

# Phonetische Merkmale prosodischer Phrasierung in deutscher Spontansprache

Benno Peters, Klaus J. Kohler, Thomas Wesener

## 1 Fragestellungen

Eine wichtige Funktion der Prosodie ist die syntagmatische Strukturierung des Gesprochenen. Hierbei entstehen prosodische Phrasen, die nicht mit syntaktischer Phrasierung übereinzustimmen brauchen und deren Grenzen sich auf der signalphonetischen Ebene durch Bündel phonetischer Merkmale manifestieren. Die folgenden Merkmale stehen im Zentrum der hier vorgelegten empirischen Untersuchung:

- segmentelle Längung im Vergleich zur lautlichen Umgebung
- Intonationsmuster, die stark fallen, hoch steigen oder Fallen und Steigen in einer Satzakkzentposition kombinieren
- F0-Reset, d.h. Neueinsatz der F0-Deklination für Gipfelkonturen
- Pausen und Atmen

Daneben können eine Reihe weiterer phonetischer Merkmale von Bedeutung sein: Phonationsphänomene (Laryngalisierung, Glottalisierung), Segmentrealisierungen an Wortgrenzen (Glottalverschluß, Unterbleiben von Lautassimilation und -reduktion), Register- und Sprechgeschwindigkeitswechsel, Veränderungen der Intensität. Diese Merkmale gehen nicht in die statistische Untersuchung ein, werden aber im Zusammenhang mit der phonetischen Beschreibung verschiedener Korpusbeispiele diskutiert.

Alle genannten Merkmale bzw. ihre Bündelung können zu wahrnehmbaren Zäsuren im Sprechablauf führen, die als prosodische Phrasengrenzen kategorisiert werden. Entsprechend der Ausprägung der Einzelmerkmale und ihrer unterschiedlichen Bündelung ergeben sich Grade der Separierung zwischen aufeinander folgenden prosodischen Phrasen. Diese graduelle Abstu-

fung kann auf einer Kohäsionsskala angesiedelt werden, die von starker Separierung bis zu vollständiger Integration aufeinander folgender Sprechabschnitte reicht. Die Grade der Kohäsion stehen in enger Beziehung zur Planung und inhaltlichen Gliederung von Äußerungen.

Ziel der vorliegenden Untersuchung ist die Beschreibung der multifaktoriellen phonetischen Kodierung von Phrasengrenzen und ihre statistische Bearbeitung anhand eines umfangreichen, manuell etikettierten Korpus zur deutschen Spontansprache (IPDS 1995, IPDS 1996). Des Weiteren wird die theoretische Modellierung von Phrasengrenzen unterschiedlicher Kohäsionsgrade mit Bezug auf das Kieler Intonationsmodell (KIM, Kohler 1991) und dessen Implementation im Etikettiersystem PROLAB (Kohler 1997) thematisiert.

Die Untersuchung der Kohäsion zwischen prosodischen Phrasen *innerhalb* eines Redebeitrags (Turn) ist das zentrale Ziel; deshalb werden vorrangig turninterne Phrasengrenzen analysiert. Erst in einem zweiten Schritt werden phonetische Merkmale turnfinaler Phrasengrenzen mit denen turninterner Phrasengrenzen verglichen.

Die Untersuchung konzentriert sich außerdem auf flüssige prosodische Phrasierung und vergleicht sie mit der phonetischen Signalisierung von Phrasengrenzen bei gestörter Sprechflüssigkeit <sup>1</sup>, die bereits von Kohler, Peters und Wesener (2001) diskutiert wurde.

Bisherige Beiträge zur phonetischen Kodierung prosodischer Phrasengrenzen im Deutschen (z.B. Batliner et al. 1995, Kompe et al. 1997, Gallwitz et al. 1998) berücksichtigen zwar eine Reihe phonetischer Merkmale, beschreiben und diskutieren die Interaktion dieser Merkmale aber nicht vor dem Hintergrund phonetischer Grundlagenforschung; das Ziel ist die Steigerung der Performanz automatischer Erkennungssysteme. Schon Wang und Hirschberg (1992) bemerken allerdings, daß bei der Detektion von Phrasengrenzen in automatischen Erkennungssystemen die Gefahr besteht, die Interaktion der phonetischen Faktoren untereinander nicht ausreichend zu erfassen. Die vorliegende Arbeit ist Grundlagenforschung auf dem Gebiet der phonetischen Deskription und Modellierung von Phrasengrenzen. Wir hoffen aber, daß sich die Erkenntnisse zur Interaktion phonetischer Merkmale an Phrasengrenzen für die Verbesserung automatischer Erkennungssysteme verwenden lassen.

Price et al. (1991) erreichen in einer lesesprachlichen Datenbasis von Einzelsätzen eine Korrelation von 0.86 zwischen automatischer Etikettierung und Handetikettierung von Phrasengrenzen, wenn lediglich die Merkmale

---

<sup>1</sup>Für eine Trennung zwischen flüssiger und unflüssiger Phrasierung plädieren auch Wang und Hirschberg (1992), allerdings unter dem Aspekt der Erhöhung der Erkennungsrate automatischer Systeme.

Längung und Pause als Kriterien der automatischen Etikettierung verwendet werden. Wir gehen von der Wichtigkeit dieser Faktoren auch in unserer Datengrundlage aus, erwarten aber, daß in der phonetischen Signalisierung von Phrasengrenzen tonale Merkmale ebenfalls einen hohen Stellenwert haben. Eine unterschiedliche Gewichtung der phonetischen Faktoren ist allein schon wegen der grundsätzlich andersartigen Datenbasen (von professionellen Radiosprechern gelesene ambige Einzelsätze vs. spontansprachliche Dialoge) zu erwarten.

Die Untersuchung der prosodischen Phrasierung geht von folgenden Hypothesen aus:

1. Wenn im Datenkorpus Phrasengrenzen symbolisiert worden sind, dann liegt mindestens eines der genannten phonetischen Merkmale vor.
2. Finale Längung
  - (a) Segmentelle Längung ist das Merkmal mit der größten Vorkommenshäufigkeit.
  - (b) Sie ist turnfinal stärker ausgeprägt als turnintern.
3. Die genannten phonetischen Merkmale zur Signalisierung von Phrasengrenzen treten gebündelt auf. Diese Merkmalsbündel kodieren in Abhängigkeit von Anzahl und Ausprägung der gebündelten Komponenten den Grad der Separierung bzw. Kohäsion zwischen aufeinander folgenden prosodischen Phrasen. Diese Gradierung der Kohäsion bildet ihrerseits die inhaltliche Gliederung und die Äußerungsplanung ab.
4. Phrasengrenzen mit Störungen des Sprechflusses (Zögerungslängung/syntaktische Abbrüche) unterscheiden sich in der Ausprägung der phonetischen Merkmale von flüssiger Phrasierung.

## 2 Methode

Die Untersuchung basiert auf den handetikettierten Daten des *Kiel Corpus of Spontaneous Speech* (IPDS 1995, IPDS 1996). Darin wurden von den Etikettierern Zäsuren im Sprechfluss auditiv als Phrasengrenztypen erfaßt und mit folgenden Symbolen festgehalten:

- flüssige Phrasengrenze **PGn**
- Phrasengrenze mit Zögerungslängung **z: PGn**
- Phrasengrenze an syntaktischen Abbrüchen **PG/**

Die Stärke der Phrasengrenze wurde im Stadium der Datenaufbereitung (Etikettierung) nicht spezifiziert. Um einerseits die ohnehin schon sehr komplexe Aufgabe der Etikettierer nicht zusätzlich zu erschweren und um Validität und Reliabilität hoch zu halten, wurde nur das Vorhandensein einer Phrasengrenze symbolisiert und die Markierung mit dem unbestimmten Index  $n$  als Platzhalter für spätere Subklassifikation in einer umfangreichen signal- und symbolphonetischen Analyse des Datenkorpus versehen. Diese Analyse wird nunmehr vorgelegt.

## 2.1 Numerische Verarbeitung: Extraktion der Längungsinformation

Für die Erfassung der finalen Längung müssen phrasenfinale mit phraseninternen Segmentdauern verglichen werden. Die Segmentdauern lassen sich aus den Etikettierdateien extrahieren. Um den Vergleich durchführen zu können, muß zunächst der Bereich am Ende der Phrase ausgewählt werden, in dem die Längung bestimmt werden soll. Messungen von Wightman et al. (1992, 1714) weisen darauf hin, daß der Reim der finalen Silbe und nicht der Takt oder die letzte betonte Silbe die Domäne ist, in der sich die phrasenfinale Längung hauptsächlich manifestiert. Deshalb wird der letzte Reim vor der Phrasengrenze für die Untersuchung ausgewählt. Für den Vergleich mit Segmentdauern im Phraseninneren werden die phraseninternen wortfinalen Reime herangezogen.

Ziel der numerischen Verarbeitung ist ein von Sprecher und Segmenttyp unabhängiger Vergleich der Lautdauer in den beiden Phrasenpositionen. Hierzu muß eine Normalisierung durchgeführt werden. In Anlehnung an Price et al. (1991) wurde folgendes Normalisierungsverfahren gewählt:

- Für jeden Sprecher und jeden Lauttyp werden in *phraseninternen* wortfinalen Reimen Mittelwert und Standardabweichung der Lautdauer berechnet. In die Berechnung gehen nur reguläre Wörter ein, in denen keine Zögerungslängung auftritt.
- Jede Lautdauer im *phrasenfinalen* Reim wird mit Bezug auf die so gewonnenen laut- und sprecherspezifischen Mittelwerte und Standardabweichungen nach folgender Formel normalisiert:

$$\tilde{d} = (d - \mu_\alpha) / \sigma_\alpha$$

mit

$d$  = Dauer des jeweiligen Segments im phrasenfinalen Reim,

$\mu_\alpha$  = sprecherspezifischer Mittelwert des jeweiligen Lauttyps in phraseninternen Reimen,

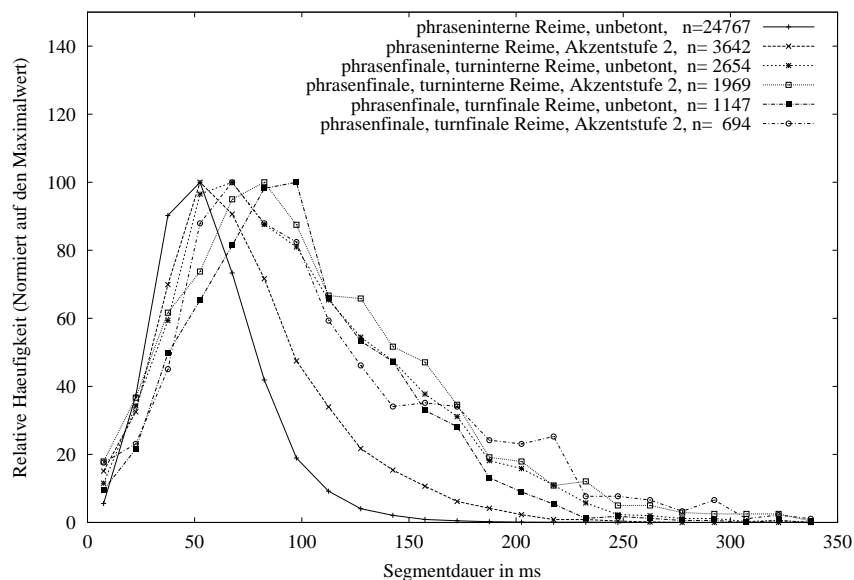
$\sigma_\alpha$  = zugehörige Standardabweichung.

- Der resultierende Wert  $\tilde{d}$  erfaßt die relative Dauer eines Segments als Anzahl der Standardabweichungen bezogen auf die mittlere Segmentdauer. Ist die Segmentdauer  $d$  genauso groß wie der Mittelwert  $\mu_\alpha$ , so ist die normalisierte Dauer  $\tilde{d} = 0$ . Ist die Segmentdauer größer (kleiner) als der Mittelwert, so ist die normalisierte Dauer positiv (negativ). Für Segmentdauern innerhalb der Standardabweichung ergeben sich Werte von -1 bis 1.
- Akzentuierung ist ebenfalls mit Längung verbunden. Um diesen Faktor zu kontrollieren, werden Mittelwertbildung und Normalisierung für drei in der prosodischen Etikettierung symbolisierte Akzentstufen im wortfinalen Reim getrennt durchgeführt: unbetont, schwach akzentuiert (Akzentstufe 1), akzentuiert (Akzentstufe 2). Die ebenfalls markierte emphatische Akzentuierung (Akzentstufe 3) wird aufgrund des seltenen Vorkommens von der Analyse ausgeschlossen. Die Kategorie 'unbetont' umfaßt neben Reimen in nicht satzakzentuierten Wörtern auch lexikalisch unbetonte Reime in satzakzentuierten Wörtern.
- Um für jeden phrasenfinalen Reim einen Kennwert zu erhalten, der dessen normalisierte Dauer insgesamt beschreibt, wird das arithmetische Mittel der normalisierten Dauern der Reimsegmente gebildet.
- Diese Kennwerte werden in drei Klassen eingeteilt:
  - $N0 < -1 \rightarrow$  Kürzung
  - $-1 \leq N1 \leq 1$
  - $N2 > 1 \rightarrow$  Längung

Da für die Einteilung in Dauerklassen keine in Perzeptionsexperimenten ermittelten Schwellenwerte zur Verfügung stehen, beruhen die Grenzwerte unserer Klassen auf einem Vergleich der Verteilungen der Segmentdauern in phraseninterner und phrasenfinaler Position. Abbildung 1 zeigt für 'unbetont' und 'Akzentstufe 2' die Verteilungen der Segmentdauern in phraseninternen bzw. -finalen Reimen, letztere getrennt nach turninterner und -finaler Position. Bei den dargestellten Daten handelt es sich um absolute Dauerwerte aller Segmenttypen sämtlicher Sprecher in Klassen mit einer Breite von 15 ms. Für die bessere Vergleichbarkeit der Verteilungen wurden die Häufigkeiten in jeder Klasse auf den Maximalwert der jeweiligen Verteilung normiert.

Aus Abbildung 1 geht hervor, daß sich die Verteilungen der phraseninternen Segmentdauern einer Normalverteilung annähern, während sich die Verteilungen aller phrasenfinalen Segmentdauern sehr stark mit einer Rechtsschiefe von der Normalverteilung entfernen. D.h. in der phrasenfinalen Position wird ein Längungsfaktor wirksam.

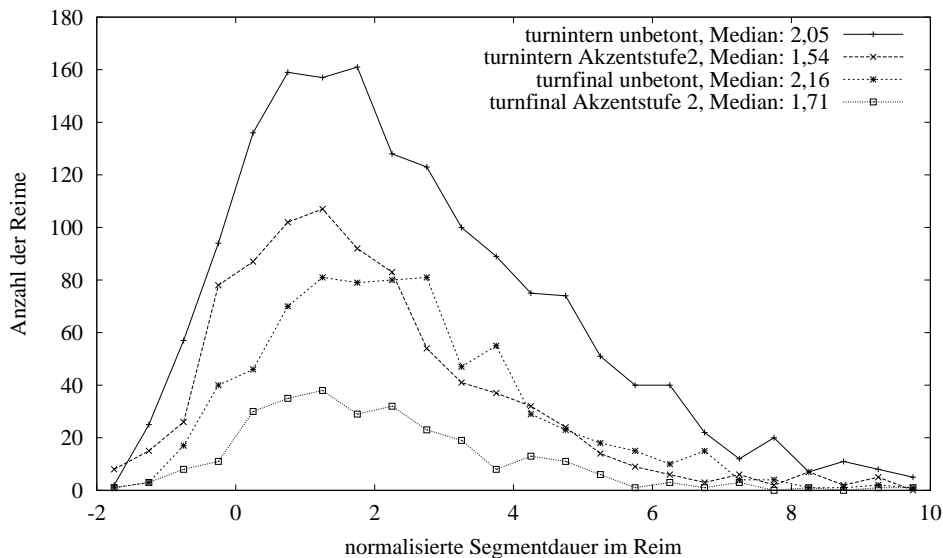
Abbildung 1: Häufigkeitsverteilungen turninterner und turnfinaler Segmentdauern im Reim in Klassen von 15 ms Breite, normiert auf den Maximalwert der jeweiligen Verteilung



In Abbildung 2 sind die Verteilungen der gemittelten, normalisierten Segmentdauern im Reim für die turninterne bzw. -finale Position in den Kategorien 'unbetont' und 'Akzentstufe 2' dargestellt. Die Verteilungen basieren auf einer Klassenbreite von 0,5 Standardabweichungen. Hier zeigt sich wieder eine deutliche Rechtsschiefe. Wären die Verteilungen in phrasenfinaler Position vergleichbar mit denen in phraseninterner Position, dann hätte sich in der Verteilung der normalisierten Segmentdauern angenähert eine Normalverteilung um den Mittelwert 0 ergeben müssen. In den Verteilungen der normalisierten Segmentdauern, in denen die finalen auf die internen Dauern bezogen wurden, schlägt sich also der Längungsfaktor nieder.

In einer Standardnormalverteilung liegen 68% der Werte innerhalb des Bereichs von  $\pm 1$  Standardabweichung; die verbleibenden 16% des rechten Astes der Verteilung sind untypisch. Die empirisch vorgefundenen Verteilungen der normalisierten phrasenfinalen Dauern haben jedoch in allen Fällen

Abbildung 2: Verteilungen der gemittelten normalisierten Segmentdauern phrasenfinaler Reime, turnintern und -final, in den Kategorien 'unbetont' und 'Akzentstufe 2'; Klassenbreite 0,5 Standardabweichungen



über 60% der Werte oberhalb einer Standardabweichung. Es ist daher plausibel, Werte oberhalb einer Standardabweichung als Indices für finale Längung anzusetzen.

## 2.2 Symbolbasierte Verarbeitung: Extraktion der übrigen Merkmale

Die segmentelle und prosodische Etikettierung der Datenbasis bietet direkten symbolbasierten Zugriff auf die phonetischen Merkmale *phrasenfinale Intonationskontur*, *Pause/Atmen* und *F0-Reset* an Phrasengrenzen. Für jede einzelne Phrasengrenze wird aus der manuellen Etikettierung automatisch ein Label abgeleitet, das die phrasenfinale Intonationskontur und das Vorhandensein/Fehlen von Pausen/Atmen sowie F0-Reset erfaßt. Die Information über die Dauerklassen N0 - N2 wird in dieses Etikett integriert. Dabei wird auch die Akzentstufe des Reims erfaßt. Auf diese Weise ist das jeweilige Bündel phonetischer Merkmale an jeder Phrasengrenze in einem Etikett symbolisiert. Die Gesamtheit dieser Etiketten bildet die Basis für die statistische Auswertung der Merkmalsbündel.

## 2.3 Grundgesamtheiten

In der Analyse der phonetischen Merkmale prosodischer Phrasierung werden vier Typen von Phrasengrenzen getrennt analysiert. Es ergeben sich die folgenden vier Grundgesamtheiten:

1. reguläre turninterne Phrasengrenzen ( $n = 2470$ )  
Phrasengrenzen im Inneren eines Redebeitrags mit Information zur Dauerklasse, ohne Zögerungslängung oder syntaktische Abbrüche
2. reguläre turnfinale Phrasengrenzen ( $n = 1092$ )  
Phrasengrenzen am Ende eines Redebeitrags mit Information zur Längungsklasse, ohne Zögerungslängung oder syntaktische Abbrüche
3. Phrasengrenzen mit Zögerungslängung ( $n = 871$ )  
Phrasengrenzen nach Wörtern mit Zögerungslängung im Inneren eines Redebeitrags, ohne syntaktische Abbrüche
4. Phrasengrenzen an syntaktischen Abbrüche ( $n = 150$ )  
Phrasengrenzen an syntaktischen Abbrüchen mit und ohne Wiederaufnahme im Inneren eines Redebeitrags, ohne Zögerungslängung

In allen Grundgesamtheiten sind Phrasengrenzen mit Häsitationspartikeln wie *ähm* und *äh* ausgeschlossen. In Bezug auf die Kohäsion benachbarter Phrasen bilden Phrasengrenzen mit Häsitationspartikeln einen Komplex, der einer gesonderten Untersuchung bedarf.

## 3 Ergebnisse

Im folgenden wird zuerst die Ausprägung und Bündelung der phonetischen Merkmale an regulären turninternen Phrasengrenzen dargestellt, die die größte Gruppe von Phrasengrenzen bilden. Darauf werden die Ergebnisse mit den anderen Typen von Phrasengrenzen verglichen.

### 3.1 Ausprägung der einzelnen Merkmale an regulären turninternen Phrasengrenzen

#### 3.1.1 Finale Längung

Tabelle 1 zeigt die Verteilung der gemittelten normalisierten Lautdauern über die drei Klassen N0-N2 für Reime aller Akzentstufen an turninternen Phrasengrenzen. Es ergibt sich eine relative Häufigkeit finaler Längung von 66.2%.

Tabelle 1: Verteilung der gemittelten normalisierten Segmentdauern über die Klassen N0-N2 in Reimen aller Akzentstufen an turninternen Phrasengrenzen (absolute und relative Häufigkeiten).

N0	N1	N2	ges.
59	777	1634	2470
2.4	31.5	66.2	100

Tabelle 2: Verteilung der drei Dauerklassen in turninternen Reimen in Abhängigkeit von der Akzentstufe (U unakzentuiert, A1-3 Akzentstufe 1-3)

Klasse	U	%	A1	%	A2	%	A3	ges.	%
N0	28	1.9	2	1.1	29	3.6	0	59	2.4
N1	430	29.0	65	35.1	282	35.2	0	777	31.5
N2	1026	69.1	118	63.7	490	61.2	0	1634	66.2
ges.	1484	100	185	100	801	100	0	2470	100

In einem Drittel der Fälle zeigt sich nach dem hier angelegten Kriterium keine finale Längung; in 2.4% liegt sogar eine Kürzung vor.

In Tabelle 1 ist die Dauerinformation unabhängig von der Akzentstufe der Reime dargestellt. Um den möglichen Einfluß der Akzentstufe auf die Verteilung der relativen Dauern zu erfassen, zeigt Tabelle 2 die Verteilung der drei Dauerklassen in turninternen Reimen in Abhängigkeit von der Akzentstufe. Die Verteilungen für betonte Reime sind von der Klasse mit der größten relativen Dauer (N2) leicht zu niedrigeren Klassen hin verschoben, und zwar für Akzentstufe 2 stärker als für Akzentstufe 1. Es sind 69.1% der unbetonten phrasenfinalen Reime gelangt, aber nur 63.7% der Akzentstufe 1 und nur 61.2% der Akzentstufe 2.

Für die Erklärung der Verschiebung in den Häufigkeitsverteilungen der verschiedenen Akzentstufen ist die Untersuchung einer Reihe möglicher Faktoren notwendig:

- unterschiedliche Reimstrukturen (Anzahl der Segmente pro Reim)
- segmentabhängige stärkere vs. schwächere Dauervariation über die beiden Phrasenpositionen (Schwa vs. Vollvokal, silbische Nasale)
- starke Dauerreduktion der Reime in unbetonten phraseninternen Funktionswörtern, die phrasenfinal nicht auftritt.

Tabelle 3: Häufigkeitsverteilung über die drei Dauerklassen in turninterner (TI) und turnfinaler (TF) Position

Klasse	N0	N1	N2	ges.
TI	59	777	1634	2470
%	2.4	31.5	66.2	100
TF	12	268	812	1092
%	1.1	24.5	74.3	100
ges.	71	1045	2446	3562
%	2.0	29.3	68.6	100

Zur Überprüfung der Hypothese, daß phrasenfinale Längung turnfinal stärker ausfällt als turnintern, wurde die Häufigkeitsverteilung über die Dauerklassen auch an turnfinalen Phrasengrenzen ermittelt. Tabelle 3 stellt die Verteilungen über die drei Klassen in turninterner und turnfinaler Position gegenüber. An turnfinalen Phrasengrenzen zeigt sich eine leichte Verschiebung zur Längungsklasse; insgesamt sind hier 74.3% der phrasenfinalen Reime gelängt statt nur 66.2% an turninternen Phrasengrenzen. Darüber hinaus zeigt Abbildung 2, daß in turnfinaler Position nicht nur die Häufigkeit in der Klasse N2 zunimmt, sondern daß auch der Median der Dauerverteilung höher liegt und damit die Segmentdauern generell länger sind als turnintern. Phrasenfinale Längung ist in unseren Daten also tatsächlich turnfinal stärker ausgeprägt.

Die 59 Fälle von Kürzung des phrasenfinalen Reims (2.4%) sind ein unerwartetes Ergebnis. Diese Fälle wurden deshalb einer auditiven und visuellen Einzelanalyse unterzogen. Die an Sonagrammen durchgeführte Beurteilung der relativen Segmentdauern im lautlichen Kontext auch außerhalb des Reims zeigte, daß in etwa der Hälfte dieser Fälle die segmentelle Längung aus dem Reim in das Wortinnere verschoben ist. In der anderen Hälfte findet sich auch im Wortinneren keine segmentelle Längung; die Phrasengrenzen sind hier vor allem durch stark fallende Konturen (vgl. Abschnitt 3.1.2), Reset und Pause/Atmen kodiert.

In vielen Fällen wurde ein Zusammenhang zwischen finaler Kürzung und Informationsstruktur der Äußerungen festgestellt. Wenn Phrasen Information enthalten, die inhaltlich für das Gespräch redundant ist, z.B. in Floskeln, Selbstgesprächen oder weil sie in der vorhergehenden Phrase bereits erwähnt wurde bzw. in der nachfolgenden Phrase präzisiert wird, kann finale Kürzung auftreten: *es tut mir leid*, | *da* (g092a020, g096a009) *oh*, *entschuldigen Sie*,

| *da* (g212a010)<sup>2</sup>. Die genannten Beispiele zeigen noch eine weitere Gemeinsamkeit: In allen Fällen trennt die Phrasengrenze mit Kürzung die initiale Phrase eines Redebeitrags von der zweiten Phrase.

### 3.1.2 Phrasenfinale Intonationskonturen

In der PROLAB-Etikettierung werden 12 phrasenfinale Intonationskonturen unterschieden. Diese Konturen verlaufen von der letzten satzakzentuierten Silbe einer Phrase bis zum Phrasenende. Tabelle 4 zeigt die Verteilung der turninternen phrasenfinalen Konturen. In Bezug auf die Separierung von Phrasen sind einige dieser Konturen besonders prominent, da sie einen perzeptorischen Einschnitt im Übergang zur folgenden Phrase verursachen. Dies sind z.B. stark fallende Konturen, die besonders im Zusammenwirken mit Laryngalisierung einen deutlichen perzeptiven Einschnitt bewirken (Kohler, Peters und Wesener 2001). Auch Intonationsverläufe mit starkem Anstieg sowie kombinierte Konturen, die sich aus einer fallenden und einer deutlich steigenden Bewegung zusammensetzen, verursachen solche Einschnitte und sind somit als tonale Indikatoren von Phrasierung zu betrachten.

Einen besonderen Status hat die mit /**2**;/ gekennzeichnete Kontur. Sie verursacht einen ähnlichen terminalen Eindruck wie eine stark fallende Kontur /**2**./, schwächt allerdings die Terminalität durch einen sehr leichten F0-Anstieg ab (Peters 2000a). Darin unterscheidet sie sich nicht nur von der terminalen Kontur /**2**./, sondern auch von einer kombinierten Kontur /**2**./ (oder gar /**2**?/), wo der F0-Anstieg sehr viel stärker ausgeprägt ist. /**2**;/ nimmt eine Zwischenstellung zwischen den separierenden Konturen /**2**./ und /**2**./ bzw. /**2**?/ ein und wird deshalb auch der Klasse der separierenden Konturen zugeordnet.

Insgesamt handelt es sich um 7 der 12 Intonationsverläufe, die in der weiteren Untersuchung als tonale Merkmale der Phrasierung eingestuft werden. Die übrigen 5 sind ebene und nur schwach steigende oder leicht fallende Muster, die in der Regel nicht zu einem perzeptorischen Einschnitt führen.

### 3.1.3 F0-Reset

In der PROLAB-Etikettierung wird ein F0-Reset etikettiert, wenn der Maximalwert der ersten Gipfelkontur einer Phrase höher liegt als der des letzten Gipfels der vorangehenden Phrase. In der Analyse von spontansprachlichen Daten hat sich gezeigt, daß das Auftreten von Talkonturen den Downstep zwischen den Gipfeln innerhalb einer Phrase stört. Dies gilt besonders, wenn Talkonturen und kombinierte Konturen am Ende von prosodischen Phrasen

---

<sup>2</sup>Das Symbol | markiert die Position der Phrasengrenze.

Tabelle 4: Verteilung der turninternen phrasenfinalen Konturen

separierende Konturen		n	%
2.	stark fallend	731	29.6
?	hoch steigend	44	1.8
1.,	leicht fallend + mittelhoch steigend	52	2.1
1.?	leicht fallend + hoch steigend	4	0.2
2.,	stark fallend + mittelhoch steigend	18	0.7
2.?	stark fallend + hoch steigend	0	0.0
2;	stark fallend + sehr leicht steigend	159	6.4
nicht separierende Konturen			
0.	eben	368	14.9
1.	leicht fallend	653	26.4
,	mittelhoch steigend	344	13.9
0;	eben + sehr leicht steigend	15	0.6
1;	leicht fallend + sehr leicht steigend	82	3.3
ges.		2470	100

Tabelle 5: F0-Reset an turninternen Phrasengrenzen

Merkmal	Reset	kein Reset	keine Entscheidung	ges.
	1557	451	462	2470
%	63.0	18.3	18.7	100

stehen; PROLAB sieht in diesen Fällen keine Entscheidung über das Auftreten von Reset vor. Tabelle 5 zeigt die Vorkommenshäufigkeit von Reset. An den 2470 Phrasengrenzen folgt in 63.0% der Fälle ein Reset, während sich in 18.3% der Downstep über die Phrasengrenze hinweg fortsetzt. In 18.7% wurde keine Entscheidung über Reset getroffen.

### 3.1.4 Pause/Atmen

An 945 der 2470 Phrasengrenzen tritt Pause oder Atmen auf (38.3%), damit sind Pausen und Atmen häufige Korrelate prosodischer Phrasierung. Das Auftreten von Pause/Atmen ist dabei in 84% der Fälle mit finaler Längung verbunden. Die in der Datengrundlage etikettierten Pausen bzw. Atmungsphasen weisen allerdings große Dauervariation auf, die in Bezug auf die separierende Wirkung dieser Merkmale von Bedeutung ist, wie Untersuchungen von Boomer und Dittman (1962) und Butcher (1981) zeigen. Außerdem beinhalten die im Kieler Korpus als Atmen etikettierten Signalabschnitte

Tabelle 6: Überblick über das Vorkommen der einzelnen Merkmale an turn-internen Phrasengrenzen

Merkmal	Längung	Reset	Kontur	Pause/Atmen
	1634	1557	1008	945
%	66.2	63.0	40.8	38.3

häufig auch Pausen. Weiterhin variiert die Intensität des Atmungsgeräuschs stark, wobei lautes Atmen zur Signalisierung von beabsichtigten Turnübernahmen eingesetzt werden kann. Die Trennung von gefüllten und ungefüllten Pausen sowie die Variation im Dauer- und Intensitätsbereich bedarf einer eigenständigen Untersuchung.

Tabelle 6 zeigt noch einmal zusammenfassend die Verteilung der einzelnen Merkmale. In 23.3% der Fälle wird die Phrasengrenze nur durch eines der untersuchten Merkmale signalisiert, und zwar in der Rangordnung Reset 10.3%, Längung 9.4%, separierende Kontur 2.3%. Pause/Atmen 1.3%.

### 3.2 Bündelung der Merkmale an regulären turninternen Phrasengrenzen

Phrasengrenzen werden meistens durch Kombinationen verschiedener phonetischer Merkmale kodiert. Es können drei Klassen phonetischer Indikatoren unterschieden werden, die in unterschiedlichen Kombinationen an Phrasengrenzen auftreten:

- Tonale Indikatoren, die durch einen diskontinuierlichen Tonhöhenverlauf Zäsuren markieren. Dieser Klasse können bestimmte F0-Konturen zugeordnet werden, die lokal zu einem tonalen Einschnitt führen, aber auch der F0-Reset, der durch einen Neueinsatz des Downstep eine globalere Separierung zweier Phrasen bewirkt.
- Segmentelle Längung, die durch eine deutliche Veränderung der Dauerstruktur im Vergleich zum Phraseninneren eine prosodische Phrasierungsgrenze signalisiert.
- Pause und Atmen, die einen deutlichen Einschnitt im Signal markieren, der die Kohäsion benachbarter Einheiten (normalerweise Wörter) mindert.

Tabelle 7: Überblick über 16 Merkmalsbündel an turninternen Phrasengrenzen (R Reset, L finale Längung, P Pause/Atmen, K separierende Konturen)

	-	L	P	K	LP	PK	LK	LPK	ges.
+R	255	286	44	202	235	60	210	265	1557
%	10.3	11.6	1.8	8.2	9.5	2.4	8.5	10.7	63.0
-R	171	232	31	57	208	16	112	86	913
%	6.9	9.4	1.3	2.3	8.4	0.6	4.5	3.5	36.9

Tabelle 8: Überblick über 8 Merkmalsbündel an turninternen Phrasengrenzen (F0 Reset und/oder separierende Konturen, L finale Längung, P Pause/Atmen)

	-	L	P	LP	ges.
+F0	514	608	120	586	1828
%	20.8	24.6	4.9	23.7	74.0
-F0	171	232	31	208	642
%	6.9	9.4	1.3	8.4	26.0

Tabelle 7 zeigt die Verteilung der phonetischen Indikatoren in turninterner Position und gibt einen Überblick über die verschiedenen Zusammensetzungen der Merkmalsbündel. In dieser Übersicht ist die Klasse der tonalen Indikatoren nicht zusammengefaßt, sondern getrennt nach Reset und einschneidender phrasenfinaler Kontur angegeben. Tabelle 8 faßt die tonalen Merkmale zu einer Klasse zusammen, da sowohl Reset als auch phrasenfinale Melodiemuster mit starker F0-Bewegung separierende Wirkung haben. Mit dieser Zusammengruppierung ist die tonale Kodierung an 74% aller Phrasengrenzen vertreten und damit das häufigste Merkmal. Beide Tabellen zeigen, daß in ungefähr zwei Drittel der Fälle mindestens zwei Merkmale der Phrasierung auftreten.

Mit den untersuchten Merkmalen lassen sich 93.1% der Phrasengrenzen erfassen. Unter den 642 (26%) Phrasengrenzen ohne tonale Kodierung weisen 440 (17.8%) Längung und 239 (9.7%) Pause/Atmen auf, wobei Pause/Atmen und Längung in 208 Fällen (8.4%) gleichzeitig auftreten. Damit bleiben 171 Fälle (6.9%), in denen sich weder Pause/Atmen noch Längung finden und in denen die automatische Klassifikation somit keine Hinweise auf phonetische Merkmale der Phrasierung liefert, obwohl eine Phrasengrenze etikettiert wurde. Auf der Suche nach weiteren möglichen Merkmalen von Phrasierung wurden diese Fälle einer auditiven und visuellen Einzelanalyse unterzogen.

Dabei stellte sich heraus, daß 75 von diesen 171 Fällen mit einer steigenden Kontur an der Phrasengrenze realisiert sind. Obwohl es sich in diesen Fällen nicht um einen starken Anstieg handelt, bewirkt der Unterschied zwischen dem Endpunkt des Anstiegs und dem nachfolgenden niedrigeren F<sub>0</sub>-Verlauf einen deutlichen tonalen Einschnitt und damit die Wahrnehmung einer Phrasengrenze. Auch hier wird also die strukturierende Funktion von Tonbrüchen noch einmal sehr deutlich.

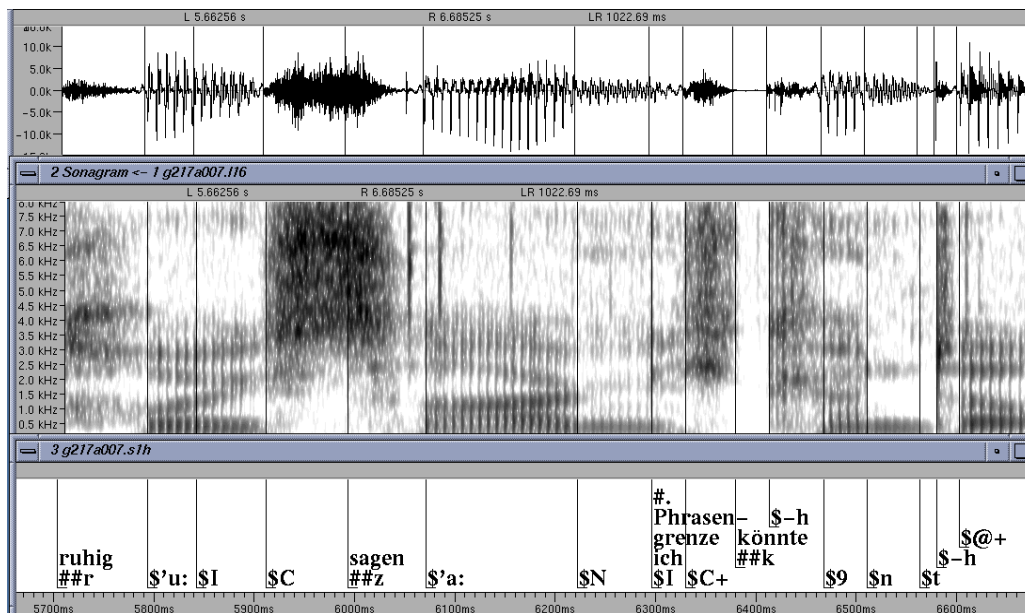
Unter den verbleibenden 96 Fällen fanden sich in 36 auch bei eingehender Signalanalyse keine Indikatoren einer phonetischen Phrasierung. Allerdings treten in allen 36 Fällen starke syntaktische Einschnitte auf, z.B. zwischen Haupt- und Nebensatz oder zwischen zwei Hauptsätzen. Diese syntaktischen Einschnitte haben offensichtlich mit der Wahrnehmung phonetisch indizierter Phrasengrenzen interagiert und zur Interpretation einer syntaktischen Grenze als prosodische Grenze geführt.

In den übrigen 60 Fällen ohne automatisch erfaßte Merkmale der Phrasierung ist eine Phrasengrenze wahrnehmbar. Dabei spielen z.B. Änderungen des Sprechtempos und des Registers eine Rolle, die in der automatischen Klassifikation der Phrasengrenzen nicht berücksichtigt wurden, weil die entsprechenden Symbolisierungen keine eindeutige zeitliche Zuordnung zu den relevanten Phrasengrenzen erlaubten. Auch bestimmte artikulatorische Merkmale verursachen die Wahrnehmung einer Phrasengrenze, so z.B. ein initialer Glottalverschluß mit nachfolgender Knarrstimme in unbetonten Funktionswörtern wie *am* (g106a003); vgl. in diesem Zusammenhang die hohe Auftretenshäufigkeit des wortinitialen Glottalverschlusses in phraseninitialer Position (Rodgers 1999), im Gegensatz zu seiner Reduktion in Hutmustern (Peters, Kohler und Wesener 2003).

Des weiteren beginnt finale Längung in einigen der mehrsilbigen Wörter schon ab der letzten akzentuierten Silbe. In anderen Fällen wiederum macht sich die Längung vor allem am Nukleus der akzentuierten Silbe fest, während die Koda wieder kürzer ausfällt, z.B. im einsilbig produzierten *sagen* [za:ŋ] (g217a007, g255a002). In *Dienstreise* (g195a003) ist weder der akzentuierte Vokal noch die letzte Silbe besonders lang, dafür aber der alveolare Frikativ in der Koda der akzentuierten Silbe und der Diphthong der zweiten Silbe. In diesen Fällen ermittelt der verwendete Algorithmus für den wortfinalen Reim keine Längung oberhalb der Standardabweichung; in einer visuell und auditiv orientierten Signalanalyse der segmentell etikettierten Daten zeigen sich solche Fälle von Längungsverschiebung hingegen recht deutlich. Außerdem bestätigen die Messungen der betreffenden Segmente und der Vergleich dieser Dauerwerte mit den berechneten mittleren phraseninternen Segmentdauern diesen Eindruck. Abbildung 3 illustriert das mit Bezug auf *sagen* [za:ŋ] in g217a007. Die gemessene Dauer für [ŋ] ist 76 ms, was dem Mittelwert für das

Segment [ŋ] in phraseninterner Position (80 ms) entspricht. Demgegenüber ist die gemessene Dauer für [a:] 152 ms und liegt deutlich über dem phraseninternen Mittelwert von 113 ms

Abbildung 3: Oszillogramm, Sonagramm und orthographische Wort- und phonetische Lautetikettierung (mit Angabe der Phrasengrenze) der Äußerung *...ruhig Sagen. ich Könnte...* in g217a007 (Großbuchstaben = satzakzentuierte Wörter)



Die Einzeluntersuchung der Phrasengrenzen ohne automatisch extrahierte phonetische Merkmale hat gezeigt, daß außer diesen Merkmalen weitere Parameter eine Rolle bei der Wahrnehmung von Phrasengrenzen spielen können. Damit können auch an den übrigen Phrasengrenzen phonetische Parameter beteiligt sein, die in den Merkmalsbündeln nicht erfaßt sind. Der Einzelfall ist somit nicht selten komplexer, als die symbolischen Bündel vermuten lassen. Die Vereinfachung der Klassifikation anhand weniger automatisch extrahierbarer Merkmale erlaubt aber Generalisierungen über eine umfangreiche Datengrundlage.

Wie bereits in der Einleitung erwähnt wurde, erreichen Price et al. (1991) eine Korrelation von 0.86 zwischen automatischer Etikettierung und Handetikettierung von Phrasengrenzen, wenn nur Längung und Pause als Kriterien der automatischen Etikettierung verwendet werden. Eine Übertragung dieser Klassifikationskriterien auf die vorliegende Datenbasis liefert folgende Ergebnisse: von den 2470 turninternen Phrasengrenzen weisen 1634 (66.2%)

gelängte Reime, 945 (38.3%) Pause/Atmen auf. Eine Kombination von Pause/Atmen und Längung tritt in 794 Fällen (32.1%) auf. Damit sind in 1785 Fällen (72.3%) die Phrasengrenzen durch Längung und/oder Pause/Atmen markiert. Noch mehr Phrasengrenzen lassen sich in unserem Korpus anhand der beiden Kriterien F0-Einschnitt und Pause erfassen (83.7%).

### **3.3 Vergleich von phonetischen Merkmalen an verschiedenen Typen von Phrasengrenzen**

In den folgenden Abschnitten werden Unterschiede in der phonetischen Signalisierung von vier verschiedenen Grenztypen aufgezeigt, nämlich der Phrasengrenzen im Inneren eines Gesprächsbeitrags, am Ende eines Gesprächsbeitrags, an syntaktischen Abbrüchen und an Grenzen mit Zögerungslängung (vgl. die in Abschnitt 2.3 beschriebenen Grundgesamtheiten).

#### **3.3.1 Phrasenfinale Konturen an den vier Phrasengrenztypen**

Tabelle 9 gibt einen Überblick über die unterschiedlichen Verteilungen der F0-Verläufe am Ende von vier durch die Etikettierung differenzierbaren Grenztypen. Die Verteilung der separierenden bzw. nicht separierenden phrasenfinalen Konturen an den vier verschiedenen Grenztypen ist in Abbildung 4 veranschaulicht.

Vor allem der Anteil der separierenden Konturen an den vier Grenztypen ist sehr unterschiedlich. Während in der turnfinalen Position 78.3% aller Phrasen mit einer separierenden Kontur enden, sind es turnintern nur 40.8%. An den beiden Phrasengrenztypen, die im Zusammenhang mit Störungen der Sprechflüssigkeit bzw. der Äußerungsplanung stehen, tritt der Anteil der Konturen mit einer starken tonalen Bewegung in den Hintergrund: An Grenzen mit Zögerungslängung liegt der Anteil der separierenden Konturen bei 22.6%, an syntaktischen Abbrüchen nur bei 15.3%. Der hohe Anteil der separierenden Konturen in der turnfinalen Position steht in unmittelbarem Zusammenhang mit der pragmatischen Funktion der Melodiemuster. Besonders die stark fallende Kontur wird häufig eingesetzt, um den Abschluß eines inhaltlichen Abschnitts anzuzeigen und dem Gesprächspartner das Wort zu übergeben. Auch der mit Abstand höchste Anteil von 10.3% der stark steigenden Konturen in turnfinaler Position ist mit Bezug auf die Funktion dieses Musters im Diskurs zu interpretieren: Der hoch ansteigende Melodieverlauf ist in der Regel mit einer direkten Frage an den Gesprächspartner verbunden und tritt deshalb kaum turnintern auf.

Tabelle 9: Verteilung der phrasenfinalen Konturen über die Positionen turnintern mit/ohne Zögern/Abbruch und turnfinal

	turnintern ohne Zögern		turnfinal		turnintern Abbruch		turnintern mit Zögern	
	n	%	n	%	n	%	n	%
separierende Konturen								
stark fallend	731	29.6	536	49.1	21	14.0	110	12.6
hoch steigend	44	1.8	112	10.3	0	0.0	20	2.3
leicht fallend + mittelhoch steigend	52	2.1	10	0.9	0	0.0	13	1.5
leicht fallend + hoch steigend	4	0.2	6	0.5	0	0.0	1	0.1
stark fallend + mittelhoch steigend	18	0.7	25	2.3	0	0.0	8	0.9
stark fallend + hoch steigend	0	0.0	5	0.5	0	0.0	0	0.0
stark fallend + sehr leicht steigend	159	6.4	161	14.7	2	1.3	45	5.2
<b>zusammen</b>	<b>1008</b>	<b>40.8</b>	<b>855</b>	<b>78.3</b>	<b>23</b>	<b>15.3</b>	<b>197</b>	<b>22.6</b>
nicht separierende Konturen								
eben	368	14.9	48	4.4	43	28.7	190	21.8
leicht fallend	653	26.4	45	4.1	75	50.0	348	40.0
mittelhoch steigend	344	13.9	117	10.7	2	1.3	61	7.0
eben + sehr leicht steigend	15	0.6	6	0.5	2	1.3	11	1.3
leicht fallend + sehr leicht steigend	82	3.3	21	1.9	5	3.3	64	7.3
<b>zusammen</b>	<b>1462</b>	<b>59.1</b>	<b>237</b>	<b>21.6</b>	<b>127</b>	<b>84.6</b>	<b>674</b>	<b>77.4</b>
ges.	2470	100	1092	100	150	100	871	100

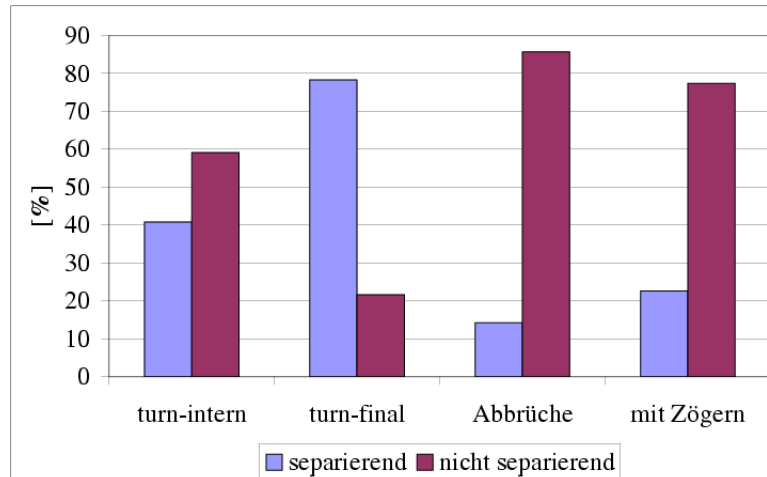


Abbildung 4: Die Verteilung von separierenden und nicht separierenden phrasenfinalen Konturen an vier verschiedenen Typen von Phrasengrenzen.

### 3.3.2 Merkmale an Phrasengrenzen mit gestörter Sprechflüssigkeit

In diesem Abschnitt werden Merkmale an Phrasengrenzen mit gestörter Sprechflüssigkeit zuerst für Grenzen mit Zögerungslängung, dann für Grenzen an syntaktischen Abbrüchen thematisiert.

Der Bereich, in dem sich Zögerungslängung manifestiert, variiert sehr stark. Einzelne gelängte Segmente können an unterschiedlichen Wortpositionen auftauchen, es können aber auch ganze Silben, Wörter oder Gruppen von aufeinander folgenden Wörtern betroffen sein. Als Domäne der finalen Längung bei flüssiger Phrasierung sind längere Wörter und Wortgruppen untypisch. Damit ist ein quantitativer Vergleich zwischen der Zögerungslängung und der finalen Längung an Phrasengrenzen ohne Zögern anhand der Dauern wortfinaler Reime nicht sinnvoll.

Auch in Bezug auf die Variationsbreite der Segmentdauern verhält sich Zögerungslängung anders als finale Längung. Die Segmente können wie an Phrasengrenzen ohne Zögern relativ kurz ausfallen, nehmen aber häufiger als bei der finalen Längung große Dauern an und erreichen oft Größen, die bei der finalen Längung nicht anzutreffen sind. In Tabelle 10 und Tabelle 11 werden die Häufigkeitsverteilungen über acht bzw. vier Merkmalsbündel an Phra-

Tabelle 10: Überblick über 8 Merkmalsbündel an turninternen Phrasengrenzen mit Zögerungslängung (R Reset, Z Zögerungslängung, K Kontur mit starkem Abfall oder Anstieg und alle kombinierten Konturen, P Pause/Atmen)

	Z	ZP	ZK	ZPK	ges.
+R	271	202	58	69	600
%	31.1	23.2	6.7	7.9	68.9
-R	106	95	34	36	271
%	12.2	10.9	3.9	4.1	31.1

Tabelle 11: Überblick über 4 Merkmalsbündel an turninternen Phrasengrenzen mit Zögerungslängung (F0 Reset und/oder Kontur mit starkem Abfall oder Anstieg und alle kombinierten Konturen, Z Zögerungslängung, P Pause/Atmen)

	Z	ZP	ges.
+F0	363	307	670
%	41.7	35.2	76.9
-F0	106	95	201
%	12.2	10.9	23.1

sengrenzen mit Zögerungslängung aufgeschlüsselt. Der Vergleich der Häufigkeiten an Phrasengrenzen mit und ohne Zögerungslängung bringt folgende Ergebnisse.

- 600 Fälle (68.9%) von Reset entsprechen der Häufigkeit an Phrasengrenzen ohne Zögerungslängung
- 402 Fälle (46.2%) von Pause/Atmen sind mit der Häufigkeit an Phrasengrenzen ohne Zögerungslängung vergleichbar.
- 197 Fälle (22.6%) mit einschneidender Kontur sind deutlich seltener als bei den Phrasengrenzen ohne Zögerungslängung
- 106 Fälle (12.2%) ohne eines der drei Merkmale sind zahlreicher als bei den Phrasengrenzen ohne Zögerungslängung. Vermutlich wurde die Phrasengrenze häufig nur wegen der Zögerungslängung vergeben

Der Vergleich zwischen Phrasengrenzen mit und ohne Zögerungslängung zeigt, daß sich die Merkmalsbündel vor allen Dingen hinsichtlich des phrasenfinalen Melodiemusters unterscheiden: Nach Tabelle 9 sind 40.8% der turn-

internen Phrasengrenzen ohne Zögerungslängung durch separierende Konturen markiert, während es bei den turninternen Phrasengrenzen mit Zögern nur 22.6% sind. Wegen des sehr ähnlichen Anteils von Reset in beiden Kategorien der Phrasengrenze (68.9% vs. 63.0%, vgl. Tabelle 5), ist eine Zusammengruppierung der tonalen Merkmale Reset und phrasenfinale Kontur wie in Tabelle 11 nicht unproblematisch; denn in der verschmolzenen Klasse der tonalen Merkmale (+F0) sind die prozentualen Anteile infolge der Überlagerung des Konturmerkmals durch das Reset-Merkmal wiederum sehr ähnlich: 76.9% bei Zögerungslängung gegenüber 74.0% bei turninternen Phrasengrenzen ohne Zögern (vgl. Tabelle 8). Es ist im Zusammenhang mit der Zögerungslängung also sinnvoll, eine klare Trennung zwischen dem Parameter Reset und dem phrasenfinalen Melodiemuster beizubehalten, obwohl beides tonale Phänomene sind.

Auch Phrasengrenzen an syntaktischen Abbrüchen zeigen in ihrer phonetischen Ausprägung deutliche Unterschiede zu flüssigen Phrasengrenzen. Der Anteil separierender Konturen ist wie bei den Phrasengrenzen mit Zögerungslängung deutlich geringer als bei flüssiger Phrasierung (vgl. 3.3.1). Wie sich schon bei der Bearbeitung der phonetischen Merkmale gestörter Sprechflüssigkeit ergeben hat, unterscheidet sich die glottale Aktivität an syntaktischen Abbrüchen von der an flüssigen Phrasengrenzen (Kohler, Peters und Wesener 2001). An syntaktischen Abbrüchen findet sich häufig kurze und hart klingende Knarrstimme (Glottalisierung), die mit Glottalverschlüssen alterniert, während sich Knarrstimme am Ende von flüssigen Phrasen (Laryngalisierung) über einen größeren zeitlichen Bereich ausdehnt, weicher klingt und nicht mit Glottalverschlüssen alterniert. Laryngalisierung tritt nur in Verbindung mit einer stark fallenden tonalen Bewegung auf. Dagegen schneidet Glottalisierung den F0-Verlauf an Abbruchstellen unabhängig vom momentanen F0-Niveau ab und ist nicht von einer stark fallenden Kontur abhängig. Dies zeigt sich darin, daß an mehr als einem Viertel der Abbruchstellen Glottalisierung oder ein Glottalverschluß auftritt, stark fallende Konturen aber nur in 14% der Phrasengrenzen an syntaktischen Abbrüchen vorkommen (vgl. Tabelle 9).

## 4 Grade prosodischer Kohäsion und ihre kommunikative Funktion

Die Datenanalyse hat drei Typen von phonetischen Merkmalen an prosodischen Phrasengrenzen herausgestellt:

- tonale Merkmale - Reset und separierende Konturen
- temporale Merkmale - segmentelle Längung
- Pausen und/oder Atmen

### 4.1 Klassifikation von Phrasengrenzen nach phonetischen Kriterien

Die Merkmale treten allein oder in verschiedenen Bündelungen auf, wobei Pause und/oder Atmen praktisch immer zusammen mit mindestens einem der anderen Merkmalstypen vorkommt. Verschiedene Bündelungen kodieren unterschiedliche Grade der Kohäsion bzw. Separierung aufeinander folgender Einheiten. Erreicht die Separierung eine bestimmte Stärke, dann führt sie zur Wahrnehmung prosodischer Phrasengrenzen, die entsprechend dem perzeptorischen Gewicht verschiedener Bündelungen in ihrer Trennstärke skaliert sind. Diese Gradierung der Trennstärke zwischen prosodischen Phrasen reflektiert die inhaltliche Gliederung des Gesprochenen, die der Sprecher bewußt oder unbewußt vornimmt.

Aufgrund der statistischen Befunde des untersuchten Datenkorpus lassen sich turninterne Phrasengrenzen in zwei große Klassen einteilen, nämlich die nur mit tonalen und/oder temporalen Merkmalen verbundenen einerseits und diejenigen, bei denen Pause/Atmen als zusätzliches Merkmal hinzutritt. Die erstgenannte Klasse hat den geringsten Trennungsgrad und wird als **PG1** bezeichnet.

Nach den Untersuchungen von Butcher (1981), der drei Perzeptionsklassen von Pausendauern festgestellt hat ('lang' bei einem Median von 1400 ms, 'kurz' bei einem Median von 486 ms und 'keine Pausenwahrnehmung' bei einem Median von 126 ms), setzen wir für eine Subklassifikation der Phrasengrenzen mit Pause/Atmen einen Grenzwert von 700 ms zwischen 'kurzer' und 'langer' Pause und/oder Atmungsphase an. Die mit kurzer Pause/Atmen verbundene Phrasengrenze wird als **PG2** bezeichnet.

Lange Pausen und/oder Einatmungsphasen können nach nicht-separierenden Konturen auftreten (siehe Tabelle 4) und markieren dann Phrasengrenzen der Klasse **PG3**. Die verbleibenden turninternen Phrasengrenzen, an

denen alle drei Merkmalstypen auftreten (tonal, temporal und lange Pause/Atmen), bilden die Klasse **PG4**. Hier haben stark fallende, in Laryngalisierung endende Konturen /2./ vor nachfolgenden langen Pausen/Atmungsphasen eine besonders starke Trennfunktion. Die Separierung wird noch zusätzlich unterstützt, wenn die fallende Kontur ein früher Gipfel ist, der den Abschluß eines Argumentationsgangs signalisiert. Die langen Pausen bzw. Atmungsphasen der Klassen **PG3** und **PG4** können auch paralinguistische Phänomene wie Schmatzen, Räuspern, Schlucken einschließen.

Die turninternen Phrasengrenzen **PG1 - PG4** können an sämtlichen syntaktischen Strukturstellen vorkommen, also auch an solchen niedriger Ordnung, wie z.B. zwischen Präposition/Artikel/Possesivpronomen und nachfolgendem Adjektiv/Substantiv oder zwischen Personalpronomen und Verb. In diesen Fällen erhalten die Phrasengrenzen zusätzliches perzeptorisches Gewicht (Boomer und Dittman 1962, Butcher 1981), werden aber in der phonetisch orientierten Klassifizierung nicht getrennt.

Turnfinale Phrasengrenzen leiten einen Sprecherwechsel ein und werden der Klasse **PGt** zugeordnet. In dieser Position treten separierende Konturen in 78.3% der Fälle auf, darunter sind die stark fallenden Konturen /2./ und /2;/ mit zusammen 63,8% am häufigsten vertreten (Tabelle 9). Wie bei **PG4** signalisieren stark fallende frühe Gipfelkonturen auch bei **PGt** Finalität und inhaltliche Abgeschlossenheit, bei **PGt** jedoch bezogen auf den gesamten Redebeitrag eines Sprechers. Finale Längung zeigt sich in 74.3% der turnfinalen Reime.

Im Aufnahmeszenario der Terminabsprache mit Knopfdruck (Kohler, Pätzold und Simpson 1995), auf dem die hier untersuchte Datenbasis beruht, konnten die Dialogpartner nur abwechselnd sprechen, solange sie einen Knopf gedrückt hielten. Damit war die Turnübergabe technisch festgelegt. Die in die Klasse **PGt** eingereichten Phrasengrenzen dieses Korpus repräsentieren also nicht das gesamte phonetische Spektrum der Turnübergabe im freien Dialog. Deshalb sind weitere Untersuchungen an Daten aus einem Dialogszenario notwendig, in dem Sprecherwechsel frei gestaltet werden können und damit auch verschiedene Grade überlappenden Sprechens auftreten.

Neben den flüssigen turninternen Phrasengrenzen **PG1 - PG4** sowie den turnfinalen **PGt** treten Phrasengrenzen mit Zögerungslängung auf. In diesen Fällen ist die segmentelle Längung nicht nur an den finalen Reim gekoppelt, sondern kann sich an verschiedenen Positionen festmachen oder sich sogar über mehrere Silben erstrecken. Außerdem ist die Längung oft wesentlich extremer und der Melodierverlauf eine nichtseparierende Kontur (siehe Tabelle 9: 77.4% nicht separierende vs. 22.6% separierende Konturen). Infolge dieser unterschiedlichen Manifestation von Phrasengrenzen mit Zögerungslängung werden diese als **PGh** von **PG1 - PG4** differenziert.

## 4.2 Illustration verschiedener Phrasengrenzen anhand von Korpusbeispielen

Die vorgestellte Klassifizierung wird nun anhand von einschlägigen Beispielen aus dem Datenkorpus illustriert. Die Korpusbeispiele befinden sich als Audiodateien auf der Webseite des Instituts für Phonetik und digitale Sprachverarbeitung der Universität Kiel unter folgendem Link:

<http://www.ipds.uni-kiel.de/publikationen/audiobsp.en.html>

In der typographischen Präsentation wird jede einzelne Phrase durch eine Textzeile und zwei nachfolgende Merkmalszeilen dargestellt. Die erste Merkmalszeile gibt Information über Vorhandensein (+) bzw. Fehlen (-) folgender phonetischer Merkmale:

- tonale Merkmale: Reset (+R/-R) und separierende Konturen (+K/-K)
- temporale Merkmale: segmentelle Längung (+L/-L), bei +L zusätzlich +h für Zögerungslängung und -h für normale phrasenfinale Längung
- Pausen und/oder Atmen: (+P/-P), bei +P zusätzlich +l bzw. -l für Pausendauern über bzw. unter 700 ms

Des Weiteren wird turnfinales Vorkommen markiert. In diesem Fall werden die binären Merkmale Pause/Atmen und Reset nicht definiert: P0, R0. Da Reset ein mindestens gleich hohes F0-Maximum in einer Gipfelkontur nach der Phrasengrenze im Vergleich mit der letzten Gipfelkontur davor voraussetzt, muss er auch turnintern vor und nach Talkonturen undefiniert bleiben (siehe 3.1.3). Aus der jeweiligen Merkmalskombination ergibt sich die Notation des Phrasengrenztyps.

Die zweite Merkmalszeile spezifiziert auftretende intonatorische, segmentelle und phonatorische Besonderheiten, u.a. stark fallende frühe Gipfelkonturen /**2**)/+/**2**./ sowie hoch steigende Talkonturen /**?**/, Segmentelisionen, Pausen-/Atmungsdauern und glottale Phänomene im Bereich der Phrasengrenze. Der wortinitiale Vokaleinsatz am Beginn einer Phrase hat vier Realisierungsmöglichkeiten: Glottalverschluß und Glottalisierung /**Qq**/, Glottalverschluß /**Q**/, Glottalisierung /**q**/, weicher Einsatz /**Q-**/. Damit verbunden sind verschiedene Grade der Separierung bzw. Kohäsion zwischen den aufeinander folgenden Phrasen (Rodgers 1999).

#### 4.2.1 Beispiel 1: g192a005, Sprecherin: KAE

Text Phrase 1: *ja gut.*

Merkmale: -R, +K, +L -h, -P → **PG1**

Spezifikation: /2)/+/2./, Elision von /t/ in *gut*

Text Phrase 2: *dann komm ich da vorbei.*

Merkmale: +R, +K, +L -h, -P → **PG1**

Spezifikation: /2)/+/2./ im Hutmuster, /Q-/ im nachfolgenden *um*

Text Phrase 3: *um wieviel Uhr denn so etwa?*

Merkmale: R0, +K, +L -h, P0, turnfinal → **PGt**

Spezifikation: Talkontur mit /?/

Die erste Phrasengrenze hat in der Datenanalyse zwei positiv spezifizierte Merkmale bekommen (+K, +L), die zweite hingegen drei (+R, +K, +L). Außerdem findet sich an der ersten Phrasengrenze eine segmentelle Reduktion, die dem Faktor Längung entgegenwirkt. Schließlich führt der frühe Gipfel in einem Hutmuster am Ende der zweiten Phrase zu einer Klammerwirkung zwischen erster und zweiter Phrase im Kontrast zur dritten. Dies alles zusammen führt zum perzeptorischen Eindruck einer leicht stärkeren Separierung zwischen Phrase 2 und 3. Das weist darauf hin, daß der Phrasengrenztyp PG1 noch weiter subklassifiziert werden sollte. Dafür sind aber perzeptiv orientierte Experimente notwendig.

In diesem Beitrag spricht Sprecherin KAE sehr schnell und flüssig und prägt die prosodische Phrasierung phonetisch eher schwach aus, reflektiert aber dennoch die inhaltliche Gliederung. Sie gibt in der ersten Phrase einen Kommentar zu einem Vorschlag der Gesprächspartnerin, spezifiziert ihn in der zweiten Phrase und setzt in der dritten Phrase mit einer Frage neu ein, auf die die Turnübergabe folgt.

#### 4.2.2 Beispiel 2: g212a007, Sprecher: ANL

Text Phrase 1: *ja,*

Merkmale: R0, -K, +L -h, +P -l → **PG2**

Spezifikation: Ausatmen 190 ms

Text Phrase 2: *gerne.*

Merkmale: R0, -K, +L -h, +P -l → **PG2**

Spezifikation: Ausatmen 230 ms, /q/ im nachfolgenden *ich*

Text Phrase 3: *ich habe also Zeit vom Donnerstag, den zweiten*

### ***Juni***

Merkmale: R0, -K, +L -h, +P -l → **PG2**

Spezifikation: Schlucken, Pause 600ms

Text Phrase 4: ***bis Mittwoch, den achten,***

Merkmale: R0, -K, +L -h, +P +l → **PG3**

Spezifikation: Schmatzen, Einatmen 1460 ms, /**Q**-/ im nachfolgenden *und*

Text Phrase 5: ***und von Samstag, dem achtzehnten,***

Merkmale: -R, -K, +L -h, +P -l → **PG2**

Spezifikation: Ausatmen 100 ms

Text Phrase 6: ***bis Donnerstag,***

Merkmale: -R, -K, +L -h, -P → **PG1**

Spezifikation: keine

Text Phrase 7: ***den dreiundzwanzigsten,***

Merkmale: R0, -K, +L -h, +P +l → **PG3**

Spezifikation: Einatmen 1070 ms

Text Phrase 8: ***und dann wieder vom siebenundzwanzigsten bis zum dreißigsten.***

Merkmale: R0, +K, +L -h, P0, turnfinal → **PGt**

Spezifikation: /**2**)/ + /**2**./ im Hutmuster, Laryngalisierung

In diesem Beitrag spricht Sprecher ANL relativ langsam mit stark markierter hierarchischer Phrasengliederung. Der Beitrag enthält drei große Blöcke, die sich inhaltlich auf verschiedene Perioden beziehen, in denen er Zeit hat. Diese Blöcke sind durch **PG3** getrennt. Innerhalb der ersten beiden Blöcke gliedert er die Zeitspannen durch **PG2** in Beginn *von...* und Ende *bis...* Den ersten Block leitet er mit zwei affirmativen Anknüpfungen an den vorangegangenen Beitrag des Dialogpartners ein und markiert auch sie mit **PG2** und damit als Teil des ersten Blocks. Im Endteil des zweiten Blocks separiert er die Nennung des Wochentags von der des Datums durch eine schwächere **PG1**. Im dritten Block integriert er die beiden Datumsangaben der Zeitspanne, diesmal ohne Nennung der Wochentage, in einem Hutmuster, das auf eine stark fallende, frühe Gipfelkontur mit Übergang in Laryngalisierung endet und mit dem er signalisiert, daß der Redebeitrag auch inhaltlich zum Abschluß kommt.

### 4.2.3 Beispiel 3: g192a001, Sprecherin: KAE

Text Phrase 1: *da bin ich leider schon völlig ausgebucht,*

Merkmale: +R, -K, +L -h, -P → **PG1**

Spezifikation: keine

Text Phrase 2: *da bin ich leider verreist*

Merkmale: +R, +K, -L -h, -P → **PG1**

Spezifikation: /2)/ + /2; / im Hutmuster

Text Phrase 3: *bis zum dritten.*

Merkmale: +R, -K, +L -h, -P → **PG1**

Spezifikation: nachfolgender High Prehead, /Q-/ im nachfolgenden *ich*

Text Phrase 4: *ich könnte höchstens*

Merkmale: -R, -K, +L -h, +P -l → **PG2**

Spezifikation: Einatmen 600 ms

Text Phrase 5: *nach den Feiertagen*

Merkmale: +R, -K, +L -h, +P -l → **PG2**

Spezifikation: /Q/ im nachfolgenden *am*, Pause 140 ms

Text Phrase 6: *am vierten.*

Merkmale: R0, +K, +L -h, P0,turnfinal → **PGt**

Spezifikation: keine

In diesem Beispiel handelt es sich um dieselbe Sprecherin KAE wie in Bsp. 1, sie spricht jedoch hier nicht so flüssig. Das Beispiel veranschaulicht sehr schön, wie die prosodische Phrasierung weniger die inhaltliche Gliederung als vielmehr den Prozeß der Äußerungsplanung widerspiegelt. Das prägt sich darin aus, daß die Gesamtinformation durch die prosodische Gliederung in viele kleine Informationsstücke unterteilt wird. An sich wäre als inhaltliche Gliederung in Übereinstimmung mit der syntaktischen Form folgende Phrasierung zu erwarten:

1. *Da bin ich leider schon völlig ausgebucht.*
2. *Da bin ich leider verreist bis zum dritten.*
3. *Ich könnte höchstens nach den Feiertagen am vierten.*

Zum Ende des Beitrags nehmen die Planungsschwierigkeiten zu, was sich in der Verwendung der stärker separierenden **PG2** (mit Pause bzw. Einatmen) zeigt, dies ganz besonders an der inhaltlich und syntaktisch nicht

motivierten Grenze zwischen Phrase 4 und Phrase 5. Diese konkurrierenden Erfordernisse der inhaltlichen Strukturierung einerseits und der Äußerungsplanung andererseits führen hier zu einer Divergenz zwischen prosodischer Phrasierung und syntaktischer Gliederung.

Dennoch entsteht hier nicht der Eindruck von Sprechflußstörung. Es handelt sich vielmehr um sukzessives Nachliefern von Information, wie sie der Sprecherin einfällt. So kommt die Sprecherin nach Phrase 5 schon zu einem Abschluß, was sie tonal durch eine fallende separierende Kontur signalisiert. Dann schiebt sie jedoch nach kurzer Pause die Datumsinformation nach. Dieses Beispiel legt nahe, daß eventuell auch **PG2** weiter subklassifiziert werden muss, je nachdem, ob die kurze Pause nach separierenden Konturen auftritt oder nicht. Auch hier sind zusätzliche, vor allem perzeptorisch orientierte Untersuchungen notwendig.

#### 4.2.4 Beispiel 4: g251a005, Sprecherin: SIK

Text Phrase 1: *das ginge.*

Merkmale: -R, +K, +L -h, +P -l → **PG2**

Spezifikation: /1)/ + /2;/ im Hutmuster, Ausatmen 250 ms

Text Phrase 2: *das ist kein Problem.*

Merkmale: +R, +K, +L -h, +P +l → **PG4**

Spezifikation: Aus- und Einatmen 2000 ms, Schmatzen

Text Phrase 3: *wie ist es mit*

Merkmale: R0, -K, +L -h, -P → **PG1**

Spezifikation: /Qq/ im nachfolgenden *einer*

Text Phrase 4: *einer eintägigen Arbeitssitzung?*

Merkmale: R0, +K, +L -h, +P +l → **PG4**

Spezifikation: Ausatmen 1000 ms

Text Phrase 5: *wenn Sie*

Merkmale: R0, -K, +L +h, -P → **PGh**

Spezifikation: /q/ im nachfolgenden *am*

Text Phrase 6: *am Donnerstag auch frei hätten,*

Merkmale: +R, -K, +L -h, -P → **PG1**

Spezifikation: keine

Text Phrase 7: *dann könnten wir uns noch einen anderen*

***Donnerstag auskucken.***

Merkmale: R0, +K, +L -h, P0, turnfinal → **PGt**

Spezifikation: /1)/ + /2;/ im Hutmuster

Sprecherin SIK gliedert diesen Beitrag in drei inhaltliche Blöcke:

1. kommentierende Antwort auf den Vorschlag der Gesprächspartnerin
2. neue Frage an die Gesprächspartnerin
3. konkreter Vorschlag für eine Antwort.

Diese Einteilung ist typisch für Redebeiträge im Terminabspracheszenario und wird im vorliegenden Beispiel durch sehr starke Phrasengrenzen (**PG4**) zwischen den Blöcken markiert.

Innerhalb der Blöcke treten einerseits die schwächeren Grenzen **PG1** und **PG2** auf, andererseits findet sich zwischen Phrase 5 und Phrase 6 eine **PGh** mit leichter Zögerungslängung, die perzeptuell durch untypische Vokallängung auf ebenem F0-Muster in einem unakzentuierten Funktionswort hervorgerufen wird. **PG1** zwischen Phrase 3 und Phrase 4 illustriert den Fall, wo die üblichen phonetischen Merkmale dieses Phrasengrenztyps durch ihr Auftreten an einer syntaktisch niedrigen Strukturstelle (nach der Präposition in einer Präpositionalphrase) perzeptorisch stärker gewichtet werden. Dies wird durch Glottalverschluß und Glottalisierung am Beginn des nächsten Wortes noch verstärkt. Bezüglich der Wahrnehmung interagiert hier also die syntaktische Struktur mit der prosodischen Gliederung.

**4.2.5 Beispiel 5: g191a018, Sprecherin: KAE**

Text Phrase 1: ***ja, dann***

Merkmale: -R, -K, +L +h, -P → **PGh**

Spezifikation: [n] in *dann*: 365 ms

Text Phrase 2: ***vielleicht***

Merkmale: +R, -K, +L +h, -P → **PGh**

Spezifikation: *vielleicht* insgesamt gelängt, Elision des [i]

Text Phrase 3: ***das Wochenende***

Merkmale: +R, -K, +L +h, -P → **PGh**

Spezifikation: [ə] in *Wochenende*: 300 ms

Text Phrase 4: ***vom elften bis zwölften Dezember?***

Merkmale: R0, +K, +L -h, P0 → **PGt**

Spezifikation: nur noch leicht gezögertes [f] in *vom*, sonst wieder sehr flüssig

In diesem Beispiel ist die Sprecherin wiederum KAE, wie in den Beispielen 1 und 3. Sie hat in den Phrasen 1 bis 3 starke Zögerungslängung: in den Phrasen 1 und 3 im letztem Segment, in Phrase 2 ausgedehnt über das ganze Wort. Diese Unflüssigkeit reflektiert inhaltliche Planungsprobleme der Sprecherin bei der Terminfindung. Erst Phrase 4 wird wieder flüssig.

## 4.3 Skalierung prosodischer Kohäsion

### 4.3.1 Rangskala prosodischer Kohäsion in flüssiger Phrasierung

Eine grundlegende Funktion der Prosodie ist die syntagmatische Gliederung des Gesprochenen durch eine phonetisch signalisierte hierarchische Klammerung sprachlicher Einheiten. Diese prosodische Gliederung spiegelt eine inhaltliche, semantisch-pragmatische Strukturierung sprachlicher Interaktion wider und verläuft parallel zu einer formal-syntaktischen Phrasierung, mit der sie koinzidieren oder interagieren oder konfligieren kann. Das in 4.1 aufgestellte und in 4.2 durch spontansprachliche Dialogbeispiele illustrierte System von Phrasengrenzen **PG1** - **PG4** repräsentiert eine solche Hierarchie syntagmatischer Gliederung mit progressiv abnehmender Kohäsion von **PG1** bis **PG4** zwischen aufeinander folgenden prosodischen Phrasen. Entsprechend dem durch diese phonetischen Grenzmarkierungen signalisierten unterschiedlichen Kohäsionsgrad wird eine hierarchische inhaltliche Strukturierung vorgenommen, wie sie Sprecher Hörern vermitteln wollen. Diese Grade abnehmender Kohäsion zwischen prosodischen Phrasen sollen als K-1, K-2, K-3, K-4 entsprechend **PG1** - **PG4** bezeichnet werden.

Innerhalb prosodischer Phrasen manifestieren sich ebenfalls verschiedene Kohäsionsgrade durch prosodische Phrasierung. So können phrasenintern globale Melodiemuster (z.B. Hutmuster sowie kontinuierlich steigende oder fallende Konturen) Sequenzen sprachlicher Einheiten (Wörter, Silben) integrieren. Hierbei hat insbesondere das mit einem frühen Gipfel endende Hutmuster eine Klammerfunktion (siehe Phrase 8 in Bsp. 2). Dieser Kohäsionsgrad soll mit K1 bezeichnet werden. Wenn innerhalb einer prosodischen Phrase statt solcher globaler Melodiemuster Abfolgen von Gipfeln und Tälern auftreten, dann werden die darin verbundenen sprachlichen Einheiten weniger stark integriert (Peters, Kohler und Wesener 2003). Dieser Kohäsionsgrad wird als Grenzmarke zwischen phrasenintern zunehmender und phrasenextern abnehmender Kohäsion als K0 kategorisiert. Es ergibt sich somit folgende Rangskala abnehmender prosodischer Kohäsion zwischen sprachlichen Einheiten:

K1 | K0 | K-1 | K-2 | K-3 | K-4

Sie wird einerseits eingesetzt, um inhaltliche Strukturen phonetisch hierarchisch abzubilden, und sie definiert andererseits Domänen für lautliche Veränderungsprozesse auf satzphonetischer Ebene (Kohler 2001b). K1 setzt z.B. einen prosodischen Rahmen für sehr häufiges Auftreten segmenteller Assimilationen und Elisionen über Wortgrenzen (Kohler 2001a, 2001c). Je niedriger der Kohäsionsgrad wird, um so weniger wahrscheinlich werden diese Anpassungsprozesse. Intervenieren Pausen und andere paralinguistische Phänomene, also ab K-2, dann wird die artikulatorische und phonatorische Angleichung an Wortgrenzen blockiert.

In prosodisch definierten phraseninternen Klammerungen der Stufe K1 können starke lautliche Abschwächungen auftreten, die zu einer Nivellierung aufeinander folgender artikulatorischer Öffnungs- und Schließbewegungen führen und so ihre Anzahl reduzieren, wenn bestimmte syntaktische, lexikalische und phonetische Voraussetzungen gegeben sind. So kann die komplette Fusion von Wörtern zu neuen Worteinheiten eintreten, z.B. in der Klitisierung (Kohler 1998, Wesener 2001). Beispiele liefern Verbindungen von Funktionswörtern, unter anderem von Hilfsverb und Personalpronomen:

*haben wir* [hame]  
*sollen wir* [zɔmɐ]  
*hat er* [hadɐ]

Aber auch wortintern finden sich derartige extreme Reduktionsprozesse bei semantischer Verflachung, wie z.B. in dem zur Floskel gewordenen *eigentlich* [amɪ] (Kohler 2001a, Kohler 2001c). Wir können also die zunächst *prosodisch* festgelegte Rangskala K1 - K-4 phrasenintern zu noch höheren Kohäsionsgraden fortsetzen, die durch zunehmende *artikulatorische* Integration definiert sind.

#### 4.3.2 Prosodische Kohäsion in vier Phrasierungstypen

Die Interaktion zwischen prosodischer Gliederung und formal-syntaktischer Struktur wird deutlich, wenn **PG1** - **PG4** an syntaktischen Knoten niedriger Ordnung aufgrund von Planungsschwierigkeiten des Sprechers auftreten. Die gleichen phonetischen Parameter in vergleichbarer Dauer werden perzeptorisch durch die Diskrepanz zur syntaktischen Struktur gewichtet und als Häsitationsphänomene wahrgenommen. Die Beispiele 4 und 5 illustrieren dies. Daneben treten, unabhängig von der syntaktischen Struktur, wiederum aufgrund von Planungsschwierigkeiten, aber auch zur Verstärkung Phrasengrenzen mit Zögerungslängung **PGh** auf, die ebenfalls als gestörter Sprechfluß perzipiert werden können. Die Beziehung zwischen **PG1** - **PG4**

an niedrigen syntaktischen Knoten einer- und **PGh** andererseits bedarf noch eingehender Untersuchung.

Es sind somit für den einzelnen Redebetrag im Dialog vier Phänomenbereiche der Phrasierung zu unterscheiden:

- flüssige Phrasengrenzen **PG1 - PG4**, die auf einer Rangskala abnehmender Kohäsion gradiert werden können
- Phrasengrenzen **PG1 - PG4**, die durch Interaktion mit der syntaktischen Gliederung in ihrer Trennstärke höher gewichtet werden
- Phrasengrenzen mit Zögerungslängung **PGh**
- Phrasengrenzen an syntaktischen Abbrüchen **PG/**

Durch **PG1 - PG4** mit syntaktischer Interferenz, durch **PGh** und durch **PG/** werden unterschiedliche Grade der Unflüssigkeit aufgrund von inhaltlichen und formalen Planungsschwierigkeiten vermittelt. Ob und in welcher Weise sich **PGh** und **PG/** ebenfalls in ihrer Trennstärke skalieren lassen, wie sie sich mit syntaktisch interagierenden **PG1 - PG4** zu einem Gesamtsystem des unflüssigen Sprechens zusammenfügen und welches die Beziehungen zur Kohäsionsskala der flüssigen Phrasengrenzen sind, bleiben offene Fragen, zu deren Beantwortung weitere Datenanalysen, vor allem in stärker spontan-sprachlichen, natürlicheren Szenarien als der Terminabsprache erforderlich sind.

#### 4.3.3 Kohäsion im Dialog

Schließlich setzt sich die Skala der Kohäsion/Separierung über den Einzelbeitrag in die verschiedenen Arten der Turnübergabe fort. Dies ist in der vorliegenden Untersuchung bereits mit der generellen Kategorie der turnfinalen Phrasengrenzen **PGt** angelegt worden.

### 4.4 Anpassung des Kieler Intonationsmodells KIM und des Symbolisierungssystems PROLAB

Die Ergebnisse der vorliegenden Untersuchung erlauben nunmehr post hoc eine Subklassifikation der KIM-Kategorie 'Phrasengrenze' in vier unterschiedlich stark separierende Klassen und den Ersatz der PROLAB-Notation **PGn** durch **PG1 - PG4**. Dies geschieht aufgrund von Merkmalsbündeln, die sich aus Messwerten und Symbolisierungen der Datenbank ergeben. Es fehlt im Bereich der Messwerte noch eine umfassende Untersuchung der Pausen- und

Atmungsdauern, außerdem eine Trennung von Aus- und Einatmen, was jedoch aus der vorhandenen Datenbank ableitbar ist. Dann wird es auch möglich, die Subklassifikation auf empirischer Basis automatisch durchzuführen.

Getrennt von den flüssigen Phrasengrenzen **PG1** - **PG4** kann aus **z:PGn** die Phrasengrenze mit Zögerungslängung **PGh** automatisch gewonnen werden. Für automatische Subklassifizierungen von **PGh** und **PG/** sind noch weitere Korpusanalysen erforderlich. Die Einführung einer Gewichtung von **PG1** - **PG4** aufgrund syntaktischer Interaktion setzt syntaktische Annotation der Datenbank voraus und ist deshalb mit dem *Kiel Corpus of Spontaneous Speech* nur exemplarisch möglich.

## 5 Diskussion und Ausblick

An turninternen Phrasengrenzen liegt in 93.1% der Fälle mindestens eines der phonetischen Merkmale *tonaler Einschnitt*, *finale Längung* oder *Pause/Atmen* vor. Unter den verbleibenden 6.9% der Fälle ohne eines der genannten Merkmale ist fast die Hälfte ebenfalls durch tonale Einschnitte markiert, allerdings nur nach mittelhoch steigenden Melodiemustern, die nicht in die Klasse separierender Konturen eingereiht wurden. In einem weiteren Drittel der Restfälle haben zusätzliche Merkmale wie Glottalverschuß/Glottalisierung, Änderung des Sprechtempos und des Sprechregisters sowie Verlagerung der Längung aus dem finalen Reim heraus in andere Teile eines Wortes eine trennende Wirkung, also Merkmale, die von der automatischen Klassifikation nicht erfaßt wurden und z.T. nicht automatisch aus der Datenbank extrahierbar sind. In etwa 20% der Restfälle liegen offenbar durch die syntaktische Gliederung gesteuerte Etikettierfehler vor. Auf dieser statistischen Basis ist Hypothese 1 in Bezug auf turninterne Phrasengrenzen prinzipiell bestätigt.

Finale Längung kommt turnintern mit großer Häufigkeit vor, häufiger als Reset, separierende Intonationskonturen und Pause/Atmen. Wird jedoch die Phrasengrenze nur durch eines der genannten phonetischen Merkmale markiert, dann ist Reset (10.3%) bzw. Reset oder separierende Kontur, d.h. die tonale Gruppe (12.6%), häufiger als Längung 9.4%. Faßt man außerdem die tonalen Merkmale Reset und separierende Kontur auch bei gemeinsamen Auftreten zusammen, dann ist diese Klasse der durch F0 kodierten Merkmale die größte. Unter diesen Bedingungen wird Hypothese 2(a) verworfen.

Finale Längung ist turnfinal stärker ausgeprägt als turnintern. Damit ist Hypothese 2 (b) bestätigt. Aufgrund dieses Ergebnisses wird die turnfinale Phrasengrenze aus der Skalierung der flüssigen turninternen Phrasengrenzen (siehe Diskussion der Hypothese 3) herausgenommen und als **PGt** kategorisiert.

Die phonetischen Merkmale zur Signalisierung von Phrasengrenzen treten im untersuchten Datenkorpus in vielfältiger Bündelung auf. Die häufigste Bündelung ist die von finaler Längung und Pause/Atmen. Der mit solchen unterschiedlichen Bündelungen verbindbare Grad der Separierung bzw. Kohäsion zwischen aufeinander folgenden prosodischen Phrasen ist modelliert und anhand ausgewählter Beispiele exemplifiziert worden. Die flüssigen Phrasengrenzen, d.h. diejenigen ohne Zögerungslängung und ohne syntaktischen Abbruch, lassen sich in Abhängigkeit von Zahl und Ausprägung der phonetischen Merkmale auf einer Rangskala abnehmender Kohäsion zwischen prosodischen Phrasen als **PG1** (K-1), **PG2** (K-2), **PG3** (K-3) und **PG4** (K-4) anordnen. Die unterschiedlichen Kohäsionsgrade bilden ihrerseits die inhaltliche Gliederung und die Äußerungsplanung ab. Damit ist Hypothese 3 teilweise bestätigt.

Zur weiteren Stützung der Hypothese 3 sind noch folgende sukzessive Arbeitsschritte erforderlich:

- Trennung von Pause, Eintatmen und Ausatmen
- automatische Messung der jeweiligen Pausen-/Atmungsdauer und Einfügen des Wertes in die Klassifikationsmatrix jeder Phrasengrenze **PGn**.
- Verteilung der Pausen- und Atmungsdauern im Gesamtkorpus
- perzeptorische Überprüfung des Schwellenwertes zur Trennung kurzer und langer Pausen mit Bezug auf die Häufigkeitsverteilung
- automatische Zuordnung der flüssigen Phrasengrenzen zu den Klassen **PG1 - PG4**
- Häufigkeitsverteilung über die Klassen **PG1 - PG4**
- Überprüfung einer eventuellen Subklassifikation von **PG1** in zwei Klassen mit unterschiedlicher Zahl und Art definierender Merkmale sowie von **PG2** in zwei Klassen mit separierender bzw. nicht separierender Kontur
- Validierung des Klassifikationssystems flüssiger Phrasengrenzen durch phonetische und inhaltliche Detailbeschreibung einer großen Anzahl zufällig aus dem Korpus herausgegriffener Dialogbeiträge.

Auch die Hypothese 4, daß sich Phrasengrenzen mit Störungen des Sprechflusses (Zögerungslängung/syntaktische Abbrüche) in verschiedenen phonetischen Merkmalen von flüssiger Phrasierung unterscheiden, wird von den

Ergebnissen bestätigt. Phrasengrenzen mit Zögerungslängung oder syntaktischen Abbrüchen unterscheiden sich besonders in Bezug auf die Ausprägung der phrasenfinalen Kontur von denen der flüssigen Phrasierung. An den unflüssigen Phrasengrenzen treten deutlich weniger separierende Konturen auf, stattdessen kommt es in etwa vier Fünftel der Fälle nur zu geringen tonalen Bewegungen. Weiterhin hat ein qualitativer Vergleich der Dauerstrukturen von Zögerungslängung und finaler Längung gezeigt, daß der Bereich, in dem sich Zögerungslängung manifestiert, stark variiert und sich im Gegensatz zur finalen Längung häufig über mehrere Silben oder sogar Wortgruppen erstreckt. Phrasengrenzen an syntaktischen Abbrüchen wiederum unterscheiden sich von Phrasengrenzen ohne Störungen der Sprechflüssigkeit durch das gehäufte Auftreten von Glottalisierung und Glottalverschluß im Unterschied zu finaler Laryngalisierung.

Ziel der weiteren Untersuchungen wird es sein, ein Gesamtmodell der syntagmatischen prosodischen Phrasierung zu entwickeln, das neben flüssigen Phrasengrenzen und flüssigen phraseninternen Gruppierungen unterschiedlicher Kohäsionsgrade auch Phrasierung mit Zögerungslängung und syntaktische Abbrüche einbezieht. Schließlich muß dieses Gesamtmodell auch über den einzelnen Redebeitrag hinausgehen und die Kohäsion in Dialogstrukturen thematisieren. Eine solche Modellierung basiert auf der Bestimmung prosodischer Merkmalsbündel am Ende von Redebeiträgen und ihrer Zuordnung zu einer pragmatischen Klassifikation verschiedener Turnübergaben. Eine derartige Untersuchung setzt spontansprachliche Daten aus einem anderen Szenario als dem der Terminabsprache mit Knopfdruck voraus. Es kann dafür die Datenbasis 'Videotask-Szenario Lindenstraße' (Peters 2000b) zugrunde gelegt werden, die inzwischen komplett segmentell und prosodisch etikettiert vorliegt.

## Danksagung

Wir danken stud. phil. Nadeshda Krutisch für Datenauswertung und Dipl. Psych. Ernst Dombrowski für Beratung zur Statistik.

## Literatur

Batliner, A., R. Kompe, A. Kießling, E. Nöth, H. Niemann und U. Kilian (1995). The prosodic marking of phrase boundaries: Expectations and

- results. In A. J. Rubio Ayuso und J. M. López Soler (Eds.), *Speech Recognition and Coding: New Advances and Trends*, Volume 147 of *NATO ASI Series F*, pp. 325–328. Berlin, Heidelberg, New York: Springer.
- Boomer, D. S. und A. T. Dittman (1962). Hesitation pauses and juncture pauses in speech. *Language and Speech* 5, 215–220.
- Butcher, A. R. (1981). *Aspects of the speech pause: Phonetic correlates and communicative functions*. AIPUK 15.
- Gallwitz, F., A. Batliner, J. Buckow, R. Huber, H. Niemann und E. Nöth (1998). Integrated recognition of words and phrase boundaries. In *Proceedings ICSLP'98, Sydney*, Volume 7, pp. 2883–2886.
- IPDS (1995). *The Kiel Corpus of Spontaneous Speech*, Volume 1, CD-ROM#2. Kiel: Institut für Phonetik und digitale Sprachverarbeitung.
- IPDS (1996). *The Kiel Corpus of Spontaneous Speech*, Volume 2, CD-ROM#3. Kiel: Institut für Phonetik und digitale Sprachverarbeitung.
- Kohler, K. J. (1991). A model of German intonation. In K. J. Kohler (Ed.), *Studies in German Intonation*, AIPUK 25, pp. 295–368.
- Kohler, K. J. (1997). Modelling prosody in spontaneous speech. In Y. Sagisaka, N. Campbell und N. Higuchi (Eds.), *Computing Prosody*, pp. 187–210. Berlin/Heidelberg/New York/Tokyo: Springer.
- Kohler, K. J. (1998). The disappearance of words in connected speech. *ZAS Papers in Linguistics* 11, 21–33. Conference on ‘The word as a phonetic unit’, ZAS, Berlin, October 22–23, 1997.
- Kohler, K. J. (2001a). Articulatory dynamics of vowels and consonants in speech communication. In *Journal of the International Phonetic Association, JIPA*, pp. 1–16.
- Kohler, K. J. (2001b). The investigation of connected speech processes: Theory, method, hypotheses and empirical data. In K. J. Kohler (Ed.), *Sound Patterns of Spontaneous German Speech: Symbolic Structures and Gestural Dynamics*, AIPUK 35, pp. 1–32.
- Kohler, K. J. (2001c). Variability of opening and closing gestures in speech communication. In K. J. Kohler (Ed.), *Sound Patterns of Spontaneous German Speech: Symbolic Structures and Gestural Dynamics*, AIPUK 35, pp. 33–96.
- Kohler, K. J., M. Pätzold und A. Simpson (1995). From scenario to segment: the controlled elicitation, transcription, segmentation and labeling of spontaneous speech. *Arbeitsberichte des Instituts für Phonetik*

*und digitale Sprachverarbeitung der Universität Kiel (AIPUK) 29*, 1–141.

- Kohler, K. J., B. Peters und T. Wesener (2001). Interruption glottalization in German spontaneous speech. In *Disfluency in Spontaneous Speech, DISS 01*, Edinburgh, pp. 45–48.
- Kompe, R., A. Kießling, H. Niemann, E. Nöth, A. Batliner, S. Schachtl, T. Ruland und H. U. Block (1997). Improving parsing of spontaneous speech with the help of prosodic boundaries. In *1997 IEEE International Conference on Acoustics, Speech, and Signal Processing, April 21-24, Munich, Germany*, Volume 2, pp. 811–814.
- Peters, B. (2000a). Individuelle und geschlechtsspezifische Unterschiede in der prosodischen Gestaltung deutscher Lese- und Spontansprache. In H. D. Schlosser (Ed.), *Sprache und Kultur*, Forum Angewandte Linguistik 38. Frankfurt am Main: Lang.
- Peters, B. (2000b). ‘Video Task’ oder ‘Daily Soap Szenario’: Ein neues Verfahren zur kontrollierten Elizitation von Spontansprache. URL [http://www.ipds.uni-kiel.de/pub\\_exx/bp2001\\_1/Linda21.html](http://www.ipds.uni-kiel.de/pub_exx/bp2001_1/Linda21.html).
- Peters, B., T. Wesener und K. Kohler (2003). Melodische Satzakkentmuster in prosodischen Phrasen deutscher Spontansprache: Statistische Verteilung und sprachliche Funktion. Manuskript.
- Price, P. J., M. Ostendorf, S. Shattuck-Hufnagel und C. Fong (1991). The use of prosody in syntactic disambiguation. *Journal of the Acoustical Society of America* 90, 2956–2970.
- Rodgers, J. E. J. (1999). Three influences on glottalization in read and spontaneous German speech. In K. J. Kohler (Ed.), *Phrase-level Phonetics and Phonology of German*, AIPUK 34, pp. 177–284.
- Wang, M. Q. und J. Hirschberg (1992). Automatic classification of intonational phrase boundaries. *Computer Speech and Language* 6, 175–196.
- Wesener, T. (2001). Aspects of ‘clitization’ and style in German function words. In K. J. Kohler (Ed.), *Sound Patterns of Spontaneous German Speech: Symbolic Structures and Gestural Dynamics*, AIPUK 35, pp. 125–148.
- Wightman, C. W., S. Shattuck-Hufnagel, M. Ostendorf und P. J. Price (1992). Segmental durations in the vicinity of prosodic phrase boundaries. *Journal of the Acoustical Society of America* 91, 1707–1717.