

Spekulationen zum Anfang von Sprache

Dieter Wunderlich, August 2007

Die Anfänge der Sprache liegen in grauer Vorzeit; lange Zeit galt das Spekulieren darüber als unwissenschaftlich. Inzwischen treffen sich Anthropologen, Archäologen, Primatologen, Genetiker, Neurophysiologen, Informatiker, Philosophen und Linguisten alle zwei Jahre zu einer *International Conference on the Evolution of Language* (zuletzt dokumentiert in Christiansen & Kirby 2003, Tallerman 2005). Dieses interdisziplinäre Interesse ist schon außergewöhnlich, und die Entdeckungen in diesem Gebiet sind es auch. Sie ermöglichen uns, wenn auch noch manchmal sehr schemenhaft, die Naturgeschichte der Sprache zu entziffern. Einige der wichtigeren Faktoren werde ich hier kurz erwähnen.

Den Anlaß, diesen Aufsatz zu schreiben, gibt eine Beobachtung, der meistens nur wenig Aufmerksamkeit gilt. Das Berechnungssystem der Sprache verteilt sich auf Morphologie und Syntax; Morphologie ist in manchem eine unfertige Syntax: sie ist komplizierter (z.B. schwieriger zu lernen), aber auch beschränkter (weil sie keine Stellungsalternativen erlaubt). Warum gibt es sie? Viele Linguisten sind überzeugt davon, dass es zuerst eine Art von Syntax gab, bevor sich Morphologie aufgrund von Verkürzungsprozessen hat entwickeln können: selbständige Wörter wurden zu unselbständigen Affixen. In dieser Perspektive müsste Morphologie leichter und nicht schwieriger als Syntax sein. Der Anteil an Morphologie stellt seit Anbeginn einen Parameter für linguistische Typologie dar; einige Sprachen bevorzugen morphologische, andere Sprachen bevorzugen syntaktische Konstruktionen; warum ist das so?

Bei dem Versuch, eine Erklärung für diese Tatsachen zu finden, kommt man unweigerlich auf Vermutungen über Umstände, die sie beeinflußt haben. So ist dann auch meine Vermutung einzuordnen, daß sich Morphologie vielleicht schon hat entwickeln können, bevor es im engeren Sinne Syntax gab. Wegen eines leicht zu antizipierenden Missverständnisses will ich anmerken, dass ich die Kombination bedeutungshaltiger Einheiten zu einem größeren Komplex nicht automatisch als Syntax (im engeren Sinne) ansehe. Für fast alle Sprachen gibt es relativ gute Tests, die die Abgrenzung von wortinternen und wortexternen Konstruktionen erlauben; zum Teil beziehen sie sich darauf, dass die Teile eines Wortes nicht bewegt werden dürfen, zum Teil darauf, dass die Teile eines Wortes besonderen phonologischen Regeln unterliegen (was ja offensichtlich zusammenhängt, denn was bewegt werden darf, kann nicht durch phonologische Nachbarschaft beschränkt werden). Und so nenne ich Syntax nicht jegliche Umstände von Komposition, sondern nur diejenigen, die von besonderen phonologischen Zwängen frei sind und im Prinzip zu Stellungsalternativen berechtigen.

Um die Umstände zu klären, unter denen sich die bemerkenswerte Aufteilung in morphologie-reiche Sprachen einerseits und isolierende Sprachen andererseits hat entwickeln können, oder jedenfalls die Tendenz zu erkennen, die sich diachronisch fortgesetzt hat, braucht es mehr als die letzten gut belegten Stadien in der Entwicklung der Sprachen. Daher kommt hier sporadisch auch die Naturgeschichte der Sprache als ganze in den Blick. Ich betrachte in Abschnitt 1 mögliche historische Kandidaten für den Anfang von Sprache von zeitlich nah zu zeitlich fern, sodann in umgekehrter Richtung verschiedene der in meiner Sicht einschneidenden evolutionären Schritte in Abschnitt 2. In Abschnitt 3 diskutiere ich das Konzept von Sprachrevolution im minimalistischen Programm, und in Abschnitt 4 begründe ich, warum Morphologie historisch älter sein kann als Syntax (im engeren Sinn).

1. Kandidaten für den Anfang

Die Frage nach dem Anfang von Sprache muß natürlich sofort eingeeengt werden, denn je nach dem, was wir unter Sprache verstehen, ergibt sich eine veränderte Perspektive. Meinen wir vornehmlich kombinatorische Syntax, meinen wir morphologische Variation von Worteinheiten, meinen wir die Kombination digitaler Grundeinheiten, meinen wir vokalische Artikulation? Es wird sich zeigen, dass die im folgenden erörterten Szenarien jeweils einen anderen Aspekt von Sprache im Vorder-

grund haben. Insofern kann Sprache evolutionär als ein Bündel aufeinander aufbauender Fortschritte betrachtet werden.

In der Suche nach einschneidenden Ereignissen der Menschheitsgeschichte, in denen sich die menschliche Sprachfähigkeit hat wesentlich ausformen können, lassen sich wenigstens vier Zeitperioden benennen, in denen eine veränderte Lebensweise mit einer entsprechenden Sprach- und Kommunikationsfähigkeit hat korrelieren können. In einem ersten Durchgang betrachte ich diese Zeitperioden chronologisch rückwärts, um sie dann in einem zweiten Durchgang in ihrer tatsächlichen zeitlichen Abfolge anzuordnen. Die Ereignisse dieser Perioden lassen sich als mögliche äußere Szenarien für die Sprachentstehung verstehen (Szenarien A bis D); sie liefern uns damit vier mögliche Kandidaten für den Anfang von Sprache.

Als **erstes** wäre der neolithische Übergang (**Szenario A**) zu nennen, eine Periode, in der die nomadische Lebensweise durch Sesshaftigkeit abgelöst wurde. Diese Periode bildet die graue Vorzeit, bis in die historische Kenntnisse über einzelne Völker zurückreichen und die archäologisch relativ gut erforscht ist. Bis zum Ende der Eiszeit vor etwa 12.000 Jahren lebten die Menschen in kleinen nomadischen Jäger- und Sammlergemeinschaften in der Größenordnung von 50 bis max. 500 Personen. Erst infolge des eintretenden Klimawandels wurden die Menschen sesshaft mit Vorratshaltung, Ackerbau und nachfolgender Viehzucht, wuchs die Bevölkerungszahl, kam es (einige tausend Jahre später) in den Zentren zu Städten und an den Peripherien zu Expansionen, also Völkerwanderungen (Diamond & Bellwood 2003). Die meisten der besser bekannten Sprachfamilien können bis in diese Zeit zurückdatiert werden, aber niemand wird annehmen, dass Sprache damals erst entstanden sei. Sprache könnte in dieser Phase jedoch eine substantielle Änderung erfahren haben.

In einer kleinen isolierten Nomadengruppe mit wenigen externen Kontakten gab es nur minimale Arbeitsteilung, weitgehend gemeinsame Voraussetzungen zur Kommunikation und oft wiederkehrende, strukturell ähnliche Kommunikationsanlässe. Was eine solche Gruppe auch als Sprache besessen haben mag, sie war gezeichnet durch einen hohen Grad an Memorisierung und Kontextualität – dies ermöglichte schnelle Sprachverarbeitung: je häufiger eine memorisierte Konstruktion benutzt wurde, desto schneller konnte sie abgerufen werden. Das konnte überlebenswichtig sein. Die Sprachen dieser nomadischen Gruppen können wir uns am besten als morphologiereich mit eher wenigen Generalisierungen, und vielleicht als kopfmarkierend vorstellen, dies auch, wenn man die Struktur heute noch gesprochener Sprachen von kleinen nomadischen Völkern zum Vergleich heranzieht.

Eine größere Bevölkerung mit wachsender Arbeitsteilung mußte zu enormer Variation der Kommunikationsanlässe führen, zu mehr externen Kontakten, zur Anwesenheit von Menschen verschiedener Sprachen auf engem Raum, damit zur *lingua-franca*-Bildung und größerer Syntaktifizierung; die notwendige Infrastruktur (Verwaltung usw.) führte schließlich zur Erfindung von Schrift. Dies mag zu vielen graduellen Veränderungen geführt haben, aber auch einen ganz anderen Typus von Sprache ermöglicht haben.

Als **zweites** wäre die jungpaläolithische Revolution (**Szenario B**) zu nennen, eine Periode, in der eine enorme Mobilität, neue Fähigkeiten und Ausdrucksbedürfnisse des Menschen zutage traten. Vor 80 - 60.000 Jahren (möglicherweise, aufgrund genetischer Untersuchungen, schon vorher) begannen Gruppen des modernen *homo sapiens* Afrika zu verlassen und sich nach und nach auf alle Kontinente zu verbreiten; so erreichten die ersten Menschen Neuguinea und Australien (damals noch ein einziger Kontinent) vor ungefähr 50.000 Jahren, und nur wenig später sind schon Menschen am mittleren Don nachgewiesen (Anikovich 2007). Um fast dieselbe Zeit findet man, an etlichen Orten in den verschiedensten Regionen der Welt, verbesserte Werkzeuge, Begräbnisspuren, schmückende Ornamente und erste Zeichen von Kunst, insb. Höhlenmalereien, in denen Tiere, Jäger, Zahlen, Sternbilder etc. dargestellt wurden. Nicht wenige datieren den Anfang von Sprache im engeren Sinn auf diese Zeit (Hauser et al. 2002, Ott 2007). Später könnte er nicht liegen, weil dann nicht erklärbar wäre, warum Australier, Afrikaner, Asiaten und Amerikaner dieselbe Art von Sprachlernfähigkeit haben; für die Kinder überall auf der Welt gibt es keine Beschränkung darin, welche Sprache sie lernen können.

Drittens könnten die Anfänge der Sprache mit der Entstehung des anatomisch modernen Menschen (*homo sapiens*) zusammenfallen, die aufgrund von Skelettfunden auf die Zeit vor knapp 200.000 Jahren datiert wird (**Szenario C**). Falls es eine genetische Disposition zur Sprachfähigkeit gibt, so sollte sie sich beim *homo sapiens* generell entwickelt haben (bzw. bei den Gruppen des *homo sapiens*, die bis heute überlebt haben). In diesem Zusammenhang ist interessant, dass in einer Londoner Familie mit erblicher Sprachschwäche (insb. Artikulationsschwäche) bestimmte charakteristische Defekte des FoxP2-Gens gefunden wurden (Lai 2001); also könnte dies eines der für Sprachfähigkeit verantwortlichen Gene sein. FoxP2 ist ein sog. Regulator-Gen, das die Aktivität anderer Gene steuert; das von FoxP2 ausbuchstabierte Protein umfasst 715 Aminosäuren. Forscher des Leipziger MPI für evolutionäre Anthropologie haben herausgefunden, dass sich das FoxP2-Protein des Schimpansen von dem der Maus in nur einer Aminosäure unterscheidet, aber in zwei weiteren Aminosäuren beim Menschen (Enard u.a. 2002). Aufgrund der beobachtbaren Variation liegt die Mutation 160.000 Jahre (nicht mehr als 200.000) zurück, und fällt damit in den Zeitraum der Entstehung des *homo sapiens*. Die Forscher vermuten, dass ein Vorteil der Mutation in der Kontrolle der Mund- und Gesichtsmuskeln (Kiefer, Kehlkopf) bestanden haben könnte. Somit könnte es sich um ein Ereignis handeln, das unmittelbar an der Entstehung oder Verbesserung der Lautsprache beteiligt war.

Viertens schließlich könnten wir die Anfänge der Sprache der Gattung *homo* generell zuschreiben, also der Periode, in der sich das menschliche Gehirn ausdifferenzierte (**Szenario D**). Die wesentliche Vergrößerung des Kortex (Frontallappen) begann vor 2 Mill. Jahren.¹ Die Frühmenschen (*homo ergaster/erectus*) konnten komplexe technische und soziale Leistungen erbringen: das Feuer behüten (seit 1,5 Mill. Jahren),² Faustkeile herstellen (seit 1,4 Mill. Jahren) und mit ihnen werfen (Calvin 1990), und bewegten sich in Gruppen aus Afrika heraus bis ins heutige Indonesien.³ Die späteren Neandertaler gehörten zu ihnen. Mußten diese Lebewesen nicht zugleich auch komplexe Denk- und Kommunikationsleistungen vollbringen, also Sprachfähigkeit besitzen? Vieles scheint dafür sprechen, dass zumindest die später entwickelten Linien des *homo erectus* über Gestenzeichen verfügten, vielleicht auch über rudimentäre Lautsprache.

2. Evolutionäre Schritte für Sprache

Hockett (1960, 1966) hat eine Reihe von Merkmalen benannt, in denen sich Menschensprache von den möglichen Kommunikationssystemen anderer Lebewesen unterscheidet; in der einen oder anderen Weise werden diese Merkmale nach wie vor akzeptiert.

Die wichtigsten der Merkmale Hocketts: vocal-auditory channel, interchangeability (of speaker and hearer), semanticity, arbitrariness, discreteness, displacement (beyond the here-and-now), reflexiveness, productivity, traditional transmission/learnability, duality of patterning.

Die meisten dieser Merkmale sind unverzichtbar; allerdings kann statt des vokal-auditorischen Kanals auch der gestisch-visuelle Kanal (und möglicherweise sogar der taktile Kanal) verwendet werden. Die genannten Merkmale sind grobenteils unabhängig voneinander oder in nicht schon offensichtlicher Weise miteinander verbunden; deshalb kann man annehmen, dass sie jeweils für sich entstanden sind. Nicht nur jedes der Merkmale wird sich schrittweise entwickelt haben, die Merkmale werden sich auch nebeneinander, aus unterschiedlichen Vorstufen heraus und zu

¹ Der *homo rudolfensis* vor 2,5 bis 1,8 Mill. Jahren hatte ein Gehirnvolumen von 600-900 ccm, der frühe *homo erectus* bis zu 1200 ccm, der *homo heidelbergensis* vor 600 bis 300.000 Jahren bis zu 1350 ccm, und der Neandertaler vor 130 bis 33.000 Jahren bis zu 1600 ccm.

² Die frühesten Funde über das Herstellen von Feuer sind 790.000 Jahre alt; sie kommen aus Israel.

³ Der 1969 in Sangiran auf Java gefundene Schädel wird auf ein Alter von 800.000 Jahren geschätzt. - Das wohl früheste menschliche Fossil außerhalb Afrikas ist ein 1999 im Kaukasus-Ort Dmanisi entdeckter 1,8 Mill. Jahre alter Schädel; sein Schädelvolumen betrug bereits 1300 ccm. Die einzigen Hilfsmittel dieses *homo erectus* waren scharfkantig zugeschlagene Flussgerölle.

unterschiedlichen Zeiten entwickelt haben. Für Menschensprache charakteristisch ist, dass alle diese Merkmale (und möglicherweise noch weitere oder in anderer Aufteilung) zusammenkommen. Einzelne dieser Merkmale finden sich auch in anderen Lebewesen ausgeprägt; es wäre also unangemessen, den Anfang von Sprache an einem der Merkmale festzumachen.

Dennoch ist die Frage berechtigt, ob es *ein* Merkmal gibt, das gewissermaßen den Schlussstein setzt und die verschiedenen teilweise unabhängigen Evolutionslinien integriert. Von nicht wenigen wird Syntax als dieses Merkmal angesehen, als ein kombinatorisches System, das die Interface-Strukturen miteinander verbindet: auf der einen Seite das Begriffssystem (konzeptuell-intentionale Strukturen), auf der anderen Seite das Lautsystem (vokal-auditorische bzw., allgemeiner, motorisch-perzeptuelle Strukturen). Nach der üblichen Vorstellung erzeugt Syntax abstrakte hierarchische Strukturen, die sowohl in das Begriffssystem als auch in das Lautsystem abgebildet werden. Das Charakteristikum von Syntax ist allerdings nicht so sehr, dass sie selbst die hierarchischen Strukturen schafft, sondern dass sie komplexe hierarchische Strukturen des Begriffssystems mit komplexen linearen Strukturen des Lautsystems verbindet (Bierwisch 2001a, 2007a).

Im folgenden werde ich nicht jedem einzelnen der Hockettschen Merkmale gesondert nachgehen, sondern einige zusammenfassen. Man wird unschwer erkennen können, dass die meisten dieser Merkmale bereits beim *homo erectus* bestanden haben dürften, aber zwei der wichtigsten möglicherweise erst beim *homo sapiens* hinzu gekommen sind.

Wie oben angedeutet, besaß schon der Frühmensch (*homo ergaster/erectus*) vor knapp 2 Millionen Jahren hohe technische Fertigkeiten. Das deutet an, daß er die Fähigkeit zum intentionalen Planen über das hier-und-jetzt hinaus hatte. Ebenso wird er die Fähigkeit zu komplexem propositionalem Denken besessen haben. Er konnte eine Proposition, die einen Sachverhalt repräsentierte, in ein Prädikat und einen referentiellen Anker zerlegen. Beim Herstellen eines Faustkeils beispielsweise bleibt der Referent (ein Stück Stein) im wesentlichen gleich, während sich das Prädikat (die jeweilige Form des Steins) bis zu einem intendierten Wert ändern soll.⁴ Das bedeutet auch, dass die Fähigkeit zur Prädikation 2. Grades bestanden hat: die jeweilige Prädikation konnte qualifiziert werden.

Für das Produzieren und Verwenden von Faustkeilen (z.B. um sie auf Wildherden zu werfen, was Calvin 1990 als eine ihrer Hauptfunktionen beschrieb) war ein genaues zeitliches Justieren (*timing*) der manuellen Handlungen nötig; eine hierarchische Struktur war möglichst schnell auf eine zeitliche Abfolge abzubilden. Jemand, der einen Stein auf ein womöglich bewegliches Ziel werfen will, übersetzt den wahrgenommenen Ort in eine intendierte Flugbahn und diese in eine gezielte Abfolge motorischer Bewegungen; andere Primaten können längst nicht so erfolgreich werfen wie der Mensch. Die notwendigen neuronalen Verknüpfungen konnten später vielleicht ausgeweitet werden zu anderen Zwecken.

In ähnlicher Weise beruft sich auch die Spiegelneuronen-Hypothese (*mirror system hypothesis*, kurz MSH) auf notwendige neuronale Strukturen und deren Ausweitung für Zwecke der Kommunikation (Rizzolatti & Arbib 1998). Die Hypothese besagt, dass das zuerst bei Affen beobachtete Spiegelsystem für Greifbewegungen (Rizzolatti et al. 1995) graduell auf andere Arten manueller und facialer Gesten ausgedehnt wurde, schließlich auch auf vokalische Gesten. Spiegelneuronen gehören zum motorischen System, aber werden bereits aktiviert, wenn man beobachtet, wie ein Partner eine einschlägige motorische Handlung vollzieht. Der Affe konnte beobachten, wie ein

⁴ Möglicherweise hielt man den Stein fest in der einen Hand, während man mit einem weiteren Stein in der anderen Hand Schläge ausführte. Krifka (2007) äußerte die Vermutung, dass die Unterscheidung von Topik und Fokus auf diese steinzeitlichen Umstände zurückgeht. Es könnte natürlich auch sein, dass sich die Händigkeit des Menschen unter diesen Umständen ausprägte (dominante vs. nichtdominante Hand). Immerhin hat die steinzeitliche Technologie den Menschen über weit mehr als die Hälfte seiner Geschichte begleitet. Es gibt die Meinung, dass man Sprache brauchte, um jemanden in einer solch anspruchsvollen Technologie zu unterweisen, aber auch die Gegenposition, dass die Herstellung von Faustkeilen eher durch Nachahmung als Unterrichtung erlernt wird, und über zehntausende von Jahren derart konstant blieb, dass eigentlich auch nicht viel zu lernen war. Ich nehme hier nicht an, dass die Frühmenschen vor 1 bis 2 Millionen Jahren sprechen konnten; aber sie konnten komplexe Propositionen bilden.

anderer Affe nach der essbaren Frucht griff und sie sich zurecht machte (letztlich, um sie sich in den Mund zu stopfen); dies aktivierte einen Teil der für solche Bewegungen zuständigen Region seines Kortex.

Diese Entdeckung war gleich in mehrfacher Hinsicht interessant. Zum ersten Mal konnte die rätselhafte Interaktivität der menschlichen Kommunikation verstanden werden. Wenn jemand ein Zeichen Z mit der Bedeutung B äußert, so intendiert er, im Hörer H einen Effekt hervorzurufen, der der Bedeutung B entspricht (Grice 1969), wieso gelingt ihm dies? Die MSH sagt, es gelingt ihm, weil gewisse Spiegelneuronen in H aktiviert werden, die ihrerseits dabei beteiligt wären, wenn H das Zeichen Z mit der Bedeutung B äußern würde (interne Reproduktion als Quelle für das Erkennen einer Intention). Im Verlauf einer Interaktion sind Sprecher- und Hörerrollen schnell austauschbar (Parität bzw. *interchangeability* bei Hockett), z.B. ergänzen sich Fragesatz und Antwort zu einer einzigen Proposition (*Who ate the cake? - I did*: dass der zweite Sprecher den Kuchen aß, wird hier zur gemeinsamen Auffassung beider Sprecher). Wenn der erste Sprecher einen möglichen Übergangspunkt im Gespräch blockiert (Sacks, Schegloff & Jefferson 1974), dann wird der zweite Sprecher an dieser Stelle in der Regel nicht einsetzen; wenn der erste Sprecher eine intendierte lexikalische Einheit momentan nicht finden kann, um sie zu produzieren, fügt der zweite Sprecher die passende Einheit ein; wenn der erste Sprecher einen sprachlichen Fehler produziert, wird der zweite Sprecher ihn in der Regel ignorieren, oft nicht einmal bemerken. Alle diese Interaktionstatsachen finden eine verhältnismäßig einfache Erklärung durch die MSH: der jeweils zweite Sprecher scannt das Wahrgenommene, indem er laufend eigene motorische Zellen aktiviert, als sei er dabei, die entsprechende Äußerung selbst zu produzieren.

Die MSH macht es plausibel, daß menschliche Kommunikation zuerst über manuell-faciale Gesten erfolgte – denn im motorischen Bereich der Hand (und, eng benachbart, dem der Gesichtsmuskeln) begann sich das System der Spiegelneuronen zu entwickeln. Für den gestischen Ursprung der Sprache sprechen im übrigen eine Reihe von Gründen (Corballis 2003, Arbib 2005). Schimpansen können eine einfache Gebärdensprache lernen, aber keine Lautsprache. In der Regel ergänzen sprechbegleitende Gesten den aktuellen Äußerungsinhalt in sowohl räumlicher (teilweise metaphorischer) als auch emotionaler Hinsicht. Wenn Sprechen (aus welchen Gründen auch immer) behindert ist, übernehmen Gesten auch morphosyntaktische Leistungen. Gebärdensprachen gibt es in religiösen Gemeinschaften und bei Taubstummen. Gebärdensprache kann mindestens genau so schnell wie Lautsprache erworben werden, möglicherweise schneller (Meier & Newport 1990). Manuelle Gebärden sind tendenziell eher ikonisch (d. h. übernehmen Merkmale der dargestellten Objekte) und darum relativ einfach zu verstehen. Insbesondere können sie leicht ein Prädikat mit einem Referenten verbinden, z.B. wenn die eine Hand das Objekt an einem Ort einführt und die andere Hand dazu das Prädikat.⁵ Durch die Kombination einzelner manueller oder facialer Gesten ergibt sich Kompositionalität, z.B. wenn die Hand an einem Ort kurz verbleibt und dann von diesem aus mit einer anderen Handform zu einem anderen Ort gleitet.

Auch wenn Gebärden Merkmale des Objektes analog nachbilden, so ist doch die Auswahl solcher Merkmale nicht zwingend, also gibt es immer ein Stück arbiträrer Konvention.⁶ Gebärden können natürlich ritualisiert und in diesem Sinne zu Symbolen abstrahiert werden. Und sie erlauben noch einen ganz anderen Weg zur Symbolentstehung. Wenn in einem späteren Entwicklungsstadium eine ikonische Gebärde von einer vokalischen Äußerung begleitet wurde, konnte die Äußerung sozusagen die Bedeutung der Gebärde übernehmen – aber was für die Gebärde ikonisch

⁵ Dies erinnert natürlich sofort an die Bemerkung oben, s. Fußnote 4. Es ist schon möglich, dass sich Gebärdensprache so wie das Faustkeil-Schlagen entwickelte. Goldin-Meadow (2005) bemerkt, dass neu entwickelte Gebärdensprachen immer die Abfolge Objekt - Verb (OV) haben, selbst wenn hörende Sprecher einer SVO-Sprache anfangen, mit den Händen zu kommunizieren.

⁶ Viele heute noch existierende Sprachen bilden den Plural eines Nomens oder den Iterativ eines Verbs durch Reduplikation, ein offensichtlich ikonisches Verfahren. Die Reduplikation ist aber in der Regel nur partiell, und Verben werden oft anders redupliziert als Nomina, so gibt es trotz Ikonizität auch einen arbiträren Anteil.

war, muß für die vokalische Äußerung notwendigerweise symbolisch (arbiträr) sein. Es bedurfte somit keiner unüberwindbarer Verabredungen für Lexikonelemente.

Die Notwendigkeit, Gruppen von Spiegelneuronen im Gehirn auszubilden, erklärt auch den kindlichen Imitationsinstinkt, bezogen auf Gebärden und Lautäußerungen; Menschenaffen ahmen zielgerichtetes Verhalten wie z.B. das Herstellen oder Benutzen eines Werkzeugs für die Nahrungsbeschaffung, nach, aber nicht scheinbar funktionsloses Verhalten. Imitation wird zum Motor des Sprachlernens und damit auch zum Motor der weiteren Sprachevolution. Iteratives Lernen durch Mitglieder aufeinanderfolgender Generationen mit jeweils leicht modifiziertem Input kann als kulturelle Evolution gekennzeichnet werden; sie erfolgt wesentlich schneller als genetische Evolution, und sie produziert Ergebnisse, die sich aus biologischen Gründen nicht vorhersagen lassen. Falls bestimmte Entwicklungsstufen des *homo erectus* nach und nach Gebärden- und Sprachsprache ausgeprägt haben und sich womöglich auch in der vokalischen Artikulation verbesserten, müssen wir annehmen, dass die kulturelle Evolution bereits mit dem *homo erectus* begann.

Damit sind lediglich folgende der Hockettschen Merkmale noch offen: *vocal-auditory channel*, *discreteness*, und eventuell *duality of patterning*.

Rätselhaft bleibt, warum visuelle Gesten durch vokalische Gesten abgelöst werden konnten. Jedenfalls in der bisher skizzierten Naturgeschichte der Sprache bietet sich nichts an, das diesen Übergang erklären könnte. Die im Bereich von 2 bis 5 KHz deutlich verbesserte Hörfähigkeit des Menschen (im Vergleich zum Schimpansen) wurde indirekt, aufgrund der Skelettanatomie von Außen- und Innenohr, schon für den Neandertaler nachgewiesen (Martinez 2004), also muß sie seit mindestens 500.000 Jahren bestehen. Dieser Frequenzbereich ist für die Wahrnehmung von Konsonanten und höheren Vokalformanten wichtig; somit waren die perzeptuellen Voraussetzungen für Lautsprache gegeben.⁷ Auf der Seite der Lautproduktion könnte die Absenkung des Kehlkopfes unabhängig aufgrund des aufrechten Gangs eingeleitet worden sein, nur war damit die willentliche Benutzung des Kehlkopfes noch nicht gegeben, vielleicht aber aufgrund der neuronalen Nachbarschaft im motorischen Areal im Prinzip erreichbar. Hier schließt die oben erwähnte Entdeckung der FOXP2-Mutation eine Lücke; diese Mutation könnte die Kontrolle des Vokalapparats so verbessert haben, daß die vokal-auditorische Modalität ebenso leicht ausführbar war wie die manuell-visuelle; aufgrund offensichtlicher Vorteile konnte sie dann auch Priorität über letztere gewinnen. Wenn man Laute produzierte, konnte man kommunizieren, auch wenn man sich nicht sah (im Dunkeln oder über Entfernungen), und man brauchte auch nicht die Beschäftigung der Hände zu unterbrechen. Die Nachteile (z.B. wer beim Essen spricht, kann sich leicht verschlucken) fielen wohl weniger ins Gewicht.

Die übrigen Merkmale könnten sich als Folge der vokalischen Artikulation eingestellt haben. Wie oben bereits erwähnt, sind vokalische Äußerungen für die meisten Bereiche auf arbiträre Symbolisierung angewiesen, und es gibt auch einen Weg, wie man sie, von einer Gebärden- und Sprachsprache ausgehend, erreichen kann. Nach Bierwisch (2007b) müssen Symbole (qua arbiträre Zuordnungen zwischen Form und Bedeutung) zwingend diskret sein; die Diskretheit der phonologischen Merkmale wird danach sozusagen extern erzwungen. Andere Autoren meinen, dass sich diskrete Artikulationsmerkmale durch automatische Selbstorganisation ergeben (Studdert-Kennedy 2005). Dies konnte Oudeyer 2005 mithilfe von Computersimulationen verifizieren: Im Zuge zahlreicher zufälliger Interaktionen von Sprechern und Hörern ergaben sich "regelmäßige Attraktoren", die man als solche diskreten Merkmale deuten kann. Studdert-Kennedy 2005 argumentiert auch, dass sich die *duality of patterning* (nicht die Merkmale selbst, sondern nur die Kombinationen aus ihnen tragen Bedeutung) in einem System von diskreten Merkmalen als optimale Lösung ergibt; bereits Abler 1989 habe gezeigt, dass, wenn immer mit endlichen Mitteln ein unendlicher 'Gebrauch' gemacht wird, ein zweistufiges System benutzt wird. Es gibt nur wenige Bausteine für Atome, aber ein experimentell immer noch erweiterbares periodisches System der Elemente; es gibt nur sehr

⁷ Wenn man sich fragt, warum Mutationen, die zu den entsprechenden Skelettunterschieden am Ohr geführt haben, positiv selektiert wurden, könnte man annehmen, dass lautliche Kommunikation eine Rolle dabei spielte. Aber sie dürfte weit weniger kontrollierbar gewesen sein als die einige hunderttausend Jahre später auftretende Lautsprache.

beschränkt viele atomare Bausteine für Moleküle, aber ein fast beliebig erweiterbares System der Moleküle; es gibt nur 20 Aminosäuren, aus denen sämtliche Proteine aufgebaut sind; es gibt nur vier Basen, mit denen sämtliche genetische Information der DNS kodiert wird. Ganz entsprechend dazu gibt es nur eine geringe Anzahl für sich bedeutungsloser distinktiver phonologischer Merkmale (bzw. Phoneme), aber eine unübersichtlich große Menge möglicher Kombinationen zu bedeutungshaltigen Lexemen.

Damit sind sämtliche Kriterien von Hockett erfüllt.

Wir haben bisher nur die Szenarien D und C betrachtet. Sie machen es glaubhaft, dass sich in den verschiedenen Entwicklungslinien des homo erectus ein Kommunikationssystem herausgebildet hatte, das einer Sprache schon sehr nahe kam, und dass der Übergang zum homo sapiens diese Entwicklung in Richtung einer Lautsprache wesentlich beschleunigt hat. Von Syntax war bisher nicht die Rede, wohl aber von der für Gebärdensprache erwartbaren Kombinatorik. Im folgenden soll kurz eine Position betrachtet werden, die gerade Syntax in den Mittelpunkt stellt.

3. Sprachevolution im minimalistischen Programm

Durch das minimalistische Programm (Chomsky 1995) wurden ältere Theorien zur Syntax (z.B. die Prinzipien- und Parametertheorie) teils aufgelöst, teils neu integriert. Interessanterweise ist ein Hauptaugenmerk der Theorie nun nicht mehr auf den kindlichen Spracherwerb, sondern auf die biologische Evolution der Sprache gerichtet. Für die Erfassung des Spracherwerbs schien es nützlich zu sein, sehr viel an angeborener Struktur anzunehmen (*Universal Grammar*); umso weniger hatte das Kind aufgrund seines Inputs zu generieren. Für die Erfassung der Evolution ist es umgekehrt nützlich, nur sehr wenig an angeborener Struktur anzunehmen; umso weniger braucht der durch Mutation und Selektion zustande gekommenen genetischen Sequenz angelastet zu werden.

Syntax ist auf rekursives MERGE reduziert, eine Funktion, die zwei lexikalische Elemente aufgrund ihrer Merkmalsinformation zu einem dritten komplexeren Objekt verbindet; letzteres kann dann wiederum im Input der Funktion auftreten. Externes MERGE (auch *base merge* genannt) realisiert u.a. Prädikat-Argument-Strukturen, internes MERGE (auch *target merge* genannt) realisiert Dislokationen. ‚Extern‘ bedeutet, dass sich ein Prädikatsausdruck mit einem Komplementäusdruck verbinden muß. ‚Intern‘ bedeutet, dass ein Komplex, bestehend aus α und β , wiederum mit α (bzw. einer Kopie von α) kombiniert; dies entspricht der bisherigen Betrachtung, dass α aus der einen Position in eine andere Position bewegt wird (Move α).

Von Hauser et al. 2002 wird zwischen Sprachfähigkeit im engeren und Sprachfähigkeit im weiteren Sinn unterschieden (*faculty of language in the narrow sense* = FLN; *faculty of language in the broad sense* = FLB). FLN umfasst einerseits die beiden Varianten von MERGE und andererseits die Verbindung zu den Interface-Systemen (Phonologie und Semantik). FLB hingegen schließt die Performanzsysteme (sensorisch-motorisches System, konzeptuell-intentionales System) mit ein.

In Weiterführung dieser Überlegungen betrachtet Ott 2007 die FLN im Rahmen der evolutionären Entwicklungsbiologie. Danach ist Sprachstruktur durch die vier Faktoren (i) genetische Anlagen, (ii) Entwicklungsconstraints, (iii) neuronale Architektur, und (iv) externe sprachliche Daten bedingt. Entwicklungsconstraints regulieren den Zeitpunkt, zu dem Gene wirksam (d.h. in Form von Proteinen ausgedrückt) werden. Die jeweilige neuronale Architektur ist verantwortlich für die Art, wie externe sprachliche Daten verarbeitet und repräsentiert werden (und ist ihrerseits natürlich auch in Wechselwirkung von genetischer Anlage und Umwelt entstanden; wegen dieser Wechselwirkung muß sie nicht vollständig genetisch determiniert sein). So weit dürften die Annahmen weitgehend unstrittig sein.

Ott 2007 nimmt nun spezifischer an (aber hierin durchaus auch verschiedenen Bemerkungen Chomskys folgend), dass FLN (also insbesondere die Funktion MERGE) spontan als Nebenprodukt des Gehirnwachstums entstanden sei, in Form einer optimalen Verknüpfung zwischen Neuronenansammlungen, und zwar aufgrund einer geringfügigen Mutation innerhalb der Funktionsweise von Regulator-Genen. Bei der Evolution von MERGE handele sich nicht um eine Anpassung (Adaptation), sondern um die Ausnutzung einer vorhandenen Struktur (Exaptation) (die man wohl

als ein System neuronaler Verknüpfungen zu verstehen hat). Die Performanzsysteme müssten schon vorher bestanden haben; keines dieser Systeme sei einzigartig für den Menschen, denn es gibt sie (in dieser oder jener Form) auch bei anderen Lebewesen. Erst FLN sei einzigartig für den Menschen. Ihre Existenz könnte der Anlaß für die paläolithische Revolution vor 80 - 60.000 Jahren gewesen sein (Szenario B).

Einige Punkte in diesem Szenario sind problematisch. Es ist nicht evident, inwiefern Begräbniskultur, Ornamente und Höhlenmalereien durch ein rekursives MERGE besonders gefördert werden konnten, auch wenn sie zweifellos eine höhere Reflexionsstufe des Menschen belegen. Sie zeugen für eine stärkere Auseinandersetzung mit Tod, Umwelt und Geschichte, aber die Produkte gehören doch großenteils zum visuellen System und sind durch eine sprunghafte Verbesserung der Sprachfähigkeit nicht zu erklären.⁸

Die Mutation, aufgrund der MERGE zustande gekommen sein soll, ist möglicherweise so geringfügig, dass man sie nicht wird entdecken können. Gleichzeitig muß aber angenommen werden, dass die Gruppe, in der sie stattgefunden hat und vererbt wurde, zum Ahnen aller heute lebender Menschen geworden ist, also vielleicht schon vor mehr als 100.000 Jahren gelebt hat. Zwar brauchte es beträchtliche Zeit, bevor die Auswirkungen der neuartigen Sprachfähigkeit aufgrund kultureller Evolution sichtbar werden konnten, aber dies schwächt natürlich das Argument, weil sich rekursives MERGE dann ja auch einfach per kultureller Evolution entwickelt haben könnte, ohne dass es ein vorausgehendes Mutationsereignis gegeben haben muß.

Anlaß für die paläolithische Revolution könnte auch eine Entwicklung gewesen sein, die nach Überwindung eines ‚Flaschenhalses‘ in der menschlichen Bevölkerungszahl eingetreten ist. Zhivotovsky et al. 2003 haben aus der relativ geringen Variation im Erbgut des Menschen geschlossen, dass kurz vor der ersten nachweislichen Trennung (die ihrer Studie nach zwischen 71 und 142.000 Jahre zurückliegt) vielleicht nur 700, maximal aber 2700 menschliche Individuen lebten.⁹ Wenn es so gewesen sein sollte, könnte eine darauf folgende Bevölkerungsexplosion zu ganz neuen Leistungen geführt haben.

Hier ist der Ort, um kurz an Bierwischs Paradox zu erinnern. Es besagt, dass eine Mutation, die zu einer verbesserten Sprachfähigkeit führt, gar nicht positiv selektiert werden kann, weil die betreffende Person gar keine Gesprächspartner hätte, mit denen dieser Vorteil zum Zuge kommen kann (Bierwisch 2001b). Die Vermeidung des Paradoxes ist relativ einfach. Wie bei aller sozialen Evolution muß man annehmen, dass die Mutation zunächst über einige Generationen vererbt wird, bis sie sich über eine bestimmte Fraktion der Population verbreitet hat und dann dort auch positiv selektiert werden kann. (Das würde so natürlich auch für MERGE gelten.) Jede Form sozialer Evolution ist insofern immer an eine Abfolge von Generationen gebunden, Rein theoretisch würde es genügen, dass die Stamm-Mutter ihre neue Fähigkeit artikuliert, ohne dass sie Vorteile davon hätte, und ihre Kinder, die auch Träger der Mutation sind, diese Produktionen nachzuahmen

⁸ Nach Funden im Braunkohletagebau Nordmark-Nord bei Halle (2006) haben Neandertaler schon vor 90.000 Jahren ‚Kunstobjekte‘ (Versteinerungen, Kristalle, Korallen, fossile Zähne) gesammelt, also auch sie hatten bereits eine Vorstellung von dem, was bemerkenswert und ästhetisch anregend ist.

⁹ Die Untersuchung stützt sich auf die Variation von 377 DNA-Abschnitten, die als Kopien auftreten und besonders häufig mutieren, ohne eine direkt erkennbare Funktion zu haben, in 52 verschiedenen weltweit verteilten Bevölkerungsgruppen. Sie zeigt clusterartig, welche Gruppen sich etwa wann trennten und wieviel Variation zu späterer Zeit aufgrund des Bevölkerungswachstums hinzugekommen ist. Demzufolge trennten sich zuerst die südsaharischen San und Mbuti (die sog. Pygmäen) – die aber ihrerseits wesentlich mehr genetische Variation aufweisen als z.B. die Neuguineeer und die indigenen Amerikaner –, dann die anderen südsaharischen Völker (Bantus usw.), die Eurasier (Nordafrika, Europa, den vorderen Orient und Südasien umfassend), die Neuguineer/Australier, und schließlich die Ostasiaten und Amerikaner. – Die geringe Zahl der menschlichen Individuen zum Zeitpunkt der ersten deutlichen räumlichen Aufspaltung muß nicht heißen, dass es sich um nur einen Stamm handelte; im Gegenteil, die Aufspaltung zeigt, dass schon räumlich getrennte Unterpopulationen bestanden haben müssen. Es kann sich um 10 bis 15 verschiedene Stämme gehandelt haben, die sich alle bereits in fortgeschrittenen Stadien einer Lautsprache befunden haben. Aber sie hatten es mit erheblichen Umweltveränderungen zu tun, in denen nur wenige überlebten.

versuchen (und sie dann möglicherweise regularisieren und verfeinern). Damit sind wir bereits bei kultureller Evolution angelangt.

Die Verfechter der generativen Syntax, und so auch die des minimalistischen Programms, haben immer ihren rekursiven Charakter hervorgehoben. Nicht eigentlich MERGE ist interessant, sondern rekursives MERGE. MERGE gibt es schon in allen Vorläufern von Sprache, weil es einem Rezipienten gar nicht möglich wäre, zwei lokal verbundene bedeutungshaltige Handlungen oder Gesten *nicht* zu kombinieren. Wie steht es aber um die Rekursivität? Es ist mehr als fraglich, ob sie erst so spät in der Naturgeschichte des Menschen, nämlich im Zusammenhang mit Sprache, speziell Syntax, entstanden ist.

Für unzählige visuell-motorische Verhaltensweisen muß Rekursivität sowohl im visuellen als auch im motorischen Bereich angenommen werden. Der Jäger, der abschätzt, wie er nach einem fliehenden Fisch oder Tier werfen will, muß immer wieder räumliche Abschnitte visuell extrapolieren und in ein motorisches Programm einspeisen. Der Sammler nimmt ein Stück nach dem anderen auf, wobei jedes zwischenzeitliche Bündel wieder Teil eines Bündels werden kann, bis eine intendierte Menge (Reisig, Pilze, Blätter, Muscheln) erreicht ist. Der Werkzeugmacher schlägt einen Schlag nach dem anderen, bis durch wechselnde Abschlüge in dieser und jener Richtung der intendierte Zustand erreicht ist. Auch die Reziprozität von interaktiven Emotionen und Handlungen, die sich wechselseitig verstärken oder blockieren, beruht auf Rekursivität. Rekursivität gibt es in einem System, das zwei Basiskategorien zuläßt, z.B. Verb (V) und Nomen (N), sowie die Möglichkeit eines Kategorienwechsels, denominalen Verben und Nominalisierung von Verben. Im Lexikon findet man Rekursivität in allen Einstellungsverben mit einem propositionalen Komplement (*ich glaube, daß er denkt, daß du überzeugt bist, daß ...*). Es ist doch sehr wahrscheinlich, dass schon homo erectus, wenn er über komplexe Propositionen verfügte und vor P etwas verstecken wollte, vielleicht auch denken konnte, P würde annehmen, dass er es dort oder dort hingetan habe. Bedeutungen lassen sich also rekursiv erzeugen. Dann ist es aber auch wahrscheinlich, dass in einem ikonischen System die Ausdrucksseite rekursiv ausgestaltet werden kann. Auch wer mit den Fingern oder den Zehen zählt, und dann in 5er, 10er, 12er oder 20er Gruppen weiterzählt, folgt einem rekursiven Schema in mehreren Hierarchiestufen.

Ob es überhaupt einer speziell angeborenen Syntax-Fähigkeit bedarf, ist, wenn man Computersimulationen glauben darf, sehr fraglich. Hurford (2000) und Kirby (2002 bis 2005) konnten mithilfe von Modellen für iteriertes Lernen zeigen, dass sich syntaktische Kompositionalität im Laufe vieler Generationen von selbst einstellt, also ein Produkt kultureller Evolution ist. Unter der Voraussetzung, daß die Agenten strukturierte Bedeutungen ausdrücken wollen und zu Anfang jeder von ihnen beliebige Verknüpfungen mit Zeichenformen herstellt, führen leichte Variationen in der Produktion und Generalisierung dazu, daß nach einer Serie von Lernprozessen, die jeweils auf einem beschränkten Input der Vorgängergeneration operieren, sich nach hunderten oder tausenden von Generationen eine kompositionale Sprache herausbildet. Wenn die Agenten alle Bedeutungen mit gleich großer Häufigkeit ausdrücken, dann entsteht am Ende eine vollständig reguläre und stabile Sprache, weil es keinen Anlaß zu Irregularitäten gibt. Wenn die Agenten aber einige Bedeutungen häufiger als andere ausdrücken, dann entsteht eine Sprache mit sowohl kompositionaler Struktur (für die selteneren Bedeutungen) als auch holistischen/irregulären Äußerungen (für die häufigsten Bedeutungen). Wenn semantische Kompositionalität mit der Möglichkeit zur Rekursion als gegeben angenommen wird, kann ein iteriertes Lernmodell zeigen, daß eine Gruppe von Agenten nach N Generationen (z.B. N=1.000) eine stabile Zeichenkombinatorik zum Ausdruck dieser Rekursion entwickelt. Die Annahme einer separaten Entstehung von syntaktischer Kompositionalität ist also überflüssig. Nicht so überflüssig ist die Frage, wie syntaktische Dislokationen entstanden sein könnten.

4. Morphologie und Syntax haben ein unterschiedliches historisches Alter

Das Nebeneinander von morphologiereichen, z.B. polysynthetischen, und isolierenden Sprachen ist schwer erklärbar. Die einen Sprachen haben fast nur Morphologie mit ein wenig syntaktischer Adjunktion, die anderen Sprachen haben fast reine Syntax mit nur wenig morphologischer

Kompositabildung. Warum gibt es Morphologie, wenn Sprachen auch ohne sie auskommen können? Die landläufige Vorstellung zur Entstehung von Morphologie besagt, dass Sprachen mit Syntax beginnen und sich einstmals selbständige Wörter durch Reduktion zu stellungsfesten Klitika und schließlich zu morphologischen Affixen entwickeln. Diese Auffassung kann aber nur einen Teil von Morphologie erklären; sie versagt in Fällen von Ablaut, Mutation und Reduplikation. Sie ist auch fraglich für typische kopfmarkierende Morphologie; über welche Art von Klitik könnte sich z.B. das algonkische Inverssystem entwickelt haben? Und warum gibt es keine der so angeblich ubiquitären Klitisierung in den isolierenden Sprachen? Die Antwort, die ich vorschlagen möchte, steht quer dazu: Morphologie und Syntax könnten sich unter verschiedenen historischen Bedingungen gebildet haben.

Problematisch im Minimalistischen Programm ist die enge Verknüpfung von externem und internem Merge als zwei gleichberechtigten Instanzen von Merge. Externes Merge dürfte schon in allen Vorformen heutiger Sprache eine Rolle gespielt haben, z.B. dann, wenn komplexe Propositionen durch mehr als einen Ausdruck realisiert wurden, etwa in der Gebärdensprache für Prädikation und referentielle Verankerung jeweils verschiedene, aber miteinander verbundene Gesten verwendet wurden. Internes Merge hingegen findet sich nur in der Syntax, abgesehen davon, dass es nicht selten auch in Anspruch genommen wird, wenn Positionsalternativen in sehr abstrakter Weise diskutiert werden. In der Morphologie muß zwischen Rekursion und Dislokation streng unterschieden werden. Morphologische Objekte schließen Rekursion ein; so gibt es Suffix-Abfolgen wie Passiv-Kausativ-Passiv, propositionseinbettende Suffixe, usw., aber Dislokationen sind ausgeschlossen.¹⁰ Internes Merge spielt innerhalb der Morphologie keine Rolle.

Ein Vergleich von Morphologie und Syntax zeigt uns, dass morphologische Objekte weniger flexibel sind als syntaktische Kombinationen, sowohl positional als auch funktional. Morphologische Objekte sind durch feste Positionen ihrer Bestandteile ausgezeichnet; es gibt keinerlei Kongruenz zwischen den Elementen und auch keine Auszeichnung einzelner Affixe als Topik oder Fokus. Morphologische Objekte werden aber von mehr phonologischen Regeln tangiert als syntaktische Kombinationen und sind darum anfälliger für Irregularitäten. Wahrscheinlich ist das Ausmaß an beobachtbarer Variation nirgends so groß wie in der Morphophonologie. Für Erwachsene ist Morphologie sehr viel schwieriger zu lernen als Syntax; deshalb ist in Pidgins, allgemein in ungesteuerten Lernervarietäten erwachsener Personen, Morphologie so gut wie abwesend. Kurz gesagt, Morphologie ist unnütz, außer für das Gedächtnis.

Neben den verschiedenen Nachteilen bietet Morphologie nämlich auch einen Vorteil. Wahrscheinlich lernen Kinder die verschiedenen Wortformen paradigmatisch, also innerhalb eines Systems, dessen kategoriale Dimensionen sie nach und nach identifizieren, und erwerben so reguläre und irreguläre Wortformen nebeneinander, etwa in dem Sinne, wie es *word-and-paradigm*-Modelle beschreiben (Robins 1959). Wenn die Lerner dann zu späterer Zeit Morpheme abstrahieren (und damit die reguläre Morphologie produktiv verwenden können), bleiben die irregulären Instanzen größtenteils unanalysiert, möglicherweise aber auch die besonders häufigen und die besonders seltenen regulären Instanzen.¹¹ In memorisierten sprachlichen Objekten zählt für die Verarbeitungsgeschwindigkeit nicht die Komplexität, sondern die Häufigkeit. Sofern sie häufig

¹⁰ Gemäß der *copy*-Theorie wird Dislokation in der neueren Syntax als eine Art Kopiervorgang verstanden. Obwohl auch Reduplikation einen Kopiervorgang beinhaltet, ist sie dennoch etwas ganz anderes als syntaktische Dislokation (Bewegung): bei der Reduplikation werden einzelne Segmentsequenzen lokal verdoppelt, bei der Dislokation erhält ein Wort oder eine Phrase eine neue Position.

¹¹ Da es für Morpheme keine Positionsalternativen gibt, ist die Analyse einer gespeicherten morphologischen Form in ihre Morphembestandteile nicht zwingend, es sei denn, die Abfolge einzelner Affixe variiert je nach ihrem Skopus (z.B. Passiv vor Kausativ oder Kausativ vor Passiv). Die besonders häufigen Formen sind zwar analysierbar, aber wegen ihrer häufigen Benutzung zugleich auch memorisiert. Bei den besonders seltenen Formen kommt die Notwendigkeit der Reanalyse nicht in den Blick. Indefrey 2002 berichtet, dass oft noch Erwachsene über keine Regel der (selten belegten) schwachen n-Deklination der deutschen Nomina verfügen.

genug sind, können memorisierte morphologische Objekte also schneller verarbeitet werden als syntaktische Kombinationen. Und das ist natürlich zweifelsfrei ein kommunikativer Vorteil.¹²

Eine andere Besonderheit der Morphologie ist, dass sie sehr viele Verfahren ausgebildet hat, um innerhalb einer Dimension die markierte Variante zu kennzeichnen: Ablaut, konsonantische Mutation, Reduplikation, Affigierung, wobei das Inventar u.U. sehr klein ist. Es gibt manchmal fast generell einsetzbare Affixe, deren Effekt lediglich kontextuell gesteuert wird. Dabei entsteht z.B. die Situation, dass der Plural der einen Nomenklasse identisch markiert ist mit dem Singular der anderen, und umgekehrt (im estnischen Partitiv -i/-e, Blevins 2005), oder dass der Kausativ der einen Verben identisch markiert ist mit dem Dekausativ der anderen Verben (Japanisch -e, Comrie 2006). Solche reversativen (bzw. polaren) Verhältnisse sind aus der Syntax unbekannt. In der Regel sind mindestens in der einen Richtung nur Teilklassen des Vokabulars betroffen; so kann der Singular im estnischen Partitiv auch auf -u oder -a enden, und im Japanischen gibt es andere konkurrierende Kausativbildungen. Interessanterweise schreibt Baerman 2007 zu einem ähnlichen Fall im Neu-Aramäischen, dass ein Muster, das so aussieht wie ein *reversal*, zufällig entstanden sei und dann im Dialekt von Amadiya als Produkt eines systematischen *reversal*-Prinzips reanalysiert wurde. Offenbar ist es möglich, Wortformen so zu organisieren, dass nicht direkt Gebrauch gemacht wird von Morphemen, die hinzugefügt werden (denn Regeln wie N+i = part.sg./part.pl, N+e = part.pl/part.sg wären nicht hilfreich). Der Titel von Andersons Buch 1992 (*A-morphous morphology*) weist in diese Richtung. Die besondere Leistung in der Morphologie kann vielleicht deshalb erbracht werden, weil sie nicht nur nach Kategorien, sondern nach Kategorien mit phonologischer Form generalisiert und so automatisch zu Unterklassen führt, anders als die Syntax.

Morphologie ist möglicherweise älter als Syntax. Dafür spricht nicht nur die Art der dargestellten Asymmetrien, sondern auch, dass Sprachen, die vermutlich alt sind gegenüber jenen Sprachen, die sich in angebarer historischer Zeit entwickelt haben, fast immer morphologiereich sind, aber nie isolierend, etwa die Sprachen der letzten heute noch verbliebenen nomadischen Völker, und auch innerhalb der Morphologie oft ein besonderes Profil im Bereich der Kopfmarkierung aufweisen (Nichols 1992). Umgekehrt finden sich isolierende Sprachen besonders unter jenen Ethnien, die eine große Bevölkerungszahl oder eine große Mobilität haben. Auch die beobachtbare sprachliche Vielfalt ist besonders groß in jenen Regionen der Welt, in denen sich morphologiereiche Sprachen konzentrieren. Etwas ist am ursprünglichsten dort, wo die Variation am größten ist.

Wie kann man sich die Sonderstellung der Morphologie innerhalb des grammatischen Systems erklären, ihre Beschränkungen, ihr vermutlich höheres Alter und ihre größere Vielfalt? Die Vorstellung, sie habe sich auf dem Wege der Verschmelzung auf Basis einer ursprünglichen Syntax entwickelt, kann hierbei nicht helfen.

Sprachliche Vielfalt begann wahrscheinlich schon bei der Organisation der artikulatorischen Merkmale (und darauf basierend, der Organisation des Vokabulars), mindestens vor der Zeit, als die Menschen Afrika verließen, vermutlich aber lange davor.¹³ Wie wir wissen, kann das Inventar der Phoneme beträchtlich variieren, dies besonders im Bereich der Merkmale, die nur in einzelnen Sprachen zu finden sind. Z.B. nur die Khoisansprachen verfügen über Clicks; Nama unterscheidet im Bereich der Clicks insgesamt 20 Phoneme, in der südlichen Khoisansprache !Xóo sind es hingegen sogar 80 Phoneme (Traill 1985).¹⁴ Diese Phoneme sind distinktiv im Bereich des Lexikons, haben zum Teil aber auch grammatische Funktionen. Da beim Reden, insbesondere beim

¹² Natürlich können auch syntaktische Kollokationen oder Konstruktionen memorisiert werden und werden in diesem Sinne dann auch „morphologieähnlich“. Wegen im ganzen geringerer Irregularität können sie aber auch leichter reanalysiert werden.

¹³ Die Konstruktion einer „ersten Trennung“ muß nicht heißen, dass erst dann eine Variation eingesetzt hat. Insbesondere wenn die Trennung mit einem „Flaschenhals“ in der Bevölkerungszahl verbunden war, ist zu vermuten, dass es schon viele verschiedene Gruppen gab, aber nur einige überlebten.

¹⁴ Neben den 5 Artikulationspositionen für Clicks werden in !Xóo jeweils 16 zusätzliche Merkmalskombinationen unterschieden, basierend auf Aspiriertheit, Stimmhaftigkeit, Nasalisierung, Glottalisierung, Spirantisierung usw.

schnellen Sprechen, ständig Aussprachealternationen entstehen, können besonders auffällige z.B. als Ausdruck einer speziellen modalen oder temporalen Einordnung verstanden und dementsprechend vielleicht generalisiert werden. Soweit einzelne Phoneme bzw. artikulatorische Merkmale grammatische Funktionen erhalten, spricht man von Ablaut, Mutation, allgemein von modulatorischer Morphologie. Carstairs-McCarthy (2005) meint, dass Morphologie durch Reinterpretation allomorphischer Alternationen schon sehr früh in der Sprachgeschichte hat entstehen können. Aussprachealternationen an den Grenzen eines Wortes könnten zur Generalisierung von Präfixen oder Suffixen geführt haben. Wenn eine morphologische Kategorie einmal entstanden ist, kann sie nachträglich durch weitere phonologische Kontrastierung oder durch Verschmelzung mit zusätzlichen Elementen auch deutlicher hervorgehoben werden; auf diese Weise könnten echte Affixe entstanden sein.

Die Entstehung von Morphologie wäre somit unmittelbar an die Herausbildung artikulatorischer Merkmale geknüpft. Je mehr Merkmalsdistinktionen zugelassen sind, umso mehr morphologische Distinktionen werden ermöglicht. Zugleich gibt es eine Bindung an Frequenz einerseits und an Auffälligkeit (z.B. Sonorität) andererseits. Besonders häufige und besonders auffällige Alternationen erleichtern eine Generalisierung. Wie die Analysen von Müller (2002, 2007) zeigen, spielt Sonorität eine wesentliche Rolle in der Wahl von Flexionsexponenten; generell tritt ein Affix in umso mehr Kontexten auf, je höher der Sonoritätswert ist; mit anderen Worten, das System der Affixe tendiert dazu, Sonorität zu maximieren.

Auch die Präferenz für morphologische Kopfmarkierung könnte auf diesem Wege erklärt werden. Ein Verb ist dadurch charakterisiert, dass es ein, zwei oder vielleicht auch drei Argumente von einem bestimmten Typus zu sich nimmt. Ein Verb mit einem pluralischen Subjekt oder einem belebten Objekt drückt etwas anderes aus als das entsprechende Verb mit singularischem Subjekt bzw. einem unbelebten Objekt. Diese Klassifikation kann am besten am Verb selbst vollzogen werden, und dazu genügen u.U. wenige modulatorische Elemente.

Wir haben nun nur noch zu klären, warum eine Sprache, die einer Population von vielleicht tausend Individuen gut gerecht wird, bei einem rapiden Anwachsen der Bevölkerung und einem rapiden Anwachsen externer Kontakte verändert wird. Mindestens zwei Faktoren spielen eine Rolle: die Kontakte sind weder vorwiegend mit kontextuell ergänzbaren Elementen noch vorwiegend mit memorisierten modulatorischen Elementen zu leisten.

Das Anwachsen der Bevölkerung führt zu einer größeren Variation der Interaktionssettings. Je weniger eine Interaktion eingespielt ist, desto mehr müssen die Ausdrucksmittel auf die speziell vorliegende Interaktion hin spezifiziert werden. Die Kombinationen müssen transparenter sein und weniger anfällig für Irregularitäten; es werden mehr Spezifikationen durch unabhängige Nominalphrasen und mehr Spezifikationen von Topik und Fokus nötig. Natürlicherweise werden Topiks vorangestellt und fokussierte Elemente nachgestellt. Die Berücksichtigung der Informationsstruktur führt deshalb notwendigerweise zu Stellungsalternativen.

Aufgrund von Bevölkerungsbewegungen entstehen mehr externe Sprachkontakte, und es gibt eine zunehmende Anzahl erwachsener L2-Lerner. Diese neigen dazu, einfachere als auch transparentere Formen zu bilden; statt durch morphologische Mittel werden Argumente aufgrund ihrer Position oder ihres Informationsgehalts identifiziert; Tempus wird durch Zeitadverbien ersetzt, Numerus durch Zahlwörter usw. Dabei können dann auch semantisch relevante Wortstellungsvariationen entstehen, die sich im Sinne von Topik und Fokus deuten lassen, sowie Satzverknüpfungen, die anzeigen, wie die ausgedrückten Ereignisse zueinander geordnet sind. Die Herausbildung eines Systems von Komplementierern findet sich tatsächlich oft erst in historisch dokumentierter Zeit.

Der am ersten in Frage kommende Kandidat für die Phase der Syntaktifizierung der Sprachen ist der neolithische Übergang, das anfangs genannte Szenario A. Erst zu dieser Zeit entstanden Populationen mit einer Größe von mehr als 1000 Individuen. In einem gewissen Sinne ist reiche Morphologie ein Übergangsphänomen von Sprache, das im Zuge der Bevölkerungszunahme zunächst um eine transparente Syntax ergänzt und später vielfach von ihr abgelöst wurde. Das schließt nicht aus, dass eine eng umgrenzte Art von Morphologie weiterhin neu entstehen kann.

Dieses Szenario rückt die These von Klein und Perdue (1997), dass sprachliche Kompositionalität mit einfacher Syntax begann, ähnlich der *Basic Variety* der ausländischen Arbeiter in Europa, in den richtigen Zusammenhang: eine solche Syntax wird auch von den vielen Arbeitsklaven im Altertum produziert worden sein und somit jedenfalls mittelbaren Anteil an der Syntaktifizierung der Sprachen gehabt haben. Schon Bickerton (1981 und später) vertrat die Meinung, dass die aus Pidgins entstandenen Kreolsprachen die Wirkungsweise der Universalen Grammatik am deutlichsten zeigen und deshalb als Modell für die Sprachentstehung angesehen werden können. Wenn das wirklich so wäre, bliebe die Existenz vieler Arten von hochkomplexer (aber gleichzeitig auch beschränkter) Morphologie ein unlösliches Rätsel. Warum sollte eine Sprachgemeinschaft den Weg von entwickelter Syntax zurück zu einem beschränkteren morphologischen System gehen?

Wie oben schon angedeutet, bedarf es keiner zusätzlichen genetischen Anlagen, um die Prinzipien der Syntax zum Tragen zu bringen. Wenn Bedeutungen rekursiv geformt werden können, dann wird sich auch ein dementsprechendes rekursives morphosyntaktisches System herausbilden. Wird ein System ohne Positionsalternativen (Dislokation) zu einem System mit Positionsalternativen erweitert, dann können die erforderlichen Extraktionsbeschränkungen vermutlich aus dem geometrischen System rekrutiert werden; geometrische Transformationen lassen bestimmte Verhältnisse konstant (siehe z.B. PARALLEL MOVEMENT, Müller 2000); wird eine Figur gegenüber einem Hintergrund bewegt, müssen gewisse Lokalitätsbedingungen erfüllt sein (siehe z.B. PHASE IMPENETRABILITY, Chomsky 2000).

Wenn es richtig ist, dass einige der gegenwärtig gesprochenen Sprachen außer minimaler Adjunktion keine relevante Syntax im Sinne von internem MERGE aufweisen, dann kann dieser Typ von Syntax nicht für alle Sprachen gelten. Es ist vielmehr wahrscheinlich, dass er unter anderen Umständen als Morphologie gebildet wurde, und dies vermutlich zu späterer Zeit. Mein hier vorgestellter Vorschlag beinhaltet zwei verschiedene Szenarien: ursprüngliche Morphologie entstand unmittelbar im Zusammenhang mit Lautsprache,¹⁵ davon getrennte Syntax (internes MERGE) entstand oder verbreitete sich unter dem Druck größerer Populationsdichte, unter dem Einfluß einer *lingua franca* oder aufgrund noch anderer Faktoren.¹⁶

Literatur:

- Abler, W. L. 1989. On the particulate principle of self-diversifying systems. *Journal of Social and Biological Structures* 12: 1-13.
- Anderson, Stephen R. 1992. *A-morphous morphology*. Cambridge University Press.
- Anikovich, Mikhail et al. 2007. *Science* 315: 223.

¹⁵ Um dem erwartbaren Einwand vorzubeugen: Nein, nicht alle Morphologie ist damals schon entstanden. Neue Morphologie kann auch über die Klitisierung phonologisch reduzierter Elemente entstehen.

¹⁶ Es ist z.B. nicht evident, warum beim ‚Inselhopping‘ der Ozeanier die ursprüngliche austronesische Morphologie verloren ging und die jüngsten der ozeanischen Sprachen fast ausschließlich isolierend sind. Angesichts der allgemein geteilten Auffassung, dass die Austronesier den Reisanbau kannten, als sie vor etwa 5.000 Jahren Südchina und speziell Taiwan verließen, waren Faktoren der Syntaktifizierung natürlich schon lange wirksam. Die Frage ist deshalb nur, warum im Zuge der weiteren Expansion die Morphologie verloren ging. Der Ursprung der ozeanischen Sprachen wird in Südhalmahera und der Cenderawasih-Bucht in Nordwest-Neuguinea angesetzt (Blust 1978); von dort aus wanderten die Austronesier längs der Nordküste Neuguineas bis zum Bismarck-Archipel, wo sie jeweils auf eine relativ dichte ursprüngliche Papua-Bevölkerung trafen. Etliche Generationen von L2-Austronesisch-Lernern aus den unterschiedlichsten Papuastämmen sind nicht ausgeschlossen. Ab 1300 v. Chr. verbreitete sich die im Bismarck-Archipel entstandene Lapita-Kultur (u.a. durch Merkmale der Töpferware charakterisiert) auf weitere, bisher unbewohnte Inseln; wir wissen nicht, wie die Bootsbesatzungen ausgesehen haben, sie könnten gemischt gewesen sein. (Einschlägige genetische Untersuchungen zur Ausbreitung der Polynesier sind mir nicht bekannt.) Ein anderer Faktor zur Reduzierung der Morphologie lag darin, dass sich die Zahl der Konsonantenphoneme teilweise rapide reduzierte (bis auf 7 im heutigen Hawaiianischen), wortfinale Konsonanten und Konsonantencluster verschwanden; das lässt nicht mehr viel Raum zu modulatorischer Morphologie.

- Arbib, Michael. 2005. The mirror system hypothesis: How did protolanguage evolve? In Tallerman, 21-47.
- Baerman, Matthew. 2007. Morphological reversals. *Journal of Linguistics* 43: 33-61.
- Bickerton, Derek. 1981. *Roots of language*. Ann Arbor, Mich.: Karoma.
- Bierwisch, Manfred. 2001a. Repertoires of primitive elements: Prerequisite or result of acquisition? In Jürgen Weisenborn (ed.) *Approaches to bootstrapping*, Vol. 2, 281-307. Amsterdam: Benjamins
- Bierwisch, Manfred. 2001b. The apparent paradox of language evolution: Can Universal Grammar be explained by adaptive selection? In Jürgen Trabant & Sean Ward (eds.) *New essays on the origin of language*, 55-79. Berlin: Mouton de Gruyter.
- Bierwisch, Manfred. 2007a. Semantic form as interface, in. Andreas Späth (ed.) *Interfaces and interface conditions*, 1-32, Berlin: de Gruyter
- Bierwisch, Manfred. 2007b. Bedeuten die Grenzen meiner Sprache die Grenzen meiner Welt? Ms.
- Blevins, James P. 2006. Word-based declensions in Estonian. *Yearbook of Morphology* 2005: 1-25.
- Blust, Robert. 1978. Eastern Malayo-Polynesian: a subgrouping argument. *Pacific Linguistics* C61:181-234.
- Calvin, William H. 1991. *The ascent of mind. Ice age climates and the evolution of intelligence*. New York: Bantam Books.
- Carstairs-McCarthy, Andrew. 2005. The evolutionary origin of morphology. In Tallerman, 166-184.
- Chomsky, Noam. 1995. *The minimalist program*. Cambridge, Mass.: MIT Press.
- Chomsky, Noam. 2000. Minimalist inquiries: the framework. In R. Martin, D. Michaels, J. Uriagereka (eds.) *Step by Step*, 89-115. Cambridge, Mass.: MIT Press.
- Christensen, Morton H. & Simon Kirby (eds.) 2003. *Language evolution*. Oxford UP.
- Comrie, Bernard. 2006. Transitivity pairs, markedness, and diachronic stability. *Linguistics* 44: 303-318.
- Corballis, Michael C. 2003. From hand to mouth: The gestural origins of language. In Christensen & Kirby, 201-218.
- Diamond, Jared & Peter Bellwood. 2003. Farmers and their languages: the first expansions. *Science* 300: 597-603.
- Enard, Wolfgang et al. 2002a. Intra- and interspecific variation in primate gene expression patterns. *Science* 296: 340-343.
- Enard, Wolfgang et al. 2002b. Molecular evolution of FOXP2, a gene involved in speech and language. *Nature* 418: 869-872.
- Goldin-Meadow, Susan. 2005. Watching language grow. *Proceedings of the National Academy of Science* 102: 2271-2272.
- Grice, H. Paul. 1969. Utterer's meaning and intentions. *Philosophical Review* 78: 147-177.
- Hauser, Marc D., Noam Chomsky & W. Tecumseh Fitch. 2002. The faculty of language: what it is, who has it, and how did it evolve? *Science* 298: 1569-1579.
- Hockett, Charles. 1960. The origin of speech. *Scientific American* 203: 88-96.
- Hockett, Charles. 1966. The problem of universals in language. In Greenberg (ed.) *Universals of language*, 1-29. MIT Press.
- Hurford, James R. 2000. The emergence of syntax. In C. Knight, M. Studdert-Kennedy, and J.R. Hurford (eds.) *The evolutionary emergence of language: Social function and the origins of linguistic form*, 219-230. Cambridge UP
- Indefrey, Peter. 2002. *Listen und Regeln. Erwerb und Repräsentation der schwachen Substantiv-Deklination des Deutschen*. Dr.phil. Dissertation, Universität Düsseldorf.
- Klein, Wolfgang & Clive Perdue. 1997. The Basic Variety, or: Couldn't language be much simpler? *Second Language Research* 13: 301-347.

- Kirby, Simon. 2002. Learning, bottlenecks, and the evolution of recursive syntax. In E. Briscoe (ed.) *Linguistic evolution through language acquisition: Formal and computational models*, 173-204. Cambridge UP.
- Krifka, Manfred. 2007. Functional similarities between bimanual coordination and topic/comment structure. In S. Ishihara, S. Jannedy & A. Schwarz (eds.) *Working Papers of the SFB 632, Interdisciplinary Studies on Information Structure (ISIS) 8*, 39 - 59. Potsdam: Universitätsverlag.
- Lai, Cecilia S. et al. 2001. A forkhead-domain gene is mutated in a severe speech and language disorder. *Nature* 413: 519-523.
- Martínez, Ignacio et al. 2004. Auditory capacities in Middle Pleistocene humans from the Sierra de Atapuerca in Spain. *Proceedings of the National Academy of Sciences* 101: 9976-9981.
- Meier, Richard P. & Elissa L. Newport. 1990. Out of the hands of babes: On a possible sign advantage in language acquisition. *Language* 66: 1-23.
- Müller, Gereon. 2000. Optimality, markedness, and word order in German. *Linguistics* 37: 377-818.
- Müller, Gereon. 2002. Remarks on nominal inflection in German. In Ingrid Kaufmann & Barbara Stiebels (eds.) *More than Words*, 113-145. Berlin: Akademie Verlag.
- Müller, Gereon. 2007. A radically non-morphemic approach to bidirectional syncretism. *Workshop on Theoretical Morphology 3*, Leipzig.
- Nichols, Johanna. 1992. *Linguistic diversity in space and time*. Chicago UP.
- Ott, Dennis. 2007. Reverse-engineering the language faculty. Origins and implications of the minimalist program, forthcoming, *Harvard Working Papers in Linguistics*.
- Oudeyer, Pierre-Yves. 2005. From holistic to discrete speech sounds: the blind snowflake-maker hypothesis. In Tallerman, 68-99.
- Rizzolatti, Giacomo & M. A. Arbib. 1998. Language within our grasp. *Trends in Neuroscience* 21: 188-194.
- Rizzolatti, Giacomo, Luciani Fadiga, Vittorio Gallese, & Leonardo Fogassi. 1996. Premotor cortex and the recognition of motor actions. *Cognitive Brain Research* 3: 131-141.
- Robins, R. H. 1959. In defense of WP. *Transactions of the Philological Society* 57:116-144.
- Sacks, Harvey, Emanuel Schegloff & ?? Jefferson. 1974. A simplest systematics for the organization of turn-taking for conversation. *Language* 50: 696-735.
- Studdert-Kennedy, Michael. 2005. How did language go discrete? In Tallerman, 48-67.
- Tallerman, Maggie (ed.) 2005. *Language origins. Perspectives on evolution*. Oxford UP.
- Traill, Anthony. 1985. *Phonetic and phonological studies of !Xóo Bushman*. Hamburg: Buske.
- Zhivotovsky, Lev A., Noah A. Rosenberg & Marcus W. Feldman. 2003. Features of evolution and expansion of modern humans, inferred from genomwide microsatellite markers. *American Journal of Human Genetics* 72: 1171-1186.