

Experimentelle Untersuchungen
zu phonetischen und semantischen Aspekten
phrasenfinaler Melodiemuster
im Deutschen

Schriftliche Hausarbeit zur Erlangung des Grades
eines Magister Artium (M.A.)
der Philosophischen Fakultät
der Christian-Albrechts-Universität
zu Kiel

vorgelegt von
Gilbert Ambrazaitis

Kiel
2004

überarbeitete Fassung vom Juni 2006

Referent: Prof. Dr. Jonathan M. Harrington

Koreferent: Prof. Dr. Klaus J. Kohler

Tag der mündlichen Prüfung: 01.12.2004

Dekan: Prof. Dr. Siegfried Oechsle

© 2004, 2006 Gilbert Ambrazaitis

Anmerkungen / *Comments*

Die vorliegende Version (Juni 2006) weicht vom Original (2004) lediglich dahingehend ab, dass Tippfehler behoben, sowie Bildunterschriften o.ä. korrigiert worden sind.

Die Ergebnisse der vorliegenden Arbeit sind in Form eines Zeitschriftenartikels in englischer Sprache erschienen:

A summary of this M.A. thesis has also been published as a journal article (in English):

Ambrazaitis, Gilbert (2005). Between Fall and Fall–Rise: Substance–Function Relations in German Phrase–Final Intonation Contours. *Phonetica* 62 (2–4), 196–214.

Inhaltsverzeichnis

1	Einleitung	1
I	Theoretische, terminologische und technische Grundlagen	5
2	Aspekte der Akustik, Produktion und Perzeption der Sprechmelodie	7
2.1	Definitionsmöglichkeiten der Intonation	7
2.2	Akustische und technische Grundlagen	9
2.3	Grundlagen zur Produktion	12
2.3.1	Die Systeme des Sprechapparats	13
2.3.2	Der Aufbau des Larynx	14
2.3.3	Phonation und Steuerung der Grundfrequenz	15
2.3.4	Nicht-gesteuerte myoelastische und aerodynamische Einflüsse auf die Grundfrequenz	16
2.4	Grundlagen zur Perzeption	18
2.4.1	Tonhöhe, Grundfrequenz und ihre Skalierung	18
2.4.2	Wahrnehmbare F_0 -Unterschiede und F_0 -Konturen	20
2.4.3	Kategoriale Wahrnehmung?	22
3	Phrasenfinale Melodiemuster des Deutschen und ihre Beschreibung	25
3.1	Drei Intonationstypen	27
3.2	Zur Modellierung der Intonation	27
3.2.1	Die Zielsetzung eines Beschreibungsrahmens	28
3.2.2	Der Phänomenbereich eines Modells	28
3.2.3	Die Kategorien eines Modells	30
3.2.4	Die methodische Basis eines Modells	30
3.3	Zwei Beschreibungsrahmen – GToBI und KIM/PROLAB	31
3.3.1	KIM/PROLAB	31
3.3.2	GToBI	33
3.3.3	KIM/PROLAB vs. GToBI im Überblick	34
3.4	Phonetische Aspekte phrasenfinaler Melodiemuster	35

3.5	Semantische Aspekte phrasenfinaler Melodiemuster	36
3.6	Hörbeispiel	37
4	Problemstellung	39
4.1	Das pseudoterminale Melodiemuster	39
4.1.1	Aspekte der Produktion	39
4.1.2	Aspekte der Perzeption	40
4.1.3	Funktionale Aspekte	42
4.2	Hypothesen und methodisches Vorgehen im Überblick	43
II	Eigene korpusbasierte und experimentelle Untersuchungen	47
5	Korpusanalyse	49
5.1	Ziele	49
5.2	Methode	50
5.2.1	Datenbasis	50
5.2.2	Voruntersuchung	54
5.2.3	Datenerhebung	55
5.2.4	Berechnungen	58
5.2.5	Datenaufbereitung und -auswertung	58
5.3	Ergebnisse	59
5.3.1	Konturgestalt	59
5.3.2	Prüfung der Hypothesen 1, 2 und 4	60
6	Perzeptionsexperiment	63
6.1	Voraussetzungen und Ziele	63
6.2	Methode	64
6.2.1	Einleitendes zum semantischen Differential	64
6.2.2	Kritische Bemerkungen zur Methode und Konfigurationen im vorliegenden Experiment	66
6.2.3	Auswahl der semantischen Skalen	69
6.2.4	Auswahlkriterien für das Sprachmaterial	71
6.2.5	Erzeugung der Stimuli	72
6.2.6	Technische Konzeption, Fragebogen und Durchführung des Experiments	78
6.2.7	Versuchspersonen	80
6.3	Ergebnisse	81
6.3.1	Postexperimentelle Befragung	81
6.3.2	Kontrollstimulus Sko1	82

6.3.3	Prüfung der Hypothese 3	84
6.3.4	Kontrollstimulus Sk00	93
6.3.5	Prüfung der Hypothese 4	95
7	Zusammenfassende Diskussion	97
7.1	Interpretation der wesentlichen Ergebnisse	97
7.1.1	Phonetische Eigenschaften der untersuchten Melodiemuster	98
7.1.2	Semantische Eigenschaften der untersuchten Melodiemuster	99
7.2	Ist das pseudoterminal Melodiemuster eine Kategorie der deutschen Intona- tionsphonologie?	101
7.3	Ein methodisch orientierter Ausblick	103
	Anhang	115
A	Korpusanalyse	115
A.1	Erläuterungen zu den Rohdatentabellen	115
A.2	Rohdaten und Informationen zu [2.,]	116
A.3	Rohdaten und Informationen zu [2;]	118
B	Perzeptionsexperiment	125
B.1	Fragebogen (Beispielseiten)	125
B.2	Freie Kommentare der Probanden zur Natürlichkeit der Stimuli	129
B.3	Deskriptive Statistik	130
C	CDROM	133

Kapitel 1

Einleitung

Diese Arbeit untersucht einen ausgewählten Aspekt der deutschen *Intonation*. In Abschnitt 2.1 wird eine Definition der Intonation erarbeitet. Vorerst kann der Terminus mit *Sprechmelodie* paraphrasiert werden, womit bereits eine entscheidende Information gegeben ist: Essentielle definierende Parameter einer Melodie sind der Tonhöhenverlauf wie auch seine zeitliche Struktur. Die Begriffe *Melodie(muster)* und *Kontur* werden in dieser Arbeit synonym verwendet.

Die Sprechmelodie kann in einer Sprache wie dem Deutschen eine Reihe unterschiedlicher kommunikativer Funktionen erfüllen. Erstens kann sie einzelne Wörter oder Silben hervorheben, was zu einer Fokussierung ausgewählter inhaltlicher Aspekte der Äußerung führt. Eine solche Hervorhebung wird als *Akzent* bezeichnet. Zum Beispiel unterscheiden sich die folgenden beiden Äußerungen nur in ihrer Akzentuierung und damit in der Fokussierung unterschiedlicher inhaltlicher Aspekte (akzentuierte Silben sind durch GROSSbuchstaben markiert): “Sie ist geGANGen” vs. “SIE ist gegangen”. Ein prototypischer sprechmelodischer Akzent entsteht durch eine lokale (etwa auf die Umgebung der akzentuierten Silbe beschränkte) deutliche Bewegung im Tonhöhenverlauf. Dabei kann die Melodie entweder ein Tonhöhenmaximum oder ein Tonhöhenminimum aufweisen. Der erste Fall wird auch als *Gipfel(kontur)* bezeichnet, der zweite als *Talkontur*.

Eine zweite Funktion besteht in der syntagmatischen Gliederung einer Äußerung. Derartige durch intonatorische Mittel erzeugte Abschnitte können als (*prosodische*) *Phrasen*¹ bezeichnet werden. Drittens können sich Äußerungen bezüglich ihrer Gesamtbedeutung allein durch den Intonationsverlauf am Ende einer Phrase (= *phrasenfinale(s)* Intonation, *Kontur* oder *Melodiemuster*) unterscheiden. Zum Beispiel kann “Sie ist geGANGen” (mit einem Gipfel auf GANG) bei fallender phrasenfinaler Intonation eine Aussage bedeuten und bei einer fallend-steigenden Intonation eine (erstaunte) Frage.

Die Gesamtbedeutung der Äußerung ist also nicht allein von den Bedeutungen der einzel-

¹Der Begriff *Prosodie* wird – abgesehen von diesem adjektivischen Gebrauch – in der vorliegenden Arbeit nicht verwendet, vgl. 2.1.

nen Wörter der Äußerung abhängig. Sie kann auch als *Pragmatik* der Äußerung bezeichnet werden, während die Bedeutung einer Einzelkomponente der Äußerung *Semantik* genannt werden kann. Im genannten Beispiel stellen nicht nur die einzelnen Wörter, sondern auch die Akzentverhältnisse sowie die phrasenfinale Intonation der Äußerung derartige Komponenten dar. Insofern hat sowohl das Wort “sie” eine eigene Semantik, als auch die Akzentuierung auf “geGANGen” sowie der fallend-steigende Melodieverlauf — und alle drei tragen zur Gesamtbedeutung (Pragmatik) der Äußerung bei. Es wird in dieser Arbeit keine Unterscheidung zwischen den Begriffen *Semantik*, *Bedeutung* und (*kommunikative*) *Funktion* vorgenommen: Die Semantik oder Bedeutung eines Melodiemusters ist seine kommunikative Funktion. Diese Funktion beeinflusst, wie auch die Funktionen der einzelnen Wörter, die Pragmatik der Äußerung.

Die hier aufgeführten Funktionen der Intonation stellen keinesfalls eine erschöpfende Liste dar, ebenso wenig kann die Intonation als alleiniges Mittel zur Erfüllung dieser Funktionen betrachtet werden. Diese Arbeit beschränkt sich auf ausgewählte Aspekte aus dem Bereich der Funktion von phrasenfinalen Melodiemustern. Dabei ist die Intonation das einzige Mittel, welches berücksichtigt wird.

Es gibt einige phrasenfinale Melodiemuster, deren Existenz und Relevanz für das Deutsche nicht angezweifelt werden: Handelt es sich bei dem phrasenfinalen Akzentmuster um eine Talkontur, so kann die phrasenfinale Melodie nur steigend sein. Nach einem phrasenfinalen Melodiegipfel gibt es global betrachtet drei Möglichkeiten für ein phrasenfinales Muster: (a) Die Melodie kann auf einem hohen Niveau auslaufen, (b) die Tonhöhe kann bis zum Ende der Phrase abfallen, oder (c) die Melodie kann einem fallend-steigenden Verlauf folgen. Jede dieser Möglichkeiten kann im Detail modifiziert werden, so dass in Fall (b) beispielsweise unterschiedliche Tiefpunkte der Melodie erreicht werden können. In Kapitel 3 werden zwei verschiedene Beschreibungsmodelle vorgestellt, die sich in der Handhabung solcher Details unterscheiden. Ein Modell sollte dabei berücksichtigen, welche Muster für die untersuchte Sprache tatsächlich relevant sind. Entscheidungen lassen sich vereinfachen, wenn empirische Befunde vorliegen. Daher sind gezielte Experimente zu sinnvoll gewählten Teilfragestellungen hilfreich. Die in Kapitel 3 vorgestellten Modelle basieren bezüglich einiger Teilprobleme auf den Ergebnissen von gezielten Experimenten, nicht aber bezüglich der Vielfalt der angenommenen phrasenfinalen Konturen.

In der vorliegenden Arbeit soll eine solche Teilproblematik gewählt und experimentell untersucht werden. Es wurde vorgeschlagen (Peters 1999), dass ein weiteres phrasenfinales Muster für das Deutsche relevant sein könnte, welches sich nicht ohne Weiteres einer der vorgestellten Möglichkeiten (a, b, c) zuordnen lässt. Es zeichnet sich durch einen tiefen Fall der Melodie aus, auf den ein sehr kleiner Wiederanstieg auf der letzten Silbe der Phrase folgt. Dieses Muster unterscheidet sich laut Peters (1999) in der Gesamtheit perceptiv und funktional von einem tief fallenden, kann aber nicht als Variante der fallend-steigenden Muster verstanden werden, da der leichte Wiederanstieg nicht unbedingt als ein solcher wahrnehm-

bar sein muss. Peters bezeichnet dieses Muster als “*pseudoterninale*” Kontur. In Kapitel 3 wird detaillierter auf Peters Hypothesen, insbesondere bezüglich der Funktion dieses Musters eingegangen.

Warum ist die Untersuchung gerade dieses Melodiemusters besonders interessant? Peters (1999) gibt an, dass der leichte Wiederanstieg der pseudoterninalen Kontur etwa 5–15 Hz² beträgt. In Kapitel 2 werden einige Grundlagen zur Spracherzeugung und -wahrnehmung vorgestellt. Es zeigt sich, dass ein so geringer Tonhöhenanstieg wie der des pseudoterninalen Musters gerade im Grenzbereich dessen liegt, was im Rahmen der intendierten Sprechmelodie produzierbar und perzipierbar ist. Ziel der Arbeit ist es festzustellen, ob die pseudoterninale Kontur von deutschen Muttersprachlern als eine eigenständige Kategorie (neben einer fallenden und einer fallend-steigenden) erkannt und einer eigenen Funktion zugeordnet wird. Ein Nebenziel besteht darin, die Ergebnisse vor dem Hintergrund psychophonetischer Untersuchungen zur Frage nach den kleinsten perzipierbaren bzw. sprachlich relevanten Unterschieden im Bereich der Intonation zu diskutieren.

Im ersten Teil der Arbeit wird in die nötigen Grundlagen aus den Bereichen der Akustik, Produktion und Perzeption der Sprechmelodie eingeführt (Kapitel 2), sowie ein Überblick über zwei Intonationsmodelle des Deutschen und ihre Behandlung der phrasenfinalen Melodiemuster gegeben (Kapitel 3). Danach werden diese Grundlagen zusammenfassend diskutiert und daraus Hypothesen zum pseudoterninalen Muster abgeleitet (Kapitel 4). Im Verlauf dieses ersten Teils werden einige für die Untersuchung wesentliche Termini definiert. Einige Grundbegriffe werden allerdings als Voraussetzungen betrachtet und nicht definiert, so etwa die Terminologie und Symbolik des *Internationalen Phonetischen Alphabets* (IPA 1999) oder auch Konzepte wie *Phonologie* oder *Silbe*. Letztere sind zwar ausgedehnter Diskussionen würdig, stehen aber nicht im Mittelpunkt dieser Arbeit. Im zweiten Teil werden zwei Untersuchungen beschrieben, die zum Testen der formulierten Hypothesen durchgeführt wurden: eine Korpusanalyse (Kapitel 5) und ein Perzeptionsexperiment (Kapitel 6). Eine zusammenfassende Diskussion (Kapitel 7) interpretiert die wesentlichen Ergebnisse und schließt mit einem Ausblick die Arbeit ab.

²Die Einheit Hz (*Hertz*) wird in 2.2 erläutert.

Teil I

Theoretische, terminologische und technische Grundlagen

Kapitel 2

Aspekte der Akustik, Produktion und Perzeption der Sprechmelodie

Das untersuchte Melodiemuster zeichnet sich durch eine tonale Bewegung aus, die im Vergleich zu den übrigen bekannten Melodiemustern relativ gering ausfällt. Die Analyse dieses Musters enthält damit eine besondere Problematik: Zum einen können Abläufe der Spracherzeugung (*Produktion*) zu kleinen *nicht intendierten* Perturbationen in der *intendierten* Sprechmelodie führen. Zum anderen existieren Beschränkungen in der Wahrnehmung (*Perzeption*) hinsichtlich tonaler Konturen von kurzer Dauer und geringem Tonhöhenumfang. Dieses Kapitel bietet eine Kurzeinführung in diese Aspekte der Sprachproduktion (2.3) und -perzeption (2.4). Vereinfacht dargestellt wird bei der Spracherzeugung die Nachricht des Sprechers in ein *akustisches Signal* umgewandelt, welches bei der Sprachwahrnehmung durch den Hörer wieder in eine Nachricht zurückgewandelt wird. Zwischen den ver- und entschlüsselnden Mechanismen der Produktion und der Perzeption steht also der übertragende Mechanismus der *Akustik*. Da die Sprechmelodie im allgemeinen anhand eines akustischen Parameters analysiert wird, nämlich der *Grundfrequenz* des Sprachsignals, werden zunächst die wesentlichen akustischen Grundlagen für das Verständnis der vorliegenden Arbeit erläutert (Abschnitt 2.2). Dabei wird bereits Bezug genommen auf praktische Aspekte der Datenerhebung und -verarbeitung, z.B. auf die Ermittlung und Manipulation der Grundfrequenz.

2.1 Definitionsmöglichkeiten der Intonation

Die gesprochene Sprache kann anhand zweier Ebenen modelliert werden. Die *segmentelle* Ebene repräsentiert in abstrakter Weise die Verkettung der einzelnen Sprachlaute. Die *suprasegmentelle* beschreibt diejenigen Eigenschaften der gesprochenen Sprache, die sich über

längere Abschnitte der segmentellen Ebene legen können.¹ Die Sprechmelodie wird der suprasegmentellen Ebene zugerechnet. Weitere suprasegmentelle Phänomene sind z.B. Variationen in der segmentellen Dauerstrukturierung, der Intensität, der Sprechgeschwindigkeit oder der Stimmqualität.

Der Terminus *Intonation* wird in der Literatur unterschiedlich definiert. Eine Möglichkeit besteht darin, sämtliche suprasegmentellen Phänomene als Teil der Intonation zu verstehen (vgl. Bußmann 2002). Als Name einer derartigen Sammelkategorie wird auch der Terminus *Prosodie* verwendet. In der Literatur besteht diesbezüglich keine terminologische Klarheit (vgl. die z.T. unklaren und widersprüchlichen Definitionen in Bußmann (2002), in denen zum einen die Intonation mit der Prosodie gleichgesetzt, zum anderen die Intonation als eine Komponente der Prosodie bezeichnet wird). Eine klare Trennung zwischen Prosodie und Intonation wird z.B. von Selting (1987) und Möbius (1993) formuliert und soll auch hier übernommen werden.

Jones (1969:275) definiert *Intonation* als “*the variations which take place in the pitch of the voice in connected speech, i.e. the variations in the pitch of the musical note produced by the vibration of the vocal cords*”. Die zeitliche Strukturierung des Tonhöhenverlaufs (im Folgenden auch *Timing*) wird nicht explizit in die Definition eingeschlossen². Selting (1987:779) hingegen definiert *Intonation* als “*the contour or melody of speech in terms of the temporal organization of perceived pitch of utterances*” und schließt damit das *Timing* explizit ein. Dies erscheint sinnvoll, da gezeigt wurde, dass ein Gipfel in der Melodie unterschiedliche Intonationskategorien repräsentiert, je nachdem, ob er früh oder später im Bereich des Vokals der akzentuierten Silbe auftritt (Kohler 1987). Ebenso können unterschiedliche Formen des Gipfels (Dauer bzw. Steilheit des Anstiegs oder des Abstiegs) in unterschiedlichen Intonationskategorien resultieren (Niebuhr 2003).

In der vorliegenden Arbeit soll die Intonation im Wesentlichen nach Selting (1987) verstanden werden, wie es bereits in der Einleitung inhaltlich angedeutet wurde. Allerdings folgen aus der Formulierung “*perceived pitch*” (s.o. Definition von Selting 1987) methodologische Schwierigkeiten (s. Abschnitt 2.4). Vereinfachend wird daher mit der *erzeugten Tonhöhe* (vgl. die Definition von Jones 1969) anstelle der *wahrgenommenen Tonhöhe* gearbeitet (die Diskrepanz wird in 2.4 thematisiert). Als akustisches Korrelat dieser erzeugten Tonhöhe kann die Grundfrequenz des Sprachsignals betrachtet werden (vgl. Abschnitt 2.3).

¹Der Terminus *Segment* bezeichnet keine phonetische Realität, sondern vielmehr ein Konstrukt, welches ein hilfreiches Konzept für die Phonologie, sowie für die Etikettierung und Auswertung phonetischer Datenbanken darstellt.

²Gewissermaßen schließt Jones (1969) die zeitliche Komponente nicht aus, denn “Variationen der Tonhöhe” erstrecken sich notwendigerweise über die Zeit.

2.2 Akustische und technische Grundlagen

Das akustische Sprachsignal besteht aus Luftdruckschwankungen und ist im physikalischen Sinne eine Schallwelle (z.B. Reetz 1999). Ein akustisches Signal ist ein *Zeitsignal*: Es enthält Informationen darüber, wie sich der Luftdruck über die Zeit ändert. Es kann über ein Mikrofon aufgezeichnet werden. Ein Mikrofon wandelt das akustische Signal in ein elektrisches Signal um, welches auf unterschiedlichen Datenträgern, z.B. in digitaler Form auf Rechnerfestplatten, gespeichert werden kann. Auf diese Weise kann es unter Zuhilfenahme von geeigneten Werkzeugen analysiert und manipuliert werden (s.u.).

In einem *Oszillogramm* können Sprachsignale graphisch dargestellt werden. Abbildung 2.1 zeigt das Oszillogramm eines einfachen Zeitsignals (vertikale Achse: Luftdruck, relativ zum atmosphärischen Druck; horizontale Achse: Zeit). Es handelt sich um den Ausschnitt eines *periodischen* Signals, d.h. das Signal besteht aus einer Schwingung, die sich in exakter Weise wiederholt (theoretisch ins Unendliche). Die Dauer einer *Periode* (also einer Schwingung) kann im Oszillogramm zwischen zwei einander entsprechenden Zeitpunkten zweier aufeinander folgender Perioden auf der Zeitachse gemessen werden. In Abbildung 2.1 sind für diesen Zweck besonders markante Zeitpunkte, nämlich die Maxima zweier aufeinander folgender Schwingungen, gewählt worden. Ein akustisches Signal, welches aus einer einfachen periodischen Schwingung erzeugt wird, heißt *Sinuston*.

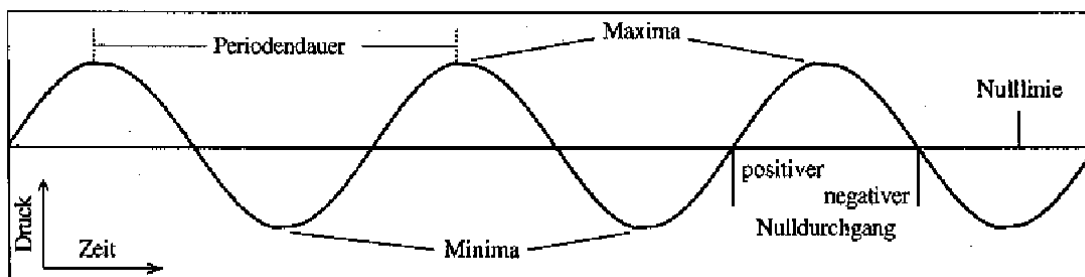


Abbildung 2.1: Oszillogramm eines Sinustons. Horizontale Achse: Zeit; vertikale Achse: Amplitude des Luftdrucks. Nulllinie = atmosphärischer Druck. (Quelle: Reetz 1999)

Jeder Sinuston ist durch drei Parameter eindeutig bestimmt: seine (*Perioden*)*frequenz*, seine *Amplitude* und seine *Phase*. Auf eine weitere Behandlung der letzten beiden kann in dieser Arbeit verzichtet werden. Die Phase spielt in der sprachlichen Kommunikation eine nebensächliche Rolle. Die Amplitude wäre interessant, wenn z.B. die Lautstärke von Sprachsignalen thematisiert würde; der Begriff wird hier verwendet, um die Auslenkungen des Signals auf der vertikalen Achse zu bezeichnen³. Die (*Perioden*)*frequenz* beschreibt, wie häufig

³Die Auslenkungen werden korrekterweise als *Elongationen* bezeichnet, während *Amplitude* eigentlich die *maximale Elongation* bezeichnet (vgl. Reetz 1999:18f). Zur Vereinfachung wird hier auf den zusätzlichen Terminus *Elongation* verzichtet und der Terminus *Amplitude* verallgemeinert, da diese Arbeit keinen Boden für diesbezügliche Missverständnisse bietet.

sich die Schwingung in einer Zeiteinheit wiederholt. Rechnerisch ist die Frequenz (F) der Kehrwert der Periodendauer (T , in s):

$$F = \frac{1}{T} \quad |F| = \frac{1}{s} = \text{Hz} \quad (2.1)$$

Die Maßeinheit der Frequenz ist damit rechnerisch 1/s, wird aber üblicherweise in Hz (= *Hertz*) ausgedrückt. Neben einfachen Sinustönen gibt es auch *komplexe* periodische Signale. Jedes periodische Signal lässt sich – laut der *Fourier-Theorie* – in eine Summe einfacher Sinustöne zerlegen (*Fourier-Analyse*), bzw. durch die Addition mehrerer Sinustöne lässt sich jedes beliebige periodische Signal erzeugen (*Fourier-Synthese*). Die resultierende Periodenfrequenz des komplexen Signals heißt *Grundfrequenz* (F_0), die einzelnen Frequenzkomponenten, aus denen sich das komplexe Signal zusammensetzt, sind die *Harmonischen*. Im einfachsten Fall entspricht die tiefste Frequenzkomponente der Grundfrequenz, und die Harmonischen sind ganzzahlige Vielfache der Grundfrequenz; zusammen bilden sie das *Spektrum*.

In der Natur existieren rein periodische Signale nicht, da sie erstens nie eine unendliche Dauer aufweisen und zweitens immer gewissen Unregelmäßigkeiten unterliegen, so dass sich zwei aufeinander folgende Schwingungen in der Regel mehr oder weniger stark unterscheiden. Diese Signale werden *quasiperiodisch* genannt. Signale, die keinerlei Anzeichen von Periodizität aufweisen, sind *aperiodisch*.

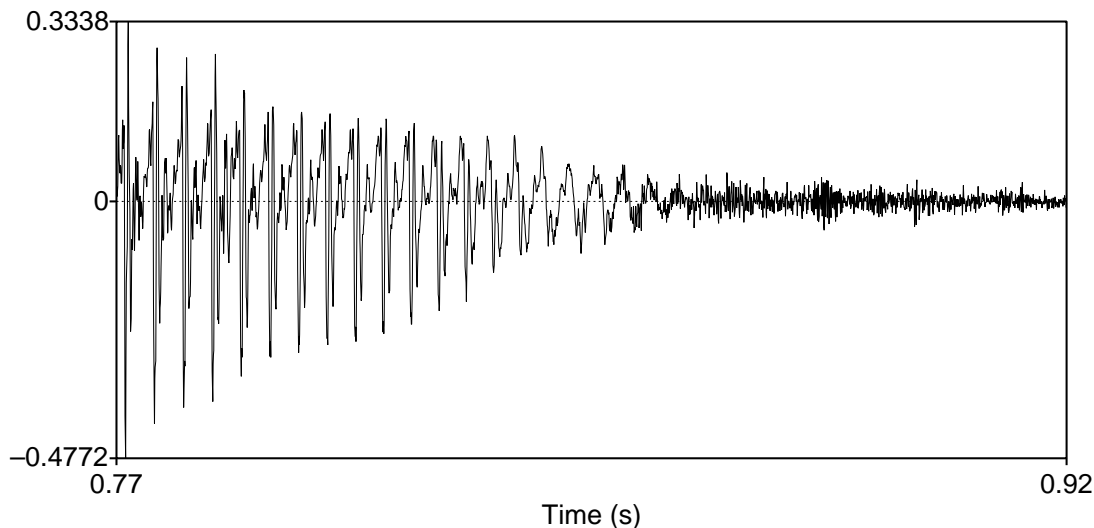


Abbildung 2.2: Beispiel für das Oszillogramm eines Sprachsignals: Ausschnitt aus der Vokal-Konsonant-Verbindung [aχ]. Im linken Teil ist das Signal quasiperiodisch [a]: Es sind regelmäßige Muster (Perioden) zu erkennen, die sich aber deutlich, hier v.a. in der Amplitude, voneinander unterscheiden. Im rechten Teil ist das Signal aperiodisch [χ]: Es sind keine regelmäßigen Muster zu erkennen. (Quelle: Datei *g102a002.116* aus dem *Kiel Corpus of Spontaneous Speech*, Vol. 2, IPDS 1996)

Sprachsignale setzen sich nur aus quasiperiodischen sowie aperiodischen Abschnitten zu-

sammen (vgl. Abbildung 2.2), wobei die quasiperiodischen Abschnitte *stimmhaft*, die aperioidischen Abschnitte *stimmlos* genannt werden. Die Änderung der Grundfrequenz ist *per definitionem* nur in quasiperiodischen Signalen möglich (und nicht etwa in periodischen). Der zeitliche Verlauf der (sich meistens ändernden) Grundfrequenz wird *Grundfrequenz-*, bzw. *F₀-Verlauf* oder *-Kontur* genannt.

Bei der Speicherung eines Sprachsignals in einem Rechner muss das analoge Signal digitalisiert werden. In einem analogen Signal existiert zu jedem Zeitpunkt ein Amplitudenwert (der im Fall eines Sprachsignals den Luftdruck angibt). Es stehen also unendlich viele Werte (auf einer kontinuierlichen Skala) im Zeit-, sowie Amplitudenbereich zur Verfügung. Im digitalen Format existieren nur endlich viele diskrete Werte, so dass bei der Digitalisierung beide Dimensionen durch endlich viele diskrete Werte approximiert werden müssen. Die zeitliche Auflösung (*Abtastrate* in Hz) sowie die *Amplitudenauflösung* (in bit) sollten so hoch gewählt werden, dass der für die Sprache relevante Frequenzbereich sowie der dynamische Bereich ausreichend scharf erfasst werden. Eine Abtastrate von 16 kHz und eine Amplitudenauflösung von 16 bit entsprechen diesen Anforderungen. Diese Werte liegen der Digitalisierung des in dieser Arbeit verwendeten Sprachmaterials zugrunde.

Anhand eines aufgenommenen und digital gespeicherten Sprachsignals kann der Grundfrequenzverlauf der stimmhaften Abschnitte mit Hilfe von Computerprogrammen berechnet werden. In dieser Arbeit wird zur Berechnung des *F₀-Verlaufs* und zur Messung von *F₀-Werten* die Software *xassp* (IPDS 1997b) verwendet. Der implementierte *F₀-Algorithmus* (Schäfer-Vincent 1982; Schäfer-Vincent 1983) arbeitet im Zeitbereich. Er vereinfacht das Signal zuerst auf die Extrema und sucht dann im vereinfachten Signal (bestehend aus den Amplitudenmaxima und -minima) nach Periodizität.

Analysierte *F₀-Verläufe* lassen sich mit geeigneten Computerprogrammen manipulieren. Die manipulierten Werte können mit Hilfe einer Resynthese auf das ursprüngliche Sprachmaterial angewendet werden. Damit steht ein nützliches Werkzeug für die phonetische Forschung zur Verfügung, welches z.B. für die systematische Erstellung von Stimuli (die sich in Parametern des *F₀-Verlaufs* unterscheiden) verwendet werden kann. Davon wird auch in dieser Arbeit Gebrauch gemacht. Es wird hierfür die Software *praat* (Boersma und Weenink URL) verwendet, die mit einer *PSOLA (Pitch-Synchronous Overlap and Add)*-Resynthese arbeitet (Moulines und Charpentier 1990). Beim PSOLA-Verfahren wird das Sprachsignal in Bausteine zerlegt, die jeweils zwei Quasiperioden des Signals umfassen, wobei die zeitliche Mitte eines solchen Bausteins mit einem Amplitudenmaximum (der *führenden Amplitude*) des Signals synchronisiert ist. Die Zerlegung in Bausteine erfolgt mit Hilfe eines *Analysefensters* (Fensterfunktion: \cos^2 , auch *Hanning* genannt). Abbildung 2.3 stellt diese Zerlegung schematisch dar. Bei der *F₀-Manipulation* werden die Bausteine neu zusammengesetzt (verschoben und addiert), wodurch die Periodendauer vergrößert oder verkleinert (bzw. *F₀* verkleinert oder vergrößert) werden kann.

Es muss beachtet werden, dass auch die *F₀-Manipulation* in *praat* auf einer vorheri-

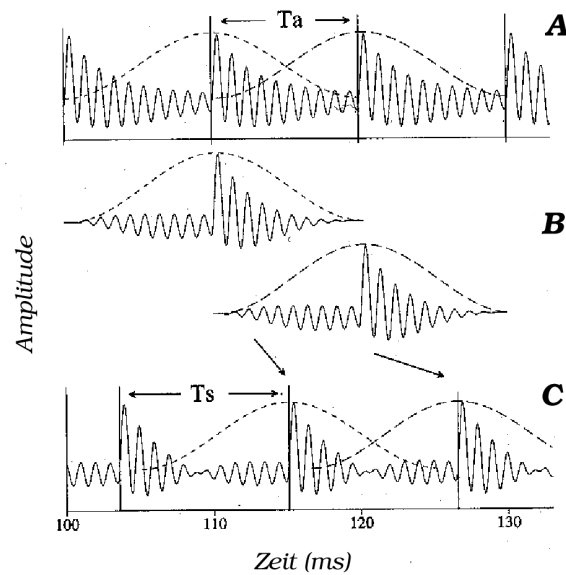


Abbildung 2.3: Prinzipielle Funktionsweise des PSOLA-Verfahrens: (A) \cos^2 -Analysefenster werden mit Maxima des Signals synchronisiert \rightarrow (B) Bausteine (je zwei Quasiperioden), die neu zusammengesetzt werden können \rightarrow (C) Manipulation mit größerer Periodendauer, bzw. kleinerem F_0 als Original. T_a = analysierte Periodendauer; T_s = resynthetisierte Periodendauer.

(Quelle: Kortekaas und Kohlrausch 1997, hier leicht modifiziert)

gen F_0 -Analyse beruht. *praat* verwendet einen anderen F_0 -Algorithmus als *xassp*. Kein F_0 -Analyseverfahren ist perfekt, so dass die Analyseergebnisse zweier Verfahren in der Regel im Detail unterschiedliche Werte liefern. Treten in dieser Arbeit leichte Diskrepanzen auf, zwischen (a) (in *xassp*) gemessenen Werten in der Korpusanalyse (Kapitel 5), z.B. für den F_0 -Wert des finalen Gipfels, und (b) nachmessbaren Werten im resynthetisierten Sprachmaterial des Perceptionsexperiments (Kapitel 6), so beruhen diese mit großer Wahrscheinlichkeit auf den unterschiedlichen F_0 -Analyseverfahren von *xassp* und *praat* (und weniger auf Messfehlern bei der Datenerhebung).

Die akustischen und die technischen Grundlagen des Sprachschalls und seiner computergestützten Verarbeitung sind hier in sehr verkürzter Weise dargestellt worden, gerade so detailliert, wie es für das Verständnis dieser Arbeit sinnvoll und nötig erscheint. Für eine ausführlichere Darstellung der akustischen Grundlagen sei ausdrücklich auf entsprechende Einführungswerke verwiesen, z.B. Johnson (1997) oder Reetz (1999).

2.3 Grundlagen zur Produktion

Der (akustischen) Grundfrequenz des Sprachsignals liegt in der Sprachproduktion die Schwingungsfrequenz der Stimmlippen zugrunde. Dieser Abschnitt führt in die wesentlichen Prin-

zipien der Stimmlippenschwingung ein. Dabei wird zunächst ein Abriss der beteiligten Anatomie gegeben, soweit sie hier von Bedeutung ist (Abschnitte 2.3.1 und 2.3.2). Danach wird auf die Steuerungsmechanismen der Schwingungsfrequenz eingegangen (2.3.3), und schließlich werden einige nicht-intendierte Variationen von F_0 vorgestellt, die als ‘‘Artefakte’’ des Sprachproduktionsprozesses betrachtet werden können (2.3.4). Letztere werden in Anlehnung an Kohler (1990) zusammenfassend als *Mikro- F_0* oder *Mikrointonation*⁴ bezeichnet, während die intendierten Variationen von F_0 kontrastierend *Makrointonation* genannt werden können. Die Ausführungen in den Abschnitten 2.3.1, 2.3.2 und 2.3.3 basieren weitgehend auf Reetz (1999) und Pompino-Marschall (2003).

2.3.1 Die Systeme des Sprechapparats

Die Physiologie des Sprechapparats lässt sich grob in drei Bereiche unterteilen: (1) das *sublaryngale System* (bei Reetz 1999: *das sub-glottale System*), bestehend aus Luftröhre, Bronchien und Lunge, (2) das *laryngale System* (der Kehlkopf) und (3) das *supra-laryngale System*, bestehend aus Rachen-, Nasen- und Mundraum. Alle drei Systeme besitzen primär vegetative Funktionen: Das subglottale System dient der Atmung, das supraglottale System der Atmung sowie der Nahrungsaufnahme, und das laryngale System verhindert, dass Fremdkörper in die Lunge eindringen können. Die meisten Sprachlaute⁵ (bzw. alle Sprachlaute des Deutschen) werden beim Ausatmen produziert. In seiner sekundären Funktion liefert das subglottale System also Energie für die Sprachproduktion, indem es einen *exhalatorischen* (oder *egressiven pulmonalen*) Luftstrom zur Verfügung stellt. Dabei verhält es sich beim Sprechen anders als beim vegetativen Atmen: Die Dauer des Ausatmens wird verlängert und die Dauer des Einatmens verkürzt, so dass längere Redepassagen möglich sind, als es bei dem Dauerverhältnis von Ein- und Ausatmen beim vegetativen Atmen der Fall wäre. Außerdem kann der subglottale Luftdruck variiert werden, was zu einer Variation der Lautstärke, aber auch der Grundfrequenz führen kann. Es wird allerdings angenommen, dass der Effekt auf F_0 nur mittelbar und teilweise abhängig von der laryngalen Aktivität ist, und nicht besonders stark zur Steuerung von F_0 beiträgt (Ohala 1978). Im weiteren Verlauf der Arbeit wird dieses erste System daher vernachlässigt.

Das zweite System dient der *Phonation*. Mit diesem Terminus wird die Erzeugung eines stimmhaften Quellensignals bezeichnet, die im Kehlkopf mit der hindurch strömenden Luft passiert (= *Stimmgebung*). Damit steuert dieses System den Parameter ‘stimmhaft/ stimmlos’ sowie die Stimmqualität — und es ist das wichtigste System zur Steuerung von F_0 . Daher soll es hier etwas ausführlicher behandelt werden (Abschnitt 2.3.2). Das dritte System ist in erster Linie für die *Artikulation* von Bedeutung. Dieser Terminus bezeichnet die Bewegungsabläufe

⁴Die parallele Verwendung dieser Terme suggeriert eine Gleichsetzung von F_0 und Tonhöhe; dieser Aspekt wird in Abschnitt 2.4 diskutiert.

⁵In einigen Sprachen werden auch andere Luftströme verwendet, z.B. ein *glottaler* zur Bildung von Ejektiven.

im supralaryngalen Bereich, die die akustischen Eigenschaften des *Ansatzrohrs* (Hohlraum im supralaryngalen System) steuern und damit das (wechselnd stimmhafte oder stimmlose) Quellensignal filtern, bzw. den Sprachschall “formen”. Dieser (für viele Bereiche der Phonetik wichtige) Prozess ist nicht Gegenstand dieser Arbeit. Es wird daher nur auf einige ausgewählte Aspekte der Artikulation eingegangen, die mit der Entstehung von Mikrointonation zusammenhängen (2.3.4).

2.3.2 Der Aufbau des Larynx

Die Ausführungen in diesem Abschnitt stellen eine Zusammenfassung der wesentlichen Komponenten des *Larynx* (*Kehlkopf*) dar, die der Stimmtoneerzeugung und -konfiguration dienen. Der Kehlkopf ist an seiner Unterseite mit der *Luftröhre* (*Trachea*) verbunden, an seiner Oberseite mit dem *Zungenbein* (*Hyoid*), einem Ringknochen, der die Zungenwurzel stützt (vgl. für den gesamten Abschnitt Abbildung 2.4). Der Kehlkopf besteht aus Knorpeln, Muskeln und Bändern und enthält u.a. die *Stimm lippen*.

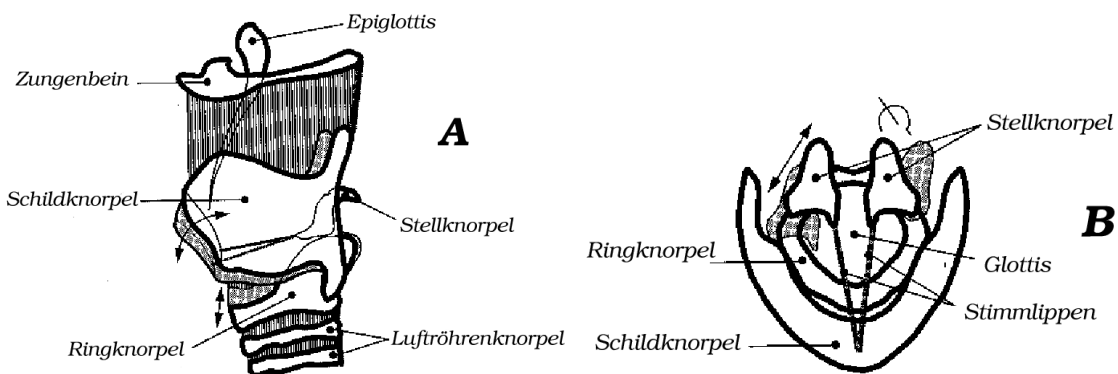


Abbildung 2.4: Schematische Darstellung des Larynx (A) von der Seite, (B) von oben. (Quelle: Pompino-Marschall 2003, hier leicht modifiziert)

Das Grundgerüst des Kehlkopfes bildet der *Schildknorpel* (*Thyroid*), der aus zwei Knorpelplatten zusammengesetzt ist. Unterhalb des Schildknorpels befindet sich der *Ringknorpel* (*Cricoid*), der das Verbindungsstück zwischen Kehlkopf und Luftröhre darstellt. *Cricoid* und *Thyroid* sind verbunden durch die *äußeren Kehlkopfmuskeln* (*Cricothyroid-Muskeln*). Diese rufen bei Kontraktion ein Vorwärtsskippen des Schildknorpels relativ zum Ringknorpel hervor.

Innerhalb des Raumes, der von den beiden Platten des Schildknorpels aufgespannt wird, sitzen auf der Hinterseite des Ringknorpels die paarigen *Stellknorpel* (*Aryknorpel*, *Arytenoïden*). Sie haben je einen Fortsatz (*Processus vocalis*), an dem je eine *Stimm lippe* (*Plica vocalis*) angewachsen ist. Die Stimmlippen bestehen sowohl aus Bändern (*Ligamenti vocali*) als auch Muskeln (*Vocalis-Muskeln*) und erstrecken sich von den Stellknorpeln bis zur Vorderseite, bzw. Spitze des Schildknorpels (die Stelle, an der die beiden Platten zusammengewachsen

sind). Der Raum zwischen den Stimmlippen wird in dieser Arbeit als *Glottis*⁶ bezeichnet.

Parallel zu den *Vocalis-Muskeln* verlaufen an deren Außenseiten die *Thyroarytenoid-Muskeln*, ebenfalls von den Stellknorpeln bis zur Spitze des Schildknorpels. Zusätzlich sind die Aryknorpel über zwei Muskelpaare mit dem Ringknorpel verbunden (*Cricoarytenoideus posterior* und *Cricoarytenoideus lateralis*). Auch die beiden Aryknorpel untereinander sind durch Muskeln miteinander verbunden (*Arytenoideus transversus* und *Arytenoideus obliquus*). Durch ihre muskulären Verbindungen können die Aryknorpel – relativ zum Ringknorpel – eine vorwärts-rückwärts-gerichtete Gleit- sowie eine rotierende Seitwärtsbewegung vollführen.

2.3.3 Phonation und Steuerung der Grundfrequenz

Damit die Phonation einsetzen kann, müssen zwei Bedingungen erfüllt sein. Erstens muss ein relativ konstanter pulmonaler Luftstrom zur Verfügung stehen, zweitens müssen die Stimmlippen in eine bestimmte Position zueinander gebracht werden. Dies geschieht im Wesentlichen durch ein Zusammenbringen (*Adduktion*) der Stimmlippen, welches durch eine entsprechende Stellung der Aryknorpel bei Kontraktion der Muskeln *Arytenoideus transversus* und *obliquus* erreicht wird⁷. Für die *Abduktion* der Stimmlippen, also das Öffnen der Glottis (z.B. zur Erzeugung von stimmlosen Frikativen), wird der *Cricoarytenoideus posterior* verwendet.

Neben einer *modalen* Phonation, die sich im Wesentlichen durch Quasiperiodizität mit vollständigen Verschlussphasen bei der Stimmlippenschwingung auszeichnet, gibt es weitere Phonationsmodi (oder Stimmqualitäten). Diesen liegen modifizierte Einstellungen der Phona-tionsanatomie zugrunde, die hier nicht im Einzelnen behandelt werden sollen. Im experimen-tellen Teil der Arbeit (Kapitel 6) spielt nur die modale Phonation eine Rolle. Erwähnt werden in dieser Arbeit jedoch auch eine *behauchte* sowie eine *geknarrte* Stimmqualität. Die erste zeichnet sich durch eine unvollständige Schließung bei der Stimmlippenschwingung aus, so dass ständig Luft durch die Glottis entweicht. Die zweite weist nur noch annäherungsweise quasiperiodische Schwingungen auf; die Periodendauer kann sehr stark variieren, ebenso die Form und Amplituden der Perioden. Johnson (1997) beschreibt diese drei Modi anhand der anteiligen Dauer der glottalen Öffnungsphase an einer Schwingungsperiode (*Open quotient*). Dabei ist der *Open quotient* für die geknarrte Stimme am kleinsten, gefolgt von der modalen Stimme und schließlich der behauchten Stimme mit dem größten *Open quotient*.

Die (modale) Schwingung der Stimmlippen ist ein komplexer Prozess, der von mehreren Faktoren abhängt. Auf das Wesentliche zusammengefasst lässt er sich wie folgt charakteri-

⁶*Glottis* bezeichnet in dieser Verwendung also kein Organ; es sind auch andere Verwendungsweisen möglich, vgl. Reetz (1999).

⁷Der *Cricoarytenoideus lateralis* führt die Stimmlippen ebenfalls zusammen, allerdings durch eine Rotation der Aryknorpel, wodurch eine kleine Öffnung zwischen den Aryknorpeln entsteht. Diese Stellung wird v.a. für das Flüstern verwendet.

sieren: Durch den subglottalen Luftdruck wird die verschlossene⁸ Glottis “gesprengt”. Die dadurch leicht geöffnete Glottis bildet für den pulmonalen Luftstrom eine Verengung, in der die Luft schneller strömen muss. Dadurch verringert sich der Luftdruck in der Verengung, also zwischen den Stimmlippen, was zur Folge hat, dass diese zusammengesogen werden. Dieser Effekt wird als *Bernoulli-Effekt* bezeichnet. Liegen die Stimmlippen wieder aneinander, beginnt der Prozess von vorn. Dieser Vorgang ist die Basis für die *aerodynamische* Theorie der Stimmlippenschwingung. Allein kann sie jedoch den Prozess nicht vollständig erklären, da auch Eigenschaften der Stimmlippen selbst, nämlich ihre Länge, ihre Elastizität und ihre schwingende Masse die Stimmlippenschwingung maßgeblich beeinflussen, was als *myoelastischer* Effekt bezeichnet wird. Diese beiden Effekte wurden von van den Berg (1958) in der *myoelastisch-aerodynamischen Theorie* zusammengefasst. Hinzu kommen weitere beeinflussende Effekte, z.B. die Tatsache, dass die Stimmlippen keine einheitliche Masse bilden, sondern der obere und der untere Teil sich zeitlich versetzt öffnen und schließen (*Zwei-Massen-Theorie*, Ishizaka und Flanagan 1972).

Zur intendierten Steuerung der Schwingungsfrequenz ist vor allem der myoelastische Effekt von Bedeutung. Bei der modalen Phonation kann die Schwingungs-, bzw. die Grundfrequenz durch die Veränderung der Länge, Elastizität und der schwingenden Masse der Stimmlippen verändert werden (sowie durch eine Veränderung des subglottalen Luftdrucks, vgl. Abschnitt 2.3.1). Dies erfolgt im Wesentlichen durch die Aktion von drei Muskelpaaren, nämlich den *Cricothyroid*-Muskeln, den *Vocalis*-Muskeln und den *Thyroarytenoid*-Muskeln. Diese drei bewirken, z.T. in Kooperation, entsprechende Verformungen der Stimmlippen. Je nach Form (z.B. “gespannt”, “verkürzt”, “dick”) wird der Bernoulli-Effekt unterstützt oder abgeschwächt, so dass die Stimmlippen entsprechend schneller oder langsamer schwingen.

Die *Cricothyroid*-Muskeln bewirken z.B. ein Vorwärtsskippen des Schildknorpels und damit eine erhöhte Längsspannung sowie eine Masseverminderung der Stimmlippen. Dadurch wird der Bernoulli-Effekt unterstützt und F_0 heraufgesetzt. Als F_0 -Senker können dagegen die *Thyroarytenoid*-Muskeln betrachtet werden. Bei ihrer Kontraktion wird die Spannung der Stimmlippen reduziert und die schwingende Masse erhöht. Letztlich dient der intendierten Variation von F_0 stets ein hochdynamisches Zusammenspiel der beteiligten Muskeln (vgl. Sundberg 1979).

2.3.4 Nicht-gesteuerte myoelastische und aerodynamische Einflüsse auf die Grundfrequenz

Aufgrund der bestehenden anatomischen Verbindungen zwischen Larynx und dem supralaryngalen System können Artikulationsbewegungen sich mechanisch auf die laryngale Struk-

⁸Die Stimmlippen werden nicht besonders fest geschlossen, dies geht aber aus Pompino-Marschall (2003) nicht deutlich hervor. Reetz (1999) nennt die zur Phonation nötige Konfiguration der Stimmlippen “*etwas geöffnet und gespannt*”.

tur auswirken (vgl. z.B. 't Hart et al. 1990). Es wurde gezeigt, dass geschlossene Vokale (z.B. [i,u]) ein signifikant höheres F_0 aufweisen können als offene Vokale (z.B. [a]), selbst wenn sie in zusammenhängender Sprache untersucht werden (vgl. für das Deutsche Antoniadis und Strube 1981; Ladd und Silverman 1984). Derartige Phänomene wurden auch für weitere Sprachen gefunden (z.B. Lehiste und Peterson 1961; Reinholt Petersen 1978). Für die F_0 -Differenz (in Halbtönen (*semitones*) = st)⁹ zwischen geschlossenen und offenen Vokalen in verschiedenen prosodischen Kontexten (aus einer vorgelesenen Geschichte) ergaben sich bei Ladd und Silverman (1984) für die einzelnen Sprecher Mittelwerte von 0,69 bis 1,37 st.

Nach der *tongue pull hypothesis* (Lehiste 1970; Ohala 1973) üben geschlossene Vokale eine vertikale Zugkraft auf den Kehlkopf aus, die sich auf die vertikale Spannung der Stimmlippen auswirkt. Dadurch wird F_0 heraufgesetzt. Bei offenen Vokalen wird hingegen ein Druck auf den Kehlkopf ausgeübt, so dass die schwingende Masse der Stimmlippen sich vergrößert und F_0 herabgesetzt wird.

Es wurde ebenfalls gezeigt, dass der F_0 -Verlauf am Beginn eines Vokals vom vorangehenden stimmlosen Konsonanten beeinflusst wird, welcher sowohl eine Affrikate, ein Frikativ oder ein (aspirierter oder unaspirierter) Plosiv sein kann (Löfqvist et al. 1989). Nach einem stimmlosen Konsonanten lässt sich ein höheres Einsetzen von F_0 beobachten als nach stimmhaften Konsonanten. Dieses angehobene F_0 kann sich in einem mikrointonatorisch fallenden F_0 -Verlauf am Vokalbeginn manifestieren, der sich bei Löfqvist et al. (1989) in signifikanter Weise etwa über die ersten sechs Quasiperioden des Vokals erstreckt. Die F_0 -Differenz zwischen dem Beginn und dem Ende des Falls liegt in der Größenordnung von 15 Hz (z.T. auch deutlich höher). Nach einem stimmhaften Konsonanten treten dagegen keine derartig auffälligen F_0 -Perturbationen auf: F_0 kann relativ konstant sein, oder aber leicht fallen oder leicht steigen. Löfqvist et al. (1989) führen diesen Effekt auf eine höhere Aktivität der *Cricothyroid*-Muskeln bei stimmlosen als bei stimmhaften Konsonanten zurück.

Die beiden bisher vorgestellten mikrointonatorischen Effekte lassen sich als myoelastische Effekte zusammenfassen. In anderen Studien (vgl. Hombert et al. 1979) wurden auch aerodynamisch basierte Erklärungsansätze für das angehobene F_0 nach stimmlosen Konsonanten gegeben. Auch Löfqvist et al. (1989) weisen darauf hin, dass die Aktivität des *Cricothyroid*-Muskels nicht die einzige Ursache für diesen Effekt ist. Zudem lassen sich weitere aerodynamische Effekte nennen, die F_0 mikrointonatorisch beeinflussen. Zum Beispiel lässt sich in stimmhaften Konsonanten ein abgesenktes F_0 finden (vgl. Laver 1994). Dieser Effekt lässt sich durch die lokale Verengung im supralaryngalen System und einen damit einhergehenden Luftstau vor dieser Verengung erklären. Der supraglottale Luftdruck steigt an, so dass sich die Druckdifferenz zwischen supra- und subglottalem Luftdruck vermindert. Der Luftstrom durch die Glottis wird damit "abgeschwächt" und die Schwingungsfrequenz herabgesetzt.

⁹Auf die Umrechnung von Hz-Differenzen in st-Differenzen wird in 2.4 eingegangen.

Die mikrointonatorischen Phänomene (von denen hier eine Auswahl präsentiert wurde) überlagern sich mit der Makrointonation, der gesteuerten F_0 -Variation. Es resultiert also ein F_0 -Verlauf, der nicht als akustisches Korrelat der Intonation allein interpretiert werden darf, da neben dem suprasegmentellen Phänomen der Sprechmelodie auch segmentelle Informationen in den F_0 -Verlauf eingehen.

2.4 Grundlagen zur Perzeption

Dieser Abschnitt führt in einige wesentliche Aspekte der Sprachperzeption ein, die im Rahmen der Intonationsforschung von Bedeutung sind. An erster Stelle soll das Konzept der Tonhöhenwahrnehmung in Relation zu der (akustischen, bzw. artikulatorischen) Grundfrequenz behandelt werden (Abschnitt 2.4.1). Zweitens unterliegt die Wahrnehmung von Tonhöhen und Tonhöhenbewegungen gewissen Beschränkungen, die hinsichtlich der Fragestellung von besonderem Interesse sind (2.4.2). Zum Schluss wird eine aktuelle Diskussion zum Phänomen der *kategorialen Wahrnehmung* in Bezug auf die Intonation zusammenfassend präsentiert (2.4.3). Für umfassendere Darstellungen zur Sprachwahrnehmung wird auf einführende Literatur verwiesen, die auch als Basis für die folgenden Abschnitte (2.4.1, 2.4.2) diene, z.B. Borden et al. (1994), Johnson (1997) und Reetz (1999).

2.4.1 Tonhöhe, Grundfrequenz und ihre Skalierung

Die Gleichsetzung der beiden Konzepte Grundfrequenz (F_0) und (wahrgenommene) Tonhöhe (*pitch*) ist unzulässig¹⁰. In einer ersten Annäherung lässt sich lediglich behaupten, dass ein (quasi)periodisches akustisches Signal eine Tonhöhenwahrnehmung hervorruft, wobei die Grundfrequenz des Signals und die Tonhöhe in einer relativ engen Beziehung zueinander stehen: Eine höhere Grundfrequenz resultiert in der Wahrnehmung eines höheren Tons. Bei detaillierterer Betrachtung ist eine Trennung dieser Konzepte jedoch unentbehrlich. Zunächst ist zu erwähnen, dass nicht jede beliebige Grundfrequenz als Ton wahrnehmbar ist: Die tiefste wahrnehmbare Grundfrequenz liegt bei 40 Hz. Bei tieferen Frequenzen werden statt eines Tons die einzelnen Perioden als Impulse wahrgenommen. Ab einer oberen Grenze von etwa 4000 Hz kann die resultierende Tonhöhe nicht mehr eingeschätzt werden (vgl. 't Hart et al. 1990). Außerdem muss ein (quasi)periodisches Signal eine gewisse Mindestdauer aufweisen, damit eine Tonhöhenwahrnehmung erfolgen kann. 't Hart et al. (1990) schlagen einen Wert von 30 ms vor, argumentieren aber, dass dieser Wert auf psychoakustische Untersuchungen zurückzuführen ist. Es könne nicht gefolgert werden, dass bei der Perzeption gesprochener Sprache für jeden Abschnitt von 30 ms eine neue Tonhöhe perzipiert würde.

¹⁰Im Folgenden meint *Tonhöhe* stets das Perzept. Was in 2.1 vereinfachend *erzeugte Tonhöhe* genannt wurde, ist dagegen die Grundfrequenz.

Des Weiteren beeinflusst die *Art* des akustischen Signals die Tonhöhenwahrnehmung. Ein einfacher Sinuston mit einer Grundfrequenz von z.B. $F_0 = 200$ Hz muss nicht dieselbe Tonhöhenwahrnehmung hervorrufen wie ein komplexes (quasi)periodisches Signal mit $F_0 = 200$ Hz. Es ist z.B. möglich, dass bestimmte Harmonische sich gegenüber der Grundfrequenz in der Wahrnehmung als Tonhöhe durchsetzen. Ebenso kann in einem Signal die Grundfrequenz fehlen, z.B. wenn sie, wie bei der Telefonübertragung, herausgefiltert wird. Dennoch kann dieselbe Tonhöhenwahrnehmung entstehen wie bei einer Variante desselben Signals, in der die Grundfrequenz noch vorhanden ist. Terhardt (1974) hat Phänomene dieser Art als *virtual pitch* in einem Wahrnehmungsmodell zusammengefasst. Neben der virtuellen Tonhöhe gibt es eine spektrale Tonhöhe, die als ein perzeptives Abbild der akustischen Grundfrequenz betrachtet werden kann. Für die beiden Arten der Tonhöhe werden grundlegend verschiedene Verarbeitungsprozesse bei der Wahrnehmung angenommen, die hier nicht näher erläutert werden sollen. Eine Folge dieser Annahme verdeutlicht aber die Diskrepanz zwischen Tonhöhe und F_0 : Es ist nach Terhardts (1974) Modell möglich, dass ein einziges komplexes Signal verschiedene Tonhöhenwahrnehmungen hervorruft.

Der Unterschied zwischen den beiden Größen Tonhöhe und F_0 liegt mit anderen Worten in ihrer Natur: Die Grundfrequenz ist eine physikalische, die Tonhöhe eine perzeptorische Größe. Erstere kann objektiv gemessen bzw. anhand eines akustischen Sprachsignals errechnet werden, letztere ist eine subjektive Empfindung, die nur durch Befragen von Versuchspersonen ermittelt werden kann. Ein Maß für die Tonhöhe, welches ausschließlich perzeptorisch definiert ist, ist die mel-Skala. Allerdings basiert sie auf Wahrnehmungsexperimenten mit Sinustönen und berücksichtigt daher nicht die Besonderheiten der Wahrnehmung von komplexen Signalen, wie z.B. Sprachsignalen. Eine weitere perzeptorisch definierte Skala, die aber auf komplexen Signalen basiert, ist die Bark-Skala. Diese Skala berücksichtigt außerdem eine weitere Eigenschaft der Frequenzwahrnehmung: Wir nehmen Frequenzen in *Filterbändern* gruppiert wahr, die für niedrige Frequenzen schmaler sind als für höhere. Umrechnungen von Hz-Werten in diese perzeptorischen Skalen sind annäherungsweise möglich, da anhand der experimentell erhobenen Werte empirische Formeln zu diesen beiden Skalen entwickelt worden sind. Die Abbildung von Hz-Werten auf der Bark-Skala erfolgt dabei in zwei Schritten: Erstens müssen die Hz-Werte in kritische Bänder zusammengefasst, zweitens die derart gruppierten Werte mit der empirischen Formel umgerechnet werden. Der erste Schritt kann dabei auf unterschiedliche Weise vorgenommen werden (Bark_{SPL} und Bark_{ERB} , s. Reetz (1999) für nähere Erläuterungen), was in unterschiedlichen Ergebnissen resultiert, die mehr oder weniger gut die wirkliche Perzeption beschreiben. Letztlich wird durch den Prozess der Umrechnung von Hz-Werten (selbst mit einer empirischen Formel) stets nur eine Annäherung an die Tonhöhenwahrnehmung erreicht.

Der wesentliche Unterschied zwischen der Grundfrequenz und der Tonhöhenwahrnehmung lässt sich auch von einer anderen Seite betrachten: Es kann angenommen werden, dass die Frequenzwahrnehmung des menschlichen Hörapparats wie eine Fourier-Analyse (vgl.

Abschnitt 2.2) funktioniert, mit dem Unterschied, dass die Frequenzwerte (Grundfrequenz und Harmonische) beim Hören nicht auf einer *linearen* Skala, sondern auf einer *logarithmischen* Skala abgebildet werden. Das heißt zwei Töne mit einem konstanten Tonhöhenunterschied werden, wenn es sich um relativ tiefe Töne handelt, einer kleineren F_0 -Differenz entsprechen, als wenn es sich um relativ hohe Töne handelt. Anhand eines konkreten Beispiels bedeutet dies: Zwei Töne mit $F_{0a} = 100$ Hz und $F_{0b} = 200$ Hz weisen denselben Tonhöhenunterschied auf wie zwei weitere Töne $F_{0c} = 600$ Hz und $F_{0d} = 1200$ Hz, obwohl sich für die F_0 -Differenzen ergibt: $\Delta F_{0ba} = 100$ Hz \neq $\Delta F_{0dc} = 600$ Hz. Der Tonhöhenunterschied (auch *Intervall*), der sich wie in diesem Beispiel aus einer Verdoppelung der Grundfrequenz ergibt, wird *Oktave* genannt. Die Oktave wird (in der westlichen Musik) in zwölf Halbtonschritte eingeteilt, ein Halbton in 100 cent.

Die Halbton-Skala stellt also eine dritte Möglichkeit der Beschreibung von Tonhöhen dar. Die wesentlichen Unterschiede zwischen der Halbton-Skala auf der einen Seite und der mel- sowie der Bark-Skala auf der anderen Seite bestehen darin, dass (a) mit der Halbtonskala nur *relative* Tonhöhen (Intervalle), nicht aber *absolute* Tonhöhen dargestellt werden können, und (b) die Halbtonskala nicht rein perceptiv, sondern lediglich durch eine logarithmische Umrechnung von F_0 -Differenzen definiert ist. Zur Umrechnung einer F_0 -Differenz in ein Tonintervall (in Halbtönen) dient folgende Formel:

$$d_{st} = 12 \times \log_2 \left(\frac{M}{R} \right) \quad (2.2)$$

Dabei ist d_{st} das Intervall in st, M (“Messfrequenz”) und R (“Referenzfrequenz”) bezeichnen die beiden beteiligten Grundfrequenzen. In der Intonationsforschung werden die vorgestellten perceptiv definierten Skalen in der Regel nicht verwendet. Gängig ist vielmehr die simple Betrachtung des F_0 -Verlaufs, bzw. die Umrechnung in Halbtöne, wenn F_0 -Differenzen unabhängig von einem absoluten F_0 -Wert von Interesse sind (z.B. Uldall 1960; Hart et al. 1990; Dombrowski 2003). Nolan (2003) führte einen experimentellen Vergleich der hier vorgestellten Tonhöhen-skalen durch. Laut der Ergebnisse können die Halbton- sowie die Bark_{ERB}-Skala die Wahrnehmung der Intonation am besten abbilden.

2.4.2 Wahrnehmbare F_0 -Unterschiede und F_0 -Konturen

Im Rahmen psychoakustischer Experimente wurde in einer Vielzahl von Studien untersucht, wie groß der Unterschied zwischen den Frequenzen zweier einfacher Sinustöne sein muss, damit er gerade noch wahrnehmbar ist. Für solche gerade wahrnehmbaren Unterschiede (*just noticeable difference, JND*) ergeben sich Werte um 1 Hz bei einem F_0 bis etwa 1000 Hz (vgl. Hart 1981; Reetz 1999). Einfache Sinustöne werden anders perzipiert als komplexe Töne (vgl. Abschnitt 2.4.1), so dass derartige Ergebnisse nicht ohne Weiteres auf die Sprachwahrnehmung übertragbar sind. Daher wurden vergleichbare Experimente mit “sprachähnlichem”

Material durchgeführt, in denen z.B. der Sinuston mit einem synthetischen Vokal ersetzt wurde. Lehiste (1970) schlägt auch im Fall von synthetischen Vokalen einen JND von 1 Hz vor. Isačenko und Schädlich (1970) sowie Rossi und Chafcouloff (1972) kommen auf deutlich höhere Werte: 5% von 150 Hz (also 7,5 Hz), bzw. 4% von 195 Hz (also 7,8 Hz). Zusammenfassend beziehen sich all diese Ergebnisse auf den Vergleich zweier stationärer Töne, die mehr (synthetischer Vokal) oder weniger (Sinuston) komplex sind. Die Studien, die mit “sprachähnlichem” Material arbeiten, unterscheiden sich in dieser Hinsicht nicht wesentlich von den rein psychoakustisch orientierten Studien.

Ein weiterer Schritt in Richtung auf die Untersuchung von Tonhöhenwahrnehmung im sprachlichen Bereich besteht in der Einbeziehung von Tonhöhenbewegungen (also *dynamischer* anstelle von *stationärer* Tonhöhe). Ein sich objektiv ändernder F_0 -Verlauf muss nicht als dynamischer, sondern kann u.U. als stationärer Ton wahrgenommen werden. Bedingende Faktoren für die Wahrnehmung eines dynamischen Tons sind die Dauer sowie der Umfang des F_0 -Verlaufs. Die unterste Grenze dieser Parameter für die Wahrnehmung einer Tonhöhenbewegung wird auch *glissando threshold* genannt. In der Studie von Rossi (1971) konnte ein steigender F_0 -Verlauf von 200 ms Dauer erst dann als tonale Aufwärtsbewegung wahrgenommen werden, wenn der Umfang des gesamten Verlaufs mindestens 19 Hz betrug (wobei die Ausgangsfrequenz bei 135 Hz lag). Klatts (1973) Probanden konnten jedoch einen steigenden von einem fallenden F_0 -Verlauf auf einem synthetischen Vokal von 250 ms Dauer unterscheiden, obwohl die Differenz zwischen Start und Ende der (Aufwärts- oder Abwärts-) Bewegung nur 1,5 Hz betrug. Laut der Ergebnisse von Rossi (1971) wäre zu erwarten gewesen, dass derart kleine F_0 -Bewegungen nicht als dynamische Töne wahrgenommen werden und folglich keine Unterscheidung möglich wäre. Andererseits bezogen sich die Ergebnisse von Rossi (1971) nur auf steigende, nicht auf fallende Verläufe. Möglicherweise treten also bei der Kontrastierung eines fallenden vs. eines steigenden Verlaufs (wie bei Klatt 1973) deutlich kleinere JNDs auf als bei der Kontrastierung eines steigenden vs. eines ebenen Verlaufs.

't Hart et al. (1990) fassten die Ergebnisse von mehreren Studien, die auf unterschiedlichen Dauern sowie F_0 -Differenzen basierten, zusammen und fanden eine empirische Formel, die die verschiedenen Ergebnisse dieser heterogenen Studien recht gut erfasst¹¹:

$$g_{thr} = \frac{0,16 \text{ st} \times \text{s}}{T^2} \quad |g_{thr}| = \frac{\text{st}}{\text{s}} \quad (2.3)$$

Dabei ist g_{thr} der *glissando threshold* in st/s (also die “ F_0 -Geschwindigkeit” oder F_0 -Rate in “Halbtöne pro Sekunde”). Aus dieser Formel folgt, dass bei einer Dauer von z.B. $T = 50$ ms eine F_0 -Bewegung von mindestens 64 st/s oder ein F_0 -Umfang von 3,2 st nötig ist, damit die Bewegung wahrgenommen werden kann $(0,16 \text{ st} \times \text{s}/(0,05 \text{ s})^2 = 64 \text{ st/s}$; bzw. $0,05 \text{ s} \times$

¹¹Die Formel ist hier leicht modifiziert, da sie im Original aus mathematischer Perspektive unsauber dargestellt ist: Es würde sich bei korrekter Anwendung der in 't Hart et al. (1990) gelieferten Formel nicht die korrekte Einheit st/s für g_{thr} ergeben, sondern s^{-2} .

64 st/s = 3,2 st). Wird die Dauer verdoppelt, so ist für die Wahrnehmung eines dynamischen Tons nur die Hälfte des F_0 -Umfangs (in st) nötig. Tabelle 2.1 führt die Ergebnisse dreier Beispielrechnungen auf, die diesen Umstand verdeutlichen. Das Ergebnis von Rossi (1971) (19 Hz von 135 Hz bei 200 ms) fällt, gemessen an dieser Formel, zu hoch aus: Es müssten eigentlich kleinere Anstiege bereits als dynamischer Ton wahrnehmbar sein (vgl. 't Hart et al. 1990).

T [ms]	ΔF_0 [st]	g_{thr} [st/s]
50	3,2	64
100	1,6	16
200	0,8	4

Tabelle 2.1: *Glissando thresholds* für drei verschiedene Dauern (T) eines F_0 -Verlaufs. Die F_0 -Rate (g_{thr}) wurde errechnet mit Formel 2.3; die F_0 -Differenz in Halbtönen (ΔF_0) ergibt sich aus $\Delta F_0 = g_{thr} \times T$, z.B. (für $T = 100$ ms) $\Delta F_0 = 16$ st/s $\times 0,1$ s = 1,6 st.

Ein wesentlicher weiterer Schritt in Richtung auf die Untersuchung der Tonhöhenwahrnehmung im sprachlichen Bereich besteht in der Einbeziehung von echtem sprachlichen Material. 't Hart (1981) verwendete viersilbige niederländische Äußerungen (Zahlwörter, z.B. /enən'twintəχ/ 'einundzwanzig'), deren dritte Silbe akzentuiert war (mit einem Gipfel im F_0 -Verlauf). Durch eine systematische Manipulation des F_0 -Verlaufs wurden Stimuli erstellt, die sich (auf das Wesentliche beschränkt dargestellt) im Intervall zwischen dem Minimum und dem Maximum der Gipfelgestalt unterschieden. Es gab 6 verschiedene Intervalle in Schritten von 1 st (1 bis 6 st). Die Probanden sollten beurteilen, welcher von zwei nacheinander präsentierten Stimuli die größere Tonhöhenbewegung enthält. Die Versuchspersonen zeigten drastische Unterschiede bezüglich (a) der Fähigkeit, das größere Intervall zu erkennen und (b) der Entscheidungsstrategie (einige Probanden beurteilten scheinbar nicht das Intervall, sondern die finale Tonhöhe). Insgesamt schließt 't Hart (1981) aus den Ergebnissen, dass nur Unterschiede von mehr als 3 st für die sprachliche Kommunikation relevant sein können.

't Hart (1981) hat in seiner Untersuchung nur den Umfang von Tonhöhenbewegungen berücksichtigt, nicht aber die Bewegungsrate (oder Steilheit der Bewegung, in st/s). Untersuchungen von 't Hart et al. (1990) zur perceptiven Unterscheidbarkeit zweier unterschiedlich steiler F_0 -Verläufe ergaben für den Quotienten aus den beiden F_0 -Raten (g_1, g_2) einen JND von etwa 2. Das heißt eine F_0 -Bewegung muss mindestens "doppelt so steil" wie eine zweite sein, um perceptiv von ihr unterschieden werden zu können.

2.4.3 Kategoriale Wahrnehmung?

Wird (mit Hilfe der Sprachsynthese) ein akustischer Parameter schrittweise verändert, so muss sich ein jeder Schritt nicht notwendigerweise in der Perzeption widerspiegeln. Das klassische Experiment zur Wahrnehmung der Phoneme /b, d, g/ im amerikanischen Englisch von

Liberman et al. (1957) zeigte, dass Hörer das akustische Kontinuum des manipulierten Parameters (F_2)¹² nicht graduell wahrnehmen, sondern eingeteilt in Abschnitte, innerhalb derer sich die Stimuli perzeptiv nicht unterscheiden. Dieses Phänomen wird als *kategoriale Wahrnehmung* bezeichnet.

Der klassische Experimentaufbau enthält zwei Teile (vgl. auch Borden et al. 1994): einen *Identifikationstest* und einen *Diskriminationstest*. In einem Identifikationstest sollen die Probanden je einen Stimulus einer der vorgegebenen Kategorien zuordnen (z.B. /b, d, g/), im Diskriminationstest sollen sie entscheiden, ob zwei nacheinander präsentierte Stimuli gleich oder ungleich sind. Kategoriale Wahrnehmung liegt vor, wenn (a) die Ergebnisse des Identifikationstests einen raschen Wechsel der Hörerurteile von einer Kategorie zur nächsten aufweisen, der sich nur über ein oder zwei Stimuli erstreckt, und zusätzlich (b) sich über diese Kategoriengrenze eine maximale Diskrimination der Stimuli ergibt.

Kohler (1987) wandte das experimentelle Design auf einen Aspekt der deutschen Intonation an und testete die Diskrimination und Identifikation der drei Gipfeltypen ‘früh’, ‘mittel’ und ‘spät’. Bei einer systematischen zeitlichen Verschiebung eines F_0 -Gipfels im Bereich des Vokals der akzentuierten Silbe ergab sich in den Diskriminationstests nur ein Diskriminationsmaximum zwischen dem ‘frühen Gipfel’ und dem ‘mittleren Gipfel’, obwohl Kontextualisierungsexperimente (Kohler 1991b) das Vorhandensein der dritten Kategorie (‘später Gipfel’) bestätigten. Auch in Experimenten zu anderen Bereichen der Intonation blieben Diskriminationsmaxima aus, so etwa bei der Wahrnehmung verschieden hoher F_0 -Gipfel (Ladd und Morten 1997) oder in einem Kontinuum phrasenfinaler Konturen von ‘fallend’ bis ‘steigend’ (Kohler 2004). Ebenso zeigten sich bei der Perzeption von Talkonturen in einer Reihe von Versuchen am IPDS keine Diskriminationsmaxima.

Niebuhr und Kohler (2004) führten ein vergleichendes Experiment zu Gipfel- und Talkonturen durch, in dem sich alle früheren Befunde zur Wahrnehmung dieser beiden Konturtypen bestätigten. Sie argumentieren, dass Identifikationstests (in Form von Kontextualisierungsexperimenten) und Diskriminationstests grundlegend verschiedene Aufgaben für die Versuchsperson darstellen. Während im ersten Fall das Perzept auf kognitive Kategorien semantischer Interpretation abgebildet werden muss, ist der Diskriminationstest enger mit der Wahrnehmung selbst verbunden. Alle drei Gipfelpositionen sind also mit kognitiven Kategorien verbunden, ebenso zwei Talpositionen. Niebuhr und Kohler (2004) schließen, dass das Konzept der kategorialen Wahrnehmung im Bereich der untersuchten Intonationskonturen nicht anwendbar ist. Nur im Falle einer einzigen Kategoriengrenze (von ‘früher Gipfel’ zu ‘später Gipfel’) konnte ein Diskriminationsmaximum festgestellt werden. In diesem Fall findet ein Tonhöhenbruch von ‘hoch’ nach ‘tief’ synchronisiert mit einem spektralen Bruch von ‘Konsonant’ nach ‘Vokal’ statt. Niebuhr und Kohler (2004) schlagen diese besondere Verknüpfung zwischen der Intonation und der segmentellen Ebene als Grund für die kate-

¹² F_2 ist der *zweite Formant*. Formanten sind Resonanzfrequenzen des Ansatzrohrs. Sie werden in dieser Arbeit nicht thematisiert.

goriale Wahrnehmung der beiden Gipfel vor. Kohler (2004) schließt, dass das Konzept der kategorialen Wahrnehmung in Bezug auf seinen Anspruch, phonologische Kontraste in der Wahrnehmung erfassen zu können, revidiert werden muss. Vielmehr werde bereits durch die Identifikationstests eine Kategorienbildung erfasst. Die Einbeziehung der Diskriminationskomponente erfasse einen zusätzlichen, psychophonetischen Aspekt.

Kapitel 3

Phrasenfinale Melodiemuster des deutschen und ihre Beschreibung

Als phrasenfinale Kontur wird der Melodieverlauf zwischen dem letzten Akzent und dem Ende einer Phrase bezeichnet. Um einen Überblick über die phrasenfinalen Melodiemuster des Deutschen geben zu können, ist es notwendig, einen theoretischen Rahmen zur Beschreibung von Melodiemustern zu definieren. Ein solcher Beschreibungsrahmen kann unterschiedliche Qualitäten besitzen. Zum einen kann es sich um ein *Modell* handeln, welches versucht, ein Phänomen (hier die Intonation des Deutschen) in seinen relevanten Aspekten zu erfassen, d.h. mit Hilfe von *Modellkategorien* zu beschreiben und zu erklären. Zum anderen kann es sich um ein *Werkzeug* handeln, welches *Labels* (=Etiketten) zur Verfügung stellt, mit denen Aspekte des Phänomens (hier z.B. Tonhöhen, Melodien, F_0 -Konturen) abgebildet werden können. Die beiden Formen schließen sich nicht aus: Für ein Modell kann ein Etikettierwerkzeug entwickelt werden, welches die Kategorien des Modells repräsentiert und praktisch handhabbar macht, z.B. für die Etikettierung von Sprachdatenbanken.

Es sollen hier zwei verschiedene solcher Rahmen vorgestellt werden, die sich in wesentlichen Aspekten unterscheiden. Der erste Beschreibungsrahmen ist das *Kieler Intonationsmodell* (= KIM, Kohler 1991a) mit dem zugehörigen Etikettierwerkzeug PROLAB (= *Prosodic Labelling*, Kohler 1997). Der zweite Beschreibungsrahmen ist GToBI (= *German Tones and Break Indices*, Grice und Baumann 2002), ein Etikettierwerkzeug, das auf der *Autosegmentell-metrischen Phonologie* (AM-Phonologie, vgl. Ladd 1996) basiert. Während eine Zuordnung ‘Modell → Werkzeug’ im Falle von ‘KIM → PROLAB’ relativ klar vorhanden ist (vgl. 3.3.1), kann eine entsprechende Zuordnung ‘AM-Phonologie → GToBI’ weniger klar postuliert werden (vgl. 3.3.2), da sich in letzterem Fall wesentlich größere Diskrepanzen zwischen Modell und Werkzeug ergeben. Daher wird die AM-Phonologie nicht als Korrelat zu KIM betrachtet. Zur terminologischen Vereinfachung soll im Folgenden das *Werkzeug* GToBI auch als ein *Modell* bezeichnet werden, wenn es mit KIM kontrastiert wird.

Da gewisse Details der beiden Ansätze KIM/PROLAB und GToBI (z.B. das genaue Re-

pertoire an Beschreibungskategorien oder Symboldiakritika) für das Thema der Arbeit nicht relevant sind, wird hier lediglich auf die generellen Unterschiede sowie auf die Behandlung der phrasenfinalen Melodien eingegangen. Dabei wird im Fall von KIM/PROLAB auf eine Beschreibung des eigentlichen Modells (KIM) weitgehend verzichtet und stattdessen vorrangig mit den PROLAB-Labels gearbeitet, damit eine vergleichbare Beschreibungsebene (die "Etiketten"-Ebene) vorliegt wie bei GToBI. Es wird jedoch versucht, den wesentlichen Unterschied zwischen einer KIM-Kategorie einerseits und einem PROLAB-Label andererseits darzustellen.

Die vorliegende Arbeit untersucht ein Melodiemuster, welches weder in GToBI noch in das KIM integriert, sondern lediglich mit PROLAB etikettierbar ist. Daher versteht diese Untersuchung sich prinzipiell losgelöst von einem speziellen Intonationsmodell. Es gibt dennoch drei Motivationen für eine Kurzvorstellung der beiden genannten Modelle an dieser Stelle. Erstens basieren die bisherigen Beschreibungen der pseudoterminalen Kontur auf Sprachdaten, die mit PROLAB etikettiert worden sind (vgl. Peters 1999; Peters 2000). Ein grundlegender Einblick in KIM und Kenntnis der PROLAB-Etiketten ist also wichtig für das Verständnis sowohl der Hypothesen (Kapitel 4) als auch der Korpusanalyse (Kapitel 5). Das am weitesten verbreitete Etikettierwerkzeug für die deutsche Intonation ist GToBI. Es soll daher zweitens versucht werden, die in den hier vorgenommenen Untersuchungen berücksichtigten Melodiemuster – so weit wie möglich – auch durch GToBI-Labels abzubilden. Drittens unterscheiden sich KIM/PROLAB und GToBI sowohl in grundlegenden Annahmen, als auch in ihren Label-repertoires bezüglich der phrasenfinalen Muster. Eine kurze vergleichende Diskussion dieser Aspekte ist interessant, weil sich in Bezug auf die Möglichkeiten, die die beiden Modelle bieten, das pseudoterminale Muster adäquat zu beschreiben und gegebenenfalls zu integrieren, wichtige Unterschiede ergeben. Die Notwendigkeit einer Integration hängt allerdings nur zum Teil von der tatsächlichen Relevanz des pseudoterminalen Musters ab, zum anderen nämlich auch von der Zielsetzung des Modells.

Das Kapitel beginnt mit einem Überblick über die klassische Einteilung der phrasenfinalen Konturen des Deutschen in drei *Intonationstypen*, die von Essen (1964) vorgeschlagen hat (Abschnitt 3.1). Danach sollen in Abschnitt 3.2 einige grundsätzliche Erwägungen zusammengefasst werden, die bei der Modellierung der Intonation eine Rolle spielen und für die Interpretation der Ergebnisse dieser Arbeit im Rahmen der (beiden hier vorgestellten) Intonationsmodelle von unmittelbarer Bedeutung sind. Die Kurzvorstellung der beiden Modelle erfolgt in Abschnitt 3.3. Anschließend wird präsentiert, welche phrasenfinalen Melodiemuster von den beiden Modellen erfasst und welche funktionalen und formalen Eigenschaften für sie angenommen werden (3.4 und 3.5).

3.1 Drei Intonationstypen

In seiner impressionistischen Beschreibung der deutschen Intonation unterscheidet von Essen (1964) drei verschiedene Intonationstypen, in deren Bezeichnungen sich ihre Funktionen widerspiegeln: der *terminale*, der *progređiente* und der *interrogative*. Diese Intonationstypen können nicht ohne Weiteres als phrasenfinale Konturen im hier definierten Sinne bezeichnet werden, da sie sich auch im Melodieverlauf *vor* dem finalen Akzent unterscheiden können. Die auffälligsten Unterschiede ergeben sich aber in der Tat *nach* dem letzten Akzent der Phrase (bei von Essen 1964 im *Nachlauf*). Kurz zusammengefasst wird der terminale Typus charakterisiert durch einen “geschlossen tieftönige[n] Nachlauf oder schnelle[n] Tonabfall in der Schwerpunktsilbe selbst”, *progređiente* Äußerungen “laufen mit gehobener Stimme aus” und bei der *interrogativen* Intonation folgen auf den finalen Akzent “nur noch ansteigende Silbentöne als Nachlauf” (von Essen 1964). Die Bezeichnungen der drei Typen dürfen allerdings nicht zu wörtlich genommen werden: Die pragmatische Funktion der Frage kann einer Äußerung z.B. nicht nur durch eine *interrogative* Intonation verliehen werden. Die Anwendungen der drei Typen fasst von Essen (1964: 65f) wie folgt zusammen:

“Die *terminale* Intonation wird angewandt in Aussagen, Aufforderungen, Ausrufen, vorangestellten Anreden, Ergänzungsfragen, indirekten Reden, den zweiten Gliedern von Doppelfragen. Die *progređiente* Intonation wird angewandt in allen unvollendeten Redeteilen, denen die Schwerpunktsbildung erst folgen soll. [...] Die *interrogative* Intonation wird angewandt in Entscheidungsfragen, Nachfragen [...], ferner – als Variante – in Aussagen, Aufforderungen und Ergänzungsfragen, wenn sie Warnungen oder Drohungen darstellen sollen, in geringer melodischer Bewegung auch bei Aufforderungen als Ausdruck der Höflichkeit. [...]” (von Essen 1964, Hervorhebungen im Original)

3.2 Zur Modellierung der Intonation

Ein Rahmen zur Beschreibung der Intonation muss den Terminus Intonation definieren. Es können unterschiedliche suprasegmentelle Parameter (vgl. Abschnitt 2.1) sowie unterschiedliche sprachliche Phänomenbereiche der Intonation zugerechnet werden. Auch kann der wichtigste Aspekt der Intonation, die tonale Komponente, hinsichtlich der Zielsetzung des Beschreibungsrahmens sowie seiner methodischen Grundlage unterschiedlich analysiert werden. Ansätze der Analyse von Intonation scheinen sich in einer groben Annäherung in Bezug auf vier Grundfragen zu unterscheiden:

- A Wie lautet die genaue Zielsetzung des Beschreibungsrahmens?
- B Wie wird der zu beschreibende Phänomenbereich eingegrenzt?

C Wie werden die Beschreibungskategorien definiert?

D Wie wird der Rahmen erarbeitet?

Durch diese vier Fragen soll keine vollständige Klassifizierung aller Intonationsmodelle ermöglicht werden, sondern es sollen anhand dieser Fragen diejenigen Aspekte von Intonationsmodellen diskutiert werden, die für das hier untersuchte Phänomen von Bedeutung sind. Mit dieser Grundlage, die in den folgenden vier Abschnitten geschaffen wird, kann versucht werden, erstens die beiden vorgestellten Modelle in diesen Rahmen einzuordnen, und zweitens das untersuchte Phänomen im Rahmen der beiden Modelle zu bewerten.

3.2.1 Die Zielsetzung eines Beschreibungsrahmens

Im einfachsten Fall kann es sich bei dem Beschreibungsrahmen um ein simples Werkzeug, z.B. für die Etikettierung der Intonation in Sprachdatenbanken handeln. Stünde hinter diesem Werkzeug kein theoretisches Modell, so wäre es vermutlich in der Lage, sehr feine Unterschiede in der Intonation aufzuzeichnen, unabhängig davon, ob diese Unterschiede eine Relevanz für die sprachliche Kommunikation besitzen. Damit wäre es ein suprasegmentelles Korrelat zu dem schwerpunktmäßig segmentell orientierten Internationalen Phonetischen Alphabet (IPA 1999).

Eine andere mögliche Zielsetzung eines Beschreibungsrahmens besteht darin, genau diejenigen Muster der Intonation einer Sprache (oder auch aller Sprachen) beschreiben zu können, die für die *sprachliche Kommunikation relevant* oder aber *in linguistischer Weise strukturierbar* sind (vgl. 3.2.2). Derartige Rahmen sollen hier als (*phonologische*) Modelle bezeichnet werden. Sie stehen im Mittelpunkt der folgenden Ausführungen.

3.2.2 Der Phänomenbereich eines Modells

Die genaue Zielsetzung des Modells hängt z.T. zusammen mit der Eingrenzung des Phänomenbereichs des Modells: Den Phänomenbereich eines simplen Werkzeugs bilden sämtliche, oder zumindest sehr viele Ereignisse der Intonation; der Phänomenbereich eines phonologischen Intonationsmodells hingegen besteht nur aus einem kleinen Ausschnitt daraus. Ladd (1996:6) definiert Intonation entsprechend als “the use of *suprasegmental* phonetic features to convey ‘postlexical’ or *sentence level* pragmatic meanings in a *linguistically structured* way” (Hervorhebungen im Original; es folgt bei Ladd (1996) eine ausführliche Definition zu jedem kursiven Terminus).

Der wesentliche Unterschied zu den übrigen präsentierten Definitionen (vgl. 2.1) besteht in der Einschränkung, dass nur diejenigen Aspekte der Sprechmelodie zum Untersuchungsgegenstand zählen, die in einer linguistischen Weise strukturiert sind. Ladd (1996) versucht

zu diesem Zweck eine klare Trennung zwischen *Linguistischem* auf der einen Seite und *Paralinguistischem* auf der anderen Seite. Linguistisch sind demnach nur diejenigen Merkmale der Intonation, die in einer kategorialen Weise strukturiert sind, z.B. können die Merkmale ‘hoher Ton’–‘tiefer Ton’ solche Kategorien bilden, oder auch ‘fallend’–‘steigend’. Paralinguistisch sind dagegen diejenigen Merkmale, die durch kontinuierlich variierbare physikalische Parameter definierbar sind. Beispiele sind die Lautstärke, das Sprechtempo oder auch der Tonhöhenumfang, z.B. einer Gipfelkontur. Entsprechend lassen sich die pragmatischen Funktionen von linguistischen Merkmalen als *funktionale Kategorien* verstehen, die pragmatischen Funktionen von paralinguistischen Merkmalen hingegen als Signalisierung von *kontinuierlich variierbaren Sprecherattitüden* oder *Emotionen*. Beispiele für Paare von funktionalen Kategorien sind ‘Frage’–‘Aussage’ oder ‘fokussiert’–‘nicht fokussiert’. Beispiele für kontinuierlich variierbare Sprecherattitüden sind ‘wütend’, ‘freundlich’, ‘überrascht’, usw. Mit anderen Worten gilt also: Sowohl hinsichtlich der Form als auch hinsichtlich der Funktion sind linguistische Merkmale definiert durch ein “Entweder-Oder”, paralinguistische Merkmale hingegen durch ein “Mehr-Oder-Weniger”.

Ein Ansatz, der eine strikte Trennung zwischen Linguistischem und Paralinguistischem versucht, ist keineswegs unproblematisch. Ladd (1996) selbst räumt ein, dass Unklarheit und Uneinigkeit darüber aufkommen kann, welche Merkmale linguistischer und welche paralinguistischer Natur sind. So werden zwei Melodiemuster des Englischen, ‘high fall’ und ‘low fall’ in einigen Beschreibungen als zwei verschiedene Intonationskategorien betrachtet (z.B. O’Connor und Arnold 1973), in anderen Beschreibungen (z.B. Crystal 1969) als paralinguistische Variationen ein und derselben Kategorie ‘fall’. Ladd (1996:39) merkt zu diesem Beispiel an:

“Since it is genuinely difficult to tell paralinguistic and [linguistic] intonational messages apart, and since there is no obvious independent body of theory nor standard of evidence to which one might appeal in order to decide issues like this, these disagreements have tended to remain unresolved.”

Es stellt sich die Frage, inwiefern eine strikte Trennung zwischen Linguistischem und Paralinguistischem möglich bzw. nötig ist. Ein Modell, welches sich als phonologisch betrachtet, kann durchaus einen breiteren Phänomenbereich definieren. In KIM werden z.B. drei verschiedene Gipfelpositionen unterschieden (‘früh’, ‘mittel’, ‘spät’), wobei der Übergang von ‘mittel’ zu ‘spät’ kontinuierlich ist (vgl. 2.4.3). Die Funktion des Merkmals ‘spät’ ist im Wesentlichen der Ausdruck von Überraschung (Kohler 1991a; Kohler 1991b), die nach Ladd (1996) paralinguistischer Natur ist (s.o.). Ein Modell wie das KIM arbeitet also nicht auf der Basis von *linguistischen* Funktionen im Sinne von Ladd (1996), sondern vielmehr auf der Basis *sprachlich-kommunikativer Funktionen*.

Es sei erwähnt, dass auch dieser Ansatz nicht frei von Problemen ist. Als Beispiel dienen die Variation des Tonhöhenumfangs einer Gipfelkontur, die nach Ladd (1996) klar pa-

ralinguistisch ist: Durch einen größeren oder kleineren Umfang wird der Akzent mit mehr oder weniger *Emphase* versehen. Die *Emphase* ist also aus einer linguistischen Sicht nach Ladd (1996) nicht relevant und kein Merkmal der (phonologischen) Intonation. Andererseits besteht kein Zweifel darin, dass ein deutlich emphatischer Akzent – realisiert durch einen sehr großen Tonhöhenumfang – eine andere sprachlich-kommunikative Funktion erfüllt als ein nicht-emphatischer Akzent – realisiert durch einen deutlich kleineren Tonhöhenumfang. Es ist daher plausibel, diesen Umstand durch ein Merkmal in einem Intonationsmodell zu verankern (z.B. \pm EMPH in KIM). Dann stellt sich die Frage, inwiefern sich unterschiedliche *Grade* der *Emphase* – die sich sowohl messtechnisch als auch auditiv im F_0 - bzw. Tonhöhenumfang manifestieren – in ihrer sprachlich-kommunikativen Funktion unterscheiden. KIM setzte ursprünglich sieben Stufen an, in das später entwickelte PROLAB ist diese Feindifferenzierung innerhalb der emphatischen Kategorie allerdings nicht eingegangen. Es ist natürlich denkbar, dass sich – ausgehend von einem emphatischen Akzent – durch eine weitere leichte Verstärkung der *Emphase* die kommunikative Funktion der Äußerung ebenso leicht verändert. Die Schwierigkeit liegt darin, zu entscheiden, wie groß derartige Veränderungen sein müssen, damit sie als “sprachlich-kommunikativ relevant” bezeichnet werden dürfen. Da keine empirischen Befunde zu einer möglichen sprachlich-kommunikativ relevanten Differenzierung verschiedener *Emphasegrade* vorliegen, erscheint die Lösung in PROLAB (gegenüber der Lösung sowohl im ursprünglichen KIM als auch bei Ladd 1996) am plausibelsten.

3.2.3 Die Kategorien eines Modells

Die Beschreibungskategorien eines Intonationsmodells können auf der Basis verschiedener Eigenschaften der Intonation definiert werden. Es seien zwei Möglichkeiten exemplifiziert. Die erste nimmt an, dass *Zieltöne* die Bausteine der Intonation bilden (z.B. Ladd 1996). Jedem Melodiemuster liegt in der phonologischen Repräsentation eine Abfolge der verfügbaren *Zieltöne* zugrunde, zwischen denen bei der phonetischen Umsetzung interpoliert wird, so dass an der phonetischen Oberfläche ein Tonhöhenverlauf erscheint. Eine andere Möglichkeit ist die Auffassung, dass ganzheitliche Intonationskonturen wie ‘Gipfelkonturen’ oder ‘Talkonturen’ die zugrundeliegenden Einheiten sind.

3.2.4 Die methodische Basis eines Modells

Ladd (1996) unterscheidet zwischen zwei traditionellen, essentiell verschiedenen Ansätzen, in die sich die Intonationsforschung bis in die späten 1970er Jahre einteilen ließ: der *instrumentelle* oder *phonetische* vs. der *impressionistische* oder *proto-phonologische*¹. Diese Einteilung bezieht sich auf die Methoden zur Analyse von gesprochenen (bzw. aufgenommenen)

¹Keiner der beiden Ansätze war nach Ladd (1996) phonologisch.

menen) Intonationskonturen: Intonationsmuster lassen sich sowohl durch analytisches Hören (*impressionistischer Ansatz*) als auch durch das Vermessen von F_0 -Konturen (*instrumenteller Ansatz*) auffinden und beschreiben. Dabei werden unterschiedliche definierende Parameter der Intonation angesetzt: die (wahrgenommene) Tonhöhe im einen, die Grundfrequenz des akustischen Signals im anderen Fall (vgl. Abschnitt 2.4).

Sind – auf die eine oder andere Weise – Hypothesen zur Intonation einer Sprache oder auch zu bestimmten Melodiemustern aufgestellt worden, so ist es sinnvoll, diese Hypothesen zu testen. Hier kommt ein dritter, auf den ersten beiden Ansätzen aufbauender, methodischer Ansatz ins Spiel: der *experimentelle*. Typischerweise werden Perzeptionsexperimente mit einer Reihe naiver Versuchspersonen durchgeführt, deren Aufgabe darin besteht, auditive Stimuli kommunikativen Funktionen zuzuordnen, oder zu entscheiden, ob sich zwei Stimuli identisch anhören (z.B. Uldall 1960; Kohler 1991b).

In der Intonationsforschung ist ein experimentelles Vorgehen keineswegs der Regelfall. Häufig werden Intonationsmodelle in erster Linie auf Basis der muttersprachlichen Intuition der Modellierenden, oder auch kombiniert mit empirischen Befunden aus Sprachdatenbanken aufgestellt, bzw. modifiziert. Letztlich können aus einem solchen Vorgehen stets nur Hypothesen darüber resultieren, welche Muster im Rahmen der Zielsetzung und des Phänomenbereichs des Modells relevant sind. Ein zusätzliches experimentelles Vorgehen zum Testen dieser Hypothesen ist daher nicht nur sinnvoll, sondern notwendig, wenn den Kategorien eines Modells eine objektive Basis zugrunde liegen soll.

3.3 Zwei Beschreibungsrahmen – GToBI und KIM/PROLAB

Dieser Abschnitt bietet zuerst je eine Kurzeinführung in GToBI und KIM/PROLAB (Abschnitte 3.3.1 und 3.3.2), sowie einen direkten Vergleich bezüglich einiger grundlegender Konzepte (3.3.3). Dies geschieht im Wesentlichen durch die Anwendung der Klassifikationskriterien aus 3.2.

3.3.1 KIM/PROLAB

Das *Kieler Intonationsmodell* (KIM) (Kohler 1991a) versteht sich als ein phonologisches Modell. Der Phänomenbereich ist nicht das rein Linguistische im Sinne von Ladd (1996), sondern vielmehr das sprachlich-kommunikativ Relevante (vgl. 3.2.2). KIM beinhaltet die Auffassung, dass ganzheitliche Intonationskonturen die zugrundeliegenden Bausteine der Intonation bilden. Sie werden in dieser Arbeit auch als *KIM-Kategorien* bezeichnet. In der phonologischen Beschreibung sind diese Kategorien durch distinktive Merkmale (mit zwei oder mehr Ausprägungen) repräsentiert. Das Merkmal <±VALLEY> unterscheidet z.B. eine Gipfel- von einer Talkontur, mit den Merkmalen <±EARLY> und <±LATE> können frühe, mittlere und späte Gipfel unterschieden werden. Charakteristisch für den frühen Gipfel ist

die Synchronisation des F_0 -Gipfels *vor* dem Vokal der akzentuierten Silbe, während der F_0 -Verlauf über diesem Vokal bereits fällt. Im mittleren Gipfel wird das F_0 -Maximum erst während des akzentuierten Vokals erreicht, im späten erst gegen Ende des akzentuierten Vokals oder später.

Ein Beispiel für eine KIM-Kategorie, repräsentiert durch Merkmale, wäre $\langle -\text{EMPH}, -\text{TERMIN}, -\text{VALLEY}, -\text{QUEST}, -\text{EARLY}, -\text{LATE} \rangle$ ('nicht-emphatischer Akzent mit mittlerem Gipfel und finaler fallend-tief-steigender Kontur'). Vereinfachend werden auch unvollständige Spezifikationen wie 'mittlerer Gipfel' als KIM-Kategorien bezeichnet. Die Hypothesen zu den KIM-Kategorien basieren auf einer Kombination von instrumenteller und impressionistischer Herangehensweise. Die wahrgenommenen Tonhöhenverläufe können also einerseits als definierender Parameter der Intonation genannt werden, andererseits wird auch der F_0 -Verlauf als instrumentell erfassbare Annäherung verwendet. Dabei wird eine strikte Trennung zwischen Makrointonation und Mikrointonation angestrebt, die auch in eine Rechnerimplementierung des KIM eingegangen ist (Kohler 1990). Die Hypothesen zu den KIM-Kategorien wurden teilweise mit experimentellen Befunden untermauert (Kohler 1991b).

Mit dem Ziel, eine prosodische Etikettierung von Sprachdatenbanken mit den Kategorien des KIM zu ermöglichen, wurde das Notationssystem PROLAB entwickelt (Kohler 1992; Kohler 1997). Grundsätzlich muss eine strikte Trennung zwischen den *Kategorien des KIM* auf der einen Seite und den *Etiketten des PROLAB-Systems* auf der anderen Seite vorgenommen werden. Eine *KIM-Kategorie* wird nicht durch genau ein *PROLAB-Etikett* repräsentiert, sondern durch ein Bündel von Etiketten, wie auch in KIM selbst eine Kategorie durch ein Bündel von Merkmalen repräsentiert wird. Allerdings entspricht auch ein *Merkmal* des KIM nicht einem *PROLAB-Etikett*: Zum Beispiel gibt es in KIM ein binäres Merkmal zur Unterscheidung von 'Gipfel' und 'Tal' und zwei weitere zur Unterscheidung von 'früh', 'mittel' und 'spät'; in PROLAB gibt es jedoch für jede der resultierenden Spezifikationen ein eigenes Etikett, z.B. $\text{KIM} \langle -\text{VALLEY}, -\text{EARLY}, -\text{LATE} \rangle = \text{PROLAB} [\hat{\quad}]^2$. PROLAB wurde nach der Formulierung des KIM schrittweise ausgebaut. Dies hat den weiteren Unterschied zwischen KIM und PROLAB zur Folge, dass durch PROLAB mehr Melodiemuster etikettiert werden können, als in KIM formuliert sind. In dieser Arbeit wird weitgehend mit PROLAB-Symbolisierungen gearbeitet, die – sofern Entsprechungen in KIM existieren – vereinfachend auch gleichbedeutend mit den KIM-Kategorien verwendet werden.

Für akzentuierte Wörter unterscheidet PROLAB drei Stufen der Akzentuierung. Es gibt eine Defaultakzentuierung, die als Stufe 2 bezeichnet wird (PROLAB-Symbol [2]). Von diesem Defaultwert kann ein Akzent in zwei Richtungen abweichen. Erstens kann es sich um einen emphatischen Akzent handeln (Akzentstufe [3]), zweitens kann das akzentuierte Wort partiell deakzentuiert sein (Akzentstufe [1]). Im letzteren Fall muss sich der Akzent der Stufe

²PROLAB und GToBI-Labels werden in dieser Arbeit durch eckige Klammern [] markiert.

[1] einem anderen Akzent in derselben prosodischen Phrase unterordnen, kann also nur relativ zu einem Defaultakzent auftreten. Ein Akzent der Stufe [1] wird daher in dieser Arbeit auch als *Nebenakzent* bezeichnet.

Ein Akzent (jeder Akzentstufe) ist mit einem tonalen Muster assoziiert. Der *frühe*, der *mittlere* und der *späte Gipfel* werden in PROLAB durch [)], [^] und [(] symbolisiert. In Kombination mit einem Defaultakzent ergeben sich also folgende Etiketten: für einen frühen Gipfel [2)], einen mittleren Gipfel [2 ^] und einen späten Gipfel [2 (]. Tritt ein Akzent als letzter in der prosodischen Phrase auf, so folgt auf die PROLAB-Etikettierung der Akzentstärke und des Konturtyps (z.B. [2 ()] ein Etikett für die phrasenfinale F_0 -Kontur. Diese werden in Abschnitt 3.4 thematisiert. Es gibt verschiedene PROLAB-Etiketten, die eine Phrasengrenze symbolisieren; die Unterschiede sind aber technischer, nicht phänomenaler Natur, das Default-Label ist [pgn]. Grundsätzlich unterscheidet KIM nicht zwischen verschiedenen Arten oder Graden der prosodischen Phrasierung.

3.3.2 GToBI

GToBI ist eine Variante des Etikettierwerkzeugs ToBI (= *Tones and Break Indices*, Beckman und Ayers Elam 1997), das ursprünglich für die Etikettierung der Intonation des Englischen entwickelt wurde und auf der AM-Phonologie basiert. Den Ausgangspunkt für die Intonationsforschung im AM-Rahmen bilden klar phonologisch orientierte Studien (Lieberman 1975; Bruce 1977; Pierrehumbert 1980). Damit kann GToBI als ein phonologisch orientiertes Werkzeug betrachtet werden. Der Phänomenbereich des GToBI geht aus der Literatur allerdings nicht klar hervor. Einerseits handelt es sich um ein Werkzeug zur Etikettierung der Intonation im Rahmen der AM-Phonologie. Es ließe sich also vermuten, dass eine ähnliche Definition von Intonation zugrunde liegt wie bei Ladd (1996). Andererseits stellt das GToBI-System durchaus Etiketten zur Verfügung, durch die ein später Gipfel von einem mittleren Gipfel³ differenziert werden kann — obwohl eine funktionale Differenzierung ‘überrascht/ nicht-überrascht’ nach Ladd (1996) paralinguistisch ist (vgl. 3.2.2). Auch Aspekte der emphatischen Sprechweise können mit GToBI etikettiert werden (Grice und Baumann 2002), sowie ein Melodiemuster, mit dem Höflichkeit zum Ausdruck gebracht wird (vgl. 3.5). GToBI kann daher nicht als eine direkte Umsetzung der AM-Phonologie betrachtet werden (vgl. Anfang des Kapitels).

Versteht man GToBI in der Tradition von Pierrehumbert (1980), so basiert es auf einer instrumentellen Herangehensweise an die Intonation, in der F_0 den definierenden Parameter darstellt. Nach Grice und Baumann (2002) werden allerdings *wahrgenommene Konturen* etikettiert, wobei dennoch mit F_0 -Verläufen gearbeitet wird. Als zugrundeliegende Bausteine der Intonation werden in der AM-Phonologie zwei Zieltöne angesehen: H (‘high’) und L

³Im Rahmen der AM-Phonologie darf *Gipfel* nicht als *Gipfelkontur* verstanden werden, sondern ist als phonetische Umsetzung des phonologischen H-Zieltons zu betrachten.

(‘low’). Sämtliche Melodiemuster werden durch diese Töne beschrieben, wobei angenommen wird, dass der Melodieverlauf zwischen diesen Tönen per Regel (meist linear) interpoliert wird.

Sowohl Akzentkonturen als auch phrasenfinale Konturen setzen sich aus den zwei Ziel-tönen [H] und [L] zusammen (wobei Konturen nur das Resultat der Interpolation darstellen; vereinfachend wird dennoch von Konturen gesprochen). In Akzentkonturen ist genau ein Ton mit der akzentuierten Silbe assoziiert, was in der Etikettierung durch das Symbol [*] markiert wird. Einfache Akzente enthalten nur einen Ton: [H*] oder [L*]. Es sind verschiedene Kombinationen von Tönen möglich, die in der Symbolisierung mit einem [+] verbunden werden: z.B. [L+H*], [L*+H]. In GToBI lassen sich zwei verschiedene Arten von Phrasen etikettieren: eine *Intonationsphrase* (IP) und eine *Intermediärphrase* (ip). Eine IP enthält immer eine oder mehrere ip. Für die Phrasengrenzen werden ebenfalls Zieltöne angesetzt (*Grenztöne*). Der Grenztöne einer ip wird mit einem [-] symbolisiert (z.B. [H-]), der Grenztöne einer IP mit einem [%] (z.B. [L-%]). Am Ende einer IP tritt immer eine Kombination von ip-Grenze und IP-Grenze und damit auch eine Kombination zweier Grenztöne auf. Die ursprünglichen ToBI-Labels sind in dieser Hinsicht transparent (z.B. [L-L-%]). In GToBI wurden diesbezüglich Veränderungen vorgenommen, so gilt z.B. ToBI [L-L-%] = GToBI [L-%]. Auf die Kombinationen, die für die Fragestellung von Interesse sind, wird in Abschnitt 3.4 eingegangen.

3.3.3 KIM/PROLAB vs. GToBI im Überblick

Das KIM entstammt eher einer perzeptorisch orientierten Sichtweise auf die Intonation, GToBI eher einer theoretisch-linguistischen. Bezüglich der Zielsetzung von KIM/PROLAB und GToBI sind die beiden Modelle vergleichbar: Sie verstehen sich als phonologische Modelle, bzw. phonologisch orientierte Werkzeuge. Der Phänomenbereich, der als relevant für eine phonologische Beschreibung der deutschen Intonation betrachtet wird, ist in den beiden Modellen offenbar nicht derselbe, zumindest wenn GToBI im Rahmen der AM-Phonologie, wie Ladd (1996) sie darstellt, verstanden wird. Ob GToBI auf diese Weise verstanden werden soll, geht aus den Beschreibungen von GToBI (z.B. Grice und Baumann 2002) nicht klar hervor, wird hier jedoch angenommen. Die als relevant betrachteten Phänomenbereiche können in einer groben Annäherung als die *kommunikativen* Aspekte (KIM/PROLAB), bzw. die *linguistischen* Aspekte (GToBI) der Sprechmelodie bezeichnet werden. Die beiden Modelle unterscheiden sich grundlegend in der phonologischen Repräsentation von Intonationskonturen. In KIM werden zugrundeliegende ganzheitliche Konturen angesetzt, die sich durch distinktive Merkmale beschreiben lassen. In GToBI liegen zwei Zieltöne zugrunde, die sich auf unterschiedliche Weise kombinieren und modifizieren lassen. Beide Modelle arbeiten weitgehend instrumentell mit F_0 -Verläufen als Korrelat für Tonhöhenverläufe. In das KIM sind *a priori* Erkenntnisse aus Experimenten eingeflossen.

3.4 Phonetische Aspekte phrasenfinaler Melodiemuster

Da in GToBI zwei Typen von Phrasen (ip und IP) angenommen werden, in KIM/PROLAB aber diese Unterscheidung nicht vorgenommen wird, muss versucht werden, eine Vergleichsbasis für die phrasenfinalen Muster aus GToBI und KIM/PROLAB zu schaffen. Im Folgenden wird daher die ip vernachlässigt und die IP als gleichwertig mit der prosodischen Phrase aus KIM/PROLAB betrachtet.

Im Anschluss an das Maximum der finalen Gipfelkontur kann F_0 entweder auf dem Niveau des Maximums verbleiben (*ebener* Verlauf) oder aber fallen. Ein weiteres Ansteigen ist nach einer Gipfelkontur *per definitionem* nicht möglich, da es sich dann um eine Talkontur handeln würde⁴. Der ebene Verlauf wird in PROLAB mit [0.] symbolisiert, in GToBI mit [H-%]. Der fallende Verlauf wird in PROLAB entweder mit [1.] oder [2.] markiert, wobei [1.] ein leichtes bis mittleres Fallen bezeichnet, und [2.] ein Absinken der Tonhöhe bis in den unteren Bereich der Sprechstimme. In GToBI gibt es nur ein Etikett für ein tiefes Fallen: [L-%].

Außerdem kann F_0 nach dem finalen Absinken wieder steigen. Solche fallend-steigenden Konturen werden in PROLAB in Bezug auf die Höhe des Wiederanstiegs unterschieden: [,] bezeichnet einen leichten bis mittleren Wiederanstieg, [?] einen Wiederanstieg bis in den oberen Bereich der Sprechstimme. In Kombination mit den beiden möglichen Arten des F_0 -Abstiegs ergeben sich also vier verschiedene Etiketten für fallend-steigende Konturen: [1.,], [1.?], [2.,] und [2.?]. In GToBI gibt es auch für die fallend-steigende Kontur nur ein Etikett: [L-H%]. Tabelle 3.1 (Nr. 1, 2 und 4) fasst die Etiketten für die genannten phrasenfinalen Melodieverläufe zusammen. Die leicht-fallende Kontur [1.] sowie ihre Kombinationen stellen eine Erweiterung von PROLAB gegenüber KIM dar, d.h. sie sind mit den Merkmalen des KIM nicht darstellbar. In Tabelle 3.1 sind sie daher in Klammern gesetzt und im Folgenden werden sie nicht weiter behandelt.

Melodiemuster	GToBI	PROLAB
1. eben	H-%	0.
2. fallend	L-%	(1.) 2.
3. fallend-minimal-steigend		(1;) 2;
4. fallend-steigend	L-H%	(1.,) (1.?) 2., 2.?

Tabelle 3.1: Phrasenfinale Melodiemuster, die nach einer Gipfelkontur auftreten können, in GToBI und PROLAB.

⁴In GToBI besteht die Möglichkeit, nach einem [H*]-Akzent eine steigende Kontur zu etikettieren. Auch dieser Fall wird hier aufgrund seiner phonetischen Form (steigend) als Talkontur betrachtet.

Zusätzlich wird in PROLAB eine weitere Art von Melodiemuster etikettiert, für die Peters (1999) eine formal, sowie eine funktional orientierte Bezeichnung vorgeschlagen hat: die *Semikolon-*, bzw. die *pseudoterninale* Kontur [2;] (vgl. Tabelle 3.1). Dieses Muster zeichnet sich durch einen finalen F_0 -Abstieg ([2.]) aus, auf den ein sehr leichter Wiederanstieg von F_0 im Bereich der letzten Silbe der Phrase folgt. Dieser Wiederanstieg hat ein Ausmaß zwischen etwa 5 Hz und 15 Hz (Peters 1999).

Die [2;]-Kontur bildet hinsichtlich des phonetischen Parameters F_0 eine Zwischenstufe zwischen [2.] und [2..]. Bei ihrer Untersuchung muss also in erster Linie getestet werden, ob sie sich von der fallenden Kontur einerseits und von den fallend-steigenden Konturen andererseits abgrenzen lässt. Die ebene Kontur ([0.] bzw. [H-%]) ist daher für die Fragestellung der Arbeit nicht von Bedeutung und wird im Weiteren vernachlässigt.

3.5 Semantische Aspekte phrasenfinaler Melodiemuster

Peters (1999; 2000) stellt auf der Basis von Beispielen aus dem *Kiel Corpus of Read and Spontaneous Speech* (IPDS 1994; IPDS 1995; IPDS 1996; IPDS 1997a) Hypothesen über die Funktionen der pseudoterninalen Kontur, sowie der fallenden und der fallend-steigenden Konturen auf. Demnach *“wird durch einen Abfall der Grundfrequenz bis an die Untergrenze der Sprechstimme Abgeschlossenheit der Argumentation signalisiert”* (Peters 1999:64). Diese Kontur *“wird in der Prosodie eingesetzt, um kategorischen Behauptungen oder Feststellungen Nachdruck zu verleihen und diese als klares Faktum darzustellen, über das nicht weiter diskutiert werden soll und wird deshalb ‘terminal’ genannt”* (Peters 1999:64). Die Bezeichnung *terminale* Kontur für das tief-fallende Muster ([L-%], [2.]) wird im Folgenden übernommen. Grice und Baumann (2002) beschreiben die Funktionen der phrasenfinalen Muster in Abhängigkeit vom vorangehenden Akzentmuster. Alle resultierenden Kontexte sollen hier nicht wiedergegeben werden. Zusammenfassend lässt sich der [L-%]-Kontur die Funktion einer neutralen Aussage oder W-Frage oder verschiedener Arten von Feststellungen zuschreiben.

Fallend-steigende Konturen *“fungieren [...] als intonatorische Markierung von Fragen oder werden in warnenden oder drohenden Äußerungen verwendet”* (Peters 1999:68). Die [2.]-Kontur wird also als eine Form des terminalen Intonationstyps bei von Essen (1964) betrachtet, die fallend-steigenden Konturen als Formen des interrogativen Typs (vgl. 3.1). Kohler (1991a) differenziert zwischen dem *low rise* [2..] und dem *high rise* [2.?] in einer fallend-steigenden Kontur und nennt [2..] *continuation type* und [2.?] *question type*. Kohler (1992) schreibt der [2..]-Kontur auch eine warnende Funktion zu (z.B. “Du [2..]”), während [2.?] verwendet wird, um großes Erstaunen zu signalisieren (“Du [2.?]”). Nach Grice und Baumann (2002) wird die fallend-steigende Kontur [L-H%] in höflichen Angeboten verwendet.

Durch die pseudoterminalen Kontur *“zeigt der Sprecher sehr subtil eine Distanzierung vom Gesagten oder eine gewisse Verhandlungsbereitschaft an”* (Peters 1999:66). Insbesondere Frauen verwenden diese Kontur, *“um zu starke Dominanz und Kategorialität im Dialog zu vermeiden”* und *“um dialogbereit zu erscheinen, ohne zu unentschlossen oder unsicher zu wirken”* (Peters 1999:66). Die Verwendung der pseudoterminalen Kontur kann auch bedeuten, dass *“Frauen in der Kommunikation dem Hörer häufiger das Gefühl vermitteln mit einbezogen zu werden, indem sie durch die Intonation eine gewisse Unsicherheit vermitteln”* (Peters 2000:161).

Melodiemuster	pragmatische Funktion
[2.] [L-%]	Terminalität: Abgeschlossenheit, Kategorialität Aussage, W-Frage
[2;]	Pseudoterminalität: Vermeidung von Dominanz und Kategorialität, Unsicherheit, Verhandlungsbereitschaft
[2.,] [2.? [L-H%]	Weiterweisung, Warnung, Drohung Frage, Erstaunen höfliches Angebot

Tabelle 3.2: Überblick über die Funktionen von phrasenfinalen Melodiemustern.

Die Funktionen, die den hier behandelten phrasenfinalen Melodiemustern von Kohler (1991a; 1992), Peters (1999; 2000) und Grice und Baumann (2002) zugeschrieben werden, sind in Tabelle 3.2 in zusammengefasster Form aufgelistet.

Peters (1999; 2000) maß Häufigkeiten von phrasenfinalen Konturen auf der Basis des *Kiel Corpus of Read Speech* (IPDS 1994). In seinem Korpus aus gelesenen Sätzen sowie Geschichten treten die (mit PROLAB etikettierten) phrasenfinalen Konturen sehr sprecherspezifisch auf. Peters (2000) präsentiert die Ergebnisse zweier Sprecher aus dem Satz-Korpus: Einer von beiden verwendete sämtliche mit PROLAB etikettierbaren Muster, der andere (von den hier berücksichtigten Mustern) nur die [2.]-Kontur. Neben dieser generellen Sprecherabhängigkeit für phrasenfinale Melodiemuster trat die [2;] zudem deutlich häufiger bei weiblichen Sprechern auf als bei männlichen.

3.6 Hörbeispiel

Für einen ersten Höreindruck zur [2;]-Kontur im Vergleich zu der [2.]- und der [2.,]-Kontur sei an dieser Stelle auf die beiliegende CDROM verwiesen. Im Verzeichnis *Beispiele* befindet sich je ein durch *PSOLA*-Resynthese erzeugtes Beispiel für [2.] (*punkt.wav*), [2;] (*semikolon.wav*) und [2.,] (*punktkomma.wav*). Diesen Resynthesen liegt eine authentische Realisierung von [2;] zugrunde.

Kapitel 4

Problemstellung

Peters (1999) beschreibt die pseudoterminale Kontur als ein Muster, welches global der terminal fallenden Kontur ähnelt, allerdings im Bereich der phrasenfinalen Silbe einen geringen Wiederanstieg der Grundfrequenz aufweist. Als ungefähres Ausmaß dieses Wiederanstiegs werden 5–15 Hz angegeben (vgl. 3.4). Es stellen sich hinsichtlich der vorgestellten Grundlagen zur Produktion, Perzeption und zur phonologischen Beschreibung phrasenfinaler Konturen folgende Fragen:

1. Stellt das [2;]-Muster ein *makrointonatorisches* Phänomen dar oder ist es vielmehr Folge eines *mikrointonatorischen* Effekts?
2. Kann das [2;]-Muster *perzeptiv* von den phonetisch ähnlichen Mustern ‘fallend’ einerseits und ‘fallend-steigend’ andererseits unterschieden werden?
3. Unterscheidet sich das [2;]-Muster *funktional* von den phonetisch ähnlichen Mustern ‘fallend’ einerseits und ‘fallend-steigend’ andererseits?

Dieses Kapitel stellt eine zusammenfassende Diskussion des ersten Teils dieser Arbeit und eine Einleitung des zweiten Teils gleichermaßen dar: Zuerst werden Konsequenzen aus den Einführungen in den Kapiteln 2 und 3 hinsichtlich der drei hier formulierten Fragen diskutiert (Abschnitt 4.1). Daraus werden an zweiter Stelle Hypothesen für die eigenen Untersuchungen abgeleitet (4.2).

4.1 Das pseudoterminale Melodiemuster

4.1.1 Aspekte der Produktion

Die Frage 1 stellt sich aufgrund folgender Überlegungen: Mikrointonatorische Effekte wurden als Artefakte der supralaryngalen Artikulation vorgestellt (2.3.4). Sie beinhalten segmentelle Informationen und können gegenüber der intendierten Sprechmelodie (der Makrointona-

tion) als nicht-intendierte F_0 -Perturbationen betrachtet werden. Mikrointonation und Makrointonation überlagern sich und formen zusammen den physikalisch messbaren F_0 -Verlauf. Dieser Sachverhalt muss bei der Analyse der Intonation stets beachtet werden, besonders wenn ein grafisch dargestellter F_0 -Verlauf zur visuellen Inspektion hinzugezogen wird.

Die Größenordnung der mikrointonatorischen Effekte liegt bei 15 Hz (oder mehr) im Falle der F_0 -Anhebung nach stimmlosen Konsonanten und bei etwa 1 st im Falle der geschlossenen vs. offenen Vokale (2.3.4). Nimmt man z.B. eine männliche Stimme mit einem durchschnittlichen F_0 -Wert von etwa 120 Hz als Grundlage, so lassen sich für einen groben Vergleich die angegebenen Hz-Werte in st umrechnen (Formel 2.2). Für den Effekt der stimmlosen Konsonanten ergibt sich für die st-Differenz aus 120 Hz und 135 Hz (= 120 Hz + 15 Hz) ein Wert von etwa 2 st.

Für die pseudoterminalen Kontur sei die Ausgangsfrequenz deutlich tiefer gewählt als 120 Hz, z.B. 80 Hz, da der Wiederanstieg auf einen vorherigen tiefen Abfall folgt. Ein Anstieg von 5-15 Hz bei $F_0 = 80$ Hz entspricht einem Anstieg von 1-3 st. Die pseudoterminalen Kontur zeichnet sich nach dieser groben Rechnung durch einen Wiederanstieg in der Größenordnung von mikrointonatorischen Effekten aus. Dieser Umstand muss in der Untersuchung des pseudoterminalen Musters berücksichtigt werden.

4.1.2 Aspekte der Perzeption

Die Frage 2 kann in zwei Teilfragen zerlegt werden. Die erste Teilfrage beschäftigt sich mit der Abgrenzung des pseudoterminalen Musters vom terminalen. Wenn durch die mikrointonatorischen Effekte nur segmentelle Information übermittelt wird, dann ist die Annahme plausibel, dass eine F_0 -Schwankung in der Größenordnung von 2 st vom Hörer in der Regel nicht als ein relevanter Beitrag zur intendierten Sprechmelodie wahrgenommen wird. Diese Annahme stimmt bei einer groben Betrachtung mit 't Harts (1981) Vorschlag überein, dass der Unterschied zwischen zwei Intervallen mindestens 3 st betragen muss, um makrointonatorisch von Bedeutung zu sein. Der Vergleich mit 't Harts Hypothese kann aber bestenfalls sehr vorsichtig vorgenommen werden. Sie stellt eine Verallgemeinerung eines speziellen intonatorischen Kontextes dar. 't Hart (1981) untersuchte Gipfelkonturen mit unterschiedlich hohen Anstiegen, d.h. es werden zwei *global ähnliche* Muster verglichen. Im Fall der pseudoterminalen Kontur steht ein (tiefer) ebener F_0 -Verlauf oder ein Abfall ([2.], [L-%]) im Kontrast zu einem steigenden Verlauf ([2;]), d.h. es werden zwei *global verschiedene* Muster verglichen.

Die Frage, ob das pseudoterminalen Muster sich in der Wahrnehmung von dem terminalen Muster unterscheidet, sollte also auf eine andere Weise formuliert werden. Eine terminale Kontur wird in der Regel als *durchgehend* fallend beschrieben (z.B. Kohler 1991a). Es soll für ein kurzes Rechenexperiment vereinfachend angenommen werden, dass eine [2.]/[L-%]-Kontur zuerst fällt und dann auf einem tiefen Ton bei 80 Hz *eben* ausläuft. Ausgehend von

diesem ebenen Verlauf kann mit der Formel 2.3 der *glissando threshold* (vgl. 2.4.2) berechnet werden, für den ein steigender Verlauf gerade als dynamischer Ton wahrgenommen werden müsste. Ist die durchschnittliche Dauer (T) des Wiederanstiegs bekannt, so kann mit der Formel die nötige Anstiegsrate (g_{thr}) berechnet werden. Bisher liegt nur für den F_0 -Umfang des Wiederanstiegs eine hypothetische Angabe von 5-15 Hz vor. Wird in der Formel 2.3 $g_{thr} = \Delta F_0/T$ gesetzt (mit $\Delta F_0 = F_0$ -Differenz in st und $T =$ Dauer in s), so ergibt sich:

$$(g_{thr} =) \frac{0,16 \text{ st} \times \text{s}}{T^2} = \frac{\Delta F_0}{T} \quad (4.1)$$

Aus einer einfachen Umformung folgt:

$$T = \frac{0,16 \text{ st} \times \text{s}}{\Delta F_0} \quad (4.2)$$

In diese Formel können die Werte 1 st, bzw. 3 st (s.o.) für ΔF_0 eingesetzt werden. Das Ergebnis für T liefert dann die Dauer, welche ein Wiederanstieg von 1 st bzw. 3 st aufweisen müsste, damit dieser Wiederanstieg nach der Theorie des *glissando threshold* als ein dynamischer Ton wahrgenommen werden kann. Es ergibt sich für $\Delta F_0 = 1$ st eine nötige Wiederanstiegsdauer von $T = 160$ ms, für $\Delta F_0 = 3$ st ist $T = 53$ ms.

Wenn also gezeigt werden könnte, dass der Wiederanstieg im Mittel etwa 3 st beträgt und etwa 50 ms dauert, so könnte vermutet werden, dass der Anstieg als ein solcher wahrnehmbar ist und sich somit auch von einem tiefen ebenen Auslaufen von F_0 abgrenzen lassen müsste. Daraus ergibt sich aber noch keine direkte Antwort auf Frage 2. Erstens könnte nicht ohne Weiteres geschlossen werden, dass der Wiederanstieg der pseudoterminalen Kontur als ein dynamischer Ton perzipiert werden kann: Der *glissando threshold* beruht auf eher psychoakustischen als psychophonetischen Experimenten (2.4.2), so dass er vermutlich für F_0 -Konturen in der gesprochenen Sprache größer ausfällt. Zweitens muss der Wiederanstieg der [2;]-Kontur für die Unterscheidung von einer [2.]-Kontur möglicherweise gar nicht als dynamischer Ton perzipiert werden können: Wird die eingangs gemachte Vereinfachung aufgehoben, so zeichnet sich die [2.]-Kontur durch einen durchgehenden Fall aus. Die (eher psychoakustischen) Ergebnisse von Klatt (1973) zeigten, dass ein deutlich geringeres Steigen, als hier bisher betrachtet wurde, von einem ebenso geringen Fallen perzeptiv gut unterschieden werden kann (vgl. 2.4.2).

Es wird klar, dass auf dieser rechnerischen Basis nicht eindeutig geklärt werden kann, ob das pseudoterminaler Muster – ausgehend von den bisherigen hypothetischen Angaben zu seiner phonetischen Gestalt – von einem terminalen perzeptiv unterscheidbar sein müsste. Der erschwerende Faktor liegt darin, dass in erster Linie psychoakustische Ergebnisse diskutiert werden können, deren Übertragbarkeit auf die sprachliche Kommunikation in Frage zu stellen ist. Die hier vorgestellte Rechnung kann und soll also lediglich als ein grober Anhaltspunkt verstanden werden.

Eine weitere Frage besteht darin, ob das pseudoterminaler Muster sich von einer fallend-

steigenden Kontur abgrenzen lässt. 't Hart et al. (1990) schlagen vor, dass zwei ansteigende F_0 -Verläufe sich in ihrer Anstiegsrate g mindestens um den Faktor 2 unterscheiden müssen, damit sie als unterschiedliche Konturen wahrgenommen werden können. Das Rechenbeispiel kann diesbezüglich fortgeführt werden: Sei für das [2;]-Muster dieses Mal $\Delta F_0 = 2$ st gesetzt. Dann ergibt sich aus Formel 4.2 für einen minimalen perzipierbaren Wiederanstieg eine Anstiegsdauer von $T = 80$ ms. Für die Anstiegsrate selbst folgt dann $g_{thr} = g_1 = 25$ st/s. Ein zweiter, perzeptiv verschiedener Wiederanstieg müsste also – bei einer vergleichbaren Anstiegsdauer – mindestens eine Rate von $g_2 = 2 \times g_1 = 50$ st/s aufweisen. Ein derart steiler Wiederanstieg ist nach Xu und Sun (2002) in der Sprachproduktion durchaus realisierbar.

Die Anstiegsbewegung selbst muss für die perzeptive Unterscheidung der beiden Muster allerdings nicht von Bedeutung sein. Ebenso ist es – gerade im Fall der phrasenfinalen Konturen – möglich, dass die Muster anhand der *final erreichten* Tonhöhe unterschieden werden (vgl. 't Hart et al. 1990). Es ist also denkbar, dass für [2;] und [2.,] zwei Anstiegsraten gefunden werden, die sich um weniger als den Faktor 2 unterscheiden und dennoch perzeptiv klar verschieden sind.

4.1.3 Funktionale Aspekte

Zur Diskussion der Frage 3 muss versucht werden, die vielfältigen Hypothesen zur Funktion der phrasenfinalen Muster (vgl. 3.5) auf einige wesentliche Aspekte zu reduzieren. Im Vordergrund steht dabei die Frage, wie das Konzept der *Pseudoterminalität* zu verstehen ist. Hierzu wiederum ist eine Klärung des Begriffs der *Terminalität* vonnöten. Aus Peters (1999) Vorschlägen lässt sich ableiten, dass sich die Terminalität aus zwei Teilkonzepten zusammensetzt: *Terminalität* = *Abgeschlossenheit* + *Kategorialität*. Dabei bezieht sich Abgeschlossenheit auf die *inhaltliche* Struktur der Äußerung, und Kategorialität kann als eine Verstärkung der Abgeschlossenheit betrachtet werden. Durch die terminale Kontur kann also zum Ausdruck gebracht werden:

- *Abgeschlossenheit*: “Meine Argumentation endet hier.”
- *Kategorialität*: “Meine Argumentation endet hier und muss von niemandem fortgeführt werden.”

Das wesentliche funktionale Merkmal der Pseudoterminalität scheint im Vergleich zur Terminalität nach Peters (1999) Beschreibungen in der Vermeidung von Kategorialität und damit einhergehend in der Signalisierung einer größeren Verhandlungsbereitschaft zu bestehen. Zur Funktion der fallend-steigenden Konturen liegen auf den ersten Blick viele heterogene Hypothesen vor (vgl. Tabelle 3.2). Es soll versucht werden, sie anhand der bisher diskutierten funktionalen Merkmale Abgeschlossenheit und Kategorialität zusammenzufassen. Keine der Funktionen in Tabelle 3.2 für [2.,], [2.?] und [L-H%] kann mit einer inhaltlichen Abgeschlossenheit assoziiert werden: In jedem Fall verlangt der Sprecher eine Fortführung des

Inhalts oder eine Reaktion auf seine Äußerung, sei es durch sich selbst (Weiterweisung zum nächsten Abschnitt des Redebeitrags) oder durch den Gesprächspartner (z.B. Frage, Angebot, Drohung). Damit geht einher, dass auch Kategorialität (Verstärkung der Abgeschlossenheit) nicht durch eine fallend-steigende Kontur zum Ausdruck gebracht werden kann.

Zusammenfassend lassen sich also die drei Konturtypen [2.], [2;] und [2.,]/[2.?] durch das Konzept der Terminalität klassifizieren, wobei gilt: [2.] = terminal, [2;] = pseudoterminal und [2.,]/[2.?] = nicht-terminal. Tabelle 4.1 stellt die Beziehung dieser drei Muster anhand der beiden definierenden Merkmale ‘abgeschlossen’ und ‘kategorisch’ in einer Merkmalsmatrix dar.

Melodiemuster:	[2.]/[L-%]	[2;]	[2.,]/[2.?]/[L-H%]
Merkmal: <i>abgeschlossen</i>	+	+	–
<i>kategorisch</i>	+	–	–
Funktion:	<i>terminal</i>	<i>pseudoterminal</i>	<i>nicht-terminal</i>

Tabelle 4.1: Ausprägungen der funktionalen Merkmale ‘abgeschlossen’ und ‘kategorisch’ für das terminale, das pseudoterminal und das nicht-terminal Melodiemuster.

Der Unterschied zwischen dem pseudoterminalen und dem terminalen Muster ist nach Ladd (1996) klar paralinguistisch, da die Funktion des einen als eine Verstärkung der Funktion des anderen betrachtet werden kann. Anhand dieses Beispiels wird erneut die Problematik einer Trennung zwischen Linguistischem und Paralinguistischem deutlich: Die hier vorgenommene Klassifizierung der drei Muster suggeriert, dass das pseudoterminal Muster der “nicht-verstärkte” Fall ist, während durch das terminale Muster eine Verstärkung zum Ausdruck gebracht wird. Ebenso gut könnte aber das pseudoterminal Muster als eine Abschwächung des terminalen betrachtet werden. Bereits die Entscheidung, welches der beiden Muster eine paralinguistische Variation des anderen ist, erscheint problematisch. Die strikte Trennung zwischen Linguistischem und Paralinguistischem nach Ladd (1996) scheint nach der hier geführten Diskussion (vgl. auch 3.2.2) ein Vorhaben zu sein, dem kein erkennbarer Nutzen hinsichtlich der Beschreibung der sprachlichen Kommunikation via Intonation im Deutschen abzugewinnen ist und das vielmehr unnötige Probleme mit sich bringt.

4.2 Hypothesen und methodisches Vorgehen im Überblick

Die Grundhypothese kann wie folgt formuliert werden:

- **Das pseudoterminal Melodiemuster unterscheidet sich hinsichtlich seiner phonetischen und seiner semantischen Eigenschaften von allen anderen Melodiemustern des Deutschen.**

Die Untersuchungen konzentrieren sich auf den Vergleich des pseudoterminalen Musters mit dem terminalen und dem nicht-terminalen. Bezüglich der nicht-terminalen Muster wird hier-

bei nicht zwischen den beiden unterschiedlichen PROLAB-Formen [2.] und [2.?] unterschieden. Stattdessen wird wie in GToBI nur eine zusammenfassende Kategorie [L-H%] angesetzt. Eine mögliche feinere Unterscheidung der Anstiegshöhe verschiedener Varianten von [L-H%] stellt einen gesonderten Untersuchungsgegenstand dar. Der hier behandelten Fragestellung wird in zwei getrennten Untersuchungen nachgegangen: einer Analyse eines mit PROLAB etikettierten Datenkorpus (Kapitel 5) und einem Perzeptionsexperiment (Kapitel 6).

Die dargestellten Grundlagen in Kapitel 2 und 3 haben Konsequenzen für die Untersuchungen in Kapitel 5 und 6. Als akustisches Korrelat der Tonhöhe wird die Grundfrequenz des Sprachsignals betrachtet. Der grundsätzliche Fehler eines solchen Vorgehens wurde in Abschnitt 2.4.1 erläutert. Dieser Fehler wird zu großen Teilen abgemildert, indem Intervalle in Halbtöne umgerechnet werden. In der Korpusanalyse werden Messungen vorgenommen, die sich am F_0 -Verlauf orientieren. Damit enthalten die Ergebnisse mikrointonatorisch bedingte Fehler und müssen entsprechend vorsichtig interpretiert werden. Für das Perzeptionsexperiment scheint das klassische Paradigma der kategorialen Wahrnehmung in diesem Rahmen keine adäquate Methode zu sein (vgl. 2.4.3). Stattdessen wird das Experiment mit Hilfe eines *semantischen Differentials* gestaltet.

Die Grundhypothese lässt sich in Teilhypothesen bzw. Teilziele zerlegen:

1. Das pseudoterminaler Muster ist nicht Folge eines mikrointonatorischen Effekts. Diese Hypothese wird nicht explizit getestet. In der Korpusanalyse wird u.a. die Beschaffenheit des segmentellen Materials untersucht, über dem die pseudoterminaler Kontur auftritt. Auf dieser Basis kann die Wahrscheinlichkeit eines mikrointonatorischen Effekts eingeschätzt werden. Würde sie z.B. in fast allen Fällen über einem schließenden Diphthong wie [ae] auftreten, so wäre die Wahrscheinlichkeit hoch, dass der leichte Wiederanstieg in der Tat mikrointonatorisch ist. Tritt sie dagegen in verschiedenen segmentellen Kontexten auf, so handelt es sich mit größerer Wahrscheinlichkeit um ein makrointonatorisches Phänomen. Denkbar wäre allerdings, dass der leichte finale Wiederanstieg ein Artefakt eines anderen, in den Grundlagen nicht vorgestellten physiologischen Vorgangs ist. Auf diese Interpretationsmöglichkeit kann in der Diskussion der Ergebnisse zurückgegriffen werden, falls die Grundhypothese insgesamt nicht bestätigt werden kann.
2. Es muss gezeigt werden, dass sich die pseudoterminaler Kontur hinsichtlich der phonetischen Parameter von einer fallend-steigenden Kontur unterscheidet, da zwischen diesen beiden Mustern bezüglich der Verlaufsform der Melodie Ähnlichkeit besteht: Beide weisen einen fallenden und einen steigenden Abschnitt auf. Im Rahmen der Korpusanalyse sollen die Muster hinsichtlich der zwei phonetischen Parameter verglichen werden, die sich direkt aus der Definition der Intonation ableiten lassen: die Höhe und das *Timing* des Wiederanstiegs. Ein Vergleich mit der terminalen Kontur ist nicht nötig,

da sich die terminale und die pseudoternale Kontur hinsichtlich des F_0 -Verlaufs *per definitionem* unterscheiden.

3. Mit einem Perzeptionsexperiment (Kapitel 6) wird untersucht, ob das pseudoternale Muster perzeptiv von dem terminalen und dem nicht-terminalen Muster unterschieden wird, und ob Hörer allen drei Mustern unterschiedliche Funktionen zuweisen. Die Hypothese lautet, dass die drei Muster sich v.a. durch die funktionalen Merkmale ‘abgeschlossen’ und ‘kategorisch’ unterscheiden lassen (vgl. Tabelle 4.1). In die Konstruktion der Stimuli gehen u.a. die Ergebnisse der Korpusanalyse (Kapitel 5) ein.
4. Basierend auf den psychoakustischen und psychophonetischen Grundlagen könnten folgende Hypothesen aufgestellt werden:
 - (a) Die Dauer sowie die Höhe des Wiederanstiegs der pseudoternalen Kontur bilden eine Anstiegsrate g_p , die in der Größenordnung des sich ergebenden *glissando threshold* g_{thr} liegt. Also $g_p \approx g_{thr}$.
 - (b) Die Anstiegsrate der nicht-terminalen Kontur g_n ist mindestens doppelt so groß wie die der pseudoternalen Kontur g_p . Also: $2 \times g_p \leq g_n$.

Wie aus Abschnitt 4.1.2 hervorgeht, lässt sich im Vorfeld nicht ohne Weiteres klären, ob die Hypothesen auf diese Weise formuliert werden können. Anhand der Ergebnisse aus der Korpusanalyse und des Perzeptionsexperiments werden Berechnungen angestellt, mit denen diese Hypothesen getestet werden können. Auf dieser Basis kann diskutiert werden, ob das Perzept des Wiederanstiegs einer pseudoternalen Kontur eher ein dynamischer Ton ist oder eher ein statischer, der sich dann möglicherweise von dem Perzept des finalen Tons einer terminalen Kontur in seiner Höhe unterscheidet.

Wie aus dieser groben Übersicht hervorgeht, werden die Hypothesen 1 und 2 mit Hilfe der Korpusanalyse, die Hypothese 3 mit Hilfe des Perzeptionsexperiments und die Hypothese 4 basierend auf den Ergebnissen beider Untersuchungen getestet. Die Hypothesen werden in den Beschreibungen der beiden Untersuchungen wieder aufgegriffen und dort detaillierter formuliert.

Teil II

Eigene korpusbasierte und experimentelle Untersuchungen

Kapitel 5

Korpusanalyse

5.1 Ziele

Zur phonetischen Gestalt der untersuchten Melodiemuster ([2.] vs. [2;] vs. [2.,]) liegen nur fragmentale, nicht auf systematischen Untersuchungen basierende Angaben vor. Ein erstes Ziel dieser Analyse besteht darin, anhand von Messungen in einem etikettierten Datenkorpus einige dieser Lücken zu schließen. Konkret können die dazu gehörenden Teilziele wie folgt formuliert werden:

- Überprüfung der Angabe von Peters (1999) bezüglich der Wiederanstiegshöhe von [2;].
- Ermittlung der Wiederanstiegshöhe von [2.,].
- Auffinden von Unterschieden zwischen [2;] und [2.,] hinsichtlich des *Timing* des Wiederanstiegs.

Die Untersuchung konzentriert sich also auf phonetische Parameter des Wiederanstiegs, nicht auf die Form des vorherigen Abstiegs. Daher wird die [2.]-Kontur in dieser Analyse nicht berücksichtigt.

Im Hinblick auf die Erzeugung der Stimuli des Perceptionsexperiments ist die Kenntnis einiger Richtwerte für die phonetischen Parameter von [2;] und [2.,] vonnöten: Zum Beispiel muss der Unterschied des Wiederanstiegs zwischen zwei Stimuli in einer Größenordnung liegen, die kleiner ist als der tatsächlich auftretende Wiederanstieg von [2;]. Die Messergebnisse der Korpusanalyse werden daher einen direkten Nutzen für die Konzeption des Perceptionsexperiments haben. Ein zweites Ziel besteht darin, anhand der Vermessungen von [2;] und [2.,] die Hypothesen 1, 2 und 4 (vgl. 4.2) zu testen. Eine detaillierte Formulierung der Hypothesen wird bei der Präsentation der entsprechenden Ergebnisse geliefert.

Es sei betont, dass mit dieser Analyse keine Aussagen zur Auftretenshäufigkeit der beiden Muster [2;] und [2.,] gemacht werden sollen. Die Verteilung der phrasenfinalen Konturen hängt stark von den Daten selbst ab, bzw. vom Aufnahmeszenario (vgl. Lese- vs.

Spontansprache bei Peters 1999; 2000), so dass für allgemeine Aussagen über die Verteilung möglichst viele verschiedene Szenarien herangezogen werden müssten. In dieser Arbeit werden aber Daten aus nur einem Aufnahmeszenario gewählt.

5.2 Methode

5.2.1 Datenbasis

Da Einflüsse verschiedener Aufnahmeszenarien, und damit Sprechsituationen und -stile nicht den Untersuchungsgegenstand der Arbeit darstellen, wurden diese explizit so weit wie möglich ausgeschaltet, indem das Korpus auf ein einheitliches Szenario beschränkt wurde: Als Datenbasis wurde das *Kiel Corpus of Spontaneous Speech Vol. 1 & 2* (IPDS 1995; 1996) gewählt¹. Dieses Korpus enthält elf Dialoge aus dem *Terminabspracheszenario*². In diesem Szenario sitzen zwei Probanden in getrennten Aufnahmekabinen und sind über Kopfhörer und Mikrofon miteinander verbunden. Die Aufnahmekanäle sind getrennt. Es kann nur auf einem Kanal zur Zeit aufgenommen werden, und die Aufnahme wird durch die Probanden selbst per Knopfdruck gesteuert. Jeder erhält einen Terminkalender mit fiktiven Terminen. Die Aufgabe besteht darin, eine Reihe weiterer Termine abzusprechen. Auf Vor- und Nachteile spontansprachlicher Daten dieser Art wird in 7.3 eingegangen.

Es gingen durch diese Auswahl Daten von 22 Sprechern (13 männlichen, 9 weiblichen) in die Untersuchung ein. Tabelle 5.1 listet die Bezeichnungen dieser Dialoge zusammen mit den beteiligten Sprechern auf.

<i>Dialog</i>	<i>Sprecher, links</i>	<i>(Geschlecht)</i>	<i>Sprecher, rechts</i>	<i>(Geschlecht)</i>
g07	HAH	(m)	TIS	(m)
g08	THS	(m)	KAK	(m)
g09	ANS	(w)	FRS	(w)
g10	UTB	(w)	NAR	(w)
g11	JAK	(m)	REK	(m)
g12	OLV	(m)	KAP	(w)
g14	BAC	(m)	SOK	(m)
g19	KAE	(w)	SAR	(w)
g21	CHD	(m)	ANL	(m)
g25	WEM	(w)	SIK	(w)
g31	AME	(m)	SVA	(m)

Tabelle 5.1: Liste der elf Dialoge aus dem *Kiel Corpus*, die für die Korpusanalyse herangezogen wurden. “links” = linker Aufnahmekanal; “rechts” = rechter Aufnahmekanal; w = weiblich; m = männlich.

¹Die PROLAB Etikettierungen, an denen sich diese Untersuchung orientiert, sind allerdings nicht auf diesen CD-ROMs enthalten, da sie erst nach der Veröffentlichung der CD-ROMs erstellt wurden.

²*Vol. 1* enthält noch zusätzliche “Testdaten” (Ausschnitte aus weiteren Dialogen), die hier nicht berücksichtigt wurden.

Anhand der PROLAB-Etikettierungen wurden alle Vorkommen von [2.] und [2;] aus dem Korpus extrahiert. Dabei wurden die als “unsicher” markierten Fälle ([%2;], [2;%], [%2.], und [2.,%]) nicht berücksichtigt. Es lagen 426 Fälle von [2;] und 49 Fälle von [2.] vor, anhand derer eine informelle Voruntersuchung durchgeführt wurde. Diese bestand in der Sichtung, bzw. dem Abhören der Daten und dem Notieren von Beobachtungen zur Gestalt der Konturen. Sowohl für die Voruntersuchung als auch für alle weiteren Messungen im Rahmen dieser Korpusanalyse wurden die Oszillogrammdarstellung und die F_0 -Analyse der Software *xas-sp* verwendet (vgl. 2.2). Ein Teil der Ergebnisse dieser Voruntersuchung floss in die weitere Einschränkung des gewählten Korpus ein (s.u.). Ein weiterer Teil hat Einfluss auf die Auswahl der (in der Hauptuntersuchung) gemessenen Parameter und wird im nächsten Abschnitt (5.2.2) präsentiert.

Von dem gewählten Korpus wurde eine Reihe von [2;]- sowie [2.]-Fällen aus der Analyse ausgeschlossen. Diese Einschränkungen erschienen aus Gründen, die im Folgenden erläutert werden, sinnvoll und wurden anhand von drei Kriterien(komplexen) vorgenommen.

1. Phrasengrenze ohne Pause:

In der Voruntersuchung wurde folgende Beobachtung gemacht: In vielen Fällen schließt sich die nachfolgende Phrase ohne eine (segmentelle) Pause an die vorherige Phrase, welche die [2;]- bzw. [2.]-Etikettierung enthält, an. In diesen Fällen ist es recht wahrscheinlich, dass der phrasenfinale Wiederanstieg durch eine Assimilation an die folgende Phrase begründet ist und bereits den Beginn des Anstiegs für den ersten F_0 -Gipfel der zweiten Phrase darstellt.

Diese plausible Überlegung konnte in vielen Fällen beim Abhören des gesamten Abschnittes mit beiden Phrasen durch den perzeptiven Eindruck des Autors bestätigt werden (und zwar sowohl für [2;]- als auch [2.]-Fälle). Es würde sich dann also um inadäquate Etikettierungen handeln, die vermutlich darauf zurückzuführen sind, dass die Etikettierer sich zu stark visuell an der F_0 -Darstellung oder aber perzeptiv zu sehr lokal (auf die einzelne Phrase beschränkt) orientiert haben. Abbildung 5.1 zeigt ein solches Beispiel aus dem gewählten Korpus.

Es wurde darauf verzichtet, für jeden Einzelfall zu entscheiden, ob die Etikettierung angemessen erscheint, da diese Beurteilung durch den Autor selbst hätte stattfinden müssen, was die Objektivität der Korpusauswahl unterbunden hätte. Stattdessen wurden alle Fälle aus der Untersuchung ausgeschlossen, denen in der Etikettierung unmittelbar das erste Wort der nächsten Phrase folgte. Mit anderen Worten: Es gingen nur diejenigen Fälle in die Untersuchung ein, in denen unmittelbar nach dem letzten Segment der untersuchten Phrase entweder eine Pause oder ein Atmen etikettiert worden ist, oder aber solche Fälle, die am Ende eines Turns oder Dialogs stehen.

2. Akzentmuster, die nicht dem “Default” entsprechen:

Da das Korpus spontansprachlich und daher wenig kontrolliert ist (s. 7.3), tritt fast jeder Fall von [2;] und [2.] in einem eigenen segmentellen und suprasegmentellen Kontext auf, der sich von den meisten anderen Kontexten unterscheidet. Bezüglich des suprasegmentel-

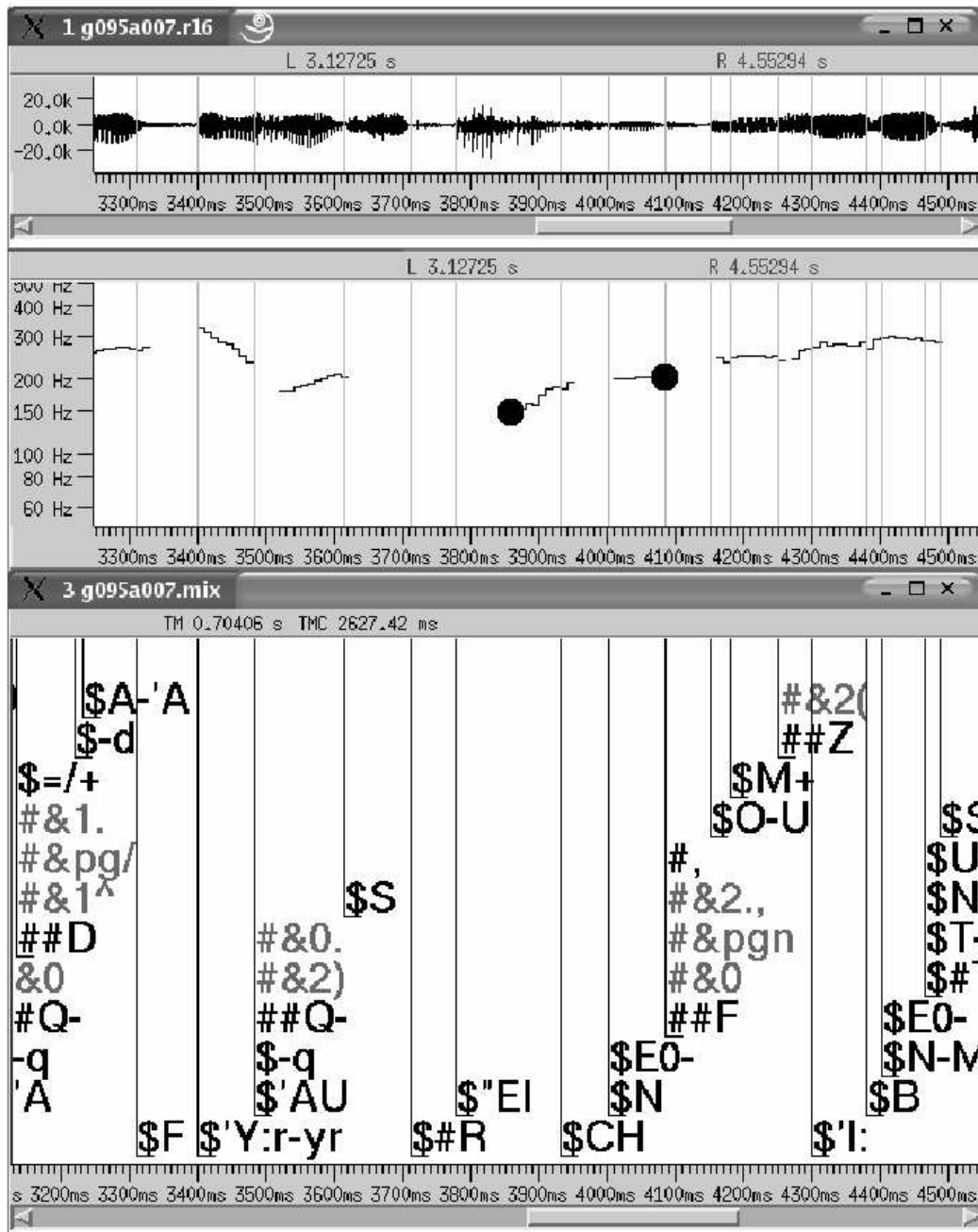


Abbildung 5.1: Beispiel einer [2.,]-Etikettierung, auf welche ohne eine etikettierte segmentelle Pause die nachfolgende Phrase folgt. Die eingezeichneten Punkte im F_0 -Verlauf markieren den Wiederanstieg, der bei der Etikettierung als fallend-steigende Kontur [2.,] eingestuft wurde. Die beiliegende CDROM (s. Anhang C) enthält dieses Beispiel als wav-Datei. (Quelle: Datei *g095a007.r16* aus dem *Kiel Corpus of Spontaneous Speech, Vol. 1*, IPDS 1995)

len Kontexts kann das Korpus aber so weit beschränkt werden, dass eine Vergleichbarkeit der Fälle annäherungsweise vorhanden ist. Als suprasegmenteller Kontext der untersuchten phrasenfinalen Konturen soll hier das phrasenfinale Akzentmuster verstanden werden. Dieses Akzentmuster kann unterschiedlich beschaffen sein. Zum Beispiel kann es aus einem einzelnen Melodiegipfel bestehen, der mit genau einer akzentuierten Silbe assoziiert ist. In diesem Fall ist in der Regel ein eindeutig zu bestimmendes Gipfelmaximum vorhanden. Eine andere Möglichkeit besteht darin, dass sich der Akzent durch einen fallenden F_0 -Verlauf manifestiert, der sich an ein vorheriges hohes F_0 -Plateau anschließt, welches auch leicht geneigt sein kann. In diesem Fall muss kein eindeutig zu bestimmendes Gipfelmaximum vorhanden sein. Eine dritte Möglichkeit besteht darin, dass ein einziges Gipfelmaximum mit mehr als nur einem akzentuierten Wort (wie in einem *Hutmuster* ('t Hart et al. 1990), z.B. [2(0. 2)] oder [L*+H H+L*]) assoziiert ist. (Das Hutmuster kann ebenfalls mit einem hohen Plateau realisiert werden.) Der erstgenannte Fall – das Vorhandensein eines Gipfels, der mit genau einem akzentuierten Wort assoziiert ist und ein eindeutig zu bestimmendes Gipfelmaximum aufweist – soll als der einfachste Fall eines Akzentmusters bezeichnet werden. Um den suprasegmentellen Kontext, in dem die untersuchten Muster auftreten, möglichst konstant zu halten, wurden nur die Fälle der hier beschriebenen einfachsten Art in der Analyse berücksichtigt. Während diese Einschränkung einerseits phänomenal motiviert ist, bringt sie andererseits auch aus messtechnischer Sicht einen Vorteil mit sich, denn bei den Messungen wurde auch das Gipfelmaximum selbst erfasst (vgl. 5.2.3).

Die Zuordnung der Fälle zu der Kategorie des einfachen Akzentmusters erfolgte nicht durch eine subjektive Beurteilung eines jeden Einzelfalls im Korpus, sondern durch die konsistente Anwendung von Kriterien, die in der Etikettierung erfüllt sein mussten. In der PRO-LAB-Etikettierung sind diese Fälle durch eine der folgenden Symbolketten markiert: [2^], [2(], [2)]. Wenn diesem phrasenfinalen Akzent noch ein weiterer Gipfelakzent vorangeht, muss auch der Akzentetikettierung ein weiteres Etikett vorangehen, welches anzeigt, wie weit die Grundfrequenz zwischen dem phrasenfinalen Akzent und dem vorhergehenden Gipfelakzent abfällt: starker Abfall ([2.]), geringerer Abfall ([1.]) oder kein Abfall (= hohes Plateau zwischen den beiden Akzenten: [0.]). Zu den einfachsten Akzentmustern zählen nach obiger Definition nur diejenigen mit einer [2.- oder [1.-]Konkatenation. Dieses Konkatenationsetikett fehlt allerdings, wenn die Phrase nur einen einzigen Akzent enthält. Falls der akzentuierten Silbe weitere unakzentuierte Silben in derselben Phrase vorausgehen (was häufig der Fall ist), so liegen diese defaultmäßig auf einem tieferen F_0 -Niveau als der F_0 -Gipfel des Akzents. Tritt dieser Defaultfall nicht ein, so muss in der Etikettierung ein *high prehead* ([hp1] oder [hp2]) etikettiert werden. Berücksichtigt wurde in der Analyse nur der Defaultfall, da im Falle eines *high prehead* Schwierigkeiten bei der Bestimmung des Gipfelmaximums auftreten können. Das phrasenfinale Akzentmuster kann einen Default-Akzent (Stufe [2]), einen Nebenakzent (Stufe [1]) oder einen emphatischen Akzent (Stufe [3]) enthalten. Die drei Akzentstufen prägen sich typischerweise u.a. durch unterschiedlich hohe F_0 -Gipfel aus. In der

Analyse wird der Wiederanstieg der [2.,]- bzw. der [2;]-Kontur auch in Relation zum vorherigen Abstieg betrachtet (s. 5.2.3, 5.2.4). Um einen Teil der zu erwartenden Variation in der Gipfelhöhe auszuschließen, wurde nur die Defaultstufe [2] berücksichtigt.

3. Störungen:

Des Weiteren können zwei Arten von Störungen in Kombination mit phrasenfinalen Konturen auftreten. Erstens kann die Phrase aus technischen Gründen abgebrochen sein [pg;], bedingt durch ein verfrühtes Loslassen des Knopfes während der Aufnahme. Zweitens treten im Korpus Abschnitte auf, die von Lachen überlagert sind. Das Lachen kann den Grundfrequenzverlauf beeinflussen. Daher werden auch diejenigen Fälle von [2;] und [2.,] ausgeschlossen, die zusammen mit einem Etikett für Lachen (-l:) auftreten.

Durch die Anwendung dieser drei Kriterienkomplexe reduziert sich das untersuchte Korpus für [2;] auf 119 Fälle (von 426) und für [2.,] auf 23 Fälle (von 49).

5.2.2 Voruntersuchung

Ein fallend-steigender F_0 -Verlauf enthält notwendigerweise ein F_0 -Minimum zwischen dem Abstieg und dem Wiederanstieg. Dieses Minimum kann (sowohl bei [2;] als auch bei [2.,]) zwei verschiedene Formen annehmen, die in Abbildung 5.2 schematisch dargestellt sind: Nach dem Ende des Abstiegs kann der Wiederanstieg sofort einsetzen, oder es kann sich an dieser Stelle ein tiefes F_0 -Plateau ausbilden.

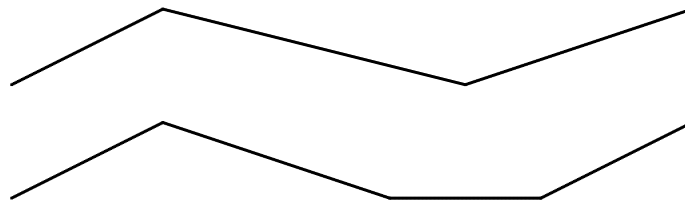


Abbildung 5.2: Schematische Darstellung eines phrasenfinalen fallend-steigenden Melodiemusters: mit (*unten*) und ohne (*oben*) tiefes Plateau.

Drei weitere Beobachtungen beziehen sich auf die Synchronisation des Wiederanstiegs mit der segmentellen Ebene. In nahezu allen Fällen von [2;] und [2.,] findet der Wiederanstieg nur auf der letzten Silbe der Phrase statt. Die Dauer des Wiederanstiegs ist bei [2.,] häufig größer als bei [2;]. Der Wiederanstieg findet bei den [2;]-Konturen in nahezu allen Fällen nur auf dem letzten (stimmhaften) Segment der Phrase statt, während er bei [2.,] häufiger etwas früher einsetzt.

Eine weitere Beobachtung betrifft die Relation zwischen Umfang des Wiederanstiegs und des vorherigen Abstiegs: In der Regel steigt F_0 in einer [2.,]-Kontur auf einen Wert, der höher liegt als das Gipfelmaximum des phrasenfinalen Akzents, während der Wiederanstieg von [2;] häufiger unter dem Gipfelniveau bleibt.

Eine letzte Beobachtung bezieht sich auf den F_0 -Verlauf am Ende der Phrase, bzw. des letzten stimmhaften Segments: Das Ende des von *xassp* analysierten F_0 -Verlaufs stimmt in der Regel nicht mit der etikettierten Segmentgrenze überein. In einigen Fällen existiert auch nach dieser Grenze (also in einer Pause oder auf einem Atmen; vgl. Kriterium 1 in 5.2.1) noch eine F_0 -Anzeige, die meist auf einen F_0 -Analysefehler zurückzuführen ist. In anderen Fällen bricht die F_0 -Anzeige vor der etikettierten Segmentgrenze ab, was teilweise ebenfalls auf Analysefehler zurückgeführt werden kann, wenn der Algorithmus aufgrund einer behauchten oder geknarrten Stimmqualität die Quasiperiodizität nicht erkennen kann. Ebenso ist es aber möglich, dass gegen Ende eines (stimmhaften) Segments Entstimmung einsetzt und somit keine Quasiperiodizität mehr vorhanden ist. Dieser letzte Abschnitt des finalen Segments wird im Folgenden als *nicht vermessbarer Abschnitt* bezeichnet.

5.2.3 Datenerhebung

Im Hinblick auf die Frage nach dem absoluten und relativen (vgl. 5.2.2) Umfang des Wiederanstiegs von [2;] vs. [2..] sowie nach dem *Timing* von [2;] vs. [2..] wurden folgende F_0 - und Dauer-Werte gemessen. Abbildung 5.3 stellt diese neun Messparameter schematisch dar.

FaOn = F_0 -Wert des Gipfelmaximums, bzw. *Onset* des fallenden Verlaufs.

FaOf = F_0 -Wert des Minimums, bzw. *Offset* des fallenden Verlaufs.

RiOn = F_0 -Wert des Minimums, bzw. *Onset* des steigenden Verlaufs.

RiOf = F_0 -Wert des Maximums des Wiederanstiegs, bzw. *Offset* des steigenden Verlaufs.

DuFa = Dauer des fallenden Verlaufs.

DuL = Dauer des tiefen Plateaus.

DuRi = Dauer des steigenden Verlaufs.

DuLS = Dauer des letzten stimmhaften Segments der Phrase.

DuE = Dauer des nicht vermessbaren Abschnitts des letzten stimmhaften Segments.

Die Messungen erfolgten nach folgenden Kriterien. Die F_0 -Werte wurden streng dem in *xassp* angezeigten Analyseergebnis entnommen. Dabei wurde explizit nicht versucht, mikrointonatorische Effekte zu berücksichtigen, bzw. allein die Makrointonation aus dem messbaren F_0 -Verlauf zu extrahieren. Ein solcher Versuch würde – auf der Basis des aktuellen Forschungsstandes – nur ein unvollständiges und unpräzises Resultat liefern. Zum Beispiel könnte ein höherer F_0 -Einsatz nach stimmlosen Konsonanten (vgl. 2.3.4) zwar in vielen Fällen erkannt werden — unklar bliebe jedoch, wie weit der mikrointonatorische Effekt in den

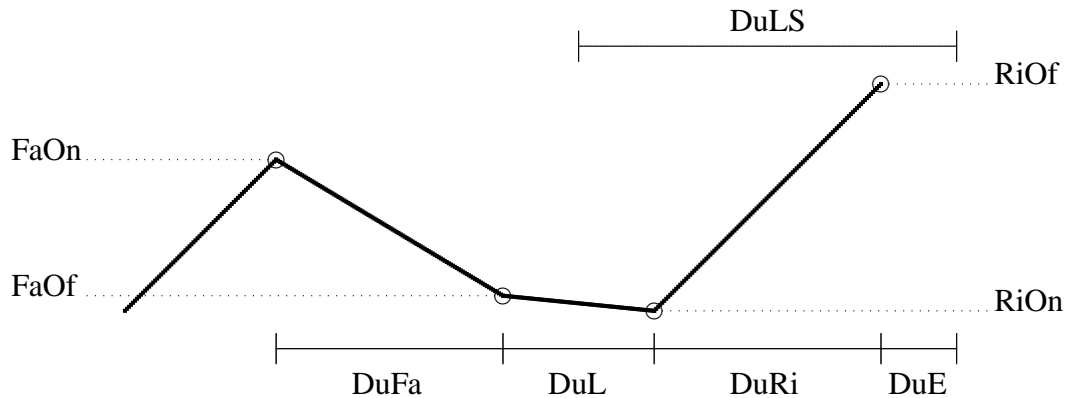


Abbildung 5.3: Illustration der gemessenen Parameter. Fette Linie = schematischer F_0 -Verlauf eines phrasenfiniten Gipfelakzents mit fallend-steigender Kontur. Erläuterung der Parameter: s. Text.

postkonsonantischen Vokal hinein reicht. Die Anwendung einer denkbaren pauschalen Regel (z.B. “miss nach stimmlosen Konsonanten den F_0 -Wert erst x ms nach Vokalonset”) wäre für den Einzelfall nicht immer adäquat. Durch einen solchen Versuch würden also möglicherweise einige der mikrointonatorisch bedingten Messfehler vermieden, dafür aber neue, methodisch bedingte Messfehler eingebracht werden. Durch ein explizites Ignorieren des mikrointonatorischen Einflusses auf die gemessene F_0 -Kontur wird ebenso explizit ein Fehler begangen, wobei die Fehlerquelle aber konstant und überschaubar gehalten wird.

Analysefehler im angezeigten F_0 -Verlauf wurden so weit wie möglich beachtet und korrigiert. Eine häufige Art ist der *Oktavfehler*, der entsteht, wenn der Algorithmus die Hälfte oder das Doppelte der korrekten Periodendauer erkennt. Dieser Fehler ist visuell leicht an einem abrupten Sprung im F_0 -Verlauf zu erkennen und lässt sich häufig durch eine Rekonfiguration der Analyseparameter korrigieren.

In Fällen, in denen der Algorithmus keinen F_0 -Wert liefert (vgl. 5.2.2), musste für die betroffenen Parameter ein fehlender Wert vermerkt werden. Betroffen sind dadurch nicht nur F_0 -, sondern auch Dauerparameter: Falls z.B. der F_0 -Verlauf im Bereich des Minimums nicht messbar ist, können auch für die Dauerparameter DuFa und ggf. DuL oder DuRi keine Werte ermittelt werden, da die Grenze zwischen Abstieg und Wiederanstieg nicht erkennbar ist.

Die Dauer DuLS wurde anhand der etikettierten Segmentgrenzen bestimmt. Zur Messung von DuL musste für jeden Einzelfall geklärt werden, ob ein tiefes Plateau vorliegt oder nicht. Kein F_0 -Verlauf in der gesprochenen Sprache enthält so deutliche und abrupte Änderungen wie in der schematischen Abbildung 5.3. Daher ist ein (kurzes) tiefes Plateau eigentlich in jedem Fall vorhanden, so dass sich die Frage stellt, wie deutlich es ausgeprägt sein muss, um als solches gewertet zu werden. Als notwendiges Kriterium wurde eine Dauer von 20 ms gesetzt, d.h. für schwach ausgeprägte tiefe Plateaus von unter 20 ms Dauer wurde DuL = 0 ms gemessen. Ist das Plateau stark geneigt, so ist es von einem Abstieg ohne tiefes Plateau

nicht gut abgrenzbar. Ein Maß für die Neigung ist die Differenz aus FaOf und RiOn. Als Grenzwert für diese Differenz wurde die Größenordnung von ± 5 Hz festgesetzt: Wenn also ein F_0 -Verlauf als plateauähnlich vorgefunden wurde, so wurde der (potentielle) Anfangs- und Endpunkt des Plateaus gesucht und überprüft, ob die Differenz der zugehörigen F_0 -Werte maximal etwa 5 Hz betrug. Andernfalls wurde $DuL = 0$ ms gemessen.

Die Messung $DuL = 0$ ms bedeutet, dass nur ein einziger F_0 -Wert für das Minimum zwischen Abstieg und Wiederanstieg ermittelt werden kann: Für $DuL = 0$ ms gilt also stets $FaOf = RiOn$. Enthält die Kontur aber ein tiefes Plateau, so können sich die Werte (um bis zu 5 Hz, s.o.) unterscheiden, wenn das tiefe Plateau geneigt ist (wie in Abbildung 5.3).

Neben diesen neun Messparametern wurden folgende weitere Informationen zu jedem Einzelfall erhoben:

Spr = Sprecherkürzel (vgl. Tabelle 5.1).

m/f = Geschlecht des Sprechers.

Wort = Das Wortmaterial von dem letzten akzentuierten Wort der Phrase bis zum Ende der Phrase, also das Wortmaterial, auf dem sich der phrasenfinale Akzent und das [2;]- oder [2.,]-Muster manifestieren, in orthografischer Repräsentation mit Markierung der akzentuierten Silbe.

Acc = Das Akzentmuster (also [2^], [2(] oder [2]).

S1 = Das segmentelle Material, auf dem sich der Wiederanstieg des [2;]- oder [2.,]-Musters manifestiert.

S2 = Das segmentelle Material, welches (in derselben Phrase) auf S1 folgt.

Ri [seg] = Anzahl der Segmente, über die sich der Wiederanstieg ausbreitet.

Ri [syl] = Anzahl der Silben, über die sich der Wiederanstieg ausbreitet.

Pa [syl] = Anzahl der Silben, die am gesamten fallend-steigenden Muster (gerechnet ab einschließlich der Akzentsilbe) beteiligt sind.

Die Repräsentation von 'Wort' ist pseudo-orthografisch, d.h. es werden elidierte Vokale gekennzeichnet. Dennoch ist die Silbenzahl des gesamten Musters Pa[syl] nicht in jedem Fall aus 'Wort' ableitbar, so dass Pa[syl] separat aufgeführt wird. Die beiden Variablen 'Wort' und Pa[syl] dienen der Orientierung bei der Sichtung der Rohdaten und werden in der Analyse nicht weiter berücksichtigt.

Die Variablen Ri[seg/syl] dienen der Prüfung der in der Voruntersuchung gemachten Beobachtung über das *Timing* der phrasenfinalen Konturen (vgl. 5.2.2). Mit Hilfe der Variablen S1 und S2 soll versucht werden, abzuschätzen, inwiefern das [2;]-Muster nur Folge eines

mikrointonatorischen Effekts ist. Die Variable Acc gibt Aufschluss darüber, ob innerhalb des bereits eingeschränkten suprasegmentellen Kontexts (vgl. 5.2.1) für die Betrachtung von [2:] vs. [2.] feiner differenziert werden muss (zwischen dem frühen, mittleren und späten Gipfel).

5.2.4 Berechnungen

Anhand der gemessenen Parameter wurden für jeden Fall folgende Größen errechnet. Zur Umrechnung von Hz-Werten in st-Werte diente Formel 2.2.

ΔFa = Umfang des F_0 -Abstiegs in st.

ΔRi = Umfang des Wiederanstiegs in st.

$\Delta RiRel$ = $\Delta Ri / \Delta Fa$, also Umfang des Wiederanstiegs in Relation zum Abstieg.

gRi = $\Delta Ri / DuRi$, also Rate des Wiederanstiegs in st/s.

$DuRiRel$ = $(DuRi + DuE) / DuLS$, also Beginn des Wiederanstiegs in Relation zum Beginn des letzten stimmhaften Segments.

Das Maß $DuRiRel$ wurde konstruiert, um das *Timing* des Wiederanstiegs zu beschreiben. Es könnte auch als "Anteil der Dauer des Wiederanstiegs an der Dauer des letzten stimmhaften Segments" bezeichnet werden, wobei der nicht vermessbare Abschnitt (DuE) als Teil der Anstiegsdauer gewertet würde. $DuRiRel = 1$ bedeutet, dass der Wiederanstieg genau am Beginn des letzten stimmhaften Segments einsetzt. $DuRiRel = 0,5$ würde einen Einsatz in der Mitte des Segments bedeuten, $DuRiRel = 0,1$ kurz vor Segmentende usw. Ist $DuRiRel > 1$, so beginnt der Wiederanstieg bereits vor dem letzten stimmhaften Segment.

5.2.5 Datenaufbereitung und -auswertung

Die Berechnungen (5.2.4) wurden für jeden Einzelfall durchgeführt. Für jeden gemessenen und errechneten Wert wurden dann der Mittelwert und die Standardabweichung über alle Fälle ($n \leq 119$ für [2:] und $n \leq 23$ für [2.,]) gebildet³. Die beiden Stichproben sind sehr heterogen. Ein Grund ist technischer Natur: Die Stichprobenumfänge unterscheiden sich drastisch. Ein zweiter Grund ist phänomenaler Natur: Mit der [2.,]-Etikettierung werden Konturen mit unterschiedlich hohen Wiederanstiegen erfasst, während die [2;]-Etikettierung nur für minimale Wiederanstiege verwendet wird. Dieser Umstand wirkt sich in einer wesentlich kleineren Standardabweichung für ΔRi bei [2:] aus als bei [2.,]. Aufgrund der Inhomogenität der beiden Stichproben wurde es nicht für sinnvoll erachtet, prüfstatische Tests durchzuführen. Die Interpretation der Ergebnisse orientiert sich daher nur an den Mittelwerten und ist mit Vorsicht auszuführen.

³Es gilt in der Regel nicht $n = 119$ bzw. $n = 23$, da einige Parameter in einigen Fällen nicht messbar waren, vgl. 5.2.3.

Ein weiterer Anlass zur Vorsicht bei der Datenauswertung ist durch die Heterogenität der einzelnen Fälle auch innerhalb einer Stichprobe gegeben. Zum Beispiel hängt der Parameter DuLS natürlich stark von der Art des letzten stimmhaften Segments ab: So ist für ein Schwa z.B. eine kürzere Dauer zu erwarten als für einen Diphthong. Es sollte also hinsichtlich der Interpretation von Dauerwerten die Verteilung der [2;-] vs. der [2.,]-Fälle auf unterschiedliche Segmentklassen (z.B. Vokal, Nasal) nicht vernachlässigt werden.

5.3 Ergebnisse

Die Rohdaten können im Anhang A eingesehen werden. Wie bereits erwähnt (vgl. 5.1) sollen im Rahmen dieser Analyse keine Häufigkeitsaussagen getroffen werden. Es soll aber eine Betrachtung der Verteilung von [2;-] und [2.,] auf die verschiedenen einbezogenen Akzentkontexte vorangestellt werden, um zu prüfen, ob eine feinere Differenzierung des suprasegmentellen Kontexts für die weiteren Analysen notwendig ist. Tabelle 5.2 zeigt diese Verteilung. [2;-] und [2.,] scheinen sich hinsichtlich ihrer Auftretenshäufigkeit mit den drei Gipfeltypen global nicht zu unterscheiden: Am häufigsten (in über 50% der Fälle) geht beiden Mustern der mittlere Gipfel voran, gefolgt von dem frühen und dem späten. Die weiteren Analysen beziehen stets die gesamte Stichprobe, nicht getrennt nach Akzentkontext, ein.

Melodiemuster	[2)]	[2 ^]	[2 (]	Summe
[2.,]	7 (30,4%)	12 (52,2%)	4 (17,4%)	23 (100,0%)
[2;-]	33 (27,7%)	68 (57,1%)	18 (15,1%)	119 (99,9%)

Tabelle 5.2: Verteilung der phrasenfinalen Muster [2;-] und [2.,] auf die drei Gipfelkonturtypen ‘früh’ [2)], ‘mittel’ [2 ^] und ‘spät’ [2 (] .

5.3.1 Konturgestalt

Aus den Beobachtungen der Voruntersuchung folgt die Frage, ob sich die [2;-] und die [2.,]-Kontur hinsichtlich der Ausbildung eines tiefen Plateaus unterscheiden. Tabelle 5.3 zeigt die Häufigkeit eines tiefen Plateaus für die beiden Konturen. Sie scheinen sich diesbezüglich nicht zu unterscheiden: Sowohl für [2.,] als auch für [2;-] tritt in etwas mehr als der Hälfte der Fälle kein tiefes Plateau auf. Für die Fälle, in denen ein tiefes Plateau auftritt, könnte die Dauer dieses Plateaus für [2;-] vs. [2.,] verglichen werden. Die Tabellen A.2 und A.4 (Anhang) zeigen, dass für DuL bei [2.,] Werte von 42–118 ms gemessen wurden, bei [2;-] Werte von 23–552 ms. DuL wird entscheidend von der Dauer, bzw. der Silbenzahl zwischen der Akzentsilbe und dem Phrasenende beeinflusst: Für den Wert DuL = 552 ms wurde eine hohe Silbenzahl Pa[syl] = 6 erhoben (vgl. Nr. 92 in Tabelle A.3). Die Bildung von Mittelwerten, bzw. ein direkter Vergleich von DuL für [2;-] vs. [2.,] erscheint daher unangebracht.

<i>Melodiemuster</i>	DuL = 0 ms	DuL > 0 ms	n.m.	<i>Summe</i>
[2..]	13 (56,5%)	8 (34,8%)	2 (8,7%)	23 (100,0%)
[2;]	64 (53,8%)	49 (41,2%)	6 (5,0%)	119 (100,0%)

Tabelle 5.3: Häufigkeiten von Konturen mit (DuL > 0 ms) und ohne (DuL = 0 ms) tiefes Plateau. n.m. = nicht messbar.

5.3.2 Prüfung der Hypothesen 1, 2 und 4

5.3.2.1 [2;] und Mikrointonation

Die Hypothese 1 (vgl. 4.2) soll und kann im Rahmen dieser Arbeit nicht erschöpfend behandelt werden. Es wurde argumentiert, dass die Wahrscheinlichkeit, dass es sich bei [2;] nur um ein mikrointonatorisches Phänomen handelt, als gering betrachtet werden kann, wenn [2;] in vielen verschiedenen segmentellen Kontexten auftritt. Es sei hierzu auf die Rohdaten verwiesen (Spalten S1 und S2 in Tabelle A.3): In den meisten Fällen (101 von 119, 84,9%) findet der Wiederanstieg von [2;] über finalen stimmhaften Segmenten statt, auf die keine weiteren stimmlosen Konsonanten folgen, welche einen mikrointonatorischen Wiederanstieg bedingen könnten. Insgesamt zeigt Spalte S1 sehr verschiedene segmentelle Grundlagen, über die sich die [2;]-Kontur legt. Diese seien hier nicht im Einzelnen aufgeführt. Hinsichtlich der Möglichkeit, dass das intrinsische F_0 von Vokalen zu leichten F_0 -Variationen führen kann, die wiederum inkorrekt als eine [2;]-Kontur klassifiziert werden könnten, seien kurz folgende Zahlen angeführt: In nur 33 Fällen (27,7%) findet der Wiederanstieg auf einem finalen Vokal statt. Darunter gibt es 2 geschlossene und 17 offene Vokale (einschließlich einiger der r-Vokalisierung), sowie 7 Diphthonge (einschließlich einiger r-Vokalisierung) und 7 mal Schwa (vgl. Tabelle A.3). Es erscheint aufgrund dieser Korpusdaten unwahrscheinlich, dass die pseudoterminale Kontur nur ein fehlinterpretiertes mikrointonatorisches Phänomen ist.

Zu erwähnen ist allerdings, dass sich die vorliegenden Daten hinsichtlich des segmentellen Materials für [2;] und [2..] unterscheiden: Der Wiederanstieg von [2;] findet in 52,1% der Fälle auf einem finalen Nasal statt und in nur 27,7% der Fälle auf einem finalen Vokal. Bei [2..] ist dieses Verhältnis gerade umgekehrt. In 65,2% der Fälle findet der Wiederanstieg von [2..] auf einem finalen Vokal statt und in nur 21,7% der Fälle auf einem finalen Nasal. Diese Verteilung kann zum Anlass für weitere Untersuchungen genommen werden, wird aber im weiteren Verlauf dieser Arbeit ignoriert.

5.3.2.2 [2;] vs. [2..]

Die Hypothese 2 kann in Bezug auf die gemessenen und errechneten Parameter sowie basierend auf den Beobachtungen der Voruntersuchung wie folgt konkretisiert werden:

- Die [2..]-Kontur unterscheidet sich von der [2;]-Kontur durch einen absolut und relativ (im Vergleich zum Gipfelmaximum) höheren Wert für ΔR_i .

- Die [2..]-Kontur unterscheidet sich von der [2;]-Kontur in *Timing*- sowie Dauerparametern durch
 - eine größere Wiederanstiegsdauer
 - einen früheren Einsatz des Wiederanstiegs in Relation zum letzten stimmhaften Segment der Phrase

Tabelle 5.4 zeigt die Mittelwerte und Standardabweichungen der Ergebnisse zur Wiederanstiegshöhe. Mit 6,60 st vs. 1,66 st liegt der Wiederanstieg für die [2..]-Konturen deutlich höher als für die [2;]-Konturen. Die Standardabweichung ist für [2..] sehr hoch, was eine phänomenale Ursache hat (vgl. 5.2.5). Der unterschiedliche absolute Wiederanstieg manifestiert sich auch relativ in Bezug auf die Gipfelhöhe: Im Mittel steigt F_0 in einer [2..]-Kontur auf ein deutlich höheres Niveau als das Gipfelmaximum des finalen Akzents (142% der Gipfelhöhe), während der Wiederanstieg von [2;] deutlich unter dem Niveau des Gipfelmaximums bleibt (38% der Gipfelhöhe).

<i>Melodiemuster</i>	<i>Parameter</i>				
	ΔRi [st]	DuRi [ms]	$g Ri$ [st/s]	DuRiRel	$\Delta RiRel$
[2..] (n=23)	6,60 (3,47)	137 (64)	48,9 (15,8)	1,07 (0,74)	1,42 (0,84)
[2;] (n=119)	1,66 (0,76)	71 (41)	28,1 (15,8)	0,62 (0,41)	0,38 (0,33)

Tabelle 5.4: Mittelwerte und Standardabweichungen (in Klammern) zu gemessenen und errechneten Parametern für [2;] vs. [2..].

Zur Frage nach dem *Timing* des Wiederanstiegs kann in einem ersten Schritt gezählt werden, in wie vielen Fällen der Wiederanstieg bereits vor dem letzten stimmhaften Segment einsetzt. Tabelle 5.5 listet die Ergebnisse dieser Zählung auf. Sowohl für [2;] als auch für [2..] beginnt der Wiederanstieg in den meisten Fällen erst auf dem letzten stimmhaften Segment ($Ri[seg]=1$). Ein früherer Einsatz wurde aber für [2..] etwas häufiger festgestellt als für [2;] (17,4% vs. 10,9%). Unter diesen Fällen finden sich für [2..] 3–5 Fälle, in denen der Wiederanstieg bereits in der vorletzten *Silbe* einsetzt, während sich für [2;] nur 1 solcher Fall findet (vgl. Tabelle A.1 und A.3).

Tabelle 5.4 zeigt die Mittelwerte und Standardabweichungen für die gemessenen bzw. errechneten Dauer- und *Timing*-Parameter DuRi und DuRiRel. Die Wiederanstiegsdauer von [2..] ist im Mittel fast doppelt so groß wie von [2;], wobei sich die Standardabweichungen aber stark überlappen. Der Wiederanstieg beginnt bei [2..] durchschnittlich am Beginn des letzten stimmhaften Segments, bei [2;] erst nach 40% der Segmentdauer.

Die Ergebnisse bestätigen also die eingangs formulierten Hypothesen, sowohl hinsichtlich der Wiederanstiegshöhe, als auch des *Timing*. Die Standardabweichungen überlappen sich im Falle der Wiederanstiegsdauer und des *Timing*-Parameters DuRiRel, nicht aber im Falle der

Wiederanstiegshöhe. Durch letztere scheinen die beiden untersuchten Melodiemuster weitaus schärfer getrennt zu sein als durch das *Timing* des Wiederanstiegs.

<i>Melodiemuster</i>	Ri[seg] = 1	Ri[seg] > 1	n.m.	<i>Summe</i>
[2..]	17 (73,9%)	4 (17,4%)	2 (8,7%)	23 (100,0%)
[2;]	103 (86,8%)	13 (10,9%)	3 (2,5%)	119 (100,0%)

Tabelle 5.5: Häufigkeiten des Wiederanstiegsbeginns *auf* dem letzten stimmhaften Segment (Ri[seg] = 1) und *vor* dem letzten stimmhaften Segment (Ri[seg] > 1). n.m. = nicht messbar.

5.3.2.3 Psychophonetische Aspekte

Im Hinblick auf Hypothese 4 wurde erstens die Rate des Wiederanstiegs (gRi) für jeden Einzelfall berechnet. Tabelle 5.4 zeigt die Mittelwerte von gRi für [2..] und [2;]⁴. Aus dem Mittelwert von $DuRi = 71$ ms kann der *glissando threshold* berechnet werden, für den der Wiederanstieg von [2;] gerade als dynamischer Ton wahrnehmbar sein müsste. Es ergibt sich mit Formel 2.3 eine nötige Anstiegsrate von $g_{thr} = 31,7$ st/s. Die gemessene und gemittelte Rate von gRi liegt mit 28,1 st/s in der Größenordnung des *glissando threshold*: $gRi \approx g_{thr}$. Damit wird Hypothese 4(a) bestätigt.

Ein Vergleich der Anstiegsraten von [2;] vs. [2..] ergibt einen Quotienten von $gRi_{2..}/gRi_{2;} = 48,9/28,1 \approx 1,74$. Dieser Wert ist deutlich kleiner als 2. Jedoch darf die Angabe von 't Hart et al. (1990) für den Mindestunterschied zweier diskriminierbarer Anstiegsraten nicht überinterpretiert werden, vielmehr beschreibt der Faktor 2 nur eine ungefähre Größenordnung. Unter diesem Aspekt stimmt der gefundene Faktor von 1,74 für [2;] vs. [2..] ungefähr mit der Angabe von 't Hart et al. (1990) überein. Die Hypothese 4(b) wird damit ebenfalls bestätigt. In der Diskussion (Kapitel 7) wird dieser Aspekt wieder aufgegriffen und mit den Ergebnissen aus dem Perzeptionsexperiment verglichen.

⁴Diese Werte sind das Resultat einer Mittelwertbildung über fallweise errechnete Anstiegsraten. Sie sind daher nicht aus den ebenfalls in der Tabelle 5.4 angegebenen Mittelwerten für ΔRi und $DuRi$ ableitbar!

Kapitel 6

Perzeptionsexperiment

6.1 Voraussetzungen und Ziele

Die Korpusanalyse (Kapitel 5) hat gezeigt, dass die mit [2;] etikettierten Konturen sich von [2..]-Konturen v.a. durch den deutlich kleineren Wiederanstieg unterscheiden. Ein weiterer Unterschied ist das *Timing*: [2;]-Konturen zeichnen sich im Vergleich zu den [2..]-Konturen im Mittel durch einen späteren und kürzeren Wiederanstieg aus. Anhand der deskriptiven Ergebnisse konnte abgeschätzt werden, dass die Höhe des Wiederanstiegs für die Unterscheidung von [2;] vs. [2..] von wesentlich größerer Bedeutung ist als das *Timing*. Diese Ergebnisse flossen in die Konstruktion des Perzeptionsexperiments ein: Es wurde ein Stimuluskontinuum generiert, in dem sich die Stimuli nur in einem Parameter, nämlich der Höhe des Wiederanstiegs unterscheiden. Laut der Korpusanalyse beträgt der durchschnittliche Wiederanstieg von [2;]-Konturen 1,66 st. Damit im Perzeptionsexperiment gewährleistet ist, dass mögliche Übergänge von einem terminalen zu einem pseudoterminalen und einem nicht-terminalen Muster erfasst werden können, sollte für den Unterschied zwischen zwei benachbarten Stimuli im Kontinuum ein Wert gewählt werden, der deutlich unter 1,66 st liegt. Es wurde mit 1 st ein Wert gewählt, der dieses Kriterium erfüllt, der aus methodischen Gründen allerdings noch relativ hoch ist (vgl. 6.2.5). Mögliche *Timing*-Unterschiede werden im Experiment weitgehend vernachlässigt: Sie werden nicht explizit untersucht, gehen aber implizit in die Manipulation des Sprachmaterials ein.

Ziel des Experiments ist die Prüfung der Hypothesen 3 und 4 (vgl. 4.2). In der Ergebnispräsentation (6.3) werden sie in Teilhypothesen zerlegt, die entsprechend der konkreten Analysen formuliert werden.

Die Beschreibung der Methode (6.2) enthält neben der Darstellung der Konstruktion, des Aufbaus und der Durchführung des Experiments (Abschnitte 6.2.3 bis 6.2.7) zusätzlich eine Einführung in die verwendete Methode sowie einige kritische methodische Vorüberlegungen (Abschnitte 6.2.1 und 6.2.2). Die statistische Auswertung der Ergebnisse ist *a priori* im Versuchsplan verankert.

6.2 Methode

6.2.1 Einleitendes zum semantischen Differential

Die Technik des semantischen Differentials geht zurück auf Osgood et al. (1957), die mit der Entwicklung dieser Methode das Ziel verfolgten, ein Werkzeug zum “Messen von Bedeutungen” zur Verfügung zu stellen. Die Methode ist keineswegs auf Forschungen im Bereich der Phonetik beschränkt. In der Intonationsforschung wurde sie bisher nur selten angewandt, z.B. von Uldall (1960; 1964) und Dombrowski (2003).

In den Grundzügen funktioniert die Methode (im Rahmen der Intonationsforschung) wie folgt. Es müssen Stimuli erzeugt werden, die hinsichtlich der untersuchten Parameter Unterschiede aufweisen. Bestehen diese Unterschiede nur bezüglich weniger Parameter (möglichst eines einzigen), und zwar derart, dass (a) die Stimuli sich nach der Maßzahl dieses Parameters ordnen lassen und (b) die Differenz dieser Maßzahl zwischen zwei (in der Ordnung benachbarten) Stimuli für alle dieser Stimuluspaare konstant ist, so liegt eine gewisse Systematik in der Menge der Stimuli vor. Eine derart geordnete Menge an Stimuli soll hier als ein *Stimuluskontinuum* bezeichnet werden. Die Methode kann – abhängig von der spezifischen Fragestellung – sowohl mit einem Stimuluskontinuum angewandt werden (vgl. Dombrowski 2003) als auch mit einer weniger systematischen Stimulusmenge (vgl. Uldall 1960; 1964).

Ferner muss eine Liste an semantischen Skalen erstellt werden. Eine solche Skala wird definiert durch zwei gegensätzliche Eigenschaftswörter, wie z.B. ‘abgeschlossen/ ungeschlossen’ oder ‘selbstbewusst/ schüchtern’, die die Randpunkte einer solchen Skala ausmachen. Die Anzahl der Stufen in dieser Skala (inkl. der beiden Randpunkte) ist grundsätzlich bei der Versuchsplanung frei wählbar und betrug sowohl bei Uldall (1960; 1964) als auch bei Dombrowski (2003) sieben. Je feiner die Abstufung wird, desto eher lässt sich in der statistischen Auswertung der Daten die semantische Skala angenähert als intervallskaliert interpretieren¹, was sowohl von Uldall als auch von Dombrowski vorgenommen wurde. Wesentlich mehr Skalenpunkte als sieben sollten aus Gründen der Übersichtlichkeit des Fragebogens vermieden werden. Der mittlere Wert der Skala repräsentiert eine Unentschiedenheit zwischen den beiden semantischen Kategorien an den Rändern der Skala. Wird eine gerade Anzahl an Skalenpunkten (z.B. sechs oder acht) gewählt, so enthält der Versuchsplan keinen solchen mittleren Wert; es handelt sich dann also um ein *forced-choice*-Experiment. Eine *forced-choice*-Komponente hat den Nachteil, dass bei Ergebnissen, die nur als Tendenz auftreten (z.B. “Stimulus 3 wird von den Hörern eher als ‘abgeschlossen’ eingestuft”) unklar bleibt, ob diese Tendenz nur ein Artefakt der Methode ist. Zudem bedeutet sie einen zusätzlichen Stressfaktor für die Probanden, da sie sich stets für eine Richtung entscheiden müssen, auch wenn es ihnen nicht möglich erscheint.

In der grafischen Darstellung einer semantischen Skala wird jede Stufe durch ein Feld

¹Dieser Aspekt wird in Abschnitt 6.2.2 etwas ausführlicher diskutiert.

repräsentiert, wobei diese Felder waagerecht angeordnet und die äußeren beiden Felder mit den Eigenschaftswörtern beschriftet sind (vgl. Abb. 6.1). Die einzelnen Stufen können verbalisiert werden: Uldall (1960) wies ihre Probanden darauf hin, dass die äußeren Felder (in Abb. 6.1 Felder mit der Nr. 3) *extremely* (z.B. *'bored/ interested'*) bedeuten, die Felder eine Stufe weiter zur Mitte hin (Nr. 2) *quite*, die Felder, die an die Mitte angrenzen (Nr. 1) *slightly* und die Mitte selbst *neutral* oder *neither*. Auch die Verbalisierung der Skalen selbst kann über die Nennung der Antonyme an den Rändern der Skala hinausgehen. Dombrowski (2003) bettete die Skalen in je einen (schriftlich präsentierten, auf die jeweilige Skala angepassten) Trägersatz ein, z.B. *"The utterance is stating something new/ something known"*.

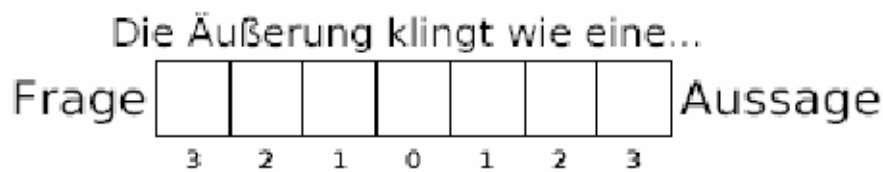


Abbildung 6.1: Beispiel für die grafische Präsentation einer semantischen Skala, die in einen Trägersatz eingebettet ist.

Bei der Durchführung des Experiments wird den Versuchspersonen jeweils eine Skala zusammen mit einem auditiven Stimulus präsentiert, der auf dieser Skala einzuordnen ist. In der Praxis kann die Präsentation der Skala z.B. als Fragebogen auf Papier erfolgen (vgl. Uldall) oder auch durch ein Computerprogramm (vgl. Dombrowski). Jeder Versuchsperson wird jede mögliche Kombination aus semantischen Skalen und Stimuli jeweils einmal vorgelegt (bei z.B. 17 Stimuli und 9 semantischen Skalen muss jede Versuchsperson also 153 Urteile abgeben)².

Uldall (1960) versuchte, die Bedeutungen von einer Reihe sehr verschiedener Intonationsverläufe über verschiedenen Äußerungen (Aussage, zwei Typen von Fragen, Befehl) zu vermessen. Dabei stellte die Menge der Stimuli allerdings kein Stimuluskontinuum (s.o.) dar. Die Auswahl der verwendeten semantischen Skalen deckte ein sehr breites Spektrum von kommunikativen Funktionen sowie Einstellungen und Emotionen des Sprechers ab. Es ging hier also zunächst einmal darum, eine neue Methode für die Intonationsforschung vorzustellen und in einem ersten Versuch anzuwenden. Dabei war das Vorgehen noch wenig hypothesesgeleitet. Im Rahmen der Datenauswertung wurde eine Faktorenanalyse durchgeführt, um zu testen, welche Gruppierungen von semantisch verwandten Skalen für den größten Teil der Varianz in den Daten verantwortlich sind. Jeder dieser Gruppierungen wurde ein Name gegeben, der wie auch die Skalen selbst aus zwei Antonymen besteht und der versucht, die in

²Uldall (1964) untersuchte eine relativ kleine Probandenzahl und ließ jeden Stimulus auf jeder Skala zweimal beurteilen, u.a. um die Stichprobe zu vergrößern. Sie arbeitete damit methodisch unsauberer als Uldall (1960) und Dombrowski (2003), deren Stichprobengröße (N) tatsächlich auf N Versuchspersonen beruht.

der Gruppe enthaltenen Skalen geeignet zusammenzufassen³. Als die stärksten drei Faktoren ergaben sich die folgenden (in Klammern ist angegeben, zu welchem Anteil der jeweilige Faktor für die Varianz in den Daten verantwortlich ist):

- Faktor I: *pleasant/ unpleasant* (mehr als 50%)
- Faktor II: *interest/ lack of interest* (etwa 20%)
- Faktor III: *authoritative/ submissive* (8–13%)

Uldall (1960) stellt fest, dass dieses Ergebnis vergleichbar ist mit dem Ergebnis von Osgood et al. (1957), die die Faktoren mit den Termini *evaluation*, *activity* und *potency* bezeichnen⁴. Diese drei stärksten Faktoren werden auch als *Dimensionen der emotionalen Bedeutung*⁵ bezeichnet.

Dombrowskis (2003) Versuchsplan baut auf diesen Ergebnissen auf, unterscheidet sich jedoch dadurch von Uldall (1960; 1964), dass er wesentlich stärker hypothesengeleitet ist. Dementsprechend wählte Dombrowski (2003) einen Teil der semantischen Skalen basierend auf den Hypothesen aus, einen anderen Teil aber derart, dass die drei Dimensionen *a priori* in seinem Skalenset enthalten sind. Folglich kann ein Vorgehen wie in Dombrowskis Experiment als Weiterentwicklung der Methode betrachtet werden, in der die Durchführung einer Faktorenanalyse nicht mehr nötig ist. Die Auswertung beschränkt sich darauf, durch geeignete Verfahren für Mittelwertvergleiche (Varianzanalysen und t-Tests) zu prüfen, ob sich die laut der Hypothesen erwarteten Unterschiede zwischen den Mittelwerten als signifikant bestätigen lassen.

6.2.2 Kritische Bemerkungen zur Methode und Konfigurationen im vorliegenden Experiment

Dombrowski (2003) schlägt vor, dass in weiteren Forschungen, die sich semantischer Differentiale bedienen, versucht werden sollte, die Methode zu standardisieren, so dass ein optimales Set an semantischen Skalen für die Intonationsforschung erarbeitet wird. Eine totale Standardisierung aber würde im Widerspruch stehen zu einer hypothesengeleiteten Auswahl der semantischen Skalen. Ein Set, welches für jedes Experiment in der Intonationsforschung verwendet werden könnte, müsste vermutlich sehr umfangreich sein und für die spezielle Fragestellung eine Menge an redundanten Skalen beinhalten. Zum Beispiel wurden die von

³Aus Uldall (1960) geht nicht deutlich hervor, welche Skalen jeweils zusammengefasst wurden, so dass hier kein Beispiel gegeben werden kann.

⁴Die Gleichsetzung von Uldalls (1960) Faktor II mit *activity* erscheint möglicherweise weniger plausibel als in den anderen beiden Fällen, verdeutlicht aber, wie frei und unscharf semantische Kategorien im Rahmen dieser Methode interpretiert werden.

⁵Im Folgenden auch (*semantische*) *Dimensionen* oder *Dimensionen von semantischen Differentialen* genannt.

Dombrowski (2003) verwendeten Skalen ‘neu/ bekannt’ und ‘überrascht/ nicht überrascht’ aufgrund der speziellen Hypothesen zu einer ganz anderen Fragestellung (nämlich bezüglich der Semantik der drei von Kohler (1991b) vorgeschlagenen Positionen von F_0 -Gipfeln ‘früh’, ‘mittel’ und ‘spät’) ausgewählt und erscheinen im Hinblick auf die hypothetische Semantik der hier untersuchten Intonationsmuster nicht notwendig.

Daher wurde im hier beschriebenen Experiment kein Versuch unternommen, zu einer Standardisierung eines semantischen Differentials für die Intonationsforschung beizutragen. Ansonsten orientiert sich das hier beschriebene Experiment methodisch weitgehend an Dombrowski (2003): Der Versuchsplan ist hypothesengeleitet, es wird ein systematisch erzeugtes Stimuluskontinuum verwendet und die Auswahl der semantischen Skalen basiert (a) auf den Hypothesen und (b) auf Vollständigkeit hinsichtlich der drei semantischen Dimensionen. Es ist allerdings zu berücksichtigen, dass das hypothesengeleitete Vorgehen dieser Untersuchung sich in einem Aspekt von Dombrowskis Vorgehen unterscheidet: Die Qualität der zugrundeliegenden Hypothesen in den beiden Untersuchungen ist verschieden. Dombrowskis Hypothesen basieren auf früheren experimentellen Untersuchungen (Kohler 1991b) und sind damit fundierter als die Hypothesen der vorliegenden Arbeit, da letztere nur auf Einzelbeobachtungen basieren (Peters 1999).

Die statistische Auswertung orientiert sich in den Grundzügen an Dombrowski (2003) und beinhaltet:

- eine *multivariate* Varianzanalyse im Messwiederholungsdesign (MANOVA)
- für jede semantische Skala eine *univariate* Varianzanalyse im Messwiederholungsdesign (ANOVA)

Die ANOVAs testen, ob sich zwischen den Antworten für die Stimuli auf der gewählten semantischen Skala in der Gesamtheit signifikante Mittelwertunterschiede ergeben. Die MANOVA leistet grundsätzlich dasselbe, jedoch unter Berücksichtigung aller Skalen. Für nähere Erläuterungen zu den verwendeten statistischen Verfahren wird auf Bortz (1979) verwiesen. Bei der Ergebnispräsentation (Abschnitt 6.3) wird etwas detaillierter auf die statistische Auswertung eingegangen. Hinzu kommen an geeigneten Stellen Einzelvergleiche mit Hilfe von t-Tests. Die statistischen Auswertungen werden mit Hilfe der Software SPSS (z.B. Bühl und Zöfel 2002) durchgeführt.

Ebenso in Anlehnung an Dombrowski (2003) wird jeder Stimulus auf jeder Skala genau einmal von jeder Versuchsperson beurteilt. Diese Paare aus Stimulus und Skala werden in randomisierter Reihenfolge präsentiert. Um das Experiment für die Versuchspersonen so klar wie möglich zu gestalten, werden sowohl die einzelnen Stufen der Skala verbalisiert (vgl. Uldall 1960), als auch die Skala selbst in einen Trägersatz eingebettet (vgl. Dombrowski 2003).

Im Vorfeld der Untersuchung muss der Versuchsplan (d.h. die Organisation der Stimuli, der semantischen Skalen und der statistischen Verarbeitung) hinsichtlich der Interpretationsmöglichkeiten der erhobenen Daten kritisch betrachtet werden. Aus den Ergebnissen von Dombrowski (2003) kann gefolgert werden, dass die Methode sehr sensitiv, möglicherweise sogar *zu* sensitiv ist. Denn relativ geringe Mittelwertunterschiede (etwa 0,7 Skalenpunkte auf einer siebenstufigen Skala) sind sogar bei einer Irrtumswahrscheinlichkeit von $\alpha = 0,01$ als statistisch signifikante Unterschiede bestätigt worden.

Es muss hierbei also unabhängig von der Prüfstatistik erwogen werden, wie groß ein Mittelwertunterschied mindestens sein muss, damit es plausibel erscheint, diesen als relevant zu interpretieren. Dombrowski (2003) setzt $\alpha = 0,01$ und interpretiert die sich dabei als signifikant ergebenden Unterschiede als relevant, was in seinen Daten einem nötigen Mindestunterschied von etwa 0,7 Skalenpunkten entspricht.

Hier soll argumentiert werden, dass ein solcher Unterschied zu gering ist, um als relevant betrachtet zu werden. Beträgt z.B. der Mittelwert für die Beurteilung eines Stimulus *A* (auf einer Skala *S* mit sieben Stufen von -3 bis $+3$) $a = 1,3$ Skalenpunkte, der Mittelwert eines zweiten Stimulus *B* (auf derselben Skala *S*) $b = 0,6$ Skalenpunkte, so beträgt der Mittelwertunterschied $a - b = 0,7$ Skalenpunkte. Es erscheint jedoch plausibel, beide Werte in derselben Weise zu interpretieren, nämlich: "Weder Stimulus *A* noch *B* wird deutlich in Richtung auf einen der Randpunkte der Skala beurteilt. Beide aber werden tendenziell in Richtung auf den rechten Rand der Skala beurteilt, jedoch mit einer deutlich stärkeren Tendenz zur Unentschiedenheit (Mitte der Skala)." Mit anderen Worten: Es sind (für jede Richtung der Skala) nur drei Werte verbalisiert (1 = "ein wenig", 2 = "ziemlich", 3 = "sehr"). Werden die Mittelwerte auf einen dieser verbalisierten Werte gerundet, so ergibt sich in beiden Fällen der Wert 1 ("ein wenig").

Es wird hier ein grundsätzliches Problem des Verfahrens deutlich: Die siebenstufige Skala wird als intervallskaliert betrachtet (ansonsten könnten keine Mittelwerte gebildet werden). Natürlich kann es sich hierbei nur um eine angenäherte bzw. geschätzte Intervallskalierung handeln, denn es fehlt eine objektive Möglichkeit zu prüfen, ob der "Abstand" auf der semantischen Skala zwischen zwei benachbarten Skalenpunkten wirklich konstant ist⁶. Die Versuchspersonen haben zudem nur die Möglichkeit, aus den sieben gegebenen Stufen zu wählen. Werte wie z.B. 1,4, die rechnerisch zwischen diesen Stufen liegen, existieren auf der dargebotenen Skala nicht, und stellen somit auch keine möglichen Werte in den Rohdaten dar. Die Mittelwertbildung (und damit die Einschätzung der siebenstufigen semantischen Skala als intervallskaliert) bedeutet also eine Konstruktion empirisch unmöglicher Werte.

Dieses Problem könnte umgangen werden, wenn die Skala nicht als intervall- sondern als ordinalskaliert behandelt werden würde und entsprechend nicht-parametrische Testverfahren herangezogen würden. Ein globaler multivariater Test könnte in diesem Fall allerdings nicht

⁶Bei einer "echten" Intervallskala (z.B. der Temperatur in °C) existieren in der Regel geeichte Messverfahren.

durchgeführt werden⁷. Außerdem haben sich in der Praxis die parametrischen Verfahren für derartige Skalenmessungen – nicht nur im Bereich der Intonationsforschung – durchgesetzt (z.B. Bergler 1975). Eine grundlegende Meta-Analyse zur Methode selbst kann im Rahmen dieser Arbeit nicht geleistet werden.

Daher soll auch in der vorliegenden Arbeit die Intervallskalierung für die semantischen Skalen angenommen werden. Allerdings sollte bei der Interpretation der Daten und der Prüfstatistik konservativ vorgegangen werden. Es erscheint als Konsequenz dieser Diskussion sinnvoll, den relevanten Mindestunterschied zwischen zwei Mittelwerten auf 1 Skaleneinheit zu setzen. Auf diese Weise kann ausgeschlossen werden, dass zwei Mittelwerte als relevant verschieden interpretiert werden, obwohl sie tendenziell dasselbe aussagen, da sich theoretisch beim Runden auf ganze Zahlen (also die verbalisierten Werte) stets verschiedene Ergebnisse für die beiden Mittelwerte ergeben würden. Zusätzlich soll sich ein Mittelwertunterschied natürlich auch (auf dem 1%-Niveau, wie bei Dombrowski 2003) als statistisch signifikant erweisen, ehe dieser als relevant betrachtet wird. Die Varianzanalysen werden dabei möglichst konservativ gerechnet: In der MANOVA wird als Prüfgröße Wilks *Lambda* gewählt; in den ANOVAs wird eine Korrektur der Freiheitsgrade nach Greenhouse-Geisser vorgenommen (vgl. Bortz 1979). In der Ergebnispräsentation (Abschnitt 6.3) wird keine Unterscheidung zwischen verschiedenen Signifikanzniveaus vorgenommen, sondern es werden nur die auf dem 1%-Niveau signifikanten Ergebnisse mit ‘*’ markiert.

6.2.3 Auswahl der semantischen Skalen

Hinsichtlich der vorgeschlagenen Funktionen der pseudoterminalen Kontur (vgl. 3.5) sollte die Auswahl der semantischen Skalen Urteile über die Stimuli in Bezug auf die Eigenschaften ‘Abgeschlossenheit’ und ‘Kategorialität’ erfassen können, die hier als die wichtigsten Aspekte der ‘Terminalität’ verstanden werden (vgl. 4.1.3). Außerdem müssten sich unterschiedliche Beurteilungen in Bezug auf die ‘Dominanz’ des Sprechers ergeben.

Der Aspekt der ‘Sicherheit’ des Sprechers wird von Peters aus zwei Richtungen betrachtet: Zum einen kann die Sprecherin durch Verwendung der pseudoterminalen Kontur vermeiden, “zu [...] *unsicher zu wirken*” (Peters 1999), andererseits signalisiert sie durch diese Kontur “*eine gewisse Unsicherheit*” (Peters 2000). Eine eindeutige Hypothese über die Beurteilung der Stimuli in Bezug auf diese Eigenschaft kann also nicht formuliert werden; jedoch scheint dieser Aspekt für die Charakterisierung der pseudoterminalen Kontur möglicherweise relevant zu sein und sollte daher in der Auswahl der Skalen in zweifacher Weise berücksichtigt werden. Zum einen kann der Sprecher Sicherheit in Bezug auf das Gesagte signalisieren, zum anderen aber auch Sicherheit als Teil seiner Sprecherpersönlichkeit.

Ebenfalls muss das semantische Differential erfassen können, ob der Stimulus wie ei-

⁷Es konnte im Rahmen der Recherchen zu dieser Arbeit zumindest kein solches Verfahren ausfindig gemacht werden.

ne Frage klingt. Es ist zu erwarten, dass nur deutlich fallend-steigende Konturen als Frage beurteilt werden. Hypothesengeleitet ergeben sich also sechs semantische Skalen, die durch folgende Antonyme beschrieben werden können⁸:

- abgeschlossen/ unabgeschlossen
- kategorisch/ verhandlungsbereit
- dominant/ nicht dominant
- nicht sicher/ sicher
- selbstbewusst/ schüchtern
- Frage/ Aussage

Es sollen neben den hypothesengeleiteten Skalen auch solche berücksichtigt werden, die die drei grundlegenden Dimensionen semantischer Differentiale repräsentieren (vgl. Abschnitt 6.2.1). Die Dimension der *Potenz* ist bereits durch die Skala ‘dominant/ nicht dominant’ vertreten. Die Dimension der *Aktivität* soll durch die zusätzliche Skala ‘interessiert/ nicht interessiert’ repräsentiert werden. Die Dimension der *Evaluation* ist möglicherweise durch die Skalen ‘kategorisch/ verhandlungsbereit’ und ‘selbstbewusst/ schüchtern’ bereits vertreten, da ‘Verhandlungsbereitschaft’ und ‘Selbstsicherheit’ wohl im Allgemeinen als positiv zu bewertende Eigenschaften aufgefasst werden. Um aber hier eine eindeutigere Skala bezüglich der Evaluation zu berücksichtigen, wird zusätzlich die Skala ‘unhöflich/ höflich’ aufgenommen.

Bei der Wahl dieser beiden zusätzlichen Skalen wäre es denkbar, die Skalen derart zu wählen, dass sie wesentlich deutlicher die Dimensionen repräsentieren (z.B. ‘aktiv/ passiv’ für *Aktivität*). Es wurde jedoch erstens versucht, die Skalen möglichst konkret und leicht nachvollziehbar zu gestalten — eine Formulierung wie “die Sprecherin klingt aktiv” würde vermutlich Verwirrung bei den Versuchspersonen auslösen, während die These “die Sprecherin wirkt interessiert” deutlich greifbarer und damit leichter zu beurteilen sein sollte. (Analoges gilt für die *Evaluation*: Hier hätte auch die Skala ‘angenehm/ unangenehm’ (vgl. Dombrowski 2003) gewählt werden können.) Zweitens wurde versucht, auch bei der Wahl dieser zusätzlichen Skalen möglichst hypothesenorientiert vorzugehen (selbst wenn sie nicht direkt aus den Hypothesen abgeleitet sind, wie die ersten sechs Skalen). Die Skalen sollten also zumindest mittelbar aus den Hypothesen abgeleitet sein: Zeigt der Sprecher an, dass er verhandlungsbereit ist und weniger dominant erscheinen möchte, so könnte ihm ein gewisses Interesse am Gesprächspartner unterstellt werden; wirkt der Sprecher weniger kategorisch, so könnte er möglicherweise als höflicher beurteilt werden.

⁸Die Anordnung der Antonyme bezüglich der linken und der rechten Seite der Skala ist hier bereits so gewählt, wie sie tatsächlich verwendet wurde.

Die Methode des semantischen Differentials ist ursprünglich entwickelt worden, um insbesondere die emotionalen Aspekte der Bedeutung zu vermessen. Sie muss jedoch nicht diesbezüglich eingeschränkt verwendet werden. Sowohl bei Dombrowski (2003) als auch in dieser Untersuchung werden auch nicht-emotionale Aspekte der Bedeutung von Intonationskonturen berücksichtigt: Signalisiert der Sprecher durch seine Intonation z.B., dass seine Äußerung noch nicht abgeschlossen ist, sondern fortgesetzt wird, so kann hier nicht von einer emotionalen Funktion der Intonation gesprochen werden. Eine klare Trennung von emotionaler vs. nicht-emotionaler Bedeutung ist jedoch nicht immer eindeutig (vgl. 3.2.2). Mit einer Art Kontrollskala ‘emotional/ sachlich’ soll daher versucht werden, die Stimuli in Bezug auf die Emotionalität, die sie ausdrücken, beurteilen zu lassen. Wird dann z.B. ein Stimulus x von den Hörern sowohl als ‘verhandlungsbereit’, als auch als ‘sachlich’ klassifiziert, so kann gefolgert werden, dass die Hörer die Verhandlungsbereitschaft, die x ausdrückt, nicht mit etwas Emotionalem in Verbindung gebracht haben.

Es ergibt sich also für die vorliegende Untersuchung ein Set aus neun semantischen Skalen, vgl. Tabelle 6.1. Die Anordnung der Antonyme bezüglich der linken und der rechten Seite der Skala ist so gewählt, dass negierte Wörter (‘un-’ oder ‘nicht’) auf beiden Seiten erscheinen. Den Probanden soll nicht nahegelegt werden, dass eines der Antonyme “negativer” ist als das andere. Daher werden auch in der grafischen Präsentation der Skalen keine Vorzeichen (‘+’, ‘-’) bei der Beschriftung der Felder verwendet (vgl. Abbildung 6.1). In der Auswertung werden die Urteile zu den links stehenden Begriffen allerdings durch negative Zahlen kodiert, z.B. ‘emotional’ = $\{-1, -2, -3\}$. Die Reihenfolge der Skalen selbst ist hier willkürlich.

Nr.	<i>Polung</i>	
	–	+
1	abgeschlossen	unabgeschlossen
2	dominant	nicht dominant
3	emotional	sachlich
4	Frage	Aussage
5	interessiert	nicht interessiert
6	kategorisch	verhandlungsbereit
7	nicht sicher	sicher
8	selbstbewusst	schüchtern
9	unhöflich	höflich

Tabelle 6.1: Die neun semantischen Skalen des vorliegenden Experiments.

6.2.4 Auswahlkriterien für das Sprachmaterial

Das Stimuluskontinuum basiert auf einem Korpusbeispiel mit einer [2;]-Etikettierung, nämlich der Äußerung “ab Montag, dem achten” (Datei *g101a002*, vgl. Anhang A), produziert von einem weiblichen Sprecher. Das Beispiel wurde aufgrund folgender Kriterien ausgewählt.

- Es handelt sich bei der gewählten Äußerung um einen authentischen Fall eines Musters, welches als pseudoternale Kontur etikettiert wurde. Dieser repräsentiert dabei in Bezug auf die prosodische als auch die segmentelle Struktur die am häufigsten vorkommende Manifestation dieses Musters (finaler Akzent: [2[^]]⁹; finales Segment: /n/ (vgl. 5.3.2.1).
- Es kann ausgeschlossen werden, dass der leichte Wiederanstieg in segmentell bedingter mikroprosodischer Variation begründet ist. Diese Annahme könnte laut Abschnitt 2.3.4 nicht ohne Weiteres aufgestellt werden, wenn auf den Wiederanstieg ein finaler stimmloser Konsonant folgte (z.B. /t/), oder wenn der Wiederanstieg auf einem schließenden Diphthong stattfände (z.B. /ae/). Ein Beispiel mit dem finalen Wort “Zeit” wäre also besonders ungünstig. Im gewählten Beispiel findet der Wiederanstieg jedoch auf einem finalen /n/ statt.
- Das gewählte Material weist im Bereich des finalen Segments keine (besonders auffällige) Abweichung der Stimmqualität von der Modalstimme auf (in einigen Beispielen wurde eine geknarrte Stimmqualität beobachtet). So kann eine problemlose Stimulusgenerierung gewährleistet werden.
- Die Äußerung kann aufgrund ihrer syntaktischen und semantischen Struktur (es handelt sich um eine Ellipse) pragmatisch sowohl als (abschließende) Aussage, als auch als Frage verstanden werden. Diese Eigenschaft ist von besonderer Bedeutung für die Prüfung der Hypothesen über die Semantik der untersuchten Melodiemuster, insbesondere für die Abgrenzung des pseudoternalen vom nicht-terminalen Muster.

6.2.5 Erzeugung der Stimuli

6.2.5.1 Vorbemerkungen zur Konzeption und Erzeugung der Stimuli

Zur Erzeugung der Stimuli wird die PSOLA-Resynthese von *praat* verwendet, die in Abschnitt 2.2 erläutert wurde. In *praat* wird der F_0 -Verlauf durch *pitch points* definiert, die sich durch eine exakte Eingabe von Zeitpunkt und F_0 -Wert präzise positionieren lassen.

Die generelle Idee zum Aufbau des Stimuluskontinuums sieht wie folgt aus. Die Stimuli unterscheiden sich lediglich in dem Parameter F_0 fin, der den F_0 -Wert des finalen *pitch points* beschreibt. Dazu wird, ausgehend vom präparierten und stilisierten Original¹⁰, F_0 fin in beide Richtungen der vertikalen Achse verschoben, so dass der vorletzte *pitch point* als *Angelpunkt* des Kontinuums bezeichnet werden kann. So entstehen beim Verschieben nach oben (a) Stimuli mit einer deutlich fallend-steigenden Kontur und beim Verschieben nach

⁹Das Beispiel stellt allerdings keinen eindeutigen Fall eines mittleren Gipfels dar. Ich halte die Etikettierung nicht für inadäquat, jedoch wäre auch eine Etikettierung als früher Gipfel [2] denkbar.

¹⁰Dieses wird in den folgenden Abschnitten näher erläutert.

unterschiedlich stark ausgeprägten fallenden Konturen. Dabei werden die Werte für F_0 fin derart gewählt, dass der Unterschied in der Höhe des Wiederanstiegs für zwei aufeinanderfolgende Konturen im Kontinuum einen konstanten Wert in Halbtonschritten aufweist.

Die Stimuli konnten nicht direkt basierend auf dem gewählten Sprachmaterial erzeugt werden. Erstens ist es in Experimenten dieser Art hilfreich, den F_0 -Verlauf auf dem zu manipulierenden Abschnitt zu stilisieren (z.B. 't Hart et al. 1990). Zweitens können während der Interaktion zwischen Experimentator und dem Resynthesesystem Hindernisse zum Vorschein kommen, die weitere Präparationen des Sprachmaterials erforderlich machen. Jene ersten Manipulationsversuche, die schließlich zum eigentlichen Ausgangsmaterial der Stimulusreihe führten, werden hier als "informelle Vorversuche" beschrieben. Dieser Abschnitt ist derart untergliedert, dass die Progression der verschiedenen Schritte der Vorverarbeitung deutlich wird. Entsprechend gibt es zwei Abschnitte "Informelle Vorversuche und Präparation I/II", die sich auf unterschiedliche Schritte in dieser Progression beziehen. Obwohl diese nötige Vorverarbeitung die Erstellung der Stimuli verkomplizierte, wurde anstelle des Beispiels "ab Montag, dem achten" (*g101a002*) kein anderes Beispiel gewählt (welches vielleicht einfacher zu verarbeiten gewesen wäre), da unter den vorhandenen Korpusbeispielen kein zweites Beispiel die Gesamtheit der Bedingungen im letzten Abschnitt (6.2.4) so gut erfüllte wie der gewählte Fall.

6.2.5.2 Informelle Vorversuche und Präparation I

Das Ausgangsbeispiel weist einen auffällig starken Glottalverschluss am Beginn der Phrase und eine auffällig kurze Vokaldauer (21ms) im ersten Wort ("ab") auf. Es wurde testweise eine Stimulusreihe erzeugt und informell drei Personen¹¹ mehrfach vorgespielt. Es bestand Übereinstimmung, dass die Auffälligkeit des ersten Wortes das Experiment stören könnte. Daher wurde in einem ersten Verarbeitungsschritt der erste Vokal des Sprachmaterials synthetisch manipuliert: Mit Hilfe der Software *wavesurfer* (Sjölander und Beskow URL) wurde die Intensität des Vokals auf 60% reduziert; mit *praat* wurde zusätzlich die Vokaldauer auf 219% des ursprünglichen Wertes erhöht (21ms → 46ms). Bei diesen Schritten wurde Wert darauf gelegt, dass das manipulierte Material ein möglichst natürliches Perzept hervorrief. Daher wurde möglicherweise keine extreme Reduzierung der Auffälligkeit des Wortes "ab" erreicht. Das Original und die manipulierte Version befinden sich auf der CDROM.

6.2.5.3 Stilisierung

Das präparierte Material wurde in einem zweiten Schritt als direkte Vorbereitung für die Erzeugung der Stimuli ein weiteres Mal manipuliert: Die F_0 -Kontur wurde (mit *praat*) stilisiert.

¹¹Darunter waren zwei Phonetiker des IPDS.

Hier wurde allerdings keine automatische Stilisierung von *praat* vorgenommen¹², sondern die stilisierte Kontur wurde manuell erzeugt. Dabei wurde ausschließlich der F_0 -Verlauf auf dem finalen Nasal manipuliert, da erstens nur dieser für die Erzeugung der Stimuli relevant war und zweitens der Ausgangsstimulus auf diese Weise zu großen Teilen authentisch verbleiben konnte.

Die manuelle Stilisierung wurde wie folgt vorgenommen. In einer nicht-stilisierten Kontur gibt es für jeden PSOLA-Baustein (vgl. 2.2) einen *pitch point*. Bei der Stilisierung wird versucht, die ursprüngliche Kontur durch möglichst wenige *pitch points* zu ersetzen, ohne dass die wesentlichen Kontureigenschaften verändert werden. Es wurden also im vorliegenden Beispiel sämtliche *pitch points* auf dem finalen Nasal gelöscht und manuell durch vier neue *pitch points* ersetzt. Die Anzahl vier ergab sich hierbei aus den gemessenen Daten für dieses Beispiel (vgl. Kapitel 5), denn es handelt sich um einen Fall mit tiefem Plateau. Es sind also *pitch points* zu folgenden Zeitpunkten nötig: (1) zu Beginn des Nasals, (2) zu Beginn des tiefen Plateaus, (3) am Ende des tiefen Plateaus und (4) am Ende des leichten Wiederanstiegs. Die Werte für (2), (3) und (4) (= FaOf, RiOn und RiOf aus der Korpusanalyse) konnten der Tabelle A.4 (Anhang A) entnommen werden; der Wert für (1) musste zum Zweck dieser Stilisierung nachträglich erhoben werden (191 Hz, gemessen 20 ms nach Beginn der Periodizität im Oszillogramm). Die genauen Zeitpunkte für die Positionierung der *pitch points* ließen sich aus den Dauern DuLS, DuL, DuRi und DuE (vgl. 5.2.3) bestimmen. Die resultierenden Zeitpunkte sowie die F_0 -Werte der *pitch points* werden in Tabelle 6.2 zusammengefasst.

<i>Pitch point:</i>	(1)	(2)	(3)	(4)
F_0 [Hz]	191	172	170	197
Zeitpunkt [ms]	20	114	168	234

<i>Parameter:</i>	DuLS	0–PP1	PP1–PP2	DuL	DuRi	DuE
[ms]	256	20	94	54	66	22

Tabelle 6.2: **Oben:** F_0 -Werte und Zeitpunkte (ausgehend von 0 ms = Beginn der Periodizität des Nasals) der vier *pitch points* für die manuelle Stilisierung des Sprachmaterials. **Unten:** Zugehörige Dauerwerte (z.T. aus der Korpusanalyse): DuLS = Dauer des Nasals; 0–PP1 = Nasal onset bis *pitch point* (= PP) 1; PP1–PP2 = PP 1 bis PP 2; DuL = D. des tiefen Plateaus bzw. PP 2 bis PP 3; DuRi = D. des Wiederanstiegs bzw. PP 3 bis PP 4; DuE = D. des nicht vermessbaren Abschnitts bzw. PP 4 bis Nasaloffset.

Eine stilisierte F_0 -Kontur sollte sich perceptiv nicht vom Original unterscheiden. Dies sollte im vorliegenden Fall zutreffen. Auf der CD-ROM (Anhang C) befindet sich neben dem Original und der manipulierten Version (vgl. 6.2.5.2) auch diese stilisierte Version. Abbildung 6.2 zeigt das Manipulationsfenster von *praat*, in dem die hier beschriebenen Operationen erkennbar sind (mit Ausnahme der Intensitätsmanipulation auf “ab”, welche mit *wavesurfer* vorgenommen wurde).

¹²Diese Option von *praat* liefert nicht immer ein wünschenswertes Ergebnis, so dass häufig manuelle Korrekturen vorgenommen werden müssen.

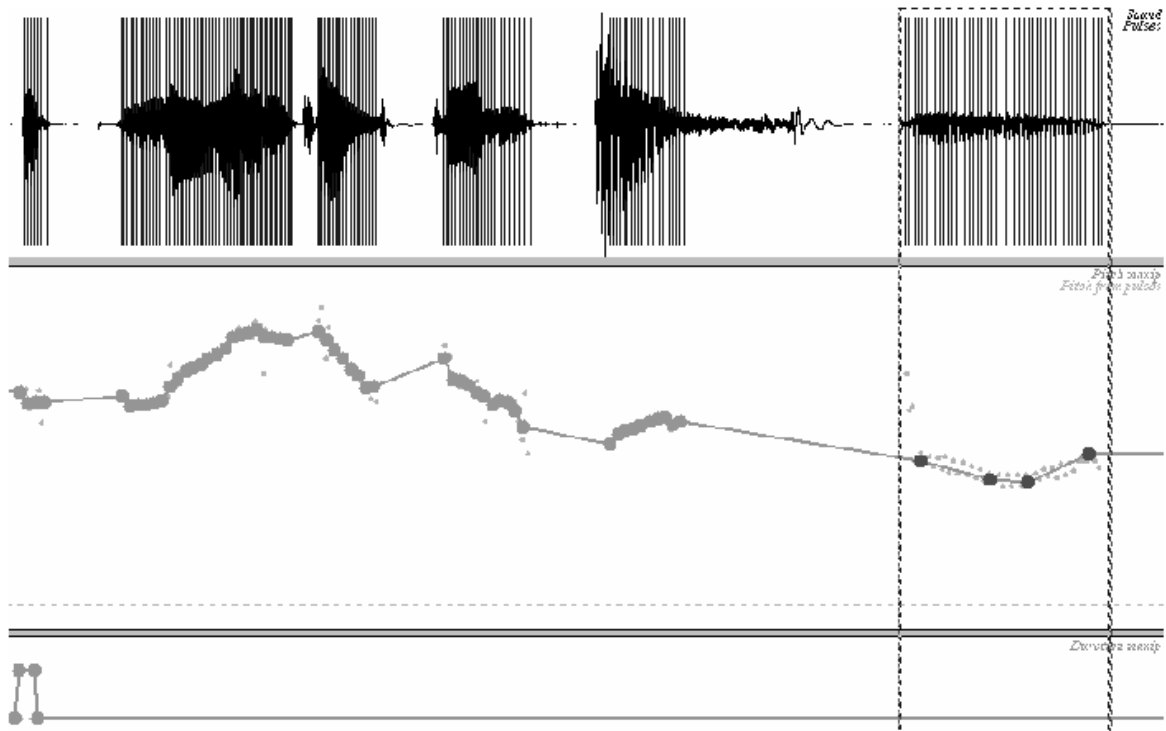


Abbildung 6.2: Manipulationsfenster aus *praat*: **oben**: Oszillogramm, durchsetzt von Pulsen (senkrechte Linien); **Mitte**: F_0 -Verlauf, bestehend aus vielen *pitch points*; **unten**: Dauermanipulation mit Hilfe von vier *duration points* auf “ab”. Die senkrechten gestrichelten Linien markieren den finalen Nasal. Im F_0 -Verlauf sind die vier *pitch points* der Stilisierung zu erkennen.

6.2.5.4 Informelle Vorversuche und Präparation II

Das Resultat der Manipulationen im letzten Abschnitt (6.2.5.3) wird im Folgenden als *Original-1* bezeichnet¹³. Bei einem ersten Versuch, ausgehend vom *Original-1* ein Stimuluskontinuum zu erzeugen, zeigte sich, dass Konturen mit einer deutlich fallend-steigenden Kontur sich auf der sehr kurzen Wiederanstiegsdauer von 66 ms (vgl. Tabelle 6.2) nicht derart herstellen ließen, dass das Perzept ausreichend natürlich wirkte. Es wurde daher eine zusätzliche Manipulation des *Original-1* durchgeführt, indem der dritte *pitch point* der stilisierten Kontur auf dem Nasal entfernt wurde. Dabei wurde der zweite auf den F_0 -Wert des dritten korrigiert (172 Hz → 170 Hz), damit das ursprüngliche Wiederanstiegsintervall exakt erhalten blieb. Abbildung 6.3 zeigt das Resultat (im Folgenden *Original-1pl*) im Vergleich zum *Original-1*. Mit dieser Operation wurde also das tiefe Plateau entfernt und die Wiederanstiegsdauer um die Dauer des tiefen Plateaus vergrößert (neue Wiederanstiegsdauer = 66 ms + 54 ms = 120 ms; vgl. Tabellen 6.2 und 6.3). Der Unterschied zwischen *Original-1* und *Original-1pl* ist möglicherweise perzipierbar (s. 6.2.5.5); *Original-1pl* ist ebenfalls auf der CDROM enthalten (Anhang C).

¹³Es handelt sich schließlich um eine Manipulation und Stilisierung des Originals, die sich im Bereich der phrasenfinalen Kontur perzeptiv nicht vom Original unterscheidet.

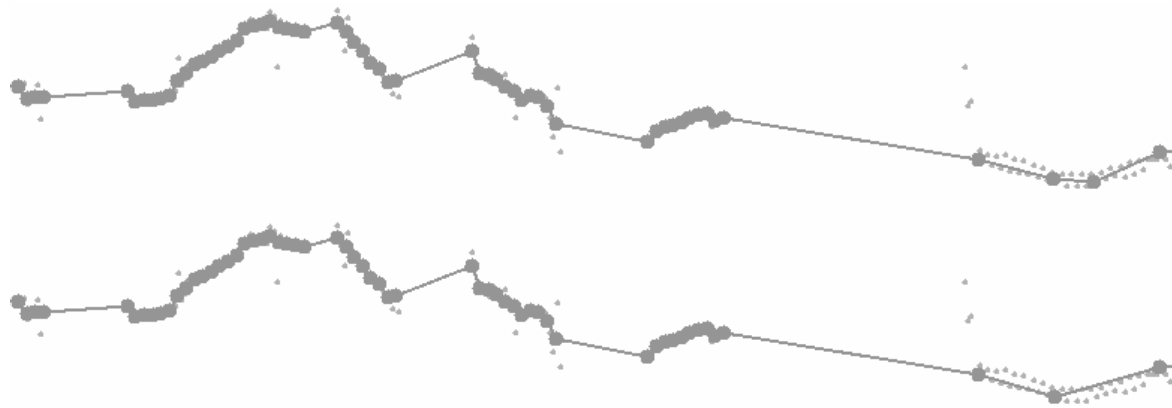


Abbildung 6.3: F_0 -Verläufe von Original-1 (*oben*) und der Stilisierung ohne tiefes Plateau (*unten*).

<i>Pitch point:</i>	(1)	(2)	(3)
F_0 [Hz]	191	170	197
Zeitpunkt [ms]	20	114	234

<i>Parameter:</i>	DuLS	0–PP1	PP1–PP2	DuRi	DuE
[ms]	256	20	94	120	22

Tabelle 6.3: Neuer Ausgangsstimulus des Kontinuums: stilisierte Kontur ohne tiefes Plateau. **Oben:** F_0 -Werte und Zeitpunkte (ausgehend von 0 ms = Beginn der Periodizität des Nasals) der **drei pitch points**. **Unten:** Zugehörige Dauerwerte: DuLS = Dauer des Nasals; 0–PP1 = Nasal onset bis *pitch point* (= PP) 1; PP1–PP2 = PP 1 bis PP 2; DuRi = D. des Wiederanstiegs bzw. PP 2 bis PP 3; DuE = D. des nicht vermessbaren Abschnitts bzw. PP 3 bis Nasal offset.

Durch die vergrößerte Anstiegsdauer klingt auch ein hoher Wiederanstieg sehr natürlich. Allerdings muss hier ein recht großer Wiederanstieg (ca. 10 st) gewählt werden, damit das Resultat eine deutlich perzipierbare fallend-steigende Kontur ist. Ein Grund dafür könnte die relativ geringe Intensität des Materials sein, auf dem der Wiederanstieg stattfindet (Nasal).

6.2.5.5 Das Stimuluskontinuum und Kontrollstimuli

Ausgehend von der Stilisierung ohne tiefes Plateau (Original-1pl) wurde das Stimuluskontinuum nach der generellen Idee zum Aufbau des Kontinuums (6.2.5.1) erzeugt. Als Differenz zwischen den F_0 fin-Werten für zwei aufeinanderfolgende Stimuli im Kontinuum wurde 1 st gewählt. Zwar könnten laut den Ergebnissen der Korpusanalyse (Tabelle A.4, Anhang A) auch kleinere Unterschiede relevant sein; es wurde allerdings als Konsequenz des letzten Abschnitts (6.2.5.4) eine relativ hohe obere Begrenzung für das Kontinuum gewählt (Wiederanstieg von 10 st über dem Wert von Original-1pl), so dass eine kleinere Schrittgröße als 1 st zu einem sehr umfangreichen Stimuluskontinuum geführt hätte. Als Untergrenze wurde ein möglichst tiefes Absinken gewählt, welches aber für eine weibliche Stimme gerade noch natürlich erscheint. Hier ergab sich ein F_0 fin von 4 st unter dem Wert von Original-1pl. Es

resultierte also ein Stimuluskontinuum aus 15 Stimuli, welches in Abbildung 6.4 schematisch dargestellt ist. Tabelle 6.4 enthält dazu den Wert des letzten *pitch points* in Hz für jeden Stimulus und die resultierenden Anstiegsdifferenzen, ausgehend vom Angelpunkt des Kontinuums. Auf die Stimuli wird im Folgenden durch “S01”, “S02” (für Stimulus 01, 02) usw. verwiesen. Für S01 und S02 ergeben sich also objektiv fallende Konturen, für S03, S04 und S05 fallend-leicht-steigende Konturen, wobei der Stimulus S05 identisch ist mit Original-1pl und somit den authentischen Fall der als pseudoterminal etikettierten Kontur repräsentiert.

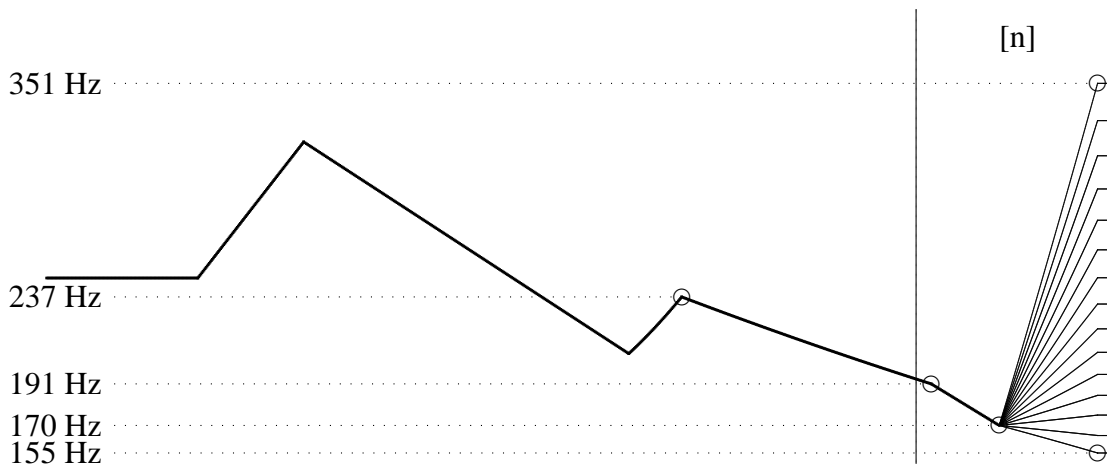


Abbildung 6.4: Stimuluskontinuum, schematische Darstellung: Gezeigt ist der F_0 -Verlauf über der Äußerung “ab Montag, dem achten” mit der systematischen Variation über dem finalen [n]. Zur Orientierung sind besondere F_0 -Werte markiert: Gipfelhöhe des finalen Akzents auf “achten” (237 Hz); *pitch point* 1, 20 ms nach Onset des [n] (191 Hz); *pitch point* 2, Angelpunkt des Kontinuums (170 Hz); die extremen Werte für *pitch point* 3 (untere Grenze des Kontinuums = 155 Hz, obere Grenze = 351 Hz).

Stimulus:	01	02	03	04	05	06	07	08	...
Diff.3-2 [st]:	-1,55	-0,55	0,55	1,55	2,55	3,55	4,55	5,55	...
F_0 fin [Hz]:	155,4	164,7	175,5	185,9	197,0	208,7	221,1	234,3	...
<i>Fortsetzung</i>	...	09	10	11	12	13	14	15	
<i>Fortsetzung</i>	...	6,55	7,55	8,55	9,55	10,55	11,55	12,55	
<i>Fortsetzung</i>	...	248,2	262,9	278,6	295,1	312,7	331,3	351,0	

Tabelle 6.4: Stimuluskontinuum: Diff.3-2 = Intervall zwischen *pitch point* 3 und 2, bzw. Wiederanstiegshöhe in st, ausgehend von 170 Hz; F_0 fin = F_0 -Wert für *pitch point* 3. Fett markiert ist der Ausgangsstimulus **Original-1pl**.

Aufgrund der Präparation in Abschnitt 6.2.5.4 basiert die Stimulusreihe nicht auf dem eigentlichen, authentischen Sprachmaterial, da eine, wenn auch möglicherweise subtile, Eigenschaft des Konturverlaufs verändert wurde. Dieser Unterschied zwischen dem Original-1 und Original-1pl erscheint gerade noch perzipierbar, wobei zu vermuten wäre, dass ein derart geringfügiger Unterschied nicht kommunikativ relevant sein kann. Demnach sollte das Experiment durch diese Präparation nicht beeinträchtigt werden. Um diese Vermutung aber abzusi-

chern, wurde der Menge der Stimuli auch das Original-1 als Kontrollstimulus (Bezeichnung: Sko1) hinzugefügt. Die Mittelwerte für die beiden Vergleichsstimuli S05 und Sko1 können durch t-Tests für gepaarte Stichproben für alle neun Skalen miteinander verglichen werden. Treten relevante signifikante Unterschiede in der Beurteilung dieser beiden Stimuli auf, so würde dies bedeuten, dass die Präparation in 6.2.5.4 tatsächlich einen zu großen Eingriff in die Authentizität des Sprachmaterials dargestellt hat.

Das Kontinuum enthält keinen Stimulus, der am Phrasenende ein glatt auslaufendes F_0 -Minimum aufweist. Stattdessen gibt es zwei Stimuli (S02, S03), die einen sehr geringen Weiterabfall, bzw. Wiederanstieg aufweisen. Eine interessante Zusatzfragestellung besteht darin, ob eine dieser beiden Konturen, bzw. beide, funktional eher mit einem glatten Auslaufen gleichgesetzt werden kann. Sollte sich z.B. herausstellen, dass die hypothetische Funktion einer pseudoterminalen Kontur auch mit einem glatten tiefen Auslaufen von F_0 assoziiert ist (vs. einem noch tieferen Fallen für eine terminale Kontur), so würde dies bedeuten, dass der leichte Wiederanstieg für eine pseudoterminalen Kontur nicht essentiell ist. Aufgrund dieser Überlegungen wurde ein weiterer Kontrollstimulus (Bezeichnung: Sk00) resynthetisiert, der dieselben Eigenschaften hat wie die Stimuli des Kontinuums, wobei F_0 fin auf 170 Hz gesetzt ist, was ein glattes Auslaufen von F_0 auf 170 Hz zur Folge hat. Auch hier können die Mittelwerte von S02 vs. Sk00 einerseits, und Sk00 vs. S03 andererseits durch t-Tests für gepaarte Stichproben für alle neun Skalen miteinander verglichen werden.

6.2.6 Technische Konzeption, Fragebogen und Durchführung des Experiments

Aus organisatorischen Gründen wurde das vorliegende Experiment mit Hilfe eines Fragebogens durchgeführt¹⁴. Auf diesen wird weiter unten im Abschnitt näher eingegangen, nachdem der generelle Aufbau des Experiments erläutert worden ist.

Mit Hilfe der Programmiersprache *awk* (z.B. Herold 1999) wurde aus den einzelnen Stimuli eine *wav*-Datei erzeugt, die den gesamten Perzeptionstest beinhaltet. Hierzu standen im IPDS *awk*-Skripte¹⁵ zur Verfügung, von denen eines an die vorliegende Versuchsplannung angepasst wurde. Diese Skripte ermöglichen ein automatisches Randomisieren, ggf. Vervielfältigen und schließlich Zusammenschneiden der Stimuli, sowie Einfügen von Pausen und Signaltönen an vom Anwender definierten Stellen.

Da jeder Stimulus genau einmal auf jeder Skala beurteilt werden sollte (vgl. Abschnitt 6.2.2), musste jeder Stimulus neun Mal in der *wav*-Datei vorkommen. Es musste dafür gesorgt werden, dass die randomisierten Stimuli (mit neunfacher Wiederholung) für die Konzeption des Fragebogens auf möglichst einfache Weise den Skalen (möglichst ebenso in einer

¹⁴Auf diese Weise konnten mehr Probanden zur gleichen Zeit an dem Experiment teilnehmen, als bei der Stimuluspräsentation über ein Computerprogramm möglich gewesen wäre.

¹⁵Geschrieben von M.T.M. Scheffers.

randomisierten Reihenfolge) zugeordnet werden konnten. Hierzu wurde bereits vor der Anwendung der *awk*-Skripte jeder Stimulus neun Mal manuell vervielfältigt und in den resultierenden 153 *wav*-Dateien die Zuordnung von Stimulus und Skala im Dateinamen verankert (z.B. *S11x5.wav* → Stimulus 11, kombiniert mit Skala 5). In der technischen Umsetzung handelt es sich also um 153 verschiedene Stimuli, die mit Hilfe der *awk*-Skripte randomisiert und zu einem Perzeptionstest verarbeitet wurden. Die resultierende *wav*-Datei ist wie folgt strukturiert.

Jeder Stimulus wird für die Beurteilung auf einer Skala zweimal direkt hintereinander (mit einer Pause von 4 s) präsentiert. Ein jeder Stimulus-Block dieser Art wird von einem Signalton eingeleitet (F_0 : 500 Hz; Dauer: 500 ms); zwischen diesem Signalton und der ersten Präsentation des Stimulus wurde eine Pause von 750 ms eingefügt. Die Dauer eines Stimulus beträgt 1502 ms, so dass die Präsentation eines Stimulus (mit Wiederholung) für die Bearbeitung einer Skala insgesamt $500 \text{ ms} + 750 \text{ ms} + 1502 \text{ ms} + 4000 \text{ ms} + 1502 \text{ ms} = 8,254 \text{ s}$ dauerte. Ein Block von der eben beschriebenen Struktur wird im Folgenden als Stimuluspräsentation bezeichnet. Das gesamte Experiment enthält also 153 Stimuluspräsentationen (wobei jede einzelne eine Wiederholung ein und desselben Stimulus enthält).

Zwischen zwei Stimuluspräsentationen wurde jeweils eine Pause von 4 s eingefügt. Zur Orientierung für die Versuchspersonen wurde nach jeder neunten Stimuluspräsentation zusätzlich ein doppelter Signalton (jeweils F_0 : 500 Hz, Dauer: 200 ms; dazwischen 200 ms Pause) und eine Pause von 2 s eingefügt. Ein solcher Block aus 9 Stimuluspräsentationen entsprach stets einer Seite auf dem Fragebogen.

Der gesamte Perzeptionstest (17 Blöcke) wurde nicht in Form einer einzigen *wav*-Datei belassen, sondern in zwei Teile geschnitten (Teil 1: 9 Blöcke, Teil 2: 8 Blöcke), um den Probanden eine kurze Pause während des Experiments zu ermöglichen. Der erste Teil des Experiments dauerte somit ca. 17 Minuten, der zweite Teil ca. 15 Minuten.

Der Fragebogen umfasste folglich 17 Seiten mit je 9 Skalen. Dabei konnte, aufgrund der globalen Randomisierung der Paare aus Stimulus und Skala, dieselbe Skala mehrfach hintereinander auf dem Fragebogen erscheinen¹⁶. Im Anhang findet sich als Beispiel die Kopie einer dieser 17 Seiten (Abbildung B.2). Zusätzlich enthielt der Fragebogen eine Einleitungsseite (Anhang, Abbildung B.1) und eine Schlussseite (Anhang, Abbildung B.3). Die Einleitungsseite enthielt Instruktionen für die Versuchsperson, Hinweise zum Ablauf des Experiments und drei Beispiel-Skalen. Zu diesen drei Beispiel-Skalen wurde ein Beispiel-Experiment erzeugt, welches genauso gestaltet war wie das eigentliche Experiment — mit der Einschränkung, dass es nur einen Block mit nur drei Stimuluspräsentationen enthielt. Die verwendeten Stimuli für diesen Testlauf wurden nicht zufällig gewählt, sondern stellten drei Landmarken im Kontinuum dar: Es handelte sich um die Stimuli mit dem tiefsten Fall (S01), mit der (stilisierten) originalen [2;]-Kontur (S05) und mit dem höchsten Wiederanstieg

¹⁶Bei der verwendeten Randomisierung erscheint ein und dieselbe Skala höchstens dreimal hintereinander.

(S15). Diese Trainingsphase sollte die Probanden erstens an die Aufgabe sowie die Stimuli gewöhnen und zweitens testen, ob noch Unklarheiten zum Ablauf bestehen.

Die letzte Seite des Fragebogens enthielt – wiederum in Form von siebenstufigen Skalen – zwei Kontrollfragen, deren Auswertung bei der Interpretation der Daten helfen sollte (vgl. Anhang, Abbildung B.3):

A Wie schwierig war die Aufgabe für Dich?

B Wie natürlich klangen die Stimuli?

Im Falle einer Antwort in Richtung ‘unnatürlich’ wurden die Probanden durch eine Zusatzfrage gebeten, kurz zu notieren, was genau ihnen unnatürlich erschien. Außerdem wurden auf der letzten Seite Daten zur Person erhoben (vgl. Anhang, Abbildung B.3), die v.a. zur Beschreibung der Versuchspersonengruppe benötigt wurden.

Das Experiment wurde 11 mal mit unterschiedlich großen Gruppen von Versuchspersonen (zwischen 1 und 9 Personen) durchgeführt. Alle Durchläufe fanden in einem akustisch behandelten Raum (Studio des IPDS) statt, so dass einwandfreie akustische Bedingungen gegeben waren. Aus den drei *wav*-Dateien für das Experiment (*Testlauf.wav*, *Teil1.wav*, *Teil2.wav*) wurde zuvor eine Audio-CD erzeugt, mit der die Stimuli den Probanden über einen CD-Spieler präsentiert wurden. Die drei Dateien sowie alle Stimuli als einzelne Dateien befinden sich auf der beiliegenden CDROM (Anhang C). Es wurde bei allen Versuchspersonen dieselbe Lautstärke verwendet.

6.2.7 Versuchspersonen

Es wurden insgesamt 49 Versuchspersonen herangezogen (24 weibliche und 25 männliche). Sechs Versuchspersonen (zwei weibliche, vier männliche) wurden aus technischen Gründen in der Datenauswertung nicht berücksichtigt. Damit reduziert sich die Zahl der Versuchspersonen auf 43 (22 weibliche, 21 männliche). Im Folgenden wird stets nur diese reduzierte Probandengruppe berücksichtigt (wobei auch Formulierungen wie “alle Hörer ...” sich auf diese 43 Probanden beziehen).

Sämtliche Probanden waren Studierende der Universität Kiel, mit einer Ausnahme (Universität Hamburg). Aus organisatorischen Gründen machten Studierende der Psychologie den größten Teil der Versuchspersonen aus (72,1 %)¹⁷. Dazu gehören alle 22 weiblichen Versuchspersonen. Unter den 21 männlichen Probanden sind weitere neun Studierende der Psychologie, fünf Studierende linguistischer oder philologischer Fächer – allerdings nicht mit dem Hauptfach Phonetik und ohne besondere Kenntnisse im Bereich der Intonationsforschung – und sieben aus verschiedenen anderen Fachrichtungen.

¹⁷Für Studierende der Psychologie an der Universität Kiel ist die Teilnahme an einer gewissen Anzahl von Experimenten dieser Art obligatorisch.

Das Alter der Versuchspersonen variierte zwischen 19 und 45 Jahren, wobei der größte Teil (30 Personen; 69,8 %) zwischen 20 und 25 Jahre alt waren. Die meisten Versuchspersonen (27 Personen; 62,8 %) hatten bereits vor der Teilnahme an diesem Experiment etwas, aber wenig Erfahrung als Versuchsperson in Experimenten zur Sprachwahrnehmung (bis zu fünfmal teilgenommen). Nur eine Person gab an, bereits öfter als fünfmal an einem solchen Experiment teilgenommen zu haben, und 15 Personen (34,9 %) hatten noch gar keine Erfahrung mit Experimenten zur Sprachwahrnehmung. Alle Probanden können als naiv hinsichtlich der untersuchten Fragestellung betrachtet werden. Die (geringe) Erfahrung als Versuchsperson beim größten Teil der Gruppe ist daher vermutlich kein störender Faktor.

Alle Probanden sind Muttersprachler des Deutschen. Es wurden keine Daten über die dialektale Zugehörigkeit erhoben.

6.3 Ergebnisse

In den folgenden Abschnitten werden die Ergebnisse größtenteils anhand von grafischen Darstellungen der Mittelwerte präsentiert. Eine numerische Präsentation der Mittelwerte sowie der Standardabweichungen bietet Tabelle B.1 (Anhang B.3). Auf die Rohdaten wird nicht Bezug genommen. Sie sind daher auch im Anhang nicht abgedruckt, können aber auf der CDROM in einem *pdf*-Dokument eingesehen werden.

Die folgenden beiden Abschnitte behandeln zwei Voraussetzungen, die für die Interpretation der wesentlichen Ergebnisse gegeben sein müssen: die Natürlichkeit der verwendeten Stimuli (6.3.1) und der Einfluss der methodisch bedingten Manipulation der Konturgestalt auf die Ergebnisse (6.3.2). Danach widmet sich Abschnitt 6.3.3 der Prüfung der Hypothese 3 (vgl. 4.2). In Abschnitt 6.3.4 wird die zusätzliche Fragestellung bezüglich der glatt auslaufenden Kontur (Sk00) behandelt. Abschließend widmet sich Abschnitt 6.3.5 der Prüfung der Hypothese 4 (vgl. 4.2).

6.3.1 Postexperimentelle Befragung

Die Auswertung der letzten Seite des Fragebogens wird den Ergebnissen vorangestellt, da die Natürlichkeit der Stimuli eine gewisse Grundvoraussetzung für die Interpretation der Ergebnisse darstellt. Ebenso macht es Sinn, sich der Aufgabe der Probanden stets bewusst zu sein, so dass die empfundene Schwierigkeit eine interessante Zusatzinformation darstellt. Der zweite Teil dieser Befragung (persönliche Daten, s. Fragebogen, Abbildung B.3) soll hier nicht behandelt werden, da die dort erhobenen Daten bereits in die Beschreibung der Versuchspersonengruppe eingegangen sind (vgl. 6.2.7).

Die Balkendiagramme in Abbildung 6.5 zeigen die Verteilungen der Antworten auf die beiden Fragen A und B (vgl. 6.2.6). Für beide Fragen streuen die Antworten über die gesamte Skala, wobei die extremen Werte (+/-3) bei beiden Fragen nur von jeweils zwei Personen

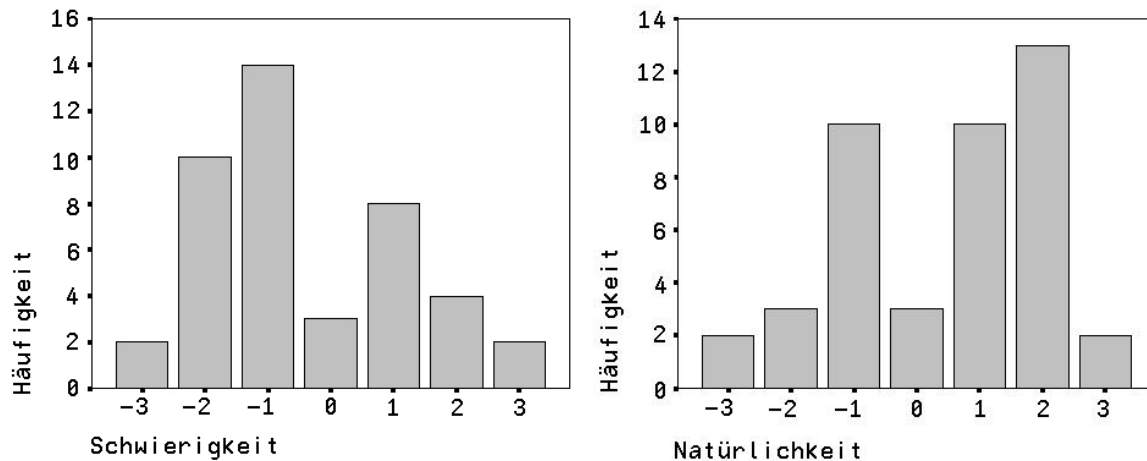


Abbildung 6.5: Verteilung der Antworten aller 43 Probanden auf die Fragen A ('Schwierigkeit'; linke Seite) und B ('Natürlichkeit'; rechte Seite) (vgl. Abschnitt 6.2.6) über die sieben Antwortmöglichkeiten (horizontale Achsen). Polung: negativ = 'unnatürlich', bzw. 'schwierig'; positiv = 'natürlich', bzw. 'einfach'.

gewählt wurden. Der größte Teil der Probanden (39 Personen; 90,1%) beurteilte die Stimuli weder als 'sehr natürlich' noch als 'sehr unnatürlich'. Analoges trifft auf die zweite Frage zu. Bei jeder Frage waren nur 3 Personen unentschlossen. Die Verteilung für Frage B zeigt, dass mehr Versuchspersonen die Stimuli für 'natürlich' halten als für 'unnatürlich'. In Anhang B.2 sind die freien Kommentare der Probanden aufgelistet, die als Antwort auf die Frage "Was klang unnatürlich?" notiert wurden. Mit einer Ausnahme lassen die Kommentare nicht schließen, dass die Stimuli wie synthetische Sprache klangen oder unnatürlich klingende Tonhöhenbewegungen enthielten. Die Tauglichkeit der Stimuli kann daher sehr hoch eingeschätzt werden. Die Kommentare scheinen hingegen eher auf die Schwierigkeit der Aufgabe abzielen. Die Verteilung für Frage A zeigt, dass in der Tat die meisten Versuchspersonen die Aufgabe als schwierig empfunden haben.

6.3.2 Kontrollstimulus Sko1

Als eine weitere Voraussetzung für die Interpretation der Ergebnisse muss geprüft werden, inwiefern die Manipulation des Sprachmaterials (vgl. 6.2.5.4) einen Einfluss auf die Ergebnisse haben könnte. Zu diesem Zweck wurde der Ausgangsstimulus des Kontinuums (S05) mit dem Kontrollstimulus Sko1 verglichen. Letzterer enthält im Gegensatz zu S05 das im Original aufgetretene tiefe Plateau. Der Vergleich wurde mittels zweiseitiger t-Tests über die Mittelwerte für jede Skala durchgeführt. Tabelle 6.5 zeigt die Mittelwerte und Standardabweichungen der Urteile für beide Stimuli auf jeder der neun Skalen sowie die Ergebnisse der t-Tests. Die Mittelwerte sind auch in Abbildung 6.6 als Balkendiagramm dargestellt.

Die Unterschiede in den Beurteilungen von S05 vs. Sko1 sind für sieben der neun Skalen minimal und nicht signifikant. Für zwei Skalen treten signifikante Unterschiede auf: Skala 3

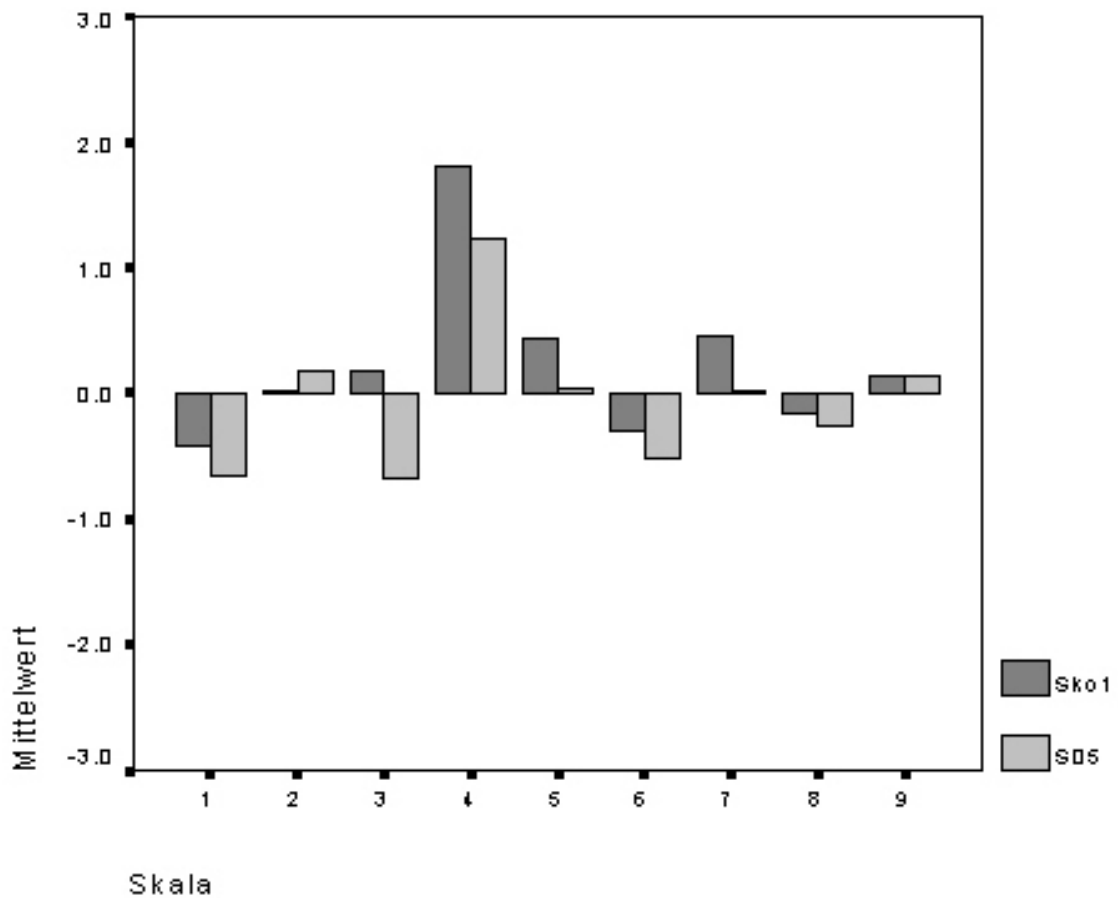


Abbildung 6.6: Mittelwerte der Urteile für Sko1 und S05. Zur Bezeichnung der Skalen s. Tabelle 6.1, Abschnitt 6.2.3

Skala	S05	Sko1	t	df	p	Sig.
1	-0,65	-0,42	0,721	42	0,475	
2	0,19	0,02	-0,507	42	0,614	
3	-0,67	0,19	2,863	42	0,007	*
4	1,23	1,81	2,943	42	0,005	*
5	0,05	0,44	1,429	42	0,160	
6	-0,51	-0,30	0,988	42	0,329	
7	0,02	0,47	1,951	42	0,058	
8	-0,26	-0,16	0,330	42	0,743	
9	0,14	0,14	0,000	42	1,000	

Tabelle 6.5: Einzelvergleiche der Urteile für Sko1 vs. S05: Ergebnisse der zweiseitigen t-Tests: df = Anzahl der Freiheitsgrade, p = Irrtumswahrscheinlichkeit, Sig. = Signifikanz.

(‘emotional/ sachlich’) und Skala 4 (‘Frage/ Aussage’). Trotz der Signifikanzen sind die Unterschiede nicht groß: Sowohl S05 als auch Sko1 werden deutlich als Aussagen beurteilt. Beide Mittelwerte liegen zwischen +1 und +2, d.h. ihre Differenz ist kleiner als 1 Skalenpunkt. Damit kann der Unterschied – nach den Ergebnissen der Diskussion in Abschnitt 6.2.2 – als nicht relevant eingestuft werden. Analoges gilt für Skala 3. Hier ist allerdings zu beachten, dass – in der Tendenz – S05 eher als emotional eingestuft wurde, während Sko1 eher neutral ist (weder emotional noch sachlich). Insgesamt zeigen diese Ergebnisse, dass die beiden Stimuli S05 und Sko1 relativ gleich beurteilt worden sind. Die Manipulation der Konturgestalt übt damit keinen relevanten Einfluss auf die Ergebnisse aus.

6.3.3 Prüfung der Hypothese 3

Die folgenden vier Abschnitte widmen sich der Prüfung von Hypothese 3 (vgl. 4.2). Die Hypothese 3 lässt sich grob in zwei Teilhypothesen zerlegen:

- Die *pseudoterminale* Kontur lässt sich perzeptiv und funktional von einer *fallend-steigenden* Kontur abgrenzen.
- Die *pseudoterminale* Kontur lässt sich perzeptiv und funktional von einer *fallenden* Kontur abgrenzen.

In Abschnitt 6.3.3.1 wird ein globaler Test mittels einer MANOVA durchgeführt. In den folgenden Abschnitten werden die zwei genannten Teilhypothesen mit Hilfe von ANOVAs geprüft. Dabei werden die ANOVAs nicht über das gesamte Stimuluskontinuum gerechnet, sondern getrennt über den “unteren” oder “oberen” Bereich des Kontinuums. Denn je nach getesteter Teilhypothese ist zu erwarten, dass signifikante Unterschiede nur oder insbesondere in *einem* der beiden Bereiche auftreten: Zur Abgrenzung von [2;] vs. [2.] sind z.B. auf der Skala ‘kategorisch/ verhandlungsbereit’ große Unterschiede im unteren Bereich zu erwarten, aber keine Unterschiede im oberen. Der obere und der untere Bereich des Kontinuums sind wie folgt definiert: Laut der Korpusanalyse beträgt der Wiederanstieg von [2;] im Mittel 1,66 st. Das verwendete Sprachmaterial weist einen deutlich höheren Wiederanstieg von 2,55 st auf. Wenn sich ein Übergang zwischen [2.] und [2;] zeigt, dann müsste er also *spätestens* bei einem Wiederanstieg von etwa 2 bis 3 st auftreten. Daher wurde die Grenze für den “unteren” Bereich des Kontinuums auf S05 (Originalanstieg von 2,55 st) gesetzt. Da dieser Wiederanstieg für eine [2;]-Kontur relativ groß ist, wurde die Untergrenze des “oberen” Bereichs etwas tiefer – auf S04 (1,55 st) – festgesetzt, damit sichergestellt ist, dass noch mindestens ein bis zwei Stimuli des oberen Bereichs typische Vertreter der [2;]-Kontur darstellen. Eine ANOVA über das gesamte Stimuluskontinuum hätte den Nachteil, dass – falls signifikante Mittelwertunterschiede auftreten – unklar bliebe, in welchem Bereich des Kontinuums diese auftreten. Eine Möglichkeit, nachträglich zu ermitteln, an welcher Stelle im Kontinuum

(z.B. “zwischen S03 und S04”) ein Übergang von einem Melodiemuster zum anderen auftritt, bieten *Post-Hoc*-Tests. Von dieser Möglichkeit wird in den folgenden Analysen jedoch kein Gebrauch gemacht. Es soll in dieser Arbeit nicht postuliert werden, dass eine Grenze zwischen zwei verschiedenen phrasenfinalen Melodiemustern sich in einem “scharfen Kategorienwechsel” (wie etwa in der kategorialen Wahrnehmung von /b, d, g/) ausprägen muss (vgl. 2.4.3). Ein Übergang kann sich ebenso gut über mehrere Stimuli hinweg ausdehnen. Paarweise Einzelvergleiche von Stimuli mittels *Post-Hoc*-Tests erscheinen daher an dieser Stelle nicht notwendig. Es genügt die Kenntnis, dass sich “irgendwo” im oberen oder unteren Bereich des Kontinuums Unterschiede ergeben. Durch die hypothesenorientierte Trennung des Kontinuums in zwei Bereiche geht diese Überlegung *a priori* in die Analyse ein.

6.3.3.1 Globale Prüfung von Einflussfaktoren

Aus der Grundhypothese lässt sich eine simple Teilhypothese hinsichtlich dieses Perzeptionsexperiments formulieren:

- Die Wiederanstiegshöhe der phrasenfinalen Kontur hat einen signifikanten Einfluss auf die Beurteilung der Stimuli durch die Probanden.

Mit Hilfe einer MANOVA wurde diese Teilhypothese global getestet. Dabei stellt das Urteil auf einer Skala je eine abhängige Variable dar; es gibt also 9 abhängige Variablen. Der Innersubjektfaktor ist die Wiederanstiegshöhe; er hat also 15 Stufen (in diesem Test werden die beiden Kontrollstimuli nicht berücksichtigt).

Peters (1999) beobachtete die pseudoterminalen Kontur vermehrt bei weiblichen Sprechern; das Geschlecht des *Sprechers* scheint hinsichtlich der Verwendung des Musters einen Einfluss zu haben. Auf diesen Aspekt wird im Rahmen dieser Arbeit nicht explizit eingegangen, da keine Aussagen über die Auftretenshäufigkeit getroffen werden (vgl. 5.1). Als eine zusätzliche Fragestellung kann aber geprüft werden, ob das Geschlecht des *Hörers* im Rahmen der Wahrnehmung und der semantischen Interpretation von phrasenfinalen Konturen einen Einfluss hat. Daher wurde für die MANOVA das Geschlecht der Versuchsperson als ein zweiter Faktor (Zwischensubjektfaktor) hinzugenommen. Selbst wenn die pseudoterminalen Kontur eher von weiblichen als von männlichen Sprechern verwendet werden sollte, erscheint es eher unplausibel, dass Frauen und Männer die Kontur unterschiedlich wahrnehmen oder unterschiedlich interpretieren. Die Hypothese bezüglich des zweiten Faktors lautet daher:

- Das Geschlecht der Versuchsperson hat keinen signifikanten Einfluss auf die Ergebnisse des Perzeptionsexperiments.

Die Ergebnisse der MANOVA sind in Tabelle 6.6 zusammengefasst. Beide Hypothesen werden deutlich bestätigt: Der Effekt der Wiederanstiegshöhe ist hoch signifikant ($p = 0,000$),

<i>Innersubjekteffekt</i>	F	df	p	Sig.
Wiederanstiegshöhe	10,822	126	0,000	*
Interaktion aus Wiederanstiegshöhe und Geschlecht	1,196	126	0,070	

Tabelle 6.6: Globaler Test: Ergebnisse der zweifaktoriellen MANOVA: df = Anzahl der Freiheitsgrade, p = Irrtumswahrscheinlichkeit, Sig. = Signifikanz.

der Effekt der Interaktion aus Wiederanstiegshöhe und Geschlecht der Versuchsperson ist nicht signifikant ($p = 0,070$). Die folgenden Analysen werden daher stets für die gesamte Stichprobe, nicht getrennt nach weiblichen und männlichen Probanden, durchgeführt.

6.3.3.2 Abgrenzung von [2;] vs. [2.]

Die Hypothesen zur Unterscheidung zwischen der terminalen und der pseudoterminalen Kontur lassen sich hinsichtlich der verwendeten Skalen wie folgt reformulieren:

- Es gibt signifikante Mittelwertunterschiede für die Beurteilungen der Stimuli S01–S05 auf den Skalen
 - ‘kategorisch/ verhandlungsbereit’
 - ‘dominant/ nicht dominant’
- Diese Unterschiede sind wie folgt gerichtet: Mit der pseudoterminalen Kontur werden eher die Merkmale ‘verhandlungsbereit’ und ‘nicht dominant’ assoziiert.
- Es gibt möglicherweise auch Unterschiede auf den Skalen:
 - ‘nicht sicher/ sicher’
 - ‘selbstbewusst/ schüchtern’

Abbildung 6.7 zeigt die Mittelwerte der Urteile für jede dieser vier Skalen, geplottet als Verlauf über das gesamte Stimuluskontinuum. Tabelle 6.7 zeigt die Ergebnisse der ANOVAs für den unteren (S01–S05) und den oberen (S04–S15) Bereich des Kontinuums für jede der vier Skalen.

Für die Skalen ‘kategorisch/ verhandlungsbereit’ und ‘dominant/ nicht dominant’ (Abbildung 6.7, oben) zeigt sich ein sehr paralleler Verlauf, der sich grob in zwei Abschnitte einteilen lässt, die sich bei etwa S03–S05 überlappen: (1.) Im unteren Bereich (S01–S05) ergibt sich eine große Transition (um ca. 2 Skaleneinheiten). Eine fallende Kontur (S01, S02) wird deutlich als ‘dominant’ und ‘kategorisch’ beurteilt. Es gibt einen graduellen Übergang von einem deutlichen Urteil ‘kategorisch’/ ‘dominant’ (S01) zu einem ‘neutralen’ Urteil (etwa S04, S05). (2.) Die Stimuli mit Wiederanstieg (S03–S15) werden auf beiden Skalen zwischen -1

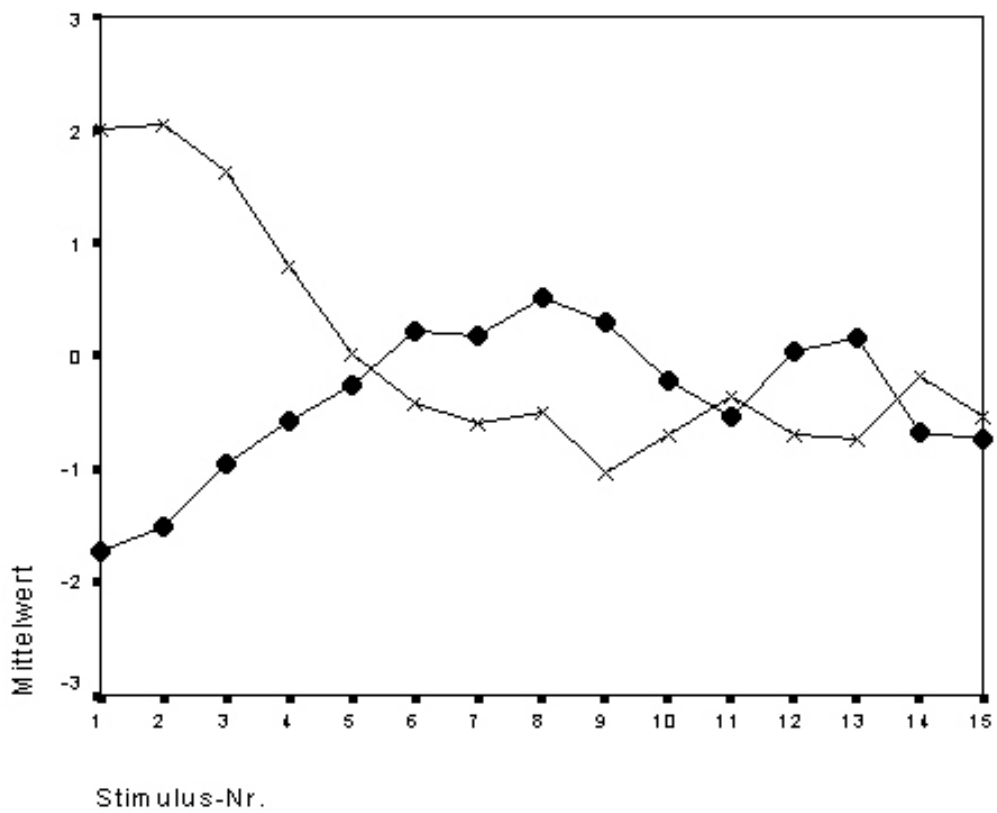
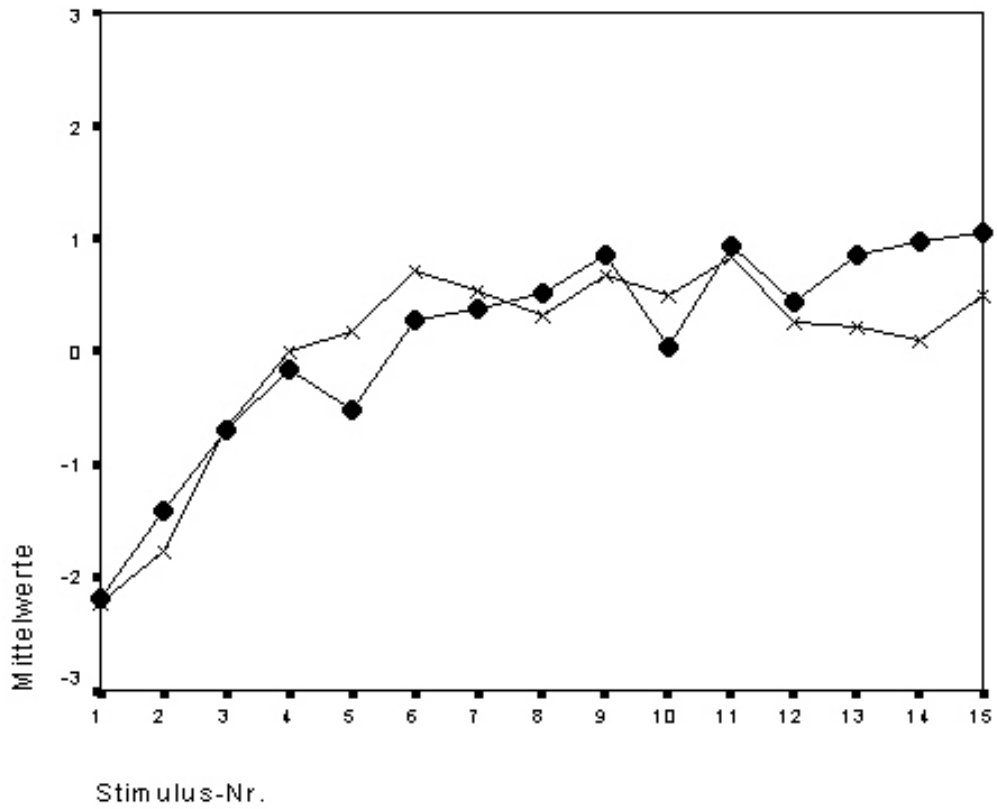


Abbildung 6.7: Mittelwerte der Beurteilungen von Stimuli S01–S15 für die Skalen: **oben:** 'kategorisch/verhandlungsbereit' (●) und 'dominant/nicht dominant' (×); **unten:** 'nicht sicher/sicher' (×) und 'selbstbewusst/schüchtern' (●).

Skala	S01–S05				S04–S15			
	df	F	p	Sig	df	F	p	Sig
‘kategorisch/ verhandlungsbereit’	3,467	17,446	0,000	*	6,575	5,401	0,000	*
‘dominant/ nicht dominant’	3,159	35,388	0,000	*	6,271	1,465	0,187	
‘nicht sicher/ sicher’	2,996	20,340	0,000	*	6,191	4,602	0,000	*
‘selbstbewusst/ schüchtern’	3,447	10,698	0,000	*	5,335	3,720	0,002	*

Tabelle 6.7: ANOVAs für vier Skalen und zwei Bereiche des Kontinuums (unterer: S01–S05; oberer: S04–S15). df = Anzahl der Freiheitsgrade, p = Irrtumswahrscheinlichkeit, Sig. = Signifikanz.

und 1, tendenziell also als ‘neutral’ eingestuft, wobei ein leichter Wiederanstieg (ca. S03–S05) zur ‘–’ Kategorie tendiert, ein größerer Wiederanstieg (ca. ab S05) zur ‘+’ Kategorie.

Im Abschnitt (1) ist also ein sehr viel deutlicherer Wechsel in der Beurteilung der Stimuli zu erkennen als in Abschnitt (2). Dies wird durch das Ergebnis der ANOVA für ‘dominant/ nicht dominant’ bestätigt: Für S01–S05 ergeben sich signifikante Unterschiede, nicht aber für S04–S15 (vgl. Tabelle 6.7). Für ‘kategorisch/ verhandlungsbereit’ ist das Ergebnis weniger klar, da für *beide* Abschnitte des Kontinuums signifikante Unterschiede auftreten. Die Ergebnisse der Prüfstatistik müssen stets im Zusammenhang mit der deskriptiven Statistik interpretiert werden. Die (deskriptiven) Ergebnisse für die beiden betrachteten Skalen sind parallel (s.o.), dennoch unterscheiden sich die prüfstatistischen Ergebnisse. Dies mag darin begründet sein, dass die Schwankung der Urteile für S04–S15 um den Wert 0 sich im Falle von ‘kategorisch/ verhandlungsbereit’ durch größere Auslenkungen in Richtung auf die Werte +1 und -1 auszeichnet, als es bei ‘dominant/ nicht dominant’ der Fall ist. Der gefundene signifikante Unterschied bei ersterer Skala für S04–S15 ist daher weniger von Bedeutung als die Unterschiede für S01–S05. Eine Betrachtung der zugehörigen F-Werte bestätigt diese Interpretation (vgl. Tabelle 6.7): Ein kleinerer F-Wert kann als Korrelat eines kleineren Unterschieds aufgefasst werden.

Für die beiden Skalen ‘nicht sicher/ sicher’ und ‘selbstbewusst/ schüchtern’ ergeben sich ebenfalls recht parallele Ergebnisse (die in der Darstellung in Abbildung 6.7, unten, aufgrund der unterschiedlichen Polung der Skalen allerdings eher als inverse Ergebnisse erscheinen). Analog zu den ersten beiden Skalen lässt sich auch hier der Verlauf der Urteile über das Kontinuum in zwei Abschnitte einteilen: Der erste reicht wiederum von S01 bis S05, der zweite beginnt etwa bei S04–S05. In Abschnitt (1) ergeben sich große Unterschiede (von etwa 1,5 bis 2 Skalenpunkten): Die fallenden Konturen (S01, S02) werden deutlich als ‘sicher’ und ‘selbstbewusst’ beurteilt. Zwischen S03 und S05 werden die Urteile kontinuierlich neutraler. Eine Ausnahme bildet hier die Kontur mit dem kleinsten Wiederanstieg (S03) auf der Skala ‘nicht sicher/ sicher’: S03 wird wie S01 und S02 noch recht deutlich als ‘sicher’ beurteilt. Abschnitt (2) zeichnet sich für ‘selbstbewusst/ schüchtern’ wieder durch ein Schwanken der Urteile um den neutralen Wert 0 aus, während sich die Urteile für ‘nicht sicher/ sicher’ in der Tendenz zwischen -1 und 0 bewegen. Fallend-steigende Konturen werden also tendenziell

als ‘nicht sicher’ beurteilt, während sich auf der Skala ‘selbstbewusst/ schüchtern’ keine derartige Tendenz ausprägt. Die ANOVAs für beide Skalen und beide Stimulusbereiche zeigen signifikante Ergebnisse (vgl. Tabelle 6.7). Wie auch im Falle der ersten beiden Skalen zeigen sich aber die wesentlichen Unterschiede in der Beurteilung nur im Bereich S01–S05. Zur Interpretation der Prüfstatistik kann daher eine analoge Argumentation herangezogen werden. Auch wird diese Argumentation wieder durch kleine F-Werte für den Bereich S04–S15 im Vergleich zu S01–S05 bestätigt.

Bezüglich der Abgrenzung eines pseudoterminalen Melodiemusters von einem terminalen konnten die Hypothesen also bestätigt werden.

6.3.3.3 Abgrenzung von [2;] vs. [2.,]

Zur Unterscheidung zwischen der pseudoterminalen und der nicht-terminalen fallend-steigenden Kontur können die Hypothesen bezüglich der verwendeten Skalen wie folgt formuliert werden:

- Es gibt signifikante Mittelwertunterschiede in den Beurteilungen der Stimuli S04–S15 auf den Skalen
 - ‘abgeschlossen/ ungeschlossen’
 - ‘Frage/ Aussage’
- Diese Unterschiede sind wie folgt gerichtet: Mit der pseudoterminalen Kontur werden eher die Merkmale ‘abgeschlossen’ und ‘Aussage’ assoziiert.

Abbildung 6.8 zeigt die Mittelwerte der Urteile für jede dieser beiden Skalen, geplottet als Verlauf über das gesamte Stimuluskontinuum. Tabelle 6.8 zeigt die Ergebnisse der ANOVAs für den unteren (S01–S05) und den oberen (S04–S15) Bereich des Kontinuums für beide Skalen.

<i>Skala</i>	<i>S01–S05</i>				<i>S04–S15</i>			
	df	F	p	Sig	df	F	p	Sig
‘abgeschlossen/ ungeschlossen’	3,358	16,120	0,000	*	7,544	29,074	0,000	*
‘Frage/ Aussage’	3,222	10,642	0,000	*	6,458	20,242	0,000	*

Tabelle 6.8: ANOVAs für zwei Skalen und zwei Bereiche des Kontinuums (unterer: S01–S05; oberer: S04–S15). df = Anzahl der Freiheitsgrade, p = Irrtumswahrscheinlichkeit, Sig. = Signifikanz.

Bei einer groben Betrachtung lässt sich auch im Falle der beiden Skalen ‘abgeschlossen/ ungeschlossen’ und ‘Frage/ Aussage’ der Verlauf der Beurteilungen über das Stimuluskontinuum in zwei Abschnitte einteilen. Der erste Abschnitt erstreckt sich über S01–S06 und beinhaltet damit sowohl die fallenden Konturen (S01, S02) als auch diejenigen mit einem

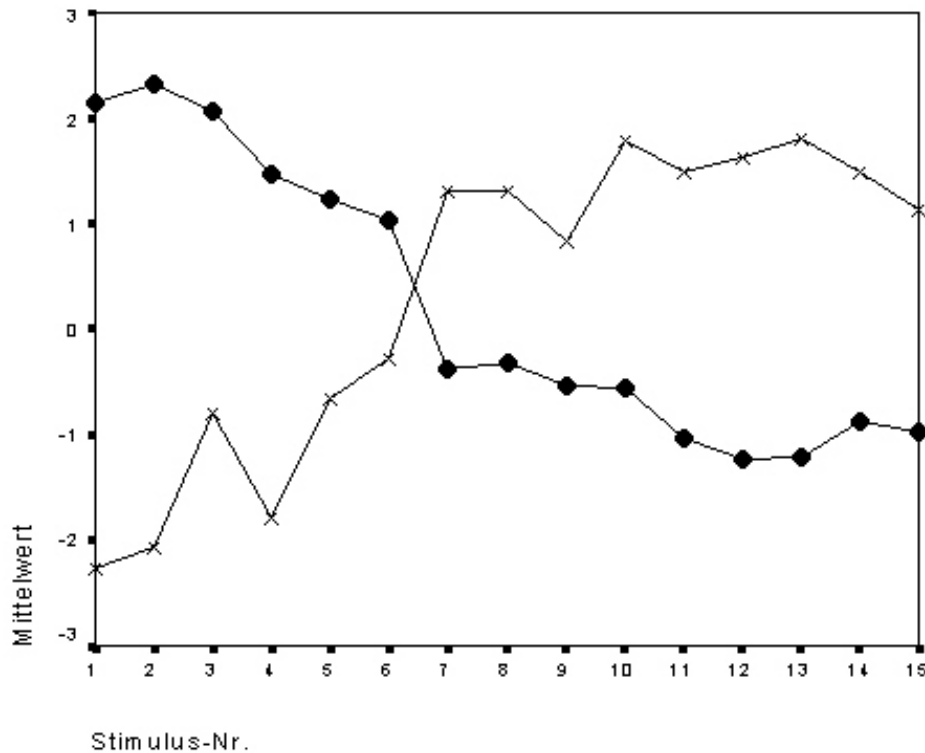


Abbildung 6.8: Mittelwerte der Beurteilungen von Stimuli S01–S15 für die Skalen: ‘abgeschlossen/ ungeschlossen’ (×) und ‘Frage/ Aussage’ (●).

leichten Wiederanstieg bis 3,55 st (S03–S06). Der zweite Abschnitt beinhaltet nur deutlich fallend-steigende Konturen (S07–S15). Die Stimuli des ersten Abschnitts werden deutlich als ‘Aussage’ und tendenziell auch als ‘abgeschlossen’ beurteilt. Die Stimuli des zweiten Abschnitts werden recht deutlich als ‘unabgeschlossen’ und tendenziell auch als ‘Frage’ beurteilt. Diese beiden Abschnitte werden durch eine relativ scharfe Grenze voneinander getrennt: Zwischen S06 und S07 tritt ein Mittelwertunterschied von mehr als 1 Skaleneinheit auf. Für beide Skalen entsteht eine recht klare Zuordnung je eines der beiden Abschnitte mit je einer semantischen Kategorie. Diese Grenze verläuft innerhalb des oberen Bereichs (S04–S15), und die ANOVAs bestätigen für diesen Bereich den gefundenen Mittelwertunterschied (vgl. Tabelle 6.8). Nach dieser groben Betrachtung der Ergebnisse werden die Hypothesen bezüglich der Abgrenzung eines pseudoterminalen Melodiemusters von einem nicht-terminalen also klar bestätigt.

Bei einer detaillierten Betrachtung wird ein zusätzlicher Effekt deutlich: Innerhalb des ersten Abschnitts (S01–S06) zeigen sich auf beiden Skalen Fluktuationen in den Urteilen. Auf der Skala ‘Frage/ Aussage’ scheinen sich zwei – nicht besonders scharf getrennte – Abschnitte herauszubilden: S01–S03 vs. S04–S06. Ähnliches – wenn auch weniger klar – kann auf der Skala ‘abgeschlossen/ ungeschlossen’ beobachtet werden. Hier bilden S01–S02 vs. S05–S06 zwei deutlich getrennte Stufen. Die Urteile für S03 und S04 fügen sich jedoch nicht in

diese Stufen ein: Einer von beiden scheint ein Ausreißer zu sein. Diese Stufenbildung innerhalb des Abschnitts S01–S06 ist für keine der beiden Skalen zu vernachlässigen, da zwischen diesen beiden Stufen jeweils ein Mittelwertunterschied von mehr als 1 Skalenpunkt auftritt. Die ANOVAs für diesen Abschnitt bestätigen diese Unterschiede als signifikant (vgl. Tabelle 6.8). Die F-Werte sind zwar kleiner als in den ANOVAs für den Bereich S04–S15, liegen jedoch in einer Größenordnung, die in Abschnitt 6.3.3.2 als relevant gewertet wurde.

6.3.3.4 Zusätzliche Skalen

Die drei verbleibenden Skalen

- ‘interessiert/ nicht interessiert’
- ‘unhöflich/ höflich’
- ‘emotional/ sachlich’

wurden nicht strikt hypothesenorientiert in die Untersuchung aufgenommen, sondern vielmehr aus methodischen Gründen (vgl. 6.2.3). Es sollen daher keine konkreten Hypothesen formuliert werden. Abbildung 6.9 und Tabelle 6.9 zeigen die Ergebnisse zu diesen drei Skalen.

Skala	S01–S05				S04–S15			
	df	F	p	Sig	df	F	p	Sig
‘interessiert/ nicht interessiert’	3,695	0,857	0,484		7,711	16,249	0,000	*
‘unhöflich/ höflich’	3,550	11,089	0,000	*	6,980	2,400	0,021	
‘emotional/ sachlich’	3,211	30,916	0,000	*	7,723	4,357	0,000	*

Tabelle 6.9: ANOVAs für drei Skalen und zwei Bereiche des Kontinuums (unterer: S01–S05; oberer: S04–S15). df = Anzahl der Freiheitsgrade, p = Irrtumswahrscheinlichkeit, Sig. = Signifikanz.

Die Stimuli S01–S08 wurden auf der Skala ‘interessiert/ nicht interessiert’ sehr deutlich als ‘neutral’ beurteilt. Die Skala trägt also nicht zur Unterscheidung einer terminalen von einer pseudoterminalen Kontur bei. Die ANOVA bestätigt, dass im Bereich S01–S05 keine signifikanten Unterschiede in den Urteilen auftreten. Stimuli mit zunehmender Wiederanstiegshöhe (ab S09 mit 6,55 st) werden zunehmend als ‘interessiert’ beurteilt. Es resultieren (etwa zwischen S08 und S11) Unterschiede von mehr als 1 Skalenpunkt, die von der ANOVA für S04–S15 als signifikant bestätigt werden.

Die Skala ‘unhöflich/ höflich’ bildet offenbar die Unterscheidung von terminalen Konturen (S01, S02) einerseits und pseudo- sowie nicht-terminalen Konturen andererseits ab. S01 und S02 werden tendenziell als ‘unhöflich’ beurteilt, für die restlichen Stimuli schwankt die Beurteilung um den Wert 0. Entsprechend zeigt die ANOVA nur für den unteren Bereich (S01–S05) signifikante Unterschiede.

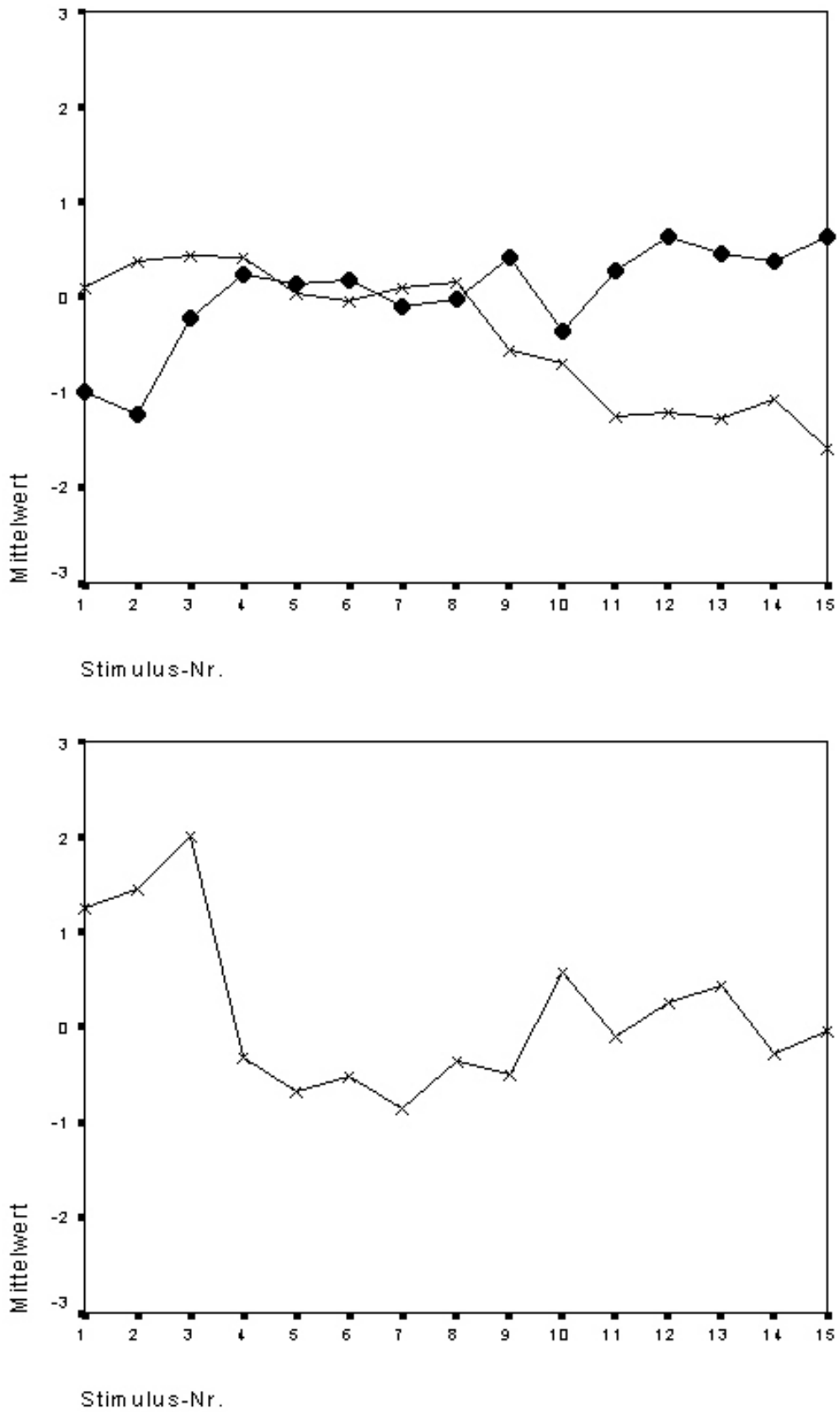


Abbildung 6.9: Mittelwerte der Beurteilungen von Stimuli S01–S15 für die Skalen: **oben:** ‘unhöflich/ höflich’ (●) und ‘interessiert/ nicht interessiert’ (×); **unten:** ‘emotional/ sachlich’.

Die Beurteilung auf der Skala ‘emotional/ sachlich’ fällt für die Stimuli S04–S15 in der Tendenz neutral aus. Die Urteile schwanken um den Wert 0, allerdings treten innerhalb dieser Schwankungen recht große Mittelwertunterschiede auf (z.B. S07 vs. S10), die von der ANOVA – wenn auch mit einem recht kleinen F-Wert – als signifikant bestätigt werden. Innerhalb dieser Schwankungen wird eine leichte Tendenz erkennbar: Die Stimuli S04–S09 werden eher als ‘emotional’ beurteilt. Deutlich hingegen werden die Stimuli S01–S03 als ‘sachlich’ beurteilt. Zwischen S03 und S04 entsteht eine recht scharfe Grenze, die sich in einem Mittelwertunterschied von mehr als 2 Skaleneinheiten manifestiert. Die ANOVA für S01–S05 bestätigt diesen Unterschied (mit einem recht hohen F-Wert) als signifikant.

6.3.4 Kontrollstimulus Sk00

Der Kontrollstimulus Sk00 wurde dem Experiment hinzugefügt, weil kein Stimulus aus dem systematisch erzeugten Kontinuum einen Fall mit einem glatten tiefen Auslaufen aufweist. Daraus folgt eine zusätzliche Fragestellung, die in diesem Rahmen berücksichtigt werden kann:

- Wird ein phrasenfinale Fallen mit einem glatten tiefen Auslaufen (Sk00) eher wie eine durchgehend fallende Kontur (S02) oder wie eine fallende Kontur mit leichtem Wiederanstieg (S03) beurteilt?

Um dieser Frage nachzugehen, wurden mit Hilfe von t-Tests die Urteile für Stimuli S02 vs. Sk00 einerseits und Sk00 vs. S03 andererseits auf jeder der neun Skalen verglichen. Tabelle 6.10 zeigt die Ergebnisse. Abbildung 6.10 veranschaulicht die Mittelwerte der drei Stimuli in einem Balkendiagramm.

Skala	Mittelwerte			t-Test S02 vs. Sk00				t-Test S03 vs. Sk00			
	S02	Sk00	S03	t	df	p	Sig.	t	df	p	Sig.
1	-2,07	-2,14	-0,79	-0,305	42	0,762		-5,739	42	0,000	*
2	-1,77	-0,93	-0,67	4,384	42	0,000	*	-1,145	42	0,259	
3	1,44	1,35	2,00	-0,388	42	0,700		-2,921	42	0,006	*
4	2,33	1,79	2,07	-2,630	42	0,012		-1,391	42	0,171	
5	0,37	0,98	0,44	2,324	42	0,025		1,816	42	0,077	
6	-1,42	-0,72	-0,70	2,496	42	0,017		-0,078	42	0,938	
7	2,05	1,58	1,63	-1,896	42	0,065		-0,221	42	0,826	
8	-1,51	-0,51	-0,95	4,383	42	0,000	*	1,710	42	0,095	
9	-1,23	0,35	-0,21	5,839	42	0,000	*	1,785	42	0,082	

Tabelle 6.10: Einzelvergleiche der Urteile für S02 vs. Sk00 und S03 vs. Sk00: Ergebnisse der zweiseitigen t-Tests: df = Anzahl der Freiheitsgrade, p = Irrtumswahrscheinlichkeit, Sig. = Signifikanz.

Wenn Sk00 grundsätzlich wie die durchgehend fallende Kontur beurteilt würde, so dürften sich keine signifikanten Unterschiede in den Mittelwertvergleichen für Sk00 und S02

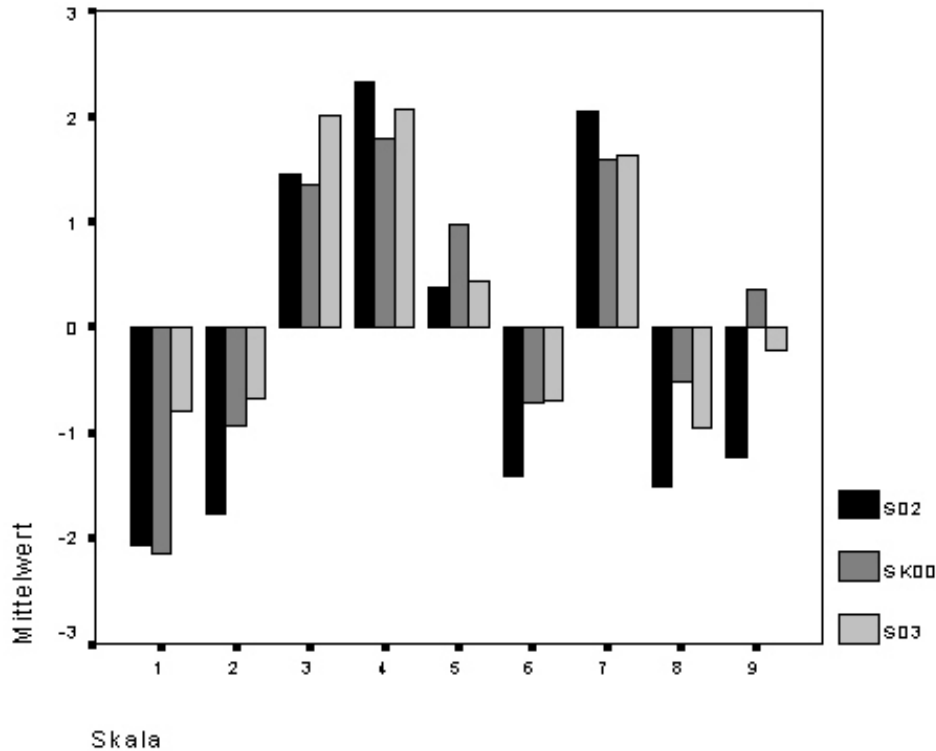


Abbildung 6.10: Mittelwerte der Urteile für S02, Sk00 und S03. Zur Bezeichnung der Skalen s. Tabelle 6.1, Abschnitt 6.2.3.

ergeben. Es treten jedoch für drei Skalen signifikante Unterschiede auf. In einem dieser drei Fälle (Skala 8 'selbstbewusst/ schüchtern') beträgt der Unterschied 1 Skalenpunkt und ist daher als relevant zu betrachten.

Im Vergleich Sk00 vs. S03 treten für zwei Skalen signifikante Unterschiede auf. In einem Fall (Skala 1 'abgeschlossen/ ungeschlossen') beträgt dieser Unterschied mehr als 1 Skalenpunkt, ist also ebenfalls relevant. In Abbildung 6.10 sind keine Regelmäßigkeiten in den Urteilen zu erkennen: Auf einigen Skalen werden S02 und Sk00 sehr ähnlich beurteilt (Skalen 1, 3), auf anderen Skalen werden S03 und Sk00 ähnlich beurteilt (Skalen 2, 6, 7), auf wieder anderen Skalen scheinen alle drei Urteile voneinander abzuweichen (Skalen 8, 9), und auf einer Skala sind sich nur die Urteile für S02 und S03 sehr ähnlich (Skala 5).

Insgesamt lässt sich als Ergebnis festhalten, dass ein tiefes glattes Auslaufen sich weder mit einer durchgehend fallenden noch mit einer fallenden und leicht wiederansteigenden Kontur gleichsetzen lässt. Die beiden relevanten Mittelwertunterschiede deuten erstens an, dass mit der glatt auslaufenden Kontur weniger Selbstbewusstsein assoziiert wird als mit der durchgehend fallenden, und zweitens, dass mit der glatt auslaufenden Kontur mehr Abgeschlossenheit assoziiert wird als mit der fallend-leicht-steigenden.

6.3.5 Prüfung der Hypothese 4

Analog zu den Berechnungen in Abschnitt 5.3.2.3 können auch anhand der Ergebnisse des Perzeptionsexperiments Berechnungen hinsichtlich der Hypothese 4 (vgl. 4.2) durchgeführt werden. Die Ergebnisse aus 6.3.3 und 6.3.4 legen nahe, dass die Stimuli S04–S06 recht gut eine pseudoternale Kontur repräsentieren, während der Status von S03 eher unklar ist. Die Dauer des Wiederanstiegs beträgt für alle Stimuli 120 ms (vgl. 6.2.5.4). Für den *glissando threshold* ergibt sich aus Formel 2.3 $g_{thr} = 11,1$ st/s. Als pseudoternale Kontur mit dem kleinsten Wiederanstieg hat sich in diesem Experiment S04 (1,55 st) herausgestellt (s.o.). Für S04 ergibt sich eine Anstiegsrate von $g_{S04} = 12,9$ st/s. Es ist also $g_{thr} \approx g_{S04}$, d.h. die Anstiegsrate der pseudoternalen Kontur mit dem kleinsten Wiederanstieg liegt gerade im Bereich des *glissando threshold*. (Die Raten der beiden anderen Stimuli S05 und S06 liegen folglich deutlich darüber: $g_{S05} = 21,3$ st/s; $g_{S06} = 29,6$ st/s). Hypothese 4(a) kann also auch durch das Perzeptionsexperiment bestätigt werden.

Zur Prüfung von Hypothese 4(b) stehen hier mehrere mögliche Werte für das nicht-terminale Muster zur Verfügung. In einem ersten Versuch soll der gefundene Grenzwert gewählt werden: Die Ergebnisse suggerieren, dass ab dem Stimulus S07 der Wiederanstieg mit einem nicht-terminalen Muster assoziiert wird. Eine Rechnung mit diesem Stimulus S07 (Wiederanstieg 4,55 st) ergibt eine Anstiegsrate von $g_{S07} = 37,9$ st/s. Ein Vergleich mit dem pseudoternalen Muster mit mittlerem Wiederanstieg (S05) ergibt einen Quotienten von $g_{S07}/g_{S05} \approx 1,78$. Damit kann, selbst bei der Rechnung mit dem unteren Grenzwert für die nicht-terminale Kontur, Hypothese 4(b) bestätigt werden.

Kapitel 7

Zusammenfassende Diskussion

Im Mittelpunkt der vorliegenden Arbeit stand ein Melodiemuster des Deutschen, welches in zweierlei Hinsicht einen interessanten Untersuchungsgegenstand darstellt: Zum einen ist es bisher nicht in Intonationsmodelle für das Deutsche integriert, zum anderen kann es auch aus einer psychophonetischen Perspektive diskutiert werden. Dieses Kapitel fasst die wesentlichen Ergebnisse der Arbeit zusammen und diskutiert diese im Rahmen der vorgestellten Grundlagen zur Produktion und Perzeption der Sprechmelodie (7.1) sowie zum Status des untersuchten Musters im Rahmen der Intonationsmodellierung (7.2). Abschließend werden aus einer methodischen Perspektive Konsequenzen für weitere Forschungen zusammengefasst (7.3).

7.1 Interpretation der wesentlichen Ergebnisse

Laut der Beobachtungen von Peters (1999) gibt es ein phrasenfinale Melodiemuster ([2;]), welches sich durch eine fallende Kontur mit leichtem finalen Wiederanstieg auszeichnet. Es unterscheidet sich in seiner phonetischen Form von einem tief-fallenden ([2.]) sowie einem (deutlich) fallend-steigenden Muster ([2.,]). Es wurde die Hypothese formuliert, dass diese drei Muster sich auch in ihrer Semantik unterscheiden. Zusammenfassend wurden die hypothetischen Funktionen der drei Muster als ‘terminal’ ([2.]), ‘nicht-terminal’ ([2.,]) und ‘pseudoterminal’ ([2;]) bezeichnet. Es existierten bisher keine auf experimentellen Befunden basierenden vergleichenden Angaben zur phonetischen Form oder zur Semantik der drei Muster. Es wurden zwei Untersuchungen durchgeführt (Kapitel 5 und 6), durch deren Ergebnisse die Grundhypothese (vgl. 4.2) bestätigt werden konnte:

- **Das pseudoterminal Melodiemuster unterscheidet sich hinsichtlich seiner phonetischen und seiner semantischen Eigenschaften von einem tief-fallenden sowie von einem fallend-steigenden Melodiemuster.**

In den folgenden zwei Abschnitten werden einige Aspekte zur Phonetik und Semantik der untersuchten Muster resümiert und interpretiert.

7.1.1 Phonetische Eigenschaften der untersuchten Melodiemuster

Die Korpusanalyse hat in den Untersuchungen zur Hypothese 2 (4.2) gezeigt, dass sich das [2;]-Muster und das [2..]-Muster in der Höhe und im *Timing* des Wiederanstiegs unterscheiden. Die Wiederanstiegshöhe hat sich als der bedeutendere Faktor herausgestellt. Im Perceptionsexperiment wurde das *Timing* nur implizit berücksichtigt: Bei der Erzeugung eines Stimuluskontinuums aus einem authentischen Fall eines mit [2;] etikettierten Musters musste das *Timing* und somit die Dauer des Wiederanstiegs synthetisch manipuliert werden, damit auch ein höherer Wiederanstieg (also ein [2..]-Muster) resynthetisierbar wurde. Mit Hilfe eines Kontrollstimulus (Sko1) wurde geprüft, ob das Ausgangsbeispiel mit spät beginnendem und kurzem Wiederanstieg (Sko1) anders beurteilt wird, als mit einem früher beginnenden und längerem Wiederanstieg (S05). Die Beurteilungen fielen für fast alle Skalen gleich aus.

Die Ergebnisse aus beiden Untersuchungen lassen vermuten, dass das *Timing* keinen wesentlichen Beitrag zur Unterscheidung der Konturen leistet. Dennoch sind in der Korpusanalyse *Timing*-Unterschiede aufgetreten. Eine mögliche Erklärung hierfür bietet folgende Überlegung. Zwei Wiederanstiege, die eine exakt gleiche Wiederanstiegsdauer bei unterschiedlicher Anstiegshöhe aufweisen, unterscheiden sich in ihrer Anstiegsrate. Seien die Mittelwerte der Wiederanstiegshöhe aus der Korpusanalyse betrachtet: 1,66 st für [2;] und 6,60 st für [2..]. Hätten zwei Anstiege mit diesen Werten eine Dauer, die typisch ist für das [2;]-Muster (71 ms), so würden sich folgende Anstiegsraten ergeben: 23 st/s für [2;] und 93 st/s für [2..]. Ein derart steiler Wiederanstieg von 93 st/s liegt nach Xu und Sun (2002) nicht im Bereich des Produzierbaren. Damit ein Wiederanstieg von 6,60 st erreicht werden kann, muss er also aufgrund von Beschränkungen des Produktionsapparats früher einsetzen bzw. eine größere Dauer als 71 ms aufweisen. Es erscheint aufgrund dieser Überlegung sowie der Ergebnisse der Arbeit naheliegend, dass gefundene Unterschiede im *Timing* von [2;] vs. [2..] kompensatorischer Natur sind, während die Wiederanstiegshöhe den maßgeblichen Unterscheidungsfaktor darstellt.

Anhand der Ergebnisse des Perceptionsexperiments zur Hypothese 3 können perzeptorisch fundierte Grenzen im Kontinuum der Anstiegshöhe angegeben werden, welche die drei Konturen voneinander trennen. Offenbar können die Stimuli S01 und S02 klar zu einer Kategorie zusammengefasst werden: Diese beiden Stimuli mit durchgehend fallender Kontur wurden auf sechs semantischen Skalen deutlich von den anderen Stimuli unterschieden. Unklar sind die Ergebnisse hinsichtlich der Kontur mit dem kleinsten Wiederanstieg (S03). Die Stimuli S04, S05 und S06 setzten sich auf zwei Skalen deutlich von den Stimuli mit höherem Wiederanstieg ab. Es ist also naheliegend, die Grenze zwischen der terminalen und der pseudoterminalen Kontur im Bereich von S03 und die Grenze zwischen der pseudoterminalen

Kontur und der nicht-terminalen zwischen S06 und S07 anzusetzen.

Dies würde einen Wiederanstieg von etwa 1,5–3,5 st für das pseudoterminalen Muster, und einen Wiederanstieg ab etwa 4,5 st für das nicht-terminale Muster bedeuten. Diese Werte decken sich mit denen aus der Korpusanalyse recht gut. Die nicht-terminale Kontur kann ohnehin sehr unterschiedlich hohe Anstiege aufweisen. Der Mittelwert für den Wiederanstieg der pseudoterminalen Kontur liegt mit 1,66 st allerdings in der Größenordnung des Wiederanstiegs von S04 (1,55 st), welcher sich hier als untere Grenze für eine pseudoterminalen Kontur herausgestellt hat.

Die perzeptorisch basierte Grenze zwischen dem terminalen und dem pseudoterminalen Muster ist keinesfalls scharf, sondern kontinuierlich ausgeprägt. Dagegen konnte für die Abgrenzung des fallend-leicht-steigenden Musters gegen das deutlich fallend-steigende Muster eine relativ scharfe Grenze gefunden werden: Auf zwei Skalen ergab sich ein abrupter Übergang in der Beurteilung der Stimuli von einer semantischen Kategorie in Richtung auf die andere.

Die Berechnungen zur Hypothese 4 (vgl. 4.2) lieferten mit den Daten aus beiden Untersuchungen grundsätzlich dasselbe Ergebnis: Die Anstiegsrate einer pseudoterminalen Kontur liegt im Bereich des bzw. über dem *glissando threshold*. Es liegt der Schluss nahe, dass der leichte Wiederanstieg von F_0 tatsächlich als ein dynamischer Ton wahrgenommen werden kann. Die Vergleichsrechnung zur Wiederanstiegshöhe von [2:] vs. [2.,] hat – wiederum basierend auf den Daten beider Untersuchungen – gezeigt, dass sich die Wiederanstiege um einen Faktor unterscheiden, der nach 't Hart et al. (1990) zur Diskrimination zweier (relevant) verschiedener Konturen nötig ist. Die Ergebnisse zur Hypothese 4 unterstützen also das generelle Ergebnis, da sie es aus einer psychophonetischen Perspektive plausibel erscheinen lassen.

Es stellte sich eingangs die Frage, ob ein derart kleiner Wiederanstieg von F_0 wie im [2;]-Muster überhaupt im Rahmen der intendierten Sprechmelodie produzierbar ist: Ein Wiederanstieg um nur ca. 2 st liegt in der Tat in der Größenordnung von mikrointonatorischen Effekten. Die Ergebnisse der Korpusanalyse suggerieren jedoch, dass das pseudoterminalen Muster ein makrointonatorisches Phänomen ist, da es im Datenkorpus in einer Vielzahl verschiedener segmenteller Kontexte auftritt.

7.1.2 Semantische Eigenschaften der untersuchten Melodiemuster

Es ist bei der Interpretation der Ergebnisse zu berücksichtigen, dass – wie in der Einleitung erwähnt – die Pragmatik einer Äußerung von der Semantik jedes ihrer Bestandteile beeinflusst werden kann. Selbst ein prototypisches Intonationsmuster der 'Höflichkeit' könnte eine Äußerung, deren lexikalische Bestandteile Beschimpfungen *par excellence* darstellen, vermutlich nicht als 'höflich' erscheinen lassen. Unter diesem Gesichtspunkt kann nicht erwartet werden, dass die Stimuli im Perzeptionsexperiment auf einer semantischen Skala in

extremer Weise verschieden beurteilt werden, selbst wenn die phrasenfinalen Muster sich extrem unterscheiden (z.B. S01 vs. S15). Denn das phrasenfinale Muster ist nur *ein* Bestandteil der Äußerung. Es wurde zwar versucht, aus semantischer Perspektive möglichst ‘neutrales’ Sprachmaterial für die Stimuli zu wählen, doch können nicht alle Bestandteile einer Äußerung kontrolliert werden. Konkret ausgewirkt hat sich dieser Umstand auf diversen Skalen in der Weise, dass die Melodiemuster nicht durch die Zuordnung zu je einer Randkategorie der Skala unterschieden wurden, sondern vielmehr durch die Zuordnung zur Randkategorie in einem Fall und zur neutralen Mitte der Skala im anderen. Zum Beispiel wurde global betrachtet keiner der Stimuli als ‘höflich’ beurteilt.

Im Perzeptionsexperiment ist das terminale Muster als ‘kategorisch’, ‘dominant’, ‘sicher’, ‘selbstbewusst’, ‘unhöflich’ und ‘sachlich’ beurteilt worden. Die pseudoterminalen Muster werden auf diesen Skalen allerdings nicht (deutlich) mit den entsprechenden Antonymen assoziiert, sondern eher mit der neutralen Mitte der Skala (s.o.). Dieser Umstand vermag das generelle Ergebnis jedoch nicht abzuschwächen: Die terminalen wurden klar von pseudoterminalen Konturen unterschieden, wenn auch nicht auf die größt mögliche Weise. Zum Beispiel wird ein leichter Wiederanstieg zwar nicht mit ‘höflich’ assoziiert, wohl aber unentschieden beurteilt, was sich durchaus als ‘nicht unhöflich’ interpretieren lässt. Hier kann ein interessanter Blick zurück auf die Beschreibung der drei Intonationstypen bei von Essen (1964) geworfen werden: Innerhalb des interrogativen Typus nennt von Essen eine Variante “mit geringer melodischer Bewegung”, welche “als Ausdruck der Höflichkeit” Verwendung findet (vgl. 3.1).

Die terminalen sowie die pseudoterminalen Konturen wurden relativ klar als ‘abgeschlossen’ und als ‘Aussage’ beurteilt, die Stimuli S07–S15 als ‘unabgeschlossen’ und ‘Frage’. Auf der Skala ‘interessiert/ nicht interessiert’ ergab sich ebenfalls eine, wenn auch deutlich weniger scharfe Trennung. Die Grenze lag hier bei einem deutlich höheren Wiederanstieg, nämlich zwischen S08 und S09, und damit bei 6,55 st. Laut der Korpusanalyse liegt ein solcher Wiederanstieg dicht am Mittelwert für die fallend-steigende Kontur (6,60 st). Die Trennung auf der Skala ‘interessiert/ nicht interessiert’ stellt daher mit großer Wahrscheinlichkeit keine Grenze zwischen den pseudoterminalen und nicht-terminalen Konturen dar. Es handelt sich dabei vermutlich eher um eine weitere Bedeutungsnuance innerhalb der deutlich fallend-steigenden Konturen: ‘Frage’ vs. ‘interessierte’ (oder verstärkte) ‘Frage’.

In Abschnitt 4.1.3 wurde eine Hypothese bezüglich der Semantik der ‘Terminalität’ formuliert, aus der zwei Merkmale (‘abgeschlossen’, ‘kategorisch’) abgeleitet wurden. Es wurde vorgeschlagen, dass sich die drei untersuchten Muster anhand dieser beiden Merkmale unterscheiden lassen. Dies konnte (als Teil der Grundhypothese) bestätigt werden.

Die gefundenen Ergebnisse lassen sich im Rahmen einer universellen Eigenschaft des sprachlichen Gebrauchs von Tonhöhen interpretieren. Ohala (1984) fasst zusammen, dass in der Funktion von hohem vs. tiefem F_0 sowohl im Vergleich der menschlichen Sprachen untereinander als auch im Vergleich von Vokalisierung von verschiedenen Spezies eine wesent-

liche Gemeinsamkeit zu beobachten ist: Ein (von Mensch oder Tier produziertes) hohes F_0 ist primär assoziiert mit einem kleinen Menschen oder Tier, ein tiefes F_0 mit einem großen Menschen oder Tier. In einer sekundären Funktion kann ein hohes F_0 Bedeutungen wie ‘unterordnend, nicht bedrohlich usw.’ übermitteln, ein tiefes F_0 hingegen ‘dominant, aggressiv, bedrohlich usw.’. In den Sprachen der Welt manifestiert sich dieses Phänomen auf unterschiedliche Weise, z.B. in einer universellen Tendenz zur Verwendung von hohem F_0 zum Ausdruck von Fragen: Ein Fragender ordnet sich dem Befragten unter. Ohala (1984) führt diese und weitere Beobachtungen zurück auf einen angeborenen *frequency code*.

Das pseudoterminal Melodiemuster ähnelt in seinem global fallenden Verlauf dem terminalen. Damit geht grob eine ähnliche Bedeutung der beiden Muster einher: Sie signalisieren eine ‘abgeschlossene Aussage’. Im Detail unterscheiden sich die Muster jedoch formal und funktional: Mit dem tiefen Fallen der terminalen Kontur wird laut der Ergebnisse ‘Dominanz’ und ‘Kategorialität’ ausgedrückt. Durch den leichten Wiederanstieg der pseudoterminalen Kontur werden gerade diese Bedeutungskomponenten unterdrückt (vgl. 4.1.3). Mit dem vergleichsweise hohen F_0 (bzw. dem weniger tiefen F_0 als bei der terminalen Kontur) ordnet sich der Sprecher dem Hörer unter. Der Unterschied zwischen den F_0 -Konturen ist nicht groß; damit einhergehend ist auch der Bedeutungsunterschied zwischen diesen Mustern nicht total, sondern partiell. Wird aber der finale F_0 -Anstieg deutlich größer, so ändert sich auch die Bedeutung in einer drastischeren Weise: Das deutlich fallend-steigende Muster wird – im Gegensatz zu den beiden ersten Mustern – nicht mit einer ‘abgeschlossenen Aussage’, sondern vielmehr mit einer ‘Frage’ assoziiert. Es erscheint naheliegend, dass die gefundenen Ergebnisse zur Form und Funktion der drei untersuchten Melodiemuster auf Ohalas (1984) *frequency code* zurückzuführen sind.

7.2 Ist das pseudoterminal Melodiemuster eine Kategorie der deutschen Intonationsphonologie?

In Abschnitt 3.2 wurden vier Kriterien vorgestellt, die (im Rahmen dieser Arbeit) zur Klassifizierung von Intonationsmodellen verwendet werden können. Zur Beantwortung der Frage, ob die Ergebnisse dieser Arbeit die Eingliederung des pseudoterminalen Musters in phonologische Modelle der deutschen Intonation gebieten, müssen die Ergebnisse hinsichtlich der Grundannahmen der Intonationsmodelle diskutiert werden. Die beiden vorgestellten Beschreibungsrahmen KIM/PROLAB und GToBI wurden in dieses Schema eingeordnet. Der wesentliche Unterschied zwischen den beiden Modellen besteht in der Definition der zugrundeliegenden Einheiten der Intonation: In KIM werden *ganzheitliche Tonkonturen*, in GToBI *Zieltöne* angesetzt.

Es stellt sich also die Frage, ob die beiden vorgestellten Modelle aufgrund ihrer Grundannahme das pseudoterminal Melodiemuster beschreiben *können*. Im KIM sind keinerlei Einschränk-

kungen bezüglich der phonetischen Form von Konturen verankert. Es wäre daher grundsätzlich denkbar, ein zusätzliches Merkmal in das Modell zu integrieren, so dass die Kontur eines phrasenfinalen Gipfels mit final fallend-leicht-steigender Kontur beschrieben werden könnte. Allerdings stellt sich die Frage, wie dies im Detail geschehen soll. Aufgrund der Ergebnisse zur Semantik wäre es sinnvoll, durch Merkmale die drei Klassen ‘terminal’, ‘pseudoterminal’ und ‘nicht-terminal’ unterscheiden zu können. KIM hat ein binäres Merkmal $\langle \pm \text{TERMIN} \rangle$ (‘terminal’). Eine Möglichkeit bestünde darin, statt dieses binären Merkmals ein Merkmal mit drei Ausprägungen anzusetzen: $\langle \alpha \text{TERMIN} \rangle$. Eine andere Möglichkeit wäre die Einführung eines zweiten binären Merkmals, z.B. $\langle \pm \text{PSEUDO} \rangle$. In der Merkmals-hierarchie (vgl. Kohler 1991a:306) würde dieses Merkmal unter $\langle + \text{TERMIN} \rangle$ erscheinen. Es würden sich also folgende Merkmalspezifikationen für die drei Muster ergeben:

- ‘terminal’: $\langle + \text{TERMIN}, - \text{PSEUDO} \rangle$
- ‘pseudoterminal’: $\langle + \text{TERMIN}, + \text{PSEUDO} \rangle$
- ‘nicht-terminal’: $\langle - \text{TERMIN} \rangle$

Eine Eingliederung des pseudoterminalen Musters in das GToBI-System erscheint dagegen schwieriger. Alle Melodiemuster müssen durch die zwei Zieltöne [H] und [L] beschrieben werden. Es gibt daher nur eine mögliche Label-Kombination um einen Wiederanstieg nach einem fallenden Verlauf zu beschreiben: [H L H] (z.B. [H* L-H%]). Diese Kombination wird jedoch bereits für das nicht-terminale Muster verwendet. Zielton-Labels können in GToBI mit Hilfe von Diakritika modifiziert werden: [!] symbolisiert einen ‘Downstep’, [^] einen ‘Upstep’. Es könnte vorgeschlagen werden, dass der leichte Wiederanstieg einer pseudoterminalen Kontur auf einen [!H]-Ton zurückzuführen ist. Dann könnte das Muster mit [L-!H%] beschrieben werden. Aus einer phänomenalen Perspektive wäre eine solche Beschreibung allerdings wenig plausibel: Ein [!H] ist eine Art von [H] — der Wiederanstieg von [2:] zeichnet sich aber gerade *nicht* durch einen hohen Ton aus. Eine Unterscheidung von [L-H%] und [L-!H%] könnte vielmehr als eine GToBI-Übersetzung der PROLAB-Labels [2.?] vs. [2.,] (die sich durch einen mehr oder weniger *hohen* Wiederanstieg auszeichnen) betrachtet werden, nicht aber als [2.,] vs. [2:].

Ein Alternativvorschlag könnte derart lauten, dass es sich bei dem leichten Wiederanstieg um einen [L%]-Ton mit *Upstep* handelt: [L-^ L%]. Diese Beschreibung wäre allerdings systembedingt inadäquat: Per Regel erfahren [L%]-Töne einen *Upstep* nach [H-]-Tönen (z.B. [H-L%] → [H-%]). Ein derartiger [^ L]-Ton wird mit einem “*mittleren oder sogar hohen Tonniveau*” (Grice und Baumann 2002:282) assoziiert, so dass z.B. [H-%] ein hohes Plateau beschreibt. Wenn eine [L-^ L%]-Etikettierung für ein pseudoterminalen Muster vorgeschlagen würde, so müsste eine Erklärung gefunden werden, warum ein [^ L] in diesem Fall einen (relativ) tiefen Ton bezeichnet, im Falle des “*Upstep per Regel*” aber einen hohen.

Aufgrund der unterschiedlichen Grundannahmen in KIM vs. GToBI ergeben sich also unterschiedliche Möglichkeiten, das pseudoterminalen Muster zu beschreiben: Für einen konturbasierten Ansatz wie KIM ergeben sich dabei keine grundsätzlichen Probleme. Für einen tonbasierten Ansatz wie GToBI wurden zwei Möglichkeiten diskutiert, deren Plausibilität aber in Frage zu stellen ist. Zur Beantwortung der zentralen Frage dieses Abschnitts muss jedoch nicht nur geklärt werden, ob ein Modell das untersuchte Muster beschreiben *kann*, sondern auch, ob das Muster in den zu beschreibenden Phänomenbereich des Modells fällt.

In der Terminologie von Ladd (1996) ist die Unterscheidung zwischen dem terminalen und dem pseudoterminalen Muster eindeutig ein paralinguistisches Phänomen und daher nicht relevant für die deutsche Intonationsphonologie. Der drastische Bruch auf der Skala 'emotional/ sachlich' untermauert diese Behauptung. Wie in 3.2.2 argumentiert wurde, erscheint eine strikte Trennung zwischen Paralinguistischem und Linguistischem aber wenig plausibel. In KIM wird eine solche Trennung nicht vorgenommen und auch im Rahmen von GToBI scheint diese Trennung zumindest nicht besonders konsequent durchgeführt zu werden. Die Ergebnisse dieser Arbeit liefern deutliche Hinweise darauf, dass das pseudoterminalen Muster sich in kommunikativ relevanter Weise von den anderen beiden untersuchten Mustern absetzt. Wird die Intonationsphonologie im Sinne von KIM verstanden, so scheint das pseudoterminalen Muster also durchaus eine Kategorie der deutschen Intonation zu sein.

7.3 Ein methodisch orientierter Ausblick

Durch die Kombination einer Korpusanalyse mit einem Perzeptionsexperiment konnte das untersuchte Melodiemuster aus zwei Blickwinkeln analysiert werden, die nur zusammen eine adäquate Beschreibung und Bewertung dieses Musters ermöglichen können: Die Korpusdaten geben Aufschluss über Aspekte der Produktion des Musters und bilden eine Basis zur Formulierung von Hypothesen, während das Perzeptionsexperiment überprüft, inwiefern das Muster von Hörern erkannt und interpretiert wird.

Weitere Untersuchungen sollten an diejenigen Aspekte anknüpfen, die im Rahmen dieser Arbeit offen geblieben sind. Zum einen zeigte sich in der Korpusanalyse eine sehr ungleiche Verteilung der [2;]- vs. der [2.,]-Kontur über segmentelles Material: Der Wiederanstieg trat bei [2;] in den meisten Fällen über einem Nasal auf, bei [2.,] in den meisten Fällen über einem Vokal. Im Rahmen weiterer Korpusanalysen könnte diese Beobachtung überprüft werden. Eine Voraussetzung wären größere Datenkorpora und möglichst auch unterschiedliche Szenarien. Auf diese Weise könnten zudem weitaus mehr Teilfragestellungen untersucht werden, als es im Rahmen dieser Arbeit möglich war. Zum Beispiel könnten brauchbare Aussagen zur Auftretenshäufigkeit der unterschiedlichen Melodiemuster – auch in Abhängigkeit vom Aufnahmeszenario und vom Geschlecht des Sprechers – getroffen werden. Zweitens könnten bei großen Stichproben auch verschiedene suprasegmentelle Kontexte berücksichtigt werden, um

z.B. die Abhängigkeit der phonetischen Parameter phrasenfinaler Konturen vom vorangehenden Gipfelakzent zu untersuchen. Ein Ziel derartiger Untersuchungen würde darin bestehen, die Schlussfolgerung der vorliegenden Arbeit über das *Timing* des Wiederanstiegs zu überprüfen (vgl. 7.1.1).

Die Einbeziehung spontansprachlicher Datenkorpora ist mit Vor- und Nachteilen behaftet. Ein wesentlicher Vorteil besteht in der Natürlichkeit des untersuchten Materials. Wenn ein Melodiemuster im Rahmen spontaner Dialoge in den unterschiedlichsten Kontexten auftritt, so kann mit relativ großer Wahrscheinlichkeit ausgeschlossen werden, dass das analysierte Phänomen ein Artefakt der Aufnahmesituation oder ein – segmentell oder suprasegmentell bedingter – kontextueller, also nicht-gesteuerter Effekt ist. Zu bedenken ist dabei jedoch die Methode, mit der die Spontansprache elizitiert wird. Die hier verwendeten Daten entstammen einem Aufnahmeszenario, in dem die Versuchspersonen gewissermaßen schauspielerisch tätig werden: Die abzusprechenden Termine sind fiktiv. Inwiefern die Art des Aufnahmeszenarios einen Einfluss auf die Verwendung von Melodiemustern hat, müsste durch einen exhaustiven Vergleich mehrerer verschiedener Szenarien untersucht werden. Ein Ansatz in diese Richtung ist bereits bei Peters (1999) zu erkennen, der die Auftretenshäufigkeit verschiedener Muster in spontansprachlichen vs. lesesprachlichen Daten vergleicht.

Ein Nachteil der Spontansprache besteht in der nahezu fehlenden Kontrollierbarkeit der Daten. Die Konsequenz sind sehr inhomogene Stichproben, die konservative Auswertungsmethoden bedingen. Daher müssen Aussagen, die aufgrund der Korpusanalyse gemacht werden, mit Vorsicht formuliert werden. In diesem Fall kam das generelle Ergebnis zur Anstiegshöhe von [2;] vs. [2.,] aber sehr deutlich heraus.

Das Design des Perzeptionsexperiments hat sich im Ganzen als sehr tauglich erwiesen. Im Detail sollten aber Änderungen vorgenommen werden, falls das Experiment ein zweites Mal durchgeführt würde: Es hätte ein wesentlich kleineres Stimuluskontinuum verwendet werden können. Entgegen der Folgerung aus den informellen Vorversuchen (vgl. 6.2.5.4) genügte ein Wiederanstieg von nur 4,55 st (S07), um das Perzept einer deutlich fallend-steigenden Kontur hervorzurufen. Als Stimulus mit dem größten Wiederanstieg im Kontinuum hätte also z.B. S10 ausgereicht. Dies wird unterstützt durch das Ergebnis auf der Skala ‘interessiert/nicht interessiert’ (7.1.2). Hier trat eine Tendenz zu einer weiteren semantischen Nuancierung ab der Wiederanstiegshöhe von 6,55 st (S09) auf. Dieses Ergebnis könnte ein Hinweis auf eine feinere Unterscheidung nach dem Grad der Anstiegshöhe *innerhalb* der deutlich fallend-steigenden Konturen sein, die explizit nicht den Untersuchungsgegenstand dieser Arbeit darstellten. Weiterhin sollte bei der Stimulusgenerierung bedacht werden, wie hoch ein Wiederanstieg bei einer konstanten Anstiegsdauer sein darf, um aus Sicht der Sprachproduktion eine realistische Melodiebewegung darzustellen. Für den Stimulus mit maximalem Wiederanstieg (12,55 st bei 120 ms Dauer) im vorliegenden Experiment ergibt sich eine Anstiegsrate von 105 st/s, also ein unrealistischer Wert (vgl. Xu und Sun 2002). Auch unter diesem Aspekt hätte bei der verwendeten Anstiegsdauer von 120 ms ein maximaler Wieder-

anstieg von etwa 7,55 st (S10) gewählt werden sollen.

Es sei bei dieser Kritik aber bemerkt, dass die hier diskutierten Detailfehler das Experiment im Ganzen nicht gestört haben. Die Stimuli wurden insgesamt als 'natürlich' bezeichnet. Es kann lediglich das obere Drittel des Kontinuums im Nachhinein als überflüssig bezeichnet werden. Die Dauer eines Experimentdurchlaufs war (laut Gesprächen mit einigen Probanden nach dem Experiment) akzeptabel. Würden die Stimuli S11-S15 aus dem Experiment entfernt werden, ließen sich die frei werdenden Kapazitäten sinnvoller nutzen: z.B. durch die Wahl einer kleineren Schrittgröße für den Wiederanstieg bei der Konstruktion des Kontinuums. Ein solches Kontinuum, z.B. bestehend aus 15 Stimuli mit einer Schrittgröße von 0,66 st, würde ein feineres Messwerkzeug darstellen, welches für eine präzisere Bestimmung der Grenzen zwischen den Melodiemustern verwendet werden könnte. Dies wäre sinnvoll, da auf Basis der vorliegenden Ergebnisse hinsichtlich des Übergangs von einer fallenden (S02), über eine glatt auslaufende (Sk00) zu einer leicht steigenden Kontur (S03, S04) keine klaren Aussagen getroffen werden können.

Das hier durchgeführte Experiment zeigt, dass die Technik der semantischen Skalen ein hilfreiches und gut funktionierendes Werkzeug für die Intonationsforschung darstellt. Allerdings erscheint der hier erhobene Anspruch auf Vollständigkeit des Differentials hinsichtlich der drei semantischen Dimensionen *Potenz*, *Aktivität* und *Evaluation* nicht notwendig zu sein — die zusätzlichen Skalen haben wenig zum Gesamtergebnis der Untersuchung beigetragen. Hinsichtlich der statistischen Auswertung der Daten sollte versucht werden, ein Verfahren zu finden, welches sich als ausreichend konservativ erweist.

Grundsätzlich hat sich das methodische Vorgehen der Arbeit als erfolgreich erwiesen. Als wichtigste Konsequenz auf weitere Forschungen sollte daher festgehalten werden, dass Untersuchungen dieser Art auch auf weitere Melodiemuster übertragen werden sollten. Gerade das PROLAB-System bietet eine Vielzahl von Etiketten für phrasenfinale Konturen. Eine interessante Fragestellung, die mit einer ähnlichen wie der hier verwendeten Methode bearbeitet werden könnte, lautet: Lassen sich innerhalb der nicht-terminalen fallend-steigenden Muster ([L-H%]) feinere formale und funktionale Differenzierungen finden? Die Ergebnisse haben eine leichte Tendenz in diese Richtung aufgezeigt (s.o.). Aber repräsentieren die vier PROLAB-Labels [1.], [1.?], [2.] und [2.?] wirklich vier verschiedene Kategorien der deutschen Intonation?

Literatur

- Antoniadis, Z. und H. W. Strube (1981). Untersuchungen zum ‘intrinsic pitch’ deutscher Vokale. *Phonetica* 38, 277–290.
- Beckman, M. E. und G. Ayers Elam (1997). *Guidelines for ToBI labelling, version 3.0*. The Ohio State University.
- Bergler, R. (Hrsg.) (1975). *Das Eindrucksdifferential. Theorie und Technik*. Bern; Stuttgart; Wien: Hans Huber.
- Bühl, A. und P. Zöfel (2002). *SPSS 11. Einführung in die moderne Datenanalyse unter Windows* (8. Aufl.). München: Addison-Wesley/ Pearson Studium.
- Boersma, P. und D. Weenink (URL). *Praat: Doing phonetics by computer*. <http://www.fon.hum.uva.nl/praat/>.
- Borden, G. J., K. S. Harris, und R. J. Lawrence (1994). *Speech science primer. Physiology, acoustics and perception of speech* (3. Aufl.). Baltimore; Philadelphia: Williams & Wilkins.
- Bortz, J. (1979). *Lehrbuch der Statistik: Für Sozialwissenschaftler*. Berlin; Heidelberg; New York: Springer.
- Bruce, G. (1977). *Swedish word accents in sentence perspective*. Lund: Gleerup.
- Bußmann, H. (2002). *Lexikon der Sprachwissenschaft* (3. Aufl.). Stuttgart: Alfred Kröner.
- Crystal, D. (1969). *Prosodic systems and intonation in English*. Cambridge: Cambridge University Press.
- Dombrowski, E. (2003). Semantic features of accent contours: Effects of f0 peak position and f0 time shape. In *Proc. 15th ICPhS*, Barcelona, S. 1217–1220.
- Grice, M. und S. Baumann (2002). Deutsche Intonation und GToBI. *Linguistische Berichte* 191, 267–298.
- Herold, H. (1999). *Linux-Unix-Profertools. Awk, sed, lex, yacc und make* (3. Aufl.). Bonn: Addison-Wesley.
- Hombert, J.-M., J. J. Ohala, und W. Ewan (1979). Phonetic explanations for the development of tones. *Language* 55, 37–58.

- IPA (1999). *Handbook of the International Phonetic Association: A guide to the use of the International Phonetic Alphabet*. Cambridge: Cambridge University Press.
- IPDS (1994). *The Kiel Corpus of Read Speech*, Band 1, CD-ROM#1. Kiel: IPdS.
- IPDS (1995). *The Kiel Corpus of Spontaneous Speech*, Band 1, CD-ROM#2. Kiel: IPdS.
- IPDS (1996). *The Kiel Corpus of Spontaneous Speech*, Band 2, CD-ROM#3. Kiel: IPdS.
- IPDS (1997a). *The Kiel Corpus of Spontaneous Speech*, Band 3, CD-ROM#4. Kiel: IPdS.
- IPDS (1997b). *xassp user's manual (Advanced Speech Signal Processor under the X window system) – user's manual version 1.2.15*. *AIPUK* 32, 31–115.
- Isačenko, A. V. und H.-J. Schädlich (1970). *A model of standard German intonation*. The Hague; Paris: Mouton.
- Ishizaka, K. und J. L. Flanagan (1972). Synthesis of voiced sounds from a two-mass model of the vocal cords. *The Bell System Technical Journal* 51, 1233–1268.
- Johnson, K. (1997). *Acoustic & auditory phonetics*. Cambridge, Massachusetts; Oxford: Blackwell.
- Jones, D. (1969). *An outline of English phonetics* (9. Aufl.). Cambridge: Heffer & Sons.
- Klatt, D. H. (1973). Discrimination of fundamental frequency contours in synthetic speech: Implications for models of pitch perception. *JASA* 53, 8–16.
- Kohler, K. J. (1987). Categorical pitch perception. In *Proc. 11th ICPPhS*, Tallinn, S. 331–333.
- Kohler, K. J. (1990). Macro and micro f_0 in the synthesis of intonation. In J. Kingston und M. E. Beckman (Hrsg.), *Papers in Laboratory Phonology I. Between the grammar and the physics of speech*, S. 115–138. Cambridge; New York; Port Chester; Melbourne; Sydney: Cambridge University Press.
- Kohler, K. J. (1991a). A model of German intonation. *AIPUK* 25, 295–360.
- Kohler, K. J. (1991b). Terminal intonation patterns in single-accent utterances of German: Phonetics, phonology and semantics. *AIPUK* 25, 115–185.
- Kohler, K. J. (1992). Prosodisches Transkriptionssystem für die Etikettierung von Sprachsignalen. *AIPUK* 26, 239–252.
- Kohler, K. J. (1997). Modelling prosody in spontaneous speech. In Y. Sagisaka, N. Campbell, und N. Higuchi (Hrsg.), *Computing prosody*, S. 187–210. Berlin/Heidelberg/New York/Tokyo: Springer.
- Kohler, K. J. (2004). Categorical speech perception revisited. In *Proc. From sound to sense: 50+ years of discoveries in speech communication*, Cambridge. M.I.T.: CD-ROM; <http://www.ipds.uni-kiel.de/kjk/forschung/lautmuster.de.html>.

- Kortekaas, R. W. L. und A. Kohlrausch (1997). Psychoacoustical evaluation of the pitch-synchronous overlap-and-add speech-waveform manipulation technique using single-formant stimuli. *JASA* 101, 2202–2213.
- Ladd, D. R. (1996). *Intonational phonology*. Cambridge: Cambridge University Press.
- Ladd, D. R. und R. Morten (1997). The perception of intonational emphasis: Continuous or categorical? *Journal of Phonetics* 25, 313–342.
- Ladd, D. R. und K. E. A. Silverman (1984). Vowel intrinsic pitch in connected speech. *Phonetica* 41, 31–40.
- Laver, J. (1994). *Principles of phonetics*. Cambridge; New York: Cambridge University Press.
- Lehiste, I. (1970). *Suprasegmentals*. Cambridge, Massachusetts: M.I.T. Press.
- Lehiste, I. und G. E. Peterson (1961). Some basic considerations in the analysis of intonation. *JASA* 33, 419–425.
- Löfqvist, A., T. Baer, N. S. McGarr, und R. Seider Story (1989). The cricothyroid muscle in voicing control. *JASA* 85, 1314–1321.
- Lieberman, A. M., K. S. Harris, H. S. Hoffman, und B. C. Griffith (1957). The discrimination of speech sounds within and across phoneme boundaries. *Journal of Experimental Psychology* 54, 358–368.
- Lieberman, M. (1975). *The intonational system of English*. Dissertation, M.I.T.
- Möbius, B. (1993). *Ein quantitatives Modell der deutschen Intonation. Analyse und Synthese von Grundfrequenzverläufen*. Tübingen: Niemeyer.
- Moulines, E. und F. Charpentier (1990). Pitch-synchronous waveform processing techniques for text-to-speech synthesis using diphones. *Speech Communication* 9, 453–467.
- Niebuhr, O. (2003). Perceptual study of timing variables in f₀ peaks. In *Proc. 15th ICPhS*, Barcelona, S. 1225–1228.
- Niebuhr, O. und K. J. Kohler (2004). Perception and cognitive processing of tonal alignment in German. In *Proc. International Symposium on Tonal Aspects of Languages: Emphasis on Tone Languages*, Beijing, S. 155–158.
- Nolan, F. (2003). Intonational equivalence: An experimental evaluation of pitch scales. In *Proc. 15th ICPhS*, Barcelona, S. 771–774.
- O'Connor, J. D. und G. F. Arnold (1973). *Intonation of colloquial English* (2. Aufl.). London: Longman.
- Ohala, J. J. (1973). Explanations for the intrinsic pitch of vowels. *Monthly Internal Memorandum, Phonology Laboratory University of California, Berkeley January 1973*, 9–26.

- Ohala, J. J. (1978). Production of tone. In V. A. Fromkin (Hrsg.), *Tone. A Linguistic Survey*, S. 5–39. New York; San Francisco; London: Academic Press.
- Ohala, J. J. (1984). An ethological perspective on common cross-language utilization of f_0 of voice. *Phonetica* 41, 1–16.
- Osgood, C. E., G. J. Suci, und P. H. Tannenbaum (1957). *The measurement of meaning*. Urbana; Chicago; London: University of Illinois Press.
- Peters, B. (1999). Prototypische Intonationsmuster in deutscher Lese- und Spontansprache. *AIPUK* 34, 1–175.
- Peters, B. (2000). Individuelle und geschlechtsspezifische Unterschiede in der prosodischen Gestaltung deutscher Lese- und Spontansprache. In H. D. Schlosser (Hrsg.), *Sprache und Kultur*, Forum Angewandte Linguistik 38. Frankfurt am Main: Lang.
- Pierrehumbert, J. B. (1980). *The phonology and phonetics of English intonation*. Dissertation, M.I.T.
- Pompino-Marschall, B. (2003). *Einführung in die Phonetik* (2. Aufl.). Berlin; New York: Walter de Gruyter.
- Reetz, H. (1999). *Artikulatorische und akustische Phonetik*. Trier: Wissenschaftlicher Verlag Trier.
- Reinholt Petersen, N. (1978). Intrinsic fundamental frequency of Danish vowels. *Journal of Phonetics* 6, 177–189.
- Rossi, M. (1971). Le seuil de glissando ou seuil de perception des variations tonales pour les sons de la parole. *Phonetica* 23, 1–33.
- Rossi, M. und M. Chafcouloff (1972). Recherche sur de le seuil différentiel de fréquence fondamentale dans la parole. *Travaux de l'Institut de Phonétique d'Aix* 1, 179–185.
- Schäfer-Vincent, K. (1982). Significant points: Pitch period detection as a problem of segmentation. *Phonetica* 39, 241–253.
- Schäfer-Vincent, K. (1983). Pitch period detection and chaining: Method and evaluation. *Phonetica* 40, 177–202.
- Selting, M. (1987). Descriptive categories for the auditive analysis of intonation in conversation. *Journal of Pragmatics* 11, 777–791.
- Sjölander, K. und J. Beskow (URL). *Wavesurfer*. <http://www.speech.kth.se/wavesurfer/>.
- Sundberg, J. (1979). Maximum speed of pitch changes in singers and untrained subjects. *Journal of Phonetics* 7, 71–79.
- 't Hart, J. (1981). Differential sensitivity to pitch distance, particularly in speech. *JASA* 69, 811–821.

- 't Hart, J., R. Collier, und A. Cohen (1990). *A perceptual study of intonation. An experimental-phonetic approach to speech melody*. Cambridge: Cambridge University Press.
- Terhardt, E. (1974). Pitch, consonance, and harmony. *JASA* 55, 1061–1069.
- Uldall, E. T. (1960). Attitudinal meanings conveyed by intonation contours. *Language and Speech* 3, 223–234.
- Uldall, E. T. (1964). Dimensions of meaning in intonation. In D. Abercrombie, D. B. Fry, P. A. D. MacCarthy, N. C. Scott, und J. L. M. Trim (Hrsg.), *In Honour of Daniel Jones*, S. 271–279. London: Longmans.
- van den Berg, J. (1958). Myoelastic-aerodynamic theory of voice production. *Journal of Speech and Hearing Research* 1, 227–244.
- von Essen, O. (1964). *Grundzüge der hochdeutschen Satzintonation* (2. Aufl.). Ratingen; Düsseldorf: A. Henn Verlag.
- Xu, Y. und X. Sun (2002). Maximum speed of pitch change and how it may relate to speech. *JASA* 111, 1399–1413.

Anhang

Anhang A

Korpusanalyse

A.1 Erläuterungen zu den Rohdatentabellen

In den Tabellen A.1, A.2, A.3 und A.4 werden folgende Abkürzungen verwendet:

Spr = Sprecher

m/f = Geschlecht des Sprechers (m = männlich / f = weiblich)

Wort = Äußerungsabschnitt vom finalen Akzent bis zum Phrasenende

Acc = Muster des finalen Akzents in PROLAB-Notation

S1 = segmentelles Material, über welchem der Wiederanstieg realisiert wird, in weiter IPA-Transkription
(r-Vokalisierungen sind mit /r/ transkribiert)

S2 = segmentelles Material, welches (in derselben Phrase) auf S1 folgt

Ri [seg/syl] = Anzahl der Segmente/ Silben, über denen der Wiederanstieg realisiert wird

Pa [syl] = Anzahl der Silben des gesamten Äußerungsabschnitts ab der finalen Akzentsilbe

FaOn = F_0 -Wert zu Beginn des fallenden Konturabschnitts

FaOf = F_0 -Wert am Ende des fallenden Konturabschnitts

RiOn = F_0 -Wert zu Beginn des steigenden Konturabschnitts

RiOf = F_0 -Wert am Ende des steigenden Konturabschnitts

DuFa = Dauer des fallenden Konturabschnitts

DuL = Dauer des tiefen Plateaus

DuRi = Dauer des steigenden Konturabschnitts

DuLS = Dauer des letzten stimmhaften Segments der Phrase

DuE = Dauer des nicht vermessbaren Abschnitts des letzten stimmhaften Segments

A.2 Rohdaten und Informationen zu [2.,]

Tabelle A.1: Informationen zu den [2.,]-Fällen

Nr.	Datei	Spr	m/f	Wort	Acc	S1	S2	Ri [seg]	Ri [syl]	Pa [syl]
1	G074a007	Tis	M	MACHen	2^	/n/	–	1	1	2
2	G081a001	Ths	M	JA	2)	/a:/	–	1	1	1
3	G081a004	Kak	M	ZUschlag'n	2^	/ŋ/	–	–	1 bis 2	3
4	G082a009	Ths	M	beSCHÄFtigt	2^	/ɪ/	/kt/	1	1	2
5	G085a003	Ths	M	JA	2)	/a:/	–	1	1	1
6	G086a016	Kak	M	DIENSTreise	2(/aɪzə/	–	3	2	3
7	G092a009	Frs	F	ANbiet'n	2(/n/	–	1	1	3
8	G093a000	Frs	F	ABzusprech'n	2^	/ɛçn/	–	3	2	4
9	G094a015	Frs	F	WO	2^	/o:/	–	1	1	1
10	G115a011	Jak	M	MITnehm'm	2^	/m/	–	1	1	2
11	G126a020	Olv	M	JANuar	2)	/ua:r/	–	2	2	3
12	G145a004	Sok	M	DIENSTreise	2^	/zə/	–	2	1	3
13	G191a001	Sar	F	JA	2)	/a:/	–	1	1	1
14	G191a010	Kae	F	BETtag	2^	/a:/	/k/	1	1	2
15	G191a018	Kae	F	deZEMber	2^	/ər/	–	1	1	2
16	G191a029	Sar	F	silVEStEr	2^	/ər/	–	1	1	2
17	G191a031	Sar	F	noVEMber	2(/ər/	–	1	1	2
18	G193a009	Sar	F	JA	2)	/a:/	–	1	1	1
19	G196a007	Sar	F	GEHT	2)	/e:/	/t/	1	1	1
20	G196a019	Kae	F	JANuar	2(/ua:r/	–	1	1	3
21	G212a007	Anl	M	JA	2)	/a:ə/	–	1	1	1
22	G213a002	Anl	M	PASST das ja	2^	/a:/	–	–	1 bis 2	3
23	G251a031	Sik	F	UHR	2^	/u:r/	–	1	1	1

Tabelle A.2: Rohdaten zu den [2.,]-Fällen

Nr.	Datei	F_0 [Hz]				Dauer [ms]				
		FaOn	FaOf	RiOn	RiOf	DuFa	DuL	DuRi	DuLS	DuE
1	G074a007	128	90	90	128	173	0	93	92	7
2	G081a001	197	119	119	–	270	0	–	311	–
3	G081a004	212	–	–	241	–	–	–	143	20
4	G082a009	151	130	130	154	344	0	50	104	15
5	G085a003	204	121	121	154	222	42	127	320	33
6	G086a016	217	140	140	272	245	0	276	86	0
7	G092a009	182	152	152	187	136	0	153	104	34
8	G093a000	196	160	160	314	480	0	242	180	56
9	G094a015	232	169	169	308	108	0	169	303	0
10	G115a011	122	101	102	124	200	53	92	60	0
11	G126a020	125	113	113	157	214	74	146	130	26
12	G145a004	221	151	151	241	421	0	221	137	0
13	G191a001	185	154	151	172	103	118	58	279	0
14	G191a010	276	173	172	243	148	47	107	181	11
15	G191a018	256	180	181	329	135	55	171	216	0
16	G191a029	249	185	185	235	220	50	101	222	66
17	G191a031	209	152	152	219	140	0	115	187	41
18	G193a009	217	154	153	340	109	52	144	238	0
19	G196a007	287	193	193	240	86	0	64	150	98
20	G196a019	247	188	188	308	154	0	197	310	71
21	G212a007	102	84	84	127	195	0	161	95	37
22	G213a002	104	–	–	111	–	–	–	155	82
23	G251a031	262	194	194	220	224	0	55	443	40

A.3 Rohdaten und Informationen zu [2;]

Tabelle A.3: Informationen zu den [2;-]Fällen

Nr.	Datei	Spr	m/f	Wort	Acc	S1	S2	Ri [seg]	Ri [syl]	Pa [syl]
1	G071a014	Tis	M	UNterbring'n	2^	/ŋ/	–	1	1	4
2	G074a005	Tis	M	TREFF'n um	2^	/m/	–	1	1	3
3	G075a000	Tis	M	HINbegeb'm	2^	/bm/	–	2	1	4
4	G081a000	Kak	M	verEINbar'n	2^	/n/	–	–	1	2
5	G085a002	Kak	M	SECHzehnte	2^	/ə/	–	1	1	3
6	G085a005	Ths	M	ORDnung	2)	/ŋ/	/k/	1	1	2
7	G085a014	Kak	M	FEIertag ist	2^	/ɪ/	/st/	–	1	4
8	G085a017	Ths	M	ORDnung	2)	/ŋ/	/k/	1	1	2
9	G086a008	Kak	M	VIERzehnt'n	2^	/n/	–	1	1	3
10	G086a008	Kak	M	verEINbar'n	2^	/n/	–	–	1	2
11	G086a016	Kak	M	SECHzehnt'n	2^	/n/	–	1	1	3
12	G086a020	Kak	M	WIEderhör'n	2^	/n/	–	1	1	3
13	G091a028	Ans	F	VIERt'n	2^	/n/	–	1	1	2
14	G092a000	Ans	F	SCHINdel	2^	/l/	–	1	1	2
15	G094a000	Ans	F	SCHINdel	2^	/l/	–	1	1	2
16	G094a029	Frs	F	SONNtag dann	2^	/n/	–	1	1	3
17	G095a001	Frs	F	beMESS'n	2^	/n/	–	1	1	2
18	G097a005	Ans	F	WUNderbar	2(/ar/	–	1	1	3
19	G101a002	Utb	F	BIN	2^	/n/	–	1	1	1
20	G101a002	Utb	F	ACHt'n	2^	/n/	–	1	1	2
21	G101a006	Utb	F	vorzubeREIt'n	2)	/aɪn/	–	3	2	2
22	G101a015	Utb	F	ACHTzehnt'n	2^	/n/	–	1	1	3
23	G102a002	Utb	F	KÖNNt'n	2^	/n/	–	1	1	2
24	G102a004	Utb	F	verREIST	2)	/aɪ/	/st/	1	1	1
25	G102a014	Nar	F	FREI	2)	/aɪ/	–	1	1	1
26	G103a000	Utb	F	GEH'N	2^	/n/	–	1	1	1
27	G103a002	Utb	F	STUND'nplan habe	2^	/bə/	–	2	1	5
28	G103a006	Utb	F	okTOber	2^	/ər/	–	1	1	2
29	G103a012	Utb	F	PASs'n	2^	/n/	–	1	1	2
30	G103a014	Utb	F	SONNa'md nehm'm	2(/mm/	–	2	1	3
31	G104a001	Utb	F	DRESd'n	2(/n/	–	1	1	2
32	G104a005	Utb	F	LOSfahr'n	2^	/n/	–	1	1	2

Fortsetzung auf nächster Seite

<i>Fortsetzung</i>										
Nr.	Datei	Spr	m/f	Wort	Acc	S1	S2	Ri [seg]	Ri [syl]	Pa [syl]
33	G104a011	Utb	F	RAHmen	2^	/n/	—	1	1	2
34	G104a013	Utb	F	ZWÖLFt'n	2(/n/	—	1	1	2
35	G105a000	Utb	F	KÖNN'n	2^	/n/	—	1	1	2
36	G105a015	Nar	F	MESSE	2^	/ə/	—	1	1	2
37	G105a016	Utb	F	FREI	2)	/aɪ/	—	1	1	1
38	G105a018	Utb	F	TAGe sind	2^	/n/	/t/	1	1	3
39	G105a018	Utb	F	beRICHtes	2^	/ə/	/s/	1	1	2
40	G106a001	Utb	F	beGRÜß'n	2^	/n/	—	1	1	2
41	G106a005	Utb	F	KUchen	2^	/n/	—	1	1	2
42	G106a007	Utb	F	überREd'n	2)	/dn/	—	2	1	2
43	G106a010	Nar	F	FEIern	2)	/n/	—	1	1	2
44	G107a000	Nar	F	zuSAMM'setz'n	2(/n/	—	1	1	3
45	G107a001	Utb	F	MÖGlich	2)	/i/	/ç/	1	1	2
46	G107a004	Nar	F	büRO	2)	/o:/	—	1	1	1
47	G107a005	Utb	F	NETTer	2^	/ɛr/	—	1	1	2
48	G107a006	Nar	F	GERne	2^	/ə/	—	1	1	2
49	G107a010	Nar	F	geBUND'n bin	2(/n/	—	1	1	3
50	G111a011	Jak	M	PASs'n	2^	/n/	—	1	1	2
51	G113a011	Jak	M	FINd'n lass'n	2(/n/	—	1	1	4
52	G115a007	Jak	M	FÜNFT'n	2(/n/	—	1	1	2
53	G115a010	Rek	M	stockHOLM	2^	/m/	—	1	1	1
54	G121a012	Kap	F	terMIN	2)	/n/	—	1	1	1
55	G121a014	Kap	F	verZICHT'n	2^	/n/	—	1	1	2
56	G121a023	Kap	F	MÜNchen	2(/n/	—	1	1	2
57	G122a009	Olv	M	IHN'n	2^	/n/	—	1	1	2
58	G123a005	Kap	F	geNAU	2)	/aʊ/	—	1	1	1
59	G123a008	Kap	F	ABmach'n	2)	/ɲ/	—	1	1	3
60	G123a010	Kap	F	ZWEIt'n	2^	/n/	—	1	1	2
61	G123a012	Kap	F	DA	2^	/a:/	—	1	1	1
62	G124a014	Olv	M	ZWÖLF	2)	/œl/	/f/	2	1	1
63	G126a014	Olv	M	KAFFe	2^	/ə/	—	1	1	2
64	G142a002	Bac	M	JA	2)	/m/	—	1	1	1
65	G143a000	Sok	M	JA	2)	/a:/	—	1	1	1
66	G144a003	Bac	M	JA	2)	/m/	—	1	1	1
67	G191a002	Kae	F	KÖLN	2^	/n/	—	1	1	1
68	G191a011	Sar	F	UHR	2^	/u:r/	—	1	1	1
69	G191a020	Kae	F	SCHADE	2^	/əə/	—	2	1	2

Fortsetzung auf nächster Seite

<i>Fortsetzung</i>											
Nr.	Datei	Spr	m/f	Wort	Acc	S1	S2	Ri	Ri	Pa	
								[seg]	[syl]	[syl]	
70	G191a021	Sar	F	VIERT'n zember	De-	2^	/ər/	–	1	1	5
71	G191a026	Kae	F	proBLEM		2)	/m/	–	1	1	1
72	G191a036	Kae	F	JA		2)	/a:/	–	1	1	1
73	G192a001	Kae	F	VIERT'n		2^	/n/	–	1	1	2
74	G192a003	Kae	F	MITTwoch		2^	/ɔ/	/x/	1	1	2
75	G192a014	Kae	F	DA(-m)		2^	/a:m/	–	2	1	1
76	G192a017	Sar	F	JA		2)	/a:/	–	1	1	1
77	G193a004	Kae	F	GUT		2)	/u:/	/t/	1	1	1
78	G194a003	Kae	F	KNAPP		2)	/a/	/p/	1	1	1
79	G194a014	Kae	F	GUT dann		2^	/n/	–	1	1	2
80	G195a019	Kae	F	SCHÖN		2^	/ø:n/	–	2	1	1
81	G196a002	Kae	F	GUT		2)	/u:/	/t/	1	1	1
82	G196a005	Sar	F	PRIma		2^	/a:/	–	1	1	2
83	G196a011	Sar	F	REICHT		2)	/a:r/	/çt/	1	1	1
84	G211a000	Anl	M	LOSleg'n		2^	/gʝ/	–	2	1	3
85	G211a000	Anl	M	noVEMber		2(/ər/	–	1	1	2
86	G211a003	Chd	M	MEInetwegen dann		2(/n/	–	1	1	5
87	G211a005	Chd	M	ZWÖLft'n		2(/n/	–	1	1	2
88	G212a003	Chd	M	MAI		2)	/a:r/	–	1	1	1
89	G213a003	Chd	M	woch'nends- ENDseminar		2^	/a:r/	–	1	1	4
90	G213a003	Chd	M	ZEIT		2^	/a:r/	/t/	1	1	1
91	G214a007	Anl	M	ORDnung		2)	/ŋ/	–	1	1	2
92	G214a009	Chd	M	EINunddreißigst'n nehm'm		2(/mm/	–	2	1	6
93	G216a000	Chd	M	JA		2)	/a:/	–	1	1	1
94	G216a003	Anl	M	DOCH für den		2^	/n/	–	1	1	3
95	G216a007	Anl	M	FEBruar		2^	/a:r/	–	1	1	3
96	G216a010	Chd	M	JA		2)	/a:/	–	1	1	1
97	G216a014	Chd	M	PRIma		2^	/a:/	–	1	1	2
98	G217a001	Chd	M	geBUNd'n		2^	/dn/	–	2	1	2
99	G217a005	Chd	M	DANN		2)	/n/	–	1	1	1
100	G251a001	Sik	F	FREItag geh'n		2^	/n/	–	1	1	3
101	G251a021	Sik	F	deZEMber		2(/ər/	–	1	1	2
102	G251a029	Sik	F	WEG		2)	/ɛ/	/k/	1	1	1

Fortsetzung auf nächster Seite

<i>Fortsetzung</i>										
Nr.	Datei	Spr	m/f	Wort	Acc	S1	S2	Ri	Ri	Pa
								[seg]	[syl]	[syl]
103	G252a007	Sik	F	JUni	2^	/i:/	–	1	1	2
104	G253a005	Sik	F	WOCH'nend- seminar	2(/a:r/	–	1	1	6
105	G254a000	Sik	F	SCHULze	2(/ə/	–	1	1	2
106	G254a004	Sik	F	TREFF'n	2^	/n/	–	1	1	2
107	G254a010	Sik	F	MAL	2)	/l/	–	1	1	1
108	G255a001	Sik	F	SECHzehnt'n	2^	/n/	–	1	1	3
109	G255a011	Sik	F	sieb'mundZWAN- zigst'n bis zum	2(/m/	–	1	1	5
110	G256a000	Sik	F	SONNtag	2(/a:/	/k/	1	1	2
111	G311a005	Sva	M	DEN schon mal	2^	/l/	–	1	1	3
112	G312a006	Sva	M	DREIzehnt'n	2)	/e:ntn/	–	4	2	3
113	G313a006	Sva	M	ORDnung	2^	/ŋ/	/k/	1	1	2
114	G313a008	Sva	M	PASS'n	2^	/n/	–	1	1	2
115	G314a001	Sva	M	LOSfahr'n	2^	/n/	–	1	1	2
116	G314a007	Sva	M	vorzubeREIt'n	2^	/n/	–	1	1	5
117	G314a013	Sva	M	geNAU	2)	/aʊ/	–	1	1	2
118	G315a008	Ame	M	einundDREIßigst'n2^ März	2^	/ɛr/	/ts/	1	1	4
119	G317a005	Sva	M	UHR	2^	/ur/	–	1	1	1

Tabelle A.4: Rohdaten zu den [2;]-Fällen

Nr.	Datei	F_0 [Hz]				Dauer [ms]				
		FaOn	FaOf	RiOn	RiOf	DuFa	DuL	DuRi	DuLS	DuE
1	G071a014	–	87	90	98	–	75	80	123	43
2	G074a005	106	–	–	104	–	–	–	218	59
3	G075a000	148	96	96	110	334	0	103	94	18
4	G081a000	173	–	–	–	–	–	–	91	22
5	G085a002	189	–	–	–	–	–	–	188	39
6	G085a005	–	145	146	156	–	203	26	59	19
7	G085a014	170	–	–	154	–	–	–	81	0
8	G085a017	222	139	139	156	296	37	34	41	0
9	G086a008	–	131	134	153	–	225	60	238	13
10	G086a008	–	–	–	154	–	–	–	118	23
11	G086a016	208	143	142	163	634	78	65	319	25
12	G086a020	–	128	130	150	–	79	64	95	0

Fortsetzung auf nächster Seite

<i>Fortsetzung</i>										
Nr.	Datei	F_0 [Hz]				Dauer [ms]				
		FaOn	FaOf	RiOn	RiOf	DuFa	DuL	DuRi	DuLS	DuE
13	G091a028	301	172	172	189	251	0	36	148	22
14	G092a000	329	172	172	184	269	0	48	151	11
15	G094a000	296	153	153	187	309	0	79	216	0
16	G094a029	368	161	161	185	438	0	49	83	0
17	G095a001	268	164	164	181	216	0	69	180	0
18	G097a005	347	177	177	196	229	0	94	204	16
19	G101a002	210	169	169	191	179	0	76	171	0
20	G101a002	237	172	170	197	411	54	66	256	22
21	G101a006	204	166	166	196	241	0	188	135	73
22	G101a015	383	179	179	186	566	34	77	159	0
23	G102a002	204	168	169	184	318	53	38	190	0
24	G102a004	239	192	192	201	240	0	55	182	0
25	G102a014	268	180	180	189	384	0	56	196	21
26	G103a000	195	172	172	192	78	92	50	150	16
27	G103a002	273	168	168	196	752	0	169	82	0
28	G103a006	334	–	–	199	–	–	–	235	36
29	G103a012	380	183	182	202	251	59	67	257	25
30	G103a014	371	199	199	224	440	0	120	65	0
31	G104a001	355	176	176	194	351	0	60	188	10
32	G104a005	379	184	184	202	406	0	103	135	20
33	G104a011	334	189	189	215	215	0	70	142	0
34	G104a013	349	192	192	208	341	109	40	296	29
35	G105a000	247	164	164	190	270	0	73	143	20
36	G105a015	283	194	193	199	187	43	19	122	11
37	G105a016	215	186	186	216	204	0	92	320	23
38	G105a018	353	198	197	202	223	149	60	71	0
39	G105a018	384	186	185	200	280	54	37	135	0
40	G106a001	420	167	167	195	312	0	68	256	22
41	G106a005	309	199	199	206	269	0	40	160	35
42	G106a007	247	192	192	221	221	0	158	117	14
43	G106a010	291	178	178	197	452	0	56	108	52
44	G107a000	262	181	182	204	530	62	112	232	0
45	G107a001	201	173	173	184	82	225	40	103	22
46	G107a004	238	193	193	200	153	0	94	217	76
47	G107a005	265	188	188	197	226	0	26	171	25
48	G107a006	315	200	200	225	238	43	71	217	0
49	G107a010	334	181	181	199	393	0	116	179	63

Fortsetzung auf nächster Seite

<i>Fortsetzung</i>										
Nr.	Datei	F_0 [Hz]				Dauer [ms]				
		FaOn	FaOf	RiOn	RiOf	DuFa	DuL	DuRi	DuLS	DuE
50	G111a011	103	94	94	109	157	0	26	52	26
51	G113a011	156	100	100	106	202	298	21	136	24
52	G115a007	162	99	99	107	183	89	27	147	0
53	G115a010	167	84	84	95	167	0	116	162	42
54	G121a012	195	171	171	190	168	79	99	107	0
55	G121a014	201	177	177	199	252	0	52	152	15
56	G121a023	269	187	189	218	227	74	65	143	14
57	G122a009	116	98	98	110	195	0	70	108	0
58	G123a005	254	190	191	204	173	41	50	193	36
59	G123a008	–	162	162	201	–	0	76	196	31
60	G123a010	294	175	174	196	219	23	119	293	31
61	G123a012	236	163	164	176	232	39	35	353	36
62	G124a014	142	100	100	113	136	0	93	79	0
63	G126a014	127	102	102	107	208	0	30	109	0
64	G142a002	182	120	120	128	150	0	24	56	32
65	G143a000	207	128	128	142	183	0	44	213	0
66	G144a003	195	115	115	144	106	0	70	70	0
67	G191a002	229	178	178	187	200	75	102	110	0
68	G191a011	246	160	160	167	157	0	72	363	19
69	G191a020	244	180	180	194	340	0	107	130	47
70	G191a021	187	142	142	155	490	59	55	85	13
71	G191a026	229	175	175	183	369	0	155	155	0
72	G191a036	214	167	167	177	81	0	99	145	0
73	G192a001	235	169	169	185	347	33	85	197	11
74	G192a003	192	163	163	179	169	0	73	130	23
75	G192a014	243	194	194	222	179	0	100	75	9
76	G192a017	195	145	145	160	127	0	35	143	0
77	G193a004	240	167	167	171	124	0	45	191	18
78	G194a003	249	166	166	176	110	23	44	140	0
79	G194a014	223	179	179	194	215	0	168	168	0
80	G195a019	176	153	153	177	117	0	144	131	0
81	G196a002	294	163	163	172	159	0	105	156	15
82	G196a005	254	155	155	175	159	0	48	119	39
83	G196a011	166	145	145	151	226	157	39	265	20
84	G211a000	106	87	87	93	296	0	117	119	32
85	G211a000	98	89	89	103	114	0	111	184	15
86	G211a003	112	94	94	104	594	0	109	109	0

Fortsetzung auf nächster Seite

<i>Fortsetzung</i>										
Nr.	Datei	F_0 [Hz]				Dauer [ms]				
		FaOn	FaOf	RiOn	RiOf	DuFa	DuL	DuRi	DuLS	DuE
87	G211a005	113	96	96	101	278	61	64	297	52
88	G212a003	108	95	95	112	210	40	66	222	18
89	G213a003	117	99	99	107	661	113	63	304	17
90	G213a003	111	88	88	91	164	0	31	214	13
91	G214a007	101	82	82	90	389	0	54	70	0
92	G214a009	116	97	97	104	159	552	105	103	13
93	G216a000	115	93	93	105	231	0	79	334	24
94	G216a003	89	85	85	91	386	0	40	103	10
95	G216a007	113	89	89	102	233	118	49	186	13
96	G216a010	100	83	83	93	205	0	41	257	27
97	G216a014	102	82	81	86	175	126	30	219	22
98	G217a001	105	91	91	113	89	0	80	140	96
99	G217a005	110	97	97	101	238	122	91	208	10
100	G251a001	272	197	197	209	359	91	75	140	0
101	G251a021	242	202	202	216	70	62	73	185	0
102	G251a029	226	198	198	208	167	0	48	118	0
103	G252a007	283	192	192	204	217	0	50	177	0
104	G253a005	281	203	203	222	626	176	49	282	12
105	G254a000	274	204	204	234	207	0	58	172	23
106	G254a004	285	203	200	216	209	89	48	224	14
107	G254a010	229	203	201	214	189	107	51	96	34
108	G255a001	235	197	195	206	357	259	80	184	0
109	G255a011	224	199	199	207	605	180	54	239	15
110	G256a000	253	188	188	201	210	0	68	205	0
111	G311a005	155	113	112	118	249	299	36	74	0
112	G312a006	138	110	110	127	240	0	296	159	55
113	G313a006	179	115	115	125	235	0	25	59	0
114	G313a008	122	109	109	115	239	0	26	133	0
115	G314a001	176	112	112	124	337	144	108	164	48
116	G314a007	–	109	111	121	–	407	63	191	18
117	G314a013	144	110	110	124	161	79	48	246	47
118	G315a008	141	114	115	139	388	64	45	96	0
119	G317a005	–	111	111	121	–	0	23	175	16

Anhang B

Perzeptionsexperiment

B.1 Fragebogen (Beispielseiten)

Siehe Abbildungen B.1, B.2 und B.3 auf den folgenden Seiten.

Fragebogen zum Wahrnehmungsexperiment

Du wirst gleich mehrfach die Äußerung „**ab Montag, dem achten**“ hören. Die Äußerung wird nicht jedesmal gleich klingen, sondern einige Versionen werden sich in der **Sprechmelodie** von anderen unterscheiden.

Deine Aufgabe besteht darin, die Äußerungen in Bezug auf verschiedene Eigenschaften zu beurteilen, z.B. ob die Äußerung für Dich wie eine Frage klingt, oder aber ob die Sprecherin eher schüchtern wirkt usw. Zu jeder Äußerung hast Du auf diesem Fragebogen eine Skala mit sieben Feldern. Am linken und rechten Ende der Skala stehen Eigenschaftswörter wie z.B. „selbstbewusst“ - „schüchtern“ oder „ unhöflich“ - „höflich“. Beispiel:

Die Äußerung klingt...

unhöflich								höflich
	3	2	1	0	1	2	3	

Bitte kreuze auf diesen Skalen an, wie die Äußerung für Dich klingt. Du kannst Dir dabei vorstellen, dass die Zahlen unter den Feldern folgende Bedeutungen haben:

0 = weder noch / weiß nicht

2 = ziemlich

1 = ein wenig

3 = sehr

Bitte kreuze immer nur ein Kästchen auf jeder Skala an!

Wenn das Experiment beginnt, wirst Du einen dreifachen Piepton hören. Du wirst jede Äußerung zweimal mit einer kurze Pause dazwischen hören, bevor Du Dein Urteil abgibst. Ein kurzer Piepton zeigt an, dass gleich die nächste Äußerung gespielt wird; gehe dann bitte auch zur nächsten Skala.

Nachdem Du jeweils neun Äußerungen beurteilt hast, folgt ein doppelter Piepton und eine etwas längere Pause, so dass Du genug Zeit hast, die Seite umzublättern. Nach etwa 17 Minuten ist der erste Teil des Versuchs vorbei und wir machen eine kurze Entspannungspause. Danach folgt ein zweiter Teil, der genauso abläuft wie der erste.

Bitte versuche, auf jeder Skala ein Urteil abzugeben und keine Skala auszulassen!

Vielen Dank!

Die Äußerung klingt wie eine...

Frage								Aussage
	3	2	1	0	1	2	3	

Die Sprecherin wirkt...

kategorisch								verhandlungsbereit
	3	2	1	0	1	2	3	

Die Sprecherin ist sich...

nicht sicher								sicher
	3	2	1	0	1	2	3	

Abbildung B.1: Erste Seite des Fragebogens mit Instruktionen und Testlauf.

Die Äußerung klingt...

unhöflich höflich

3 2 1 0 1 2 3

Die Sprecherin wirkt...

selbstbewusst schüchtern

3 2 1 0 1 2 3

Die Sprecherin wirkt...

dominant nicht dominant

3 2 1 0 1 2 3

Die Äußerung klingt...

unhöflich höflich

3 2 1 0 1 2 3

Die Äußerung klingt...

abgeschlossen unabgeschlossen

3 2 1 0 1 2 3

Die Äußerung klingt wie eine...

Frage Aussage

3 2 1 0 1 2 3

Die Sprecherin klingt...

emotional sachlich

3 2 1 0 1 2 3

Die Sprecherin wirkt...

interessiert nicht interessiert

3 2 1 0 1 2 3

Die Sprecherin wirkt...

selbstbewusst schüchtern

3 2 1 0 1 2 3

Abbildung B.2: Beispiel-Seite aus dem Fragebogen.

Vielen Dank, dass Du so lange durchgehalten hast!

Zum Schluss...

möchte ich Dich bitten, mir noch ein paar Fragen zu beantworten.

A. Zum Experiment

Wie **schwierig** war die Aufgabe für Dich, die Äußerungen anhand der Skalen zu beurteilen?

Die Aufgabe war...

schwierig								einfach
	3	2	1	0	1	2	3	

Wie **natürlich** klangen die Äußerungen?

Die Äußerungen klangen...

unnatürlich								natürlich
	3	2	1	0	1	2	3	

Falls Du eher **unnatürlich** angekreuzt hast:

Kannst in wenigen Worten ausdrücken, was Dir unnatürlich erschien?

B. Zu Dir

(Diese Angaben werden ausschließlich für die Auswertung dieser Untersuchung verwendet!)

Wie lauten die ersten beiden Buchstaben Deines Vornamens? _____

Wie lauten die ersten beiden Buchstaben Deines Nachnamens? _____

Wie alt bist Du? _____ Dein Geschlecht: O w O m

Dein Studienfach (Hauptfach): _____

Ist Deine Muttersprache Deutsch? O Ja. O Nein, sondern: _____

Hast Du vorher schon mal als Versuchsperson an einem Experiment zur Sprachwahrnehmung teilgenommen?

O Noch nie. O 1-2 mal. O 3-5 mal. O öfter als 5 mal.

Abbildung B.3: Letzte Seite des Fragebogens.

B.2 Freie Kommentare der Probanden zur Natürlichkeit der Stimuli

- Die Stimme klang leicht synthetisch.
- aus dem Zusammenhang gerissen
- die (meistens) Unabgeschlossenheit der Aussagen und die geringfügigen Unterschiede zu den sehr natürlichen Aussagen
- Kontext fehlt, Betonung lag immer auf dem Anfang (Montag), daher auf emotionaler Ebene schwer zu vergleichen
- Ich hatte bei der Stimme immer eine Keifzicken-Sekretärin bei 'ner Besprechung im Kopf.
- Es fehlt die unterschiedliche Betonung der ersten drei Wörter.
- Durch die Art der Stimme kam es z.B. sehr oft unhöflich rüber.
- Z.T. klangen die Ausdrücke wie keine der angegebenen Möglichkeiten.
- Eher die Abfolge der immer gleichen Sätze klang synthetisch.
- In normalen Gesprächen gibt noch mehr "tönend" in der Stimme kund, "was los ist" (→ emotional, höflich). Ich fand hier fast nicht rauszuhören, ob es verhandlungsbereit klang oder nicht.
- bis auf das "n" am Ende alles gleich angehört
- Man hatte deutlich das Gefühl, dass die Frau wusste, dass sie aufgenommen wurde. Das fiel meiner Meinung nach deutlich auf.
- Die formulierte Äußerung muss im Kontext betrachtet werden um sie besser beantworten zu können.

B.3 Deskriptive Statistik

Tabelle B.1: Deskriptive Statistik zum Perzeptionsexperiment:
Stim = Stimulus; **Min/ Max** = kleinster/ größter angekrenzter Wert; \bar{x} = Mittelwert; s = Standardabweichung. Die Namen der Stimuli sind wie folgt kodiert: z.B. "Stimulus S03 gepaart mit Skala 7" = S03X7. Polung der Skalen: linker Term (z.B. 'emotional'): negativ; rechter Term (z.B. 'sachlich'): positiv.

Stim	Min	Max	\bar{x}	s	Stim	Min	Max	\bar{x}	s
<i>abgeschlossen – unabgeschlossen</i>					<i>dominant – nicht dominant</i>				
S01X1	-3	1	-2,26	0,93	S01X2	-3	1	-2,23	0,95
S02X1	-3	2	-2,07	1,26	S02X2	-3	2	-1,77	1,02
S03X1	-3	3	-0,79	1,55	S03X2	-3	3	-0,67	1,43
S04X1	-3	2	-1,79	1,25	S04X2	-3	2	0	1,36
S05X1	-3	2	-0,65	1,4	S05X2	-3	3	0,19	1,64
S06X1	-3	3	-0,28	1,59	S06X2	-3	3	0,72	1,47
S07X1	-2	3	1,3	1,41	S07X2	-2	3	0,53	1,39
S08X1	-3	3	1,3	1,19	S08X2	-2	2	0,33	1,27
S09X1	-3	3	0,84	1,74	S09X2	-3	3	0,67	1,76
S10X1	-3	3	1,79	1,6	S10X2	-2	3	0,49	1,3
S11X1	-3	3	1,49	1,32	S11X2	-2	2	0,84	1,19
S12X1	-2	3	1,63	1,25	S12X2	-2	3	0,26	1,38
S13X1	-2	3	1,81	1,07	S13X2	-2	3	0,21	1,57
S14X1	-3	3	1,49	1,89	S14X2	-3	3	0,09	1,74
S15X1	-3	3	1,14	1,71	S15X2	-3	3	0,49	1,55
SK00X1	-3	2	-2,14	0,97	SK00X2	-3	2	-0,93	1,5
SKo1X1	-3	3	-0,42	1,72	SKo1X2	-2	3	0,02	1,42
<i>emotional – sachlich</i>					<i>Frage – Aussage</i>				
S01X3	-3	3	1,26	1,9	S01X4	-2	3	2,14	1,08
S02X3	-3	3	1,44	1,35	S02X4	-2	3	2,33	1,08
S03X3	-1	3	2	1,05	S03X4	1	3	2,07	0,7
S04X3	-2	2	-0,33	1,38	S04X4	-2	3	1,47	1,49
S05X3	-3	2	-0,67	1,51	S05X4	-2	3	1,23	1,15
S06X3	-3	2	-0,51	1,56	S06X4	-2	3	1,02	1,26
S07X3	-3	3	-0,86	1,36	S07X4	-3	3	-0,37	1,65
S08X3	-3	3	-0,35	1,38	S08X4	-3	3	-0,33	1,58
S09X3	-3	2	-0,49	1,28	S09X4	-3	3	-0,53	1,56
S10X3	-2	3	0,58	1,52	S10X4	-3	3	-0,56	1,83
S11X3	-3	3	-0,09	1,51	S11X4	-3	2	-1,02	1,54
<i>Fortsetzung auf nächster Seite</i>									

<i>Fortsetzung</i>									
Stim	Min	Max	\bar{x}	<i>s</i>	Stim	Min	Max	\bar{x}	<i>s</i>
S12X3	-2	3	0,26	1,42	S12X4	-3	2	-1,23	1,66
S13X3	-2	3	0,44	1,37	S13X4	-3	3	-1,21	1,68
S14X3	-3	3	-0,28	1,61	S14X4	-3	3	-0,88	1,85
S15X3	-3	3	-0,05	1,84	S15X4	-3	2	-0,98	1,9
SK00X3	-3	3	1,35	1,36	SK00X4	-2	3	1,79	1,19
SKo1X3	-2	3	0,19	1,47	SKo1X4	-2	3	1,81	1,22
<i>interessiert – nicht interessiert</i>					<i>kategorisch – verhandlungsbereit</i>				
S01X5	-3	3	0,09	1,64	S01X6	-3	-1	-2,19	0,76
S02X5	-3	2	0,37	1,56	S02X6	-3	2	-1,42	1,31
S03X5	-3	3	0,44	1,4	S03X6	-3	2	-0,7	1,57
S04X5	-2	3	0,42	1,26	S04X6	-2	3	-0,16	1,45
S05X5	-2	2	0,05	1,25	S05X6	-3	3	-0,51	1,62
S06X5	-2	2	-0,05	1,34	S06X6	-2	2	0,28	1,39
S07X5	-2	3	0,09	1,31	S07X6	-2	2	0,37	1,23
S08X5	-2	2	0,16	1,17	S08X6	-3	3	0,51	1,2
S09X5	-3	1	-0,56	1,2	S09X6	-3	3	0,86	1,39
S10X5	-2	2	-0,7	1,06	S10X6	-2	3	0,05	1,38
S11X5	-3	2	-1,26	1,22	S11X6	-2	3	0,93	1,49
S12X5	-2	2	-1,21	0,94	S12X6	-2	3	0,44	1,82
S13X5	-3	2	-1,28	1,2	S13X6	-3	2	0,86	1,34
S14X5	-3	2	-1,07	1,3	S14X6	-3	3	0,98	1,81
S15X5	-3	2	-1,58	1,16	S15X6	-3	3	1,05	1,57
SK00X5	-2	3	0,98	1,26	SK00X6	-3	2	-0,72	1,42
SKo1X5	-3	3	0,44	1,53	SKo1X6	-2	3	-0,3	1,52
<i>nicht sicher – sicher</i>					<i>selbstbewusst – schüchtern</i>				
S01X7	-2	3	2	1,43	S01X8	-3	2	-1,72	1,22
S02X7	-2	3	2,05	1,11	S02X8	-3	2	-1,51	1,32
S03X7	-2	3	1,63	1,2	S03X8	-3	2	-0,95	1,38
S04X7	-2	3	0,79	1,32	S04X8	-3	2	-0,58	1,53
S05X7	-3	2	0,02	1,55	S05X8	-3	2	-0,26	1,53
S06X7	-3	3	-0,42	1,61	S06X8	-2	2	0,21	1,37
S07X7	-3	3	-0,6	1,47	S07X8	-3	3	0,19	1,64
S08X7	-3	3	-0,49	1,5	S08X8	-3	3	0,51	1,49
S09X7	-3	2	-1,02	1,49	S09X8	-2	2	0,3	1,37
S10X7	-3	3	-0,7	1,75	S10X8	-3	3	-0,21	1,6
S11X7	-3	3	-0,35	1,54	S11X8	-3	2	-0,53	1,67
S12X7	-3	2	-0,7	1,32	S12X8	-2	3	0,05	1,38
S13X7	-3	3	-0,74	1,62	S13X8	-2	3	0,16	1,65
<i>Fortsetzung auf nächster Seite</i>									

<i>Fortsetzung</i>									
Stim	Min	Max	\bar{x}	<i>s</i>	Stim	Min	Max	\bar{x}	<i>s</i>
S14X7	-3	3	-0,19	1,89	S14X8	-3	3	-0,67	1,61
S15X7	-3	3	-0,53	1,64	S15X8	-3	2	-0,74	1,42
SK00X7	-2	3	1,58	1,38	SK00X8	-3	3	-0,51	1,59
SKo1X7	-3	3	0,47	1,52	SKo1X8	-3	2	-0,16	1,48
<i>unhöflich – höflich</i>									
S01X9	-3	3	-1	1,53					
S02X9	-3	3	-1,23	1,63					
S03X9	-2	3	-0,21	1,47					
S04X9	-3	3	0,23	1,32					
S05X9	-3	3	0,14	1,41					
S06X9	-2	3	0,19	1,4					
S07X9	-3	2	-0,09	1,19					
S08X9	-3	3	-0,02	1,3					
S09X9	-2	3	0,42	1,4					
S10X9	-3	2	-0,35	1,48					
S11X9	-3	3	0,28	1,44					
S12X9	-3	3	0,63	1,29					
S13X9	-2	3	0,47	1,39					
S14X9	-3	3	0,37	1,54					
S15X9	-3	3	0,63	1,8					
SK00X9	-3	3	0,35	1,43					
SKo1X9	-2	3	0,14	1,25					

Anhang C

CDROM

Die CDROM befindet sich im hinteren Einband und enthält:

<i>Datei(en)</i>	<i>Beschreibung</i>
/Beispiele/Kap3/*	Hörbeispiele zu Abschnitt 3.6
/Beispiele/Kap5/g095a007.wav	Hörbeispiel zu Abbildung 5.1
/Stimuli/Manipulationen/Original.wav	Original des Sprachmaterials für die Stimulusgenerierung
/Stimuli/Manipulationen/Original-m.wav	Sprachmaterial mit manipuliertem /a/ in “ab”
/Stimuli/Manipulationen/Original-1.wav	Sprachmaterial, stilisiert auf dem finalen /n/
/Stimuli/Manipulationen/Original-pl.wav	Sprachmaterial, stilisiert ohne tiefes Plateau
/Stimuli/einzeln/*	Stimuli als einzelne <i>wav</i> -Dateien
/Stimuli/Kontinuum.wav	Das Stimuluskontinuum im Ganzen (ohne Kontrollstimuli)
/Stimuli/Test/*	Die drei Teile des Perzeptionstests
/Fragebogen.pdf	Der vollständige Fragebogen
/Rohdaten-pe.pdf	Die Rohdaten zum Perzeptionsexperiment

Erklärung

Hiermit erkläre ich, dass ich die vorliegende Arbeit selbständig und ohne fremde Hilfe angefertigt habe. Ferner versichere ich, dass diese Arbeit noch nicht zum Zwecke der Erlangung der Magisterwürde an anderer Stelle vorgelegen hat.

Kiel, d. 26. August 2004

(Gilbert Ambrazaitis)

Lebenslauf

Name:	Gilbert Ingwer Ambrazaitis
Geburtstag:	15. März 1979
Geburtsort:	Kiel
Eltern:	Marianne und Brian Ambrazaitis
Staatsangehörigkeiten:	deutsch, britisch
1985–1989	Friedrich-Junge-Grundschule Kiel
1989–1998	Ricarda-Huch-Gymnasium Kiel
Sommer 1998	Abitur
August 1998 – August 1999	Zivildienst in Kiel
September 1999	Immatrikulation an der Christian-Albrechts-Universität zu Kiel für den Magisterstudiengang: <ul style="list-style-type: none">• Phonetik u. digitale Sprachverarbeitung (Hauptfach)• Allgemeine Sprachwissenschaft (Nebenfach)• Informatik (Nebenfach)
Oktober 2000	Zwischenprüfung im Nebenfach Informatik
Juli 2001	Zwischenprüfung im Hauptfach Phonetik u. digitale Sprachverarbeitung
September 2001	Zwischenprüfung im Nebenfach Allgemeine Sprachwissenschaft
Januar – März 2003	ERASMUS-Studienaufenthalt an der Universität Stockholm
November – Dezember 2004	mündliche Prüfungen im Hauptfach und den beiden Nebenfächern (s.o.)
Dezember 2004	Erlangung des Grades eines Magister Artium (M.A.) mit der vorliegenden Arbeit
Akademische Lehrer:	Prof. Dr. J.M. Harrington (Phonetik) Prof. Dr. K.J. Kohler (Phonetik) PD Dr. A. Simpson (Phonetik) Prof. Dr. U. Mosel (Allgemeine Sprachwissenschaft) Prof. Dr. H.-J. Klein (Informatik)