

Das Dekodieren der Intonation von andauernden Intonationsphrasen

Dario Maric (Sarajevo)

Abstract

This article re-examines models of phonetic realization of tones offered by certain autosegmental-metrical models of intonation, taking into consideration knowledge of conversation analysis and approaches derived from conversation analysis applied on phonetic and phonological researches. These autosegmental-metric models of intonation record H and L tones usually according to its position within pitch range of individual speaker in an intonational phrase. According to this, listener's decoding of intonational realization is possible only in fully realized intonational phrase, since the highest and the lowest tone values that determine pitch range used in an intonational phrase can be realized within any syllable of intonational phrase. This kind of decoding of the intonational realization would mean that speaker cannot use intonation to make transition relevance place projectable before it is realized, which is contrary with intonational research knowledge within conversation analysis and approaches derived from conversation analysis applied on phonetic and phonological researches.

1 Einleitung

Nachdem begonnen wurde, die Erkenntnisse der Konversationsanalyse als relevante methodische Ausgangspunkte innerhalb phonetischer und phonologischer Forschung anzuwenden, entstand der Bedarf an Überprüfung einzelner Analyseergebnisse innerhalb der Phonetik und der Phonologie. Aus der Konversationsanalyse hervorgegangene, an phonetischen und phonologischen Analysen angewandten Ansätze ('A Phonology of Conversation' Local/Kelly/Wells 1986, 'Interaktionale Phonologie der Konversation' Selting 1995) fordern außerdem die Erforschung phonetischer und phonologischer Realisation in spontanen Gesprächen, und nicht in vorgelesenen Wörtern oder Sätzen, die außerhalb des Kontextes konstruiert sind.

Eine der für diese phonologischen Ansätze einflussreichsten Thesen der Konversationsanalyse besagt, dass sich die soziale Interaktion in Gesprächen in Form eines kooperativen Bestrebens der Gesprächsteilnehmer äußert, eigene Kommunikationsvorhaben zu verwirklichen und anderen zu deuten sowie fremde Kommunikationsvorhaben zu interpretieren. Dies erreichen die Gesprächsteilnehmer, indem sie Gespräche organisieren und steuern. Spuren dieses Bestrebens können in der Gesprächsstruktur selbst erkannt werden, was die Gesprächsanalytiker für die Beschreibung der Regelmäßigkeiten nutzen, die allerdings nicht als korrekte sondern als unmarkierte oder präferierte Realisationen gelten. Unmarkierte Realisationen sind nämlich erwartete verbale Reaktionen der Gesprächsteilnehmer auf die gedeuteten Kommunikationsvorhaben, während die markierten Realisationen nicht unkorrekt sind. Sie stellen vielmehr Abweichungen von erwarteten Reaktionen dar und lenken auf andere Interpretationen im Prozess des Folgerns. Die Verfahren, mit denen Gesprächsteilnehmer Konversationsziele verwirklichen, werden auf konkreten Kontext hin verwendet, d. h. ihre Funktion üben sie gänzlich nur an der Stelle innerhalb der Sequenz in der konkreten Kommunikationssituation aus, für die sie speziell konstruiert sind.

Durch die Sequenzorganisation im Gespräch interpretieren die Gesprächsteilnehmer die Funktion eigener zuvor realisierter Sequenzen, der Sequenzen, die gerade im Entstehen sind, oder der noch zu realisierenden, angekündigten Sequenzen.

Die autosegmental-metrischen Intonationsmodelle sehen in ihren endgültigen Beschreibungen der Intonation bestimmter Sprachen die intonatorischen Funktionen der Gesprächsorganisation, deren Existenz nach Anwendung konversationsanalytischer Ansätze in Untersuchungen bezeugt werden konnte, jedoch nicht vor. Diese Beschreibungen in Form der 'Finite-State'-Grammatiken der Intonation bestimmen alle möglichen Intonationskonturen, bzw. alle möglichen Kombinationen phonologisch relevanter Ton- und Phrasenakzente, sowie Grenztöne, die im Toninventar enthalten sind, ohne davor die phonologische Repräsentation der Intonation aller in der betreffenden Sprache bekannten Funktionen der Intonation festgelegt zu haben.

Local/Kelly/Wells (1986: 412–414) warnen vor solchen verfrühten Phonologisierungen des Beschreibungsapparats ins Besondere im Hinblick auf die Tatsache, dass die Relevanz dieser phonologischen Kategorien durch die Evidenz des Orientierens der Gesprächsteilnehmer an ihnen bewiesen werden muss.

Gilles (2005: 14) bemängelt am autosegmental-metrischen Intonationsmodell für das Deutsche GToBI genau die Folge verfrühter Phonologisierung, nämlich die Existenz der phonologischen Kategorien, die "keine Entsprechung auf der Formseite haben, sondern vielmehr aufgrund theorie-immanenter Annahmen postuliert werden". An GToBI setzt Gilles (2005: 15) außerdem eine begrenzte Anzahl der Analysen von Funktionen aus, die diese Konturen übernehmen.

Das GToBI-System bietet außerdem das Modell der phonetischen Realisierung und somit implizit des Dekodierens der Töne, nach welchem das Dekodieren der intonatorischen Realisation vor der gänzlich verwirklichten Intonationsphrase nicht möglich ist, was im Widerspruch zu Erkenntnissen der Forschung des Sprecherwechsels in Gesprächen steht.

2 Die phonetische Realisierung und das Dekodieren der Töne nach autosegmental-metrischen Intonationsmodellen

Die Grundeinteilung der Intonationsmodelle gründet auf unterschiedlichen Auffassungen bedeutungsrelevanter Bestandteile der intonatorischen Realisation. Die konturbasierten Modelle sind der Ansicht, dass phonologisch relevant nur Tonbewegungen sind, während die Ebenen-Modelle die Intonationskonturen als das Ergebnis des Verbindens der Zielpunkte von phonologisch relevanten Tönen betrachten. Die Einteilung der Intonationsmodelle auf der Basis unterschiedlicher Auffassungen der kleinsten bedeutungstragenden Einheiten der Intonation ist auch möglich. Dabei geht es um das Akzeptieren der Existenz der Intonationsmorpheme, gleichviel ob es sich bei Intonationsmorphemen um Konturen oder Töne handelt.

Die konturbasierten Modelle sowie die Intonationsmodelle zur Sprachsynthese (z. B. Fujisaki, TILT und KIM) werden hier nicht thematisiert, weil die Probleme bezüglich des Dekodierens der intonatorischen Realisation an möglichen Satzenden, die innerhalb der Intonationsphrase eventuell vorkommen, auf sie nicht zutreffen. Auf diese Modelle wird im Kapitel 4 näher eingegangen.

Der autosegmentale Ansatz mit zwei, eventuell drei abstrakten Tönen im Paradigma ist als Versuch entstanden, tonale Erscheinungen in afrikanischen Sprachen zu beschreiben, die die damaligen phonologischen Theorien nicht beschreiben konnten.

Die Töne konnten bei der Beschreibung einzelnen Segmenten (Phonemen) nicht zugeordnet werden, so dass innerhalb der phonologischen Repräsentation der segmentalen Ebene zusätzlich die prosodische Ebene in mehreren getrennten Unterebenen (eng. 'tier') hinzugefügt wurde. Das Sprechsignal war auf diese Weise in mindestens zwei autonomen aber parallelen

Ebenen (segmentale und tonale Ebene) in lineare Abfolgen der Einheiten unterteilt, wovon auch die Bezeichnung autosegmentale Phonologie abgeleitet ist. Die tonale Ebene stellte die Abfolge der abstrakten hohen (H – 'high') und tiefen (L – 'low') Töne dar.

Da in Beschreibungen der Intonation im Rahmen der autosegmentalen Phonologie auch die metrische Struktur die phonologischen Prozesse beeinflusst, fügt Ladd (1996) der Bezeichnung autosegmentale metrische bei, um die Ansätze, die sich an die Arbeit der amerikanischen Wissenschaftlerin Pierrehumbert (1980) anlehnen, zu benennen.

Die autosegmental-metrischen Modelle verwenden für die Beschreibung der Intonationskonturen hauptsächlich die Tonabfolgen mit nur zwei abstrakten Tönen (H und L) im Paradigma. Diese Töne können Tonakzente (z. B. L+H*), Grenztöne (z. B. H%) sowie Phrasenakzente (z. B. L oder L-) darstellen, je nachdem welchem konkreten Bestandteil der segmentalen Ebene, dem sogenannten tontragendem Element (eng. 'Tone Bearing Unit' oder abgekürzt 'TBU'), sie zugeordnet werden. Das tontragende Element kann in verschiedenen Sprachen oder Dialekten die phonologische Dauereinheit Mora, die Silbe, der Vokal oder der Sonorant sein. Wenn Töne betonten Silben der hervorgehobenen Wörter zugeordnet werden, handelt es sich um Tonakzente. Sind sie der ersten oder der letzten Silbe der Intonationsphrase zugeordnet, handelt es sich um Grenztöne. Die Töne, die einer lexikalisch betonbaren Silbe, die nach dem hervorgehobenen Wort in der Intonationsphrase vorkommt, zugeordnet sind, realisieren Phrasenakzente.

Nach dem autosegmental-metrischen Modell von Pierrehumbert (1980) werden die Tonhöhenwerte der Töne immer in Bezug auf die Basislinie ('baseline') festgelegt. Die Basislinie stellt die hypothetisch angenommene, infolge der Deklination in einer Äußerung konstant fallende Linie dar, die die untere Grenze des Sprechstimmumfangs des Sprechers ausmacht. Die Tonhöhenwerte der Basislinie müssen nicht aus den konkreten niedrigsten Tonhöhenwerten des Intonationsverlaufs erschlossen werden. Sie stehen nämlich in einer konstanten Beziehung zu Tonhöhenwerten der Zielpunkte über der Basislinie auch bei denjenigen Intonationsverläufen, die keine niedrigen Tonhöhenwerte für eine Basislinie aufweisen. Festgelegt werden die Tonhöhenwerte der Töne im Pierrehumberts Modell außerdem in Bezug auf die Prominenzabstufung des Fokus sowie in Bezug auf die phonetischen und phonologischen Werte des vorausgehenden Tons. Nach diesem Modell sind also die Referenzwerte, die der Hörer für die Festlegung der Töne benötigt, teilweise schon vor dem Ansetzen der Intonationsphrase (im Falle der hypothetisch angenommenen Basislinie) zugänglich, teilweise befinden sie sich im vorausgehenden Ton. Die Zugänglichkeit der Referenzwerte für den Hörer in jedem Moment der Entstehung der Intonationsphrase macht das Dekodieren der Töne auch vor übergaberelevanten Stellen möglich.

Pierrehumbert beweist im Experiment mit Veränderungen des Stimmumfangs infolge sukzessiver Erhöhung der Emphase in Äußerungen (1980: 48) die Konstanz der Tonhöhenwerte der Basislinie. Die niedrigsten Tonhöhenwerte in H* L⁻ L% und H% L⁻ H% Teilverläufen, die innerhalb verschiedener Stimmumfangs eines Sprechers realisiert wurden, blieben nämlich immer konstant, während die höheren Werte variierten.

In Pierrehumberts Modell wird aber außer Acht gelassen, dass das Tonhöhenregister und somit auch die Tonhöhenwerte der Basislinie für kommunikative Zwecke verändert werden können (Cruttenden 1997: 124; Gussenhoven 2004: 76; für das Deutsche Selting 1995: 155–177), so dass von einem hypothetisch angenommenen, für einen Sprecher konstanten Tonhöhenwert der Basislinie nicht ausgegangen werden kann.

Gussenhoven (2004: 76) spricht im Zusammenhang mit Veränderungen des Tonhöhenregisters von der Variation der 'Referenzlinie' mit "constant value for every contour that enters into the calculation of the targets for the Hs and Ls". Die Wahl des Tonhöhenregisters ist also

für die ganze Intonationsphrase gültig: "A choice of pitch register will remain valid for the whole intonational phrase" (Gussenhoven 2004: 76).

Die Vertreter der Ansätze, die auf der Vorstellung der Intonation als einer Folge kleinerer Intonationskonturen basiert, insbesondere die Befürworter der IPO (Institut für Perzeptionsforschung in Eindhoven) Theorie setzen dem autosegmental-metrischen Ansatz hinsichtlich der Festlegung der Töne die Vernachlässigung der Tatsache aus, dass der Ton nur im Fall, dass er verglichen wird mit einem Ton, der im gleichen Kontext tiefer ist als er, als hoch wahrgenommen wird, beziehungsweise, dass er als tief wahrgenommen wird, wenn er mit einem Ton verglichen wird, der höher ist als er.

Aufgrund der Erkenntnisse von Pierrehumbert und Beckmann über die Intonationsstruktur im Englischen wurde das autosegmental-metrische Transkriptionssystem MAE_ToBI (Mainstream American English Tones and Break Indices, ursprünglich nur ToBI) konzipiert. Im Vergleich zur Notationsweise, die in ihren Studien (Pierrehumbert 1980; Beckmann/Pierrehumbert 1986) vorgestellt wurde, ist dieses Transkriptionssystem durch eine geringfügig umgeänderte Notationsweise der Intonation gekennzeichnet.

Dank der Erkenntnis, dass das Prinzip der Notation von Prosodie des allgemeinen amerikanischen Englisch mit dem Transkriptionssystem MAE_ToBI auch auf andere Sprachen und Varietäten angewendet werden kann, ist das ToBI Bezugssystem für die Entwicklung der prosodischen Transkriptionssysteme verschiedener Sprachen und Varietäten des Englischen (Beckmann/Hirschberg/Shattuck-Hufnagel 2006) ausgearbeitet worden. ToBI bietet die Notation der prosodischen Realisation innerhalb mehrerer Ebenen: der tonalen Ebene, der Ebene für die Notation der Grenzen von Intonationsphrasen, Grenzen der Intermediärphrasen, Grenzen der Wörter innerhalb Intermediärphrasen und Grenzen innerhalb der Wörter, der Ebene für die Notation der Kommentare zu außer- und parasprachlichen prosodischen Erscheinungen sowie der Ebene für die orthographische Transkription.

Auf der Basis des ToBI Bezugssystems ist auch das autosegmental-metrische Transkriptionssystem GtoBI – German Tones and Break Indices (Grice/Benzmüller 1995; Grice/Baumann 2002; Grice/Baumann/Benzmüller 2005) entwickelt worden.

GtoBI legt die H- und L-Töne hauptsächlich nach ihrer Position im Stimmumfang, der in einer Intonationskontur genutzt wird ('pitch span', Ladd 1996: 260), fest, und in geringerem Maße nach dem Vergleich zum vorausgehendem Ton. Die Töne, die innerhalb der oberen drei Viertel des in einer Intonationskontur genutzten Stimmumfangs realisiert werden, werden als H-Töne notiert. Als L-Töne werden diejenigen Töne notiert, die innerhalb des unteren Viertels des in einer Intonationskontur genutzten Stimmumfangs realisiert werden:

The H and L tones are described as high or low relative to a speaker's pitch range which can be thought of as having a topline as an upper limit and a baseline as a lower limit. As a rule of thumb, tones which are perceived as high are roughly in the top three quarters of the range, those perceived as low in the bottom quarter. (Grice/Baumann/Benzmüller 2005: 13).

Diese Notations- bzw. Generierungsweise der Tonabfolgen beim Sprechen setzt im Voraus festgelegte ("vorausgeplante") phonetische Realisierung der ganzen Intonationskontur ('look-ahead' Ladd 1996: 29) vor dem Ansetzen der Intonationsphrase voraus. Denn, um wissen zu können, in welchem Viertel des Stimmumfangs der Intonationsphrase der Ton realisiert wird, ist es notwendig, die höchsten und die tiefsten Tonhöhenwerte, die den Stimmumfang direkt bestimmen, in der Intonationskontur festzulegen.

3 Die Rolle der Intonation beim Sprecherwechsel

In diesem Kapitel werden einige Forschungen zur Steuerung des Sprecherwechsels in Gesprächen dargestellt, die diesbezüglich unter anderem auch die Intonation für relevant halten.

Anschließend werden Auers Ausführung der Kontextualisierungstheorie (1986) und die Einordnung der intonatorischen Funktion beim Sprecherwechsel in diese Theorie dargestellt.

Obwohl auch in früheren Arbeiten die intonatorische Funktion beim Signalisieren des Fortsetzens ('continuation') bzw. des Abschließens der Sprechaktivität ('completion') erforscht wurde, wird erst mit dem Aufkommen der ethnomethodologischen Konversationsanalyse begonnen, über den Hörer als möglichen nächsten Sprecher sowie darüber, wie man zum Sprecher oder zum Hörer wird, zu denken. Die Tatsache in Betracht ziehend, dass der Sprecherwechsel in spontanen Gesprächen mit im Durchschnitt sehr wenigen Überlappungen und mit sehr kurzen Pausen zwischen den Redebeiträgen vollzogen wird, gehen die amerikanischen Soziologen Sacks/Schegloff/Jefferson (1974) von einem präzisen Mechanismus aus, der lokal in jedem einzelnen Redebeitrag den Sprecherwechsel steuert. Dieser Mechanismus beruht auf Regeln, die an übergaberelevanten Stellen greifen, und auf der Möglichkeit der Projizierung der übergaberelevanten Stellen vor ihrer Realisierung. Die Berechenbarkeit ist nämlich teilweise dank der Syntax möglich, deren jede einzelne Vollständigkeit vom Anfang des Redebeitrags die möglichen übergaberelevanten Stellen festlegt. Die übergaberelevanten Stellen werden nach dem Modell von Sacks, Schegloff und Jefferson jedoch endgültig intonatorisch vorausbestimmt. Die Intonation signalisiert die Kohäsie zwischen den vor den möglichen übergaberelevanten Stellen vorkommenden Sätzen, Teilsätzen, Phrasen oder einzelnen Wörtern, und Sätzen, Teilsätzen, Phrasen oder einzelnen Wörtern, die nach den möglichen übergaberelevanten Stellen vorkommen, oder lässt eben das Beenden des Redebeitrags erkennen: "For example, discriminations between *what* as one-word question and as the start of a sentential (or clausal or phrasal) construction are made not syntactically, but intonationally" (Sacks/Schegloff/Jefferson 1974: 721). Der erwähnte Satz, die Phrase oder das einzelne Wort kann dementsprechend durch ein- oder mehrmaliges Vorkommen einen Redebeitrag (Turn) konstituieren, weshalb diese Äußerungen als Turnkonstruktionseinheiten (auch: beitragsbildende Einheiten) definiert wurden: "There are various unit-types with which a speaker may set out to construct a turn. Unit types for English include sentential, clausal, phrasal, and lexical constructions" (Sacks/Schegloff/Jefferson 1974: 702). Die Turnkonstruktionseinheit setzt sich also aus mindestens einem Einwortsatz und höchstens aus einem komplexen Satz zusammen. Auf der Basis der syntaktischen Kriterien kann das mögliche Ende der Turnkonstruktionseinheit nur retrospektiv als ihr wirkliches Ende gekennzeichnet werden. Solche Bestimmung im Nachhinein ist die Folge der Möglichkeit der Grundstrukturweiterung um freie Angaben, der eventuellen Nachstellung der Attribute oder der eventuellen Besetzung der fakultativen, von der Valenz des Verbs eröffneten Leerstellen im Satz mit fakultativen Ergänzungen. Alle diese Erweiterungen, seien sie freie Angaben, Attribute oder fakultative Ergänzungen, können auch satzartig vorkommen. In diesem Zusammenhang ist auch die mögliche Anknüpfung eines neuen Hauptsatzes zur Entstehung einer Satzverbindung zu erwähnen. Sacks/Schegloff/Jefferson (1974) präzisieren jedoch nicht, auf welche Art und Weise die Turnkonstruktionseinheit strukturiert sein kann, d. h. was unter Vollständigkeit verstanden werden soll, sowie wodurch und auf welche Art und Weise die Grundstrukturen erweitert werden können:

How projection of unit-types is accomplished, so as to allow such 'no gap' starts by next speakers, is an important question on which linguists can make major contributions. (Sacks/Schegloff/Jefferson 1974: 703)

Der Hörer kann also mit Hilfe der syntaktischen Kompetenz nur mit eingeschränkter Sicherheit die übergaberelevanten Stellen, d. h. die Enden der Turnkonstruktionseinheiten und somit die Enden der Redebeiträge voraussehen. Die Rolle der Intonation in diesem Zusammenhang besteht darin, die Zweifel beim erwähnten Voraussehen lokal zu beseitigen.

Nach dem Sprecherwechsel bzw. im ersten Redebeitrag des Gesprächs haben die Sprecher das Recht auf nur eine Turnkonstruktionseinheit. Nach dem Abschluss der Turnkonstrukti-

onseinheit und dem Erreichen der übergaberelevanten Stelle entweder vergibt der Sprecher das Rederecht, indem er den nächsten Sprecher auswählt bzw. nicht auswählt, oder behält er das Rederecht, die Selbstwahl anwendend. Der nächste Sprecher kann im Fall, dass er nicht ausgewählt wurde, das Wort ergreifen, muss es aber nicht tun. Falls der nächste Sprecher vom letzten Sprecher nicht ausgewählt wurde, darf jeder Gesprächsteilnehmer, einschließlich den letzten Sprecher, gleichberechtigt um das Rederecht wetteifern, wobei die Schnelligkeit ausschlaggebend ist: "First starter acquires rights to a turn, and transfer occurs at that place" (Sacks/Schegloff/Jefferson 1974: 704).

Das Turnvergeben bzw. das Turnhalten durch die Selbstwahl nach der übergaberelevanten Stelle und die Orientierung des Hörers an ihrer Projizierung kann durch die konkrete Reaktion des Hörers, der sich das Rederecht nimmt bzw. sich das Rederecht nicht nimmt, ohne dabei Überlappungen oder relativ lange Pausen zu verursachen, bewiesen werden.

Diese Bestätigungsweise, die aus dem ethnomethodologischen Kategorisieren der Daten nach ihrer Anwendung bei Interpretation und Handeln von Seiten der Gesellschaftsmitglieder hervorgeht, ist ein wichtiger Ausgangspunkt im beschriebenen Modell der Gesprächsanalyse.

Sowie die grammatische Struktur der Turnkonstruktionseinheit unbeschrieben geblieben ist, so ist auch die intonatorische Realisation, die die Kohäsie zwischen den vor und nach den möglichen übergaberelevanten Stellen vorkommenden textuellen Elementen signalisiert, sowohl phonetisch als auch phonologisch nicht beschrieben worden. Dasselbe gilt auch für den Intonationsverlauf, der das Ende des Redebeitrags signalisiert.

Die Intonation beim Sprecherwechsel ist unter anderem Gegenstand der Arbeit der Koautorin des oben besprochenen Aufsatzes, Jefferson (1986). Sie erforscht darin die Überlappungsstellen in Gesprächen sowie die für den Sprecherwechsel relevante syntaktische, pragmatische und intonatorische Vollständigkeit. Die Beobachtungen der Intonation wurden vor den möglichen übergaberelevanten Stellen ausgeführt, nach denen der Sprecher entweder das Rederecht vergibt oder das Rederecht behält, indem er, nach Jefferson, diese zwei Vorgehen mittels Intonation ankündigt.

Die gewonnenen Prozentzahlen der Sprecherwechsel ohne Überlappung nach der Intonation, die das Turnhalten signalisiert (30% der Fälle der Sprecherwechsel ohne Überlappungen) und die Prozentzahlen der Sprecherwechselversuche mit Überlappung nach der Intonation, die das Turnvergeben signalisiert (25% der Fälle der Sprecherwechsel mit Überlappung) bewogen die Autorin jedoch dazu, die Funktion der Intonation beim Turnhalten und Turnvergeben anzuzweifeln. Problematisch scheinen aber introspektiv gewonnene Vorurteile der Autorin über die Intonationskonturen, die Turnvergeben ('completion intonation') bzw. Turnhalten ('non-completion intonation') signalisieren, was sie teilweise auch selbst einräumt:

I immediately ran into two problems. One was that I have only the grossest idea what 'completion intonation' is, or could be defended to be. But there were some contours I supposed anybody would agree to be completion or non-completion intonation, so I focussed on those. (Jefferson 1986: 180).

Die Erkenntnisse der ethnomethodologischen Konversationsanalyse wenden Local/Kelly/Wells (1986) erstmals innerhalb der Phonetik und Phonologie in ihrer Studie an, die die phonetischen Aspekte des Sprecherwechsels im Tyneside Englischen, einer Nicht-Standardvarietät des Englischen, thematisiert. In Anlehnung an J. R. Firths (1957) These der Gebundenheit der Sprache an den Situationskontext und an seine Trennung der phonologischen von phonetischen Kategorien erforschen sie die Funktionspotentiale der gesamten phonetischen Realisation im authentischen Sprachgebrauch. Die Methoden der ethnomethodologischen Konversationsanalyse anwendend legen Local/Kelly/Wells das Fundament zu einem neuen Ansatz zur phonologischen Analyse, der 'Phonologie der Konversation' ('phonology of

conversation'). Die Methoden der Konversationsanalyse auf phonologische Analysen anzuwenden, bedeutete, im Widerspruch zu damaligen Ansätzen zur Intonationsforschung im Diskurs zu stehen. Diese Ansätze stützen sich nämlich im Gegensatz zur "Phonologie der Konversation" beim Beweisvorgang im für die Zwecke der Analyse künstlich erzeugten Diskurs auf die eigene Intuition. Local/Kelly/Wells lenken außerdem die Aufmerksamkeit auf die Erforschung der kompletten phonetischen Realisation in Gesprächen. Sie erforschen nicht nur die intonatorische oder die prosodische Realisation (Lautstärke, Sprechgeschwindigkeit, Rhythmus und Pausen), sondern auch die Realisation der segmentalen, lautlichen Ebene. Sie vermeiden verfrühte 'Phonologisierung' der prosodischen Mittel: der Intonation, der Lautstärke, der Sprechgeschwindigkeit und des Rhythmus in die holistischen Akzente, indem sie gerade die prosodischen Mittel, die die Akzente ausmachen, impressionistisch notieren. Somit verschieben sie die Festlegung der phonologischen Kategorien auf die spätere Phase der Analyse.

Die phonetische Realisation wird in phonetischen Wörtern untersucht, die einer möglichen übergaberelevanten Stelle vorausgehen, an der es nicht zum Sprecherwechsel kommt, sowie in phonetischen Wörtern, die einer wirklichen übergaberelevanten Stelle mit dem Sprecherwechsel im Anschluss bzw. ohne den Sprecherwechsel im Anschluss vorausgehen.

Das Ziel ist es also, die phonetischen Realisationen, die der Sprecher zur Ankündigung des Turnhaltens bzw. des Turnvergebens verwendet, zu beschreiben. Untersucht wurden die Sprecherwechsel ohne Überlappungen, die Sprecherwechsel mit Überlappungen und das Wettstreifen um das Rederecht. Die Kriterien für die Bestimmung der unabhängigen Variable – des Turnhaltens bzw. des Turnvergebens sind die syntaktische Vollständigkeit, d. h. das Erreichen der übergaberelevanten Stelle, sowie retrospektiv die Reaktion des Gesprächspartners im Anschluss. Unter Reaktion des Gesprächspartners im Anschluss sind nicht nur das Turnhalten bzw. Turnvergeben 'in the clear' gemeint, sondern auch eine eventuelle Reaktion auf den vorgesehenen aber nicht verwirklichten Sprecherwechsel und auch eine eventuelle Reaktion auf die Unterbrechung.

Der Sprecher reagiert nämlich auf den vorgesehenen aber nicht verwirklichten Sprecherwechsel, indem er "simply re-cycles the syntax and content of the preceding talk" (Local/Kelly/Wells 1986: 433).

Local/Kelly/Wells führen implizit ein neues Kriterium bei der Steuerung des Sprecherwechsels – die pragmatische Vollständigkeit – ein. Dieses Kriterium verwenden sie ursprünglich als Beweis für die Reaktion des Sprechers auf den vorgesehenen aber nicht verwirklichten Sprecherwechsel: "We suggest this because she then goes on to make repeated attempts to give up the turn and because, as her subsequent talk shows, she had said all she wanted to say" (Local/Kelly/Wells 1986: 432).

Die Existenz bestimmter segmentaler und prosodischer Merkmale wie zum Beispiel der größeren Zentralisiertheit der Vokale, die Turnvergeben signalisieren, ist in den zwei letzten phonetischen Worten der Turnkonstruktionseinheit festgestellt worden, gleichviel ob sie die letzte Turnkonstruktionseinheit im Redebeitrag ist oder nicht, d. h. ungeachtet dessen, ob es im Anschluss an sie zum Sprecherwechsel kommt oder nicht. Diese Merkmale bleiben jedoch vor der möglichen übergaberelevanten Stelle ohne den Sprecherwechsel im Anschluss aus. Eine spezifische intonatorische Realisation des Turnvergebens konnte jedoch nicht ermittelt werden: "Unlike voiceless plosion, pitch must always be present at final position; at the same time, pitch differs from the features in (a)–(d) in that there is no unique pitch pattern associated with turn final position" (Local/Kelly/Wells 1986: 423). Es konnten schließlich zwei für das Turnvergeben charakteristische Intonationsmuster beobachtet werden: der turnfinale Tonhöhenanstieg, der größer als jeder davor im Turn realisierte Tonhöhenanstieg ist, und der

Tonhöhenabfall bis zur untersten Grenze des Sprechstimmumfangs in der letzten Silbe des Turns.

Local/Kelly/Wells meinen außerdem, dass neben den erwähnten phonetischen Merkmalen auch die Syntax, der Augenkontakt und die Gestikulation das Turnvergeben signalisieren können.

Die Beteiligung der Intonation an der Konstitution der Turnkonstruktionseinheit wird in den Arbeiten der amerikanischen Linguistinnen Ford/Thompson (1996) und Ford/Fox/Thompson (1996) behandelt, wobei neben der syntaktischen und intonatorischen Vollständigkeit auch die pragmatische Vollständigkeit zur Ermittlung der übergaberelevanten Stellen in Gesprächen herangezogen wird. Obwohl in Arbeiten, die den Sprecherwechsel oder die Konstitution der Turnkonstruktionseinheit thematisieren, außer der syntaktischen und intonatorischen Vollständigkeit auch die pragmatische Vollständigkeit als wichtig betont wird, wird sie innerhalb der Linguistik zum Gegenstand ernsthafterer und gewissermaßen umfangreicherer systematischer Auseinandersetzung zum ersten Mal in den Arbeiten von Ford/Thompson (1996) und Ford/Fox/Thompson (1996). Die Turnkonstruktionseinheit ist nach diesem Ansatz nicht nur durch die syntaktische und intonatorische Vollständigkeit gekennzeichnet, sondern auch durch die pragmatische Vollständigkeit, d. h. durch ein geäußertes Kommunikationsvorhaben: "Whether in speech act framework or, as in the present case, in the framework of conversation analysis, action completion is conceived of based on the potential that any utterance has for constituting an action in an interactional sequence." (Ford/Thompson 1996: 148). Das Kriterium für die Definition und die Ermittlung der pragmatischen Vollständigkeit ist nicht nur "complete conversational action within its specific sequential context" (Ford/Thompson 1996: 150), sondern auch ein fallender Intonationsverlauf am Ende der Äußerung. Methodische Schwierigkeit bei der Ermittlung der pragmatischen Vollständigkeit besteht im Vergleich zur Bestimmung der syntaktischen und intonatorischen Vollständigkeit, die "are easily operationalized and replicated" (Ford/Thompson 1996: 148) darin, dass die Beurteilung, wie die Autorinnen selbst anführen, in diesem Fall auf der Intuition des Analytikers gründet und dass sie provisorisch ist. Dieses Problem kann nach Ford/Thompson (1996) teilweise durch die Unterscheidung zweier pragmatischer Vollständigkeitsarten – der lokalen und der globalen, die ungefähr mit der Äußerung und dem Redebeitrag gleichgesetzt werden können, gelöst werden. Das Hauptkriterium für diese Unterscheidung stellt die Realisation zusätzlicher Rede von Seiten des aktuellen Sprechers nach dem Erreichen der Stelle der pragmatischen Vollständigkeit dar.

Die Annahme der Wichtigkeit pragmatischer Vollständigkeit für die Konstitution des Redebeitrags und die Steuerung des Sprecherwechsels hat sich als beachtenswert bewährt, weil fast alle Stellen der erreichten pragmatischen Vollständigkeit an Stellen der erreichten intonatorischen und syntaktischen Vollständigkeit ermittelt wurden. Auf diese Weise bilden sie gemeinsam mit ihnen komplexe übergaberelevante Stellen ("Complex Transition Relevance Places" abgek. CTRPs). Die überwiegende Mehrheit der Gesamtanzahl der intonatorischen bzw. der pragmatischen Vollständigkeitsarten kommt an komplexen übergaberelevanten Stellen vor, während nur die Hälfte der Gesamtanzahl der syntaktischen Vollständigkeitsarten an diesen Stellen anzutreffen sind. Dies lässt sich durch die Tatsache erklären, dass die übrigen syntaktischen Vollständigkeitsarten an möglichen übergaberelevanten Stellen, die schließlich keine übergaberelevanten Stellen geworden sind, vorkamen.

Dass die CTRPs wirklich relevant für die Konstitution des Redebeitrags und die Steuerung des Sprecherwechsels sind, zeigen die Prozentzahlen der Sprecherwechsel, die nach diesen Stellen erfolgt sind – 71% – sowie die Prozentzahlen der Sprecherwechsel mit Überlappungen vor dem Erreichen dieser Stelle – 29%. Die letzt genannte Prozentzahl bestätigt eben die vorläufige Behauptung über die Relevanz der CTRPs für den Sprecherwechsel. Diese Sprecher-

wechsel treten nämlich mit Überlappungen auf, was ein Anzeichen dafür ist, dass der Sprecher offensichtlich nicht vorhat, das Rederecht vor dem Erreichen der komplexen übergaberelevanten Stelle zu vergeben. Er kämpft um das Rederecht, indem er weiterspricht, was wiederum Überlappungen zur Folge hat.

Die Tatsache, dass es nur an ungefähr einer Hälfte der gesamten komplexen übergaberelevanten Stellen zum Sprecherwechsel kommt, mindert nicht ihre Wichtigkeit bei der Steuerung des Sprecherwechsels. Wie schon hervorgehoben wurde, werden 29% der Sprecherwechsel vor dem Erreichen der komplexen übergaberelevanten Stellen mit Überlappungen realisiert. Der Rest der komplexen übergaberelevanten Stellen ohne Sprecherwechsel kann der Selbstwahl des Sprechers nach einer abgeschlossenen Turnkonstruktionseinheit beigemessen werden, weil "transition places (here CTRPs) are option points, and that by one of the possible options, a current speaker may continue to talk." (Ford/Thompson 1996: 165). Unberücksichtigt bleibt in diesen Arbeiten jedoch die phonetische und die phonologische Struktur des fallenden Intonationsverlaufs am Ende der Äußerung, der die intonatorische Vollständigkeit und das Turnvergeben charakterisiert.

In seiner Analyse der distinktiven Intonationsmuster hinsichtlich der Steuerung des Sprecherwechsels im Englischen kommt Schegloff (1987) zum Schluss, dass im Unterschied zur syntaktischen oder pragmatischen Struktur die Prosodie das Turnvergeben bzw. das Turnhalten lokal, von der letzten akzentuierten Silbe vor der möglichen übergaberelevanten Stelle bis zur möglichen übergaberelevanten Stelle signalisiert. Dieser Bereich stellt den übergaberelevanten Raum ('transition relevance space', Schegloff 1987) dar.

Nach Schegloff (1987) hängt der Wert der Tonhöhenveränderung in der akzentuierten Silbe oder in einer benachbart an der akzentuierten Silbe liegenden Silbe im Vergleich zu Tonhöhenwerten der anderen Silben im gleichen phonetischen Wort vom Turnhalten bzw. Turnvergeben ab. So signalisiert eine relativ kleine Tonhöhenveränderung das Turnhalten, während eine relativ große Tonhöhenveränderung das Turnvergeben signalisiert.

Die intonatorische Funktion der Steuerung des Sprecherwechsels im Deutschen wird unter anderem erstmals von Selting (1995) in ihrer Studie zur Prosodie in Gesprächen der nordwestdeutschen Varietät des Deutschen untersucht. Außer der Kontextualisierung des Turnvergebens bzw. des Turnhaltens ohne Unterbrechungen und Überlappungen, wird in dieser Studie auch die Kontextualisierung der legitimen bzw. der nichtlegitimen Unterbrechungen, die Kontextualisierung des Turnhaltens bei einem Versuch des Unterbrechens und die Kontextualisierung der Hintergrundkommentare des Hörers analysiert.

Ein fallender und steigender Intonationsverlauf in unakzentuierten Silben, die der letzten akzentuierten Silbe vor der möglichen übergaberelevanten Stelle folgen, signalisieren in der nordwestdeutschen Varietät des Deutschen nach Selting das Turnvergeben, während ein gleichbleibender und ein leicht steigender Intonationsverlauf das Turnhalten signalisieren.

Die Kontextualisierungstheorie ('contextualisation', Gumperz 1999) behandelt die intonatorische Funktion der Steuerung des Sprecherwechsels als den Kontextualisierungshinweis, der die Schemata des Sprecherwechsels indiziert. Innerhalb dieser Theorie wird die Intonation unter anderen prosodischen und nicht-prosodischen Mitteln nämlich als Kontextualisierungshinweis aufgefasst (engl. 'contextualisation cue'). Kontextualisierungshinweise indizieren vorwiegend zusammen mit Kontextualisierungshinweisen anderer Kontextualisierungstypen (Kinetik, Proxemik, Blickverhalten, Varietäten-/Sprachwahl, lexikalische Variation, Variation der sprachlichen Formulierung und andere prosodische Kontextualisierungshinweise) in Kontextualisierungsverfahren Schemata, die die Struktur des ganzen Wissensinventars des Menschen ausmachen. Auf diese Weise wird das Geäußerte einem bestimmten Kontext zugeordnet, was dem Hörer das Verstehen ermöglicht. Im Gegensatz zu sprachlichen Zeichen haben

die Kontextualisierungshinweise keine inhärente Bedeutung, vielmehr ist ihre Bedeutung "conveyed as part of the interactive process" (Gumperz 1999: 131). Ihre Bedeutung ist ein holistisches Produkt gemeinsamen Indizierens der Schemata mittels mehrerer Kontextualisierungshinweistypen.

Auer (1986: 27) unterscheidet in seiner Ausführung der Kontextualisierungstheorie (Auer/Di Lucio 1992; Auer 1996; Auer 1986) Schemata auf fünf Ebenen: 1. die generellen Schemata des fokussierten Interagierens (üben die kommunikative Funktion, die als Beantwortung der Frage: Reden wir (gerade) miteinander? aufgefasst werden könnte), 2. die Schemata des Sprecherwechsels (Wer spricht gerade mit wem?), 3. Handlungsschemata (Was tun wir gerade?), 4. thematische Schemata (Worüber sprechen wir gerade?), 5. Beziehungsschemata (Wie stehen wir gerade zueinander?).

Der Kontextualisierungshinweis kann im Kontextualisierungsverfahren auf mehreren Ebenen Schemata indizieren, was er vorwiegend zusammen mit anderen Kontextualisierungshinweistypen erzielt.

4 Das Dekodieren der Intonation hinsichtlich der Funktion der Steuerung des Sprecherwechsels

Couper-Kuhlen und Selting (1996: 28–30) postulieren in Anlehnung an die Methoden der Konversationsanalyse den speziellen Umgang mit Daten bei der Erforschung der Prosodie in Gesprächen. Gespräche werden nämlich interaktiv hergestellt, was die Möglichkeit des Entscheidens und des Interpretierens dieser Entscheidungen im Laufe der noch andauernden Einheiten voraussetzt:

The notion of emergence also transfers to prosody: just as a multi-unit turns, in the default case, not produced as a monolithic, preconceived whole but grows incrementally and anew at each transition relevance place (TRP) as the result of an interactive achievement, so prosodic units carried by turns or turn-constructive units may emerge as a contingency of the moment-to-moment decisions speakers make to continue or not and, if so, *how* to continue. (Couper-Kuhlen/Selting 1996: 29f.).

Der in zahlreichen konversationsanalytischen Studien bezeugte Sprecherwechselmechanismus setzt die Möglichkeit des Dekodierens der Intonation nicht nur nach vollständig realisierten Intonationsphrasen voraus, sondern auch vor dem Erreichen der möglichen übergaberelevanten Stelle, falls eine in der Intonationsphrase vorhanden ist. Nach Schegloff (1987) beginnt das Dekodieren der intonatorischen Realisation hinsichtlich der Funktion der Steuerung des Sprecherwechsels von der letzten akzentuierten Silbe vor der möglichen übergaberelevanten Stelle. Die Intonation legt also die Stellen der möglichen Satzenden als übergaberelevante Stellen und Enden der Turnkonstruktionseinheiten fest, oder signalisiert an diesen Stellen die Kohäsie zwischen den vor und nach diesen Stellen vorkommenden Strukturen der textuellen Schicht der Rede.

Die konturbasierten Modelle gehen von phonologisch relevanten Einheiten aus, deren Gestalt aufgrund der diesen Einheiten immanenten Tonhöhenänderungen festgelegt wird. Die Referenzwerte für die Wahrnehmung dieser Teilverläufe werden also im Laufe ihrer Verwirklichung zugänglich. Das Intonationsmodell für das Deutsche von von Essen (1964) lässt die Möglichkeit für die Bestimmung der Melodeme gleich nach der ersten Nachlaufsilbe zu, weil die zwei von insgesamt drei vorgestellten Intonationstypen durch zwei unterschiedliche Tonlagen der ersten Nachlaufsilbe im Vergleich zur Tonlage der Nucleussilbe charakterisiert sind. Ist der Ton der ersten Nachlaufsilbe tiefer, dann liegt der terminale Intonationstyp vor, während bei höherem Ton der ersten Nachlaufsilbe ein interrogativer Intonationstyp vorliegt.

Der progrediente Intonationstyp ist durch das Ausbleiben der Nucleussilbe gekennzeichnet.

Im Falle eines einsilbigen Nachlaufs oder einer Intonationsphrase ohne Nachlauf, in der dann die Tonbewegung in der Nucleussilbe relevant wird, beginnt der übergaberelevante Raum jedoch erst mit der letzten Silbe der Intonationsphrase.

Bei konturbasierten Modellen, die mit etwas komplexeren Tonbewegungen des Nukleus oder der nuklearen Kontur operieren wie z. B. mit der fallend-steigenden oder steigend-fallenden, bleibt die Möglichkeit der relativ frühen Identifikation ihrer Gestalt, wie sie bei von Essen bei mehrsilbigen Nachläufen anzutreffen ist, aus.

Das superpositionale Fujisaki-Intonationsmodell zur Sprachsynthese (Fujisaki 1983) behandelt die Intonationskontur einer Äußerung als eine komplexe Funktion, die in kleinere konstitutive Funktionen zerlegt werden kann. Die zwei Komponenten, die Phrasen- und die Akzentkomponente, werden in diesem Modell der Intonationskontur zugrunde gelegt. Sie stellen die Reaktionen auf diskrete Phrasen- bzw. Akzentkommandos im Prozess der Generierung dar.

Das Akzentkommando ist als eine Stufenfunktion gestaltet, die die lokalen Tonanstiege und Tonabfälle in akzentuierten Silben erzeugt. Das Phrasenkommando ist hingegen als eine Impulsfunktion konzipiert, die einen zum Tonhöhengipfel ansteigenden und danach allmählich zum niedrigsten Wert der Grundfrequenz, F_{min} , abfallenden globalen Intonationsverlauf gestaltet. Für die Phrasierung der kleineren syntaktischen Einheiten in der Äußerung werden weitere Phrasenkommandos mit einer Stufenfunktion angenommen. Zusammengefügt ergeben die Akzent- und die Phrasenkomponenten die konkrete Intonationskontur der Äußerung.

Dieses Modell impliziert, dass der niedrigste Tonhöhenwert, den ein Sprecher realisieren kann, das F_{min} , vom Hörer erfasst werden kann, ohne dass er auf die zuvor realisierten Tonhöhenwerte des Intonationsverlaufs Bezug nehmen muss. Im Experiment mit drei unterschiedlich langen japanischen Aussagesätzen (Fujisaki 1983), die von einem Sprecher geäußert wurden, betragen die Werte des F_{min} jeweils 76, 83 und 84 Hz. Da diese Werte vom Hörer erst in der letzten Silbe der Intonationsphrase wahrgenommen werden können, müsste im Hinblick auf die intonatorische Funktion der Steuerung des Sprecherwechsels präzisiert werden, bis zu welchem relativen Wert die Intonationskontur der Äußerung abfallen muss, um die nächste mögliche übergaberelevante Stelle als Turnende vorauszubestimmen.

Das Modell zur automatischen Analyse und Synthese der Intonation, TILT (Taylor 2000), betrachtet die Intonationskontur als eine Abfolge intonatorischer Ereignisse, die durch die dazwischen liegenden interpolierten Tonhöhenwerte verbunden werden. Es werden zwei Typen von intonatorischen Ereignissen unterschieden, die Akzent- und die Grenztöne, deren Form durch steigende und/oder fallende Tonbewegungen gekennzeichnet ist. Die Anteile der steigenden bzw. der fallenden Tonbewegung in der gesamten Form des intonatorischen Ereignisses werden durch den zwischen +1 und -1 liegenden Wert notiert, wobei +1 einen reinen Tonanstieg und -1 einen reinen Tonabfall bezeichnet.

In diesem Modell wird die Form der intonatorischen Ereignisse aufgrund der ihnen immanenten Tonhöhenänderungen festgelegt, so dass auf Stimmumfang kein Bezug genommen wird.

Die Möglichkeit, das Turnende vorauszubestimmen, ist jedoch in diesem Modell dem Hörer erst mit dem in der letzten Silbe der Intonationsphrase vorkommenden Ton, dem Grenzton gegeben.

Das Kieler Intonationsmodell, KIM (Kohler 1990), gestaltet die Intonationskontur als eine Aufeinanderfolge holistischer intonatorischer Einheiten, der Tonhöhengipfel und der Tonhöhentäler der akzentuierten Silben. Tonhöhengipfel ('peaks') sind charakterisiert durch einen steilen Tonhöhenanstieg, während die Tonhöhentäler ('valleys') eher sukzessive Tonhöhenanstiege aufweisen. Die sukzessiven Tonhöhenbewegungen können hoch oder leicht ansteigen, je nachdem, ob sie Frage oder Weiterweisung signalisieren. Die prosodischen Grenzsignale werden in diesem Modell abgestuft, wofür drei phonologisch relevante Stufen vorgesehen

sind. Die am Wortende vorkommenden F0-Endpunkte können entweder die Endpunkte eines größeren (notiert mit 2) oder mittleren Tonabfalls (notiert mit 1) oder einer Tonbewegung, die keinen Tonabfall aufweist (notiert mit 0), darstellen.

Innerhalb dieses Modells müsste hinsichtlich der intonatorischen Funktion der Steuerung des Sprecherwechsels präzisiert werden, welche Stufe des F0-Endpunktes in einem, dem letzten phonetischen Wort vor der möglichen übergaberelevanten Stelle vorausgehenden phonetischen Wort erreicht werden muss, damit die folgende mögliche übergaberelevante Stelle als Turnende vorausbestimmt werden kann. Dafür müsste die Möglichkeit der Erweiterung der Stufen der F0-Endpunkte offen gehalten werden.

Das autosegmental-metrische Transkriptionssystem für das Deutsche, GToBI, geht von der Festlegung der relativen Tonhöhenwerte der Töne (hoch bzw. tief) größtenteils nach ihrer Position im Stimmumfang aus, der in einer Intonationskontur genutzt wird. Die Position innerhalb der oberen drei Viertel des Stimmumfangs ist nämlich die Position des H-Tones, die innerhalb des unteren einen Viertels des Stimmumfangs die Position des L-Tones.

Die relativen Tonhöhenwerte der Töne hängen nach diesen Modellen unmittelbar von höchsten und von tiefsten Tonhöhenwerten im Stimmumfang des Sprechers ab, die auch nach der möglichen übergaberelevanten Stelle selbst in letzten zwei Silben der Intonationsphrase realisiert werden können.

Nach Gussenhoven (2004: 76f.) haben die zweiten Tonakzente mit höchsten Tonhöhenwerten in einer Intonationsphrase keine Auswirkung auf die schon vorher getroffene Wahl des Stimmumfangs der Äußerung: "[...] the second may also be spoken with the substantially higher peak than the first, without changing the phonological identity of the contour." Für diese Behauptung werden aber keine Beweise geliefert.

Die phonetische Realisierung und das Dekodieren der Töne im Bezug auf die Referenzwerte, die erst nach der Realisierung der ganzen Intonationsphrase zugänglich sind, zu akzeptieren, würde bedeuten, dass der Hörer die Intonationskontur nicht dekodiert, bevor sie vollständig verwirklicht ist. Dies führt zum Schluss, dass der Hörer nur retrospektiv das Turnvergeben des Sprechers erkennen kann, d. h. erst nachdem es der Sprecher schon realisiert hat. Es wäre also unmöglich, das Turnvergeben bzw. das Turnhalten des Sprechers vor dem Erreichen der möglichen übergaberelevanten Stelle mittels Intonation zu projizieren. Demzufolge wäre die interne Struktur der Intonationskontur vor ihrem Ende für den Hörer nicht erkennbar und somit zum Projizieren nicht fähig, was die Gesprächsteilnehmer daran hindern würde, den Kommunikationsfluss zu lenken:

Communication without projection would be restricted to behavioural segments which are either independent events, or chained to each other as stimulus-response sequences, beyond the control of a speaker and recipient. (Auer 2002: 2)

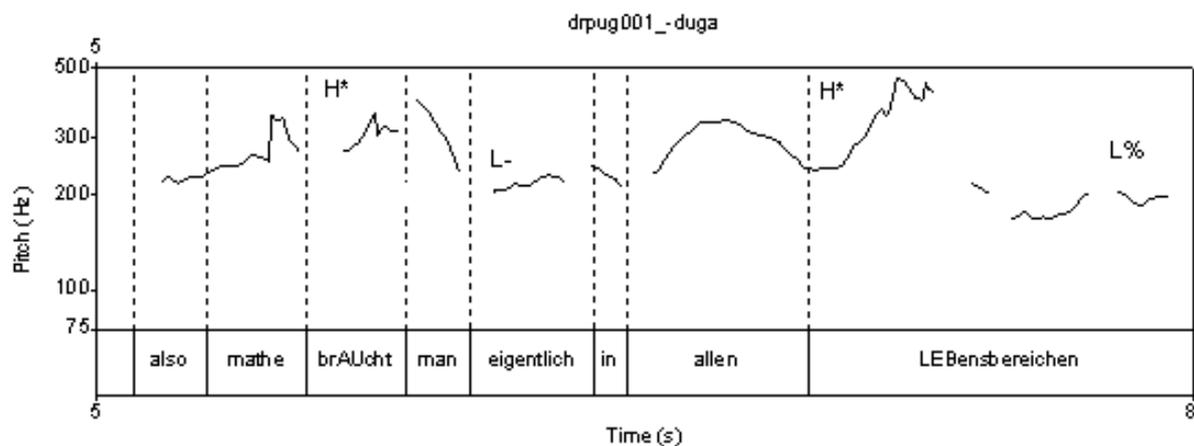


Abb. 1 Beispiel einer Intonationsphrase mit der Erweiterung des Stimmumfangs in ihrem letzten phonetischen Wort in einem Diagramm mit Grundfrequenzänderungen

Auf Abbildung 1. ist ein Beispiel einer Intonationsphrase dargestellt, innerhalb derer es nach der möglichen übergaberelevanten Stelle zur Erweiterung des Stimmumfangs kommt. Dies kommt dadurch zustande, dass der höchste Grundfrequenzwert in der Intonationsphrase in der Silbe *leb* erreicht wird, und der tiefste in der Silbe *reich*. Angenommen, dass die intonatorische Realisation vor der möglichen übergaberelevanten Stelle, d. h. vor dem Ende des möglichen Satzes – *Mathe braucht man* dekodiert wird, wäre der Tonakzent in der Silbe *braucht* nach seiner Position innerhalb der Viertel des vor der möglichen übergaberelevanten Stelle realisierten Stimmumfangs als L+H* Tonakzent dekodiert worden, weil er sich zum Teil im ersten Viertel des Stimmumfangs befindet.

Nachdem die Intonationsphrase und zugleich die Turnkonstruktionseinheit vollständig realisiert wurden, befindet sich dieser Tonakzent nicht mehr innerhalb des unteren Viertels des Stimmumfangs, sondern innerhalb des dritten Viertels von unten, was dazu führt, dass dieser Tonakzent diesmal als H* festgelegt wird.

Demzufolge ist die Skalierung der Töne der ganzen Intonationsphrase vor ihrem Ansetzen vom Gesichtspunkt des Sprechers denkbar und unproblematisch, was aber für das Dekodieren der Intonationskontur erst nach ihrer vollständigen Realisation nicht behauptet werden kann. Eine solche Dekodierungsweise geht von der Signalisierung des Turnvergebens bzw. des Turnhaltens vor der möglichen übergaberelevanten Stelle mittels Intonation nicht aus, obwohl diese Funktion der Intonation in mehreren Sprachen, unter anderen auch im Deutschen, bewiesen wurde.

Eine Lösung bietet sich in Intonationsmodellen an, die im Hinblick auf die intonatorische Funktion der Steuerung des Sprecherwechsels von Referenzwerten für das Festlegen bzw. das Dekodieren der Töne ausgehen, die für den Hörer spätestens von der letzten akzentuierten Silbe vor der möglichen übergaberelevanten Stelle zugänglich sind. Dass diese Werte wirklich Referenzwerte für die phonetische Realisierung und das Dekodieren der Töne darstellen, müsste natürlich durch weitergehende Analysen bewiesen werden.

Die oben analysierte Intonationsphrase im Ausschnitt aus einem Gespräch zwischen dem Hörer und dem Moderator einer Radio-Telefonsendung transkribiert gemäß der Notationskonvention des GATs – Gesprächsalytischen Transkriptionssystems (Selting et al. 1998).

Drpug001; Deutschlandradio, Deutschlandradio Kultur; Was erwarten Sie vom Jahr der Mathematik? 5:15, 2:35-2:53; Moderator, Anruferin.

```

01 S: und äh also ↑EIGentlich macht `mAthe spaß;=
02 =und öh ich ↓FINde auch öh mann `kAnn in- (-)
→ 03 also mathe `brAUcht man eigentlich in allen `LEBensbereichen; (--
04 aso mathe kann man `Immer(.)`IMmer <<dim>ver`wEndn>;
05 V: ein plädo`yER für die ↑mathema`tIk,=
06 =<<all> ich bedanke mich ganz ↑`hErzlich bei `IHnen,=>
07 =<<all> und viel spaß noch beim `HOBByrechnen,=>
08 =<<all> sagma ma ↑`SO? >
09 S: `JA: (h) [danke,
10 V: [schönen `dAnk für ihren `ANruf,

```

5 Fazit

Die Überprüfung des Modells der phonetischen Realisierung und des Dekodierens der intonatorischen Realisation nach dem Transkriptionssystem für das Deutsche, GToBI, nach dem die H- und L-Töne vorwiegend nach ihrer Position im Stimmumfang, der in einer Intonationskon-

turn genutzt wird, festgelegt werden, führte zum Schluss, dass eine solche Generierungsweise unakzeptabel ist, wenn die Tatsache berücksichtigt wird, dass der Stimmumfang des Sprechers erst nach der vollständig realisierten Intonationsphrase festgelegt werden kann. Diese Generierungsweise der Intonation impliziert, dass der Hörer nur retrospektiv das Turnhalten bzw. das Turnvergeben erkennen kann. Demzufolge ist es nicht möglich, den Sprecherwechsel vor dem Erreichen der möglichen übergaberelevanten Stelle zu steuern. Dies ist aber nach Befunden der Intonationsforschung, die auf Erkenntnissen der Konversationsanalyse gründet, durchaus möglich.

Eine Lösung wäre die Annahme der Referenzwerte für die Festlegung der Töne in Intonationsmodellen, die für den Hörer spätestens von der letzten akzentuierten Silbe vor der möglichen übergaberelevanten Stelle zugänglich sind.

Literatur

- Auer, Peter (1986): "Kontextualisierung". *Studium Linguistik* 19: 22–47.
- Auer, Peter (1992): "Introduction: John Gumperz' Approach to Contextualization". In: Auer, Peter/Di Lucio, Aldo (eds.): *The Contextualization of Language*. Amsterdam, John Benjamins 1–38.
- Auer, Peter (1996): "From Context to Contextualization". *Links & Letters 3: Pragmatics*: 11–28.
- Auer, Peter (2002): "Projection in Interaction and Projection in Grammar". *Interaction and Linguistic Structures (InLiSt)* 33.
- Beckmann, Mary E./Pierrehumbert, Janet (1986): "Intonational Structure in Japanese and English". *Phonology Yearbook* 3: 255–309.
- Beckmann, Mary E./Hirschberg, Julia/Shattuck-Hufnagel, Stefanie (2006): "The Original ToBI System and the Evolution of the ToBI Framework". In: Jun, Sun-Ah (ed.): *Prosodic Typology. The Phonology of Intonation and Phrasing*. Oxford, Oxford University Press: 9–54.
- Couper-Kuhlen, Elizabeth/Selting, Margret (1996): "Towards an Interactional Perspective on Prosody and a Prosodic Perspective on Interaction". In: Couper-Kuhlen, Elizabeth/Selting, Margret (ed.): *Prosody in Conversation. Interactional Studies*. Cambridge, Cambridge University Press: 11–56.
- Cruttenden, Alan (1997): *Intonation*. 2nd ed. Cambridge: Cambridge University Press.
- Essen, Otto von (1964): *Grundzüge der hochdeutschen Satzintonation*. 2nd ed. Ratingen/Düsseldorf: A. Henn Verlag.
- Firth, John R. (1957): *Papers in Linguistics 1934–1951*. London: Oxford University Press.
- Ford, Cecilia/Fox, Barbara A./Thompson, Sandra A. (1996): "Practices in the Construction of Turns: The 'TCU' Revisited". *Pragmatics* 6/3: 427–454.
- Ford, Cecilia E./Thompson, Sandra A. (1996): "Interactional Units in Conversation. Syntactic, Intonational, and Pragmatic Resources for the Management of Turns". In: Ochs, Elinor/Schegloff, Emanuel A./Thompson, Sandra A. (eds.): *Interaction and Grammar*. Cambridge, Cambridge University Press: 134–184.
- Fujisaki, Hiroya (1983): "Dynamic Characteristics of Voice Fundamental Frequency in Speech and Singing". In: MacNeilage, Peter F. (ed.): *The Production of Speech*. New York, Springer: 39–55.
- Gilles, Peter (2005): *Regionale Prosodie im Deutschen. Variabilität in der Intonation von Abschluss und Weiterweisung*. Berlin/New York: de Gruyter. (= *Linguistik – Impulse & Tendenzen* 6).
- Grice, Martine/Benzmüller, Ralf (1995): "Transcription of German Intonation Using ToBI-Tones – The Saarbrücken System". *Phonus* 1: 33–51.

- Grice, Martine/Baumann, Stefan (2002): "Deutsche Intonation und GToBI". *Linguistische Berichte* 191: 267–298.
- Grice, Martine/Baumann, Stefan/Benzmüller, Ralf (2005): "German Intonation in Autosegmental-Metrical Phonology". In: Jun, Sun-Ah (ed.): *Prosodic Typology. The Phonology of Intonation and Phrasing*. Oxford, Oxford University Press: 55–83.
- Gumperz, John (1999): *Discourse strategies*. Cambridge: Cambridge University Press.
- Jefferson, Gail (1986): "Notes on 'Latency' in Overlap Onset". *Human Studies* 9: 153–183.
- Kohler, Klaus (1990): "Macro and Micro F0 in the Synthesis of Intonation". In: Kingston, John/Beckman, Mary E. (eds.): *Papers in Laboratory Phonology I*. Cambridge, Cambridge University Press: 115–138.
- Ladd, Robert D. (1996): *Intonational Phonology*. Cambridge: Cambridge University Press.
- Local, John K./Kelly John/Wells, William H. G. (1986): "Towards a Phonology of Conversation. Turn-taking in Tyneside English". *Journal of Linguistics* 22: 411–437.
- Pierrehumbert, Janet B. (1980): *The Phonology and Phonetics of English Intonation*. Dissertation: Bloomington.
- Sacks, Harvey/Schegloff, Emanuel A./Jefferson, Gail (1974): "A Simplest Systematics for the Organization of Turn-Taking for Conversation". *Language* 50: 696–735.
- Schegloff, Emanuel A. (1987): "Analyzing Single Episodes of Interaction. An Exercise in Conversation Analysis". *Social Psychological Quarterly* 50/2: 101–114.
- Selting, Margret (1995): *Prosodie im Gespräch. Aspekte einer interaktionalen Phonologie der Konversation*. Tübingen: Niemeyer. (= *Linguistische Arbeiten* 329).
- Selting, Margret et al. (1998): "Gesprächsanalytisches Transkriptionssystem (GAT)". *Linguistische Berichte* 173: 91–122.
- Taylor, Paul (2000): "Analysis and Synthesis of Intonation Using the Tilt Model". *Journal of the Acoustical Society of America* 107/3: 1697–1714.