

„Early immediate constituents“ – ein kognitiv-funktionales Prinzip der Wortstellung(svariation)

Christiane Hoffmann

1. Einleitung oder vom „Gesetz der wachsenden Glieder“

Im folgenden wird ein Mechanismus vorgestellt, der sprachenübergreifend Wortstellung und Wortstellungsvariation beeinflusst und in quantitativen, empirischen, sprachenübergreifenden Untersuchungen bereits mehrfach getestet wurde, vgl. Hawkins (1994). Das Prinzip der „Early Immediate Constituents“ (EIC) formuliert einen an Grundlagen der sprachlichen Informationsverarbeitung orientierten Mechanismus, der bestimmend dafür ist, daß bestimmte Linearisierungsmöglichkeiten von Äußerungsteilen anderen gegenüber präferiert werden. Der gesamte Argumentationszusammenhang beruht auf Annahmen zur menschlichen Sprachverarbeitung und zu formalen Grundlagen der syntaktischen Struktur von Sprache und wird anhand sprachenübergreifender Untersuchungen überprüft. Einige der dafür relevanten Beobachtungen und Hypothesen haben bereits eine längere Geschichte:

Ich nenne ein zweites Gesetz von weitgreifender Bedeutung: *das Gesetz der wachsenden Glieder*, das ich aufgedeckt habe und dessen Gültigkeit ich zunächst für die Erweiterungsgruppen des Deutschen, Lateinischen und Griechischen aufgezeigt habe, d.h. für die Gruppen aus gleichberechtigten, durch *und*, *oder* verknüpften Gliedern. Das Gesetz lautet: von zwei Gliedern von verschiedenem Umfang steht das umfangreichere nach. Es ist das ein altüberliefertes Gesetz, aber auch hier wird in letzter Linie die Bedarfsstellung vorliegen: der Begriff des Umfangreichen wird sich weithin mit dem des Gewichtigeren decken, aber nachdem die zweckmäßige Stellung zur überlieferten geworden ist, entsteht daraus ein unbewußtes rhythmisches Gefühl, das gebietet, einem kürzeren Glied einen längeren Abklang folgen zu lassen. [...] Aber auch das Germanische und Deutsche beschränkt die Wirksamkeit des Gesetzes keineswegs auf die Erweiterungsgruppen. [...] Der Mannigfaltigkeit der überlieferten Stellungen steht die Gruppe der Bedarfsstellungen gegenüber mit minderer Zahl der Möglichkeiten, wenn auch durchaus nicht der Einzelfälle. (Behagel, 1930:86ff.)

Behaghels Beobachtungen zur Linearisierung von Wörtern und Satzteilen finden sich zum größten Teil in den Ausführungen von Hawkins wieder. Auch die von Behaghel unterschiedenen Termini „überlieferte Stellung“ und „Bedarfsstellung“ diskutiert Hawkins unter dem Stichwort „grammatikalisierte Wortstellung und performanzorientierte Umstellungen“. Behaghel liefert eine extern motivierte, funktionale Erklärung der von ihm beobachteten Linearisierungspräferenzen: Er macht das „rhythmische Gefühl“ der Sprecherinnen und Sprecher für Linearisierungsvorzüge verantwortlich, während Hawkins eine kognitionsbasierte Motivation vorstellt.

Hawkins' „Performance theory of order and constituency“ vereinigt den methodischen Ansatz aus der funktionalen Typologie mit jenem der modernen strukturalistischen Syntaxforschung, deren Begriffsbildung und Methodik Behaghel nicht zur Verfügung standen. Der Terminus „Gesetz“ wird bei Behaghel in einer Lesart verwendet, die als „häufig beobachtete Regularität“ paraphrasiert werden könnte, und deckt sich nicht mit den Sprachgesetzen, wie sie heute im Rahmen der quantitativen und synergetischen Linguistik erforscht werden.

2. Funktionale Typologie und Wortstellungsforschung

2.1 Funktionale Typologie

Die (funktionale) Typologie unternimmt die „universal characterization of typological variation“ (Croft, 1990:193). Typologie und Universalienforschung ergänzen sich: Die Analyse sprachenübergreifender Variation zielt auf die und führt zu der Ermittlung sprachlicher Invarianzen (Universalien). Diese – Variation restrinierenden – Aussagen stecken den Rahmen ab, der eine Antwort auf die Frage „What is a possible human language?“ geben kann:¹

In carrying out a typology of languages on some parameter, one establishes a certain number of logically possible types, and then assigns each language of the sample to one or other of these types. If all the logical possibilities have actual representatives, and there is no marked skewing of membership among the various types, then this result, though perhaps of typological interest, is not particularly interesting from the viewpoint of universals: it demonstrates that there are no restrictions on language variation with respect to the chosen parameter. Where, however, some of the logical possibilities are not represented, or are represented by a statistically significant low or high number of representatives, then the typological result does become of importance for the statement of language universals. (Comrie, 1981:31)

¹ Andere linguistische Schulen versuchen ebenfalls, innerhalb ihres Ansatzes diese Frage und die Frage nach Universalien zu beantworten, z.B. die formale Syntaxforschung. Zur Abgrenzung vgl. Comrie (1981), Croft (1990); interessante Überlegungen zum Stichwort „grammaire normative“ finden sich bereits in Frei (1929).

Die Typologie definiert sprachliche Typen („linguistic types“), die in sprachenübergreifenden Untersuchungen systematisch analysiert werden und zu einer typologischen Klassifikation führen, folgendermaßen:

Typological classification is the process of describing the various linguistic types found across languages for some grammatical parameter, such as grammatical number or the formation of relative clauses. (Croft, 1990:27)

Dabei wird ein grammatisches Phänomen definiert (z.B. grammatischer Numerus), welches in verschiedenen Sprachen durch unterschiedliche sprachliche Strukturen realisiert wird. Oft ist es notwendig, einen sprachlichen Typ nach weiteren strukturellen Merkmalen zu differenzieren. Sprachliche Typen oder Strategien, eine grammatische Kategorie auszudrücken, werden klassifiziert nach Sprachen, in denen sie vorkommen:

The final definition of a linguistic structural type is *a particular structural feature associated with a particular construction in a particular language*. (Croft, 1990:38)

Jedes strukturelle Merkmal repräsentiert eine unabhängige Variable, auf der sprachenübergreifende Generalisierungen basieren können.

2.1.1 Implikative Universalien

Unabhängige, strukturelle Merkmale (auch: grammatische Parameter) werden zu implikativen Universalien² zusammengefaßt, um der Frage nach den Variationsgrenzen für die Merkmalsausprägungen einen Schritt näher zu kommen. Diese Universalien werden als Implikationen der Propositionalen Logik formuliert, die zwei strukturelle Merkmale als Abhängigkeit zwischen zwei logisch unabhängigen Parametern beschreiben. Die resultierenden vier logischen Typen, von denen einer den Wahrheitswert „falsch“ besitzt, können Muster von sprachlicher Variation genau dann erfassen, wenn der logisch falsche Typ tatsächlich in den Sprachen der Welt nicht belegt ist.

Croft betont, daß es notwendig ist, ein tieferes Verständnis von dem Zusammenhang zu haben, durch den unabhängige Variablen implikativ in Bezug zueinander gesetzt werden (vgl. Croft, 1990:51).

² Vgl. Croft (1990), Kap. 3.

2.1.2 Markiertheit und Markiertheitskriterien

Ein wichtiges Konzept in der Typologie ist das der Markiertheit.³ Es geht zurück auf Arbeiten der Prager Schule, vgl. z.B. Trubetzkoy zur phonologischen Markiertheit und Jakobson für den Bereich der Morphosyntax. Dieses fundamentale Konzept besagt, daß sprachliche Größen wie bspw. Numerus *Asymmetrien* hinsichtlich ihrer Belegung aufweisen, welche aufgrund verschiedener Kriterien bestimmt werden können. Markiertheitsphänomene stellen in vielen Fällen die Grundlage für grammatische Hierarchien und implikative Universalien dar. Das Konzept der Markiertheit erlaubt es, Bezüge zwischen den Strukturen sprachlicher Einheiten sprachenübergreifend herzustellen und somit sowohl „basic structures“ zu identifizieren, als auch bisher als Irregularitäten eingestufte Beschreibungen als „manifestations of regular typological patterns“ (Croft, 1990:65) zu verstehen.

Dazu muß die ursprüngliche („klassische“) Fassung der Prager Schule revidiert werden. Sie ging von einer zweiwertigen Charakterisierung von Markiertheitsstrukturen aus. Bei binärer Belegungsmöglichkeit eines Attributs ist ein Wert markiert, der andere nicht markiert. So kann z.B. im Deutschen die stimmlose Variante von Plosiven als die unmarkierte identifiziert werden, die stimmhafte als die markierte, da erstere aufgrund der Auslautverhärtung in diversifizierteren Distributionen auftaucht und somit universeller, der „Default-Fall“, ist. Im Bereich der Morphosyntax wird oft das Beispiel der Numeruskategorien angeführt. Im allgemeinen wird die Singularkategorie mit weniger als oder gleichviel Morphemen ausgedrückt wie die Pluralkategorie und stellt somit als den prototypischen Fall dar.⁴

Oft haben grammatische Kategorien jedoch keine binäre Wertbelegung. Dennoch ist es möglich, markiertheitstheoretisch zu argumentieren. Viele Sprachen verfügen neben Singular und Plural über weitere Numeruskategorien wie Dual, Trial oder Paucal. Anhand der Markiertheitskriterien können Aussagen formuliert werden, die jeweils zwei Teile dieser Skala betreffen: der Singular ist unmarkiert in bezug auf die (markierten) Kategorieausprägungen Plural, Dual, Trial, Paucal. Der Plural ist unmarkiert in bezug auf den Dual, Trial, Paucal, etc. Die möglichen Wertbelegungen einer Kategorie können somit auf einer Skala oder Hierarchie angeordnet werden, wobei zwischen den einzelnen Einheiten der Präferenz-/grammatischen Hierarchie die Relation „ist unmarkiert in bezug auf“ besteht.

³ Vgl. Croft (1990), Kap. 4; Greenberg (1966).

⁴ Hier treten jedoch z.B. Phänomene der Markiertheitsumkehrung („markedness reversal“) auf, die an dieser Stelle nicht weiter ausgeführt werden, vgl. Croft (1990), Kap. 6.

Diese Relation wird üblicherweise durch „<“ in einer Markiertheitshierarchie symbolisiert:⁵

<i>Numerus (morphosynt.)</i> Singular < Plural < Dual < Trial
--

Grammatische bzw. Markiertheitshierarchien charakterisieren Muster sprachenübergreifender Variation und spezifizieren, welche Kookkurrenzen von Sprachstrukturen belegt sein sollten und welche nicht. Andererseits können Markiertheitsmuster auch als implikative Universalien paradigmatischer Auswahlmöglichkeiten ausgedrückt werden, obwohl sie dieselbe Kategorie betreffen. Entlang einer grammatischen Hierarchie können aufeinanderfolgende Merkmalsbelegungen als alternierende Implikans-Implikatum Beziehungen aufgefaßt werden, vgl. z.B. Greenbergs Universale 34 zur Kategorie Numerus:

No language has a trial number unless it has a dual. No language has a dual unless it has a plural. (Greenberg, 1963:94, 112)

Dieses Universale kann formuliert werden als „Dual < Plural“ und „Trial < Dual“, wobei das Implikatum der zweiten Proposition zum Implikans der ersteren wird und die grammatische Hierarchie zur Kategorie Numerus vom hinteren Ende, der „markiertesten“ Kategorie, beschrieben wird.

2.1.2.1 Markiertheitskriterien

Verschiedene Analyse Kriterien dienen der Ermittlung von Markiertheitsbeziehungen innerhalb von Kategorien. Im Idealfall wird eine Markiertheitsbeziehung durch mehrere Kriterien bestätigt, diese Maximalforderung kann jedoch im Einzelfall nicht immer erfüllt werden. Ausgehend von der Prager Schule und den Überlegungen Greenbergs resümiert Croft die folgende Markiertheitskriterien für grammatische Kategorien: Struktur, flexivisches Verhalten, distributionelles Verhalten, sprachenübergreifendes Verhalten, Textfrequenz, sprachenübergreifende Frequenz (vgl. Croft, 1990:73-85).

Von den aufgeführten Markiertheitskriterien sind besonders jene, die sich auf Frequenzen beziehen, für den in dieser Arbeit vorgestellten Ansatz von Bedeutung. Anhand dieser Markiertheitskriterien interpretiert Hawkins die Resultate der Frequenzuntersuchungen zu den aus dem Prinzip der „Early Immediate Consti-

⁵ Givón verwendet „>“ in seinen Veröffentlichungen, wobei die Reihenfolge unmarkiert vor markiert von links nach rechts bestehen bleibt.

tuentis“ gewonnenen Hypothesen zur Distribution von syntaktischen Konstruktionen in Texten und Grammatiken, vgl. Abschnitt 3.

An dieser Stelle wird deutlich, daß die Typologie zu einem bedeutenden Teil auf quantitativen Untersuchungen empirisch gewonnener Sprachdaten beruht, sowohl in der Form von textuellem Material, als auch in der Form von Strukturbeschreibungen verschiedener Sprachen, die im Grunde als typologische Häufigkeitsverteilungen grammatischer Strategien bewertet werden können. Das Konzept der Markiertheit arbeitet mit Markiertheitskriterien, die allesamt als relative quantitative Werte ausgedrückt werden, und setzt gerade voraus, daß die Wertbelegung der untersuchten sprachlichen Merkmale und Strukturen *nicht* gleichverteilt ist. Eine statistisch abgesicherte Diskussion der quantitativen Auswertung und Beurteilung der numerischen Ergebnisse bleibt in den meisten typologischen Untersuchungen jedoch aus. Dies stellt ein methodisches und methodologisches Manko dar, welches durch eine engere Verknüpfung der Typologie mit dem epistemologisch fundierten Analyse-Instrumentarium der Quantitativen Linguistik (QL) behoben werden könnte, vgl. auch Abschnitt 3.5.

2.2 Funktionale Linguistik

Der Zweig der Typologie, der sich als funktional begreift, begnügt sich nicht mit dem bloßen Datensammeln, dem Aufstellen von empirischen Generalisierungen und dem Angeben von Korrelationen zwischen grammatischen Parametern. In einem zweiten Schritt, nachdem durch empirische Vorarbeit grammatische Hierarchien und implikative Universalien aufgestellt wurden, muß gefragt werden, durch welche Anforderungen oder zugrundeliegenden Mechanismen gerade diese typologischen Variationsmuster zustande kommen. Das zuvor – durchaus angemessen für die Zwecke der detaillierten Forschung – abstrahierte und isolierte Forschungsobjekt „Sprache“ (*langage*) muß nun in seine „natürliche“, originäre Umgebung zurückversetzt werden. Sprachen werden von Menschen zu bestimmten Zwecken benutzt. Die Sprachbenutzer und -benutzerinnen formen Sprache durch den ständigen Gebrauch, wobei die Randbedingungen durch die soziale Umwelt, die kognitiven Möglichkeiten und die physische Lebenswelt der Menschen selbst gegeben sind. Somit lassen sich die sprachlichen Strukturen und ihre Interaktion als Antwort auf diese diversen Anforderungen und Randbedingungen verstehen.

Die funktionale Linguistik⁶ bezieht ihre Forschungsmetapher (ursprünglich) aus der Biologie und Soziologie. Organismen adaptieren ihre anatomischen Strukturen möglichst optimal gemäß den an sie gestellten Umweltanforderungen,

⁶ Eine klare, von einzelnen Syntaxansätzen unabhängige, beispielreiche Einführung bietet: Stockwell, R. (1977). *Foundations of syntactic theory*. Englewood Cliffs, NJ: Prentice-Hall.

verschiedene Spezies entwickeln unterschiedliche Strukturen zur Ausübung ähnlicher Funktionen. Anthropologen wie Malinowski⁷ machen z.B. anhand der Übersetzungsproblematik deutlich, daß die menschliche Sprachtätigkeit als *kontextuell* eingebundenes Verhalten untersucht werden muß. Dieser Ansatz wird vor allen Dingen von Firth aufgenommen⁸ und findet seinen aktuellen Niederschlag in der Systemisch-funktionalen Grammatik Hallidays.⁹ Bühlers Organonmodell und die funktionalen Arbeiten der Prager¹⁰ modellieren Sprache als System, welches den kommunikativen Anforderungen der Sprachproduktion und -rezeption genügen muß. Grundsätzlich zu unterscheiden sind unterschiedliche Definitionen des Funktionsbegriffs. Zum einen werden *systeminterne* Funktionen identifiziert, die die Rolle definieren, die ein Systemelement für ein Subsystem spielt, so z.B. die bedeutungsdifferenzierende Funktion von Phonemen. Zum anderen werden *externe* Funktionen beschrieben, die das System als strukturiertes Ganzes zu erfüllen hat, wie z.B. die von Jakobson angeführte „ästhetische Funktion“ von Sprache.¹¹

Deskriptiv ausgerichtete Grammatiken wie z.B. der Ansatz Hallidays oder der funktionale Ansatz Diks¹² betonen eher die systeminternen Funktionen innerhalb von Sprachen. Die Rechtfertigung ihres Modellierungsansatzes beruht jedoch letztlich auf einer Bedienung externer Bedürfnisse durch das sprachliche System, und sie verwenden die Ergebnisse der funktionalen Typologie, z.B. Hierarchien, in ihren Modellierungen.

⁷ Vgl. Malinowski, B. (1923). The problem of meaning in primitive languages. In Ch. K. Ogden & I.A. Richards (Hg.), *The meaning of meaning* (296-336), (= Supplement 1), 10. Aufl. London, 1969 (1923).

⁸ Vgl. Firth, J.R. (1935). The technique of semantics. Kapitel 3 aus: ders.: *Papers in linguistics* (S. 7-33), 1934-1951, London: Oxford University Press, 1964 (1957).

⁹ Vgl. Halliday, M.A.K., & Hasan, R. (1985). *Language, context, and text: aspects of language in a social-semiotic perspective*. Second edition. Oxford: University Press, 1989 (1985). (Language education) und Halliday, M.A.K. (1985). *An introduction to functional grammar*. London: Arnold.

¹⁰ Einen Überblick liefert z.B. Daneš, F. (1987). On Prague school functionalism in linguistics. In R. Dirven, u.a. (Hrsg.), *Functionalism in linguistics* (S. 3-38), Amsterdam u.a.: John Benjamins. (Linguistic and literary studies in Eastern Europe – LLSEE ; 20).

¹¹ Für einen Überblick über diese und weitere Differenzierungen des Funktionsbegriffs vgl. z.B. Nuyts, Jan (1992). *A cognitive-pragmatic approach to language*. Amsterdam u.a.: John Benjamins Publ..

¹² Für eine Zusammenfassung vgl.: Siewierska, A. (1991). *Functional grammar*. London: Routledge.

Funktionale Typologie und synergetische Linguistik¹³ versuchen *explanativ* die in beide Richtungen nicht eindeutige Beziehung zwischen Struktur und Funktion zu analysieren. Letztere verwendet dazu die wissenschaftstheoretische Basis der Funktionalanalyse, wie sie in der Biologie und in den Gesellschaftswissenschaften als explanativer Ansatz, der dem Naturwissenschaften an Erklärungskraft nicht nachstehen sollte, entwickelt wurde. Fundamental ist dazu die Modellierung des Forschungsobjekts als System mit einer bestimmten Organisation, welches mit der Umgebung interagiert, und die deduktive Erklärung mittels stochastischer Gesetze, unter Axiomatisierung des Prinzips der Selbstregulation/-organisation und Postulierung funktionaler Äquivalente.

2.2.1 Funktionalerklärung in der Typologie

The medium of expression is a linear string of sounds, and the primary constraint of human beings is their cognitive-processing abilities, from neuromuscular articulatory control and auditory perception to inferential abilities from past experience and current context. (Next in importance as a constraint on human beings are sociocultural constraints on interpersonal interaction.) Thus a proper functional-typological explanation accounts for a particular universal as the best means of encoding, transmitting and decoding information in a linear medium by human beings in real time. (Croft, 1990:253)

Die funktionale Typologie versucht in einem ersten Schritt, für die von ihr z.B. als grammatische Hierarchien formulierten typologischen Muster externe Motivationen zu finden (vgl. Croft, 1990, Kap. 7).

Als Paradebeispiel für die Motivierung einer typologischen Hierarchie durch externe Anforderungen/Gegebenheiten gilt die Hierarchie der „Basic colour terms“ (vgl. Comrie, 1981:34):

Hierarchie der „Basic colour terms“
weiß, schwarz < rot < grün, gelb < blau < braun

Traditionelle Untersuchungen von Farbsystemen haben sich darauf beschränkt, die unterschiedliche Anzahl der Farbtermini und ihre exakten Grenzen zueinander zu beschreiben. In einem typologischen Ansatz fragten Berlin und Kay¹⁴ jedoch nicht nach den (unscharfen) Grenzen zwischen den Farben, sondern nach den prototypischen Vertretern für eine Farbe, wie sie als Terminus in einer Sprache verwandt wird. Diese neue (typologische) Fragestellung ermöglichte eine andere Klassifikation der Ergebnisse und führte zu der oben dargestellten Hierarchie.

¹³ Für einen methodologischen Überblick und Literaturhinweise vgl. auch Hoffmann & Krott in diesem Band.

¹⁴ Vgl. Berlin, B., & Kay, P. (1969). *Basic color terms: Their universality and evolution*. Berkeley: University Press.

Des weiteren scheint es Hinweise zu geben, daß die Hierarchie der Farbfokussierungen mit der menschlichen Farbwahrnehmung korreliert werden kann und somit eine wahrnehmungspsychologische Erklärung bzw. Motivation eines sprachlichen Universales möglich ist.

Croft führt hauptsächlich zwei externe Motivationen für die von ihm beschriebenen typologischen Muster und Universalien an: Ökonomie und Ikonizität (vgl. Croft, 1990, Kap. 7). Basierend auf Zipfs „Principle of least effort“ können die meisten Markiertheitsmuster als ökonomisch motiviert verstanden werden. Der unmarkierte Fall ist in den meisten Fällen jener, der am häufigsten vorkommt, der Default-Fall, der prototypisch ist. Hier ist es am ökonomischsten, diesen z.B. durch Null-Morpheme zu kodieren. Croft bezweifelt jedoch, daß diese Motivation z.B. für die Wortstellungsuniversalien herangezogen werden kann, die eher anderen Faktoren unterliegen. Welche externen Motivationen hier zugrunde liegen könnten, wird in Kapitel 2.3 diskutiert. Diskursorientierte Faktoren wie die „cognitive salience“ faßt Croft ebenfalls unter die ökonomische Motivation: das, was für die menschliche Perzeption, das menschliche Überleben und somit die menschliche Kommunikation häufig am vorrangigsten ist, wird durch den unmarkierten Fall repräsentiert. Diese motivationelle Argumentation wird oft auch mit den Begriffen „foregrounding“ und „backgrounding“ geführt, wenn es um pragmatische Unterscheidungen zur Informationsstruktur geht.

Als weitere externe Motivation identifiziert Croft Ikonizität.¹⁵ Mit dem Begriff der Ikonizität verbindet sich – kurz gesagt – die Auffassung, daß die sprachliche Struktur die menschliche Erfahrungsstruktur reflektiert. Unterschieden werden können „Isomorphismus“ – die direkte Korrespondenz von Begriffen und sprachlichen Ausdrücken – und „Ikonische Motivation“ – die Korrespondenz von Relationen. Hier handelt es sich nicht darum, die simplizistische These aufzustellen, daß sich die Struktur der menschlichen Erfahrung in der sprachlichen Struktur wiederfinden lassen müsse. Im Gegenteil, diese Arbeitshypothese hat erst dazu geführt, typologische Ähnlichkeiten zwischen sprachlichen Strukturen zu untersuchen, deren Systematizität ansonsten unentdeckt geblieben wäre, vgl. z.B. Haimans sprachenübergreifende Ergebnisse zu der Ähnlichkeit zwischen Konditionalkonstruktionen und Entscheidungsfragen. Croft argumentiert überzeugend (vgl. Croft, 1990:171f.), daß es sich bei der ikonischen Motivation nicht um eine „self-fulfilling hypothesis“ handelt, fordert jedoch zur unabhängigen Stützung

¹⁵ Vgl. hierzu vor allem die Arbeiten von John Haiman, z.B.: Haiman, J. (1985). *Natural syntax*. Cambridge: University Press. Haiman, J. (1983). Iconic and economic motivation. *Language*, 59, 781-819. Aber auch: Givón, T. (1985). Iconicity, isomorphism and non-arbitrary coding in syntax. In J. Haiman (Hg.), *Iconicity in syntax*. Amsterdam u.a.: Benjamins. Givón, T. (1991). Markedness in grammar: distributional, communicative and cognitive correlates of syntactic structure. *Studies in language*, 15,2, 335-370.

experimental-psychologische Untersuchungen zur menschlichen kognitiven Verarbeitung.

Es muß betont werden, daß die beiden hier angeführten externen Motivationen für die Ausprägung sprachlicher Strukturen nicht unabhängig voneinander wirken. In einer Art „rudimentärer“ synergetischer Argumentation, von Croft „Competing motivations model“ genannt, werden sie als z.T. konkurrierende Bedürfnisse identifiziert, wobei die sprachenübergreifende Variation verschiedene mögliche Instanzen der Lösung des Konflikts zwischen den Bedürfnissen darstellt, je nachdem, welche Motivation sich in einem sprachlichen Typ stärker durchsetzt, wobei jedoch die Nichtberücksichtigung beider Motivationen als typologische Möglichkeit ausgeschlossen ist (vgl. Croft, 1990:192ff.).¹⁶

In seinen Ausführungen zur Erklärung von Universalien und typologischer Variation resümiert Croft (1990), Kap. 9, daß sich sowohl Ökonomie als auch Ikonizität letztlich auf Überlegungen zu den Anforderungen an das kognitive Sprachverarbeitungssystem des Menschen zurückführen lassen: „[...] economy and iconicity are basically different manifestations of processing efficiency.“ (S. 254).

2.3 Wortstellungsforschung¹⁷

Die typologische Wortstellungsforschung untersucht die Linearisierung unterschiedlicher Satz-/Äußerungsteile sprachenübergreifend.¹⁸ Dabei werden sprachliche Einheiten sowohl auf der Ebene der grammatischen Funktion wie S und O, der Ebene der Wortklassen wie N, A und P und neuerdings auch der Ebene der syntaktischen Konstituenten wie NP und PP verwendet. Problematisch gestaltet sich die einheitliche, sprachenübergreifende Zuordnung sprachlicher Einheiten zu den gewählten linguistischen Beschreibungskategorien. Ist das, was ich im Warlpiri als Subjekt identifiziere, wirklich einem Subjekt im Englischen ähnlich?¹⁹ Sind diese rein deskriptiven Probleme überwunden, geht es darum, die Frage nach der Motivation für die gefundenen Muster und ihre Verbreitung in den Sprachen zu beantworten.

¹⁶ Zur Erklärung in der Typologie vgl. auch die Beiträge in dem Sammelband: Hawkins, J.A. (1988). *Explaining language universals*. Oxford u.a.: Basil Blackwell.

¹⁷ Im folgenden werden nur die vornehmlich syntaktisch orientierten Forschungen erwähnt, der Diskurs-pragmatische Bereich bleibt zunächst ausgeklammert.

¹⁸ Für einen Überblick über die Ansätze zur Wortstellungsanalyse in der Typologie, der funktionalen Linguistik und der formalen Syntaxtheorie vgl. u.a.: Siewierska, A. (1988). *Word order rules*. London u.a.: Croom Helm.

¹⁹ Zur Diskussion vgl. u.a. Croft (1990), Comrie (1981) oder Hopper, Paul; Thompson, S.A. (1980). Transitivity in grammar and discourse. *Language*, 56, 251-299.

Ausgehend von den Wortstellungsuniversalien diskutiert Greenberg²⁰ generelle Prinzipien, die den gefundenen Mustern gemein sind. Er identifiziert *Harmonie* („harmony“) und *Dominanz* („dominance“) als generelle Mechanismen, die die Distribution von Wortstellungsmustern zu regulieren scheinen.

Croft illustriert, daß die beiden Prinzipien Dominanz und Harmonie gegenläufig wirken können und somit unter das typologische Erklärungsprinzip des „Competing motivations model“ fallen:

In Greenberg's word-order analysis, dominance favors some word orders, such as NA, absolutely, while harmony will favor an alignment of the adjective with other modifiers. Since for some modifiers, modifier-noun order is dominant, and for others, noun-modifier order is dominant, a language cannot be harmonic without having some recessive orders. However, it cannot be both recessive and disharmonic at the same time. (Croft, 1990:57)

Mir erscheint es schwierig, diese beiden Konzepte epistemologisch einzuordnen: sie sind zum einen empirische Generalisierungen, da sie die Korrelationen, die in implikativen Universalien formuliert sind, verallgemeinern, zum anderen indizieren sie bereits kognitive Anforderungen, denen Sprachen genügen müssen, um als Kommunikationsmittel effizient eingesetzt werden zu können, welche jedoch präzisiert werden müssen.

Typologen²¹ nach Greenberg haben versucht, gerade dieses Wechselspiel einander widerstrebender Prinzipien zu ignorieren und Sprachen typologisch entweder als VO- oder OV-Sprachen zu klassifizieren, je nachdem, welchem der beiden Harmoniemuster sie entsprechen, und dies durch eine allgemeine Tendenz innerhalb einer Sprache zur analogen und konsistenten Head-modifier-Linearisierung zu motivieren. Dies stellt jedoch eine grobe Verallgemeinerung dar, die eine große Anzahl von Sprachen nicht erfaßt.

Hawkins (1983) löst die krude Dichotomie der beiden Harmonieprinzipien auf, hält jedoch an dem Grundsatz der konsistenten Head-modifier-Linearisierung fest. Er führt das graduelle Prinzip der „Cross-categorical harmony“ (CCH) ein, welches besagt, daß es um so weniger attestierte Sprachen gibt, je stärker die konsistente Linearisierung durchbrochen wird. Dabei betont er, daß die Linearisierung VO bzw. OV ein schlechter typologischer Indikator ist und verwendet Adpositionen als typologische Indikatoren. Hier kann er weitere Generalisierungen über Kookkurrenzen vornehmen, die sich auf die Beziehung zwischen dem Typ der Adposition und die Art der Linearisierung der Substantivmodifikatoren wie Adjektiv, Demonstrativpronomen, Relativsatz, etc. beziehen: die „Prepositional“ bzw. „Postpositional Noun Modifier Hierarchie“. Diese beiden Hierarchien

²⁰ Vgl. Greenberg, 1961:96ff. und Croft, 1990: Kap. 3.3

²¹ So z.B. Vennemann, Lehmann, Keenan, für eine ausführliche Diskussion und Kritik vgl. z.B. Hawkins, 1983.

stellen Umformungen der bereits bekannten Wortstellungsuniversalien dar, wobei gerade diese Art der Formulierung die graduelle Instabilität der Operator-Operand-Serialisierungen erfassen soll, indem sie eine Hierarchie der „Abweichungen“ definiert. Diese kann Hawkins an einer Stichprobe von ca. 300 Sprachen dadurch bestätigen, daß es umso weniger belegte Sprachen gibt, je stärker diese abweichen. Um die Ausprägung dieser Hierarchien zu motivieren, führt er unter anderem das Prinzip der „Heaviness“ an.

Während das Prinzip der Heaviness, welches letztlich durch die kognitiven Mechanismen der Sprachverarbeitung erklärt werden könne, vgl. Hawkins (1983), S. 292, in Hawkins' Veröffentlichung von 1983 auf die Linearisierung innerhalb von Nominalphrasen beschränkt bleibt und vor allem ein Resultat des Versuchs darstellt, eine konsistentere typologische Klassifizierung der Sprachen nach Wortstellungskookkurrenzen zu motivieren, bildet es in seinem 1994 veröffentlichten Ansatz, wie er in Abschnitt 3 vorgestellt wird, ein wichtiges Teilstück in der Diskussion von Konstituenz und Linearisierung in Sprachen insgesamt.

3. Ein performanzbasierter Ansatz zu Konstituenz und Wortstellung

3.1 Motivation

Hawkins (1994), vgl. aber auch Hawkins (1990), Hawkins (1992), stellt einen umfassenden, auf durch Sprachdaten empirisch gestützten Ansatz vor, der Wortstellungsphänomene, die für die Sprachen der Welt beschrieben wurden, typologisch und sprachenübergreifend erklären soll. Dabei bedient er sich der in der funktionalen Typologie entwickelten Methodologie, die gekennzeichnet ist durch die Beschreibung von sprachenübergreifenden Daten. Ihre Ausprägung und Variation sollen durch universale sprachformende Prinzipien in letzter Instanz *erklärt* werden.

Mit dem CCH entwickelte Hawkins ein erstes Prinzip, die zu beobachtenden Wortstellungsphänomene durch ein syntaktisches Komplexitätsmaß zu verstehen und die „Head-modifier-Struktur“ als allgemeines kognitives Verarbeitungsprinzip zugrunde zu legen (vgl. Hawkins, 1983:297f.). Der vorliegende Ansatz stellt einen im Sinne der Erklärungsprinzipien und der empirischen Fundierung umfassenderen Versuch dar, syntaktische Strukturen und Linearisierungspräferenzen durch funktionale Anforderungen der menschlichen Informationsverarbeitung zu motivieren.

Bei der Sprachverarbeitung muß gefragt werden, wie es möglich ist, daß Sprachbenutzerinnen und -benutzer erfolgreich sprachliche Strukturen in Echtzeit produzieren und verstehen.

Hawkins postuliert, daß die Anforderungen der Informationsverarbeitung, die in der Performanz ersichtlich werden, sprachliche Strukturen so geprägt haben, daß sie eben diesen Anforderungen möglichst gut entsprechen können. Diese

Argumentation unterstellt, daß Menschen weltweit unter ähnlichen kognitiven Bedingungen Sprache verwenden. Um genau zu untersuchen, in welcher Weise Grammatiken den an sie gestellten Anforderungen genügen, müssen möglichst viele sprachenübergreifende Daten erhoben und analysiert werden. Ebenfalls muß geprüft werden, ob sie auf ein quantifizierbares Prinzip zurückgeführt werden können, welches es darüber hinaus ermöglicht, ihre (graduelle) Variation zu erfassen:

With regard to the question of whether performance has or has not influenced grammars, we can then juxtapose these metrics, and the performance evidence for them, with comparative data from grammars. The cross-linguistic generalizations that we should concentrate on are, correspondingly, not the absolute universals that are supposedly present in or absent from every grammar, but variation-defining universals, such as implicational regularities, hierarchies [...] and distributional universals determining language frequencies. I will argue that these grammatical data correlate very precisely with degrees of processing ease in performance, and provide evidence for the profound influence of performance on grammars, of a type that has not yet received the systematic attention it deserves. (Hawkins, 1994:13f.)

Bei dem vorgestellten Verarbeitungsprinzip handelt es sich um das Prinzip der „Early immediate constituents“ (EIC), welches – kurz ausgedrückt – besagt, daß es für die syntaktische Verarbeitung während des Sprachverstehens am vorteilhaftesten ist, wenn die Tochterkonstituenten einer Mutterkonstituente so schnell wie möglich erkannt werden. Voraussetzung für die Gültigkeit von EIC ist zunächst, daß die online-Verarbeitung einer linearisierten Äußerung von links nach rechts angenommen wird. Eine weitere – grundlegendere – Prämisse dieses Prinzips besteht darin, daß in Äußerungen syntaktische Strukturen kodiert sind – es sich nicht nur um lineare Nebeneinanderstellungen von lexikalischen Einheiten handelt –, deren Erkennen für die Interpretation von semantischen und pragmatischen Relationen notwendig ist. Bei der menschlichen Sprachverarbeitung wird davon ausgegangen, daß ein „Prozessor“²² eine eingehende Lautkette phonologisch und syntaktisch analysiert („parst“) und im weiteren lexikalische, semantische und pragmatische Interpretationen vornimmt. Der Gedächtnisaufwand ist umso geringer, je eher die syntaktische Struktur einer (Teil-)Äußerung vollkommen verarbeitet ist und nicht weiter im syntaktischen Prozessor präsent gehalten werden muß. Je länger es jedoch dauert, bis eine Teilstruktur vollständig erkannt wurde, desto höher ist der Verarbeitungsaufwand für die syntaktische Komponente. Diese muß nun gleichzeitig die noch abzuarbeitende Struktur offen halten,

²² Im folgenden wird „Prozessor“ als (laxe) Kurzform für „menschliche Sprachverarbeitungs-komponente“ benutzt, ohne daß eine allzu enge Analogie zur Computemetapher intendiert ist. Der „syntaktische Prozessor“ ist somit der syntaktische Teil der menschlichen Sprachverarbeitungs-komponente.

während weiterhin neue, ebenfalls zu verarbeitende Äußerungsteile eingehen. Dies erklärt z.B., warum komplexe und lange Strukturen oft nicht als Einbettungen zugelassen sind (vgl. Hawkins, 1994:23):

Bsp. 1. **Did*_S[*that John failed the exam*] *surprise Mary*?

Bsp. 2. *Did*_{NP}[*this fact*] *surprise Mary*?

Im Englischen ist die Einbettung von Nominalphrasen in Interrogativkonstruktionen grammatikalisiert, während die Einbettung von Nebensätzen nicht erlaubt ist. Die Grammatikalisierung der Einbettung von NPn im Gegensatz zu Nebensätzen läßt sich mit ihrer durchschnittlich geringeren Länge erklären, die weniger „Parallelverarbeitung“ im Sinne des obigen Prinzips erfordert. Dies bedeutet nicht, daß es in allen Sprachen so geregelt sein muß, denn die genaue Art der Einbettung, die zugelassen wird, ist abhängig von der sprachlichen Struktur insgesamt und anderen, EIC-sensitiven Elementen.

Ebenso wird der vielfach diskutierte „Heavy NP-Shift“ durch das EIC-Prinzip motiviert (vgl. Hawkins, 1994:57):

Bsp. 3 *I*_{VP}[*gave*_{NP}[*the valuable book that was extremely difficult to find*]_{PP}[*to Mary*]

Bsp. 4. *I*_{VP}[*gave*_{PP}[*to Mary*]_{NP}[*the valuable book that was extremely difficult to find*]]

Die komplette Struktur der VP ist in Bsp. 4 bereits mit Beginn der sehr langen NP bekannt und kann so von dem Prozessor abgeschlossen werden, während in Bsp. 3 die Struktur der VP solange offengehalten werden muß, bis nach der Verarbeitung der NP die letzte Tochterkonstituente der VP gelesen wird.

3.2 Operationalisierung

3.2.1 Syntaktische Komplexität

Um die Analogie zwischen syntaktischer Komplexität und Verarbeitungsaufwand verfolgen zu können, müssen die Definition des strukturellen Bereichs, dessen Komplexität gemessen wird, und die Art der Quantifizierung möglichst genau dem postulierten Verarbeitungsprozeß angepaßt werden. Hawkins verwirft bestehende Komplexitätsmaße wie das von Miller und Chomsky und jenes von Frazer²³ zugunsten eines verbesserten Komplexitätsmaßes. Dazu muß zunächst der Bereich der „Structural Domain“ definiert werden:

²³ Für Diskussion und Literatur vgl. Hawkins 1994:24ff.

Structural Domain (SD): A Structural Domain consists of a grammatically and/or psycholinguistically significant subset of structurally related nodes in a tree dominated by a constituent C. (Hawkins, 1994:25)

Hawkins betont, daß ein Komplexitätsmaß – sowie auch die in seinen weiteren Untersuchungen angenommenen Strukturen – weitestgehend unabhängig von der tatsächlichen Formulierung der Grammatik sei, fordert jedoch als „Minimalbedingungen“ grundsätzlich eher flache Strukturen in Phrasenstrukturbeschreibungen und setzt das aus der X-Bar-Theorie bekannte Prinzip der „Head-modifier-Beziehung“ voraus, wobei er letzteres jedoch modifiziert.

Daß die Definition des strukturellen Bereichs zur Bestimmung der Komplexität den Verarbeitungshypothesen gerecht werden muß, läßt sich an einem Beispiel illustrieren:

Die Tatsache, daß – im Sinne von EIC – Bsp. 3 syntaktisch schwieriger als Bsp. 4 zu verarbeiten ist, läßt sich nicht mit einem Komplexitätsmaß erfassen, das sich auf die Komplexität der *gesamten* Äußerung bezieht und *sämtliche* terminale Knoten umfaßt, z.B. als Proportion der Anzahl der nicht-terminalen Konstituenten zur Anzahl der terminalen Knoten. Die Komplexität der beiden Äußerungen wäre in diesem Fall identisch. Um ein der Verarbeitung angemessenes Maß zu finden, muß u.a. der Prämisse der „online-Verarbeitung von links nach rechts“ Rechnung getragen werden.

Eine weitere grundlegende Annahme für die syntaktische Sprachverarbeitung bezieht Hawkins aus der X-Bar-Theorie. Dieses von Chomsky, Jackendoff, Lightfoot u.a.²⁴ in unterschiedlichen Varianten vorgestellte Prinzip zur syntaktischen Konstituenz formuliert, daß die im Strukturbaum höchste phrasale Kategorie einer Phrase die maximale Projektion ihres lexikalischen Kopfes/Bezugselements („head“) ist. Aus der Perspektive der syntaktischen Sprachverarbeitung ist diese Forderung unmittelbar einsichtig: Sobald der lexikalische Kopf einer Phrase eindeutig erkannt wurde, ist die Identität einer gesamten syntaktischen Teilstruktur einer Äußerung bekannt. Sobald z.B. eine Präposition, im Englischen das Bezugselement der PP, vom Prozessor verarbeitet wurde, kann dieser eindeutig die Existenz einer PP in der eingehenden Äußerung postulieren und die folgende NP als Komplement des Bezugselements syntaktisch in die offene PP-Struktur eingliedern. Ist diese PP wiederum Tochterkonstituente einer VP, so trägt allein das eindeutige Erkennen des Bezugselements der PP dazu bei, die VP-Struktur zu vervollständigen, da nun wiederum ebenfalls eindeutig die offene VP-Struktur ergänzt werden kann. Im Grunde geht es im vorliegenden Zusammenhang also nicht so sehr um die „Head“-Eigenschaft von lexikalischen

²⁴ Für Diskussion und Literatur vgl. Hawkins (1983) und Hawkins (1994).

Elementen, die noch weitere syntaktische Eigenschaften wie die der Subkategorisierung umfaßt, sondern darum, inwieweit eine lexikalische/terminale Kategorie eindeutig die syntaktische Struktur, deren Konstituente sie ist, bestimmt. Hawkins definiert den Status von „Mother Node Constructing Categories“ (MNCCs):

Axiom of MNCC Existence: For each phrasal mother node M there will be at least one daughter category C that can construct M on each occasion of use.

Axiom of MNCC Uniqueness: Each MNCC will consistently construct a unique M on each occasion of use. (Hawkins, 1994:63)

Von Relevanz für die syntaktische Verarbeitung ist, daß es Kategorien gibt, die unmißverständlich die Identität der Mutterkonstituente anzeigen; wäre dies nicht der Fall, so käme es zu ständigen Ambiguitäten und damit oft zu notwendigen Revisionen bei der syntaktischen Verarbeitung:

Hence, reliable mother node construction is a prerequisite both for recognizing the abstract syntactic structure and for attaching non-uniquely determined ICs to higher phrases. (Hawkins, 1994:61)

Es können zwei Typen von MNCCs identifiziert werden: Einheiten aus geschlossenen Wortklassen und Funktionswörter sowie lexikalische Bezugselemente (N, A, V, P). Im Sinne eines deterministischen Parsingvorgangs, so wie er von Marcus²⁵ beschrieben wurde, nimmt Hawkins an, daß ein Mutterknoten konstruiert werden muß, wenn die ihn eindeutig bestimmende Kategorie verarbeitet wird, er aber nicht konstruiert wird, bis eine solche angetroffen wurde:

Mother Node Construction (MNC): In the left-to-right parsing of a sentence, if any word of syntactic category C uniquely determines a phrasal mother node M, in accordance with the PS rules of the grammar, then M is immediately constructed over C. (Hawkins, 1994:62)

Für Kategorien, die keine MNCCs sind, wird folgendes Verarbeitungsprinzip postuliert, welches Kimballs „Right association“ ähnelt (vgl. Kimball, 1973:23ff.):

Immediate Constituent Attachment (ICA): In the left-to-right parsing of a sentence, if an IC does not construct, but can be attached to, a given mother node M, in accordance with the PS rules of the grammar, then attach it, as rapidly as possible. Such ICs may be encountered *after* the category that constructs M, or *before* it, in which case they are placed in a look-ahead buffer. (Hawkins, 1994:62)

²⁵ Für Diskussion und Literatur vgl. z.B. Gibson, E.A. (1991). *A computational theory of human linguistic processing. Memory limitations and processing breakdown*. PhD., Carnegie Mellon Univ., Dept. of Phil., Pittsburgh, Penn., und Hawkins (1994).

Nach diesen Prämissen wird nun die „Constituent Recognition Domain“ (CRD) definiert:

Constituent Recognition Domain (CRD): The CRD for a phrasal mother M consists of the set of terminal and non-terminal nodes that must be parsed in order to recognize M and all ICs of M, proceeding from the terminal node in the parse string that constructs the first IC on the left, to the terminal node that constructs the last IC on the right, and including all intervening terminal nodes that they construct. (Hawkins, 1994:58f.)

Der Bereich zur Bestimmung der syntaktischen Komplexität ist somit genau so gewählt, daß er sich nur auf das sprachliche Material bezieht, welches zum Erkennen einer Konstituente notwendig ist, ausgehend von der Grundannahme, daß, sobald diese vollständig bekannt ist, der Prozessor entlastet werden kann:

[...] ICs of a phrase can be recognized on the basis of a proper subset of words and categories dominated by that phrase, and [...] the size of this proper subset can vary. Some orderings shorten the number of dominated elements that must be scanned for IC recognition, making this process faster and also more efficient, since the same amount of information is extracted from less input. (Hawkins, 1994:57f.)

Die Strukturbäume in Abb. 1 veranschaulichen das Konzept der CRD und seine Motivation anhand der Beispiele Bsp. 3 und Bsp. 4. Die CRDs sind durch parallele Linien angedeutet.

Es zeigt sich, daß die CRD für die VP abgeschlossen ist, sobald die letzte MNCC der letzten IC erkannt wurde. In der oberen Struktur wird deutlich, daß, exklusive der VP und ihrer ICs, 28 terminale und nicht-terminale Knoten gelesen werden müssen, um die Struktur der VP komplett zu analysieren. Nach der Umstellung reicht es aus, acht terminale und nicht-terminale Knoten zu lesen, um die VP komplett zu erkennen. Insgesamt verringert sich also die Anzahl der Knoten, die simultan verarbeitet werden müssen:

The different CRD scores [...] appear to correspond to differences in processing load, therefore, involving the size of working memory and the number of computations performed simultaneously on elements within this working memory. (Hawkins, 1994:60)

Diese kursorische Quantifizierung des Verarbeitungsaufwands wird im folgenden durch verschiedene differenzierte Maße verfeinert.

Doch zunächst soll noch eine weitere Frage angesprochen werden. Das vorgestellte Verfahren ist bestens geeignet, wenn es auf rechts-verzweigende Sprachen angewandt wird, deren MNCCs an der linken Peripherie liegen. Wie sieht es jedoch bei links-verzweigenden Sprachen aus, in deren Phrasenstrukturbeschrei-

bungen MNCCs typischerweise rechts zu finden sind? Hawkins illustriert diese Frage an einem japanischen Beispiel (vgl. Hawkins, 1994:66):²⁶

Bsp. 5: $S_1[NP[Mary-ga]] VP[S[S_2 kinoo John-ga kekkonsi-ta] to] it-ta]]$
Mary yesterday John married that said
Mary said that John got married yesterday

Bsp. 6: $S_2[S[S_1 Kinoo John-ga kekkonsi-ta] to] [NP Mary-ga] [VP it-ta]]$

Bei links-verzweigenden Sprachen wird als Parsing-Hypothese angenommen, daß momentan strukturell nicht einfügbare Teilstrukturen oder Einheiten in einem „look-ahead“-Puffer gespeichert werden, bis die strukturelle Verarbeitung eine Integration ermöglicht. Daß diese Tatsache den Prozessor nicht in dem Maße belastet wie das verzögerte Erkennen der kompletten Teilstruktur in rechts-verzweigenden Sprachen, ist einsichtig, da zu diesem Zeitpunkt, beim Lesen der ersten IC einer Phrase, noch gar keine übergeordnete Struktur vorhanden ist oder begonnen hat, die der Vervollständigung harren könnte.

Die lange Einbettung S2 in Bsp. 5 verzögert das vollständige Erkennen von S1 durch starke Vergrößerung der CRD²⁷. Die unvollständige Struktur des Matrixsatzes muß solange im Prozessor gehalten werden, bis die Einbettung vollständig abgearbeitet ist, was eine starke Erhöhung des Verarbeitungsaufwands bedeutet.

Bis der Subordinator („Complementizer“) *to* in Beispiel Bsp. 6 gelesen wird, wird der Nebensatz zunächst als S konstruiert. Erst der Subordinator dient als MNCC für S', so daß erst jetzt die Mutterkonstituente S2 konstruiert wird, in die der Nebensatz und die folgende NP und VP sofort integriert werden können. Die Umstellung des Nebensatzes verkürzt also die CRD drastisch, da nun die für die komplette Erkennung von S2 notwendigen Knoten *to*, *Mary-ga*, *it-ta* innerhalb eines sehr kleinen Fensters erfaßt werden können.

Die so definierte CRD erlaubt es, einen psycholinguistisch relevanten strukturellen Bereich aller Knoten im Baum zu definieren und eine Aussage darüber zu machen, daß unterschiedliche Wortstellungen in diesem Bereich in verschiedenen großen SDs resultieren. Sie stellen die Basis dar, die letztlich auf weniger sprachlichem Material beruht, auf der Wortstellungsvarianten als weniger komplex und effizienter und damit präferiert beurteilt werden können. Aus kleineren SDs resultieren

²⁶ Eine leichte Verwirrung entsteht in diesem Beispiel dadurch, daß der eingebettete Nebensatz (ohne Comp) in Bsp. 5. als S2 annotiert ist, in Bsp. 4 jedoch als S1.

²⁷ Die Satzteile im Bereich der CRD sind in den Beispielen durch Unterstreichung angedeutet.

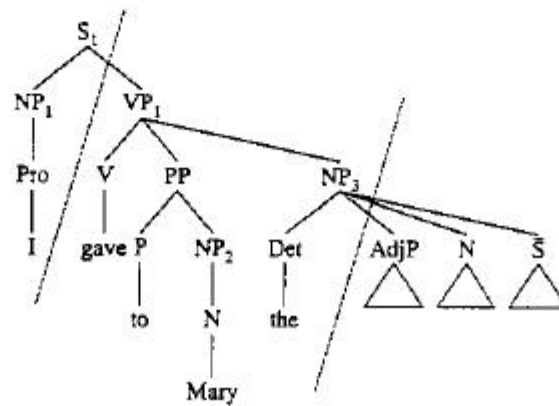
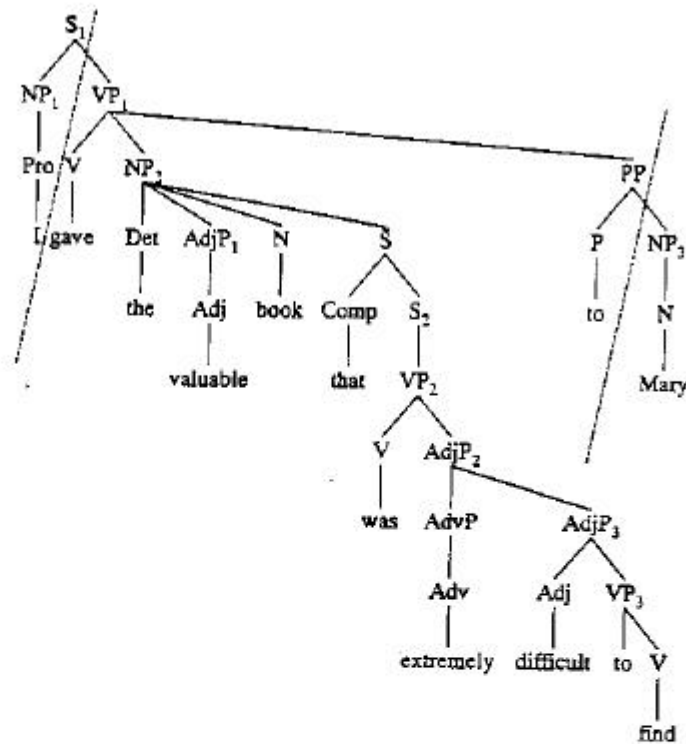


Abb. 1: Illustration der CRD nach Hawkins (1994:59)

- ◆ weniger Elemente im Arbeitsspeicher,
- ◆ eine reduzierte Anzahl an phrasalen Knoten, deren Struktur simultan erkannt werden muß,
- ◆ weniger hohe Verarbeitungsanforderungen an den Prozessor.

3.2.2 EIC-Komplexitätsmaße

Aus den oben vorgestellten Überlegungen zur Bestimmung der Komplexität von syntaktischen Strukturen bei der Sprachverarbeitung wird ersichtlich, daß es sich bei der Quantifizierung von Präferenzen bei der Wortstellungsvariation um ein relatives Maß handeln muss, welches die Komplexität von syntaktischen Strukturen gegeneinander abwägt. Des weiteren muß betont werden, daß die syntaktische Komplexität nur in dem Bereich bestimmt werden sollte, der auch für die aktuell interessierende Wortstellungsvariation relevant ist. So haben wir an den Beispielen Bsp. 3 und Bsp. 4 gesehen, daß es sich um einen Bereich innerhalb der VP handelte, während bei Bsp. 5 und Bsp. 6 der gesamte Satz diskutiert wurde.

Für Hawkins scheint die interessanteste Proportion zur Bestimmung der Komplexität in der Proportion zwischen ICs und Nicht-ICs zu liegen: „how many non-ICs need to be processed in order to recognize the ICs (at least one of which will construct the mother)?“ (Hawkins, 1994:69f.)

In Abb. 1 ergibt diese Quantifizierung für den oberen Strukturbaum $3/28$, also 10,7%, während die durch Umstellung verbesserte SD im unteren Strukturbaum $3/8$, also 37,5% erzielt, wenn die IC - Nicht-IC-Proportion berechnet wird. Diese Quantifizierung wird wie folgt definiert:

Calculating IC-to-non-IC ratios: The IC-to-non-IC ratio for a CRD is calculated by dividing the number of ICs in the domain by the total number of non-ICs (or words alone) in that domain, expressing the result as a percentage. The ratio for a whole sentence is the aggregate of the scores for all CRDs within the sentence. (Hawkins, 1994:76f.)

Hawkins betont aber auch, daß es in manchen Fällen genügen kann, die Anzahl der ICs der untersuchten Phrase zur Anzahl der terminalen Knoten, die benötigt werden, um diese insgesamt zu erkennen, in Bezug zu setzen. Im folgenden operiert Hawkins zumeist mit dieser IC-zu-Wort-Proportion, die für die vorliegenden Beispiele $3/11 = 27\%$ bzw. $3/4 = 75\%$ beträgt.²⁸

Aus einer Menge an Äußerungen mit unterschiedlichen möglichen Wortstellungen wie z.B. Bsp. 3 und Bsp. 4 kann mit dem EIC-Maß die präferierte bestimmt werden. Im Falle des Heavy NP-Shift kann also gesagt werden, daß eine Umstellung in jedem Fall von Vorteil ist und meistens präferiert wird. Diese Tatsache sollte sich in empirischen Untersuchungen bestätigen. Hier handelt es sich um einen stochastischen Zusammenhang und keineswegs um einen deterministischen Mechanismus. Somit sind auch Äußerungen attestiert, in denen keine Um-

²⁸ Diese Meßwerte für die IC-zu-Wort-Proportion nenne ich im folgenden EIC-Werte, während die IC-zu-Wort-Proportion EIC-Maß heißt.

stellung vollzogen wurde, obwohl dies im Sinne von EIC von Vorteil gewesen wäre. Es sollte sich gleichzeitig bestätigen, daß innerhalb der nicht-präferierten Varianten, die also nicht von einer Umstellung betroffen sind, jene häufiger sind, bei denen die Umstellung nur eine geringe Verbesserung des EIC-Wertes brächte. Hier müßte sich also eine graduelle Dispräferenz finden lassen: je höher der Gewinn durch eine Umstellung, desto infrequenter die Fälle, in denen keine Umstellung stattfindet. Und wenn keine Umstellung in Fällen mit sehr ungünstigen EIC-Werten erfolgt, so sollte es auf jeden Fall auch Instanzen geben, in denen keine Umstellung bei wenig ungünstigen EIC-Werten erfolgt.

Aufgrund dieser Überlegungen differenziert Hawkins das EIC-Prinzip anhand des wissenschaftlichen Analyseinstrumentariums der funktionalen Typologie, die sowohl mit Markiertheitskriterien als auch mit grammatischen Hierarchien arbeitet, vgl. Abschnitt 2.1.2. Im unmarkierten Fall wird eine Wortstellungsvariante gewählt, die einen möglichst günstigen Wert hat. Im markierten Fall lassen sich die einzelnen Instanzen der Nicht-Umstellung so anordnen, daß sie eine Präferenz- bzw. Zugänglichkeitshierarchie bilden. Nun kann das Prinzip der „Early Immediate Constituents“ endgültig begründet werden:

EIC (Expanded): The human parser prefers linear orders that maximize the IC-to-non-IC ratios of constituent recognition domains. Orders with the most optimal ratios will be preferred over their non-optimal counterparts in the unmarked case; orders with non-optimal ratios will be more or equally preferred in direct proportion to the magnitude of their ratios. For finer discriminations, IC-to-non-IC ratios can be measured left-to-right. (Hawkins, 1994:78f.)

Dieses, die lineare Abfolge betreffende Maß, der Aggregatwert, wird berechnet, indem kumulativ die EIC-Werte für von links nach rechts aufeinanderfolgende ICs innerhalb einer CRD bestimmt werden und anschließend ihr Durchschnitt gebildet wird. Dieses Maß ist interessant für solche Strukturen, in denen nicht nur Präferenzen in einem strukturellen Teilbereich einer Äußerung bestimmt werden sollen, sondern in denen verschiedene Varianten, die denselben EIC-Wert erzielen, beurteilt werden sollen.

3.3 Hypothesen nach EIC und Grammatikalisierung

Aus dem EIC-Prinzip können Hypothesen zu verschiedenen Bereichen innerhalb der Wortstellungsforschung abgeleitet werden:

1. Optionale Umstellungen in der sprachlichen Verwendung, wie z.B. die Linearisierung unterschiedlich langer PPn.
2. Wortstellungen, die durch grammatische Regeln einer Sprache konventionalisiert sind. Dabei ist anzunehmen, daß diese Konventionalisierung ein diachroner Reflex des EIC-Prinzips ist:

„In a free-word-order language EIC makes the unmarked case prediction that the optimal orderings will be most frequent in performance. If such a language changes into a fixed-order-language, it makes sense that the ordering that is fixed by grammatical convention will be the one that occurred in the unmarked case anyway, prior to the acquisition of the ordering convention.“ (Hawkins, 1994:95)

3. Grammatikalisierte und damit oft obligatorische Umstellungsregeln, die dazu dienen, gemäß EIC nicht optimale durch grammatische Regeln entstandene Varianten zu verbessern.

Heavy NP-Shift ist danach ein (optionales) Performanzprinzip, das auf dem grammatisch schwer faßbaren Konzept „Heaviness“ arbeitet und bereits nach Durchschnittswerten als optimal grammatikalisierte Strukturen „nachbessert“, falls dies in einer aktuellen performativen Verwendung notwendig sein sollte. Dies wird deutlich an den Beispielen Bsp. 3 und Bsp. 4. Hawkins argumentiert (vgl. Hawkins, 1994:89ff., 182ff.), daß in der englischen VP die Abfolge [V NP PP] grammatikalisiert ist, da somit EIC in den meisten Fällen genügt wird. Empirisch gewonnene Durchschnittswerte von NP- und PP-Länge zeigen, daß englische PPn durchschnittlich länger sind als NPn. In 85% der untersuchten Fälle entsprach die Wortstellung [V NP PP] EIC, eine Umstellung war ausgeschlossen, da sie den EIC-Wert nur noch hätte verschlechtern können. Diese Tatsache ist auch plausibel, da die PP immer eine NP und die Präposition enthält. Wird in einer konkreten Realisierung die NP jedoch ausnahmsweise einmal länger als die PP, so können Umstellungen vorgenommen werden. Interessanterweise läßt sich hier eine „Latenz“ feststellen: die Umstellung erfolgt mit desto höherer Wahrscheinlichkeit, je größer der EIC-Wert-Gewinn durch sie wird, während eine große Zahl der geringfügigen Verbesserungen nicht durchgeführt wird. Hawkins führt dieses Phänomen auf eine „Trägheit“ zurück, die aus der Grammatikalisierung resultiert, anders ausgedrückt: es existiert eine *produktive* grammatische Regel, die sich gegen die EIC-Umstellungsvoraussagen in vielen Fällen durchsetzt. Diese Art der „Nachbesserung“ in der Performanz findet sich auch bei anderen Extrapositionsvorgängen, wie z.B. der sententiellen Subjektextraposition sowie der Particle-Bewegung.

3.4 Empirische Überprüfung der EIC-Hypothesen

Hawkins und seine Arbeitsgruppe haben in umfangreichen Untersuchungen empirisch überprüft, wie die verschiedenen, durch EIC motivierten Hypothesen, vgl. Abschnitt 3.3, zu Wortstellungsphänomenen in der Performanz, in der Grammatik und typologisch ihre Voraussagen erfüllen.

Die empirische Überprüfung zu Performanz und Grammatik wurde an Material von neun typologisch sehr unterschiedlichen Sprachen durchgeführt. Gemäß der verschiedenen Hypothesen, die aus EIC abgeleitet werden können, umfaßte

die Untersuchung freie Wortstellung in mehrfach verzweigten Strukturen, binär verzweigende Strukturen und Umstellungen in der Performanz. Exemplarisch wird im folgenden eine Studie zum Einfluß von EIC im Bereich der freien Wortstellung vorgestellt.

Im Ungarischen kann davon ausgegangen werden, daß das Verb eine Phrase eröffnet, in der NP und PP beliebig linearisiert werden können. Im Ungarischen ist die Verbalphrase rechts-verzweigend (das Verb als MNCC findet sich an der linken Peripherie), die PP ist links-verzweigend, enthält also eine Postposition, während es NPn beider Arten gibt. In der vorliegenden Untersuchung wurden jedoch nur rechts-verzweigende NPn mit der MNCC am linken Rand berücksichtigt. Der untersuchte strukturelle Bereich hat somit entweder die Form [V PP_m _mNP] oder die Form [V _mNP PP_m]. Von den 55 Instanzen dieser Strukturen in den Daten haben 17 die Linearisierung [V NP PP] und 34 die Linearisierung [V PP NP]. Werden zunächst die Fälle betrachtet, in denen die NP gleich lang oder länger als die relativ kurze PP ist, so wird nach EIC die Linearisierung PP NP klar präferiert, da die NP an der linken Peripherie konstruiert wird und somit der lange Teil der NP für die CRD der VP unberücksichtigt bleiben kann. Je länger die NP, desto größer, gemessen am EIC-Wertdifferential, wird die Präferenz für die Linearisierung [V PP NP], vgl. Abb. 2.

*IC-to-word ratios for Hungarian [V {PP_m _mNP}] where
_mNP ≥ PP_m*

Assume V = 1 word.

I	_m NP = 2 PP _m = 2	[V _m NP PP _m]	[V PP _m _m NP]	CRD: 3/5 = 60%	CRD: 3/4 = 75%	Differential = 15%
II	_m NP = 3 PP _m = 2	[V _m NP PP _m]	[V PP _m _m NP]	CRD: 3/6 = 50%	CRD: 3/4 = 75%	Differential = 25%
III	_m NP = 4 PP _m = 2	[V _m NP PP _m]	[V PP _m _m NP]	CRD: 3/7 = 43%	CRD: 3/4 = 75%	Differential = 32%

Abb. 2: EIC-Wert-Differenziale, illustriert am Ungarischen, nach Hawkins, 1994:165

Wenn die PP jedoch länger ist als die NP, erweisen sich die Vorhersagen als etwas komplexer. Je größer die PP, bei kurzer NP, desto geringer wird der Vor-

teil, sie nach vorne zu stellen, um ihre MNCC an der rechten Peripherie und die NP an der linken Peripherie möglichst schnell zu erkennen. Hier scheint sich wiederum die generelle Links-rechts-Asymmetrie durchzusetzen, kürzere ICs, also die NP, nach vorne zu stellen und die lange PP nach hinten, auch wenn diese somit erst mit dem letzten Wort der PP, und damit auch der VP, erkannt werden kann, wodurch die CRD ihre maximale Komplexität für dieses Beispiel erreicht, aber auch der Aggregatwert für die CRD verbessert wird. Nimmt jedoch die Größe der NP ebenfalls zu, so hebt sich dieser Vorteil auf, da nun das Prinzip der Links-rechts-Asymmetrie, welches durch den Aggregatwert erfaßt wird, nicht mehr greifen kann, da es sich bei der NP nicht mehr um eine kurze/leichte IC handelt. Sind beide ICs relativ schwer (NP = vier Wörter und PP = 5 bis 8 Wörter), oder relativ leicht, so läßt sich kaum eine Präferenz für die eine oder andere Linearisierung angeben. Zusammenfassend heißt dies: ist die NP gleich lang oder länger als eine relativ kurze PP, so ist bei steigender NP-Länge eine starke Präferenz für die Linearisierung [V PP NP] gegeben. Diese Tendenz läßt sich aus der Abb. 3 ablesen. Die Tabelle gibt eine Übersicht über alle VP-Instanzen, die die ICs V, NP und PP enthalten. Zwei Reihen (X und Y) geben die Art der Linearisierung an, während die Spalten die Länge von PP und NP relativ zueinander anzeigen. Die Zellen der Matrix zeigen die absoluten Frequenzen der einzelnen Konfigurationen. Unter der Tabelle finden sich unter I. die Auswertungen für den ersten unmarkierten Fall, dessen Werte im rechten Teil der Tabelle aufgeführt sind. In 90,9% der Fälle, in denen die NP gleich lang oder länger als die PP ist, findet sich die Linearisierung [PP NP]. Für den markierten Fall muß gelten, daß, wenn die NP sehr viel länger als die PP ist, es auch markierte Fälle geben muß, in denen relativ häufiger Umstellungen erfolgen, wenn die NP nur wenig länger ist als die PP. Dieses Markiertheitskriterium ist insofern erfüllt, als es nur markierte Fälle gibt, in denen beide Phrasen gleich lang sind, also keine markierten Fälle mit großen Längendifferenzen auftreten.

Ist die PP länger als die relativ kurze NP, so ist bei steigender PP-Länge eine Präferenz für die Linearisierung [V NP PP] gegeben. Dieser Zusammenhang ist im linken Teil der Tabelle ersichtlich: Je größer die Längendifferenz bei insgesamt längerer PP, desto höher ist die relative Häufigkeit der Linearisierung [NP PP]. Hawkins setzt hier eine empirisch ermittelte Grenze für den unmarkierten Fall: Ist die PP um zwei oder weniger Wörter länger, so sind beide Linearisierungen möglich. Ist die PP jedoch sehr viel länger, so handelt es sich um den markierten Fall, und es sollte nur die Linearisierung [NP PP] zu finden sein.

Ist der Längenunterschied zwischen NP und PP, bei größerer PP, jedoch sehr gering, oder aber sind beide Konstituenten jeweils sehr lang, bei größerer PP, so läßt sich keine Präferenz für eine der beiden Varianten angeben.

Bei der empirischen Überprüfung erweist sich EIC als Formulierung kognitiv-funktionaler Anforderungen an sprachliche Strukturen in allen drei Hypothesenbereichen als durchweg plausibel.

Hungarian {V {PP_m mNP}}

112 pages of data; cf. table 4.4.

s[. . . V {PP_m mNP} . . .] (the only other ICs permitted in the clause are AdvP and/or NP-Nominative in the positions marked “ . . . ”)

where mNP = any NP constructed on its left periphery, i.e. with no left-branching genitive NP or relative \bar{S} ;
 PP_m = PP[NP P]

n = 55	PP _m > mNP :3+	:2	:1	mNP = PP _m	mNP > PP _m :1	:2
X [PP _m mNP]	0	2	6	10	17	3
Y [mNP PP _m]	2	5	7	3	0	0
Ratio of Y/X:	100%	71%	54%	23%	0%	0%

I *EIC predictions*: mNP ≥ PP_m
 Unmarked case: 30/33 = most optimal (X), i.e. 90.9% Correct
 Marked case: ratio of Y/X for mNP = PP_m ≥ (mNP > PP_m :1) ≥ 2 All correct

II *EIC predictions*: PP_m > mNP
 PP_m > mNP :3+ words: Y only Correct
 PP_m > mNP :2 words: Y ≥ X Correct
 PP_m > mNP :1 word: either X or Y Correct

Unmarked case: at least 20/22 = most optimal, i.e. 90.9% Correct
 Marked case: ratio of X/Y for PP_m > mNP :2 (= 29%) ≥ 3+ (= 0%) Correct

Abb. 3: Ergebnis der Untersuchungen am Ungarischen nach Hawkins, 1994:168

3.5 Kritik an der Methodologie aus Sicht der QL

Grundsätzlich problematisch hinsichtlich der EIC-Hypothesen ist die Art der Quantifizierung bzw. die Bewertung der durch sie entstandenen Meßergebnisse. Zwar ist es zunächst nachvollziehbar, daß z.B. die Ergebnisse der Untersuchungen zur Performanz als „sehr gut“ bewertet werden können: in den meisten Fällen werden Werte um 90% festgestellt, in denen sich die (unmarkierten) Wortstellungsvarianten optimal gemäß EIC darstellen. Auch bei den verbleibenden 10% der markierten Varianten lassen sich die postulierten Präferenzbeziehungen in den meisten Fällen nachweisen. Grundsätzlich stellt sich jedoch die Frage, welches Ergebnis als EIC-gegenläufig bewertet werden könnte. Hier ist sicher ein erster Scheidepunkt die 0%-Marke: verhielten sich die untersuchten Varianten genau entgegengesetzt zu EIC, so müßte dieses ohne Zweifel verworfen oder ein Spiegel-Prinzip aufgestellt werden. Verhielten sich die Daten zu 50% nach EIC, so wäre das Prinzip ebenfalls in Frage zu stellen, denn dann könnte von einer zufälligen Verteilung ausgegangen werden. Die verbleibenden Werte zwischen 51%- und 100%-Verteilung im Sinne von EIC stellen jedoch eine große theoretische

Grauzone dar, die von Hawkins nicht zufriedenstellend elaboriert wird. Zwar bewertet er Ergebnisse, die nur zu 75% EIC entsprechen, schon als problematisch und versucht dementsprechend, Hypothesen zur Interaktion mit anderen Prinzipien heranzuziehen, doch verbleibt das Unbehagen über den Mangel eines methodischen Instruments zur eindeutigen Bewertung der Ergebnisse. Nach dem in der Typologie definierten Markiertheitskriterium der Frequenz, Croft, 1990:85:

Frequency (textual): if a marked value occurs a certain number of times in frequency in a given text sample, then the unmarked value will occur at least as many times in a comparable text sample.

reicht es, daß der unmarkierte Fall mindestens genauso häufig auftritt wie der markierte Fall. Diese Definition läuft jedoch einer intuitiven Einschätzung, daß eine 1:1-Verteilung zufällig sei, zuwider. Der funktionale Zusammenhang sollte mathematisch als Verteilung oder Funktion mit entsprechenden Parametern definiert werden, damit signifikante Schlüsse gezogen werden können.

In der Quantitativen bzw. Synergetischen Linguistik,²⁹ die sich mit Sprachgesetzen und ihrer empirischen Fundierung/Überprüfung befaßt, ist eben das methodische Instrumentarium vorhanden, welches es erlaubt, eindeutige Aussagen zu empirischen Ergebnissen vor dem Hintergrund eines quantifizierten Zusammenhangs zu machen. Dies ist jedoch nur möglich, wenn dieser Zusammenhang auch theoretisch aufgearbeitet ist und z.B. als Verteilung oder Funktion angegeben werden kann. Dazu müssen die Variablen und Parameter, die vor allen Dingen sprachlich relevante und identifizierbare Größen darstellen sollten, bekannt sein. Liegt ein solcher theoretischer Zusammenhang vor, können die gewonnenen empirischen Daten mit Hilfe der in der Statistik entwickelten Testtheorie im Hinblick auf ihre Konformität zur theoretischen Annahme beurteilt werden.

Wie könnte nun eine theoretische Formulierung des vorliegenden Zusammenhangs aussehen? Zunächst haben wir zwei Maße: den EIC-Wert und den Aggregatwert. Die relative Frequenz (RFRQ) von Wortstellungsvarianten einer bestimmten syntaktischen Struktur in einem Text oder typologisch, die als durch EIC motiviert angenommen werden können, ist proportional zum EIC-Wert-Differential bzw. Aggregatwert-Differential:

$$\text{RFRQ} \sim \text{EIC-Wert-Differential}$$

$$\text{RFRQ} \sim \text{Aggregatwert-Differential}$$

Je höher der Gewinn gemäß EIC (sei es nun EIC-Wert oder Aggregatwert), der durch eine Variante der Linearisierung von mehreren möglichen erzielt werden kann, desto größer ist die Wahrscheinlichkeit (hier abgeleitet aus der relativen

²⁹ Vgl. Köhler (1986) als grundlegendes Werk und illustratives Beispiel.

Häufigkeit), daß diese vorteilhafte Linearisierungsvariante in Performanzdaten (aber auch in Form grammatischer Regeln) anzutreffen ist.

Soll dieser Zusammenhang nun funktional ausgedrückt werden, so muß ein Parameter hinzugenommen werden, der die Beziehung zwischen den Größen angemessen abbildet. Die empirischen Daten lassen zunächst einen linearen Zusammenhang vermuten, obwohl betont werden muß, daß zumeist nicht-lineare Zusammenhänge am besten geeignet sind, um sprachliche Modelle aufzustellen.

$$\text{RFRQ} \sim \text{EIC-Wert-Differential} * b$$

$$\text{RFRQ} \sim \text{Aggregatwert-Differential} * a$$

Es ist momentan jedoch schwer, über die sprachliche Realität dieses möglichen Parameters eine Aussage zu machen, was jedoch unbedingt notwendig wäre. Möglicherweise variieren die Parameter, je nachdem, welche syntaktische Konstruktion oder welche Art der Umstellung (grammatisch oder Verwendungsorientiert) vorliegt.

Ein weiterer wichtiger Schritt in Richtung Theoriebildung besteht darin, die so quantifizierten Hypothesen in Beziehung zu anderen Hypothesen/Gesetzen in der Linguistik zu setzen. Naheliegend ist ein Zusammenhang zu dem in der Sprachwissenschaft auf verschiedenen Beschreibungsebenen bestätigten (besser: bislang nicht verworfenem) Menzerathschen Gesetz, welches besagt, daß die Länge der Komponenten eines Konstrukts eine inverse Funktion der Länge des aus ihnen gebildeten sprachlichen Konstrukts ist. Obwohl über einen solchen Zusammenhang zwischen Menzerathschem Gesetz und EIC momentan nur spekuliert werden kann, ist grundsätzlich zumindest eine Interaktion oder, genauer gesagt, eine Komplementation anzunehmen. Die sprachlichen Variablen, die für die Bestimmung der EIC-Maße (EIC-Wert und Aggregatwert) eine Rolle spielen, sind solche, auf die auch auf das Menzerathsche Gesetz angewandt werden könnte, das eine Funktion zwischen Eigenschaften von Einheiten direkt benachbarter linguistischer Beschreibungsebenen angibt. Dazu müßte es sich als plausibel erweisen, eine Phrase als Konstrukt und ihre einzelnen ICs als Komponenten im Sinne des Menzerathschen Gesetzes zu postulieren. Ergebnisse in diese Richtung liefert u.a. Köhler (1982), der das Menzerathsche Gesetz auf seine Gültigkeit auf Satz-Teilsatzebene erfolgreich testete. Während bei dieser Metrisierung die Linearisierung der verschieden langen ICs aufgehoben wird (zur Bestimmung der Abhängigkeit wird der durchschnittliche Mittelwert der Komponentenlängen des Konstrukts gebildet), macht EIC gerade für den grammatischen Bereich Aussagen, den das Menzerathsche Gesetz in seiner vorliegenden Formulierung ausklammert. Ein weiteres Problem bei dem Vergleich von EIC und Menzerathschem Gesetz besteht darin, daß EIC Voraussagen bei einer konstanten Komponentenanzahl liefert, während das Menzerathsche Gesetz gerade die Konstanz in der Variation der Komponentenanzahl formuliert.

Von Interesse ist auch Köhlers Versuch, vgl. Köhler (1984), das Menzethsche Gesetz anhand eines (psycholinguistischen) Sprachverarbeitungsmodells zu motivieren. Seine Grundannahmen ähneln denen von Hawkins:

1. Die menschliche Sprachverarbeitung ist – im Sinne der Linearität von Äußerungen – ein sequentieller Prozeß.
2. Ein Register mit fester endlicher Größe steht als Arbeitsspeicher zur Verfügung.
3. Dieses Register hat zwei Aufgaben: 1. Präsenhaltung der zu bearbeitenden Größe und 2. Speicherung der erforderlichen Strukturinformation über die Verknüpfung der Komponenten untereinander.
4. Das beschriebene hierarchische Organisationsprinzip der linguistischen Beschreibungsebenen betreffend hat zur Folge, daß die Strukturinformation jeweils zusammen mit der zugehörigen Ausdruckskette zu speichern ist. (Vgl. Köhler, 1984:171f.)

Der wichtige Unterschied zu Hawkins' Annahmen zur Sprachverarbeitung liegt in der CRD begründet. Während Köhler davon ausgeht, daß die Strukturinformation zusammen mit dem *gesamten* sprachlichen Material gespeichert wird, schließt Hawkins durch die gelieferte CRD-Definition ja gerade jenes sprachliche Material, das z.B. mit der letzten PP-IC einer englischen VP-CRD verbunden ist, bis auf die Präposition als MNCC, aus seinen Komplexitätsbetrachtungen aus. (Daß es später verarbeitet wird, bestreitet er natürlich nicht.) Köhler geht davon aus, daß die gesamte sprachliche Einheit – zumeist eine Äußerung oder eine eng textuell zusammenhängende Einheit wie ein Textabschnitt –, hierarchisch gesehen über alle Beschreibungsebenen hinweg, Platz im Register finden muß und stellt in diesem Zusammenhang auch Überlegungen zur Kapazitätsbeschränkung des Registers an. Köhlers Überlegungen lassen sich durch jene von Hawkins ergänzen. Hawkins postuliert den Transfer eines vollständig analysierten Teilstücks des zu verarbeitenden sprachlichen Konstrukts in einen permanenteren Speicher, während nur eine Identitätsmarkierung für das gesamte Teilstück im Register verbleibt, was zu einer Entlastung des Registers führt und es gleichzeitig ermöglicht, dieses Konstrukt wieder als Komponente eines höheren Konstrukts zu integrieren, wobei zunächst seine (interne) Struktur dafür keine Rolle spielt. Mit diesen Überlegungen ist das von Köhler aufgeworfene Problem des Registerumfangs weiterhin ungelöst, denn nun stellt sich die Frage, wie komplex ein Teilstück sein darf, bevor es als komplett verarbeitet auf eine Identitätsmarkierung reduziert werden wird.

Zusammenfassend scheint es sinnvoll, beide Ansätze miteinander zu verbinden: Während der Verarbeitung einer noch nicht komplett erkannten Teilstruktur treffen Köhlers Annahmen über die Speicherung von Strukturinformation und sprachlichem Material zu, ist eine Teilstruktur jedoch komplett erkannt, können

die nun nicht mehr interessierenden Details hierarchisch tiefer liegender Ebenen aus dem Register ausgelagert werden, um dieses somit zu entlasten.

3.6 Zusammenfassung

Die in diesem Abschnitt referierten Überlegungen und Untersuchungen zum EIC-Prinzip stellen, neben einer Vorstellung des grundlegenden Ansatzes, nur einen geringen Teil der von Hawkins aus dem EIC-Prinzip gezogenen Schlußfolgerungen zur Syntax dar. Besonders interessant und verfolgenswert sind z.B. seine Motivierung der typologischen Frequenzverteilung der Grundwortstellungen SVO, SOV, VSO, VOS und OVS und seine grundsätzlichen, aus Verarbeitungs- und Effizienzüberlegungen resultierenden Annahmen zur Syntax im Kapitel „Grammaticalized node construction“.

Diese Themen erfordern es jedoch, in einem besonderen Rahmen detailliert behandelt zu werden, da besonders die Klärung der Begriffe Subjekt und Objekt einer ausführlicheren Behandlung bedarf.

4. Empirische Untersuchungen

Einige wenige der von Hawkins durchgeführten Untersuchungen an Performanzdaten wurden nachvollzogen. Im Gegensatz zu den Untersuchungen von Hawkins sollten diese jedoch weitestgehend automatisiert werden. So waren syntaktisch analysierte Texte notwendig, deren bereits vorhandene Aufbereitung es erlaubte, für EIC interessante syntaktische Strukturen maschinell zu verarbeiten

Es wäre wünschenswert gewesen, das EIC-Prinzip an einer Sprache zu untersuchen, die von der Projektgruppe um Hawkins noch nicht berücksichtigt worden ist. Leider standen jedoch nur syntaktisch analysierte Daten des Englischen zur Verfügung, welche zur Überprüfung der EIC-Hypothese herangezogen werden konnten.

Als Datenmaterial diente zum einen die Datenbank, die mit dem linguistischen Analyse-Tool LDB (Linguistic Database),³⁰ mitgeliefert wird. Dieses Korpus besteht aus 4917 Sätzen, aufgeteilt in fünf Teilkorpora sehr unterschiedlicher Textsorten: Tenniskommentar, Krimi, Psychologie, Literaturwissenschaft, Biologie. Bei dem ersten Teilkorpus handelt es sich um gesprochene Sprache.

Zum anderen wurden Daten der Penn Treebank, die im Rahmen der ACL/DCI-Initiative³¹ der Forschung zur Verfügung gestellt werden, verwandt.

³⁰ Vgl. Halteren (1990).

³¹ Diese „Data Collection Initiative“ der Association for Computational Linguistics hat es sich zur Aufgabe gemacht, möglichst viele Sprachdaten der Forschung zu geringen Kosten und ohne Copyright-Gebühren zur Verfügung zu stellen.

Diese Daten stammen aus dem Treebank-Projekt (1990/91) der University of Pennsylvania.³² Obwohl die syntaktische Analyse hier sehr rudimentär ist, konnte das Korpus (ca. 13.000 Sätze des Wall Street Journal) dennoch im Hinblick auf eine syntaktische Konstruktion untersucht werden. Dies war von Vorteil, da so das EIC-Prinzip an einer Datenmenge überprüft werden konnte, die den Hawkins'schen Korpusumfang für die betreffende Struktur um eine Größenordnung übertrifft.

Bei den folgenden Untersuchungen wurde die einfachste Quantifizierung des EIC-Prinzips gewählt, indem die Anzahl der Wörter derjenigen Konstituenten gezählt wurde, die in den untersuchten syntaktischen Konstruktionen für das EIC-Prinzip eine Rolle spielen.³³

In den Abbildungen finden sich die Längen(differenz)angaben auf der X-Achse (in Wörtern gemessen, soweit nicht anders angegeben), während auf der Y-Achse die relative Frequenz (RFRQ) jener Linearisierung abgetragen wurde, die der EIC-Hypothese entspricht. Die RFRQ wird in jeder Untersuchung als die Anzahl der Linearisierungen gemäß der EIC-Hypothese (PFRQ) dividiert durch die Anzahl der Linearisierungen insgesamt (AFRQ) berechnet. Gibt es z.B. in einem Teilkorpus zehn Instanzen von zwei Präpositionalphrasen (innerhalb einer VP) mit der Längendifferenz von fünf Wörtern zwischen den beiden PPn und ist die kürzere PP in sieben Fällen (PFRQ) vor der längeren PP linearisiert, so trägt für die Längendifferenzklasse 5 die RFRQ 0,7.

4.1 Untersuchungen an den LDB-Daten

4.1.1 Subjektextrapolation

Gemäß EIC kann die Hypothese formuliert werden, daß „schwere“ Subjekte im Englischen an das Satzende verlagert werden, da dies die CRD für die gesamte Äußerung verringert und somit ein schnelleres Erkennen der ICs von S gewährleistet wird. Ein extrapoliertes Subjekt wird im Englischen durch ein expletives *it* in der Linearisierung repräsentiert.³⁴

An den fünf Teilkorpora der LDB wurden diese syntaktischen Konstruktionen und die Häufigkeit von Extrapolation versus Nicht-Extrapolation bei unter-

³² Bei den Daten aus dem Penn Treebank Corpus handelt es sich um eine vorläufige Version ihrer Aufbereitung. Für Informationen zu diesem und dem Nachfolgeprojekt vgl. Marcus (1993) und <http://www.cis.upenn.edu/~ldc/treebank2/treebank2.index>.

³³ Bei den eigentlichen Untersuchungen verwendet Hawkins ebenfalls dieses Maß, das oft auf die Heranziehung einer Längendifferenz hinausläuft. Nur für seine theoretischen Überlegungen verwendet er das EIC-Maß bzw. das Aggregatmaß.

³⁴ Zur genauen Diskussion vgl. Hawkins (1994), S.190ff.

schiedlicher Subjektkategorie und unterschiedlicher Subjektlänge in Wörtern (als Quantifizierung von „heavy“) untersucht.

Die aus den LDB-Korpora erhobenen Datenmengen für die Teilkorpora und die einzelnen Subjektkategorien sind zu gering, als daß eine statistisch verlässliche Analyse aus ihnen resultieren könnte. Deshalb wurden die Daten für die Teilkorpora und die verschiedenen Subjektkategorien zusammengefaßt. In

Abb. 4 sind die absoluten Frequenzen der jeweiligen Subjektkategorien durch die Größe der Datenpunkte angedeutet, wobei der Punkt umso größer ist, je mehr absolute Instanzen dieser Kategorie (AFRQ) das Teilkorpus enthält, vgl. auch Tab. 1. Die Tendenz ist eindeutig: Je länger und damit „schwerer“ die Subjektkategorie in Anzahl der Wörter, desto stärker wird von der Subjektex-traposition Gebrauch gemacht.

Tabelle 1

Kenndaten zur LDB-Subjektextraposition mit der Subjektlänge in Anzahl der Wörter, PFRQ: Anzahl der Extrapositionen, AFRQ: Anzahl der möglichen Extrapositionen, RFRQ: Verhältnis von PFRQ zu AFRQ

<i>Subjekt- länge</i>	<i>PFRQ</i>	<i>AFRQ</i>	<i>RFRQ</i>	<i>Subjekt- länge</i>	<i>PFRQ</i>	<i>AFRQ</i>	<i>RFRQ</i>
2	4	36	0,111111	21	4	4	1
3	4	4	1	22	4	4	1
4	5	7	0,714286	23	1	1	1
5	17	18	0,944444	24	4	4	1
6	12	15	0,8	25	5	5	1
7	14	18	0,777778	26	5	5	1
8	13	14	0,928571	27	2	2	1
9	12	18	0,666667	28	3	3	1
10	12	12	1	29	7	7	1
11	12	15	0,8	30	3	3	1
12	11	13	0,846154	31	4	4	1
13	7	7	1	32	2	2	1
14	8	8	1	34	4	4	1
15	13	13	1	35	1	1	1
16	14	14	1	38	1	1	1
17	7	7	1	41	2	2	1
18	9	9	1	42	1	1	1
19	9	9	1	43	1	1	1
20	6	6	1	44	1	1	1

Eine Verfeinerung des Maßes für die Tendenz zur Extraposition, die in den vorliegenden Untersuchungen nur anhand der Größe der Subjekte vorgenommen

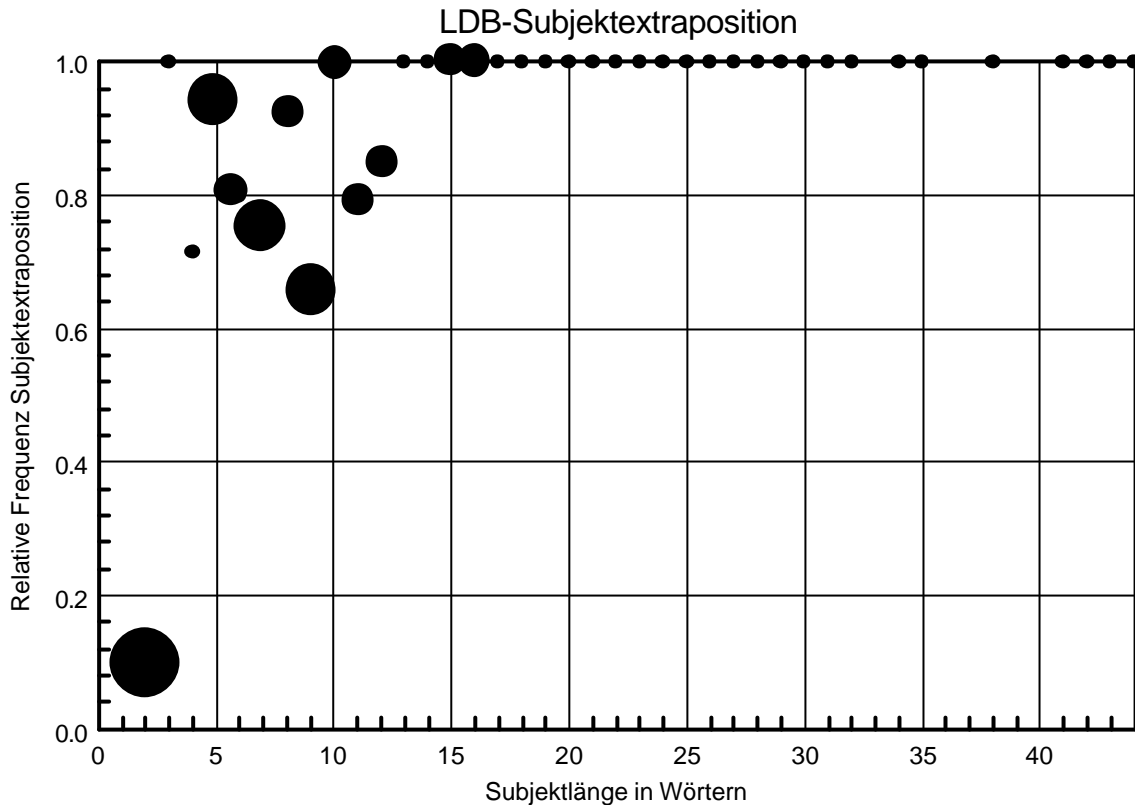


Abb. 4: LDB - Subjektextrapolation

wurde, ließe sich dadurch erreichen, daß die Länge bzw. Größe des Prädikats bei der Quantifizierung berücksichtigt wird, so wie dies z.B. Hawkins (1994) und Erdmann (1988) vorgeschlagen haben.

Zusammenfassend läßt sich festhalten, daß die gemäß EIC gebildete Hypothese zur Extrapolation deutliche Tendenzen zu erfassen scheint. Die hier durchgeführten Untersuchungen sollten auf jeden Fall an einem größeren Korpus wiederholt werden, um für eine statistische Analyse angemessen große Datenmengen hinsichtlich der verschiedenen Subjektkategorien und der Textsorten zu erhalten.

4.1.2 Präpositionalphrasen

Die Untersuchungen zur Linearisierung von PPn in Verbalphrasen anhand der LDB-Daten ergaben ebenfalls zu geringe AFRQen, als daß sie statistisch hätten analysiert werden können. Sie weisen jedoch prinzipiell eine ähnliche Tendenz auf wie die im folgenden Abschnitt diskutierten Untersuchungen zu PPn, die auf einer größeren Datenanzahl beruhen: bei der Linearisierung von zwei Präpositionalphrasen besteht gemäß EIC die Präferenz, die kürzere PP vor der längeren PP zu realisieren, um die CRD der VP möglichst zu verringern.

4.2 Untersuchungen zu Präpositionalphrasen an den Penn-Treebank-Daten

Anhand der syntaktisch analysierten Penn-Treebank konnten Hawkins' Untersuchungen (vgl. Hawkins, 1994:123ff.; Hawkins, 1992:204f.) über die Linearisierung von PPn innerhalb einer VP an einer größeren Stichprobe nachvollzogen werden.

Nach EIC kann die Hypothese aufgestellt werden, daß innerhalb einer VP die

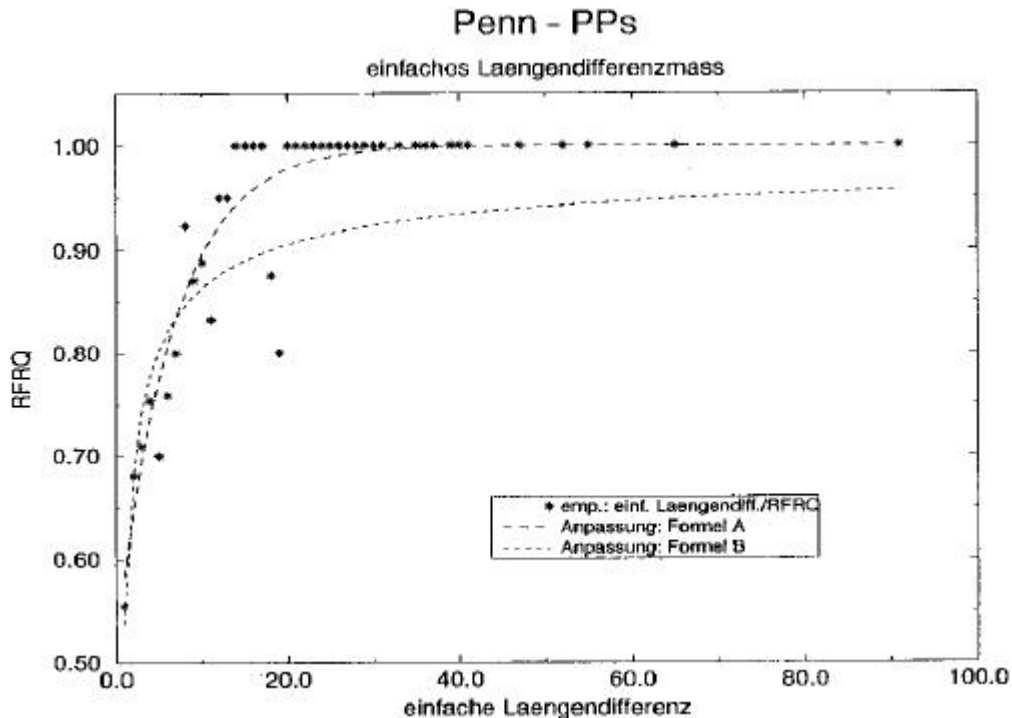


Abb. 5: Penn-PP - einfaches Maß

kürzere PP vor der längeren PP linearisiert werden sollte, um die CRD für die VP zu verkleinern. Des weiteren sollte diese Kurz-vor-lang-Linearisierung umso häufiger auftreten, je größer die Längendifferenz ist.

Insgesamt wurden 1657 Verbalphrasen untersucht. Dabei wurde das Merkmal der Längendifferenz (in Wörtern) im Hinblick auf die Frequenz der Kurz-vor-lang-Linearisierung untersucht. Von allen PPn innerhalb einer Längendifferenzkategorie wurde ermittelt, in wievielen Fällen die längere PP auf die kürzere folgte. So konnte für jede Ausprägung der Längendifferenz eine relative Frequenz ermittelt werden. 287 Fälle, in denen die PPn gleich lang waren, wurden nicht berücksichtigt.

4.2.1.1 Theoretisches Modell

Zunächst kann rein deskriptiv versucht werden, eine möglichst gute Anpassung einer mathematischen Funktion an die empirischen Werte zu erzielen. Die Anpassung eines nicht-linearen Zusammenhangs erfordert die Vorgabe einer Funktionsgleichung, deren Parameter z.B. iterativ angepaßt werden. Im vorliegenden Zusammenhang wird eine nicht-lineare Funktionsgleichung gesucht, die asymptotisch (konvex) gegen 1 läuft und monoton ansteigt.

Als Modell bieten sich Wachstumsfunktionen, z.B. folgende Gleichungen, an:

$$\text{Formel A: } y = 1 - \frac{1}{a e^{bx}}$$

$$\text{Formel B: } y = 1 - \frac{1}{ax^b}$$

Diese beiden Modelle wurden mit dem Statistik-Paket SPSS durch nicht-lineare Regressionen angepaßt, die Parameteranpassung erfolgte in einem iterativen Verfahren der kleinsten Quadrate.

4.2.1.2 Ergebnis der Anpassung

Der Determinationskoeffizient kann Aufschluß über die Güte der Anpassung geben. Er gibt den Anteil der Gesamtvariation von Y an, der durch X linear erklärt werden kann. Sein Maximalwert liegt bei 1, eine Anpassung ist umso besser, je näher sich der Determinationskoeffizient 1 annähert. Gemeinhin geht man davon aus, daß das nicht-lineare deskriptive Modell für empirische Daten zunächst aufrecht erhalten werden kann, wenn der Determinationskoeffizient bei mindestens 0,9 liegt. Der Determinationskoeffizient³⁵ ist definiert als

$$r^2 = 1 - \frac{\sum (y - \hat{y})^2}{\sum (y - \bar{y})^2}$$

mit y = empir. Wert, \hat{y} = theor. Wert, \bar{y} = Mittelwert.

Bei einer linearen Anpassung gibt der Determinationskoeffizient nicht nur Aufschluß über die Güte der Anpassung an die empirischen Daten. Er kann des weiteren in die Berechnung des Standardfehlers einfließen, welcher mit dem t -Test auf Signifikanz überprüft werden kann. Das Ergebnis des t -Tests gibt an,

³⁵ Die Summe aller quadrierten Abweichungen der empirischen Werte vom theoretischen Wert wird dividiert durch die Summe aller quadrierten Abweichungen des empirischen Werts vom Mittelwert (Varianz von y); das Ergebnis wird von 1 subtrahiert.

wie signifikant die Abweichung der Regressionsgeraden von der (fiktiven) der Grundgesamtheit zugrundeliegenden Regressionsgeraden ist.

Der *t*-Test verbietet sich bei der nicht-linearen Anpassung, da prinzipiell nicht ausgeschlossen werden kann, daß es neben dem vorgegebenen Modell noch weitere gibt, die die Daten ebensogut beschreiben. (Dies ist bei der linearen Anpassung nicht der Fall, die nur von einem Modell – eben dem linearen – ausgeht.) Von einem nicht-linearen Zusammenhang kann also nur unter großen Vorbehalten auf das nicht-lineare Modell der Grundgesamtheit geschlossen werden. Manchmal ist es möglich, Daten und Modell durch Logarithmierung monoton zu transformieren, um somit die Vorteile der stärkeren Aussagekraft der linearen Regression zu nutzen. Andererseits muß in diesen Fällen betont werden, daß der Schluß auf die Grundgesamtheit nur für die transformierten Daten zulässig ist.

4.2.1.3 Einfaches Längendifferenzmaß

In Abb. 5 sind die Ergebnisse der nicht-linearen Regression mit den Formeln A und B dargestellt. Die empirischen Werte der RFRQ beruhen auf dem einfachen Längendifferenzmaß, welches Häufigkeitsklassen auf der Basis der Längendifferenz der Anzahl der Wörter zweier PPn bildet. Die in der Tabelle aufgeführten Determinationskoeffizienten wurden ermittelt.

Formel	Det.Koeff. r^2
A	0,91973
B	0,86024

Ausgehend von der Güte der Anpassung sollte das durch Formel A ausgedrückte deskriptive Modell vorgezogen werden. Zwei Datenpunkte der empirischen Ergebnisse können als Ausreißer betrachtet

werden und sollten, ebenso wie alle Datenpunkte der Längendifferenzen > 16 , von der statistischen Analyse ausgeschlossen werden, da ihre AFRQen unter zehn Datenpunkten liegen. Dieses Verfahren ist hier durchaus legitim, da der Ausschluß nicht einzelne Datenpunkte herausgreift, sondern den kompletten Endbereich betrifft.

4.2.1.4 Gewichtetes Längendifferenzmaß

Aufgrund der kritischen Anmerkungen von Hawkins³⁶ bezüglich seiner eigenen Quantifizierung, die für ihn jedoch keinerlei Konsequenzen für die Operationalisierung haben, wurde für die hier vorliegenden Untersuchungen ein gewichtetes Längendifferenzmaß entwickelt. Es beruht auf der Überlegung, daß die Umstellung zweier PPn gemäß EIC weniger dringend/präferiert ist, wenn es sich bei den beiden PPn um sehr lange handelt und die Längendifferenz nur sehr gering ist. Folgendes Maß wurde entwickelt:

³⁶ Vgl. Hawkins (1992), FN 3.

$$\frac{|L(PP_1) - L(PP_2)|}{DL_{PP}} \text{ mit } L(PP_x) = \text{Länge einer PP in Wörtern}$$

DL_{PP} = durchschnittliche Länge der PPn

wobei $DL_{PP_{1,2}} = \frac{L(PP_1) + L(PP_2)}{2}$.

Die absolute Längendifferenz zwischen zwei PPn wird mit der durchschnittlichen Länge der beteiligten PPn gewichtet. Von der Hypothesenbildung her funktioniert diese Operationalisierung analog zum einfachen Längen-differenzmaß: Je höher die gewichtete Längendifferenz, desto häufiger sollte die kürzere PP vor der längeren linearisiert werden. Tabelle 2 illustriert, wie sich die geschilderte Zusatzannahme über die absolute Länge der PPn niederschlägt.

Tabelle 2
Illustration der gewichteten Längendifferenz

$L(PP_1)$	$L(PP_2)$	gew. Längendiff.
3	4	0,133
19	20	0,051
1	5	1,33
20	24	0,18

Eine Längendifferenz von einem Wort führt bei relativ kurzen PPn zu einem insgesamt höheren Wert als bei relativ langen PPn. Ebenso wirkt sich eine größere Längendifferenz bei relativ kurzen PPn stärker aus als bei relativ langen PPn und führt gleichzeitig zu einer größeren hypothetischen Umstellungspräferenz als eine geringe Längendifferenz. Sind beide PPn gleich lang, so führt diese differenzierte Operationalisierung zu einem Wert von 0, analog zum einfachen Differenzmaß. Abb. 6 illustriert die empirischen und deskriptiven Ergebnisse aus dem gewichteten Längendifferenzmaß. Der Versuch der (gewichteten) Anpassung der Formeln A und B muß als gescheitert betrachtet werden, da die Determinationskoeffizienten sehr unbefriedigend ausfallen: Eine zu große Streuung der Datenpunkte um die deskriptiv möglichen Y-Werte verhinderte eine gute nicht-lineare Anpassung.

	Det.Koeff. r^2
A	0,52604
B	0,38298

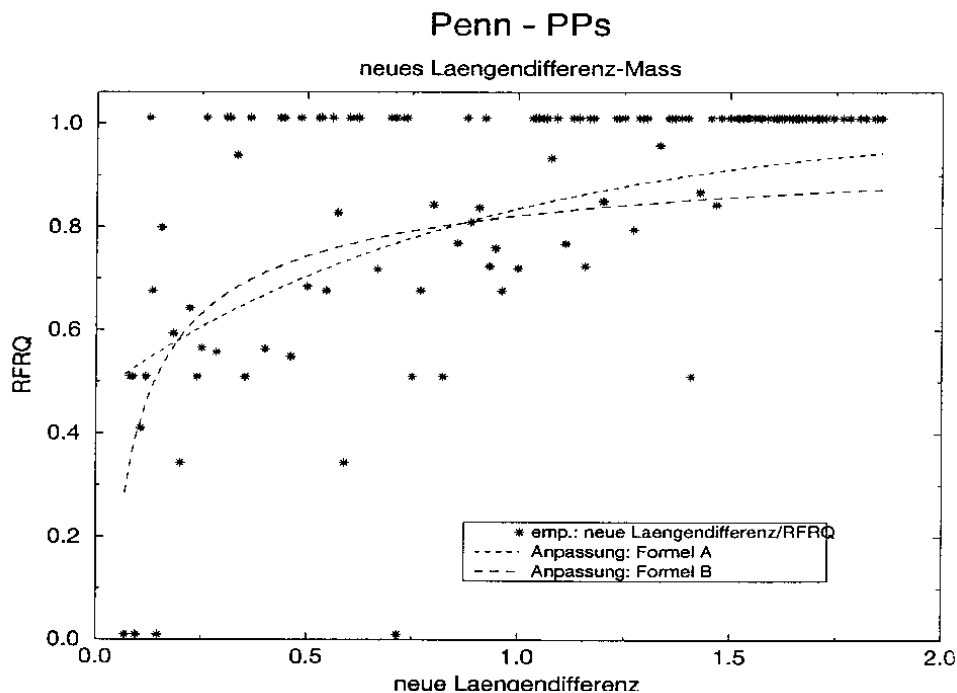


Abb. 6: Penn-PP - gewichtetes Maß

Zunächst muß festgehalten werden, daß das neue Maß zu einer Erhöhung der Anzahl der Merkmalsklassen führt: bei der einfachen Längendifferenz entstehen 44 Klassen, während das neue Maß zu 125 Klassen führt. Wenn dieselben Daten auf mehr Klassen aufgeteilt werden, wird die durchschnittliche Klassenbelegung geringer. Dies führt dazu, daß es kaum noch Klassen gibt, die mindestens zehn Datenpunkte enthalten, wie es eigentlich für eine statistische Analyse notwendig wäre.³⁷

Trotz der abschlägigen Ergebnisse aus der Anpassung sollte die Operationalisierung des neuen Maßes nicht aufgegeben werden, bevor nicht Untersuchungen an größeren Datenmengen durchgeführt wurden, die für eine statistische Analyse ausreichen.

Des weiteren könnte eine differenziertere Quantifizierung durch Verwendung der Proportion der beiden Längen zueinander erzielt werden.

Bleiben die Ergebnisse mit der differenzierten Operationalisierung weiterhin unbefriedigend, so muß überlegt werden, ob entweder die Quantifizierung den Hypothesen nicht angemessen ist oder gar die Hypothesen an sich modifiziert oder revidiert werden müssen.

³⁷ Dies war auch der Hauptgrund dafür, die Extrapositions-Daten nicht statistisch zu analysieren.

4.3 Generelle Linearisierungstendenzen

Ausgehend von Hawkins' theoretischen Überlegungen und seinen Untersuchungen könnte die Maximalhypothese gewagt werden, daß in der gesamten Links-rechts-Linearisierung von Konstituenten in flachen Strukturbeschreibungen die relative Anordnung von ICs der Wurzel zueinander ebenfalls dem Kriterium der Länge folgt. In rechts-verzweigenden Sprachen wie dem Englischen, in denen die MNCC zumeist an der linken Peripherie der Konstituenten zu finden ist, bedeutet dies, daß die durchschnittliche Länge der unmittelbaren Konstituenten des Wurzelknotens relativ zueinander steigen sollte.³⁸ Diese Hypothese ist im Grunde „resultativer“ Art und kann nur indirekt (unter Umkehrung der Abhängigkeiten: die Länge ist abhängig von der Position) aus dem EIC-Prinzip motiviert werden, entspricht jedoch Yngves „Principle of right branching“.

Diese Hypothese wurde an den LDB-Daten überprüft. IC-Kategorien-unabhängig wurden die Längen der unmittelbaren ICs der Wurzel in Wörtern gemessen und zueinander in Bezug gesetzt. Pro Teilkorpus wurde der durchschnittliche Längenunterschied zwischen ICs ermittelt.

Das Ergebnis dieser Exploration wird durch

Abb. 7 illustriert. Auf der X-Achse finden sich die Positionen zwischen zwei direkten Konstituenten der Wurzel, wobei Position 1 die Längendifferenz zwischen erster und zweiter Konstituente bezeichnet, Position 2 die Längendifferenz zwischen zweiter und dritter Konstituente, etc. Auf der Y-Achse finden sich die Mittelwerte der einzelnen Längendifferenzen pro Position. Die gerahmten Datenpunkte stellen Ausreißer dar, die insgesamt nur über eine sehr geringe AFRQ verfügen und somit vernachlässigt werden können.

Bis zu Position 2 (also etwa bis zum Verb) ergeben sich negative Mittelwerte, was bedeutet, daß bis zur dritten Konstituente nachfolgende Konstituenten durchschnittlich kürzer sind als vorhergehende. Bei den folgenden Positionen ist in zweifacher Hinsicht ein positiver Trend zu verzeichnen: Zum einen liegen die Differenzwerte zwischen den ICs im positiven Bereich, was bedeutet, daß eine folgende Konstituente durchschnittlich länger ist als eine vorhergehende. Zum anderen erweisen sich die positiven Differenzwerte als inkrementell. Somit sind folgende Konstituenten nicht nur länger als vorhergehende: folgende Konstituenten sind *umso länger* als die ihnen vorausgehenden, *je weiter hinten* in der Äußerung sie sich finden.

³⁸ Aus Sprachverarbeitungsperspektive sollte dieser Sprachbau präferiert sein, vgl. Yngves „Principle of right branching“ (vgl. Yngve (1960) und Köhler (1999) und Köhler in diesem Band.

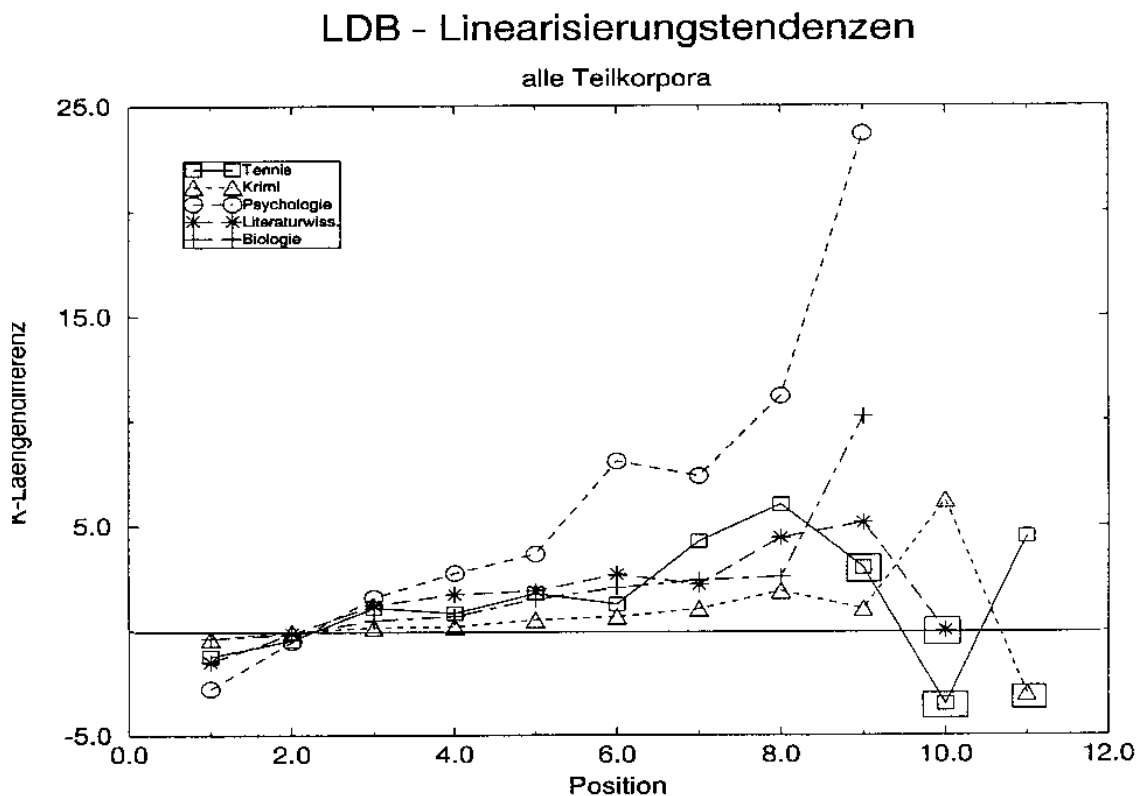


Abb. 7: LDB - Linearisierungstendenzen

Ähnliche Tendenzen zeigt Abb. 8. Ausgehend von der Überlegung, daß auf der IC-Ebene unter der Wurzel häufig koordinierte Strukturen zu finden sind, wurden Koordinationen aufgelöst und als zwei Äußerungen gewertet. Somit wurde der „Verzerrungsfaktor“ Koordination aus den Untersuchungen entfernt. Dieser bestand darin, daß die relativ kurze koordinierende IC (z.B. nach einem Teilsatz) das vermutete Wachstum der ICs empfindlich stören mußte. Durch die Auflösung der Koordinationen wurde dieser Faktor beseitigt, komplexe Satzgefüge auf einfachere reduziert. Diese Reduktion bringt es jedoch mit sich, daß sie eine verzerrte sprachliche Wirklichkeit darstellt: Durch die Tilgung der Koordination in einer Äußerung können, müssen aber nicht, vollständige einfache Sätze entstehen. Durch Tilgung der Koordination zweier Verbalphrasen, die sich eine Subjekt-NP „teilen“, entstehen z.B. ein grammatisch vollständiger und ein grammatisch unvollständiger Satz. Dieser Vorbehalt im Hinterkopf ist es dennoch interessant zu sehen, vgl. Abb. 8, daß sich die für

Abb. 7 festgestellten Tendenzen zu bestätigen scheinen. Die mit Kästchen markierten Ausreißer entstehen wiederum nur durch Mittelwerte aus sehr geringen AFRQen (hier jeweils 1).

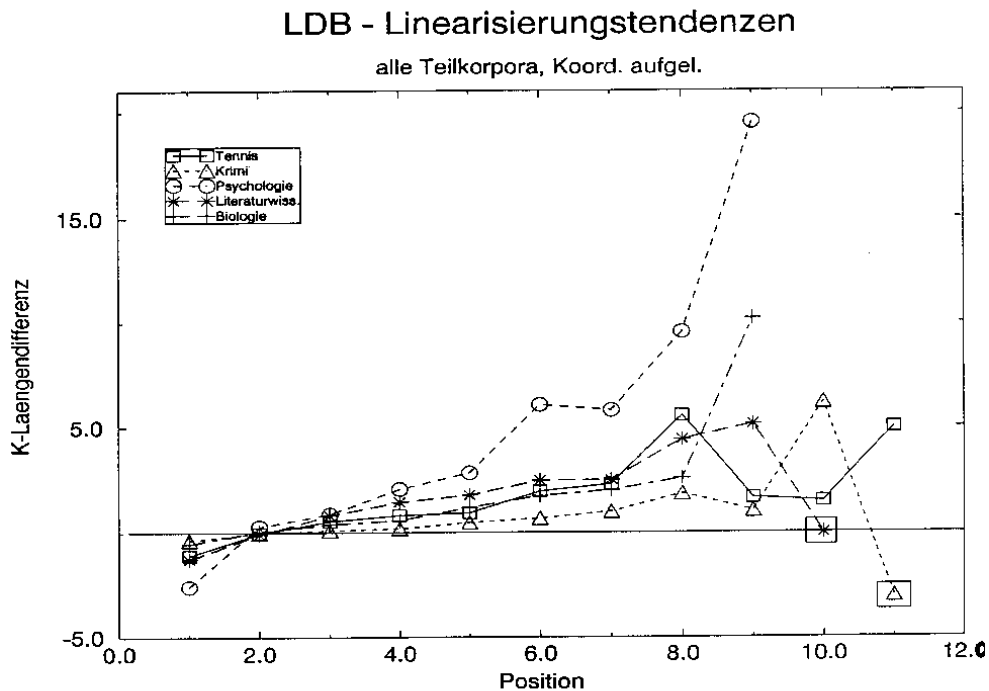


Abb. 8: LDB - Linearisierungstendenzen, Koord. aufgelöst

4.4 Zusammenfassung

Die Untersuchungen des Englischen hinsichtlich einiger syntaktischer Strukturen, innerhalb derer die Linearisierung dem EIC-Prinzip folgen sollte, können im Ergebnis zunächst als positiv bewertet werden. Das EIC-Prinzip scheint durchaus generelle Linearisierungstendenzen zu erfassen.

Die Ergebnisse der Analysen an den LDB-Daten ließen es leider nicht zu, textsortenspezifische Charakteristika der Linearisierungstendenzen aufzuspüren. Zwar verlaufen die Belegungen der Häufigkeitsklassen durchaus unterschiedlich, doch kann aufgrund der dünnen Datenlage keine Hypothese zu einem textsorten-abhängigen Parameter gewagt werden. Hier müssen umfangreichere Daten, die statistisch ausgewertet werden können, erhoben werden.

In weiteren Untersuchungen sollte die neue Operationalisierung des Längendifferenzmaßes an größeren Korpora verwendet werden, um es als Maß besser beurteilen zu können. Des weiteren könnte es sich als Maß für andere syntaktische Strukturen (nicht nur PPs) anbieten.

Die Untersuchungen zu den generellen Linearisierungstendenzen sollten differenzierter verfolgt werden, um z.B. IC-Kategorien-abhängige Aussagen machen zu können. Auch ist die Mittelwertbildung zur Bestimmung der Längendifferenz

adjazenter ICs eine äußerst vergrößernde Reduktion, da sie die Verteilung der Längendifferenzen einer einzelnen Position nur durch einen einzigen statistischen Kennwert erfaßt. Hier ist es notwendig, genauere Analysen durchzuführen.

Bei den PPN-Analysen zum einfachen Längendifferenzmaß konnte eine Beschreibung der empirischen Daten auf heuristischer Basis erzielt werden, welche den Zusammenhang mathematisch faßt. Wünschenswert wäre die deduktive Gewinnung einer entsprechenden mathematischen Funktion, um den empirischen Befund, der sich aus dem von Hawkins postulierten Mechanismus ergibt, explanativ begründen zu können.

5. Ausblick

Im vorliegenden Aufsatz wurde der Hawkinssche Ansatz zur Wortstellung und Konstituenz incl. seiner methodologischen Begrifflichkeit (aus der funktionalen Typologie) vorgestellt und anhand weiterer Untersuchungen mittels der Methodologie der synergetischen Linguistik verfolgt. Die Ergebnisse der empirischen Untersuchungen konnten Hawkins' Hypothesen vorläufig bestätigen.

Ein grundlegendes Problem bleibt bei der Hawkinsschen Argumentation zu den kognitiven Sprachverarbeitungsstrategien, die die Basis für das Prinzip der „Early immediate constituents“ bilden, bestehen. Zunächst ist das Argument durchaus einleuchtend, „längere“ bzw. „schwerere“ Konstituenten an das Ende (bzw. den Anfang, je nach Sprachtyp und syntaktischer Argumentation) zu stellen, um die CRD zu verkürzen. Bei weiterer Überlegung stellt sich jedoch die Frage, ob nicht auch bis zum Ende eben dieser Konstituente abgewartet und damit „gelesen“ werden muß, um Klarheit darüber zu erzielen, ob nicht noch weitere Konstituenten folgen, was oft grammatisch möglich ist. Somit würde der Verarbeitungsgewinn, möglichst früh ein vollständiges Erkennen der IC-Struktur einer Konstituente zu erzielen, hinfällig. Andererseits scheinen die empirischen, sprachenübergreifenden Untersuchungen zu den aus dem EIC-Prinzip gewonnenen Hypothesen durchaus für seine Plausibilität zu sprechen. Behaghels *Gesetz der wachsenden Glieder* bzw. der „Heavy NP-Shift“ oder auch das EIC-Prinzip scheinen sich tatsächlich sprachlich zu manifestieren. Eine vollkommen befriedigende Erklärung über den zugrundeliegenden (kognitiven oder auch pragmatischen) Mechanismus steht jedoch noch aus.

Ein noch nicht eingelöstes Desideratum betrifft den methodologischen Ansatz der empirischen Hypothesenprüfung. Die Hypothesen zu den Linearisierungspräferenzen sollten formal so gefaßt werden, daß statistische Tests möglich sind, wenn diese auch zunächst nur die deskriptive Validität der Hypothesen betreffen. In einem weiteren Schritt sollte die linguistische Interpretation der beteiligten Parameter erfolgen sowie eine konkrete numerische Einbindung der beteiligten funktionalen Anforderungen. Eine weitere Aufgabe bestünde darin, die so formulierten Hypothesen in Zusammenhang zu bereits bestehenden Hypothe-

sen in der Quantitativen Linguistik zu setzen und sie in ein Modell der sprachlichen Selbstorganisation einzubetten. Doch um diesen Schritt zu tun, müssen zunächst die funktionalen Anforderungen, denen Wortstellungsphänomene unterliegen, klarer formuliert werden, als dies bisher der Fall ist, wie die Diskussion im vorherigen Abschnitt zeigte.

Andere Teilfragen sind noch offen. Es müßte z.B. überprüft werden, ob das EIC-Prinzip wirklich formalismenunabhängig ist, wie zunächst angenommen. Wie können des weiteren die syntaktischen Strukturen identifiziert werden, über die EIC-Hypothesen überhaupt gebildet werden dürfen? Wie könnte eine klarere Meßvorschrift bei einzelnen syntaktischen Strukturen zur Hypothesenprüfung aussehen? Wie können die von Hawkins angedeuteten Motivationen für die Linearisierung von Äußerungsteilen in die Ergebnisse der eher diskursorientierten Typologie integriert werden: widersprechen sie sich oder ergänzen sie sich gar?

Die vorliegende Arbeit hat sich im Grunde nur auf den Teil des von Hawkins vorgestellten Ansatzes konzentriert, der sich mit der Wortstellung befaßt. Seine Überlegungen zur Wortstellung sind jedoch eng mit den Thesen zur Konstituenz verknüpft, einem Bereich, der noch intensiver berücksichtigt werden müßte.

6. Literatur

- Behaghel, O.** (1930). Von deutscher Wortstellung. *Zeitschrift für Deutschkunde*, 44, 81-89.
- Comrie, B.** (1981). *Language universals and linguistic typology. Syntax and morphology*. Oxford: Blackwell.
- Croft, W.** (1990). *Typology and universals*. Cambridge u.a.: University Press.
- Erdmann, P.** (1988). On the principle of 'weight' in English. In C. Duncan-Rose u.a. (Hg.): *On language. Rhetorica - Phonologica - Syntactica*. A Festschrift for Robert P. Stockwell from his friends and colleagues. London: Routledge.
- Frei, H.** (1929/1971). *La grammaire des fautes. Introduction à la linguistique fonctionnelle, assimilation et différenciation, brièveté et invariabilité, expressivité*. Genève: Slatkine reprints.
- Givón, T.** (1984). *Syntax. A functional-typological introduction. Band 1*. Amsterdam: Benjamins.
- Greenberg, J.H.** (1963/1966). Some universals of grammar with particular reference to the order of meaningful elements. In J.H. Greenberg (Hg.): *Universals of language. Report of a conference held at Dobbs Ferry, New York, April 13-15, 1961*. 2. Aufl. Cambridge, Mass.: (MIT).
- Greenberg, J.H.** (1966). *Language universals. With special reference to feature hierarchies*. The Hague u.a.: Mouton. (Janua linguarum; Series minor; 59).
- Hawkins, J.A.** (1983/1988). *Word order universals*. 2. pr. San Diego u.a.: Academic Press (Quantitative analyses of linguistic structure).
- Hawkins, J.A.** (1990). A parsing theory of word order universals. *Linguistic inquiry*, 21(2), 223-261.
- Hawkins, J.A.** (1992). Syntactic weight versus information structure in word order variation. In J. Jacobs (Hg.), *Informationsstruktur und Grammatik*. Opladen: Westdeutscher Verlag (Linguistische Berichte / Sonderheft; 4). 196-219
- Hawkins, J.A.** (1994). *A performance theory of order and constituency*. Cambridge: University Press (Cambridge studies in linguistics; 73).
- Hoffmann, Ch.** (1999). Word order and the principle of „Early Immediate Constituents“ (EIC). Proceedings of the third international conference on quantitative linguistics, August, 26-29, 1997, Helsinki, Finland, *Journal of quantitative linguistics*, 6,2, 108-116.
- Kimball, J.** (1973). Seven principles of surface structure parsing in natural language. *Cognition*, 2,1, 15-47

- Köhler, R.** (1984). Zur Interpretation des Menzerathschen Gesetzes. In J. Boy (Hg.), *Glottometrika* 6 (S. 177-183), Bochum: Brockmeyer.
- Köhler, R.** (1986). *Zur linguistischen Synergetik. Struktur und Dynamik der Lexik*. Bochum: Brockmeyer. (Quantitative linguistics; 31).
- Köhler, R.** (1990). Synergetik und sprachliche Dynamik. In W. Koch (Hg.), *Natürlichkeit der Sprache und Kultur* (S. 96-112), Bochum: Brockmeyer.
- Köhler, R.** (1993). Synergetic linguistics. In R. Köhler et al. (Hg.), *Contributions to quantitative linguistics*. Proceedings of the first international conference on quantitative linguistics, Qualico, Trier, 1991 (S. 41-51), Dordrecht u.a.: Kluwer.
- Köhler, R.** (1999). Syntactic structures. Properties and interrelations. Proceedings of the third international conference on quantitative linguistics, August, 26-29, 1997, Helsinki, Finland, *Journal of quantitative linguistics*, 6,2, 46-57.
- Marcus, M.P., Santorini, B., & Marcinkiewicz, M.A.** (ca. 1993) *Building a large annotated corpus of English*. Technical report. University of Pennsylvania, Department of Computer and Information science. (<http://www.cis.upenn.edu/~ldc/treebank2/cl93.html>).
- Siewierska, A.** (1993). Syntactic weight vs information structure and word order variation in Polish. *Journal of linguistics*, 29, 233-265.
- Uhlírová, L.** (1997). Length vs order. Word length and clause length from the perspective of word order. *Journal of quantitative linguistics*, 4, 266-275.
- Yngve, V.H.** (1960). A model and an hypothesis for language structure. *Proceedings of the American philosophical society*, 104,5. 444-466.