

Peter Auer (Freiburg)

Syntax als Prozess^{*}

1. Welche Syntax braucht die Gesprochene Sprache?

In den ersten Jahrzehnten der Mündlichkeits-/Schriftlichkeitsforschung ist oft die Frage gestellt worden, ob wir beim mündlichen Kommunizieren eine andere Grammatik verwenden, als wenn wir in zeitlich zerdehnter Form schriftlich kommunizieren. Es wurde darüber diskutiert, ob mündliche und schriftliche Syntax zwei unterschiedliche Systeme darstellen, es wurden syntaktische Phänomene beschrieben, die nur oder präferenziell im einen oder anderen Modus vorkommen, und es wurden diese Phänomene aus den spezifischen Bedingungen des jeweiligen Mediums erklärt. Es ging also darum, ob die Gesprochene Sprache¹ eine eigene Syntax *hat*. Im Gegensatz dazu stellt sich der vorliegende Beitrag die Frage nach der *Form* der syntaktischen Beschreibung, zielt also auf eine *Theorie* der gesprochenen Syntax ab. Wie müsste eine syntaktische Theorie aussehen, die möglichst realitätsnah ('realistisch'; Auer 2004) die vorgefundenen Phänomene der Mündlichkeit erfassen kann? Die Frage präsupponiert, dass die Erforschung der Gesprochenen Sprache eine spezifische grammatische Beschreibungstechnik braucht, die nicht immer schon in der 'allgemeinen' syntaktischen Theorie enthalten ist. Eine solche modalitätsspezifische Grammatiktheorie geht über die gängige Grammatikschreibung hinaus, die eine vom Realisierungsmodus unabhängige sprachlich-grammatische Kompetenz der Sprecher im Auge (quasi eine *offline*-Grammatik) im Auge

* Mein Dank für zahlreiche Hinweise zu Abschnitt 4 geht an die Teilnehmer und Teilnehmerinnen eines Forschungsseminars zum Thema Construction Grammar, besonders an Karin Birkner, Christa Lenz, Göz Kaufmann, Dagmar Frohning und Fabian Overlach. Dem Herausgeber dieses Bands, Heiko Hausendorf, danke ich für zahlreiche Hinweise zum gesamten Text.

¹ Die Großbuchstaben bei Gesprochene Sprache indizieren, dass es sich um das handelt, was oft auch mit « konzeptioneller Mündlichkeit » bezeichnet wird.

hat. Ihre Modellierung, wie sie heute von fast allen Grammatikern, gleich welcher theoretischer Ausrichtung, angestrebt wird, ist für viele Zwecke ausreichend, auch wenn ihr der Vorwurf gemacht werden kann, in der Praxis doch eher an schrift- und daher standardsprachlichen als an mündlichen Strukturen orientiert zu sein. Sie reicht jedoch nicht aus, um die spezifische Differenz der mündlichen Realisierungsform von Sprache (wie übrigens auch nicht die spezifische Differenz schriftlicher Sprache) zu erfassen.

Im Folgenden soll gezeigt werden, dass eine modalitätsangemessene Syntaxbeschreibung der mündlichen Sprache entsprechend deren spezifischen Produktions- und Rezeptionsformen insbesondere drei Eigenschaften haben muss:

– sie muss ‘inkrementell’ sein, d.h. den syntaktischen Strukturaufbau in der ‘Echtzeit’ erfassen; denn mündliche Sprache wird linear in der Zeit produziert und rezipiert (und unterscheidet sich damit von der Zweidimensionalität des geschriebenen Textes);

– sie muss dialogisch orientiert sein, d.h. sie muss die Kooperation der Gesprächsteilnehmer bei der Ko-Konstruktion syntaktischer Einheiten berücksichtigen; denn mündliche Sprache in direkten Interaktionen ist durch maximale Synchronisiertheit von Produktion und Rezeption gekennzeichnet, die Rückkopplungen zwischen beiden nicht nur möglich macht, sondern immer schon impliziert;

– sie muss der Tatsache Rechnung tragen, dass mündliches Kommunizieren auch unter hohem Zeit- und Handlungsdruck deshalb funktioniert, weil viele, auch scheinbar komplexe Syntaxstrukturen bereits mehr oder weniger stark musterhaft festgelegt sind. Statt sie kompositionell (‘generativ’) anhand relativ allgemeinen Regeln aufzubauen, muss eine realistische Beschreibung mündlicher Syntax dieser Verfestigung häufig gebrauchter syntaktischer Muster (im Sinne der *construction grammar*; siehe unten, Abschn. 4) Rechnung tragen.

Die drei Anforderungen an die Modellierung mündlicher Syntaxstrukturen entsprechen also drei Merkmalen der prototypischen mündlichen Kommunikation: Linearität in der Zeit, Synchronisierung der

Handlungsabläufe zwischen Sprecherin und Hörer in der face-to-face-Interaktion und an Mustern (*constructions*) orientierter Ablauf. Sie werden in den folgenden drei Kapiteln genauer behandelt. Besonders der inkrementelle, emergente Charakter mündlicher Syntaxkonstruktionen (*online*-Grammatik, vgl. Auer 2000) macht Syntax zu einem Prozess, der unter bestimmten kognitiven und interaktiven Bedingungen (in der Regel dialogisch) abläuft.

2. Inkrementelle Syntax

Inkrementelle Syntax beschreibt die fortlaufenden Projektionen über den weiteren Verlauf der emergenten syntaktischen Struktur, die es den Hörern erlauben, den entstehenden Redebeitrag ohne Verzögerung zu prozessieren. Syntaktische Projektionen bauen auf syntaktischen ‘Gestalten’ auf, die, sobald sie identifiziert sind, nach dem gestaltpsychologischen Prinzip der „guten Fortsetzung“ durch die Produktion einer mehr oder weniger präzise vorhersagbaren Abschlussstruktur geschlossen werden müssen. Syntaktische Projektionen sind für die Interaktion nicht zuletzt deshalb von großer Wichtigkeit, weil sie die Vorhersage von möglichen Redezug-Abschlusspunkten ermöglichen. Sie kommen aber auch innerhalb von Turnkonstruktionseinheit zum Tragen.²

Inkrementelle syntaktische Analyse impliziert die Abkehr von der reinen *top-down*-Analyse nach dem Modell der *immediate constituents* zugunsten einer Kombination von linearer und hierarchischer Beschreibung (vgl. dazu Kindt 2003). Die grundlegenden Eigenschaften von Projektionsverfahren (sowie die komplementären retraktiven syntaktischen Verfahren) sind bereits andernorts (Auer 2000, Auer im Druck, Auer 1996; vgl. auch Stein 2003) ausführlich dargestellt worden (vgl. auch Couper-Kuhlen, in diesem Band). Hier sollen lediglich einige sprachvergleichende Aspekte genannt werden. Syntaxen von Sprachen unterscheidet sich in ihrem Projektionspotential. Zum Beispiel gilt das Japanische als eine projektionsarme (vgl. Ford, Fox &

² Psycholinguistische Evidenz für *online*-Prozessierung von Sprache, aber auch für deren projektiven Charakter, liefern indirekt Forschungen im Paradigma des *shadowing*, also des Mitsprechens während der spontanen Sprachproduktion eines Anderen (vgl. den klassischen Aufsatz von Marslen-Wilson 1985). Dies ist nur möglich, wenn fortlaufend Fortsetzungserwartungen aufgebaut und in der Regel auch eingelöst werden.

Thompson 2003:130f mit weiteren Literaturhinweisen), das Deutsche als eine projektionsstarke Sprache. Bei einem solchen Sprachvergleich muss man allerdings berücksichtigen, dass Interaktionsteilnehmer neben syntaktischen über andere Möglichkeiten der Projektion (semantisch-pragmatische, prosodische und non-verbale) verfügen. Hat eine Sprache also weniger gute syntaktische Projektionsmöglichkeiten, so kann daraus nicht geschlossen werden, dass keinerlei Projektionen über den weiteren Verlauf von Äußerungen möglich sind – sonst wäre Interaktion gar nicht möglich. Sie können auch auf anderen sprachlichen Ressourcen aufbauen als den syntaktischen.

Für den Sprachvergleich spielen unter anderem die folgenden Parameter eine Rolle:

- das Ausmaß der Serialisierungsvorschriften (entsprechend dem Unterschied zwischen sog. konfigurationellen und nicht-konfigurationellen Sprachen; vgl. z.B. Maracz & Muysken (Hrsg.) 1989). Es ist unmittelbar einsichtig, dass eine Sprache mit rigiden Serialisierungsvorschriften präzisere syntaktische Projektionen erlaubt als eine mit freier Wortstellung. Dies betrifft vor allem Adjazenzprojektionen (vgl. unten) von einer syntaktischen Position zur nächsten. Im klassischen Latein lässt sich z.B. aus dem Konstruktionsbeginn *Gallia ...* keine Fortsetzungserwartung für die nächste Position ableiten (*Gallia est...*, *Gallia omnis ...*, *Gallia divisa...*, *Gallia in...*); im Deutschen folgt hingegen auf den Konstruktionsbeginn *Frankreich ...* – sieht man von immer möglichen parenthetischen Erweiterungen und Prolepsen ab – mit großer Sicherheit ein finites Verb in der nächsten (d.h. der ersten Klammer-) Position.
- ‘Links- und rechtsverzweigende’ Strukturen stehen für völlig unterschiedliche Produktions- und Rezeptionstechniken und sind deshalb ein wichtiger Parameter für den Sprachvergleich aus der Perspektive der *online*-Syntax. Er interagiert mit dem Parameter ‘head marking’ vs. ‘dependent marking’ (Nichols 1992). Dies lässt sich am Beispiel der Attribution zeigen. Für Sprachen mit Modifikatormarkierung (*dependent marking*) erlaubt die Serialisierung Modifikator vor Kopf die Projektion des Kopfs; etwa ermöglicht im Deutschen ein flektiertes Adjektiv die Vorhersage, dass (immer abgesehen von kontextuellen ‘Ellipsenlizenzen’) ein Nomen folgen

wird. Rechtsverzweigende Strukturen, also die Serialisierung Kopf vor Modifikator (z.B. etwa im Deutschen Erweiterungen von Nominalphrasen durch Relativsätze), sind hingegen in Sprachen dieses Typs überhaupt nicht vorhersagbar, weil am Kopf die Attribution nicht markiert wird. (Dementsprechend ist dt. A + N: *ein schönes* → N_{neutr} stärker projizierend als franz. N + A: *un livre* → ?). Für kopfmarkierende Sprachen gilt hingegen das umgekehrte.

Ein Sonderfall der Kopf/Modifikator-Beziehung ist die Rektion. Voranstehende regierende Elemente ermöglichen die beste Vorhersage über nachfolgende regierte Elemente (deren Abstand vom regierenden Kopf und deren Reihenfolge – wenn es mehrere Ergänzungen gibt – natürlich variabel sein kann und dann nicht vorhersagbar ist). So ermöglichen im Deutschen verbinitiale Syntagmen (wiederum unter Berücksichtigung lokaler ‘Ellipsenlizenzen’) die Vorhersage des Kasus und der Anzahl nachfolgender nominaler Ergänzungen (*kommt* → *ein Mann in einen Laden...*), Präpositionen ermöglichen die Vorhersage von Anzahl und Kasus der folgenden Nominalphrasen (*zwischen* → *Himmel und Erde*), etc. Bei umgekehrter Serialisierungsrichtung ist die Vorhersage des Kopfes aus der Ergänzung nur unter bestimmten Bedingungen möglich (*(dass) ein Mann in einen Laden* → *kommt*), nämlich dann, wenn die Ergänzungen morphologisch markiert sind und so ihre syntaktische Funktion erkennbar wird. Syntaxen unterscheiden sich aber in der Durchsichtigkeit der Abbildung der semantischen Rollen auf die morphologischen Kasus. Deshalb ist *(dass) dem Linguisten die neue Grammatik* → *gefällt* stärker projizierend als *(dass) Fritz Maria* ? → *gefällt, liebt, ...* Morphologisch reichere Sprachen sind also bei der Serialisierung ‘regiertes Element vor regierendem Element’ projektionsstärker als morphologisch ärmere. Man kann dies auch daran erkennen, dass sog. *garden path sentences* (vgl. z.B. Ferreira & Henderson 1991), bei denen die syntaktische *online*-Prozessierung besonders schwierig ist, im Englischen viel häufiger sind als im Deutschen. Die deutschen psycholinguistischen Lehrbücher greifen deshalb meist zu recht gekünstelten Beispielen wie *ich glaube dass Fritz zugunsten / von Maria nie etwas unternommen* → *worden wäre* vs *ich glaube dass Fritz / zugunsten von Maria nie etwas unternommen* → *hätte* um das Phänomen zu illustrieren. Wenn man im Beispielsatz die Eigennamen durch

kasusmarkierte NPs ersetzt, lässt sich diese Sprachabhängigkeit sofort erkennen.

– wie sich schon am Beispiel der Verbalrektion zeigen ließ, ergibt sich aus der hierarchischen Struktur von Syntax, dass sich Projektion nicht auf Übergangswahrscheinlichkeiten zwischen adjazenten Elementen reduzieren lässt. Vielmehr können auch nicht-adjazente Strukturen projiziert werden. Ein wichtiges Projektionsverfahren sind daher diskontinuierliche Konstituenten. Die deutsche Satzklammer, ein Distanzprojektionsverfahren par excellence, ist dafür allerdings nur dann geeignet, wenn das linke Klammerelement das rechte erwartbar macht. Deshalb sind zwar Strukturen wie

ich komme morgen um 8 Uhr am Hauptbahnhof → an

wir werden zu spät → kommen

projizierend, nicht aber

ich will noch ein Brötchen → haben

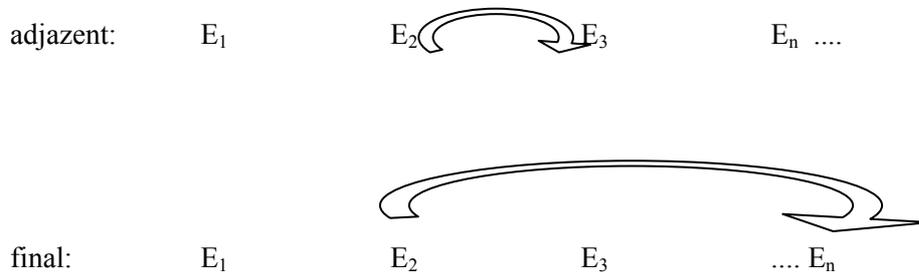
du hast Besuch → bekommen

ich komme erst morgen → an.

In der zweiten Beispielgruppe ist schon vor dem letzten Klammerelement ein möglicher syntaktischer Abgeschlossenheitspunkt erreicht.³

– Adjazenprojektionen beziehen sich auf den Übergang von einem syntaktischen Strukturelement zum nächsten, also von Wort zu Wort oder Phrase zu Phrase, Distanzprojektionen auf die Vorhersagbarkeit nicht nebeneinander stehender Elemente. Ein Sonderfall der Distanzprojektion ist die Endprojektion, d.h. die Vorhersage des Abschlusselementes einer selbständigen syntaktischen Einheit von einem bestimmten Punkt an:

³ Selbstverständlich ist die Vorhersagbarkeit des rechten Klammerelements hier auch von den im Vor- und Mittelfeld stehenden Ergänzungen bestimmt.



(für n 'Ereignisse', etwa Produktionen eines Wortes, die sich miteinander verknüpfen, bis die syntaktische Struktur/Gestalt mit dem n -ten Ereignis abgeschlossen ist).

Die kognitiven Prozessierungsbedingungen, die in der direkten Kommunikation relevant sind, sprechen gegen zu weit ausgreifende Projektionsbögen und für Projektionen, die relativ nah am Adjazenzpol liegen. Neue zu prozessierende Elemente werden im unmarkierten Fall als projektionseinlösend für die unmittelbaren Vorgängerelemente verstanden (vgl. Frazier & Rayner 1982), wobei diese Adjazenz allerdings auf verschiedenen syntaktischen Hierarchieebenen definiert werden kann (z.B. auch auf Phrasenniveau). Größere (längere) Syntagmen werden in der Gesprochenen Sprache gern durch mehrfache lokale Übergänge hergestellt. Andererseits sprechen Gesichtspunkte der Gesprächsorganisation dafür, den *Abschluss* von Turnkonstruktionseinheiten (TCUs) projizierbar zu machen. Dadurch lässt sich nicht nur das Rederecht für die Sprecherin langfristig sichern. Dem Hörer bleibt auch ausreichend Zeit, um den nächsten Redebeitrag zu planen und exakt zu platzieren. Interaktiven Vorteilen der Finalprojektion können also kognitive Prozessierungsnachteile gegenüberstehen. Dieser Konflikt erklärt, warum selbst Sprachen, in denen Finalprojektionen eine große Rolle spielen (etwa das Deutsche), für lange, z.B. satzwertige Konstruktionen wie Relativsätze dieses Verfahren aufgeben und stattdessen mehr lokal operierende Erweiterungen zulassen (cf. Antinucci, Duranti & Gerbert 1979).

Strikte SOV-Sprachen, in denen dem finiten Verb in Endstellung keine Konstituenten mehr folgen können, benutzen dieses, um den Abschluss einer

syntaktischen Gestalt zu markieren. Allerdings gilt das wohl für viele SOV-Sprachen nur tendenziell: auch im Türkischen (vgl. Auer 1990) oder Japanischen (Couper-Kuhlen & Ono, im Dr.) sind Erweiterungen über das Finitum hinaus gerade in der Gesprochenen Sprache möglich und üblich.

Wie sieht nun das grammatische Wissen über Projektionen aus? Relativ leicht ist diese Frage für solche Konstruktionen einer Sprache zu beantworten, die bestimmte Konstituentenabfolgen kontextfrei erzwingen. Etwa sind die deutschen Artikelformen grundsätzlich projizierend (was ihrer Eigenschaft entspricht, ohne Ausnahme vor dem Bezugsnomen und außerdem vor allen anderen Modifikatoren des Nomens zu stehen). Dasselbe gilt für die nebenordnenden Konjunktionen des Deutschen, die nachfolgende satz- oder phrasenwertige koordinierte Strukturen projizieren, und für die subordinierende Konjunktionen, die nachfolgende Verbletztsyntaxen erwartbar machen. Dass solche Information kontextfrei gespeichert werden kann, schließt natürlich Ambiguitäten nicht aus; eine definite Artikelform kann sich im Laufe der *online*-Prozessierung auch als Relativpronomen oder rhematisches Personalpronomen entpuppen, die subordinierende Konjunktion *weil* als koordinierende, etc. Hier ändert sich die grammatische Kategorie und damit die syntaktische Funktion des Wortes.

Andere Strukturen entfalten jedoch ihr Projektionspotential erst im syntaktischen Zusammenhang, d.h. die Abgeschlossenheit einer Konstruktion (und damit Syntax insgesamt) ist kontextabhängig. Vgl. dazu das folgende Beispiel:

(Big Brother, Zlatko und Jürgen)

((wenige Sekunden vorher hat eine andere Containerbewohnerin mit Zlatko darüber gesprochen, dass Kerstin bei den Männern übernachtet hat; alles fast flüsternd))

```
01  Jrg:      <<pp>die KERStin?
02  Zlt:      <<pp>>ja,>
03  Jrg:      <<pp, presto>wo WAR die denn;>
04  Zlt:      <<pp>beim alex im BETT;>
05  Jrg:      <<mp>quatsch;>
```

- 06 Zlt: <<spöttisch lachend> ha natÜrlich;>
 07 Jrg: → hab ich nicht geSEHN.
 08 Zlt: <<spöttisch lachend> Logo.>

Die kontextfreie Valenzstruktur des Verbs *gesehen* in Z. 07 erfordert eine Objektergänzung, die der Sprecher nicht liefert. Dennoch bildet seine Äußerung im vorliegenden sequenziellen Zusammenhang zweifelsohne eine vollständige Turnkonstruktionseinheit, d.h. sie projiziert keine syntaktische Fortsetzung. Die syntaktische Gestalt ist abgeschlossen, nachdem das distanzprojizierte Abschlusselement (das Partizip *gesehen* in der 'rechten' Klammerposition) erreicht worden ist. Dies ist möglich, weil die Objektergänzung in der vorliegenden Konstruktion (verbinitiales Syntagma mit Personalpronomen in der zweiten Position) aus dem vorausgehenden Kontext inferiert werden kann.

Ein anderer Ausschnitt aus 'Big Brother' beginnt wie folgt:

(Big Brother)

- 01 Friseur: jetzt gehn ma mal zu jürgen,
 02 Sbr: ((lacht los))
 03 → zu jürgen;

Die bestätigende Wiederholung eines Teils der Äußerung des ersten Sprechers durch die zweite (*zu jürgen*) ist syntaktisch gesehen von dieser abhängig. Nur im gegebenen Kontext ist sie eine vollständige Turnkonstruktionseinheit. Stünde dieselbe Äußerung am Beginn einer Gesprächsepisode, würden sich die Projektionsverhältnisse sofort verändern: die Präpositionalphrase *zu Jürgen* würde dann weitere Äußerungsteile erwartbar machen.

Hier - wie in der traditionelleren Forschung zur Gesprochenen Sprache - von Ellipsen zu sprechen, suggeriert fälschlicherweise, dass solche Äußerung aus ihren vollständigen Pendants sozusagen durch Wegstreichen von Strukturelementen abgeleitet werden. Wesentlich sinnvoller ist eine andere Sichtweise: syntaktische Strukturen bleiben nach ihrer Produktion/Rezeption

für eine gewisse Zeit verfügbar und halten für den nächsten (oder denselben) Sprecher ein ‘Angebot’ bereit, sie nach bestimmten Regeln für die Nachfolgeäußerung zu nutzen. Trotz der offensichtlichen Flüchtigkeit mündlicher Sprache ist das produzierte Syntagma also nicht schon im Augenblick seiner vollständigen Produktion irrelevant; es bleibt vielmehr zunächst noch strukturell (und natürlich auch inhaltlich) verfügbar. In diesem Sinn gehören im Beispiel die Äußerung des Friseurs in Z. 01 und die Sabrinas in Z. 03 zusammen zu einer syntaktischen Struktur.⁴

Syntaktische Latenzen haben, wie Alles in der Gesprochenen Sprache, eine Zeitstruktur. Sie verlieren ihre Wirksamkeit rapide, wenn sie nicht unmittelbar nach dem Bezugssyntagma ausgenutzt werden. Dennoch sind auch weitläufigere Latenzen möglich. Betrachten wir die Äußerung

05 Adr: die uta au nich.

Es ist klar, dass wir es hier trotz prosodischer Abgeschlossenheit syntaktisch gesehen mit einem Syntagma zu tun haben, das auf ein früheres zurückverweist und dessen noch latent vorhandene Syntax ausnutzt. Die unmittelbare Vorgängeräußerung ist jedoch:

04 Jrg: da werd=isch au fuchsteufelswild.

05 Adr: die uta au nich.

Jürgens Äußerung kann nicht der Bezugspunkt für Andreas Äußerung sein, denn es gibt keine strukturelle Beziehung zwischen den beiden, was im vorliegenden Fall an der Negationspartikel im Zusammenhang mit dem Adverb *au(ch)* liegt. Geht man im Gesprächsausschnitt weiter zurück, so findet man jedoch einen solchen Bezugspunkt:

(Big Brother, Sabrina und Jürgen)

01 Sbr: lebensmittel weg das is ne SÜNde.⁵

02 Jrg: also das mag=isch au nisch.

⁴ Vgl zu dieser Herangehensweise an sog. Ellipse aus psycholinguistischer Sicht auch Kindt 2003.

- 03 Sbr: nee-
04 Jrg: da werd=isch au fuchsteufelswild.
05 ___ Adr: die uta au nich.

Andreas Äußerung baut also auf der Struktur einer Äußerung Jürgens auf, die zumindest über zwei andere Äußerungen hinweg latent verfügbar bleibt.

Projektionen sind die Basis der inkrementellen Syntax, die für die Gesprochene Sprache so zentral ist. Der Wechsel zwischen strukturaufbauenden (Projektionsbögen herstellenden) und strukturarbeitenden (Projektionen einlösenden, Projektionsbögen abschließenden) Phasen konstituiert einen kognitiven Rhythmus, der die Prozessierung erleichtert, indem er in den strukturarbeitenden Phasen Entlastungen für andere mentale und interaktive Aktivitäten schafft; zum Beispiel die Vorbereitung des nächsten Turns. Entsprechend ist die Möglichkeit syntaktischer (wie auch anderer) Projektionen in der Zeit grundsätzlich ein Vorteil; problematisch sind hingegen Strukturen, die lange Zeit ambig bleiben und keine Vorhersagen über den weiteren syntaktischen Verlauf ermöglichen. Ein Extremfall dafür sind die schon genannten *garden path sentences*.

Man könnte einwenden, dass solche Projektionen ja auch in der schriftlichen Kommunikation, also beim Lesen und beim Schreiben, gelten. Bei sehr ungeübten Lesern und Leserinnen ist das in einem bestimmten Sinn richtig; Erstklässler arbeiten beim Lesen in einem Satz Wort für Wort bis zu dem Punkt ab, an dem sie seine grammatische Struktur erkennen und daher Projektionen über den weiteren syntaktischen Verlauf des Satzes machen können. Dies erkennt man oft daran, dass sie dann das bisher Gelesene mit anderer (passenderer) Intonation und rhythmisch integriert wiederholen. Geübte Leser prozessieren Texte hingegen nicht linear, sondern flächig, d.h. sie können in der Regel mindestens mehrere Wörter, wenn nicht mehrere Zeilen des Textes zugleich überschauen und relevante syntaktische Bezugspunkte erkennen. Die Prozessierung erfolgt nicht notwendigerweise linear in der Zeit, sondern kann mit großflächigen Vor- und Rückgriffen arbeiten. Für solche Leser ist Lesen deutlich schneller als Zuhören, und die

⁵ Gemeint ist: 'Lebensmittel wegzuerwerfen ist eine Sünde.'

02 Jrg: das sind wir ALle.
 02 → Adr: wenn ich zu hause sehe;
 03 IHR ja noch nich ma;
 04 → (was) ICH alles wegschmeiß,
 05 (3.0)
 06 → Sbr: das=s ne SÜnde.

Die zweite Sprecherin vervollständigt also die begonnene, aber unvollständige⁶ Struktur der ersten. Anders im folgenden Beispiel: hier schließt sich ein syntaktisch nicht selbständiger, kausaler Nebensatz Jürgens (*weil de no nie in der situation warst*) an eine schon abgeschlossene Konstruktion Andreas (*auf son gedanken wär ich nie gekommen*) an:

(Big Brother)

01 Adr: <<h>da:nn (-) nimmt er sogar die
 früchte aus dem tee noch;>
 02 IN sein müsli.
 03 (-)<<stimmlos;all>sa ma auf so=n>
 gedanken[wär ich NIE:] gekommen.
 04 → Sbr: [aus dem beutel];
 05 Adr: nee:.,
 06 frischer is das.
 07 Jrg: ja;
 08 → wahrscheinlich weil de no nie in der
 situation warst
 09 dass de sowas machen MUSStest.
 10 Adr: ja

⁶ Dass Andrea nach dem *wenn*-Satz eine längere Pause entstehen lässt, legt nahe, dass sie selbst keine Fortsetzungsabsicht hatte. Solche Fälle von Aposiopese laden zur tatsächlichen oder 'gedanklichen' Fortsetzung der begonnenen Struktur durch die Rezipienten ein und können strategisch verwendet werden.

Jürgens *wahrscheinlich weil de no nie in der situation warst* ist strukturell auf die Vorgängeräußerung *auf son gedanken wär ich nie gekommen* bezogen und wäre ohne sie nicht in der Lage, eine Turnkonstruktionseinheit zu bilden. Allerdings zeigt das Beispiel auch, dass vom Rezipienten kookonstruierte syntaktische Konstruktionen nicht mit allein vom Sprecher konstruierten identisch sind: denn Jürgens Beitrag weist trotz aller Abhängigkeit von Andreas Äußerung strukturelle Merkmale auf, die ihn zu seiner eigenen ‘Stimme’ machen. Dazu gehört neben dem epistemisch-modalen Adverb *wahrscheinlich*, das nur aus der Perspektive des Anderen, nicht der Sprecherin sinnvoll ist, vor allem die pronominale Verschiebung *ich* → *du (de)*.

Dialogisch konstruierte Strukturen liefern oft gute Evidenz für die interaktionale Realität der im Abschnitt 2 beschriebenen Projektionsverfahren. Im folgenden Ausschnitt nutzt Manuela ein Zögern in Johns Redebeitrag, um ihn zuende zu führen. Simultan vervollständigt John selbst seinen Redebeitrag mit fast denselben Worten und der annähernd gleichen syntaktischen Struktur:

(Big Brother)

01 Jhn: et jibt halt wirklich ooch
missverständnisse;=ja
02 also WEESS ick nich
03 wenn de dir .h
04 mal irgendwie ne BRAvo kuckst
05 und da fragen halt heutz[ta]ge IMmer
noch jugendliche
06 Adr: [ja]
07 Jhn: kam=man sich beim ersten mal schon
ANstecken
08 oder kann man beim (.)
→ beim beim [ersten geschlechtsverkehr]
09 → Man: [ersten mal SCHWANger werden]

10 Jhn: SCHWANger werden
 11 weeß ICke wat .h

Solche Kollaborationen sind nur möglich, weil die Rezipienten eng synchronisiert mit den Sprechern Projektionen über den weiteren Verlauf der emergenten Struktur vornehmen, die es ihnen ermöglichen, schnell und strukturell korrekt in diese Emergenz einzugreifen.

Sogar ein schweigender Rezipient kann die emergente Syntax des Sprechers beeinflussen. Etwa hält der Rezipient im folgenden Beispiel in Zeile 03 jedwede Reaktion auf die Äußerung des ersten Sprechers zurück:

(Telefongespräch)

01 A: ja der muss früh wieder HEIM
 02 weil der=hat abns terMIne ne,
 03 → (o.7)
 04 → mit em be BE da:

Erwartbar wäre nach der Nachlaufpartikel *ne* in Zeile 02 zumindest ein Rezipientensignal gewesen, das die Erklärung, warum 'er' nicht mitkommen kann, ratifiziert und dem augenblicklichen Sprecher A die Erlaubnis zum Weitersprechen gibt. K schweigt jedoch, und es entsteht eine 0.7-sekündige Pause. In solchen sequenziellen Positionen, also als Reaktion auf das Fehlen einer erwartbaren Folgehandlung des Anderen, produzieren erste Sprecher systematisch Elaborierungen oder Reparaturen ihrer bisherigen Äußerungen. Dies geschieht auch im vorliegenden Fall in Zeile 04, wo A seine bisherige, schon vollständige Äußerung syntaktisch im Format einer Ausklammerung erweitert. Man kann also mit Fug und Recht sagen, dass die erweiterte Struktur der letztendlich produzierten Äußerung von der Nicht-Handlung des Gesprächspartners - als Handlung! - (mit) gesteuert wird.

4. Konstruktionen (*constructions*)

In generativer Denkweise sind viele Grammatiker gewohnt, unter alternativen grammatischen Beschreibungen derjenigen den Vorzug zu geben, die die stärksten Generalisierungen erlaubt. Die Struktur von Einzeläußerungen wird deshalb so weit wie möglich aus der Anwendung und Interaktion allgemeiner Strukturregularitäten ('Regeln', 'Beschränkungen') erklärt. Aus der Sicht der Sprachproduktion und -rezeption unter den Bedingungen der direkten mündlichen Kommunikation ist jedoch eine solche, abstrakte Syntax wenig wahrscheinlich. Die oben skizzierte inkrementelle Herangehensweise baut hingegen auf sprachlichem Wissen über relativ detaillierte, oberflächennahe und redundante Konstruktionschemata auf. Je größer der Detaillierungsgrad und damit auch die Anzahl solcher Konstruktionschemata ist, um so leichter kann der Rezipient die entstehende Struktur einem solchen Schema zuordnen und ihren weiteren Verlauf vorhersagen. Eine Grammatiktheorie, die sich mit solchen Formen grammatischen Wissens systematisch beschäftigt, ist die *construction grammar*. Es gibt sie in zahlreichen Ausprägungen (Croft 2001, Deppermann 2004, Fillmore, Kay & O'Connor 1998, Goldberg 1995, Michaelis & Lambrecht 1996), von denen diejenige für die Erforschung der Gesprochenen Sprache am attraktivsten ist, die Häufigkeiten mit einbezieht (z.B. Ford, Fox & Thompson 2003, Bybee 1998). Dahinter steckt die Idee, dass die Sprachbenutzer rekurrente Anwendungen allgemeinerer Regeln mit der Zeit getrennt speichern und direkt abrufen. *Constructions* sind also oft mehr oder weniger idiosynkratische Instantiierungen allgemeinerer syntaktischer Strukturen, die meist bestimmte interaktive Funktionen haben und die als solche im Erstspracherwerb noch vor den allgemeineren Strukturschemata gelernt werden (Tomasello 1998). Der Grad der Idiosynkrasie kann sehr unterschiedlich sein; manche Konstruktionen sind vollständig – auch lexikalisch – spezifiziert, andere lassen bestimmte Abänderungen zu oder enthalten Leerstellen und sind daher abstrakter. Per definitionem ist die Syntax und/oder die Semantik einer Konstruktion nicht kompositionell aus ihren Teilen ableitbar. Die kategorische Trennung zwischen Syntax und Lexikon löst sich auf.

Das Vorgehen der *construction grammar*, nämlich allgemeine Strukturbeschreibungen (Regeln) in stärker restringierte und dafür präziser

in ihrer interaktionalen Funktion beschreibbare Einzelkonstruktionen aufzulösen, soll hier kurz an einem Beispiel vorgeführt werden, nämlich den projizierenden Konstruktionen mit *so*, einem der hochfrequenten Wörter des Gesprochenen Deutsch. Thurmair (2001: 27ff) bezeichnet *so* in einer der wenigen einschlägigen Arbeiten als deiktisches Adverb, das auf Eigenschaften oder „Aspekte“ (Ehlich 1987) verweist. Weinrichs Textgrammatik widmet *so* immerhin einen ganzen Abschnitt und versucht, seine verschiedenen Vorkommensweisen unter den Begriff der Bedeutungsrahmung zu fassen (2003:583ff); darunter ist offenbar *so* etwas wie Hervorhebung gemeint (vgl.: „[...] wird jeweils ein bestimmter Bedeutungsinhalt [...] umrahmt und dadurch emphatisch hervorgehoben“, 584). Diese Beschreibungen resultieren ganz offensichtlich aus dem ehrenwerten, aber vielleicht von vorne herein zum Scheitern verurteilten Versuch, aus einer sehr heterogenen Gruppe von Verwendungskontexten eine möglichst allgemeine Beschreibung abzuleiten. Sie wird aber dadurch *so* abstrakt, dass sie an der Sprachrealität vorbei geht. Schaut man sich die Vorkommnisse von *so* in einem (hier: mündlichen) Corpus genauer an, so löst sich die allgemeine Beschreibung alsbald in eine Vielzahl von wesentlich spezifischeren Konstruktionsschemata auf, die eigene syntaktische, semantische und pragmatische Eigenschaften haben. Der Verweis auf Eigenschaften spielt bei vielen von ihnen keine Rolle; nicht einmal die deiktische Komponente von *so* bleibt immer erhalten. Auch eine allgemeine Rahmungsfunktion lässt sich nicht nachweisen.

Im folgenden gehe ich lediglich auf diejenigen Konstruktionen ein, in denen *so* eine satzwertige Struktur im Folgesyntagma projiziert. (Die zahlreichen Verwendungsformen als alleinstehende Partikel, als alleinstehendes Adverb, in Vergleichsstrukturen des Typs *so ... wie*, als fakultative Einleitung der Apodosis von Konditionalstrukturen sowie als modifizierendes Element vor einem einzelnen Wort werden also nicht berücksichtigt. *Sodass* sehe ich als eigenständige Konjunktion an.⁷) Darunter finden sich mindestens die folgenden:

(a) Konsekutivkonstruktion: ... X ist *so* Y, *dass* S

7 Sie werden teilweise bei Thurmair 2001 sowie bei Ehlich 1987, Burkhardt 1987 und Sandig 1987 beschrieben.

Semantik: Ein Prädikat trifft in einem solchen Maß auf ein Argument zu, dass daraus *q* folgt. Das Prädikat ist oft, aber nicht zwingend evaluativ. Typisch sind Beispiele ohne *complementiser* im Folgesyntagma. (Das projizierende *so* ist in den folgenden Ausschnitten durch Fettdruck hervorgehoben, die projizierte satzwertige Folgestruktur durch einen Pfeil links vor der Zeile.)

(psychotherapeutische Gruppensitzung)

01 U: bei MIR ischs jetzt (-) etwa en jAhr HER
 02 oder des jährt sich jetzt äh (1.0)
 03 ja in ' in der KLInik wa:r
 04 und beim: psychiAter war:
 05 und was weiss ICH wo war
 06 .hh un:d dA GMERKT hab,
 07 =ich bin:- ich bin **SO** weit UNtn;
 08 → ich KANN einfach gar nimmer weiter
 runter,

(Big Brother)

01 Adr: =dann wusst=ich nich;
 02 dann hat er mir gestern erzählt
 03 da:nn hat er so=n früchtete:,
 04 .h ((schluckt))der is **SO** intensiv und
 stark,
 05 → den kann er sich auch dreimal sogar
 aufgießen?

Die Beziehung zwischen den beiden in Folgerelation stehenden Syntagmen kann aber auch in Form eines *dass*-Satzes stärker grammatikalisiert sein:

(psychotherapeutische Gruppensitzung)

01 TM: und (1.0) es KANN sein dass sie

		HILfestellung dafür BRAUch(en)
02		dass sie sich dem loch STELLN.
03	P:	<<p>mhm.> (2.0)
04	TW:	wenn der zweifel SO stark isch
05	→	dass=se denken des bringt mich UM;
06		wenn ich des tue. (1.5)
07	P:	<<p>mhm> (15.5)

In dieser Verwendung trägt *so* immer einen Fokusakzent. Syntaktisch modifiziert es ein Adjektiv/Adverb. Semantisch bezieht es sich nicht auf die Art und Weise, in der etwas zutrifft, sondern auf den Grad, zu dem das Prädikat zutrifft. *So* wird hier also nicht modal, sondern skalar verwendet. Hieraus ergibt sich die konsekutive Interpretation.⁸ Oft ist die Konstruktion emphatisch, was die Nähe zu der Evaluativkonstruktion ohne Folgesyntagma ausmacht (vgl. Alex aus Big Brother: *hatte null bock, also konnten mich SO am arsch lecken*). Die Konstruktion kommt in der Gesprochenen Sprache häufig vor.

(b) Konzessivkonstruktionen: *so* ADJ/ADVB *es/das/der (auch)* V: S

Semantik: Obwohl die in S ausgedrückte Proposition unter den Bedingungen des Vordersatzes erwartbar ist, trifft sie nicht zu. *So* ist in diesem Fall immer unbetont. Das Folgesyntagma ist immer satzwertig, es kann jedoch nicht durch *dass* angeschlossen werden.

Diese Konstruktion kommt in meinem mündlichen Material nur einmal vor, allerdings in einer parenthetischen Variante (die Rahmung der Parenthese wird durch *er is ... hier für mich die größte Herausforderung* hergestellt):

(Big Brother)

01 K: er is **so** blöd das (.) äh KLINGT, oder so .h,

⁸ Die in meinem Corpus nicht belegte Konstruktion mit Infinitiv (*sind Sie so freundlich mich am Bahnhof abzuholen?* bzw. mit identischer Funktion *sind Sie so freundlich und holen mich am Bahnhof ab?*) hat damit nichts zu tun, wie neben der gänzlich anderen Funktion schon die Tatsache zeigt, dass in diesem Fall *so* nicht betonbar ist.

02 aber von diesem (.) ehm (1.2) MENSCHlichen
her

03 ↓ehm; .h

04 wo du wirklich (.) AUStarieren musst und
(0.8) und ne geFÜHL haben musst,

05 für eh wo is (.) die richtige weite; (.)

06 also wo is (.) der richtige punkt oder so?

07 .h (.) .h (.) ehm

08 HIER für mich die GRÖSSte herausforderung.

Die Konstruktion ist im heutigen Deutsch auf ein dem so folgendes Adjektiv oder Adverb beschränkt. Als konzessive Konjunktion (also satzeinleitend), ist die Verwendung von *so... (auch)* archaisch (*so er auch klagte, es wurde ihm nicht geholfen*).

(c) katadeiktischer Verweis auf die Art und Weise eines Zustands, Sachverhalts oder einer Tätigkeit, der/die im Folgesyntagma formuliert (expliziert) wird: V *so, dass* S

Diese Konstruktion entspricht am ehesten der Beschreibung Thurmairs (2001). Sie kommt mit niedriger Frequenz vor. *So* tritt meist in der syntaktischen Funktion eines Satzglieds auf. Der Anschluss erfolgt immer durch einen *dass*-Satz. Wenn der Vordersatz überhaupt eine eigene Intonationsphrase bildet, liegt auf *so* der Fokusakzent (es sei denn, der Negator *nicht* zieht diesen, wie im folgenden Beispiel, auf sich):

(Bewerbungsgespräch, Rollenspiel)

01 B: das kaRAte wird ja (-) .h NICHT **so**
betrieben

02 → dass man (-) seinen partner TRIFFT,
 ((etc.))

(Big Brother)

01 Sbr: ich bin NICHT **so**
 02 → dass ich=s hinter=m [RÜCken] mache
 02 Alx: [gut so]

So kann jedoch auch ein Nomen modifizieren und verschmilzt dann mit einer Artikelform. Diese verwandte, aber formal und pragmatisch nicht identische Konstruktion verweist auf einen selbständigen Nachsatz mit koreferentem Subjekt:

$$X_i \text{ ist } \textit{son} \text{ Y} \rightarrow \textit{s}[\text{Pron}_i \dots]_s$$

Das finite Verb ist immer *sein*, das Folgesyntagma wird als Relativsatz asyndetisch in Hauptsatzform angeschlossen:

(Big Brother)

01 Sbr: <<all>ich bin ja auch **so=n** mensch>
 02 → ich verGESS ja NICHTS (.) ne?
 03 das=s fURchtbar (-)

In einer Variante wird *ich vergess ja nichts* durch einen Relativsatz ... *der nichts vergisst* ersetzt. Hier kann das *so* fehlen. Schließlich kommt auch die wohl innovativste Variante

$$X_i \text{ ist } n \text{ Y} \rightarrow \textit{s}[\text{Pron}_i \dots]_s$$

vor:

(Big Brother)

01 Adr: ich muss schon SAGN
 02 ich trag GERne äh wäh äh (-) schöne
 KLEIdung .h (-)
 03 aber oftmals ist mir das was mir
 [geFÄLLT einfach zu äh]

04 Jrg: [warum MACHST=es dann nicht.]
 05 Adr: es ist (-) mir einfach zu TEUer.
 (0.75)
 06 also=es kann ich mir dann nich
 das kann ich mir nicht [LEISten;
 07 Jrg: [hm=hm
 08 → Adr: .h <<all>aber ich BIN auch n=mensch
 → ich> kann in JEANS rumlaufen (-)
 09 mit SWEATshirt,

Die Pragmatik der Konstruktion (c) besteht darin, eine einfache Äußerung in zwei syntaktische, semantisch-pragmatische und teils auch prosodische Komponenten zu zerlegen. Die rhematische Komponente wird abgespalten und als sozialen Typisierung ausformuliert; das Relevanzzentrum und oft auch das prosodische Zentrum liegen im Nachsatz. Auf diese Weise dient die Konstruktion der (Selbst-)positionierung einer Person, in der Regel des Sprechers bzw. der Sprecherin.

(d) Klassisches korrelierendes *so* mit *wie* als Einleitung des Folgesyntagmas kommt ebenfalls nur mit geringer Frequenz vor. Hier wird im Folgesyntagma ein Vergleich mit der durch *so* indizierten Eigenschaft ausgedrückt (*er ist so (dumm), wie du ihn mir beschrieben hast*).

(psychotherapeutisches Gruppengespräch, eine der Klientinnen richtet sich an eine andere)

01 M: d' du beLEHRST hier grad ALle;
 02 =dass es eigentlich **SO** is,
 03 → <<schneller> wie DU des denkst.=>
 04 und d gibst eigentlich KEInem andern
 ne CHANCE,

Der Vordersatz kann auf das Adverb reduziert werden:

(psychotherapeutische Gruppensitzung)

01 U: des war (da) immer so n ZWANG
 02 wo i no=it net net RAUSkonnt
 03 aso selber nix entSCHEI:dn und selber
 net .hh
 04 (1.0)
 05 immer **so**
 → wie die ANDren halt wolln.
 06 au(ch) mit Andren net nur mit mein
 Eltern;

Dieses eigenschaftskorrelierende *so* kann auch als Modifikator eines Adverbs oder Adjektivs auftreten, so dass das allgemeine Konstruktionsschema (X V) *SO* (Adj/Advb) *wie* S ist.

(e) Sehr häufig ist *so* als Quotativ-Partikel, die eine oder mehrere nachfolgende, asyndetisch angeschlossene Turnkonstruktionseinheiten (die Redewiedergabe) projiziert. Der Vordersatz enthält in dieser Konstruktion oft kein Verbum dicendi, sondern lediglich ein Personalpronomen der 1. oder 3. Person (vgl. Golato 2000):

(Big Brother, Erzählungen vom 'ersten Mal')

01 Man: liegst einfach nur schön ja
 02 und dann meinte er im ernst (--)
 03 warte mal ich kann ja mal kurz
 04 also seinen besten freund (-) em
 05 ich ruf den kurz an dass er kommt und
 kondome bring(h)t
 06 Ker: ((lacht))
 07 Man: und ich [**so**
 08 Alx: [((lacht))
 09 Ker: [((lacht))
 10 → Man: [wie bitte]

11 Alx: [((lacht)]
 12 Ker: [((lacht))]
 13 Man: dann meint ich **so**
 14 → nee (--)
 15 → das meinste jetzt nicht ernst ja .h
 16 ja doch der wohnt ja gleich in=n
 straße weiter
 17 das geht ja schnell und so .h
 18 ich **so**
 19 → na des kannst=e vergessen .h
 20 dann ist der wirklich raus gegangen
 (--) zum telefonieren
 21 in dem moment hab ich mich schon
 wieder angezogen (1)
 22 ja der ist gleich da (-)
 23 ich **so**
 24 → nee s=kannst jetzt echt vergessen (--)
 25 will ich nicht

Auch diese Konstruktion hat nur noch wenig mit der allgemeinen Beschreibung von *so* als deiktischem Adverb zu tun, das auf Eigenschaften oder „Aspekte“ verweist. Es ist immer unakzentuiert

(f) *es ist so, (dass) S*

Bei weitem am häufigsten ist jedoch ein anderer Typ von katadeiktischem *so*. *So* verweist hier auf einen folgenden Komplementsatz, der durch *dass* eingeführt werden kann, aber nicht muss. Es ist immer akzentuiert. Das syntaktische Format des Vordersatzes ist stark formelhaft, d.h. es kommt nur als *es KOPULA so* vor. Zunächst zwei Beispiele ohne *dass*-Anschluss:

(Telefonat)

01 A: <<nach hinten zu Theo> theo geht der

thomas zum WEINFest?>
 02 (2.5)
 03 B: weiß ers NEDde?=
 04 A: =der theo meint er glaubt es NICHT
 bei dem wetter.
 05 B: bei dem WETter
 06 weil bei UNS isches **SO**:
 07 → (-) wir ham doch n SCHIFFSnachbar.(-)
 08 → und der hat uns jetzt beSTIMMT schon
 das ZEHNTemal zum ESsen eingeladen;
 09 → [und IMmer hatten wir was ANDres vor;
 10 A: [mHM
 11 → B: jetzt ham mir gsa(gd) mir gehn heut
 Abend mit DEM ä: (-) nach BEburg.

(Telefonberatung in einer Radiosendung, A = Anruferin, B = Therapeutin;
 die Anruferin hat sich darüber beklagt, von ihrer Umwelt nicht ausreichend
 wahrgenommen und gewürdigt zu werden)

01 B unserem geSPRÄCH jetzt hier=
 02 des dauert vielleicht jetzt so
 03 fünf miNUTen ungefähr
 04 [hh ham sie] doch AUCh das gefühl
 05 A [ja wir (wa)]
 06 B h dass ich sie nicht Ausreichend zu
 KENNTnis genommen habe; (-)
 07 A ich ähm::
 08 (1.5)
 09 wis[sen sie]
 10 B [dass:] zwischen uns das GLEIChe
 passiert ist;=
 11 A =ja:, ja:, schon: auch weil ich
 ähm des äh'

26

Syntax

12 na des kann ich ihnen jetzt nicht
sagen was ich jetzt denke=
13 des möcht ich auch nicht unbedingt
14 aussprechen- [.hh
15 B [mhm,=
16 =aber es könnte genAu das GLEIChe sein
was ihnen sonst h mit der Umwelt AUCH
passiert;
17 A ja: wissen sie es ist einfach **SO:**,
18 → ich hab diese GRUNDSätzliche erFAHrung
→ hier gemacht (-)
19 ä:hm: tut mir leid wenn ich jetzt
dieses wort HIER sagen muss=
20 =also damit meine ich diese
geSELLschaft hier
21 → .hhh dass äh: die pluralität MEHR gilt
→ als: der EINzelne mensch; ((etc.))

Hier einige Beispiele für dieselbe Konstruktion mit *dass*-Anschluss:

(Rollenspiel Bewerbungsgespräch; der Bewerber ist nach seinen Erfahrungen im Zertifizierungswesen gefragt worden)

01 IF: Isosystem (-) [isch ihnen ein beGRIFF.
02 B: [des Isosystem is mir auf
jeden fall ein beGRIFF,
03 un ich hAb ja auch in meim ABSchreiben
geschrieben,
04 .h dass es für mich n NAheliegenes
ziel is,=
05 =diesen de ge quu quu EM schein
zu machen, [hh ehm
06 IF: [ja,

07 B: es is **SO**
 08 → dass ich über meinen VATER der
 jahrelang (-) bei der i: be EM,
 09 → als qualitätsSICherungsbeauftragter
 geARbeitet hat^h,
 10 → jetzt mittlerweile (-) .h selber
 AUditor is (-) von der dee quu ES,
 11 → =also er (-) .h zErtifiziert als
 EXterner (-) ehm mittelständische
 u' unterNEHmen,
 12 → hh hAb ich also (.) über mein vATER
 (-) EINige informationen über dieses
 (.) äh über diese Iso neunTAUsend eben
 AUCh erhalten, ((etc.))

(psychotherapeutische Gruppensitzung)

01 M : un dann DENK ich mir
 02 wie' wie SOLL ich denn JEmals LEben,
 03 wenn sich DES bei mir nich ändert;
 04 .h wie SOLLn des SEIN; (0.5)
 05 dann kann ich ja überhaupt nich mehr
 FRÖHlich sein;=
 06 =wenn=ch wenn ich SOWas nich mal mehr
 schAFF, .h
 07 wenn ich nich mal normAl LEben kann;
 08 ((schnieft))
 09 (8.5)
 10 TM: t .h dann is=es **SO**
 11 → dass diese SCHEISSWoche, (1.0)
 12 M: ((schneuz[t]))
 13 TM: → [AU zu ihnen gehört;=mh? (2.0)
 14 → so wie diese ANdere.(1.5)
 15 → (worin sie:) (0.5)

16 → und des wär ja n erschter SchRITT,
 17 → sich sogn können des ghört
 au[ch zu mir. (1.5)
 18 M: [((schluchzt))
 19 aber warum?
 20 ((schluchzt))

Es scheint zunächst, als handele es sich hier nicht um eine eigene Konstruktion, sondern lediglich um einen lexikalisch restringierten Sonderfall des katadeiktischen ‘Grundschemas’ (Typ c). Dagegen spricht aber schon, dass die ‘gefrorene’ Konstruktion (f) wesentlich häufiger vorkommt als die Konstruktion, aus der sie in dieser Sicht ‘abgeleitet’ würde. Dies legt nahe, dass die häufigere Konstruktion als Ganze gespeichert und prozessiert wird. Darüber hinaus gibt es auch semantische Gründe, Typ (f) nicht unter Typ (c) zu subsumieren. Im Gegensatz zu (etwa)

das kaRAtE wird ja (-) .h NICHT **so** betrieben dass man
 (-)seinen partner TRIFFT,

handelt es sich (etwa) in

dann is=es **so** dass diese SCHEISSwoche AU zu ihnen gehört

aufgrund der semantischen Leere des grammatischen Subjekts (expletives *es*) und der Kopula (*sein*) nicht wirklich um eine korrelative Konstruktion: das Folgesyntagma erläutert nicht die Art und Weise, in der das Prädikat des ersten Syntagmas zutrifft. Thema und Rhema werden nicht auf Vordersatz und Nachsatz verteilt, sondern die gesamte Proposition ist im Nachsatz enthalten. Das lässt sich auch daran zeigen, dass nur in der ersten Konstruktion *so* durch ein explizit und eindeutig modales Adverbiale wie *auf diese Weise*, *dergestalt* oder ähnliche ersetzt werden kann.

Während also die Semantik der Konstruktion *es ist so*, (*dass*) völlig entleert ist, hat sie doch andererseits eine einheitliche interaktive Funktion, die der Konstruktion (c) abgeht. Diese interaktive Funktion betrifft einerseits die Strukturierung des Redebeitrags: oft führt die Floskel komplexe Argumente

ein (*big packages*), die nicht in einer Turnkonstruktionseinheit verpackt werden können und für die sich der Sprecher oder die Sprecherin projektiv das Rederecht sichern. Entsprechend ist die Turnkonstruktionseinheit, die dem *es ist so* unmittelbar folgt, nicht kohärent mit der dieser vorausgehenden Turnkonstruktionseinheit. Die Einleitung mit der Floskel signalisiert der Rezipientin, dass erst im weiteren Verlauf des Redebeitrags eine solche Kohärenz zu erwarten ist. Andererseits hat die Konstruktion (f) oft auch die Funktion, gesichtsbedrohende Äußerungen zu markieren. Im Telefonat über die Schiffsnachbarn besteht die Gesichtsbedrohung z.B. darin, dass die Anruferin ein schon arrangiertes Treffen mit 'Thomas' beim Weinfest für den heutigen Abend zugunsten eines Restaurantbesuchs mit ihrem Schiffsnachbarn absagen möchte. Das Radiotherapiegespräch hat vor dem durch die Floskel eingeleiteten Redebeitrag der Anruferin seinen kritischen Punkt erreicht, an dem die Therapeutin die Anruferin zu einer situierten Stellungnahme zwingt, indem sie deren allgemeine Klage über die Verständnislosigkeit der Welt auf die Therapiesituation selbst bezieht. Die Anruferin sucht dieser Stellungnahme auszuweichen, weil die offenbar zutreffende Antwort (nämlich dass sie auch das gerade laufende Gespräch als Evidenz für ihre Grundthese sieht) der Therapeutin gegenüber massiv gesichtsbedrohend wäre. Im Ausschnitt aus dem Bewerbungsgespräch führt der Bewerber die einschlägigen Kompetenzen seines Vaters im Zertifizierungswesen als Argument ein - ein äußerst riskantes Manöver. Schließlich ist in der Bulimie-Therapiesituation der Therapeut in der kritischen Situation, der Klientin erklären zu müssen, dass ihre psychischen Probleme Teil ihres Lebens bleiben werden.

Es gibt Übergangsbeispiele, die zeigen, dass die formale Verfestigung der Konstruktion (Kopula und expletives Pronomen) allein nicht ausreicht, um die Konstruktion (f) von (c) zu unterscheiden, sondern dass dazu auch die genannten interaktiven und semantischen Gesichtspunkte berücksichtigt werden müssen. So handelt es sich im folgenden Ausschnitt zwar formal um eine Variante der gefrorenen Konstruktion *es ist so dass* (hier im Konjunktiv II), die überdies der Anforderung an Konstruktion (f) entspricht, die gesamte Proposition in den Nachsatz zu verlagern. Jedoch ist sie weder semantisch entleert (was man an der Austauschbarkeit mit einer modalen Vollform wie *dergestalt* erkennen kann), noch erfüllt sie eine ihrer typischen Funktionen

(d.h., es wird kein komplexes Argument und auch keine gesichtsbedrohende Handlung eingeleitet):

(Telefonat. Thema: Wetterbedingungen für den Flug mit einem Sportflugzeug)

01 A: ä: (-) du HÖR mal;
 02 ich hab mal angerufen (n) ALtenbach,=
 03 B: =ja,
 04 A: ä::hsch also s wäre **SO**
 05 dass man rüberfliegen KÖNNte,=

Hier wird also trotz formaler Ähnlichkeit mit (f) aus der Semantik und interaktiven Funktion deutlich, dass es sich um eine Instantiierung der Konstruktion (c) handelt.

Zusammenfassend lassen sich mindestens die folgenden Konstruktionen (*constructions*) identifizieren, in denen *so* ein nachfolgendes satzwertiges Syntagma projiziert:

(a) Konsekutivkonstruktion:

... X ist *SO* Y → (*dass*) S

(b) Konzessivkonstruktion:

so ADJ *es/das/der auch* V → S

(c) katadeiktische korrelative *dass*-Konstruktion (Topikalisierung):

... V *SO* → *dass* ... S

Sonderfall adnominale Verwendung:

X_i ist *son* Y → $s[\text{Pron}_i \dots]_s$

(d) katadeiktische korrelative *wie*-Konstruktion:

(X V) SO (Adj/Advb) → wie S

(e) Quotativ-Konstruktion:

(und) ich/er so → S

(f) es ist so-Konstruktion

es ist SO → (dass) S

Wie schon angedeutet, unterscheiden sich die einzelnen Konstruktionen sehr stark in ihrer Häufigkeit. Bei weitem am häufigsten sind die Konstruktionen (e) und (f), Konsekutivkonstruktionen (a) kommen weniger häufig vor, die übrigen Konstruktionen selten. Eine Beschreibung im Rahmen der *construction grammar* impliziert, dass die sechs Konstruktionen einzeln gespeichert und prozessiert werden, auch wenn zwischen ihnen beschreibbare Ähnlichkeiten bestehen. Die Verfügbarkeit einer Vielzahl solcher Konstruktionen ermöglicht es den Gesprächsteilnehmern, mit großer Geschwindigkeit und trotzdem großer Präzision Projektionen aufzubauen und Gestaltschlüsse vorherzusagen.

Die hohe Relevanz, die verfestigte Syntaxkonstellationen für die mündliche Interaktion haben, steht nicht im Widerspruch zu der für diese postulierten *on-line*-Prozessierung und dialogischer Ko-Konstruktion. Im Gegenteil: gerade weil wir ein immenses Arsenal solcher Syntaxkonstellationen (mit der dazu gehörigen semantischen, pragmatischen und prosodischen Information) gespeichert haben, können wir unter den Bedingungen der inkrementellen Syntaxverarbeitung und des turn-taking in der direkten (face-to-face) Interaktionssituation effektiv kommunizieren.

4. Abschließende Bemerkungen

In diesem Beitrag wurden drei Merkmale einer Syntaxtheorie dargestellt, die die *online*-Syntax der Gesprochenen Sprache braucht, um deren spezifischen Produktions- und Rezeptionsbedingungen gerecht zu werden.

Ausgangspunkt war die spezifische Zeitstruktur der gesprochenen Sprache. Sie erfordert die permanente Bearbeitung von Projektionen syntaktischer Art. Eine Syntax der Gesprochenen Sprache muss das Projektionspotential syntaktischer Strukturen in die Beschreibung integrieren und die verschiedenen Projektionstypen bewerten.

Zum zweiten muss eine Theorie der Gesprochenen Syntax die interaktive Natur emergenter syntaktischer Strukturen berücksichtigen. Die Äußerungen (*turn constructional units*) eines Sprechers sind oft nicht autark, sondern von den Äußerungen der vorausgehenden Sprecherin syntaktisch abhängig. Umgekehrt stellen sie selbst Strukturlatenzen zur Verfügung, die nächste Sprecher für die Syntax ihrer Äußerungen nutzen können. Emergente syntaktische Strukturen werden oft kollaborativ hergestellt. Rezipientinnen beeinflussen überdies durch ihre (fehlenden oder produzierten) Rezipientensignale die syntaktische Form der emergenten Äußerung.

Zum dritten wurde für die Integration des Konstruktionsbegriffs der *construction grammar* in die Syntax der Gesprochenen Sprache argumentiert, weil diese das Funktionieren von Projektionen in der *face-to-face*-Interaktion besser erklären kann als die üblichen, generalisierteren syntaktischen Beschreibungen.

Ich habe in der Diskussion mehrfach auf die kognitiven *und* interaktiven Aspekte der Prozessierung von *on-line*-Syntax hingewiesen. In der Tat erscheint es mir kaum denkbar, dass eine adäquate modalitätsspezifische Grammatik des Mündlichen auf einen der beiden Aspekte verzichten könnte. Es gibt wohl auch keinen Grund, warum sie sich widersprechen müssten. Im Gegenteil: bis zum Beweis des Gegenteils erscheint es plausibel, dass unser Denken so funktioniert, dass es mit der fundamentalen Tatsache, dass Sprache ihren Ort in der Interaktion hat, optimal zurecht kommt; wie es auch umgekehrt plausibel ist, dass Interaktionsabläufe so strukturiert werden, dass die Interaktionsteilnehmer sie mit ihrer mitgebrachten kognitiven Ausstattung bewältigen können.

Zitierte Literatur

- Antinucci, Francesco, Duranti, Alessandro & Gebert, Lucana (1979): Relative clause structure, relative clause perception, and the change from SOV to SVO. *Cognition* 7, 145-176.
- Auer, Peter (1990): „Einige umgangssprachliche Phänomene der türkischen Syntax und Möglichkeiten ihrer Erklärung aus 'natürlichen' Prinzipien". In: N. Boretzky, W. Enninger, Th. Stolz (Hrsg.), *Spielarten der Natürlichkeit - Spielarten der Ökonomie. Beiträge zum 5. Essener Kolloquium*, 2. Bd., 2. Halbbd., Bochum: Brockmeyer, 1990, 271-298.
- Auer, Peter (1996): „On the prosody and syntax of turn-continuations“. In: E. Couper-Kuhlen & M. Selting (Hrsg.), *Prosody in Conversation*, Cambridge: CUP, S. 57-100.
- Auer, Peter (2000): „On line-syntax – oder: was es bedeuten könnte, die Zeitlichkeit der mündlichen Sprache ernst zu nehmen“ *Sprache und Literatur* 85 (2000), 43-56. (Themenheft: Die Medialität der Gesprochenen Sprache)
- Auer, Peter (im Druck): „Projection in interaction and projection in grammar“, erscheint in *Text*.
- Burkhardt, Armin (1987): „SOSO? Kritik und weiterführende Überlegungen zu Konrad Ehlich's Aufsatz über die Funktionen des deutschen so.“ In: I. Rosengren (Hrsg.), *Sprache und Pragmatik*, Lund, S. 299-313.
- Bybee, Joan (1998): The emergent lexicon. *CLS* 34, 421-435.
- Croft, William (2001): *Radical construction grammar: syntactic theory in typological perspective*. Oxford: Oxford Univ. Press.
- Couper-Kuhlen, Elizabeth and Tsuyoshi Ono (in print), 'Incrementing' in conversation: A comparison of practices in English, German and Japanese. To appear in *Pragmatics*, Special issue on 'Turn continuation in cross-linguistic perspective'
- Deppermann, Arnulf (2004), *Grammatik und Semantik aus gesprächsanalytischer Sicht. Methodologischer Rahmen und exemplarische Untersuchungen*. Unveröffentl. Habilitationsschrift, Univ. Freiburg/Br.
- Ehlich, Konrad (1987): *so - Überlegungen zum Verhältnis sprachlicher Formen und sprachlichen Handelns, allgemein und an einem widerspenstigen Beispiel*. In: I.

- Rosengren (Hg.): Sprache und Pragmatik. Lund (= Lunder germanistische Forschungen 55). S. 279-298.
- Ferrara, Kathleen (1992): The interactive achievement of a sentence: joint productions in therapeutic discourse. *Discourse Processes* 15, 207-228.
- Ferreira, F. & Henderson, J.M. (1991): Recovery from misanalyses of garden-path sentences. *Journal of Memory and Language* 25, 725-745.
- Fillmore, Ch., Kay P., O'Connor Mary Catherine (1988): Regularity and idiomaticity in grammatical constructions: the case of *let alone*. *Language* 64, 3, 501-538.
- Ford, Cecilia F./Fox, Barbara A./Thompson, Sandra A. (2003): Social interaction and grammar. In: Tomasello, Michael (Hg.), *The New Psychology of Language*, Bd. 2. Mahwah N.J., S. 119-143.
- Falk, Jane, 1979. *The Duet as a Conversational Process*. Ph. Diss, Princeton University. Frankfurt, Dorothea
- Frazier, L./Rayner, K. 1982: Making and correcting errors during sentence comprehension: Eye movements in the analysis of structurally ambiguous sentences. *Cognitive Psychology* 14, 178-210.
- Galato, Andrea (2000), An innovative German quotative for reporting on embodied actions: und ich so/und er so 'and I'm like/ and he's like'. *Journal of Pragmatics* 32, 29-54.
- Goldberg, Adele (1995), *Constructions: a construction grammar approach to argument structure*. Chicago: UP
- Hayashi, Makoto; Mori, Junko, Takagi, Tomoyo (2002): Contingent achievement of co-tellership in a Japanese conversation: On analysis of talk, gaze, and gesture. in C. Ford/B. Fox/S. Thompson eds. *The Language of Turn and Sequence*, Oxford: OUP 2002,81-122
- Kindt, Walther (2003): Ellipsen und andere syntagmatische Aspekte. In: Rickheit, Gert/Herrmann, Theo/Deutsch, Werner (Hg.): *Psycholinguistik/Psycholinguistics - Ein internationales Handbuch/An International Handbook*. Berlin. S. 306-316.
- Lerner, Gene H. (1991): On the syntax of sentences-in-progress. *Language in Society* 20 (3), S. 441-458.
- Maracz, Laci K & Muysken, Pieter (Hg.), (1989): *Configurationality: the Typology of Asymmetries*. Dordrecht: Foris.
- Marslen-Wilson, W.D. (1985): Speech shadowing and speech comprehension. In: *Speech Communication* 4, S. 55-73.

- Michaelis, Laura A. & Lambrecht, Knud (1996), Toward a construction-based theory of language function: the case of nominal extraposition. *Language* 72, 2, 215-247.
- Nichols, Johanna (1992): *Linguistic Diversity in Space and Time*. Chicago.
- Szczepek, Beatrice (2000): Formal Aspects of Collaborative Productions in English Conversation = InLiSt - Interaction and Linguistic Structures, No. 17, URL: <http://inlist.uni-konstanz.de/issues/17/index.htm>.
- Sandig, Barbara (1987), Kontextualisierungshinweise: Verwendungen von *so* im Prozeß sprachlichen Handelns. In: I. Rosengren (Hrsg.), *Sprache und Pragmatik*, Lund.
- Stein, Stephan (2003), *Textgliederung. Projektierende Kraft syntaktischer Konstruktionsschemata*. Berlin: de Gruyter.
- Thurmair, Maria (2002): *Vergleiche und Vergleichen*. Tübingen.
- Tomasello, Michael (1998), The return of constructions. *J. Child Lg.* 25, 3, 431-443.
- Weinrich, Harald (2003), *Textgrammatik der deutschen Sprache*. Hildesheim: Olms (2. Aufl.)