

Sequenzen phonologischer Intonationsgipfel

**Theoretische Möglichkeiten und
empirische Realisierung im
segmentellen Kontext**

Oliver Niebuhr

1 Einleitung

Die folgende Untersuchung beschäftigt sich mit den Kombinationsmöglichkeiten phonologischer Intonationseinheiten in prosodischen Phrasen mit zwei vergleichbar prominenten Satzakkentsilben. Die Grundlage hierzu bildet das Kieler Intonationsmodell KIM (Kohler 1991a). Das KIM ist ein Konturtonmodell. Folglich werden die phonologisch kontrastierten Einheiten durch ganzheitliche intonatorische Gestalten repräsentiert.

Hierzu zählen beispielsweise die Gipfelkonturen. Sie sind in ihrer phonetischen Realisierung häufig durch konvex geformte auf- und absteigende F₀-Bewegungen im Bereich der Satzakkentsilbe gekennzeichnet. Sie werden daher auch als Akzentkonturen bezeichnet. Im KIM werden drei solcher Gipfelkonturen unterschieden: Frühe, mittlere und späte Gipfel. Vereinfacht dargestellt, belegen Studien hierzu (Kohler 1991b; Gartenberg und Panzlaff-Reuter 1991; Niebuhr 2003), dass die Synchronisation des Gipfelmaximums relativ zum Beginn des akzentuierten Vokals die ausschlaggebende phonetische Dimension für ihre Unterscheidung ist. In frühen Gipfeln liegt das F₀-Maximum vor dem Beginn des akzentuierten Vokals. Die Grundfrequenz fällt dadurch in den Vokal hinein. Mittlere und späte Gipfel sind durch einen Anstieg im Vokal gekennzeichnet, wobei der Anstieg für späte Gipfel erst gegen Ende des Vokals beginnt. Die F₀-Maxima dieser Intonationseinheit sind daher in der finalen Peripherie der Akzentsilbe oder im nächsten stimmhaften Abschnitt der nachfolgenden Silbe lokalisiert. Abbildung 1 stellt die drei Gipfelkonturen schematisch dar.

Für die Satzakkentsilbe, die mit der Akzentkontur verbunden ist, werden im KIM drei phonologische Akzentstufen differenziert. Von einer Akzentuierung mittlerer Stärke kann in jeweils einer Stufe nach oben und unten abgewichen werden. Relativ zu dieser Verstärkung und Reduktion wird die mittlere Stufe nachfolgend als neutrale Akzentuierung bezeichnet.

Folgen in einer Phrase zwei Gipfelkonturen aufeinander, so bestehen für den Sprecher mehrere Möglichkeiten, sie melodisch untereinander zu verbinden. Das KIM betrachtet solche Konkatenationsunterschiede als phonologisch relevant, behandelt sie jedoch nicht als (eigenständige) Intonationseinheiten, sondern als weitere Eigenschaften der ganzheitlichen Gipfelkonturen. Sie manifestieren sich meist über den der Akzentsilbe nachfolgenden unakzentuierten Silben.

Das KIM kontrastiert zwei grundsätzliche Konkatenationsmöglichkeiten. Einerseits kann eine Abfolge zweier Gipfelkonturen durch eine hörba-

re Einbuchtung miteinander konkateniert sein. Andererseits besteht die Möglichkeit, die (Hochpunkte beider) Gipfel ohne eine hörbare Einbuchtung miteinander zu verbinden. Durch Letzteres entsteht eine neue, übergeordnete Intonationseinheit, die in Anlehnung an die Terminologie der holländischen Intonationsschule (vgl. 't Hart et al. 1990) als Hutmuster bezeichnet wird.

Die im KIM angelegten Kategorisierungen werden durch das Etikettierungssystem PROLAB (*prosodic labelling*, Kohler 1997) in eine etikettierbare Syntax überführt, die auf dem 7bit-ASCII Format beruht. Dabei werden die phonologisch kontrastierten ganzheitlichen Gipfelkonturen formal in eine Sequenz aus Gipfelkategorie und ihrer dazugehörigen Konkatenationseigenschaft aufgespalten. Zusätzlich zu den im KIM kontrastierten Konkatenationseigenschaften werden in PROLAB zwei unterschiedliche Grade der Konkatenation zweier Gipfelkonturen mit Einbuchtung differenziert: Eine starke Einbuchtung bis an den unteren Bereich des sprecherindividuellen Stimmumfangs und eine leichtere (aber weiterhin hörbare) Einbuchtung. Abbildung 2 stellt die beschriebenen Konkatenationsmöglichkeiten zweier Gipfelkategorien schematisch dar.

Der frühe, mittlere und späte Gipfel wird in PROLAB durch /)/, /[^]/ und /(/, die starke, leichte und fehlende Einbuchtung durch /2./, /1./ und /0./ symbolisiert. Zusätzlich zu diesen intonatorischen Etiketten wird die Akzentstufe der dazugehörigen Satzakkentsilbe durch die Ziffern /1, 2, 3/ angezeigt. Die Ziffer /2/ steht für die neutrale Akzentuierung, /1/ und /3/ stellen dementsprechend die Reduktion oder Verstärkung dieser neutralen Akzentuierung dar. Die erläuterten Ziffern stehen in der Etikettiersyntax vor dem Symbol der Gipfelkategorie. Anhand der beschriebenen, im KIM differenzierten Gipfelkonturen verfolgt die vorliegende Untersuchung zwei benachbarte Zielsetzungen.

Das erste Ziel besteht in der Aufstellung phonotaktischer Regularien für die Kombination von Gipfelkonturen. Das KIM stellt eine enge Verbindung zwischen den phonologischen Intonationseinheiten und ihrer syntaktischen und/oder semantisch-pragmatischen Funktion her, die als Hauptursache kombinatorischer Restriktionen zwischen phonologischen Intonationseinheiten zu sehen ist. Darüber hinaus sieht die Phonotaktik des KIM nur wenige strukturell bedingte Restriktionen in der Kombinierbarkeit phonologischer Intonationseinheiten in Phrasen vor. Zu den wenigen strukturellen Restriktionen gehören die Kombinationsmöglichkeiten von Gipfelkategorien in Hutmustern. Da der Abstieg bzw. der Anstieg essentielle

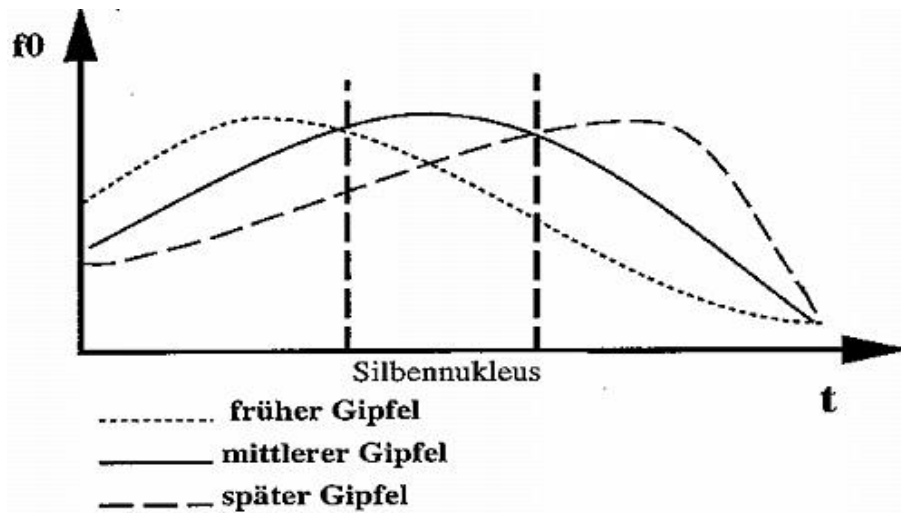


Abbildung 1: Idealisierte Darstellung physikalischer Grundfrequenzverläufe für die drei Gipfelkategorien des KIM: Früh, mittel und spät. Die Abbildung stammt aus Peters und Kohler (2004).

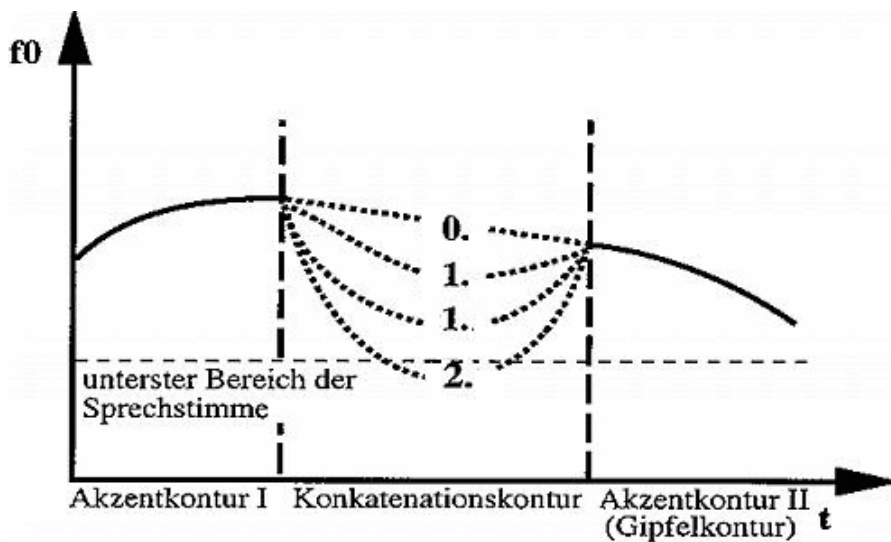


Abbildung 2: Idealisierte Darstellung physikalischer Grundfrequenzverläufe relativ zum akzentuierten Silbennukleus für die drei in PROLAB unterschiedenen Konkatinationmöglichkeiten /0./, /1./ und /2./.. Abbildung modifiziert nach Peters und Kohler (2004).

Merkmale des frühen bzw. späten Gipfels sind, können frühe Gipfel nicht am Beginn und späte Gipfel nicht am Ende eines Hutmusters stehen. Folgende 5 Kombinationen von Gipfelkonturen werden daher als unmöglich angesehen:

$/\hat{0}.$ ($/$, $/(\hat{0}.$ ($/$, $/) \hat{0}.$)/ $,$ $/) \hat{0}.$ \wedge $/$, $/) \hat{0}.$ ($/$

Dem gegenüber steht, dass durch die Verkettung paradigmatischer Oppositionen die Menge an theoretisch beschreibbaren Intonationsmustern faktoriell anwächst, je mehr Strukturstellen (d.h. Intonationseinheiten) innerhalb einer Phrase existieren. Durch die weitere Unterteilung der Gipfelkonkationen mit Einbuchtung in $/1./$ und $/2./$ durch PROLAB wird diese Menge zusätzlich vergrößert.

Dies soll an einem Beispiel verdeutlicht werden: Ausgehend von einer Phrase mit zwei neutral akzentuierten Silben können bereits mindestens 22 verschiedene Intonationsmuster differenziert werden, die durch die Kombination von 3 Gipfelkategorien inklusive je 3 Konkatenationseigenschaften mit weiteren 3 Gipfelkategorien (abzüglich 5 unmöglicher Gipfelsequenzen mit $/0./$, siehe oben) zustande kommen. Tritt ein weiterer Akzent hinzu, erhöht sich diese Zahl bereits auf 160. Hierin ist die Variabilität in den Akzentstufen und phraseninitialen (vor dem ersten Satzakzent) und -finalen (nach dem letzten Satzakzent) Intonationskonturen (siehe hierzu Peters und Kohler 2004) noch nicht enthalten.

Es ist davon auszugehen, dass dieses intonatorische Potential vom Sprecher weder phonetisch noch funktional ausgeschöpft werden kann. Insofern ist der Ausbau der Phonotaktik notwendig für eine übersichtlichere Struktur des Systems und für eine adäquatere Abbildung der phonologischen Realität.

Das zweite Ziel dieser Arbeit ist es, weitere Hinweise darauf zu erhalten, inwieweit dem Unterschied zwischen einem Hutmuster und einer Gipfelsequenz mit Einbuchtung eine funktional oder zeitlich motivierte Steuerung zugrundeliegt. Das KIM nimmt für diesen Unterschied eine funktional motivierte Steuerung an (vgl. Kohler 1991a). Neuere Untersuchungen von Peters, Kohler, und Wesener (2005) anhand einer phonetischen Analyse und funktionalen Interpretation von Beispielen aus dem *Kiel Corpus of Spontaneous Speech* (IPDS 1995; IPDS 1996; IPDS 1997) ergaben keine disjunkten Verteilungstendenzen der Gesamtdauern für beide globalen Muster. Hutmuster sind also nicht an minimale Dauern gebunden, deren Überschreitung zwangsläufig und nur dann zur Realisierung einer Gipfelsequenz mit dazwischen liegender Einbuchtung führt. Dies wäre jedoch die

Voraussetzung für die Interpretation des Hutmusters als zeitlich gesteuerte Reduktion einer solchen Gipfelsequenz. Insofern bestätigen Peters et al. die Vorstellung einer funktional motivierten Steuerung. Dennoch bleibt die Frage bestehen, inwieweit der Unterschied zwischen Hutmuster und Gipfelsequenz mit Einbuchtung *ausschließlich* aus einer funktional motivierten Steuerung hervorgeht. Hierfür müsste dieser Unterschied, der an die Konkatenationseigenschaften von /2./ bzw. /1./ gegenüber /0./ gekoppelt ist, kontextunabhängig aufrecht erhalten werden können. Dies soll nachfolgend überprüft werden.

Um die formulierten Zielsetzungen zu erreichen, wird eine zweistufige experimentelle Strategie verfolgt: In einem ersten Schritt wird ein interaktives Perzeptionsexperiment mit resynthetisierten Intonationsmustern über systematisch ausgewählten Kontexten durchgeführt. Die hieraus gewonnenen Ergebnisse werden in einem folgenden Schritt in einer korpusempirischen Analyse auf der Grundlage des *Kiel Corpus of Spontaneous Speech* (IPDS 1995; IPDS 1996) überprüft. Letzteres soll dazu beitragen, die Validität der Ergebnisse zu unterstützen.

2 Das Perzeptionsexperiment

2.1 Aufbau des Experimentes

Im Rahmen des Perzeptionsexperimentes wurde unter Konstanthaltung aller sonstigen Faktoren der zeitliche Abstand aufeinanderfolgender Gipfelkonturen stufenweise reduziert. Unter den jeweiligen zeitlichen Voraussetzungen wurden die phonologischen Eigenschaften der Gipfelkonturen systematisch variiert. Der Einfluss dieser beiden unabhängigen Variablen auf die abhängige Variable, die Realisierbarkeit der Gipfelkonturen (insb. deren perzeptorische Identifikation), wurde untersucht. Dies geschah durch zwei geschulte Hörer, einer hiervon war der Autor selbst.

2.2 Methode des Perzeptionsexperimentes

2.2.1 Die Auswahl der kombinierten Intonationskategorien

Die Auswahl der zu konstruierenden Intonationsmuster ergab sich aus einem Kompromiss zwischen dem zu erwartenden Zeitaufwand einerseits und der mit steigender Anzahl an Mustern wachsenden Möglichkeit, Restriktionen aufzuzeigen, andererseits. Es wurde in diesem Zusammenhang

Tabelle 1: In die systematische Resynthese eingehende Kombinationen aus Gipfelkategorien mit variierenden Konkatenationseigenschaften (Intonationsmuster) über neutral satz-akzentuierten Silben.

'mittel' initial	'spät' initial	'früh' initial
2^ 0. 2)	2(0. 2)	2) 1. 2)
2^ 1. 2)	2(1. 2)	2) 2. 2)
2^ 2. 2)	2(2. 2)	2) 1. 2^
2^ 0. 2^	2(0. 2^	2) 2. 2^
2^ 1. 2^	2(1. 2^	2) 1. 2(
2^ 2. 2^	2(2. 2^	2) 2. 2(
2^ 1. 2(2(1. 2(
2^ 2. 2(2(2. 2(

entschieden, alle drei Gipfelkategorien, früh /)/, mittel /^ / und spät /(/ , über zwei neutralen Satzakkenten mit allen drei Konkatenationsmöglichkeiten /0./, /1./ und /2./ zu kombinieren. Daraus resultieren 27 verschiedene Intonationsmuster.

Von diesen 27 Intonationsmustern müssen die 5 auf Seite 3 bereits genannten Kombinationen ausgeschlossen werden. Die 22 verbleibenden Intonationsmuster sind in Tabelle 1 unter Verwendung der PROLAB-Symbolisierung aufgeführt.

2.2.2 Das Äußerungsmaterial

Die Grundlage der Resyntheseriehe wird durch vier Äußerungen gebildet. Sie sollten den Variationsspielraum der intonatorischen Kategorien stufenweise einschränken und dadurch kombinatorische Restriktionen provozieren. Hierfür wurden die Äußerungen mit einem konstanten Rahmen konstruiert, innerhalb dessen die Dauer der Stimmhaftigkeit kontrolliert herabgesetzt wurde.

Tabelle 2 stellt die vier Sätze anhand einer orthographischen und phonetisch-allophonischen Transkription vor. Der Einfachheit halber wird nachfolgend nur noch auf die Satznummern Bezug genommen.

Der konstante Rahmen der vier Äußerungen bezieht sich auf folgende Punkte: Die Äußerungen bestehen syntaktisch aus einem Aussagesatz der Form "Er hat [...] gesagt.". Zwischen "Er hat" und "gesagt", deren Intonationskontur den *pre-head* bzw. *tail* (in der Terminologie von Kingdon 1958) der prosodischen Phrase bilden, wurden jeweils zwei vergleichbar

Tabelle 2: Orthographische und phonetische Transkription der die Basis der systematischen Resyntheseriehe etablierenden Sätze 1-4. Akzentuierte Silben sind kursiv dargestellt.

Satz	Orthographie	Phonetik
1	Er hat <i>immer</i> wieder <i>nein</i> gesagt	ɛfiət ^h 'imøvide'namgəza:kt ^h
2	Er hat <i>immer</i> <i>nein</i> gesagt	ɛfiət ^h 'imø'namgəza:kt ^h
3	Er hat <i>oft</i> <i>nein</i> gesagt	?ɛfiət'ɔft ⁿ 'namgəza:kt ^h
4	Er hat <i>oft</i> <i>nichts</i> gesagt	?ɛɛfiət ^h 'ɔft ⁿ 'niçtsçəza:kt ^h

prominente Satzakkentsilben der mittleren (neutralen) Akzentstufe realisiert. Die intonatorischen Kategorien des *pre-heads* bzw. des *tails* waren in allen vier Äußerungen identisch. Die Äußerungen wurden zudem von demselben männlichen Sprecher (kk) mit vergleichbarer Sprechgeschwindigkeit, rhythmischer Struktur und finaler Längung produziert.

Die stufenweise Herabsetzung der Dauer der Stimmhaftigkeit im experimentell relevanten Bereich zwischen *pre-head* und *tail* über die vier Sätze (siehe Tabelle 3) basiert auf einer zweidimensionalen Strategie. Zum einen wurde die Dauer der Stimmhaftigkeit zwischen den beiden Satzakkenten durch eine geringere Anzahl unbetonter Silben reduziert. Zum anderen wurde (gleichzeitig) der Anteil phonetisch stimmloser Segmente erhöht.

Während beispielsweise im ersten Satz “Er hat immer wieder nein gesagt” (vgl. Tabelle 2) zwischen der ersten Silbe von “immer” und “nein” noch drei unbetonte Silben stehen, folgen im dritten und vierten Satz beide akzentuierte Silben unmittelbar aufeinander. Des Weiteren wird vom ersten zum vierten Satz der Diphthong der Wortform “nein” /ai/ gegen den monophthongalen Kurzvokal der Wortform “nichts” /ɪ/ substituiert, und beide akzentuierte Silben erhalten eine stimmlose Koda.

Tabelle 3 fasst die Gesamtdauer an Stimmhaftigkeit in jeder Äußerung zusammen, die zur Realisierung der in Tabelle 1 angeführten Intonationsmuster bzw. zu deren Signalisierung gegenüber einem Hörer verbleibt. Die Dauern der akzentuierten Silbennuklei sind aufgrund ihrer Relevanz für die Realisierung der phonologisch relevanten Eigenschaften der drei Gipfelkonturen gesondert aufgeführt. Dauern und Segmentgrenzen wurden anhand von Sonagramm und Oszillogramm ermittelt.

Tabelle 3: Dauern (in ms) a) der akzentuierten Silbennuklei aus “immer”, “oft”, “nein” und “nichts” sowie b) der gesamten modalen Stimmhaftigkeit zwischen *pre-head* und *tail*, jeweils gerundet auf 5ms.

Satz	“immer” / ”oft”	“nein” / ”nichts”	Sth. Gesamt
1	55	190	785
2	65	205	560
3	120	170	410
4	105	66	195

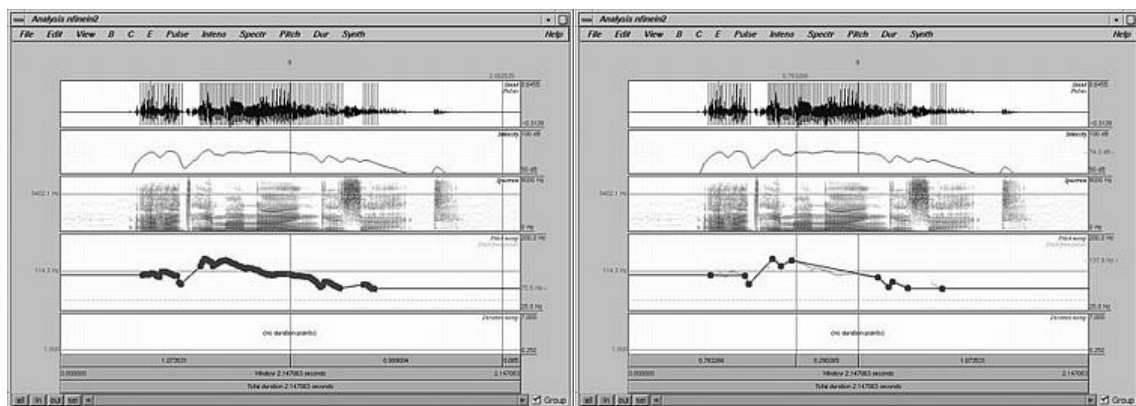


Abbildung 3: Darstellung des unter *praat* analysierten und stilisierten (rechts) Satzes 2. Von oben nach unten sind folgende Fenster dargestellt: Oszillogramm, Intensitätsverlauf, Spektrogramm und F0-Verlauf.

2.2.3 Die systematische Konstruktion der Intonationsmuster

Für die Konstruktion der in Tabelle 1 aufgeführten globalen Intonationsmuster wurde zunächst der Grundfrequenzverlauf der jeweiligen Originaläußerung (Satz 1-4) mit *praat* (Boersma und Weenik) analysiert. Die Grundfrequenzanalyse wurde in 10ms Schritten in einem Frequenzbereich zwischen 75Hz und 250Hz durchgeführt. Das daraus entstandene F0-Muster wurde mit einer Genauigkeit von zwei Halbtonschritten stilisiert und so auf eine handhabbare Anzahl an Konturpunkten reduziert. Die Konturpunkte repräsentieren die Wendepunkte im F0-Verlauf. Abbildung 3 illustriert das Ergebnis der Analyse- und Stilisierungsprozedur anhand einer Vorher-Nachher-Gegenüberstellung von Satz 2.

Im Bereich zwischen *pre-head* und *tail* wurde der Intonationsverlauf systematisch durch das Verschieben der Konturpunkte in Zeit- und Frequenzdimension manipuliert, um die notwendigen Eigenschaften der Gipfelkonturen zu generieren. Die zum *pre-head* und zum *tail* gehörigen Abschnitte

des Gesamtmusters blieben bei der systematischen Resynthese unangetastet. Sie sind phonetisch durch einen annähernd ebenen Verlauf auf tiefem Frequenz- bzw. Tonhöheniveau gekennzeichnet.

Für die /1./-Konkatenation musste lediglich eine (hörbare) Einbuchtung zwischen den F0-Gipfelmaxima durch einen Konturpunkt konstruiert werden. Für die Konstruktion einer /2./-Konkatenation indes (vgl. Abb. 2) wurde ein sprecherindividueller Tiefpunkt in der Einbuchtung von 80Hz festgelegt, der unabhängig von der Lage der umliegenden Konturpunkte über alle Sätze konstant gehalten wurde.

Die Werte der F0-Maxima der beiden Gipfel wurden aus der jeweiligen Originaläußerung als Richtwerte für die Konstruktion der Intonationsmuster berücksichtigt. Als verbindlich für die Konstruktion der Intonationsmuster galt hingegen die vorgegebene phonologische Akzentstruktur (Stufe /2/ auf beiden Satzakzenten), und dass das zweite Gipfelmaximum gegenüber dem ersten hörbar tiefer liegt (siehe *downstep*, Kohler 1991a).

Die zu kombinierenden Gipfelkonturen wurden immer durch ein Minimum an Konturpunkten repräsentiert und durch einen linear interpolierten F0-Abschnitt miteinander verbunden. In Abhängigkeit von den erforderlichen Eigenschaften der zu verkettenden Gipfelkonturen wurden dem Intonationsmuster entweder Konturpunkte hinzugefügt oder entfernt. Ein zweigipfliges Muster mit Einbuchtung (/1./ oder /2./) ist entsprechend durch fünf Konturpunkte darstellbar. Für ein Hutmuster (/0./) reichen bereits vier Konturpunkte aus.

Auf diese Weise wurde versucht, genau *eine* phonetische Variante jedes in Tabelle 1 genannten Intonationsmusters durch die in *praat* implementierte *TD-psola* (vgl. Moulines und Charpentier 1990) zu resynthetisieren. Bei diesem Resyntheseverfahren bleiben die laut- und stimmlichen Eigenschaften von der Grundfrequenzmanipulation nahezu unbetroffen, was eine sehr gute Resynthesequalität sicherstellt, solange die Grundfrequenz in solchen Grenzen manipuliert wird, die eine Überlappung der Fensterfunktionen gewährleisten. Dies wurde bei der Durchführung der Resynthese berücksichtigt.

Nach den beschriebenen Prinzipien und Verfahrensweisen konnten maximal 88 Intonationsmuster generiert werden (22 Intonationsmuster x 4 Sätze).

2.2.4 Die Beurteilung hinsichtlich kombinatorischer Restriktionen

Die Resynthese der Sprachsignale sowie die Konstruktion der Grundfrequenzmuster wurde durch einen geschulten Phonetiker (on) durchgeführt, der mit den Belangen des KIM zuvor weitgehend vertraut gemacht wurde. Im Anschluss an ein generiertes Muster über einem der Sätze wurde das hörbare Resultat durch eine weitere geschulte und KIM-erfahrene Phonetikerin (tkh) im Hinblick auf die Übereinstimmung mit den intendierten Eigenschaften der Gipfelkonturen überprüft und wenn notwendig erneut modifiziert.

Dies wurde solange durchgeführt, bis *beide* Hörer das phonetische Intonationsmuster zweifelsfrei den intendierten phonologischen Kategorien zuordnen konnten. Konnte eine solche Zuordnung nicht erfolgen, wurde dies entsprechend vermerkt und mit dem nächsten Intonationsmuster fortgefahren.

Die Entscheidung des Kontrollhörers sowie des Exploranden, ob eine kombinatorische Restriktion vorliegt oder nicht, wurde auf der Basis der folgenden drei Kriterien getroffen:

1. Akustisches Kriterium (A): Die segmentelle Grundlage lässt eine Konstruktion der für die Wahrnehmung der jeweiligen Gipfelkonturen notwendigen akustischen Eigenschaften nicht zu.
2. Perzeptorisches Kriterium (P): Die notwendigen Eigenschaften für die Zuordnung einer Intonationseinheit zu einer bestimmten Gipfelkontur waren zwar im akustischen Signal angelegt, konnten jedoch nicht perzipiert werden. Dieser Aspekt weist auf die Möglichkeit einer Diskrepanz zwischen Perzeption und Akustik hin. Es ist in diesem Kontext wichtig hervorzuheben, dass *Wahrnehmung* immer die Interpretation der physikalischen Realität vor dem Hintergrund bestimmter Kriterien bedeutet. Sie ist somit kein Abbild der physikalischen, in diesem Fall akustischen Realität. Ihren interpretativen Charakter bekommt die Wahrnehmung insbesondere durch die Kriterien der Gewohnheit bzw. einer zugrundeliegenden Wissensbasis und der daraus resultierenden Modifikation der Akustik durch Erwartungshaltungen des Wahrnehmungsapparates (vgl. z.B. "Worteffekt", Gannon 1983; "Deklinaton", Gussenhoven und Rietveld 1988), multimodale Interferenzen (z.B. "McGurk-Effekt", McGurk und MacDonald 1976), aufgabenspezifische Limitationen des Wahrnehmungsapparates (z.B. "absolute threshold" und "differential threshold", 't Hart

et al. 1990) sowie durch das Kriterium des Aufbaus des Wahrnehmungsapparates, das sich in die physiologisch determinierten Aspekte der Signaltransformation (z.B. “bark scale”, vgl. Ladefoged 1996) und die prozessoralen Aspekte (z.B. “virtual pitch theory”, Terhardt 1974) subkategorisieren lässt und letztlich auch als Limitation zu betrachten ist. Eine weiterführende Diskussion zum interpretativen Charakter der Wahrnehmung im allgemeinen findet sich in Wendt 1989).

3. Natürlichsprachliches Kriterium (N): Die vorangehenden Ausschlusskriterien werden nicht erfüllt, aber das perzipierte Intonationsmuster klingt aus muttersprachlicher Sicht des Exploranden unnatürlich und technisch. Mit diesem Urteil gehen meist berechtigte Zweifel einher, ob das konstruierte Intonationsmuster von einem menschlich-biologischen Sprechapparat produzierbar wäre (vgl. Sundberg 1979; Xu und Sun 2002). Hierbei handelt es sich auch um ein perzeptorisch angelegtes Kriterium. Angesichts dieses Kriteriums ist das Problembewusstsein des Exploranden, die durch die Technik theoretisch gegebenen Manipulationsmöglichkeiten nicht über das Maß biologischer Fähigkeiten hinaus anzuwenden von immenser Wichtigkeit.

Für das Vorhandensein einer kombinatorischen Restriktion wurde es als ausreichend angesehen, wenn *eines* der genannten drei Kriterien (A, P oder N) erfüllt wurde.

2.2.5 Hinweise zur Vergleichbarkeit der generierten Intonationsmuster

Als unter 2.2.3 die systematische Konstruktion der einzelnen Intonationsmuster beschrieben wurde, ist nicht näher auf die (In-)Konsistenz in der Platzierung der einzelnen relevanten Konturpunkte im Zeit- und Frequenzraum eingegangen worden.

Hierzu muss hervorgehoben werden, dass eines der Ziele dieser Untersuchung darin besteht, Restriktionen in der Kombination *phonologischer* Intonationseinheiten (Gipfelkonturen) aufzudecken. *Bestimmte* phonetische Ausprägungen einer phonologischen Intonationseinheit (Gipfelkontur) sind damit nicht Gegenstand dieser Untersuchung.

Dementsprechend wurden im Vorfeld der Konstruktion und Resynthese der 22 Intonationsmuster weder zeitliche (bzw. an Eigenschaften der Lautsegmente orientierte) noch frequenzbezogene Referenzpunkte (außer der Berücksichtigung des *downstep*, siehe 2.2.3) für die Platzierung der einzelnen Konturpunkte festgelegt. Es wurde lediglich bestimmt, dass keiner der

Konturpunkte des Intonationsmusters in einen stimmlosen Signalabschnitt fallen sollte, um die Perzipierbarkeit der phonologisch relevanten Eigenschaften einer Gipfelkontur nicht herabzusetzen. In der Sequenz der weitgehend stimmlosen Akzentsilben “oft nichts”, musste beispielsweise der Konturpunkt für eine Einbuchtung daher entweder in den vokalischen Nukleus der ersten Silbe gesetzt werden oder in den initialen Nasal der folgenden Silbe. Je nach zu realisierender Gipfelkategorie konnte sich dadurch eine extreme zeitliche Nähe zu den Konturpunkten der F0-Gipfelmaxima ergeben.

Eine Möglichkeit zu entscheiden, wohin der Tiefpunkt der Einbuchtung letztlich zu setzen ist, ergab sich daher durch die Position der ihn umgebenden Konturpunkte in Zeit und Frequenz (den phonetischen Eigenschaften des gesamten Intonationsmusters) sowie durch die Beschaffenheit der artikulatorischen Grundlage. Genauso gut konnte es jedoch sein, dass zugunsten der Positionierung eines Tiefpunktes an einer bestimmten Stelle die Konturpunkte der umliegenden F0-Gipfelmaxima in Zeit und/oder Frequenz verschoben wurden.

Somit wurde für jedes der 22 Intonationsmuster in jedem der 4 Kontexte ein individuelles Gesamtarrangement der Konturpunkte hergestellt, das vor dem Hintergrund der drei Restriktionskriterien aus 2.2.4 ein bestmögliches Perzept ergab. Dies führte in der Konsequenz jedoch dazu, dass Realisierungen der erforderlichen Eigenschaften einer Gipfelkontur in verschiedenen intonatorischen und artikulatorischen Kontexten sowie Realisierungen desselben phonologischen Intonationsmusters über verschiedenen Sätzen nur bedingt miteinander verglichen werden können.

Diese Vorgehensweise hat jedoch einen essenziellen Vorteil: Sie gestattet eine flexible und kreative Adaption der phonetischen Manifestation einer Gipfelkontur an bestimmte kritische Momente, die aus dem Konflikt zwischen der Kodierung ihrer phonologisch relevanten Eigenschaften und der beschränkten Dauer der Stimmhaftigkeit entstehen. Analytische Beobachtungen verschiedener Strategien haben darüber hinaus zu weiteren Erkenntnissen geführt.

Die Vorgehensweise ist somit weiterhin als systematisch zu charakterisieren, da sich dieser Terminus auf die Alternation von *phonologischen* Kategorien und nicht auf die Alternation von *a priori* definierten *phonetischen* Varianten einer Gipfelkontur bezieht. Die beschriebene Methodik darf ferner nicht als willkürlich fehlinterpretiert werden, da zumindest *pre-head* und *tail* jeder Äußerung konstant gehalten wurden (vgl. 2.2.3) und

sich der übrige Bereich am parametrischen Raum der dazugehörigen Originaläußerung orientierte.

2.3 Ergebnisse des Perzeptionsexperimentes

Die erzeugten Resynthesen finden sich zusammen mit den Analyseabbildungen unter www.ipds.uni-kiel.de/kjk/publikationen/audiobsp.de.html. Die Ergebnisse zur Resynthese der 22 möglichen Intonationsmuster sind in Tabelle 4 zusammengefasst. Konnte ein Muster in einem der 4 Sätze nicht resynthetisiert werden, ist die Ursache dafür durch den Anfangsbuchstaben des erfüllten Restriktionskriteriums (vgl. 2.2.4) angegeben. Werden mehrere Restriktionskriterien gleichzeitig erfüllt, sind die Anfangsbuchstaben durch ein '+'-Zeichen getrennt aufgelistet.

Aus Tabelle 4 sind drei wesentliche Ergebnisse erkennbar. Erstens konnten nicht alle 88 Resynthesen (22 Intonationsmuster über 4 Sätzen) erfolgreich abgeschlossen werden. In einer Reihe von Fällen waren die Abfolgen der Gipfelkonturen nicht realisierbar.

Zweitens zeigt Tabelle 4 zwei globale Trends. Zum einen nimmt die Anzahl der kombinatorischen Restriktionen mit zurückgehendem Variationsspielraum für die Realisierung der erforderlichen Eigenschaften der einzelnen Gipfelkonturen durch die stufenweise Reduktion von Dauer und Stimmhaftigkeit in und zwischen den beiden Satzakkentsilben zu. Zum zweiten erhöht sich die Anzahl der kombinatorischen Restriktionen mit wachsender Komplexität der Intonationsmuster. Die Komplexität eines Intonationsmusters hängt unter anderem mit der Stärke der Einbuchtung zwischen den Hochpunkten beider Gipfelkonturen zusammen. So verlangen /1./ und /2./ gegenüber /0./ (Hutmuster) zwei zusätzliche Richtungsänderungen im F₀-Verlauf (Ab- und erneuter Anstieg), wodurch das Gesamtmuster komplexer wird. Durch die besonders ausgeprägte Einbuchtung ist eine /2./-Konkatenation zudem komplexer als eine /1./-Konkatenation. Den beiden globalen Trends der Ergebnisse entsprechend, weist die Ergebnistabelle unten mehr Einträge erfüllter Restriktionskriterien auf als oben und rechts mehr als links.

Werden diese globalen Trends auf einzelne Intonationsmuster bezogen, zeigt Tabelle 4 beispielsweise, dass mit /0./ konkatenierte Gipfelkonturen (Hutmuster) in allen 4 Sätzen resynthetisierbar gewesen sind, während Verbindungen aus Gipfelkonturen mit einer starken Einbuchtung (/2./) dazwischen bereits ab dem zweiten Satz ("immer nein") überwiegend in eine

Tabelle 4: Ergebnisse der systematischen Resynthese der 22 Intonationsmuster (zwischen zwei neutral akzentuierten Silben) in den 4 Sätzen. Die Ursachen kombinatorischer Restriktionen sind durch den Anfangsbuchstaben des erfüllten Restriktionskriteriums angezeigt (vgl. 2.2.4). **A**= Akustisches Kriterium, **P**= Perzeptorisches Kriterium, **N**= Natürlichsprachliches Kriterium.

Konkatenation mit	Intonation	Satz 1	Satz 2	Satz 3	Satz 4
0.	2^ 0. 2)				
	2^ 0. 2^				
	2(0. 2)				
	2(0. 2^				
1.	2) 1. 2)				
	2) 1. 2^				
	2) 1. 2(
	2^ 1. 2)				
	2^ 1. 2^				
	2^ 1. 2(
	2(1. 2)		N	P+N	A+P+N
	2(1. 2^				
2.	2(1. 2(
	2) 2. 2)		N	N	N
	2) 2. 2^			N	N
	2) 2. 2(N
	2^ 2. 2)		P+N	P+N	P+N
	2^ 2. 2^		N	P+N	P+N
	2^ 2. 2(P+N
	2(2. 2)	P+N	P+N	P+N	A+P+N
	2(2. 2^		P+N	P+N	P+N
	2(2. 2(N

kombinatorische Restriktion gemündet sind. Im vierten Satz (“oft nichts”) liegt für jede Paarung der drei Gipfelkonturen, wenn sie durch ein /2./ konkateniert wurden, eine kombinatorische Restriktion vor. Des weiteren sind Intonationsmuster mit dem späten Gipfel initial und dem frühen Gipfel final am häufigsten von kombinatorischen Restriktionen betroffen gewesen und haben als einzige in zwei Fällen alle drei Restriktionskriterien gleichzeitig erfüllt (/2(1. 2)/ und /2(2. 2)/ in Satz 4). Entsprechend konnte das Intonationsmuster, das eine starke Einbuchtung und zugleich eine Abfolge aus spätem und frühem Gipfel enthalten sollte, als Einziges nicht einmal im Satz 1 (“immer wieder nein”) generiert werden.

Umgekehrt sind die Intonationsmuster, in denen ein früher und ein später

Gipfel bzw. zwei späte Gipfel aufeinander folgen besonders selten von kombinatorischen Restriktionen betroffen gewesen. Gleiches gilt für die heterogenen Paarungen aus frühem und mittlerem Gipfel. Selbst bei einer Konkatenation durch /2./ haben sich für diese Intonationsmuster erst im vierten Satz kombinatorische Restriktionen ergeben.

Drittens wird aus Tabelle 4 ersichtlich, dass das Verfehlen eines natürlichsprachlichen Eindrucks (Kriterium N) die häufigste Ursache für das Zustandekommen einer kombinatorischen Restriktion gewesen ist. Auf der anderen Seite waren bis auf die beiden Muster /2(1. 2)/ und /2(2. 2)/ in Satz 4 alle Intonationsmuster akustisch herstellbar.

3 Die korpusempirische Analyse

3.1 Warum eine zusätzliche Korpusanalyse?

Die Ergebnisse des interaktiven Perzeptionsexperimentes haben gezeigt, dass für Gipfelsequenzen mit Einbuchtung eine Reihe kontextsensitiver kombinatorischer Restriktionen existieren, die sich auf die Konkatenationsmöglichkeit mit /2./ konzentrieren. Obwohl für jede experimentelle Untersuchung grundsätzlich davon ausgegangen werden muss, dass die gefundenen Ergebnisse keinen absoluten Wert besitzen, gilt dies für die vorliegenden Ergebnisse aus folgenden Gründen in besonderer Weise.

Es ist davon auszugehen, dass bei der Sprachproduktion der artikulatorische Strang (die überlappende Erzeugung lautspezifischer Ansatzrohrkonfigurationen) und der intonatorische Strang (die Erzeugung und Verbindung globaler intonatorischer Einheiten) prinzipiell unabhängig voneinander koordiniert und kontrolliert werden können. Die unterschiedliche Synchronisation phonologisch kontrastierender Gipfel- und Talkonturen mit akzentuierten Silben unterstützt diese Hypothese ebenso wie die mögliche Realisierung unterschiedlicher F₀-Gipfelgestalten und -höhen über konstanten Kontexten.

Trotz dieser prinzipiellen Unabhängigkeit besteht jedoch die Tendenz beider Stränge, sich z.B. in ihrer Dauerstruktur gegenseitig zu beeinflussen. Gartenberg und Panzlaff-Reuter (1991) zeigen hierzu, dass die Silbennuklei in Verbindung mit späten Gipfeln unter sonst vergleichbaren Bedingungen eine größere mittlere Dauer aufweisen als Nuklei, über denen frühe und mittlere Gipfel realisiert worden sind. Andere Untersuchungen belegen, dass die Synchronisation einer phonologischen Intonationseinheit über unterschiedlichen Silbenstrukturen starken Schwankungen unterlie-

gen kann (z.B. van Santen 2002), die auch zu Überlappungen der Synchronisationsbereiche angrenzender Gipfelkonturen führen können.

Die Perzeptionsuntersuchung konnte aus Gründen experimenteller Systematik diese Interaktionen nicht berücksichtigen. Die zugrundeliegenden Segmentdauern der Sätze 1-4 blieben für alle generierten Intonationsmuster konstant. Die einleitend beschriebenen, diskreten Synchronisationsbereiche des frühen, mittleren und späten Gipfels (vgl. Abb. 1) wurden kontextunabhängig aufrecht erhalten. Beide unflexiblen Vorgehensweisen schränken die Dauer für die Realisierung einer Gipfelkontur stärker ein, als dies bei der Produktion von Sprache gegeben sein kann. Da die zur Verfügung stehende Dauer (insb. stimmhafter Silbennuklei) zur Realisierung von Gipfelkonturen aber ausschlaggebend für die Feststellung kombinatorischer Restriktionen sein kann, sind die Ergebnisse aus Tabelle 4 nur bedingt auf tatsächliche Sprechereignisse übertragbar.

Kohler (1991b) kommt ferner zu der Auffassung, dass das Intensitätsmuster einer Äußerung für die Wahrnehmung angrenzender phonologischer Gipfelkonturen relevant sein kann. Auch dieser Parameter wurde bei der Generierung und perzeptorischen Beurteilung der Intonationsmuster nicht gipfelkonturspezifisch verändert.

Ein abschließender Grund, warum die Ergebnisse des Perzeptionsexperimentes nicht für sich allein stehen können, betrifft die unterschiedlichen Strukturen der beiden Akzentsilben in den Sätzen 1-4. Im Abschnitt 2.2.3 wurde darauf hingewiesen, dass die Satzakkzentsilben der generierten Äußerungen unabhängig vom Intonationsmuster eine neutrale Akzentuierung aufweisen sollten. Die Beurteilung hinsichtlich der phonologischen Akzentstufe orientiert sich an der perzeptorischen Prominenz der jeweiligen Silbe in Relation zu den benachbarten Silben und anderen akzentuierten Silben derselben intonatorischen Phrase (jeder Satz bildete eine intonatorische Phrase). In die Wahrnehmung der Prominenz fließen auch die Parameter der Intensität und Dauer einer Silbe bzw. deren Nukleus ein.

Daraus resultiert ein inhärentes Ungleichgewicht in der Prominenz der Satzakkzentsilben “im-” (aus “immer”) und “oft” gegenüber “nein” (siehe Tabelle 3). Vergleichbar starke F₀-bedingte Prominenzindikatoren vorausgesetzt, waren die ersten Satzakkzentsilben der Sätze 1-3 (“im-” und “oft”) der Akzentsilbe “nein” in ihrer perzeptorischen Prominenz unterlegen und tendierten dazu, in Relation zu “nein” als teilweise deakzentuiert (Akzentstufe /1/) wahrgenommen zu werden. Diese Diskrepanz musste über F₀-Gipfeldynamik und -höhe der Gipfelkonturen kompensiert wer-

den. Hierdurch wurden die Möglichkeiten der phonetischen Variation der Gipfelkonturen in Zeit- und Frequenzdimension eingeschränkt, was die Feststellung kombinatorischer Restriktionen in einigen Fällen begünstigt haben könnte. Erst ab Satz 4 (“Er hat oft nichts gesagt”) waren die Prominenzen beider Satzakzentsilben “oft” und “nichts” bei vergleichbaren F0-Indikatoren ausgeglichener.

Zusammenfassend tendieren die Ergebnisse der Perzeptionsuntersuchung somit dazu, das Vorkommen kombinatorischer Restriktionen in der Verbindung von Gipfelkonturen zu *überschätzen*. Dies bedeutet zum einen, dass nicht davon auszugehen ist, dass im realen Sprachgebrauch außer den kombinatorischen Restriktionen, die aus der Perzeptionsuntersuchung hervorgetreten sind, noch weitere strukturelle Beschränkungen in der Verbindung phonologischer Gipfelkategorien existieren (dies schließt semantisch bedingte Restriktion nicht mit ein). Zum anderen macht es die Möglichkeit der Überschätzung erforderlich, die im Perzeptionsexperiment gefundenen kombinatorischen Restriktionen anhand einer umfassenden Korpusanalyse zu bestätigen. Nachfolgend soll das Korpus beschrieben werden, das für diese Analyse herangezogen wurde.

3.2 Das *Kiel Corpus of Spontaneous Speech*

Das *Kiel Corpus of Spontaneous Speech* (IPDS 1995; IPDS 1996; IPDS 1997) umfasst spontansprachliche (d.h. nicht schriftlich vorgegebene) Dialoge zwischen Sprecherinnen und Sprechern der Standardvarietät des Deutschen. Diese sollten anhand eines vorgegebenen Kalenders mehrere Termine für geschäftliche Treffen vereinbaren und wurden dabei in einer schallgedämpften Kabine unter Studiobedingungen aufgenommen (siehe Kohler et al. 1995 sowie Peters 2005 für eine detaillierte Beschreibung des Verfahrens). Das mehrstündige Material liegt in einer segmentell und prosodisch etikettierten Form vor. Die prosodische Etikettierung erfolgte im PROLAB-System.

Die nachfolgend erläuterte Korpusanalyse basiert auf einem Ausschnitt dieses Datenkorpus, der die ersten beiden CDs umfasst (IPDS 1995; IPDS 1996). In Zahlen bedeutet dies folgendes:

- 32 Sprecher
- 82 Dialoge
- 1387 Dialog-Turns

Der skizzierte Ausschnitt des *Kiel Corpus* wurde aus zweierlei Gründen für die Überprüfung der Ergebnisse des Perzeptionsexperimentes herangezogen: Erstens ist es aufgrund seines Umfanges eine repräsentative Stichprobe spontansprachlichen Materials des Deutschen, die den qualitativen Ansprüchen dieser Untersuchung gerecht wird und zweitens ist durch die PROLAB-Etikettierung die phonetische Kompatibilität (rechnergestützt auffindbarer) intonatorischer Muster zu denen des Perzeptionsexperimentes gewährleistet.

3.3 Methode der Korpusanalyse

Wenn die Ergebnisse des interaktiven Perzeptionsexperimentes dieser Untersuchung valide sind, dann finden sich genau an den Stellen Lücken in der Produktion der phonologisch kontrastiven Intonationsmuster, wo unter vergleichbaren kontextuellen Bedingungen in der Perzeption kombinatorische Restriktionen vermerkt wurden. Um dies in einer Korpusanalyse zu überprüfen, ist es ausreichend, das Augenmerk auf Kombinationen phonologischer Intonationsgipfel zu richten, die im Perzeptionsexperiment in die Sackgasse einer kombinatorischen Restriktion gemündet sind. Gemäß Tabelle 4 betrifft dies alle 9 durch ein /2./ verbundenen Gipfelkategorien sowie die eine durch /1./ konkatenierte Abfolge aus spätem und frühem Gipfel, jeweils verbunden mit neutral akzentuierten Satzakkentsilben. Somit waren insgesamt 10 Intonationsmustertypen Gegenstand der Korpusanalyse. Die übrigen 12 der in Tabelle 1 dargestellten Mustertypen wurden in der Analyse nicht berücksichtigt; zum einen, weil dies die eingeschränkte Aufgabe dieser Analyse nicht vorsieht und zum zweiten, weil auf Basis der Argumentation in Abschnitt 3.1 kombinatorische Restriktionen für diese Mustertypen unwahrscheinlich sind.

Im Rahmen der Methode sind zwei grundlegende Arbeitsschritte anzusprechen: Zum einen, wie diese Muster aus dem *Kiel Corpus* extrahiert wurden und zum anderen, auf der Basis welcher Kriterien sie zu den Sätzen 1-4 der Resyntheseriehe zugeordnet wurden.

Das Auffinden von Produktionen der 10 Intonationsmustertypen im ersten Arbeitsschritt geschah automatisch mittels eines Suchalgorithmus in *awk* (Close, Robbins, Rubin, und Stallman 1993) über einer aufbereiteten Textdatei, in der die prosodische Etikettierung aller hier berücksichtigten prosodischen Phrasen des *Kiel Corpus* (IPDS 1995; IPDS 1996) zusam-

mengefasst ist.¹ In der gefundenen Fallmenge sollten nur diejenigen Produktionen von Interesse sein, die wie alle Muster der Resynthesereihe sowohl durch einen Anstieg zum ersten Gipfel als auch durch einen Abstieg vom zweiten Gipfel gekennzeichnet waren (vgl. *pre-head* und *tail* in 2.2.3). Daher wurden alle Fälle aussortiert, die durch eine /0./-Konkatenation mit der vorangehenden und/oder nachfolgenden Akzentkontur verbunden waren. Des Weiteren blieben - im Einklang mit den Konstruktionsbedingungen in der Resynthesereihe (vgl. 2.2.3) - alle Fälle unberücksichtigt, die durch einen *upstep* (vgl. Peters und Kohler 2004) des zweiten Gipfels gekennzeichnet waren und deren Gipfelkategorien nicht mit neutralen Satzakkzenten verbunden waren. Abschließend wurden alle übrigen Fälle durch den Untersuchungsleiter (on) auditiv auf eine korrekte prosodische Etikettierung hin überprüft. Insgesamt verblieben aus dem *Kiel Corpus* 110 relevante Intonationsmuster, die sich auf die genannten Mustertypen verteilten.

Um eine sinnvolle Überprüfung der Ergebnisse des Perzeptionsexperimentes anhand der Sprachdaten des *Kiel Corpus* durchführen zu können, musste im Rahmen des zweiten Arbeitsschrittes eine Zuordnung der 110 Produktionen der 10 Mustertypen aus dem *Kiel Corpus* zu den Sätzen 1-4 des Perzeptionsexperimentes erfolgen. Hierfür bedurfte es eines Kriteriums, das festlegt, welche Eigenschaften diese 110 Äußerungsabschnitte aus dem *Kiel Corpus* aufweisen mussten, damit die kontextuellen Rahmenbedingungen für die Realisierung der 10 Mustertypen mit einem der Sätze 1-4 aus der Resynthesereihe als vergleichbar gelten konnten.

Die Suche nach dem dafür geeigneten Kriterium profitierte von besonderen Verteilungseigenschaften kombinatorischer Restriktionen in den Ergebnissen des Perzeptionsexperimentes: Wenn kombinatorische Restriktionen für Satz 3 ("oft nein") vorlagen, waren sie in der Regel auch in Satz 4 ("oft nichts") vorhanden (vgl. Tabelle 4). Zudem war diese Verbindung aus Satz 3 und 4 weitaus häufiger von Restriktionen betroffen als Satz 2 ("immer nein"). Nun ist es ein wesentlicher Unterschied von Satz 2 gegenüber Satz 3 und 4, dass er über eine unbetonte Silbe zwischen den beiden Satzakkzentsilben verfügt, während in den anderen beiden Sätzen die Akzentsilben direkt aufeinander folgen. Dies deutet darauf hin, dass die Anzahl unbetonter Silben zwischen den beiden Akzenten von größerer Bedeutung für das Auftreten kombinatorischer Restriktionen war, als die Struktur (bzw. die Dauer der Stimmhaftigkeit) innerhalb der

¹ Für das Verfassen des *awk*-Skriptes danke ich Benno Peters.

(Akzent-)Silben. Es wurden daher diejenigen kontextuellen Bedingungen als vergleichbar für die Realisierung der 10 Mustertypen definiert, die eine gleiche Anzahl unbetonter Silben zwischen den umliegenden Akzentsilben aufwiesen.

Dieses Kriterium bietet eine *hinreichende* Vergleichbarkeit der kontextuellen Rahmenbedingungen. Es strebt keine absolute Vergleichbarkeit an und lässt daher genug Freiraum, um (kompensatorische) Eigenschaften der realen Sprachproduktion aufzeigen zu können, die im Perzeptionsexperiment nicht abgedeckt werden konnten (z.B. Veränderungen in der Dauerstruktur, vgl. 3.1) und die den Sinn dieser Korpusanalyse ausmachen. Anhand dieses Kriteriums ergaben sich die folgenden Zuweisungen:

- Alle im *Kiel Corpus* gefundenen Realisierungen der relevanten 10 Intonationsmustertypen, die mehr als eine unbetonte Silbe zwischen den Akzentsilben aufwiesen, wurden Satz 1 (*“immer wieder nein”*) der Resyntheseriehe zugewiesen (d.h. sie wurden als *”unter vergleichbaren kontextuellen Bedingungen wie das entsprechende Muster in der Resyntheseriehe realisiert”* angesehen).
- Alle im *Kiel Corpus* gefundenen Produktionen der 10 Mustertypen mit genau einer unbetonten Silbe zwischen den beiden Satzakkenten wurden Satz 2 der Resyntheseriehe (*“immer nein”*) zugeordnet.
- Alle im *Kiel Corpus* gefundenen relevanten Muster ohne unbetonte Silbe zwischen beiden Satzakkenten wurden Satz 3 und 4 der Resyntheseriehe (*“oft nein”* bzw. *“oft nichts”*) zugeordnet.

Die Ermittlung der Anzahl unbetonter Silben in den Produktionsdaten des *Kiel Corpus* erfolgte anhand des Perzeptionseindrucks der jeweiligen Äußerung und nicht auf Basis der (kanonischen) Transkription der Lautsegmente (siehe hierfür Kohler et al. 1995). Dadurch konnten die Einflüsse jeglicher Form von Elision, Assimilation und Reduktion als besondere Charakteristika spontansprachlicher Daten auf die Anzahl und Grenzen von Silben berücksichtigt werden.

3.4 Ergebnisse der Korpusanalyse

Tabelle 5 zeigt die zweidimensionale, absolute und relative Häufigkeitsverteilung der im *Kiel Corpus of Spontaneous Speech* (IPDS 1995; IPDS 1996) enthaltenen 110 Realisierungen der 10 Intonationsmustertypen auf

Tabelle 5: Ergebnisse der Korpusanalyse: Absolute und relative (in %, gerundet auf ganze Zahlen) Häufigkeitsverteilung der 110 aus dem *Kiel Corpus of Spontaneous Speech* extrahierten Intonationsmuster über die relevanten 10 Mustertypen. Gesternte Werte indizieren Abweichungen gegenüber den Ergebnissen der Resynthesereihe aus Tabelle 4.

Intonation	Satz 1	Satz 2	Satz 3 und 4	Fälle gesamt
2(1. 2)	21 (19%)			21 (19%)
2) 2. 2)	8 (7%)			8 (7%)
2) 2. 2^	5 (5%)	9 (8%)	2* (2%)	16 (15%)
2) 2. 2(2 (2%)	2 (2%)		4 (4%)
2^ 2. 2)	6 (5%)			6 (5%)
2^ 2. 2^	24 (22%)	3* (3%)		27 (25%)
2^ 2. 2(8 (7%)	1 (1%)		9 (8%)
2(2. 2)				0 (0%)
2(2. 2^	9 (8%)			9 (8%)
2(2. 2(7 (6%)	3 (3%)		10 (9%)
Gesamt	90 (82%)	18 (16%)	2 (2%)	110 (100%)

die Sätze 1-4 des Perzeptionsexperimentes. Gesternte Zahlen markieren von Sprechern produzierte Gipfelsequenzen, die unter vergleichbaren Bedingungen in der systematischen Resynthesereihe als nicht realisierbar eingestuft worden sind. Letzteres wird unter 4.1 ausführlich diskutiert.

Von primärem Interesse im Rahmen dieser Untersuchung ist die horizontale Häufigkeitsverteilung. Dies ist die Verteilung der 110 erfassten Gipfelsequenzen über die drei Restriktionsbedingungen, die hinsichtlich der integrierten unakzentuierten Silben mit den Sätzen 1, 2 und 3-4 der vorangehenden Resynthesereihe korrespondieren. Hierfür ergibt sich aus Tabelle 5, dass in der Mehrzahl von 90 der 110 zweigipfligen Muster (dies entspricht 82%) die beiden Akzentsilben durch mindestens zwei unakzentuierte Silben getrennt waren. Lediglich für die Abfolge aus frühem und mittlerem Gipfel (/2) 2. 2^ /) überwiegen diejenigen Realisierungen, in denen die beiden Akzentsilben durch weniger als zwei unbetonte Silben separiert waren und die dementsprechend mit Satz 2 und 3-4 der Resynthesereihe korrespondieren. Dieses Muster ist zudem das einzige, für das in zwei Fällen unmittelbar aufeinander folgende Akzentsilben vorgefunden werden konnten. Generell traten Fälle, in denen zwischen den Satzakkentsilben der Stufe /2/ weniger als zwei unakzentuierte Silben standen, für Intonationsmuster mit frühem Gipfel an zweiter Stelle in dieser Analyse nicht auf.

In Verbindung mit der vertikalen Verteilung der Häufigkeiten über die 10 Mustertypen ist von Bedeutung, dass für Abfolgen aus spätem und frühem Gipfel, die mit /2./ verbunden sind, unabhängig von der Anzahl unbetonter Silben keine einzige Realisierung gefunden werden konnte. Für die gleiche Sequenz von Gipfelkategorien, jedoch mit /1./ konkateniert, konnten hingegen 21 Fälle im ausgewerteten Bereich des *Kiel Corpus of Spontaneous Speech* gefunden werden, die allesamt mindestens zwei unakzentuierte Silben zwischen den zwei Satzakkenten einschließen und daher mit Satz 1 der Resynthesereihe korrespondieren.

Diese asymmetrische Häufigkeitsverteilung ist auch deswegen bemerkenswert, weil eine weitere, hier nicht dargestellte Suchabfrage ergeben hat, dass die in dieser Untersuchung relevanten Gipfelsequenzen (über Satzakkenten der Stufe /2/) mit /1./-Konkationen generell viel häufiger vorkamen als mit /2./-Konkationen. Ungeachtet der Anzahl der unbetonten Silben zwischen den beiden Satzakkenten enthält der betrachtete Ausschnitt des *Kiel Corpus* z.B. 1.163 Realisierungen aller 9 Gipfelverbindungen mit /1./ (einschließlich der 21 Fälle von /2(1. 2)/ aus Tabelle 5), aber nur 89 Produktionen aller 9 Gipfelkombinationen mit /2./. Ferner konnten 684 Gipfelverbindungen mit /0./ gefunden werden, wobei zu berücksichtigen ist, dass diese Zahl nur mit knapp der Hälfte der Gipfelkombinationen (4 Stück, vgl. Tabelle 1) zustande gekommen ist, die für Muster mit Einbuchtung theoretisch möglich sind.

Auch innerhalb der durch /2./ verbundenen Gipfelkategorien ist die Verteilung auf die 9 Mustertypen sehr heterogen. So kamen durch Abfolgen von zwei mittleren sowie von frühen und mittleren Gipfeln mit 27 bzw. 16 Realisierungen (dies entspricht 25% bzw. 15%) am häufigsten vor. Die umgekehrte Abfolge aus mittlerem und frühem Gipfel wurde hingegen mit insgesamt 6 Fällen (6%) vergleichsweise selten realisiert. Noch seltener waren Abfolgen aus frühem und spätem Gipfel mit insgesamt nur 4 Fällen, dies entspricht 4%.

4 Interpretation und Diskussion der Ergebnisse

4.1 Vergleich der Ergebnisse aus Perzeptionsexperiment und Korpusanalyse

Als wichtigster Befund eines Vergleichs ist festzuhalten, dass die Anzahl der in Tabelle 5 durch Sterne markierten Werte, die Diskrepanzen zwischen

den Ergebnissen des Perzeptionsexperimentes und der Korpusanalyse repräsentieren, sehr gering ausfällt. Insofern ist innerhalb der 10 betrachteten Gipfelkategorien mit Einbuchtung die Bandbreite tatsächlicher Produktionen bis auf wenige Ausnahmen genau für die Kombinationen aus Mustertyp und Realisierungskontext lückenhaft, die bereits im Perzeptionsexperiment als nicht umsetzbar herausgekommen sind.

Ein besonders nennenswertes Beispiel für diese Übereinstimmungen ist, dass die Abfolge aus spätem und frühem Gipfel mit einer Konkatenation bis an den unteren Bereich des sprecherindividuellen Stimmumfangs, /2(2. 2)/, die in der Resyntheseriehe als kontextunabhängig nicht realisierbar hervorgegangen ist (vgl. Tabelle 4), auch in den mehrstündigen Produktionsdaten des *Kiel Corpus* nicht ein einziges Mal vorkommt. Werden der späte und der frühe Gipfel hingegen durch eine /1./-Konkatenation verbunden, wurden sie im Perzeptionsexperiment unter bestimmten Bedingungen als realisierbar eingestuft und traten auch in der Korpusanalyse unter eben diesen Bedingungen auf.

Die Diskrepanzen zwischen den Ergebnissen der Experimente beziehen sich auf zwei durch /2./ verbundene Gipfelkategorien, nämlich die Abfolgen aus frühem und mittlerem Gipfel sowie aus zwei mittleren Gipfeln. Für den ersten Mustertyp (/2) 2. 2[^] /) konnten im Korpus zwei Realisierungen gefunden werden, in denen die satzakzentuierten Silben unmittelbar aufeinander folgen. Eine derartige Realisierung hat im Perzeptionsexperiment das natürlichsprachliche Kriterium erfüllt (vgl. Tabelle 4). Gleiches gilt für die Abfolge zweier mittlerer Gipfel (/2[^] 2. 2[^] /), wenn dazwischen lediglich eine unbetonte Silbe liegt (wie in Satz 2). Dennoch kamen im analysierten Bereich des *Kiel Corpus* 3 solcher Fälle vor.

Dass die genannten 5 Intonationsmuster (zwei Fälle von /2[^] 2. 2[^] / und 3 Fälle von /2) 2. 2[^] /) die auf der Resyntheseriehe begründeten kombinatorischen Restriktionen durchbrochen haben, kann zumindest teilweise auf Besonderheiten ihrer Distribution zurückzuführen sein. So stehen alle 5 Gipfelsequenzen in turnfinaler Position. Durch damit verbundene Längungsphänomene (siehe hierzu z.B. Cruttenden 1986) sind die Dauern der Gesamtmuster, gemessen anhand der F0-Wendepunkte vom Beginn des Anstiegs zum ersten Gipfel bis zum Ende des Abstiegs vom zweiten Gipfel, bedeutend länger als diejenigen in den Sätzen 2-4 der Resyntheseriehe, wobei die Längung in besonderem Maße, aber nicht ausschließlich, die Realisierung der zweiten Akzentkontur betrifft. Dass sich diese Längung hier weiter als in den meisten anderen Fällen in die Phrase

hinein ausdehnt, kann mit zusätzlichen Kompensationseffekten zur Realisierung der Intonationsmuster zu tun haben. Darüber hinaus sind die F0-Maxima in dieser turnfinalen Position bereits so tief herabgesunken, dass es im Gegensatz zu den Mustern in der Resyntheserie, in denen die Werte der F0-Maxima für alle Muster in etwa konstant gehalten wurden (vgl. 2.2.3), keiner umfangreichen Einbuchtung zwischen den beiden Gipfeln mehr bedurfte, um zu einer /2./-Konkatenation im Sinne der derzeitigen PROLAB-Definition (Peters und Kohler 2004) zu kommen. Tatsächlich war die F0-Differenz zu beiden Gipfelmaxima in diesen Fällen so gering (10-20Hz), dass jede Anhebung des F0-Minimums die Gesamtkontur direkt von einer Gipfelsequenz (mit /2./-Konkatenation) in ein Hutmuster (mit /0./-Konkatenation) übergehen ließ (Diese Aussage basiert auf dem Gehörseindruck informeller Resynthesen). Die reduzierten F0-Differenzen und die gleichzeitig erhöhten Dauerwerte setzen die nötige F0-Dynamik zur Realisierung der Intonationsmuster gegenüber den Bedingungen in der Resyntheserie deutlich herab und begünstigen so das Auftreten dieser Diskrepanzen.

Diese Beispiele führen die Rigidität vor Augen, die für die Parametermanipulation in der Resyntheserie galt, und zeigen so die Notwendigkeit der nachgeschalteten korpusempirischen Analyse. Dennoch ist es wichtig zu betonen, dass der hohe Grad an Kongruenz zwischen den Ergebnissen beider Experimente die Validität der in der Resyntheserie gefundenen kombinatorischen Restriktionen erhärtet.

Die 5 davon abweichenden Intonationsmuster stehen zudem nur kombinatorischen Restriktionen des Perzeptionsexperimentes entgegen, die auf dem schwächsten aller Kriterien basieren, dem natürlichsprachlichen Kriterium (N, vgl. 2.2.4 und Tabelle 4). In der Triade aus akustischem, perzeptorischem und natürlichsprachlichem Kriterium ist das akustische das schwerwiegendste, da es das perzeptorische und das natürlichsprachliche Kriterium mit einschließt. Hierauf folgt das perzeptorische Kriterium, welches das natürlichsprachliche beinhaltet. Letzteres steht ganz unten in der skizzierten Hierarchie und bedeutet darüber hinaus, dass die angelegten phonologischen Kategorien bereits perzipiert wurden, nur nicht in einer angemessenen Weise. Es ist anzunehmen, dass derartige Restriktionen am leichtesten von einem Sprecher z.B. durch Veränderungen in der Dauerstruktur (wie etwa die Dehnung stimmhafter Segmente) aufgelöst werden können.

Auf der Basis dieser Validität können nun Aussagen hinsichtlich der zwei Zielsetzungen dieser Untersuchung in den nachfolgenden Abschnitten getroffen werden.

4.2 Der Beitrag der Ergebnisse zur ersten Zielsetzung

Die Ergebnisse dieser Untersuchung können die vorhandenen strukturellen Restriktionen in der Phonotaktik des KIM bestätigen und ergänzen. So besagen die zu den strukturellen Restriktionen gehörigen *timing constraints* (vgl. Kohler 1991a), dass eine Kombination aus mittlerem oder spätem Gipfel an erster Akzentstelle und frühem Gipfel an zweiter Akzentstelle bei weniger als zwei unbetonten Silben dazwischen (in dieser Untersuchung ab Satz 2, “immer nein”, gegeben) nicht durch eine Einbuchtung (/1./ oder /2./) verbunden werden kann. Hinsichtlich der Abfolge aus spätem und frühem Gipfel wird dies durch die Ergebnisse der Resyntheseriehe und der Korpusanalyse bestätigt. Eine Abfolge aus mittlerem und frühem Gipfel konnte bei einer /1./-Konkatenation in der Resyntheseriehe jedoch weiterhin realisiert werden (Tab. 4). Dies zeigt, dass die vorhandenen strukturellen Restriktionen des KIM teilweise zu eng gefasst sind und revidiert werden müssen.

Zusätzlich enthalten die Ergebnisse Hinweise auf weitere kontextsensitive Beschränkungen. Tabelle 4 zeigt hierzu, dass bei einer /2./-Konkatenation und weniger als zwei unbetonten Silben zwischen den beiden Satzakkenten alle Kombinationen mit frühem Gipfel an der zweiten Akzentposition ausgeschlossen werden konnten. Des Weiteren waren unter denselben Bedingungen alle Intonationsmuster mit mittlerem Gipfel an zweiter Akzentposition auszuschließen, wenn die Gipfelkategorie an der ersten Akzentposition nicht *früh* war (vgl. Tab. 4). Folgten die zwei Satzakkentsilben unmittelbar aufeinander, musste jede Kombination zweier Gipfelkategorien mit /2./ ausgeschlossen werden. Mit Ausnahme der zuvor unter 4.1 diskutierten 3 Ausnahmen in der Abfolge zweier mittlerer Gipfel, werden diese Hinweise durch die Korpusanalyse weiter erhärtet.

Die Diskrepanz zwischen den bislang im KIM erfassten und den in dieser Untersuchung aufgezeigten strukturellen Restriktionen weist darauf hin, dass die in das KIM integrierte Phonotaktik die Kombinationsmöglichkeiten von Gipfelkonturen kontextsensitiv nicht ausreichend reduziert. Folglich wird möglicherweise die intonatorische Vielfalt (und damit ebenfalls die funktionale Vielfalt), die dem Sprecher unter bestimmten Bedin-

gungen zur Verfügung steht, überschätzt. Sollten diese neuen Erkenntnisse in die phonotaktischen Regularien des KIM aufgenommen werden können (auf die Voraussetzungen hierfür gehe ich unter 4.4 und 4.5 ein), stellt dies eine beträchtliche Erweiterung dieser Regularien dar. Diese Erweiterung führt zu einer adäquateren Abbildung des bedeutungsunterscheidenden Potentials der Intonation in bestimmten Kontexten und ist somit im Sinne der ersten Zielsetzung dieser Arbeit zu bewerten.

4.3 Der Beitrag der Ergebnisse zur zweiten Zielsetzung

Die Ergebnisse der vorliegenden Untersuchung haben eine Reihe von Restriktionen in der Kombination unterschiedlicher intonatorischer Muster über den vier Sätzen dieser Untersuchung aufzeigen können. Diese kombinatorischen Restriktionen beschränken sich auf Muster mit einer Einbuchtung. Hutmuster sind, abgesehen von den *a priori* Ausschlüssen (die im Rahmen der Zielsetzung in der Einleitung genannt wurden), von kombinatorischen Restriktionen nicht betroffen. Da der phonologische Unterschied zwischen zwei Gipfeln mit Einbuchtung (/1./ und /2./) und einem Hutmuster (/0./) nicht in allen Kontexten aufrecht zu erhalten ist, kann ihm keine (ausschließlich) funktional motivierte Steuerung zugrundeliegen. Insofern widersprechen die Ergebnisse dieser Untersuchung der Hypothese von Peters et al. (2005), die eine solche funktionale Steuerung annimmt. Dass im Zusammenhang mit Hutmustern keine kombinatorischen Restriktionen aufgezeigt werden konnten, ermöglicht gleichzeitig die Annahme eines Hutmusters als zeitlich gesteuerte Reduktionsform einer Gipfelsequenz mit Einbuchtung.

Die Ergebnisse zeigen jedoch auch, dass sich die meisten der gefundenen kombinatorischen Restriktionen auf die durch /2./ konkatinierten Gipfelkategorien beschränken. Der Kontrast zwischen Intonationsmustern, deren Gipfelkonturen durch /0./ oder /1./ miteinander verbunden werden, konnte mit Ausnahme der Abfolge von späten und frühen Gipfeln (/2(1. 2)/, Tab. 4 und Tab. 5) aufrecht erhalten werden. Dies gilt selbst für Satz 4, in dem die Intonationsmuster unter der extremen Bedingung unmittelbar aufeinanderfolgender Satzakkzentsilben mit stimmloser Koda (“oft nichts”) generiert wurden.

Insofern lassen die Ergebnisse dieser Untersuchung - mit Einschränkung auf die Konkatenationskategorien /0./ und /1./ - Spielraum für die Annahme eines funktional gesteuerten Unterschiedes zwischen Hutmustern und

Gipfelsequenzen mit Einbuchtung im Sinne der Hypothese von Peters et al. (2005) und den Annahmen des KIM (Kohler 1991a).

Vor dem Hintergrund der zweiten Zielsetzung sind die Ergebnisse nicht geeignet, eine funktionale Steuerung von Hutmustern und Gipfelsequenzen mit Einbuchtung in der gesamten Breite der involvierten Gipfelkonturen eindeutig zu untermauern oder zu widerlegen. In den Ergebnissen deutet sich vielmehr eine differenziertere Sichtweise an, nach der die Vielzahl kombinatorischer Restriktionen in Verbindung mit /2./ einen funktionalen Einsatz dieser Kategorie unwahrscheinlich macht, während die funktionale Steuerung eines Unterschiedes zwischen /0./ und /1./ weiterhin angenommen werden kann.

4.4 Weiterführende Erkenntnisse der Untersuchung zu den Konkatenationsmöglichkeiten

Im Zusammenhang mit der Interpretation der Ergebnisse der Korpusanalyse, die in Tabelle 5 dargestellt sind, muss darauf hingewiesen werden, dass die Verteilung der erfassten Intonationsmuster über die 10 Mustertypen (vertikale Verteilung gemäß 3.4) nicht nur zeitlich bedingte Restriktionen widerspiegelt. Vielmehr ist wahrscheinlich, dass sie auf eine Interaktion zeitlicher und semantischer Faktoren zurückzuführen ist.

Dass das Muster /2(2. 2)/ beispielsweise im analysierten Ausschnitt des *Kiel Corpus of Spontaneous Speech* (IPDS 1995; IPDS 1996) nicht gefunden werden konnte, während für die gleichen Gipfelkategorien mit einer /1./-Einbuchtung 21 Produktionen erfasst wurden, weist deutlich in Richtung zeitlich bedingter kombinatorischer Restriktionen.

In der gleichen Weise ist jedoch nicht erklärbar, warum z.B. eine Kombination aus frühem und spätem Gipfel mit /2./ insgesamt nur 4 mal im Korpus (über Akzenten der Stufe /2/) produziert wurde, während eine Abfolge aus zwei mittleren Gipfeln mit /2./ insgesamt 27 mal im Korpus gefunden werden konnte. Hinsichtlich der Realisierung von /2) 2. 2(/ bilden die phonologischen Merkmale der Gipfel - Abstieg als perzeptorisch ausschlaggebendes Merkmal des frühen Gipfels und Anstieg als perzeptorisch essentielles Merkmal des späten Gipfels - mit dem Vorhandensein einer dazwischenliegenden Einbuchtung eine optimale Verkettung. Durch diese Synergieeffekte ist davon auszugehen, dass das Muster /2) 2. 2(/ nur sehr wenig Zeit in der Realisierung braucht. Damit konform geht, dass /2) 2. 2(/ in der Resynthesereihe erst bei Satz 4 für nicht realisierbar erklärt werden

musste. Dem gegenüber wurde die Sequenz zweier mittlerer Gipfel mit /2./, für die dieses optimale Ineinandergreifen phonologischer Merkmale nicht gilt, in der Resynthesereihe bereits ab Satz 2 als nicht mehr realisierbar eingestuft. Trotz des daraus abzuleitenden größeren Zeitbedarfs war dieses Muster im analysierten Ausschnitt des *Kiel Corpus* (IPDS 1995; IPDS 1996) knapp 7 mal häufiger anzutreffen. Diese divergenten Häufigkeiten in der Realisierung beider Muster im *Kiel Corpus* müssen durch die unterschiedliche Kompatibilität der Semantik der beteiligten Gipfelkonturen erklärt werden.

Kohler (1991a) beschreibt die Bedeutung des frühen, mittleren und späten Gipfels wie folgt: Der frühe Gipfel drückt die Abgeschlossenheit einer Argumentation bzw. einer Handlung aus oder kennzeichnet bekannte Information im Rahmen des jeweiligen Dialoges. Der mittlere Gipfel hingegen betont das Neue und eignet sich so zur Einleitung einer Argumentation oder Handlung bzw. zum Anzeigen einer bislang unbekanntes Information. Der späte Gipfel bildet keine antonyme Beziehung zum mittleren Gipfel aus, sondern ergänzt dessen Bedeutung, indem er das Neue in den Gegensatz zur Erwartung des Sprechers stellt.

Aus dieser Perspektive erscheint eine abschließende Feststellung, gefolgt von Überraschung (in derselben Phrase) keine sinnvolle Sequenz zu bilden, wodurch sich die geringen Fallzahlen für /2) 2. 2(/ trotz der phonetischen Synergien (indiziert in Tab. 4) innerhalb des Musters begründen lassen. Umgekehrt bilden Sequenzen entweder aus Überraschung oder Feststellung eine inhaltliche Kohärenz (in derselben Phrase) aus, was erklärt, warum Muster wie /2(2. 2(/ und /2) 2. 2)/ trotz eines erhöhten Zeitbedarfs in ihrer phonetischen Realisierung, zu größeren Fallzahlen im Korpus geführt haben (siehe Tabelle 5). Letzteres ergab sich auch in der exhaustiveren Korpusuntersuchung von Peters et al. (2005).

Die skizzierten Einflüsse auf die vertikale Häufigkeitsverteilung in der Korpusanalyse (Tabelle 5) werden primär durch die Semantik der Gipfelkategorie bestimmt und gelten daher prinzipiell unabhängig von der Stärke der Einbuchtung (/1./ oder /2./). Infolge der Konzentration der Korpusanalyse auf /2./-Fälle können ähnliche Beispiele für /1./ an dieser Stelle jedoch nicht präsentiert werden. Eine informelle Gegenüberstellung der relativen vertikalen Häufigkeitsverteilungen der /1./- und /2./-Fälle aus dem *Kiel Corpus* (IPDS 1995; IPDS 1996) hat jedoch signifikante Parallelen ergeben, die von der relativen Häufigkeitsverteilung in Hutmuster (/0./) integrierter Gipfelkategorien abweicht.

Bezüglich der Häufigkeitsverteilung der 10 in der Korpusanalyse betrachteten Mustertypen über die Sätze 1, 2 und 3-4 (horizontale Häufigkeitsverteilung gemäß 3.4) sind für Verbindungen zweier Gipfel durch /2./ zahlreiche Lücken zu beobachten (siehe Tabelle 5). Damit konform, haben die Ergebnisse der Resynthesereihe (Tab. 4) im Zusammenhang mit der /2./-Konkatenationsmöglichkeit umfangreiche kombinatorische Restriktionen ergeben. Gleichzeitig konnten jedoch phonologische Unterschiede zwischen Hutmustern und zwei Gipfelkonturen mit einer /1./-Einbuchtung in nahezu allen Kontexten der Sätze 1-4 aufrecht erhalten werden.

Dies wirft die Frage auf, inwieweit die von PROLAB eingeführte Konkatenationsmöglichkeit /2./, die solchen umfangreichen Beschränkungen unterliegt, als phonologisch relevante Eigenschaft einer Gipfelkontur neben /1./ geführt werden muss.

Perzeptionsexperimente zur Kategorisierung phonologisch relevanter Eigenschaften von Intonationskonturen (wie es sie für die Gipfelkonturen gibt, siehe Kohler 1991b; Niebuhr 2003) wurden für die Konkatenationsmöglichkeiten von Gipfelkonturen im Deutschen bislang nicht durchgeführt. Daher müssen für die Diskussion der zuvor aufgeworfenen Frage zum einen qualitative Unterschiede in den semantisch-pragmatischen Funktionen der Konkatenationsmöglichkeiten /0./, /1./ und /2./ herangezogen werden. Hierbei können einschneidende funktionale Differenzen als Argument für eine unterschiedliche phonologische Einordnung verwendet werden. Zum anderen sollte das Potential der paradigmatischen Kontrastierung von /1./ und /2./ Berücksichtigung finden.

Was die funktionale Seite dieser Frage anlangt, so haben sich sowohl Kohler (1991a), als auch Peters et al. (2005) mit den semantisch-pragmatischen Eigenschaften der Konkatenationsmöglichkeiten auseinandergesetzt. Kohler (1991a) assoziiert die stärkere phonetische Kohäsion eines Hutmusters (direkte und kürzeste Verbindung beider Gipfelkonturen) gegenüber einer Sequenz zweier Gipfelkonturen mit Einbuchtung mit einer stärkeren inhaltlichen Kohäsion zwischen den beiden akzentuierten Komponenten der Äußerung. Dies lässt sich am Beispiel der Äußerung “Samstag oder Sonntag” veranschaulichen. Bei der Realisierung eines Hutmusters über den satzakzentuierten lexikalischen Akzentsilben in “Samstag” und “Sonntag” wird die Äußerung semantisch ähnlich mit der Äußerung “Wochenende”, während der Sprecher bei der Realisierung zweier Gipfelsequenzen mit Einbuchtung über “Samstag” und “Sonntag” beide Tage als gleichrangige Alternativen hervorhebt und damit inhaltlich trennt.

Verläuft die Akzentstruktur einer Äußerung parallel zu dessen Thema-Rhema-Struktur, so integriert das Hutmuster Thema und Rhema in eine größere semantisch-pragmatische Einheit. Dabei unterstützt die Verbindung beider Gipfelkonturen auf hohem F0- bzw. Tonhöhenniveau, ebenso wie der Anstieg zum ersten Gipfel, die offene Einstellung des Sprechers gegenüber dem Thema (dem ersten Satzakkzent). Der Wegfall des erneuten Anstiegs zum zweiten Satzakkzent (mittlerer oder früherer Gipfel) verstärkt die Bedeutung des Abgeschlossenen für das Rhema. Das Hutmuster ist dadurch ideal geeignet, um ein Thema aufzugreifen und - aus Sicht des Sprechers - durch offensichtliche, nicht diskussionsbedürftige Informationen zu kommentieren. Bei einer Sequenz zweier Gipfelkonturen mit Einbuchtung werden Thema und Rhema hingegen phonetisch und inhaltlich eigenständiger als im Hutmuster. Diese Eigenständigkeit wird nicht zuletzt dadurch etabliert, dass die Einbuchtung einen F0- bzw. Tonhöhenabstieg in der ersten Gipfelkontur kreiert, der das Thema als gegeben und fixiert hinstellt. Ein gleichzeitiger F0- bzw. Tonhöhenanstieg in der zweiten Gipfelkontur verstärkt die Bedeutung des Offenen für das Rhema. Der Sprecher erhält dadurch im Gegensatz zum Hutmuster die Möglichkeit, das Rhema einer Äußerung mit etwas Vorangehendem inhaltlich zu kontrastieren.

Die Interpretationen der Sprechmelodie von Peters et al. (2005) gehen mit diesen durch Kohler formulierten Unterschieden zwischen Hutmuster und Gipfelsequenz mit Einbuchtung konform. Sie ergänzen diesen funktionalen Unterschied der "Klammerung der Komponenten" um zwei weitere funktionale Unterschiede. So signalisiert ein Sprecher durch ein Hutmuster, dass er die Information als bekannt und selbstverständlich einstuft, während über Sequenzen zweier Gipfel mit Einbuchtung das Gegenteil zum Ausdruck gebracht werden kann. Peters et al. bemerken des Weiteren, dass Hutmuster eher eine emotionale Neutralität des Sprechers zum Gesagten hervorheben, während Gipfelsequenzen mit Einbuchtung nachdrücklich und auffordernd wirken.

Die skizzierten funktionalen Unterschiede von Kohler (1991a) und Peters et al. (2005) sind als gewichtige Argumente für die Unterscheidung von /0./ gegenüber /1./ und /2./ auf phonologischer Ebene auch ohne die Durchführung von Perzeptionsexperimenten anzusehen. Die genannten Autoren nennen jedoch keine gleichermaßen einschneidenden funktionalen Unterschiede hinsichtlich der Stärke der Einbuchtung (/1./ vs. /2./). Die Aussagen von Peters et al. (2005) deuten im Gegenteil darauf hin, dass graduelle Unterschiede in der Stärke der Einbuchtung lediglich zu gra-

duellen Abstufungen im beschriebenen funktionalen Gegensatz von Hutmustern und Gipfelsequenzen mit Einbuchtung führen. Die Konkatenationsmöglichkeiten /1./ und /2./ sind somit nicht mit unterschiedlichen semantisch-pragmatischen Funktionen assoziierbar, sondern erfüllen in unterschiedlicher Deutlichkeit ein und dieselbe Funktion. Im Zusammenhang mit dem semantischen Einfluss auf die Häufigkeit der Produktion der 22 in dieser Arbeit behandelten Intonationsmuster (siehe oben in diesem Abschnitt und 3.4) wurde zudem erwähnt, dass die Verteilung der in Hutmuster integrierten Gipfelkategorien signifikant von den vergleichbar verlaufenden Häufigkeitsverteilungen der Gipfelkategorien mit Einbuchtung abweicht. Diese (informelle) Beobachtung deutet ebenfalls an, dass eine einschneidende funktionale Differenzierung lediglich zwischen /0./ gegenüber /1./ und /2./ besteht.

Ein weiteres Argument gegen eine phonologische Differenzierung zwischen /1./ und /2./ ist ihr stark restriktives Potential zur paradigmatischen Kontrastierung. In vielen Fällen ist die Ursache hierfür das im Vergleich zu /1./ stark eingeschränkte Vorkommen von /2./, wie es sowohl die Perzeptionsexperimente als auch die Korpusanalyse gezeigt haben (vgl. 2.3 und 3.4). Dies wird besonders deutlich bei einem Vergleich der Gesamthäufigkeiten produzierter Muster mit /1./ oder /2./ aus dem *Kiel Corpus* (IPDS 1995; IPDS 1996), bei dem 89 gefundene Gipfelverbindungen mit /2./ 1.163 Gipfelverbindungen mit /1./ gegenüberstehen (vgl. 3.4).

Zahlreiche Einzelfallstudien im Rahmen der Korpusanalyse weisen jedoch darauf hin, dass diese Sichtweise, in der /2./ der limitierende Faktor ist, zu einseitig ist. Ebenso gibt es den umgekehrten Fall: In den wenigen Kontexten, in denen /2./ vorkommt, kann oft eine Kontrastierung mit /1./ nicht erfolgen. Dies hängt damit zusammen, dass die überwiegende Zahl der im *Kiel Corpus* gefundenen Gipfelsequenzen mit /2./ in phrasen- oder turnfinaler Position produziert wurde. Dies gilt insbesondere für die Mustertypen /2) 2. 2(/, /2) 2. 2^ /, /2^ 2. 2^ / und /2^ 2. 2)/. In dieser Distribution ist das globale F0-Niveau auf ein /2./ Niveau abgesunken und der globale F0-Verlauf stark eingeebnet. Beides führt dazu, dass das F0-Minimum zwischen den Gipfeln nicht auf ein /1./ Niveau angehoben werden kann, bevor das Gesamtmuster in ein Hutmuster übergeht. Folglich können in diesen Fällen /1./ und /2./ keine paradigmatische Opposition bilden, die eine Grundvoraussetzung für die Ausprägung eines funktionalen Unterschiedes und damit für das Postulat eigenständiger phonologischer Kategorien ist.

Zusätzlich zeigen diese Fälle, dass /2./ nicht mit einer stärkeren Einbuchtung zwischen den umliegenden Gipfeln als /1./ - das heisst mit einer größeren F0-Differenz - einher gehen muss. Im Gegenteil kann /2./, in Abhängigkeit von der Höhe der angrenzenden F0-Gipfel, sogar einen sehr viel geringeren F0-Umfang kennzeichnen als /1./. Vorausgesetzt, dass die zuvor skizzierten skalaren funktionalen Unterschiede in Gipfelsequenzen mit Einbuchtung an der Stärke der Einbuchtung zu den umliegenden Gipfeln hängen und nicht an deren absoluter Tiefe im Frequenzraum, dann erscheint eine funktional orientierte Etikettierung, wie sie das KIM für PROLAB vorgibt, auf der Basis der derzeitigen Definition von /1./ und /2./ nicht durchführbar.

Die von beiden Seiten entscheidend eingeschränkte Möglichkeit der Bildung paradigmatischer Oppositionen zwischen /1./ und /2./ und die Tatsache, dass diese Unterscheidung nicht mit einer funktionalen Divergenz einher geht, sondern eher skalare Unterschiede in derselben Funktion wiedergibt, legt somit nahe, die mit PROLAB eingeführte Subkategorisierung der Konkatenation zweier Gipfelkonturen mit Einbuchtung in /1./ und /2./ zu revidieren und damit zum Ausgangspunkt der im KIM angelegten phonologisch relevanten Eigenschaften von Gipfelkonturen zurückzukehren. Zumindest jedoch sollten die Kriterien für die Unterscheidung zwischen /1./ und /2./ überprüft werden. So erscheint es wenig sinnvoll, phonologische Eigenschaften der Gipfelkontur(en) losgelöst davon an der absoluten Lage von F0 im Frequenzraum zu definieren. Eine phrasenmediale Definition von /1./ und /2./, die sich unmittelbar an den Merkmalen der zugehörigen Gipfel orientiert (z.B. an der zurückgelegten Distanz zum Gipfelmaximum), ist nicht nur aus der phonologischen Theorie des KIM heraus adäquater, sondern kann auch eine funktional eineindeutige Etikettierung der relevanten Intonationsmerkmale gewährleisten.

Durch die Aufgabe der Differenzierung zwischen /1./ und /2./ würde der unter 4.2 geforderte Ausbau der Phonotaktik weitgehend obsolet werden. Ferner würden sich hierdurch die Evidenzen gegen die Annahme einer funktional motivierten Steuerung des Unterschiedes zwischen Hutmuster und Gipfelsequenz mit Einbuchtung auf ein einziges Intonationsmuster (/2(1. 2)/ in Satz 2-4) reduzieren.

4.5 Hinweise zur Aussagekraft der Ergebnisse

Im Rahmen von 3.1 wurden bereits viele Aspekte angesprochen, vor deren Hintergrund die Gültigkeit der Ergebnisse des Perzeptionsexperimentes kritisch hinterfragt werden sollten. Ein weiterer Aspekt, der an dieser Stelle ergänzt werden muss, besteht in der Beurteilung der Intonationsmuster durch nur einen geschulten Hörer (on). Obwohl dessen Ergebnisse durch einen zweiten Hörer (tkh) im *interobserver agreement* bestätigt wurden, wäre die Durchführung formeller Perzeptionsexperimente mit einer ausreichend umfangreiche Stichprobengröße nötig gewesen.

Dies gilt vor allem im Zusammenhang mit der Wahrnehmung und Beurteilung prosodischer Phänomene (Akzentstufen und phonologische Intonationseinheiten). Die Erfahrung des Autors zeigt hierzu, dass die Beurteilung von Akzent und Intonation großen interindividuellen Streuungen unterlegen sein kann und dass die Urteile geschulter und naiver Hörer beträchtlich voneinander abweichen können. Dies mag damit zusammenhängen, dass individuelle Fokussierungen auf bestimmte *cues* einer multifaktoriell kodierten linguistischen Einheit im Bereich der Prosodie anders zum Tragen kommen als im segmentellen Bereich.

Hinsichtlich der Validität bzw. Reliabilität der Korpusanalyse müssen die kleinen Fallzahlen für die jeweiligen Kombinationen von Mustertyp und Realisierungskontext angesprochen werden, die zusammengefasst in Tabelle 5 dargestellt sind. Selbst über alle Satzkontexte hinweg betrachtet sind für die 10 relevanten Mustertypen nur in einigen Fällen insgesamt mehr als 10 Beispiele im Korpus gefunden worden. Bei deren Verteilung auf die Sätze 1-4 werden die Häufigkeiten in den Zellen entsprechend noch einmal geringer. Von diesen Häufigkeiten ausgehend, kann es nicht ausgeschlossen werden, dass eine deutlich größere Gesamtstichprobe als der hierfür verwendete Ausschnitt aus dem *Kiel Corpus of Spontaneous Speech* (IPDS 1995; IPDS 1996) Fallbeispiele für bislang leere Zellen hervorbringt und somit die Diskrepanzen zu den Ergebnissen der Resyntheserie erhöht. Andererseits zeigen die 21 gefundenen Realisierungen von /2(1. 2)/, dass eine relativ große Fallzahl nicht zwangsläufig zu einer heterogenen Verteilung über die Sätze 1-4 führen muss (siehe Tabelle 5). Dennoch besteht die Möglichkeit, dass die durchgeführte Korpusanalyse die Anzahl kombinatorischer Restriktionen unterschätzt.

In diesem Zusammenhang muss auch angesprochen werden, dass 12 Intonationsmustertypen nicht in die Korpusanalyse mit einbezogen wur-

den. Theoretisch ist es denkbar, dass nicht für alle Kontextbedingungen, die mit den Sätzen 1-4 der Resynthesereihe vergleichbar sind, Realisierungen dieser Muster im Korpus gefunden werden können. Dies wäre ein starkes Indiz auf bislang unentdeckte kombinatorische Restriktionen. Die Resultate der Resynthesereihe haben jedoch für diese 12 Mustertypen keine kombinatorischen Restriktionen ergeben. Da die Resynthesereihe durch das *ceteris paribus*-Prinzip wissenschaftlicher Experimente in ihrer rigiden Parameterstruktur als *worst case*-Szenario für die Realisierung der Intonationsmuster zu bewerten ist, sind *a priori* für die 12 nicht analysierten Mustertypen keine kombinatorischen Restriktionen zu erwarten. Hinzu kommt, dass die (informell beobachtete) besonders hohe Fallzahl für jeden dieser 12 Mustertypen, die mindestens einige Dutzend, oft jedoch einige hundert Produktionen umfasst, dagegen spricht, dass leere Zellen für die Kombinationen aus Mustertyp und Realisierungskontext (Satz 1-4) zurückbleiben.

Ebenfalls sollte bei all der genannten Kritik nicht vernachlässigt werden, dass die Ergebnisse in dieser Untersuchung ein sehr konsistentes Bild zum kontextsensitiven Auftreten kombinatorischer Restriktionen zeichnet, das mit zahlreichen funktionalen und modellbezogenen Überlegungen zur Intonation konform geht. Dies sollte als ein entscheidender Beitrag zur Validität und Reliabilität der Ergebnisse dieser Untersuchung interpretiert werden.

5 Zusammenfassung und Ausblick

Die Ergebnisse dieser Untersuchung sind wie folgt zusammenzufassen:

- (a) Restriktionen bei der Kombination zweier Gipfelkonturen beschränken sich kontextunabhängig auf Konkationen mit Einbuchtung. Hutmuster sind von Restriktionen nicht betroffen, abgesehen von den *a priori* vorgenommenen Ausschlüssen.
- (b) Die kombinatorischen Restriktionen konzentrieren sich auf Gipfelkonturen, die durch eine Einbuchtung an den unteren Bereich des sprecherindividuellen Stimmumfangs (/2./) miteinander konkateniert sind. Die sich daraus ergebenden umfangreichen Einschränkungen für die Ausbildung paradigmatischer Oppositionen zwischen /1./ und /2./ in Verbindung mit nicht signifikanten funktionalen Unterschieden

zwischen diesen Kategorien sprechen dafür, diese durch PROLAB eingeführte Subkategorisierung einer Konkatenation mit Einbuchtung phrasenmedial aufzugeben und zu der ursprünglich binären phonologischen Opposition des KIM (Einbuchtung ja oder nein) zurückzukehren.

- (c) Wenn die sich aus (b) ergebende Konsequenz umgesetzt wird, dann unterstützen die Ergebnisse die von Kohler (1991a) und Peters et al. (2005) angenommene funktional motivierte Steuerung zwischen Hutmuster und Gipfelsequenz mit Einbuchtung. Hutmuster sind somit nicht als zeitlich bedingte Residuen eines zweigipfligen Musters zu sehen, sondern phänomenal und funktional eigenständig. Die einzige Gegenevidenz hierzu beschränkt sich auf das Muster /2(1. 2)/ in den Sätzen 2-4.
- (d) Die Phonotaktik des KIM ist an einigen Stellen zu restriktiv und an anderen Stellen nicht restriktiv genug. Die Annahme, dass eine Kombination aus mittlerem und frühem Gipfel, mit Einbuchtung konkateniert, bei weniger als zwei unbetonten Silben dazwischen nicht möglich ist, hat sich in dieser Untersuchung nicht bestätigt. Stattdessen hat sich gezeigt, dass bei einer /2./-Konkatenation und weniger als zwei unbetonten Silben zwischen den beiden Satzakzenten alle Kombinationen mit frühem Gipfel an zweiter Akzentstelle auszuschließen sind. Die Ergebnisse der Resynthesereihe enthalten ferner Hinweise darauf, dass dies für alle Kombinationen mit mittlerem Gipfel an zweiter Akzentstelle ebenfalls gelten könnte, wenn die Gipfelkontur an der ersten Akzentstelle gleichzeitig nicht früh ist. Sollten die Schlussfolgerungen aus (b) gelten, entfallen diese strukturellen Restriktionen jedoch.
- (e) Insgesamt betrachtet neigt das Perzeptionsexperiment dazu, die Bandbreite kombinatorischer Restriktionen zu überschätzen, während die Korpusanalyse möglicherweise gerade umgekehrt zu wenige dieser Restriktionen erfasst.

Zum Ausblick gehört es, auf die auffälligen Parallelitäten der Entwicklung der Forschungsgegenstände im Bereich der segmentellen und supra-segmentellen (hier insbesondere intonatorischen) Phonetik hinzuweisen.

Die segmentelle phonetische Forschung ist von einer rein deskriptiven Beschreibung des phonetischen Materials einer Sprache und seiner Pro-

duktion zu einer Systematisierung dieses Materials vor dem Hintergrund diverser phonologischer Theorien gekommen, die überwiegend auf dem Kontrast von Segmenten in isolierten Wortformen gegründet sind. Erst auf dieser Basis sind die aktuellen Bestrebungen möglich, auf der nächsten Metaebene den Blickwinkel über die konkrete Umsetzung einzelner phonematischer Segmente in isolierten Wortformen hinaus zu erweitern und zu einer Beschreibung phonologischer Prozesse auf Äußerungsebene zu gelangen, die zwar auf die isolierten Wortformen der darunter liegenden Ebene referiert, die aber zugleich das Konstrukt diskreter segmenteller Einheiten zunehmend aufgeben muss. Für das Deutsche wurde hierzu bereits umfangreiche Forschungsarbeit geleistet (z.B. Kohler 1990; Kohler 2001; Kohler und Rodgers 2001; Wesener 2001). Hieran sind Theorien gebunden, die die vorgefundene phonetische Variabilität als Resultat der diametralen Faktoren artikulatorischer Ökonomie und perzeptorischer Diskontinuität sehen, der eine soziale Komponente nachgeschaltet sein kann (Kohler 1979; Lindblom 1990).

Die intonatorische Forschung (zum Deutschen) ist noch lange nicht auf diesem Wissensstand angekommen. Allerdings sind die grundlegenden kontrastiven Intonationseinheiten (z.B. Gipfelkonturen) und ihre phonologisch relevanten Eigenschaften weitgehend identifiziert. Auf dieser Basis kann - parallel zu der Entwicklung der segmentellen Forschung - der nächste Schritt getan werden, der die Prozesse und Steuerungsmechanismen untersucht, die bei der Kombination der kontrastiven Intonationseinheiten auf der Äußerungsebene wirksam werden. Wichtige Fragen hierbei sind:

1. Wie sieht die phonetische Variabilität kontrastiver Intonationseinheiten bei ihrer Verkettung auf Äußerungsebene aus?
2. Wie interagieren die einzelnen Intonationseinheiten miteinander (z.B. tonale Sandhi)? Kommt es zu Überlappungen bzw. Interferenzen ihrer phonologisch relevanten Eigenschaften? Welche neuen übergeordneten Einheiten entstehen hierbei?
3. Wie kann die gefundene phonetische Variabilität erklärt werden? Gibt es für die Erklärung Parallelen zu den Theorien der segmentellen Satzphonetik?
4. Inwieweit lassen sich Wechselwirkungen zwischen den phonetischen Variationen der segmentellen und intonatorischen Ebene aufzeigen

und inwieweit reflektieren sie allgemeine Prinzipien der Sprachproduktion?

Diese und weitere Fragen zu stellen und zu untersuchen wird eine große Herausforderung für die Zukunft sein. Die vorliegende Untersuchung kann als ein kleiner Schritt in die genannte Richtung gesehen werden, die Interaktion kontrastiver Intonationseinheiten zu untersuchen, jetzt, wo diese Einheiten selber weitgehend feststehen. Weitere Untersuchungen, die zum Teil auf den Ergebnissen dieser Arbeit aufbauen können, müssen folgen, denn es hat sich im Bereich der segmentellen Forschung eindrucksvoll gezeigt, dass die Erkenntnisse, die aus der Betrachtung von Einzelwörtern gewonnen wurden, für die Beschreibung satzphonetischer Muster nicht ausreichend sind.

Literatur

- Boersma, P. und D. Weenik. *Praat: doing phonetics by computer*. URL: <http://www.fon.hum.uva.nl/praat/>.
- Close, D. B., A. D. Robbins, P. H. Rubin, und R. Stallman (1993). *The GAWK manual*. Free Software Foundation.
- Cruttenden, A. (1986). *Intonation*. Cambridge Textbooks in Linguistics. Cambridge, New York, Oakleigh: Cambridge University Press.
- Ganong, W. (1983). Phonetic categorization in auditory word perception. *Journal of Experimental Psychology* 6, 110–125.
- Gartenberg, R. und C. Panzlaff-Reuter (1991). Production and perception of F0 peak patterns in German. In K. J. Kohler (Hrsg.), *Studies in German Intonation*, AIPUK 25, pp. 29–113. Kiel: IPDS.
- Gussenhoven, C. und A. C. M. Rietveld (1988). Fundamental frequency declination in Dutch: testing three hypotheses. *Journal of Phonetics* 16(3), 355–369.
- IPDS (1995). *The Kiel Corpus of Spontaneous Speech*, Volume 1, CD-ROM#2. Kiel: IPDS.
- IPDS (1996). *The Kiel Corpus of Spontaneous Speech*, Volume 2, CD-ROM#3. Kiel: IPDS.
- IPDS (1997). *The Kiel Corpus of Spontaneous Speech*, Volume 3, CD-ROM#4. Kiel: IPDS.

- Kingdon, R. (1958). *The Groundwork of English Intonation*. London: Longmans.
- Kohler, K. J. (1979). Kommunikative Aspekte satzphonetischer Prozesse im Deutschen. In H. Vater (Hrsg.), *Phonologische Probleme des Deutschen*, Studien zur deutschen Grammatik 10, pp. 13–39. Tübingen: Narr.
- Kohler, K. J. (1990). Segmental reduction in connected speech in German: phonological facts and phonetic explanations. In W. J. Hardcastle und A. Marchal (Hrsg.), *Speech Production and Speech Modelling*, pp. 69–92. Dordrecht: Kluwer.
- Kohler, K. J. (1991a). A model of German Intonation. In K. J. Kohler (Hrsg.), *Studies in German Intonation*, AIPUK 25, pp. 295–360. Kiel: IPDS.
- Kohler, K. J. (1991b). Terminal intonation patterns in single-accent utterances of German: Phonetics, phonology, and semantics. In K. J. Kohler (Hrsg.), *Studies in German Intonation*, AIPUK 25, pp. 115–185. Kiel: IPDS.
- Kohler, K. J. (1997). Modelling prosody in spontaneous speech. In Y. Sagisaka, N. Campbell, und N. Higuchi (Hrsg.), *Computing Prosody: Computational Models for Processing Spontaneous Speech*, pp. 187–210. New York: Springer.
- Kohler, K. J. (2001). Articulatory dynamics of vowels and consonants in speech communication. *JIPA* 31(1), 1–16.
- Kohler, K. J., M. Pätzold, und A. Simpson (1995). *From scenario to segment: the controlled elicitation, transcription, segmentation and labelling of spontaneous speech*. AIPUK 29. Kiel: IPDS.
- Kohler, K. J. und J. Rodgers (2001). Schwa deletion in German read and spontaneous speech. In K. J. Kohler (Hrsg.), *Sound Patterns of Spontaneous German Speech: Symbolic Structures and Gestural Dynamics*, AIPUK 35, pp. 97–123. Kiel: IPDS.
- Ladefoged, P. (1996). *Elements of Acoustic Phonetics*. Chicago, London: UCP.
- Lindblom, B. (1990). Explaining phonetic variation: A sketch of the H&H theory. In W. J. Hardcastle und A. Marchal (Hrsg.), *Speech Production and Speech Modelling*, pp. 403–439. Dordrecht: Kluwer Academic Publishers.

- McGurk, H. und J. MacDonald (1976). Hearing lips and seeing voices. *Nature* 264, 746–748.
- Moulines, E. und F. Charpentier (1990). Pitch-synchronous waveform processing techniques for text-to-speech synthesis using diphones. *Speech Communication* 9(5-6), 453–467.
- Niebuhr, O. (2003). Perceptual study of timing variables in F0 peaks. In *Proc. 15th ICPHS, Barcelona*, pp. 1225–1228.
- Peters, B. (2005). The database - The Kiel Corpus of Spontaneous Speech. In K. J. Kohler, F. Kleber, und B. Peters (Hrsg.), *Prosodic Structures in German Spontaneous Speech*, AIPUK 35a, pp. 1–6. Kiel: IPDS.
- Peters, B. und K. J. Kohler (2004). *Trainingsmaterialien zur prosodischen Etikettierung mit dem Kieler Intonationsmodell KIM*. URL www.ipds.uni-kiel.de/kjk/forschung/lautmuster.en.html.
- Peters, B., K. J. Kohler, und T. Wesener (2005). Melodische Satzakkentmuster in prosodischen Phrasen deutscher Spontansprache - Statistische Verteilung und sprachliche Funktion. In K. J. Kohler, F. Kleber, und B. Peters (Hrsg.), *Prosodic Structures in German Spontaneous Speech*, AIPUK 35a, pp. 7–54. Kiel: IPDS.
- Sundberg, J. (1979). Maximum speed of pitch changes in singers and untrained subjects. *Journal of Phonetics* 7, 71–79.
- 't Hart, J., R. Collier, und A. Cohen (1990). *A perceptual study of intonation: An experimental-phonetic approach to speech melody*. Cambridge: Cambridge University Press.
- Terhardt, E. (1974). Pitch, consonance, and harmony. *JASA* 55(5), 1061–1069.
- van Santen, J. P. H. (2002). Quantative modelling of pitch accent alignment. In *Proc. Speech Prosody, Aix-en-Provence*. URL: <http://lpl.univ-aix.fr/sp2002/pdf/vansanten.pdf>.
- Wendt, D. (1989). *Allgemeine Psychologie: Eine Einführung*. Stuttgart: Kohlhammer.
- Wesener, T. (2001). Some nonsequential phenomena in German function words. *JIPA* 31(1), 17–27.
- Xu, Y. und X. Sun (2002). Maximum speed of pitch change and how it may relate to speech. *JASA* 111(3), 1399–1413.