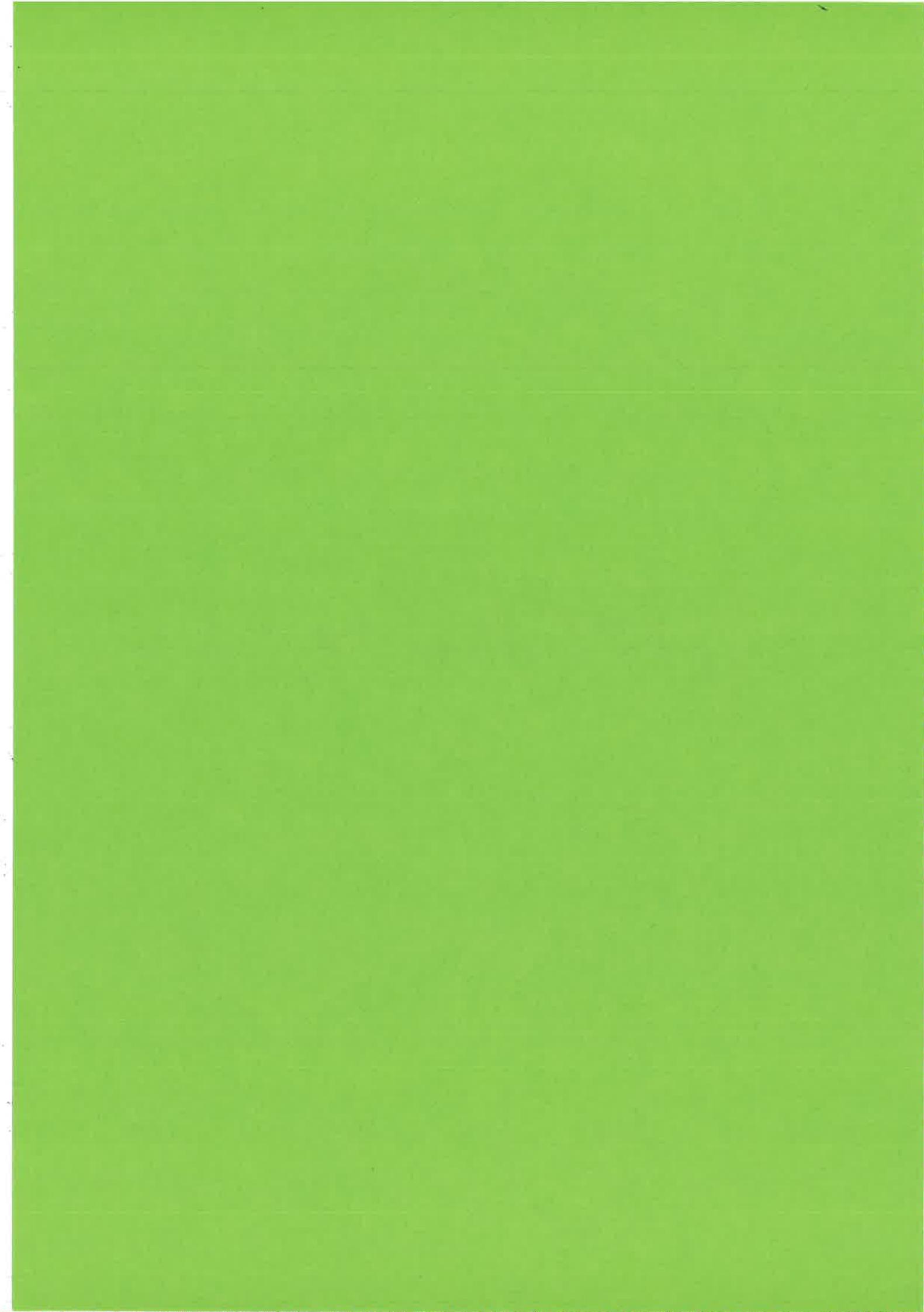


WPFG

working papers in functional grammar

wpfg no. 53
September 1993

Texterzeugung am Beispiel kuschitischer Narrativtexte: Grundlagen und Ergebnisse
Klaus Wedekind
Addis Abeba



Texterzeugung am Beispiel kuschitischer Narrativtexte

1. Einleitung

Das Hauptziel der vorliegenden Untersuchung ist es, am Beispiel einiger kuschitischer Texte zu beleuchten, worin der nicht weiter reduzierbare Kern narrativer Texte besteht, und wie ein Erzähler mit dem umgeht, was er erzählt.¹ Die Untersuchung soll also die gedanklichen Vorgänge erhellen, die zur Erzeugung eines Textes führen - wobei das Konzept eines gegebenen Textkerns von Interesse ist. Es wird hier nicht vorausgesetzt, daß ein Text Syntax erzeugt, wohl aber, daß ein Text den Umgang mit der vorgegebenen Syntax rigoros steuert.²

Unter "Texterzeugung" soll die konkrete linguistische Ausformung einer vom Sprecher erinnerten oder intendierten Erzählung verstanden werden. Das sprachliche Generieren von Aussagen ist auch als Übersetzen definiert worden, nämlich als das Übersetzen einer intendierten Aussage in seine tatsächliche Form hinein (Garrett 1975, zitiert nach Tanenhaus 1988: 27). Im Gegensatz zu den meisten bisherigen Untersuchungen geht es hier allerdings nicht um kurze, isolierte Aussagen, sondern um längere Texte; und es geht nicht um Texterzeugung in einer linguistisch beschriebenen Literatursprache, sondern um eine bisher kaum beschriebene Sprache mit rein oraler Tradition. Das *genre* der hier untersuchten Texte ist auf Erzählungen (Narrativtexte) eingegrenzt. Narrativtexte stellen allerdings insofern ein sehr inklusives *genre* dar, als sie Passagen sehr verschiedener Textsorten einschließen.

Der hohe Stand äthiopischer Rhetorik und die für Äthiopien zentrale Rolle oraler Kunst ist wiederholt beschrieben worden (z.B. Levine 1965). Bei den hier vorgestellten Narrativtexten handelt es sich um kuschitische Texte; also nicht um Texte der besser beschriebenen äthiosemitischen Tradition. Was allerdings den Wert angeht, welcher der Rhetorik in kuschitischen Sprachen beigemessen wird, so stehen sie darin anderen Sprachen Äthiopiens nicht nach; das hat Cerulli schon sehr früh belegt (Cerulli 1922). Die linguistischen Mittel und der kühle Abstand einer Untersuchung mögen den Texten nicht gerecht werden. Aber jeder Versuch, sich der Kunst äthiopischer Rhetorik anzunähern, trägt seinen Lohn in sich.

2. Darstellung der Methode und Terminologie

Ein Narrativtext wird hier gesehen als das, was der Erzähler zu erzählen hat, und wie er es erzählt. Die vorliegende Studie sucht adäquate Darstellungen dessen, was aus dem Repertoire eines Erzählers kommt, und wie es zu dem tatsächlich erzählten Text wird.

Die Trennung in "was" und "wie" ist zwar der maschinellen Texterzeugung entlehnt (McKeown 1985: 2); sie läßt sich aber auf Konzepte der antiken Rhetorik zurückverfolgen (z.B. *res* und *dispositio*) und entspricht psycholinguistischen Dichotomien (z.B. *message* vs. *strategy*). Das "Was" hat seinen Ort im Repertoire des Erzählers, das "Wie" in der Sprachkompetenz.

Die vorliegende Darstellung ist *text-based* (Levy 1979: 184), insofern als sie auf der Untersuchung von Textsammlungen beruht und Behauptungen aufstellt darüber, was den Kern eines Narrativtextes ausmacht. Die Darstellung ist *mind-based*, insofern als sie aus der Perspektive des Sprechers beobachtet, wie es zur Synthese eines Textes kommt. Das "Was" wird hier dargestellt als die Partitur eines Textkerns, und das "Wie" als eine Serie von Regeln.

¹. In den vorliegenden Ausführungen geht es vor allem darum, unter neuen Perspektiven zusammenzufassen und zu ergänzen, was an Einsichten über die Synthese von Gedeo-Texten vorliegt (Wedekind 1990). Der Apparat syntaktischer, morphologischer und phonologischer Regeln wird in diesem Zusammenhang nur illustrierend gestreift: die einschlägigen Regeln der Gedeo-Sprache sind in einer zusammenhängenden Darstellung (a.a.O.) - zusammen mit Burji und Sidamo - bereits behandelt worden.

². Zur Frage der "Erzeugung" syntaktischer Muster aus dem Diskurs (z.B. García 1979: 46; Givón 1979) nehmen wir die schwächere Position ein, daß Syntaxregeln fest vorgegeben sind, aber unter Textbedingungen abgerufen werden. Daß ein in Texten bevorzugtes Muster langfristig auf die Syntax zurückwirken wird - wie von Dik vorausgesetzt (z.B. Dik 1978: 177, 1989: 378) - ist für diese HEC-Sprachen in bezug auf die häufigen pragmatischen Umstellungen von Interesse (z.B. Wedekind 1990: 651 zu den Satzmustern O[-THEMA], SV und OV, S[-CODA]).

2.1 Übersicht

Die hier eingenommene Sicht von Texterzeugung wird im folgenden in einer kurzen Übersicht dargestellt. Dabei geht es - in statischer Sicht - um die sprachlichen Mittel, die einem Sprecher zur Verfügung stehen, und - in dynamischer Sicht - um die Schritte, die einen Text bis zu seiner Darstellung entwickeln.

In der statischen Sichtweise (A) wird vorausgesetzt, daß der Sprecher über verschiedene Bereiche sprachlichen Wissens und sprachspezifischer Kompetenz verfügt. Der Zugang zu diesen Bereichen wird erst dann zu dem (weiter unten beschriebenen) dynamischen Ablauf aktiviert, wenn der Sprecher im Blick auf ein gegebenes Umfeld eine bestimmte Absicht verfolgt. In allen hier untersuchten Texten ist die Absicht übrigens die, aus einem Repertoire einen Narrativtext mitzuteilen: etwas zu erzählen.

Das Repertoire eines Sprechers (1.1) enthält die Texte, die erzählt oder erzeugt werden können. Bei den zweisprachigen Sprechern der hier untersuchten kuschitischen Sprachen ist der Großteil dieses Repertoires nicht sprachspezifisch (einige Texte wurden in zwei verwandten Sprachen aufgenommen). Zur sprachspezifischen Kompetenz werden das Lexikon (2.1) und die Grammatik (2.2) gerechnet.

Einige Grenzen zwischen den hier behaupteten Modulen sind aber deutlich durchlässig: Zwischen Erzählrepertoire und Lexikon zum Beispiel sind sie insofern offen, als im Repertoire eines Sprechers neben den Texten auch Wendungen oder Versatzstücke (*frozen expressions*) vorhanden sind, mit denen der Sprecher umgeht, als wären sie lexikalische Einheiten - von Idiomen des Lexikons vielleicht nur historisch zu unterscheiden. Ähnlich verhält es sich mit der Grenze zwischen den Modulen "Lexikon" und "Morphologie": Das mentale Lexikon hält neben elementaren Wurzeln auch gewisse morphologische Derivationen in der Weise bereit, als wären sie elementare morphemische Einträge. Die diffuse Durchlässigkeit der hier genannten Modulen ist aber kein hinreichender Grund, solche Modulen überhaupt - wegen ihres ontologisch bisher nur "nützlichen" Daseins - wegzubehaupten.

Schema 1

(A) Statische Darstellung (Inventar des Erzählers)

1. Repertoire "Was"	2. Sprachspezifische Kompetenz "Wie"		
1.1 Texte Narrativtexte [...]	2.1 Lexikon Idiome Wörter Morpheme	2.2 Sprachmuster Syntax Wörter Morpheme	Morphologie Phonologie
Wendungen			

In dynamischer Sicht (B) wird vorausgesetzt, daß es zwei Arten von Vorgehensweisen gibt: einerseits gesteuerte, relativ langsame, und andererseits automatische, relativ schnelle.

Einerseits trifft ein Sprecher Entscheidungen über den vor ihm liegenden Text - das heißt, er unternimmt relativ langsame, kontrollierte Schritte, um zum Beispiel die Darstellung der kommenden Szene vorzubereiten. Andererseits verfügt er über unmittelbare Zugriffe zu den Lexikoneinträgen und den Automatismen seiner sprachlichen Kompetenz, die normalerweise nicht bewußt gesteuert werden müssen. Sie sind so schnell, daß sie den Redefluß nicht spürbar unterbrechen - selbst bei simultaner Vorbereitung korrigierter Versionen oder stilistischer Varianten nicht.

Texterzeugung am Beispiel kuschitischer Narrativtexte

Schema 2

(B) Dynamische Darstellung der Vorgänge des Erzählens

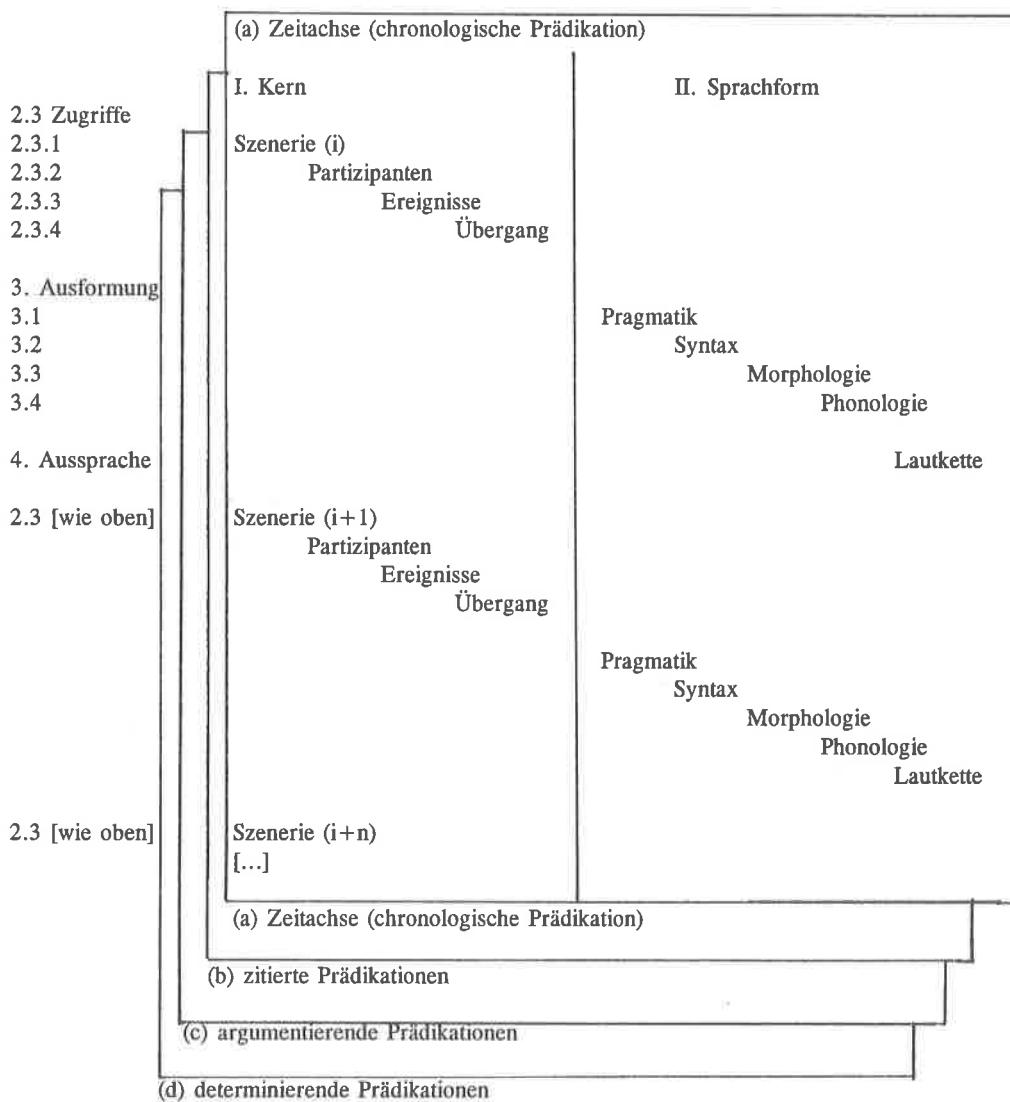
1. Absicht für das Umfeld

2. Entscheidungen

2.1 Textentscheidungen



2.2 Prädikationsentscheidungen



2.2 Die Schritte im einzelnen

Die Schritte (B) des Erzählers - von der Absicht (1), etwas mitzuteilen, bis zu den tatsächlichen Lautketten des erzählten Textes (4) - sind nach Entscheidungen (2) und Automatismen (3) im einzelnen wie folgt zu unterscheiden:

Anfänglich (2.1) wählt der Sprecher im Hinblick auf eine (holistische) Übersicht den gesamten Text. Dabei wird keine der Szenen bis zur ihrer linguistischen Form ausgestaltet.

Dann (2.2) werden die Prädikationen einer Szene - eine nach der anderen - konzeptuell aufbereitet. Das entspricht zwar nicht einem Trend des Diskursfundierten Generierens, wie er von Tanenhaus kürzlich dargestellt worden ist, aber es widerspricht jedenfalls nicht den hier vorliegenden Erkenntnissen der Psycholinguistik (Tanenhaus 1988: 13, vgl. Johnson-Laird 1983). Jeder Narrativinhalt läßt sich nämlich in wechselnde Szenen (2.3.1) eingliedern, in denen verschiedene Partizipanten (2.3.2) eine Abfolge von Ereignissen (2.3.4) durchlaufen. Dabei werden die Konzepte Szene, Partizipant und Ereignis bewußt in dieser bevorzugten Reihenfolge gegeben. Es soll dem Konzept Ereignis (Prädikation) kein Primat und keine zeitliche Priorität über die anderen Konzepte (Partizipant oder Szene) zuerkannt werden, und es soll dem Konzept Partizipant kein Primat und keine zeitliche Priorität über das Konzept Szene zuerkannt werden.

Für verschiedene Arten von Prädikationen gelten dabei verschiedene Strategien: (a) die Prädikationen der Zeitachse werden anders miteinander verknüpft als (b) Zitate oder (c) argumentierende oder (d) determinierende Prädikationen. Auch hier sind die Grenzen zwischen den verschiedenen Teilen wieder durchlässig: In jede Prädikation der Zeitachse zum Beispiel gehen Teile der vorausgehenden Prädikation mit ein: Partizipanten bleiben als Topik bestehen, Zeit und Modus werden unverändert übernommen, usw. Darum werden auch in jeder Prädikation nur wenige Phrasen grammatisch ausgeformt - vielleicht zwei oder drei - die anderen werden impliziert. Das entspricht der psycholinguistisch optimalen Menge von zwei aufzuarbeitenden Konzepten, ehe eine sprachliche Äußerung artikuliert wird. In kuschitischen Sprachen verlangt das Ende einer Prädikation dann eine explizite Entscheidung darüber, welche Perspektiven wechseln und welche bestehen bleiben.

Etwas breiter ausgeführt heißt das: Die Szenen - d.h. die Konstellationen von zeitlichen, räumlichen und anderen Umständen - bieten den jeweils stabilen Rahmen dafür, daß die Partizipanten - d.h. die wechselnden Gruppierungen von Personen und Dingen - bestimmte Ereignisse durchlaufen - d.h. an einer chronologischen oder logischen Abfolge von Aktionen und Entwicklungen teilnehmen. (Was für die narrative Welt konkreter Dinge und Bewegungen gilt, läßt sich metaphorisch auch für die Welt der Ideen und Argumentationen sagen: auch in argumentativen Texten z.B. gibt es Szenen, Partizipanten und Ereignisse: *an argument is a journey*. Narrativtexte sind aber besser zu "fassen"; cf. Lakoff & Johnson 1980: 89.) In der Terminologie der Funktionalen Grammatik gilt das oben Gesagte wie folgt: Alle Szenen sind darstellbar als Konstellationen von temporalen und lokalen (seltener: modalen etc.) Satelliten, die für mehrere Prädikationen gelten. Alle Partizipanten sind darstellbar als Gruppierungen von belebten (seltener: unbelebten) Termen oder Argumenten, die in mehreren Prädikatsrahmen präsent bleiben. Alle Ereignisfolgen sind darstellbar als chronologische (seltener: logische) Abfolgen von Prädikationen, deren jeweiliger Kern ein Verbrahmen ist.

Nach diesen Entscheidungen über Umstände, Partizipanten und Ereignisse einer Szene wird (3) die Sprachform dieser Prädikation bis (4) zur Phonologie hin fertig ausgebaut. In den hier untersuchten kuschitischen Sprachen ergibt eine solche Szene syntaktisch jeweils einen komplexen Satz oder Paragraphen. Diese linguistische Ausformung geschieht weitgehend automatisch. Sobald dann jedes Konzept seine phonologische Form gefunden hat und ausgesprochen worden ist, kann ein Großteil der Sprachformen de-aktiviert oder aus dem Frischgedächtnis gelöscht werden. Es gilt wieder (2) der übergreifende Plan des Repertoires, und die Aufmerksamkeit wendet sich der nächsten, kommenden Szene zu.

Bei gleichbleibender Szenerie werden am Anfang einer kommenden Szene die schon aufgebauten Gegebenheiten noch einmal genannt (*tail-head link*), und der Erzähler gewinnt Zeit, die neuen Teile der Prädikationen zu entwickeln. Bei wechselnder Szenerie braucht der Sprecher - und mit ihm der Zuhörer - eine Pause oder ein Signal, um den zu erwartenden neuen Perspektiven und Partizipanten gerecht zu werden. (Es versteht sich, daß die Spuren dieses planenden Vorgehens - Wiederholungen, Pausenmorpheme etc. - getilgt werden, wenn der Sprecher den oralen Text als "Literatur" ediert; Wedekind 1990: 392, 462.)

Der eben vorgestellte Vorgang wiederholt sich so oft, bis alle Propositionen der Erzählung in Lautketten verwandelt sind. Schließlich wird der Ausstieg aus dem Narrativtext dadurch signalisiert, daß Zeit und Aspekt ausdrücklich auf das "Hier" und "Jetzt" des Erzählers zurückgeschaltet werden.

Texterzeugung am Beispiel kuschitischer Narrativtexte

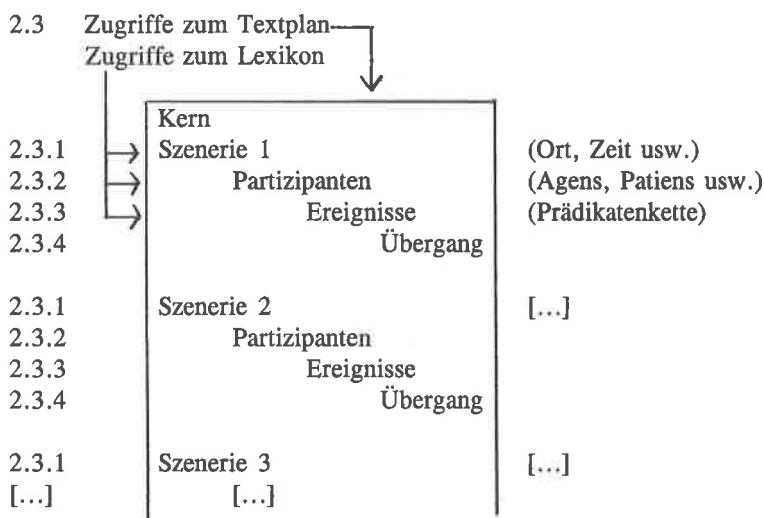
2.2.1 Schritte, die den Textkern betreffen

Die mentale Aufbereitung aller Propositionen (2.3) läßt sich als Teil des oben gegebenen Gesamtschemas (B) darstellen.

Wenn der Sprecher sich seinem Repertoire auswählend zugewendet hat (2.1), ruft er Teile des Textkerns ab - Teile von ungefähr der Größenordnung einer Proposition - wobei einige der aktivierten Konzepte bis in die nachfolgende Proposition hinein vorgehalten werden, und wobei der übergreifende Plan des gesamten Narrativtextes einige Entscheidungen steuert.

Die Zugriffe zum Lexikon (2.3) betreffen während dieses Schrittes nur die semantische Seite der Einträge. Die phonologische Seite der Einträge wird hier nicht integriert.

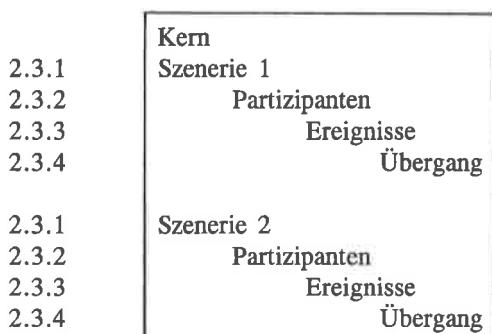
Schema 3



Festzuhalten ist hier, daß Partizipanten über mehrere Ereignisse hin präsent bleiben können. Das Lexikon wird nicht jedesmal neu dazu befragt. Übergänge (2.3.4) ergeben sich aus der Summe aller Veränderungen: neue Szene, neuer Agens, neue Ereigniskette usw. Das könnte schematisch etwa wie folgt angedeutet werden:

Schema 4

2.3 Zugriffe zum Lexikon und zum Textplan (Repertoire)



In einer "Textpartitur", die alle zugrundeliegenden Konzepte als lexikalische Einträge andeutet, wird ein solcher Textkern für einen konkret zu erzählenden Text zum Beispiel wie folgt ausgefüllt. Das Beispiel ist dem weiter unten ganz vorgestellten Text entnommen (Teil 8: Gedeo-Text K).

Schema 5
Beispiel eines Textkerns (Text K)

Zeitachse

Szenerie 1

Partizipanten

	Agens Patiens	andere Rollen	Ereignisse Prädikate	Übergänge
2	[1]'ani	soolaama-DIR-SZE		
2	ich	(Ortsname)	ke'- anfang-	
3		DET rakk-FOK giddo DEM-DIR	ke'-	
3		Mühe in	anfang-	
4		makiina-INSTR		dag-
4		Wagen		komm-

"Von S. kommend, bin ich unter großen Mühen mit dem Bus gekommen."

Szenerie 2

Partizipanten

	Agens Patiens	andere Rollen	Ereignisse Prädikate	Übergänge
6	[2]berhaanu tasamma, DET-belto,		dag-	
6	(Name)	Sohn		komm-
7	[1]	nafse-DET-DIR	'alayy-	
7	(ich)	Leben	fürcht-	
10		> > 8- 9 < <	hiyy-	
10			sag-	
11			'alayy-	
11			fürcht-	
12			dag-	
12			komm-	

"B.S., der Sohn des Z., kam, ich fürchtete um mein Leben, ich dachte (8-9) [...] und fürchtete mich, als ich kam."

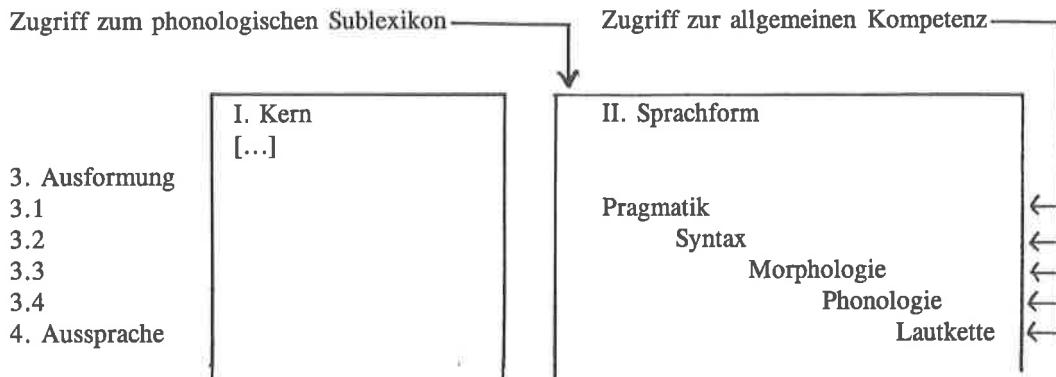
2.2.2 Schritte, die sprachspezifische Prozesse betreffen

Nachdem feststeht und bereitsteht, "was" zu sagen ist, wird dieser festgestellte Inhalt linguistisch aufbereitet (vgl. oben 3.1-3.4 in Schema 2). Die lexikalischen und textlichen Konzepte durchlaufen alle sprachspezifischen Prozesse pragmatischer, syntaktischer, morphologischer und phonologischer Art. In einem zweiten Zugriff zum Lexikon werden die semantischen Konzepte zu phonologischen Formen spezifiziert; die für die Repräsentation ausgewählten Wörter und Phrasen werden sequenziert, und in einem letzten Schritt werden auf benachbarte Laute die betreffenden morphophonologischen Regeln angewandt.

Das Ergebnis ist der Text in seiner tatsächlichen phonologischen Form (4), so wie er gesprochen wird.

Texterzeugung am Beispiel kuschitischer Narrativtexte

Schema 6 Prädikationen in ihrer Sprachform



3. Der Textkern

3.1 Zum Status des Textkerns

Es ist noch festzustellen, in welcher Weise der Kern einen Narrativtext repräsentiert. Der Kern (Schema 2, Kolumne I) soll die Kenntnis abbilden, die ein Sprecher von einem Narrativtext hat, und die ihn befähigt, diesen Text zu erzählen - allgemeine Sprachkompetenz vorausgesetzt. Die folgenden Überlegungen und Methoden stehen hinter der Identifizierung des Kerns eines Textes.

Aus der Tatsache, daß ein Erzähler eine Erzählung kennt, folgt nicht, daß dieser Text in der tatsächlich vorgeführten Form erinnert worden ist. Die Abbildung eines Textes im Gedächtnis des Sprechers muß ja nicht den ganzen Text direkt und unverkürzt enthalten, wenn der Sprecher nachweislich über Automatismen verfügt, die mühelos jede tatsächliche morphologische und phonologische Ausformung übernehmen kann. Aus der Art, wie ein Text bei Wiederholungen variiert (Wedekind 1990: 322-347), läßt sich zum Beispiel schließen, daß das Gedächtnis eines Erzählers nicht jede Wendung syntaktisch und phonologisch "fertig" zum Abruf bereitgehalten hat. Was tatsächlich gespeichert und bereitgehalten wird, muß also eine solche Form des Textes sein, die weitgehende syntaktische und phonologische Ausformungen erlaubt oder erfordert.

Für bestimmte Teile eines Textes gilt es allerdings, daß sie nicht neu aufgebaut werden müssen. Sie werden vielmehr in einer phonologisch fertigen Form aus dem Gedächtnis abgelesen, als wären sie das Echo einer früheren Performanz. Solche Versatzstücke (1.1 in Schema 1) treten bei jeder Wiederholung in der gleichen Form auf. Ihr Platz und ihre Funktion in einem Text ist weitgehend voraussagbar: Mit großer Wahrscheinlichkeit finden sie sich an den Höhepunkten einer tradierten Erzählung, oft sind sie in direkter Rede gehalten; viele sind metrisch oder lautmalerisch ausgeformt und in dieser prägnanten Form leicht memorierbar (Wedekind 1990: 338-339).

Dieser streng ausgeformte, unter der Umgebung von Syntax und Morphologie abrufbare Status solcher Versatzstücke kann aber nicht für alle Teile aller erinnerten Texte zutreffen. Das läßt sich schon daraus schließen, daß verschiedene Passagen beim Vortrag auf verschiedene Weise variieren - manche stark und manche gar nicht - und daß sie auch verschiedene Spuren strategischer Aufbereitung zeigen: Verschiedene Häufigkeiten von Versprechern, Fehlstarts oder Pausen (Wedekind 1990: 392-396). Ganz deutlich liegen also nicht alle Passagen eines Narrativtexte in gleicher Weise phonologisch fertig zur Performanz bereit. Was aber nicht fertig abrufbar vorgelegen hat und trotzdem zum Teil eines Narrativtextes geworden ist, muß aus einer verkürzt oder minimal gegebenen Repräsentierung aufbereitet worden sein.

In welcher Weise ein Narrativinhalt dem Sprecher vorliegt - wenn nicht in seiner vollständigen phonologischen Form - das steht nicht fest. Aber die Narrativinhalte sollen hier in einer solchen Form dargestellt werden, daß ihre Darstellung (*representation*) folgende Bedingungen erfüllt:

- (a) bei psycholinguistisch bereits belegten Aspekten der Texterzeugung sollen diese Erkenntnisse integriert werden, sofern sie direkt relevant sind;
- (b) darüberhinaus sollen mögliche Widersprüche zu solchen Erkenntnissen vermieden werden;

- (c) bei psycholinguistisch nicht belegten Aspekten der Texterzeugung sollen Erkenntnisse integriert werden, die aus der maschinellen Texterzeugung belegt sind;
- (d) als linguistischer Rahmen und als Schreibweise soll weitgehend das gelten, was die Funktionale Grammatik bereitstellt (Dik 1978 und 1989): FG-Zielsetzungen liegen in der Richtung der gegenwärtigen Untersuchung.

Welche Wirklichkeit ein solcher Textkern hat, und welcher ontologische Status diesem Kern und der darauf bezogenen Textstrategie zuerkannt wird, muß hier in mehreren Punkten offen bleiben - besonders was die Beziehungen zwischen allgemeinen mentalen Vorgängen und sprachlich-syntaktischen angeht (Vgl. z.B. die extreme Position J.A. Fodors, alle zu einer Aussage führenden Überlegungen seien "syntaktischer" Art; 1983: 40, zitiert nach ter Meulen 1988: 435.).

Nachweisbar ist einerseits die Auswirkung eines - wie auch immer zu repräsentierenden - Textwissens ("Repertoires"), und logisch notwendig jedenfalls ist andererseits die Reduktion der ausgeführten ("performierten") Texte auf weniger extensive Repräsentationen, welche sich durchgängig auf sprachliche Kompetenz bezieht. Die Konzepte Textkern und Textstrategie werden hier an psycholinguistisch harten Fakten und an texttheoretisch schmieg samen Konzepten geprüft und bis auf Widerlegungen im einzelnen als heuristische Werkzeuge eingesetzt.

3.2 Analytische und synthetische Sicht

Für die Darstellung von Texten wurde hier also die Synthese gewählt, nicht die Analyse - obwohl die Darstellung auf den vorausgegangenen Analysen mehrerer Textsammlungen beruht. (Vor allem ist dies eine Sammlung von etwa vierzig Gedeo-Texten; dazu kommen kleinere Textsammlungen in den Schwester sprachen Burji und Sidamo.)

Es gibt mehrere Gründe für die synthetische Darstellung. Im Gegensatz zur analytischen Darstellung eines Textes ist seine Synthese ein Prozeß, bei dem die Komplexität anfänglich gering ist und dann schrittweise und kontrolliert zunimmt. (Das gilt besonders für eine additive Synthese, aber auch für die hier vorgelegte "entfaltende" Synthese.) Sie ist deshalb besser nachvollziehbar als eine Analyse - und sie ist ihrem Gegenstand auch angemessener: dem Erzählen. Schließlich ist das Erzählen ein Vorgang, der für den Zuhörer eine Welt aufbaut und entfaltet. Ein weiterer Grund für die Wahl der synthetischen Darstellung ist die Nähe zur maschinellen "Texterzeugung" (*text generation*). In den letzten Jahren ist Texterzeugung in mehreren Sprachen maschinell praktiziert worden. Zwar sind die bisher behandelten Sprachen reich analysierte Sprachen, und die Texte sind zumeist kürzer und aus einem anderen *genre* als die Narrativtexte dieser Studie - jedenfalls liegen aber Erfahrungen vor, die zu Vergleichen herausfordern.

3.3 Eingrenzung des Textkerns

Die methodischen Schritte und Überlegungen, die von der Analyse einer Textsammlung zur Definition eines Textkerns führen, sehen folgendermaßen aus:

Die Texte der gegenwärtigen Untersuchung wurden zunächst alles dessen entblößt, was durch irgendwelche erkennbaren Bedingungen "voraussagbar" war. Für jedes vielleicht zu streichende Morphem wurde die Bedingung gesucht, die es vielleicht redundant machte, und für jedes genuin voraussagbare Phänomen wurde die Regel formuliert, durch die es wirklich definiert war. Erst der (praktisch) nicht mehr reduzierbare Kern textlicher Informationen galt als die "Minimalform" oder der "Kern" eines Textes - nach seinem Ausmaß ein Bruchteil der entfalteten Form (vgl. Wedekind 1990: 164-165; 147-180; 323-336). Auf der Suche nach einer angemessenen Darstellung des Kerns wurden die gemeinsamen Charakteristika und Regeln solcher Minimalformen über die analysierten Texte und Sprachen hinweg verglichen.

Notwendigerweise muß eine so definierte Minimalform eines Textes das enthalten, was ein Sprecher "weiß", wenn er einen Text "kennt". Es fehlt darin nichts, was die Kenntnis dieses Textes konstituiert. (Und in der Sammlung der Regeln fehlt nichts, was für die rhetorische und linguistische Entfaltung des ganzen Textes unabdingbar wäre.)

Im Idealfall repräsentiert der Textkern also genau das, was ein Erzähler hinzulernt, wenn er eine neue Erzählung in sein Repertoire aufnimmt.

Texterzeugung am Beispiel kuschitischer Narrativtexte

Möglicherweise aber enthält die Minimalform mehr als das. Es ist sogar sehr wahrscheinlich, daß sie mehr enthält: ein so entwickelter Textkern wird einen Rest an kulturell vorgegebenen "Voraussagbarkeiten" enthalten, oder einen Rest an Formeln und Formulierungen, die von ähnlichen Texten her als Erzählroutine voraussagbar wären. Weiterhin ist einiges aus morphologischer Redundanz voraussagbar und offensichtlich gibt es dazu noch die phonologische Redundanz lexikalischer Einträge, die universale Redundanz der ikonischen Abbildung kognitiver Vorgänge, oder die kulturell gegebene Voraussagbarkeit gewisser Handlungsketten oder Charakterkonstellationen. Aber die Streichung aller solcher Aspekte war entweder nicht im Interesse der vorliegenden Darstellung oder einfach unmöglich: Die Streichung aller phonologischen Redundanz im Lexikon z.B. könnte die Darstellung weitgehend unlesbar machen; die Streichung (oder Explikation) aller kulturell voraussagbaren Konstellationen wäre für einen Außenstehenden letztlich unmöglich.

Die Reduktion jedes Textes auf sein Skelett sollte möglichst nicht dazu führen, daß abstrakte oder zugrundeliegende Formen postuliert werden, die "nicht nachweislich" vorkommen. strikt ausgeschlossen sind jedenfalls Formen, die "nachweislich nicht" vorkommen. Was also in der betreffenden Sprache nie gesagt oder gedacht wird, soll auch nicht im Teil der Repräsentierung eines Textkerns postuliert werden. Das entspricht den Forderungen der Funktionalen Grammatik (Dik 1980b: 50).

3.4 Darstellung des Textkerns

Der semantische Kern ist das, "was" erzählt wird, ohne das grammatische "Wie" - vgl. die Darstellung in Kolumne I, Schema 2. Alle Einträge der Kolumne I zusammengenommen stellen, auf ein nicht weiter reduzierbares Minimum gebracht, das dar, was hier als die Kenntnis einer Erzählung verstanden wird.

Der Kern baut sich auf wie folgt. Für alle Inhalte der Szenerie, der Partizipanten und der Ereignisse (2.3.1-2.3.4 im Schema 2) werden die Konzepte aus dem Lexikon abgerufen, bis alles in die Repräsentierung einer Szene integriert ist, "was" mitzuteilen ist: Ort oder Zeit, Agens oder Patiens, Handlung oder Zustand. Nachdem diese Schritte zurückgelegt sind, ist eine Prädikation voll spezifiziert - um hier schon einmal die FG-Ausdrucksweise zu verwenden (vgl. dazu Teil 5 unten) - und ihr gesamter Inhalt wird in dieser Weise bereitgehalten.

Was die Schreibweise angeht, verwenden wir die FG-Darstellung von Prädikationen, um dieses erste Stadium eines zu entfaltenden Textes linguistisch zu repräsentieren. Für jedes Stadium der Entfaltung des Textes werden - von Schritt 2.2 bis 3.4 des Schemas 2 - vier Arten von Prädikationen getrennt geschrieben:

- (a) die chronologisch verknüpften Prädikationen der Zeitachse,
- (b) die direkten Reden,
- (c) die argumentativ verknüpften Prädikationen und
- (d) die determininierenden und beschreibenden Prädikationen.

Diese vier Arten von Prädikationen sind ineinander textlich eingebettet: (a) und (b) enthalten alle anderen, während (c) und (d) in die vorher genannten eingebettet sind. Die enge Verbindung der vier Arten von Prädikationen ist in der schematischen Darstellung so angedeutet, daß sie als dicht beieinander liegend abgebildet werden (Schema 2).

3.5 Dimensionen und Entfaltung eines Textkerns

Die Darstellung von Texten in ihrer Gesamtheit verlangt eine Trennung verschiedener Dimensionen. Im einzelnen muß die Darstellung folgende Bedingungen erfüllen:

- (a) Szeneneinheiten müssen in ihren Abgrenzungen erkennbar sein. Grammatisch gesprochen heißt das: die temporalen Satelliten und die Grenzen der Satzgefüge oder Paragraphen müssen sichtbar werden.
- (b) Partizipanten müssen nach ihrem Auf- und Abtreten, nach ihrer Rollenidentität und nach ihrer Bedeutung durch den gesamten Text zu verfolgen sein. Grammatisch gesprochen heißt das: Alle neu eingeführten oder implizit weitergeführten Terme müssen nach ihrem referentiellen Index (x_1/x_2), ihrer semantischen und syntaktischen Funktion (Agens/Subjekt) und nach ihren pragmatischen Funktionen (Topik/Fokus) zu verfolgen sein.

Klaus Wedekind

- (c) Ereignisse müssen in ihrer Zeitachse oder in ihren logischen Beziehungen erkennbar sein. Grammatisch gesprochen heißt das: tempusorientierte Prädikate (Verben) müssen deutlich konsekutiv angeordnet sein, statische Attributionen müssen außerhalb der Zeitachse ihren Termen zugeordnet werden.

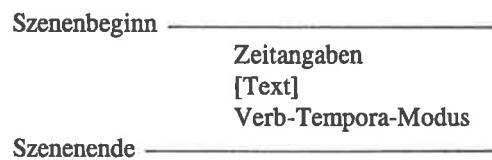
Diese Bedingungen verlangen eine Textpartitur.

Eine solche simultane Darstellung verschiedener Dimensionen wird hier der Kern genannt oder, in Anlehnung an die afro-asiatische Wurzel SFR "Register, Wissen, Rolle, Evidenz", *Semantic Frames Roster* (Wedekind 1990: 1). Die hier verwendete Schreibweise für einen Textkern bezieht sich auf Dik (1978 und 1989), soweit es um einzelne Prädikationen geht, und soweit es um übergreifende Textabläufe geht, folgt sie Grimes (1975). Um darüberhinaus einige der oben genannten Textaspekte zu integrieren, wurde die hier vorgestellte Kern- oder SFR-Darstellung weiter entwickelt. Die Textdimensionen finden sich in der Darstellung wie folgt wieder.

(a) Szeneneinheiten

Die Szenen - grammatisch gesprochen: die komplexen Sätze - sind als Paragraphen durch Linien voneinander abgetrennt. Die Zeitangaben - grammatisch gesprochen: die Tempussatelliten oder Tempus- und Modusangaben am Verb - markieren besonders den Anfang und das Ende einer Szene.

Schema 7 Szeneneinheit



(b) Partizipanten

Alle Partizipanten (grammatisch: Terme, Argumente) sind durch den gesamten Text nach Indexnummern ($x_1/x_2/x_3$) identifiziert. Dabei spiegelt die Indexnummer nicht einfach die Reihenfolge des Auftretens wieder, sondern auch Rang: Typischerweise hat der Held Index 1, der Gegenspieler Index 2, und die widrigen Elemente und Dinge Index 3 usw. Die Verbindung zu folkloristisch oder dramatisch analysierbaren Konstellationen liegt damit auf der Hand (cf. Dundes 1964, Lowe 1969). Außerdem sind die verschiedenen semantischen oder syntaktischen Funktionen auf verschiedene Kolumnen verteilt. Pragmatische Funktionen werden durch Abkürzungen (TOP, FOK) angegeben.

Schema 8 Partizipanten

Agens/Subjekt	Patiens/Objekt
Die Frau x_1	den Wagen x_2
sie TOP x_1	ihn x_2
Er x_2	sie FOK x_1

(c) Ereignisse und andere Prädikationen

Die dramatischen Ereignisse und Zustände sind Prädikationen (grammatisch: verbale und nominale Prädikate). Prädikationen werden bei chronologischer Folge (a) als Zeitachse angeordnet; bei (b) zitierten, (c)

Texterzeugung am Beispiel kuschitischer Narrativtexte

logisch-argumentativ verknüpften oder (d) determinierenden Relationen werden diese Prädikationen außerhalb der Zeitachse gehalten.

Schema 9

Zeitachsenprädikate

Er	x ₁	kam
	x ₁	sah
	x ₁	siegte

Schema 10

Determinierende Präd.

Er ist Cäsar

Schema 11 Relationen

Relationen

regnet deshalb
feucht

(d) Relationen

In SOV-Sprachen sind alle Beziehungen zwischen Propositionen als "Übergänge" zwischen Propositionen markiert. Sie signalisieren Veränderungen betreffend Ort, Zeit, Partizipanten und Ereignisketten; morphologisch sind sie Verbsuffixe.

Schema 12

Propositionen

Übergänge:

[Cäsar x ₁]	kam	-er-danach
	sah	-er-danach
	siegte	-er-Bestätigung

3.6 Die Entfaltung des Textes bis zur phonologischen Form

Schließlich bleibt darzustellen, wie sich schrittweise aus der minimalen Repräsentierung eines Textes die volle phonologische Form entfaltet.

In der Ausgangsphase weist der Textkern nur semantische Rollen auf. Jeder Term hat eine Rolle wie z.B. Agens, Patiens, Benefiziens, oder Richtung. (Die Anordnung als eine Sequenz von Prädikationen unterscheidet diese Darstellung des Textes von Rosenbergs ebenfalls semantisch orientierter *story data base SDB*, Rosenberg 1979: 97). Später, bei der Ausformung oder Entfaltung des Textes werden den Termini syntaktische und pragmatische Rollen zugewiesen. In den meisten Fällen ist diese Zuweisung unauffällig und unmarkiert (fast jedes Agens wird z.B. zum Subjekt). In einigen Fällen aber greifen die Textregeln ein, um einer NP eine abweichende Funktion zuzuweisen; so wird ein Patiens z.B. zum Subjekt. Das ist üblicherweise so, wenn der Textablauf verlangt, daß ein Term als Topik erhalten bleibt, obwohl seine semantische Rolle vom Agens zum Patiens gewechselt hat - wie etwa in einer direkten Rede (Text A, Wedekind 1990), wo es heißt:

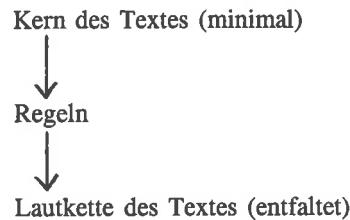
Klaus Wedekind

"Wir (AGENS) tun dies - wir (AGENS TOPIK) tun jenes - und wir (PATIENS TOPIK) werden dafür verachtet!"

Von der ersten, abstrakten Darstellung eines Texts bis zu seiner endgültigen, voll entfalteten Darstellung führt eine begrenzte Anzahl von Regeln. Im Prinzip wird die Anzahl der Regeln offen gehalten; in der analytischen Praxis hat sich allerdings für die vorliegende Textsammlung ein schnelles Konvergieren auf etwa hundert häufige Regeln herausgestellt (s. Teil 7.6.1 unten).

Der Text durchläuft zu seiner Entfaltung verschiedene Arten von Regeln, vornehmlich Textregeln, Regeln der Funktionszuweisung, Regeln der syntaktischen Anordnung und Regeln der Phonologie. Graphisch lässt sich das wie folgt zeigen:

Schema 13



4. Zur Psycholinguistik der Texterzeugung

Was geht in einem Erzähler psycholinguistisch vor, wenn er erzählt? Gibt es Spuren der Schritte, die er zurücklegt, wenn er von der Kenntnis einer Geschichte zu deren Darstellung übergeht? Welche Schritte sind unterscheidbar? Lassen sie sich simulieren? Kann überhaupt von "Schritten" die Rede sein?

Wenn der Kern einer Erzählung sich isolieren lässt und wenn die Schritte eines Erzählers sich identifizieren lassen - ist es dann möglich, einen perfekten *griot* zu simulieren?

4.1 Psycholinguistische Voraussetzungen der Texterzeugung

Wie in den meisten Publikationen zur Texttheorie wird hier vorausgesetzt, daß die Form eines Textes von drei oder vier Hauptfaktoren bestimmt ist: (1) von den Inhalten, die der Erzähler vermitteln möchte, (2) von der Art, wie der Erzähler diese Vermittlung gedanklich steuert, (3) von der Wirkung, die der Sprecher unter den gegebenen Umständen erwartet, (4) und schließlich von den Formen, welche eine Sprache anbietet; vgl. Levy (1979) zu Punkt 1-4 und Garrett (1988) zu 1 und 3-4.

In der vorliegenden Untersuchung geht es um Belege für den Weg von (1) nach (4): Wie gehen (1) die vorgegebenen Inhalte in (4) die vorgegebenen Sprachformen ein? Auf der Basis des (1) inhaltlichen Kerns einer Erzählung sollen auch Voraussagen versucht werden darüber, in welcher Weise dieser Kern unter psycholinguistischen Gesichtspunkten zu einer (4) Erzählung entfaltet wird. Es werden also auch Einsichten in (2) die gedanklichen Vermittlungsvorgänge gesucht, soweit sich deren Spuren in dieser Untersuchung andeuten.

4.2 Vorliegende Untersuchungen

Psycholinguistische Untersuchungen zu diesen Fragen sind nicht zahlreich; wir werden uns hier vor allem auf die zusammenfassende Darstellung durch Garrett beziehen (Garrett 1988).

Besonders die Psycholinguistik der Erzeugung von Texten ist ein Gebiet, über das wenig gearbeitet worden ist - weniger jedenfalls als über die Psycholinguistik der Erzeugung von Sätzen. Für Texterzeugung gilt weitgehend noch, was Fodor im Hinblick auf das Produzieren von Sätzen beklagte: "Es ist ja so wenig, was wir an detaillierterem Wissen über das Produzieren von Sätzen besitzen." (Fodor 1983: 10)

Im Vergleich zu den psycholinguistischen Studien des Verstehens sind Studien des Erzeugens linguistischer Aussagen in der Psycholinguistik erst sehr spät aufgenommen worden. "Unter den drei Hauptgebieten der Psycholinguistik ist das Gebiet der (linguistischen) 'Produzierens' bisher am wenigsten untersucht und vielleicht am wenigsten verstanden worden. Erst seit kurzer Zeit nimmt es in der experimentellen

Texterzeugung am Beispiel kuschitischer Narrativtexte

Psycholinguistik den gleichen Status ein wie das linguistische 'Verstehen'" (Tanenhaus 1988: 25).

Es läge nahe, einschlägige Erkenntnisse zum Gebiet der Texterzeugung (*text generation*) besonders in der TG-Schule oder ihren Nachfolgern zu suchen. Aber gerade hier ist die Psycholinguistik des "Generierens" von Aussagen "im Gegensatz zum Sprachverstehen und zum Spracherwerb" ein vernachlässigtes Gebiet. Diese Schule hat hier "auf die Theoriebildung und Forschung nur einen begrenzten Einfluß ausgeübt" (Tanenhaus 1988: 26).

4.3 Zur Modularität

Zur Modularität der Texterzeugung, oder jedenfalls zur ausschließlich modularen Darstellung dessen, was beim Reden vorgeht, gibt es Untersuchungen sowohl aus den Gebieten der neueren Psycholinguistik als auch aus der maschinellen Texterzeugung.

Die Ergebnisse der psycholinguistischen Untersuchungen von sprachlichen Irrtümern (Versprechern, Pausen, sprachlichen Reparaturen) legen es nahe, den Vorgang des Erzählens als eine Schrittfolge zu verstehen: Es werden verschiedene mentale Bereiche oder "Moduln" einzeln aktiviert, und es gibt Indizien, daß dies in bestimmten Schritten geschieht - sogar in einer zeitlichen Reihenfolge.

Andererseits sind aus dem Gebiet der maschinellen Texterzeugung (*text generation*) Überlegungen veröffentlicht worden, die einer solchen "modularen" Sicht zu widersprechen scheinen. Überlegungen dieser Art sind von Danlos (1987) angestellt worden, die zu dem Schluß kommt, daß es keine absolute Priorität gewisser Moduln geben darf. Es sei nicht denkbar, daß etwa die Informationsanordnung dem Lexikon immer vorzuordnen sei, oder daß der Zugriff zum Lexikon den syntaktischen Entscheidungen immer vorausgehe.

Danlos gibt ein Beispiel, das eine Entscheidung zwischen den Lexikoneinträgen *kill*, *murder* und *assassinate* erzwingt. Diese Entscheidung, sagt Danlos (1987: 197), ist nicht zu treffen ohne Zugang zu (1) pragmatischem Wissen (Verbrechen vs. Unfall), (2) semantischem Wissen (politisches vs. apolitisches Opfer) und (3) syntaktischem Wissen (*x shot y*, *killing him* vs. *x shot y*, *murdering him*). Wenn diese Aussage aufbereitet wird - mental oder maschinell - dann ist es nach Danlos also undenkbar, daß der Lexikoneintrag *assassinate* erst ausgewählt wird, nachdem die mentale Konstruktion fertig vorliegt, aber die syntaktische Verwendung noch nicht.

Die oft vorausgesetzte Schrittfolge 1. Idee, 2. Lexikon, 3. Syntax wird also von Danlos bestritten - und ebenso deren Umkehrung:

*Anordnung der Ideen -> Wahl der lexikalischen Konzepte -> syntaktische Form

*Wahl der lexikalischen Konzepte -> Anordnung der Ideen -> syntaktische Form

Sie plädiert für die Interdependenz der verschiedenen Arten von Entscheidungen. Für Danlos scheint diese Interdependenz überhaupt jede Modularität in Frage zu stellen (Danlos 1987: 96). Die gleiche Art von Interdependenz gilt nach Danlos auch im Bereich ganzer Texte: Die Sequenzierung der Information, die Aufteilung auf einen oder mehrere Sätze und die Form dieser Sätze sind keine voneinander isolierten Schritte (Danlos 1987: 95).

Die Frage ist, ob diese Ergebnisse aus der maschinellen Texterzeugung den Ergebnissen der Psycholinguistik so widersprechen, daß "Interaktion" tatsächlich "Modularität" ausschließt. Es scheint uns vielmehr denkbar, daß die mentale Vorbereitung einer Aussage "modular" an Bereiche und Schrittfolgen gebunden ist, und daß die tatsächlichen Entscheidungen für eine linguistische Aussage trotzdem, wie von Danlos gefordert, durch (schnelle) Folgen von interaktiven, tastenden Versuchen erreicht werden können.

Um beim Beispiel zu bleiben, das Danlos verwendet: Die Absicht, eine Aussage zu machen, die den Begriff *assassinate* verwendet, kann durchaus im Modul "Idee" mit einer Idee beginnen, die das Konzept *assassinate* voraussetzt - und dann daran scheitern, daß der Lexikoneintrag selbst dem Sprecher zwar "auf der Zunge" liegt, aber von seinem Lexikon her - phonologisch orientiert - nicht erfüllt wird. Diese Schwierigkeit, daß das Wort bekannt ist, aber nicht über die Zungenspitze kommt (vgl. die Untersuchungen zum Phänomen *tip of the tongue*) muß aber den Sprecher nicht hindern, mit einer nun semantisch orientierten Suche im Lexikon noch einmal neu anzusetzen (also Interaktion trotz Modularität). Psycholinguistisch ist jedenfalls nachgewiesen, daß ein Lexikon sowohl von der phonologischen als auch von der semantischen Komponente her zugänglich ist (Emmorey & Fromkin 1988: 145-146). Nach einem solchen teils erfolglosen "Hin und Her" zwischen Idee,

Lexikon und syntaktischem Phrasenaufbau wird der Sprecher seine Idee schließlich doch ohne großen Zeitverlust in eine dem ersten Entwurf angemessene lexikalische und syntaktische Form bringen.

Ein Prozeß dieser Art, der erfolglose Strecken, Irrtümer und sich verweigernde Modulen zuläßt, mag für maschinelle Texterzeugung ausgeschlossen sein - der Erfahrung eines Erzählers widerspricht so etwas nicht.

4.4 Psycholinguistische Fakten zur Texterzeugung

In dieser Studie sollen die Annahmen über den Vorgang des Texterzeugens möglichst psycholinguistisch belegte Vorgänge widerspiegeln; zumindest aber sollen sie den psycholinguistisch belegten Fakten nicht widersprechen.

Es werden deshalb im folgenden relevante Ergebnisse aus der Psycholinguistik herangezogen, um die hier eingenommene Sichtweise daran zu überprüfen. Die einschlägigen Ergebnisse stammen vorwiegend aus den Untersuchungen von Versprechern beim Erzeugen von Sätzen (*error analysis*), aber auch aus anderen Bereichen, wie zum Beispiel dem des Verstehens von Aussagen (Garrett 1975, Tanenbaum 1988). Sie werden im folgenden skizziert, und zwar in der modularen Reihenfolge, die dem Ablauf des Texterzeugens entspricht.

4.5 Garretts Modell

Zuerst soll als Hintergrund für die darauffolgende Diskussion das von Garrett entwickelte modulare Modell vorgestellt werden.

In Garretts psycholinguistischem Modell geht es zwar um die Erzeugung von Sätzen. Garrett nimmt dabei aber ausdrücklich an, daß Sätze, die in Texte eingebettet sind, keine grundlegend anderen Überlegungen erforderlich machen als isolierte Sätze (Garrett 1988: 72).

Garrett führt von abstrakten Lexikoneinträgen zur tatsächlichen Aussprache von Phrasen und Sätzen (Garrett 1975, modifiziert 1988). Wir werden davon ausgehen, daß die hier angenommene Modularität (1988: 88) durch Überlegungen wie die von Danlos nicht widerlegt ist, und werden im folgenden dieses Modell weitgehend voraussetzen, soweit es um Sätze geht.

1. Das, "was" zu sagen ist, ist bei Garrett das Konstrukt "M" (message).

Nach unserer Sichtweise entspricht "M" einer Prädikation des Textkerns.

2. In einem ersten Modul (*functional level*) werden Aussagen der gedanklichen Größenordnung von ein bis zwei Sätzen geplant. Garrett nennt dies *multi-phrase syntactic planning*.

Nach unserer Sichtweise entspricht dieser Schritt dem Planen eines Verbrahmens samt allen semantisch definierten Termen wie Agens, Patiens usw. Der erste Zugriff auf das Lexikon ist nur semantisch orientiert; es wird hier noch keine phonologische Ausformung erwartet. Zu beachten ist, daß nur wenige Phrasen bearbeitet werden - was nach unserer Beobachtung einer Prädikation eines normal voranschreitenden Textes entspricht.

3. In einem weiteren Modul werden Segmente und Prosodien einbezogen. Aus dem Lexikon ist der vollständige, phonologisch ausgeformte Eintrag abgelesen worden.

Nach unserer Sichtweise entspricht dieser Schritt der ersten, syntaktischen und morphologischen Ausformung einer voll spezifizierten Prädikation.

4. Im nächsten Modul werden alle phonologischen Details eingefügt.

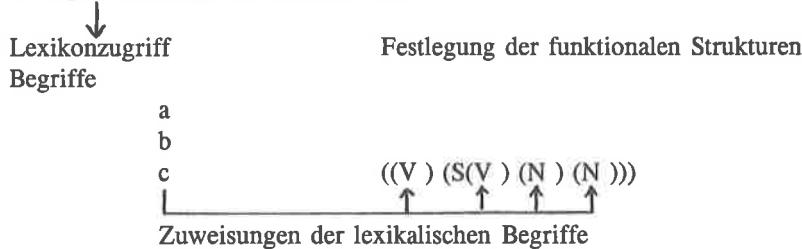
Nach unserer Sichtweise entspricht das den letzten, morphophonologischen Schritten der Ausformung einer Prädikation - der Anwendung von FG-Ausdrucksregeln.

Texterzeugung am Beispiel kuschitischer Narrativtexte

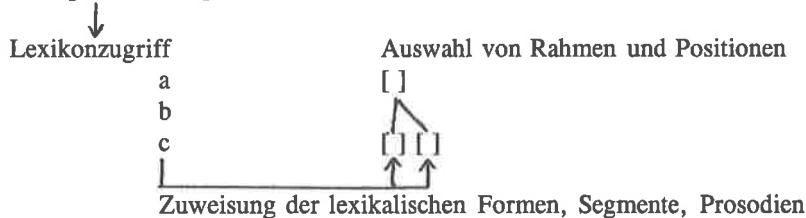
Schema 14

Garretts Modell im Überblick

1. Repräsentierung der *Message* (M)



2. Repräsentierung nach Funktion (F)



3. Repräsentierung nach Position (P)

4.6 Aufbau von Texten

Im Blick auf ganze Texte - mehr als im Blick auf Phrasen und Sätze - ist es berechtigt und hilfreich, nach kommunikativen "Zielen" eines Sprechers zu suchen (Levy 1979). Die Wahl von syntaktischen Mustern ist zum Beispiel aufgrund vorher getroffener strategischer Entscheidungen des Sprechers zu erklären: Ein Sprecher "drückt aus", "spezifiziert", und "strebt Ziele an" (Levy 1979: 204, 207-209). Diese Ziele steuern den Aufbau einzelner Sätze, und in dieser Hinsicht läuft die Textstrategie jedem Satz voraus (vgl. Schema 1, 1.1).

Die im Hinblick auf Text und Paragraphen "gesteuerten" Vorgänge sind deutlich langsamer als z.B. die Automatismen der morphologischen Ausformung. Sie sind auch der inspizierenden Reflexion eher zugänglich und können getrennt von den Automatismen der Sprachkompetenz betrachtet werden. Schwierigkeiten in diesen Vorgängen sind durch besondere Arten des Zögerns und Pausierens angezeigt. Wenn Schwierigkeit und Planungslast zunehmen, wachsen nachweislich auch die Pausen und Verzögerungen, die zur strategischen Planung eines Textes eingesetzt werden müssen. Allgemeine kognitive Schwierigkeiten und Schwierigkeiten der Struktur verursachen Zögern (Garrett 1988: 79).

In unserer Untersuchung (Gedeo-Text B, Wedekind 1990) treten Pausen des Diskurs- "Steuerns" zum Beispiel besonders bei Szenenübergängen auf. Chafe arbeitet mit den Konzepten "Selbst", "Bewußtsein" und "Information", um die Prozesse zu beschreiben, die er beim vorbereitenden "Denken" von Texten beobachtet (Chafe 1980: 11).

Die Textstrategie bringt "höher angeordnete Entscheidungen" mit sich; sie greift, nach Garrett, nicht nur an wenigen kritischen Punkten - zum Beispiel bei Paragraphenanfängen - in den Ablauf des Erzählens ein - vielmehr lässt sich feststellen, daß expressive Variationen wie zum Beispiel die des Überredens, Unterhaltens usw., unmittelbar die phonologischen Ergebnisse steuern können - zum Beispiel die Intonation (Garrett 1988: 88). Wir gehen davon aus, daß der Text als übergreifende Einheit immer wieder die Ausformung der einzelnen Szenen mitbestimmt.

Stilistische Varianten werden so erklärt, daß ein Sprecher den "gleichen" Inhalt ohne Mühe doppelt aufbereiten kann. Im Phänomen des irrtümlichen Überblendens zweier Versionen gibt es recht klare Belege dafür, daß zwei verschiedene, phonologisch voll ausgeformte Versionen des gleichen Inhalts bis kurz vor den Zeitpunkt der tatsächlichen Aussprache bereithalten werden können (Garrett 1988: 92). Vgl. dazu die vier Varianten des Gedeo-Textes A unserer Sammlung.

Zur Menge dessen, was ein Sprecher zu einem bestimmten Zeitpunkt des Erzählens bereithält, gibt es verschiedene Aussagen. Nach Givón füllt ein Abschnitt von etwa 20 Sätzen das gesamte zu einem gegebenen Zeitpunkt "aktive" Gedächtnis (Givón 1983: 245). Was aber vorausgeplant wird, hat nach den vorliegenden psycholinguistischen Ergebnissen wohl nur das Ausmaß einiger weniger Phrasen. Daß die übergreifenden Strukturen immer wieder für ein paar Phrasen vorausgedacht werden müssen, ist für den Bereich der Syntax schon durch Phänomene wie die grammatische Übereinstimmung belegt (Tanenhaus 1988: 26). Das Ausmaß vorbereiteter Texteinheiten korreliert, nach Chafe, mit den verschiedenen Kapazitäten verschiedener Regionen des Gedächtnisses. Das bei der Textherstellung arbeitende "Bewußtsein" gruppiert, nach Chafe (1980: 26), alle zu bearbeitende Information jeweils um ein "Zentrum" herum, und es geht dabei um Informationsmengen vom Ausmaß eines komplexen Satzes. In unserer Sichtweise wird der Einheit einer "Szene" - syntaktisch als komplexer Satz oder Paragraph identifiziert - ein besonderer Status zugeschrieben.

4.7 Zur internen Organisation des mentalen Lexikons

Das interne Lexikon ist nach verschiedenen Wortklassen, nach Bedeutungsarten und nach phonologischen Formen auf verschiedene Weise zugänglich. Untersuchungen von Sprechern zeigen zum Beispiel, daß beim irrtümlichen Ersetzen von Wörtern die Wortklasse erhalten bleibt. Außerdem sind Wörter großer Texthäufigkeit im mentalen Lexikon wahrscheinlich zusammengefaßt und sind mit Sicherheit anders verfügbar, als es die selteneren Einträge sind. Morpheme der großen, offenen Klassen (Verben, Substantive) werden nicht mit den Funktoren vermischt gespeichert, sondern sie werden, wie sich aus den vorliegenden psycholinguistischen Daten schließen läßt, in verschiedenen Teilen des Gedächtnisses oder auf verschiedene Art bereitgehalten. Verschiedene Arten des Umgangs mit verschiedenen Arten von Morphemen lassen sich also in psycholinguistischen Analysen deutlich unterscheiden (Emmorey & Fromkin 1988: 145-146). Wir versuchen dieser Trennung gerecht zu werden, indem wir verschiedene Wortklassen im Lexikon getrennt aufführen, und indem wir Funktoren nur in den Ausdrucksregeln zulassen.

Was die Beziehung des Lexikons zur Morphologie angeht (vgl. unsere früheren Beobachtungen zu deren Durchlässigkeit), ist nach Emmorey und Fromkin (1988: 145-146) folgendes gesichert: Morphologisch verwandte Wörter werden enger beieinander gespeichert als Wörter, denen keine morphologische Beziehung zugekannt wurde. Morphophonologisch veränderliche Wörter werden anders bereitgehalten als solche, deren Form konstant ist. Derivationen werden nicht einfach in der Form von Wurzeln erinnert, auf die dann noch die Derivationsregeln angewendet werden müssen; psycholinguistische Untersuchungen legen es vielmehr nahe, daß zumindest einige der derivierten Wörter - die häufigeren - als fertige Einträge gespeichert sind (Emmorey & Fromkin 1988). Das entspricht unseren Beobachtungen über derivierte Verben, die von Sprechern der HEC-Sprachen nicht als Ableitungen empfunden werden: es gibt eine ansteigende Skala von (morphologisch gleichen) Verben, die - in Relation zu ihrer Häufigkeit - zunehmend als monomorphematisch empfunden werden.

Erstarnte Wortfolge, Häufigkeit und Greifbarkeit einer Phrase spielen bei der Auswahl von Textpassagen eine beschleunigende Rolle (Garrett 1988: 80). Es wird hier angenommen, daß nicht jede Passage eines Textes vom Lexikon her in der gleichen Weise ganz neu entwickelt werden muß; vgl. unsere Unterscheidung von Texten, "Versatzstücken", Idiomen, Wörtern und Morphemen (Schema 1, 1.1).

4.8 Das Entfalten einer Prädikation

Das, "was" gesagt werden soll, läßt sich nach Garrett als "Message" (M) vom allgemeinen Wissen eines Sprechers isolieren. Die Integration jedes zu entwickelnden Satzes wird von M gesteuert (Garrett 1988: 70). Wir versuchen hier, mit dem Konzept Textkern das zu identifizieren, was einen Text integriert (Schema 1).

Ein lexikalischer Eintrag wird mit Hilfe konzeptueller Faktoren gewählt und mit Rollen wie Agens oder Patiens gekennzeichnet, um sich entsprechend den aktivierten Einträgen mit dem Verbrahmen zu verbinden. Je höher die Aktivierung, um so schneller wird der lexikalische Eintrag mit der Oberflächenrolle in der Struktur verbunden (Garrett 1988: 79). Bei den Mechanismen, die ein Topik "erhalten" oder einen Fokus bewirken, setzen wir diese Vorgänge voraus (vgl. Schema 2 oben).

Die Zeit, die es braucht, bis ein Sprecher zu einer völlig neuen Äußerung ansetzt, erlaubt - nach Bock und Garrett - Rückschlüsse darauf, daß die Prozesse des konzeptuellen und syntaktischen Sequenzierens und die des Aufsuchens von Wortformen verschiedene Prozesse sind (Garrett zitiert hierzu unveröffentlichte Beiträge

Texterzeugung am Beispiel kuschitischer Narrativtexte

von Bock, 1988: 79). Der Zugang zum Lexikon und die Ausformung einer Prädikation sind verbundene aber verschiedene Vorgänge. Kurz vor dem Umschalten auf eine alternative Satzform gibt es nachweislich Momente des Zögerns. Eine solche alternative Satzform kann notwendig werden, wenn eine vorbereitete Struktur wegen des Fehlens eines lexikalischen Eintrags schnell durch eine andere ersetzt werden muß (Garrett 1988: 90; vgl. besonders die Diskussion der Einwände von Danlos im vorausgehenden Teil). Wenn ein Sprecher zu einem neuen Topik und zu einer neuen Szene ansetzt, braucht er, nach unseren Untersuchungen, oft Pausen der strategischen Vorbereitung (Gedeo-Text B).

Was verschiedene Bedeutungen und phonologische Formen angeht, so ist schon darauf hingewiesen worden, daß nach den vorliegenden Ergebnissen das Lexikon sozusagen sowohl von links als auch von rechts zugänglich ist: von der Seite der Bedeutung her und ebenso von der Seite der phonologischen Form. Zum Beispiel können Wörter genauso nach ihrer ersten Silbe oder ihrer Länge erinnert werden, wie auch nach ihrer Bedeutung (Garrett zitiert hierzu Butterworth 1979, Fromkin 1971, Garrett 1978 und Emmorey & Fromkin 1988: 145-146). Zugang zu verschiedenen lexikalischen Formen und Bedeutungen verschafft sich ein Sprecher also auf verschiedenen Wegen. Das ist nicht nur für die Suche nach lexikalischen Einträgen nachgewiesen, sondern interessanterweise auch für den Aufbau syntaktischer Formen (Garrett 1988: 74). Es gibt also einen wiederholten Zugriff zum Lexikon (vgl. Schema 6 des vorausgehenden Teils).

4.9 Syntaktische Ausformung

Psycholinguistisch ist es nachweisbar, daß die sequentielle Anordnung syntaktischer Komponenten unabhängig vom Zugriff zum Lexikon geschieht. In unserer Sichtweise folgt die Ausformung eines Satzes erst nach dem Zugriff zum Lexikon und nach der vollen Spezifizierung einer Prädikation. In unserer Darstellung kopiert eine Prädikation einen Teil von dem, was vorausgeht. Daß das aktive Gedächtnis die Wahl der syntaktischen Form mitbestimmt, läßt sich schon daraus ablesen, daß zum Beispiel in Frage- Antwort-Gesprächen die Syntax der Frage von der Antwort oft übernommen wird. Ähnliche Entsprechungen sind in Narrativtexten häufig.

Irrtümliche Verwechslungen von Wörtern und Lauten sind strikt auf verschiedene linguistische Einheiten begrenzt. Daraus ergeben sich die folgenden Schlüsse: Es muß ein Stadium der Satzaufbereitung geben, zu dem einerseits zwar schon die syntaktischen Beziehungen zwischen Wörtern bestehen, aber andererseits die phonologische Repräsentierung noch nicht. Diese Stufe der Satzaufbereitung ist eine Stufe, auf der die fertig aufbereitete Menge zwar mehr als eine Phrase, aber nicht mehr als zwei Sätze umfaßt: Wortvertauschungen überspringen maximal die Grenze zum Nachbarsatz, mehr nicht (Garrett 1988: 76). Das entspricht unserer Trennung von Ausdrucksregeln der Syntax von Ausdrucksregeln der Phonologie. Sogar das Aufsuchen von Wörtern im mentalen Lexikon geht in zwei Stufen vor sich (Fromkin 1971, Butterworth 1980, Garrett 1978): Information über die Form eines Wortes ist bei der anfänglichen Auswahl aus dem Lexikon noch nicht vorhanden. Ein erneuter Zugriff ist nötig zu dem Zeitpunkt, wenn die segmentale Struktur spezifiziert wird (Garrett 1988: 73).

Syntaktische Phrasen werden zunächst ohne Belastung durch segmentales Gepäck gewählt und vorbereitet. Garrett (1988: 81) trägt Belege dafür zusammen, daß das Einsetzen konkreter phonologischer Formen in einem zweiten Schritt erfolgt. Vgl. unsere Trennung von Kern und dessen grammatischer Ausformung.

Was die Gestaltung von Topik und Subjekt angeht, ist - wiederum durch Tests - belegt, daß für ein bereits genanntes Nomen die Wahrscheinlichkeit steigt, in einem später folgenden Satz zum Subjekt und Topik erhoben zu werden (Garrett 1988, vgl. auch 1975). In jeder normal voranschreitenden Kommunikation wird die neue Information dem laufenden Text in der Größenordnung von einem Konzept pro Satz beigegeben (Givón: *one chunk per clause*). Die Art und Größenordnung einer solchen *chunk*-Einheit ist etwa die eines Lexikoneintrags (Givón 1983: 262). Dies scheint auch in normaler mündlicher Kommunikation das Optimum des Fortschreibens zu sein, welches einen Hörer weder über- noch unterfordert (Levy 1979). Dies wird in unserer Darstellung durch weitgehendes Kopieren vorausgehender Prädikationen imitiert, so daß jeweils pro Satz nur ein bis zwei Phrasen wirklich neu hinzutreten.

5. Aus der Funktionalen Grammatik übernommene Konzepte

Die "Funktionale Grammatik" (FG) wurde von Simon Dik zum erstenmal 1978 im Zusammenhang vorgestellt, und sie ist 1989 mit einem Ausblick auf eine Text-FG erweitert worden. Einige grundlegende FG-Konzepte und Sichtweisen sind hier übernommen worden - zum Teil für die Anwendung auf Texte modifiziert, vereinfacht

oder erweitert - und deshalb empfiehlt es sich, das FG-Modell hier wenigstens im Abriß darzustellen, und in einigen ausgewählten Aspekten, die für die vorliegende Untersuchung von Texten relevant werden. Dazu gehören (a) das Lexikon als Fundus dessen, "was" gegeben ist, (b) die pragmatischen Funktionen und deren Platz in Texten, (c) die Rolle von Satelliten in bezug auf Texte, und (d) die Steuerung einiger Zuweisungsregeln durch den Text.

5.1 Abriß der Funktionalen Grammatik

Die FG erkennt als Hauptkomponenten die folgenden Teile an: (a) ein Lexikon mit Termen und Prädikaten, (b) einen Apparat von Regeln, der aus den Lexikon Prädikationen aufbaut und spezifiziert, und (c) schließlich die Ausdrucksregeln, die jede Prädikation zu ihrer tatsächlichen phonologischen Form entwickeln. Etwas weiter ausgeführt bedeutet das folgendes:

- (a) Das Lexikon enthält einen Fundus von Termen und Prädikatsrahmen; lexikalische Regeln bauen daraus einen Fundus von komplexen Termen und komplexen Prädikatsrahmen auf. Die Einträge sind konkret gehalten.
- (b) Eine erste Sequenz von Regeln setzt die Terme in Prädikatsrahmen ein, aber mit Rücksicht auf explizite Bedingungen - zum Beispiel wird ein Argument Patiens oder ein belebter Agens nur von bestimmten Verben akzeptiert.

Die so entwickelten Prädikationen werden weiter spezifiziert, indem Angaben über syntaktische und pragmatische Funktionen integriert werden: Ein Agens zum Beispiel wird syntaktisch zum Subjekt und pragmatisch zum Topik oder zur Coda eines Satzes (Dik 1980b: 46).

- (c) In den Ausdrucksregeln werden dann - auf dem Hintergrund universaler Regeln - alle die sprachspezifischen Prozesse angeordnet, die sich nicht aus den funktional begründeten Annahmen über menschliche Sprachen von selbst ergeben. Hier finden sich also alle Regeln der Syntax, Morphologie und Phonologie, soweit sie unabdingbar sind für den Aufbau der tatsächlichen phonologischen Satzform.

Schema 15

FG: Eine Schematische Darstellung ausgewählter Modulen

(a) Lexikon: Lexikalische Prädikatsrahmen und Terme

(b) Spezifizierende Regeln: Regeln, welche Terme in Prädikaträumen einpassen Regeln, welche syntaktische Funktionen zuweisen (SUBJ OBJ) Regeln, welche pragmatische Funktionen zuweisen (TOP FOC) Semantisch voll spezifizierte Prädikation

(c) Ausdrucksregeln: Ausdrucksregeln der syntaktischen Sequenzierung Ausdrucksregeln der Morphologie Ausdrucksregeln der Phonologie Phonologisch voll ausgeformte Äußerung

Texterzeugung am Beispiel kuschitischer Narrativtexte

Schema 16

Beispiele aus der Entwicklung einer Prädikation

(a) Lexikon

Terme: ‘I’ Pronoun
 ‘John’ Name <belebt>
 ‘book’ Nomen usw.
Verbrahmen: ‘give’ Verb ()_{AGENS} ()_{PATIENS} ()_{REZIPIENT}

(b) Spezifizierende Regeln

Einpassung der Terme in den Verbrahmen:

give (I)_{AGENS} (book)_{PATIENS} (John)_{REZIPIENT}

Zuweisung von syntaktischen und pragmatischen Funktionen, Operanden usw.

[VERGG]give (I)_{AG/SUBJ} [1DET](book)_{PAT/OBJ/COD} (John)_{REZ/FOK}

(c) Ausdrucksregeln

SVO-Sequenzierung, Coda-Parenthese:

(I)_{AG/SUBJ} [VERGG](give) (John)_{REZ/FOK} [1DET](book)_{PAT/OBJ/COD}

Morphologie, Fokusphonologie usw.:

‘I gave it to JOHN, the book’

5.2 Die FG als Texterzeugungsmodell betrachtet

Das FG-Modell versteht sich nicht als ein ausschließlich texterzeugendes oder generierendes Modell. Es hat aber mehrere Untersuchungen gegeben, die diese Perspektive einnehmen; die Tendenz, Sprache synthetisch zu sehen, ist im FG-Modell vielleicht angelegt (vgl. z.B. schon Dik 1980b). Diese Tendenz ist schon in der Anordnung der Komponenten augenfällig. Sie kommt der vorliegenden Untersuchung entgegen, und sie wird hier selektiv betont und genutzt.

5.3 Psychologische Adäquatheit und Funktionale Grammatik

Was das Selbstverständnis der Funktionalen Grammatik zur Frage der psychologischen Adäquatheit angeht, so erhebt FG nicht den Anspruch, mentale Vorgänge des Sprechens und Verstehens abzubilden. Es wird in der FG-Grammatik aber angestrebt, keine Konflikte mit gesicherten psychologischen Erkenntnissen zu erlauben und im Modell festzuschreiben (Dik 1978: 7).

In der vorliegenden Untersuchung wird die Abfolge von Schritten aus dem FG-Modell direkt übernommen, soweit sie nicht aufgrund anderer Überlegungen modifiziert oder expliziert gemacht werden mußte. Das übergeordnete Prinzip bleibt, keine Widersprüche zu festen Einsichten aus der Psycholinguistik zu erlauben.

5.4 Die pragmatischen Funktionen der Funktionalen Grammatik

Die FG-Darstellung von pragmatischen Funktionen wie Thema und Fokus sind hier im Hinblick auf Texte von besonderem Interesse. In FG werden vier solche Funktionen anerkannt; die Thema-, Topik- und Codafunktionen sind wie folgt charakterisiert.

(a) Thema

Der Begriff Thema wird in der FG strikt nur auf solche Nominalphrasen angewendet, die - wie eine Parenthese - durch eine phonologische Pause abgetrennt sind vom darauffolgenden syntaktisch vollständigen Satz. Eine solche Nominalphrase muß aber eine übergreifende semantische Beziehung zu dem nachfolgenden Satz haben - eben die, daß sie die Idee des Satzes thematisch vorausnehmen.

Für die vorliegende Untersuchung ist erstens zu bemerken, daß der FG-Begriff Thema hier insofern erweitert wird, als er auch für ganze Paragraphen gelten soll. Zweitens ist der Begriff Thema nicht scharf von unserem Begriff Szenerie (*setting*) abzugrenzen: Die Voranstellung vor den syntaktisch vollständigen Satz und die Abgrenzung durch Pausen gelten in gleicher Weise für Thema und Szenerie. (Auch eine semantische Eingrenzung von Szenerie auf Ort- und Zeitsatelliten ergäbe keine scharfe Abgrenzung gegen den Begriff Thema.)

(b) Topik

Der Begriff Topik wird hier wie in der Funktionalen Grammatik verstanden - nämlich als der "springende Punkt" einer Aussage; der Teil, über den prädiziert werden soll. Typischerweise ist ein Topik etwas Vorausgesetztes, Bekanntes und fällt mit der Subjektsyntax zusammen.

Allerdings wird im FG-Modell auch eine solche Phrase schon als Topik anerkannt, die im vorausgehenden Text noch keinerlei Bekanntheitsgrad erreicht hatte und deren Topikalität durch die betreffende Nominalphrase erst aufgebaut werden soll.

In unserer Analyse wird diese pragmatische Konstellation als Fokus und noch nicht als Topik identifiziert. Der Grund, warum eine solche erst im Aufbau begriffene Nominalphrase hier als Fokus gilt, ist (a) die noch fehlende Bekanntheit und (b) die Tatsache, daß eine solche Nominalphrase in kuschitischen Sprachen genau so markiert ist wie ein Fokus: Bevor ein Topik etabliert ist, muß es die Aufmerksamkeit erregen, die für eine fokussierte Phrase typisch ist - aber nicht für ein gegebenes Topik. Ein weiterer Grund ist mit (b) verwandt; er gilt ebenfalls für die hier untersuchten kuschitischen Sprachen: Aufgrund der Form, die ein kommendes Topik beim ersten Auftreten noch hat, kann nicht vorausgesagt werden, ob es sich zum Topik etablieren wird. Es ist zu diesem Zeitpunkt noch "im Fokus".

(c) Coda

Coda wird hier ganz in Anlehnung an FG definiert als die "zu spät" kommende Phrase, die einem syntaktisch schon abgeschlossenen Satz noch angehängt wird, als Ergänzung, Korrektur oder Erweiterung. Wie das Thema ist die Coda eine strikt durch Pausen abgetrennte Parenthese - nur folgt sie dem Satz.

5.5 FG-Satelliten, Adverbien und Szenerien

Satelliten in der Funktionalen Grammatik sind die dem Satz beigefügten adverbialen Bestimmungen (*adjuncts*), die dem Verb gar nicht oder nur lose verbunden sind - im Gegensatz zu solchen Adverbien, die den Charakter eines Verbs mitbestimmen. Die Grenzen zwischen beiden sind aber fließend und lassen sich über Sprachfamilien hin schwer festlegen. Es ist zum Beispiel aus der Perspektive der kuschitischen Sprachen schwer nachzuvollziehen, daß in der FG-Literatur der Benefaktiv als Satellit *par excellence* gilt: in kuschitischen Sprachen wird gerade der Benefaktiv bevorzugt ins Verb inkorporiert.

In Erweiterung der FG-Auffassung werden hier Satelliten auch für ganze Paragraphen angenommen: sie funktionieren als die Szenerie (*setting*) für mehrere Sätze bzw. Verben.

5.6 Steuerung der FG-Zuweisungsregeln

Im Unterschied zum bisher vorgelegten FG-Modell, in welchem die verschiedenen Arten von Regeln nicht durch textliche Notwendigkeiten motiviert und aktiviert sind, werden nach unserer Sichtweise eine ganze Anzahl von Regeln vom Text her gesteuert.

Abgesehen von häufig beschriebenen Entscheidungen wie Pronominalisierung oder Subjekt- und

Texterzeugung am Beispiel kuschitischer Narrativtexte

Topikzuweisung, gehören hierher auch die Fokuszuweisung, die Wahl der Aspekte und Tempora und die Inkorporation - besonders häufig als Kausativ oder als Benefaktiv.

6. Anlehnungen an maschinelle Texterzeugung

McKeown stellt für maschinelle Texterzeugung fest (McKeown 1985: 6), daß der Großteil der vorausgegangenen Arbeiten sich mit der Erzeugung einfacher Sätze befaßt. Die Forschungen zur Computeranalyse und Computerinterpretation natürlicher Sprachen seien den Forschungen zur Texterzeugung um viele Jahre voraus (McKeown 1985: 2; Danlos 1987: 1). Die Leitfragen bei der maschinellen Texterzeugung sind aber nach McKeown auch die der Diskursforschung, nämlich: "was" es zu sagen gibt, und "wie" es sich als Text zusammenfügt (McKeown 1985: 3).

Hier ist im Hinblick auf unser Konzept Textkern von besonderem Interesse, was die maschinelle Texterzeugung als Ausgangspunkt einer zu entfaltenden Aussage betrachtet. In den frühen textgenerierenden Ansätzen wurde meist ein inhaltsleeres Ausgangssymbol als gegeben betrachtet, etwa nach Art des S der Satzsyntax. In späteren Arbeiten erhält das "was" wirkliche Inhalte: In Meehans Texterzeugungsmodell ist eine zufällige Folge von Aussagen "gegeben". Entfaltet wird sie durch die Grammatik des *Tale-Spin* (Meehan 1981). In Daveys Modell des Texterzeugens wiederum ist das, "was" gegeben ist, der Bewegungsablauf eines Spieles. Daraus werden alle Aussagen mit Hilfe von Hallidays *systemic grammar* entfaltet (Davey 1978). In McKeowns Modell ist ein "Diskursschema" samt einer Abfolge von "Fokusselektionen" gegeben. Mit einer entsprechenden Syntax werden die daraus folgenden Aussagen entfaltet (McKeown 1985).

In Anlehnung an Erkenntnisse aus der maschinellen Texterzeugung versuchen wir, die Eigenschaften dessen, was vorgegeben ist (*message*) mit Hilfe des Konzeptes Kern näher zu bestimmen.

7. Zur Repräsentativität eines Korpus narrativer Texte

In welchem Sinne ist eine Textsammlung repräsentativ oder authentisch, wenn sie aus Tonbandaufnahmen von Narrativtexten besteht, und inwieweit erlaubt eine solche Textsammlung, daß auf ihrer Basis verallgemeinernde Aussagen gemacht werden? Welche Vorbehalte sind solchen Aussagen voranzustellen?

Diese Fragen werden hier im Hinblick auf verschiedene Faktoren behandelt, welche die vorliegenden Texte charakterisieren; vor allem sind dies Unabhängigkeit, Inklusivität, Häufigkeit von Satztypen, allgemeine Statistik und Authentizität.

7.1 Die relative Unabhängigkeit vom situativen Kontext

Wenige Texte erklären sich ganz und gar selbst. Es kann aber für einen Text schon als Vorteil gelten, wenn er einen großen Teil dessen enthält, was für seine eigene Interpretation nötig ist. Für Narrativtexte gilt das noch am ehesten.

Der Hörer geht zwar auch hier nie ganz durch die gleichen Vorstellungen und Szenen wie der Sprecher, der sie ihm vermittelt: Die Welt des Zuhörers sieht anders aus und mag auch - gemessen an der des Sprechers - sehr unvollständig sein. Aber Narrativtexte sind in dieser Hinsicht der Tendenz nach "vollständig".

Für andere Textarten ist das durchaus nicht der Fall. Der situative Kontext eines zufälligen Zwiegesprächs zum Beispiel ist im Text des Zwiegesprächs selbst nicht noch einmal vorhanden und kann ohne zusätzliches Protokoll später nicht rekonstruiert werden. Für die Erzählungen der vorliegenden Narrativtexte gilt aber, daß die Welten, in denen sie spielen, in Anwesenheit des Sprechers aufgebaut werden. Für alles, was es in einer solchen Welt geben muß, werden vom Sprecher Wörter mitgegeben. Jeder Umstand und jeder Partizipant wird an seinem Platz und zu seiner Zeit ganz explizit mit Hilfe von Wörtern aufgestellt. In dieser Hinsicht sind die Narrativtexte vergleichsweise eigenständig.

7.2 Die Inklusivität von Narrativtexten

Narrativtexte enthalten Passagen anderer Textsorten, oder sogar vollständige Texte eines anderen *genres*. Das ist nicht zufällig der Fall, sondern es gehört zum Typ eines Narrativtextes: In jeder "direkten Rede" öffnet sich nämlich ein Narrativtext, um andere Arten von Texten zu integrieren - nicht nur rekursiv weitere Narrativtexte - sondern im Prinzip Satztypen und Paragraphen jeder Art.

Klaus Wedekind

In der vorliegenden Textsammlung zum Beispiel steht ungefähr jedes dritte Prädikat außerhalb des zentralen narrativen Zeitablaufs. Neben dem eigentlichen Ablauf der Geschichte gibt es also eine große Anzahl nicht narrativer Erklärungen, direkter Reden, argumentativer Satzfolgen. Hier und da steht auch ein prägnanter Spruch.

Mit bezug auf die Diskurstypologie Longacres (1983: 5), welche von einer sehr breiten Basis verschiedener Textdaten abstrahiert ist, kann die Inklusivität von Narrativtexten wie folgt dargestellt werden. Unter der Benutzung von zwei Parametern, nämlich plus/minus Zeitorientierung und plus/minus Agensorientierung ergeben sich für Longacre vier Typen von Texten. Das sind Narrativtexte, vorgangsbeschreibende Texte, verhaltensbeschreibende Texte und Expositionen.

Während nicht-narrative Texte nur selten Passagen eines anderen Texttyps enthalten, gehört es bei Narrativtexten typischerweise dazu, daß sie genau dies tun, nämlich in den direkten Reden.

Schema 17

Integrierte Passagen anderer *genres*

	Narrativ	Prozedural	Verhalten	Exposition
Narrativ	-	j	j	j
Prozedural	?	-	?	?
Verhalten	?	j	-	?
Exposition	?	?	?	-

In der untersuchten Sammlung von Texten der Gedeo- und Burji-Sprachen (Wedekind 1990) finden sich Passagen verschiedener genres eingebettet in den Narrativtext, wie die folgende Tabelle zeigt:

Schema 18

Integrierte Passagen anderer *genres*

	Narrativ	Prozedural	Verhalten	Exposition
Narrativ	Gedeo-Text A Burji-Text A	Gedeo-Text C Burji-Text A	Gedeo-Text A Burji-Text A	Gedeo-Text B Burji-Text C

7.3 Die Dominanz bestimmter Satztypen

Die Mehrzahl der in den Narrativtexten vorkommenden Sätze gehört zum Typ der positiven Aussagen im Perfekt der Vergangenheit, und sie sind "deklarativ".

Das könnte beklagt werden - denn dies ist wieder einmal der Satztyp, der sowieso schon in linguistischen Beschreibungen bevorzugt dargestellt wird. Zwar enthalten die Zitate eine beträchtliche Anzahl von Fragen und Imperativen (wobei nicht vergessen werden darf, daß alle nicht finiten Verben in dieser Hinsicht neutral sind) - aber unter den finiten Verben sind die deklarativen ganz deutlich die Mehrheit. Givón stellt im Blick auf Satztypen fest, daß für die ersten Lebensjahre eines Kindes zwar die Kommunikation im manipulativen (imperativen) und fragenden (interrogativen) Sprachgebrauch besteht; es gibt auch keine Sprache, die ohne diese Satztypen auskommt. Aber für Erwachsene nimmt in jeder Kultur der "deklarative" Satztyp sehr schnell den ersten Platz ein; für Givón ist Information überhaupt der Prototyp von Kommunikation (Givón 1979).

Unter dem Gesichtspunkt der Repräsentativität kann also gesagt werden, daß Aussagen nicht nur in dem hier untersuchten Text-*genre* ein häufiger Satztyp sind - sie sind der zentrale Satztyp erwachsener Kommunikation überhaupt: Erwachsene informieren.

7.4 Authentizität magnetisch aufgenommener Texte

Inwieweit sind die hier untersuchten Texte authentisch?

Die Frage ist insofern berechtigt, als alle Transkriptionen auf Tonbandaufnahmen beruhen. Während ältere, noch ohne Tonbandgerät vorgenommene Transkriptionen wegen der Verfälschungen des verlangsamen Diktats verdächtig waren, sind es Tonbandaufnahmen gerade wegen der Anwesenheit der Maschine.

Bei den hier zugrundeliegenden Aufnahmen waren zwar Gesprächspartner zugegen, die die Voraussetzung dafür schafften, daß die Kommunikation nicht nur in Richtung Maschine laufen mußte - das traditionelle Wiederholen der letzten betonten Silbe durch den Gesprächspartner ist zum Beispiel nicht eliminiert - aber die bloße Sichtbarkeit eines Mikrofons zum Beispiel bleibt ein Faktor, der die gesamte Situation mitbestimmt. Sobald ein Sprecher sich bereiterklärt, eine Bandaufzeichnung zuzulassen, ist die Kommunikationsrichtung und -funktion gegenüber der traditionellen Funktion verändert. Bei Narrativtexten steht das "Mitteilen" immer noch im Vordergrund, aber es muß zugegeben werden, daß zwischen dem Sprecher und dem primären Adressaten eine Maschine steht.

Kleine Arrangements zwecks größerer Natürlichkeit (*creating the context*, wie etwa Goldstein vorschlägt, Goldstein 1964: 128) machen diese Veränderung nicht einfach rückgängig. Für dieses "Problem", wenn es eines ist, gibt es keine andere Lösung, als die Situation zu nehmen, wie sie ist. Für Kommunikationssituationen, die ein Mikrofon einbeziehen, postulieren wir eine neue Adäquatheit. Die so aufgenommenen Texte - sie mögen traditionelle Texte sein oder extemporierte - werden als Texte akzeptiert, die auf ihre Art einer neuen Situation kommunikativ gerecht werden. Authentizität ist an der gelungenen oder nicht gelungenen Einpassung in den situativen Kontext zu messen. Für die Transkription, für die Analyse und für das Verständnis solcher "verfremdeten" Texte heißt das: Bemerkungen wie zum Beispiel am Ende eines Textes *c'ufi!* "stell das Mikrofon ab!" werden nicht gelöscht.

Nach den uns möglichen Vergleichen verhält es sich so, daß die genannten Verfremdungen die Text-"Ränder" betreffen. Den zentralen Teil eines tradierten Narrativtextes tangieren sie kaum.

7.5 Quantität und Repräsentativität eines Textkorpus

In der Untersuchung von Texterzeugung werden fünf Gedeo-Texte im Detail dargestellt, einmal abgesehen von den Varianten des ersten Gedeo-Textes (er wird in vier Varianten dargestellt). Die Gedeo-Texte stammen aus einer Sammlung von etwa 40 analysierten Texten, während die hier und da zum Vergleich herangezogenen Burji- und Sidamo-Texte aus kleineren Sammlungen entnommen sind.

Für Textstudien besteht in der linguistischen Praxis ein großes Hindernis darin, daß es bei Texten um große Mengen von Daten geht - und daß diese Daten möglichst unverkürzt und im Zusammenhang dargestellt werden müssen. Schon aus diesem Grund liegt es nahe, eine Auswahl zu treffen; aber die Frage ist bei jeder Art von Korpus, wann und in welcher Weise ein solcher Korpus schon als repräsentativ gelten kann. Unter dem Gesichtspunkt einer Feldlinguistik mit deskriptiven Zielen zum Beispiel empfiehlt Samarin allgemein, daß eine Datensammlung verschiedene Sprecher, Textsorten und Paradigmen einschließt (Samarin 1967: Kap. 4); ein absolutes Maß, das für jeden denkbaren Fall eine geringste und hinreichende Menge garantiert, kann aber nicht angegeben werden.

Nach dem Prinzip des verminderten Auftretens neuer Ereignisse gibt es natürlich auch für linguistische Phänomene und deren Auftreten in einem Korpus wie dem vorliegenden eine kritische Grenze. Wenn sie erreicht ist, kann ein Korpus durch noch mehr Daten nicht vollkommen werden. (Die ersten sechs Wörter der Swadesh-Liste zum Beispiel enthalten schon alle Gedeo-Vokale, und die ersten vierzig das gesamte Phoneminventar.) Der hier zugrundegelegte Text-Korpus wurde deshalb nach verschiedenen quantifizierbaren Phänomenen hin durchsucht mit der Frage, ob sie im Korpus vollständig, stark oder schwach repräsentiert sind. Eine solche Statistik ist für die Frage der Repräsentativität von Texten zwar nicht schlüssig, soweit es um Regeln geht - aber sie ist auch nicht ohne orientierende Bedeutung.

7.6 Statistik der untersuchten Texte

Die vorliegende Untersuchung von Texterzeugung akzeptiert eine im Prinzip offene Menge von Regeln, von Textregeln bis zur Morphophonemik. Die Frage ist, wie weit - und wohin - ein Ansatz führt, der sozusagen empirisch abwartend immer weitere Regeln aufnimmt, solange sie in neuen Narrativtexten auftreten.

Klaus Wedekind

Es wurde keine prinzipielle Begrenzung gesetzt darüber, welche Arten von Regeln und Parametern möglich sind. Auch die Eingrenzung der Daten auf Narrativtexte stellt keine scharfe Begrenzung der möglicherweise auftretenden sprachlichen Phänomene dar; das wurde oben bei der Charakterisierung der Textsammlung schon vermerkt: Narrativtexte schließen im Prinzip alle anderen Arten von Texten ein. Das Feld ist also mit jedem neuen Text weit offen für neue sprachliche Phänomene.

Die Daten bestehen aus 40 Gedeo-Texten und kleineren Textsammlungen der Sprachen Burji und Sidamo. Nur fünf der Gedeo-Texte sind bisher ausführlich vorgestellt worden: Es sind Narrativtexte, und sie werden hier als Gedeo-Text A, B, C1, C2 und C3 identifiziert. Die Texte der Sprachen Burji und Sidamo sind vergleichbar; sie sind ebenfalls Narrativtexte, und die Sprachen gehören der gleichen Sprachfamilie an (Wedekind 1990: 139-460, 463-649).

Sowohl offene als auch geschlossene Morphemklassen wurden untersucht. Streng genommen sind für diese Sprache nur die Verben und Nomen wirklich offene Klassen. Die Adjektive und Adverbien bilden geschlossene Klassen von wenigen Dutzend Mitgliedern, wenn man von Derivationen absieht.

Das statistische Ergebnis dieses Ansatzes sieht wie folgt aus (vgl. auch die Kurven der nächsten Seiten). Wie zu erwarten, nimmt die Menge neuer Phänomene in einer fallenden Kurve ab: nach einer anfänglich großen Anzahl neuer Phänomene strebt die Kurve zürst schnell - und dann immer langsamer - auf null zu. Für die kleinen geschlossenen Morphemklassen allerdings erreicht die Kurve schon im Bereich von fünf untersuchten Narrativtexten null. Das gilt besonders für einige Arten von Funktoren: nach den ersten zwei oder drei Texten treten keine neuen mehr auf. Für größere geschlossene Klassen streben die Kurven ähnlich auf null zu. Die offenen lexikalischen Wortklassen sind in diesen Texten mit nur knapp hundert Verben und hundert Nomen vertreten. Ihr statistisches Verhalten ist hier nicht repräsentativ und nicht von Interesse. (Eine gewisse "Kontrolle" des allgemeinen Vokabulars war beabsichtigt und wurde durch die Eingrenzung auf wenige Themen herbeigeführt.) Neben den Verben und Nomen gibt es keine anderen offenen Wortklassen.

7.6.1 Ergebnisse im einzelnen

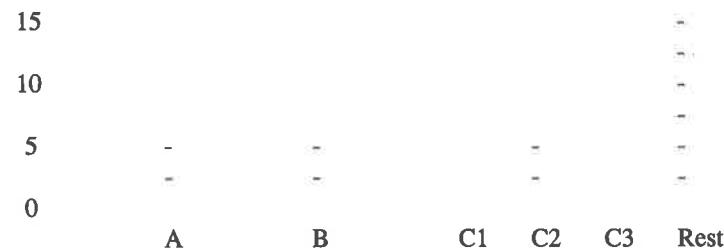
Im einzelnen ergeben sich folgende Ergebnisse. Die Statistiken umfassen geschlossene Klassen und Regeln.

Adverbien und Adjektive:

In erschöpfenden Wortlisten (Wedekind 1978) sind nicht mehr als etwa dreißig Adjektive und ebensoviele Adverbien belegt. In den vorliegenden fünf Texten kommt die Hälfte aller Adverbien vor, und ein Sechstel aller Adjektive.

Schema 19

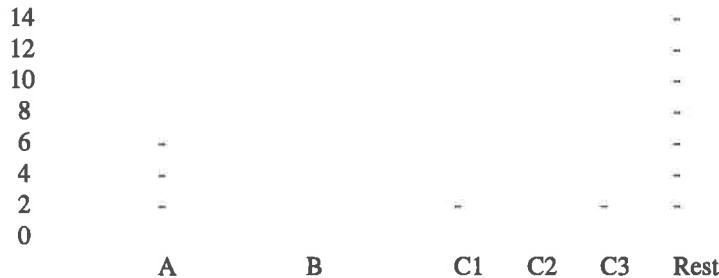
Adverbien (insges. ca. 30)



Texterzeugung am Beispiel kuschitischer Narrativtexte

Schema 20

Adjektive (insges. ca.30)

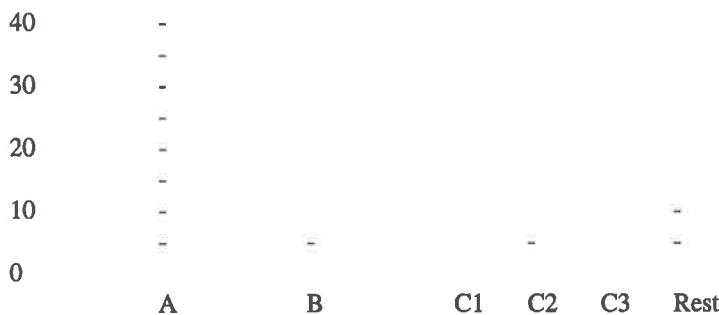


Suffixe:

Die fünf dargestellten Texte enthalten 80% aller Suffixe. Die Gesamtzahl der Gedeo-Suffixe ist aus einer Textkonkordanz von 30.000 Morphemen und aus verschiedenen anderen Untersuchungen bekannt.

Schema 21

Suffixe (insges. ca.75)



Andere Funktoren:

Alle Zeit/Aspektarten, alle Pronomina und alle Postpositionen kommen in den angegebenen fünf Texten mindestens einmal vor. Es gibt keinen Rest.

Klaus Wedekind

Schema 22

Postpositionen (insges. 9)

5						
4						
3	-					
2	-	-		-		
1	-	-	-	-	-	
0	A	B	C1	C2	C3	kein Rest

Schema 23

Pronomina (insges. 9)

5	-					
4	-	-				
3	-	-				
2	-	-	.			
1	-	-				
0	A	B	C1	C2	C3	kein Rest

Schema 24

Zeiten/Aspekte (insges. 8)

7	-					
6	-					
5	-					
4	-					
3	-					
2	-					
1	-	-				
0	A	B	C1	C2	C3	kein Rest

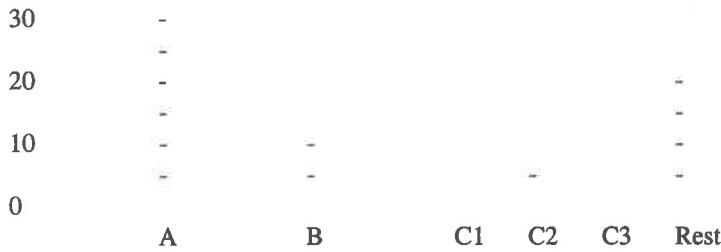
Morphophonologische Regeln:

Ähnlich wie mit geschlossenen Klassen verhält es sich mit den Regeln, die eine geschlossene Klasse betreffen, z.B. den morphophonologischen Regeln. Es gibt insgesamt ca. 60 solcher Regeln, wenn die Analyse sich auf alle Phoneme statt auf alle phonologischen Komponenten bezieht (Wedekind 1985). Die Kurve, die das Auftreten dieser Regeln in den fünf Texten anzeigt, nähert sich Null ähnlich wie die Kurve der Morphemklassen.

Texterzeugung am Beispiel kuschitischer Narrativtexte

Schema 25

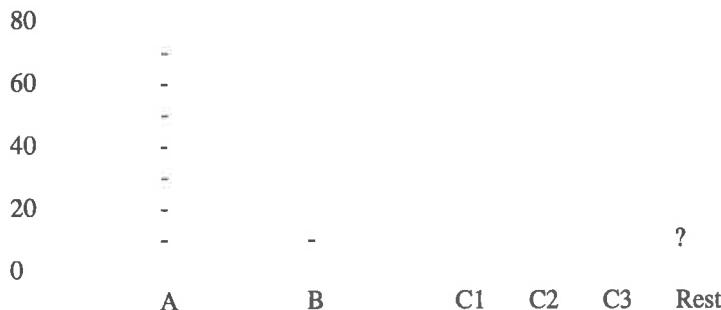
Morphophonologische Regeln (insges. ca. 60)



Die Verteilung neu auftretender grammatischer Regeln kann nicht in der gleichen Weise streng dargestellt werden wie zum Beispiel die Verteilung neu auftretender Funktoren oder phonologischer Regeln. Die hier gegebene Kurve ist also nicht exakt wie die Kurven, die sich auf abzählbare Morpheme beziehen; sie darf nur als Skizze gelesen werden.

Schema 26

Grammatische Regeln (soweit einzeln genannt)



Einer der Gründe dafür, warum die Verteilung der grammatischen Regeln nicht als repräsentativ dargestellt gelten kann, liegt darin, daß beim ersten Auftreten eines neuen grammatischen Phänomens auch solche Aspekte vollständig aufgeführt wurden - zum Beispiel in Paradigmen - die in dem betreffenden Text noch gar nicht aufgetreten waren. Diese Art der Zusammenfassung hat zur Folge, daß die Kurve stärker als repräsentativ abfällt - von 70 auf 7 auf 1.

7.6.2 Zusammenfassung der statistischen Ergebnisse

Zusammenfassend läßt sich aus den quantitativen Beobachtungen zur Repräsentativität der Texte in phonologischer, morphologischer, und syntaktischer Sicht folgendes sagen:

- Die produktivsten Funktoren und Regeln sind schon in den ersten zwei oder drei Texten vorgekommen, und die produktivsten syntaktischen Regeln vermutlich auch.
- Die Morpheme aller kleinen Klassen sind in den ersten fünf Texten vollständig erschienen.
- Die Neigung der entsprechenden Kurven läßt vermuten, daß mit fünf bis zehn weiteren Texten alle alltäglichen morphophonologischen Regeln, alle Funktoren und fast alle syntaktischen Regeln erscheinen könnten.

Aber schon der eine zusätzliche Text, der dieser Untersuchung beigefügt ist (s. Gedeo-Text K in Teil 8), bringt fast keine Erweiterung des Lexikons oder des Regelapparates. Daraus kann geschlossen werden, daß die typischen Aspekte von Narrativtexten der Gedeo-Sprache mit dem vorgelegten Material weitgehend beschrieben sind. Vergleiche mit anderen HEC-Texten tragen zur Festigung der Ergebnisse bei.

8. Beispiel eines Textes

In diesem Abschnitt soll an einem Beispiel illustriert werden, wie ein Text sich zu seinem Kern verhält. Die Wortliste ist unter 8.1 gegeben, und es werden keine anderen als die genannten Regeln angewendet (Wedekind 1990). Die Schreibweise der verschiedenen Aspekte eines Textes ist in den Schemen 7-13 oben dargestellt worden.

8.1 Das lexikalische Material des Textes

Alle Morpheme des Beispieldtextes sind unten im Lexikon dieses Textes aufgeführt. Außerdem sind alle Funktoren dieses Textes aufgeführt; sie werden durch syntaktische oder morphologische Regeln inseriert. Die morphophonologischen Regeln bilden den Schluß dieser Listen.

Ein Asterisk (*) weist die Einträge aus, die noch nicht in den früher vorgestellten Texten aufgetreten sind (Gedeo-Text A, B, C1, C2 und C3, Wedekind 1990). Die Anzahl neuer Einträge ist also gering: Neu für diesen Text sind nur einige Einträge der offenen Klassen (nämlich ein Drittel der Nomina und Verben - darunter viele Lehnwörter - und, wie zu erwarten, alle Namen). In den kleinen und geschlossenen Morphemklassen (Adverbien, Adjektive, Demonstrative, Pronomen, Zahlen) und ebenso bei den Suffixen tritt nichts Neues auf. Ähnliches gilt auch für die morphophonologischen Regeln. Nur eine der vier Regeln ist neu (und selbst diese nicht, wenn man auf phonologische Komponenten statt auf Phoneme sieht).

Schema 27

Wortliste des Textes K

Adverbien, Nomina des Ortes und der Zeit

'okko	"dort"
giddo	"innen"
giddo/i	"das Innere" Acc./Nom.
kaba	"hier"
tecco	"heute"
tecco/i	"der heutige Tag" Acc./Nom.
kayya	"hierher"
kayya/i	"diese Richtung" Acc./Nom.

Adjektiv

lumo	"groß, zahlreich"
------	-------------------

Ausruf

*yoona/i	"na, also"
----------	------------

Demonstrativum

tenne	"diese" (Fem.)
-------	----------------

Texterzeugung am Beispiel kuschitischer Narrativtexte

Namen

*berhaanu	"Personenname" (Amhar. Urspr.)
*dilla/ii	"Ortsname" Acc./Nom.
*soolaama/ii	"Ortsname" Acc./Nom.
*tafarra	"Personenname" (Amhar. Urspr.)
*tasamma	"Personenname" (Amhar. Urspr.)
*zike	"Personenname"
*wondo	"Ortsname"

Nomina

'ora/i	"Weg" Masc. Acc./Nom.
*bariisa	"früher Morgen"
*belto/i	"Kind, Sohn" Masc. Acc./Nom.
belto	"Tochter" Fem.
boga/i	"Platz, Gegend" Masc. Acc./Nom.
hado	"Familie, Verwandte" Fem.
hado/i	"Heim, Zuhause" Masc. Acc./Nom.
*hizbe	"Leute, Volk" Fem.
k'ane	"Tag" (Amhar. Urspr.) Fem.
*makiina/i	"Fahrzeug, Bus, Maschine" Masc. Acc./Nom.
manjo/i	"Person, Mann" Masc. Acc./Nom.
manjo	"Frau" Fem.
mine/i	"Haus" Masc. Acc./Nom.
*nafse	"Leben, Seele" (Amhar. Urspr.) Fem.
rakko	"Mühe, Problem" Fem.
*sha'e	"Tee" Fem.
yaane	"Angelegenheit" Fem.

Numerale

lame/i	"zwei" Acc./Nom.
--------	------------------

Pronomen

'ane/i/-	"ich" Acc./Nom.
----------	-----------------

Verben

*'alayy-	"fürchten"
*'azzan-	"trauern" (Amhar. Urspr.)
'odeess-	"mitteilen, Neuigkeiten ankündigen" (irreg. Kausat.)
*'ood'-	"weinen" (Intrans./Autobenef.)
dag-	"kommen"
dassi hiyy-	"glücklich sein" (Adverbiales Verb)
gal-	"übernachten"
gat-	"bleiben, übrigbleiben"
gey-	"erreichen, ankommen"
*gorra'-	"umbringen"
*ha'w-	"trinken" (Intrans./Autobenef.)
*haak-	"vergessen, nicht gewahr werden"
hed'-	"da sein, existieren, leben, wohnen" (irreg. Verb)

hir-	"kaufen"
hidd'- vgl.d'	"sich kaufen, mit (dem Leben) davonkommen" (Autobenef.)
hiyy-	"sagen, meinen, denken"
hun-	"wegwerfen, beseitigen"
kad-	"werden, sich ereignen"
ke'-	"aufstehen, anfangen"
*lab-	"scheinen, ähneln"
mar-	"gehen, fortgehen"
rey-	"sterben"
sodaat-	"fürchten"
*shiy-	"töten"
*tur-	"warten, verbleiben"
yaad-	"trauern"

Suffixe an Nomen bzw. an Sätzen

-'a	Adv. Suffix: "zu, für" (Richtung, indir. Obj.)
-'naa'a	Konjunktion: "um zu, damit"
-'ni	Konjunktion: "x1 während x2" (gleiche Zeit, neues Subj.)
-ba	Adv. Suffix: "bei" (allgemein lokal/temporal)
-ba	Konjunktion: "während" (gleiche Zeit, gleiches Subj.)
-cco/i	Adv. Suffix: "bei, nahe"
-n-ga/i	Derivationssuffix "Singulativ" Acc./Nom.
-ka/i	Emphat. "sogar" Masc. Artikel Acc./Nom.
-ke	Masc. Artikel Acc./Nom.
-ke-e'e	Masc. Artikel / Fokus-Suffix
-n (1)	Adv. Suffix: "wegen" Masc.
-ni	Emphase "sogar" (nominal)
-nna	Modus: Bestätigende Aussage
-nni (1)	Koordinat. Suffix "und"
-nni (2)	Adv. Suffix: "mit"
-ssha (-ffa)	Konjunktion: "dadurch daß"
-tt'e-e'e	Emphatisches Suffix "sogar"
-te	Adv. Suffix: "wie, als"
-tt'a	Konjunktion: "damit, als, wie wenn"
-tt'e	Adv. Suffix: "wegen" (Fem.)
-tt'i	Konjunktion: "weil"
-wodda	Kopula: Identifizierung / Fokussuffix
	Fem. Artikel Acc.
	Fem. attributives Suffix, Genitiv
	Kasussuffix: Genitiv
	Konjunktion: "danach" (Zeitfolge, gleiches Subj.)
	Adv. Suffix "x1, danach x2" (Zeitfolge, neues Subj.)

Suffixe an Verben (Aspekt/Tempus, Person, Derivation)

-a	Infinitiv
-a'e-nne	Imperf. 1.Ps.
-a'ne	Imperf. 3.Ps./2.Ps.Pl.
-aa	Imperf. Futur 3.Ps./2.Ps.Pl.
-anno	Imperf. Futur 1.Ps.
-e	Perf. 3.Ps./2.Ps.Pl.
-e'e	Perf. 3.Ps./2.Ps.Pl.

Texterzeugung am Beispiel kuschitischer Narrativtexte

-e'e-nne	Perf. 1.Ps.
-e-baa	Perf. 3.Ps./2.Ps.Pl. Neg.
-ed'	Autobenefaktiv
-ee	Perf. 3.Ps./2.Ps.Pl.
-e-mma	Perf. Vergght. 3.Ps./2.Ps.Pl.
-e-nne	Perf. 1.Ps.
-mma	Vergght. 3.Ps./2.Ps.Pl.
-n (2)	Pl.
-t	2.Ps. oder Fem.Sg.

Morphophonologische Regeln

*-b+ -t -> -bb	2.Ps. oder Fem.
-d + -t -> -dd	2.Ps. oder Fem.
-r + -d' -> -dd'	Autobenefaktiv
-r + -n -> -rr	Pl.

8.2 Beispiel eines Textkerns

Aus den oben gegebenen lexikalischen Einträgen sind die Prädikationen des folgenden Textes aufgebaut.

Schema 28
Gedeo-Text K

28a: Zeit-Achse

	Agens	Patiens	andere Rollen	Prädikate	Übergänge
2	[1] <i>'ani</i> ich		<i>soolaama</i> -DIR-SZE (Ortsname)	<i>ke'</i> -	
3			DET <i>rakko</i> -FOC <i>giddo</i> DEM-DIR <i>ke'</i> - Mühe in		<i>anfang</i> -
4			<i>makiina</i> -INSTR Wagen	<i>dag-</i> <i>komm-</i>	
"Von S. kommend, bin ich unter großen Mühen mit dem Bus gekommen."					
6	[2] <i>berhaanu tasamma</i> , (Name)	DET- <i>belto</i> ,	Sohn	<i>REP</i> <i>dag-</i> <i>komm-</i>	
7	[1] (ich)		<i>nafse</i> -DET-DIR Leben	<i>'alayy</i> -SS fürcht-	
10			>> 8- 9<<	<i>hiyy-</i> <i>sag-</i>	
11				<i>'alayy-</i> fürcht-	
12				<i>dag-</i> <i>komm-</i>	
"B.S. der Sohn des Z., kam, ich fürchtete um mein Leben, Ich dachte (8-9) [...] und fürchtete mich, als ich kam."					
14			<i>wondo</i> (Ortsname)	<i>REP</i> <i>gey-</i> <i>erreich-</i>	
15		<i>nafse</i> Leben		<i>hir'</i> -BEN rett-	
16		<i>sha'e</i> Tee		<i>ha'w-</i> trink-	

"Als ich W. erreicht hatte, hatte ich mein Leben gerettet; ich trank etwas Tee."

Texterzeugung am Beispiel kuschitischer Narrativtexte

Agens	Patiens	andere Rollen	Prädikate	Übergänge
17		DEM-LOC <i>kayy</i> -DIR hin-	<i>dag-</i> <i>komm-</i>	
20		>>18-19<<	<i>hiyy</i> -LOC <i>sag-</i>	
21		<i>sodaata</i> -LOC-COM <i>dilla</i> Furcht (Ort)	<i>dag-</i> <i>komm-</i>	
22	<i>boga</i> -EMPH Gegend		<i>haak-</i> <i>vergeß-</i>	

"(Von dorther) kam ich, ich dachte (18-19) und fürchtete mich sehr; ich kam und wußte nicht einmal, wo ich war."

24 [1] [ich]	S-GRND	<i>kaba</i> hier	<i>dag</i> -SS <i>komm-</i>	
25		<i>tafarra</i> -PROX (Name)	<i>gal</i> -SS <i>übernacht-</i>	
30	[4]	<i>hado</i> heim	<i>mar-</i> <i>geh-</i>	
31			<i>gey</i> -CONCL <i>erreich-</i>	

"Ich kam hier an, übernachtete bei T. und ging bis zum Haus."

35 [4] <i>hizbe</i>	'an-THM S-DET- <i>manjo</i> -ART DET- <i>mine</i> -LOC DET- <i>k'ane</i> -SZE yaat-CONCL	REP
Leute	mich Mann Haus Tage	trauer-

"Die Leute nahmen mich auf wie einen, der für tot gehalten war und trauerten zwei Tage mit mir."

41 [1] ' <i>ani</i>	S-DET-COOR S-DET-CONN DEM-DET- <i>yaane</i> -THM [4] <i>hado</i> -DET-LOC ' <i>odeess</i> -CONCL	REP
ich	Sache Familie	ansag-

"Ich teilte [...] die Sache der Familie mit."

42 [4] <i>hado</i> -DET Familie	[5] S-DET <i>hado</i> -DIR heim	'ood'-LOC wein-
46	[5] S-DET-DIR	<i>dassi</i> -LOC freu-
48	[5] S-DET	'azzan-LOC trauer-
49		<i>tur-</i> verbleib-

"Die Familie weinte wegen der Verwandten, freute sich über [...] und trauerte noch lange."

28b: Zitate (Wörtliche Reden)

	Agens	Patiens	andere Rollen	Prädikate	Übergänge
8			>>yooni tecco oh heute		
9 [3]	[1] 'an-			shiy-	
	[sie] mich			töt-	
			<<		
			"Oh! Heute werden sie mich töten."		
			>>		
18 [3]	[1] 'an-			gora '-SS	
	[sie] mich			umbring-	
19				hun-	
				beseitig-	
			<<		
			"Sie werden mich umbringen und beseitigen."		

28c: Argumentative Relationen

	Agens	Patiens	andere Rollen	Prädikate	Übergänge
23 DEM-FOC				kad-GRUND	
(dies)				ereign-	
24				[dag-] komm-	
			"Weil sich dies ereignet hatte, kam [...]"		
28 [1]			bariisa-SZE DEM-LOC-SZE	dag-SG	
[ich]			früh	ankomm-	
29				[lab-NEG-EMPH-DET] schein-	
			"Daß ich frühmorgens ankommen würde, war unwahrscheinlich."		

28d: Determinierende Prädikationen

- Vorangestelltes Determinans

3	lumott '-DET	[rakko]	
	groß-	Mühe	
6	zikek-DET	[belti]	
	(Name)	Sohn	
36	'an-DET	[mine]	
	mein	Haus	

Texterzeugung am Beispiel kuschitischer Narrativtexte

36 lame [k'ane]
zwei Tage

38 'alaya-DET [yaane]
Furcht Sache

- Nachgestelltes Determinans

Agens	Patiens	Ort etc.	Prädikat	Übergang
7 [nafse]	'an-DET			
Leben	mein			
"mein Leben"				
29 [S-SG]			lab-NEG-EMPH-DET	
(Sache)			erschein-	
"die Sache war unwahrscheinlich"				
33 [manj-]			rey-DET	
Mann			sterb-	
"ein Mann, der gestorben war"				
39 [hado]			rey-DET	
Verwandte			sterb-	
"gestorbene Verwandte"				
40 [hado]			hed'-DET	
Verwandte			leb-	
"lebende Verwandte"				
43 [hado]			rey-DET	
Verwandte			sterb-	
"gestorbene Verwandte"				
45 [hado]			gat-DET	
Verwandte			übrigbleib-	
"übriggebliebene Verwandte"				
47 [6] [5] [hado]-SUBJ DEM			gorra'-VERGG-DET	
[sie] Verwandte			umbring-	
"umgebrachte Verwandete"				

Text-Regeln, syntaktische, morphologische und phonologische Regeln stellen die Beziehung zwischen dem Kern eines Textes und seiner vollen phonologischen Form her.

Alle hier vorkommenden Regeln sind in den früher dargestellten enthalten (Wedekind 1990: 141-460). Der volle Text wird hier unter 8.3 dargestellt.

8.3 Beispiel eines entfalteten Textes

Schema 29
Gedeo-Text K

	Subjekt Objekt	Adverbien	Verben	Übergänge
1		<i>soolaamii'ni</i> (Ortsname) -von		
2 [1] 'ani ich			<i>ke'enne</i> -tt'i, aufgestanden -SS	
3		<i>lumott'i rakkote giddo 'okko'ni ke'enne-tt'i,</i> viel-DET Mühe-FOC in dort-LOC aufgestanden-SS		
4		<i>makiiniki</i> Wagen-INST	<i>dage'enne</i> -ni. kam-ich -CONCLUS.	
"Ich kam von S.; ich kam unter viel Mühen von dort und bin mit dem Bus gekommen."				
5 [1] [ich]		<i>makiiniki 'ora</i> Wagen-INST Weg	<i>daganno</i> kam-	-wodda, -als-SUBJ->
6 [2]berhaanu tasamma, zikekiki belti, B.T., Z.-s Sohn-SUBJ			<i>daga</i> kam-	-'naa'a -SUBJ->
7 [1] <i>nafse</i> [1] 'ant'ee'e [ich] Leben mein-OBJ			<i>'alayy</i> fürchtend	-e'e -SS
8		<i>>>yooni!</i> >>oh!		tecco heute
9 [3] [1] 'anee'e [sie] mich			<i>shinnee</i> töten-	ni. CONCLUS.
		<<		
10 [1] [ich]			<i>hiyyenne</i> meinend	-tt'i -SS
11			<i>'alayya'enne</i> fürchtend	-tt'i -SS
12			<i>dage'enne</i> kam-ich	-ni. -CONCLUS.
"Als ich also mit dem Bus gekommen war, kam B.T., Z.s Sohn und ich fürchtete mich und dachte "Heute werden sie mich töten"; so fürchtete ich mich und kam (hierher)."				
13 [1] [ich]			<i>dagenne</i> gekommen	-tt'i, -SS
14		<i>wondo</i> (Ortsname)	<i>geyenne</i> erreicht	-tt'i -SS
15 <i>nafse</i> Leben			<i>hidd'enne</i> erhalten	-tt'i, -SS

Texterzeugung am Beispiel kuschitischer Narrativtexte

16	sha'e Tee	ha'we'enne trank-ich	-ni. -CONCLUS.
----	--------------	-------------------------	-------------------

"Als ich angekommen war und W. erreicht hatte, hatte ich mein Leben gerettet. Ich trank etwas Tee.

17 [1] [ich]	'okko'ni kayyi'a dort-her von >>	dagenne gekommen	-tt'i, -SS
18 [3] [1] 'anee'e [sie] mich		gorraan mordend	-e'e -SS
19		hunnaa	-ni. beseitigen-sie-CONCL.
	<<		
20 [1] [ich]		hiyya meinend	-'ni -SS
21	sodaata'ninni dilla Furcht-mit (Ort)	dagenne gekommen	-tt'i -SS
22	boganga Umgebung-sogar	haake'enne vergaß-ich	-ni. -CONCLUS.

"Als ich also von dort gekommen war, dachte ich "sie werden mich ermorden und beseitigen"; so kam ich voller Furcht nach D. und wußte nicht einmal mehr, wo ich war."

23 tennete dies-FOKUS		kaddemma ereignete	-tt'ee'e, -weil
24 [1] [ich]	kaba hier	dag kommend	-e'e -SS
25	tafarraba T.-bei	gal	-e'e übernachtend -SS
26	'okko'ni dort-her		
27			bariisa frühmorgens
28		dagenne ankomme-ich	-cco -daß
29 S		labbebaa scheinend-NEG-EMPH-DET	-n-gi
30 [1] [ich]	[4] hado Familie[-zu]	marenne gegangen	-tt'i -SS
31		geyenne	-ni. erreichte-ich-CONCLUS.

"Weil das so war, bin ich hierher gekommen und habe bei T. übernachtet und von dort weiter war es mir unwahrscheinlich, daß ich frühmorgens ankommen würde, so ging ich zur Familie und kam dort an."

Klaus Wedekind

	Subjekt Objekt	Adverbien	Verben	Übergänge
32	[1] [ich]	[4] <i>hado</i> Familie	<i>maranno</i> gehend	-wodda, -als-SUBJ->
33	[4] [Fam.]	[1] 'ane'e' mich	<i>reyed'ee</i> gestorbenen	-ke -DET
34		<i>manjissha</i> Mann-gleich		
35	[4] <i>hizbe</i> Leute			<i>lame k'ane</i> zwei Tage
36		[1] 'anke mine'ni mein Haus-in	<i>yaatee</i> trauerten-sie-CONCLUS.	-ni.
37	[4] [Fam.]		<i>yaata</i> trauernd	-'naa'a -SUBJ->
38	[1] [ich]	tennee'e 'alayint'a yaane diese Furcht-DET Sache		
39	[5] [Verwdt.]		<i>reyee</i> gestorbene	-ka-nna -DET-und
40			<i>hed'ee</i> lebende	-ka-nna -DET-und
41	[1] 'ani [5] ich [Verwdt.]	[4] <i>hadike'ni</i> Familie-zu	'odeessenne kündete-ich	-ni. -CONCLUS.
	"Als sie trauerten, teilte ich ihnen diese furchtbaren Dinge über die gestorbenen Verwandten und über die, welche überlebt hatten.			
42	[4] <i>hadiki</i> , Familie-DET			
43	[5] [Verwdt.]		<i>reyee</i> gestorbenen	-ke -DET
44	[4] [Fam.]	[5] <i>hadi'a</i> Verwandt.-zu	'ood'a weinend	-'ni -SS
45	[5] [Verwdt.]		<i>gate</i> geblieben	-ke-e'e -DET-wegen
46	[4] [Fam.]	[5] [Verwdt.]	<i>dassi hiyya</i> froh seiend	-'ni -SS

Texterzeugung am Beispiel kuschitischer Narrativtexte

47 [5] [Verwdt.]	'okko dort	gorraanemma ermordet	-ke-e 'e -DET-wegen
48 [4] [Fam.]	[5] [Verwdt.]	'azzana trauernd	-'ni -ss
49		turree blieben-sie	-ni. -CONCLUS.

"Die Familie weinte über die Verwandten, freute sich über die, welche überlebt hatten, und trauerte noch lange über die, die umgebracht worden waren."

9. Vergleiche und Ergebnisse

Die Ergebnisse aus den Untersuchungen von HEC-Texten - sowohl aus der vorliegenden kurzen Studie als auch aus der ausführlichen Darstellung (Wedekind 1990) - werden hier in zwei Gruppen zusammengefaßt: Zuerst nach Zügen, die vor allem eine der Sprachen charakterisieren (entweder Gedeo oder Burji oder Sidamo), dann nach Ergebnissen, die für alle drei HEC-Sprachen typisch sind.

9.1 Gedeo

Die für die Gedeo-Sprache typischen Eigenheiten lassen sich unter wenigen Gesichtspunkten zusammenfassen: Definitheit (Artikel), pragmatische Funktionen (Topik, Fokus), Adjektivsemantik und Modus.

Definitheit und Topik: In Gedeo ist der sogenannte "definite" Artikel -k/-t sehr häufig und in vielen adverbialen Verbindungen sogar obligatorisch. Außerdem kann dieser sogenannte "Artikel" - widersinnigerweise - zusammen mit dem indefiniten Morphem "ein" auftreten. Seine "determinierende" Funktion tritt also hinter anderen Funktionen zurück - zum Beispiel der, das Ende einer Nominalphrase zu markieren. Zusammenhängend damit ist über Topikmarkierung folgendes zu bemerken: In Gedeo tritt das meist als "und" übersetzte gemeinäthiopische Suffix -nna nicht ohne vorausgehenden "Artikel" auf. Sehr selten koordiniert dieses "und" einfach zwei Nominalphasen, wie seine Übersetzung nahelegen würde. Die Textkonstellationen, in denen -ke/-te -nna auftritt, weisen diese Kombination "-der/-die -und" als Topikmarkierung aus. (Vgl. ähnlich Gasser für Amharisch 1983: 99).

Fokus: In Gedeo funktioniert die Kopula als Fokussuffix -te/-ke. Pro Paragraph wird nur eine einzige Phrase mit diesem Fokus markiert. Was Fokus angeht, ist in Gedeo auch die Anordnung der Elemente in der Nominalphrase weitgehend von pragmatischer Bedeutung: das nachgestellte Element zieht den Fokus der Aufmerksamkeit auf sich..

Semantik der Adjektive: Es gibt in Gedeo nur wenige ursprüngliche Adjektive. Mediale Verben, Partizipien und Qualitätsnomina übernehmen die semantisch typischen Adjektivaufgaben. Ähnliches trifft wahrscheinlich für die anderen HEC-Sprachen zu; aber Untersuchungen dazu liegen uns nicht vor.

Modus und Aspekt: In Gedeo schließt ein Modussuffix oder "Validator" eine Szene ab - zum Beispiel das "bestätigende" Suffix -ni. Ähnlich wie in Burji geht die Wahl der Aspekte mit der Wahl des Modussuffixes einher - im Falle des -ni ist es zum Beispiel vorwiegend Perfekt und Vergangenheit.

9.2 Burji

Burji unterscheidet sich von den Sprachen des kuschitischen "Hochlands" vor allem durch sein ausgeprägtes Fokussystem, das so eigentlich nur in den Sprachen des Tieflandes bekannt war. Es geht bei dieser kontrastiven

Charakterisierung also vor allem um das Fokussystem dieser Sprache.

Fokus, Topik und "Pause": In Burji ist die Entscheidung über die Fokussuffixe -k'aa/-naa nicht nur mit Entscheidungen über Aspekt und Modus verbunden, sondern auch mit der Definitität von Nominalphrasen. Jede Szene hat einen - und mit wenigen Ausnahmen nur einen - Fokus. Der Zuhörer erwartet, daß der zu besonders zu beachtende Teil eines Paragraphen sozusagen mit einem Aufkleber - dem Fokussuffix -k'aa/-naa - dem Blick entgegenkommt. Fokussierung von Nomen und Pronomen läßt sich in dieser Sprache von der Wahl des Verbaspakts und des gesamten Verb- und Pronomenparadigmas nicht trennen. Neue Information wird in Burji entweder durch eine Pause oder durch ein besonderes Morphem 'ora "Pause" angekündigt. Dieses Morphem institutionalisiert sozusagen das Zögern, welches einem neu ausgreifenden Ansatz vorausgeht. Je nach Medium kann es als gesprochene "Interpunktions" oder geschriebene "Sprechpause" betrachtet werden: auch in geschriebenen Texten hat es seinen Platz. Die Wahl der absoluten Pronomina und der determinierenden Suffixe shi -gu -ngoo "fem. masc. belebt" usw. hängt von der Topikalisierung und Belebtheit des Referenten ab, und korreliert ebenfalls mit Fokussierung. Eine Übersicht über diese Korrelationen gibt Wedekind (1990: 540-547).

Pronomen und "Switch Reference": In Burji - als der einzigen kuschitischen Sprache des äthiopischen Hochlands - ist Subjektwechsel im Pronomen angezeigt: 'iss- "die gleiche Person" vs. 'is- "eine andere Person". Ein Pronomen der 3. Person Plural (oder eine entsprechende Pluralnominalphrase) darf in dieser Sprache nur ausgelassen werden, wenn das Verb ein entsprechendes Suffix erhält - und umgekehrt. Diese Bedingung gilt tatsächlich nur für die 3. Person Plural, nur in wenigen Aspekten und nur in dieser Sprache.

9.3 Sidamo

Sidamo steht nur am Rande dieser vergleichenden Darstellung - trotzdem gibt es einige Punkte zu bemerken, in denen diese für HEC-Sprachen zentrale Sprache sich unterscheidet: über Pronomina, Subjektwechsel, nominale Suffixe und ungewöhnliche Präfixe.

Subjekte und Personen: In Sidamo gibt es - wie in Gedeo - verschiedene unterordnende Satzsuffixe, die einerseits einen Wechsel des Subjektes anzeigen - das sind die längeren, auf -a endenden Suffixe - und die andererseits das Weiterbestehen des gleichen Subjekts anzeigen - das sind die kürzeren, auf -i endenden Suffixe. Es gibt in beiden Sprachen nicht nur je ein solches Suffix, sondern mehrere mit den hier beschriebenen Eigenschaften. Was die Markierung von Personen am Verb angeht, ist Sidamo ebenfalls bemerkenswert, weil Objektsuffixe am Verb und Possessivsuffixe am Nomen auftreten - wie in den äthiosemitischen Sprachen des Nordens. In bezug auf die Anzahl der Nominalphrasen pro Verb unterscheidet sich Sidamo trotzdem nicht von den anderen HEC-Sprachen: Wie in Gedeo oder Burji treten in den hier untersuchten Texten nicht mehr als zwei Nominalphrasen pro Verb auf. Was allerdings direkte Reden angeht, ist Sidamo in dieser Hinsicht bemerkenswert - gegenüber den anderen Sprachen - weil vor dem Beginn einer direkten Rede der Sprecher selten genannt wird. Das ist in Gedeo und Burji fast immer der Fall. Schließlich ist Sidamo in bezug auf Personenmarkierung am Verb noch deshalb auffällig, weil - jedenfalls in der älteren Generation und in einigen Dialekt - die 1. Person nach masc. und fem. unterschieden wird. (Das ist für äthiopische Sprachen überhaupt ungewöhnlich. Die anderen HEC-Sprachen unterscheiden hier nur bei Pronomen im Vokativ.)

Negation: In Sidamo ist Negation ungewöhnlich, insofern als hier ein Präfix di- "nicht" auftritt. Die HEC-Sprachen sind reine Suffixsprachen.

9.4 Charakteristika der HEC-Sprachen

Weitere Ergebnisse der Textstudien gelten für alle drei hier untersuchten Sprachen. Diese Ergebnisse werden hier nach verschiedenen Gesichtspunkten zusammenfassend dargestellt.

SOV-Charakteristika und -Besonderheiten: Die untersuchten Sprachen sind SOV-Sprachen; aber nur in 6 Prozent aller untersuchten Sätze des Textkorpus ist tatsächlich eine volle SOV-Folge zu finden. Bei der Charakterisierung der HEC-Sprachen als "SOV-Sprachen" sollte die Tatsache nicht übersehen werden, daß die Mehrzahl aller einfach HEC-Sätze nur aus einem Verb besteht, und daß die vom Text her bestimmten Umordnungen recht zahlreich sind, wie zum Beispiel O,SV oder OV,S. Alle drei Sprachen erlauben auch, daß zwischen Objekt und Verb andere Konstituenten treten (Modifizierer oder Adverbien). Das gilt für SOV-Sprachen als ungewöhnlich (Lehmann 1978: 19), läßt sich aber aus pragmatischen Notwendigkeiten erklären.

Texterzeugung am Beispiel kuschitischer Narrativtexte

Thema und Szenerie: In allen drei untersuchten Sprachen ähneln die formalen Mittel, ein Thema zu etablieren, denen, eine Szenerie aufzubauen: So wie thematisierte Nominalphrasen vorangestellt und durch eine Pause vom syntaktisch vollständigen Satz abgetrennt werden, um das Thema zu wechseln, so werden auch Orts- oder Zeitangaben vorangestellt und von der entsprechenden Szene oder dem Paragraphen durch eine Pause abgetrennt, um die Szenerie zu wechseln. Was in der Diskurstheorie *setting* genannt wird (Longacre, Grimes), konvergiert in den HEC-Sprachen mit dem, was in der Funktionalen Grammatik *theme* genannt wird.

Topik- und Subjektfunktionen: Kein Partizipant tritt anders in einen Text ein als durch einen Namen oder ein Nomen. Im Text anwesend bleibt ein einmal eingeführter Partizipant durch kurze Verbsignale, typischerweise [...] -i "gleiches Subjekt (SS)" vs. [...] -a "anderes Subjekt (DS)". Wenn das Subjekt und Topik einmal eingeführt ist, wird es fast nie wieder durch Pronomen oder Nominalphrasen ausgedrückt. Dafür gibt es verschiedene Gründe: Einer ist die allgemeine Vermeidung von Nominalphrasen, ein anderer ist die Markierung des Subjektwechsels durch - relativ aufwendige - Verbsuffixe. Was in diesen HEC-Sprachen zum Subjekt und Topik werden soll, wird oft vorher in anderen semantischen Funktionen eingeführt. (Zum Subjekt und Topik wird es durch Wendungen erhoben wie etwa -ke/-te-naa "diese(r) und".) Einmal eingeführt, gilt das gleiche Subjekt oder Topik dann im Durchschnitt für etwa fünf Verben oder "Ereignisse". In mehreren der untersuchten Texte treten in der Zeitachse strikt nur "belebte" Subjekte als Topik auf.

Fokus: Zur Fokussierung innerhalb der Nominalphrase läßt sich folgendes feststellen: Die hier untersuchten Sprachen haben trotz der typischen SOV-Tendenz zur Konstituentenfolge "Genitiv - Nomen" (Givón 1979: 215) auch die umgekehrte Folge "Nomen - Genitiv", sofern der Fokus des Textes dies verlangt.

Verben des Sagens und direkte Rede: Für äthiopische Sprachen sind Verben des Sagens und Tuns wichtig und typisch und eine starke Quelle von verbalen Neubildungen. Was Zitate angeht, muß für die hier untersuchten HEC-Sprachen die Frage, wie direkte Reden (Zitate) sich universal zum Verb des Zitierens verhalten (Longacre 1983: 131), so beantwortet werden: Das Verb des Zitierens bezieht sich - soweit es die HEC-Regeln der Übereinstimmung anzeigen - nicht auf die letztgenannte Nominalphrase, auch nicht auf seinen letzten Satz; es wird vielmehr das ganze Zitat als maskulines oder feminines abstraktes Objekt angesehen (meist feminin) - und auf dieses Objekt bezieht sich das abschließende Verb **hiyy-** "sagen".

Formative und Verteilung des semantischen Raums: In diesen HEC-Sprachen bilden relativ wenige Silben sehr viele Formative. Dabei sind viele Konzepte über mehrere Dimensionen hinweg in gleicher Weise benannt - ob es sich um Konzepte wie -a "hin zu" in Raum, Zeit oder abstraktem Rang handelt, um -k/-t "Definität, Vokativ, Deixis" (vgl. Pike 1965), oder um das Formativ [...]aa "Frage (nach der Identität)" wie im englischen wh-. In den HEC-Sprachen wird der semantische Raum, den die wenigen Dutzend Adjektive sozusagen brächliegen lassen, durch die folgenden Wortarten gefüllt: durch autobenefaktive und intransitive Verben, Partizipien, Genitive und vor allem durch die Subklasse der Nomina, welche Qualitäten ausdrückt (sogenannte adjektivische Nomina).

Komplexität von Phrasen und Sätzen: In den Texten dieser Sprachen kommen komplexe Nominalphrasen nicht vor. Ein normaler Textablauf führt nicht zu Situationen, wo etwa eine definite Phrase noch mit Attributen beladen würde wie zum Beispiel "die kleinen zerfallenen Häuser am Wegrand". Während HEC-Phrasen kurz sind, sind die Sätze - in traditioneller syntaktischer Sichtweise jedenfalls - sehr komplex: Die Texte dieser HEC-Sprachen enthalten viele - nach einfachen syntaktischen Kriterien - "komplexe" Sätze. Außer dem letzten Verb eines Paragraphen ist tatsächlich keines der anderen Verben selbständige. Die abhängigen Verben sind zwar zueinander koordiniert, lassen sich aber nicht als "finit" deuten. Die gleiche Komplexität von Verbfolgen gilt für mehrere äthiopische Sprachfamilien, einschließlich äthiopsemitsch (zu Amharisch vgl. Gasser 1983: 101).

Tempus und Aspekt: In diesen HEC-Texten sind Zeit und Aspekt durch ihren Platz im Text bestimmt. Tempus, Aspekt und Modus werden nicht Verben zugewiesen, sondern Verbserien oder Szenen (Wolfson 1978, Schiffрин 1981: 56). Diese weitgehend voraussagbare Zuweisung von Tempus und Aspekt kann in Form einer Tabelle angegeben werden (vgl. Wedekind 1990: 437ff. für Gedeo; 498ff. für Burji; und 625 für Sidamo). Neben dem "Platz im Text" spielt auch das Medium eine Rolle (mündlich vs. schriftlich), die stilistische Ambition und die Determiniertheit der Partizipanten.

Allgemeine Textstrategien: Die hier untersuchten Sprachen lassen sich zusammenfassend aus der Textperspektive durch folgende Eigenschaften oder Textstrategien charakterisieren: Fokussierung bestimmter Konstituenten der Nominalphrase, Verbverkettung, starke Markierung der Satzränder als Schaltpunkte - und Meidung von Nominalphrasen. Nicht nur Nominalphrasen werden zugunsten von Verbkonstruktionen vermieden - auch die Pronomina. (In den hier untersuchten Sprachen werden Pronomina vermieden; sie werden selten anders gebraucht als possessiv. Unter den Pronomen sind nur die Possessivpronomen häufig; oft funktionieren

sie als definite Artikel, wie auch im Amharischen.)

Im Ablauf eines HEC-Narrativtextes ist die Anzahl der tatsächlich zu einem Verb tretenden Nominalphrasen nicht von der Valenz des Verbs bestimmt - der Text erlaubt weniger als die Verbvalenz erlauben würde. In jedem HEC-Text überwiegt die Anzahl der Verben die der Nomina. In SOV-Sprachen wie den hier untersuchten HEC-Sprachen ist es zwar theoretisch - von der Syntax und der Valenz der Verben her - möglich, daß ein syntaktisch korrekter Satz fünf Nominalphrasen hintereinander aufreihst, ohne daß ein Verb dazwischen treten müßte. Mehrere Nominalphrasen könnten also einander folgen, zum Beispiel mit den Funktionen "Zeitangabe, Ortsangabe, Subjekt, Objekt, Adverb": NP NP NP NP NP. Außerdem könnten zu diesen Phrasen noch Genitivattribute treten, ohne daß bisher irgendein Verb dazwischengetreten wäre. Daraus wird deutlich, daß in einer solchen Sprache die Grenze und Funktion jeder Nominalphrase morphologisch deutlich markiert sein muß. (Nicht-SOV-Sprachen brauchen das nicht und tun es auch nicht: vgl. die Belege in Heine & Voßen 1981). In den untersuchten Texten treten solche monströsen Sätze mit fünf bis sieben Nominalphrasen nie auf - aber ohne Kasussuffixe wäre selbst die Interpretation von Sätzen mit nur zwei oder drei Nominalphrasen vieldeutig.

Eine typische SOV-Strategie braucht also, aus der Sicht eines Textes, folgende Mittel: Kasussuffixe; Abgrenzung des Subjekts als Topik; Unterscheidung verschiedener Nomina durch Genus, Verwendung vieler Verben (z.B. Relativsätze) anstelle von Nominalphrasen. Es treten zwischen zwei Subjekte oder zwei Nominalphrasen meist mehrere Verben. Sätze mit einer oder keiner Nominalphrase sind die häufigsten.

Weitere Charakteristika der allgemeinen Textstrategie: Verben treten eng zu "Serien" oder Ketten zusammen; das letzte Verb einer Kette ist semantisch unbelastet und "ausgebleicht" und markiert die Satz-"Richtung". Bei den anderen Verben werden Personen und Aspektsuffixe minimal gehalten: nur Veränderungen - nicht was konstant bleibt - werden markiert. Was den Hauptverben vorausgeht, sind nicht Nominalphrasen, sondern ebenfalls Verben.

Übergänge und Veränderungen sind die Last der Verben oder ihrer Suffixe: In allen untersuchten HEC-Texten ist das positive, deklarative Verb im Perfekt die häufigste Verbform in der Zeitachse. Andere Formen bilden ungefähr die Hälfte aller Formen und treten außerhalb der Zeitachse auf. Verbsuffixe, die keine Veränderungen ankündigen, unterscheiden sich typisch von solchen Suffixen, die weitgehende Veränderungen ankündigen. Wo keine Veränderung anzukündigen sind, ist das Suffix kurz und leicht (z.B. nur -e "Gerundium"); wo aber Veränderungen anzukündigen sind - z.B. Veränderungen der Handlungskette, Veränderungen der Zeit, des Ortes, oder/und der Partizipanten - werden schwere, auffällige Suffixe oder Postpositionen benutzt - z.B. -cci -nni -kayyi'a "Singulativ -mit -nachdem", etwa: "nachdem x1 dies getan hatte, tat x2 [...]").

In allen untersuchten Texten finden sich die irregulären, negativen, seltenen und morphologisch "umständlichen" Aspekte und Verbkonstruktionen außerhalb der Zeitachse des Textes: in den direkten Reden mit ihren phatischen Elementen, in Rückblenden der Zeit - etwa als -mra'a "Plusquamperfekt", in Negationen falscher Annahmen - etwa als -bo'no -bott'o "1. 2.Ps. nicht", oder in Entwürfen imaginierter Welten - etwa als das sehr selten verwendete (und deshalb in HEC-Sprachen lange übersehene) -u'unnu -u'uttu "1. 2.Ps. unwirklich". Eine der am meisten "unregelmäßigen" und "aufwendigen" Formen ist der negative Imperativ: -tott'e'e, -tinokk'e'e "Nicht! (sg., pl.)". Hier tritt auch das einzige Präfix der Gedeo-Sprache auf: mee- "Nicht!". In den einfachen, "normalen" Abläufen der Zeitachse dagegen finden sich die "normalen" (regulären, unauffälligen, kurzen) Verbparadigmen. Die Abläufe spiegeln in dieser Sicht also den psychologischen Aufwand des Sprechers ab.

Über switch reference hinaus: Ein System, von dem *switch reference* nur ein Teil ist, steuert die Textabläufe: Die HEC-Verbsuffixe verfolgen ja nicht nur das Subjekt. Wie Givón für andere SOV-Sprachen beobachtet hat (Givón 1979: 276, Fußnote), sind es auch in den hier untersuchten SOV-Sprachen Verbsuffixe, die den Fluß des Textes regulieren: sie klären, wo das Topik ist, und sie lenken die Aufmerksamkeit des Zuhörers auch durch Modussuffixe. In den HEC-Sprachen sind Verbsuffixe die Schaltpunkte (*switches*) eines übergreifenden Systems. Je mehr sich von Szene zu Szene ändert - oder von Verb zu Verb - desto stärker muß das entsprechende Signal sein. Bei größeren Veränderungen der Szene wird ein Verb ohne semantische Eigenlast gewählt (z.B. hed'- "dasein", ke'- "anfangen"), um solche schweren Suffixsignale zu tragen.

Im Ablauf der hier untersuchten Texte gilt es, nicht nur durch eine Beschreibung eines Systems von *switch reference* die Identität der Subjekte oder der dominierenden Partizipanten zu verfolgen - vielmehr muß *switch reference* als Subsystem einer übergreifenden Steuerung betrachtet werden: In HEC-Texten wechseln mit den dominierenden Partizipanten auch Zeiten, Orte, und Perspektiven des Sprechers (cf. Schiffрин 1981: 51):

Texterzeugung am Beispiel kuschitischer Narrativtexte

ganze Diskurswelten können sich ändern.

Literaturangaben

- Butterworth, Brian (1979).
Hesitation and the production of verbal paraphasias and neologisms in jargon aphasia, *Brain and Language* 8. 133-161.
- Cerulli, Enrico (1922).
The Folk Literature of the Galla of southern Abyssinia. Harvard African Studies 3. Cambridge, MA: Harvard University Press.
- Chafe, Wallace L. (1980).
The Deployment of Consciousness in the Production of a Narrative. In Wallace L. Chafe (ed.), *The Pear Stories: cognitive, cultural, and linguistic aspects of a narrative production*. Norwood: Ablex. 9-50.
- Danlos, Laurence (1987).
The linguistic basis of text generation. Cambridge etc.: Cambridge University Press.
- Davey, Anthony (1978).
Discourse production: a computer model of some aspects of a speaker. Edinburgh: Edinburgh University Press.
- Dik, Simon C. (1978).
Functional Grammar. Amsterdam: North-Holland.
- Dik, Simon C. (1980a)
Studies in Functional Grammar. New York etc.: Academic Press.
- Dik, Simon C. (1980b)
Seventeen sentences. In Edith A. Moravcsik & Jessica R. Wirth (eds.), *Current approaches to syntax* (Syntax and semantics 13), New York etc.: Academic Press, 45-75.
- Dik, Simon C. (1989)
The Theory of Functional Grammar. Dordrecht: Foris.
- Dundes, Alan (1964)
The morphology of North American folktale. Helsinki.
- Emmorey, Karen D. and Victoria A. Fromkin (1988).
The mental lexicon. In Frederick P. Newmeyer (ed.), *Linguistics: the Cambridge survey*, Vol. 3, Cambridge: Cambridge University Press, 124-149.
- Fodor, J. A. (1983)
The modularity of mind: an essay on faculty psychology. Cambridge, Ma.: MIT Press.
- Fromkin, Victoria (1971).
The non-anomalous nature of anomalous utterances. *Language* 47: 27-52.
- Garrett, Merrill F. (1975).
The analysis of sentence production, in: G. Bower (ed.), *The psychology of learning and motivation*, Bd. 9, New York: Academic Press.
- Garrett, Merrill F. (1984).
The organization of processing structure for language production: applications to aphasic speech. In D. Caplan (ed.), *Biological perspectives on language*, Cambridge MA: MIT Press.
- Garrett, Merrill F. (1988).
Processes in language production. In Frederick P. Newmeyer (ed.), *Linguistics: the Cambridge survey*, Vol. 3, Cambridge: Cambridge University Press, 69-96.
- Gasser, Michael (1983).
Topic continuity in written Amharic narrative. In Talmy Givon (ed.), *Topic continuity in discourse: a quantitative cross-language study*, Amsterdam etc.: Benjamins, 95-140.
- Givon, Talmy (1979).
From discourse to syntax: grammar as a processing strategy. In Talmy Givon (ed.), *Discourse and syntax* (Syntax and Semantics 12.) New York etc.: Academic Press.
- Goldstein, Kenneth S. (1964).
A guide for field workers in folklore. Hatboro: Folklore Associates.

Klaus Wedekind

- Grimes, Joseph E. (1975).
The Thread of Discourse. (Janua Linguarum, series minor 207.) The Hague: Mouton.
- Heine, Bernd & Rainer Voßen (1981).
Sprachtypologie. In Heine, Bernd, Thilo C. Schadeberg & Ekkehard Wolff (eds), *Die Sprachen Afrikas*, Bd. 5, Hamburg: Buske, 407-444.
- Johnson-Laird, P.N. (1983).
Mental models: towards a cognitive science of language, inference and consciousness. Cambridge: Cambridge University Press.
- Lakoff, George, & Mark Johnson (1980).
Metaphors we live by. Chicago: University of Chicago Press.
- Lehmann, Winfred P. (ed.) (1978).
Syntactic typology. Hassox: Spiers.
- Levine, Donald N. (1965).
Wax and gold: tradition and innovation in Ethiopian culture. Chicago: University of Chicago Press.
- Levy, David M. (1979).
Communicative goals and strategies: between discourse and syntax. In Talmy Givon (ed.), *Discourse and syntax* (Syntax and Semantics 12.) New York etc.: Academic Press, 183-212.
- Longacre, Robert E. (1983).
The grammar of discourse. New York: Plenum.
- Lowe, Ivan (1969).
An algebraic theory of English pronominal reference. *Semiotica* 1: 397-422.
- McKeown, Kathleen R. (1985).
Text generation: using discourse strategies and focus constraints to generate natural language text. Cambridge: Cambridge University Press.
- Meehan, J. (1981).
TALE-SPIN. In R.C. Schank & C.K. Riesbeck (eds), *Inside computer understanding: five programs plus miniatures*, Hillsdale NJ: LEA, 197-258.
- Meulen, Alice ter (1988).
Linguistics and the philosophy of language. In Frederick P. Newmeyer (ed.), *Linguistics: the Cambridge survey*, Vol. 3, Cambridge: Cambridge University Press, 320-446.
- Pike, Kenneth Lee (1965).
Non-linear order and anti-redundancy in German morphological matrices, *Zeitschrift für Mundartforschung* 32: 193-221.
- Rosenberg, St. T. (1979).
Frame-based text processing. In Dieter Metzing (ed.), *Frame conceptions and text understanding*, Berlin: de Gruyter, 62-82.
- Samarin, William (1967).
Field Linguistics: A Guide to Linguistic Field Work. New York/London.
- Schiffrin, Deborah (1981).
Tense Variation in Narrative. *Language* 57: 45-62.
- Tanenhaus, Michael K. (1988).
Psycholinguistics: an overview. In Frederick P. Newmeyer (ed.), *Linguistics: the Cambridge survey*, Vol. 3, Cambridge: Cambridge University Press, 1-37.
- Wedekind, Klaus (1978).
A Short Trilingual Dictionary of the Gedeo Language. [Manuscr., 180 pp., Addis Abeba University.]
- Wedekind, Klaus (1985).
Gedeo verb morphophonemics. *Afrikanistische Arbeitspapiere* 2, 1985: 82-109.
- Wedekind, Klaus (1990).
Generating narratives: interrelations of knowledge, text variants, and Cushitic focus strategies. Berlin etc.: Mouton de Gruyter.
- Wolfson, Nessa (1978).
A feature of performed narrative, *Society* 7: 215-237.