Linguistische Arbeiten 439

Herausgegeben von Hans Altmann, Peter Blumenthal,
Hans Jürgen Heringer, Ingo Plag, Heinz Vater und Richard Wiese

Stefan Rabanus

Intonatorische Verfahren im Deutschen und Italienischen

Gesprächsanalyse und autosegmentale Phonologie

Max Niemeyer Verlag
Tübingen 2001

Die Deutsche Bibliothek – CIP-Einheitsaufnahme

Rabanus, Stefan: Intonatorische Verfahren im Deutschen und Italienischen: Gesprächsanalyse und autosegmentale Phonologie / Stefan Rabanus. – Tübingen : Niemeyer, 2001
(Linguistische Arbeiten ; 439)
Zugl.: Greifswald, Univ., Diss., 2000

ISBN 3-484-30439-1 ISSN 0344-6727

© Max Niemeyer Verlag GmbH, Tübingen 2001
Das Werk einschließlich aller seiner Teile ist urheberrechtlich geschützt. Jede Verwertung außerhalb der engen Grenzen des Urheberrechtsgesetzes ist ohne Zustimmung des Verlages unzulässig und strafbar. Das gilt insbesondere für Vervielfältigungen, Übersetzungen, Mikroverfilmungen und die Einspeicherung und Verarbeitung in elektronischen Systemen. Printed in Germany.
Gedruckt auf alterungsbeständigem Papier.
Gesetzt mit LaTeX 2^{ε}
Druck: Weihert-Druck GmbH, Darmstadt
Einband: Industriebuchbinderei Nädele, Nehren

Danksagung

Die vorliegende Intonationsstudie ist von September 1996 bis September 1999 im Rahmen des interdisziplinären Projekts „Neue Wege der Intonationsforschung“ an der Universität Greifswald und während eines längeren Studienaufenthaltes am Istituto di Fonetica e Dialettologia (IFD) des Consiglio Nazionale delle Ricerche in Padua entstanden. Mein Dank gilt deshalb zunächst allen Kolleginnen und Kollegen, die dort zum Gelingen der Untersuchung beigetragen haben, besonders Cinzia Avesani, Christoph Bandt, Piero Cosi, Christina Gansel, Frank Jürgens, Roland Kehrein, Emanuela Magno Caldognetto, Bernd Pompe, Gisela Ros und Jürgen Erich Schmidt, meinem Doktorvater. Für wertvolle Hinweise und kritische Gutachten zur Arbeit danke ich Hans Altmann (München), Harald Thun (Kiel) und Richard Wiese (Marburg). Anregende Gespräche konnte ich mit Peter Auer, Peter Gilles (beide Freiburg), Jörg Peters und Margret Selting (beide Potsdam) führen, sowie mit vielen Mitgliedern der Gruppe des Freiburger Pragmatischen Kolloquiums und des Gruppo di Fonetica Sperimentale der Associazione Italiana di Acustica. Bei der Transkription sehr umgangs- oder regionalsprachlicher Passagen in den italienischen Gesprächen haben mir Edda Farnetani (Padua), Anna dal Negro (Crema) und Carmela Scilipoti (Reggio Calabria) geholfen. Für Hilfe bei analyse- und programmiertechnischen Problemen bin ich Paul Boersma (Amsterdam), Christian Rabanus (Hannover) und Markus Walther (Marburg) zu Dank verpflichtet. Korrekturgelesen haben Sylvia Setzkorn, Gabriele Diederich-Schmidt und meine Eltern.

Viele Personen, von denen ich Anregungen oder Hilfe erfahren habe, sind hier umgenannt geblieben, vor allem diejenigen, die anonym bleiben müssen: meine Gesprächsteilnehmer. Unter ihnen befindet sich mancher gute Freund und manche gute Freundin, deren Bedeutung für mich weit über den wissenschaftlichen Bereich hinausgeht.

Marburg an der Lahn, September 2000

Inhaltsverzeichnis

Abkürzungen XI

Transkription XIII

Gesprächsteilnehmer XV

1 Einleitung 1
1.1 Ziele 2
1.2 Korpus 3
1.3 Gliederung 3

2 Phonologie und Phonetik der Intonation 5
2.1 Prosodische Merkmale und Einheiten 5
2.2 Silbe 9
2.2.1 Die Silbe als zentrale Analysekategorie 9
2.2.2 Die Isochronie-Hypothese 13
2.3 Intonationsmodelle 15
2.3.1 Das Intonationsmodell der Britischen Schule 16
2.3.1.1 Intonationsgruppen und ihre Begrenzung 16
2.3.1.2 Akzenttöne und Nukleus 18
2.3.2 Das Intonationsmodell der autosegmentalen Phonologie 21
2.3.2.1 Metrische Komponente 24
2.3.2.2 Intonationskomponente: Das Pierrehumbert-Modell . 29
2.3.2.3 ToBI 33
2.3.2.4 Funktionen der Intonation 34
2.3.3 Theorie und Empirie: Die Modelle im Vergleich 36
2.4 Akzentstruktur und Fokussierung 39
2.5 Phonetik der Intonation 43
2.5.1 Akustik der Intonation 44
2.5.2 Perzeption der Intonation 45
2.5.3 Artikulation der Intonation 47
2.5.4 Das IPO-Intonationsmodell 49
2.5.5 Wichtige phonetische Phänomene 50
2.5.5.1 Deklination 50
2.5.5.2 Downstep 54
2.5.5.3 Alinierung 54
2.5.5.4 Register und Umfang 55
2.5.5.5 Durchschnittswert 56
2.6 Beschreibungskategorien 56
2.6.1 Toninventar 56
2.6.2 Aufnahme- und Messtechnik 59

3 Intonation im Deutschen und Italienischen 63
3.1 Nicht-intonatorische prosodische Merkmale 63
3.2 Untersuchungen der Intonation des Deutschen 64
3.2.1 Forschung der 50er/60er Jahre 65
3.2.1.1 Otto v. Essen 65
3.2.1.2 Isačenko und Schädlich 66
3.2.2 Grammatisch orientierte Intonationsforschung 68
3.2.2.1 John Pheby 68
3.2.2.2 Das Münchner Projekt von Hans Altmann 70
3.2.3 Das „Kiel Intonation Model“ 75
3.2.4 Konversationsanalytische Intonationsforschung 79
3.2.5 Autosegmentale Intonationsforschung 80
3.2.5.1 Dieter Wunderlich 80
3.2.5.2 Susanne Uhmann 81
3.2.5.3 Caroline Féry 82
3.2.6 Zwischenbilanz 84
3.3 Untersuchungen der Intonation des Italienischen 85
3.3.1 Britische Schule 86
3.3.1.1 Giulio Lepschy 86
3.3.1.2 Luciano Canepari 88
3.3.1.3 Amedeo De Dominicis 91
3.3.2 Autosegmentale Intonationsforschung 93
3.3.2.1 Martine Grice 93
3.3.2.2 Cinzia Avesani 95
3.3.3 Zwischenbilanz 97
3.4 Sprechstile: gelesen vs. spontansprachlich 98
3.5 Kontrastive Untersuchungen 100
3.5.1 Italienisch – Englisch 100
3.5.2 Italienisch – Französisch 101
3.5.3 Deutsch – Französisch 101
3.5.4 Englisch – Deutsch – Italienisch 102
3.5.5 Resümee 103

4 Gesprächsanalytische Intonationsforschung 105
4.1 Methodische Grundlegung 105
4.1.1 System des Verhaltens in Gesprächen 105
4.1.2 Äußerungspaare 106
4.1.3 Das „richtige Korpus“ 109
4.1.3.1 Experiment 109
4.1.3.2 Teilnehmende Beobachtung 111
4.1.3.3 Voruntersuchung: Interview 111
4.1.4 Vineta-Korpus: Die Gespräche im Varietätengefüge 112
4.1.4.1 Gesprächssituation, Diaphasik, Diastratik 112
4.1.4.2 Diatopie 114
4.1.5 Transkription 118

4.2 Funktionskategorien . . . 120
4.2.1 Ethnomethodologische Konversationsanalyse . . . 121
4.2.1.1 Aushandlungsprozesse in Gesprächen . . . 122
4.2.1.2 Untersuchungsfelder der Konversationsanalyse . . . 125
4.2.2 Sprechakttheoretische Dialoganalyse . . . 128
4.2.2.1 Verlaufsorientierte Beschreibung . . . 129
4.2.2.2 Musterorientierte Beschreibung . . . 130
4.2.3 Dialoganalyse und/oder Konversationsanalyse? . . . 134
4.2.4 Verhaltenskategorien: Übersicht . . . 137
4.3 Intonatorische Verfahren . . . 139

5 Empirische Untersuchung . . . 145
5.1 Einzelsprachliche Untersuchungen . . . 145
5.1.1 Selbstinitiierte Selbstreparaturen . . . 145
5.1.1.1 Selbstinitiierte Selbstreparaturen im Deutschen . . . 145
5.1.1.2 Selbstinitiierte Selbstreparaturen im Italienischen . . . 155
5.1.2 Bearbeitung akustischer Verstehensprobleme . . . 158
5.1.2.1 Akustische Verstehensprobleme im Deutschen . . . 158
5.1.2.2 Akustische Verstehensprobleme im Italienischen . . . 162
5.1.3 Lokale Erwartungsprobleme . . . 166
5.1.3.1 Lokale Erwartungsprobleme im Deutschen . . . 166
5.1.3.2 Lokale Erwartungsprobleme im Italienischen . . . 169
5.1.4 Turnbeanspruchung . . . 174
5.1.4.1 Turnbeanspruchung im Deutschen . . . 174
5.1.4.2 Turnbeanspruchung im Italienischen . . . 176
5.1.5 Turnverteidigung . . . 182
5.1.5.1 Turnverteidigung im Deutschen . . . 182
5.1.5.2 Turnverteidigung im Italienischen . . . 186
5.1.6 Widerspruch . . . 198
5.1.6.1 Widerspruch im Deutschen . . . 198
5.1.6.2 Widerspruch im Italienischen . . . 210
5.1.7 Insistieren . . . 218
5.1.7.1 Insistieren im Deutschen . . . 218
5.1.7.2 Insistieren im Italienischen . . . 228
5.2 Vergleich . . . 237
5.2.1 Form-Funktions-Zuordnungen . . . 237
5.2.1.1 Konversationelle Reparaturen . . . 237
5.2.1.2 Kampf um das Rederecht . . . 238
5.2.1.3 Inhaltlicher Dissens . . . 240
5.2.2 Funktions-Form-Zuordnungen . . . 240

6 Zusammenfassung und Ausblick . . . 243

Bibliographie . . . 247

Sachregister . . . 266

Abkürzungen

A.I.A.	*Associazione Italiana di Acustica*
dB	Dezibel
C	Strukturstelle auf der Skelett-Ebene: „Konsonant“
Durchschnittswert	durchschnittliche F_0 des tiefen Onsets eines Sprechers
F_0	Grundfrequenz
FHG	Fokus-Hintergrund-Gliederung
GAT	Gesprächsanalytisches Transkriptionssystem
GToBI	*German Tone and Break Indices*
H*	hoher Akzentton
H-	hoher Phrasenton
H%	hoher Grenzton
Hz	Hertz
IdS	Institut für deutsche Sprache
IFD	*Istituto di Fonetica e Dialettologia*
IPA	*International Phonetic Alphabet*
IPO	*Instituut voor Perceptie Onderzoek*
KIM	*Kiel Intonation Model*
λ	Lambda-Operator
MIT	*Massachusetts Institute of Technology*
ms	Millisekunde
σ	Silbe
SK	Steigungskoeffizient
ST	*Semitone*, Halbton
T* (L*)	tiefer Akzentton
T- (L-)	tiefer Phrasenton
T% (L%)	tiefer Grenzton
ToBI	*Tone and Break Indices*
ToBIt	*Tone and Break Indices italiani*
V	Strukturstelle auf der Skelett-Ebene: „Vokal“

Transkription

Die Transkription der Gespräche erfolgt in Blöcken aus je drei Zeilen. Die Basistranskription in der ersten Zeile richtet sich nach den Konventionen des Gesprächsanalytischen Transkriptionssytems (GAT). Die Feintranskription der Intonationskontur wird mit dem in Kap. 2.6.1 eingeführten Toninventar in der zweiten Zeile vorgenommen. Akzenten, denen kein Ton sicher zugewiesen werden kann, wird in dieser Zeile das Zeichen »?« zugeordnet. In der dritten Zeile werden – soweit messbar – die Hz-Werte der Zielpunkte der Akzent- und Grenztonbewegungen angegeben. Der Wert des „frühen Gipfels" in H+T* wird in Klammern notiert. Für Äußerungen des teilnehmenden Beobachters SF, die in der Intonationsanalyse nicht berücksichtigt werden, erfolgt weder Feintranskription, noch Messung. Wenn die akustische Qualität der Äußerungen nicht ausreicht, fallen diese Zeilen auch bei den Redebeiträgen anderer Interaktanten weg.

In der Arbeit werden folgende GAT-Konventionen verwendet:

`Wo[rt` `[Wort`	Überlappung, Simultansprechen von Wörtern
`Turn=` `=Turn`	schneller Anschluss eines Turns an einen anderen
`(.)`	kurze Pause, kleiner als 0,2 s
`(-)`	Pause von 0,2 bis 0,5 s
`(2.0)`	gemessene Pause in s
`:, :: etc.`	Dehnung, je nach Länge, z.B. `so::`
`'`	Abbruch durch Glottisverschluss, z.B. `ich de'`
`.hh, .hhh`	Einatmen, je nach Dauer
`hh, hhh`	Ausatmen, je nach Dauer
`((lacht))`	Deskription von Nonverbalem
`so(h)o`	Lachpartikel beim Reden
`hm, ja, äh, eh etc.`	einsilbige Diskurspartikeln
`hm=m, ja=a etc.`	zweisilbige Diskurspartikeln
`'hm'hm`	redupliziertes Signal mit Glottisverschlüssen
`((...))`	Auslassung im Transkript
→	Verweis auf im Text behandelte Transkriptzeilen
`( )`	unverständliche Passage
`(solche)`	vermuteter Wortlaut
`al(s)o`	nicht mit Sicherheit identifizierbare Laute/Silben
`akZENT`	prominentester Akzent der IP, „Hauptakzent" (Akzentsilbe in Majuskeln)
`akzEnt`	anderer Akzent, „Nebenakzent" (Vokal(e) der Akzentsilbe als Majuskel(n))
`ak!ZENT!`	extra starker Akzent

?	finale Tonhöhenbewegung: hoch steigend
,	mittel steigend
-	gleichbleibend
;	mittel fallend
.	tief fallend
↑	Tonhöhensprung nach oben
↓	Tonhöhensprung nach unten
`<<t>...>`	(global) tiefes Tonhöhenregister
`<<h>...>`	(global) hohes Tonhöhenregister
`<<f>...>`	Lautstärke: forte
`<<p>...>`	piano
`<<all>...>`	Sprechgeschwindigkeit: allegro, schnell
{okay}	Äußerung, die vom jeweiligen Standard abweicht und dialektalen oder fremdsprachigen Aussprachekonventionen folgt (in GAT nicht vorgesehen)

Gesprächsteilnehmer

Die Tabelle informiert über die Sozialdaten der Sprecher (Geschlecht, Muttersprache, Alter, Beruf, Heimatort) und die Durchschnittswerte (Ø) des tiefen Offsets ihrer Intonationsphrasen: n_1 = Gesamtzahl der Äußerungen mit messbarem tiefen Offset, n_2 = längere Äußerungen, n_3 = Kurzäußerungen/Diskurspartikeln.

Sigle		n_1	Ø	n_2	Ø	n_3	Ø
BB	weibl., dt., 23 Studentin, Bergen/Rügen	27	205 Hz	18	197 Hz	9	220 Hz
SK	weibl., dt., 30 wiss. Mitarbeiterin, Köln	77	179 Hz	57	178 Hz	20	181 Hz
RM	weibl., it., 40 Lehrerin, Mailand	58	173 Hz	16	166 Hz	19	188 Hz
RR	weibl., it., 27 Lehrerin, Reggio C.	17	173 Hz	16	173 Hz	1	175 Hz
MR	weibl., it., 30 Lehrerin, Reggio C.	14	171 Hz	14	171 Hz		
SC	weibl., it., 48 Hausfrau, Colico (LC)	36	159 Hz	30	158 Hz	6	164 Hz
NH	weibl., dt., 27 Studentin, Hochheim/Main	90	153 Hz	65	149 Hz	25	163 Hz
MM	weibl., it., 26 Studentin, Messina/Reggio C.	9	132 Hz	6	126 Hz	3	144 Hz
PA	männl., dt., 23 Student, Anklam/Vorpommern	19	122 Hz	16	118 Hz	3	144 Hz
IH	männl., dt., 27 Student, Hückelhoven (Kreis Heinsberg/NRW)	20	110 Hz	15	104 Hz	5	128 Hz
GP	männl., it., 59 Restaurator, Palermo	11	104 Hz	7	101 Hz	4	109 Hz
JB	männl., dt., 30 wiss. Mitarbeiter, Bielefeld	44	102 Hz	30	103 Hz	14	100 Hz
KB	männl. dt., 24 Student, Bansin/Usedom	9	100 Hz	7	102 Hz	2	95 Hz
IC	männl., it., 50 Handwerker, Colico (LC)	52	96 Hz	43	96 Hz	9	98 Hz

1 Einleitung

Intonation, der Gebrauch des Tonhöhenverlaufs, ist ein Phänomen, dessen Erforschung umfassende Möglichkeiten des Verstehens von Sprechen und Kommunikation eröffnet. An der Erforschung der Intonation wird stärker als an anderen Phänomenen deutlich, dass Kommunikation über ein Zusammenwirken sprachlicher und nichtsprachlicher Signalisierungssysteme funktioniert. Die Bedeutung der Intonation für die Kommunikation lässt sich nur bestimmen, wenn gleichzeitig die anderen Signalisierungssysteme - allen voran Grammatik und Lexik - berücksichtigt werden. In der Intonationsforschung wird außerdem evident, dass die phonetischen Merkmale der Rede, die auf Artikulation, Transmission und Perzeption des Sprachsignals beruhen, Produkte entsprechender (sprach-)wissenschaftlicher Untersuchungsmethoden sind und nicht die *Phänomene an sich* abbilden. Im Mittelpunkt einer Studie, die die Zuordnung von Formen zu Funktionen in natürlicher Interaktion zum Thema hat, steht die Perzeption: Funktion hat nicht das, was der Analytiker misst, sondern das, was die Interaktionsteilnehmer hören.

Aufgrund der vielen Schwierigkeiten, die der Analyse in dieser Situation entstehen, ist das Wissen über das Phänomen Intonation trotz einer über hundertjährigen Forschungsgeschichte und zahlreicher großer Projekte in den letzten Jahren noch immer gering. Die vorliegende Studie versteht sich deshalb als Beitrag zur Grundlagenforschung. Sie versucht, die analysepraktischen Schwierigkeiten zu überwinden, indem sie die Interferenzen der anderen Signalisierungssysteme nicht als „Störungen" auffasst, die es zu beseitigen gelte. Die Komplexität wird, im Gegenteil, als Chance begriffen, an der Intonation beispielhaft das Funktionieren von Sprechen und Kommunikation als Ergebnis des Zusammenspiels verschiedener Signalisierungssysteme zu zeigen. Aus diesem Grund wird die Intonation in ihrer natürlichen Umgebung untersucht: in ungesteuerten Alltagsgesprächen. Mit der Herauslösung der zu untersuchenden Phänomene und der Versetzung in die künstliche Umgebung des Experiments würden die interaktiven Funktionen verlorengehen, die untersucht werden sollen.

Mit der Verbindung von Gesprächsanalyse und Intonationsforschung wird daher einem Desiderat entsprochen, das trotz erster konversationsanalytischer Arbeiten in dieser Richtung noch weitgehend unerfüllt ist. Die vorliegende Studie setzt das in Rabanus (1998) skizzierte Programm um. Sie ist den methodischen Prinzipien der ethnomethodologischen Konversationsanalyse verpflichtet.

Die Analyse wird an einem Korpus aus deutschen und italienischen Gesprächen vorgenommen. Zur Intonation dieser beiden bedeutenden europäischen Sprachen liegt bisher keine umfassendere kontrastive Studie vor. Der Blick auf zwei Sprachen gibt außerdem Aufschluss darüber, was von der These zu halten ist, dass die Intonation ein universales, übereinzelsprachliches Signalisierungssystem ist.

1.1 Ziele

Die Untersuchung orientiert sich an den folgenden vier Leitfragen:

1. Welche *Funktionen* erfüllt Intonation in der konversationellen Interaktion und welche Funktionskategorien sind für die kontrastive Analyse geeignet?
2. Wie lässt sich Intonation *formal* beschreiben und welchen Anforderungen müssen die Beschreibungskategorien in der kontrastiven Analyse genügen?
3. Wie sind die Formen mit den Funktionen *gekoppelt*?
4. Welche *Unterschiede* in der Form-Funktions-Zuordnung bestehen zwischen Deutsch und Italienisch?

Die Funktionskategorien, die gleichzeitig die Vergleichskategorien der kontrastiven Analyse bilden, sind einem abstrakten und übereinzelsprachlichen *System des Verhaltens in Gesprächen* entnommen (Kap. 4.1.1). In Fortführung von Arbeiten der Konversations- und Dialoganalyse werden sieben Funktionskategorien als Teil dieses Systems definiert (Kap. 4.2): Selbstinitiierte Selbstreparatur, Bearbeitung akustischer Verstehensprobleme, lokales Erwartungsproblem, Turnbeanspruchung, Turnverteidigung, Widerspruch, Insistieren. Der Ausdruck der syntaktischen und diskurssemantischen Funktionen der Intonation, die in den Laboruntersuchungen bevorzugt betrachtet werden (Kap. 3), tritt in der konversationellen Interaktion hinter den Ausdruck diese Funktionen zurück.

Die formale Beschreibung der Intonationskonturen erfolgt im Rahmen des Intonationsmodells der autosegmental-metrischen Phonologie (Kap. 2.3.2) mit einer an die Erfordernisse dieser Studie angepassten Fassung des Transkriptionssystems ToBI (Kap. 2.6.1). ToBI wird als System zur auditiven phonetischen Beschreibung von Tonhöhenverläufen rekonstruiert. Phonologische Einheiten (Phänomene, die das Kriterium der phonologischen Distinktivität erfüllen) sind hinsichtlich der betrachteten Funktionskategorien allerdings nicht die Töne von ToBI, sondern *intonatorische Verfahren* (Kap. 4.3), mit denen Konturen zum Zweck der Unterscheidung der genannten Funktionskategorien bearbeitet werden. Intonationskonturen werden als Ressource verstanden, mit der sich die Interaktionsteilnehmer non-verbal über ihre lokalen Handlungsziele verständigen.

Welche Zuordnungsverhältnisse zwischen Formen (intonatorischen Verfahren) und Funktionen (Kategorien aus dem System des Verhaltens in Gesprächen) bestehen, wird zunächst für Deutsch und Italienisch getrennt untersucht (Kap. 5.1). Anschließend werden die Ergebnisse der einzelsprachlichen Untersuchungen einander gegenübergestellt und interpretiert (Kap. 5.2).

1.2 Korpus

Die empirische Basis der Untersuchung ist ein Korpus aus drei deutschen und drei italienischen Alltagsgesprächen aus den Jahren 1997 und 1998. Um Verwechslungen mit den Korpora referierter Untersuchungen zu vermeiden, habe ich mein Korpus „Vineta-Korpus" genannt.[1] Die deutschen Gespräche wurden in Greifswald und Mainz aufgezeichnet, die italienischen in Reggio di Calabria und in Colico am Comer See. Die Anzahl der Teilnehmer pro Gespräch variierte zwischen drei und fünf Personen, womit abgesehen von mir selbst – ich habe als teilnehmender Beobachter an allen Gesprächen teilgenommen – sieben deutsche und sieben italienische Muttersprachler aktiv an den Gesprächen beteiligt waren. An den italienische Gesprächen nahmen mehr Frauen als Männer teil, in den deutschen ist das Geschlechterverhältnis zahlenmäßig ausgeglichen. Alle deutschen Sprecher waren zum Zeitpunkt der Aufnahmen zwischen 23 und 30 Jahren alt und gehörten als Studenten oder wissenschaftliche Mitarbeiter dem akademischen Milieu an. Auch die italienischen Sprecher waren größtenteils Akademiker, nur ein Gespräch fand im Haus eines Handwerkerehepaars statt. Das Alter der italienischen Muttersprachler lag zur Aufnahmezeit zwischen 27 und 59 Jahren. Die italienischen Sprecher repräsentieren die größere Bandbreite an sozialen Beziehungen. Regionalsprachliche Variation gibt es sowohl in den deutschen als auch in den italienischen Gesprächen. Allerdings orientieren sich alle Sprecher an der Standardnorm. Äußerungen im Basisdialekt kommen nur in Gesprächen homogener Gruppen mit regionalsprachlicher Kompetenz vor (in zwei italienischen Gesprächen) und werden von den Teilnehmern als Abweichungen von der Norm kontextualisiert.

Das Korpus wird ausführlich in Kap. 4.1.4 beschrieben. Die Transkription von insgesamt 46 Minuten aus den Gesprächen ist in der Digitalen Bibliothek Marburg, online im Internet unter http://archiv.ub.uni-marburg.de/sonst/2000/0003/ [Stand: 23.09.2000] abgelegt. Die Tonaufnahmen der in Kap. 5 untersuchten Äußerungspaare sind dort mit der Transkription verknüpft und können angehört werden.

1.3 Gliederung

In Kap. 2 werden Terminologie und Begrifflichkeit der Intonationsforschung eingeführt und erläutert. In den Kapiteln 2.1 bis 2.4 werden die phonologischen Aspekte der Intonationsforschung diskutiert. Dazu zählen die Unterscheidung von prosodischen Merkmalen und Einheiten, sprachtypologische Fragen hinsichtlich Silbe und rhythmischer Struktur, Intonationsmodelle und der Zusammenhang von Akzentstruktur und Fokussierung. Kap. 2.5 beleuchtet die phonetischen Aspekte der Intonationsforschung und führt die Intonation als Untersuchungsgegenstand der akustischen, auditiven und

[1] *Vineta* ist der Name einer sagenhaften, reichen und prächtigen Stadt, die vor der Küste der Insel Usedom, 50 Kilometer von Greifswald entfernt, auf dem Grund der Ostsee liegen soll.

artikulatorischen Phonetik vor. In Kap. 2.6 werden das für diese Studie entworfene Toninventar sowie die Aufnahme- und Messtechnik vorgestellt.

In Kap. 3 wird die Forschungsgeschichte anhand der wichtigsten Arbeiten zur Intonation im Deutschen und im Italienischen referiert. Weil sich die vorliegende Arbeit – bei aller Exaktheit der phonetischen Analysen – als phonologische Grundlagenforschung versteht, werden hier zwar auch experimentell-phonetische und anwendungsorientierte Untersuchungen referiert, der Schwerpunkt liegt aber auf sprachwissenschaftlichen Studien, in deren Zentrum der Form-Funktions-Zusammenhang steht. Kap. 3.2 referiert in einem Forschungsüberblick ausgewählte Untersuchungen der deutschen Intonation. Die Geltung der dort postulierten Form-Funktions-Zuordnungen wird im Hinblick auf konversationelle Kontexte relativiert. In Kap. 3.3 werden analog dazu Untersuchungen der italienischen Intonation referiert. In Kap. 3.4 wird gezeigt, dass nicht nur die Form-Funktions-Zuordnungen, sondern auch die phonetischen Merkmale von gelesenen Sätzen und spontansprachlichen Äußerungen differieren. Merkmale, die das lautliche Erscheinungsbild des Deutschen von dem des Italienischen unterscheiden, ohne intonatorisch zu sein, werden in Kap. 3.1 aufgelistet. Kap. 3.5 stellt einige Ansätze zur kontrastiven Intonationsanalyse vor.

In Kap. 4 wird der Ansatz zur *gesprächsanalystischen Intonationsforschung* entwickelt. Er ist Konsequenz der Überlegungen hinsichtlich der Eignung der in Kap. 3 vorgestellten Modelle für die (kontrastive) Analyse der Intonation im Gespräch. In Kap. 4.1 wird die Methode begründet. Kap. 4.2 expliziert die Funktionskategorien vor dem Hintergrund von Konversations- und Dialoganalyse. In Kap. 4.3 werden die intonatorischen Verfahren eingeführt.

Kap. 5 enthält die empirische Untersuchung. In Kap. 5.1 wird die Analyse, nach Funktionskategorien geordnet, für Deutsch und Italienisch getrennt durchgeführt. In Kap. 5.2 werden die Ergebnisse der einzelsprachlichen Untersuchungen zusammengefasst, verglichen und interpretiert.

In Kap. 6 werden die in Kap. 1 gestellten Leitfragen abschließend beantwortet.

2 Phonologie und Phonetik der Intonation

In den folgenden Kapiteln werden die phonetischen und phonologischen Konzepte der Intonationsforschung diskutiert. Die Phonologie (Kap. 2.1 bis 2.4) wird mit einer doppelten Begründung vor der Phonetik (Kap. 2.5) behandelt: Einerseits werden perzipierte phonetische Merkmale vom Sprachteilhaber direkt und unreflektiert funktional interpretiert. Andererseits finden phonetische Untersuchungen im Rahmen der Sprachwissenschaft immer vor dem Hintergrund einer bestimmten phonologischen Theorie oder zur Überprüfung einer phonologischen Hypothese statt. Beides spricht dafür, die phonologische Aspekte vor den phonetischen zu behandeln. In Kap. 2.6 werden die in Auseinandersetzung mit den phonetischen und phonologischen Aspekten entwickelten Beschreibungskategorien erläutert: das Toninventar und die Art und Weise, wie die Zuweisung der Töne durch phonetische Messwerte gestützt wird.

2.1 Prosodische Merkmale und Einheiten

> Prosodie wird verstanden als Oberbegriff für diejenigen suprasegmentalen Aspekte der Rede, die sich aus dem Zusammenspiel der akustischen Parameter Grundfrequenz (F_0), Intensität und Dauer in silbengroßen oder größeren Domänen ergeben.[1]

Dieses Zitat gibt die in der Forschung weithin akzeptierte Definition für sprachliche Phänomene an, die oberhalb der Ebene der Lautsegmente angesiedelt sind. Mit dem Terminus *Prosodie* wird heute wieder auf all jene Phänomene referiert, die man in den 70er und 80er Jahren des 20. Jahrhunderts als *Suprasegmentalia* bezeichnet hat.[2] Bezugseinheit für die in dieser Arbeit behandelten prosodischen Phänomene ist die *Silbe*. Zur Systematisierung der zahlreichen prosodischen Phänomene wird hier die auf Kohler und Schmidt zurückgehende Unterscheidung von prosodischen bzw. suprasegmentellen *Merkmalen* und prosodischen bzw. suprasegmentellen *Einheiten* aufgegriffen.[3]

---

[1] Selting (1995), S. 1.

[2] Vgl. Schmidt (1986), S. 17, Fn. 4. *Prosodie* wird vorgezogen unter anderem von Zifonun et al. (1997), Bd. 1, S. 189 in der Grammatikschreibung, Wiese (1996), S. 26 in der generativen Phonologie, Cruttenden (^{2}1997), S. 1 in der Phonologie der Britischen Schule, Bertinetto/Magno Caldognetto (1993), S. 143 in der italienischen Terminologie und Selting (1995), S. 1 wie zitiert in der Konversationsanalyse. In der generativen Phonologie ist diese Präferenz mit der Einführung der autosegmentalen Phonologie begründet, siehe S. 22.

[3] Kohler (^{1}1977), S. 118f. oder (^{2}1995), S. 110 spricht nur von *prosodischen Merkmalen* und verzichtet auf einen eigenen Terminus für prosodische Merkmale mit sprachlicher Funktion. Schmidt (1986), S. 16-38 schließt an Kohlers Überlegungen an und führt den Terminus *suprasegmentelle Einheiten* für die suprasegmentellen bzw. prosodischen Merkmale mit sprachlicher Funktion ein. *Prosodische Einheit* entspricht in etwa *Prosodem*, vgl. dazu z.B. Hammarström (1963).

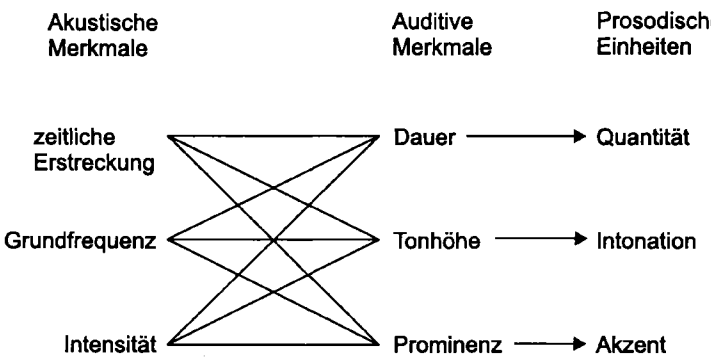

Abb. 2.1: Der artikulatorische Zusammenhang

Prosodische Merkmale sind phonetische Phänomene, die in einem Dreiebenenmodell als auditive, akustische und artikulatorische Merkmale beschrieben werden können. Dazu zählen:

> 1. die auditiven Merkmale Dauer, Tonhöhe und Prominenz, 2. die akustischen Merkmale zeitliche Erstreckung, Grundfrequenz und Intensität und 3. die artikulatorischen Merkmale zeitliche Steuerung der Artikulationsbewegung, Schwingungsverhalten der Stimmlippen, Erzeugung und Zustand des Ausatmungsluftstroms, soweit sie nicht intrinsische Eigenschaften kleinster syntagmatischer Segmente sind.[4]

Prosodische Einheiten sind dagegen phonologische Phänomene. Von prosodischen Einheiten spricht man dann, wenn die Phänomene das Kriterium der phonologischen Distinktivität erfüllen, das heißt, wenn sie sprachliche Funktionen distinguieren. Prosodische Einheiten unterscheiden eine Fülle von syntaktischen, semantischen und pragmatischen Funktionen. Kohler unterscheidet ausgehend von den auditiven prosodischen Merkmale drei prosodische Einheiten, nämlich *Quantität*, *Akzent* und *Intonation*:

> Neben die phonetische D a u e r tritt die phonologische Q u a n t i t ä t, neben die phonetische P r o m i n e n z der phonologische W o r t - und S a t z a k z e n t und neben die phonetische T o n h ö h e der phonologische T o n und die phonologische I n t o n a t i o n.[5]

Prosodische Einheiten sind auditive prosodische Merkmale mit sprachlicher Funktion. Zwischen prosodischen Einheiten und akustischen Merkmalen gibt es dagegen keine eineindeutige Beziehung. Die akustischen Merkmale stehen in einem komplexen artikulatorischen Zusammenhang. Das bedeutet, dass mit einer prosodischen Erscheinung (und einem auditiven Merkmal) jeweils ein ganzes Bündel akustischer Merkmale

---

[4] Schmidt (1986), S. 17f.

[5] Kohler (^{2}1995), S. 110. Gibbon (1995), S. 458 unterscheidet in seinem Modell die linguistischen Domänen *word level* mit den Einheiten *tone*, *stress* und *length* und *supra-word level* mit *intonation*, *accent* und *rhythm*.

korreliert.[6] Abb. 2.1 verdeutlicht den artikulatorischen Zusammenhang. Besonders augenfällig ist der artikulatorische Zusammenhang beim Akzent. Es ist umstritten, welches der akustischen Merkmale Intensität, Grundfrequenz und zeitlicher Erstreckung primär ist bzw. ob es überhaupt ein primäres Merkmal gibt.[7] Aber auch für die Intonation besteht ein artikulatorischer Zusammenhang.

In der vorliegenden Arbeit werden alle prosodischen Merkmale auf die Kategorien Quantität, Akzent und Intonation bezogen. Auf eine ihnen nebengeordnete „Restkategorie" für Phänomene wie Sprechtempo, Rhythmus, Stimmqualität und Pausen, wie sie Möbius (1993, S. 9) vorschlägt, wird verzichtet. Sprechtempo, Rhythmus und Pausen gehen partiell in die Analyse der metrischen Struktur ein. Die Funktionen von Stimmqualität und Klangfarbe werden hier nicht untersucht.

Die prosodische Einheit *Intonation* steht im Mittelpunkt der vorliegenden Untersuchung. In der Forschungsliteratur kommen unterschiedliche Konzeptualisierungen von *Intonation/intonazione* vor.[8] *Intonation* wird mit dem auditiven Merkmal 'Tonhöhe' gleichgesetzt,[9] als 'Intonation im weiteren Sinn' und damit als Synonym zu *Prosodie* verstanden[10] oder als prosodische Erscheinung wie oben ausgeführt neben Akzent und Quantität gestellt.[11] In der vorliegenden Untersuchung wird Intonation in der letztgenannten Konzeptualisierung verwendet: Intonation ist der „Gebrauch des Melodieverlaufs" und korreliert mit der Tonhöhe (auditiv) und dem Schwingungsverhalten der Stimmlippen am Kehlkopf (artikulatorisch).[12] Unter den korrelierenden akustischen Merkmalen kommt der Grundfrequenz besondere Bedeutung zu. Intonation meint den Melodieverlauf der Gesamtäußerungen und beschränkt sich nicht auf die letzte Tonhöhenbewegung.

Intonation wird in der autosegmentalen Phonologie durch Akzent-, Phrasen- und Grenztöne auf der Ton-Ebene repräsentiert. Die Intonationskontur wird als Summe lokaler Töne aufgefasst. Die in Kap. 3 referierten Studien zeigen, welche syntaktischen, diskurssemantischen und einstellungsbezogenen Funktionen durch Töne unterschieden werden können. Die in dieser Arbeit untersuchten Funktionen aus dem *System des Verhaltens in Gesprächen* werden allerdings nicht durch einzelne Töne oder Intonationskonturen ausgdrückt, sondern durch *intonatorische Verfahren.*[13] „Gebrauch des Tonhöhenverlaufs" heißt deshalb auch: 'Verwendung eines bestimmten intonatorischen Verfahrens'.

Lexikalische Töne oder Tonakzente kommen in den untersuchten Varietäten des Deutschen und Italienischen nicht vor.[14]

---

[6] Die Ergebnisse des Greifswalder Projekts zeigen, dass für die Perzeption von Funktionen Merkmalskombinationen signifikanter sind als Einzelmerkmale, vgl. Bandt et al. (2001).

[7] Siehe unten, S. 8f.

[8] Zur Begrifflichkeit in der englischsprachigen Debatte vgl. Ladd (1996), 6ff.

[9] Z.B. von Selting (1995), S. 1.

[10] Z.B. von Altmann et al. (1989), S. 2; Canepari (1985), S. 31; Pheby (1981), S. 839.

[11] Z.B. von Kohler (^{2}1995), S. 121 und Dominicis (1992), S. VIII.

[12] Vgl. im Detail Kap. 2.5.

[13] Siehe dazu Kap. 4.2 und 4.3.

[14] Zu Tonakzenten in deutschen Dialekten vgl. Schmidt (1986).

Auch *Akzent* wird in der Forschung unterschiedlich konzeptualisiert. Die wichtigste Unterscheidung ist die zwischen *Wortakzent* und *Satzakzent*.[15] In der vorliegenden Studie wird Akzent als Satzakzent oder, zutreffender, als *Äußerungsakzent* verstanden und folgendermaßen definiert: Akzent ist die Hervorhebung einer oder mehrerer sprachlicher Einheiten (Silbe, Wort, Konstituente) durch phonetische Mittel. Primäres auditives Merkmal des Akzents ist die Prominenz. Für die akustischen Merkmale des Akzents gilt das oben zum artikulatorischen Zusammenhang Ausgeführte ganz besonders: Zur Prominenzsignalisierung können alle phonetischen Merkmale beitragen. Der (Äußerungs-)Akzent ist durch kommunikative Erfordernisse (zum Beispiel Markierung der Informationsstruktur) gesteuert.[16]

Auch der *Wortakzent* zeichnet sich durch auditive Prominenz aus. Weil der Wortakzent aber weitgehend unabhängig von kommunikativen Erfordernissen ist und stattdessen nach abstrakten Regeln zugewiesen wird,[17] sind akustische Korrelate nicht in jedem Fall nachweisbar. Für die Perzeption von Wortakzenten mag in manchen Fällen das Sprachwissen ausreichend sein.[18] Wortakzente werden in der vorliegenden Untersuchung nicht empirisch untersucht.[19]

Dem Deutschen wird traditionell ein auf Grundfrequenzvariation gegründeter melodischer Akzent zugeschrieben.[20] Neuere Untersuchungen sehen dagegen in der zeitlichen Erstreckung der prominenten Silbe das primäre akustische Merkmal des Akzents.[21] Unter bestimmten Bedingungen (zum Beispiel beim Flüstern) ist Prominenzsignalisierung durch Grundfrequenzvariation sogar unmöglich.[22] Auch im Italienischen wird der Akzent primär durch zeitliche Erstreckung realisiert.[23] Unabhängig davon, welchen Stellenwert die Grundfrequenzvariation für die Prominenzwahrnehmung hat, lassen meine Analysen die Feststellung zu, dass auf Akzenten in den allermeisten Fällen Grundfrequenzvariation stattfindet.

In der autosegmentalen Phonologie wird die Akzentstruktur auf der metrischen Ebene repräsentiert. Die metrische Analyse[24] ist notwendige Voraussetzung für die Analyse der Intonationsstruktur, weil nur Akzenten (und Grenzstellen) Töne zugewiesen werden.

---

[15] Vgl. Kohler (^{2}1995), S. 114-120.

[16] Vgl. Kap. 2.4.

[17] Schmidt (1986), S. 27 spricht deshalb vom „normativen Wortakzent".

[18] In der englischsprachigen Literatur wird deshalb im Anschluß an Bolinger (1972b), S. 22 zwischen realisiertem *accent* und abstraktem *stress* unterschieden, wobei *stress* notwendige Bedingung für *accent* ist. Vgl. Ladd (1996), S. 48f.; Uhmann (1991), S. 21f.; Bertinetto (1981), S. 50-53. Anders konzeptualisiert Cruttenden (^{2}1997), S. 13.

[19] Vgl. aber die Ausführungen zu den Wortakzentregeln in Kap. 2.3.2, S. 26ff.

[20] Vgl. Isačenko/Schädlich (^{2}1971), S. 20ff.

[21] Vgl. Dogil (1999), S. 291-299, bes. S. 292f. und Jessen et al. (1995), passim. Kontrovers dazu Möbius (1993), S. 10-16. Kohler (1991), S. 298-305 unterscheidet im „Kiel Intonation Model" einen auf zeitlicher Erstreckung basierenden von einem auf Grundfrequenzvariation basierenden Akzent. Siehe Kap. 3.2.3, S. 75ff.

[22] Vgl. Kohler (^{2}1995), S. 114.

[23] Vgl. Voghera (1992), S. 96 und den umfassenden Überblick in Bertinetto (1981), S. 41-90.

[24] Vgl. Kap. 2.3.2, S. 24ff.

Quantität ist eine prosodische Einheit, deren primäres auditives Merkmal die Dauer ist. Im Standarddeutschen können Daueruntersch iede als redundante Merkmale von Vokalqualitäten aufgefasst werden können.[25] Im Standaritalienischen und in verschiedenen deutschen Dialekten sind Quantitäten dagegen prosodische Einheiten, weil sie nicht nur die Vokale, sondern die gesamte Silbe betreffen.[26]

Quantitäten spielen in der vorliegenden Untersuchung nur eine untergeordnete Rolle in der metrischen Analyse. Sie werden auf der CV-Ebene wiedergegeben, die in den Analysen nicht systematisch berücksichtigt wird. Die Intonation wird direkt an die Silbenstruktur angebunden.

2.2 Silbe

2.2.1 Die Silbe als zentrale Analysekategorie

Die Bezugseinheit prosodischer Merkmale und Einheiten ist in dieser Arbeit die *Silbe*. Der Vorschlag der (linearen) generativen Phonologie der „Sound Pattern of English", die Silbe als Trägerin prosodischer Eigenschaften abzuschaffen, indem in die Merkmalskomplexe von Lautsegmenten auch abstrakte Akzentstärken (*stress*) integriert werden, ist nicht zuletzt von der späteren generativen Phonologie selbst zurückgewiesen worden.[27] Zahlreiche phonotaktische Regeln beziehen sich auf die Silbe. Auch die Regeln zu Akzentverteilung und Tonassoziation lassen sich mit Bezug auf die Silbe einfacher formulieren als mit Bezug nur auf Lautsegmente und morphologische Grenzsymbole.[28] Außerdem besitzt die Silbe für die meisten Muttersprachler mentale Realität, was sich darin ausdrückt, dass den meisten Muttersprachlern die Zerlegung von Äußerungen in Silben auch ohne linguistische Kenntnisse keine Schwierigkeiten bereitet.[29] Die Silbe ist also eine perzeptiv relevante Einheit und eine wichtige phonologische Kategorie. Im generativen Modell von Nespor und Vogel (1986, S. 11) ist die Silbe eine von sieben prosodischen Konstituenten in hierarchischer Stufung:

---

[25] Vgl. Schmidt (1986), S. 22-24.

[26] Bannert (1976), S. 25 belegt, dass die Domäne der Quantität im Mittelbairischen „aus der Sequenz von betontem Vokal und dem folgenden Konsonanten besteht." Es sind nur die kontrastierenden Sequenzen 'Langvokal + Kurzkonsonant' und 'Kurzvokal + Langkonsonant' möglich. Im Mittelfränkischen sind vokalische und konsonantische Dauer unter Tonakzenten komplementär, vgl. Schmidt (1986), S. 185-191. Zum Standarditalienischen vgl. Solari (1997), S. 225.

[27] Der Verzicht auf die Silbe ist auch in den „Sound Pattern of English" nicht vollkommen, weil Chomsky und Halle (1968), S. 354 für Vokale das Merkmal [+syllabic] anstelle von [+vocalic] ansetzen. Siehe dazu in der vorliegenden Arbeit auch S. 21f.

[28] Vgl. Auer (1994), S. 56.

[29] Vgl. Wiese (1996), S. 33. Selbst in der exotischen Sprache Fijian kommt „syllabic oral spelling" vor, vgl. Blevins (1995), S. 209f. Interessant ist auch das Ergebnis einer Untersuchung von Burani und Cafiero (1991), die für die Silbe im Italienischen eine Rolle bei der Erkennung graphisch repräsentierter Wörter nachweist.

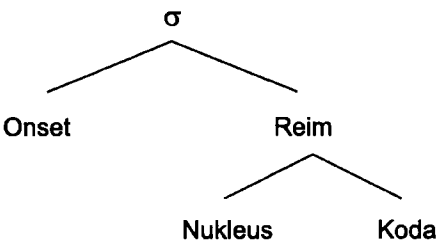

Abb. 2.2: Konstituentenstruktur der Silbe

> These seven units, from large to small, are: the phonological utterance (U), the intonational phrase (I), the phonological phrase (ϕ), the clitic group (C), the phonological word (ω), the foot (Σ) und the syllable (σ).[30]

Als phonetische Einheit ist die Silbe problematisch, weil Silbengrenzen exakte Korrelate weder in akustischer noch in artikulatorischer Hinsicht haben.[31]

Für die Silbe (symbolisiert durch $\gg\sigma\ll$) wird in der generativen Phonologie eine Konstituentenstruktur angenommen. Für Englisch, Deutsch und Italienisch gilt das in Abb. 2.2 dargestellte Modell.[32] Wiese (1996, S. 44) ergänzt dieses Modell um eine für das Deutsche gültige Skelett-Ebene. Die Annahme eines Silbenmodells mit einer

---

[30] Vgl. dazu auch Nespor (1999), S. 117-126. Die Konstituente *Fuß* lässt sich im Deutschen mit Gesetzmäßigkeiten in der Wortbildung und der Distribution von [ʔ] begründen, vgl. Wiese (1996), S. 56-61. Darüber hinaus ist der Fuß eine wichtige metrische Kategorie in den sog. akzentzählenden Sprache, siehe Kap. 2.2.2, S. 13. Das *phonologische Wort* – nach Auer (1994), S. 71 zusammengesetzt „aus dem Stamm, den Präfixen, den meisten Suffixen (ausgenommen *-heit*, *-keit*, *-bar*, *-lich* etc.) und den enklitischen Erweiterungen" – wird als Domäne der Silbifizierung, zahlreicher phonologischer Regeln (z.B. der Assimilation) und des Wortakzents angesehen, vgl. Wiese (1996), S. 65ff. Auer (1994), S. 76 hält das phonologische Wort für die prosodische Hauptkategorie des Standarddeutschen. Die Bedeutung der *klitischen Gruppe* ist umstritten, die der *phonologischen Phrase* im Deutschen als Domäne für bestimmte Akzentverschiebungen sehr begrenzt. Vgl. Wiese (1996), S. 74-77 zur phonologischen Phrase im Deutschen, und Nespor (1985), passim zur phonologischen Phrase im Italienischen. Die *More* (μ) ist nach einer traditionellen Auffassung – vgl. Auer (1991), S. 10 – für das Standard-Neuhochdeutsche empirisch nicht gerechtfertigt und wird auch für das Italienische selten operationalisiert. In modernen autosegmentalen Ansätzen zeichnet aber eine Änderung dieser Auffassung ab. Van der Hulst (1999), S. 10-14 verwendet die More als universales Konzept zur Bestimmung des Silbengewichts. Peters (in Vorb.) unterscheidet durch Zerlegung von Silben in Moren unterschiedliche Realisierung der Fokusintonation in verschiedenen regionalen Varietäten des Deutschen.

[31] Vgl. dazu den Forschungsüberblick von Bertinetto (1981), S. 148-156. Heike (1992), S. 26 schreibt: „Zusammenfassend kann man sagen, daß die Silbe sich im (natürlichen oder simulierten) lautsprachlichen Produktionsprozess mit dem zeitlichen Erstreckungsbereich (ko-)artikulativer Steuerungsprozesse deckt [...], daß sie aber entgegen einer diskret-segmentalen Vorstellung keine scharfen Grenzen hat."

[32] Graphik nach Eisenberg/Ramers/Vater (1992), S. V. Zum Englischen vgl. Blevins (1995), S. 213, zum Italienischen Nespor (1993), S. 156 und Mioni (1993), S. 128, zum Deutschen Vater (1992), S. 100.

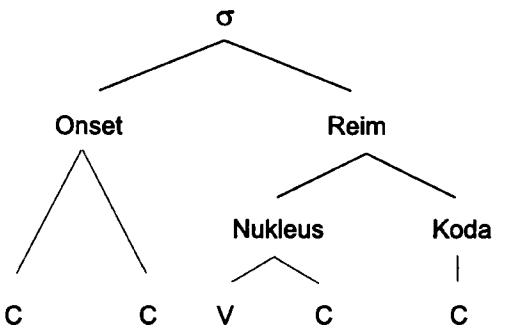

Abb. 2.3: Konstituentenstruktur der Silbe und CV-Struktur

Konstituentenstruktur – gegenüber einem „flachen" Modell ohne Konstituenten – ist sinnvoll, weil sich viele phonotaktische und phonologische Regeln auf genau eine dieser Konstituenten beziehen (Abb. 2.3).[33] So sind zum Beispiel homorgane Obstruent-Sonorant-Cluster wie [pm], [fm], [kŋ] im Onset ausgeschlossen, während sie in der Koda als [mp], [mf], [ŋk] auftreten.[34] Nur im Reim findet dagegen die Vokalisierung von [ʀ] statt.[35] Regressive Nasalassimilationen scheinen auf die Koda beschränkt zu sein.[36]

Dagegen ist die Annahme einer Skala abnehmender Schallfülle oder *Sonorität*[37] vom Zentrum der Silbe (»V« des Nukleus) zu ihren Rändern sowohl theoretisch als auch empirisch problematisch. Heike (1992, S. 9) sieht in der Konstruktion der Sonoritätshierarchie ein typisches Beispiel für zirkuläre Argumentation, weil einerseits behauptet wird, „die Sonoritätshierarchie steuere die phonotaktischen Gegebenheiten, und andererseits [...] die Beoachtungen an der Phonotaktik als Basis für die Formulierung des Sonoritätskonzepts angenommen" werden. Heike weist auch darauf hin, dass der Vorwurf zirkulärer Argumentation bei der Bestimmung der Sonoritätshierarchie bereits von de Saussure (1972, S. 88-90) gegen Sievers (^{2}1881) erhoben wurde.

Die Silbe ist eine auditive Einheit. In der Intonationsanalyse – im unten vorgeschlagenen autosegmentalen Modell ist die Silben-Ebene die zentrale Achse, mit der die Einträge auf der metrischen, der Ton- und der Segment-Ebene assoziiert werden – muss die Zahl der Silben einer Äußerung sicher bestimmt werden. Für das Italienische sieht Bertinetto (1981, S. 162) weder große Probleme bei der Bestimmung der Silbenzahl noch bei derjenigen der Silbengrenzen. Für die Bestimmung der Silbenzahl

---

[33] Die Symbole »C« und »V« bezeichnen Strukturstellen und sind nicht als Abkürzungen für 'Konsonant' und 'Vokal' misszuverstehen. Z.B. wird in diesem Modell ein langer Vokal oder Diphthong mit der Folge VC assoziiert, ein Beispiel findet sich unten, S. 13. Vgl. dazu Wiese (1996), S. 37-43.

[34] Vgl. Vater (1992), S. 105f., außerdem Wiese (1996), S. 44 und S. 234ff.

[35] Wiese (1996), S. 252-258.

[36] Vgl. Wiese (1996), S. 218-224; Vater (1992), S. 111f.

[37] Kohler (^{2}1995), S. 73-75 nennt folgende Klassen von Lauten mit abnehmender Sonorität: offene Vokoide, geschlossene Vokoide, Liquide, Nasale, stimmhafte Frikative, stimmlose Frikative, Plosive. Vgl. zur Sonoritätshierarchie auch Wiese (1996), S. 258-261.

seien lediglich Laute, die bei schnellem Sprechtempo zwischen [i] und [j] bzw. zwischen [u] und [w] oszillieren, problematisch.[38] Für die Bestimmung der Silbengrenzen zeigt ein Perzeptionsexperiment differierende Urteile von Informanten nur bei Clustern aus [s] oder [z] und Konsonant bzw. bei [tm] und [tl].[39] Auch im Deutschen ist die Bestimmung der Silbenzahl nur dort problematisch, wo [r]- oder [i]-Glides auftreten, außerdem in Silben, die [ə] als Kern haben und in der fließenden Rede oft zu den silbischen Konsonanten [m̩], [n̩], [l̩] und [ʁ̩] reduziert werden.[40] Die Silbengrenzen sind im Deutschen mit seinen komplizierten Konsonantenclustern dagegen in vielen Fällen nur sehr schwer zu bestimmen.[41] In der phonologischen Theorie lassen sich Bestimmungsschwierigkeiten jedoch überwinden: Beispielsweise wird das Problem der Festlegung der Silbengrenze nach akzentuiertem Kurzvokal durch die Annahme *ambisilbischer* Konsonanten gelöst.[42]

Die unterschiedlichen Schwierigkeitsgrade in der Bestimmung von Silbenzahl und Silbengrenzen im Deutschen und Italienischen werden üblicherweise mit der unterschiedlichen typologischen Einordnung der Sprachen als silbenzählend (Italienisch) und akzentzählend (Deutsch) erklärt.[43] Die italienische Silbe hat in der Skelett-Ebene eine starke Tendenz zu CV.[44] Untersuchungen zeigen, dass 55% aller italienischen Silben die CV-Skelettstruktur aufweisen.[45] Die maximal möglichen dreifachen Konsonantencluster wie CCCV in [stra] sind im Onset sehr selten, in der Koda kommen Zwei- und Dreifachcluster – CVCC in /'gɔlf/ oder CVCCC in /'films/ – ausschließlich in Fremdwörtern und Neologismen vor.[46] Assimilations- und Reduktionsprozesse verstärken im Italienischen diese Tendenz. Der potentiell viersilbigen Form *lo ha detto* mit der Skelettstruktur CVVCVCCV entspricht gesprochen und geschrieben die dreisilbige Standardrealisierung *l'ha detto* mit CVCVCCV. Das rhythmisch ungünstige Zusammentreffen zweier V wird vermieden. Dagegen haben Assimilations- und Reduktionsprozesse im Deutschen oft gegenläufige Tendenz: Sie verstärken die ohnehin starke Neigung zu komplizierten Konsonantenclustern. In der Skelett-Struktur sind im Onset Dreifachcluster wie CCCVCC in /'ʃtroːm/ und in der Koda sogar Fünffachcluster vom Typ CVCCCCC in /'hɛʁpsts/ möglich.[47] Die Tilgung von [ə], die in der gesprochenen Sprache häufig auftritt, aber auch im Konjugationsparadigma bestimmter Verben vorgesehen ist, führt dazu, dass zum Beispiel das Wort *Barren* mit der

---

[38] Vgl. Bertinetto (1981), S. 158.

[39] Vgl. Bertinetto (1981), S. 161-163.

[40] Vgl. Kohler (1996), passim; Wiese (1996), S. 49-51; Auer (1994), S. 68; Vater (1992), S. 106-110; Bertinetto (1981), S. 157f.

[41] Zur Phonotaktik des Deutschen vgl. Kohler (^{2}1995), S. 175-186.

[42] Vgl. dazu Ramers (1992). Wiese (1996), S. 51 ist deshalb der Meinung, „that very few clear cases for non-predictable syllabification are available" und skizziert deshalb ein (vorläufiges) Regelwerk der Silbifizierung im Deutschen (S. 51-56).

[43] Siehe Kap. 2.2.2.

[44] Was nach Prince/Smolensky (1993), S. 89 die Struktur von „universally optimal syllables" und in allen Sprachen der Welt möglich ist.

[45] Bortolini (1976), S. 12. Vgl. auch Auer/Uhmann (1988), S. 244-249.

[46] Vgl. Bortolini (1976), S. 8ff.

[47] Wieses Modell sieht einen solchen Fünffachcluster nicht vor. Das Wort /'hɛʁpst/ wird in eine Silbe mit der Struktur CVCC und zwei von einem extrasilbischen C abhängende Konsonanten zerlegt. Vgl. Wiese (1996), S. 47-49.

potentiellen Skelett-Struktur CVCVC in der Regel als /'baʀn̩/ mit CVCC realisiert wird.

Die Fehlerforschung spiegelt die Relevanz der Unterschiede in der Skelett-Ebene von Deutsch und Italienisch wider. Italienische Deutschlerner tendieren zum „Segmentieren" langer oder unbekannter Konsonantencluster durch Sprossvokalbildung. Deutsch /'pfanə/ mit der im Italienischen unbekannten Affrikate [pf] und der Struktur CCVCV wird zu /pə'fannə/ mit CVCVCCV. Der im Italienischen seltene konsonantische Wortauslaut wird oft durch Tilgung des finalen C oder Ergänzung von [ə] beseitigt. Aus dem Infinitiv /'vapnən/ mit CVCCVC wird /'vapnə/ mit CVCCV; das Substantiv /'boːt/ mit CVCC wird zu /'boːtə/ mit CVCCV.[48] In beiden Fällen wird die Skelettstruktur der im Italienischen idealen CV-Alternation angenähert.

2.2.2 Die Isochronie-Hypothese

In der *Isochronie-Hypothese* wird die Beobachtung formuliert, dass der Mensch dazu tendiert, akustische Eindrücke in rhythmischen Gruppen zusammenzufassen. Die Tendenz ist relativ unabhängig von der Genauigkeit der Folge der akustischen Stimuli (beispielsweise in der musikalischen Aufführung), und sie kann sogar dort beobachtet werden, wo akustisch keine rhythmische Gliederung nachweisbar ist (zum Beispiel kann auch das gleichmäßige Ticken der Uhr als rhythmisch gegliedert perzipiert werden). Obwohl die Isochronie-Hypothese nicht bewiesen ist, kann man mit ihr unter bestimmten Bedingungen sinnvoll operieren.

In der sprachwissenschaftlichen Tradition werden zwei Typen von Isochronie und entsprechend zwei Typen von Sprachen unterschieden: silbenzählende Sprachen (Sprachen mit *isocronia sillabica*) und akzentzählende Sprachen (Sprachen mit *isocronia accentuale*).[49] Standarditalienisch wird weithin als Prototyp einer silbenzählenden Sprache genannt, während Deutsch und Englisch als akzentzählende Sprachen gelten. Konkret lässt sich die Isochronie-Hypothese wie folgt formulieren: Silbenzählende Sprachen erhalten die Isochronie ihrer Silben, das heißt, die Silbendauer ist unabhängig von der Akzentverteilung. Akzentzählende Sprachen erhalten dagegen die Isochronie ihrer Füße,[50] das heißt, die Silbendauer nimmt mit zunehmender Silbenanzahl zwischen zwei Iktus ab.[51] Gemäß dieser These müsste eine zweisilbige italienische Äußerung exakt die doppelte Dauer einer einsilbigen haben, eine dreisilbige exakt die dreifache. Im Deutschen wären solche Vorhersagen nicht auf der Ebene der Silben, aber auf derjenigen der Füße möglich: Der Abstand zwischen zwei Iktus müsste konstant bleiben, unabhängig davon, ob zwischen ihnen eine, zwei oder drei unbetonte Silben liegen. Auer und Uhmann (1988, S. 220-237) geben einen Überblick über Überprüfungen der Hypothese hinsichtlich verschiedener, auch nicht-indogermanischer Sprachen und stel-

---

[48] Beispiele aus Zuanelli Sonino (1975), S. 93 und S. 99. Die Notierung des langen Vokals [oː] durch VC auf der Skelett-Ebene folgt dem Vorschlag von Wiese (1996), S. 38f.

[49] Vgl. Bertinetto/Magno Caldognetto (1993), S. 147ff.

[50] Wiese (1996), S. 56 definiert *Fuß* wie folgt: „[...] the foot consists of the string of syllables starting from one stressed syllable up to (but not including) the next one." Damit ist *Fuß* synonym mit *Takt* in der Definition von Pheby (1981), S. 852.

[51] Vgl. Auer/Uhmann (1988), S. 217. Zu *Iktus* vgl. Pheby (1981), S. 852ff.