
Deutsche Intonation und GToBI

Martine Grice und Stefan Baumann

ABSTRACT

In this paper we provide an overview of work carried out on the intonation of Standard German, both in auditory phonetic studies and in the instrumentally-based phonological accounts within the autosegmental-metrical framework. We examine how far the different accounts shed light on controversial issues such as leading tones, levels of phrasing, and phrase accents, and propose a surface-oriented annotation framework, GToBI, which aims to capture all empirically observed distinctive intonation patterns. For illustration purposes, the contours which are reported to occur most commonly are given in schematic form, along with their GToBI transcription and examples of their usage.

1 Einleitung

Dem Englischen, Niederländischen und Deutschen wird eine große Ähnlichkeit auf dem Gebiet der Prosodie und Intonation nachgesagt, ein Umstand, der auf ihre gemeinsame Zugehörigkeit zur Westgermanischen Sprach(en)familie zurückzuführen sein mag. Alle drei Sprachen gelten als akzentzählend, sind rhythmisch in linksköpfige Füße bzw. Takte gegliedert, heben Information mit Hilfe einer Reihe verschiedener Akzenttöne¹ hervor und verwenden Grenztöne, um Äußerungen in Phrasen zu unterteilen. In Strukturen mit weitem Fokus fällt der nukleare Akzentton in allen drei Sprachen zumeist auf das letzte Argument der Intonationsphrase.

Im Rahmen der Autosegmental-Metrischen (AM) Phonologie sind im wesentlichen zwei Beschreibungsmodelle für die Intonation des Deutschen entwickelt worden. Auf der einen Seite handelt es sich hierbei um die Ansätze von Féry (1993) und Grabe (1998), die sich an Gussenhovens (1984) Analyse des Niederländischen anlehnen, auf der anderen Seite um GToBI (*German Tones*

¹ *Akzentton* und *Tonakzent* werden in diesem Aufsatz synonym verwendet. Beide Begriffe entsprechen dem englischen *Pitch Accent*.

and *Break Indices*), einem gemeinschaftlich von Martine Grice, Matthias Reyelt, Ralf Benzmüller, Anton Batliner und Jörg Mayer (Grice et al 1996; Reyelt et al 1996) erarbeiteten System, das eng verwandt ist mit dem ursprünglichen englischen ToBI (im folgenden als *Mainstream American English ToBI* (MAE_ToBI) bezeichnet: Beckman & Hirschberg 1994; Beckman & Ayers-Elam 1997, Beckman, Hirschberg & Shattuck-Hufnagel (im Druck)), welches seinerseits auf der Analyse des Englischen von Pierrehumbert (1980) und Beckman & Pierrehumbert (1986) basiert.

Die hier relevanten Unterschiede zwischen den Ansätzen bestehen in folgenden zwei Punkten: Erstens sind Tonakzente in Gussenhovens Modell stets linksköpfig, analog zur Fußstruktur. Dies bedeutet, daß die Tonbewegung unmittelbar vor der akzentuierten Silbe, mit der der Tonakzent assoziiert ist, nicht als Teil dieses Tonakzents beschrieben wird. Dagegen postulieren GToBI und MAE_ToBI sowohl links- als auch rechtsköpfige Tonakzente. Letztere tragen zum einen der Tonbewegung auf der akzentuierten Silbe, zum anderen der Tonbewegung unmittelbar vor der Akzentsilbe Rechnung. Der Ton zur Linken wird als *Leading*-Ton bezeichnet. Zweitens analysiert Gussenhoven nukleare Konturen als eine Kombination aus einem Tonakzent und dem Grenzton einer Intonationsphrase. GToBI und MAE_ToBI nehmen einen zusätzlichen Ton zwischen Akzent und Grenzton an, der *Phrasenakzent* genannt wird.

Auf den ersten Blick erscheint es, als wären diese Unterschiede darauf zurückzuführen, daß eine Forschergruppe das Deutsche eher dem Englischen anlehnt, die andere eher dem Niederländischen. Da aber Gussenhoven die Existenz von *Leading*-Tönen und Phrasenakzenten auch für das Englische anzweifelt, wird deutlich, daß die erwähnten Unterschiede eher theoretischer als typologischer Natur sind.

Zwei weitere deutsche AM-Ansätze sind die von Wunderlich (1988) und Uhmann (1991), deren Zuordnung zu den genannten Gruppen weniger eindeutig ist. Im Hinblick auf Struktur und Inventar der Akzenttöne und Phrasierungseinheiten sind diese Ansätze mit dem Modell Gussenhovens vergleichbar, wenn sie sich auch nicht explizit darauf berufen.²

Im folgenden Kapitel werden wir zunächst kurz auf die traditionelle Literatur zur deutschen Intonation eingehen, die in der Regel theorie-neutral war und sich auf auditive Eindrücke, z.T. phonetisch detailliert beschrieben, stützte. Danach geben wir einen Überblick über die Autosegmental-Metrische Literatur, die zum Teil auf den Arbeiten Gussenhovens aufbaut, gefolgt von einer ausführlichen Beschreibung GToBIs. Abschließend wird GToBI einem Vergleich mit anderen Autosegmental-Metrischen Modellen zur deutschen Intonation unterzogen sowie unsere Motivation für abweichende Analysen erläutert.

² In bezug auf die verwendeten Darstellungsebenen hingegen ähneln beide Ansätze dem Modell von Pierrehumbert, bei dem eine abstrakt-phonologische Repräsentation ohne Umweg über phonologische Modifikationen oder *Adjustment Rules* (wie bei Gussenhoven) in eine phonetische Oberflächenform übertragen wird.

2 Beschreibungsmodelle deutscher Intonation

Intonation ist traditionell entweder in Form von Konturen, d.h. als Tonbewegungen in bestimmte Richtungen, oder in Form von Ebenen, in die sich der Stimmumfang eines Sprechers einteilen läßt, beschrieben worden. In letzterem Ansatz werden Intonationsmuster als Abfolge verschiedener Tonstufen verstanden. Zunächst werden wir die kontur-basierten Modelle untersuchen, dann einen Blick auf die frühen Ebenen-Modelle werfen, bevor wir uns mit den neueren Autosegmental-Metrischen Ansätzen beschäftigen, die ebenfalls Ebenen-Modelle darstellen.

2.1 Kontur-basierte Modelle

Frühe Beschreibungen deutscher Intonation, wie etwa durch von Essen (1964), Pheby (1975), Kohler (1977) und Fox (1984), gründen sich größtenteils auf auditive Analysen und sind didaktisch ausgerichtet. Sie stellen Intonationsmuster mittels einer detaillierten interlinearen Transkription der Tonhöhe jeder einzelnen Silbe einer Äußerung dar. Alle diese Arbeiten folgen den Ideen der Britischen Schule, z.B. Crystal (1969) und Halliday (1967), die Intonation in Form von dynamischen Tonkonturen beschreiben und die dem Nukleus (Hallidays *Tonic*, Phebys *Tonstelle*) als der prominentesten Silbe der Äußerung besondere Bedeutung beimessen. Für Pheby und Fox beginnt die Tonbewegung, die für die Klassifizierung eines Intonationsmusters entscheidend ist, auf der nuklearen Silbe und erstreckt sich bis zum Ende der Phrase. In der Britischen Schule wird diese Tonbewegung als *Nuclear Tone* bezeichnet.

Bei Kohler beschränkt sich die distinktive Kontur einer Intonationseinheit nicht auf die Akzentsilbe und das ihr nachfolgende Material. Vielmehr kann die entscheidende Tonbewegung *vor* der Akzentsilbe beginnen. Kohler versteht Akzenttöne als Tonhöhengipfel, deren zeitliche Koordination mit dem Text (*Alignment*) nach vorne oder nach hinten variieren kann. In neueren, instrumentalphonetisch unterstützten Arbeiten (z.B. Kohler 1991), differenziert er drei Arten von Gipfeln (früh, mittel, spät) und weist ihnen verschiedene pragmatische Interpretationen zu: Ein früher Gipfel, bei dem das Tonhöhenmaximum auf der prä nuklearen Silbe liegt, markiert selbstverständliche oder bereits gegebene Information. Ein mittlerer Gipfel - mit der größten Tonhöhe etwa in der Mitte der akzentuierten Silbe - markiert neue Information. Ein später Gipfel, der gegen Ende der akzentuierten oder erst in der folgenden Silbe erreicht wird, drückt Emphase bzw. ein größeres Engagement auf seiten des Sprechers aus, als es bei einem mittleren Gipfel der Fall ist (1991:160).

Ein etwas anderer kontur-basierter Ansatz wurde von Selting (1995) vorgeschlagen mit dem Ziel, ein Beschreibungssystem für die Analyse spontaner Dialoge zu entwickeln. Seltings Modell ist größtenteils auditiv, unterstützt durch instrumentale Analysen. Es unterscheidet sich von den Ansätzen der Britischen

Schule darin, daß Intonationskonturen als holistische Einheiten betrachtet werden und der Nukleus keinen Sonderstatus besitzt. Im Gegensatz zur gängigen Auffassung in den frühen auditiven Studien zur deutschen Intonation spiegeln für Selting Konturen nicht die grammatische Struktur von Äußerungen wider, sie signalisieren vielmehr das Zusammenspiel 'als kohäsiv wahrgenommener prosodischer Einheiten' (Selting 1995:39) in der Diskursorganisation, z.B. bei der Konstruktion von Turns. Seltings Ansatz kann als *Overlay Approach* (Ladd 1996) bezeichnet werden, weil sie nicht nur die Form einzelner lokaler Akzenttöne bestimmt, sondern auch Angaben zum globalen Tonhöhenverlauf und zu Veränderungen des Stimmumfangs macht.

Wir werden auf einige dieser Modelle in Kapitel 4 näher eingehen.

2.2 Frühe Ebenen-Modelle

Nicht alle Arbeiten zur Intonation des Deutschen waren kontur-basiert. In den frühen sechziger Jahren beschrieb z.B. Moulton (1962), der Tradition der amerikanischen Strukturalisten Pike (1945) und Trager & Smith (1951) folgend, die deutsche Intonation mit Hilfe von distinktiven Tonebenen (Moulton verwendete drei statt der bis dahin üblichen vier Ebenen; er verzichtete auf die vierte Ebene für emphatische Sprechweise). Ebenso wie Trager & Smith hatte auch Moulton sogenannte 'terminale Konturen' in seinem Inventar, die die Tonbewegung am Ende einer Phrase als steigend, fallend oder gleichbleibend auswiesen. Aus diesem Grund ist Moultons Ansatz eher eine Mischung aus Ebenen und Konturen als ein reines Ebenen-Modell.

Im ersten reinen Ebenen-Modell des Deutschen reduzierten Isačenko & Schädlich (1966) die Anzahl der Ebenen auf zwei. Hierzu haben sie Äußerungen auf einer hohen und einer tiefen Frequenz monotonisiert und Teststimuli mit einem Schritt von der einen Ebene zur anderen, entweder vor der akzentuierten Silbe (praktisch³) oder danach (postiktisch), zusammengestellt. Praktisch fallende Tonhöhe ist äquivalent zu Kohlers frühem Gipfel, postiktisch fallende Tonhöhe entspricht mittleren oder späten Gipfeln. Auf die frühen Gipfel werden wir bei der Diskussion der GToBI-Akzenttöne zurückkommen (Abschnitt 4.1.1).

2.3 Autosegmental-Metrische Modelle

Neuere Ebenen-Modelle sind im theoretischen Rahmen der Autosegmental-Metrischen Phonologie entwickelt worden. Der Terminus 'Autosegmental-Metrisch' (AM) geht auf Ladd (1996) zurück und bezeichnet alle Intonationsmodelle, die der richtungsweisenden Arbeit von Pierrehumbert (1980) folgen. Alle diese Ansätze postulieren mindestens zwei (H (*high*) und L (*low*)), höch-

³ Der Wortakzent wird auch als *Iktus* bezeichnet.

stens drei (H, L und M (*mid*)) Ebenen für die Beschreibung von Intonation. Töne bzw. Tonbewegungen werden als Abfolge von Zielpunkten verstanden, die auf diesen Ebenen liegen. Zum einen dienen Töne dazu, bestimmte Silben hervorzuheben. Diese werden dann als Tonakzente (bzw. Akzenttöne) bezeichnet und sind entweder monotonal (z.B. H*) oder bitonal (z.B. L*+H). Der gesternte Ton ist mit der Akzentsilbe assoziiert (und in der Regel auch zeitlich koordiniert, wengleich sich die Frage des *Alignment* in neueren Untersuchungen als äußerst komplex erwiesen hat (Arvaniti et al 1999)). Ungesternte Töne, die der akzentuierten Silbe vorangehen, nennt man *Leading*-Töne, der Akzentsilbe nachfolgende Töne *Trailing*-Töne. Wie bereits erwähnt, besteht ein Unterschied zwischen Pierrehumberts Modell und GToBI einerseits und den übrigen AM-Modellen für das Deutsche andererseits darin, daß erstere sowohl *Leading*-Töne als auch *Trailing*-Töne in ihrem Inventar haben, während letztere lediglich *Trailing*-Töne erlauben. In Abschnitt 4.1 werden wir ausführlicher auf diesen Aspekt eingehen.

Des weiteren können Töne als initiale oder finale Grenzmarkierungen für intonatorisch relevante Phrasen fungieren. In den unten beschriebenen Modellen werden Töne als perzeptiv-phonetische Kategorien verstanden, die sich akustisch in Veränderungen des Grundfrequenz(F0)-Verlaufs manifestieren. Allerdings ist der konkrete F0-Wert dieser Töne in Akzentfolgen oder in Kombinationen aus Akzenten und Grenzönen nicht immer eindeutig, d.h. die phonetische Umsetzung der phonologisch definierten H und L Töne kann sich je nach Modell sehr unterscheiden. Ferner wird die Höhe der (paradigmatisch ausgewählten) Töne von syntagmatischen Phänomenen wie *Upstep* oder *Downstep* beeinflusst, bei denen die Grundfrequenz in Relation zu einem vorangehenden tonalen Ereignis entweder angehoben oder abgesenkt wird.

Wir werden im folgenden eines der beiden erwähnten, an Gussenhoven angelehnten Modelle (Féry) und die beiden unabhängigeren Modelle (Wunderlich, Uhmann) kurz vorstellen, ihr Inventar an Tonakzenten und Grenzönen auflisten sowie ihre Beschreibungen gängiger Intonationsmuster des Deutschen untersuchen, um sie in Kapitel 4 einem Vergleich mit GToBI zu unterziehen.

Wir beginnen mit der Arbeit von Wunderlich (1988). Wie die meisten deutschen Intonologen konzentriert er sich auf die 'grammatischen' Funktionen von Intonation, vor allem Satzmodusunterscheidung und Fokus-Hintergrundgliederung (vgl. z.B. Altmann, Batliner & Oppenrieder (1989) für eine Vielzahl experimenteller Daten). Wunderlich nimmt nur eine Phrasierungseinheit an, die Intonationsphrase. Sein tonales Inventar besteht aus einzelnen Akzenten, Akzent-Akzent-Sequenzen und Akzent-Grenzton-Sequenzen. Sie sind nachfolgend zusammen mit den Kontexten, in denen sie vorkommen, aufgelistet (soweit vom Autoren angegeben):

H*	Gipfelakzent – Standard-Akzent
H* H L*	Brückenakzent – mehrfache Foki, Kontrast
%H L*	Fallend-Tiefakzent – Ausrufe
L* H%	Tiefakzent-Steigend

L* H (H%)	Echoakzent – Echofragen
H* H	linker Brückenpfeiler – Anfänge von Listen

Wunderlich nimmt für jedes Muster mehrere Verwendungsmöglichkeiten an, wobei er die Funktionen des Brückenakzents genauer untersucht. Diese Akzentfolge gilt als typisch für deutsche Sätze mit mehrfachem Fokus: Nach einem Gipfelakzent (H*) bleibt der Ton auf einem hohen Niveau, bis er auf der oder zur akzentuierten Silbe steil abfällt (L*). Das ungesternte H beschreibt das hohe Plateau zwischen den Akzenten. Es könnte gleichzeitig als *Trailing*-Ton des ersten und als *Leading*-Ton des zweiten Akzents interpretiert werden. Wenn der Brückenakzent durch eine Phrasengrenze in zwei Teile gespalten wird, beginnt der zweite Teil mit dem Tonmuster %H L*, wobei %H für einen phraseninitialen Grenzton steht. GToBI würde den letzten Akzent in einem Brückenakzent mit einem hohen *Leading*-Ton (H+L*) notieren, was dem ungesternten H in H* H L* sowie dem %H in %H L* entspräche.

Uhmans (1991) Arbeit über deutsche Intonation ist sehr viel ausführlicher und detaillierter, allerdings beschränkt sie sich, wie Wunderlich, auf eine Untersuchung der prosodischen Markierung grammatischer Funktionen, insbesondere auf das Verhältnis zwischen Intonation und Fokus-Hintergrund-Struktur. Uhmans Inventar besteht aus einem optionalen Grenzton L% oder H% zu Beginn einer Intonationsphrase (eine mittlere Tonhöhe ist der Normalfall und bleibt daher unmarkiert), einem obligatorischen finalen Grenzton L% oder H% sowie vier Akzenttönen L*, H*, L*+H und H*+L. Der nukleare Akzent ist immer bitonal, pränukeare Akzente können mono- oder bitonal sein. Tonale Zielpunkte direkt vor einer akzentuierten Silbe (d.h. *Leading*-Töne) gelten als unnötig für eine adäquate Beschreibung der Tonmuster.

Uhmans betrachtet Grenztonen als phonologische Korrelate der (syntaktischen) Phrasierung und Akzenttöne als phonologische Korrelate des Fokusmerkmals. Sie ordnet Akzenttönen mehr oder weniger distinktive Bedeutungen zu: Ein pränukeares L*+H fungiert als Topikmarker, L* hebt Hintergrundkonstituenten, H* Fokus- oder Hintergrundkonstituenten hervor, und H*+L repräsentiert den *Default*-Fokusakzent.

Folgende Beziehungen zwischen nuklearem Tonmuster und Satzmodus werden postuliert:

H*+L	L%	Deklarativsätze, W-Fragen
L*+H	H%	Echo-Fragen, Ja/nein-Fragen
H*+L	H%	Ja/nein-Fragen (nuklearer Akzentton markiert)
L*+H	L%	W-Fragen (nuklearer Akzentton markiert)

Der Ansatz von Féry (1993) widmet sich ebenfalls dem Einfluß der Fokusstruktur auf die Intonationsmuster des Deutschen. Ihr Inventar weicht nur geringfügig von Uhmans ab. So nimmt Féry z.B. die gleichen vier bitonalen Akzenttöne an. Die monotonalen Akzenttypen sind kein zugrundeliegender Bestandteil ihres Inventars, sondern werden durch phonologische Regeln aus

H*+L H%	Fragen, Drohungen
L*+H+L	Implizites 'selbstverständlich', leicht drohend
H+H*+L	Fernsehreporter-Stil
H*+M	Rufkontur

Allerdings ist die Struktur L*+H+L problematisch, weil Féry hier einen bitonalen *Trailing*-Ton konstruiert, der ansonsten in ihrem System nicht vorkommt. Die Kontur wäre besser als L*+H mit folgendem L% beschrieben. Féry verwirft diese Analyse aber, vermutlich weil sie der Verteilung der Töne über die Silben in ihren Beispielen nicht entspricht. Außerdem ist die steigend-fallende Kontur nicht zwangsläufig ein Spiegelbild der fallend-steigenden Kontur, die Féry als H*+L H% repräsentiert. Nach GToBI, das, wie wir sehen werden, nicht nur *Trailing*-Töne, sondern auch zwei verschiedene Typen von Grenztonen in seinem Inventar hat, wird das steigend-fallende Muster als L*+H L-(%) transkribiert, wobei das H einen *Trailing*-Ton darstellt. Dies ist nicht als parallel zum fallend-steigenden Muster H* L-H% zu verstehen: In letzterem variiert das *Alignment* des L je nach Position metrisch starker (*stressed*) postnuklearer Silben, die als 'Anker' für den L Ton dienen. Das L ist hier ein tiefer Grenzton (vgl. Abschnitte 3.3 und 4.2). GToBI erfaßt auf diese Weise, daß die zeitliche Koordination von Tönen und Text in beiden Mustern unterschiedlich ist. Im übrigen ist die Bedeutung, die Féry diesem Muster, das sie auch als 'späten Gipfel' bezeichnet (1993:96), zuweist, eine andere als bei Kohlers spätem Gipfel: Während nach Féry die Kontur L*+H+L Selbstverständlichkeit ausdrückt, weist Kohler diese Bedeutung einer ganz anderen Kontur zu, nämlich dem frühen Gipfel (vgl. Abschnitte 2.1 und 4.1.1).

Die Beschreibung stilisierter Konturen, z.B. Rufkonturen, stellt ein weiteres Problem für Férys Modell dar. Sie repräsentiert die Rufkontur als H*+M und fügt ihrem System damit eine mittlere Tonebene hinzu. Obwohl einiges dafür spricht, Rufkonturen von anderen Intonationsphänomenen abzugrenzen, weil sie oft (geradezu) gesungen werden und daher eine musikalische Notation angebrachter sein mag, werden sie in GToBI mit Hilfe des regulären Inventars beschrieben: Die akzentuierte Silbe ist hoch (H*) und der folgende Phrasenakzent herabgestuft (*downstepped*), symbolisiert als !H-.

Obwohl Féry explizit von linksköpfigen Tonakzenten (d.h. der gesternte Ton steht links) ausgeht, nimmt sie einen Akzent in ihr Inventar auf, für dessen Beschreibung die Tonhöhe vor dem gesternten Ton essentiell ist: der frühe Gipfel H+H*+L. In Abschnitt 4.1 werden wir sehen, daß diese Kontur auch bei GToBI vorkommt, wenn auch in einer anderen Darstellung.

3 GToBI

3.1 Einführung

GToBI ist ein Transkriptionssystem für die Intonation des Deutschen, das leicht zu erlernen und flexibel für verschiedene Anwendungen einsetzbar ist. Es wurde zwischen 1995 und 1996 von Forschern aus Saarbrücken, Stuttgart, München und Braunschweig mit dem Ziel entwickelt, prosodisch annotierte Daten leichter austauschen zu können. Es wurde ein Konsistenz-Test des gemeinschaftlich vereinbarten Systems durchgeführt (Grice et al (1996), Reyelt et al (1996)), der zeigte, daß die Transkribenten oft schon nach kurzer Zeit in der Lage waren, GToBI konsistent anzuwenden. Als Hilfen standen gedruckte Trainingsmaterialien sowie begleitende Klang-Dateien (vgl. Benzmüller & Grice 1997) zur Verfügung. Individuelles Training wurde praktisch nicht gegeben. Die GToBI-Trainingsmaterialien sind im Jahr 2001 überarbeitet und erweitert worden, besonders im Hinblick auf *Leading*-Töne und Phrasenakzente. Sie sind über die offizielle ToBI-Webseite (<http://www.ling.ohio-state.edu/~tobi>) zugänglich, auf der auch Links zu ToBI-Systemen in anderen Sprachen zu finden sind. In diesem Aufsatz wird eine leicht modifizierte Version des ursprünglichen GToBI vorgestellt.

Ein Teil der Flexibilität GToBIs ist auf seine verschiedenen Beschreibungsebenen, die sogenannten *Tiers*, zurückzuführen. Das Konsens-System hat mindestens drei Tiers zur Verfügung für die Annotation von Text, Tönen und Grenzen. Prinzipiell werden Informationen nur dann notiert, wenn sie nicht (automatisch) von den Labels anderer Tiers oder vom Sprachsignal abgeleitet werden können.

Der Text-Tier liefert eine orthographische Transkription der geäußerten Wörter. Auf dem Ton-Tier wird die wahrgenommene Kontur in Form von Akzenttönen und Grentönen annotiert, mit diakritischen Zeichen für Veränderungen des Tonhöhenumfangs (*Pitch Range*) wie *Downstep* und *Upstep*, die unmittelbar vor den betreffenden Ton gesetzt werden. Informationen über die Stärke von Phrasengrenzen finden sich auf dem *Break-Index*-Tier. Eine weitere Ebene kann für zusätzliche Informationen aller Art verwendet werden (in MAE_ToBI als *Miscellaneous Tier* bezeichnet). Abbildung 1 zeigt ein Beispiel mit Sprachsignal, Ton-, Text-, Grenzstärken- und Kommentar⁵-Labels sowie F0-Kontur. Die einzelnen Labels werden in den folgenden Abschnitten erläutert.

⁵ *Creaky Voice* ist eines der gängigsten Kommentar-Labels. Es bezeichnet eine Anregungsart, bei der die Stimmbänder nur sehr langsam und/oder unregelmäßig schwingen. Aus diesem Grund kann oftmals keine Grundfrequenz berechnet werden. *Creaky Voice* ist ein häufig auftretendes Phänomen am Ende von Äußerungen.

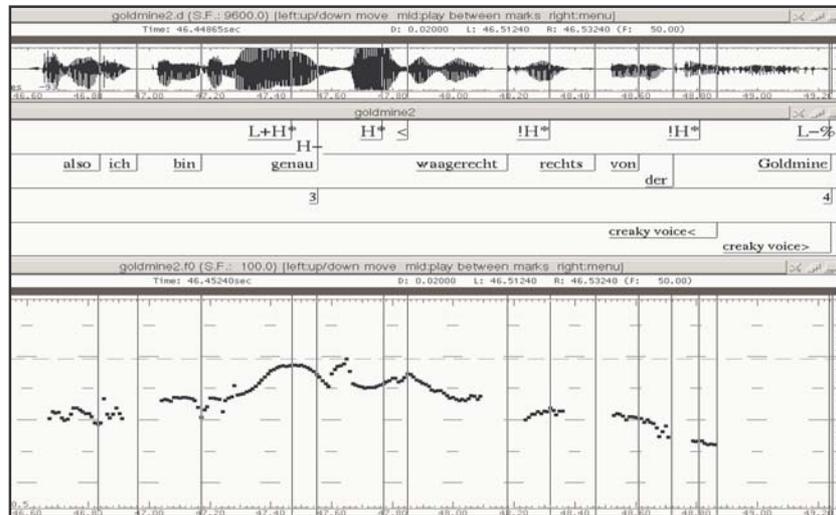


Abbildung 1: F0-Kontur und Annotationsebenen (Tiers) der Äußerung 'Also ich bin genau waagrecht rechts von der **GOLD**mine'⁶ (aus Grice & Benzmüller 1995).

Ein Großteil der kodierten Information muß in Relation zu anderen Tiers betrachtet werden. So sind z.B. die Ton- und Text-Tiers über die lexikalisch betonte (*stressed*)⁷ Silbe des akzentuierten Wortes miteinander verknüpft. Diese Assoziation wird auf der tonalen Ebene durch einen Stern ausgedrückt (z.B. L+H*). Die lexikalisch betonte Silbe eines Wortes muß gesondert identifiziert werden, weil *Stress* in der orthographischen Transkription nicht markiert wird. Es gibt zwei Ausnahmen von dieser Konvention: Zum einen wird *Stress* transkribiert, wenn die Wahl der Wortbetonung frei ist, wie in dem Wort *Kaffee*, das entweder als 'Kaffee oder Kaf 'fee (Duden 2000) realisiert werden kann. Als Kennzeichen dient das IPA-Symbol für Betonung ('), das auf der Text-Ebene unmittelbar vor der betreffenden Silbe eingefügt wird. Auf die gleiche Weise wird *Stress* markiert, wenn der Tonakzent nicht auf die lexikalisch primärbetonte Silbe eines Wortes fällt. In Beispielsatz (3) sind keine Markierungen nötig, weil die lexikalisch betonte Silbe *zwan* den Akzent trägt. In (4) ist dagegen der Akzent aus rhythmischen Gründen auf die Vorsilbe *hun* verschoben.⁸

⁶ In diesem Aufsatz verwenden wir fettgedruckte Großbuchstaben, um nukleare Silben zu kennzeichnen.

⁷ Bei abstrakten Hervorhebungen auf Wortebene sprechen wir von *Betonungen* bzw. *Stresses*, bei konkreten Hervorhebungen auf Äußerungsebene von *Akzenten*.

⁸ In ihrer Studie über das amerikanische Englisch haben Shattuck-Hufnagel et al. (1994) die Wahrnehmung von Prominzen in Äußerungen mit rhythmischen Zusammenstößen untersucht. Sie konnten zeigen, daß in vielen Fällen eine Verschiebung der wahrgenommenen Hervorhebung von der lexikalisch betonten auf eine frühere Silbe stattfindet, die sich durch eine deutliche F0-Bewegung (ein Merkmal von akzentuierten Silben) auf der prominenten Silbe manifestiert, ohne daß

- (3) hundertzwanzig
 (4) 'hundertzwanzig Mann (Giegerich 1985:218)

Da also die akzentuierte Silbe eines Wortes bis auf die erwähnten Ausnahmen nur indirekt zu bestimmen ist, ist es wichtig, daß das Label für den Tonakzent (d.h. der gesternte Ton) in der mit dem Akzent assoziierten Silbe liegt. Dies ist unproblematisch, wenn sich der Tonhöhengipfel oder das Tonhöhenal *innerhalb* der akzentuierten Silbe befindet. In diesem Fall wird das Label auf den Gipfel oder das Tal gesetzt. Wenn allerdings Gipfel oder Tal *außerhalb* der Akzentsilbe liegt, wird das Tonakzent-Label innerhalb der mit dem Akzent assoziierten Silbe platziert und der F₀-Extremwert mit einem '>' bzw. '<' Label markiert, je nachdem, ob er sich vor oder nach der Akzentsilbe befindet.⁹ Im Deutschen wird der Gipfel eines L+H* Tonakzents meist relativ spät in der Akzentsilbe erreicht, häufig erst auf der nachfolgenden Silbe (vor allem, wenn die Akzentsilbe kurz ist). Dies ist des öfteren auch bei H* Akzenten der Fall. Um Situationen zu vermeiden, in denen Labeller die genaue Position einer Silbengrenze bestimmen müssen, sollte das '<' Label immer dann verwendet werden, wenn der Gipfel in der Nähe einer Silbengrenze oder später erreicht wird (das Gleiche gilt für L* Tonakzente und Täler). Ein Beispiel für die Verwendung des '<' Symbols findet sich in Abbildung 1.

Unsichere Labels können durch ein Fragezeichen ('?'), prototypische Realisierungen einer bestimmten Kategorie durch ein Dollarzeichen ('\$') ausgewiesen werden.

3.2 Tonakzente

Im folgenden werden die sechs elementaren Tonakzente, die in GToBI vorkommen, beschrieben und schematisch dargestellt. Fettdruck markiert den Tonhöhenverlauf auf der akzentuierten Silbe, eine durchgezogene Linie steht für eine obligatorische, eine gestrichelte Linie für eine fakultative Tonbewegung vor oder nach der Akzentsilbe. Die als H bzw. L symbolisierten Töne sind nicht in absolute Grundfrequenzwerte übersetzbar, sondern in Relation zum jeweiligen Stimmumfang eines Sprechers zu sehen. Als Faustregel gilt, daß H Töne etwa innerhalb der oberen drei Viertel, L Töne im unteren Viertel des Stimmumfangs anzusiedeln sind.

diese Silbe länger würde (ein Merkmal von *stressed* Silben auf Äußerungsebene). Die lexikalisch betonte Silbe bleibt dagegen lang und zeigt keine tonale Bewegung. Daher ist es angemessen, dieses Phänomen als *Akzentverschiebung* (bzw. *Pitch Accent Shift*) zu bezeichnen und nicht als *Betonungsverschiebung* (bzw. *Stress Shift*), was zumindest im Englischen der gängigere Begriff ist. Im Deutschen wird oft nicht zwischen Akzent und Betonung unterschieden.

Féry (1986:23) weist darauf hin, daß im Deutschen Akzentzusammenstöße eher akzeptiert werden als im Englischen.

⁹ Wir folgen hiermit der Annotierungs-Konvention des MAE_ToBI.

– H*



Eine kanonisch mit H* markierte Silbe wird als relativ hoch wahrgenommen. Ihr kann ein flacher Anstieg vorangehen.

– L+H*



Wie bei H* wird die Akzentsilbe als hoch wahrgenommen. Sie folgt unmittelbar einem tiefen Zielpunkt, was zu einem steilen Tonhöhenanstieg innerhalb der akzentuierten Silbe oder zu einem Sprung auf die akzentuierte Silbe führt. Der Gipfel wird oft erst spät in der akzentuierten Silbe erreicht (vgl. Adriaens 1991, Grabe 1998).

– L*



Die mit L* assoziierte Silbe repräsentiert ein lokales Tonhöhenminimum, dem eine leicht fallende Bewegung vorangehen kann. Die akustischen Parameter Intensität und Dauer, die neben der F0-Bewegung zur Wahrnehmung eines Akzents beitragen, sind bei diesem Akzenttyp von entscheidender Bedeutung.

– L*+H



Hier folgt dem tiefen Zielpunkt innerhalb der Akzentsilbe ein (steiler) Anstieg, der spät in der Akzentsilbe beginnt und seinen Gipfel auf der nächsten Silbe (manchmal später) erreicht. Im Gegensatz zu L+H* wird der Akzent als tief wahrgenommen.

– H+L*



Die Akzentsilbe liegt im Bereich der Grundlinie des Stimmumfangs des Sprechers. Ihr geht ein hoher Zielpunkt voran, der in der Regel auf der unmittelbar vorhergehenden Silbe erreicht wird.

– H+!H*



Wie bei H+L* geht der Akzentsilbe eine Silbe mit größerer Tonhöhe unmittelbar voran. Allerdings ist die akzentuierte Silbe hier nicht tief, sondern liegt etwa in der Mitte zwischen dem Tonhöhengipfel und der Grundlinie des Stimmumfangs des Sprechers. Dieser mittlere Ton wird als herabgestufter (*downstepped*) Hochton symbolisiert: !H. Wenn diesem Akzenttyp ein tiefer Grenzton (L-) unmittelbar folgt, entspricht die Kontur einer stetig fallenden Bewegung von der

präakzentuierten Silbe durch die akzentuierte Silbe bis hin zur letzten Silbe der Phrase.

Das Fehlen des Akzenttyps H*+L im Inventar von GToBI wird ausführlich in Abschnitt 4.2 diskutiert.

Die sechs (paradigmatischen) Tonakzenttypen können je nach (syntagmatischem) Kontext modifiziert werden, d.h. die Ausrichtung der H und L Töne im Stimmumfang des Sprechers (bzw. Register¹⁰) kann variieren. Veränderungen im Stimmumfang können anhand einer konstruierten oberen und unteren Registerlinie beschrieben werden. Die gängigste Veränderung besteht im Absenken der oberen Registerlinie, dem bereits mehrfach erwähnten *Downstep*. Hierbei wird dem betreffenden H Ton ein Ausrufezeichen als Diakritikum vorangestellt, wie beim eben vorgestellten Akzenttyp H+!H*. Generell kann jeder H Ton herabgestuft werden (z.B. !H*, L*+!H), bis auf den ersten innerhalb einer Intonationsphrase, weil *Downstep* nur in Relation zu einem vorangehenden H Ton möglich ist. Es ist wichtig zu beachten, daß sich alle H Töne, die einem *Downstep* in derselben Intonationsphrase folgen, innerhalb des durch den *Downstep* verringerten Registers bewegen. Diese H Töne müssen nicht einzeln als herabgestuft markiert werden, es sei denn, es findet ein weiterer signifikanter *Downstep* statt. Eine solche Folge mehrerer *Downsteps* hintereinander zeigt die zweite Phrase in Abbildung 1, ...*WAAgerecht RECHTS von der GOLDmine*: 'H* !H* !H* L-%'. In den eher generativ ausgerichteten Autosegmental-Metrischen Modellen wird *Downstep* entweder automatisch durch einen bestimmten Akzenttyp ausgelöst, wie in Pierrehumberts Modell für das Englische (1980), oder als fakultative Operation behandelt (Gussenhoven 1984). Da GToBI oberflächenorientierter ist als die anderen AM-Modelle, wird *Downstep* einfach bei jedem Auftreten explizit markiert.

Zu Beginn jeder neuen Phrase wird die u.a. durch *Downstep* verursachte Absenkung der oberen Registerlinie in aller Regel aufgehoben (*Reset*). Ein *Reset* kann auch nach einer Folge von *Downsteps* innerhalb einer Phrase vorkommen, dann meist unmittelbar vor dem Nukleus. Das heißt: Nach einem schrittweisen Absinken der Tonhöhe findet ein Schritt nach oben, ein *Upstep*, zum Gipfel der nuklearen Silbe statt. Solche Konturen bzw. Akzentfolgen kommen im Standarddeutschen etwa bei emphatischer Sprechweise vor. Sie sind auch für süddeutsche Dialekte (Truckenbrodt, erscheint) sowie für das Englische (Ladd 1983:735, Beispiel 4b; Beckman & Pierrehumbert 1986:299¹¹) belegt. In GToBI werden diese Töne durch das diakritische Zeichen '^' markiert.¹²

¹⁰ Genaugenommen versteht man unter *Register* den in einem Äußerungsteil genutzten Sprechstimmumfang. Der *Sprechstimmumfang* ist der Teil des Stimmumfangs, der zum Sprechen genutzt wird, während *Stimmumfang* den gesamten Tonhöhenbereich bezeichnet, der vom jeweiligen Sprecher produziert werden kann.

¹¹ Im Gegensatz zu Ladd postulieren Beckman & Pierrehumbert eine Intermediärphrasengrenze unmittelbar vor dem *Upstep*. In den deutschen Beispielen, die einen *Upstep* enthalten und mit GToBI annotiert wurden, haben wir keine Hinweise auf eine Intermediärphrasengrenze gefunden.

¹² Ein Beispiel für *Upstep* bei *Leading*-Tönen wird in Abschnitt 4.1.1 (Abb. 5) diskutiert.

Die explizite Kennzeichnung des *Upstep*, die GToBI von anderen Autosegmental-Metrischen Ansätzen unterscheidet, wird nicht nur für Akzenttöne sondern auch für Grenztöne verwendet, auf die wir im nächsten Abschnitt eingehen werden.

3.3 Grenztöne

GToBI unterscheidet zwei Ebenen der Phrasierung: die (kleine) Intermediärphrase (ip) und die (große) Intonationsphrase (IP). Eine IP besteht aus mindestens einer ip, und jede ip beinhaltet mindestens einen Tonakzent.¹³

Die Grenztöne dieser Phrasen beschreiben die Kontur vom letzten Ton des nuklearen Tonakzents bis zum Ende der jeweiligen Phrase. Die schematischen Darstellungen der Grenztöne und Grenzton-Kombinationen gehen jeweils von einem vorangehenden H*-Tonakzent aus. Es gibt drei verschiedene Grenztöne für Intermediärphrasen:

- L- 

L- symbolisiert einen tiefen Zielpunkt am Ende der ip.

- H- 

H- hat in etwa den gleichen F0-Wert wie der letzte Gipfel der Phrase. Wenn mehrere Silben zwischen dem nuklearen H Ton und dem hohen ip-Grenztone auftreten, entsteht ein hohes Plateau.

- !H- 

Ein hoher ip-Grenztone kann ebenso wie jeder andere H Ton in Relation zu einem vorangehenden H herabgestuft werden. Dies kommt zum Beispiel in Rufkonturen vor (vgl. Abschnitt 3.5).

Theoretisch kann ein weiterer Grenzton einer Intermediärphrase angenommen werden, nämlich ^H-. Allerdings haben sich bisher keine Beispiele gefunden, die diesen Ton eindeutig von H- unterschieden hätten.

Der Zielpunkt eines ip-Grenztone wird oft auf einer postnuklearen, lexikalisch betonten Silbe (falls vorhanden) erreicht und erstreckt sich bis zum Anfang der letzten Silbe der Phrase. Diese Tendenz des zeitlichen *Alignment* von ip-Grenztone und lexikalisch betonter Silbe wurde in einer Untersuchung von Grice

¹³ Eine Ausnahme bilden die sogenannten *Intonational Tags* (Anhängsel wie z.B. ..., *ne?*), die als enklitische tonale Einheiten betrachtet werden können.

und Benzmüller (1998)(vgl. Abschnitt 4.2) bestätigt und zeigt, daß es sich bei dieser Art von Grenztönen um sogenannte *Phrasenakzente* im Sinne von Grice, Ladd & Arvaniti (2000) handelt. Phrasenakzente bestehen aus Tönen, die als Grenztöne fungieren, aber auch mit lexikalisch betonten Silben oder anderen tontragenden Einheiten assoziiert werden können. GToBI bietet die Möglichkeit, diese Assoziation durch die Transkription eines zusätzlichen L(*) oder H(*) auf dem Ton-Tier anzuzeigen. Der in Klammern gesetzte Stern und die Bezeichnung *Phrasenakzent* bringen die (sekundäre) Prominenz dieses Tons zum Ausdruck. Abbildung 2 illustriert ein Beispiel für diese Notation (weitere Beispiele finden sich in Abb. 3, Abschnitt 3.5, und in Abb. 6, Abschnitt 4.1.1).

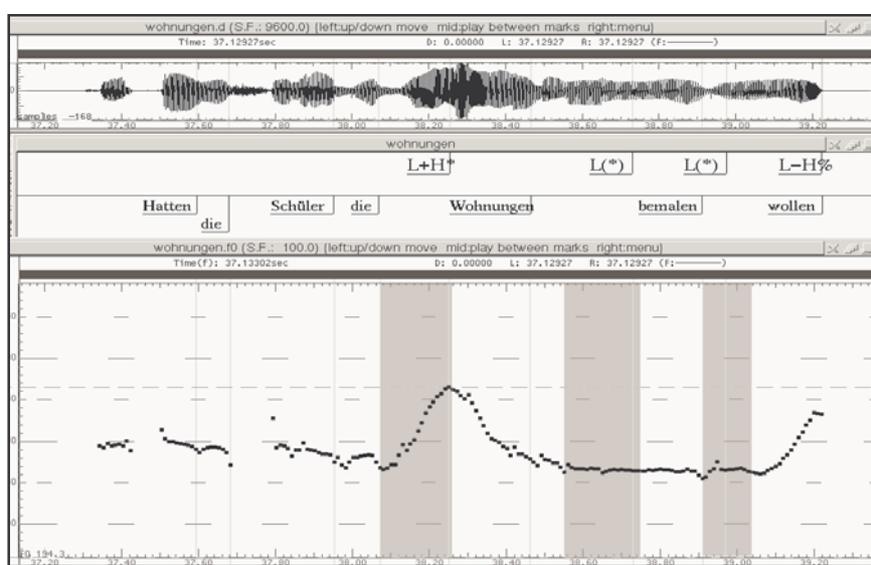


Abbildung 2: Optionale Transkription von Phrasenakzenten in der Äußerung 'Hatten die Schüler die **WO**hnungen be**MA**len **WO**llen?' (leicht abgeändert aus Grice, Ladd & Arvaniti (2000)); die schattierten Bereiche markieren die prominenten Silben *WOH*, *MA* und *WOL*.

Eine Intonationsphrasengrenze tritt immer zusammen mit einer Intermediärphrasengrenze auf. Daher werden ihre Töne als Kombinationen angegeben. Für die hier beschriebene neue Version von GToBI wurde die Notation der Grenztöne vereinfacht und phonetisch transparenter gestaltet. So wird etwa in Fällen, in denen IP- und ip-Grenztöne die gleiche Tonhöhe repräsentieren, nurmehr einer der beiden Töne transkribiert: H-% ersetzt die alte Notation H-L% (Erläuterung siehe unten) und L-% die ursprüngliche Notationsform L-L%. Die folgende

Beschreibung kanonischer Grenzton-Konturen geht von mindestens zwei Silben zwischen dem letzten Akzentton und der Phrasengrenze aus.¹⁴

– H-%



Diese Grenzton-Kombination repräsentiert ein hohes Plateau. Der Hauptunterschied zwischen H- und H-% ist nicht tonal, sondern besteht in der wahrgenommenen Stärke der Grenze, ausgedrückt durch die Labels 3 bzw. 4 auf dem parallelen *Break-Index-Tier* (vgl. Abschnitt 3.4).

Es hat bereits verschiedene Vorschläge für die Beschreibung eines Plateaus an Phrasengrenzen gegeben. Grabe (1998) verwendet zum Beispiel die Transkription 0% für eine Kontur, die vom Ende des letzten Akzenttons bis zur Grenze mehr oder weniger gleich bleibt. Das Problem dieser Notation besteht darin, daß ein solcher unmarkierter Grenzton nicht direkt anzeigt, ob die Phrase hoch oder tief endet. Sein Wert hängt vom vorangehenden Akzentton ab. Die ursprüngliche GToBI-Transkription eines Plateaus war H-L%, mit automatischem *Upstep* des L Tons (analog zur Notation bei Pierrehumbert 1980). Da aber ein L zur Repräsentation eines mittleren oder sogar hohen Tonniveaus jeder Intuition zuwiderläuft und daher nur schwer zu vermitteln ist, wird in der neuen Transkription auf diesen Ton verzichtet. Das kombinierte Label H-% hat den Vorteil, die Tonhöhe am Ende der Phrase direkt wiederzugeben, ohne daß ein syntagmatischer Rückgriff auf vorangehende Akzenttöne nötig wäre. Auf diese Weise wird das System durchsichtiger, leichter erlernbar und besser nutzbar für automatische Datenverarbeitung.

– H-^H%



Der *Upstep* des H% symbolisiert einen steilen Anstieg der Tonhöhe innerhalb der letzten Silbe der Phrase, oft sehr hoch im Stimmumfang des Sprechers.

– L-H%



Diese Grenzton-Kombination steht für eine fallend-steigende Kontur, wenn ein H Ton vorangeht, ansonsten für eine flache Kontur mit einem Anstieg auf eine mittlere Tonhöhe innerhalb der letzten Silbe der Phrase.

– L-%



Der Hauptunterschied zwischen L- und L-% liegt in der wahrgenommenen Grenzstärke. Außerdem fällt L-% im allgemeinen tiefer als L-. Das Absinken

¹⁴ Wenn zwischen der letzten Akzentsilbe und der Grenze weniger Material zur Verfügung steht, geht ein Teil der beschriebenen Form verloren, weil nicht genug Zeit für die Realisierung der einzelnen Töne vorhanden ist.

der Tonhöhe am Phrasenende ist vermutlich auf das sogenannte *Final Lowering* zurückzuführen, ein Effekt, der sich allerdings nicht auf die letzte Silbe beschränken muß. Wir unterscheiden nicht zwischen mehreren ‘Tiefegraden’ an IP-Grenzen.¹⁵

– %H



GToBI bietet die Option, einen außergewöhnlich hohen Beginn einer IP durch den initialen Grenzton %H zu markieren. Der Beginn einer IP auf mittlerem oder tiefem Tonniveau ist der *Default*-Fall und wird daher nicht explizit gekennzeichnet.

Folgende Grenzton-Kombinationen sind derzeit nicht im Inventar von GToBI enthalten:

– L-^H%



Prinzipiell könnte diese Kombination für die Beschreibung einer flachen Kontur mit einem steilen Anstieg bis in einen hohen Bereich des Stimmumfangs des Sprechers verwendet werden (L-H% steigt nur auf ein mittleres Niveau). Allerdings ist diese Kontur bislang weder durch genügend Beispiele belegt noch gibt es Hinweise auf einen Bedeutungsunterschied gegenüber L-H% (oder H-^H%), der die Einführung einer eigenen Kategorie rechtfertigen würde.

– H-L%



Diese Notation wurde in früheren GToBI-Versionen verwendet, um ein Plateau zu beschreiben (mit automatischem *Upstep* des L% nach einem H- Phrasenakzent; siehe oben). Da *Upstep* nunmehr explizit markiert wird, könnte H-L% eine fallende Kontur nach einem hohen Plateau symbolisieren. Wenn auch eine solche Kontur im Standarddeutschen nicht belegt ist, so wurde sie doch in Untersuchungen zu ostfälischen (Kerckhove 1948:63) und pfälzischen Dialekten (Peters 2001a, 2001b) erwähnt.

Die übrigen drei logisch möglichen Grenzton-Kombinationen sind durch bereits erwähnte abgedeckt: H-H% und H-^L% können mit H-% gleichgesetzt werden, und L-^L% wäre der Kontur L-H% sehr ähnlich. Weitere Nuancen bei der Beschreibung von Tonbewegungen an Intonationsphrasengrenzen haben sich bisher als nicht notwendig erwiesen, weil das GToBI-Inventar alle in der einschlägigen Literatur erwähnten charakteristischen Bedeutungen oder Funktionen phrasenfinaler Konturen abdeckt.

¹⁵ Dies bedeutet, daß wir nicht zwischen L-% und L-L% für ‘tief’ und ‘extra tief’ unterscheiden, weil unklar ist, inwiefern eine solche Differenzierung für das Standarddeutsche funktional gerechtfertigt werden könnte.

3.4 *Break*-Indizes

Die Verwendung der *Break*-Index-Ebene ist eng an MAE_ToBI angelehnt, d.h. die Indizes 3 und 4 entsprechen den Grenzen einer Intermediär- bzw. einer Intonationsphrase. Allerdings werden diese beiden *Break*-Indizes in GToBI im Normalfall nicht (hand)gelabelt, sondern können automatisch gesetzt werden, weil die entsprechenden Grenzen auch auf dem Ton-Tier gekennzeichnet sind (als – bzw. %). Bei Unklarheiten oder mangelnder Übereinstimmung zwischen den Tiers stehen folgende Labels zur Verfügung: ‘4-’ wird in den Fällen verwendet, in denen eine Phrasengrenze wahrgenommen wird, über deren Stärke allerdings nicht entschieden werden kann. Ferner unterscheidet GToBI zwei Diskrepanzen zwischen der tonalen und der rhythmischen Ebene, die in MAE_ToBI beide mit dem Index 2 markiert werden: ‘2r’ steht für einen rhythmischen Bruch bei tonaler Kontinuität, z.B. bei einer Zögerpause, während ‘2t’ eine tonale Unterbrechung bei rhythmischer Kontinuität repräsentiert, z.B. wenn eine Phrasengrenze wahrgenommen wird, aber keine Pause eintritt. Dies ist oft bei schneller Sprechweise der Fall. *Break*-Index 2 ohne zusätzliches Diakritikum findet bei GToBI keine Verwendung, ebensowenig wie Indizes unterhalb dieser Ebene.

3.5 Gängige nukleare Intonationsmuster

Tabelle 1 zeigt schematische Darstellungen von nuklearen Intonationskonturen des Deutschen, die häufig in der Literatur erwähnt werden, sowie mögliche Kontexte und konkrete Textbeispiele. In den schematischen Konturen markieren extra fettgedruckte Linien akzentuierte Silben und weniger fettgedruckte Linien postnuklear betonte Silben. Horizontale Linien repräsentieren die Grundlinie des Stimmumfangs des Sprechers. Wir gehen bei der Darstellung von maximalen Konturen aus, d.h. dem Nukleus folgen mindestens eine postnuklear betonte sowie eine weitere Silbe. Obwohl die meisten konkreten Beispiele tatsächlich eine postnukleare Betonung aufweisen, so entsprechen doch nicht alle der maximalen Kontur. Dies liegt daran, daß die Äußerungen aufgrund ihrer pragmatischen Beispielhaftigkeit und nicht wegen ihrer rhythmischen Struktur ausgewählt wurden. So wird etwa die steigende Kontur 3a mit zwei Äußerungen beschrieben: *Tauschen Sie auch BRIEFMARken?* und *Von wem ich das HAbe?*. Im ersten Satz entspricht die extra fettgedruckte Linie der nuklearen Silbe, **BRIEF**, und die fettgedruckte Linie der postnuklear betonten Silbe **MAR**. Im zweiten Satz folgt der nuklearen Silbe **HA** nur noch eine Silbe. Da für die Realisierung der schematischen Kontur (steigend-Plateau-steigend) nicht genug Segmente vorhanden sind, steigt die Tonhöhe von der Silbe **HA** direkt bis zum Ende der Phrase. Wenn die nukleare Silbe die letzte der Phrase bildet, kann die Kontur verkürzt werden (*truncation*), besonders wenn die Silbenkoda stimmlose

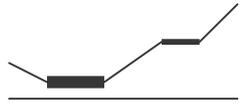
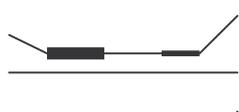
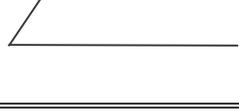
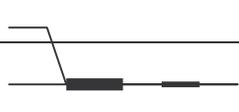
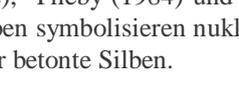
Obstruenten enthält. Dies gilt vor allem für fallende Konturen (vgl. Grabe 1998).

Die Tabelle beinhaltet kontextuelle Interpretationen, die auf die konkreten Beispiele bezogen sind. Sie sollen nicht als allgemeine Bedeutungen der Konturen verstanden werden. Wir nehmen keine strikte Trennung von rein linguistischen und eher paralinguistischen Aspekten intonatorischer Bedeutung vor. Dies geschieht im Einklang mit den Erkenntnissen von Scherer, Ladd & Silverman (1984), die in einer Reihe von Experimenten festgestellt haben, daß auch die emotionale Botschaft einer Äußerung von phonologischen Kategorien wie Akzentton-Typ oder Grenzton mitgetragen wird, stets in Verbindung mit anderen linguistischen Kategorien, wie z.B. dem Satztyp. In einem im Grundsatz phonologischen Beschreibungsmodell wie GToBI steht diese, (vermutlich) kognitiv gesteuerte Seite der emotionalen Botschaft im Vordergrund, die an anderer Stelle (Ladd et al 1985, Mozziconacci 1998) dem Begriff *Einstellungen* zugeordnet wurde, im Gegensatz zu *Emotionen* im engeren Sinne, die eher physiologisch bedingt sind und deren angemessene Beschreibung sich kontinuierlicher akustischer Parameter bedient (vorausgesetzt die Emotionen werden überhaupt über den Sprachkanal ausgedrückt).

Die Konturen in der Tabelle werden unten im einzelnen diskutiert.

Tabelle 1 - Gängige nukleare Intonationsmuster des Deutschen und Beispiele für ihre Verwendung

		GToBI	Schematische Kontur	Kontext	Beispiel
Fallend	1a	H* L-%		Neutrale Aussage Neutrale W-Frage	Mein ZAHN tut WEH. ¹ Wo hast du den WA gen ge PARKT ? ¹
	1b	L+H* L-%		Kontrastive Feststellung	Schon der Ver SUCH ist STRAFbar! ²
Steigend-fallend (Später Gipfel)	2	L*+H L-%		Sebstverständliche Feststellung Engagierte oder sarkastische Feststellung	Das WEISS ich SCHON! ⁶ Der Blick ist ja F abelhaft! ³

Steigend	3a	L* H- ^H%		Neutrale Entscheidungsfrage Echo-Frage	Tauschen Sie auch BRIEF-MARKEN ? ¹ Von wem ich das HABE ? ²
	3b	L* L- H%		Empörung Melden am Telefon	DOCH! BECKEN-BAUER ? ⁴
	3c	(L+)H* H-^H%		Anschlußfrage	...oder ist Ihr BRUDER HIER ? ⁵
Gleichbleibend	4	(L+)H* H-(%)		Weiterweisende Äußerung Floskelhafte Ausdrücke	ANDERERSEITS ... ⁶ Guten MORGEN ! ³
Fallend-Steigend	5	(L+)H* L-H%		Höfliches Angebot	Mögen Sie ROGGENBRÖTCHEN ? ¹
Früher Gipfel	6a	H+!H* L-%		Bestätigung einer Bekannten Tatsache	Hab' ich mir schon ge- DACHT ? ⁷
	6b	H+L* L-%		Beruhigende oder höfliche Aufforderung	Nun er- ZÄHLE doch MAL ! ²
Stilisierte Herabstufung	7	(L+)H* !H-%		Ausrufe	BECKEN-BAUER!

Beispiele aus ¹Féry (1993), ²von Essen (1964), ³Fox (1984) ⁴Ladd (1996), angelehnt, ⁵Moulton (1962), ⁶Pheby (1984) und ⁷Grice & Benzmüller (1995). Fettgedruckte Großbuchstaben symbolisieren nukleare Silben, normal gedruckte Großbuchstaben postnuklear betonte Silben.

Fallend – In der Autosegmental-Metrischen Literatur wird nur *ein* fallender Konturtyp (in nuklearer Position) beschrieben. GToBI unterscheidet dagegen zwischen einer einfach fallenden Bewegung, notiert als H* L-%, und einer fallenden Kontur, der ein steiler Anstieg zur Akzentsilbe oder innerhalb der Akzentsilbe vorangeht. Letztere wird mithilfe eines tiefen *Leading*-Tons symbolisiert: L+H* L-%. Diese Kontur drückt oft einen Kontrast aus (besonders bei großer Ausnutzung des Stimmumfangs eines Sprechers), wie im Beispiel in Abbildung 3.

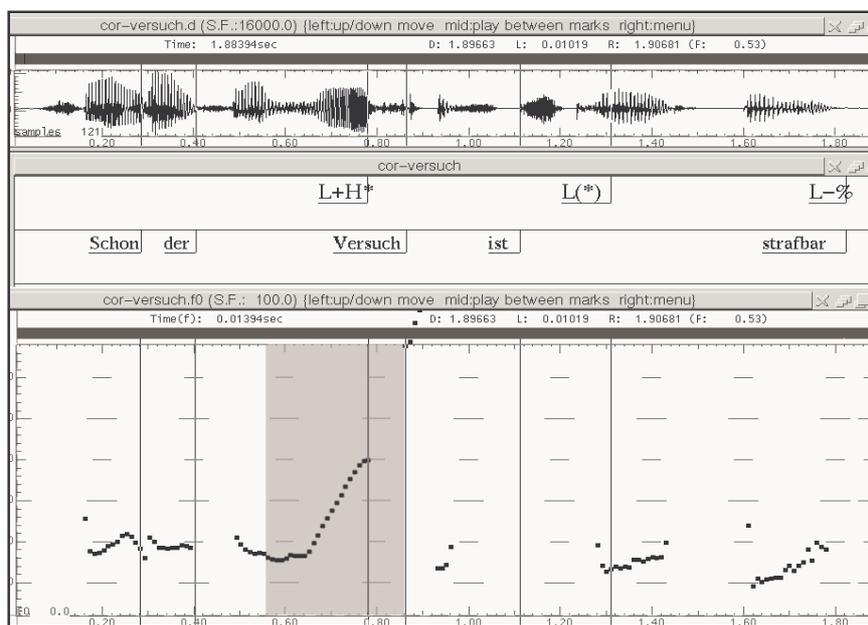


Abbildung 3: F0-Kontur von L+H* L-% auf 'Schon der VerSUCH ist STRAFbar'. Der schattierte Bereich markiert die nukleare Silbe *SUCH*; L(*) repräsentiert einen optionalen Phrasenakzent auf der postnuklear betonten Silbe *STRAF*.

Steigend-fallend – Die steigend-fallende Kontur, notiert als L*+H L-%, unterscheidet sich von der eben beschriebenen Kontur L+H* L-% in der Synchronisierung mit der akzentuierten Silbe. Bei L*+H L-% ist die Tonhöhe in der Akzentsilbe zunächst tief. Der Anstieg beginnt innerhalb der Akzentsilbe oder kurz danach (vgl. Abbildung 4). Bei L+H* L-% dagegen setzt der Anstieg früher ein, oft bereits vor der Akzentsilbe. Die akzentuierte Silbe der ersten Kontur klingt tief, die der zweiten Kontur hoch.

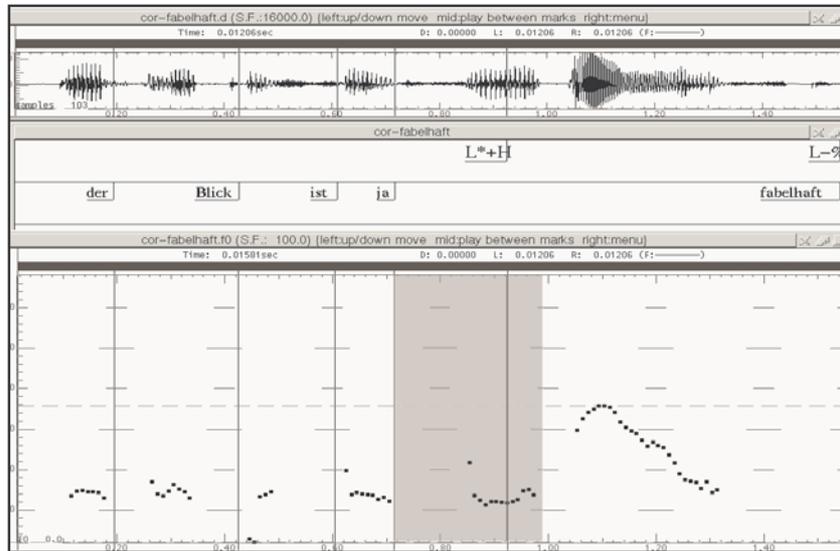


Abbildung 4: F0-Kontur von L*+H L-% auf 'Der Blick ist ja **F**abelhaft'; der schattierte Bereich markiert die nukleare Silbe **FA**.

Steigend – In der frühen Autosegmental-Metrischen Literatur finden sich höchstens zwei verschiedene Arten steigender Konturen. GToBI differenziert zwei tief einsetzende nukleare Bewegungen, L* L-H% und L* H-^H%, wobei der Endpunkt der zweiten höher liegt als der der ersten. Zusätzlich gibt es die Möglichkeit einer Kontur, die von einer mittelhohen Akzentsilbe bis zum Phrasenende steigt (mit oder ohne Anstieg zur Akzentsilbe): (L+)H* H-^H%.

Gleichbleibend – Konturen, die auf gleichbleibender oder anhaltender Tonhöhe enden, werden in der Literatur kaum behandelt. Nach Féry kann eine Kontur, die einem nuklearen L*+H folgt, bis zum Phrasenende auf der Höhe des *Trailing*-Tons bleiben. Dieser gleichbleibende (Grenz-)Ton wird von Féry nicht transkribiert. Damit verzichtet sie auf eine Unterscheidung zwischen steigenden und gleichbleibenden nuklearen Mustern: „As a matter of fact, rising tones and progradient intonation cannot be kept apart.“ (1993:89). GToBI kennzeichnet eine gleichbleibende Kontur mit oder ohne *Rising On glide* (vgl. Abschnitt 4.1.2) als L+H* H-% bzw. H* H-%.

Fallend-Steigend – Im allgemeinen werden fallend-steigende Bewegungen in GToBI als (L+)H* L-H% transkribiert. Es ist aber auch möglich, eine hohe fallend-steigende Bewegung darzustellen, bei der die Tonhöhe zwischen den

beiden Gipfeln nicht ganz tief absinkt, und zwar als (L+)H* !H-^H%. Allerdings ist fraglich, ob eine solche Unterscheidung tatsächlich notwendig ist.

Früher Gipfel – GToBI hat zwei Arten von Frühe-Gipfel-Konturen im Inventar: H+!H* und H+L*. Die erste entspricht dem von Kohler beschriebenen Muster, das eine ‘zusammenfassende, abschließende Argumentation’ (1995:198) anzeigt, sowie der Kontur, die von Féry als H+H*+L transkribiert wird. Von Essen behauptet, daß diese Tonbewegung Finalität kennzeichnet, sogar auf nicht-abgeschlossenen Teilen einer Äußerung. Das Muster werde oft von Radiosprechern beim Verlesen der Meldungen verwendet. Diese Beobachtung ähnelt derjenigen Férys, die Frühe-Gipfel-Konturen vor allem bei Fernsehreportern ausgemacht hat. Die schematische Darstellung von H+!H* in Tabelle 1 gibt Raum für eine postnuklear betonte Silbe, die in dem spontansprachlichen Beispiel nicht vorhanden ist. In einem Beispielsatz wie *Sie hätte ja LÜgen KÖNnen* wäre die zweite fettgedruckte Linie mit der Silbe *kön* assoziiert. Die andere Frühe-Gipfel-Kontur, H+L*, drückt nach von Essen eine fatalistische Stimmung aus. Sie kann auch in beruhigenden oder höflichen Aufforderungen verwendet werden, wie in dem in Tabelle 1 angegebenen Beispiel. Eine detailliertere Diskussion der Frühe-Gipfel-Konturen folgt in Abschnitt 4.1.1.

Stilisierte Herabstufung – Die stilisierte Herabstufung (oder Rufkontur) wird als (L+)H* !H- repräsentiert. Der Phrasenakzent !H- fällt auf eine betonte Silbe, falls eine solche vorhanden ist. Wie bereits in Abschnitt 3.3 erwähnt, besteht die Möglichkeit, die Prominenz dieser Silbe mit dem zusätzlichen Label (*) zu kennzeichnen. Die vielfältige Verwendung der Rufkontur im Deutschen ist ausführlich von Gibbon (1976, 1998) beschrieben worden.

4 GToBI im Vergleich mit anderen AM-Ansätzen

Tabelle 2 zeigt, daß die drei AM-Ansätze von Wunderlich (1988), Uhmann (1991) und Féry (1993) jeweils eine Reihe von nuklearen Konturen des Deutschen darstellen (können), diese Konturen aber nicht immer deckungsgleich sind. GToBI ist mit dem Ziel entwickelt worden, sämtliche bekannten Tonmuster beschreiben zu können. Zusätzlich ermöglicht GToBI feinere Differenzierungen bei der Darstellung einzelner Tonverläufe. Diese größere Vielfalt wird vor allem durch zwei Faktoren erreicht: Zum einen hat GToBI *Leading*-Töne im Inventar, die eine Unterscheidung z.B. zwischen einer fallenden Kontur und einer fallenden Kontur mit vorangehendem *Onglide* ermöglichen. Zum anderen werden in GToBI zwei Phrasierungsebenen (ip und IP) und damit zwei Arten von Grenztönen (Phrasenakzente und IP-Grenztöne) verwendet, die zusätzlich durch explizite Markierung von *Upstep* und *Downstep* differenziert werden

können. Alle anderen Ansätze gehen dagegen nur von einem Grenzton aus. Wir werden diese Aspekte im folgenden näher behandeln.

Tabelle 2 - Nukleare Tonmuster des Deutschen – drei Modelle im Vergleich mit GToBI

		Wunderlich	Uhmann	Féry	GToBI
Fallend	1 a	H* L	H*+L L%	H*+L	H* L-%
	1 b				L+H* L-%
Steigend-fallend (Später Gipfel)	2		L*+H L%	L*+H+L	L*+H L-%
Steigend	3 a	L* H H%	L*+H H%	L*+H	L*(+H) H-^H%
	3 b	L* H%			L* L-H%
	3 c				(L+)H* H-^H%
Gleichbleibend	4			L*+H	(L+)H* H-(%)
Fallend-steigend	5		H*+L H%	H*+L H%	(L+)H* L-H%
Früher Gipfel	6 a			H+H*+L	H+!H* L-%
	6 b	%H L* L			H+L* L-%
Stilisierte Herabstufung	7			H*+M	(L+)H* !H-%

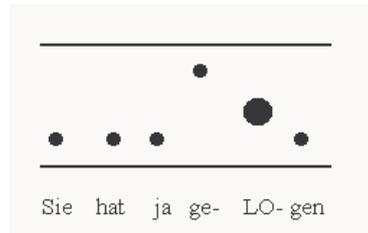
4.1 *Leading*-Töne

GToBI bietet die Möglichkeit zur Beschreibung hoher oder tiefer *Leading*-Töne. Dies bedeutet, daß die Tonhöhe vor einer Akzentsilbe entweder als H transkribiert werden kann, was als Frühe-Gipfel-Kontur bezeichnet wird, oder als L, was eine Kontur mit Anstieg auf eine akzentuierte Silbe zur Folge hätte, die auch *Rising On glide* genannt wird.

4.1.1 Frühe Gipfel

Es gibt viele Beispiele, die die Existenz von Frühe-Gipfel-Konturen (bzw. Konturen mit einem hohen *Leading*-Ton) im Deutschen nahelegen. Ein klassischer Fall ist Kohlers früher Gipfel, wiedergegeben in (5a).

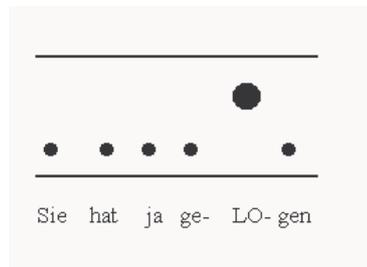
(5a)



(angelehnt an Kohler 1995:123)

Kohler hat Perzeptionstests durchgeführt, die eindeutig belegen, daß ein hoher Ton unmittelbar vor der Akzentsilbe distinktiv ist, d.h. die Frühe-Gipfel-Kontur, die gegebene Information signalisiert, repräsentiert eine andere linguistische Kategorie als ein mittlerer Gipfel, der zur Markierung neuer Information verwendet wird. In GToBI kommt diese Unterscheidung in den Notationen zum Ausdruck: Ein früher Gipfel wird durch einen H *Leading*-Ton vor der (tieferen) Akzentsilbe symbolisiert (in (5a) auf der Silbe *ge-*), während ein mittlerer Gipfel als H* auf der Akzentsilbe (möglicherweise mit einem L *Leading*-Ton) transkribiert würde (im Beispielsatz (5b) auf der Silbe *LO-*):

(5b)



(angelehnt an Kohler 1995:123)

Im Zwei-Ebenen-Modell von Isačenko & Schädlich entspricht die 'praktisch' fallende Kontur dem frühen Gipfel. Sie ist in (6a) schematisch dargestellt. (6b) zeigt eine 'postiktisch' fallende Kontur, die mittlere und späte Gipfel repräsentiert.

(6a) präiktisch fallend

die ↓ KINder

(6b) postiktisch fallend

die KIN ↓ der

(Isačenko & Schädlich 1966:60)

GToBI hat nicht nur *eine* Frühe-Gipfel-Kontur, sondern *zwei*: H+!H* und H+L*. Grabe (1998:89f.) erklärt den Unterschied zwischen beiden Formen als teilweisen bzw. vollständigen *Downstep* des zugrundeliegenden Akzenttons H*+L (allerdings hält sie den Unterschied für graduell). Wie in Abschnitt 3.5 gezeigt, beschreibt schon von Essen (1964) beide Arten der Frühe-Gipfel-Kontur und weist ihnen unterschiedliche Verwendungen zu.

Die Abbildungen 5 und 6 illustrieren spontansprachliche Äußerungen der beiden Kontur-Typen.

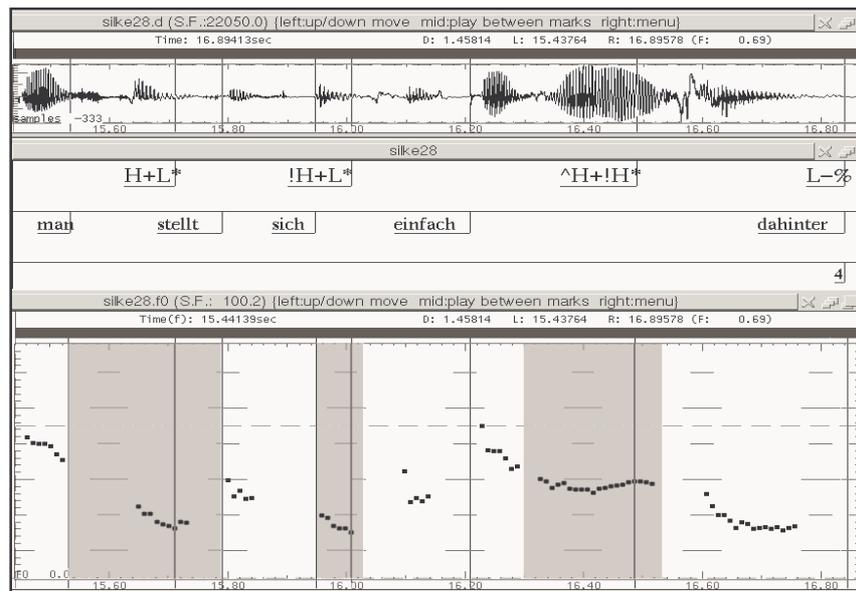


Abbildung 5: F0-Kontur eines nuklearen H+!H* L-% auf 'Man stellt sich einfach da**HIN**ter'; die schattierten Bereiche markieren die akzentuierten *STELLT*, *EIN* und *HIN*.

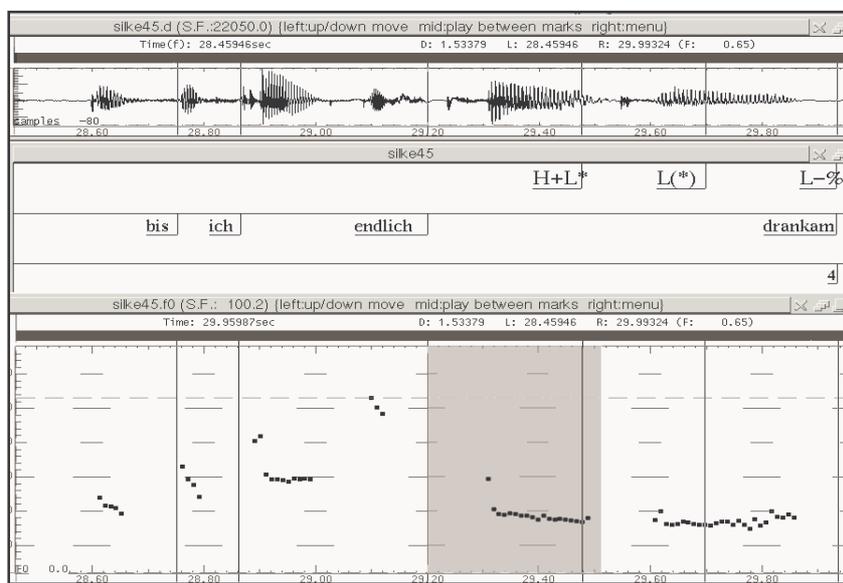


Abbildung 6: F₀-Kontur von H+L* L-% auf ‘... bis ich endlich **DRAN-**KAM’. Der schattierte Bereich markiert die nukleare Silbe **DRAN**; L(*) repräsentiert einen optionalen Phrasenakzent auf der postnuklear betonten Silbe **KAM**.

Der in Abbildung 5 dargestellte Satz liefert ein Beispiel für die Auswirkungen der syntagmatischen Prozesse *Downstep* und *Upstep* (vgl. Abschnitt 3.2) auf *H Leading-Töne*. Die ersten beiden Akzenttöne der Äußerung haben die Form H+L*, wobei der zweite H Ton auf *sich* in Relation zum ersten H Ton auf *man* herabgestuft ist. Dieser *Downstep* (!H) wird durch den nächsten H *Leading-Ton* auf der Silbe *da* aufgehoben; die Tonhöhe erreicht durch *Upstep* (^H) wieder das Niveau des initialen H Tons der Phrase.

4.1.2 Rising Onglides

Selting (1995) geht davon aus, daß ein distinktives lokales Tonmuster auf einer akzentuierten Silbe beginnt und sich über die nachfolgenden unakzentuierten Silben erstreckt, und zwar bis unmittelbar vor den nächsten Akzent. Damit ist Seltings Akzentdomäne mit Abercrombies *Fuß* bzw. dem *Takt*, wie diese Einheit in der Literatur über das Deutsche oft genannt wurde (z.B. Kohler 1977, Pheby 1975), verwandt und mit Gussenhovens (1990) *Tonal Association Domain* identisch. Sie entspricht auch der Domäne des *Nuclear Tone* in der Britischen Schule

4.2. Phrasierungsebenen und Phrasenakzente

In Autosegmental-Metrischen Ansätzen des Deutschen wird – wie in großen Teilen der traditionellen Literatur – im allgemeinen nur eine Phrasierungsebene angenommen, die Intonationsphrase (z.B. Uhmann 1991, Grabe 1998). Die Ausnahme unter den Anhängern der AM Phonologie bildet Féry (1993), die zusätzlich (kleinere) Intermediärphrasen in ihrem Inventar hat, ebenso wie MAE_ToBI für das Englische und GToBI für das Deutsche. Unter den älteren, auditiv ausgerichteten Ansätzen können lediglich in von Essens Ansatz zwei Phrasierungsebenen ausgemacht werden. Seine Analyseeinheit ist die Rhetorische Phrase. Er unterscheidet zwei Arten einer Rhetorischen Phrase: eine größere mit Nukleus (bzw. *Schwerpunkt* in von Essens Terminologie) und eine kleinere ohne Nukleus. Wenn eine Äußerung mehr als eine Phrase umfaßt, enthält die letzte einen Nukleus. Dies wird in Beispiel (9) illustriert:



(9) Ich habe geTAN | was mir beFOHlen war.

(von Essen 1964:38)

Nach GToBI wird dieser Beispielsatz in zwei Intermediärphrasen (ip) aufgeteilt, die zusammen eine Intonationsphrase (IP) bilden, dargestellt in (10):

(10) [[Ich habe geTAN]_{ip} [was mir beFOHlen war]_{ip}]_{IP}

Obwohl die Analysen von Essens und GToBIs sehr ähnlich aussehen, gibt es doch einen wichtigen Unterschied: Während von Essen die beiden Phrasentypen *aufgrund ihrer Kontur* als eigene Kategorien betrachtet (*progre-diente Tonführung* auf der einen, *terminale und interrogative Tonführung* auf der anderen Seite), postuliert GToBI zwei verschiedene, hierarchisch strukturierte Domänen der Phrasierung, die unabhängig von bestimmten Tonhöhenverläufen sind. Die einzige Restriktion, die das Tonmuster betrifft, besteht in der Anzahl der Töne, die am Ende der beiden Phrasierungsebenen auftreten: Am Ende einer Intermediärphrase steht nur ein Grenzton (der Phrasenakzent), am Ende einer Intonationsphrase dagegen immer zwei (Phrasenakzent plus IP-Grenzton). Die Tonmuster an ip-Grenzen sind also in der Regel weniger komplex als an IP-Grenzen.

Féry verzichtet in ihrem Modell auf Phrasenakzente. Sie argumentiert, daß die beiden Funktionen, die der Phrasenakzent erfüllt, nämlich Beschreibung des Tonhöhenverlaufs zwischen dem Nukleus und der Intonationsphrasengrenze sowie Markierung der Grenze einer Intermediärphrase (1993:79), vom *Trailing-*

Ton des letzten Akzenttons der Phrase übernommen werden. Allerdings wird hierdurch eine klare Differenzierung zwischen einer Sequenz prä nuklearer bitonaler Akzenttöne (die Féry (1993:120) ausdrücklich erlaubt) und einer Folge mehrerer Intermediärphrasen, die jeweils auf einen bitonalen Tonakzent enden, praktisch unmöglich.

Die Frage nach der angemessenen Transkription für nuklear fallende Konturen ist immer noch umstritten: H^*+L , wie bei Féry und Uhmman, oder $H^* L$ wie in GToBI? Grice & Benz Müller (1998) haben belegt, daß der Punkt, an dem die F_0 die Grundlinie nach einem Gipfelakzent erreicht, variiert: Je weiter die nächste (postnuklear) betonte Silbe entfernt ist, desto später liegt dieser tiefe Zielpunkt. Es zeigte sich sogar, daß in 94% aller fallend-steigenden Konturen und in 91% aller fallenden Konturen der Zielpunkt genau mit der betonten Silbe korrespondiert (vgl. Abbildung 2). Dieses Ergebnis ist ein starkes Argument für eine Analyse der beschriebenen Konturen als $H^* L-H\%$ bzw. $H^* L-\%$ im Gegensatz zu $H^*+L H\%$ bzw. $H^*+L L\%$.

Wir wollen allerdings nicht ausschließen, daß es Konturen gibt, in denen der tiefe Zielpunkt in einem relativ konstanten Abstand vom Gipfelakzent auftritt und somit das L als Teil eines bitonalen Akzenttons H^*+L analysiert werden muß.¹⁶ Wir vermuten, daß es dialektale Unterschiede in der Position des Zielpunkts gibt, ein Phänomen, das auch in anderen Sprachen zu beobachten ist. So hat etwa die Fragekontur in verschiedenen Dialekten des Griechischen, Rumänischen und Ungarischen die gleiche tonale Struktur, aber die Assoziation des Phrasenakzents mit dem Text ist jeweils von Dialekt zu Dialekt unterschiedlich. Ferner hat Fitzpatrick-Cole (1999) in ihrer Untersuchung des Berner Schweizerdeutsch eine Assoziation von Grenzton (analysiert als IP-Grenzton) und lexikalisch betonter Silbe gefunden. Für das Standarddeutsche haben wir allerdings zurzeit keine ausreichenden Hinweise auf eine distinktive Opposition von H^*+L und $H^* L$. Daher betrachten wir die beiden Notationen solange als äquivalent, bis neue Forschungsergebnisse über die genaue Position des tiefen Zielpunkts in verschiedenen (semantischen und segmentellen) Kontexten vorliegen.

5. Fazit

GToBI ist ein Transkriptionsmodell, das als Hilfsmittel für die Forschung an der phonologischen Struktur deutscher Intonation entwickelt wurde. Sein Beschreibungsinventar ist deutlich größer als diejenigen anderer Autosegmental-Metrischer Modelle zur Intonation des Deutschen, wobei alle distinktiven Kategorien, die GToBI zur Verfügung stellt, durch unabhängige Studien, seien sie auditiv oder instrumental ausgerichtet, motiviert sind. Das umfangreichere Inventar ergibt sich aus dem Ziel dieses Transkriptionsmodells, so flexibel wie möglich zu sein, um alle relevanten, empirisch belegten Muster beschreiben zu

¹⁶ Dies könnte etwa in Strukturen mit engem Fokus oder Kontrast der Fall sein, wie es einige Beispiele in Uhmman (1991) suggerieren.

können, auch wenn sich Teile dieser Beschreibungen als redundant herausstellen sollten, weil sich einige Muster als von anderen ableitbar erweisen oder eher durch graduelle Faktoren zu erklären sind als durch kategorische.

Literatur

- Adriaens, L.M.H. (1991): Ein Modell deutscher Intonation: eine experimentell-phonetische Untersuchung nach den perzeptiv relevanten Grundfrequenzänderungen in vorgelesenem Text. Diss. Technische Universität Eindhoven.
- Altmann, H., A. Batliner & W. Oppenrieder, Hrsg. (1989): Zur Intonation von Modus und Fokus im Deutschen. Tübingen: Niemeyer.
- Arvaniti, A., D.R. Ladd & I. Mennen (1999): „What is a starred tone?“. In: M. Broe & J. Pierrehumbert, Hrsg. (1999): Papers in Laboratory Phonology V, Language Acquisition and the Lexicon. Cambridge: Cambridge University Press.
- Beckman, M.E. & G. Ayers-Elam (1997): Guide to ToBI Labelling. Text und begleitende Audio-Beispiele erhältlich unter http://ling.ohio-state.edu/Phonetics/E_ToBI/etobi_homepage.html.
- Beckman, M.E. & J. Hirschberg (1994): The ToBI Annotation Conventions. Manuskript und begleitendes Sprachmaterial. Ohio State University.
- Beckman, M.E., J. Hirschberg & S. Shattuck-Hufnagel (im Druck): „The original ToBI system and the evolution of the ToBI framework“. In: S. Jun, Hrsg. (im Druck): Prosodic Typology. Oxford University Press.
- Beckman, M.E. & J. Pierrehumbert (1986): „Intonational Structure in Japanese and English“. Phonology Yearbook 3, 255-309.
- Benzmüller, R. & M. Grice (1997): „Trainingsmaterialien zur Etikettierung deutscher Intonation mit GToBI“. Saarbrücken: Phonus 3, 9-34.
- Benzmüller, R. & M. Grice (1998): „The nuclear accentual fall in the intonation of Standard German“. In: ZAS Papers in Linguistics: Papers on the conference "The word as a phonetic unit". Berlin. 79-89.
- Crystal, D. (1969): Prosodic Systems and Intonation in English. Cambridge: Cambridge University Press.
- Duden (2000): Aussprachewörterbuch (Duden Band 6). Mannheim: Dudenverlag. (4., neu bearbeitete und aktualisierte Auflage).
- Essen, O. von (1964)²: Grundzüge der hochdeutschen Satzintonation. Ratingen: Henn (1. Auflage 1956).

- Féry, C. (1986): „Metrische Phonologie und Wortakzent im Deutschen“. *Studium Linguistik* 20, 16-43.
- Féry, C. (1993): *German Intonational Patterns*. Tübingen: Niemeyer.
- Fitzpatrick-Cole, J. (1999): „The Alpine Intonation of Bern Swiss German“. *Proceedings ICPHS99, San Francisco*. 941-944.
- Fox, A. (1984): *German intonation: An outline*. Oxford: Clarendon Press.
- Gibbon, D. (1976): *Perspectives of intonation analysis*. Bern: Lang.
- Gibbon, D. (1998): „Intonation in German“. In: D. Hirst & A. Di Cristo, Hrsg. (1998): *Intonation Systems. A survey of 20 languages*. Cambridge: Cambridge University Press. 78-95.
- Giegerich, H.J. (1985): *Metrical Phonology and Phonological Structure. German and English*. Cambridge: Cambridge University Press.
- Grabe, E. (1998): *Comparative Intonational Phonology: English and German*. (MPI Series in Psycholinguistics 7). Wageningen: Ponsen and Looijen.
- Grice, M. & R. Benzmüller (1995): „Transcription of German Intonation using ToBI-Tones – The Saarbrücken System“. *Saarbrücken: Phonus* 1, 33-51.
- Grice, M. & R. Benzmüller (1998): Tonal affiliation in German falls and fall-rises. Poster presented at the 5th Conference on Laboratory Phonology, York.
- Grice, M., D.R. Ladd & A. Arvaniti (2000): „On the place of phrase accents in intonational phonology“. *Phonology* 17.2, 143-185.
- Grice, M., M. Reyelt, R. Benzmüller, J. Mayer & A. Batliner (1996): „Consistency in Transcription and Labelling of German Intonation with GToBI“. *Proceedings of the Fourth International Conference on Spoken Language Processing, Philadelphia*. 1716-1719.
- Gussenhoven, C. (1984): *On the Grammar and Semantics of Sentence Accents*. Dordrecht: Foris.
- Gussenhoven, C. (1990): „Tonal association domains and the prosodic hierarchy in English.“ In: S. Ramsaran, Hrsg. (1990): *Studies in the pronunciation of English*. London: Routledge.
- Halliday, M.A.K. (1967): *Intonation and Grammar in British English*. The Hague: Mouton.
- Isačenko, A.V. & H.J. Schädlich (1966). „Untersuchungen über die deutsche Satzintonation“. *Studia Grammatica* VII, 7-64.
- Kerckhove, M. van de (1948): „Intonationssystem einer Mundart“. *Zeitschrift für Phonetik und Allgemeine Sprachwissenschaft*, 52-65.

- Kohler, K.J. (1977): Einführung in die Phonetik des Deutschen. (Grundlagen der Germanistik 20). Berlin: Schmidt. (2., neu bearbeitete Auflage 1995).
- Kohler, K.J. (1991): „Terminal intonation patterns in single-accent utterances of German: Phonetics, phonology, and semantics“. Kiel: AIPUK 25, 115-185.
- Ladd, D.R., K.E.A. Silverman, F. Tolkmitt, G. Bergman & K.R. Scherer (1985): „Evidence for the independent function of intonation contour type, voice quality, and F0 range in signaling speaker affect“. *Journal of the Acoustical Society of America* 78, 435-444.
- Ladd, D.R. (1996): *Intonational Phonology*. Cambridge: Cambridge University Press.
- Moulton, W.G. (1962): *The Sounds of English and German*. Chicago: University of Chicago Press.
- Mozziconacci, S.J.L. (1998): *Speech variability and emotion: Production and perception*. Diss. Technische Universität Eindhoven.
- Peters, J. (2001a): *Postnukleare Tonhöhengipfel in der Vorderpfalz und in Mannheim*. Ms., Universität Potsdam.
- Peters, J. (2001b): *Frageintonation in der Pfalz. Eine Reanalyse des Güntherodt-Korpus*. Ms., Universität Potsdam.
- Pheby, J. (1975): *Intonation und Grammatik im Deutschen*. Berlin: Akademie-Verlag.
- Pheby, J. (1984)²: „Phonologie: Intonation“. (Kap. 6). In: K.E. Heidolph et al., Hrsg. (1984): *Grundzüge einer deutschen Grammatik*. Berlin: Akademie-Verlag. 839-897. (1. Auflage 1980).
- Pierrehumbert, J. (1980): *The Phonology and Phonetics of English Intonation*. Diss. MIT. Bloomington: Indiana University Linguistics Club.
- Pike, K.L. (1945): *The intonation of American English*. Ann Arbor: University of Michigan Press.
- Reyelt, M., M. Grice, R. Benzmüller, J. Mayer & A. Batliner (1996): „Prosodische Etikettierung des Deutschen mit ToBI“. In: D. Gibbon, Hrsg. (1996): *Natural Language and Speech Technology, Results of the third KONVENS conference*. Berlin, New York: Mouton de Gruyter. 144-155.
- Scherer, K.R., D.R. Ladd & K.E.A. Silverman (1984): „Vocal cues to speaker affect: Testing two models“. *Journal of the Acoustical Society of America* 76 (5), 1346-1356.
- Selting, M. (1995): *Prosodie im Gespräch. Aspekte einer interaktionalen Phonologie der Konversation*. Tübingen: Niemeyer.
- Shattuck-Hufnagel, S., M. Ostendorf & K. Ross (1994): „Stress shift and early pitch accent placement in lexical items in American English“. *Journal of Phonetics* 22, 357-388.

Trager, G.L. & H.L. Smith (1951): *An outline of English structure*. Norman, Oklahoma: Battenburg Press. (neu bearbeitete Auflage 1957, Washington, American Council of Learned Societies).

Truckenbrodt, H. (erscheint): „Upstep and Register Levels“. *Phonology*.

Uhmann, S. (1991): *Fokusphonologie. Eine Analyse deutscher Intonationskonturen im Rahmen der nicht-linearen Phonologie*. Tübingen: Niemeyer.

Wunderlich, D. (1988): „Der Ton macht die Melodie - Zur Phonologie der Intonation des Deutschen“. In: H. Altmann, Hrsg. (1988): *Intonationsforschungen*. Tübingen: Niemeyer. 1-40.

Danksagung

Ralf Benzmüller war maßgeblich an der Entwicklung der Ideen, die diesem Aufsatz zugrunde liegen, beteiligt. Ihm gilt unser besonderer Dank. Außerdem möchten wir Bistra Andreeva, Bill Barry, Caroline Féry, Jörg Peters, Jürgen Trouvain, Hubert Truckenbrodt sowie zwei anonymen Gutachtern für hilfreiche Kommentare herzlich danken.