungstatsache für alle diejenigen, die sich mit Fragen der sprechkünstlerischen Gestaltung befassen. Auch der Tempowechsel zwischen Teilen eines Gesamttextes, ein Wechsel der mittleren Geschwindigkeit zwischen Wortgruppen oder Sätzen kann, ähnlich wie in der Musik, vor allem bei literarischen, künstlerischen Texten ein notwendiges, d. h. im Dienste der adäquaten Sprechweise stehendes Ausdrucksmittel sein. Darüber hinaus ist oftmals eine

© Frank & Timme Verlag für wissenschaftliche Literatur

mit der musikalischen Agogik vergleichbare Veränderlichkeit der Geschwindigkeit auf engem Raume zu beobachten, und zwar zuweilen in Abhängigkeit vom Akzentgeschehen, also von der Struktur des „rhythmischen Körpers“ (im Sinne des Phonetikers O. v. Essen). Diese Variation oder Fluktuation der Grundgeschwindigkeit in gewissen Grenzen dürfte deshalb eine interessante Erscheinung sein, weil die Erfahrung lehrt, dass sich darin in besonderer Weise die Textsensibilität eines ausdrucksstarken Sprechers – auch eines Redners – widerspiegelt.

Von besonderer Markantheit wäre dann als weitere Erscheinung die Häufigkeit und Ausdehnung von Unterbrechungen bzw. Pausierungen in größeren Textblöcken, ein grundsätzlich andersartiges Phänomen der Zeitstruktur. Geht es bei den Geschwindigkeitsphänomenen um die mehr oder weniger große „Signaldichte“ jener Strecken, die nicht durch signallose Zwischenräume unterbrochen sind, so scheint demgegenüber die segmentfreie, schalllose Zeit ein Phänomen ganz anderer Art zu sein, auch wenn die Pause Signalcharakter hat oder haben kann. Vor allem spielen zeitliche Ausdehnung der schalllosen Zeit und die Häufigkeit des Auftretens schallloser Strecken eine Rolle als mögliche charakterisierende Größen für ganze Textblöcke. In welch hohem Maße die Funktion der Pausen als Unterbrechungen der Segmentsequenz für die gesprochene Äußerung wirkt, wird dann nicht nur erkennbar, sondern auch erlebbar, wenn Normverletzungen des zeitlichen Ablaufs – extreme Verringerungen des Sprechtempos, inadäquate Pausierung – als Störungen bzw. Normverletzungen in Erscheinung treten. Der Hörer ist anscheinend auf eine normgerechte und das heißt auch fließende oder flüssige Art der Sprechäußerung eingestellt, er erwartet eine störungsfreie Kontinuität der Darbietung. Er erwartet Unterbrechungen an dafür geeigneten, vor allem durch die syntaktische Struktur bedingten Stellen wie Satz- und Syntagmagrenzen. An anderen Stellen wirken Pausen als Störungen, wenn sie nicht durch besondere expressive Umstände motiviert sind. Auch kann eine syntaktisch nicht motivierte, im Inneren eines Syntagmas auftretende Pause “Bedeutsamkeit“ signalisieren; in diesem Zusammenhang können unerwartete Pausen ein Aufmerken, eine Orientierungsreaktion bewirken. Zudem sind Pausen auch für die rhythmische Struktur, die rhythmische „Gangart“ – den Unterschied von Legato und Staccato beispielsweise – von besonderer Wirkung.

Der Nutzen der Pausenzeit für Rezeption bzw. Verarbeitung kann als geringer bewertet werden. Der Sprechgeschwindigkeit gebührt im Hinblick auf die mögliche Verarbeitungsgeschwindigkeit beim Sprachverstehen der Vorrang.

© Frank & Timme Verlag für wissenschaftliche Literatur

2.2 Reproduzierendes und freies Sprechen

Grundsätzlich unterscheiden sich zeitlich zwei verschiedene Realisierungsweisen von Gesprochenem: einmal die ohne schriftlich fixierte Textformulierung spontan bzw. frei produzierte Äußerung, die im zeitlichen Verlauf andere Strukturformen aufweist als die zweite Form, die man auch „lautes Lesen“ nennt, also die akustische Realisierung eines geschriebenen Textes. In der Mitte zwischen diesen beiden verschiedenen Produktionsformen steht diejenige, bei der ein fertiger Text in der Art und Weise einer (fiktiven) spontanen Äußerung realisiert wird, wie dies der Fall ist, wenn der Schauspieler oder der Vortragskünstler einen (literarischen) Text oder der Redner seinen Redetext auswendig reproduzieren. In diesem Fall wird z. B. im Drama der spontansprachliche Duktus sprechkünstlerisch simuliert, typische Erscheinungen der frei gesprochenen Äußerung ausgenommen, nämlich alle Arten von Versprechern (lapsus linguae) und unwillkürlichen Iterationen sowie die embolischen Laute; auch halten sich Irregularitäten des zeitlichen Verlaufes in gewissen Grenzen. Beispielsweise kommen die Pausenpositionen nur ausnahmsweise an Stellen mit enger syntaktischer Bindung (z. B. nach Konjunktionen oder Artikeln) vor. Das wäre in frei gesprochenen Äußerungen öfter der Fall. Ansonsten treten sie vielmehr an Stellen auf, die von der syntaktischen Struktur dafür disponiert sind, nämlich als „Strukturpausen“, syntaktische Pausen (juncture pauses) an Syntagma- oder Satzgrenzen. Häsitationspausen (Zögerungspausen; hesitation pauses) kommen vorwiegend an Stellen vor, deren syntaktische Kohäsion bzw. semantische Kohärenz Unterbrechungen eigentlich nicht gestatten und somit erscheinen sie als Deviationen, Irregularismen. Im reproduzierten Text werden solche Pausen dementsprechend als Störungen vermieden. Sie sind aber Kennzeichen des freien Sprechens. Allerdings wirken sich Zögerungsphänomene auch bei den Strukturpausen aus: Ihre Dauer nimmt zu, über die Norm hinaus. Somit ist eine Häsitationspause nicht die alleinige Datenquelle für den Verzögerungsgrad, und die Klassifikation weist in dieser Hinsicht eine Unschärfe auf, die bei der Datengewinnung bedacht werden muss.

Die temporalen Strukturen frei gesprochener Äußerungen einerseits und laut gelesener Texte anderseits weisen interessante Verschiedenheiten auf. Sprecherbedingte, subjektive Faktoren kommen bei der frei gesprochenen Äußerung zur Geltung, objektive, z. B. situative oder textliche Faktoren eher im laut vorgelesenen oder memoriert vorgetragenen Text.

Subjektive, individuelle bzw. habituelle Faktoren sind solche, die in der psychischen bzw. psychophysischen Ausstattung, dem individuellen Typus (Habitus) des Sprechers ihre Ursache haben. Bei den temporalen Strukturen gibt zwei-

© Frank & Timme Verlag für wissenschaftliche Literatur

fellos so etwas wie eine persönliche, habituelle Sprechgeschwindigkeit (im Sinne einer Grundgeschwindigkeit bei syntoner psychischer Normallage); es gibt z. B. Schnellsprecher und „Langsamsprecher“. Typologisch hängt die individualtypische Grundgeschwindigkeit zweifellos mit den psychophysischen Eigenschaften des Konstitutionstyps zusammen. Folgt man Fährmann, so neigen sog. „spirituelle“ Temperamente zu einem erhöhten Sprechtempo, die „viskösen“ dagegen zu einer generell langsamen bzw. zu langsamen Sprechgeschwindigkeit. Sog. zyklothyme Temperamente können zwischen den Polen hypomanisch und subdepressiv variieren. Im hypomanischen Fall zeigt sie eine Tendenz zu rascher, gleichmäßiger Geschwindigkeit mit wenig Pausen. Subdepressiven ist demgegenüber ein langsam-bedächtiger Sprechverlauf zueigen, dazu viele, jedoch der syntaktischen Struktur angemessene Strukturpausen. Bei den schizothymen Temperamenten findet sich in der „hyperästhetischen“ Richtung eine rasche bis sehr rasche, hastige, oft überstürzte Sprechweise (Fährmann 1960, 352), in der „anästhetischen“ dagegen ein langsamer bis sehr langsamer bzw. gedehnter Sprechduktus. Zu diesen Verschiedenheiten gibt es bei den zyklothymen eher einen ausgeglichenen Ablauf, bei den schizothymen Temperamenten dagegen Unausgeglichenheit, d. h. eine Neigung zur Veränderlichkeit des temporalen Ablaufs.

Genotypische und phänotypische Momente können bei der individuellen Sprechercharakteristik zusammenkommen: Kognitive Beweglichkeit des Sprechers, hohe Verbalisierungsroutine, d. h. Ausdrucksgewandtheit und reiche, hochaktive Sprachkompetenz dürfen als Momente gelten, die trainierbar sind, also „phänotypisch“, sich auch durch intensiven Gebrauchszwang in der Sozialisation bis zu einem gewissen Grade ausbilden. Eine genetische Voraussetzung – innerer Dynamik und sprachlicher Begabung mit kognitiver Beweglichkeit – begünstigt solche Entwicklungsprozesse und bestimmt auch Grenzen der Trainierbarkeit.

Umgekehrt kann zudem von der zeitlichen Struktur des Sprechablaufes auf die innere Mobilität des Sprechers geschlossen werden. Auch bei der intuitiven Einschätzung des persönlichen Temperamentes nimmt bekanntlich die Beurteilung der Sprechweise einen vorderen Rangplatz ein (zusammen mit dem Ausdrucksgeschehen von Mimik und Gestik und dem Gesamtverhalten).

Typbedingt bzw. durch das Temperament bestimmt sind somit auch die Ausprägungsgrade der Emotionalisierungsrichtungen (mehr freudig, lebhaft, kommunikativ, extrovertiert oder auf der anderen Seite ernster, ruhig, introvertiert), damit zusammen die Erregungsgrade bzw. die Erregbarkeit (Neurotizismus). Jedoch bilden gerade die aktuellen psychischen Zustände die Klammer zu den objektiven Bedingungen, unter denen sich die Sprechäußerung vollzieht, denn je

© Frank & Timme Verlag für wissenschaftliche Literatur

nach der Situation werden bei dem Sprecher z. B. größere oder geringere Spannungsniveaus (mit sympathikotoner oder parasympathikotoner Tendenz) evoziert.

Die Wirkungen solcher gegebenenfalls durch die Situation stimulierter Spannungspotentiale, z. B. bei sympathikotonem Stress, zeigen sich auch besonders eindrucksvoll an sprachpathologischen Fällen von Redeablaufstörungen: Der zum Poltern (Tumultus sermonis, Tachylalie), einer pathologischen Tendenz zu erhöhtem, überstürztem Sprechtempo neigende Sprecher wird durch situativen, sympathikoton wirkenden Stress stets eine Verminderung seiner Symptomatik, also eine Annäherung an normales Sprechtempo und an einen eher ausgeglichenen temporalen Verlauf zeigen; im Gegensatz dazu wird der Stotterer (Balbuties), dessen Krankheitsbild die Neigung zu Stockungen, zu klonischen oder tonischen Hemmungen kennzeichnet, bei erhöhtem Situationsstress stets eine Verstärkung der störenden Symptomatik erleben. Beides gilt sowohl für das freie Produzieren als auch für das reproduzierende laute Lesen. Jedoch tritt im Allgemeinen der habituelle Sprechduktus beim reproduzierenden Sprechen nicht allzu deutlich auf; von den habituellen Besonderheiten bleiben beim reproduzierenden Sprechen allenfalls gewisse Restbestände.

Der Kompliziertheitsgrad der Sache, über die gesprochen wird, kann besonders störende Erscheinungen, nämlich Verzögerungen (hesitations), Stockungen, embolische Laute, Iterationen hervorbringen. Es ist zu erwarten, dass solche Erscheinungen umso häufiger auftreten, je höher der kognitive Aufwand und Anspruch ist, den Gegenstand und Thema erheben. Dass gerade die temporalen Abläufe davon betroffen sind, liegt an dem strengen Zeitregime des Sprechens und der zeitlichen Abstimmung mehrerer innerer, vorbereitender Teilprozesse – gedankliche Planung, innersprachliches Konzept, Ausspruchsplanung – deren inkrementelle, zeitlich versetzte Abläufe sich gegenseitig stören, wenn sie zeitlich nicht gut koordiniert sind. Auch spielen Umfang und Aufwand, die Schnelligkeit der assoziativen Verbindungen der simultan zum äußeren Sprechprozess ablaufenden Reflexionsprozesse eine erhebliche Rolle bei der gegebenenfalls beeinträchtigten zeitlichen Koordination: Das Reflexionstempo der gedanklichen Vorbereitung wird den Gesamtprozess – ebenso wie die innersprachliche Vorbereitung – beeinflussen, nämlich verzögern oder beschleunigen, sogar stoppen, unterbrechen oder stimulieren. Dementsprechend stehen die mehr oder weniger große Kompliziertheit des Gegenstandes und die sprachliche, z. B. syntaktische Kompliziertheit im Zusammenhang. Komplizierte inhaltliche Relationen können sich in tiefer syntaktischer Staffelung bzw. größerer syntaktischer Strukturtiefe äußern, deren prosodische Formierung einen größeren Aufwand an hervorhebenden und begrenzenden Mitteln erfordert.

© Frank & Timme Verlag für wissenschaftliche Literatur

In der vorliegenden Betrachtung zeitlicher Kenngrößen in Sprechäußerungen wird jedoch dem reproduzierenden Sprechen als der Sprechform, die zeitliche Normgegebenheiten am ehesten erkennen lässt, das Hauptaugenmerk zu widmen sein. Ein größeres Korpus von Sprechaufnahmen sowohl des Hörfunks als auch (vor)gelesener bzw. gestalteter literarischer Texte (in einigen Fällen auch memorierter Texte des Sprechtheaters) soll eine Vergleichsbasis liefern, die Aussagen sowohl über Normgrenzen als auch über eventuelle Deviationen vom störungsfreien, „flüssigen" Ablauf beim reproduzierenden Sprechen liefern, wie sie beim freien Sprechen auftreten können.

2.3 Entropie und Redundanz

Unter den Textcharakteristika sind auch solche zu berücksichtigen, die weder aus syntaktischen noch aus semantischen Gegebenheiten allein zu ermitteln sind, sondern im Satzkontext zusammenwirken: Dazu gehört die wechselnde Informationsdichte im Text, die mittels der Wahrscheinlichkeitsstruktur, d. h. die in der Sequenz beständig ab- oder zunehmende Unbestimmtheit der Fortsetzungserwartung zu erfassen ist. Bekanntlich wird der Informationsgehalt – im Sinne der Kontextentropie – als eine Funktion der Wahrscheinlichkeit des Auftretens der jeweils nachfolgenden Textelemente bestimmt. Es ist nun evident, dass Strecken mit größerer Fortsetzungswahrscheinlichkeit (d. h. mit geringerer Information bzw. Entropie) und dem damit verbundenen geringeren Verarbeitungsaufwand ein höheres Sprechtempo „vertragen" als Strecken mit niedriger Fortsetzungswahrscheinlichkeit (mit höherer Information bzw. geringerer Redundanz), die zudem eine höhere Verarbeitungsleistung erfordern. Strecken größerer Redundanz bzw. geringerer Information treten nun in jedem Text, sogar im (längeren) einzelnen Satz, im Wechsel mit solchen geringerer Redundanz und größerer Information auf. Tendenziell werden informationsreiche Stellen oder Passagen durch Besonderheiten ihrer prosodischen Ausstattung (vor allem durch die Hervorhebung ihrer akzenttragenden Silben im Satz) von informationsärmeren Strecken abgesetzt. Informationsgewicht und semantisches Gewicht sind übrigens einander direkt proportional: Wachsende Redundanz, was zugleich wachsende Entbehrlichkeit bedeutet, entspricht auch einer Abnahme der semantischen Dichte.

Die temporale Struktur gesprochener Texte hängt nun mit Information bzw. Redundanz in der Weise zusammen, dass redundantere Strecke eine größere Geschwindigkeit (korrespondierend mit der Abwesenheit anderer kulminativ wirkender Prosodika) gestattet. Größere Redundanz bildet also eine Disposition für

© Frank & Timme Verlag für wissenschaftliche Literatur

erhöhte Sprechgeschwindigkeit und gestattet übrigens zugleich eine gewisse Vernachlässigung der segmentalen Form, Lautverlust(e) oder partielle Assimilationen, ohne dass Verständlichkeitseinbußen eintreten. Der Sprecher muss das Informationsgewicht einer Textpassage einschätzen können, oder anders: Er signalisiert durch die temporale Gestaltung im Zusammenhang mit den anderen prosodischen Faktoren, welches Bedeutungsgewicht dem Textteil (der Wortgruppe oder auch einem einzelnen Wort) beizumessen ist.

Für das freie Sprechen gilt in diesem Zusammenhang – wie noch an anderer Stelle darzulegen ist –, dass informationsreiche Strecken einen größeren Planungsaufwand, größere kognitive Anstrengung und eben auch mehr Zeitaufwand erfordern und aus diesem Grund häufiger zeitlich retardiert oder durch Stockungen gestört oder störanfällig erscheinen. Es ist die „allmähliche Verfertigung der Gedanken beim Reden“ (Kleist), die generell zu einer Entschleunigung des Sprechprozesses beiträgt.

2.4 Grenzwerte?

Eine andere Frage im Zusammenhang mit der Zeitproblematik beim Sprechen ist die Frage der oberen und unteren Grenzwerte, damit auch der Norm im Sinne von Mittelwerten, sowohl bei der Sprechgeschwindigkeit also auch bei den Pausen, vor allem der Pausendauer. Sowohl produktiv als auch rezeptiv sind der Anzahl der in der Zeiteinheit zu übermittelnden Signale relativ enge Grenzen gesetzt. Die natürliche Trägheit der produzierenden Organe scheint hierbei das dominierende Problem zu sein, weniger das perzeptive Auflösungsvermögen, dem offenbar nicht einmal durch das sog. menschliche Moment, d. h. jene für den Menschen spezifische Minimaldistanz, in der zwei Impulse noch als getrennte Schallereignisse wahrgenommen werden, d. i. 1/16 bis 1/18 s, eine Grenze gesetzt ist: Da die Lautsequenz des Wortkörpers in der Gestalteinheit identifiziert werden kann, ohne dass alle Segmente mit gleicher Präzision und sicher perzipierbar vorhanden sind, ist es möglich, dass diese Grenze erheblich überboten werden kann, wenn man bedenkt, dass die Darbietungszeit zur Wahrnehmung eines lautlichen Segmentes nur sehr gering zu sein braucht, z. B. in bereits als extrem geltenden Fällen 0,01 s, was nicht etwa bedeutet, es könnten perzeptiv 100 Segmente pro Sekunde verarbeitet werden. Das Problem der minimalen Darbietungszeit hat überhaupt nur Gewicht im Zusammenhang mit der Erkenntnis, dass durch die “Gestaltbildung“ hinsichtlich des Informationsbetrages ein großer Teil der segmentalen Information entbehrlich ist, so dass in Abwandlung einer gestaltpsychologischen Gesetzmäßigkeit mit Recht gesagt wer-

den kann, das Ganze sei (hinsichtlich des Informationsbetrages) **weniger** als die Summe seiner Teile, allerdings nur dadurch, dass es **qualitativ mehr** als die Summe seiner Teile ist. Allein durch diese elementare Gegebenheit der Sprachwahrnehmung sind assimilatorischer Lautverlust und die kommunikative Möglichkeit phonetischer Ellipsenbildungen mit radikaler Verstümmelung der segmentalen Struktur oder segmentaler Minimaldauer im Bereich von 0,01 Sekunden zu erklären. Ähnliche Erfahrungen vermitteln Untersuchungen mit der Beschneidung der akustischen Information in technischen Kanälen. Auf der Grundlage der Stereotyp-Bildung reicht ein kleiner Bruchteil der Information für die Zeichenerkennung und -identifizierung aus. Bekanntlich spielt hierbei der Determinationszusammenhang des einbettenden Kontextes eine Rolle: je stärker die Kontextdetermination, desto geringer die Relevanz des einzelnen Segmentes, desto höher seine Entbehrlichkeit (Redundanz) und desto größer seine Disposition zur Vernachlässigung bzw. Eliminierung. An diesen Sachverhalt sei hier nur andeutungsweise erinnert.

2.5 Exkurs: die vorgetäuschte Pause (Nullpause)

Bekanntlich schwankt die Perzeption der Pause hinsichtlich der Dauer: Schalllose Strecken wirken innerhalb schneller Umgebung oder am Ende einer beschleunigten Passage länger als im langsam gesprochenen Text. Gleichfalls wirken in einer Textumgebung mit besonders starker expressiver Ladung oder innerhalb von Syntagmen – wo sie normalerweise nicht auftreten sollten – Pausen länger als im sachlich-ruhig gesprochenen Text bzw. an Stellen, wo syntaktische Grenzen vorliegen. Auch im informationsreicheren, also weniger redundanten Text wirken Pausen länger. Es entsteht also, wenn man Pausen nach ihrer Dauer in die Maßkategorien Kurz- oder Langpausen einordnet, eine nicht unbeträchtliche Grauzone dadurch, dass es eine Diskrepanz zwischen gemessener und wahrgenommener Dauer gibt. Umgekehrt kann es sein, dass sehr kurze Pausen nicht mehr als solche wahrgenommen werden; unter welchen prosodischen und syntaktischen Bedingungen dieser Fall eintritt, wurde bisher noch nicht systematisch eruiert.

Das Vorhandensein von Pausen kann vorgetäuscht werden, wenn bei Zäsuren – seien sie syntaktisch oder prosodisch (rhythmisch-metrisch bzw. melodisch) oder auch expressiv bzw. durch das Zusammenwirken mehrerer dieser Faktoren bedingt – eine Kennzeichnung des Einschnittes durch besondere Mittel (z. B. Dehnung eines Segmentes) oder prosodischen Kontrast (dynamische Sprünge, Tonbrüche, erheblicher Geschwindigkeitswechsel, Timbre-Kontrast) erfolgt.

© Frank & Timme Verlag für wissenschaftliche Literatur

Hier einige Beispiele: Zu Beginn von Goethes „Hochzeitlied“ (Zeile vier und fünf) können die syntaktischen Einschnitte, jeweils durch Komma gekennzeichnet, mit oder ohne Kurzpausen gestaltet werden: ⟨Da, wo ihr den Enkel des seligen Herrn, / den heute vermählten, beschmauset. /⟩. Fehlende Pausierung kann an den drei syntaktischen Einschnitten durch leichte Dehnung des Vokals von ⟨da⟩, durch Dehnung der /n/-Auslaute bei ⟨Herrn⟩ und ⟨vermählten⟩ kompensiert werden, sodass dennoch eine Pause wahrgenommen wird. Alle einer Zäsur vorangehenden „Dauerlaute“ können diesen Effekt erzeugen, bei Plosiven in dieser Position tritt eine Längung der Verschlussphase (Plosion) ein. Auch wenn nach der Zäsur ein Plosiv auftritt, wird die Dehnung eines vorangehenden Dauerlautes unterbleiben. Es ist dann dessen Verschlussphase (Plosion), deren Dehnung geeignet ist, die Pausierung vorzutäuschen, so z. B. bei ⟨Geht es gut aus, dann...⟩

In der ersten Strophe von Goethes Ballade „Erlkönig“ liegen jeweils am Zeilenende des ersten und zweiten Verses der ersten Strophe Stellen vor, bei denen die gelängte Plosion des Plosivs die Pause ersetzen kann:

> „Wer reitet so spät durch Nacht und Wind?
> Es ist der Vater mit seinem Kind;
> Er hat den Knaben wohl in dem Arm,
> Er faßt ihn sicher, er hält ihn warm.“

Gestaltet der Sprecher diesen epischen Einstieg rasch, mit einer Vorankündigung der späteren Beschleunigung des Geschehens, so werden am Ende der Zeilen eins bis drei keine messbaren Pausen mehr auftreten, stattdessen eben längere Plosionen oder beim Übergang von Zeile drei zu Zeile vier eine Dehnung des /m/ am Zeilenende, im Inneren der letzten Zeile vielleicht auch zur Dehnung der glottalen Plosion bei ⟨er⟩.

Expressive Intensivierung im Verein mit Beschleunigung könnte in der wörtlichen Rede: ⟨Mein Vater, mein Vater, jetzt faßt er mich an!⟩ für vorgetäuschte Pausen sorgen, weil sehr deutliche syntaktische Einschnitte – jeder Ausruf eine syntagmatische Einheit – vorliegen. Aber auch in der panischen Eile der letzten Strophe geschieht Vergleichbares:

> „Dem Vater grausets, er reitet geschwind,
> Er hält in Armen das ächzende Kind,
> Erreicht den Hof mit Müh und Not;
> [...]“

Immerhin liegen hier zwei Satzübergänge vor, die aber Goethe schon nur mit Kommata statt mit Punkten markiert. Wiederum vermeint man, kurze Pausie-

rungen wahrzunehmen, obwohl die Pause durch punktuelle segmentale Ritardandi oder Retenuti ersetzt wird.

Generell muss man sich – bei der Gewinnung von Messwerten – hüten, eine gelängte Okklusion, auch beim festen Stimmeinsatz (glottal stop), als Pausenzeit zu veranschlagen und zu messen. Es handelt sich lediglich um eine Lautdauer-Erscheinung.

In Brechts Ballade „Legende von der Entstehung des Buches Taoteking..." heißt es in Strophe sechs: ⟨Und er schrie: „He du! Halt an! / Was ist das mit diesem Wasser [...]⟩". Weder bei der Zäsur nach ⟨schrie⟩ noch nach ⟨du⟩ oder ⟨an⟩ müssen messbare Pausen auftreten, obwohl es sich um syntaktisch potenzielle Pausenstellen handelt. Dennoch werden hier Pausen wahrgenommen. Ähnlich bei der expressiv geladenen Stelle der „Bürgschaft", wo es heißt: ⟨"Was wollt ihr?" ruft er, für Schrecken bleich, „Ich habe nichts als mein Leben, [...]"⟩ und von den drei potenziellen Pausenstellen (vor ⟨ruft⟩, vor ⟨für⟩, vor ⟨ich⟩) kaum eine realisiert wird.

Stoßen zwei homorgane Plosive im Wortübergang aufeinander, wird die ihnen gemeinsame Plosion, die ohnehin bereits durch den Explosionsverlust des ersten der beiden Plosiv-Segmente eine Verdoppelung erfährt, vielleicht noch einmal zusätzlich verlängert (⟨Er kommt – doch ziemlich spät.⟩). Die akustischen Eigenschaften zur Wahrnehmung einer Pause könnten hier am ehesten gegeben sein.

Anders die vorgetäuschten Pausen, die im Zusammenhang mit engen Enjambements, vorwiegend am Verszeilenende entstehen. Es sind Zäsuren, die der syntaktischen Gliederung zuwiderlaufen und allein durch rhythmisch-metrische Gegebenheiten verursacht werden, eben durch die metrische Einheit des Verses: Die erste Zeile von Rilkes „Liebeslied" fordert nach der Konjunktion ⟨daß⟩ am Ende der Verszeile eine irgendwie deutliche Zäsur: ⟨Wie soll ich meine Seele halten, daß / sie nicht an deine rührt. [...]⟩ Vergleichbar auch der Übergang von Zeile vier zu fünf: ⟨Ach gerne möcht ich sie bei irgendwas / Verlorenem im Dunkel unterbringen /⟩, wo das gedehnte /s/ in ⟨irgendwas⟩ vor dem anlautenden /f/ in ⟨Verlorenem⟩ deutlich wahrnehmbar wird. Anders in dem Versübergang von Zeile zwei zu drei: ⟨[...] wie soll ich sie / hinheben über dich zu andern Dingen?⟩, wo trotz der engen syntaktischen Kohäsion zwischen ⟨sie⟩ und ⟨hinheben⟩ die metrische Struktur eine Zäsur fordert, die ebenfalls durch eine Pause von minimaler Dauer oder aber durch ein „Retenuto" realisiert werden könnte.

Bei dem „halbengen" Enjambement in Goethes Ballade „Hochzeitlied": ⟨Nun dappelt's und rappelt's und klappert's im Saal / von Bänken und Stühlen und

© Frank & Timme Verlag für wissenschaftliche Literatur

Tischen⟩ wird der Versübergang bzw. das Versende bei rascher Geschwindigkeit durch eine vorgetäuschte Zäsur, die der Dehnung des /l/ in ⟨Saal⟩ zuzuschreiben ist, markiert. An diesem Sonderfall der Pausierung lässt sich klar erkennen, mit welcher Relativität im Hinblick auf Pausenexistenz und Pausendauer zu rechnen ist.

Dass sich demgegenüber besonders in expressiven Kontexten eine obligatorische Deutlichkeit der Pause mit entsprechender Dauer erweist, lässt sich zum Beispiel am Ende der Ballade „Erlkönig“ erkennen, wo die Pause an einer Stelle mit hoher Kohäsion (⟨In seinen Armen das Kind war tot.⟩) als expressive Stockung vor ⟨war⟩ auftritt. (Ähnlich in der „Bürgschaft“, wo in der letzten Strophe, in der direkten Rede des Tyrannen: ⟨Ihr habt das Herz mir bezwungen.⟩, eine Pause wiederum an einer Stelle mit hoher Kohäsion, vor ⟨bezwungen⟩ sich als expressiv besonders wirksam erweist.

3 Psychophysische Bedingungen für den Zeitverlauf beim Sprechen

Obwohl die Zeitverhältnisse des freien Sprechens in diesen Untersuchungen eher beiläufig Erwähnung finden, seien doch gewisse Eigenschaften des freien Sprechens in Betracht gezogen. Denn manche Zeitphänomene des freien Sprechens – zum Beispiel Häsitationspausen (Zögerungspausen) – finden Eingang in die Ausdruckspalette der Gestaltung literarischer Texte, nicht nur des Sprechtheaters, ob versgebunden oder nicht, sondern auch bei der Gestaltung literarischer Texte schlechthin.

Der Sprechvorgang ist in ähnlicher Weise an gewisse zeitliche Grenzen gebunden, wie der Verstehensprozess mit seinen verschiedenen Teilphasen. Offensichtlich ist jedoch der zeitliche Spielraum, in dem sich innersprachliche und artikulatorische Prozesse in ihrem Zusammenwirken ereignen, an physiologisch vorgegebene temporale Bedingungen des Sprechens gebunden, besonders soweit es Grenzwerte (Maximalgeschwindigkeit beim Sprechen, Minimalisierung der Pausenzeit) betrifft.

Wie für jeden körperlichen Bewegungsablauf gibt es auch für die Artikulationswerkzeuge des Menschen, bei aller individuellen Streuung motorischer Fertigkeiten, hinsichtlich der Geschwindigkeit ein Optimum (das über längere Zeit ohne nennenswerte psychophysische Ermüdung eingehalten werden kann) und ein Maximum (das nur für kürzere Zeitdauer und sozusagen selten oder ausnahmsweise realisiert wird).

Obwohl in dieser Untersuchung das reproduzierende Sprechen gewissermaßen als Vergleichsgrundlage vorrangig betrachtet werden soll, sei auch ein Blick auf

© Frank & Timme Verlag für wissenschaftliche Literatur

jene zeitlich strukturierten Teilprozesse des Sprechens geworfen, die den Zeitablauf bei freien Sprechäußerungen beeinflussen.

Für den zeitlichen Verlauf frei gesprochener Äußerungen sind allem Anschein nach die „internen“, psychischen Zeitverhältnisse und die Koordination verschiedener Teiltätigkeiten der Planung der Sprachäußerung von Belang. Diese Prozesse sind in ihrem Zusammenwirken zwar noch keineswegs völlig geklärt, jedoch ist ihre inkrementelle Struktur (Levelt 1989) mit der zeitlichen Versetzung aufeinander folgender interner Arbeitsgänge als Ursache für Unregelmäßigkeiten bedeutsam. Auf die stets vorauseilende propositionale, gedankliche Planungsphase folgt die innersprachliche Planung und erst danach, sozusagen nachhinkend, der Prozess der artikulatorischen Realisierung. Da diese drei Teilphasen kurze Zeit nach dem Einsetzen des Sprechprozesses in zeitlicher Staffelung simultan verlaufen, ergibt sich eine Ursache für interessante Störungen. Bevor der Befehl an die Bewegungsprogramme der motorischen Sprachregion gegeben wird, die den Artikulationsprozess in seinem präzisen Zusammenwirken von Atemapparat, Kehlkopf (Stimmerzeugung) und Sprechwerkzeugen des Ansatzraumes (oberhalb der Glottis, also die Bewegungen von Zunge, Velum, Kiefer, Lippen u. a. betreffend) in Gang setzen, müssen diese drei Vorbereitungsphasen vorangegangen sein. Während also die artikulatorische Realisierung des ersten Teiles stattfindet, wird innersprachlich die nachfolgende Artikulationssequenz vorbereitet, gedanklich (propositional) aber bereits die übernächste Artikulationssequenz. In der ersten Phase wird also in einem Zusammenwirken volitionaler, emotionaler und kognitiver Vorgänge (mit von Fall zu Fall verschiedener Dominanz) das Kommunikationsmotiv, die Kommunikationsabsicht hervorgebracht. Hier wird der inhaltliche Rahmen der Äußerung konzeptuell abgesteckt, in Form einer (Teil-)Zielvorgabe, die den propositionalen Kern des zu äußernden Inhaltes in sich trägt. Innerhalb dieses „konzeptuellen Rahmens“ kommt es dann sehr rasch zu einer Detaillierung: zur Anreicherung von Information sowie zu ihrer Organisation, in hierarchischer Ordnung, sodann die „Linearisierung“ auf der Zeitachse in chronologischer Eindimensionalität, die aus der Mehrdimensionalität der Wissensstrukturen gewonnen wird, in diesem Stadium noch mit einem Minimum an innersprachlichem Material. Gegebenenfalls können Kernbegriffe als semantische „Kondensationskerne“ bereits in dieser Phase der Konzeptualisierung deutlicher ins Bewusstsein treten. Die sprachliche Explizitheit nimmt zu, wenn das zweite Planungsstadium erreicht ist, nämlich das der grammatischen und phonologischen Enkodierung, der (inneren) Formulierung. Da erfolgt – wie vielfach beschrieben – in Form des inneren Sprechens eine relativ detaillierte Vorwegnahme der endgültigen Form, jedoch nur in Ausnahmefällen mit jenem Grad von Ausführlichkeit wie bei kompletten

© Frank & Timme Verlag für wissenschaftliche Literatur

Formulierungsproben, die im Schreibprozess als Vorwegnahme oder interne Erprobung der dann aus dem Gedächtnis schreibmotorisch zu realisierenden Form auftreten. Zudem scheint der Grad der innersprachlichen Ausführlichkeit individuell höchst verschieden zu sein, auch die Grade der Bewusstheit, mit der die Zeichenkörperabbilder vor der motorischen Realisierung intern deutlich werden. Allenfalls treten die „Wortklangbilder“ (Kainz 1956) unmittelbar vor der artikulatorischen Realisierung ins Bewusstsein, nämlich in dem Augenblick, wo die motorischen Programme des Sprechbewegungsablaufes abgerufen werden. Die zeitliche Abstimmung der Teilphasen der Sprachproduktion geschieht unter einem Zeitzwang, der durch die Erwartung des Hörers ausgeübt wird: Es entspricht der Norm, ein bestimmtes Geschwindigkeitsminimum nicht zu **unterschreiten**. Der Sprecher äußert sich insofern unter Vollzugszwang. In allen internen Phasen der Vorbereitung, die ihrerseits als Teiltätigkeiten mit bestimmten, nicht ganz leicht zu beschreibenden Zwischenresultaten aufgefasst werden können, gibt es jedoch die Möglichkeit von Verzögerungen, und zwar aus verschiedenen Gründen:

(1) Es ist möglich, dass die in der konzeptuellen Vorbereitung anzureichernde Information, die aus dem Langzeitgedächtnis (Permanentspeicher) möglichst reichlich, sozusagen mit Überschuss, ins operative Gedächtnis überführt werden muss, zu spärlich fließt, also nicht mit der geringen Zugriffszeit verfügbar ist, wie für eine kontinuierliche Zulieferung der Information für den weiter verarbeitenden Prozess der sprachlichen Vorbereitung erforderlich. Wenn die Informationslieferung in der kognitiven Vorbereitung nach bereits begonnenem Sprechprozess ausbleibt, kommt es auch zu einem Versiegen der Sprachschallproduktion, der Sprecher stockt, er hat „den Faden verloren“ oder muss sich bemühen, Wissenslücken durch Mobilisierung von Information zu schließen.

(2) Es ist weiterhin möglich, dass eine ausreichende, ja übergroße Informationsmenge in das operative Gedächtnis eingespeichert wurde, dass es jedoch Schwierigkeiten gibt, das Material zu organisieren, Prioritäten der Relevanz schnell genug zu finden sowie eben die zeitliche Abfolge (Linearisierung) herzustellen. Solche Schwierigkeiten lassen sich äußerlich kaum von denen der Informationsaktivierung unterscheiden. Die Verzögerungsphänomene gleichen sich: Sie betreffen längere Sequenzen, sind weniger, wie bei den nunmehr zu erwähnenden sprachlichen Planungsschwierigkeiten, an einzelnen Punkten auszumachen.

(3) Die eigentlichen Formulierungsschwierigkeiten – also Schwierigkeiten vor allem der grammatischen Enkodierung (punktuell auch der lexikalischen Ausstattung) – sind bei Suchvorgängen als Verzögerungen an bestimmten

© Frank & Timme Verlag für wissenschaftliche Literatur

syntagmatischen Schaltstellen (z. B. nach der Konjunktion) oder vor wichtigen autosemantischen Gliedern (Adjektiven, Substantiven), also Textelementen von hoher Information mit einiger Sicherheit zu erkennen. Wortfindungsschwierigkeiten und Fehlkalkulationen der syntaktischen Ausspruchsplanung können aber gehäuft auftreten, wenn bereits die vorangehenden kognitiven Prozesse den Sprecher überfordern; es wird also durch die unter (1) und (2) genannten Störungsquellen u. U. der Gesamtprozess labilisiert. Unter einer beeinträchtigten kognitiven Vorbereitung, die die gesamte innere Aufmerksamkeit des Sprechers auf sich zieht, leidet dann das sprachliche Formniveau insgesamt, weil ihm weniger Sorgfalt (Aufmerksamkeit) gewidmet werden kann.

Die Schwierigkeiten können dadurch vervielfacht werden, weil im Operativgedächtnis bzw. im Bewusstsein (vgl. Frank: auf der „Bühne des Bewusstseins“) gleichzeitig sowohl eine längerfristige als auch eine kurzfristige kognitive Planung stattfindet und weil, während bereits für bestimmte inhaltliche Komplexe die sprachliche Detailplanung begonnen hat, simultan die Fortsetzung der inhaltlichen Planung stattfindet, also die kognitive Planung für jene Textstrecke, die der in der sprachlichen Planung befindlichen unmittelbar nachfolgt (im Sinne der erwähnten inkrementellen Struktur). Weiterhin muss gleichzeitig – vor allem bei längeren Sprachäußerungen mit komplizierterer Binnenstruktur – die Zielvorstellung in Form von mehr oder weniger sprachlich manifestierter Information gleichsam als „Dauerbeleuchtung“ des Gesamtprozesses langfristig anwesend sein. Sie verhindert, dass bei der kognitiven Planung zu viel an unwesentlicher Information mitverarbeitet wird, schaltet also, z. B. in Form von thematischer Information, als Selektionsschwelle in der Art eines Filters die nicht zum Thema gehörige oder mit ihm in weniger direktem Zusammenhang stehende Information aus.

Grundsätzlich haben wir ja in der ersten konzeptionellen propositionalen Planungsphase davon auszugehen, dass ein **Über**angebot an Information evoziert und operativ wenigstens flüchtig gespeichert wird, aus dessen Gesamtmenge durch schnellstmögliche Selektion verwendbare Information ausgewählt wird. Übrigens kann auch dieser Selektionsvorgang z. B. durch Unentschlossenheit oder Konzentrationsmangel des Sprechers eine Ursache für Verzögerungsphänomene sein.

Die Simultaneität der chronologisch versetzten Teilprozesse in der Vorbereitung des Sprechprozesses kann sich, da sie nicht total unabhängig voneinander ablaufen, sondern offenbar Informationen von dem einen in den anderen Prozess überspringen können, ebenfalls als Quelle von Störungen auswirken. Dies ge-

© Frank & Timme Verlag für wissenschaftliche Literatur

schieht entweder in Form von Sprechfehlern (Versprecher; lapsus linguae) oder wiederum in Form von Verzögerungserscheinungen, die oftmals als „abgefangene“ Fehlleistungen zu betrachten sind, ebenso wie manche unauffälligere Formulierungskorrekturen.

Ein weiterer Vorgang kommt zu den bereits genannten hinzu: Bekanntlich werden die der Vorbereitung des Sprechprozesses dienenden Informationen im operativen Gedächtnis mit der eben realisierten Information der Sprachäußerung addiert, da sich der Sprecher über den auditiv-akustischen Rückmeldekreis während des Sprechens kontrolliert.

Die Wahrnehmung der eigenen Sprachäußerung führt nicht nur im Falle verzögerter Rückkopplung („Lee-Effekt“) zu charakteristischen Störungen im zeitlichen Ablauf der Sprachäußerung, sondern auch u. U. im Normalfall, zum Beispiel durch die mögliche Tendenz verschiedener Textelemente zu perseverieren. Zumindest wäre es denkbar, dass die über den auditiven Kontrollkreis erfolgende Rückmeldung der Information perseverative Tendenzen eines Sprechers verstärkt oder dass es zu störenden Interferenzen zwischen der rückgemeldeten und der zum Ausspruch vorbereiteten Information kommt.

Im freien, spontanen Sprechen ist bei weniger geübten Sprechern das Auftreten von Störungen durch zu lange oder syntaktisch ungünstig platzierte Pausen eine häufige, doch alles in allem wohl übliche Erscheinung. Es wird bei Koordinationsmängeln mehr Zeit verbraucht, über die normale Planungszeit hinaus, die dem Sprecher mit der Sprechzeit (t_o) und der Pausenzeit (t_p) zur Verfügung steht. Der Effekt ist in solchem Fall also die Zunahme der Gesamtpausenzeit, langsameres Sprechen, d. h. Sprechen auf Zeitgewinn, um den vermehrten Planungsaufwand, dessen zeitlicher Anspruch je nach dem Schwierigkeitsgrad oder der Zahl der abzuwickelnden Teilprozesse verschieden hoch ist, Genüge zu tun.

Eine besondere Störungsvariante liegt vor, wenn die Schnelligkeit und Sicherheit sowohl der kognitiv-inhaltlichen (konzeptuellen) als auch der sprachlichen Planung durch bestimmte emotionale Prozesse beeinträchtigt wird, die mit Angst oder Misserfolgserlebnis und allgemein erhöhtem psychischem Spannungsgrad in Verbindung stehen. Der Zusammenhang zwischen Angst bzw. Angstverarbeitung und Veränderungen des zeitlichen Verlaufs ist nachgewiesen (Helfrich und Dahme 1974).

Stressfaktoren können dabei übrigens durchaus verschiedene, nämlich leistungsmindernde, andererseits aber auch leistungssteigernde Wirkung haben, depressive Zustände sind bekanntlich stets mit Verlangsamung und Kohärenzstörungen verbunden.

© Frank & Timme Verlag für wissenschaftliche Literatur

Der Vollständigkeit wegen sei der seltenere Fall erwähnt, dass nämlich der äußere Artikulationsprozess unter inneren Informationsdruck gerät, auf Grund besonders schnell verlaufender kognitiver und planerischer Prozesse, so dass in dem Bestreben, sich der hohen Geschwindigkeit der vorbereitenden Prozesse anzupassen, beschleunigte, überstürzte, pausenarme Sprechweise zustande kommt, die u. U. gewisse Anklänge an Abläufe bei pathologisch beschleunigter Sprechweise (Tachylalie, Tumultus sermonis oder Poltern) aufweist. Propulsive Sprechweise kann in solchen Fällen auch mit staccatohafter Zerrissenheit einhergehen, wenn durch häufige kurze Pausen viele kurze Pausenabschnitte entstehen. Artikulatorische Unsauberkeiten, d. h. assimilatorische Laut- und Silbenverluste, können als Folge einer motorischen Beschleunigung auftreten, die die Sprechwerkzeuge überfordert.

Andererseits ist die Zügelung umfänglichen Informationsangebotes, d. h. gute Organisationsfähigkeit der inhaltlichen Information, Voraussetzung einer optimalen sprachlichen Form, jedoch nicht die einzige: Um einen störungsarmen, hinsichtlich seiner zeitlichen Struktur von Stockungen freien Verlauf des Artikulationsprozesses zu erreichen, ist es notwendig, ein gut automatisiertes, d. h. mit minimaler Zugriffszeit zur Verfügung stehendes, reiches sprachliches Formenrepertoire zu besitzen, das souveräne und zeitsparende Anwendung gestattet. Diesbezüglich gibt es bekanntlich deutliche interindividuelle – genotypische wie phänotypische – Verschiedenheiten.

Die Glätte des artikulatorischen Prozesses (fluency of speech; Flüssigkeit der Sprechtätigkeit) ist auch für die spontane, freie Formulierung – wie schon erwähnt – ein Ideal. Soweit die Verzögerung nicht als rhetorisches oder sonstwie sprechkünstlerisches Mittel eingesetzt wird, gilt sie als Defizit. Auch bei spontansprachlichen Äußerungen wird die Qualität vom Vorbild jenes zeitlichen Gleichmaßes mitbestimmt, wie es bei der lautsprachlichen Reproduktion vorgefertigter Texte zu erleben ist. Hier gelten – wenn nicht der Einsatz sprechkünstlerischer Gestaltungsmittel bei expressiv geladenen Texten besondere Variabilität verursacht – keine oder nur geringe Geschwindigkeitsschwankungen sowie seltene und nur an syntaktisch akzeptablen Stellen realisierte Pausen mit geringer Variabilität der Dauer. Der Musterfall ungestörter Fluency of Speech liegt danach in Nachrichtenlesungen des Hörfunks vor, deren emotionsarme, abständige, gewissermaßen sterile Sprechform physiognomischer Merkmale weitgehend entkleidet ist, und zwar zugunsten einer normierten Sprech-Uniform, die allenfalls leichte Klangdifferenzen der Stimme zulässt.

Der vorgefertigte Text enthebt den Sprecher jeglicher internen kognitiven und lexikalisch-syntaktischen Planung. Deshalb ist die Möglichkeit eines gleichmä-

© Frank & Timme Verlag für wissenschaftliche Literatur

ßigen, störungsfreien Zeitverlaufs von vornherein gegeben, und alle Abweichungen von diesem Gleichmaß werden dem Sprecher als mangelnde Lesetechnik, Entzifferungsschwierigkeit, als Ausdruck eines niedrigen Bildungsgrades oder bestenfalls Unkonzentriertheit ausgelegt. Im gelesenen Text wird der lapsus linguae viel deutlicher als Störung und peinliche Entgleisung empfunden und mit viel geringerer Toleranz hingenommen, weil die Hauptursache für Verlaufsstörungen und Fehlleistungen, eben die internen Planungsphasen, weitgehend ausfallen. Nun ist zwar bekannt, dass eine adäquate Form des zu lesenden Textes nur dann zustande kommt, wenn der Inhalt vollständig mitvollzogen wird, wenn also die satzakzentuelle Gestaltung tatsächlich den syntaktisch-semantischen Gegebenheiten der Textordnung gerecht wird; doch sind auch Reproduktionen von Texten möglich, bei denen der Lesende diesen inhaltlichen Nachvollzug nicht leistet, vielleicht deshalb eine Reihe kleinerer prosodischer Fehler begeht, alles in allem aber doch eine verständliche, d. h. die Bedeutung(en) übermittelnde akustische Form realisiert. Dies wäre auch erreichbar, wenn eine Wort-für-Wort-Lesung ohne die geringste prosodisch-syntagmatische Bindung geleistet würde. Dies erschwert allenfalls die Verständlichkeit, schließt sie aber nicht aus. Von solchen Fällen abgesehen, wird eine ganz und gar dem Text entsprechende, adäquate Lesung im Idealfall erst dann entstehen, wenn auch die inhaltlichen Bezüge bis in die letzten Zusammenhänge erfasst sind. Bei literarischen Texten – innerhalb derer die rhetorischen Texte einen interessanten Grenzfall darstellen – kommt es zu zeitlichen Sprechverläufen, die der Ungleichmäßigkeit der Verläufe des spontan Gesprochenen wieder in höherem Maße angenähert sind, zumal dann, wenn die Texte eine höhere emotionale Ladung besitzen. Beispielsweise werden in solchen Fällen gelegentlich sogar Wortfindungsschwierigkeiten vor überraschenden, mit hoher Information oder (und) emotionaler Ladung versehenen Wörtern simuliert und in Form von sog. Staupausen als sprechkünstlerisches (stilisiertes) Gestaltungsmittel verwendet, ebenso wie kurzzeitige Retardationen.

Ein besonders interessantes Indiz für Koordinationsmängel zwischen verschiedenen Teilphasen des Prozesses der Sprachproduktion sind die mit embolischen Lauten gefüllten Pausen, die dadurch entstehen, dass ein differenziertes artikulatorisches Programm augenblicklich nicht aktiviert ist, sondern lediglich ein diffuser motorischer Befehl an den Stimmerzeugungsapparat und die ihn mit Atemluft speisenden Atmungsorgane erteilt wird. Man könnte diesen Vorgang, bei dem unartikulierte neutrale Vokale entstehen oder Phonation bei verschlossenem Ansatzraum interjektionsähnliche Produkte, sog. embolische Laute hervorbringt, als ein “blindes“ Agieren des Stimmerzeugungsapparates auf Grund des aus-

bleibenden Informationsangebots verstehen, ein Beweis dafür, dass differenzierte motorische Programme nur dann abgerufen werden können, wenn die Planungsphasen zu einem kontextuell eingliederbaren Resultat geführt haben. Allerdings ist zu bedenken, dass es auch idiomatische Stereotypien gibt, die an solchen Stellen, wo der innersprachliche Vorlauf zu keinem lexikalisch-semantischen bzw. syntaktischen Resultat geführt hat, „einspringen“, um längere Such- und Planungspausen zu verkürzen oder allgemein Zeit zu gewinnen („nicht wahr“, „ich will mal sagen“, „gewissermaßen“, „gleichsam“ usw.).

Offensichtlich ist es nicht allein die produktive Phase des lautsprachlichen Kommunikationsprozesses, die wichtige Voraussetzungen für zeitliche Verlaufsformen der gesprochenen Sprache schafft. Wenn auch nicht unmittelbar, so wirken rezeptorische Sachverhalte doch **mittelbar** auf die lautsprachliche Gestaltung des Sprachproduktes ein. Hierbei stehen obenan die Erfordernisse der Verständlichkeit, die dem Sprecher über den sog. großen oder kommunikativen Rückmeldekreis bewusst werden, nämlich die Rückmeldung der Informationen über das Empfängerverhalten. Allein an der Reaktion des Hörers ist ablesbar, ob eine sprachliche Äußerung den intendierten, vom Sprecher beabsichtigten Erfolg hatte oder nicht.

Die Nötigung, sich verständlich zu machen, führt z. B. in Fällen sprachpathologischer Beeinträchtigungen – wie z. B. bei insuffizienter Velumfunktion im Falle von Lippen-Kiefer-Gaumen-Segel-Spalten – zu komplizierten kompensatorischen Funktionen, z. B. zur weitgehenden Laryngalisierung der oral teilweise effektlosen Artikulationsbewegungen.

Auch andere Erfordernisse der sprachlichen Form werden in der Sprachontogenese des Menschen über den kommunikativen Rückmeldekreis gewissermaßen „interiorisiert“: Nicht nur die Einhaltung eines bestimmten Präzisionsniveaus der Artikulation, sondern eben auch die zeitlichen Eigenschaften der angemessenen Sprechgeschwindigkeit oder stimmlichen Intensität (Lautheit). Es ist nicht ungewöhnlich, dass in der Sprachontogenese durch entsprechende Hinweise und Ermahnungen Einfluss auf die prosodische, speziell die zeitliche Gestaltung der Sprechweise genommen wird (Aufforderungen, langsamer zu sprechen oder „temperamentvoller“, lebhafter, nicht so schleppend, viskös usw.). Ein ungefähres Wissen darüber, welche Geschwindigkeit oder welcher Grad von Flüssigkeit der Sprechweise bestimmten Situationen oder auch Textformen – je nach Informationsdichte – angemessen ist, liefert allem Anschein nach eine Art Norm-Maßstab, an dem die Sprechweise einzelner Menschen gemessen wird, ein Wissen, dessen Unschärfe etwa mit dem zu vergleichen ist, was Kainz “Sprachgefühl“ nannte.

© Frank & Timme Verlag für wissenschaftliche Literatur

Damit ist noch nichts zu den Ursachen für diese Norm-Maßstäbe gesagt, die offenbar in den Modalitäten der Sprachwahrnehmung und rezeptorischen Verarbeitung gesprochener Texte zu suchen sind. Hier nämlich sind in ähnlicher Weise wie bei der Sprachproduktion, sogar in noch deutlicherer Ausprägung, zeitliche Grenzen gesetzt, die durch Kapazität und Dauer des inneren Informationsflusses sowie der Informationsspeicherung vor allem im Kurzzeitgedächtnis (KZG) als Operativspeicher gegeben sind.

Zu der erst in groben Umrissen geklärten Gedächtnisproblematik sei in Kürze nur so viel erwähnt, dass die Funktion des menschlichen KZG mit seiner sehr eng beschränkten Kapazität von ca. 160–200 Bit in einer Zeitspanne von ca. zehn Sekunden Bewahrdauer, offensichtlich vor allem der Operativspeicher ist, in dem höhere, integrative, das Beziehungserfassen über größere Kontextstrecken hinweg gewährleistende Verknüpfungen arrangiert werden und bewusstes Reflektieren („denkendes Auswerten“ nach Kainz) über die semantischen Grenzen des Textes hinaus stattfindet.

Die elementareren Prozesse der Mustererkennung, die sog. sensorische Speicherung und die primäre, offenbar auch die sekundäre Identifikation, also das Erkennen des Zeichenkörpers und der Zeichenbedeutung im Sinne der bilateralen Zeichenauffassung, sind den Verarbeitungsprozessen im KZG offenbar vorgeschaltet, so dass in Anlehnung an Norman und Bobrow (1975) ein lineares Dreikomponentenmodell der Speicherung anzunehmen ist, bei dem neben KZG und LZG (Langzeitgedächtnis) als dritte Komponente die sensorische Speicherung mit sehr kurzer Bewahrdauer von ca. zwei Sekunden und einer sehr geringen Kapazität angenommen wird.

Während nun die Prozesse der Mustererkennung mit großer Geschwindigkeit verlaufen und durch sie keine Überlastung des informationsverarbeitenden Systems zu erwarten ist, sind die im KZG zu leistenden höheren Verarbeitungsphasen der sprachlichen Nachricht, von denen die Qualität der Aneignung und der längerfristigen Speicherung, der Integration aufgenommener und aufbereiteter Information in vorhandene Wissensstrukturen abhängt, vielfach nicht nur durch Zeitdruck und Zeitmangel, sondern auch durch zu wenig Speicherkapazität benachteiligt. Komplizierte Informationsstrukturen oder auch ganz einfach größere Informationsmengen beanspruchen mehr Verarbeitungszeit im KZG. Bei großer Zuflussgeschwindigkeit kann es zur Beeinträchtigung der Verarbeitungsintensität kommen oder aber im Extremfall zu Speichersperren: Der Speicher ist mit noch nicht fertig verarbeiteter Information besetzt und für neue Information nicht aufnahmebereit.

© Frank & Timme Verlag für wissenschaftliche Literatur

Doch sind hier wiederum interindividuelle Verschiedenheiten in Betracht zu ziehen: Was für den einen Hörer in hohem Maße Erstinformation ist, kann für den anderen Bestätigungsinformation sein und somit einen höheren Redundanzgrad aufweisen. Gelingt es dem Sprecher, sich – wie wünschenswert – in die Rolle des Hörers zu versetzen, der mit wenig oder keiner auf den Sachverhalt bezogenen Vorinformation versehen ist, so wird er an Stellen mit mutmaßlich geringer Redundanz die Sprechgeschwindigkeit herabsetzen, um das Informationsangebot zeitlich zu strecken, er wird demgegenüber an Stellen mutmaßlich erhöhter Redundanz die Lautproduktion beschleunigen, um die negativen Auswirkungen der Unterforderung des Hörers – Herabsetzung des Aktivationsniveaus, Aufmerksamkeitsverlust, Aversion – zu vermeiden.

Auf die zeitlichen Bedingungen der Unterforderung sei hier noch beiläufig eingegangen. Die Kürze der Speicherzeit im KZG (sechs bis zehn Sekunden) machte es erforderlich, dass bestimmte Informationsportionen, die sich über größere syntaktische Strecken ausdehnen, innerhalb der "Gegenwartsdauer" abgeschlossen sind. Das gilt beispielsweise für pränominale Attribute mit ihrem Bezugswort (dem Substantiv, dem sie zugeordnet sind), deren Ausdehnung oft erheblich sein kann. Bei einer Überdehnung der Zeitdauer für die präzise Vergegenwärtigung der Information wird es zu Verzögerungen der Recodierung kommen, die es nicht mehr gestatten, die Zuflusskapazität zum LZG (ca. 0,5 Bit/s) auszulasten. Der notgedrungen eintretende Zerfallseffekt behindert die Synthese eng zusammengehörender Informationseinheiten zu Einheiten höherer Ordnung, eben die Recodierung, vor allem die rechtzeitig erfolgende Recodierung, womit auch die das Sprachverstehen (Beziehungserfassen) begleitenden Reflexionsprozesse verlangsamt werden.

Die grundsätzliche Bedeutung der syntaktischen Struktur für die Geschwindigkeit, mit der die Komplexbildung (Recodierung) verlaufen kann, sowie für den Anspruch, der an die Speicherkapazität des KZG gestellt wird, sei hier nur erwähnt. Die Strukturtiefe von Sätzen wird anhand der Linksverzweigung (Regressivität) von IC-Strukturbäumen ersichtlich. Linksverzweigte Strukturen, z. B. pränominale Attributierungen gegenüber (rechtsverzweigten) postnominalen, belasten durch vermehrte Speicherung von Zwischenresultaten die Recodierung im KZG, während rechtsverzweigte eine Art additive Linearität erzeugen. Für Apresjan (1971, 198) ist es in diesem Zusammenhang eine Frage, ob durch die Begrenzung des menschlichen KZG nicht die Sprachstruktur in der Weise beeinflusst wird, dass grammatische Mechanismen zur Begrenzung der Satztiefe geschaffen werden. Nach Yngve sind solche Mechanismen u. a. der "Vorzug,

© Frank & Timme Verlag für wissenschaftliche Literatur

den die Sprache der Nachstellung gegenüber der Voranstellung einräumt“ oder der Mechanismus zur Veränderung der Struktur von Sätzen, deren Tiefe nahe dem kritischen Punkt von sieben bzw. mehr als sieben liegt (über die Bedeutung der Zahlenangaben vgl. Apresjan a. a. O.), weiterhin eine Begrenzung regressiver Strukturen auf eine Strukturtiefe von vier bis fünf. Übrigens ist offenbar der Rezipient geschriebener Sprache in geringerem Maße für regressive Strukturen empfindlich als der Rezipient gesprochener Sprache.

Bekanntlich lässt sich die Informationsverteilung im Text pauschal – in der summarischen Berücksichtigung von syntaktischer, semantischer und pragmatischer Information – mit relativ guter Annäherung als Wahrscheinlichkeitsstruktur (Hörmann 1981) beschreiben, und zwar ermittelt aus der Unbestimmtheit der Fortsetzung eines Textes von Element (Buchstabe, Laut oder Wort) zu Element, wie sie sich z. B. anhand des Shannonschen Ratetestes ergibt; Textelemente, deren Voraussagbarkeit (Prädiktabilität) wenig unbestimmt ist (oder völlig sicher, also mit der Wahrscheinlichkeit p = 1 erfolgt) sind Glieder minimaler Information oder maximaler Redundanz. Allgemeiner: Je höher der Grad der Prädiktabilität (zwischen 0 und 1), desto geringer die Information.

Die Wahrscheinlichkeit selbst wird mittels der Menge der falschen Rate-Antworten abgeschätzt, die der richtigen Antwort vorausgehen.

Im Allgemeinen werden im Deutschen Textglieder mit hoher Information prosodisch besser ausgestattet (mit stärkeren Kontrasten versehen, auffälliger gemacht gegenüber ihrer redundanteren Umgebung) als Textglieder mit geringer Information (hoher Redundanz). Der Satzakzent hat bekanntlich die Aufgabe, die „wesentlichen“ Elemente des Satzes herauszuheben, dementsprechend sind es auch vollsemantische Wörter, die vorzugsweise als Akzentträger fungieren. Deshalb sind bei „hervorhebendem Sprechen“ häufigere, kontrastreichere Akzentuierung, verringerte Sprechgeschwindigkeit als Eigenschaften gesteigerter Eindringlichkeit Realisierungsmittel von Textstrecken mit verdichteter Information. Der textsensible Sprecher, worunter ein Sprecher verstanden sei, der fähig ist, die Verarbeitungsleistung im aktuellen Rezipienten abzuschätzen, der also im Augenblick des Sprechens nicht nur die Sprecherrolle, sondern simulativ auch die Hörerrolle einnehmen kann, dieser Sprecher also wird die Anpassung seiner Sprechweise an die Informationsstruktur des Textes intuitiv steuern, und zwar auf dem Hintergrund seiner Erfahrung als Rezipient. Er wird einerseits hinsichtlich der Grundgeschwindigkeit mit seinem Sprechtempo die angemessene Informationsdosierung anstreben, sich andererseits durch die Variation der Geschwindigkeit auf kürzeren syntagmatischen Strecken der jeweiligen Infor-

© Frank & Timme Verlag für wissenschaftliche Literatur

mationsdichte anpassen, z. B. durch geraffte Darbietung von redundanteren Einfügungen oder beiläufigen (parenthetischen) Bemerkungen. Doch – darauf sei abschließend noch einmal hingewiesen – liegen gerade hier Hindernisse für eine störungsfreie sprachliche Kommunikation.

Neben diesen kommunikationspsychologischen Sachverhalten sind andere, interne Prozesse des Rezipienten, die zu Diskrepanzen der zeitlichen Koordination im Verstehensprozess führen können, als wenig kalkulierbar und imponderabel einzuschätzen: so vor allem jene Verzögerungen der Recodierung, die auf verlangsamte Aktivierung und vergrößerte Zugriffszeit von Informationen im LZG zurückzuführen sind. Dieser Mangel an Rückflussgeschwindigkeit macht sich sowohl bei den Prozessen der sensorischen Speicherung und Mustererkennung bemerkbar – besonders deutlich bei noch nicht ausreichend automatisierter fremdsprachiger Rezeption zu beobachten, als auch bei den im KZG zu leistenden höheren Verarbeitungsleistungen, bei denen umfänglichere, komplexe und möglicherweise sehr abstrakte Informationsensemble aus dem LZG bereitgestellt werden müssen – ebenfalls in möglichst kurzer Zeit. Müßig zu erwähnen, dass somit der Zeitfaktor auch ein entscheidender Parameter der Intelligenz im Prozess der Rezeption wird: Quantität und Qualität der in der Rezeptionszeit simultan ablaufenden kognitiven Prozesse bestimmen die Qualität der Verarbeitung neu aufzunehmender Information. Der zeitliche Rahmen dieses Geschehens wird aber bei der lautsprachlichen Kommunikation – im Gegensatz zur schriftsprachlichen – durch den Sprecher und den mittels seiner Sprechgeschwindigkeit dosierten Informationsfluss vorgegeben, im Sinne strenger objektiver Bedingungen, die der Rezipient – je nach Rückmeldemöglichkeit – allenfalls durch metakommunikative Eingriffe (Bitte um langsameres Sprechen, Rückmeldung von Verstehensschwierigkeiten usw.) beeinflussen kann. Anders bei der Rezeption von Geschriebenem, wo der Rezipient den Informationsfluss selbst bestimmt, wo er außerdem die Möglichkeit hat, durch wiederholte und partielle Aufnahme des Textes die Redundanz, die er nötig zu haben glaubt, selbst zu erhöhen.

Auch bietet die schriftsprachliche Rezeption deutliche Vorteile in Bezug auf die Möglichkeit, vor allem bei größerem Textumfang neu aufzunehmende Information mit – unvollständig gespeicherter – früherer in Verbindung zu bringen. In der lautsprachlichen Kommunikation beruht das Erkennen der Beziehungsverflechtung innerhalb des gegebenen Textes auf der Leistung des Gedächtnisses; sie hängt ab von Menge, Genauigkeit und raschem Zugriff der zuvor den längerfristigen Speichern eingegebenen Information, dementsprechend auch von der Qualität der Verarbeitung.

© Frank & Timme Verlag für wissenschaftliche Literatur

4 Sprechzeit und Pausenzeit – Pausenzeitquotient (PZQ)

Der Anteil der Zeit der Sprechruhe, der signalfreien bzw. schalllosen Zeit, an der Gesamtdauer einer gesprochenen Textsequenz variiert in relativ engen Grenzen. Dass beim Sprechen Zeiten der Sprechruhe zunächst durch psychologische und physiologische Bedingungen unvermeidlich sind, und zwar durch Bedingungen auf Seiten des Senders, also der Sprachproduktion, versteht sich von selbst.

Physiologische Notwendigkeiten bestehen zunächst nur durch den Vorgang der Einatmung. Jedoch ist nur ein sehr kleiner Teil der signalfreien Zeit der Einatmung gewidmet. Auch bei einer sogenannten Atempause wird nur ein Bruchteil der Pausenzeit für den Einatmungsvorgang benutzt, vor allem der Pausenteil kurz vor Sprechbeginn. Die Einatmungszeit selbst ist gegenüber der Einatmungszeit bei der Ruheatmung absolut erheblich verkürzt. Nicht nur das Verhältnis von Einatmungszeit zu Ausatmungszeit bzw. Phonations- und Artikulationszeit ist also bei der Sprechatmung zuungunsten der Einatmung gravierend verändert, sondern die jeweilige Prozessdauer selbst. Eine psychologische Notwendigkeit der Sprechruhe ist darin zu sehen, dass der Sprecher oftmals Planungszeit benötigt, vor allem dann, wenn größere inhaltliche und syntaktische Einheiten abgeschlossen sind, zum Beispiel am Ende eines Satzes oder eines Teiltextes (Abschnitts) innerhalb einer größeren zusammenhängend gesprochenen Textmenge. Diese Zeit dient dazu, den gehörigen innersprachlichen Verlauf, wie er bei der Phasenverschiebung zwischen dem internen Sprechdenken und dem äußeren Artikulationsprozess vorliegt, herzustellen oder auch größere inhaltliche Überblicke, Rückbesinnungen, Orientierungen zu ermöglichen, Denkprozesse abzuwickeln, Alternativen zu erwägen usw. Nicht genau abzusehen ist, in welchem Maße der Zeitaufwand bei emotionaler Erregung verschiedener Richtung erhöht oder verringert wird.

Reine Sprechzeit oder Artikulationszeit, d. h. signalerfüllte Zeit (t_0) oder Gesamtsprechzeit (t) kann zur signallosen Zeit oder Gesamtpausenzeit (t_p) auf verschiedene Weise ins Verhältnis gesetzt werden: Am einfachsten wäre es, den prozentualen Zeitanteil der signallosen Zeit anzugeben, der als Pausenzeit, dividiert durch Gesamtsprechzeit, multipliziert mit 100, also

$$\frac{t_p \cdot 100}{t}$$

zu berechnen wäre und bei einem gleichgroßen Anteil von reiner Sprechzeit und Pausenzeit – gewissermaßen einem Extremwert – 50% erreichen würde. Es

könnte jedoch auch ein Quotient angenommen werden, der keinen Prozentwert, sondern einen Dezimalbruch ergibt:

$$\frac{t_p}{t}.$$

Umgekehrt ist Cowan (zit. bei van Caspel 1955) verfahren, der einen Pausenquotienten aus dem Verhältnis t_0 dividiert durch t berechnet hat und damit freilich in Kauf nimmt, dass sich der Quotient mit zunehmender Pausenzeit verkleinert. Bei Pausenlosigkeit ($t_0 = t$) ergäbe sich dafür ein Wert von 1,0; erreicht die Pausenzeit (t_p) einen Anteil von 50 % an der Gesamtzeit (t), so ergibt sich für den Quotienten ein Wert von 0,5. Es erscheint aber günstiger, einen Pausenzeitquotienten (PZQ) aus dem Verhältnis der Gesamtsprechzeit (t) zur reinen Sprechzeit (t_0) zu bilden, so dass

$$\mathrm{PZQ} = \frac{t}{t_0} \text{ oder, da } t = t_0 + t_p$$

$$\mathrm{PZQ} = \frac{t_0 + t_p}{t_0} = 1 + \frac{t_p}{t_0}.$$

Im pausenlosen Text also wäre PZQ = 1; liegt dagegen der extreme Fall $t_0 = t_p$ vor, so erhält man

$$\mathrm{PZQ} = \frac{2t_0}{t_0} = 2.$$

Der prozentuale Anteil der Pausenzeit an der Gesamtsprechzeit (t) ist – in der Form des Dezimalbruches – leicht aus PZQ zu berechnen:

$$t_p(\%) = \frac{\mathrm{PZQ} - 1}{2},$$

so dass ein PZQ-Wert von 1,28 einem Wert von 0,14 bzw. 14% (Anteil der Pausenzeit an der Gesamtsprechzeit (t)) entspricht. Selbstverständlich erhält man als Resultat durch die Division $t_p : t$ ebenfalls diesen Dezimalbruch. Bei der Ermittlung des Pausenzeitanteils wird meisten dieser Prozentwert benutzt.

Grundsätzlich ist eine Berechnung von PZQ nur sinnvoll, wenn größere zusammenhängende Sprachäußerungen, mindestens mehrere Sätze umfassend, der Berechnung zugrunde liegen. Es handelt sich also um eine textphonetische Maßzahl, für die ein Minimalumfang der zugrunde liegenden „Stichprobe“ – analog zu statistischen Maßzahlen – kalkuliert werden muss. In den hier vorgenommenen Messungen und Berechnungen liegen die Äußerungen zwischen einer und fünf Minuten Dauer.

© Frank & Timme Verlag für wissenschaftliche Literatur

4.1 PZQ in gelesenen und frei gesprochenen Texten

Der Zeitaufwand für interne Planungsprozesse, der sich beim frei gesprochenen Text in einer Zunahme der Pausenhäufigkeit und der Pausenzeit auswirkt, fällt bekanntlich beim gelesenen, nach der vorliegenden Schriftform reproduzierten Text weg, soweit keine Entzifferungsschwierigkeiten auftreten. Die vorliegenden Inhalte bedürfen freilich der exakten Bedeutungserschließung, um die suprasegmentalen Formelemente adäquat anzuwenden und dementsprechend die richtige phonostilistische Ebene bzw. angemessene Sprecherhaltung (z. B. bei der sprechkünstlerischen Interpretation von Dichtung) zu finden. Im Allgemeinen fällt der innersprachliche Aufwand nicht als Zeit beanspruchender Prozess ins Gewicht. Insofern wird eine Kalkulation des Zeitaufwandes anderer Pausenfunktionen (z. B. der Signalfunktion der syntaktischen Binnengliederung) möglich.

Einen ersten interessanten Vergleich liefern zwei relativ gut voneinander abgrenzbare hörfunkgerechte Formen, nämlich die Nachrichtenlesung (N) und die Programmansage (Prg).

Bei den Nachrichtenaufnahmen sowie den Programmansagen handelte es sich um Nachrichtenlesungen des Rundfunks der DDR aus den späten fünfziger Jahren, die auch anderen orthoepischen Untersuchungen, z. B. von E.-M. Krech und H. Ulbrich als Korpus zugrunde lagen. Für die belletristischen Texte lagen Aufnahmen der Hörbibliothek für Blinde, Leipzig, vor. Die Hörfunkaufnahmen wurden später durch Kontrolluntersuchungen einiger Rundfunkstationen der Bundesrepublik Deutschland ergänzt, die eine hochgradige phonetische Ähnlichkeit ergaben.

Bei 18 N-Proben liegt ein Mittelwert von PZQ = 1,21 mit einer Ausdehnung des Streuungsbereiches zwischen 1,13 und 1,26 vor, das Maximum der deutlich linksverschobenen Verteilung liegt im Bereich von 1,15 und 1,20.

Nachrichtenlesungen dürften zweifellos die strengste lautsprachliche Äußerungsform überhaupt darstellen. Eine gewisse suprasegmentale Ausgeglichenheit, die Einhaltung eines glatten, gleichmäßigen Sprechduktus, der individuelle Unterschiede zwar nicht völlig eliminiert, aber doch verringert, all dies lässt die Annahme zu, dass auch die signalfreie Zeit hier auf ein streng normiertes Minimum beschränkt wird. Pausen erscheinen nur an solchen syntaktischen Stellen, wo sie mit hoher Wahrscheinlichkeit auftreten, also zwischen den Sätzen und Nebensätzen usw., wo auch meist eine Markierung durch Interpunktion erfolgt.

Gegenüber der Gleichmäßigkeit der Sprechweise bei Nachrichtenlesungen weisen Programmansagen (vorwiegend Programmvorschau) bereits eine erheb-