

Zum Segmentierungsproblem in der Gesprochenen Sprache*

Peter Auer
(FRIAS, Universität Freiburg/Br.)

Abstract	1
1. Die Frage der Grundeinheiten der Gesprochenen Sprache und ihrer Segmentierbarkeit	1
2. Theorie des Segmentierens.....	2
3. Probleme beim Segmentieren der Gesprochenen Sprache.....	3
3.1. Segmentieren in Handlungen	5
3.2. Segmentierung in syntaktische Einheiten	6
3.3. Segmentierung in Intonationsphrasen	8
3.4. ‚Kognitive‘ Grundeinheiten	10
4. Das Verhältnis von Redebeiträgen zu Sätzen	10
5. Eine Alternative zum Segmentieren.....	11
5.1. Syntaktische Zäsuren.....	12
5.2. Prosodische Zäsuren.....	14
5.3. Zäsurierung nach Handlungsgrenzen	15
5.4. Eine revidierte Analyse	15
6. Abschließende Bemerkungen.....	17
7. Literaturverzeichnis.....	17
Anhang	19

Abstract

In diesem Beitrag wird argumentiert, dass der Begriff des Segmentierens für die Analyse der Gesprochenen Sprache grundsätzlich ungeeignet ist, weil er ihre on-line-Emergenz missachtet. Versuche, Gesprochene Sprache zu segmentieren, sind dem strukturalistischen Erbe in der Linguistik geschuldet; sie ignorieren die jüngeren Entwicklungen der Erforschung von mündlicher Sprache und reduzieren diese zu einem zweidimensionalen, papierenen Gebilde. Dies gilt für alle Strukturebenen der Sprache, Syntax genauso wie Prosodie und Pragmatik. Es wird vorgeschlagen, den Begriff des Segmentierens durch den des Zäsurierens zu ersetzen, das anstelle von Einheiten auf Abschlusspunkte hin orientiert ist.

1. Die Frage der Grundeinheiten der Gesprochenen Sprache und ihrer Segmentierbarkeit

Seit den 60er Jahren lässt sich durch die verschiedenen Publikationen zur Gesprochenen Sprache hindurch ein Thema verfolgen, das für viele Autoren von zentraler Bedeutung zu sein scheint. Es ist die Frage nach den Grundeinheiten, in die sich die zu untersuchenden Materialien ‚segmentieren‘ lassen. Schon früh wurde festgestellt, dass der Satz begriff – welchen Status er auch immer in der Geschriebenen Sprache haben mag – für die Gesprochene Sprache kaum hinreichend zu operationalisieren ist. Fast allen Autoren erscheint es daher wichtig, die Grundeinheit der Gesprochenen Sprache neu zu bestimmen. Bezeichnend ist etwa die folgen-

* Für viele hilfreiche Kommentare zu einer früheren Version danke ich Karin Birkner und Harald Baßler.

de Feststellung aus einer frühen Arbeit: „Jeder linguistischen Untersuchung, der ein Textkorpus zugrunde gelegt wird, stellt sich das Problem, dieses Corpus zu segmentieren, indem durch ein operationalisiertes Verfahren die einzelnen Texte restfrei in die kleinsten relevanten Einheiten zerlegt werden“ (Caroli 1977:142). Die Palette der Grundeinheiten, die zu diesem Zweck vorgeschlagen worden sind, reicht von rein syntaktischen, wie dem sog. Teilsatz bei Leska (1965/1975), über prosodische, wie die Intonationsphrase (z.B. bei Wunderli 1979), bis zu eher pragmatischen, wie Sprechakten (so Caroli 1977:106) oder Redezügen. Dazu kommen Mischkategorien, die semantische, syntaktische, pragmatische und prosodische Kriterien kombinieren, wie z.B. Raths Äußerungseinheiten (1990), sowie kognitive Einheiten wie Chafes *idea units* (1980), die den Sinnschritten des DUDEN (z.B. 1973) nicht unähnlich sind.

In diesem Beitrag werde ich zeigen, dass die Suche nach solchen Grundeinheiten dann ein für die Forschung unergiebiges Unterfangen ist, wenn diese als Segmentierungseinheiten verstanden werden. Dazu ist es notwendig, den Prozess des Segmentierens selbst ins Blickfeld zu rücken, der in der Forschung oft völlig untheoretisch gebraucht wird. Innerhalb der Linguistik ist er jedoch Teil eines bestimmten Paradigmas, nämlich des taxonomischen Strukturalismus. Die strukturalistische Operation des Segmentierens ist, so soll gezeigt werden, keine angemessene Methodik für die Untersuchung mündlicher Sprache. Stattdessen soll die Idee des Gestaltschlusses eingeführt werden, die ihren on-line-Charakter in den Vordergrund rückt. Die anstelle des Segmentierens tretende Operation ist dann die des *Zäsurierens*.

2. Theorie des Segmentierens

Die theoretisch am besten reflektierten Aussagen zum Thema Segmentieren finden sich in der Linguistik in den Arbeiten der taxonomischen Schule, besonders bei ihrem einflussreichsten Vertreter Zellig Harris (1951, ⁸1969:25ff). Sie beziehen sich – entsprechend der doppelten Artikulation der Sprache – auf die Phonologie und Morphologie. Auf beiden Strukturebenen wird durch das Segmentieren eine Konstellation von Elementen der niedrigeren Ebene (des Lautstroms in der Phonologie, von Phonemverkettungen in der Morphologie) in Einheiten (Phoneme bzw. Morpheme) unterteilt, auf die die strukturalistischen Verfahren der Wiederholung, der Permutation und der Substitution anwendbar sind.

An die Segmentierung sind nach taxonomischer Auffassung mindestens die folgenden vier methodischen Anforderungen zu stellen: (a) Exhaustivität: Das Material muss vollständig segmentiert werden, d.h. die Segmentierung muss zu einer restlosen Bearbeitung führen. Es dürfen keine nicht-segmentierten Elemente übrig bleiben. (b) Atomismus: Die gefundenen Segmente dürfen nicht Segmente derselben Art einschließen. (c) Diskretheit: Ein Element darf nicht zugleich Bestandteil mehrerer Segmente sein, d.h. die Segmentgrenzen müssen eindeutig definiert werden. (d) Ebenenkonstanz: Das Material muss in Elemente der gleichen linguistischen Beschreibungsebene segmentiert werden. Es ist zum Beispiel in einer phonemischen Segmentierung nicht möglich, einige Bestandteile nur phonetisch zu segmentieren. Exhaustivität, Atomismus, Diskretheit und Ebenenkonstanz sind Grundanforderungen an die Segmentierung. Es sei dahingestellt, ob sie für die meist grammatikorientierten Ziele der Strukturalisten ausreichend und sinnvoll sind.¹ Sie sind jedenfalls Ausfluss der strukturalistischen Praxis, die in der Regel corpusbezogen war. Wichtiger für die Erforschung der gesprochenen Sprache ist die Frage, ob die dafür vorgeschlagenen Grundeinheiten und Segmentierungsverfahren den Kriterien der Exhaustivität, Atomismus, Diskretheit und Ebenenkonstanz standhalten; denn nur dann ist es gerechtfertigt, den Begriff des Segmentierens zu verwenden.

¹ Auch strengen Strukturalisten ist natürlich nicht entgangen, dass das Kriterium der Diskretheit von diskontinuierlichen Morphemen (etwa Zirkumpositionen) nicht erfüllt wird.

3. Probleme beim Segmentieren der Gesprochenen Sprache

Im Folgenden will ich zeigen, dass sich Exhaustivität und Atomizität in der Gesprochenen Sprache zumindest dann nicht erreichen lassen, wenn zugleich das Postulat der Ebenenkonstanz erfüllt werden soll, und dass auch das Postulat der Diskretheit zu erheblichen Problemen führt. Vorausgesetzt ist, dass das zu segmentierende Material (Transkript) die wesentlichen Strukturmerkmale der Gesprochenen Sprache nicht von vornherein ausfiltert. Dazu gehören Abbrüche und Wiederholungen, Rezipientensignale und Interjektionen, Anakoluthe und Reparaturen. Lässt man beliebig Phänomene fort (z.B. die genannten, vermeintlich verunglückten oder ‚überflüssigen‘ Äußerungselemente) oder formuliert sie in ‚vollständige Sätze‘ um, sind die genannten Postulate auf triviale Weise durch Eliminierung aller problematischen Fälle erfüllt.² Dieses Filtern oder Ausbessern kommt letztendlich einer schriftsprachlich voreingenommenen Brille gleich, durch die Gesprochene Sprache betrachtet wird: All das, was sie am meisten von der Geschriebenen unterscheidet, wird ausgeblendet. Eine solche Methodologie kann die Gesprochene Sprache nur als defizientes Derivat der Geschriebenen erfassen und ist heute zu recht nicht mehr diskutabel.

Vorausgesetzt ist des Weiteren, dass überhaupt eine operationalisierbare Definition der jeweils vorgeschlagenen Grundeinheit vorliegt. Fehlt eine solche, lassen sich die Problemfälle ad hoc durch fallweise eingeführte Segmentierungsvorschriften aus der Welt schaffen. Eine solche operationalisierbare Definition fehlt zum Beispiel für Raths ‚Äußerungseinheiten‘ (1990). Rath versteht sie als mögliche Untergliederung eines Redebeitrags; die Gliederung kann inhaltlich, prosodisch, durch Pausen oder durch Gliederungssignale erreicht werden. Lediglich die Syntax scheint als Kriterium ausgeschlossen zu werden. Eine positive Definition der Äußerungseinheiten wird aber nicht gegeben, so dass sich nicht entscheiden lässt, ob die gefundenen Einheiten dem Postulat der Ebenenkonstanz standhalten. Zur Segmentierung eignen sich die ‚Äußerungseinheiten‘ daher nicht.

Um zu zeigen, dass ernst zu nehmende Transkripte Gesprochener Sprache nicht exhaustiv, diskret und atomistisch in definierte Grundeinheiten segmentiert werden können, die auf derselben linguistischen Analyseebene angesiedelt sind, werden in diesem Abschnitt die spezifischen – allerdings auch ähnlichen – Probleme diskutiert, die sich bei der Segmentierung in Sprechakte oder sprachliche Handlungen, in Sätze und in Intonationsphrasen ergeben. Ich greife dabei auf das Transkript des Beginns einer Schlichtungsverhandlung (Schröder 1997) zurück, das konversationsanalytischen Transkriptionskonventionen (GAT; vgl. Selting et al. 1999) angepasst und auf der Grundlage der Tonaufnahme³ prosodisch verfeinert wurde.

(1) Beginn einer Schlichtungsverhandlung

01 C: leopold gegen may? (-) und abel gegen may.
 02 ((Die Beteiligten nehmen ihre Plätze ein.))
 03 C: ja. (-) herr may was sAgen se zu der KLA:ge?=
 04 die KLÄgerin sagt sie: hätt noch GELD zu kriegen;
 (2.5)
 05 B: dazu is zu sAgen, (1.5) dass äh:: (2.0) der KLÄ:ger,
 06 C: .ts <<undeutlich> (äh wir reden nur über) die FRAU
 07 leopold jetzt ne??
 08 B: ((räuspert sich)) <p>BITte??>

²Frühe Arbeiten zur Gesprochenen Sprache – etwa die von Leska (1965/1975) – sind so vorgegangen und haben ihrer Analyse von vorne herein ein bereinigtes Transkript zugrunde gelegt, in dem unbequeme Äußerungsteile auf der Basis eines vorwissenschaftlichen, intuitiven Wissens als irrelevant ausgeblendet wurden; nur so konnte Leska zu einer konsistenten quantitativen Auswertung ihres Materials kommen.

³Für die Überlassung der Tonaufnahme zu diesem Transkript sei Reinhard Fiehler vom Institut für Deutsche Sprache in Mannheim herzlich gedankt. Anwesend sind der Schlichter (C), der Beklagte (B) sowie die beiden Kläger, Leopold und May.

09 C: <<pp>wir REden (-) über die FRAU leopold jetzt.=>
10 B: =NEIN, <<pp>(red ich NICHT).>(1.0)
11 für den jungen MANN da.
12 C: NEIN, wir reden jetzt von der FRAU leopold.
13 B: ach SO:;
(0.5)
14 C: SIE sagt sie sei vom siebten ERSTen bis zwölften ZWOten
15 als AUSHilfe gegen einen STUNdenlohn von zehn mark
16 beSCHÄftigt gewesen=sie rechnet sich (vier'
17 zweihundertvierunddreißig) stunden AUS,=für
18 zweitausenddreihundertvierzig mark NETto, (-)
19 .hh hat s=es geKRIGGT?
20 wenn ja' wenn NEI:N, waru[m NICHT.
21 B: [
22 (1.0) <<mf>das stimmt NISCHT.>
23 [es war]
24 C: [WAS (.)] stimmt nicht?
25 B: es war (1.0) bei (1.0) der (2.0) frau leopold oder so
26 (-) bei FROLlein leopold (--> <<staccato>KEInerlei (-)
27 beZAHlung (-) verEINbart
28 C: sie hat UNendgeltlich gearbeitet.=
29 B: =JAwoll;
(0.5)
30 C: um GOTTeslohn.
31 B: es ging (-) <<kurz>ja. >
32 C: [<<ppp> aha, >
33 B: [(-) es ging von VORNherein (-) SO an? (- -) wenn de'
34 wann de (--> hier (1.0) äh:: (-) der junge MANN, (-)
35 zunächst mal (.) meine FRAU angesprochen hat, (-) .h ob
36 er' (-) bei uns nicht äh sich (-) über (-) die (-) proDUKte und
37 so weiter, .h inf[ormiere KÖNNte. (-) er MÖChte gErn
38 C: [<ppp>mhm.
39 B: (.) da bisschen (-) öh: (2.0) hier HERkommen, NET wahr, (-)
40 und äh sein WEIL sein WISsen, (-) hier etwas AUSzubauen,
41 (--> <<staccato>von (-) einer (-) beZAHlung;> (--> von eim
42 (.) von eim (.) STUNdenlohn oder so;= war NIE die rede
43 gewesen.
(1.0)
44 C: äh: (-) versteh ich sie rIchtig.(-) IST davon gar nicht
45 geSPROChen worden? =O:der hat man verEINbart dass nichts
46 beZAHLT wird. (-)
47 B: es: der junge MANN hat von vorn heREIN;(-) schon gar
48 nicht von einer (-) beZAHlung (-) geSPROChen;
49 dass er sie verLANGT hätte; (-)
50 wäre des der FALL gewesen dann hätten wir=s ABgele:hnt; (-)
51 C: ja herr may also ei' EINS versteh ich jetzt NICHT so
52 richtig. äh <<stottert>ge' ge' gewöhnlicherWEise,> .hh
53 äh (-) LEben die menschen,= von dem was sie
54 verDIEnen; =nur WEnige können von ihm verMÖgen leben.
55 (-) deswegen ist es doch ein wär=s doch UNgewöhnlich dass
56 er sagt ich komm jetzt mal zu IHnen und äh: (-) will kein
57 GELD haben (0.5)
58 [wie soll denn des GEHN;=
59 B: [genAu so WAR=S.
60 C: =sogar AUSzubildende kriegen doch geld;
61 und die (.) können ja AU nix.
62 [ne?
63 B: [SO war=s gewEsen=(über); (-) über beZAHlung war NICHT die
64 REde.
65 (1.0)
66 ABSolut nIcht.
67 (--> <<all> wir hAm ja auch von VORNerein g(e)SACHT,> .h
68 dass die situatiO:N, (-) äh:: (-) zunächst also jetzt mal SO is,

69 .hh im HINblick auf die: (.) Umsätze;=die also h (-) um: (1.0)
 70 (-) wie sie jetzt (.) im VIERTel letzten jahr der FALL
 71 war; .h FÜNfundzwanzig prozent NACHgelassen haben; (-)
 72 .h und wir sind NICHT in der LA:ge, (-) personAL, (-) äh z'
 73 daZUnehmen? (.) [nach dem was wir HAben, (-)
 74 C: [.mhm;
 75 nIcht WAHR?
 76 (-) und äh damit war der fall erLEdigt.
 77 =es war KEInerlei (-) FORderung gestellt worden;
 78 (-) wir wie geSACHT,= wir hätten das ABgelehnt;

3.1. Segmentieren in Handlungen

Beginnen wir mit der Grundeinheit ‚sprachliche Handlung‘ (‚Sprechakt‘). Was eine solche sprachliche Handlung ausmacht, ist schwer zu sagen; in ethnomethodologischer Tradition ließe sich jedoch mit aller Vorsicht definieren, dass Verhalten dann als eine soziale Handlung gelten kann, wenn es *accountable* ist, d.h. wenn die Teilnehmer dieses Verhalten identifizieren (vielleicht sogar benennen), begründen, kritisieren, angreifen, auf es eingehen, es bearbeiten, beantworten oder sonst wie darauf erwidern können. Da allerdings keine Theorie sprachlichen Handelns vorliegt, die hierarchisch strukturierte komplexe Handlungen (wie sie für sprachliches Handeln in Gesprächen typisch sind) erfassen könnte, ist die Segmentierung in Handlungen von vorne herein mit schwerwiegenden Problemen behaftet. Dies gilt sogar für institutionelles Handeln wie im vorliegenden Beispiel, dessen Grobstruktur zumindest seitens des Schlichters bereits in die Interaktion als vorgängiges Handlungswissen hineingetragen wird.

Relativ problemlos können zu Beginn des Transkripts die institutionelle sprachliche Handlung des Aufrufens einer Verhandlung und in der zweiten Zeile die nonverbale Folgehandlung der Teilnehmer (Plätze einnehmen) identifiziert werden. Die Äußerung des Schlichters in Z.3/4 ist als Aufforderung an den Beklagten, zum Vorwurf Stellung zu nehmen, gehört worden und wird von ihm auch so verstanden. Diese Handlung ist allerdings bereits komplex strukturiert: sie besteht aus drei Elementen, deren jeweiliger Handlungsstatus unklar ist, nämlich der Partikel *ja*, die im gegebenen Zusammenhang den Übergang in die eigentliche Verhandlung markiert, der Aufforderung *herr may was sAgen se zu der KLA:ge?* und der diese erläuternden Äußerung *die KLÄgerin sagt sie: hätt noch GELD zu kriegen*. Aber handelt es sich um drei eigenständige Handlungen? Für die zweite Einheit innerhalb des Redebeitrags lässt sich die Frage sicherlich bejahen. Die dritte Einheit wäre dann als subsidiär zur zweiten (und nicht als eigenständige Handlung) einzustufen, wenn sie ohne die Aufforderung zur Stellungnahme nicht als eigenständige Handlung zu verstehen wäre. Ist sie hingegen auch ohne die erste als dieselbe Handlung (Aufforderung zur Stellungnahme) interpretierbar, ist von zwei gleichen Handlungen auszugehen, die sich lediglich dadurch unterscheiden, dass die zweite Version spezifischer ist als die erste. Der Handlungsstatus metapragmatischer Elemente wie des einleitenden *ja* in Z.3 ist noch schwieriger zu beurteilen. Der erste größere Redebeitrag des Schlichters wirft also schon typische Probleme auf: besonders der Status von subsidiären (unselbstständigen) Einheiten und deren Abgrenzung von Handlungsketten ist theoretisch nicht gelöst. Damit wird es schwierig, das Atomismus-Kriterium der Segmentierung zu erfüllen.

Weitere Schwierigkeiten der Handlungsanalyse ergeben sich in Z.5: *dazu is zu sagen, dass der Kläger* ist sicherlich keine vollständige sprachliche Handlung. Kaum hat B nämlich eine referenzielle Nominalphrase geäußert (*der Kläger*), wird er vom Schlichter unterbrochen und korrigiert: hier sei nicht vom Kläger, sondern von der Klägerin die Rede. Es folgt nun eine komplexe Reparatursequenz, die in verschiedene Handlungen segmentiert werden kann, die vielfältig miteinander verschränkt sind und alle der übergeordneten Handlung (Stellungnahme des Beklagten) untergeordnet sind: Fremdreparatur in Z.6/7, Fremdinitiierung einer Reparatur

auf dieser Reparatur in Z.8, Bearbeitung dieser Fremddinitiiierung durch Wiederholung der Fremdreparatur in Z.9, Ablehnung der Fremdreparatur in Z.10/11, Insistieren auf der Fremdreparatur in Z.12 und Annahme der Fremdreparatur in Z.13. Nach Abschluss der Reparatursequenz wird die unvollständige Äußerung, die ihr vorausging, nicht wieder aufgenommen. *Dazu is zu Sagen dass äh:: der KLÄ:ger* kann im gegebenen Kontext höchstens als Versuch einer Erwiderung/Antwort auf den Schlichter gewertet werden; das Äußerungsfragment ist aber keine solche Erwiderung oder Antwort. Dies wird nicht zuletzt daran deutlich, dass an dieser Stelle niemand der Stellungnahme des Beklagten widersprechen oder zustimmen könnte. (Der Schlichter widerspricht nicht dem Äußerungsfragment Z.5 insgesamt, sondern lediglich der Relevanz der darin hergestellten Referenz auf ‚den Kläger‘.) Es gibt also, so lässt sich an diesem ebenso trivialen wie häufigen Fall zeigen, Äußerungen, die weder allein noch im Zusammenhang anderer Äußerungen eine sprachliche Handlung darstellen. Sie durchkreuzen alle Segmentierungsabsichten, weil sie das Exhaustivitätspostulat verletzen; sie bleiben übrig, wenn ein Text vollständig in Handlungssegmente zerlegt werden soll.

In Z.14-20 formuliert der Schlichter seine Aufforderung zur Stellungnahme ein weiteres Mal in noch elaborierterer Form. Der relativ lange Beitrag wirft ein weiteres typisches Segmentierungsproblem im Zusammenhang des Atomismuspostulats auf. Die komplexe Handlung, die der Schlichter in Z.14 beginnt, ist nämlich auch schon nach Z.16 (... *beSCHÄFtigt gewesen*), sicherlich aber nach Z.18 (*zweitausenddreihundertvierzig mark NETto*) potenziell abgeschlossen und bearbeitbar (beantwortbar); die sprachliche Handlung ‚Aufforderung zur Stellungnahme‘ könnte hier also schon als geglückt gelten. Dasselbe gilt nach dem explizierenden *hat s=es geKRIGGT?* und der Erweiterung des Redebeitrags durch eine weitere Frage, die für den Fall einer negativen Antwort auf die erste vorsorglich eine Begründung einfordert (*wenn NEI:N, warum NICHT*).⁴ Solche Erweiterungen schon abgeschlossener, d.h. als Handlung interpretierbarer Äußerungen können zwar wiederum als untergeordnete Handlungen des Elaborierens oder Reformulierens verstanden werden, sie sind für den Prozess des Segmentierens aber ganz offensichtlich problematisch: wenn wir einen frühen pragmatisch möglichen Abschlusspunkt als Segmentgrenze ansetzen, bleiben die späteren Turn-Konstruktionseinheiten als überflüssig übrig (contra Exhaustivitätspostulat), wenn wir die späteste mögliche Segmentgrenze wählen, müssen wir mit der Tatsache umgehen, dass auch schon innerhalb des Segments eine vollständige, identische sprachliche Handlung zu erkennen ist. Damit wird das Atomizitätspostulat verletzt.

Das Problem der unvollständigen Segmente und das der expandierten Segmente sind in gewisser Weise spiegelbildlich: im einen Fall liegt zu wenig segmentierbares sprachliches Material vor, im anderen Fall zu viel. Beide sind für die Segmentierung Gesprochener Sprache unüberwindbare Hindernisse. In der mündlichen Sprache lassen sich Handlungen zwar identifizieren, jedoch nicht problemlos aus dem Transkript ‚herausschneiden‘: es ist schwierig, ihre Ränder gegen die Nachbarhandlungen exakt zu bestimmen.

Andere Schwierigkeiten kommen hinzu. Die eindeutige Zuordnung von Äußerungen zu Sprechakttypen ist ja keineswegs trivial – oft lässt sich eine Äußerung mehrfach typisieren. Im Transkript ist zum Beispiel die Äußerung des Schlichters *wir REden über die FRAU leopold jetzt* (Z.9) von ihm selbst als Korrektur der vorhergehenden Äußerung Z.5 von B gemeint, sie wird aber von B als Feststellung verstanden.

3.2. Segmentierung in syntaktische Einheiten

Eine Satzdefinition für die Gesprochene Sprache ist bekanntlich außerordentlich schwierig zu finden; traditionelle Definitionen wie ‚ein Satz enthält Subjekt und Prädikat‘ oder ‚ein Satz wird durch ein Verb und seine (obligatorischen?) Ergänzungen bestimmt‘ sind für den Vor-

⁴ Zum Äußerungsfragment *wenn ja* siehe unten, Abschnitt 3.3.

gang des Segmentierens wenig hilfreich, weil sie eine große Menge von Äußerungen (sog. Ellipsen, Anakoluthen, Reparaturen etc.) unklassifiziert lassen. Etwas erfolgreicher ist Bloomfields klassische Definition (1933:170):

In any utterance, a linguistic form appears either as a constituent of some larger form, as does *John* in the utterance *John ran away*, or else as an independent form, not included in any larger (complex) linguistic form, as, for instance, *John* in the exclamation *John!* When a linguistic form occurs as part of a larger form, it is said to be in *included position*; otherwise it is said to be in *absolute position* and to constitute a *sentence*.

A form which in one utterance figures as a sentence, may in another utterance appear in included position. In the exclamation form cited, *John* is a sentence, but in the exclamation *Poor John!* the form *John* is in included position. In this latter exclamation, *poor John* is a sentence, but in the utterance *Poor John ran away*, it is in included position. (Bloomfield 1933:170)

Ein Satz ist demzufolge als größte morphologisch gebundene Einheit in der Lage, einen selbstständigen Redebeitrag zu bilden – natürlich nur im geeigneten Kontext.⁵ Der Satz ist also nach innen morphosyntaktisch kohärent, nach außen relativ selbstständig.

Dennoch ergeben sich auch mit dieser Satzdefinition Schwierigkeiten. Abgebrochene Äußerungen sind ja aus syntaktischer Sicht nicht weniger problematisch als bei der Segmentierung nach sprachlichen Handlungen: Z.5 *dazu is zu Sagen, dass äh der KLÄ:ger* ist nicht nur kein erfolgreicher Sprechakt, sondern auch kein erkennbar vollständiger Satz. Nach Bloomfields Definition, wörtlich genommen, gibt es allerdings keine Möglichkeit, solche Fragmente nicht als Satz zu bezeichnen. Wirft man allerdings Sätze und Satzfragmente in einen Topf, gibt es keine Möglichkeit mehr, eine Syntax der Gesprochenen Sprache zu schreiben; diese soll es ja gerade ermöglichen, abgebrochene oder sonst verunglückte Strukturen von wohlgeformten zu unterscheiden.

Ähnliches gilt für Konstruktionswechsel wie in Z.67, wo die komplexe syntaktische Konstruktion *wir hAm ja auch von VORnerein g(e)SACHT dass die situatiO:N zunächst also jetzt mal SO is ...* nach verschiedenen Erweiterungen nicht (wie nach dem kataphorischen *so* erwartbar; vgl. Auer 2007) durch ein *dass*-Komplement weitergeführt wird, sondern in eine Koordinationsstruktur (Z.72: *und wir sind NICHT in der LA:ge...*) übergeht. Hier findet eine Äußerungskomponente, die strukturell in inkludierter Position stehen muss, keine entsprechende Einbettung in einen größeren syntaktischen Zusammenhang. Auch solche Äußerungen bleiben bei der Segmentierung übrig und verletzen das Exhaustivitätspostulat.

Ein weiteres Problem der syntaktischen Einheitenbildung auf der Grundlage des Satzes zeigt sich schon in der ersten Transkriptzeile: *Leopold gegen May* qualifiziert sich im gegebenen Kontext zwar als eigenständige, potenziell turn-konstitutive Äußerung, da es morphologisch nicht an die Folgeäußerung *und Abel gegen May* gebunden ist und eine abgeschlossene morphosyntaktische Struktur hat. Hingegen ist *und Abel gegen May* selbst nicht selbstständig, sondern im Sinne Bloomfields inkludiert: die Koordination *und* fordert zwei koordinierte Elemente, von denen eines fehlt. *Leopold gegen may* ist also ebenso wie die größere Struktur *leopold gegen may und abel gegen may* eine maximale syntaktische Form (Satz). (Dasselbe Problem ergäbe sich, wenn man *und* zur ersten Äußerung zählen würde; noch problematischer wird das Segmentierungsergebnis, wenn die Kopula als eigener Satz behandelt wird.) Kritisch ist hier offenbar die Frage, wann von *included position* im Sinne Bloomfields zu sprechen ist. Die drei Wörter *Leopold / gegen / May* hängen morphosyntaktisch stärker zusammen als die Koordinationsstruktur *Leopold gegen May / und / Abel gegen May*; reicht im letzteren Fall das textbildende *und* aus, um beide Elemente in inkludierte Position zu verweisen?

⁵ Bemerkenswert ist, daß Bloomfield seine Satzdefinition explizit vom Kontext abhängig macht: was in manchen Kontexten (etwa elliptischen) als abgeschlossen gilt, ist in anderen Kontexten inkludiert und daher nicht satzwerdig.

Dasselbe Problem ergibt sich bei der Behandlung von Satzexpansionen (Auer 1991), die morphosyntaktisch nicht an die vorausgehende Einheit gebunden sind; etwa wird in Z.28/30 *sie hat UNendgeltlich gearbeitet* expandiert durch *um GOTTeslohn*. Die expandierende Nominalphrase kann entweder als selbstständiger (,elliptischer‘) Satz gesehen werden oder aber als Bestandteil des vorhergehenden Syntagmas (Reparatur von *unendgeltlich*). Ganze Klassen von Syntagmen führen zu denselben Problemen; erwähnt seien Vorlaufelemente wie *versteh ich sie richtig* (Z.44) oder *ja herr may* (Z.51) und Nachlaufelemente (,tags‘) wie *ne?* (etwa in Z.62). Sie sind in Bloomfields Sinn nicht selbstständig, aber auch nicht ,inkludiert‘, denn sie stehen in keiner (oberflächlich markierten) morphosyntaktischen Beziehung zu dem folgenden bzw. vorausgehenden Syntagma.

Schließlich sind für die Segmentierung in Sätze alle ,Ellipsen‘ problematisch, z.B. in *warum nicht?* (Z.20) oder *red ich nicht* (Z.10, vgl. Auer 2007). Hier besteht ganz offensichtlich ein Abhängigkeitsverhältnis zwischen den ,elliptischen‘ Äußerungen und ihrem Bezugssyntagma (wie die Unmöglichkeit der Permutation zeigt); diese Bezugssyntagmen selbst sind aber in sich abgeschlossen und daher Sätze im Sinne Bloomfields. Wird das Gesamtsyntagma als Satz verstanden (oder sowohl das gesamte als auch das eingebettete selbstständige), ist das Atomismus-Postulat verletzt, wird nur das Bezugssyntagma als Satz gewertet, ist die Segmentierung nicht exhaustiv.

Der Teilsatz (*clause*; vgl. etwa Crystal 1979) löst die genannten Probleme der Segmentierung nicht, führt aber zu weiteren Segmentierungsproblemen, weil die Grenze zwischen Neben- und Hauptsatz oft fließend ist. Dies zeigt sich z.B. im Deutschen an ,abhängigen Hauptsätzen‘ und an ,Nebensätzen‘ mit reduzierter Matrixstruktur (etwa: *glaub der kommt heute nicht, vielleicht dass er kommt*: zwei oder ein Teilsatz/*clause*?; Auer 1998).

In Teilen der Forschung wird eine Alternative zur Segmentierung in Sätze darin gesehen, dass neben autonomen syntaktischen Strukturen (z.B. klassisch geformte Einfachsätze) auch elliptische Strukturen, Freie Themen, *tags*, Rechtsexpansionen und vielleicht sogar Anakoluthe als Vertreter einer selbst nur vage definierten syntaktischen Grundeinheit angesehen werden. Es ist jedoch unmöglich, für eine solche Grundeinheit eine empirisch brauchbare syntaktische Definition zu geben, die über eine Auflistung von einzelnen Konstruktionen hinausgeht. Implizit verwenden solche Untersuchungen fast immer ein prosodisches Kriterium (prosodische Selbstständigkeit), um die Spontansprache zu segmentieren. Wir wenden uns daher nun der Prosodie zu.

3.3. Segmentierung in Intonationsphrasen

Die Grenzen von Intonationsphrasen (IPs) sind nicht trivialerweise durch Pausen gegeben; Pausen können innerhalb von Intonationsphrasen entstehen, genauso wie Intonationsphrasen ohne Pause aneinander gefügt werden können (*latching*). Üblich ist, von einer IP zu fordern, dass sie zumindest einen nuklearen (Fokus-) Akzent enthält. Da dieses Kriterium aber lediglich eine notwendige, keine hinreichende Bedingung darstellt, reicht es für eine Definition nicht aus. So werden in der Literatur außerdem phrasenbegrenzende Merkmale wie finale Dehnung, initiales Accelerando auf dem Vorlauf, finale Intonationsbewegung (fallender oder steigender Grenzton), Reset der Deklination oder/und des *pitch range* zu Beginn zur Segmentierung in Intonationsphrasen herangezogen (Selting 1995, Cruttenden 1986: 35ff für eine ausführliche Diskussion). Dabei bleibt aber unklar, ob eines oder mehrere dieser Kriterien vorliegen müssen. Statt einer Operationalisierung haben wir es also eher mit einer komplexen Heuristik für das Erkennen von IP-Grenzen zu tun, deren Anwendung auf das Datenmaterial erhebliches Geschick erfordert – und manchmal dennoch nicht zu klaren Entscheidungen führt. So muss man nicht lang nach Fällen suchen, in denen eine oder mehrere Silben nicht eindeutig der vorhergehenden oder der nachfolgenden Kontur zugeordnet werden können. Betrachten wir dazu die Zeilen 55-57 aus dem Transkript.

Wie lässt sich die Äußerung *deswegen ist es doch ein wärs doch UNgewöhnlich dass er sagt ich komm jetzt mal zu IHnen und äh (-) will kein GELD haben* prosodisch segmentieren? Sie enthält drei intonatorisch hervorgehobene (akzentuierte) Silben, nämlich *UN(gewöhnlich)*, *IH(nen)* und *GELD*. (Die ersten beiden sind durch einen Hochton, die letzte durch einen steigenden Ton markiert.) Es kann sich also um eine, zwei oder höchstens drei Intonationsphrasen handeln (denn die Segmentierungsvorschrift, nach der jede IP einen Fokusakzent enthalten muss, schließt nicht aus, dass in derselben IP noch andere durch Grundfrequenzänderungen hervorgehobene Akzente auftreten). Die übrigen Segmentierungshilfen führen zu keiner Entscheidung: an den syntaktisch nahe liegenden IP-Grenzen (*ungewöhnlich / dass ... sagt / ich ... ihnen / und*) tritt keine Tempoveränderung, kein *pitch reset*, ja noch nicht einmal ein Tonhöhenprung auf, der auf einen Phrasengrenzton hinweisen könnte. (Übrigens ist nicht einmal das Ende des Ausschnitts durch einen klaren Grenzton markiert: auf *haben* schwebt die Intonation.) Die einzige wesentliche Verzögerung tritt nach der Konjunktion *und* auf, die selbst gedehnt und durch die nachfolgende gefüllte Pause zusätzlich abgetrennt wird. So lässt sich prosodisch lediglich die Segmentierung *deswegen ist es doch ein wärs doch ungewöhnlich dass er sagt ich komm jetzt mal zu ihnen und äh // will kein Geld haben* rechtfertigen, die intuitiv natürlich wenig befriedigt.

Die schon mehrfach genannten Probleme mit der Segmentierung von Abbrüchen und Expansionen in sprachliche Handlungen und Sätze gelten natürlich auch für die Prosodie. So gibt es auch abgebrochene Intonationskonturen, die zu Konflikten mit dem Postulat der Exhaustivität führen; vgl. etwa in Z.19/20:

19 .hh hat s=es geKRIGGT?
20 wenn ja' wenn NEI:N, waru[m NICHT].

Die Extraktion der Grundfrequenz ergibt hier den in Abb. (1) wiedergegebenen Verlauf.

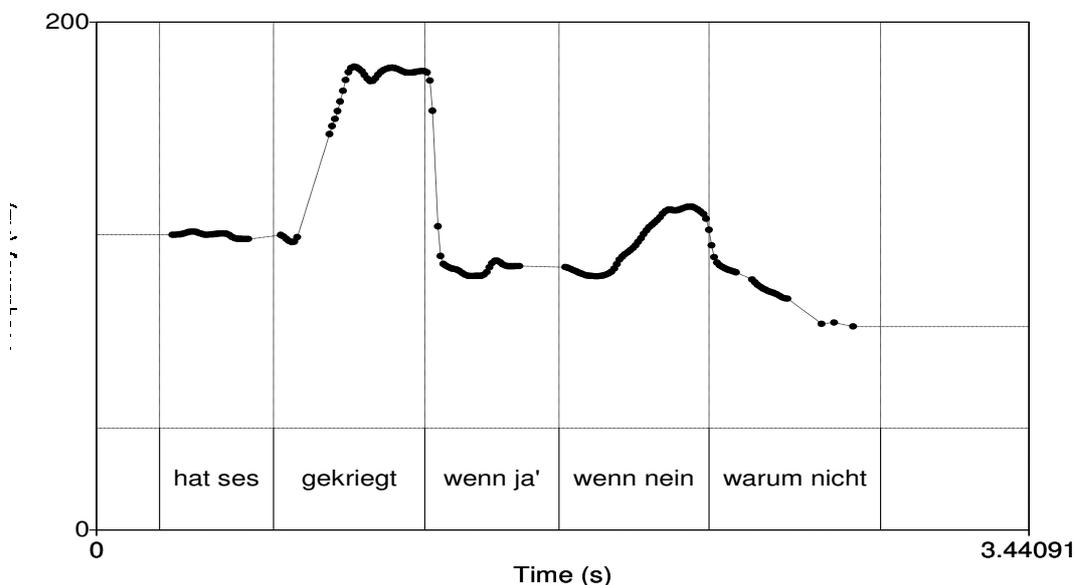


Abbildung 1

Die Tonhöhenbewegung auf der Akzentsilbe und der Neuansatz auf einer wesentlich tieferen Ebene markiert hier deutlich die Grenze von *hat ses geKRIGGT?* Wo aber beginnt die nächste IP? *Wenn NEI:N warum NICHT* ist eine vollständige Intonationsphrase, aber zwischen der ersten und der zweiten IP produziert der Sprecher zwei weitere Silben: *wenn ja*. Auf der durch Glottalverschluss abgeschnittenen Silbe *JA* fehlt der zu erwartende Fokusakzent. Wollten wir uns dafür entscheiden, auch abgebrochene Phrasen als Segmente zu betrachten, würden wir

wieder das Postulat der Ebenenkonstanz durchbrechen; abgebrochene Phrasen zeigen nämlich systematisch gerade nicht die oben genannten Indikatoren für IP-Grenzen (und manchmal, wie im Beispiel, nicht einmal eine nukleare Akzentbewegung). Sie könnten also schwerlich als IPs betrachtet werden.

3.4. ‚Kognitive‘ Grundeinheiten

Dass auch gemischte Definitionen, die auf verschiedene Ebenen sprachlicher Organisation zurückgreifen, nicht zum Ziel führen, wenn sie als Segmentierungsgrundlage verwendet werden, kann man an Chafes Vorschlag zeigen, *idea units* als Grundeinheit der Analyse zu verwenden. *Idea units* sind an sich kognitiv definiert, denn Chafe sieht sie als sprachliches Korrelat eines „single focus of consciousness“ (1985 und 1980), d.h. „the amount of information on which a person’s attention can be focused at one time“ (1985:106). Da dem Kurzzeitgedächtnis nachgesagt wird, es könne sieben Items behalten und werde alle zwei Sekunden überschrieben, vermutet Chafe, dass auch eine *idea unit* etwa sieben Wörter des Englischen umfasst und ca. zwei Sekunden lang dauert. Aber dies sind natürlich keine operativen Kriterien, sondern lediglich Durchschnittswerte. Im prototypischen Fall soll eine *idea unit* anhand der folgenden Merkmale erkennbar sein: (a) sie umfasst eine Intonationskontur, (b) sie ist von leeren oder gefüllten Pausen umgrenzt und sie stellt (c) syntaktisch einen Teilsatz (*clause*) dar.

Die prototypische Definition kombiniert damit lediglich die Probleme, die schon in der Syntax und Prosodie besprochen wurden; das Kriterium der Begrenzung durch Pausen ist zusätzlich problematisch.

4. Das Verhältnis von Redebeiträgen zu Sätzen

Im letzten Abschnitt wurde gezeigt, dass die Idee der Segmentierbarkeit dem *on-line*-Charakter (Auer 2000) der gesprochenen Sprache grundsätzlich widerspricht. Ob man auf pragmatischer, auf syntaktischer oder auf prosodischer Ebene zu segmentieren versucht: die vorgeschlagenen Einheiten können nie so definiert werden, dass die Kriterien der Exhaustivität, der Diskretheit, des Atomismus und der Ebenenkonstanz alle erfüllt sind. Natürlich kann man die Relevanz dieser Kriterien anzweifeln; solche Zweifel an der strukturalistischen Methode haben ja nicht zuletzt zum Entstehen der generativen Grammatik geführt (vgl. etwa in der Phonologie die Debatte um die Ebenenkonstanz, die Morris Halle mit einem berühmten Beitrag zur russischen Phonologie/Palatalisierung auslöste; cf. Halle 1962). Dann muss aber auch die Suche nach einer Grundeinheit der Analyse der Gesprochenen Sprache aufgegeben werden, die durch die Operation des Segmentierens nachgewiesen werden kann.

Bevor eine alternative Herangehensweise skizziert wird, ist ein mögliches Missverständnis auszuräumen: schlägt nicht auch die Konversationsanalyse eine Grundeinheit für die Segmentierung von Gesprächen vor – nämlich die Kategorie des Redebeitrags (*turns*) – und könnte nicht diese Kategorie dazu dienen, die gesprochene Sprache zu segmentieren? Tatsächlich ist jedoch der *turn* konversationsanalytisch gesehen ein uninteressantes Epiphänomen. Wie lang ein Sprecher spricht, ist eine nebensächliche Angelegenheit; wichtig sind die interaktiven Verfahren, die die Turn-Übernahme regeln. Sie bestehen bekanntlich einerseits aus bestimmten Kriterien, die sog. turnübergaberelevante Stellen zu erkennen erlauben, und andererseits den Regeln und Präferenzen, die an solchen Stellen den Sprecherwechsel regeln.⁶

⁶Die übergaberelevanten Stellen zu bestimmen ist allerdings auch in der Konversationsanalyse ein ungelöstes Problem. Der Begriff der „Turnkonstruktionseinheit“, der vor allem in Schegloff 1996 als Segmentierungseinheit verwendet wird, ist schlecht definiert und als Mischkategorie ebenso problematisch wie strittig (vgl. die ausführliche Definition bei Ford/Fox/Thompson 2002 sowie bei Selting 2000 und 2001).

Es lohnt, das konversationsanalytische Argument gegen die Segmentierung in Redebeiträge genauer zu rekonstruieren (vgl. Goodwin 1981), weil es sich mutatis mutandis genauso auf die übrigen Segmentierungsvorschläge übertragen lässt. Zunächst würde die Segmentierung in Redebeiträge ja voraussetzen, dass Gespräche durch das geordnete Nacheinander von *turns* verschiedener Sprecher gekennzeichnet sind. Jedes einigermaßen seriöse Transkript zeigt, dass das nicht der Fall ist. Gesprächsteilnehmer überlappen sich (vgl. etwa in unserem Fall die Z.62/63), sie fangen gleichzeitig an zu sprechen (vgl. Z.23/24, Z.32/33, Z.58/59), sie sprechen gelegentlich alle nicht (vgl. die Pause nach Z.43), und sie produzieren simultan zum Redebeitrag eines anderen *backchannel*-Signale (etwa Z.37/38).

Wichtiger ist jedoch ein anderer Punkt: der tatsächliche Turnwechsel ist nicht vorhersagbar, sondern wird von Sprecher und Hörer gemeinsam ausgehandelt. Ein/e Sprecher/in kann bestenfalls erreichen, dass eine Turnkonstruktionseinheit aufgrund ihrer formalen Merkmale möglichst eindeutig als Abschlusspunkt eines Redebeitrags verstanden wird und so zum Sprecherwechsel einlädt. Ob der/die Rezipient/in dieses Angebot annimmt, ist jedoch ungewiss. Tut er es nicht, wird – vielleicht nach einer Pause – der bisherige Sprecher weitersprechen, tut er/sie es, ist der vorherige Redebeitrag beendet. Da also jeder Redebeitrag theoretisch ad infinitum expandiert werden kann – ein wesentliches Merkmal von Alltagsgesprächen, das sie von manchen institutionellen Interaktionstypen unterscheidet – verlagert sich das konversationsanalytische Interesse von der Einheit *turn* zu den Verfahren, die die Teilnehmer anwenden, um *turn-taking* möglich zu machen oder zu verhindern (cf. Auer 1992, Gülich 1999).

Syntaktische, prosodische und Handlungs-Einheiten teilen mit Redebeiträgen diese Eigenschaft der Erweiterbarkeit in der Zeit. Sie können expandiert werden, eben weil auch Redebeiträge expandiert werden können (und manchmal müssen).

5. Eine Alternative zum Segmentieren

Es drängt sich als Konsequenz die Frage auf, ob das Segmentieren wirklich eine sinnvolle wissenschaftliche Aktivität ist, d.h. eine, die die interaktiv und kognitiv realen Vorgänge bei der Produktion und Interpretation von Sprache in mündlicher, direkter Kommunikation erfasst. Hier sind Zweifel erlaubt. Segmentieren bedeutet, einen Text (und sei er ein Transkript), also ein zweidimensionales Objekt, in kleine Teile zu zerschneiden. Die Schnitte müssen eindeutig sein, und es darf nichts übrig bleiben. Die Teile lassen sich dann gegebenenfalls wechselseitig ersetzen, verschieben, gruppieren, nach Umgebungen ordnen etc. Diese wissenschaftlichen Operationen reflektieren den Gestus des allmächtigen Linguisten am Schreibtisch, der das fertige Produkt sprachlichen Handelns bearbeitet. Es gibt beim Segmentieren kein Früher und Später, sondern nur ein Oben und Unten, Rechts und Links im geschriebenen/verschrifteten Text. Die Zeitlichkeit der mündlichen Sprache ist aufgehoben; dieser ist zu einem Objekt geworden, bei dem immer schon klar ist, ‚wie es weiter geht‘: der Text liegt ja fertig vor.

Teilnehmer an direkten Interaktionssituationen sind hingegen in die zeitliche Emergenz der Interaktion gezwungen; es gibt für sie kein Hin- und Herspringen zwischen später und früher. Sie müssen, um überhaupt an der Interaktion teilnehmen zu können, das Gehörte sofort danach beurteilen, ob es sich zu einer abgeschlossenen Einheit zusammenfügt oder als offene Gestalt noch einer Fortführung bedarf. Einmal getroffene Entscheidungen können revidiert werden; so erweist sich, was vielleicht zuerst als vollständig und abgeschlossen erschien, oft im nächsten Augenblick als Bruchstück einer größeren Gestalt. Es ist jedoch nicht möglich, die Interpretation eines Textes erst nach dessen vollständiger Produktion zu beginnen.

Interaktionsteilnehmer bilden also bei der *on-line*-Prozessierung der Gesprochenen Sprache keine Einheiten, sie müssen aber ständig mögliche Abschlusspunkte erkennen. Diese Abschlusspunkte können mehr oder weniger gut konturiert sein. Optimale (d.h. prägnante) Gestaltschlüsse sind dann erreicht, wenn an einem Punkt sämtliche syntaktische, prosodische und

semanto-pragmatische Projektionen abgearbeitet sind. Weniger prägnante Gestaltschlüsse werden nur auf einigen dieser drei Dimensionen eindeutig markiert oder sind innerhalb einer oder mehrerer Dimensionen nicht eindeutig markiert. Prägnante Gestaltschlüsse sind zugleich gute Kandidaten für die Übergabe des Rederechts. Das *turn taking* erfordert also eine *on-line*-Analyse der laufenden Sprachproduktion des Sprechers in formaler und inhaltlicher Hinsicht, denn gut erkennbare Gestaltschlüsse können turnübergaberelevant sein (vgl. Auer 1996). Daher muss das sprachliche Signal auf verschiedenen Dimensionen auf Projektionen und Projektionseinlösungen hin analysiert werden. Eine gute Gestalt ist aber – im Sinne der Gestaltpsychologie – immer mehr als die Summe ihrer Teile.

Im Vergleich zur Segmentierung ist die Annahme, dass die Interaktionsteilnehmer nach Gestaltschlüssen suchen, wesentlich schwächer. Man kann Abschlusspunkte suchen und finden, ohne dass die Abschnitte zwischen vollständig und exhaustiv als Einheiten beschreibbar wären. Obwohl also das ‚Zäsurieren‘ eine Bedingung des Segmentierens ist, ist es keineswegs damit identisch. Es ist auf Punkte hin orientiert, nicht auf Einheiten.

Werfen wir nun einen zweiten Blick auf den in (1) dokumentierten Interaktionsausschnitt, um den neuen Ansatz zu verdeutlichen. (In den folgenden Transkriptausschnitten sind Zäsuren in der Syntax durch ‚|‘, in der Prosodie durch ‚‡‘ und semantisch-pragmatische Abschlusspunkte durch ‚*‘ markiert.)

5.1. Syntaktische Zäsuren

Oft ist nicht einfach zu entscheiden, ob eine (fehlende) Gestaltschließung syntaktisch oder semanto-pragmatisch bedingt ist. Teilweise liegt das daran, dass Syntax normalerweise in den Grenzen des Satzes gedacht wird, jedoch auch satzübergreifend Kohäsivität auf formale Weise hergestellt wird. Eine Äußerung wie

33 B: [(-) es ging von VORNherein (-) SO an? ‡

ist im gegebenen Kontext ohne Zweifel unvollständig und daher nicht gestaltschließend. Ist diese Unvollständigkeit aber eine syntaktische (ausgelöst durch das kataphorische *so*), oder ist sie eine semantisch-pragmatische? Ich plädiere für die erste Alternative und folge damit einer weiten Auffassung von Syntax, die auch ‚textuelle‘ Verknüpfungen mit umfasst. Umgekehrt ist die Äußerung

34 wann de (--) hier (1.0) äh:: (-) der junge MANN, ‡ (-)
35 zunächst mal (.) meine FRAU angesprochen hat, |‡

formal gesehen abgeschlossen,⁷ weil der folgende Komplementsatz nicht vom Verb *ansprechen* gefordert wird (obwohl er von ihm abhängig ist):

35 .h ob
36 er' (-) bei uns nicht äh sich (-) über (-) die (-) proDUKte und
37 so weiter, ‡.h inf[ormiere KÖNNte. |‡

In beiden Fällen ist jedoch die sprachliche Handlung des Beklagten noch nicht abgeschlossen.

Syntaktische Zäsuren stimmen nicht mit syntaktischen Segmentgrenzen überein, wie sich vor allem am Ende der Syntagmen zeigen lässt. Zum Beispiel würde niemand auf die Idee kommen, das Syntagma

12 C: wir reden jetzt von der FRAU | leopold. |‡*

⁷ Der Sprecher verwendet anstelle des zu erwartenden komplementsatzeinleitenden *dass* das offenbar temporal gemeinte *wenn/wann* („als“).

in zwei Sätze oder Teilsätze zu segmentieren. Aus der *on-line*-Verarbeitungsperspektive des Rezipienten ist das Syntagma aber auch schon nach *FRAU* potenziell abgeschlossen. Allerdings ist an diesem Punkt ein denkbar schlechter Gestaltschluss erreicht: andere, v.a. prosodische Signale verhindern, dass dieser syntaktische Abschlusspunkt für die Turnübergabe relevant wird. Dennoch werden sowohl nach *Frau* als auch nach *Leopold* syntaktische Abschlusspunkte erreicht. Dies bedeutet aber natürlich nicht, dass die beiden Äußerungsteile *wir reden jetzt von der FRAU* und *Leopold* syntaktische Segmente darstellen.

Ein ähnlicher Unterschied zwischen Segmentieren und Zäsurieren ergibt sich in

```
20      wenn ja' wenn NEI:N, † waru[m | NICHT. |]*
21 B:      [<<pp>hm;> |]*
```

wo ebenfalls die schließlich entstandene syntaktische Struktur weiter ausgreift als nötig: auch nach *warum* wäre (im vorliegenden Kontext) schon ein syntaktischer Abschlusspunkt möglich gewesen. In solchen Fällen wird die Relevanz des möglichen Abschlusspunkts auf syntaktischer Ebene durch die Prosodie überdeckt und außer Kraft gesetzt: in beiden Fällen ist das syntaktisch expandierende Element (*leopold*, *NICHT*) prosodisch vollständig in die Intonationskontur integriert. Das zweite Beispiel ist auch deshalb interessant, weil der Rezipient die Prozessierung der Äußerung des Schlichters durch die Platzierung seines Rezipientensignals sichtbar macht. Der frühestmögliche Punkt für die unmarkierte Platzierung ist dort erreicht, wo der weitere Verlauf der projektierten Äußerung so stark determiniert ist, dass ihn der Rezipient mit großer Wahrscheinlichkeit vorhersagen kann. Dies ist bereits am ersten möglichen syntaktischen Abschlusspunkt der Fall.

Im folgenden Fall ist die Expansion in Z.30 hingegen sowohl syntaktisch wie auch prosodisch eindeutig als solche markiert, d.h. nach Z.28 ist bereits ein relativ guter Gestaltschluss erreicht:

```
28 C:  sie hat UNendgeltlich gearbeitet.= |]*
29 B:  =JAwoll; |]* (0.5)
→30 C: um GOTteslohn. |]*
```

Dennoch sind die Syntagmen in Z.28 und Z.30 aufeinander bezogen. Das zweite baut auf dem Muster des ersten auf, denn das in ihm enthaltene Adverb *UNendgeltlich* wird nun durch *um GOTteslohn* ersetzt. In einem segmentbasierten Ansatz ist schwer zu entscheiden, ob *um GOTteslohn* eine selbständige Einheit (Ellipse? Satz?) auf derselben Segmentierungsebene wie 28 ist. In einem auf Zäsuren hin orientierten Ansatz ist hingegen klar, dass nach beiden Syntagmen ein syntaktischer und prosodischer Abschlusspunkt erreicht ist; über den Status des expandierenden Elements *um GOTteslohn* ist damit aber noch nichts gesagt (vgl. dazu Auer 1996 und 1991).

Vermutlich hängt die Überschreibbarkeit syntaktischer Abschlüsse durch prosodische Integration der Expansion sowohl von der Art (v.a. Stärke) der syntaktischen wie auch von der Art (Stärke) der prosodischen Signalisierung ab. In

```
24 C:  [WAS |* (.)] stimmt nicht? |]*
```

liegt zwischen *was* und *stimmt nicht* syntaktisch gesehen eine relativ schwache (Phrasen-) Grenze (NP/VP), und auch die prosodische Integration der beiden Äußerungsteile ist wenig ausgeprägt. (*WAS* weist durch den Fokusakzent und die nachfolgende kurze Pause Merkmale einer selbständigen Kontur auf, gegen die lediglich das Fehlen eines finalen Grenztons spricht.) Die mittlere Kohäsivität auf der syntaktischen Ebene wird durch unklare prosodische Zäsurierung weder im Sinne eines eindeutigen, noch im Sinne eines fehlenden Gestaltschlusses entschieden. Bei ‚satzwertigen‘ Expansionen wie im folgenden Fall liegen hingegen stärkere syntaktische Grenzen (weniger Kohäsivität) vor:

47 B: der junge MANN hat von vorn heREIN; † (-) schon gar
 48 nicht von einer (-) beZAHlung (-) geSPROChen; |†*
 49 dass er sie verLANGT hätte; |†* (-)

Der ‚weiterführende‘ *dass*-Satz in 49 ist eine nicht-notwendige syntaktische Expansion des vorausgehenden Syntagmas. In diesem Fall wird die starke syntaktische Grenze durch die eindeutige prosodische Grenzmarkierung unterstützt.

Wo Abbrüche (Anakoluthen) vorkommen, etwa in

31 B: es ging (-) <<kurz>>ja. > |†*

bleiben offene Strukturen stehen: *es ging* erreicht keinen syntaktischen (und auch keinen prosodischen oder semantischen) Abschluss. Dieses Fragment muss und kann keinem Segment zugeordnet werden. Die darauf folgende Äußerungskomponente *ja* ist hingegen eine vollständige und abgeschlossene Gestalt.

5.2. Prosodische Zäsuren

Bei der prosodischen Analyse muss zwischen dem Abschluss von weiterweisenden Konturen, etwa

37 B: er MÖChte gErn
 39 (.) da bissel (-) öh: (2.0) hier HERkommen, † |

und solchen, die turnbeendend sein können, etwa

41 (--) von eim
 42 (.) von eim (.) STUNdenlohn oder so; † = war NIE die rede |
 43 gewesen. |†* (1.0)

unterschieden werden. (Im Transkript sind die weiterweisenden Konturen durch Komma, die abschließenden durch Punkt, Strichpunkt und Fragezeichen markiert.) Letztere lassen sich im Deutschen vor allem an deutlich fallenden oder deutlich steigenden finalen Tonhöhenbewegungen erkennen.⁸

Grundsätzlich sind nur wenige prosodische Gestaltschlüsse nicht auch syntaktische Abschlusspunkte. Gerade im Fall des Sprechers B, der mühsam und unter dem Stress der institutionellen Situation formuliert, kommt es aber durchaus vor, dass innerhalb einer emergenten syntaktischen Einheit (sogar innerhalb emergenter Einfachsätze) prosodisch zäsuriert wird, wie etwa in dem eben zitierten Beispiel

41 (--) von eim
 42 (.) von eim (.) STUNdenlohn oder so; † = war NIE die rede |
 43 gewesen. |†* (1.0)

nach *oder so* oder in

67 (--) <<all> wir hAm ja auch von VORnerein g(e)SACHT, > † .h
 68 dass die situatiO:N, † (-) äh:: (-) zunächst also jetzt mal SO is, |†

nach *situatioO:N*. Abschlussmarkierung allein mit prosodischen Mitteln hat allerdings nur geringe Chancen, für das *turn-taking* relevant zu werden.

Umgekehrt gibt es zahlreiche Fälle, in denen ein syntaktischer Abschluss nicht durch einen Grenzton markiert wird. Sprecher können durch eine solche prosodische ‚Camouflage‘ eines

⁸Im Detail sind die Verhältnisse viel komplexer und außerdem regional verschieden; vgl. Gilles 2005.

syntaktischen Abschlusspunkts den Gestaltschluss verschlechtern und dadurch die Turnübergabe zu verhindern suchen. So wird in

14 C: SIE sagt sie sei vom siebten ERSten bis zwölften ZWOten
 15 als AUShilfe gegen einen STUNdenlohn von zehn mark
 16 beSCHÄFtigt gewesen== |* sie rechnet sich (vier'
 17 zweihundertvierunddreißig) stunden AUS,

der syntaktische Abschlusspunkt nach *gewesen* durch unmittelbaren Anschluss (markiert durch ‚=‘) und schwebende Intonation überdeckt. Dass die syntaktische Grenze durch mangelnde prosodische Grenzmarkierung dennoch nicht immer völlig außer Kraft gesetzt werden kann, zeigen typische Fälle von Turbulenzen bei der Verteilung des Rederechts, die an solchen Stellen entstehen können:

55 C: (-) deswegen ist es doch ein wär=s doch UNgewöhnlich dass
 56 er sagt ich komm jetzt mal zu IHnen und äh: (-) will kein
 57 GELD haben |*(0.5)
 58 [wie soll denn des GEHN;= |+*
 59 B: [genAusso WAR=S. |+*

Der Rezipient setzt trotz der schwebenden finalen Intonation in Z. 57 mit seiner bestätigenden Äußerung in Z. 59 ein, während der Schlichter in Überlappung seinen Redebeitrag expandiert (Z. 58). Kongruieren syntaktische und prosodische Grenzmarkierung also nicht (weil nur auf einer Signalisierungsebene Abschluss markiert wird), ist aus der Perspektive des *turn taking* die Syntax der stärkere Indikator.

5.3. Zäsurierung nach Handlungsgrenzen

Schießlich lassen sich im Gespräch zumindest in einigen Fällen klare semanto-pragmatische Abschlusspunkte erkennen, und zwar dort, wo eine sprachliche Handlung erkennbar abgeschlossen ist. Die revidierte Transkription geht davon aus, dass die pragmatisch-inhaltliche Interpretation insofern von der formalen (syntaktischen und prosodischen) Zäsurierung abhängig ist, als Handlungsgrenzen syntaktische Abschlüsse voraussetzen. Umgekehrt ist es aber möglich, dass eine sprachliche Handlung über zahlreiche formale Abschlusspunkte hinweg ausgedehnt wird. Diese Diskrepanz zwischen Syntax und Semanto-Pragmatik wird vor allem in längeren Redebeiträgen relevant; vgl. etwa Z. 33-43. Solche längeren Redebeiträge werden oft durch initiale Vorlaufhandlungen eingeleitet, durch die der Sprecher seinen umfangreichen Redebeitrag ankündigt und lizenziert; hier etwa durch

33 B: [(-) es ging von VORNherein (-) so an? † (--)

5.4. Eine revidierte Analyse

Im Folgenden wird das Gesamttranskript noch einmal insgesamt wiedergegeben. Dabei handelt es sich natürlich nur um eine Variante der Zäsurierung: wie bereits angedeutet, sind manche Gestaltschlüsse nicht eindeutig.

(2) Vorschlag für eine Zäsurierung des Gesprächs nach syntaktischen, prosodischen und handlungsbezogenen Abschlusspunkten

01 C: leopold gegen may? |+* (-) und abel gegen may. |+*
 02 ((Die Beteiligten nehmen ihre Plätze ein.))*
 03 C: ja. |+ (-) herr may |* was sAgen se zu der KLA:ge? |+* =
 04 die KLÄgerin sagt sie: hätt noch GELD zu kriegen; |+*
 (2.5)
 05 B: dazu is zu SAgen, † (1.5) dass äh:: (2.0) der KLÄ:ger, †

06 C: .ts <<undeutlich> (äh wir reden nur über) die FRAU |
07 leopold |* jetzt |* ne?> |*
08 B: ((räuspert sich)) <p>BITte?> |*
09 C: <<pp>wir REden (-) über die FRAU | leopold |* jetzt.=> |*
10 B: =NEIN, |* <<pp>(red ich NICHT).> |*(1.0)
11 für den jungen MANN da. |*
12 C: NEIN, |* wir reden jetzt von der FRAU | leopold. |*
13 B: ach | * SO:; |*
(0.5)
14 C: SIE sagt sie sei vom siebten ERSTen bis zwölften ZWOTen
15 als AUSHilfe gegen einen STUNdenlohn von zehn mark
16 beSCHÄftigt gewesen= |* sie rechnet sich (vier'
17 zweihundertvierunddreißig) stunden AUS, | *#=für
18 zweitausenddreihundertvierzig mark NETto, |* (-)
19 .hh hat s=es geKRIGGT? |*
20 wenn ja' wenn NEI:N, † waru[m | NICHT. |*
21 B: [<<pp>hm;> |*
22 (1.0) <<mf>das stimmt NISCHT.> |*
23 [es war]
24 C: [WAS |* (.)] stimmt nicht? |*
25 B: es war (1.0) bei (1.0) der (2.0) frau leopold oder so
26 (-) bei FROLlein leopold (--)<<staccato>KEInerlei (-)
27 beZAHLung (-) verEINbart |*
28 C: sie hat UNendgeltlich gearbeitet.= |*
29 B: =JAWoll; |* (0.5)
30 C: um GOTTeslohn. |*
31 B: es ging (-) <<kurz>ja. > |*
32 C: [<<pp>aha, > |*
33 B: [(-) es ging von VORNherein (-) SO an? † (--> wenn de'
34 wann de (--> hier (1.0) äh:: (-) der junge MANN, † (-)
35 zunÄchst mal (.) meine FRAU angesprochen hat, |† (-) .h ob
36 er' (-) bei uns nicht äh sich (-) über (-) die (-) proDUKte und
37 so weiter, †.h inf[ormiere KÖNNte. |† (-) er MÖchte gErn
38 C: [<<pp>mhm. |*
39 B: (.) da bisschen (-) öh: (2.0) hier HERkommen, |† NET wahr, |† (-)
40 und äh sein WEIL sein WISSen, † (-) hier etwas AUSzubauen, |†
41 (--> <<staccato>von (-) einer (-) beZAHLung;> † (--> von eim
42 (.) von eim (.) STUNdenlohn oder so; † = war NIE die rede |
43 gewesen. |* (1.0)
44 C: äh: (-) versteh ich sie rIchtig. |† (-) IST davon gar nicht
45 geSPROChen worden? |* =0:der hat man verEINbart dass nichts
46 beZAHLT wird. |* (-)
47 B: es: der junge MANN hat von vorn heREIN; † (-) schon gar
48 nicht von einer (-) beZAHLung (-) geSPROChen; |*
49 dass er sie verLANGT hätte; |* (-)
50 wäre des der FALL gewesen dann hätten wir=s ABgele:hnt; |* (-)
51 C: ja | herr may | also ei' EINS versteh ich jetzt NICHT | so
52 richtig. |† äh <<stottert>ge' ge' gewöhnlicherWEIse,> †.hh
53 äh (-) LEben die menschen,= † von dem was sie
54 verDIEnen; |† =nur WEnige können von ihm verMÖgen leben. |†
55 (-) deswegen ist es doch ein wär=s doch UNgewöhnlich dass
56 er sagt ich komm jetzt mal zu IHnen und äh: (-) will kein
57 GELD haben |*(0.5)
58 [wie soll denn des GEHN;= |*
59 B: [genAuso WAR=S. |*
60 C: = sogar AUSzubildende kriegen doch geld; |*
61 und die (.) können ja AU nix. |*
62 [ne? |*
63 B: [SO war=s | gewEsen= |*(über);† (-) über beZAHLung war NICHT die
64 REde. |*
65 (1.0)
66 ABsolut nIcht. |*
67 (--> <<all> wir hAm ja auch von VORNerein g(e)SACHT,> † .h

68 dass die situatiO:N, † (-) äh:: (-) zunächst also jetzt mal SO is, †
 69 .hh im HINblick auf die: (.) Umsätze;= † die also h (-) um: (1.0)
 70 (-) wie sie jetzt (.) im VIERTel letzten jahr der FALL
 71 war; †.h FÜNfundzwanzig prozent NACHgelassen haben; † (-)
 72 .h und wir sind NICHT in der LA:ge, † (-) persoNAL, † (-) äh z'
 73 daZUnehmen? |†* (.) [nach dem was wir HABen, |†* (-)
 74 C: [.mhm; |†*
 75 nIcht WAHR? |†*
 76 (-) und äh damit war der fall erLEdigt. |†*
 77 =es war KEInerlei (-) FORderung gestellt worden; |†*
 78 (-) wir wie geSACHT, †= wir hätten das ABgelehnt; |†*

Insgesamt kongruieren in unserem Ausschnitt 46 Abschlüsse auf allen drei Ebenen (Rezipientensignale mitgezählt), d.h. es wird ein optimaler Gestaltschluss erreicht.⁹ 8 Abschlusspunkte sind nur syntaktisch, 23 nur prosodisch markiert; in 10 Fällen wird mit einem syntaktisch-prosodisch markierten Abschluss keine sprachliche Handlung abgeschlossen. In weiteren 10 Fällen wird der Abschluss der sprachlichen Handlung lediglich von einem syntaktischen, jedoch nicht von einem prosodischen Gestaltschluss begleitet. Zu einem tatsächlichen Wechsel des Rederechts kommt es lediglich in 23 Fällen, von denen nur drei nicht auf sämtlichen Ebenen als mögliche Abschlusspunkte markiert werden (nämlich in Z. 27 und Z. 57, wo die intonatorische Abschlussmarkierung fehlt, und in Z. 06, einer klaren Unterbrechung, wo keinerlei Abschlusspunkt erreicht wird).

6. Abschließende Bemerkungen

Es ist ein wenig eigenartig, dass der strukturalistische Begriff des Segmentierens sich noch immer in der Literatur zur Gesprochenen Sprache großer Beliebtheit erfreut, obwohl das Gros der Forschung erklärtermaßen nicht strukturalistischen Theorien folgt. Ich habe in diesem Beitrag zu zeigen versucht, dass es sich dabei nicht nur um eine terminologische Unachtsamkeit handelt. Vielmehr suggeriert der Begriff des Segmentierens eine Herangehensweise an Gesprochene Sprache, die nicht mit deren grundsätzlicher Eigenschaft der *on-line*-Prozessierung zur Deckung zu bringen ist. Es liegt daher nahe, diesen Begriff endgültig über Bord zu werfen. Dadurch bekommt die Frage nach den ‚Einheiten‘ in der Gesprochenen Sprache eine ganz andere Wendung; sie lassen sich am besten durch den Begriff der mehr oder weniger prägnanten Gestaltschlüsse ersetzen, die ihrerseits unmittelbar an Fragen des *turn taking* angeschlossen werden können.

7. Literaturverzeichnis

- Auer, Peter, 1991, Das Ende deutscher Sätze – Rechtsexpansionen im deutschen Einfachsatz. In: *ZGL* 1991, S. 139-157.
- Auer, Peter, 1992, The neverending sentence: on rightward expansion in spoken syntax. In: M. Kontra/T. Váradi (Hrsg.), *Studies in Spoken Languages: English, German, Finno-Ugric*. Budapest: Hungarian Academy of Sciences, 1992, S. 41-60.
- 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, 1997, Formen und Funktionen der Vor-Vorfeldbesetzung im Gesprochenen Deutsch. In: P. Schlobinski (Hrsg.), *Syntax des gesprochenen Deutsch*, Westdeutscher

⁹ Auch Croft (1995) kommt in einer Untersuchung zum Englischen zu einer sehr hohen Übereinstimmung prosodischer und syntaktischer Grenzen.

Verlag, S. 55-92.

- Auer, Peter, 1998, Zwischen Parataxe und Hypotaxe: 'abhängige Hauptsätze' im Gesprochenen und Geschriebenen Deutsch. In: *ZGL* 1998, S. 284-307.
- Auer, Peter, 2000, On-line-Syntax – oder: was es bedeuten könnte, die Zeitlichkeit der mündlichen Sprache ernst zu nehmen. In: *Sprache und Literatur* 85 (2000), S. 43-56 (Themenheft: Die Medialität der Gesprochenen Sprache).
- Auer, Peter, 2007, Syntax als Prozess. In: H. Hausendorf (Hrsg.), *Gespräch als Prozess. Linguistische Aspekte der Zeitlichkeit verbaler Interaktion*. Tübingen: Narr, S. 95-142.
- Bloomfield, Leonard, 1933, *Language*. New York: Holt, Rinehart & Winston.
- Caroli, Folker, 1977, Pragmatische Aspekte syntaktischer Variation in der Gesprochenen Sprache. Göppingen: Kümmerle.
- Chafe, Wallace, 1980, The deployment of consciousness in the production of a narrative. In: W. Chafe (Hrsg.), *The Pear Stories. Cognitive, cultural and linguistic aspects of narrative production*. Norwood: ALEX (= Advances in Discourse Processes, Vol III).
- Chafe, Wallace, 1985, Linguistic differences produced by differences between speaking and writing. In: D.R. Olson/N. Torrance/A. Hildyard (Hrsg.), *Literacy, Language, and Learning*. Cambridge: CUP, S. 9-50.
- Croft, William, 1995, Intonation units and grammatical structure. In: *Linguistics* 33 (1995), S. 839-882.
- Cruttenden, Alan, 1986. *Intonation*, Cambridge: CUP.
- Crystal, David, 1979, Neglected grammatical factors in conversational English. In: S. Greenbaum/G. Leech/J. Svartvik (Hrsg.), *Studies in English Linguistics for Randolph Quirk*. London, Longman, S. 153-166.
- Duden 1973 = Duden Grammatik der deutschen Gegenwartssprache, bearb. von P. Grebe u.a., 3., neu bearb. u. erw. Aufl. Mannheim: Bibliograph. Inst.
- Ford, Cecilia E./Fox, Barbara A./Thompson, Sandra A., 2002, Constituency and the grammar of turn increments. In: dies. (Hrsg.), *The Language of Turn and Sequence*. Oxford: OUP, S. 14-38.
- GAT = M. Selting u.a., Gesprächsanalytisches Transkriptionssystem. In: *Linguistische Berichte* 34 (Hft. 173, 1998), S. 91-122.
- Gilles, Peter, 2005, Regionale Prosodie im Deutschen: Variabilität in der Intonation von Abschluss und Weiterweisung. Berlin: de Gruyter
- Goodwin, Charles, 1981, *Conversational Organization*. New York: Academic Press.
- Gülich, Elisabeth, 1999, Les activités de structuration dans l'interaction verbale. In: J.-M. Barbéris (Hrsg.), *Le Français parlé*. Montpellier: Univ. Paul Valéry, S. 21-47.
- Halle, Morris, 1962, Phonology in generative grammar. In: *Word* 18 (1962), S. 54-72.
- Harris, Zellig, 1951 (⁸1969) *Structural Linguistics*. Chicago: UP.
- Leska, Christel, 1965 (1975), Vergleichende Untersuchungen zur Syntax gesprochener und geschriebener Gegenwartssprache. PBB (Ost) 87, 427-464. Erweiterte Fassung: Statistische Untersuchungen zur Syntax gesprochener und geschriebener deutscher Gegenwartssprache (C. Höhne-Leska). Berlin : Akad.-Verl..
- Rath, Rainer, 1990, ‚Satz‘ und ‚Äußerungseinheit‘. Syntaktische und interaktive Struktur in der Sprache? In: E. Leupold/Y. Petter (Hrsg.), *Interdisziplinäre Sprachforschung und*

Sprachlehre. Tübingen: Narr, S. 197-216.

Schegloff, Emanuel, 1996, Turn organization: one intersection of grammar and interaction. In: E. Ochs/E. Schegloff/S. Thompson (Hrsg.), *Interaction and Grammar*, Cambridge: CUP, S. 52-133.

Schröder, Peter, 1997, (Hrsg.) Schlichtungsgespräche: ein Textband mit einer exemplarischen Analyse. Berlin: de Gruyter (=Schriften des Instituts für deutsche Sprache Bd. 5) .

Selting, Margret, 1995, Prosodie im Gespräch. Tübingen: Niemeyer.

Selting, Margret, 2000, The construction of units in conversational talk. In: *Language in Society* 29 (2000), S. 477-517.

Selting, Margret, 2001, Fragments of units as deviant cases of unit production in conversational talk. In: M. Selting/E. Couper-Kuhlen (Hrsg.), *Studies in Interactional Linguistics*. Amsterdam: Benjamins, S. 229-258.

Wunderli, Peter, 1979, Satz, Paragraph, Text - und die Intonation. In: J.S. Petöfi (Hrsg.), *Text v. Sentence*. Hamburg: Buske, S. 319-341.

Anhang: Transkriptionskonventionen

((hustet))	Kommentar
(wort)	vermuteter Wortlaut
wo'	Abbruch (phonetisch eindeutig markiert)
[bla bla] bla	eckige Klammern markieren simultane Passagen
[worte]	
=	schneller Anschluss oder Verschleifung
.h	einatmen
`	Abbruch
(.)	Mikropause
(-)(--)(---)	Pausen bis zu 1 Sekunde
(2.0)	gezählte Pause
blaBLA	Akzent (Primärakzent)
blabLA	Sekundärakzent (nicht immer transkribiert)
wort,	leicht steigende Intonation
wort?	stark steigende Intonation
wort;	leicht fallende Intonation
wort.	stark fallende, finale Intonation
(h)wort	lachend gesprochen
hehehe	silbisches Lachen
<<f>/<<mf>/<<ff>wort>	laut/mittellaut/sehr laut gesprochene Passage, endet bei >
<<p>/<<pp>/<ppp>wort>	leise/sehr leise gesprochene Passage, endet bei >
<<all>wort>	schneller gesprochen, endet bei >
<<staccato>wort>	„abgehackt“ gesprochen, endet bei >
	Zäsuren in der Syntax
‡	Zäsuren in der Prosodie
*	semantisch-pragmatische Abschlusspunkte