

Universale phonologische Strukturen und Prozesse

Herausgegeben von
Karl Heinz Ramers, Heinz Vater
und Henning Wode

*Sonderdruck
aus LA 310*

Max Niemeyer Verlag
Tübingen 1994



Psychische Präsenzzeit und Selbstorganisation von Sprache

Gertraud Fenk-Oczlon und August Fenk

1. Der theoretische Ausgangspunkt

Wenn es trotz aller Verschiedenheit der Sprachen so etwas gibt wie sprachliche Universalien, so kann man sich diese dadurch erklärlich machen, daß man ein einheitliches und in der *conditio humana* gelegenes Erzeugungssystem (vgl. Chomskys "mental organ"!) von Sprachen unterstellt. Man kann sie aber auch -ausschließlich oder zusätzlich - als Folge von perzeptiven, kognitiven und motorischen Mechanismen ansehen, welche vor und außerhalb jeder Sprachverwendung wirksam sind und trotz aller sprachlichen Vielfalt gemeinsame strukturelle Merkmale erzwingen. Ein sehr starkes Argument für diese zweite Erklärung liegt jedenfalls dann vor, wenn es gelingt, aus der Kenntnis solcher Mechanismen *ganz bestimmte Universalien vorherzusagen* und nachzuweisen. (Über einen solchen Versuch wird in den Abschnitten 3 und 4 berichtet; zuvor wird in Abschnitt 2 gezeigt, daß der Rekurs auf kognitive Prinzipien relativ einfache und kohärente Erklärungen für bereits beschriebene Universalien erlaubt.)

Unter "Universalien" verstehen wir nicht nur ein Merkmal oder ein Prinzip, welches in jeder einzelnen Sprache auffindbar ist. "Other important classes of universals are those that rest on frequency distribution, in which the units are languages themselves /.../ the various statistical measures of central tendency, dispersion, limits, and so forth can be regarded as universal facts about language..." (Greenberg 1968:143). Wobei wir annehmen, daß die relevanten Variablen nicht unabhängig voneinander sind, sondern in gesetzlicher Weise interagieren. Ein solches Interagieren läßt sich statistisch als "cross-linguistische" Korrelation fassen - z.B. als negative Korrelation, welche systemtheoretisch im Sinne von Balance-Effekten gedeutet werden kann: Einzelne Sprachen, die zu einem oberen Extremwert in Variable X tendieren, tendieren aus kompensatorischen Gründen zu einem unteren Extremwert in Variable Y, und vice versa.

Die beschränkte Verarbeitungskapazität der involvierten perzeptiven und kognitiven "Prozessoren" sowie die aus diesen Beschränkungen resultierenden ökonomischen Zwänge werden also, so die zugrundeliegende Annahme, ihren Niederschlag in derartigen Universalien finden. So betrachtet sind jene Zwänge, denen sich Slobin zufolge alle natürlichen Sprachen fügen müssen - "to be clear, be processible, be quick and easy" (Slobin 1977) - Symptome von Ökonomieprinzipien in Kognition und Kommunikation. Und so man die begriffliche Unterscheidung zwischen einerseits "kogniti-

ven Zwängen" und andererseits "pragmatischen" bzw. "funktionalen Zwängen" beibehalten will, müssen die letztgenannten als Folge der kognitiven Zwänge erklärt werden: Die kognitiven Erklärungen müssen, wie Wode (1984:352) dies unter Bezug auf Slobin ausdrückt, den pragmatischen und Diskurszwängen "übergeordnet" werden.

Die hier versuchten "kognitiven Erklärungen" konzentrieren sich auf die psychische Präsenzzeit bzw. auf die Gegenwartsdauer, für deren Begrenzung zumeist die Beschränkungen unserer "sensorischen Speicherung" und unseres Kurzzeitgedächtnisses verantwortlich gemacht werden.

2. Psychische Präsenzzeit, sensorische Speicherprozesse und sprachlicher Rhythmus

Seit es eine "wissenschaftliche Psychologie" gibt, beschäftigt sie sich mit dem Thema der Dauer der erlebten Gegenwart sowie mit der "Spanne" des zugrundeliegenden "unmittelbaren Gedächtnisses". Heute wird die psychische Präsenzzeit bzw. die psychologische Gegenwart zunehmend mit sprachlichen Strukturen in Verbindung gebracht - und zwar nicht nur als Erklärung für die Gliederung des Sprachflusses in überschaubare Segmente (sowie für verwandte quantitative Phänomene bei der Reproduktion von Wortlisten, vgl. Cowan 1992), sondern auch als kognitives Prinzip, auf welches sich sprachübergreifende Aspekt-Tempus Kategorien zurückführen lassen (Lehmann 1992).

Fortschritte in der Quantifizierung der Präsenzzeit (z.B. in Abhängigkeit vom Alter der Probanden) brachten u.a. die Arbeiten Riedels (1967), und Beiträge zum Verständnis der zugrundeliegenden Prozesse lieferten in den letzten 20 bis 25 Jahren vor allem die experimentellen Untersuchungen zum sogenannten "Recency-Effekt": Bei einer unmittelbar der Präsentation folgenden freien Reproduktion werden die zuletzt gebotenen Elemente (z.B. Wörter) der Lernreihe deutlich häufiger wiedergegeben als jene aus dem mittleren Abschnitt der Lernreihe. Bei akustischer Präsentation der Lernreihe ist dieser Recency-Effekt stärker ausgeprägt als bei optischer Präsentation; hierfür hat sich der Terminus "Modalitäts-Effekt" eingebürgert (vgl. Crowder/Morton 1969). Der Modalitätseffekt läßt sich nicht nur bei Reihen von zusammenhanglosen Wörtern, sondern auch bei Sätzen der natürlichen Sprache beobachten (Fenk 1980a). Und die letzten Elemente der Lernreihe werden nicht nur häufiger wiedergegeben als die weiter zurückliegenden Elemente, sondern auch *vor* diesen anderen Elementen zu Papier gebracht (Fenk 1980b). Die auditive Überlegenheit kann man dahingehend deuten, daß sich im auditiven Bereich - eventuell in Ko-Evolution mit Sprache - ein besonders leistungsfähiger "Pufferspeicher" entwickelt hat, der den flüchtigen akusti-

schen Input für kurze Zeit in wahrnehmungsnaher Form verfügbar hält.

Über welche Zeitspanne und über wieviele Elemente sich Gegenwartsdauer und Recency-Effekt erstrecken, hängt also jedenfalls vom Alter der Probanden und von der Sinnesmodalität der Lernreihe ab, und vermutlich auch von deren Informationsdichte. (Die entsprechende und unseres Wissens bislang nicht systematisch geprüfte Hypothese: Je mehr Information pro Zeit verarbeitet werden soll, umso kürzer die zeitliche Spanne, innerhalb welcher das "unmittelbar" zuvor Gebotene psychisch präsent ist.) An eine solche informationstheoretische Deutung lassen auch jene Faktoren denken, von welchen Fraisse (1957/1985) zufolge die Dauer der subjektiven Gegenwart abhängt, nämlich vom *Zeitintervall* zwischen den Stimuli sowie von der *Anzahl* und der *Organisation* der Stimuli. Die Beispiele und Erläuterungen zu den drei genannten Faktoren entnehmen wir Fraisses "Psychologie der Zeit" (1985).

(a) Das Intervall zwischen den Stimuli

Wenn zwischen dem "Tick" und dem "Tack" einer Uhr mehr als etwa zwei Sekunden liegen, so verschwindet der subjektiv erlebte Rhythmus. Unterhalb dieses oberen Grenzwertes liegt nach den Ergebnissen aus dem Laboratorium Wundt's das optimale Intervall zwischen 0,3 und 0,5 sec, und in "der Musik läßt sich feststellen, daß die auf der Organisation des melodischen Themas basierende Anzahl der Noten in Intervallen erfolgt, die je nach dem Komponisten oder Musikstück zwischen 0,15 und 0,90 variieren" (Fraisse 1985:92).

(b) Die Stimulusanzahl

"Am Beispiel der Pendelschläge einer Uhr haben wir schon gezeigt, daß man, wenn man 3 oder 4 schlagen hört, sofort weiß, wie spät es ist, ohne die einzelnen Schläge zu zählen. Die Wahrnehmung des Ganzen oder die genaue Reproduktion von Schlägen kann, ohne diese zu zählen, unmittelbar symbolisch übertragen werden. Kinder die noch nicht zählen können, reproduzieren fehlerlos eine Reihe von fünf bis sechs Schlägen /.../. Wenn es jedoch Mitternacht schlägt, dann müssen wir zählen: Beim letzten Schlag von insgesamt 12 Stundenschlägen sind uns die ersten nicht mehr gegenwärtig" (Fraisse 1985:92).

Aufgrund der von Fraisse und von anderen Autoren durchgeführten Experimente kann man "sagen, daß die Dauer der wahrgenommenen Gegenwart stärker variiert als die Anzahl der wahrgenommenen Elemente, was eindeutig zeigt, daß die Gegenwart nicht einem vom Inhalt unabhängigen Zeitfeld entspricht" (ebd., S. 93).

Fünf oder sechs Elemente markieren diesen Untersuchungen zufolge die Grenze der Fähigkeit zur Sukzessionswahrnehmung. Diese Zahlenregion spielt auch außerhalb der Sukzessionswahrnehmung eine Rolle: Fraisse erinnert u.a. an die bekannten Dressurversuche mit Vögeln, welche lernten, auf Anhieb in jener Dose nach Futter zu suchen, auf deren Deckel eine bestimmte Anzahl von Körnern zu sehen war. Dieses "Experiment ist nur erfolgreich, wenn die Punkteanzahl des positiven Signals nicht 6 oder 7 überschreitet" (ebd. 94).

Am Rande sei vermerkt, daß die "magische Zahl sieben, plus minus zwei" auch in

der informationstheoretisch orientierten Psychologie ihren Platz behauptet hat: Wir merken uns rund 7 Einheiten - etwa 7 Buchstaben, oder aber 7 Gruppen ("chunks") aus Buchstaben, etwa Wörter. Der Prozeß des "chunking" und des "clustering" ermöglicht es uns, die Behaltensleistung für die basalen Elemente (Buchstaben) zu vervielfachen (Miller 1956); wenn die einzelnen "chunks" gut "eingelernt" sind, behalten wir bei annähernd gleichbleibender subjektiver Information ein Vielfaches an Elementen und an "objektiver Information" (vgl. Fenk 1985:362).

(c) Die Organisation der Stimuli

"Es ist eine bekannte Tatsache, daß weit mehr Elemente im Raum wahrgenommen werden können, wenn sie eine räumliche Gestalt oder eine Bedeutungseinheit bilden. Das gleiche gilt für die Zeit. Wenn identische Töne z.B. in Zweier-, Dreier-, Vierer- oder Fünfergruppen erfolgen, können wir vier oder fünf Tongruppen ohne zu zählen wahrnehmen, d.h. im günstigsten Fall insgesamt 20 bis 25 Töne" (Fraisie 1985:94).

"Die Merkfähigkeit wird offenbar ebenfalls begünstigt, wenn das Gesamte infolge der Organisation der Elemente eine Bedeutungseinheit erhält" (ebd. 95), sodaß sich die ansonsten mit 2-3 sec begrenzte Gegenwart auf ungefähr 5 sec erstrecken kann. "Diese Zeit ist notwendig, um einen Satz mit 20 bis 25 Silben auszusprechen. Selbst die längsten poetischen Verse und der längste Takt in der Musik dauern kaum länger als 5 sec..." (ebd. 95).

"Jede wahrgenommene Einheit ist in eine fließende Bewegung integriert, wobei die Beständigkeit unserer Einstellungen und unser Gedächtnis bestimmende Faktoren für diese Kontinuität sind. Die Bedeutungseinheit spielt in der Poesie sowie in der Alltagssprache eine große Rolle" (ebd. 100).

Fraisie bringt also die Präsenzzeit sowohl mit der Gestaltwahrnehmung als auch mit der sprachlichen Segmentierung in Verbindung. Zum Thema "Gestaltwahrnehmung" könnte man ergänzen: All die von den Gestalt-Theoretikern entdeckten "Bedingungen der guten Gestalt" sind aus informationstheoretischer Sicht (vgl. Attneave 1965) Bedingungen, welche die "interne Redundanz" von Element-Anordnungen erhöhen. Eine rhythmische Gliederung in überschaubare und formal ähnliche Sub-Gruppen ist ein ganz wesentlicher Beitrag zu dieser Redundanz und zur Senkung der kognitiven Kosten, und dies gilt insbesondere für die aus flüchtigen akustischen Ereignissen zusammengefügte *Zeit*-Gestalten.

Unsere aus Silben (und Wörtern) zusammengesetzten "clauses" und Sätze sind solche Zeitgestalten. Kein Wunder also, daß die Analyse sprachlicher Strukturen immer wieder auf die aus der Psychologie bekannten Gesetzmäßigkeiten stößt.

Einige Beispiele hierfür:

- o Pöppel (1985) stellt fest, daß bei etwa 75% der 200 von ihm untersuchten Gedichte die Versdauer zwischen zwei und drei Sekunden liegt.
- o Kaiser (1939) kommt zu dem Ergebnis, daß die durchschnittliche Länge von durch Sprechpausen getrennten sprachlichen Einheiten etwa sieben Silben beträgt.
- o Kozhevnikov und Chistovich (1966) sprechen von einer optimalen Verständlich-

- keit von Satzeinheiten bis zu einer Länge von sieben Silben.
- o Garrett (1975) kommt zu der Annahme, daß eine detaillierte phonologische Planung etwa fünf bis sieben Silben umfaßt.
 - o Martin (1972) stellt der seriellen Struktur von Sprache eine rhythmische (hierarchische) Struktur gegenüber. Zeitliche Begrenzungen der rhythmischen Sequenzen sollen zur Effizienz der Produktion und Perzeption beitragen, wobei sich das rhythmische Muster über etwa sieben Silben hinweg organisiert.

Die Phonologie hat inzwischen eine ganze Reihe verschiedener Namen parat, die im wesentlichen für *ein* Phänomen bzw. für *eine* bestimmte Lautgestalt stehen - nämlich für eine prosodische Einheit, welche eine kohärente Informationsstruktur aufweist und zwischen zwei Pausen liegt:

"Tone group" ist die in der britischen Tradition (z.B. Halliday 1967) bevorzugte Bezeichnung, Trager/Smith (1951) sprechen von "phonemic clauses". Pike (1945) verwendet den Terminus "rhythm units" und Crompton (1980) den Terminus "phonological phrase". Chafe spricht von "intonation units" und charakterisiert diese folgenderweise:

"When a speaker is speaking, he or she verbalizes one piece of temporarily active information after another. Each such piece is expressed in what I will call an 'intonation unit' /.../. An intonation unit is a sequence of words combined under a single coherent intonation contour, usually preceded by a pause. An intonation unit in English typically contains about five or six words, and new intonation units begin about two seconds apart. Evidently active information is replaced by other, partially different information at approximately two second intervals" (Chafe 1987:22)

3. Universelle Strukturen in der sprachlichen Realisierung von Propositionen. Eine empirische Untersuchung

Ziel unserer sprachstatistischen Analyse war die Extraktion sprachuniverseller Regularitäten bezüglich der Aggregationsmöglichkeiten innerhalb unmarkierter Aussagesätze bzw. innerhalb "prototypischer" Intonationseinheiten, welche laut Chafe (1987:38) jeweils eine klassische Subjekt-Prädikat-Struktur realisieren. Insbesondere galt es jene Konstruktionsprinzipien aufzuklären, nach denen in typologisch unterschiedlichen Sprachen die "Intonationseinheiten" aus ihren Elementen (z.B. Silben, Phonemen) aufgebaut sind. Daten, die eine derartige Analyse ermöglichen, stammen aus einer früheren experimentellen und sprachvergleichenden Untersuchung (Fenk-Oczlon 1983a). Deren Ergebnisse sollen (in Abschnitt 3.1) kurz vorgestellt werden.

3.1 Sieben Silben pro "Kernsatz"?

Will man die Art und Weise vergleichen, in der eine "Bedeutungseinheit" in unterschiedlichen Sprachen realisiert wird, ist man auf entsprechende Übersetzungen angewiesen. Um dem Ideal der Bedeutungsäquivalenz - trotz aller aus der Übersetzungstheorie bekannten Probleme, die bei sehr unterschiedlichen Sprachen ganz besonders virulent werden - wenigstens irgendwie zu entsprechen, wurde ein Pool von 22 deutschen "Kernsätzen" (vgl. die Systematik von Hundsnurscher et al. 1970) gebildet und von 28 "native speakers" in ihre jeweilige Muttersprache übersetzt. Der Kernsatz - er entspricht zumindest in "unseren" Sprachen einer einfachen Prädikat-Argument-Struktur - ist auch insofern "einfach" und am ehesten sprachuniversell vertreten, als in ihm eine semantische Grundeinheit (Proposition) mit einer syntaktischen Grundeinheit und einer Intonationseinheit zusammenfällt.

Derartige Einfachsätze haben Entsprechungen zu (vorsprachlichen?) kognitiven Konzepten und Propositionen. Sie weisen, so wird oft argumentiert, eine enge Verwandtschaft mit einem quasi vorsprachlichen Situationsverständnis auf - etwa mit der Art, wie wir Objekte wahrnehmen und wie wir die Wirkung unseres Handelns auf diese Objekte wahrnehmen. Die Annahme, daß derartige Einfachsätze, welche der sprachlichen Realisierung solcher "Propositionen" entsprechen, in allen Sprachen auffindbar sind, ist zumindest plausibel.

Das in unserem Zusammenhang wichtigste Ergebnis ist in Abbildung 1 dargestellt. Diese Illustration läßt auf einen Blick erkennen, daß die Silbenzahl pro einfachen unmarkierten Aussagesatz trotz aller Verschiedenheit der Sprachen im wesentlichen nur im Bereich 7 plus minus 2 schwankt, daß aber andererseits diese Schwankung nicht zufällig, sondern durch irgendein System geprägt ist. Unser Verdacht fiel auf die Silbenkomplexität: In Sprachen mit großer Silbenkomplexität wird eine Bedeutungseinheit durch eine relativ geringe Anzahl von Silben transportiert, Sprachen mit sehr einfachen und kurzen Silben - der Extremfall ist in dieser Hinsicht das Japanische - benötigen zur Codierung dieser Bedeutungseinheiten eine größere Anzahl von Silben. Diese Annahme konnte statistisch signifikant bestätigt werden (Fenk-Oczlon/Fenk 1985): Je mehr Silben pro einfachem Aussagesatz, desto weniger Phoneme pro Silbe. Der entsprechende Korrelationskoeffizient: $r = -0,77$, sign. (1%)

Holl.	5.05					
Franz.	5.3					
Chin.	5.4					
Tsch.	5.4					
Slov.	5.5					
Hebr.	5.5			$\bar{x} = 6.43$		
Dtsch.	5.5					
Isl.	5.5	Bamb.	6.45			
Estn.	5.7	Türk.	6.5			
Russ.	5.7	Alban.	6.5			
Sbkr.	5.8	Port.	6.6			
Engl.	5.8	Pers.	6.6			
Ewon.	5.8	Hindi	6.7	Ital.	7.5	
Ung.	5.9	Pen.	6.7	Griech.	7.5	Anjang 8.2
Arab.	5.9	Maz.	6.95	Span.	7.9	Korean. 8.2
<hr/>						
	5 - 5.99	6 - 6.99	7 - 7.99	8 - 8.99	9 - 9.99	10 - 10.99

Abb. 1: Die durchschnittliche Silbenzahl pro "Kernsatz" in 29 verschiedenen Sprachen

3.2 Einzelsprachen als Variationen eines durchgängigen Organisationsprinzips

Natürliche Sprachen sind vermutlich so gebaut, daß einerseits unsere Kapazität zur Aufnahme und Verarbeitung der sprachlichen Information nicht überfordert wird, und daß andererseits die verfügbare Kapazität in ökonomischer Weise ausgenutzt wird. Die pro Zeiteinheit übertragene Information wird also nur relativ wenig variieren, und vorzugsweise innerhalb einer nach oben und nach unten hin irgendwie begrenzten "Bandbreite". Aus dieser allgemeinen Annahme eines annähernd "konstanten" sprachlichen Informationsflusses kann man spezifische Annahmen bezüglich der kleinsten Artikulationseinheit, nämlich der Silbe, ableiten. Diese Postulate - unten durch Kursivstellung hervorgehoben - betreffen die Silbe als Baustein von Wörtern (A) und als Baustein von Sätzen (B):

- A *In Sprachen mit komplexer gebauten Wörtern, die auszusprechen mehr Zeit in Anspruch nimmt, wird pro Wort mehr Information übertragen als in Sprachen mit einfacher gebauten, kürzeren Wörtern.* Und tatsächlich läßt sich nachweisen (Fucks 1956/1964, Fenk/Fenk 1980), daß in Sprachen mit größerer Wortlänge, gemessen in Silben, auf das durchschnittliche Wort ein "entsprechend" höherer Informationsgehalt entfällt.

- B *Da die Länge der Silben, also der kleinsten Artikulationseinheiten, nicht sehr variabel ist, wird auch ihre Anzahl pro Proposition über die Sprachen hinweg nur wenig variieren.* Und tatsächlich konnte gezeigt werden, daß die Silbenanzahl pro "einfachen unmarkierten Aussagesatz" über die Sprachen hinweg relativ stabil ist; die Schwankungsbreite liegt bei etwa sieben plus minus zwei Silben. Innerhalb dieser Schwankungsbreite sollte aber folgendes gelten: *In Sprachen mit komplexer gebauten Silben, die auszusprechen mehr Zeit in Anspruch nimmt,¹ werden Propositionen mithilfe einer geringeren Anzahl von Silben realisiert.* Über eine entsprechende negative Korrelation zwischen der Anzahl der Silben pro Satz und der Anzahl der Phoneme pro Silbe haben wir ebenfalls bereits in Abschnitt 3.1 berichtet.

Diese zuletzt genannte "cross-linguistische" Regularität ist nun in zweifacher Hinsicht interessant:

Erstens als Hinweis darauf, daß zusätzlich zu unserem "Prinzip des konstanten Informationsflusses" noch so etwas wie ein "Prinzip der clausalen Verarbeitung" wirksam ist. *Wenn die "Größe", die "Komplexität" oder die "Länge" jener Segmente (z.B. Silben), deren Anzahl als Maßstab für die "Größe" oder "Komplexität" oder "Länge" eines umfassenderen Segmentes (z.B. Satz) herangezogen wird, ihrerseits in systematischer Weise mit dieser Anzahl zusammenhängt, dann vermutlich im Sinne einer negativen Korrelation.*

Diese Gesetzhypothese kann man als Verallgemeinerung der von uns gefundenen Regularität betrachten. Und sie folgt auch aus dem Prinzip des konstanten Informationsflusses - unter der Voraussetzung freilich, daß das umfassendere Segment irgendwelche Kapazitätsgrenzen der involvierten Prozessoren tangiert. Genau diese Voraussetzung scheint bei den sogenannten "clauses" und bei den sogenannten "Kernsätzen" gegeben zu sein: Fünf sehr komplexe oder zehn sehr kurze Silben - das ist offenbar so etwas wie eine mundgerechte Portion für die "clausale" Verarbeitung von Sprache. Die Annahme einer "clausalen" Verarbeitung steht also keineswegs im Widerspruch zur Annahme des konstanten Informationsflusses. Ein Beispiel: Wenn der durstige Biertrinker sein Glas in großen Schlucken (= "clauses") leert, dann ist dies ein Fall von "clausaler" Verarbeitung, und trotz der Diskontinuität des Bierdurchsatzes ist dieser "konstant" - pro Schluck wegen (des begrenzten Schlundquerschnittes und) der be-

¹ Die zeitliche Dauer einer Silbe hängt natürlich nicht ausschließlich von der Phonemzahl der Silbe ab. "The duration of a syllable depends, on the one hand, on its complexity and specific phonological composition and, on the other hand, on stress and overall tempo and position within a phrase or a stress group (Fant et al. 1991:356). Aber die Phonemzahl ist gewissermaßen die "durchgängig" wirksame Determinante: Ein gesteigertes Sprechtempo verkürzt die Dauer phonemreicher wie auch phonemärmer Silben, sodaß die phonemreichen nach wie vor im Durchschnitt länger dauern.

Auch liegt es uns ferne, große Phonemzahl mit großer Komplexität oder mit kompliziertem Bau der Silbe gleichzusetzen. Aber sie ist eine wesentliche Voraussetzung!

grenzten Schluckdauer, und über einen längeren Zeitabschnitt hinweg deshalb, weil innerhalb dieses Zeitabschnittes nur eine eng begrenzte Anzahl großer Schlucke Platz hat.

Offensichtlich sind der Arbeitsrhythmus der Sprachverarbeitung und der Arbeitsrhythmus der Sprachproduktion präzise aufeinander abgestimmt. (Was schon im Hinblick darauf zweckmäßig und erklärlich ist, daß jeder Sprecher immer auch Hörer und Kontrolleur seiner eigenen Sprachproduktion ist). Die einfache Silbe entspricht von der Motorik her der kleinstmöglichen sprachlichen Einheit, bei welcher ein "right-ear-advantage" Effekt festgestellt wurde (Kimura 1976). Die "Kernsätze" entsprechen, kognitionstheoretisch gesehen, der sprachlichen Realisierung einer als "Proposition" bekannten "Bedeutungseinheit", und von der Motorik her einer in Silben gegliederten "breath-group" (Lieberman 1967).

Und *zweitens* ist unsere "cross-linguistische" Regel - "je größer der Satz in Silben, desto kleiner die Silbe in Phonemen" - interessant, weil sie sich perfekt unter dem verallgemeinerten Menzerathschen Gesetz - "je größer das Ganze, umso kleiner die Teile" - zur Deckung bringen läßt.

Menzerath hatte schon sehr früh die quantitativen Beziehungen zwischen der Silbenzahl und der Phonemzahl deutscher Wörter untersucht und dabei folgende Regularität gefunden:

"I. Die relative Lautzahl nimmt mit steigender Silbenzahl ab, oder mit anderer Formel gesagt: je mehr Silben ein Wort hat, um so (relativ) kürzer (lautärmer) ist es.
 /.../ Es tritt eine 'Sparsamkeitsregel' in Erscheinung, die sich psychologisch auf eine Ganzheitsregel dieser Art gründet: je größer das Ganze, um so kleiner die Teile! Diese Regel /.../ wird aus der Tatsache verständlich, daß das Ganze jeweils 'überschbar' bleiben muß." (Menzerath 1954:100f)

Wir gehen nun davon aus, daß diese Regel auch für andere Sprachen gilt (vgl. z.B. Grotjahn 1982, Köhler 1989). Darüberhinaus nehmen wir an (Hypothese I'), daß sie auch im Sinne einer "cross-linguistischen Funktion" Gültigkeit hat. (Deren graphische Realisierung wäre ein Streudiagramm - mit eingezeichneter Regressionsfunktion, mit der Silbenzahl pro Wort als x-Achse und der Phonemzahl pro Silbe als y-Achse, dessen einzelne Punkte nicht für verschiedene Wörter einer Sprache stehen, sondern für Durchschnittswerte verschiedener Einzelsprachen.)

Aus Hypothese I', also der "cross-linguistischen" Erweiterung der Menzerathschen Regel I, in Verbindung mit unserer bereits statistisch bestätigten Hypothese (nennen wir sie "Hypothese II'!) resultiert eine Annahme ("Hypothese III"), welche den Gültigkeitsbereich von Menzeraths Verallgemeinerung einschränkt (vgl. Abb. 2 sowie Fenk/Fenk-Oczlon 1993), obwohl beide Prämissen unter diese Verallgemeinerung passen:

H I' Je mehr Silben pro Wort, umso weniger Phoneme pro Silbe

H II Je mehr Silben pro Satz, umso weniger Phoneme pro Silbe

 H III Je mehr Silben pro Satz, umso *mehr* Silben pro Wort (Präsizer: Je weniger Phoneme pro Silbe, umso mehr Silben pro Satz und pro Wort)

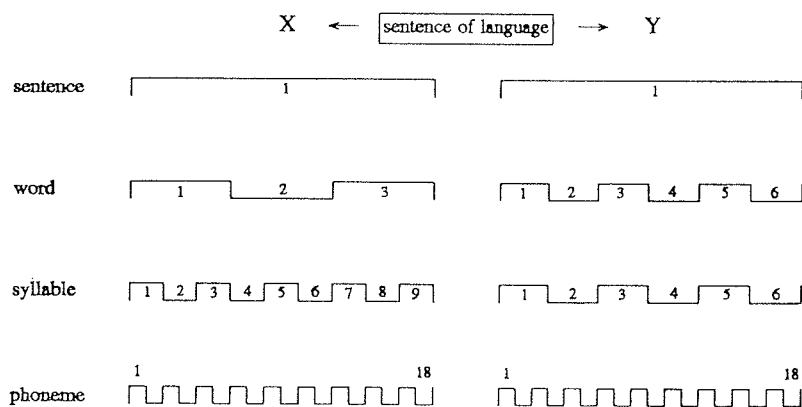


Abb. 2: Eine graphische Veranschaulichung der in unserem Modell prognostizierten Koinzidenz von geringer Phonemzahl pro Silbe mit hoher Silbenzahl pro Satz und pro Wort (aus: Fenk/Fenk-Oczlon 1993)

Wenn im "cross-linguistischen" Vergleich die Anzahl der Silben pro Satz "konstant" ist bzw. nur innerhalb eines engen Rahmens schwankt, dann müssen die Wörter von Sprachen, deren Sätze im Durchschnitt aus relativ vielen Wörtern zusammengesetzt sind, aus relativ wenigen Silben zusammengesetzt sein (vgl. auch unser in Abbildung 2 skizziertes Modell!). Dieses Postulat in anderen Worten:

H IV Je mehr Wörter pro Satz, umso weniger Silben pro Wort

Das **Ergebnis** der Auswertung:

Auch die den neuen Hypothesen I', III und IV entsprechenden Korrelationen erwiesen sich als statistisch bedeutsam. Die in der folgenden Zusammenstellung angeführten Koeffizienten erhielten wir bei Berechnung des *linearen* Zusammenhanges. (Bei

Berechnung nicht-linearer, z.B. quadratischer Regressionsfunktionen ist die Summe der quadrierten Abweichungen noch kleiner bzw. die "Korrelation" noch höher. Vgl. Fenk/Fenk-Oczlon 1993)

I'	Je mehr Silben pro Wort, umso weniger Phoneme pro Silbe.	$r = -0,45$ s
II	Je mehr Silben pro Satz, umso weniger Phoneme pro Silbe.	$r = -0,77$ s
III	Je mehr Silben pro Satz, umso <i>mehr</i> Silben pro Wort.	$r = 0,38$ s
IV	Je mehr Wörter pro Satz, umso weniger Silben pro Wort.	$r = -0,69$ s

Was sich im "cross-linguistischen" Vergleich als selbstregulierendes System darstellt, läßt sich auch im kontrastiven Vergleich zweier Sprachen demonstrieren. Die Analyse der aus einem Sprachlehrbuch ("Türkisch für Deutsche", Liebe-Harkort 1980) entnommenen Stichprobe deutscher Dialog-Passagen und ihrem türkischen (Übersetzungs-) Äquivalent brachte folgendes Resultat: Die Probe umfaßte im Deutschen, verglichen mit dem Türkischen, mehr Wörter, weniger Silben, also auch weniger Silben pro Wort, sowie mehr Phoneme pro Silbe. Der Zweisprachen-Vergleich förderte also dieselben Größenverhältnisse (vgl. Tab. 1) zutage wie der "cross-linguistische" Vergleich:

I	Je mehr Silben pro Wort, umso weniger Phoneme pro Silbe.
II	Je mehr Silben pro Text, umso weniger Phoneme pro Silbe.
III	Je mehr Silben pro Text, umso mehr Silben pro Wort.
IV	Je mehr Wörter pro Text, umso weniger Silben pro Wort.

Tab. 1: Deutsch-Türkisch im Vergleich (siehe Text!)

n der Wörter	322	211
n der Silben	499	517
Ø Silbenanzahl/Wort	1,549	2,450
Ø Phonemzahl/Silbe	2,57	2,25

Man darf natürlich nicht erwarten, daß jeder denkbare Paarvergleich zwischen Einzelsprachen sich so perfekt in das von uns skizzierte cross-linguistische Bild fügt. (Sonst hätten ja auch die entsprechenden Korrelationskoeffizienten noch höher ausfallen müssen!) Betrachtet man zum Beispiel die von Fant et al (1991:352) publizierten Ergebnisse eines Vergleichs Schwedisch/Französisch/Englisch im Hinblick auf unsere Theorie, so zeigen sich bei insgesamt zwölf Vergleichsmöglichkeiten (aus drei Sprachen und vier Gesetzeshypothesen) gleich fünf "Ausnahmen"; drei davon stammen aus Vergleichen mit dem Französischen. Die Ergebnisse dieser Pilotstudie von Fant et al.

basieren auf einer recht kurzen Textprobe, nämlich einer einminütigen Passage aus einer vom Schwedischen ins Englische und Französische übersetzten Novelle; und die Unterschiede zwischen den statistischen Kennwerten, an denen unsere Regeln "scheitern", sind so gering, daß man keine allzu weitreichenden Folgerungen daran knüpfen sollte. (Die durchschnittliche Silbenanzahl pro Wort beträgt z.B. im Französischen 1,387 und im Englischen 1,303.) Trotzdem wollen wir die Resultate als einen Hinweis auf eine mögliche Präzisierung unseres Modells ernst nehmen. Die entsprechende, vorläufig eher spekulative Annahme:

Bei in zwei verschiedenen Sprachen identischen Durchschnittswerten (z.B. gleiche Wortanzahl im Satz) können dennoch unterschiedliche Streuungswerte auftreten und bedeutsam werden. Bei Sprachen mit hohem Synthesegrad ist die Streuung der Wortlänge viel geringer als in eher isolierenden Sprachen wie dem Französischen, wo vergleichsweise lange Wörter und sehr kurze Funktionswörter einander abwechseln.

4. Sprachtypologische Perspektiven

Vermutungen über wechselseitige Abhängigkeiten zwischen phonologischen Eigenschaften, morphologischen Eigenschaften und Wortfolge-Eigenschaften von Sprachen wurden in der typologischen Literatur immer wieder geäußert: SOV-Sprachen sind z.B. nach Einschätzung von Skalička (1968) und Lehmann (1978) häufiger innerhalb der agglutinierenden Sprachen anzutreffen, und Lehmann zufolge tendieren sie zu einfachen C(C)V Silben sowie zu Vokalharmonie. Auch Donegan und Stampe (1986) sowie Gil (1986) diskutieren mögliche Zusammenhänge zwischen einerseits agglutinierendem Sprachtyp und einfacheren Silbenstrukturen und andererseits fusionierendem Sprachtyp und eher komplexen Silbenstrukturen.²

Beziehungen zwischen Charakteristika im Sprachrhythmus und in der Silbenstruktur

² Bei den zuletzt genannten Autoren spielt auch die Unterscheidung von jambischen und trochaischen Sprachen eine Rolle. Allerdings bezieht sich die Unterteilung jambisch vs. trochaisch bei Donegan & Stampe auf den "word canon", bei Gil hingegen auf größere Einheiten. So ist es nicht verwunderlich, daß jambische Sprachen nach Donegan/Stampe eher fusionierend (oder isolierend) und trochaische Sprachen eher agglutinierend sind, während es sich bei Gils Wortgebrauch umgekehrt verhält: Jambische Sprachen haben einfachere Silbenstrukturen und sind eher agglutinierend, trochaische Sprachen haben komplexere Silbenstrukturen und sind eher fusionierend. Bei Gil beruht die Unterteilung in jambische vs. trochaische Sprachen auf der Annahme, daß in jambischen Sprachen schwache Elemente vor starken stehen, also Modifier vor Head, und daß in trochaischen Sprachen die starken Elemente vor den schwachen stehen, also Head vor Modifier. Ramat (1986) wendet allerdings ein, daß die Relation zwischen Operator und Operand nicht eine Sache der Stärke, sondern der Determination sei.

werden insbesondere im Zusammenhang mit der Unterscheidung zwischen "stress-timed" und "syllable-timed" Sprachen thematisiert: Hinter der Unterteilung zwischen diesen beiden Sprachrhythmustypen verbirgt sich letztlich, so vermutet eine wachsende Anzahl von Autoren (z.B. Roach 1982, Miller 1984), nicht mehr als eine unterschiedliche Silbenkomplexität sowie eine unterschiedliche Anfälligkeit von Vokalen in unbetonten Silben gegenüber Reduktionsprozessen. "Stressed-timed" Sprachen tendieren offenbar zu Vokalreduktionen und haben - daher? - komplexere Silben als "syllable-timed" Sprachen.

Für eine solche Deutung sprechen auch die Ergebnisse aus dem von Fant et al. (1991) durchgeführten Vergleich von Betonungs- und "timing"-Charakteristika im Schwedischen, Englischen und Französischen. Im Französischen, einer "syllable-timed" Sprache, sind die Längenunterschiede zwischen betonten und unbetonten Silben geringer als in "stress-timed" Sprachen wie dem Schwedischen oder dem Englischen. Die relative Uniformität der Silbendauer im Französischen führen die Autoren auf das Überwiegen von CV-Silben und auf den geringen Ausprägungsgrad von betonungsinduzierten Längungen zurück. Hingegen erhöhen Vokalreduktionen im Englischen und Schwedischen den Längenkontrast zwischen betonten und unbetonten Silben.

Inwieweit passen nun unsere statistischen Ergebnisse in die hier vermuteten Zusammenhänge zwischen morphologischen, phonologischen und Wortfolge-Charakteristika?

4.1 Silbenreiche Wörter und phonemarme Silben als Charakteristikum agglutinierender Sprachen

In unserem Ergebnis *"Je mehr Silben pro Wort, umso weniger Phoneme pro Silbe"* kann man ein weiteres Argument für die verbreitete, aber kaum systematisch geprüfte Annahme sehen, daß agglutinierende Sprachen einfachere Silbenstrukturen haben. Der entsprechende Gedankengang:

Wir gehen davon aus, daß in Sprachen mit einer starken Agglutinationstendenz eine starke Tendenz zu vielsilbigen Wörtern zu verzeichnen ist. Agglutinationsprozesse werden, das liegt gewissermaßen in der Natur von "Agglutination", zu vielsilbigen Wörtern führen, sodaß diese Vielsilbigkeit umgekehrt auch als Symptom oder Indiz für die Zugehörigkeit der jeweiligen Sprache zum agglutinierenden Sprachtyp gewertet werden darf. Damit ist noch nicht behauptet, daß eine große Silbenzahl pro Wort *nur* durch Agglutination zustande kommt und somit ein *hinreichendes* Klassifikationskriterium ist. Immerhin könnte die Vielsilbigkeit der Wörter ja auch, so der naheliegende Einwand, durch eine Fusionierungstendenz etabliert oder begünstigt werden. Vorsicht ist auch aus Sicht der Forschungsmethodik geboten: Strenggenommen hätte man von vornherein einen Quotienten aus der Anzahl der grammatikalischen

Morpheme pro Wort und der Anzahl der sie repräsentierenden Affixe bilden und sich auf eine typologische Zuordnung zum eher agglutinierenden oder eher fusionierenden Sprachtyp festlegen müssen. (Zur Erläuterung dieses Quotienten: Ein hoher Fusionsgrad liegt dann vor, wenn viele grammatikalische Morpheme durch wenige Affixe realisiert werden. Bei einem 1:1 Verhältnis von grammatikalischen Morphemen und Affixen hätten wir es mit einer rein agglutinierenden Sprache zu tun.) Trotzdem wollen wir einen Schritt weitergehen und behaupten, daß eine große Silbenzahl pro Wort mit dem agglutinierenden Sprachtyp *enger* assoziiert ist als etwa mit dem fusionierenden Sprachtyp. Denn was den fusionierenden Sprachtyp in erster Linie kennzeichnet (vgl. Comrie 1981, der deshalb auch den Terminus "flektierend" für weniger passend hält) und graduell vom agglutinierenden Sprachtyp unterscheidet, das ist der Verschmelzungsgrad der grammatikalischen Morpheme:

In agglutinierenden Sprachen sind die Morpheme klar segmentierbar und invariant, im Unterschied zu fusionierenden Sprachen, die außerdem zu Kasussynekretizismus tendieren. Schon aus diesem Kriterium läßt sich folgendes ableiten: Fusionierende Sprachen neigen zumindest bei gleichem Synthesegrad, also bei gleicher Anzahl der Morpheme pro Wort, zu kürzeren Wörtern, da ja die einzelnen grammatikalischen Morpheme miteinander "verschmolzen" sind. Und dieses Verschmelzen könnte (in der Diachronie) durch bestimmte phonologische Prozesse wie Vokalreduktionen und -tilgungen gefördert werden, welche wiederum die Tendenz zu komplexeren Silben (Konsonantenclustern) begünstigen. Solche Tilgungen sind, wie schon erwähnt, typisch für "stress-timed" Sprachen. In agglutinierenden Sprachen hingegen dürften solche Reduktionen seltener³ sein, was durchaus eine Affinität zwischen einfacher Silbenstruktur, "agglutinierend" und "syllable-timed" vermuten läßt. Die Annahme einer solchen Affinität ist also theoretisch begründbar, und sie gewinnt an Plausibilität dadurch, daß sie eine Erklärung für unsere statistische Regularität ("Je mehr Silben pro Wort, umso weniger Phoneme pro Silbe") zu liefern mag.

4.2 Sätze mit SOV-Wortfolge und großer Silbenzahl als Charakteristikum agglutinierender Sprachen

Die Resultate der hier referierten Experimente und Auswertungen stützen Lehmanns bereits erwähnte Diagnose, wonach SOV-Sprachen zu einfachen Silben und zu einer

³ Diese wenigen auch in "rein" agglutinierenden Sprachen vorkommenden Vokalreduktionen und -tilgungen dürften, so eine weitere Hypothese, vorzugsweise bei besonders häufigen Wörtern auftreten (z.B. türk. *burada* 'hier' zu *burda*), da bei sehr häufigen Wörtern der ökonomische Vorteil einer Verkürzung am größten ist (vgl. Fenk-Oczlon 1989).

agglutinierenden Morphologie neigen.⁴ Darüberhinaus erweitern sie die Gruppe der voneinander wechselseitig abhängigen Parameter (Agglutinierung, SOV-Wortfolge, einfache Silben) um eine weitere Größe, nämlich um die *Silbenzahl* (pro Wort; pro Satz), die ja unserer Untersuchung zufolge in gesetzmäßiger Weise mit der Phonemzahl pro Silbe (also der "Einfachheit" der Silbe) interagiert.

Einfach ausgedrückt: Agglutination erfordert "einfache" (phonemarme) und "gleichbleibende" Bausteine oder wird durch solche Bausteine begünstigt. Und die Kodierung von Bedeutungseinheiten braucht mehr Silben, wenn die Silben "kleiner" und "einfacher" sind. Wenn nun SOV-Sprachen tatsächlich zum Agglutinieren und zu einfacher gebauten Silben neigen, dann müßten sie auch zu größerem Silbenaufwand neigen.

Um diese Hypothese zu prüfen, haben wir eine Zuordnung der 29 untersuchten Sprachen zur Gruppe der SOV-, SVO- und VSO-Sprachen getroffen, und zwar auf Basis des in den Versuchs-Sätzen festgestellten zahlenmäßigen Überwiegens der jeweiligen Reihenfolge von Subjekt, Objekt und Verb. Anschließend wurde die durchschnittliche Silbenzahl für jede dieser Gruppen ermittelt. Das Ergebnis: Erwartungsgemäß waren die Sätze der insgesamt acht SOV-Sprachen im Mittel aus einer relativ großen Anzahl ($\bar{x} = 7,2$) relativ einfach gebauter Silben gebildet. In den 19 SVO-Sprachen betrug der Silbenaufwand pro Satz im Mittel nur 6,2 und in den beiden VSO-Sprachen nur 5,7 Silben.

5. Eine kurze Bilanz

Abschließend kann festgehalten werden: Offenbar gibt es über die Sprachen hinweg eine Tendenz, Bedeutungseinheiten (Propositionen) mit Hilfe einer begrenzten, motorisch und kognitiv bewältigbaren Anzahl von Silben zu realisieren. Die Variationsbreite der Silbenzahl ist beschränkt, und die Variation *innerhalb* dieser beschränkten Bandbreite relevant im Hinblick auf typologische Kategorien: Die im "cross-linguistischen" Vergleich statistisch feststellbare, "typologische" Koinzidenz bestimmter Ausprägungsgrade bestimmter Variablen in bestimmten Sprachen - z.B. das Zusammenfallen von hoher Silbenzahl pro Satz und hoher Silbenzahl pro Wort mit geringer Phonemzahl pro Silbe - wird auf diese Weise systemtheoretisch erklärbar als Folge kognitiver, universell (und übrigens auch in der *Sprachproduktion*) wirksamer Mechanismen, welche erstens die Größe der sprachlichen Aggregate beschränken und welche zweitens bei Verschiebungen entlang einer Dimension eine kompensatorische Ver-

⁴ Mutmaßungen über die Ursache der Affinität zwischen SOV-Wortfolge und agglutinierendem Sprachtyp wurden bereits an anderer Stelle angestellt (Fenk-Oczlon 1983b).

schiebung in anderen Dimensionen erzwingen.

Literatur

- Attneave, F. (1965): Informationstheorie in der Psychologie. - Bern: Huber.
- Chafe, W. (1987): Cognitive constraints on information flow. - In: R.S. Tomlin (ed.): Coherence and grounding in discourse (Amsterdam: Benjamins) 21-51.
- Comrie, B. (1981): Language Universals and Linguistic Typology. - Oxford: Basil Blackwell.
- Cowan, N. (1992): Verbal Memory Span and the Timing of Spoken Recall. - In: Journal of Memory and Language 31, 668-684.
- Crowder, R.G. and Morton, J. (1969): Precategorical acoustic storage (PAS). - In: Perception & Psychophysics 5, 365-373.
- Crompton, A. (1980): Timing patterns in French. - In: Phonetica 37, 205-234.
- Donegan, P.J. and Stampe, D. (1983): Rhythm and the holistic organization of language structure. - In: J.F. Richardson, M. Karks, A. Chukerman (eds.): Papers from the parasession on the interplay of phonology morphology and syntax (Chicago linguistic society) 337-353.
- Fant, G., Kruckenberg, A. and Nord, L. (1991): Durational correlates of stress in Swedish, French and English. - In: Journal of Phonetics 19, 351-365.
- Fenk, A. (1980a): "Ein Bild sagt mehr als tausend Worte...?" Lernleistungsunterschiede bei optischer, akustischer und optisch-akustischer Präsentation von Lehrmaterial. - In: AV-Forschung 23, 3-50.
- (1980b): Temporale Struktur und Sinnesmodalität von Nachrichten als Determinanten der Gedächtnisorganisation. - In: Psychologische Beiträge 22, 3, 396-413.
- (1985): Is the reduction of subjective information (per unit of time) independent from the starting level? - In: G. d'Ydewalle (ed.): Cognition, information processing, and motivation (Amsterdam: Elsevier Science Publishers B.V.) 361-373.
- Fenk, A. und Fenk, G. (1980): Konstanz im Kurzzeitgedächtnis - Konstanz im sprachlichen Informationsfluß? - In: Zeitschrift für experimentelle und angewandte Psychologie 27, 400-414.
- Fenk, A. und Fenk-Oczlon, G. (1993): Menzerath's law and the constant flow of linguistic information. - In: R. Köhler/B. Rieger (eds): Contributions to quantitative linguistics (Kluwer Academic Publishers).
- Fenk-Oczlon, G. (1983a): Bedeutungseinheiten und sprachliche Segmentierung. Eine sprachvergleichende Untersuchung über kognitive Determinanten der Kernsatzlänge. - Tübingen: Narr.
- (1983b): Ist die SVO-Wortfolge die "natürlichste"? - In: Papiere zur Linguistik 29, 23-32.
- (1989): Geläufigkeit als Determinante von phonologischen Backgrounding-Prozessen. - In: Papiere zur Linguistik 40, 91-103.
- Fenk-Oczlon, G. and Fenk, A. (1985): The mean length of propositions is seven plus minus two syllables - but the position of languages within this range is not accidental. -In: G. d'Ydewalle (ed.): Cognition, information processing, and motivation (Amsterdam: Elsevier Science Publishers B.V.) 355-359.
- Fraisse P. (1985): Psychologie der Zeit. - München: Ernst Reinhardt (Original edition: Psychologie de temps. Presses Universitaires de France, Paris 1957).
- Fucks, W. (1956/1964): Die mathematischen Gesetze der Bildung von Sprachelementen aus ihren Bestandteilen. - In: Nachrichtentechnische Fachberichte, Beiheft der NTZ 3, 7-21.
- Garrett, M. (1975): The analysis of sentence production. - In: G. Bower (ed.): Psychology of learning and motivation 9 (New York: Academic Press) 133-177.
- Gil, D. (1986): A prosodic typology of language. - In: Folia Linguistica XX/1-2, 165-231.

- Greenberg, J.H. (1968): *Anthropological Linguistics: An Introduction*. - New York: Random House.
- Grotjahn, R. (1982): Ein statistisches Modell für die Verteilung der Wortlänge. - In: *Zeitschrift für Sprachwissenschaft* 1, 44-75.
- Halliday, M.A.K. (1967): *Intonation and grammar in British English*. - The Hague: Mouton.
- Hundsnerscher et al. (1970): *Transformationelle Schulgrammatik*. Göppinger Akademische Beiträge Nr. 7. - Göppingen.
- Kaiser, I. (1939): Some properties of speech muscles and the influence thereof on language. - In: *Arch.neerl.phonet.exper.* 15, 1-78.
- Kimura, D. (1976): The asymmetry of the human brain. - In: *Recent progress in perception*. Scientific American (San Francisco: Freeman & Co. LTD) 246-254.
- Kozhevnikov, V.A. and Chistovich, L.A. (1966): *Speech: articulation, and perception (revised)*. - Washington D.C.: U.S. Department of Commerce, Joint Publication Research Service.
- Köhler, R. (1989): Das Menzerathsche Gesetz als Resultat des Sprachverarbeitungsmechanismus. - In: G. Altmann/M.H. Schwibbe (Hrsg.): *Das Menzerathsche Gesetz in informationsverarbeitenden Systemen* (Hildesheim: Olms Verlag) 108-112.
- Lehmann, V. (1992): Grammatische Zeitkonzepte und ihre Erklärung. - In: *Kognitionswissenschaft* 2, 156-170.
- Lehmann, W.P. (1978): English: A characteristic SVO language. - In: W.P. Lehmann (ed.): *Syntactic typology* (Sussex: Harvester Press).
- Liebe-Harkort, K. (1980): *Türkisch für Deutsche*. - Frankfurt am Main: Scriptor Verlag.
- Lieberman, P. (1967): *Intonation, perception and language*. - Cambridge: MIT-Press.
- Martin, J.G. (1972): Rhythmic (hierarchical) versus serial structure in speech and other behavior. - In: *Psychological Review*, 487-509.
- Menzerath, P. (1954): *Die Architektur des deutschen Wortschatzes*. *Phonetische Studien*, Heft 3. - Bonn: Ferd. Dummlers Verlag.
- Miller, G.A. (1956): The magical number seven, plus or minus two: some limits on our capacity for processing information. - In: *Psychological Review* 63, 81-97. Dtsch.: In: D. Norman (Hg.): *Aufmerksamkeit und Gedächtnis* (Weinheim und Basel: Beltz 1973).
- Miller, M. (1984): On the perception of rhythm. - In: *Journal of Phonetics* 12, 73-83.
- Pike, K.L. (1945): *The intonation of American English*. - Ann Arbor: University of Michigan Press.
- Pöppel, E. (1985): *Grenzen des Bewußtseins*. - Stuttgart: Deutsche Verlagsanstalt.
- Ramat, P. (1986): Is a holistic typology possible? - In: *Folia Linguistica* XX/1-2, 3-14.
- Riedel, H. (1967): *Psychostruktur*. - Quickborn: Verl. Schnelle.
- Roach, P. (1982): On the distinction between "stress-timed" and "syllable-timed" languages. - In: D. Crystal (ed.): *Linguistic controversies* (London: Edward Arnold) 73-79.
- Skalička, V. (1968): Problém syntaktické typologie. - In: *Slavica Pragensia* 10, 137-142.
- Slobin, D.I. (1977): *Language change in childhood and in history*. - In: J. Macnamara (ed.): *Language learning and thought* (New York, San Francisco, London) 185-214.
- Trager, G.L. and Smith, H.L. (1951): *An outline of English structure*. - In: *Studies in Linguistics*. Occasional Papers, vol. 3.
- Wode, H. (1984): Psycholinguistische Grundlagen sprachlicher Universalien: Möglichkeiten eines empirischen Paradigmas. - In: *Folia Linguistica* XVIII/3-4, 345-377.