2. Grundeinheiten der Sprache und des Sprechens

Abstract: Der Beitrag widmet sich der Frage, nach welchen Einheiten das Sprechen in der Interaktion auf der ‚mittleren Ebene‘ strukturiert ist. Die Diskussion geht aus von für die gesprochene Sprache typischen bzw. spezifischen Phänomenen, die Kriterien normgrammatisch vollständiger Sätze zuwider laufen, wie z. B. Ellipsen, Expansionen und Diskursmarkern. Anschließend werden die in der Literatur der vergangenen etwa 40 Jahre theoretisch entwickelten und empirisch untersuchten Einheitenkonzepte diskutiert und im Hinblick auf ihre Leistungsfähigkeit zur Beschreibung und Erklärung nicht-normgrammatischer Phänomene evaluiert.

- 1 Einleitung
- 2 Die Unzulänglichkeit der Kategorie ‚Satz‘ für das Sprechen
- 3 Einheitenkonzepte
- 4 Problemfälle
- 5 Fazit
- 6 Literatur

1 Einleitung

Dieser Beitrag widmet sich der Frage, in welchen Einheiten das Sprechen in der Interaktion (ab hier: ‚Sprechen‘) auf der ‚mittleren Ebene‘ strukturiert ist. In Anlehnung an Fiehler u. a. (2004) können wir in Bezug auf das Sprechen folgende Ebenen unterscheiden:

- Phoneme,
- Wörter,
- die hier interessierende ‚mittlere Ebene‘,
- Sequenzen bzw. Themen,
- kommunikative Gattungen,
- Interaktionen.

Jeder dieser Einheitentypen ist durch eigene autonome Organisationsprinzipien charakterisiert und sie sind jeweils konstitutiv für Einheiten der nächsthöheren Ebene. In unserem Beitrag geht es um Einheiten oberhalb der Wortebene und unterhalb der Sequenz- und Themenebene. Für die traditionelle, monologische konzeptionelle Schriftsprache (Koch/Oesterreicher 1985) ist wohl unstrittig, dass die Einheiten dieser Ebene kanonisch Sätze sind bzw. dass nicht kanonische Strukturen (Ellipsen)

in Bezug auf das Modell des Satzes analysiert und verstanden werden können. Dieser Beitrag will zunächst anhand eines exemplarischen Beispiels zeigen, dass und warum die Strukturierung des Sprechens vielfach nicht der Kategorie ‚Satz‘ folgt und dass diese folglich nicht als Grundeinheit zur Beschreibung und Analyse des Sprechens taugt (2.). Anschließend stellen wir alternative Konzepte vor, die für die Beschreibung der Einheitenkonstitution im Sprechen entwickelt wurden (3.). Wir diskutieren die jeweiligen Einheitenbildungskriterien, ihre Motivation, ihre Leistungsfähigkeit und ihre Probleme. Schließlich gehen wir auf Strukturen des Sprechens in der Interaktion ein, die von keiner der bisher vorgeschlagenen Einheitenkonzeptionen befriedigend behandelt werden können (4.). Wir schließen mit einer kritischen Reflexion der Leistungsfähigkeit und Notwendigkeit von Einheitenkonzepten für die Analyse des Sprechens (5.).

2 Die Unzulänglichkeit der Kategorie ‚Satz‘ für das Sprechen

Unserer Untersuchung legen wir zugrunde, dass ‚Satz‘ dadurch definiert sei, dass eine syntaktische Struktur mindestens eine finite Verbform und in der Regel ein Subjekt sowie die Realisierung der obligatorischen verbgebundenen Ergänzungen beinhalte (Forsgren 1992). Das schließt regelhaft subjektlose Strukturen wie z. B. das unpersönliche Passiv ein. Sätze zeichnen sich dadurch aus, dass sie (etwa im Gegensatz zu Turns und Intonationsphrasen, vgl. Abschnitt 3) kontextfrei als solche feststellbar sind und ausschließlich durch Verhältnisse auf der syntaktischen Ebene definiert sind (Deppermann 2012). Ein erheblicher Teil der Strukturen des Sprechens ist nicht satzförmig. Dabei handelt es sich keineswegs nur um ‚Performanzphänomene‘ im Sinne der generativen Grammatik, d. h. um Abbrüche oder andere defektive Strukturen (Anakoluthe), die durch kommunikativ nicht systematisch eingesetzte Planänderungen, Aufmerksamkeitsprobleme oder Störungen durch nicht vom Sprecher zu verantwortende Ereignisse verursacht sind. Vielmehr ist ein großer Teil der sprachlich-kommunikativen Praktiken, die routinemäßig zur Bearbeitung bestimmter Kommunikationsaufgaben eingesetzt werden, nicht satzförmig. Diskurspartikeln (Rückmelder, Diskursmarker, Häsitiationssignale, Interjektionen usw.), Ellipsen, Analepsen, Links- und Rechtsversetzungen, freie Themen, Apokoinukonstruktionen oder Vokative sind Beispiele für nicht-satzförmige, usuelle Strukturen des Sprechens im Deutschen, die kommunikativ systematisch eingesetzt werden (im Überblick: Schwitalla 2012). Andere Konstruktionen wie z. B. viele Arten von Expansionen im Nachfeld, Selbstkorrekturen und -reformulierungen genügen ebenfalls nicht dem Satzkriterium, da sie topologisch und/oder rektionsbezogen nicht auf die zuvor etablierte Satzstruktur bezogen sind (Auer 1991). Deppermann (2012) stellte in einer Untersuchung von 6 mal 10 Minuten zufällig ausgewählter Ausschnitte von Sprechen

in unterschiedlichen Interaktionstypen (informelle Freizeitgespräche, Rettungsübungen, Sprechen während eines Tischfußballspiels, biographisches Interview, Psychotherapie, Arzt-Patient-Gespräche) fest, dass nur 18,4 % der in den 60 Minuten produzierten 1010 Gesprächsbeiträge (Turns) mit einem Satz begannen bzw. aus einem oder mehreren Sätzen bestanden. 56 % der Beiträge begannen dagegen mit einer syntaktisch nicht eingebundenen Diskurspartikel (Schwitalla 2002; Fischer 2006) bzw. bestanden lediglich aus einer oder mehreren Diskurspartikeln.

An einem Beispiel wollen wir zeigen, auf welche Probleme der Versuch, Interaktionen in Satzstrukturen (exhaustiv) zu segmentieren, stößt, und einige der nicht-satzförmigen Strukturen zeigen, die systematisch in der Interaktion eingesetzt werden. In #1 spricht PA mit JO und UD darüber, wie er seine Kopfschmerzen behandelt hat.

#1 FOLK_E_00066_SE_01_T_01, Gartengespräch, 00:12:16

```

01  PA   hab dann heut ma mein KRÄUterbuch,
02        (0.51) zu RÄte gezogen,
03        die ha:m waCHOLderte und LÖwenzahnte empfohlen,
04        (0.76) [ich hab mir dann] bEIdes gEMACHT,
05  JO           [HM_m.           ]
06  PA   und außerdem noch mit (.) fröhmesner BAdöl-
07        wo lavendel und rOsmarin drin is EIngeschmiert,
08        (0.98)
09  JO   nIch gleich die rauWOLFia.
10        (0.95)
11  PA   ä:h sagt mir jetzt NIX.
12        (1.67)
13  PA   nee aber ä:hm
14        (.)
15  UD   was IS das denn.
16        (0.59)
17  JO   die rauWOLFia?
18        (0.34)
19  UD   ne PFLANze oder [was.   ]
20  JO           [isch ne] PFLANze:-
21        so wie:-
22        (0.55)
23  PA   un is auch [bei Ischiasn]erv gut oder wie.
24  JO           [TOLLkirsche.]
25        (0.29) is einfach n: (0.35) beTÄbungsmittel.
26        (0.49)
27  PA   ah.
28        (0.38) nee dis dis waCHOLderte,
29        un LÖwenzahnte,
30        dis soll wohl irgendwie gut für_n Ischiasnerv sein,
```

In diesem Interaktionsausschnitt finden wir viele übliche Strukturen, die im Sinne der o. g. Definition nicht satzförmig sind:

- Sprecherellipse (Hoffmann 1999): Topik-*drop*: „ø hab dann heut...“, S01;
- Objektellipse (Hoffmann 1999): „und außerdem ø noch mit (...) eingeschmiert“, S06-07;
- Ellipse von Subjekt, Verbalkomplex und ggfs. Fragepronomen: „ø nlch gleich die rauWOLFia.“, S09;
- Analepse: Topik-*drop* (Helmer i. V.): „ø sagt mir jetzt NIX.“, S11, ebenso in S20 und S23; Verstehensprüfung durch Formulierung nur des kritischen Elements (Selting 1987): „die rauWOLFia?“, S17 und „ne PFLANze oder was.“, S19;
- Rückmeldepartikeln: „HM_m“, S05 (Ehlich 1986; Schmidt 1988); „ah.“, S27 (Golato/Betz 2008);
- Antwortpartikeln (Sorjonen 2001): „nee“, S13 und S28;
- Häsitationsignal (Clark/Fox-Tree 2002): „äh“, S11;
- Rückversicherungspartikel (Rehbein 1979): „oder wie.“, S23;
- Linksversetzung: In S28-30 werden zunächst zwei nominale Elemente thematisiert: „dis dis waCHOLDertee, un LÖwenzahntee,“, die anschließend pronominal im Satzrahmen wiederaufgenommen werden „dis soll wohl [...]“ („Referenz-/Thematisierungs-Aussage-Struktur“, Fiehler u. a. 2004);
- Abbruch (S13; Selting 2001);
- Diskontinuierliche Strukturen: In S17-24 produziert JO in vier inkrementellen Schüben die Aussage „die rauWOLFia? isch ne PFLANze:– sO wie:– TOLLkirsche.“. Die Produktion dieser Aussage geschieht aber nicht autonom, sondern sie reagiert auf UDs Frage in S19 „ne PLFANze oder was.“, und die Aussage wird fortgeführt, während bereits ein anderer Sprecher (PA) das Rederecht übernommen hat (S23). Die Produktion eines Satzes geschieht hier also über mehrere Turnkonstruktionseinheiten, Turns und Sprecherwechsel hinweg und bezieht Beiträge des Adressaten in die Satzkonstruktion ein (Goodwin 1979);
- Reanalyisierte Strukturen: „die rauWOLFia“ wird in S17 zunächst als Verstehensprüfung formuliert, durch die Fortsetzung „isch ne planze“ in S20 dann aber retrospektiv als Referenz einer als Antwort produzierten Aussage reinterpretiert. Wir sehen hier, dass der einheitenbezogene Status eines Elements und die Einheitenextension ambig sein kann, da sie durch folgende Interaktionsbeiträge neu interpretiert werden kann, ohne dass dadurch die vorangegangene Interpretation (etwa als auf einem Missverständnis beruhend) ungültig werden muss (Auer 2010).

Aufgrund der Usualität dieser Strukturen ist die Kategorie ‚Satz‘ als Grundkonzept für die exhaustive Analyse von Einheiten des Sprechens untauglich. Damit soll weder behauptet werden, dass Sprecher keine Sätze produzieren noch dass sie sich bei der Produktion und Rezeption von Äußerungen in der Interaktion nicht oftmals an Satzformigkeitskriterien und damit verbundenen Fortsetzungs- und Gestaltschlusser-

wartungen orientieren (Selting 1995a und in diesem Band). Im Gegenteil, es ist eine empirisch noch keineswegs gut geklärte Frage, welche spezifischen Leistungen Sätze in der Interaktion im Vergleich zu anderen interaktiv möglichen und usuellen Strukturen erbringen (Deppermann 2012). In Bezug auf die Frage nach den Grundeinheiten des Sprechens aber stellt sich die Frage, welcher Logik die Konstruktion von kommunikativ vollständigen, für die Interaktionsteilnehmer erkennbaren und als solche verstandenen Einheiten folgt. Damit verbinden sich folgende Fragen:

- Sind die relevanten Einheiten linguistisch (prosodisch, syntaktisch), pragmatisch (handlungs- oder intentionsbezogen), interaktionsorganisatorisch (mit Bezug auf den Sprecherwechsel) oder semantisch (propositional oder konzeptuell) konstituiert?
- In welchem Verhältnis stehen die verschiedenen potenziell einschlägigen Ebenen und Phänomene der Einheitenbildung zueinander, d. h. prosodische, grammatische, semantische, pragmatische, turnkonstruktionale und interaktionale (sequenziell-responsive) Strukturen?
- Wie zeigen Gesprächsteilnehmer in Produktion und Rezeption an, wann sie Äußerungsstrukturen als abgeschlossene Einheiten betrachten und welche Kriterien sie dafür ansetzen?

3 Einheitenkonzepte

Alle spezifisch für die analytische Gliederung des Sprechens entwickelten Einheitenkonzepte streben eine Segmentierung von Interaktionen in Einheiten der ‚mittleren Ebene‘ an. Anders als in diesem Artikel (vgl. Abschnitt 1) werden Gattungen und Sequenzen oft nicht als systematische Organisationsebene berücksichtigt. Stattdessen wird als „nächstniedrige Einheit unterhalb des Gesprächs“ (Fiehler u. a. 2004, 200) der *Turn* (auf Deutsch meist: (Gesprächs-/Rede-)Beitrag (zur Begriffsgeschichte vgl. Rath 2001)) bzw. der „Gesprächsschritt“ (Brinker/Sager 2010; Henne/Rehbock 1995) angesehen – eine mit Rederecht gemachte Äußerung eines Sprechers, die durch vorangehenden und folgenden Rederechtswechsel oder durch den Beginn oder das Ende des Gesprächs begrenzt wird. Äußerungen, für die kein Rederecht beansprucht wird, z. B. Rückmeldesignale, kommt kein Turn-Status zu (vgl. Schegloff 1981). Entsprechend sind Sprecherwechsel ohne Rederechtswechsel möglich. Dabei ist die Grenzziehung zwischen einem bloßen Rückmeldesignal und einem eigenständigen minimalen Turn nicht immer eindeutig (vgl. 4.2); außerdem können Turns kollaborativ entstehen (vgl. 4.3). Die gesprächsanalytische Kategorie des Turns hat sich auch in linguistischen Darstellungen wie z. B. der IDS-Grammatik etabliert, wo ihr als zusätzliche Eigenschaft propositionaler und illokutiver Gehalt zugeschrieben wird (vgl. Zifonun u. a. 1997, 469), was im Hinblick auf nicht-propositionswertige Turns wie Begrüßungen und andere Arten von Performativa allerdings problematisch ist.

Es herrscht kein Konsens darüber, wie die untergeordneten Einheiten, aus denen sich ein Turn zusammensetzt, analytisch erfasst werden sollten. Einige Einheitenkonzepte beschränken sich auf eine sprachliche Ebene (z. B. Syntax, Prosodie, Pragmatik). Sie segmentieren Turns entsprechend anhand von Grenzmerkmalen nur auf dieser Ebene in ebenenspezifische Einheiten (vgl. 3.1). Die meisten Ansätze dagegen favorisieren, explizit oder implizit, holistische, ebenenübergreifende Einheitenkonzepte. Sie beziehen bei der Segmentierung Abschlusssignale auf verschiedenen Ebenen ein (vgl. 3.2).

3.1 Ebenenspezifische Einheitenkonzepte

Obwohl der *Satz*, wenn er wie unter 2. vom finiten Verb und dem Vorhandensein obligatorischer Argumente ausgehend definiert wird, für die exhaustive Segmentierung von Interaktionen ungeeignet ist, haben auch Ansätze der Gesprochene-Sprache-Forschung häufig mit diesem als Grundeinheit gearbeitet. Dabei wurden besondere Kriterien für eigentlich nicht satzförmige Einheiten hinzugenommen. Ein Beispiel ist Kindt (1994), dessen „weitgehend grammatikunabhängiger Satzbegriff“ auch isolierte Phrasen, die kaum als Ellipsen rekonstruierbar sind, umfasst. Einen Überblick über entsprechende Konzepte geben Fiehler u. a. (2004, 175 ff.). Die Relevanz syntaktischer Fortsetzungserwartungen und die Orientierung Interagierender an syntaktischen Mustern, auch solchen unterhalb der Satzebene, betont Auer (2000, 2007a) mit seiner Konzeption der „Online-Syntax“.

Die Sprechakttheorie zieht *sprachliche Handlungen* für die Gliederung des Sprechens heran. Die funktionale Segmentierung geschieht hier unabhängig von formalen Einheiten. Auf Interaktionen angewandt hat dies v. a. die Funktionale Pragmatik (vgl. Ehlich 2010a), die auf der Sprechakttheorie aufbauend eine eigene Sprechhandlungstheorie mit sog. *Prozeduren* als kleinsten Handlungseinheiten entwirft. Entsprechend sieht das für diesen Ansatz verwendete Transkriptionssystem HIAT (vgl. Rehbein u. a. 2004) eine Segmentierung in auf Prozeduren, aber auch auf formale Merkmale bezogene Einheiten vor (vgl. 3.2), denen dann im Analyseprozess Handlungen zugewiesen werden (vgl. Ehlich 2010b). Traditionelle Sprechakttheorie wird dagegen nur vereinzelt auf Interaktionen angewandt (z. B. Hagemann/Rolf 2001; Staffeldt 2014). Problematisch ist dabei die semantische Bedingung für Sprechakte. Der illokutionäre Akt, d. h., die Handlung, die durch eine Äußerung vollzogen wird, ist fast immer an einen propositionalen Akt gebunden. Dieser wird oft durch Sätze, aber auch durch Ellipsen und Analepsen realisiert. Für Gesprächspartikeln – mit Ausnahme von Responsiven und Interjektionen, denen oft trotz der Nichtpropositionalität illokutiver Gehalt zugesprochen wird – kann Propositionalität nicht angenommen werden. Daher kann einigen mit nicht-propositionalen Einheiten (z. B. Diskursmarkern) ausgeführten Handlungen so nicht Rechnung getragen werden, ganz zu schweigen von non-verbalen Interaktionsbestandteilen.

Darüber hinaus ist es generell unklar, welches Inventar an Handlungen man ansetzen sollte, um eine exhaustive Segmentierung zu ermöglichen: Sprechakte, durch Adjazenzipaare repräsentierte Handlungen oder ein anderes (grob- oder feinkörnigeres) Inventar? Außerdem werden, wie schon Levinson (1981) aus theoretischer Perspektive problematisiert, durch eine Äußerung häufig mehrere Handlungen gleichzeitig ausgeführt (vgl. z. B. #1, S19 „ne PLFANze oder was.“ – zugleich ‚Frage‘ und ‚Verstehensprüfung‘). Unabhängig vom Inventar stellt sich die Frage, anhand welcher Kriterien Handlungen analytisch identifiziert werden können. Einen aktuellen Überblick über die Problematik gibt Levinson (2013).

Eine für das Sprechen spezifische formale Segmentierung ermöglicht die Prosodie: Die Gliederung in aufeinanderfolgende Abschnitte mit einem „als kohäsiv wahrgenommenen Tonhöhenverlauf“ (Selting u. a. 2009, 370) wird anhand des Konzepts der *Intonationsphrase* erfasst. Diese bzw. ihre Vorläufer, die nicht alle anhand authentischer Daten konzipiert worden sind, sind in der Literatur unter verschiedenen anderen Bezeichnungen wie z. B. *tone group* (z. B. Halliday 1967), *intonation-group* (Cruttenden 1997) oder *intonation unit* (z. B. Chafe 1994) bekannt (einen Überblick über die Forschungsgeschichte geben z. B. Barth-Weingarten 2011; Couper-Kuhlen 1986; Selting 1995b und Szczepek Reed 2010). Allen Konzeptualisierungen ist im Wesentlichen gemein, dass von obligatorischen und fakultativen Identifikationskriterien für das Vorhandensein einer prosodischen Einheit ausgegangen wird: Es muss mindestens ein Fokusakzent und eine finale Tonhöhenbewegung vorliegen. Darüber hinaus können Intonationsphrasen durch Tonhöhen sprünge, Tempoveränderungen (v. a. finale Dehnung und beschleunigt gesprochene Silben am Anfang), Rhythmusveränderungen, hörbares Einatmen, Glottisverschluss, Knarrstimme sowie Pausen abgegrenzt sein. Da diese Merkmale optional sind und außerdem zum Teil auch innerhalb von Intonationsphrasen vorkommen können, ist eine eindeutige Segmentierung oft schwierig. Angelehnt an Auer (2010) (vgl. 3.2) hat Barth-Weingarten (2011, 2013) daher vorgeschlagen, nicht von Einheiten mit klaren Grenzen, sondern von Zäsuren auszugehen, die mehr oder weniger stark ausgeprägt sein können, je nach Anzahl der vorhandenen Merkmale, und Einheiten als Epiphänomene zu verstehen. Für englischsprachige Telefongespräche zeigt sich, dass Sprecherwechsel an Stellen mit starken Zäsuren bzw. einem ganz bestimmten Bündel prosodischer und phonetischer Merkmale stattfinden. Die Stärke von Hörerrückmeldungen richtet sich nach der Stärke der Zäsur (vgl. Barth-Weingarten 2013), was ein emischer Beleg für das Verständnis von Einheiten bzw. Zäsuren aus Teilnehmersicht sei. Da es zu anderen Sprachen und vor allem zu anderen Settings mit anderen pragmatischen Voraussetzungen noch keine Untersuchungen gibt, soll anhand des folgenden Beispiels angedeutet werden, wie sich die Orientierung der Interagierenden an der Stärke der prosodischen Zäsuren zeigt. Das Beispiel stammt aus einer Schichtübergabe in einer psychiatrischen Klinik, in der über Patienten gesprochen wird.

#2 FOLK_E_00111_SE_01_T_01, Schichtübergabe, 00:06:56

01 MO ich finde das müsste man zeitlich EINSchränken;=
 02 =<<all> der war kam> gestern WIEder,
 03 °h der war KNACke floride psyCHotisch;=
 04 =WAHNhaft;=
 05 =als er REINKam.
 06 MA [hm:.]
 07 MO [°h] der hat die AUgen zugemacht,=
 08 =<<all> als es geREGnet hat,=
 09 =er hat gesagt er> DARF sie jetzt grad nich AUfmachen?
 10 (0.2)
 11 MO j[a,]
 12 MC [oH]A.

Alle bis auf zwei Zäsuren in MOs Erzählung sind sehr schwach. Durch Beschleunigungen nach einer finalen Tonhöhenbewegung (S01-02), schnelles Sprechen über einen intonatorischen Abschlusspunkt hinweg (S08-09, *rush through*, vgl. z. B. Schegloff 1987, 106 f.; Selting 1995b, 92) und durch *latching* (vgl. z. B. Du Bois u. a. 1993, 59), das Fehlen von (Mikro-)Pausen zwischen den (potenziellen) Intonationsphrasen (S01-02, 03-05, 07-09), werden Zäsuren minimiert. Zwei Zäsuren (S03 und 07) sind relativiert durch hörbares Einatmen, das zwar zur Zäsurenstärke beiträgt, aber zugleich die Absicht anzeigt, noch weiterzusprechen. Nur an der zweiten dieser beiden Stellen (S07), an der wir vorausgehend eine deutlich terminale, tief fallende Grenzintonation feststellen können, gibt MA in Überlappung mit dem Einatmen eine leise, nicht akzentuierte Rückmeldung („hm:“, S06). Im ersten Fall (S03) dagegen geht ein nur leicht steigender Tonhöhenverlauf voraus. Eine stärkere Zäsur als zwischen S05 und S07 findet sich in S09. Sie ist markiert u. a. durch den stark steigenden finalen Tonhöhenverlauf und die anschließende Pause (S10). Hier erfolgt eine stärkere Hörerrückmeldung: kein reines Fortsetzungssignal (vgl. 4.2), sondern die potenziell turnwertige Partikel „oHA.“ (S12). Deren Relevanz auch für den Sprecher wird durch dessen Vergewisserungssignal („ja“, S11) deutlich, das in teilweiser Überlappung mit dieser geäußert wird. Die Qualität der Rückmeldung richtet sich dabei natürlich grundsätzlich nach den semantisch-pragmatischen Merkmalen der vorausgehenden Einheit, das prosodische Format, das im Zäsurierungsansatz gezielt isoliert betrachtet wird, unterstützt diese.

Am Konzept der Intonationsphrase ist von Fiehler u. a. (2004, 193) die Formorientierung kritisiert worden, da diese zu einer Aufspaltung von funktionalen Einheiten führe. Aber auch die Behandlung von Intonationsphrasen *als* funktionale Einheiten, die aus Chafes (1994) Gleichsetzung der *intonation unit* mit einer kognitiv relevanten Verarbeitungseinheit (einer dem Arbeitsgedächtnis entsprechenden Zeitspanne, innerhalb der nur eine bestimmte Menge an Information verarbeitet werden kann) resultiert, ist problematisiert worden (vgl. Auer 2010). Andererseits ist die Formorientierung positiv gewertet worden. Sie ermöglicht eine objektivierbare Segmentierung als eine handlungs- oder TCU-basierte Gliederung. Außerdem erlaubt sie es, ausge-

hend von der genauen Repräsentation der Form, Form-Funktions-Zusammenhänge zu identifizieren und diesbezügliche Generalisierungen zu prüfen. Das Transkriptionssystem GAT2 setzt daher die Intonationsphrase als Segmentierungseinheit an (Selting u. a. 2009). Die prosodische Segmentierung verdeutlicht, an welchen formalen Merkmalen neben den syntaktischen sich Interagierende orientieren: Intonationsphrasengrenzen müssen nicht mit syntaktischen Grenzen zusammenfallen. Nicht selten verteilen sich empirisch gesehen Sätze auf mehrere Intonationsphrasen, was durch die Verteilung der Fokusakzente eine informationsstrukturelle Portionierung ermöglicht:

```
aus #1  28  UD   (0.38) nee dis dis waCHOLdertee,
        29      un Löwenzahntee,
        30      dis soll wohl irgendwie gut für_n Ischiasnerv sein,
```

Umgekehrt können mehrere (Teil-)Sätze als eine Einheit präsentiert werden, wenn syntaktische Grenzen nicht prosodisch markiert werden (vgl. #3, S01).

```
#3 FOLK_E_00066_SE_01_T_02, Gartengespräch, 00:53:11
01  UD      un KRATZt_s schon wenn de brEmsch.
02          (0.25) [n KRATZgeräu][sch.]
03  JO          [na ja pff:  ]
04          [ ne]e.
```

Manchmal ist eine rein prosodische Abgrenzung zwischen solchen Fällen einer Intonationsphrase mit zwei Akzenten und dem Anschluss einer neuen Intonationsphrase durch *rush through* wie in #2 (S08-09) schwierig: Wenn zwei (potenzielle) Fokusakzente vorliegen und dazwischen nur schwach ausgeprägte prosodische Zäsurierungsmerkmale zu finden sind, dann ist die Entscheidung zwischen der Annahme von ein oder zwei Intonationsphrasen willkürlich, zumal es auch Intonationsphrasen mit mehr als einem Fokusakzent gibt (vgl. z. B. #1, S03). Hinzu kommt, dass die Konzeptualisierung des *rush through* syntaktische (und pragmatische) Grenzen als potenzielle prosodische Grenzen voraussetzt, obwohl die Syntax für die prosodische Ebene definitorisch nicht als einheitenbildend vorgesehen ist (vgl. aber Walker 2010 zur tatsächlichen prosodischen Abgeschlossenheit der ersten Intonationsphrase beim *rush through*).

3.2 Ebenenübergreifende Einheitenkonzepte

Mit ihrem für die Konversationsanalyse grundlegenden Aufsatz zum Turn-Taking etablierten Sacks/Schegloff/Jefferson (1974) die *Turnkonstruktionseinheit* (*turn-constructio-nal unit*, kurz: TCU) als Grundbaustein für Turns. Der mögliche Umfang dieser wird zunächst anhand syntaktischer Einheiten (Sätze, Teilsätze, Phrasen, Wörter) definiert.

Jede TCU endet mit einem möglichen Abschlusspunkt bzw. einer übergaberelevanten Stelle (*transition relevance place*, kurz: TRP). Die nachfolgende Forschung betont neben der Syntax stärker auch die Rolle der Prosodie, der Handlungsebene und auch nonverbaler Aspekte (z. B. Blickverhalten und Gesten, vgl. 4.4 und 4.5) für die Identifikation von TRPs (vgl. z. B. Ford/Thompson 1996; Ford/Fox/Thompson 1996; Schegloff 1996, 2007; Selting 1995b, 2000; einen aktuellen Überblick gibt Clayman 2013). Die TCU als potenziell selbständiger Turn kann nur im Sprechen gebildet werden und ist so keine allein durchs Sprachsystem bzw. kontextfrei beschreibbare Einheit.

Da für die Bestimmung von TRPs mögliche Abschlusspunkte auf verschiedenen Ebenen herangezogen werden, stellt sich in Fällen, in denen diese Grenzschnale auf den verschiedenen Ebenen nicht zusammenfallen, die Frage, ob auch Einheiten ohne TRP, aber mit einem funktional relevanten Abschluss, eine TCU bilden können. Zum Beispiel lässt sich diskutieren, wie viele TCUs in den Segmenten 28-30 von #1 (hier mit Fortsetzung wiedergegeben) vorliegen:

#1a FOLK_E_00066_SE_01_T_01, Gartengespräch, 00:12:46

27 PA ah.
 28 (0.38) **nee dis dis waCHOLderte**,
 29 **un Löwenzahnte**,
 30 **dis soll wohl irgendwie gut für_n Ischiasnerv sein**,
 31 °hhh u:nd
 32 (0.68) n hab i noch frühmesner BAdeöl drauf geschmiert,
 33 weil da au noch stand man könnt_s mit laVENdelgeist oder
 ROSmaringeist einschmieren,
 34 (.)°hhh dis hatt ich nich zur HAND,
 35 n da hab ich frühmesner BAdeöl genommen-
 36 weil da beides DRIN is,
 37 lavEndel und ROSmarin;
 38 °hhh im moment is es auf jeden fall BESser;
 39 sa_mer ma SO.

Auf die Antwortpartikel *nee* (S28) folgen ohne prosodische Grenze zwei koordinierte linksversetzte Nominalphrasen („dis dis waCHOLderte, un Löwenzahnte“), die sich auf zwei Intonationsphrasen verteilen, sowie ein V2-Satz („dis soll wohl irgendwie gut für_n Ischiasnerv sein“), ebenfalls in einer separaten Intonationsphrase. Die linksversetzten NPs weisen zwar jeweils eine prosodische Grenze auf, sie können auch als separate Referenz-Handlungen angesehen werden (vgl. Fiehler u. a. 2004) und syntaktisch ist zumindest das Ende einer Phrase erreicht. Diese Grenzen stellen jedoch aus pragmatischen Gründen keine übergaberelevanten Stellen dar. Obwohl eine isolierte NP durchaus einen Turn konstituieren kann, z. B. als Antwort, reicht sie im gegebenen Kontext nicht aus, da der Sprecher durch sie hier nach Beendigung einer Nebensequenz (#1, S09-25) zuvor thematisierte Referenten refokussiert, so dass diese Referenz eine Prädikation projiziert. Das folgende Syntagma endet mit einer möglichen Satzgrenze (rechte Satzklammer), stellt eine vollständige Proposition dar

und weist auch eine prosodische Grenze auf. Allerdings liegt ein leicht steigender finaler Tonhöhenverlauf vor, was im Deutschen als „progre dient“ interpretiert wird (vgl. Auer/Selting 2001, 1124; s. aber Szczepek Reed 2004 für abweichende Befunde zum Englischen). Deshalb ist auch hier fraglich, ob ein TRP vorliegt. Dass PA seine Wiederaufnahme der Ausführungen aus S01-07 (#1) mit S30 noch nicht beendet hat, geht auch aus der Fortsetzung ab S31 (#1a) hervor. Hier wird bereits gegebene Information rekontextualisiert. Auch die übergreifende Gesprächsstruktur hat also mit Einfluss darauf, dass in S30 kein eindeutiger TRP vorliegt.

Mit der Problematik, dass Kriterien für das Ende einer TCU divergieren können, ist in der Forschung auf zweierlei Weise umgegangen worden: 1) Um eine exhaustive Segmentierung zu ermöglichen, werden TCUs ohne TRP angenommen. Sie liegen immer dann vor, wenn auf einer Ebene ein Abschlusspunkt fehlt; ein TRP liegt nur vor, wenn ein syntaktischer, prosodischer und semantisch-pragmatischer Abschluss zugleich vorliegt. Es können zudem auch, wie in #1a, sequenzstrukturelle Gründe das Entstehen eines TRP am TCU-Ende verhindern (vgl. Selting 2000); für TCUs ohne TRP wird z. T. eine neue Kategorie eingeführt (vgl. z. B. die *turn constructional phrase* (TCP) bei Szczepek Reed 2010). 2) Es wird keine exhaustive Segmentierung angestrebt. Stattdessen wird die TCU als Epiphänomen verstanden (vgl. Ford/Fox/Thompson 1996) und betont, dass der analytische Fokus darauf liegen müsse, an welchen möglichen Abschlusspunkten auf verschiedenen Ebenen neben klaren TRPs sich die Interagierenden bei der Konstitution von Turns orientieren. Die Frage nach der Segmentierung in untergeordnete Einheiten wird als „not ultimately the question we need to answer“ (Ford/Fox/Thompson 1996, 431) betrachtet. Dies entspricht auch dem Status des Turns als konstitutionstheoretisch relevante Einheit in der konversationsanalytischen Forschung, der trotz der zahlreichen Untersuchungen zu den ihm untergeordneten Einheiten besteht. Er zeigt sich auch in der Segmentierung von Transkripten nach Jefferson (1983, 2004), die nach Turns, nicht nach TCUs, erfolgt.

Auer (2010) verzichtet dagegen auf die Verwendung der Einheit TCU. Er geht nicht von Einheiten aus, sondern von *Zäsuren* auf prosodischer, syntaktischer und pragmatischer Ebene. Gibt es auf allen Ebenen an derselben Stelle eine Zäsur, so liegt ein besonders salienter Gestaltschluss vor. Gibt es auf nur einer Ebene einen Abschlusspunkt, ist die Zäsurierung nur schwach. Sprecherwechsel finden in der Regel an Stellen statt, an denen Abschlusspunkte auf allen Ebenen vorliegen. Empirisch zeigt sich, dass pragmatische Abschlüsse immer prosodische Abschlüsse voraussetzen, die meist syntaktische Abschlüsse voraussetzen. Letztere sind am zahlreichsten, aber für sich allein auch am wenigsten vorhersagekräftig für mögliche Sprecherwechsel, während der pragmatische Abschluss entscheidend ist (vgl. fürs Englische Ford/Thompson 1996). Der Zäsurierungsansatz ist also hilfreich für die Analyse der Handlungsstruktur in der Interaktion. Er vermeidet weiterhin ein Problem, das sich als Artefakt aus der Annahme der Einheitenkategorie TCU erst ergibt: Wenn nicht auf allen Ebenen eine Zäsur vorliegt, muss nicht entschieden werden, ob ein TRP und

damit eine TCU vorliegt, sondern die Beschreibung der vorliegenden Zäsuren auf den verschiedenen Ebenen bildet für sich die Basis für weitere Analysen.

Im Zusammenhang mit der Vorläufigkeit der Abgeschlossenheit von TCUs werden häufig Erweiterungen nach TRPs thematisiert. Schegloff (1996) schlägt eine dichotome Unterscheidung in Fortsetzungen einer TCU, *increments*, die selbst keine TCU bilden, da sie semantisch und grammatisch unselbständig sind und keine eigene Handlung konstituieren, und Fortsetzungen eines Turns durch Hinzufügen einer neuen TCU vor, geht aber auch auf Erweiterungen ein, die in keine dieser Kategorien passen (z. B. Vergewisserungssignale und Anredeformen). Im Gegensatz zur ursprünglichen Definition nach Sacks u. a. (1974) können Turns dieser Konzeption zufolge also nicht nur aus TCUs bestehen, sondern aus zwei weiteren Einheitenkategorien. *Increments* sind als lineare grammatische Erweiterungen eines vorläufig abgeschlossenen Syntagmas definiert. Ohne dass dies explizit Teil von Schegloffs Definition ist, folgen diese in allen diskutierten Beispielen zudem immer nach einer prosodischen Grenze, da sie sonst im Englischen, das keine rechte Satzklammer kennt, gar nicht als nachträgliche Erweiterungen erkennbar wären (vgl. Auer 2006, 2007b). Ford/Fox/Thompson (2002) vertreten eine breitere Auffassung von *increments* („nonmain-clause continuation after a possible point of turn completion“, Ford/Fox/Thompson 2002, 16); sie gehen im Gegensatz zu Schegloff davon aus, dass diese durchaus eigenständige Handlungen konstituieren können, und unterscheiden entsprechend zwei Typen: 1) *Extensions* „can be heard as syntactically and semantically coherent with what has come before“; diese stellen also keine syntaktisch unabhängige Einheit dar, sie setzen außerdem die vorausgehende Handlung fort. 2) *Free constituents* sind nicht als Konstituente des vorausgehenden Syntagmas interpretierbar und stellen eine neue (die vorherige bewertende oder kommentierende) Handlung dar. Nach Schegloff (1996) wären dies keine *increments*, sondern neue TCUs. Auer (2007b, 651) hebt hervor, dass sich eine dichotome Unterscheidung in Turn-Fortsetzungen und TCU-interne Fortsetzungen nicht durchhalten lasse, sondern eine graduelle Sicht sinnvoller sei, zumal die Korrelation mit dem unterschiedlichen Handlungsstatus keine vollständige sei – auch viele syntaktisch unabhängige Fortsetzungen stellten retrospektive Handlungsfortsetzungen dar und umgekehrt seien auch manchmal syntaktische Fortsetzungen neue Handlungen (z. B. bei kollaborativen Turns).

Darüber hinaus sind die Realisierungsformen von *increments* einzelsprachspezifisch unterschiedlich. Auer (1991, 2006, 2007b) hat gezeigt, dass im Deutschen als abschließendes Strukturelement vor allem die rechte Satzklammer zentral ist. Viele im Nachfeld auftretende Konstituenten sind syntaktisch abhängig (regiert, nachträglich insertierend oder substituierend). Eine umfassende Systematik von Expansionen muss jedoch auch unabhängige(re) Fortsetzungen (wie koordinierte und appositive Strukturen) erfassen. Auch für einige andere Sprachen sind mittlerweile auf deren grammatische Strukturen zugeschnittene Konzeptionen von *increments* vorgenommen worden (vgl. u. a. Couper-Kuhlen/Ono 2007; Luke/Thompson/Ono 2012; Ono/Couper-Kuhlen 2007). Vor dem Hintergrund dieser Untersuchungen und der Schwierig-

rigkeit, eine sinnvolle (über)einzel sprachliche Definition zu finden, stellt sich die Frage, ob eine Annahme dieser ebenenübergreifenden Einheitenkategorie sinnvoll ist oder ob die entsprechenden Phänomene nicht auch anhand des Zusammenspiels verschiedener ebenenspezifischer Einheitengrenzen beschrieben werden können. Der Vorteil einer solchen Konzeption läge darin, dass für Äußerungsbestandteile, deren Einheitenstatus in Bezug auf Konzepte wie Turn oder TCU unklar ist, parameterbezogen zu analysieren ist, in welchen Hinsichten und aufgrund welcher Merkmale sie als Einheiten verstanden und behandelt werden (können) und wieso sie daher von Gesprächsteilnehmern unterschiedlich hinsichtlich ihrer Abgeschlossenheit gewertet werden können.

Zur Einheitendiskussion innerhalb der gesprächsanalytischen und interaktional-linguistischen Forschung lässt sich zusammenfassend sagen, dass eine Konzeptualisierung, die TCUs als Epiphänomene betrachtet oder sie im Zäsurierungskonzept aufgehen lässt, die problematische Unterscheidung zwischen der Expansion einer TCU und dem Anschluss einer neuen TCU überflüssig macht. Damit stellt sich auch die Frage, ob bzw. wofür die Kategorie TCU selbst relevant ist. Konzipiert ist sie, anders als die meisten im Folgenden zu diskutierenden Einheitenkategorien nicht aus Analytikersicht, sondern als Teilnehmerkategorie: An einem TRP (oder einer Zäsur auf allen Ebenen) zeigen die Interagierenden einander die Möglichkeit oder Erwünschtheit des Sprecherwechsels an. D. h., in solchen klaren, prototypischen Fällen lassen sich Turns bzw. TCUs als Einheiten beschreiben, an denen sich die Gesprächsteilnehmer erkennbar orientieren. In Fällen, in denen kein Sprecherwechsel stattfindet und auch keine Rückmeldesignale o. Ä. platziert werden, aber dennoch ein TRP angenommen wird, wird auch die TCU notwendigerweise oft zur Analytikerkategorie, die auf linguistisches und pragmatisches Wissen rekurriert, so dass es notwendig ist, die für die Einheit bzw. Grenze angesetzten Kriterien zu explizieren und zu reflektieren. Darüber hinaus sollte diskutiert werden, ob und inwiefern man die Turntakingrelevanz ganz von der Frage der Einheitenkonstitution entkoppeln kann: Ein TRP sollte immer eine für die Teilnehmer interaktionsorganisatorisch relevante Stelle sein, rein linguistische, insbesondere grammatische Einheiten bieten allein durch ihre analytische Identifikation keine Erklärung dafür, welche Handlung sie im gegebenen Kontext ausführen (vgl. Ford/Fox/Thompson 2013, die aus diesem Grund gar eine neue deskriptive, handlungsbasierte Metasprache fordern). Dennoch kann auch eine differenzierende Untersuchung der tatsächlichen Beschaffenheit von ursprünglich schriftsprachlich bzw. systemlinguistisch definierten Einheiten im Sprechen sinnvoll sein, auch für die Bestimmung ihrer Rolle bei der Turnkonstitution (zu aktuellen Beiträgen zu dieser Diskussion Szczepek Reed/Raymond 2013).

Weitere Einheitenkonzepte, die Merkmale auf verschiedenen Ebenen einbeziehen sind u. a. die Äußerungseinheit nach Rath (1990), die Äußerung nach Rehbein (1995) und die funktionale Einheit nach Fiehler u. a. (2004).

Unter *funktionalen Einheiten* verstehen Fiehler u. a. (2004, 204) „die kleinsten Bestandteile des Beitrags, denen eine [...] (separate) Funktion [im und für den Kom-

munikationsprozess] zugeschrieben werden kann.“ In Abgrenzung von der zunächst strukturell orientierten TCU wird eine Segmentierung von Turns allein auf der Handlungsebene angestrebt, wobei Syntax und Prosodie bei der Identifikation der Einheiten explizit unterstützend herangezogen werden (vgl. Fiehler u. a. 2004, 203). Die Liste an möglichen kommunikativen Aufgaben, die einer funktionalen Einheit oder einer Kombination von funktionalen Einheiten zugewiesen werden kann, ist offen, z. B. „zu etwas Stellung nehmen, eine Bewertung äußern, fluchen, [...] eine Redewiedergabe ankündigen, eine Verstehensanweisung für eine andere Äußerung geben, [...] etwas verneinen, etwas ablehnen [...]“ (Fiehler u. a. 2004, 205). Es werden drei Arten von funktionalen Einheiten nach dem Grad ihrer Selbständigkeit unterschieden: 1) Projizierende funktionale Einheiten, z. B. Diskursmarker und redeeinleitende Phrasen (in der Terminologie von Fiehler u. a. 2004: ‚Operatoren‘) oder linksversetzte Konstituenten, können nicht allein auftreten, sondern lassen eine bestimmte Art von Folgeeinheit erwarten. Im Gegensatz zum konversationsanalytischen Ansatz, nach dem in solchen Fällen keine TCU vorliegt, kommt ihnen also hier aufgrund der isolierbaren Handlung, die mit ihnen ausgeführt wird, Einheitenstatus zu. 2) Selbstständige funktionale Einheiten haben keine projektive Kraft und können allein stehen oder frei mit folgenden Einheiten kombiniert werden. 3) Assoziierte funktionale Einheiten, z. B. Nebensätze, können syntaktisch nicht allein stehen, weil sie von einer selbstständigen Trägereinheit abhängen. Wie bei Ford/Thompson (1996) liegen immer, wenn eine abgeschlossene Handlung identifiziert wird, auch syntaktische und prosodische Abschlusspunkte vor. Solche treten aber auch einheitenintern auf. Problematisch ist an der Auffassung von Fiehler u. a. (2004), dass immer eine fein- oder grobkörnigere Segmentierung möglich ist, da es kein geschlossenes Inventar von möglichen Handlungen gibt. Insbesondere kommt Gesprächspartikeln hier mal Handlungs- und Einheitenstatus zu, ein andermal sind sie aber Teil größerer funktionaler Einheiten, ohne dass für die Unterscheidung Kriterien bereitgestellt werden.

Rath (1976, 1985, 1990, 1997) dagegen behandelt alle Gesprächspartikeln auf mechanische Weise gleich. Er bezeichnet die Grundbausteine von Turns als *Äußerungseinheiten* und definiert sie als Einheiten zwischen Grenz- und Gliederungssignalen. Zur Abgrenzung werden prosodische Mittel wie (gefüllte) Pausen und Intonationskonturen sowie lexikalische Mittel wie „Gliederungspartikeln“ (z. B. *ja*, *nicht (wahr)*, *und (so)*) und „syntaktische Formeln“ (z. B. *ich meine*, *verstehst du*, *ich würde sagen*) herangezogen (vgl. Rath 1985, 1658). Die Satzgrenze ist kein Kriterium: Wenn an einer solchen keine Gliederungssignale vorliegen, umfasst die Äußerungseinheit mehr als einen Satz; umgekehrt kann ein Satz durch Gliederungssignale in mehrere Äußerungseinheiten aufgespalten werden. Syntaktische und interaktive Gliederung werden damit als getrennte Ebenen angesehen: „Sprecher und Hörer ziehen innerhalb der ablaufenden Syntax eigene, interaktive Grenzen [...]“ (Rath 1990, 209). Es wird hervorgehoben, dass alle an der Interaktion Beteiligten durch das Äußern von (sprecher- oder hörerseitigen) Gliederungssignalen zur Konstitution von Einheiten beitragen, diese seien „Ergebnisse interaktiver und kognitiver Prozesse“ (Rath 1990,

214). Die genaue Definition und Rolle der prosodischen Gliederungssignale bleibt im Gegensatz zu den unter 3.1 besprochenen Intonationsphrasenansätzen unklar, weil diese Signale in der Argumentation nur dann herangezogen werden, wenn die vorausgehende Annahme einer Äußerungseinheitengrenze gestützt werden soll. Prosodische Zäsuren innerhalb der Äußerungseinheit werden nicht diskutiert.

Beispiel #4 soll zeigen, zu welchen potenziell problematischen Einheiten dieser Ansatz führt.

#4 FOLK_E_00066_SE_01_T_02, Gartengespräch, 00:53:29

```

01 UD °h es GIBT sogar:r-
02      ? ä:hm brEmsklötze die sin hÄrter als die FELge so;
03      (.) das hab ich schon ma gesSEHN,
04      °h
05 AL hm.
06 UD da hat sich einer komplEtt die FELge durchgeschrubbt,
07      weil er zu ar[ge BREMS][klötze] [drauf hatte.           ]
08 JO      [ha JA:~? ]
09                        [dis is]
10                        [<<all> wenn de die falsch]en
      blems BREMSklötze nimmst-=
11      =und die Alufelgen [hast;~>]
12 AL      [ ja g]eNAU.

```

Sowohl das „ä:hm“ als auch das „so“ in #4, S02 wären nach Rath Gliederungssignale, die eine Äußerungseinheit („brEmsklötze die sin hÄrter als die FELge“) abgrenzen. Sie bilden beide mit dieser eine prosodische Einheit, während andere Gliederungssignale wie das „hm.“ in S05 und das „ha JA:~?“ in S08 prosodisch selbständig auftreten. Die Gliederungssignale führen zur Verteilung von Sätzen auf mehrere Äußerungseinheiten (S01-02, „es GIBT sogar:r“ und „brEmsklötze die ...“), aber auch dazu, dass Kombinationen aus Haupt- und Nebensatz als eine Äußerungseinheit zu betrachten sind, weil keine Pausen oder Gliederungssignale dazwischentreten (S06-07), obwohl sie prosodisch voneinander abgegrenzt sind, da sie zwei Intonationsphrasen mit eigenen Fokusakzenten und Grenztönen bilden.

Die Annahme der Unabhängigkeit der kommunikativen Gliederung von der syntaktischen und prosodischen ist aus verschiedenen Gründen problematisch. Erstens wird der Orientierung der Interaktionsteilnehmer an syntaktischen und prosodischen Strukturen damit nicht genügend Rechnung getragen (Selting 1995a). Zweitens ist der Status der Grenzsignale selbst nicht klar: Sie sind keine Äußerungseinheiten, da sie dazwischen („innerhalb“ der Grenzen, Rath 1990, 202; in den „Gelenkstellen“, Rath 1986, 1658) auftreten; andererseits können sie, da sie als die Einheiten erst retrospektiv konstituierend angesehen werden, auch als Teil dieser aufgefasst werden, und in Raths (1997) jüngster Publikation zum Thema werden die Grenzen in Beispielen auch erst nach ihnen notiert. Drittens stellen die als Gliederungssignale bezeichneten Wörter und Phrasen eine heterogene Gruppe dar, so dass die pauschale Behand-

lung als nicht-referenzielle Elemente, die generell nicht auf der syntaktischen und semantischen Ebene operieren, nicht angemessen erscheint: Diskursmarker (*also*, (*ich*) *mein*) als projizierende Vor-Vorfeld-Einheiten sind ebenso wie koordinierende Konjunktionen (*und*, *oder*, *denn*) peripherer Teil der syntaktischen Struktur, Phrasen wie *ich würde sagen* und *ich glaube*, die oft noch als Matrixsätze interpretiert werden können, operieren in diesen Fällen auch auf der semantischen Ebene. Vergewisserungssignale wie *ne* und *weißt du* und Rückmeldesignale bzw. Responsive wie *ja* und *hm* operieren zwar durchweg auf der diskursorganisatorischen Ebene und sind (syntaktisch) selbständig, aber gerade deshalb sind sie nicht als „Einschnitte“ (Rath 1997, 16) anzusehen, zumal sie, wie Rath selbst betont, auch in Überlappung auftreten können (vgl. „ha JA:?“ in #4, S08 und „ja geNAU.“ in S12) und dann nicht linear ‚zwischen‘ den Äußerungseinheiten auftreten.

Auch Schwitalla (2012) segmentiert Turns in Äußerungseinheiten, bezieht aber zugleich aktuelle Erkenntnisse zur Gliederung durch lexikalische, syntaktische und prosodische Mittel ein, so dass sein Vorgehen sich von dem von Rath unterscheidet und im Ergebnis die meisten seiner Äußerungseinheiten Intonationsphrasen entsprechen (vgl. Schwitalla 2012, 95).

Während *Äußerung* häufig – und so auch im vorliegenden Artikel – nicht terminologisch gebraucht wird, verwendet Rehbein (1995) diese Bezeichnung für die Einheiten, aus denen sich ein Turn zusammensetzt. Er definiert sie als Realisierung mindestens einer sprachlichen Prozedur im Sinne der Funktionalen Pragmatik. Die Segmentierung in Äußerungen wird anhand morpho-syntaktischer und prosodischer Kriterien, aber im ersten Schritt auch explizit „intuitiv“ (Rehbein 2001, 932) vorgenommen und führt meist zu satzwertigen Einheiten, die aus Prozedurenkombinationen konstituiert sind. Sie bildet die Grundlage für die analytische Zuweisung von Sprechhandlungen (vgl. a. 3.1). Laut Fiehler u. a. (2004, 191) entspricht das Ergebnis der Segmentierung in Äußerungen weitgehend dem der Segmentierung in funktionale Einheiten, auch wenn die theoretischen Voraussetzungen sich nicht entsprechen. So kommt z. B. auch Interjektionen, Responsiven und anderen eigenständigen Partikeln sowie freien Themen und anderen isoliert verwendeten Konstituenten Einheitenstatus zu.

Ein mehrdimensionales Einheitenkonzept, das spezifisch für eine einzelne Untersuchung zum kanadischen Englisch entwickelt worden ist, ist die *Talk Unit* nach Halford (1996). Diese ist definiert als „minimal self-contained message unit“ – sie umfasst mindestens eine Proposition – und zugleich als „maximal unit defined by syntax and intonation“ (Halford 1996, 33). Weder prosodische noch syntaktische selbstständige Einheiten dürfen auf mehrere Talk Units aufgespalten werden. Das bedeutet, dass immer die jeweils größere Einheit auf einer Ebene den Umfang der Gesamteinheit bestimmt und die Talk Units häufig relativ groß sind (vgl. Fiehler u. a. 2004, 194), vor allem weil komplexe Sätze, die sich über mehrere prosodische Einheiten (die bei Halford nicht Intonationsphrasen entsprechen, sondern größer sein können) verteilen, als eine Einheit gelten. Die Konzeption ähnelt der der TCU, da davon ausgegangen wird, dass die Talk Units mit übergaberelevanten Stellen enden,

die sich auch hier nur dann ergeben, wenn Grenzen auf mehreren Ebenen zusammenkommen (vgl. Halford 1996, 34). Nicht-Propositionswertiges kann zwar keine Talk Unit bilden, die Möglichkeit der Bildung einer separaten Einheit wird aber zumindest einigen Partikeln, sogar dann, wenn sie nicht als propositionswertig interpretiert werden können, zugestanden (vgl. Halford 1996, 97).

Schließlich soll noch kurz ein nicht aus der Gesprochene-Sprache-Forschung stammendes Einheitenkonzept erwähnt werden, weil es im Zuge einer Differenzierung zwischen Form- und Funktionsebene dem Satz gleichberechtigt zur Seite gestellt und als auch auf die Spontansprache anwendbar eingeführt wird – die *kommunikative Minimaleinheit* (KM) (Zifonun 1987):

Kommunikative Minimaleinheiten sind die kleinsten sprachlichen Einheiten, mit denen sprachliche Handlungen vollzogen werden können. Sie verfügen über ein illokutives Potential und einen propositionalen Gehalt. In gesprochener Sprache weisen kommunikative Minimaleinheiten eine terminale Intonationskontur auf, es sei denn, sie werden mit weiteren kommunikativen Minimaleinheiten koordinativ verknüpft. (Zifonun u. a. 1997, 91)

Die KM werden untergliedert in Vollsätze, Nicht-finit-KM und Ellipsen. Responsive und Interjektionen gelten aufgrund des angenommenen fehlenden propositionalen Gehalts und illokutiven Potenzials nicht als KM, sondern als *interaktive Einheiten* (Zifonun u. a. 1997, 62 ff. und 362 ff.). Da außerdem sprachliche Handlungen sprechakttheoretisch und somit enger gefasst werden als z. B. bei Fiehler u. a. (2004), stellen auch Operatoren und herausgestellte Konstituenten keine KM dar. Ob sie als interaktive Einheiten in Frage kämen oder aufgrund der peripheren syntaktischen Eingebundenheit (über das Vor-Vorfeld) als Teile von KM gelten müssen, wird nicht thematisiert.

Die Diskussion hat gezeigt, dass keines der vorliegenden Einheitenkonzepte eine eindeutige Lösung für jegliche in der gesprochenen Sprache vorzufindende Struktur bringt (s. Abschnitt 4). Aus unserer Sicht scheint klar zu sein, dass ein Ansatz zu präferieren ist, der von den beobachtbaren Aktivitäten der Interaktionsorganisation ausgeht und damit am Einheiten-Verständnis der Gesprächsbeteiligten selbst ansetzt. Dabei haben die formbezogenen Ansätze den Vorteil, dass sie a) weitaus klarere Kriterien für die Definition von Einheitengrenzen liefern, b) erlauben aufzuweisen, dass sich Interaktionsteilnehmer tatsächlich in hohem Maße an formbezogenen Eigenschaften der Turnproduktion orientieren und c) den Beitrag formaler Eigenschaften des Turn-Designs zur Interpretation des Handelns explizit erfassen lassen.

4 Problemfälle

Im Folgenden wollen wir auf gängige Strukturen des Sprechens in der Interaktion eingehen, die allen der in Abschnitt 3 vorgeschlagenen Einheitenkonzeptionen Probleme bereiten. Im Einzelnen geht es um den Status von nichtautonomen Turnkom-

ponenten (4.1), Äußerungen ohne vollen Turnstatus (4.2), kollaborative Turnproduktionen (4.3), nonverbale Produktionen (4.4) und die Konstitution multimodaler Einheiten (4.5).

4.1 Nicht-autonome Turnkomponenten

Die klassische Definition von Turns geht davon aus, dass diese aus *turn constructional units* (TCUs) bestehen, die in einem *transition relevance place* enden (Sacks u. a. 1974; vgl. a. Abschnitt 3.2). Einige Phänomene der gesprochenen Sprache sind in Bezug auf ihren Status als Turnbestandteile besonders problematisch. Diskursmarker (Auer/Günthner 2005; Imo 2012; Schiffrin 1987) und Linksversetzungen sind aus Sicht der Turnkonstruktion keine TCUs, da sie nicht in einem TRP enden und nicht allein einen Turn ausmachen können. Prosodisch sind sie oftmals keine eigenen Intonationsphrasen. Syntaktisch sind sie dagegen autonom, denn sie sind nicht in Kongruenz- und Rektionsbeziehungen integriert, projizieren aber eine Fortsetzung durch nachfolgende sprachliche Strukturen. Somit sind sie lediglich über die topologische, aus syntaktischer Sicht fragwürdige Hilfskonstruktion eines ‚Vor-Vorfelds‘ dem Satzschema zu subsumieren. Diskursmarker machen bspw. häufig auch eigene Intonationsphrasen aus, konstituieren aber keinen TRP:

#5 FOLK_E_00066_SE_01_T_01, Gartengespräch, 00:24:08

01 JO (.) und dann is natürrch KLAR äh-
 02 (0.76) ich MEIN,
 03 PA ? hh°
 04 JO du hast die ANlage,

Einen Diskursmarker und eine Häsitationspartikel (s. 4.4), gefolgt von einem freien Thema sehen wir in #6:

#6 FOLK_E_00003_SE_01_T_01, Prüfungsgespräch, 00:02:08

01 JS also ?
 02 ä:hm
 03 diese diese bottomup und topDOWNprozesse,=
 04 =in bezug auf die artikulaTION?
 05 können sie DAzu vielleicht noch was sAgen?

Alle drei Strukturen sind syntaktisch nicht eingebunden. Die beiden Partikeln in S01 und S02 machen keine eigenen Intonationsphrasen aus, gehören aber auch nicht zur folgenden Intonationsphrase des freien Themas. Keine der drei Strukturen führt zu einem TRP.

Increments (Schegloff 1996; Auer 2006, 2007b) und andere Formen der Expansion (Auer 1991, 1996) haben ebenfalls einen schillernden Status (vgl. a. Abschnitt 3.2). *Increments* sind keine eigenen TCUs, da sie keine eigene Handlung vollziehen.

Manchmal sind sie auch prosodisch mit der Vorgängerstruktur integriert (d. h. ohne Fokusakzent und eigene Intonationskontur) und grammatisch von vorangehenden TCUs abhängig. Andererseits sind sie eigene Phrasen und werden nach einem TRP produziert. In #7, S03 wird nach einem semantisch-pragmatisch eindeutigen TRP ein grammatisch abhängiges *increment* angehängt, das keine neue Handlung vollzieht:

#7 FOLK_E_00055_SE_01_T_07, Studentisches Alltagsgespräch, 01:18:43

```
01  US  des war GANZ blöd;
02      ich hatte normAl immer so_n kleinen DISporahmen
      auf_m konto halt;
03      °h n[ormalerweise] KRIEGT ma des ja nich
      we_ma noch nich SELber verdient,
04  NH      [m_hm,      ]
05  US  aber (0.5) die ham mir des irgendwie ma SO erklÄrt,
```

4.2 Äußerungen ohne vollen Turnstatus

Die Produktion eines vollen Turns geht mit der Beanspruchung des Rederechts (*floor*) und dem lokalen Management des Rederechts am Ende des Turns durch den Turnproduzenten einher (Sacks u. a. 1974: „first speaker selects next speaker“). Für zahlreiche, regelhaft produzierte Äußerungen im Rahmen einer Interaktion gelten diese Bestimmungen aber nicht. Die Rede ist hier von Rückmeldern (Yngve 1970), *continuers* (Fortsetzungssignalen; Schegloff 1981), Interjektionen (Nübling 2005; Reber/Couper-Kuhlen 2010) und anderen *secondary speaker contributions* (Bublitz 1988). Im folgenden Ausschnitt produzieren zuerst NO (S04, 07), dann EL (S14, 16) Rückmeldesignale:

#8 FOLK_E_00039_SE_01_T_01, Pärchengespräch, 00:13:42

```
01  EL  ds jetzt schon wieder EEne woche her;=
02      =da ham_wa uns SO lange druff jefrEut,=
03      =und
04  NO  (.) ja.
05      (2.86)
06  EL  war aber wIrklich_n SCHÖner tach.
07  NO  (.) ja.
08      (3.7)
09  EL  ham sich alle echt (.) Mühe jegeben;
10      (4.17)
11  EL  <<p> hm_hm,>
12      (1.09)
13  NO  nee kannst schon echt FROH sein wenn du so ne freunde
      hast;
14  EL  ja;
15  NO  die sich so KÜMmern;
16  EL  hm_HM,
```

Mit Rückmeldern wird kein Rederecht über die mit ihnen produzierte, meist sehr kurze Äußerung hinaus beansprucht. Das Rederecht verbleibt beim vorangehenden bzw. parallel aktiven primären Sprecher. Noch ist kaum erforscht, welche Erwartungen für die Synchronisation dieser Äußerungen mit der Turnproduktion des primären Sprechers gelten (s. aber Barth-Weingarten 2009, 2013; Stivers 2008). Der Einheitenstatus dieser Äußerungsformate ist außerdem schwierig, da sie zumeist keine syntaktische Struktur aufweisen, sondern nur aus Partikeln bestehen. Sie vollziehen keine deutlich profilierte Handlung – sie zeigen Aufmerksamkeit und wohl auch Verstehen an, können aber nicht als vollwertige Zustimmungen gelten –, sondern fungieren primär als gesprächsorganisatorische *continuer*. In S11 wird ein *continuer* eingesetzt, um den Turnabschluss und den Verzicht auf weiteres Rederecht anzuzeigen. Als Intonationsphrasen sind diese Phänomene allenfalls rudimentär ausgebildet: Sie sind oft einsilbig und weisen nur selten einen Fokusakzent auf, und dies auch nur, wenn man die Akzentuierung einer Silbe bei redupliziertem *hm* so (in Bezug auf welchen Möglichkeitsraum?) verstehen möchte.

4.3 Kollaborative Turns

Kollaborative Turnproduktionen sind in pragmatischer und syntaktisch-projektionsbezogener Hinsicht Einheiten, die aber auf verschiedene Sprecher verteilt sind und sprachlich oftmals auch deutliche Züge autonomer Einheiten tragen. Ein Beispiel dafür ist:

#9 FOLK_E_00003_SE_01_T_01, Prüfungsgespräch, 00:02:08

```
01  JS    können sie DAzu vielleicht noch was sagen?
02      inwieFERN-
03      °h (.) [ähm die
04  DM      [inwiefern die funktionIE]ren,
05  JS    ja,
```

JS projiziert mit „inwieFERN“ (S02), dass sie ihren Turn mit einem indirekten Fragesatz fortsetzen wird. Nach der Produktion des Interrogativpronomens bricht sie jedoch ab und atmet ein (S03), was als Wortsuche bzw. Sprechplanungspause verstanden werden kann. DM ergänzt den von JS projizierten indirekten Fragesatz (S04) und JS bestätigt die Ergänzung als ihrer Redeintention entsprechend (S05). DMs Ergänzung ist zwar eine selbständige Intonationsphrase, syntaktisch und handlungsbezogen ist sie jedoch abhängig vom Turn der vorangehenden Sprecherin. Während faktisch von JS und DM zwei Turns produziert werden, gelten sie den Beteiligten interaktional doch als ein gemeinsamer Turn, der JS als von ihr gemeint zugerechnet wird. DM übernimmt also das Rederecht nicht für sich, sondern stellvertretend für die vorangehende Sprecherin JS, um ihren Turn zu vollenden (Lerner 1991). Während DMs kollaborative Turnfortsetzung in semantischer und pragmatischer Hinsicht nahtlos

an JS anschließt, tut sie dies in syntaktischer Hinsicht nicht: DM setzt den indirekten Fragesatz nach dem Fragepronomen erst fort, nachdem sie dieses wiederholt hat. Kollaborationen bestehen oft nicht aus bruchlosen Erfüllungen der im Vorgängerturn gestifteten syntaktischen Projektionen: Teilwiederholungen, syntaktische Strukturveränderungen und die Produktion von Diskurspartikeln, die z. B. Reformulierung, Ungewissheit, Tentativität oder Zögern anzeigen, sind übliche Merkmale von Ergänzungen (Oloff 2009). Ähnliche Probleme bezüglich des Einheitenstatus können auch diskontinuierliche Turnproduktionen des gleichen Sprechers aufwerfen (vgl. S17, 20-22, 24 in #1 in Abschnitt 2). In Bezug auf die Produktion von TCUs handelt es sich hier entweder um eigenständige TCUs oder aber um *increments*, in Bezug auf den Sprecherwechsel gar um eigenständige Turns, während die semantische, grammatische und pragmatische Struktur des diskontinuierlichen Turns eine integrierte Ganzheit konstituiert.

Ebenfalls unselbständig hinsichtlich ihres syntaktischen und oft auch hinsichtlich ihres Turnstatus, aber im Gegensatz zu Turnergänzungen nicht durch die syntaktische Struktur des Turns des vorangegangenen Sprechers projiziert, sind Fortführungen des Vorgängerturns. In #10 ergänzt der Sprecher UD in S08 eine Aussage aus JOs Erzählung.

#10 FOLK_E_00066_SE_01_T_01, Gartengespräch, 00:24:15

```

01  JO   ich glaub der kriegt im schnitt irgendwie vielleicht so
        hundertzwanzig EUro,
02      wenn er seine ANlage anschleppt,
03      und da dann MISCHT,
04      (0.44)
05  UD   hm_HM;
06      (0.5)
07  JO   u:nd dann ha[t er]
08  UD           [für ] die NACHT.
09      (0.33)
10  JO   für die für den Abend oder so?
```

UDs Fortführung „für die NACHT.“ expliziert den zeitlichen Bezug von JOs vorangegangener Aussage darüber, wie viel ein Bekannter in seiner Tätigkeit als Dee-Jay verdient. Es handelt sich um eine semantische Spezifikation, die die Vorgängeräußerung syntaktisch abhängig, in einer eigenen Intonationsphrase expandiert, ohne eigenes Rederecht zu beanspruchen. Im Gegensatz zu einer Ergänzung handelt es sich aber um eine autonomere, fortführende Struktur: Die Fortführung vervollständigt nicht eine vom vorangehenden Sprecher erzeugte, noch unvollständige Struktur durch die Erfüllung einer mit ihr etablierten Projektion, sondern sie baut auf der syntaktisch vollständigen, keine Projektionen offenlassenden Vorgängerstruktur auf, um eine Spezifikation, evtl. im Dienste einer Verstehensprüfung anzubringen.

4.4 Nonverbale kommunikative Beiträge

Face-to-face-Interaktionen bestehen nicht nur aus sprachlichen Aktivitäten. Zur Interaktionskonstitution tragen ebenso die vokalen und, im Falle der wechselseitigen visuellen Präsenz der Interaktionspartner, auch die weiteren multimodalen Ressourcen wie Blick, Mimik, Gestik, Körperpostur und –bewegung und Handhabung von Objekten bei. Wenn die wissenschaftliche Frage nach den Grundeinheiten des Sprechens in der Interaktion den Bedingungen der Interaktionskonstitution angemessen Rechnung tragen will, dann darf sie sich nicht auf ein künstlich restringiertes, linguistisch definiertes Objekt beschränken, sondern muss erkunden, wie sprachliche und nicht sprachliche Strukturen bei der kommunikativen bzw. interaktiven Einheitenkonstitution zusammenspielen. Diese Erweiterung der Perspektive über das Linguistische hinaus betrifft wenigstens drei Dimensionen:

- die Rolle von Vokalisierungen,
- nicht-vokale Handlungen, die auf die Struktur der verbalen Interaktion bezogen sind,
- die Koordination von verbalen und anderen Ressourcen in der Turnproduktion (s. 4.5).

Die Abgrenzung von Vokalisierungen und verbalen Aktivitäten ist teilweise willkürlich (Reber/Couper-Kuhlen 2010). Während eine zunehmende Anzahl von Diskurspartikeln wie Rezeptions- und Häsitations-signale oder Interjektionen Eingang in Lexika findet und somit als Lexeme und Bestandteile der Sprache kanonisiert werden (z. B. *hm*, *aha*, *äh*), kommen andere Vokalisierungen nach wie vor nicht in Lexika vor und werden auch von den meisten Linguisten nicht als Lexeme akzeptiert (z. B. Zungeschnalzen, *pf*, bilabial plosives Ausatmen). Diese Phänomene sind nicht weniger konventionell und funktional spezifisch als die kanonisierten. Dass sie nicht als Bestandteil der Sprache betrachtet werden, scheint in anderen Merkmalen begründet. Z. B. weichen phonetische Strukturen vom phonetisch-phonematischen System der Einzelsprache ab, phonologische Repräsentationen sind weniger stabil und restringiert als bei Wörtern üblich und die silbische Struktur ist abweichend (z. B. kein (Voll-)Vokal). Diese Eigenschaften sind noch deutlicher ausgeprägt bei anderen vokalen Phänomenen wie Lachen, Weinen oder Seufzen, denen gar keine definitive lautliche Typ-Repräsentation zuzuordnen ist. Natürlich sind solche Phänomene nicht Teil grammatischer Strukturen. Aber sie sind oft selbst turnwertig und Bestandteile von Turns und sie gehören manchmal auch zu Intonationsphrasen.

Nichtvokale kommunikative Beiträge sind auch in verbal dominierten Interaktionen oft mitkonstitutiv für die Handlungsorganisation. Sowohl *first pair parts* von Nachbarschaftspaaren (wie Fragen, Angebote oder Aufforderungen) als auch *second pair parts* (wie Antworten und Befolgungen von Aufforderungen) können nonverbal konstituiert werden (Clark 2012; Rauniomaa/Keisanen 2012). Rederechtsbeanspruchende Gesten können parallel zur Turnproduktion eines anderen Teilnehmers pro-

duziert werden (Schmitt 2005). Ein Interaktionsbeteiligter kann anzeigen, dass er das Rederecht weiterhin beansprucht, obwohl er aktuell aufgrund von Rederechtskonkurrenz schweigt, indem er seine Gestik (z. B. erhobene Hand, Zeigegeste) und eine vorgelehnte Körperpositur einfriert (Oloff 2013). Eine exhaustive Einheitenanalyse des Sprechens, die nicht ein künstlich disziplinär gereinigtes Objekt verbaler Produktion, sondern die faktische Praxis multimodaler Interaktionskonstitution analysieren will, muss reflektieren, wie solche nicht verbal konstituierten Phänomene selbst (Handlungs-)Einheiten konstituieren bzw. an der Konstitution verbaler Einheiten teilhaben. Dabei sind jedoch die besonderen semiotischen und interaktiven Qualitäten v. a. der nonvokalen Ressourcen zu beachten. Es hat u. E. keinen Sinn, hier etwa von ‚nonverbalen Turns‘ zu sprechen, da z. B. gestische Turnbeanspruchungen oder kinesische Erfüllungen einer konditionellen Relevanz nicht nach den für Turns distinktiven Prinzipien organisiert sind. Die Orientierung an ‚one party at a time‘ und die dementsprechende Wahrnehmung von gleichzeitigen Aktivitäten als (zu vermeidenden) Überlappungen sowie die Zusammensetzung der Aktivitäten aus *turn constructional units*, die jeweils in einem *transition relevance place* enden, gilt für nonverbale Aktivitäten nicht bzw. nicht in gleicher Weise wie für verbale Turns (Schmitt 2005; Deppermann 2013a). Vielmehr scheinen uns solche Phänomene darauf zu verweisen, dass es notwendig ist, zwischen Handlungen und Turns konzeptuell klar zu unterscheiden.

4.5 Multimodale Einheiten

In der multimodalen Interaktion werden kommunikative Einheiten als komplexe multimodale Gestalten konstruiert: Verschiedene multimodale Ressourcen wie Sprechen, Blick, Mimik, Gestik, Körperbewegung und die Handhabung von Objekten werden miteinander gestalthaft koordiniert (vgl. hierzu auch Fricke und Bressen in diesem Band). Diese Koordination besteht zum einen in der simultanen Ausführung von handlungskonstitutiven Aktivitäten mit unterschiedlichen modalen Ressourcen. Die verschiedenen Modalitäten haben dabei aber jeweils ihre eigene zeitliche Ordnung und sind in jeweils spezifischer Weise sequenziell und simultan mit anderen Ressourcen koordiniert. Besonders gut erforscht ist die Koordination von verbalen Referenzen mit vorlaufenden Zeigegesten (Streeck 2009; Schegloff 1984; Kendon 1972). Ein Beispiel solcher Koordination zeigt #11. Die Leiterin eines Rettungseinsatzes instruiert ihre Assistentinnen:

#11 FOLK_RETÜ_01_A02; 00:44-00:48

01 EL: %*du machs vital*par#ame#ter; (.)
 EL: *zeigt auf AS1-*...#zeigt auf Messgerät#,,,,,,
 EL: *schaut zu AS1-----#,,,,,,
 %Grafik 1
 02 EL: %*von dir will ich_n öh öh STÜTZverband,# *(.) mit öh öhm#
 EL: *zeigt auf Ersthilfekoffer vor AS2----->*auf Tasche--->
 EL: *schaut auf Ersthilfekoffer-----#schaut auf PATs
 Bauch#
 %Grafik 2
 03 EL: (--) zwei PÄCKchen* und MULL[binden, *
 EL: ----->*zeigt auf Ersthilfekoffer*

Abb. 1: Einsatzleiterin zeigt auf und schaut zu Assistentin 1

Abb. 2: Einsatzleiterin schaut und zeigt auf Ersthilfekoffer

Wie auch sonst üblich, beginnen hier alle Zeigegesten unmittelbar vor bzw. zu Beginn der TCU bzw. der Phrase, in der die Referenz verbalisiert wird, auf die die Zeigegeste Bezug nimmt. Der Onset der ersten Zeigegeste auf die Assistentin 1 („du“) beginnt, wie häufig, schon vor dem Turn (S01). Die folgende Zeigegeste auf das Messgerät, mit dem die Vitalparameter gemessen werden sollen, wird so produziert, dass der Apex der Geste während der verbalen Referenz „vITALparameter“ erreicht wird. In S02 antizipieren die Gesten von EL noch deutlicher die verbalen Referenzen: Zeitgleich mit dem Beginn der TCU zeigt EL auf den Ersthilfekoffer, in dem sich der im weiteren Verlauf der TCU benannte Stützverband befindet, um dann nach Abschluss der TCU auf die Tasche zu zeigen, in welcher sich die in der folgenden TCU angesprochenen „PÄCKchen“ befinden (S02-03). Am Ende der Phrase „zwei PÄCKchen“ wechselt EL erneut das Zeigziel und zeigt nun wiederum auf den Ersthilfekoffer, in dem die Mullbinden sind, auf die sie in der folgenden Phrase referiert. Da es sich um die Einführung neuer Referenten handelt, liegt auf den verbalen Referenzen zusätzlich jeweils der Fokusakzent. Obwohl selbst nicht zum verbalen Turn gehörig, sind die Zeigegesten für die mit dem Turn vollzogene Handlung konstitutiv. Er funktioniert nicht ohne sie als autonome kommunikative Einheit. Die Abgrenzung des Turnbeginns ist oft schwierig. Turns beginnen vielfach mit Vorlaufelementen wie (kommunikativ relevantem, da Turnbeanspruchung signalisierendem) Einatmen und Redeannahme oder Dispräferiertheit indizierenden Vokalisationen und die Turnproduktion erfordert multimodale Arrangements, z. B. die Schaffung eines gemeinsamen Interaktionsraums und die Koordination wechselseitiger Wahrnehmung und Aufmerksamkeit der Interaktionspartner (Mondada 2007; Deppermann 2013b).

Auch die Beendigung von Turns, vor allem solcher, die zugleich themen- oder sequenzbeendigend sind, impliziert oftmals die Koordination mehrerer multimodaler Ressourcen. So werden z. B. Zeigegesten und Objekte, die im Rahmen der Turnkonstruktion eine Rolle gespielt haben (z. B. als Referenten oder Verkörperungen des Diskurstopiks), kurz vor oder gleichzeitig mit der Turnbeendigung retrahiert, während die Abwendung des Rumpfes vom Gesprächspartner erst nach der Turnbeendigung vollendet wird (Mondada i. Dr.). Es kann aber durchaus sein, dass schon während der Turnproduktion mit dem Blick bereits die Transition zu einem nächsten Interaktionspartner bzw. Aufmerksamkeitsobjekt eingeleitet wird, also eine schrittweise und modalitätsspezifisch asynchrone Transition von einer fokussierten Interaktion zu einer nächsten stattfindet (Schmitt/Deppermann 2010).

5 Fazit

Keines der Konzepte, die für die Beschreibung der Grundeinheiten des Sprechens vorgeschlagen wurden, kann allen Phänomenen, die im Sprechen in Interaktionen regelmäßig vorkommen, vollständig Rechnung tragen. Alle Konzepte sind mit empi-

rischen Phänomenen konfrontiert, deren Einheiten-Status in Bezug auf das jeweilige Konzept unklar bzw. nicht zu erfassen ist. Dies liegt daran, dass alle Konzepte erfordern, dass Interaktionsbeiträge bestimmten Strukturierungsprinzipien folgen, die aber in der Gesprächspraxis nicht immer konsistent angewandt werden bzw. nicht für alle sprachlichen Ebenen gleichermaßen Gültigkeit haben (können). Die Diskussion hat gezeigt, dass die Suche nach Einheiten des Sprechens in der Interaktion von zwei unterschiedlichen Perspektiven motiviert ist. Zum einen interessiert die Erklärung der Handlungsorganisation, d. h., es werden gesprächs- und handlungsorganisatorisch relevante Einheiten identifiziert. Dabei sind auch multimodale Praktiken für die Einheitenkonstitution mitentscheidend. Zum anderen interessiert die Konstitution verbaler Einheiten. Die Frage ‚Aus welchen Grundeinheiten besteht ein Gespräch?‘ umfasst also aus der ersten Perspektive alle Modalitäten und fragt primär nach der Relevanz der identifizierten Einheiten für die Gesprächsteilnehmer, während die zweite Perspektive spezifisch nach der Organisation sprachlich definierter Einheiten fragt. Sie haben zwar oftmals eine gesprächsorganisatorische Funktion, doch ist dies nicht der primäre Klassifikationsaspekt. Dies geschieht notgedrungen zum Teil um den Preis, dass Phänomene, die für das situierte Handeln der Gesprächsteilnehmer relevant sind und die auch Funktion und Verständnis des Sprechens fundieren, außer Acht gelassen werden.

Wenn kein Einheitenkonzept konsistent und durchgängig angewandt werden kann, dann folgt daraus, dass eine exhaustive und eindeutige Segmentierung des Sprechens in Interaktionen nicht möglich ist. Auer (2010) fordert, daraus die Konsequenz zu ziehen, das Konzept der Segmentierung für das Sprechen überhaupt aufzugeben und stattdessen Zäsuren anzunehmen, die mehr oder weniger eindeutig sind und oftmals im Nachhinein, anhand des weiteren Interaktionsverlaufs, revidiert werden. Barth-Weingarten (2013) schlägt vor, im Rahmen eines Parameteransatzes unterschiedliche Zäsurstärken komparativ zu bestimmen. Turnproduktion ist ein dynamischer, durch apriorische Planung nicht (vollständig) determinierter Prozess. Er ist stets für situative und interaktive Kontingenzen offen, die zu Adaptationen, Reparaturen, Planänderungen, Expansionen und anderen situiert-prozessualen Phänomenen der Turngestaltung führen („Online-Syntax“, Auer 2000, 2007a). So gesehen scheint die Sicht, dass das Sprechen in Interaktionen in Einheiten organisiert sei, ein Post-hoc-Artefakt zu sein. Es verdankt sich in erheblichem Maße der linguistischen Tradition der Einheitenbildung, die zumeist an intuitiven Beispielen entwickelt wurde, und es beruht wesentlich auf der Perspektive des Analytikers, der immer erst im Nachhinein, in Form einer Aufnahme oder eines Transkripts, mit sprachlicher Praxis in gefrorener Form konfrontiert ist (Ford 2004). Ist es also deshalb besser, die Suche nach Einheiten aufzugeben und die Vorstellung von Einheiten als ideologisch abzulehnen? Sicher wäre dies zu einfach. An verschiedenen Stellen dieses Beitrags haben wir darauf hingewiesen, dass sich Interaktionsteilnehmer an routinisierten Formaten, an Projektionen von Fortsetzungen und Endpunkten von sprachlichen Strukturen und Handlungen orientieren. Mit dem Wissen um diese

koordinieren sie ihr Handeln und zeigen an, wo sie Einheiten als vollendet verstehen und behandeln. Einheiten haben also zweifellos ihren Platz in der kommunikativen Praxis für die InteraktionsteilnehmerInnen. Wenn unsere Diagnose zutrifft, dass (aus guten Gründen) keines der zur Verfügung stehenden Einheitenkonzepte für alle Phänomene des Sprechens gleichermaßen passend ist und eine vollständige und deterministische Einheitensegmentierung erlaubt, dann wird es nicht zu vermeiden sein, dass über die Logik der Segmentierung in Transkripten nach pragmatischen und theorieabhängigen Kriterien (die natürlich empirisch fundiert sein müssen) entschieden werden muss. Dennoch bleibt festzuhalten, dass Einheiten in der Interaktion noch viel flexibler und verhandelbarer sind, als dies die für die Organisation des Sprechens bisher vorgeschlagenen Konzepte annehmen lassen.

6 Literatur

- Auer, Peter (1991): Vom Ende deutscher Sätze. In: Zeitschrift für germanistische Linguistik 19, 139–157.
- Auer, Peter (1996): On the prosody and syntax of turn-continuations. In: Elizabeth Couper-Kuhlen/ Margret Selting (Hg.): Prosody in Conversation. Cambridge, 57–100.
- Auer, Peter (2000): *On line*-Syntax – Oder: was es bedeuten könnte, die Zeitlichkeit der mündlichen Sprache ernst zu nehmen. In: Sprache und Literatur 31, 43–56.
- Auer, Peter (2006): Increments and more. Anmerkungen zur augenblicklichen Diskussion über die Erweiterbarkeit von Turnkonstruktionseinheiten. In: Arnulf Deppermann/Reinhard Fiehler/ Thomas Spranz-Fogasy (Hg.): Grammatik und Interaktion. Untersuchungen zum Zusammenhang von grammatischen Strukturen und Gesprächsprozessen. Radolfzell, 279–294.
- Auer, Peter (2007a): Syntax als Prozess. In: Heiko Hausendorf (Hg.): Gespräch als Prozess. Linguistische Aspekte der Zeitlichkeit verbaler Interaktion. Tübingen, 95–124.
- Auer, Peter (2007b): Why are increments such elusive objects? An afterthought. In: Pragmatics 17 (4), 647–658.
- Auer, Peter (2010): Zum Segmentierungsproblem in der Gesprochenen Sprache. In: InLiSt – Interaction and Linguistic Structures 49. <http://www.inlist.uni-bayreuth.de/issues/49/InList49.pdf> [16.12.2014].
- Auer, Peter/Margret Selting (2001): Der Beitrag der Prosodie zur Gesprächsorganisation. In: Brinker u. a., 1122–1131.
- Auer, Peter/Susanne Günthner (2005): Die Entstehung von Diskursmarkern im Deutschen – ein Fall von Grammatikalisierung? In: Torsten Leuschner/Tanja Mortelsmans (Hg.): Grammatikalisierung im Deutschen. Berlin/New York, 335–362.
- Barth-Weingarten, Dagmar (2009): When to say something – some observations on prosodic-phonetic cues to the placement and types of responses in multi-unit turns. In: Dagmar Barth-Weingarten/Nicole Dehé/Anne Wichmann (Hg.): Where Prosody Meets Pragmatics. Bingley, 143–181.
- Barth-Weingarten, Dagmar (2011): The fuzziness of intonation units: some theoretical considerations and a practical solution. In: InLiSt – Interaction and Linguistic Structures 51. <http://www.inlist.uni-bayreuth.de/issues/51/Inlist51.pdf> [16.12.2014].
- Barth-Weingarten, Dagmar (2013): From „intonation units“ to cesuring – an alternative approach to the prosodic-phonetic structuring of talk-in-interaction. In: Szczepek Reed/Raymond, 91–124.

- Barth-Weingarten, Dagmar/Elisabeth Reber/Margret Selting (Hg.) (2010): *Prosody in Interaction*. Amsterdam/Philadelphia.
- Brinker, Klaus u. a. (Hg.) (2001): *Text- und Gesprächslinguistik. Ein internationales Handbuch zeitgenössischer Forschung*. 2. Halbband. Berlin/New York.
- Brinker, Klaus/Sven Sager (2010): *Linguistische Gesprächsanalyse. Eine Einführung*. 5., neu bearbeitete Auflage. Berlin.
- Bublitz, Wolfram (1988): Supportive Fellow-Speakers and Cooperative Conversations. Discourse Topics and Topical Actions, Participant Roles and 'Recipientaction' in a Particular Type of Everyday Conversation. Amsterdam/Philadelphia.
- Chafe, Wallace (1994): *Discourse, Consciousness, and Time: The Flow and Displacement of Conscious Experience in Speaking and Writing*. Chicago.
- Clark, Herbert H. (2012): Wordless questions, wordless answers. In: Jan de Ruiter (Hg.): *Questions. Formal, Functional and Interactional Perspectives*. Cambridge.
- Clark, Herbert H./Jean E. Fox Tree (2002): Using *uh* and *um* in spontaneous speaking. In: *Cognition* 84 (1), 73–111.
- Clayman, Steven E. (2013): Turn-constructional units and the transition-relevance place. In: Sidnell/Stivers, 150–165.
- Couper-Kuhlen, Elizabeth (1986): *An Introduction to English Prosody*. Tübingen.
- Couper-Kuhlen, Elizabeth/Tsuyoshi Ono (2007): Incrementing in conversation. A comparison of practices in English, German and Japanese. In: *Pragmatics* 17 (4), 513–552.
- Cruttenden, Alan (1997): *Intonation*. Zweite Auflage. Cambridge.
- Deppermann, Arnulf (2012): Über Sätze in Gesprächsbeiträgen – wann sie beginnen und wann man sie braucht. In: Colette Cortès (Hg.): *Satzeröffnung. Formen, Funktionen, Strategien*. Tübingen, 1–14.
- Deppermann, Arnulf (2013a): Introduction: multimodal interaction from a conversation analytic perspective. In: *Journal of Pragmatics* 46 (1), 1–7.
- Deppermann, Arnulf (2013b): Turn-design at turn-beginnings: multimodal resources to deal with tasks of turn-construction in German. In: *Journal of Pragmatics* 46 (1), 91–121.
- Du Bois, John u. a. (1993): Outline of discourse transcription. In: Jane A. Edwards/Martin D. Lampert (Hg.): *Talking Data. Transcription and Coding in Discourse Research*. Hillsdale, 45–87.
- Ehlich, Konrad (1986): *Interjektionen*. Tübingen.
- Ehlich, Konrad (2010a): Funktionale Pragmatik – Terme, Themen und Methoden. In: Hoffmann, 214–231.
- Ehlich, Konrad (2010b): Sprechhandlungsanalyse. In: Hoffmann, 242–257.
- Fiehler, Reinhard u. a. (2004): *Eigenschaften gesprochener Sprache*. Tübingen.
- Fischer, Kerstin (Hg.) (2006): *Approaches to Discourse Particles*. Amsterdam.
- Ford, Cecilia E. (2004): Contingency and units in interaction. In: *Discourse Studies* 6 (1), 27–52.
- Ford, Cecilia/Sandra A. Thompson (1996): Interactional units in conversation. Syntactic, intonational, and pragmatic resources for the management of turns. In: Ochs/Schegloff/Thompson, 134–184.
- Ford, Cecilia E./Barbara A. Fox/Sandra A. Thompson (1996): Practices in the construction of turns. The „TCU“ revisited. In: *Pragmatics* 6, 427–454.
- Ford, Cecilia E./Barbara A. Fox/Sandra A. Thompson (2002): Constituency and the grammar of turn increments. In: Cecilia E. Ford/Barbara A. Fox/Sandra A. Thompson (Hg.): *The Language of Turn and Sequence*. Oxford, 14–38.
- Ford, Cecilia E./Barbara A. Fox/Sandra A. Thompson (2013): Units and/or action trajectories? The language of grammatical categories and the language of social action. In: Szczepek Reed/Raymond, 13–55.

- Forsgren, Kjell-Åke (1992): Zum Problem des Satzbegriffs im Deutschen. In: Beiträge zur Geschichte der deutschen Sprache und Literatur 114 (1), 3–27.
- Golato, Andrea/Emma Betz (2008): German *ach* and *achso* in repair uptake: resources to sustain or remove epistemic asymmetry. In: Zeitschrift für Sprachwissenschaft 27, 7–37.
- Goodwin, Charles (1979): The interactive construction of a sentence in natural conversation. In: George Psathas (Hg.): Everyday Language. Studies in Ethnomethodology. New York, 97–121.
- Hagemann, Sven/Eckart Rolf (2001): Die Bedeutung der Sprachakttheorie für die Gesprächsforschung. In: Brinker u. a., 885–896.
- Halford, Brigitte K. (1996): Talk Units. The Structure of Spoken Canadian English. Tübingen.
- Halliday, Michael A. K. (1967): Intonation and Grammar in British English. The Hague/Paris.
- Helmer, Henrike (i. V.): Analepsen im gesprochenen Deutsch. Mannheim (Unveröffentlichte Dissertation).
- Henne, Helmut/Helmut Rehbock (1995): Einführung in die Gesprächsanalyse. 3. Auflage. Berlin/New York.
- Hoffmann, Ludger (1999): Ellipse und Analepse. In: Angelika Redder/Jochen Rehbein (Hg.): Grammatik und mentale Prozesse. Tübingen, 69–90.
- Hoffmann, Ludger (Hg.) (2010): Sprachwissenschaft. Ein Reader. 3., aktuelle und erweiterte Auflage. Berlin/New York.
- Imo, Wolfgang (2012): Wortart Diskursmarker? In: Björn Rothstein (Hg.): Nicht-flektierende Wortarten. Berlin/New York, 48–88.
- Jefferson, Gail (1983): Issues in the transcription of naturally-occurring talk: caricature versus capturing pronunciation particulars. In: Tilburg Papers in Language and Literature 34, 1–12.
- Jefferson, Gail (2004): Glossary of transcript symbols with an introduction. In: Gene H. Lerner (Hg.): Conversation Analysis: Studies From the First Generation. Amsterdam/Philadelphia, 13–23.
- Kendon, Adam (1972): Some relationships between body motion and speech. In: Aron Wolfe Siegman/Benjamin Pope (Hg.): Studies in Dyadic Communication. Elmsford/New York, 177–216.
- Kindt, Walter (1994): Satzbegriff und gesprochene Sprache. In: Lingua 94, 25–48.
- Koch, Peter/Wulf Oesterreicher (1985): Sprache der Nähe – Sprache der Distanz. Mündlichkeit und Schriftlichkeit im Spannungsfeld von Sprachtheorie und Sprachgeschichte. In: Olaf Deutschmann u. a. (Hg.): Romanistisches Jahrbuch. Berlin/New York, 15–43.
- Lerner, Gene H. (1991): On the syntax of sentences in progress. In: Language in Society 20, 441–458.
- Levinson, Stephen C. (1981): The essential inadequacies of speech act models of dialogue. In: Herman Parret/Marina Sbisa/Jef Verschueren (Hg.): Possibilities and Limitations of Pragmatics: Proceedings of the Conference on Pragmatics, Urbino, July 8–14, 1979. Amsterdam, 473–492.
- Levinson, Stephen C. (2013): Action formation and ascription. In: Sidnell/Stivers, 103–129.
- Luke, Kang-kwong/Sandra A. Thompson/Tsuyoshi Ono (2012): Turns and increments: a comparative perspective. In: Discourse Processes 49, 155–162.
- Mondada, Lorenza (2007): Interaktionsraum und Koordinierung. In: Reinhold Schmitt (Hg.): Koordination. Tübingen, 55–93.
- Mondada, Lorenza (i. Dr.): Multimodal completions. In: Arnulf Deppermann/Susanne Günthner (Hg.): Temporality in Interaction. Amsterdam/Philadelphia.
- Nübling, Damaris (2005): Die prototypische Interjektion: Ein Definitionsvorschlag. In: Zeitschrift für Semiotik 26, 11–45.
- Ochs, Elinor/Emanuel A. Schegloff/Sandra A. Thompson (Hg.) (1996): Interaction and Grammar. Cambridge.
- Oloff, Florence (2009): Contribution à l'étude systématique de l'organisation des tours de parole: Les chevauchements en français et en allemand. Mannheim (Dissertation). [https://ub-madoc.bib.uni-mannheim.de/29617/\[16.12.2014\]](https://ub-madoc.bib.uni-mannheim.de/29617/[16.12.2014]).

- Oloff, Florence (2013): Embodied withdrawal after overlap resolution. In: *Journal of Pragmatics* 46 (1), 139–156.
- Ono, Tsuyoshi/Elizabeth Couper-Kuhlen (2007): Increments in cross-linguistic perspective: introductory remarks. In: *Pragmatics* 17 (4), 505–512.
- Rath, Rainer (1976): Gesprochenes Deutsch und seine Gliederung. In: *Beiträge zu den Fortbildungskursen des Goethe-Instituts für ausländische Deutschlehrer an Schulen und Hochschulen*. München, 63–76.
- Rath, Rainer (1985): Geschriebene und gesprochene Form der heutigen Standardsprache. In: Werner Besch/Oskar Reichmann/Stefan Sonderegger (Hg.): *Sprachgeschichte. Ein Handbuch zur Geschichte der deutschen Sprache und ihrer Erforschung*. 2. Halbband. Berlin/New York, 1651–1663.
- Rath, Rainer (1990): „Satz“ und „Äußerungseinheit“. Syntaktische und interaktive Struktur in der Sprache? In: Eynar Leupold/Yvonne Petter (Hg.): *Interdisziplinäre Sprachforschung und Sprachlehre. Festschrift für Albert Raasch zum 60. Geburtstag*. Tübingen, 197–216.
- Rath, Rainer (1997): „Äußerungseinheit“ oder „möglicher Satz“? In: *Deutsche Sprache* 25, 1–20.
- Rath, Rainer (2001): Gesprächsschritt und Höreraktivitäten. In: Brinker u. a., 1213–1226.
- Rauniomaa, Mirka/Tiina Keisanen (2012): Two multimodal formats for responding to requests. In: *Journal of Pragmatics* 44, 829–842.
- Reber, Elisabeth/Elizabeth Couper-Kuhlen (2010): Interjektionen zwischen Lexikon und Vokalität: Lexem oder Lautobjekt? In: Arnulf Deppermann/Angelika Linke (Hg.): *Sprache intermedial. Stimme und Schrift, Bild und Ton*. Berlin/New York, 69–96.
- Rehbein, Jochen (1979): Sprechhandlungsaugmente. Zur Organisation der Hörersteuerung. In: Harald Weydt (Hg.): *Die Partikeln der deutschen Sprache*. Berlin/New York, 58–74.
- Rehbein, Jochen (1995): Segmentieren. *VerbMobil Memo 64*. Unveröffentlichtes Manuskript.
- Rehbein, Jochen (2001): Das Konzept der Diskursanalyse. In: Brinker u. a., 927–945.
- Rehbein, Jochen u. a. (2004): *Handbuch für das computergestützte Transkribieren nach HIAT. Arbeiten zur Mehrsprachigkeit 59*. Hamburg.
- Sacks, Harvey/Emanuel A. Schegloff/Gail Jefferson (1974): A simplest systematics for the organisation of turn taking in conversation. In: *Language* 50 (4), 696–735.
- Schegloff, Emanuel A. (1981): Discourse as an interactional achievement: some uses of ‚uh huh‘ and other things that come between sentences. In: Deborah Tannen (Hg.): *Analyzing Discourse: Text and Talk*. Washington, 71–93.
- Schegloff, Emanuel A. (1984): On some gestures' relation to talk. In: J. Maxwell Atkinson/John Heritage (Hg.): *Structures of Social Action. Studies in Conversation Analysis*. Cambridge, 299–345.
- Schegloff, Emanuel A. (1987): Analyzing single episodes of interaction. An exercise in conversation analysis. In: *Social Psychological Quarterly* 50 (2), 101–114.
- Schegloff, Emanuel A. (1996): Turn organization: one intersection of grammar and interaction. In: Ochs/Schegloff/Thompson, 52–133.
- Schegloff, Emanuel A. (2007): *Sequence Organization in Interaction. A Primer in Conversation Analysis*. Band 1. Cambridge.
- Schiffrin, Deborah (1987): *Discourse Markers*. Cambridge.
- Schmidt, Claudia (1988): ‚Typisch weiblich – typisch männlich‘. Geschlechtstypisches Kommunikationsverhalten in studentischen Kleingruppen. Tübingen.
- Schmitt, Reinhold (2005): Zur multimodalen Struktur von turn-taking. In: *Gesprächsforschung – Online-Zeitschrift zur verbalen Interaktion* 6, 17–61. <http://www.gespraechsforschung-ozs.de/fileadmin/dateien/heft2005/ga-schmitt.pdf> [16.12.2014].

- Schmitt, Reinhold/Arnulf Deppermann (2010): Die Transition von Interaktionsräumen als Eröffnung einer neuen Situation. In: Lorenza Mondada/Reinhold Schmitt (Hg.): *Situationseröffnungen. Zur multimodalen Herstellung fokussierter Interaktion*. Tübingen, 335–386.
- Schwitalla, Johannes (2002): Kleine Wörter. Partikeln im Gespräch. In: Jürgen Dittmann/Claudia Schmidt (Hg.): *Über Wörter. Grundkurs Linguistik*. Freiburg i. Br., 259–281.
- Schwitalla, Johannes (2012): *Gesprochenes Deutsch. Eine Einführung*. 4., neu bearbeitete und erweiterte Auflage. Berlin.
- Selting, Margret (1987): Reparaturen und lokale Verstehensprobleme. Oder: Zur Binnenstruktur von Reparatursequenzen. In: *Linguistische Berichte* 108, 128–149.
- Selting, Margret (1995a): Der ‚mögliche Satz‘ als interaktiv relevante syntaktische Kategorie. In: *Linguistische Berichte* 158, 298–325.
- Selting, Margret (1995b): Prosodie im Gespräch. Aspekte einer interaktionalen Phonologie der Konversation. Tübingen.
- Selting, Margret (2000): The construction of units in conversational talk. In: *Language in Society* 29 (4), 477–517.
- Selting, Margret (2001): Fragments of units as deviant cases of unit-production in conversational talk. In: Margret Selting/Elizabeth Couper-Kuhlen (Hg.): *Studies in Interactional Linguistics*. Amsterdam, 229–258.
- Selting, Margret u. a. (2009): Gesprächsanalytisches Transkriptionssystem 2 (GAT 2). In: *Gesprächsforschung – Online-Zeitschrift zur verbalen Interaktion* 10, 353–402.
<http://www.gespraechsforschung-ozs.de/heft2009/px-gat2.pdf> [16.12.2014].
- Sidnell, Jack/Tanya Stivers (Hg.) (2013): *The Handbook of Conversation Analysis*. Oxford.
- Sorjonen, Marja-Leena (2001): *Responding in Conversation: A Study of Response Particles in Finnish*. Amsterdam.
- Staffeldt, Sven (2014): Sprechakttheoretisch analysieren. In: Sven Staffeldt/Jörg Hagemann (Hg.): *Pragmatiktheorien. Analysen im Vergleich*. Tübingen, 105–148.
- Stivers, Tanya (2008): Stance, alignment and affiliation during story telling: when nodding is a token of preliminary affiliation. In: *Research on Language in Social Interaction* 41, 29–55.
- Streeck, Jürgen (2009): Forward-Gesturing. In: *Discourse Processes* 46 (2–3), 161–179.
- Szcepek Reed, Beatrice (2004): Turn-final intonation in English. In: Elizabeth Couper-Kuhlen/Cecilia E. Ford (Hg.): *Sound Patterns in Interaction: Cross-Linguistic Studies From Conversation*. Amsterdam, 97–117.
- Szcepek Reed, Beatrice (2010): Intonation phrases in natural conversation: a participant's category? In: Barth-Weingarten/Reber/Selting, 191–212.
- Szcepek Reed, Beatrice/Geoffrey Raymond (Hg.) (2013): *Units of Talk – Units of Action*. Amsterdam/Philadelphia.
- Walker, Gareth (2010): The phonetic constitution of a turn-holding practice: rush-throughs in English talk-in-interaction. In: Barth-Weingarten/Reber/Selting, 51–72.
- Yngve, Victor (1970): On getting a word in edgewise. In: *Papers from the Sixth Regional Meeting of the Chicago Linguistic Society*. Chicago.
- Zifonun, Gisela (1987): *Kommunikative Einheiten in der Grammatik*. Tübingen.
- Zifonun, Gisela u. a. (1997): *Grammatik der deutschen Sprache*. 3 Bände. Berlin/New York.