

Annotationsrichtlinien DIMA

Frank Kügler & Stefan Baumann

1. Ziel

Das Ziel dieser Richtlinien besteht darin, ein Konsenssystem zur Transkription der Intonation des Deutschen vorzuschlagen: "Deutsche Intonation – Modellierung und Annotation" (DIMA). Es basiert auf den Grundannahmen des autosegmental-metrischen Ansatzes zur Analyse von Intonation (Pierrehumbert, 1980; Beckman & Pierrehumbert, 1986; Gussenhoven, 2004; Ladd, 2008). Dieser will die phonologisch relevanten prosodischen Ereignisse von Äußerungen ermitteln, die entweder der Hervorhebung oder der Abgrenzung sprachlicher Einheiten dienen, und stützt sich dabei auf dezidiert phonetische Kriterien (*“phonetically informed phonological transcription”*). Mit dieser Integration phonetischer und phonologischer Kriterien soll eine Vergleichbarkeit von Daten/Korpora unterschiedlicher Arbeitsgruppen im deutschsprachigen Raum gewährleistet werden.

Zwei Kriterien soll das System erfüllen: (i) Es soll leicht erlernbar bzw. trainierbar und damit transparent sein sowie eine hohe Reliabilität zwischen den Annotatoren erzielen. (ii) Das Basistranskript soll grundsätzlich (automatisch) in die phonologischen Systeme der verschiedenen Intonationsmodelle des Deutschen (z.B. Féry, 1993, 2012; Mayer, 1995; Grice, Baumann & Benz Müller, 2005; Peters, 2005, 2014) übersetzbar sein, woraus sich ergibt, dass die (eigentliche) phonologische Interpretation der Daten zu einem späteren Zeitpunkt als die Annotation erfolgt. Somit ist die angestrebte Konsenstranskription als eine *phonetisch orientierte* Repräsentation einer intonatorischen Oberflächenkontur zu betrachten, die jedoch den *phonologischen Kern* der Kontur abbildet.

2. Voraussetzungen und Annahmen

2.1 Voraussetzungen

Für die Annotation von Intonation nach DIMA wird folgendes Grundwissen vorausgesetzt:

- (1) Kenntnis des autosegmental-metrischen Ansatzes der Intonationsanalyse (vgl. Gussenhoven 2004, Ladd 2008, Peters 2014).
- (2) Kenntnis mikroprosodischer Einflüsse von Vokalen und Konsonanten auf die Grundfrequenz (vgl. Möbius, Zimmermann & Hess 1987, Ohala & Eukel 1987, Kohler 1990, Whalen & Levitt 1995, Kap. 3.4.2 in Grice & Baumann 2016).
- (3) Erfahrungen mit der Perzeption von Tönen, Akzenten und Intonationsphrasen (vgl. Grundlagen der Prosodie auf der GToBI-Webseite, http://www.gtobi.uni-koeln.de/x_grundlagen_prosodie.html).

2.2 Annahmen

Die Durchführung der Annotation wird hier in *Praat* (Boersma & Weenink, 2017) illustriert. Für die Annotation gesprochener Sprache nehmen wir drei Ebenen der prosodischen Annotation sowie eine zusätzliche Ebene für Kommentare an. Wir empfehlen, auf diesen vier Ebenen *point tiers* zu verwenden. Die drei Ebenen der Prosodie kennzeichnen jeweils unabhängig voneinander Phrasengrenzen, Töne und die relative Prominenz von Wörtern innerhalb einer Phrase. Darüber hinaus soll die Textebene mindestens eine Wortebene und eine Silbenebene beinhalten (vgl. Abbildung 1). Die Annotation der Tonebene und der Prominenzebene erfolgen grundsätzlich mit Bezug auf Silben.

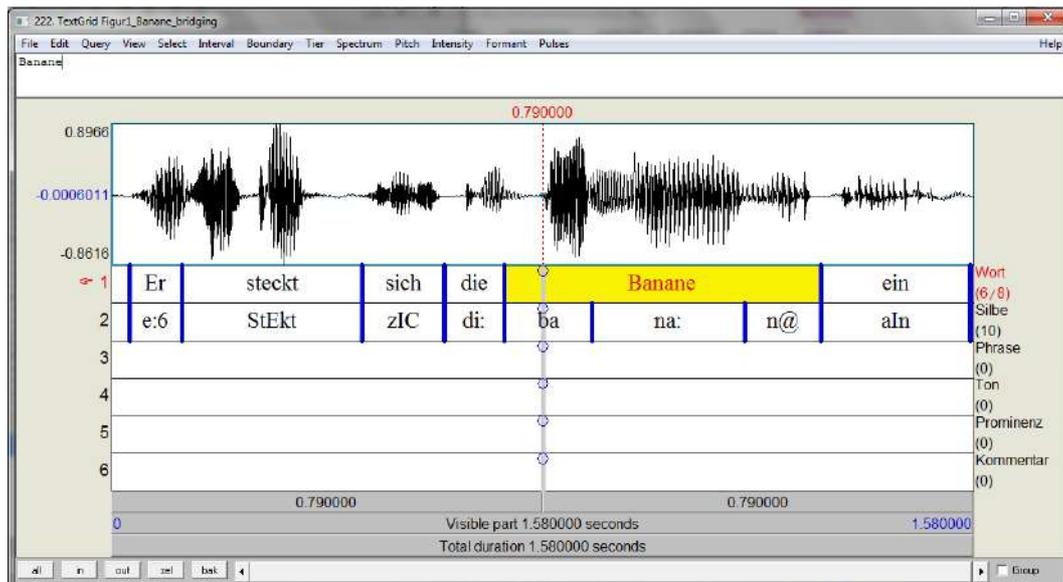


Abbildung 1. Ebenen der prosodischen Annotation nach DIMA. Die einzelnen Ebenen sind von oben nach unten folgendermaßen benannt: Wort, Silbe, Phrase, Ton, Prominenz, Kommentar. Die Silbenebene ist hier in SAMPA (Wells 1997) transkribiert; dies ist keine notwendige Voraussetzung für DIMA.

Jede Ebene wird unabhängig von den anderen Ebenen annotiert. Die Unabhängigkeit der Ebenen hat den Vorteil, dass z.B. prosodische Grenzen und Prominenzen nicht in Bezug auf Töne annotiert werden müssen, sondern allein aufgrund ihrer phonetischen Eigenschaften und perceptiven Wahrnehmung als Grenze oder Prominenz. Grenzen müssen mit Tönen einhergehen, die allerdings erst in einem späteren Schritt annotiert werden. Ein Prominenzlabel dagegen, egal welcher Stufe, muss kein korrespondierendes Tonlabel zugewiesen bekommen (vgl. unten „Tonebene“). Auf jeder Ebene der Annotation gibt es die Möglichkeit, Unsicherheiten mit einem „?“ links vor dem jeweiligen Symbol zu kennzeichnen. Generell werden Diakritika, die ein Label modifizieren, unmittelbar *links vor* das betreffende Label gesetzt.

Für die prosodische Annotation wird das Symbolinventar gemäß Tabelle 1 festgelegt. Die Symbole sind im Wesentlichen dem klassischen ToBI System (Beckman & Ayers-Elam 1997) entlehnt. Die Verwendung der einzelnen Symbole wird im Folgenden erläutert und anhand von Beispielen illustriert. Es ist zu beachten, dass die Symbole per se keine *Bedeutung* tragen, wie z.B. die Markierung von Fokus oder Kontrast. Die DIMA-Annotation dient aber als geeignetes Hilfsmittel für die Erforschung der Relation zwischen intonatorischer Form sprachlicher Einheiten und ihrer semantisch-pragmatischen Bedeutung im Diskurs.

Tabelle 1. Symbole für die prosodische Transkription auf den vier Ebenen.

Ebene	Symbole
Phrase	% -
Diakritika	! ^ & ?
Ton	H* L* H L
Diakritika	! ^ < > ?
Prominenz	1 2 3
Diakritikum	?
Kommentar	z.B. alternative Annotation

2.3 Praat-Einstellungen und praktische Hinweise für die Annotation

Für die Darstellung der Tonhöhe bei der Annotation in *Praat* sollen *speckles* verwendet, ein Tonhöhenumfang in Abhängigkeit vom Sprechergeschlecht gewählt und die gewählten F0-Einstellungen für denselben Sprecher beibehalten werden. Praktische Hinweise für die Einstellungen in Praat sind auf der DIMA-Webseite zu finden: http://dima.uni-koeln.de/?page_id=280. Zudem werden dort Hinweise zum Vorgehen bei Zweifelsfällen gegeben. Zum einen betrifft dies die Möglichkeit der Resynthese einer Audiodatei, um ein Minimalpaar für die Identifizierung möglicher funktionaler Unterschiede einer Kontur zu erzeugen. Zum anderen wird die Möglichkeit des verlangsamten Abspielens einer Audiodatei erläutert, wodurch der Höreindruck für das Auftreten der Töne auf den entsprechenden Silben geschärft werden kann.

Während der Annotation werden die folgenden Vorgehensweisen empfohlen:

- 1) Zur Identifikation der Phrasengrenzen und der Prominenzen soll die *F0-Kontur ausgeblendet* werden, um sich allein auf den Höreindruck konzentrieren zu können. Idealerweise soll die Bestimmung der Tonakzente ebenfalls zunächst ohne F0-Kontur erfolgen. Die Alignierung der Tonlabel kann dann in einem zweiten Schritt mit eingeblandeter Tonhöhe durchgeführt werden.
- 2) Zur Identifikation der Tonhöhe von Akzenttönen wird vorgeschlagen, ein Drei-Silben-Fenster zu nutzen, d.h. jeweils eine Silbe vor und nach der betreffenden Silbe. Um die Tonhöhe von Zielpunkten einschätzen zu können, muss sie in Relation zu den umgebenden Silben betrachtet bzw. angehört werden (wie bei Prominenzen).

3. Annotationsprozess

Die prosodische Annotation basiert grundsätzlich auf den relationalen *Höreindrücken* des Transkribierenden, die in phonologische Kategorien (hier: das Inventar in Tabelle 1) klassifiziert werden. Die Annotation erfolgt in mehreren Schritten auf drei eigenständigen Ebenen. Grundsätzlich wird empfohlen, zuerst die Phrasenebene, als zweites die Prominenzebene und als letztes die Tonebene zu annotieren. Dieses Vorgehen lenkt die Aufmerksamkeit auf einzelne Aspekte des komplexen prosodischen Ereignisses und erleichtert somit z.B. die Entscheidung über die An- bzw. Abwesenheit einer Grenze oder der Prominenzstufe unabhängig vom tonalen Wert einer Silbe. Außerdem soll auf Phrasenebene begonnen werden, weil die Einteilung in prosodische Phrasen mit starker Grenze die Domäne für die weitere Annotation definiert. Dies hat zur Folge, dass weitere Grenzmarkierungen, Prominenzen und Töne immer *phrasenweise* annotiert werden.

3.1 Annotation der Phrasenebene

Grundlegendes

Auf der Phrasenebene werden zwei Phrasentypen unterschieden:

- prosodische Phrase mit starker Grenze: %
- prosodische Phrase mit schwacher Grenze: -

Gemäß der prosodischen Hierarchie (Selkirk, 1984; Nespor & Vogel, 1986) nehmen wir an, dass eine prosodische Phrase mit schwacher Grenze von einer prosodischen Phrase mit starker Grenze dominiert wird. Grundsätzlich postulieren wir damit zwei Ebenen der Phrasierung, die jedoch so theorie-neutral wie möglich über die relative Stärke der Grenze ausgedrückt werden. Phonetisch-perzeptive Kriterien für eine Grenze sind eine Pause, phrasenfinale Längung, eine phrasenfinale tonale Bewegung, Tonhöhen-Reset sowie segmentale Phänomene wie Laryngalisierung oder *creaky voice*. Je nach Ausprägung dieser Kriterien wird eine starke oder schwache Grenze annotiert.

Annotation der Phrasen

Grundsätzlich werden der Beginn und das Ende einer prosodischen Phrase annotiert. Dies gilt für beide Phrasentypen. Wenn eine Pause zwischen den Phrasen vorhanden ist, werden zwei Phrasenlabel, jeweils eins für das Ende der ersten und den Beginn der zweiten Phrase annotiert. Wenn keine Pause vorhanden ist, wird nur *ein* Phrasenlabel annotiert, das zugleich das Ende der ersten und den Beginn der zweiten Phrase markiert. Abbildung 2 illustriert das Vorkommen zweier Phrasentypen in einer Äußerung. Im Annotationsprozess wird also zunächst eine Phrase mit starker Grenze identifiziert und deren linke und rechte Grenze annotiert. Innerhalb dieser Phrase werden dann gegebenenfalls Phrasengrenzen mit schwacher Grenze annotiert. Das heißt, dass eine Phrase mit schwacher Phrasengrenze immer innerhalb einer Phrase mit starker Phrasengrenze auftritt.

Hinweis zur Annotation der Phrasenebene:

Die Entscheidung, ob eine Grenze als stark oder schwach zu bewerten ist, ist sprecherabhängig, d.h. für jeden Sprecher muss ein eigenes perzeptives Bezugssystem entwickelt werden, innerhalb dessen ein Annotator die relative Stärke einer Grenze beurteilen kann. Dies setzt (idealerweise) voraus, dass genug Sprachmaterial eines Sprechers zur Verfügung steht.

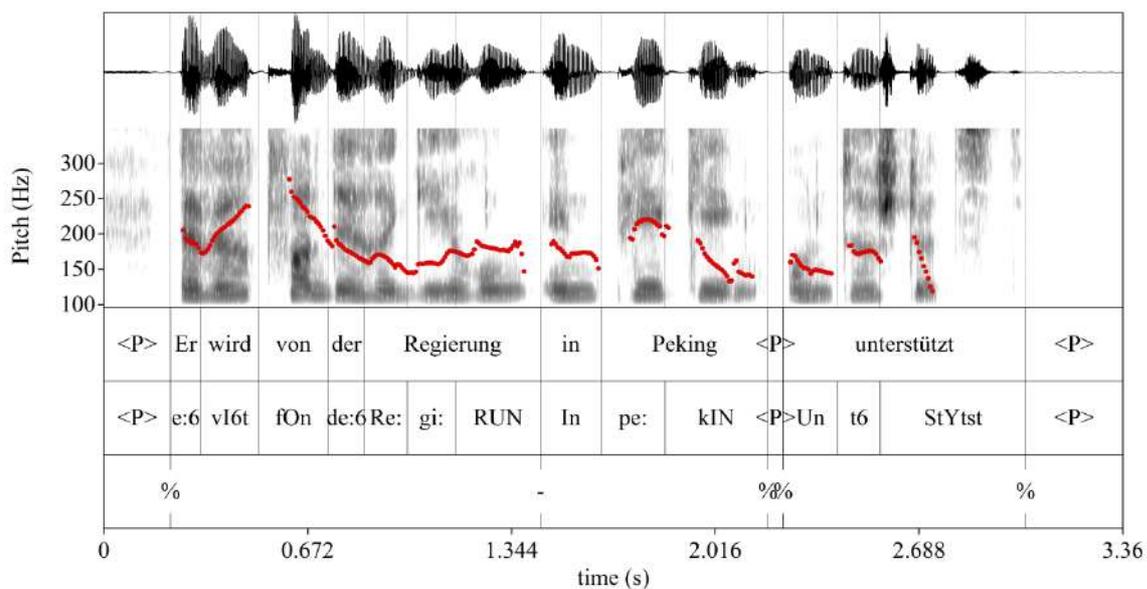


Abbildung 2. Illustration von starken und schwachen Phrasengrenzen, Beispiel „Peking“. Starke Grenzen werden hier jeweils vor und nach Pausen markiert, die schwache Phrasengrenze tritt an einer Stelle ohne Pause, aber mit deutlicher glottaler Grenzmarkierung durch den Glottalverschluss vor „in“ auf, und wird nur einfach markiert.

Diakritika: Für Unterschiede im Tonhöhenumfang (*pitch range*) zwischen zwei oder mehreren Phrasen werden zwei Arten von Diakritika verwendet: Downstep „!“ und Upstep „^“ (vgl. van den Berg, Gussenhoven & Rietveld 1992, Ladd 1988).

In DIMA wird eine Phrase mit den Diakritika für Downstep und Upstep versehen, wenn das Tonhöhenregister verschoben oder durch eine unterschiedliche Tonhöhenspanne (reduzierte bzw. erweiterte *pitch span*, Ladd 2008) gekennzeichnet ist. Nur der Beginn des Bruchs im Register – Downstep oder Upstep – wird markiert, und zwar direkt links vom Phrasenlabel (vgl. Abb. 3a und 3b, sowie das nachfolgende Beispiel in Abb. 4). Innerhalb einer vorangehenden

Phrase kann also beispielsweise die absolute F0 eines H* niedriger sein als die F0 eines L* in einer nachfolgenden heraufgestuften Phrase. An der rechten Phrasengrenze wird kein Diakritikum gesetzt. Bleibt das Tonhöhenregister verschoben, gibt es keinen Tonhöhenregisterbruch, der markiert werden muss (vgl. Abb. 5 für phrasalen Downstep und Abb. 6 für phrasalen Upstep). Kehrt das Tonhöhenregister hingegen zu seinem Ausgangslevel zurück, kann dieser Tonhöhenbruch durch ein Diakritikum gekennzeichnet werden, sofern diese Verschiebung des Tonhöhenregisters perzeptuell auffällig ist.

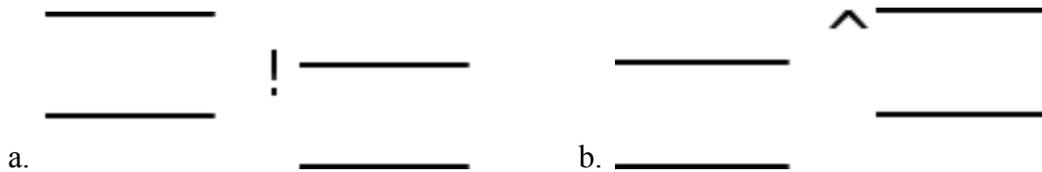


Abbildung 3. a. Downstep des Tonhöhenregisters auf Phrasenebene; b. Upstep des Tonhöhenregisters auf Phrasenebene.

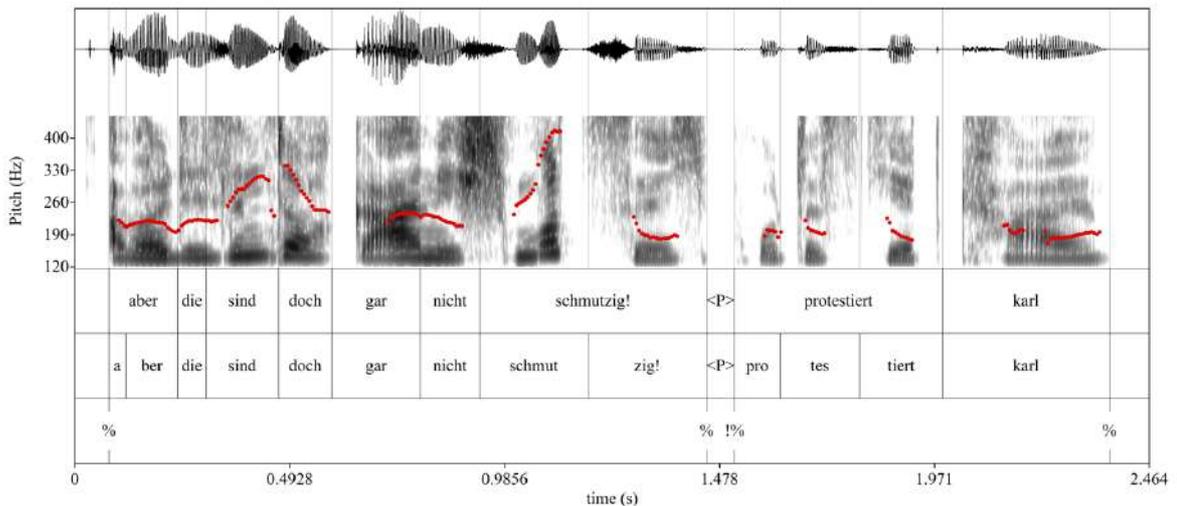


Abbildung 4. Illustration eines phrasalen Downsteps; zu Beginn der reduzierten Phrase (*protestiert Karl*) wird ein „!“ links vom Phrasenlabel annotiert.

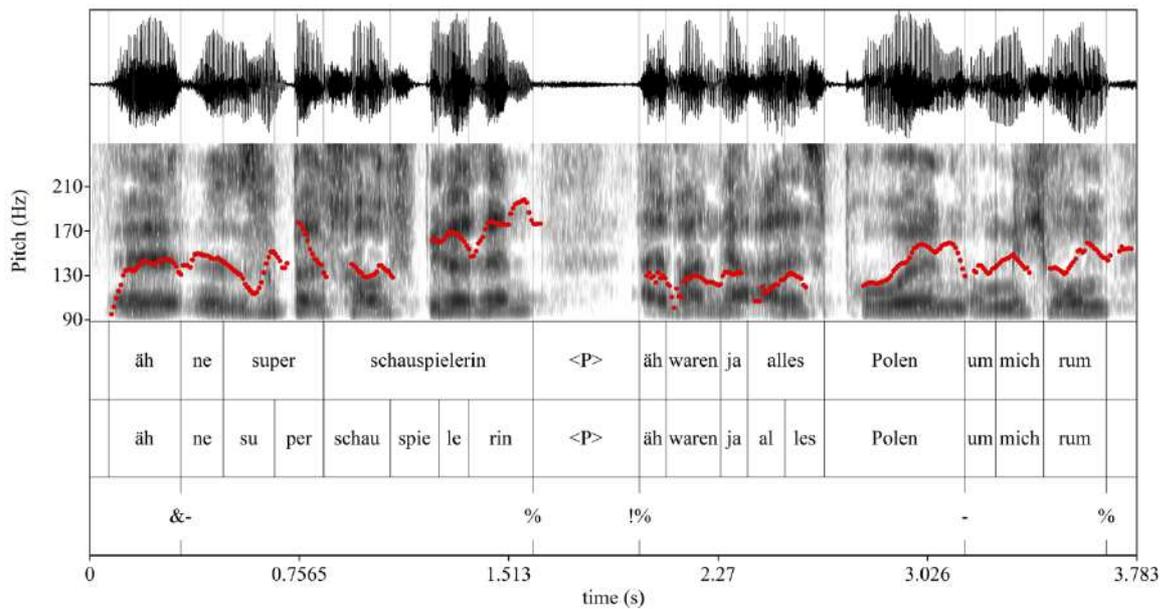


Abbildung 5. Illustration eines phrasalen Downsteps, der sich über zwei Phrasen erstreckt; zu Beginn der ersten reduzierten Phrase wird ein „!“ links vom Phrasenlabel annotiert; die folgende Phrase verläuft in etwa im selben abgesenkten Tonhöhenregister und wird nicht weiter markiert.

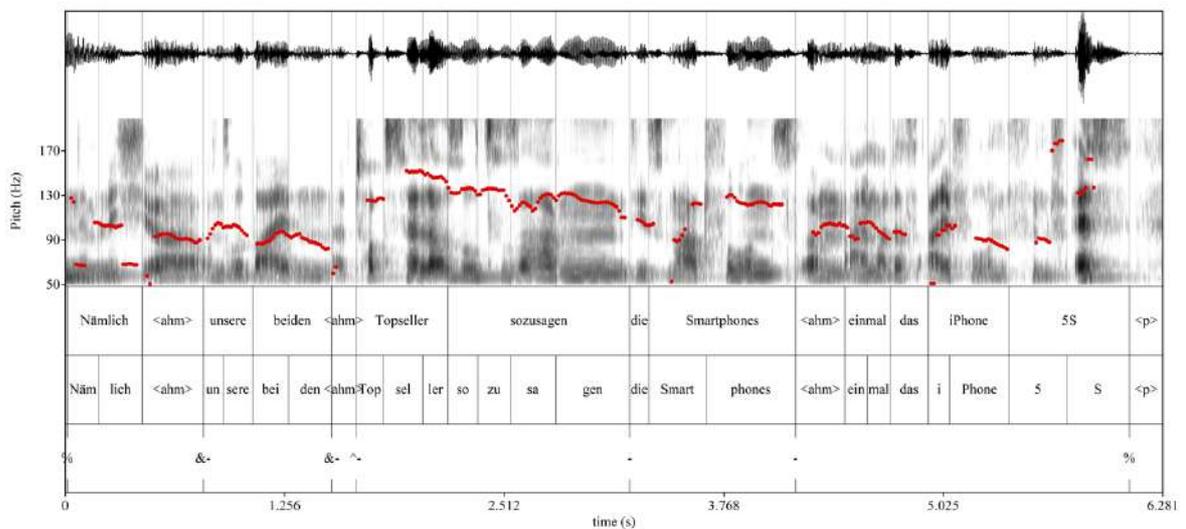


Abbildung 6. Illustration eines phrasalen Upsteps, der sich über zwei Phrasen erstreckt; zu Beginn der ersten heraufgestuften Phrase wird ein „^“ links vom Phrasenlabel annotiert.

Neben der weitgehend linguistisch determinierten Einteilung in kleine und große prosodische Phrasen können in gesprochener, insbesondere spontaner Sprache Disfluenzen vorkommen, die sich in eine Phrase integrieren oder eine Phrase unterbrechen, z.B. Häsitationen wie „ähm“. Für die Markierung von Disfluenzen wird das sogenannte ‚Kaufmanns-Und‘ (&) verwendet. Dieses Label wird als Diakritikum links an das jeweilige Symbol der Phrasengrenze gesetzt und kann sowohl am Anfang als auch Ende der Konstituente vorkommen. Abbildung 7 illustriert die Verwendung des Diakritikums.

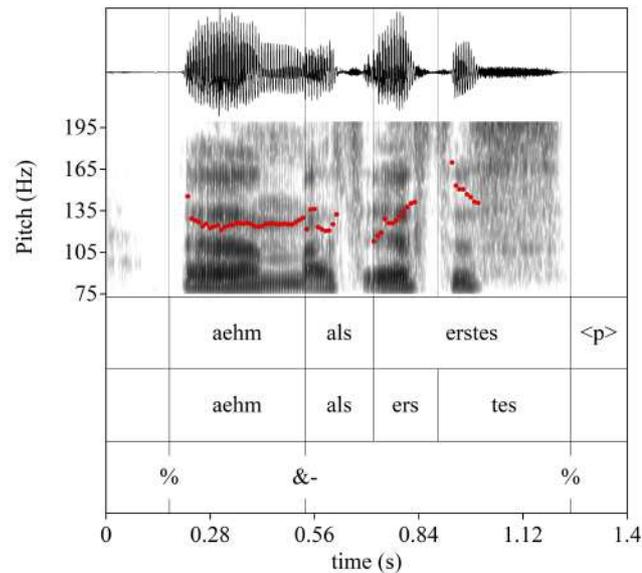


Abbildung 7. Illustration einer Disfluenz; an der prosodischen Unterbrechung wird das Label „&“ als Diakritikum zur prosodischen Phrase mit schwacher Grenze auf Phrasenebene annotiert.

3.2 Tonebene

Grundlegendes

Die Tonebene differenziert zwischen Akzenttönen und Nicht-Akzenttönen. Akzenttöne sind durch einen Stern gekennzeichnet, der die *phonologische Assoziation* eines Tons mit einer Prominenz (mindestens Stufe 1) anzeigt. Sie sind somit Ausdruck der Hervorhebungsfunktion innerhalb einer Äußerung und werden bei der Annotation mit der prominenztragenden Silbe *synchronisiert*. Nicht-Akzenttöne markieren hingegen Ziel- bzw. Wendepunkte im Tonverlauf, die nicht notwendigerweise mit einer Prominenz einhergehen.

Für die Beantwortung der Frage "Wann annotiere ich einen Stern?" ist die *Wahrnehmung* durch den Annotator entscheidend: Nimmt er/sie einen vollen Tonakzent wahr (i.d.R. gekennzeichnet durch eine Tonbewegung, erhöhte Dauer und Intensität sowie eine unreduzierte Vokalqualität), ist dem Ton ein Stern beizufügen und innerhalb der Grenzen der prominenten Silbe zu setzen (s.u.). Er markiert also – im abstrakten wie konkreten Sinn – den *phonologischen Kern* des Akzents und seiner Zugehörigkeit zum hervorgehobenen Text.

Kennzeichnet ein Ton einen Wendepunkt in der Kontur, der nicht mit der Wahrnehmung einer akzentuierten Silbe bzw. eines akzentuierten Wortes einhergeht, wird kein Stern gelabelt. Die Position eines solchen Nicht-Akzenttons richtet sich nach seinem Extremwert (bzw. dem genauen Zielpunkt) in der F0-Kurve. Eine phonologische Zuordnung zu einem Tonakzent ist hier nicht maßgeblich, sondern wird ggf. zu einem späteren Zeitpunkt durchgeführt (z.B. bei der Übertragung in ein bestimmtes Modell).

Es gilt also: **Der Stern zeigt den Kern, der bloße Ton die Position!**

Grundsätzlich wird, wie im klassischen Ansatz nach Pierrehumbert (1980), nur zwischen zwei phonologischen Tonebenen unterschieden, H und L, die jeweils relativ zueinander als hoch und tief interpretiert werden; d.h. bezüglich der absoluten F0-Werte kann ein Tiefton höher sein als ein Hochton in einer Äußerung, jedoch nicht in unmittelbarer Nachbarschaft eines Hochtons.

Annotation der Töne

Begonnen wird mit den Grenztönen (sowohl am Anfang als auch am Ende von Phrasen mit starker und schwacher Grenze). Darauf folgen nacheinander die Akzenttöne, in deren linker und rechter Umgebung Nicht-Akzenttöne gelabelt werden können. Dieses iterative Vorgehen gewährleistet, dass Töne relativ zueinander bewertet werden, was z.B. Downstep- und Upstep-Relationen einschließt.

○ Grenztöne

Als Orientierung dienen hier die Grenzen der Phrasenebene; Tonlabel unterhalb des Grenzlabels platzieren → Phrase %
Ton L

Hinweise zu Grenztönen:

1. Grundsätzlich erhalten Grenzlabel einen korrespondierenden Ton auf der Tonebene. Repräsentiert ein Grenzlabel sowohl den Anfang als auch das Ende einer Phrase, können zwei Tonlabel vergeben werden, aber nur wenn diese sich in ihrem Tonwert unterscheiden (vgl. z.B. Abb. 14 nach *Banane*); andernfalls wird nur ein Tonwert gelabelt.
2. In der Regel beginnt eine große prosodische Phrase mit einem tiefen Grenzton „L“. Dieser Wert wird als Standardwert verstanden und repräsentiert tiefe bis mittlere Phrasenanfänge. Wenn eine Phrase mit einer außergewöhnlich hohen F0 beginnt, wird ein hoher linker Grenzton „H“ festgesetzt – es sei denn, die Kontur kann durch einen hohen Ton (z.B. Akzentton) in der ersten Silbe der Phrase beschrieben werden.
3. Werden Disfluenzlabel als Grenzton-Diakritika vergeben, korrespondieren auch diese mit einem Tonlabel, weil Grenzlabel grundsätzlich mit einer Tonannotation einhergehen.

Die nächsten beiden Schritte wiederholen sich bis zum Ende der Phrase:

○ Akzenttöne

Akzenttöne sind durch den Stern „*“ gekennzeichnet (H* / L*)

Als Orientierung dienen hier die Label der Prominenzebene – allerdings muss nicht jede Prominenz mit einem Akzentton einhergehen! Umgekehrt ist aber ein Akzentton mindestens mit Prominenzstufe 1 verknüpft.

Ein Akzentton wird immer in der betonten (stressed) Silbe eines Wortes annotiert

→ H* / L*

Das Tonlabel wird auf den Wendepunkt in der F0-Kurve (Maximum, Minimum) platziert, wenn dieser *innerhalb der prominenten Silbe* auftritt. Akzenttöne als Teil einer Plateaukontur haben nicht immer einen Wendepunkt; die Annotation erfolgt dann in der Mitte des Vokals der wortbetonten Silbe. Das Tonlabel wird in die Mitte des Vokals platziert, wenn das Maximum/ Minimum außerhalb der prominenten (wortbetonten) Silbe auftritt.

Diakritika I: Für den Fall, dass der tonale Zielpunkt eines Akzenttons nicht in der prominenten Silbe liegt, werden Diakritika an dem Zielpunkt in der vorangehenden oder nachfolgenden Silbe gelabelt. Die Diakritika werden an dem phonetischen Minimum/Maximum der F0 außerhalb der prominenten (wortbetonten) Silbe platziert (s. Abb. 8, „*wird* – H*“ und „*von* – <“).

Dabei zeigt das Label auf den zugehörigen Ton:

Zielpunkt erscheint in der nachfolgenden Silbe und signalisiert Zugehörigkeit zum *vorhergehenden* Ton → H* <
 Zielpunkt erscheint in der vorhergehenden Silbe und signalisiert Zugehörigkeit zum *nachfolgenden* Ton → > H*

Diakritika II: Downstep „!“ und Upstep „^“ werden links vor dem Akzentton notiert. Downstep und Upstep sind relational zum vorangehenden Ton gleicher Qualität zu interpretieren → z.B. !H* / !H / ^H* / ^H

o Nicht-Akzenttöne

Die Nicht-Akzenttöne erscheinen in ihrem Tonwert (H / L)

Zur Identifizierung dieser Töne dienen die Wendepunkte in der F0-Kurve vor und nach einem Akzentton → H / L (ggf. inklusive ! oder ^)

Hinweis zu Akzent- und Nicht-Akzenttönen:

1. Es werden keine mikroprosodischen Perturbationen annotiert; Zielpunkte sollen ggf. unmittelbar vor (siehe 2.) bzw. nach (siehe 3.) Perturbationen gesetzt werden. Abbildung 9 illustriert eine mikroprosodische Perturbation zwischen dem akzenttragenden Wort (*machst*) und dem nachfolgenden Wort (*du*). Der Akzentgipfel wird in der akzenttragenden Silbe annotiert, obwohl die F0 zu Beginn der folgenden Silbe *du* höher ist. Die höhere F0 wird nur durch die stimmlosen Obstruenten hervorgerufen und deshalb nicht als Akzentgipfel in der nachfolgenden Silbe annotiert.
2. Wenn man einen tiefen Wendepunkt vor einem hohen Akzentton wahrnimmt (z.B. *nicht schmutzig*, Abb. 10), wird der tiefe Nicht-Akzentton an die tiefste F0 in oder vor der akzenttontragenden Silbe gesetzt. Im Falle von nicht-sonoranten Segmenten kann dieser Tiefpunkt auch deutlich vor der Akzentsilbe liegen. Ein L-Ton wird auch dann eingefügt, wenn keine Interpolation zwischen linker Grenze und hohem Akzentton vorliegt (*waschen* in Abb. 9, *sind* in Abb. 10).
3. Nach einem hohen Akzentton (und besonders häufig vor einer rechten Phrasengrenze) wird ein tiefer Nicht-Akzentton annotiert, wenn die F0 ein hör- und sichtbares Minimum bildet, d.h. wenn die F0 nicht zwischen hohem Akzentton und tiefer Grenze interpoliert. Abbildung 9 zeigt ein Beispiel für mehrere Fallbewegungen (mit jeweils deutlichem Knick) am Ende einer Phrase (*Gardinen, waschen, sonst*; siehe auch *schmutzig* in Abb. 10).

Abbildung 8 illustriert das Zusammenspiel von Akzenttönen und Nicht-Akzenttönen. Tiefe Töne erscheinen in diesem Beispiel zu Beginn und am Ende von Phrasen mit starker prosodischer Grenze. Der erste Akzentton (auf *wird*) ist perzeptiv ein hoher Ton (H*), dessen F0-Maximum erst in der nachfolgenden Silbe (*von*) erreicht wird, ausgedrückt durch „<“. Außerdem stehen zwei spätere Hochtöne in einer Downstep-Relation zu einem vorangehenden H. Dies betrifft sowohl Akzenttöne (auf *-stützt*) als auch Nicht-Akzenttöne (an der schwachen Grenze nach *Regierung*). Die Phrase nach der schwachen Grenze beginnt im unteren bis mittleren Stimmumfang der Sprecherin, so dass als Standardwert ein L-Ton annotiert wird.

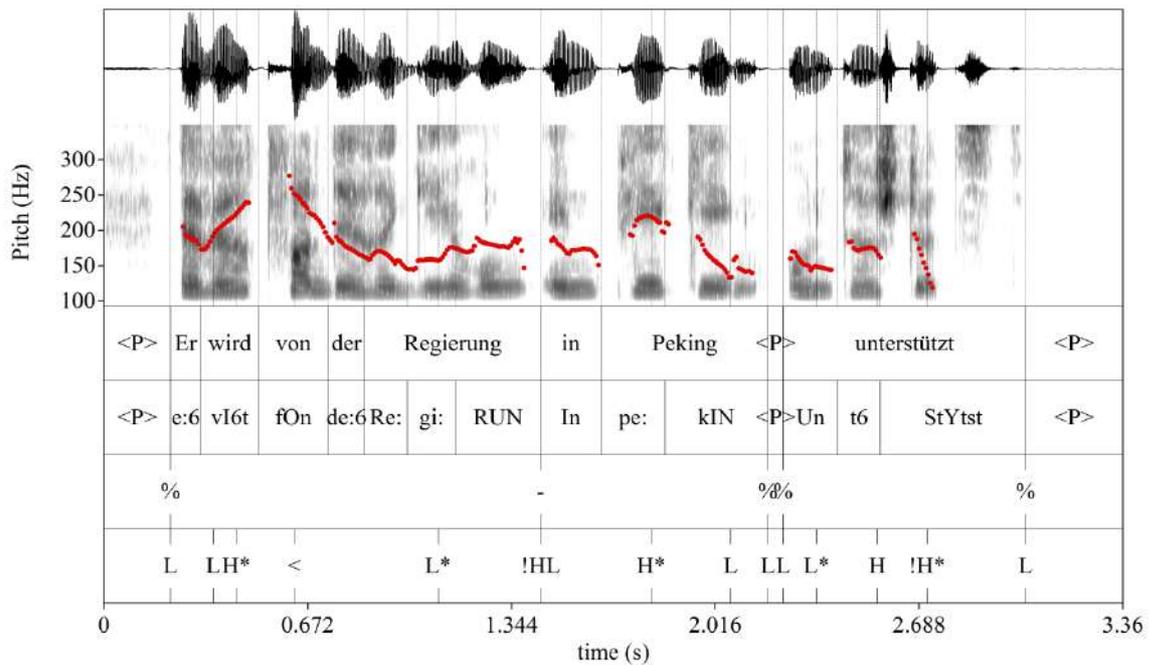


Abbildung 8. Illustration der Tonebene, Beispiel „Peking“. Akzenttöne werden mit „*“ gekennzeichnet. Nicht-Akzenttöne werden an phonologischen Grenzen gesetzt und erscheinen an F0-Wendepunkten zwischen Grenzen und Akzenttönen sowie zwischen Akzenttönen.

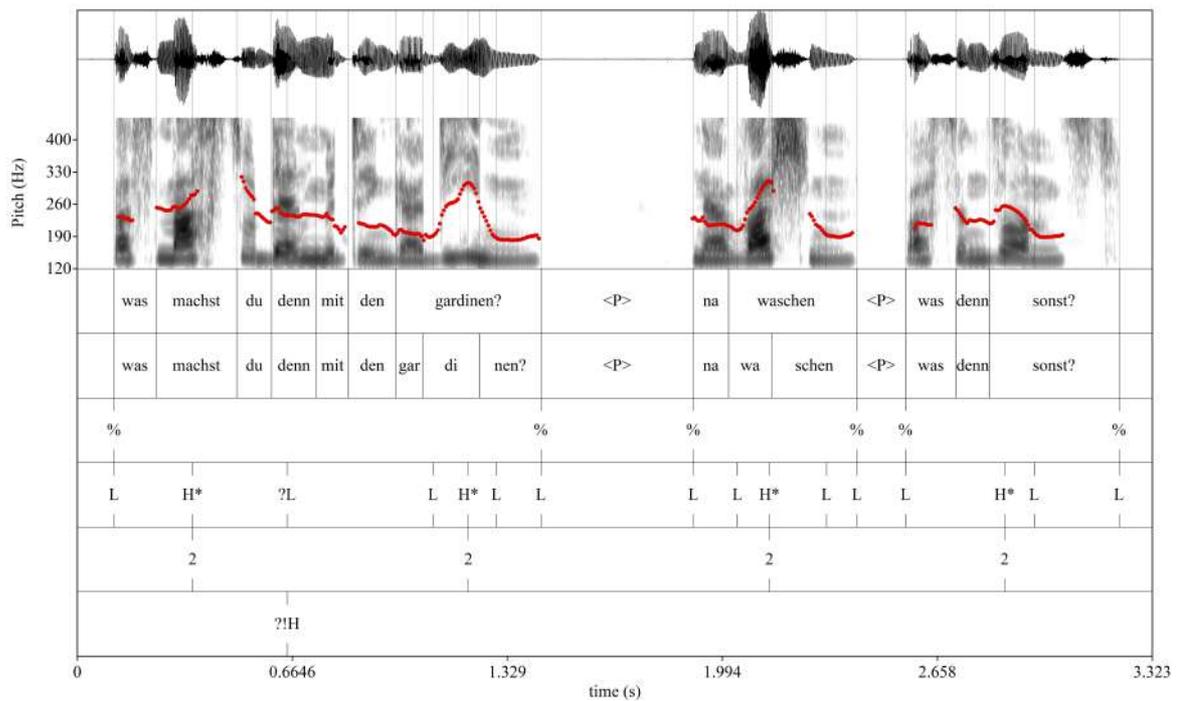


Abbildung 9. Illustration a) mikroprosodischer Perturbationen, die nicht annotiert werden, sowie b) tiefer Plateaus nach hohen Akzenttönen, die mit tiefen Nicht-Akzenttönen vor tiefen Grenzen markiert werden.

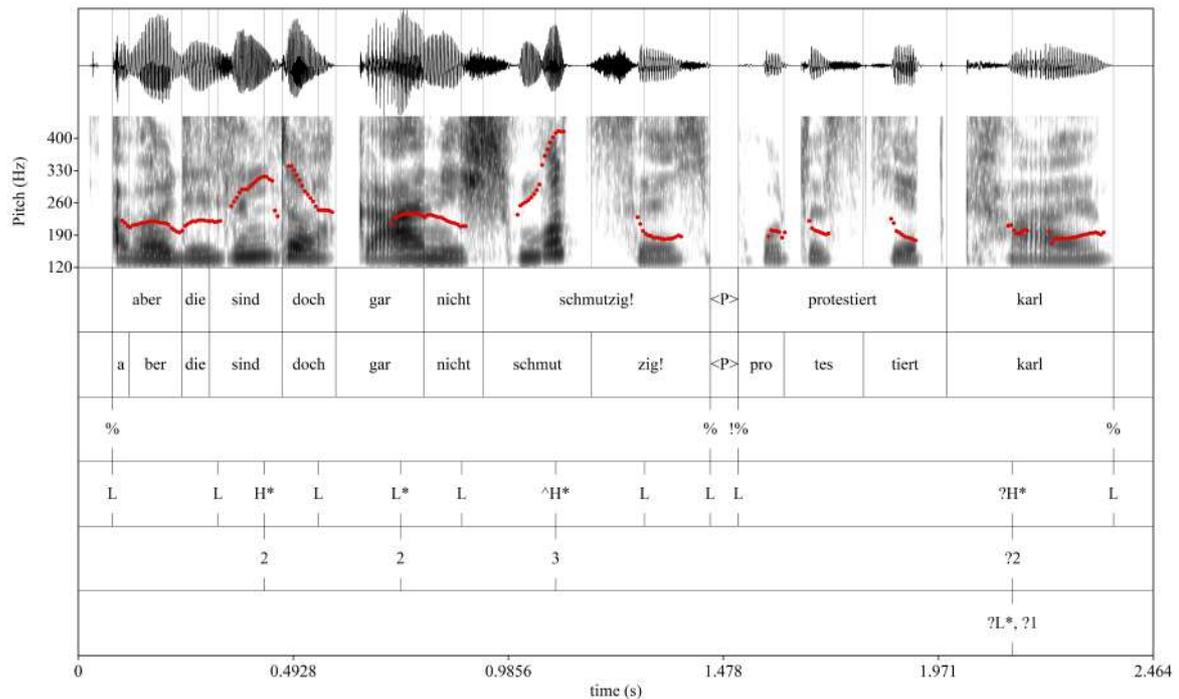
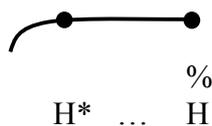


Abbildung 10. Illustration der Annotation tiefer Wendepunkte vor (und nach) hohen Akzenttönen.

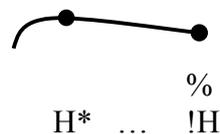
Sonderfall – Plateaukonturen:

Es gibt drei verschiedene Plateaukonturen, die in DIMA entsprechend annotiert werden können (vgl. für Plateaukonturen Niebuhr 2013, Peters 2014).

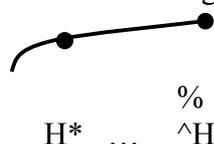
1a. Hochplateau, das auf gleicher, hoher F0 Ebene bleibt, oder von einem tiefen Akzentton zu einem hohen Nicht-Akzentton ansteigt.



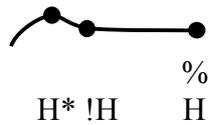
1b. In Abgrenzung zu einem Plateau gibt es eine Kontur mit gleichmäßig fallender F0 von einem hohen Akzentton oder einem hohen Nicht-Akzentton bis zu einem herabgestuften finalen Grenzton; die Perception der Grenze ist eindeutig *nicht* tief.



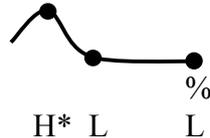
1c. Darüber hinaus gibt es eine Kontur mit gleichmäßig steigender F0 von einem hohen Akzentton oder einem hohen Nicht-Akzentton bis zum finalen hohen Grenzton. Die Grenze wird eindeutig höher wahrgenommen als der vorangehende Hochton.



2. Abgestuftes Hochplateau nach einem hohen Akzentton oder einem hohen Nicht-Akzentton; F0 bleibt konstant auf abgestuftem F0-Niveau bis zum Phrasenende.



3. Tiefplateau, das auf gleicher, tiefer F0 Ebene bleibt.



3.3 Prominzenzebene

Grundlegendes

Auf der Prominzenzebene werden drei Stufen der perzeptiven Prominenz unterschieden (vgl. Kieler Intonationsmodell KIM; Kohler, 1991), die auf Silbenebene annotiert werden. Nicht-Prominenz wird nicht annotiert, d.h. dass die drei Stufen der Prominenz alle eine relative Prominenz gegenüber nicht-prominenten Silben ausdrücken. In Abbildung 11 sind die Prominenzstufen 1 und 2 illustriert. Abbildung 12 zeigt eine Äußerung, in der mehrfach die Prominenz der Stufe 3 vorkommt.

- 1 = schwache Prominenz;
Die Stufe 1 drückt reduzierte Prominenz aus, die nicht notwendigerweise mit einer (eindeutigen) F0-Bewegung einhergehen muss. Eine Prominenz der Stufe 1 ist entweder rhythmisch oder tonal bedingt. Somit kann bei Stufe 1 auf der Tonebene ein Akzentton, ein Nicht-Akzentton oder kein Ton vorkommen. Beispiele für die Stufe 1 sind: (i) postfokale Prominenzen in stark reduziertem Tonhöhenregister (vgl. Kügler & Féry 2017), (ii) (partielle) Deakzentuierung, (iii) Prominenz in einer Hutkontur zwischen den beiden Rändern des Hutes, (iv) rhythmisch bedingte Akzente, (v) Druckakzente, (vi) Phrasenakzente.¹
- 2 = starke Prominenz;
In der Regel fällt eine Prominenz der Stufe 2 mit einem Tonakzent zusammen. Allerdings *muss* eine starke Prominenz *nicht* mit einem korrespondierenden Tonlabel einhergehen.
- 3 = extra-starke Prominenz, die deutlich über die starke Prominenz hinausgeht;
Die Stufe 3 kennzeichnet in der Regel emphatische Realisierungen von vollen Akzenten, die sich in der tonalen (phonologischen) Form nicht von Akzenten mit starker Prominenz unterscheiden (z.B. wird sich ein H* der Stufe 3 oftmals phonetisch, etwa durch eine stärkere F0-Auslenkung, von einem H* der Stufe 2 unterscheiden). Zusätzlich kann es aber vorkommen, dass sich die Emphase nicht in der F0 ausdrückt, sondern durch segmentale Hyperartikulation hervorgerufen wird, etwa durch starke Friktion oder Aspiration. Diese schlägt sich in einer erhöhten akustischen Energie der Formanten und der Dauer von Obstruenten bei reduziertem Silbennukleus nieder (vgl. *force accent*, Kohler 2005). Dies hat zur Folge, dass selbst Stufe 3 nicht zwingend mit einem Akzentton auf der Tonebene einhergehen muss.

¹ Inhaltlich überlappen sich die aufgeführten Konzepte. Zur terminologischen Unterscheidung von *Phrasenakzent*, *Phrasenton* und *Druckakzent* vgl. Grice & Baumann (2016:91ff.).

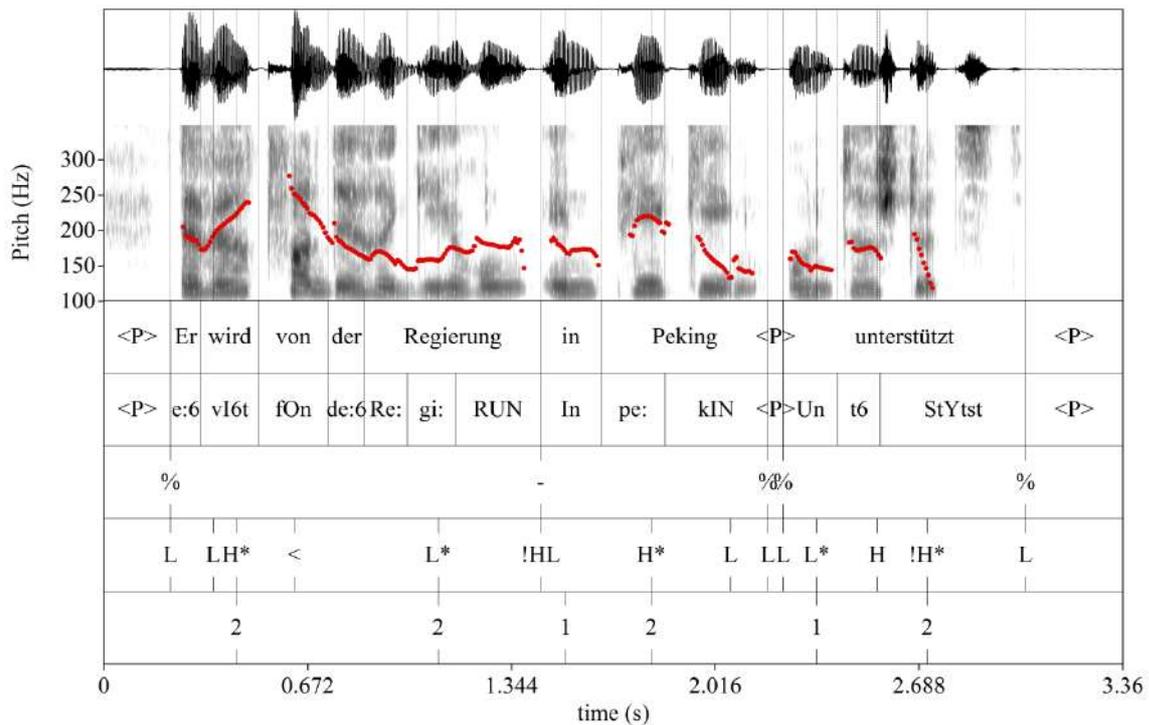


Abbildung 11. Illustration der Prominenzebene mit Prominenzstufen 1 und 2, Beispiel „Peking“.

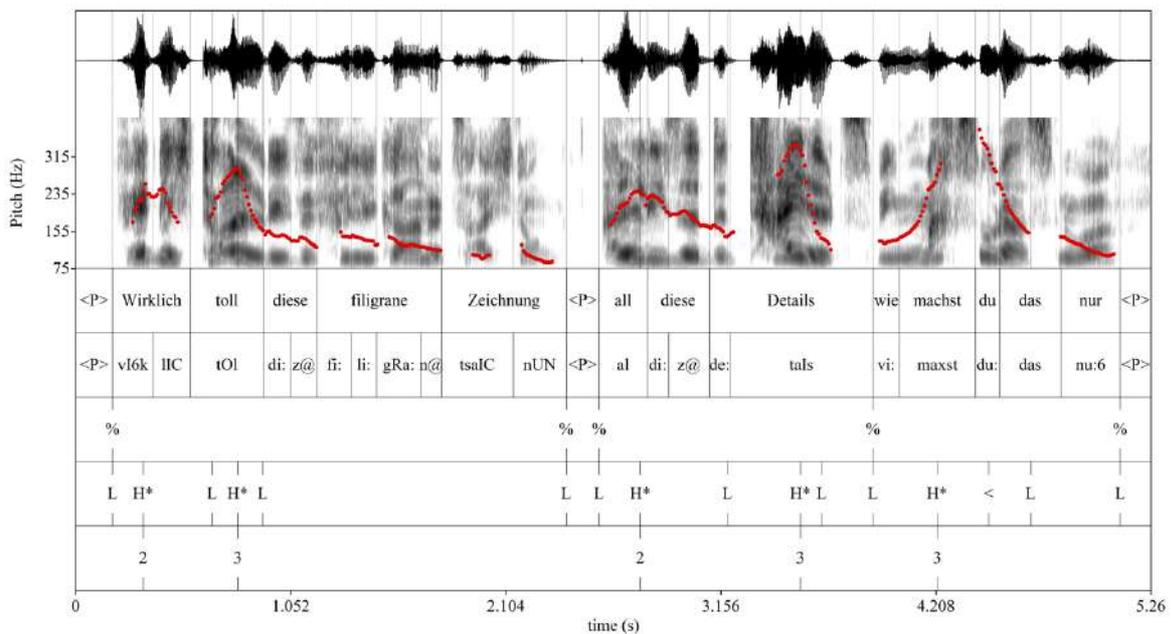


Abbildung 12. Illustration der Prominenzstufe 3, die mehrfach in dieser Äußerung vorkommt.

Annotation der Prominenz

- **Identifizieren von Prominzen und deren Stufe:**

Label im Vokal der prominenten Silbe einfügen

Prominenzstufe festlegen → schwach (=1), stark (=2), extra-stark (=3)

Hinweise zur Prominenzebene:

1. Die Wahrnehmung von Prominenzen und ihrer Stufen ist relational, d.h. nur in Bezug auf den prosodischen Kontext und den sprecherspezifischen Stil kann eine Prominenz (bzw. Nicht-Prominenz) angemessen klassifiziert werden.
2. Prominenzlabel können ohne Tonlabel vorkommen (z.B. bei *Blatt* in Abb. 14). Umgekehrt fällt ein Tonakzent (d.h. ein gesternter Ton) mindestens mit Prominenzstufe 1 zusammen.
3. In der Regel korrespondiert Prominenzstufe 2 mit einem Akzentton auf der Tonebene.
4. In der Regel enthält eine prosodische Phrase mit starker Grenze mindestens eine Prominenz der Stufe 2.
5. Die stärkste Prominenz innerhalb einer Phrase muss allerdings nicht zwangsläufig der Stufe 2 entsprechen, z.B. im Falle einer reduzierten Phrase.

Die Prominenzstufen sind grundsätzlich unabhängig von der Markierung auf der Tonebene. Abbildung 13 zeigt Beispiele für mögliche (bzw. besonders plausible) Labelkombinationen, anhand stilisierter Konturen der Äußerung *Hätten sie nach Pisa fahren sollen?*² Die Konturen unterscheiden sich in der Synchronisierung der tonalen Zielpunkte mit dem Text auf den letzten beiden Wörtern. Besonderes Augenmerk gilt hier der lexikalisch betonten, ersten Silbe in *fahren* (schattiert).

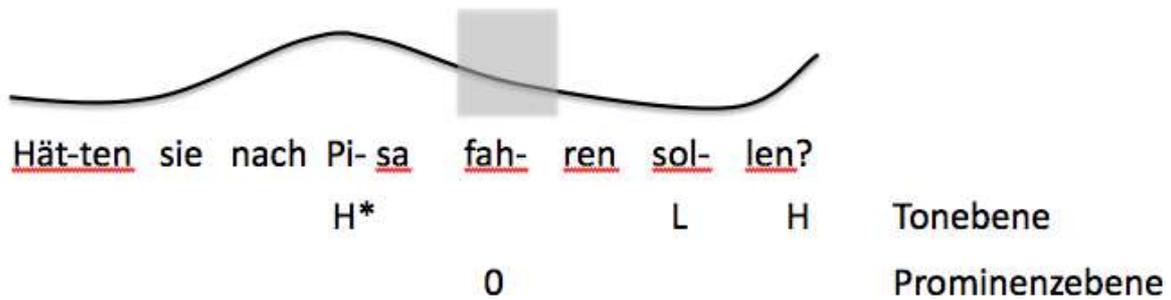
Abbildung 13a: Auf *fahren* wird kein tonaler Zielpunkt erreicht – zwischen dem H* auf *Pi-* und dem L auf *sol-* findet tonale Interpolation statt. Die lexikalisch starke Silbe *fah-* wird daher vermutlich als nicht-prominent (Stufe 0; wird in DIMA nicht annotiert) wahrgenommen. Prinzipiell sind auch höhere Prominenzstufen hier denkbar, je nachdem inwiefern Dauer, Intensität und Vokalqualität zu einem stärkeren perzeptiven Eindruck von Prominenz führen.

Abbildung 13b: Auf der Silbe *fah-* ist der tiefe Wendepunkt in der F0-Kontur nach dem letzten Tonakzent erreicht. Die Silbe kann als nicht-prominent (Stufe 0) wahrgenommen werden, aber auch als schwach prominent (Stufe 1). Stärker empfundene Prominenz wird z.B. durch längere Dauer oder Intensität hervorgerufen, oder durch zusätzliche segmentale Hyperartikulation, wie etwa bei einem *force accent* mit starker Friktion auf dem /f/. Wichtig ist, dass die Prominenzstufen hier mit einem *Nicht-Akzentton* korrespondieren.

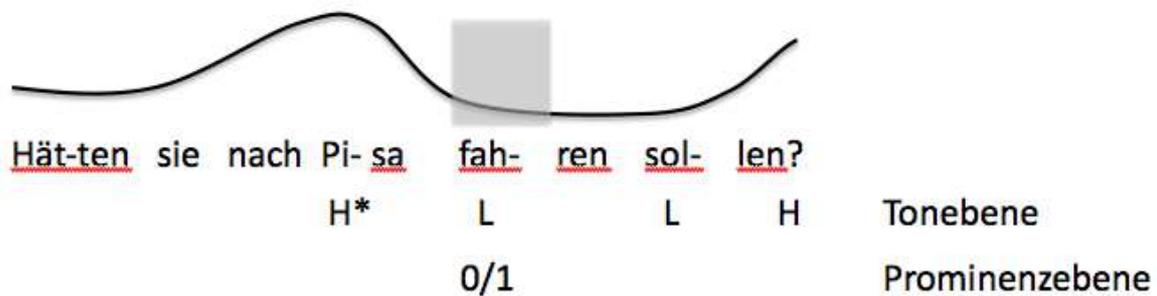
Abbildung 13c: Hier wird *fah-* als tonakzentuierte Silbe wahrgenommen und daher als L* transkribiert. Die Kontur, die diese Silbe als letzten tonalen Wendepunkt vor der hohen Grenze auszeichnet, macht ein entsprechendes Perzept wahrscheinlicher als in 13a und 13b. Hier sind die Prominenzstufen 1 oder 2 wahrscheinlich (die Markierung durch den Stern als Tonakzent widerspricht einem völligen Fehlen perzeptiver Prominenz). Ihre Wahl hängt von der Kombination mit nicht-tonalen Parametern (Dauer, Intensität) und ihrer Ausprägung ab.

² Satzauswahl und Analyse wurden inspiriert durch Benz Müller & Grice (1998).

a.



b.



c.

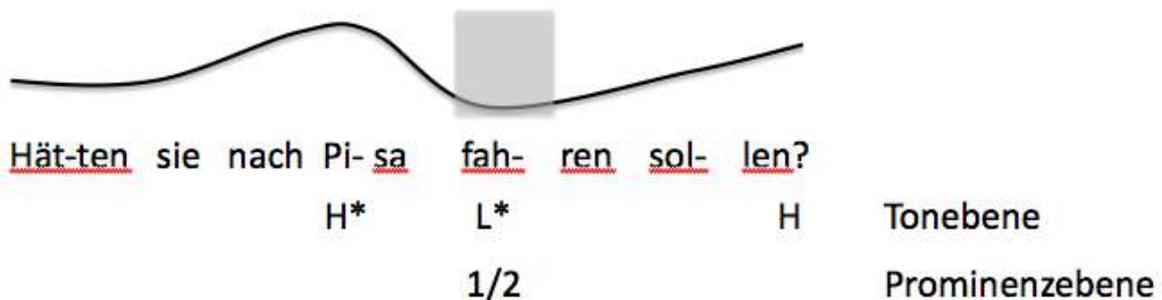


Abbildung 13. Konstruierte Beispiele zur Illustration der Unabhängigkeit der Prominenzebene von der Tonebene; Bezugspunkt ist die erste Silbe von „fahren“ (schattiert) in drei ähnlichen Konturen.

3.4 Kommentarebene

Auf dieser Ebene werden Kommentare aller Art notiert, z.B. Lachen. Daneben können auch unterschiedliche Typen von Disfluenzen, wie etwa Redeflussstörungen, Häsitationen, Unterbrechungen, Abbrüche o.ä., gekennzeichnet werden. Hierbei ist zu beachten, dass eine Disfluenz auch auf Phrasenebene durch eine entsprechende Markierung, in diesem Fall das Label „&“, angezeigt werden sollte.

Die Annotation von Intonation ist nicht immer eindeutig. Dies betrifft alle Ebenen der Annotation, d.h. Phrasen, Töne und Prominenzen. Die Kommentarebene bietet deshalb auch die Möglichkeit, eine alternative Annotation mit Hilfe des Fragezeichens „?“ und dem entsprechenden Grenz-, Ton- oder Prominenzlabel anzugeben (z.B. drückt „?%“ eine Alternative zu einer annotierten Phrase mit schwacher Grenze auf Phrasenebene aus, die dort ebenfalls mit einem vorangehenden „?“ gekennzeichnet ist). Im Folgenden ist ein Beispiel

aufgeführt, das in einer Annotatorengruppe als Zweifelsfall identifiziert wurde (Abbildung 14): Der steigende Tonhöhenverlauf auf *Banane* wurde entweder als tiefer Nicht-Akzentton mit einem nachfolgenden H* oder als L* auf der akzentuierten Silbe mit nachfolgendem hohen Nicht-Akzentton wahrgenommen und entsprechend annotiert. Diese Unklarheiten sind durch das „?“ an der linken Seite des jeweiligen Tonlabels markiert. Auf Kommentarebene wurde die alternative Annotation (L* H) ebenso mit einem vorangestellten „?“ annotiert. Die Klammern um die zwei Alternativen zeigen die Zusammengehörigkeit beider Label an, da eine Änderung des ersten fraglichen Labels (L wird L*) auch eine Anpassung des zweiten fraglichen Labels (H* wird H) nach sich zieht.

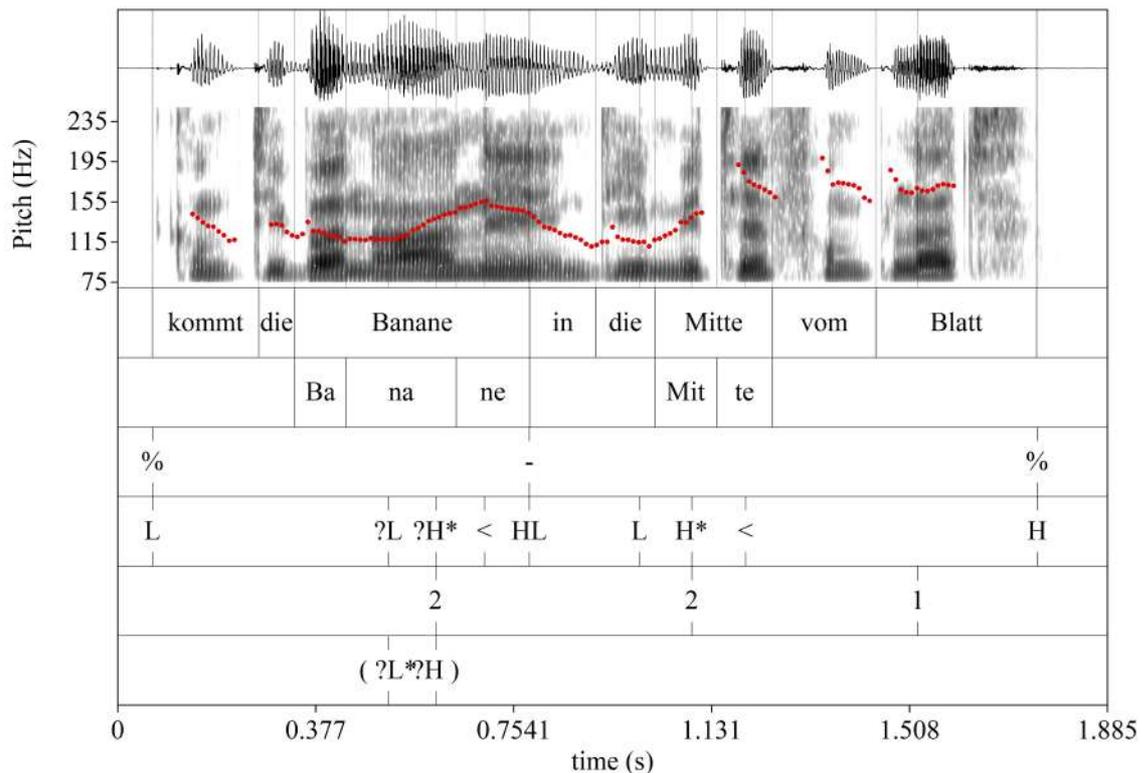


Abbildung 14. Illustration der Kommentarebene; die Unklarheit bezieht sich auf die Tonebene; eine alternative Annotation zweier Tonlabel, die sich aufeinander beziehen, ist auf Kommentarebene durch Klammerung angegeben.

Sofern Unsicherheiten auf mehreren Ebenen bestehen, werden die Alternativen als Kommentar in der Reihenfolge der Ebenen angegeben. Dies soll an einem kurzen Beispiel illustriert werden (Abb. 15): in der reduzierten Phrase kann nicht eindeutig bestimmt werden, ob es sich um einen hohen oder tiefen Ton handelt, der mit der Prominenz assoziiert ist (H*, L*). Darüber hinaus ist auch die Ausprägung der Prominenz nicht genau zu erfassen (1 oder 2). Diese Unsicherheiten haben neben einer Annotation von „?H*“ auf Tonebene und „?2“ auf Prominenzebene auch eine Markierung auf Kommentarebene in der Form „?L*, ?1“ zur Folge. Ein „?“ ohne Zusatzlabel ist auch möglich, sofern der Annotator zwischen der An- und Abwesenheit eines Ereignisses, z.B. einer Prominenz, schwankt. Die Alternative zur Nicht-Existenz einer Prominenz sollte auf der Kommentarebene vermerkt werden (siehe Abb. 16).

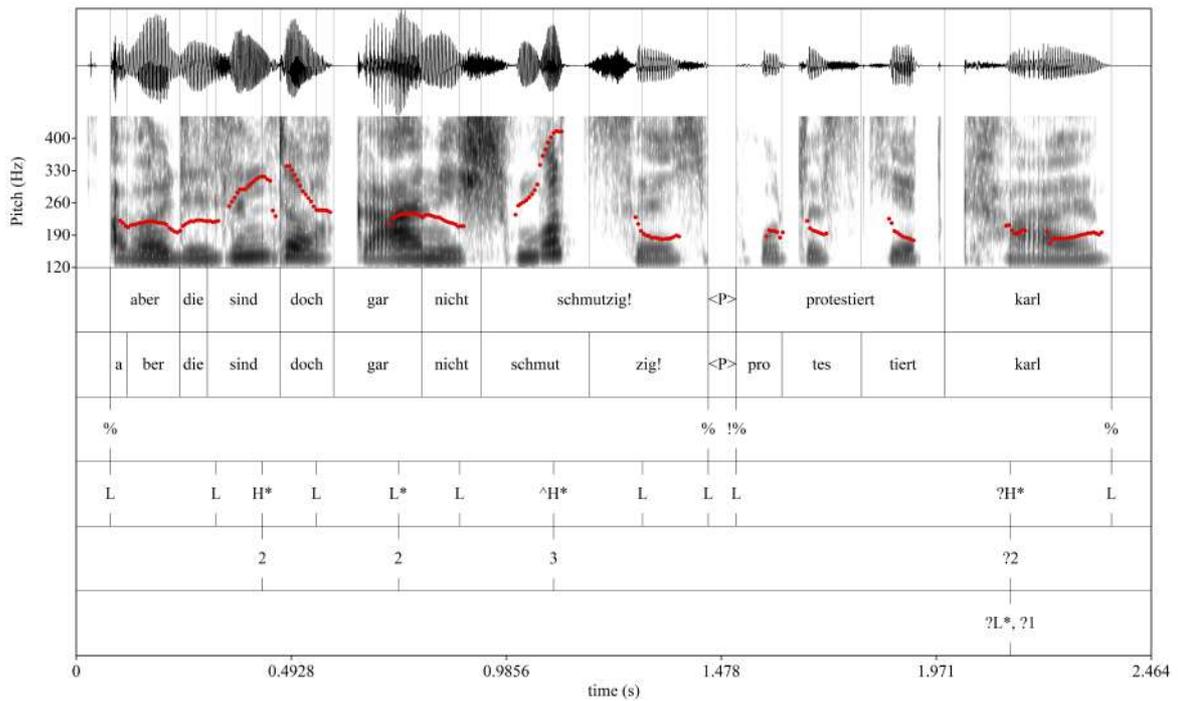


Abbildung 15. Illustration der Kommentarebene; die Unklarheit bezieht sich sowohl auf die Ton- als auch auf die Prominzenzebene. Die alternative Annotation erfolgt nacheinander in der Reihenfolge der Annotationsebenen, hier die Abfolge von Ton- und Prominzenzebene.

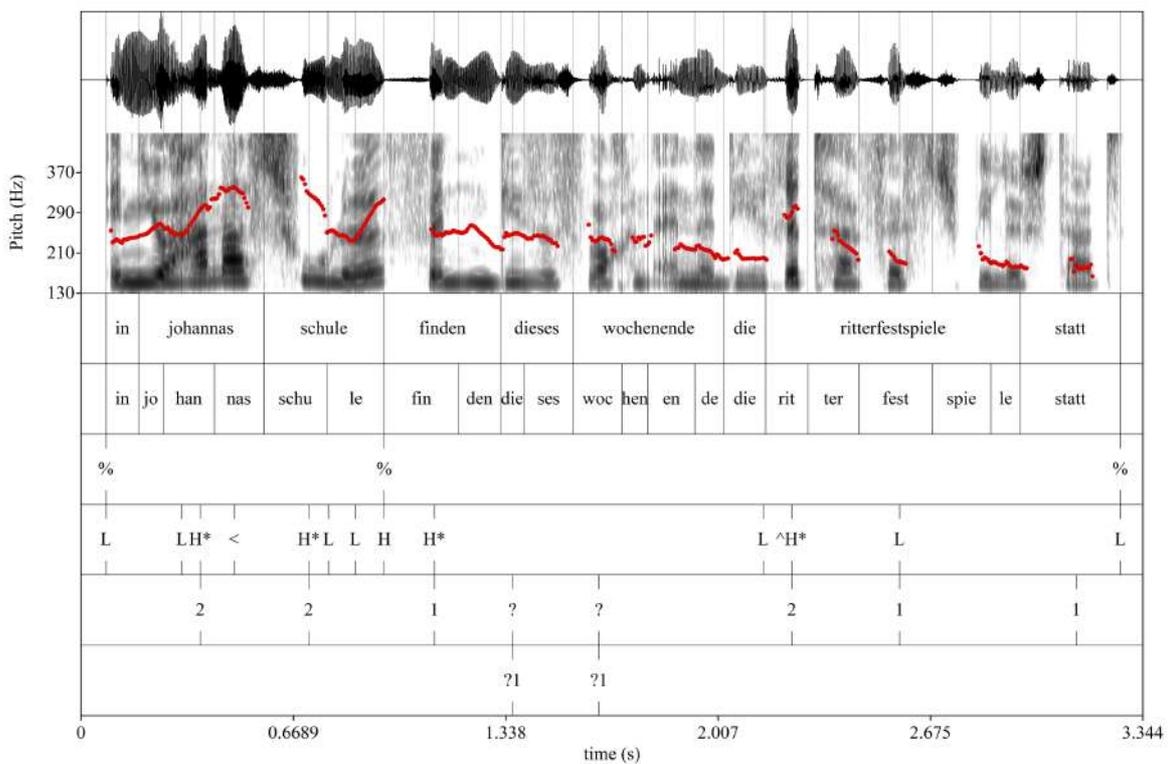


Abbildung 16. Illustration der Kommentarebene; die Unklarheit bezieht sich auf die Prominzenzebene. Mit dem „?“ auf Prominzenzebene wird die Abwesenheit einer Prominenz markiert, als Alternative dazu die Prominenzstufe 1 auf Kommentarebene.

4. Übersetzbarkeit in phonologische Systeme

Es soll grundsätzlich möglich sein, die DIMA-Annotation mit Hilfe von Regeln in eines der o.g. phonologischen Systeme zu übersetzen. Beispiele zur Übersetzung von DIMA zu GToBI, Stuttgart ToBI und dem Oldenburger Modell werden auf der DIMA-Homepage (<http://dima.uni-koeln.de>) veröffentlicht.

Literatur

- Beckman, M. E. & Ayers-Elam, G. (1997). Guidelines for ToBI Labelling, Version 3. Ohio State University. http://www.ling.ohio-state.edu/~tobi/ame_tobi/labelling_guide_v3.pdf.
- Beckman, M. E. & Pierrehumbert, J. (1986). Intonational structure in Japanese and English. *Phonology Yearbook*, 3, 255–309.
- Benzmüller, R. & Grice, M. (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.
- van den Berg, R., Gussenhoven, C. & Rietveld, T. (1992). Downstep in Dutch - implications for a model. In: G. J. Docherty & D. R. Ladd (Eds) *Papers in Laboratory Phonology II* (pp. 335–359). Cambridge: Cambridge University Press.
- Boersma, P. & Weenink, D. (2013/2017). Praat: doing phonetics by computer [Computer program]. Retrieved from <http://www.praat.org/> (May 2017).
- Féry, C. (1993). *German Intonational Patterns*. Tübingen: Niemeyer.
- Féry, C. (2012). Die Annotation der deutschen Intonation: Wie kann GToBI die Syntax und die Semantik integrieren? *Linguistische Berichte* 229, 39–97.
- Grice, M. & Baumann, S. (2016). Intonation in der Lautsprache: Tonale Analyse. In: Primus, Beatrice & Ulrike Domahs (eds.): *Handbuch Laut, Gebärde, Buchstabe*. De Gruyter, Reihe Handbücher Sprachwissen, Band 2. 84-105.
- Grice, M., Baumann, S. & Benzmüller, R. (2005). German Intonation in Autosegmental-Metrical Phonology. In S.-A. Jun (Ed.), *Prosodic Typology. The Phonology of Intonation and Phrasing* (pp. 55–83). Oxford: Oxford University Press.
- Gussenhoven, C. (2004). *The Phonology of Tone and Intonation*. Cambridge: Cambridge University Press.
- t' Hart, J., Collier, R., & Cohen, A. (1990). *A perceptual study of intonation: An experimental-phonetic approach to speech melody*. Cambridge: Cambridge University Press.
- Kohler, K. J. (1990). Macro and micro F0 in the synthesis of intonation. In Kingston, J. & Beckman, M.E. (eds), *Papers in Laboratory Phonology I*. Cambridge University Press, 115-138.
- Kohler, K. J. (1991). A Model of German Intonation. In K. J. Kohler (Ed.), *Arbeitsberichte (AIPUK): Vol. 25. Studies in German Intonation* (pp. 295–360). Kiel: Institut für Phonetik und digitale Sprachverarbeitung.
- Kohler, K. J. (2005). Form and function of non-pitch accents. *AIPUK* 35a, 97-123.
- Kügler, F. & Féry, C. (2017) Postfocal Downstep in German. *Language and Speech*, 60, 2, 260-288. DOI: 10.1177/0023830916647204
- Ladd, D. R. (1988). Declination “reset” and the hierarchical organization of utterances. *Journal of the Acoustical Society of America* 84, 530–544.
- Ladd, D. R. (2008). *Intonational Phonology*. Cambridge: Cambridge University Press.
- Mayer, J. (1995). *Transcription of German intonation: the Stuttgart System*, Ms., University of Stuttgart.

- Möbius, B., Zimmermann, A. & Hess, W. (1987). Microprosodic fundamental frequency variations in German. *Proceedings of the 11th ICPHS (Tallinn)*, volume 1, 146-149.
- Niebuhr, O. (2013). Resistance is futile - The intonation between continuation rise and calling contour in German. *Proceedings of 14th InterSpeech*, Lyon, 225–229.
- Nespor, M. & Vogel, I. (1986). *Prosodic phonology*. Berlin: Mouton De Gruyter.
- Ohala, J. J. & Eukel, B. W. (1987). Explaining the intrinsic pitch of vowels. In R. Channon & L. Shockey (eds.), *In honor of Ilse Lehiste. Ilse Lehiste Pu'hendusteos*. Dordrecht: Foris. 207–215.
- Peters, J. (2005). Intonation. In *Duden-Die Grammatik* (pp. 95–128). Mannheim: Dudenverlag [7. Auflage].
- Peters, J. (2014). *Intonation*. Heidelberg: Winter.
- Pierrehumbert, J. B. (1980). *The phonology and phonetics of English intonation* (PhD dissertation). MIT, Cambridge, Massachusetts.
- Selkirk, E. O. (1984). *Phonology and syntax: The relation between sound and structure*. Cambridge, MA: MIT Press.
- Wells, J.C. (1997). SAMPA computer readable phonetic alphabet. In: Gibbon, D., Moore, R. and Winski, R. (eds.): *Handbook of Standards and Resources for Spoken Language Systems, Part IV, section B*, Berlin and New York: Mouton de Gruyter.
- Whalen, D. H. & A. G. Levitt (1995). The universality of intrinsic F0 of vowels. *Journal of Phonetics* 23: 349-366.