

23.01.2003

Rheinische Friedrich-Wilhelms-Universität Bonn
Institut für Kommunikationsforschung und Phonetik
Prof. Dr. Wolfgang Hess

Grundlagen der Phonetik

4. Suprasegmentale Merkmale und Prosodie

4.1 Zur Terminologie

Suprasegmentale Merkmale sind gekennzeichnet dadurch, dass sie den segmentalen Merkmalen überlagert, aber nicht zeitlich auf diese begrenzt sind. Hierunter werden drei Merkmalsbereiche subsumiert (Lehiste, 1970):

- 1) *Quantität* – Phänomene, die mit der zeitlichen Dauer der Realisierung sprachlicher Einheiten zu tun haben;
- 2) *Intensität* – hierunter wird alles verstanden, was unter die Kategorie Betonung und Akzentuierung fällt;
- 3) *Intonation* – alle melodischen Phänomene, die sich in Form von Tonhöhenverläufen und -veränderungen in der Sprache äußern, also letztlich im Verlauf der Sprachgrundfrequenz F_0 .

Allen diesen Größen ist gemeinsam, dass sie in der Schriftform der Sprache, wenn überhaupt, nur unzulänglich repräsentiert sind.

Suprasegmentale Merkmale lassen sich i.a. schlecht in die segmentale Ebene eingliedern. An sich ist es notwendig, die mit der Segmentation in Abschnitt 3.2.1 postulierte strenge sprachliche Linearität (d.h., lückenlose und überlappungsfreie Abfolge) zu verlassen und weitere Stränge zu eröffnen, die zur segmentalen Beschreibung parallel verlaufen und beispielsweise die Transkription der drei suprasegmentalen Merkmale aufnehmen. [In neueren Strömungen der Phonetik, beispielsweise der autosegmentalen Phonetik (Goldsmith, 1990; Goldstein und Browman, 1989) werden solche Ansätze systematisch verfolgt.]

Die Literatur kennt zum vorliegenden Themenbereich drei Oberbegriffe:

- Suprasegmentale Eigenschaften (Suprasegmentalia);
- Prosodie;
- Intonation.

Unglücklicherweise sind diese Termini oft synonym, zirkulär oder widersprüchlich definiert; die Ebenen sollten aber – wie überall – auseinander gehalten werden. So definiert beispielsweise Bußmann (1990) Prosodie als

“Gesamtheit sprachlicher Eigenschaften wie Akzent, Intonation, Quantität, Sprechpausen. Sie beziehen sich im allgemeinen auf Einheiten, die größer sind als ein Phonem. Zur Prosodie zählt auch die Untersuchung von Sprechgeschwindigkeit, Rhythmus und Sprechpausen.” (BUSSMANN, 1990:618).

Intonation im weiteren Sinn wird an anderer Stelle dann wie folgt definiert:

“Gesamtheit der prosodischen Eigenschaften von sprachlichen Äußerungen, die nicht an den Einzellaut gebunden sind.” (BUSSMANN, 1990:352)

Wir behelfen uns an dieser Stelle mit einer Definition von Hierarchien (Bild 4.1). Demnach ist der Begriff *Prosodie* enger ausgelegt als der der Suprasegmentalia: Es gibt suprasegmentale Merkmale, die nicht zur Prosodie gehören, aber nicht umgekehrt. *Intonation* wiederum soll im

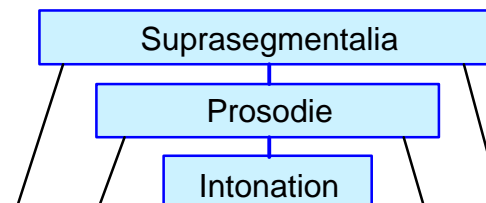


Bild 4.1. Begriffliche Hierarchie zu suprasegmentalen Merkmalen und Prosodie

Rahmen unserer Begriffshierarchie noch enger definiert werden, indem der Begriff auf die Beschreibung und Modellierung *melodischer* Phänomene beschränkt wird.

Eine Ursache für die unklare Abgrenzung der Begriffe ist wohl die Interaktion der Parameter auf akustischer Ebene.

Was ist nun *Prosodie*? Das Wort stammt aus dem (Alt-)Griechischen: “Pros-odie” – das Hinzugesungene – und beschreibt damit rhythmische und melodische Phänomene. In der Literaturwissenschaft steht Prosodie für die Untersuchung von Versdichtungen, Versmaßen und deren Rhythmen. In Linguistik und Phonetik steht Prosodie für die rhythmischen und melodischen Eigenschaften von Sprache allgemein, von einer Einzelsprache sowie von sprachlichen Äußerungen.

Nicht zur Prosodie, wohl aber zu den suprasegmentalen Merkmalen gehören (extralinguistische) sprecherspezifische Merkmale wie Stimmumfang, mittlere Sprachgrundfrequenz oder mittlere Lautstärke einer Stimme. Einige Autoren zählen auch paralinguistische Phänomene (Emotion oder Sprechstil) nicht zur Prosodie, auch wenn sie sich in prosodischen Merkmalen niederschlagen.

Suprasegmentale Merkmale können auf die segmentale Ebene durchgreifen, müssen dies aber nicht. Inwieweit sie dies tun oder nicht tun, hängt von der jeweiligen Einzelsprache ab. Wir können dies in der Aussage von Wunderlich (1988) so formulieren:

Es gibt eine begrenzte Zahl von Freiheitsgraden im rhythmischen und tonalen Bereich. Es hängt von der jeweiligen Sprache ab, sie a) nicht zu nutzen, b) vorwiegend im segmentalen Bereich zu nutzen, c) vorwiegend im suprasegmentalen Bereich zu nutzen.

In Abschnitt 4.2 beschreiben wir die Fälle, in denen suprasegmentale Merkmale im segmentalen Bereich phonologisch distinktiv sind. Abschnitt 4.3 gibt dann eine kurze Einführung in die Prosodie. Als einführende Literatur sei hier die sehr systematische Darstellung von Lehiste (1970) empfohlen.

4.2 Verwendung suprasegmentaler Merkmale im segmentalen Bereich

4.2.1 Quantität auf der segmentalen Ebene

Einige wesentliche Beispiele hierzu haben wir bereits in Abschnitt 3.1 kennengelernt. Von den suprasegmentalen Eigenschaften ist die Quantität diejenige, die am engsten mit einzelnen Segmenten assoziiert ist. Sie ist also meist auf Einzellaute beschränkt, überlagert sich dort aber den qualitativen Eigenschaften.

Ist die Quantität phonologisch distinktiv, so werden in der Regel zwei Stufen unterschieden (*kurz* vs. *lang*). Drei Stufen (*kurz – lang – überlang* oder *kurz – halblang – lang*) kommen in einigen Sprachen (z.B. dem Estnischen [e:stnɪʃ]) oder Dialekten (im deutschen Sprachraum z.B. das Trierer Moselfränkisch, vgl. Ternes, 1987:116) vor, sind aber, aufs Ganze gesehen, selten.

Die Quantität wird auf segmentaler Ebene in den verschiedenen Sprachen verschieden genutzt.

- Im Deutschen ist die Quantität im Bereich der Vokale mit zwei Stufen distinktiv, nicht aber im Bereich der Konsonanten (vgl. die Diskussion im Abschnitt 3.2.2).
- Im Italienischen ist die Quantität im Bereich der Konsonanten mit zwei Stufen distinktiv, nicht aber im Bereich der Vokale (vgl. das Beispiel in Abschnitt 3.1).
- In Sprachen wie Finnisch, Ungarisch oder Latein ist die Quantität mit jeweils zwei Stufen sowohl für Vokale als auch für Konsonanten distinktiv, und zwar jeweils unabhängig voneinander, so dass z.B. für das Finnische Minimalquadrupel existieren (Ternes, 1987:115):

muta /muta/ – Schlamm *muuta* /mu:ta/ – anderes (Partitiv Sing.)
mutta /mut:a/ – aber *muuttaa* /mu:t:a:/ – ändern

(An Stelle der Transkription /mut:a/ mit dem Längungs-Diakritikum ist für Konsonanten auch die Transkription mit einer Geminate, also hier /mutta/, üblich.)

- In Sprachen wie Spanisch, Russisch oder Französisch ist die Quantität weder für Vokale noch Konsonanten distinktiv.

Dass die Quantität in einer Sprache nicht distinktiv ist, heißt nicht, dass es in dieser Sprache keine Unterschiede in der Länge von Vokalen gibt. So gibt es im Französischen beispielsweise so etwas wie "naturlange" und "naturkurze" Vokale, aber keine Minimalpaare, die sich nur durch die Quantität unterscheiden. Ist also eine Vokalqualität kurz, so ist sie es stets; für lange Vokale gilt Gleiches. Der in älteren Lehrbüchern noch geforderte Quantitätsunterschied etwa zwischen *maître* /mɛ:tʁ/ (Meister) und *mettre* /mɛtʁ/ (stellen, setzen), der ein Minimalpaar konstituiert und damit die Quantität distinktiv gemacht hätte, existiert in der modernen französischen Aussprache nicht mehr. Im Italienischen gibt es ebenfalls lange und kurze Vokale; hier wird die Vokalquantität jedoch dazu verwendet, die Betonung zu markieren: alle unbetonten Vokale sind kurz, und die betonten erfahren – zumindest in offenen Silben – eine erhebliche Längung. Entsprechendes gilt für das Russische. Im Englischen wird die Vokaldauer u.a. dazu verwendet, die Stimmhaftigkeit eines folgenden auslautenden Plosivs (das Englische kennt keine Auslautverhärtung!) zu markieren; der Vokal [æ] in *had* /hæd/ ist wesentlich länger als in *hat* /hæt/.

Mora. Die Mora ist eine virtuelle Einheit für die Dauer von Segmenten. Mit ihr wird versucht, Quantitäten verschiedener Laute zusammenzufassen und somit eine Aussage über die Dauer größerer Segmente zu treffen. Insbesondere wird sie silbenbezogen eingesetzt (so z.B. im Japanischen).

Im Lateinischen beispielsweise unterscheiden wir zwischen *naturlangen* und *positionslangen* Silben. Naturlang ist eine Silbe dann, wenn sie einen langen Vokal enthält (z.B. zweitletzte Silbe in *legatus* /le:ga:tus/), positionslang, wenn dem Vokal (hier gleichgültig, ob lang oder kurz) noch ein Konsonant folgt (z.B. zweitletzte Silbe in *concursum* /konkursus/). Konsonanten vor dem Vokal bleiben bei dieser Betrachtung unberücksichtigt. Wir können nun bei der Zuweisung der Morae wie folgt vorgehen:

- Kurzer Vokal: 1 Mora;
- Langer Vokal: 2 Morae;
- jeder Konsonant: 1 zusätzliche Mora.

Damit haben natur- und positionslange jeweils (mindestens) 2 Morae zu verzeichnen; derartige Silben werden auch als *schwere* Silben bezeichnet, im Gegensatz zu den *leichten* Silben, die hier nur 1 Mora umfassen.

Das lateinische Versmaß ist – im Unterschied zum deutschen – ein quantifizierendes. Es kann mit dem Begriff der Mora erklärt werden. Während sich bei deutschen Versen betonte und unbetonte Silben nach den Regeln des Versmaßes abwechseln, sind es bei lateinischen Versen entsprechend den Regeln schwere und leichte Silben. Für einen deutschen Sprecher sind lateinische Verse entsprechend schwierig beim Vortrag, wenn dieser richtig, d.h., der lateinischen Sprache angemessen gestaltet werden soll.

4.2.2 Intensität

Auf der segmentalen Ebene äußert sich die Intensität durch die Wortbetonung. In einem zwei- oder mehrsilbigen Wort können grundsätzlich alle Silben gleichstark realisiert werden. In der Regel sind sie jedoch unterschiedlich gewichtet, wobei meist eine Silbe besonders stark hervortritt.

Existiert eine betonte Silbe, so wird sie stets sorgfältiger artikuliert als unbetonte Silben. Sie wird darüber hinaus durch suprasegmentale Merkmale hervorgehoben. Dies sind durchweg Längung sowie eine Tonbewegung (melodischer Akzent); die Intensität (Lautstärke) als akustischer Parameter geht auch mit, spielt aber in der Regel eine untergeordnete Rolle. Welcher Parameter hier primär zum Zuge kommt, hängt von der jeweiligen Sprache ab, insbesondere auch davon, welche Rolle die supra-segmentalen Parameter in dieser Sprache auf der segmentalen Ebene spielen. So nutzen z.B. Sprachen wie das Italienische oder das Russische den (durch die Nichtverwendung auf der segmentalen Ebene gewonnenen) Freiheitsgrad, die Quantität zur akustischen Markierung der Betonung zu verwenden. (Andere Sprachen, wie das Spanische, hätten diesen Freiheitsgrad auch, nutzen ihn aber nicht.) Das Deutsche könnte sich eine vergleichsweise starke Längung betonter Silben nicht leisten, ohne mit der distinktiven Vokalquantität in Konflikt zu geraten; dort wird hauptsächlich ein melodischer Akzent (Anhebung der Grundfrequenz) verwendet.

Wie in den vorhergehenden Fällen gehen wir auch hier von der Minimalpaaranalyse aus:

Auf segmentaler Ebene ist die Wortbetonung dann distinktiv, wenn es bedeutungsverschiedene Minimalpaare gibt, die sich nur durch die Wortbetonung unterscheiden.

Die Sprachen verhalten sich hier wieder durchaus verschieden.

- Sprachen wie Finnisch, Tschechisch, Ungarisch, Türkisch oder Französisch haben einen *gebundenen* Wortakzent; dieser liegt immer auf der gleichen Silbe und ist daher genau vorhersagbar (Finnisch, Tschechisch, Ungarisch: erste Silbe; Französisch: leichter Akzent auf der letzten Silbe; Türkisch: Akzent auf der letzten Silbe mit Ausnahme einiger weniger nicht betonbarer grammatischer Endungen).
- Das Lateinische hat einen variablen, d.h., *freien* Akzent; trotzdem ist dieser vorhersagbar und damit phonologisch nicht distinktiv. Nur ist hier die Regel (die von drei Silben aufwärts greift) komplizierter. Ein mehrsilbiges Wort wird auf der zweitletzten Silbe betont, wenn diese mehr als eine Mora umfasst. Umfasst die zweitletzte Silbe nur eine Mora, so fällt die Betonung auf die drittletzte Silbe: *habere* /ha:'be:re/ haben; *currere* /'kurre:re/ laufen; *legatus* /le:'ga:tus/ Gesandter; *conkursus* /kon'kursus/ Wettlauf.
- Im Deutschen ist der Wortakzent marginal distinktiv. Hier betrifft das Merkmal einige wenige Lehnwörter, z.B.
August /'äugust/ (Vorname) – *August* /äu'gust/ (Monat),
 daneben aber im größeren Umfang zusammengesetzte Verben wie
 ein Hindernis *umfahren* /'umfa:rən/ (es in eine liegende Position befördern) –
 ein Hindernis *umfahren* /um'fa:rən/ (einen Weg darum herum einschlagen)
 oder Minimalpaare wie *'übersetzen* (Wasser) – *über'setzen* (Sprache). Immerhin folgt hier unmittelbar aus der Wortbetonung, ob das Präfix in der Präteritumsform des Verbs abgetrennt wird oder nicht.
- Im Englischen und in einigen anderen Sprachen ist der Wortakzent syntaktisch distinktiv, d.h., er dient dazu, Wortarten oder Flexionsformen zu unterscheiden. Im Englischen betrifft dies Substantiv und Verb oder Adjektiv vom gleichen Stamm:
 the 'record (Substantiv) – to re'cord (Verb)
 the 'adult (Substantiv) – ad'ult (Adjektiv).

Dass mit dem Betonungsunterschied auch ein Ausspracheunterschied einhergeht, der die Vokalqualität betrifft, ist hier wiederum sekundär, da es auch Minimalpaare dieser Art gibt, die sich nur durch die Betonung unterscheiden. – Vergleichbares gilt im Katalanischen:

Andorra /an'dɔra/ (Substantiv) – *Andorrà* /andɔr'a/ (Adjektiv)
Julia /'dʒulia/ (weiblich) – *Julià* /dʒuli'a/ (männlich)

Die Konstellation ist hier leicht einzusehen, wenn man weiß, dass das Katalanische die lateinische Adjektivendung (in diesem Fall *-anus* /'a:nus/) zu *-à* verkürzt, die ursprüngliche Betonung aber beibehalten hat, so wie hier bei *Julià* (aus *Iulianus*).

- Im Spanischen existiert ebenso wie in einigen anderen romanischen Sprachen (außer Französisch) ein distinktiver Wortakzent im Verbalparadigma, ebenfalls mit Trennung der Wortklassen:
término /'termino/ Termin –
termino /ter'mino/ ich beende –
terminó /termin'o/ er/sie hat beendet.
- Völlig frei ist der Wortakzent im Russischen; dort stellt er eine der Hauptschwierigkeiten beim Erlernen der Sprache dar, insbesondere da er in der Orthographie nicht erscheint. Er ist auf der Ebene verschiedener Wortstämme ebenso distinktiv wie zur Unterscheidung verschiedener grammatischer Formen des gleichen Lexems (in den folgenden Beispielen ist der Wortakzent in der Transliteration auf orthographischer Ebene gekennzeichnet):
sláva /'slava/ Ehre – *slová* /sla'va/ Wörter
párnaya paarweise, paarig – *parnáya* frisch, warm (jeweils fem. sing.)
 von *ruká* die Hand: *rukí* (Genitiv Singular) – *rúki* (Nominativ Plural).

4.2.3 Tonbewegungen

Wie in den vorigen Abschnitten wollen wir auch hier nur die Fälle betrachten, wo der Verlauf der Grundfrequenz allein zu einem Bedeutungsunterschied führt. Die Hervorhebung der betonten Silbe durch Tonhöhenveränderung gehört nicht dazu; sie ist die akustische Realisierung des phonologischen Parameters *Intensität*.

Ein selbständiger phonologischer Parameter *Tonbewegungen* liegt dann vor, wenn die Distinktivität nur im Verlauf der Grundfrequenz liegt und durch kein anderes Merkmal ersetzt werden kann. Im Gegensatz zur

Musik haben wir es allerdings bei der Sprache stets mit relativen Bewegungen zu tun (relative Tonhöhen, relative Intervalle).

Bei den Sprachen, bei denen Tonbewegungen auf segmentaler Ebene eine distinktive Rolle spielen, unterscheiden wir zwischen *Tonsprachen* im engeren Sinn und *Tonakzentsprachen*.

Tonsprachen im engeren Sinn. Sprachenfamilien mit Tonsprachen finden sich vorwiegend in Ost- und Südostasien, in Westafrika und in Mittelamerika. Klassisches Beispiel im ostasiatischen Raum ist das Chinesische; jedoch sind auch Vietnamesisch und Thailändisch Tonsprachen.

Im Chinesischen unterscheiden wir vier Töne:

Ton 1 – konstant	$\overline{[ma]}$ Mutter
Ton 2 – steigend	$\underset{/}{[ma]}$ Hanf
Ton 3 – fallend-steigend	$\underset{\vee}{[ma]}$ Pferd
Ton 4 – fallend	$\underset{\backslash}{[ma]}$ schimpfen

Entsprechend der Definition des Phonems werden diese Töne, da sie phonologisch distinktiv sind, als *Toneme* bezeichnet. Jede Silbe des Chinesischen besitzt eines dieser 4 Toneme. Wir sprechen hier von *Konturtonemen*, da sich die Toneme akustisch durch die Bewegung (fallend, steigend usw.) unterscheiden.



Toneme, die sich bei gleichem zeitlichem Verlauf in der Tonhöhe unterscheiden, werden *Registertoneme* genannt. Sie treten in einigen westafrikanischen Sprachen und im Thailändischen auf.

Im allgemeinen ist die Zahl der verschiedenen Toneme in einer Sprache gering. Manche westafrikanischen Sprachen haben nur zwei. Sprachen mit mehr als 5 Tonemen sind selten.

Tonakzentsprachen. Dies sind Tonsprachen im gemäßigten Sinn. Sie haben einen besonderen Tonakzent, der jedoch stets ein ganzes Wort erfasst.

Tonakzentsprachen sind beispielsweise Schwedisch oder Norwegisch. Im Schwedischen existieren zwei Wortakzente, der *Akzent I*, der in seinem eingipfligen Grundfrequenzverlauf dem entspricht, was wir aus dem

Deutschen kennen, und den *Akzent II* mit einem zweigipfligen Grundfrequenzverlauf.

			
• <i>anden</i> (I)	Ente –	<i>anden</i> (II)	Geist
• <i>axel</i> (I)	Achsel –	<i>axel</i> (II)	Achse

Das Japanische kennt einen Tonakzent in Form einer Registertonhöhe:

—	—
• <i>hana</i> Nase –	<i>hana</i> Blüte

Hier ist anzumerken, dass zahlreiche mehrsilbige Wörter im Japanischen überhaupt keinen Wortakzent tragen, oder dass der Akzent, wenn er auftritt, mehrere aufeinanderfolgende Silben erfassen kann. (Ein postvokalisches /n/ zählt im Japanischen als eigene Silbe bzw. Mora.)

4.3 Linguistische und akustische Parameter der Prosodie

4.3.1 Die prosodische (phonologische) Hierarchie

Wie uns bereits das Beispiel mit der Entwicklung des Wortakzents vom Lateinischen hin zu anderen romanischen Sprachen gezeigt hat, können die segmentale und die suprasegmentale Ebene bei der Weiterentwicklung einer Sprache getrennte Wege gehen. In obigem Beispiel werden bei der Entwicklung von Latein zu Katalanisch Wortendungen abgeworfen; der lateinische Wortakzent jedoch, der im klassischen Latein strengen Regeln unterliegt und vorhersagbar ist, bleibt bei der Entwicklung zum Katalanischen liegen und verändert seine Stellung nicht; dadurch wird das Akzentsystem des Katalanischen ein anderes, und der Wortakzent wird nicht mehr auf so direkte Weise vorhersagbar wie im Lateinischen.

Dementsprechend gibt es gute Gründe, bei formal-phonologischen Beschreibungen, die die Prosodie mit einschließen, von einem streng linearen, einsträngigen Formalismus abzuweichen und statt dessen nichtlineare (oder besser *multilineare*) Phonologien anzuwenden, die für diesen Zweck entwickelt worden sind, wie die *autosegmentale* Phonologie oder die *metrische* Phonologie (vgl. Spencer, 1996). An dieser Stelle soll das

Problem nur erwähnt werden; im Detail diskutiert wird es an anderer Stelle (Vorlesungen PHONOLOGIE und PROSODIE).

Eine Beschreibungsebene oberhalb der segmentalen Ebene reicht jedoch für die Beschreibung prosodischer Phänomene nicht aus. Suprasegmentale Merkmale treten auf allen Ebenen auf: Silbe, Wort, Phrase, Satz und gesamte Äußerung (die aus mehreren Sätzen bestehen kann). Aus diesem Grund übernehmen wir aus den nichtlinearen phonologischen Beschreibungsschemata das Prinzip der *phonologischen* bzw. *prosodischen Hierarchie* (Bild 4.2). Auch hier sind die Begriffe nicht einheitlich definiert; beispielsweise wird *phonologisches Wort* und *prosodisches Wort* häufig synonym verwendet. Phonologische Hierarchie und prosodische Hierarchie dienen jedoch verschiedenen Zwecken. Während die phonologische Hierarchie vor allem den Geltungsbereich phonologischer Regeln festlegt, definiert die prosodische Hierarchie die Struktur und den Rhythmus einer Äußerung und sorgt für die Zuweisung von Betonung und Phrasierung auch oberhalb der Wortebene.

Phonologische Äußerung	U	Prosodische Äußerung
	IP	(volle) Intonationsphrase
	ip	Intonationsphrase
Phonologische Phrase	Φ	
Phonologisches Wort	ω	Prosodisches Wort
	Σ	Fuß
Silbe	σ	Silbe
Segment	X	(Segment)

Bild 4.2. Phonologische / prosodische Hierarchie

- Die erste Ebene oberhalb des Segments ist bei beiden Hierarchien die der *Silbe*.
- Ein *Fuß* (*foot*) umfasst eine betonte Silbe und die zugehörigen unbetonten Silben. Der Fuß ist in der Literaturwissenschaft wichtig für die Definition von Versmaßen (*Iambus*, *Trochaeus*, *Dactylus* etc.).

- Ein *prosodisches Wort* umfasst einen oder mehrere Füße.
- Eine *Intonationsphrase* umfasst eines oder mehrere prosodische Wörter. Sie ist dadurch gekennzeichnet, dass sie (in einer realisierten Äußerung) mindestens einen Phrasenakzent und am Ende eine Phrasengrenze besitzt. Die Phrasengrenze ist in der Regel durch eine Pause und/oder einen Einschnitt im Grundfrequenzverlauf (der *Intonationskontur*) markiert. In manchen Beschreibungssystemen ist die Intonationsphrase unterteilt. Ein Satz (der als Einheit in dieser Hierarchie nicht vorkommt), umfasst mindestens eine *volle Intonationsphrase* (IP); diese schließt den Satz ab und markiert auch den Satzmodus (z.B. Aussage oder Frage). Eine längere IP, die mehrere Phrasenakzente hat, kann aber in sich nochmal unterteilt sein und damit mehrere Intonationsphrasen (*ip*) umfassen.
- Die phonologische oder prosodische *Äußerung* ist die größte Einheit, innerhalb derer prosodische Abhängigkeiten definiert und nachgewiesen werden. Sie umfasst eine oder mehrere Intonationsphrasen (IP). Dies kann sein: in Dialogen ein *Turn*, also eine zusammenhängende Äußerung des gleichen Sprechers von dem Moment, in dem er zu sprechen beginnt, bis zu dem Moment, in dem er die Rolle des Sprechers wieder an einen Dialogpartner abgibt. Dies kann weiterhin bei gelesenen Texten ein Absatz oder eine andere, einen oder mehrere Sätze umfassende zusammenhängende Einheit sein.

4.3.2 Interaktion akustischer und linguistischer Parameter

Die drei linguistischen Parameter in der Prosodie (siehe Abschnitt 4.1 sowie Lehiste, 1970) sind definiert als *Quantität* (rhythmische und zeitliche Phänomene), *Intensität* (Betonung und Akzentuierung) sowie *Intonation* (melodische Phänomene). Dem gegenüber stehen drei akustische Parameter:

- *Dauer* als das benötigte Zeitintervall für die jeweilige Einheit;
- *Energie*, *Leistung*, *Amplitude*, *Lautstärke* als akustischen Maß für die *Schallintensität* einer Einheit. Man beachte, dass der Begriff *Intensität* hier doppelt verwendet werden kann, d.h., sowohl auf linguistischer als auch auf akustischer Ebene in Form der *Schallintensität*. Zur Vermeidung

derung von Verwechslungen wird in diesem Kapitel für den akustischen Parameter der Terminus *Energie* verwendet.

- *Sprachgrundfrequenz* F_0 als Trägerin der Sprechmelodie.

Diese drei akustischen Parameter sind jedoch nicht unabhängig voneinander (Bild 4.3). Es besteht eine gute Korrespondenz zwischen Quantität und Dauer sowie zwischen Intonation und Sprachgrundfrequenz. Der akustische Parameter Energie spielt jedoch eine untergeordnete Rolle: Betonung und Akzentuierung wirken vor allem über die Dauer und über die Sprachgrundfrequenz. Daneben besteht ein starker Einfluss der Intensität auf den Artikulationsaufwand: Betonte Silben werden stets sorgfältiger und präziser artikuliert als unbetonte; daher verlangen sie auch mehr Zeit und werden dementsprechend verlängert.

Wie kein anderer akustischer Sprachsignalparameter ist die Sprachgrundfrequenz F_0 funktional belastet (Bild 4.4). Mikroprosodie, Wortakzent, Satzakzent, sprecherspezifische und emotionale Randbedingungen – alle greifen sie auf diesen Parameter durch. Während die Sprachgrund-

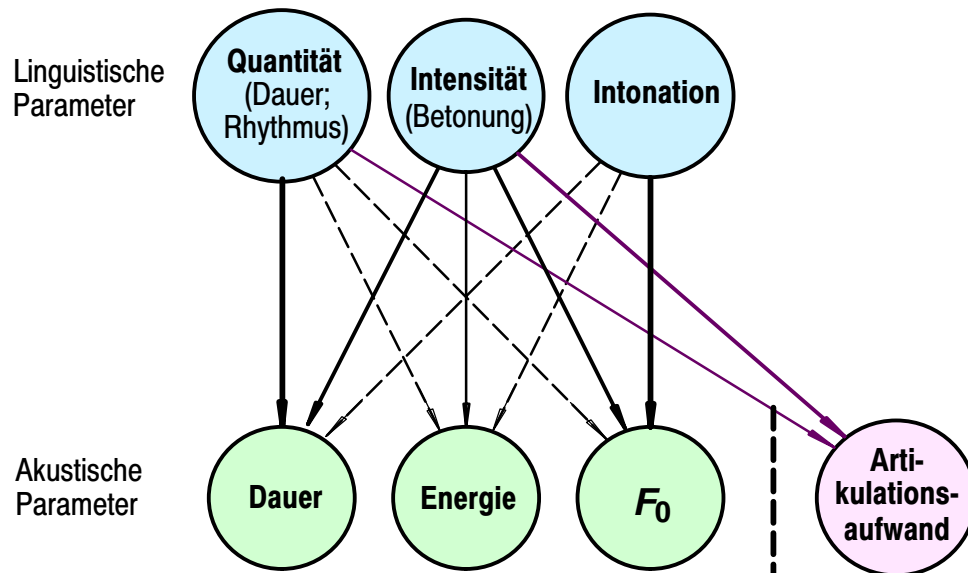


Bild 4.3. Interaktion zwischen den linguistischen und akustischen Parametern in der Prosodie

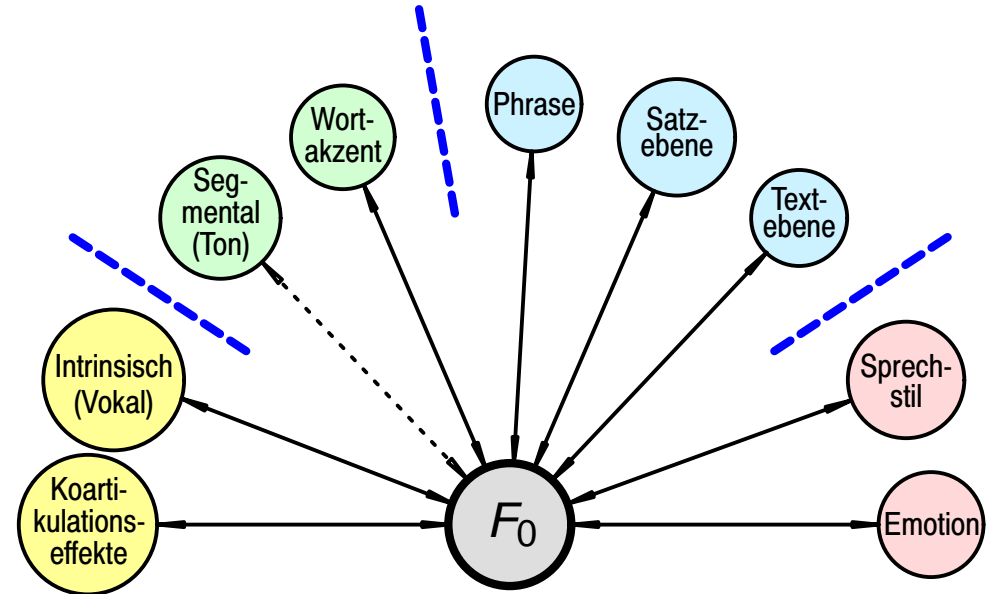


Bild 4.4. Die funktionale Belastung der Sprachgrundfrequenz

frequenz auf segmentaler Ebene so gut wie keine Rolle spielt, ist ihr Beitrag zur Prosodie damit entscheidend.

- Auf artikulatorischer Ebene treten zwei Erscheinungen auf, die (nicht unbedingt zutreffend) auch als *Mikroprosodie* bezeichnet werden. Als *intrinsische Grundfrequenz* (*intrinsic pitch*) wird die Beobachtung bezeichnet, dass die Grundfrequenz bei Hochzungenvokalen um 5-10% höher liegt als bei Tiefzungenvokalen. Weiterhin ergeben sich an Stimmhaft-Stimmlos-Übergängen kurzzeitige Schwankungen (wenige Perioden) vor dem Abreißen der Stimme und beim Wiederanschwingen, die als *Koartikulationseffekte* angesehen werden können (siehe Vorlesung ARTIKULATORISCHE UND AKUSTISCHE PHONETIK). Diese Phänomene sind artikulatorisch bedingt und unterliegen nicht dem willentlichen Einfluss des Sprechers; sie sind demzufolge auch nicht Träger linguistischer Information, sind aber deutlich zu hören und auch in Intonationskonturen deutlich zu sehen.
- Auf segmentaler Ebene spielt die Grundfrequenz eine Rolle in Ton-sprachen und Tonakzentsprachen (siehe Abschnitt 4.2).

- Der Wortakzent spielt segmental eine Rolle, soweit er distinktiv ist. Gleichgültig, ob der Wortakzent auf Wortebene bedeutungsunterscheidend ist oder nicht, er beeinflusst die Grundfrequenz auf allen Ebenen.
- Sprechstil und Emotion gehören als paralinguistische und extralinguistische Größen nach unserer Definition nicht zur Prosodie, haben aber großen Einfluss auf die akustisch-prosodischen Parameter und hier vor allem auch auf die Sprachgrundfrequenz.

4.3.3 Prosodische Parameter auf den verschiedenen Ebenen der Hierarchie

Prominenz. Unter Prominenz sei die linguistisch und perzeptiv relevante Hervorhebung eines Abschnittes des Sprachsignals verstanden (Möbius, 1993). Prominenz ist in erster Linie ein perzeptiver Parameter und graduell, nicht kategorial. Sie kann auf allen Ebenen der prosodischen Hierarchie definiert und bestimmt werden. Die Zuordnung zu linguistischen Kategorien ist dann ein zweiter Schritt.

„*Mikroprosodie*“ (s.o.) bezeichnet (ko-)artikulatorisch bedingte Änderungen der Sprachgrundfrequenz. Sie hat nur perzeptive und akustische, aber keine linguistische Relevanz und wird daher nicht zur Prosodie im strengen Sinn gerechnet. Sie wird vom Sprecher nicht bewusst gesteuert. Bei der prosodischen Analyse sprachlicher Äußerungen wirkt sie eher störend; bei der Sprachsynthese muss sie berücksichtigt werden, da sie zur Natürlichkeit synthetischer Sprache beisteuert.

Prosodie auf Wortebene. Betonung und Akzentuierung sind die wichtigsten Korrelate der Prosodie auf der Wortebene. Durch sie wird einzelnen Silben eine erhöhte Prominenz zugewiesen. Die Termini Betonung und Akzent (engl.: *stress* und *accent*) werden unterschiedlich definiert, leider auch oft synonym und widersprüchlich. In Anlehnung an MÖBIUS (1993) bezeichnen wir prominente Silben in einer Äußerung als betont; von einer *akzentuierten Silbe* (oder Akzentsilbe; *accented syllable*) soll gesprochen werden, wenn eine betonte Silbe auch tonal, also durch eine Grund-

frequenzbewegung markiert ist. Die regelhaft betonte Silbe eines Wortes trägt den *Wortakzent* bzw. *lexikalischen Akzent*.

Akustisch-phonetisch kann die Hervorhebung einer Silbe auf unterschiedliche Weise realisiert werden. Meistens sind alle prosodisch relevanten akustischen Größen betroffen: Dauer, Grundfrequenz, Energie und spektrale Eigenschaften (entsprechend erhöhtem artikulatorischen Aufwand für prominente Silben); welcher Parameter primär beeinflusst wird, ist sprachenspezifisch. TERNES (1987) unterscheidet zwischen melodischem Akzent (beeinflusst vornehmlich Grundfrequenz) und dynamischem Akzent (beeinflusst vornehmlich Energie bzw. Dauer).

Die phonologische Analyse gelangt meist zu einer kleinen Anzahl von Betonungsstufen, z.B. *unbetont – schwach betont – betont – stark betont*. Für den Wortakzent haben sich drei Stufen eingebürgert: *Silbe mit Hauptakzent – Silbe mit Nebentakzent – unbetonte Silbe*. Dem überlagern sich (aus höheren Ebenen) *Kontrastakzent* und *Emphase* (s.u.)

Kontrastakzent und Emphase. Diese in der Literatur oft synonym behandelten Begriffe sind nach WAGNER (2002) zu trennen.

- Funktion eines *Kontrastakzents* ist es, aufzuzeigen, dass eine früher gegebene Information korrigiert bzw. überschrieben werden muss. Der Kontrastakzent führt dazu, dass ein eng begrenzter Äußerungsteil (meist nur eine Silbe) sehr prominent wird.
- Emphase soll anzeigen, dass dem Sprecher ein geäußertes Sachverhalt besonders wichtig ist. Emphase führt zu erhöhter Prominenz längerer Abschnitte einer Äußerung.

Generell gilt jedoch: Silben mit Kontrastakzent oder Emphase sind prominenter als alle anderen Silben.

Prosodie oberhalb der Wortebene. Oberhalb der Wortebene dienen (linguistische) prosodische Kriterien vor allem der *Strukturierung* von Äußerungen.

- *Phrasierung:* Strukturiert eine Äußerung und gliedert sie in Phrasen, also kleine, syntaktisch oder semantisch zusammenhängende Teile der Äußerung (unterhalb oder höchstens bis zur Satzebene). Akustisches Korrelat: Pause und Grundfrequenzverlauf.

- *Satzakzent*: Kennzeichnet den als besonders wichtig empfundenen Teil eines Satzes und verleiht ihm eine höhere Prominenz.
- *Satzmodus*: Unterscheidet vor allem Aussage- und Fragesätze, bei Phrasen auch nicht-finale (progrediente) Stellung. Akustisches Korrelat: Grundfrequenzverlauf.

Phrasierung ist die Gliederung einer Äußerung in zwei oder mehr Abschnitte, die jeweils durch eine eigene Intonationskontur gekennzeichnet und häufig durch kurze Sprechpausen (oder sogar durch Atempausen) getrennt sind. An Phrasengrenzen tritt in zahlreichen Sprachen (Deutsch, Englisch) eine Dehnung der der Phrasengrenze unmittelbar vorausgehenden Silben auf (finale Längung; final lengthening). Am stärksten betroffen ist die letzte Silbe vor der Grenze; nach vorne hin nimmt der Effekt schnell ab. Finale Längung ist bei Phrasengrenzen im Äußerungsinneren ausgeprägter als äußerungsfinal (MÖBIUS, 1993).

Eine Phrasengrenze unterbricht den typischen GF-Verlauf der Akzentgruppe, die der Phrasengrenze vorausgeht. (Eine Akzentgruppe, die in der prosodischen Hierarchieebene etwa einem prosodischen Wort entspricht, bestehe hierbei aus einer betonten Silbe und allen ihr folgenden unbetonten Silben bis hin zur nächsten betonten Silbe).

Der Grundfrequenzverlauf an einer Phrasengrenze hängt von der Stellung der Phrase in der Äußerung und vom Satzmodus ab.

- *Äußerungsfinaler GF-Fall in Aussagesätzen*: Abfall der GF in der finalen Akzentgruppe (Deutsch: bis zu 10 Halbtöne).
- *Frageintonation*: Anstieg der GF in der finalen Akzentgruppe eines Fragesatzes – jedoch in der Regel nur in Echofragen, also Fragesätzen, die die Struktur eines Aussagesatzes ohne Fragewort haben.
- *Progrediente Intonation*: (Schwacher) Anstieg der GF an Phrasengrenzen im Inneren einer Äußerung (Regelfall bei einer “kleinen” Intonationsphrase ip).

Deklination und Downstepping. Ein besonders in gelesenen Äußerungen beobachtetes Phänomen ist der allmähliche Fall der GF vom Beginn der Äußerung hin zu ihrem Ende. Dieses Phänomen wird – allerdings nicht unumstritten – als *Deklination* bezeichnet (Bild 4.5a). Einige Schu-

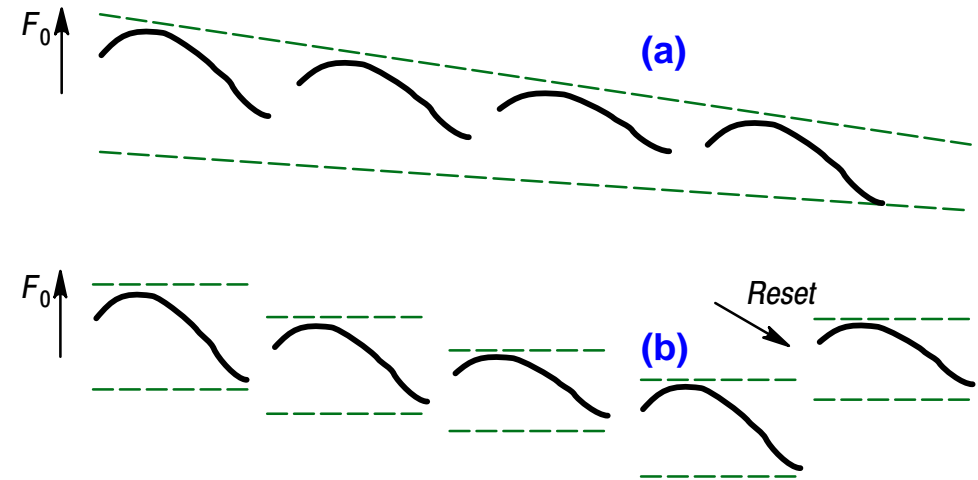


Bild 4.5a,b. Deklination (a) versus *Downstepping* (b): prinzipieller Verlauf der Grundfrequenzkontur. Eingezeichnet ist jeweils eine obere (*topline*) und untere (*baseline*) Begrenzungslinie für den Wert der Grundfrequenz.

len der Intonationsforschung lehnen allerdings die Deklination ab und fordern statt dessen ein *Downstepping* auf struktureller, nicht auf zeitlicher Basis (Bild 4.5b; vgl. MÖBIUS, 1993:25). Spontansprachliche Äußerungen zeigen oft weder Deklination noch Downstep. Verträglich mit dem Prinzip des *Downstepping* ist das Zurücksetzen einer Intonationskontur an größeren Phrasengrenzen (*Resetting*).

Satzakzent und Fokus. Die im Verlauf einer Äußerung vom Sprecher als besonders wichtig angesehene Information erhält den *Satzakzent*. Dieser Teil des Satzes steht im *Fokus*. Der Fokus hat vornehmlich mit Semantik und Pragmatik auf Äußerungsebene zu tun und steht in engem Zusammenhang mit der Frage, welcher Teil der Äußerung für den Hörer “neu” und welcher bereits “bekannt” ist (“*given*” vs. “*new*”). ALTMANN et al. (1989) definieren den Fokus wie folgt:

“Fokus und Fokussierung sind semantische Begriffe, die die informationelle Gliederung von Äußerungen betreffen: Vorausgesetzter Hintergrundinformation steht fokussierte Information gegenüber, die aus einer kontextuell festgelegten Menge von Alternativen ausgewählt wird. Gekennzeichnet wird der Fokus durch die akzentuelle Hervorhebung eines Ausdrucks,

des Fokusexponenten. Dieser Ausdruck trägt aber auch einen Teil der [...] intonatorischen Satzmodusmarkierung, so daß jede Untersuchung die gegenseitige Beeinflussung der intonatorischen Kennzeichnung von Satzmodus und Fokus berücksichtigen muß.” (ALTMANN et al., 1989:1)

Sind alle Teile der Äußerung gleich wichtig, so sprechen wir von einem *weiten Fokus*. In diesem Fall befindet sich der Satzakzent an vorhersagbarer Position, im Deutschen beispielsweise auf dem Hauptakzent der letzten Phrase der Äußerung.

Sind einzelne Teile einer Äußerung besonders wichtig (z.B. weil neu oder vom vorherigen Diskurs abweichend), so können sie prosodisch durch einen *engen Fokus* markiert werden. Ein enger Fokus lässt sich durch syntaktische Mittel und die Wortwahl, aber insbesondere auch durch die Prosodie ausdrücken. Wird ein enger Fokus prosodisch markiert, so geschieht dies durch den Satzakzent an beliebiger Stelle der Äußerung oder sogar durch Kontrastakzente und/oder Emphase.

Ein enger Fokus ist in der Schriftform der Sprache nicht repräsentiert und ist deshalb ein typisches Stilmittel gesprochener Sprache, weil die Markierung durch prosodische Mittel in der Regel ökonomischer ist als durch Wortwahl oder Wortstellung im Satz.

Literatur

Einführende Darstellungen

Kohler, Klaus J. (1977, ²1995): *Einführung in die Phonetik des Deutschen* (Erich Schmidt, Berlin)

Spencer, Andrew (1996): *Phonology* (Blackwell, Oxford)

Ternes, Elmar (1987): *Einführung in die Phonologie* (Wissenschaftliche Buchges., Darmstadt)

Zu Prosodie und suprasegmentalen Merkmalen

Altmann, Hans (Hrsg.) (1988): *Intonationsforschungen* (Niemeyer, Tübingen)

Hirst, Daniel / Di Cristo, Albert (eds.) (1998): *Intonation systems. A survey of twenty languages* (Cambridge Univ. Press, UK)

Lehiste, Ilse (1970): *Suprasegmentals* (MIT Press, Cambridge, MA)

Möbius, Bernd (1993): *Ein quantitatives Modell der deutschen Intonation. Analyse und Synthese von Grundfrequenzkonturen* (Niemeyer, Tübingen)

Rossi, Mario / Di Cristo, Albert / Hirst, Daniel / Martin, Philippe / Nishinuma, Y. (1981): *L'intonation de l'acoustique à la sémantique* (Klingsieck, Paris)

Weitere Quellenangaben

Bußmann, Hadumod (1990): *Lexikon der Sprachwissenschaft* (Stuttgart)

Chomsky, Noam / Halle, Morris (1968): *The sound pattern of English* (New York)

Wagner, Petra (2002): *Vorhersage und Wahrnehmung deutscher Betonungsmuster* (Diss., Univ. Bonn)